Parallel System Support
Programs for AIX

Diagnosis Guide

Version 3 Release 4

GA22-7350-03

Parallel System Support
Programs for AIX

Diagnosis Guide

Version 3 Release 4

GA22-7350-03

Note!

Before using this information and the product it supports, read the information in “Notices” on page 597.

Fourth Edition (December 2001)

This edition applies to version 3, release 4 of the IBM Parallel System Support Programs for AIX (PSSP) licensed program (product number 5765-D51) and to all subsequent releases and modifications until otherwise indicated in new editions. This edition replaces GA22-7350-02. Significant changes or additions to the text and illustrations are indicated by a vertical line (|) to the left of the change.

IBM welcomes your comments. A form for readers’ comments may be provided at the back of this publication, or you may address your comments to the following address:

International Business Machines Corporation
Department 55JA, Mail Station P384
2455 South Road
Poughkeepsie, NY 12601-5400
United States of America

FAX (United States & Canada): 1+845+432-9405
FAX (Other Countries):
Your International Access Code +1+845+432-9405

IBMLink (United States customers only): IBMUSM10(MHVRCFS)
Internet e-mail: mhvrcfs@us.ibm.com

If you would like a reply, be sure to include your name, address, telephone number, or FAX number.

Make sure to include the following in your comment or note:
- Title and order number of this book
- Page number or topic related to your comment

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Figures xv

Tables xvii

About this book xix
Who should use this book. xix
Typographic conventions xix

Part 1. Detecting and investigating PSSP problems 1

Chapter 1. Diagnosing SP problems overview 3
How to use this book 3
SP systems and PSSP software supported by this book 3
When to use this book 3
Special troubleshooting considerations 4
Essential documentation - Other manuals to accompany this book 4
Preparing for your first problem before it happens 7
Knowing your SP structure and setup 7
Making effective use of the IBM Support Center 9
When to contact the IBM Support Center. 9
Information to collect before contacting the IBM Support Center 10
How to contact the IBM Support Center. 13

Chapter 2. Detecting SP problems and keeping informed 15
Runtime notification methods. 15
Graphical tools - SP Perspectives 15
Command line tools 20
Asynchronous (batch) notification methods. 30
Graphics tools - SP Event Perspective 31
Command line tools - Problem Management 32
Automating your response to problems 38
Important - WHEN actions are performed 38
Important - WHERE actions are performed 39
Graphical tools - SP Event Perspective 40
Command line tools - Problem Management 40

Chapter 3. Conditions to monitor on the SP system 43
Conditions to monitor using Perspectives or Problem Management. 43
Monitor these hardware conditions. 43
Monitor these software conditions 43
Descriptions of each condition 44
Preparing to examine and monitor this information 54
SP Event Perspective - conditions that you can monitor using the default event definition 55
SP Event Perspective - conditions that you can monitor that you must define to the SP Event Perspective 55
SP Event Perspective - creating the event definitions. 60
SP Hardware Perspective 62
Problem Management 62

Chapter 4. Error logging overview 67
Classifying Error Log events 67
Effect of not having a battery on error logging 68

Managing and monitoring the error log 68
Viewing error log information in parallel 68
Summary log for SP Switch, SP Switch2, and switch adapter errors 69
Viewing SP Switch error log reports 69
Using the AIX Error Notification Facility 70
Using the SP logs 74

Chapter 5. Producing a system dump 81
Actions 81
Action 1. Produce a system dump 81
Action 2. Verify the system dump 83

Chapter 6. Diagnosing hardware and software problems 87
High-Level SP symptoms 87

Part 2. Diagnosing PSSP subsystems 89

Chapter 7. Diagnosing NIM problems 91
Related documentation 91
Requisite function 91
Error information 91
Trace information 92
NIM debug SPOT 92
NIM SPOT logs. 94
Information to collect before contacting the IBM Support Center 94
Error symptoms, responses, and recoveries 94
Actions 95

Chapter 8. Network installation progress 97

Chapter 9. Diagnosing node installation problems 101
Related documentation 101
Requisite function 101
Error information. 101
Trace information 102
Post-installation customization trace. 102
Nodecond log 102
Information to collect before contacting the IBM Support Center 102
Error symptoms, responses, and recoveries. 102
Actions 103

Chapter 10. Diagnosing boot problems. 113
Related documentation 113
Error information. 113
Node LED/LCD indicators 113
Trace information 113
Console log. 113
Information to collect before contacting the IBM Support Center 113
Error symptoms, responses, and recoveries. 114
Actions 114

Chapter 11. Diagnosing Root Volume Group problems 117
Actions 117
Action 1 - Check disks. 117
Action 2 - Check disk allocation 118
Action 3 - Force the Root Volume Group extension 118

Action 4 - Unlock the Root Volume Group 118
Action 5 - Add space to physical volumes. 118
Action 6 - Add physical volumes to the Root Volume Group 118
Action 7 - Verify the number of copies of AIX on the node for mirroring 118
Action 8 - Verify the number of copies of AIX on the node for unmirroring 119
Action 9 - Check for user logical volumes on the physical volume 119
Action 10 - Verify mirroring or unmirroring 119
Root Volume Group terminology 119

Chapter 12. Diagnosing system connectivity problems 121
Actions 121
Action 1 - Diagnose multiple nodes 121
Action 2 - Diagnose individual nodes 121
Action 3 - Diagnose a network problem 122
Action 4 - Diagnose a Topology-related problem 122

Chapter 13. Diagnosing IP routing problems 123
IP source routing. 123

Chapter 14. Diagnosing SDR problems. 125
Related documentation 125
Requisite function 125
Error information 126
Trace information 127
SDR daemon trace 127
Information to collect before contacting the IBM Support Center 128
Diagnostic procedures. 128
Check current system partition. 128
Query the state of the sdrd 128
Stop and start the sdrd 129
Check the sdrd processes using the ps command 129
Check for sdrd memory leaks and CPU utilization 129
Check for an sdrd hang 130
Check for an sdrd core dump 130
SDR verification test 130
Look for MBCS data in the SDR 131
Error symptoms, responses, and recoveries. 131
Actions 132

Chapter 15. Diagnosing SP Switch problems 137
Related documentation 137
Requisite function 137
Internal SP Switch subsystem components 137
Error information 138
AIX Error Log information 138
Adapter configuration error information. 152
SP Switch device and link error information 154
Dump information 156
errpt.out 156
css_dump.out 156
fs_dump.out 157
cssadm.debug 157
cssadm.stderr 157
cssadm.stdout. 157
regs.out 158
router.log 158

CSS_test.log 158
vdidl.out 159
spdata.out 159
netstat.out 159
scan_out.log and scan_save.log 159
Trace information 160
flt 160
rc.switch.<PID>.<date>.<time> 160
rc.switch.log 161
out.top 161
dist_topology.log 161
worm.trace 161
fs_daemon_print.file 162
daemon.stdout 162
| logevnt.out 162
| summlog.out 162
daemon.stderr. 162
dtbx.trace 162
dtbx.failed.trace 163
cable_miswire 163
css.snap.log 163
Ecommands.log 163
spd.trace 163
Missing error data warning 163
Information to collect before contacting the IBM Support Center 164
css.snap package 164
Log files within css.snap package 165
Disk space handling policy 168
Diagnostic procedures. 168
SP Switch diagnostics. 169
Clock diagnostics 170
SP Switch adapter diagnostics. 172
Cable diagnostics 173
SP Switch node diagnostics. 174
SP Switch advanced diagnostics 177
Error symptoms, responses, and recoveries. 177
SP Switch symptoms and recovery actions 177
Recover an SP Switch node 178
Worm error recovery 178
Ecommands error recovery 180
Estart error recovery 180
Unfence an SP Switch node 181
Eunfence error recovery 182
switch_responds is still on after node panic 183

Chapter 16. Diagnosing SP Switch2 problems 185
Related documentation 185
Requisite function 185
Internal SP Switch2 subsystem components 186
Error information. 186
SP Switch2 log and temporary file hierarchy 187
| plane.info file 187
AIX Error Log information 188
Adapter configuration error information. 205
SP Switch2 device and link error information 206
Dump information 210

- errpt.out 210
- cadd_dump.out 210
- ifcl_dump.out 210
- logevnt.out 210
- summlog.out 211
- col_dump.out 211
- cssadm2.debug 211
- cssadm2.stderr 211
- cssadm2.stdout 212
- regs.out 212
- router.log 212
- CSS_test.log 212
- odm.out 213
- spdata.out 213
- netstat.out 213
- scan_out.log and scan_save.log 213
- DeviceDB.dump 214

Trace information 214
- rc.switch.log 214
- daemon.stdout 214
- daemon.log 214
- adapter.log 214
- flt 214
- fs_daemon_print.file 215
- out.top 215
- topology.data 216
- cable_miswire 216
- css.snap.log 216
- colad.trace 216
- spd.trace 217
- Ecommands.log 217
- emasterd.log 217
- emasterd.stdout 217
- chgcss.log 217
- la_error.log 217
- la_event_d.trace 217

Missing error data warning 218

Information to collect before contacting the IBM Support Center 218
- css.snap package 218
- Log files within css.snap package 219
- Disk space handling policy 223

Diagnostic procedures 223
- SP Switch2 diagnostics 224
- SP Switch2 adapter diagnostics 226
- Cable diagnostics 227
- SP Switch2 node diagnostics 228
- SP Switch2 Time Of Day (TOD) diagnostics 231
- SP Switch2 advanced diagnostics 233

Error symptoms, responses, and recoveries 233
- SP Switch2 symptoms and recovery actions 233
- Recover an SP Switch2 node 234
- Worm error recovery 235
- Ecommands error recovery 239
- Estart error recovery 239
- Unfence an SP Switch2 node 240
- Eunfence error recovery 241

switch_responds is still on after node panic 242

Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools 245
Adapter Error Log Analyzer (ELA) 246
When to run the adapter ELA 246
How to run the adapter ELA 246
Interpreting the results of the adapter ELA 246
SP Switch or SP Switch2 stress test 247
When to run the SP Switch or SP Switch2 stress test 247
How to run the SP Switch or SP Switch2 stress test. 247
Interpreting the results of the SP Switch or SP Switch2 stress test 247
Multiple senders/single receiver test. 248
When to run the Multiple senders/single receiver test 248
How to run the multiple senders/single receiver test 248
Interpreting the results of the multiple senders/single receiver test 248
SP Switch or SP Switch2 wrap test 249
When to run the SP Switch or SP Switch2 wrap test 249
How to run the SP Switch or SP Switch2 wrap test 249
Interpreting the results of the SP Switch or SP Switch2 wrap test 249

Chapter 18. Diagnosing SP Security Services problems 251
Related documentation 251
Requisite function 254
Distributed Computing Environment (DCE) Version 3.1 for AIX 254
System Data Repository (SDR) 255
Error information. 256
Log files 256
Dump information 258
Information to collect before contacting the IBM Support Center 258
SP Security Services configuration errors. 258
SP trusted services authentication errors 259
SP trusted services authorization errors 259
Diagnostic procedures. 260
Find out about your configuration. 260
Verify DCE Security Services configuration 261
Verify configuration of Kerberos V4 263
Check enabled security 265
Check authentication 266
Check authorization. 268
Error symptoms, responses, and recoveries. 271
Actions 272

Chapter 19. Diagnosing Per Node Key Management (PNKM) problems 289
Related documentation 289
Requisite function 290
Error information. 290
AIX Error Log 290
Debug information 292
Information to collect before contacting the IBM Support Center 292
Diagnostic procedures. 293
Has the spnkeyman daemon been started? Is it sleeping? 293
Are key files for the SP System Services on the host where spnkeyman is running? 293
Have the password expiration times been changed in the SP DCE organization, spsec-services, to indicate when keys should expire? . . . 294
Have the SP server keys expired? 294

Was the password lifetime expiration time set to less than 24 hours? . . . 294
Has the SP DCE organization, spsec-services, been created, and have the SP service principals been added to that organization? 295
Is the DCE registry populated with the principals and accounts for SP Services? 295
Is DCE installed on the host where spnkeyman is running? 295
Are the DCE daemons running on the host where spnkeyman is running? 295
Is There a route for the DCE client daemons to access the DCE server daemons? 295
Is the DCE secd daemon running on the DCE server host? 295
Error symptoms, responses, and recoveries 296

Chapter 20. Diagnosing remote command problems on the SP System 299
| Enhanced Security Option 299
| Things to be aware of when using Restricted Root Access (RRA) 299
| Using secure remote commands instead of AIX rsh and rcp commands 300
Related documentation 301
Requisite function 302
| Using AIX rsh and rcp commands 302
Error information 304
Information to collect before contacting the IBM Support Center 305
AIX remote command diagnostics 306
Diagnostics specific to Kerberos V4 308
Diagnostics specific to Kerberos V5 309
Error symptoms, responses, and recoveries 309
Remote command (rsh/rcp) symptoms 309
Remote commands (rsh/rcp) recovery actions 310
| Using secure remote commands - symptoms 315
| Using secure remote commands - recovery actions 315

Chapter 21. Diagnosing System Monitor problems 317
Related documentation 317
Requisite function 317
Error information 318
AIX Error Logs and templates 318
System Monitor daemon log file 320
Dump information 320
Trace information 321
Information to collect before contacting the IBM Support Center 321
Diagnostic procedures 321
Installation verification tests 321
Configuration verification tests 323
Operational verification tests 327
Error symptoms, responses, and recoveries 328
Actions 329

Chapter 22. Diagnosing SP Logging daemon Problems 333
Related documentation 333
Requisite function 333
Error information 333
AIX Error Log 333
hwevents file 335
splogd.debug file 335
Dump information 336
Information to collect before contacting the IBM Support Center 336
Diagnostic procedures 336

Installation verification tests 336
Configuration verification tests 337
Operational verification tests 339
Error symptoms, responses, and recoveries 348
Actions 348

Chapter 23. Diagnosing Topology Services problems 351
Related documentation 351
Requisite function 351
Error information 352
AIX Error Logs and templates 352
Dump information 375
Core dump 375
phoenix.snap dump 376
Trace information 379
Topology Services service log 379
Topology Services user log 381
hats or topsvcs script log 382
| Network Interface Module (NIM) log 383
Information to collect before contacting the IBM Support Center 384
Diagnostic procedures 386
Installation verification test 386
Configuration verification tests 387
Operational verification tests 390
Error symptoms, responses, and recoveries 402
Actions 402

Chapter 24. Diagnosing Group Services problems 419
Related documentation 419
Requisite function 420
Error information 420
AIX Error Logs and templates 420
Dump information 426
Core dump 426
phoenix.snap dump 427
Trace information 427
GS service log trace 427
| GS service log trace - summary log 429
Group Services startup script log 430
Information to collect before contacting the IBM Support Center 430
How to find the GS nameserver (NS) node 432
How to find the Group Leader (GL) node for a specific group 433
Diagnostic procedures 434
Installation verification test 434
Configuration verification test 435
Operational verification tests 436
Error symptoms, responses, and recoveries 446
Actions 447

Chapter 25. Diagnosing Event Management problems 455
Related documentation 455
Requisite function 455
Internal components of Event Management 456
Resource monitors 456
Service information 458
Automatic method - phoenix.snap 458

Manual method of data collection. 458
Error information 460
AIX Error Log for Event Management 460
Error log files 460
Event Management daemon errors 460
EMCDB problems 461
Resource Monitor problems. 462
Dump information 463
Dump information from Event Management Resource Monitor daemons 463
Trace information 463
Trace facility built into Event Management 463
Information to collect before contacting the IBM Support Center 464
Diagnostic instructions. 464
Verify SP software installation 464
Identify the failing node 467
Verify event registration 468
Verify Resource Monitors. 472
Error symptoms, responses, and recoveries 472
Recover crashed node (Event Management Resource Monitor daemons) 472
Recover EMCDB. 472
Security errors 472

Chapter 26. Diagnosing IBM Virtual Shared Disk problems 473
Related documentation 473
Requisite function 474
Before you begin. 474
Error information 474
Errors logged by the IBM Virtual Shared Disk device driver and IBM Recoverable Virtual Shared Disk subsystem 474
Dump information 477
Trace information 478
Internal virtual shared disk device driver circular trace buffer. 478
IBM Recoverable Virtual Shared Disk subsystem tracing to the console log 478
IBM Recoverable Virtual Shared Disk subsystem logging of recovery actions 478
Information to collect before contacting the IBM Support Center 479
Diagnostic procedures. 479
Installation verification test 479
Configuration verification tests 479
Operational verification tests 481
Error symptoms, responses, and recoveries 487
Recognizing recovery 487
Planning for recovery 487
Virtual Shared Disk node failure 488
Switch failure scenarios 488
Topology Services or recovery service daemon failure 489
Disk EIO errors 490

Chapter 27. Diagnosing Job Switch Resource Table Services problems 491
Related documentation 491
Requisite function 491
Error information 492
Job Switch Resource Table Services error and information log 492
AIX Error Logs and templates for JSRT Services 492
Information to collect before contacting the IBM Support Center 493
Diagnostic procedures - Installation verification. 494
Error symptoms, responses, and recoveries 494

Actions 495

Chapter 28. Diagnosing User Access problems 501
Actions 501
Action 1 - Check the /etc/security/passwd file 501
Action 2 - Check Login Control 501
Action 3 - Verify that the automount daemon is running 501
Stopping and restarting automount 505

Chapter 29. Verifying System Management installation 507
Verification test output 507
What system management verification checks 508
Objects tested by SYSMAN_test on the control workstation only 509
Objects Tested by SYSMAN_test on the control workstation and boot/install servers 509
Objects tested by SYSMAN_test on all SP nodes (not the control workstation) 509
Objects relating to optional system management services. 510
Additional tests 510
Interpreting the test results 511

Chapter 30. Diagnosing Perspectives problems on the SP System 513
Related documentation 513
Requisite function 514
Error information 514
Information to collect before contacting the IBM Support Center 514
Diagnostics 515
Error symptoms, responses, and recoveries 515
Actions 517

Chapter 31. Diagnosing file collections problems on the SP System . . . 523
Related documentation 523
Requisite function 523
SP configuration 524
Error information 525
Information to collect before contacting the IBM Support Center 525
Diagnostics 526
File collections errors on any SP host 527
File collections configuration errors 528
File collections server (control workstation or boot/install server) errors 528
File collections client errors 530
Error symptoms, responses, and recoveries 530

Part 3. Diagnosing SP node and network attached hardware 533

Chapter 32. SP-specific LED/LCD values 535
Other LED/LCD codes. 539

Chapter 33. Diagnosing 604 and 604e High Node problems 541
604 and 604e High Node characteristics 541
Error conditions and performance considerations 541
Using SystemGuard and BUMP programs 541

Chapter 34. Diagnosing POWER3 SMP Thin and Wide Node problems 543
POWER3 SMP Thin and Wide Node characteristics. 543
Boot sequence for the POWER3 SMP Thin and Wide Node 543
Error conditions and performance considerations 544

Service Processor surveillance 545

Chapter 35. Diagnosing POWER3 SMP High Node problems 547
POWER3 SMP High Node characteristics 547
Boot sequence for the POWER3 SMP High Node 547
Error conditions and performance considerations 548
Service Processor Surveillance 548
SP Expansion I/O Unit. 548

Chapter 36. Diagnosing 332 MHz SMP Thin and Wide Node problems 551
332 MHz SMP Node characteristics. 551
Boot sequence for the 332 MHz SMP Node 551
Error conditions and performance considerations 552
Service Processor surveillance 552

Chapter 37. Diagnosing dependent node configuration problems 553
SP configuration diagnosis 553
SP Switch Router configuration diagnosis 555
SNMP configuration diagnosis 556

Chapter 38. Diagnosing SP-attached Server and Clustered Enterprise Server problems 559
Related documentation 559
Requisite function 560
Error information 560
  Logs 560
Dump information 562
Trace information 562
Information to collect before contacting the IBM Support Center 562
Diagnostic procedures. 563
  Installation verification tests 563
  Configuration verification tests 564
  Operational verification tests 566
Error symptoms, responses, and recoveries 569
  Actions 570

Chapter 39. Diagnosing IBM @server pSeries 690 problems 577
Related documentation 577
Requisite function 577
Error information 578
  Logs 578
  Dump information 579
  Trace information 580
Information to collect before contacting the IBM Support Center 580
Diagnostic procedures. 580
  Installation verification tests 580
  Configuration verification tests 582
  Operational verification tests 584
Error symptoms, responses, and recoveries 588
  Actions 588

Chapter 40. Diagnosing PSSP T/EC Event Adapter problems 593

Part 4. Appendixes 595

Notices 597

Trademarks. 598
Publicly available software 599

Glossary of Terms and Abbreviations 601

Bibliography 609
Information formats 609
Finding documentation on the World Wide Web 609
Accessing PSSP documentation online 609
Manual pages for public code 610
RS/6000 SP planning publications 610
RS/6000 SP hardware publications 610
RS/6000 SP Switch Router publications 611
| Related hardware publications 611
RS/6000 SP software publications 611
| AIX publications 613
DCE publications 613
Redbooks 613
Non-IBM publications 613

Index 615

Figures

Tables

1. Details about each condition to monitor . . . 44
2. Conditions and default event definitions . . . 55
3. SP Error Log label suffixes mapped to syslog priorities and AIX Error Log types . . . 67
4. SP log files . . . 74
5. System dump information. . . . 81
6. System dump status codes . . . 83
7. High-Level SP symptoms . . . 87
8. NIM symptoms . . . 94
9. Sample NIM installation trace . . . 97
10. Node Installation Symptoms . . . 103
11. Boot symptoms . . . 114
12. Root Volume Group symptoms . . . 117
13. Verification of mirroring or unmirroring for Root Volume Groups . . . 119
14. System connectivity symptoms . . . 121
15. System Data Repository (SDR) symptoms . . . 131
16. Resource Name failure indications - SP Switch . . . 138
17. Possible causes of adapter failures - SP Switch . . . 139
18. Possible causes of TB3MX specific adapter failures . . . 140
19. Possible causes of fault service daemon failures - SP Switch . . . 140
20. Possible causes of adapter diagnostic failures - SP Switch . . . 152
21. adapter_config_status values - SP Switch . . . 152
22. SP Switch device status and recovery actions. . . . 154
23. SP Switch link status and recovery actions . . . 155
24. AIX Error Log entries that invoke css.snap - SP Switch . . . 164
25. SP Switch log files and their location . . . 165
26. Clock problems and recovery actions . . . 171
27. SP Switch adapter Service Request Number failures and recovery actions . . . 173
28. SP Switch symptoms and recovery actions . . . 177
29. SP Switch worm return codes and analysis. . . . 179
30. SP Switch Estart problems and analysis . . . 181
31. Resource Name failure indications - SP Switch2 . . . 188
32. Possible causes of adapter failures - SP Switch2 . . . 189
33. Possible causes of fault service daemon failures - SP Switch2 . . . 193
34. Possible causes of adapter diagnostic failures - SP Switch2 . . . 203
35. Possible causes of SP Switch2 TOD management (emasterd) failures. . . . 204
36. Possible causes of SP Switch2 PCI Adapter failures . . . 204
37. adapter_config_status values - SP Switch2. . . . 205
38. SP Switch2 device status and recovery actions . . . 207
39. SP Switch2 link status and recovery actions . . . 208
40. AIX Error Log entries that invoke css.snap - SP Switch2. . . . 219
41. SP Switch2 log files . . . 220
42. SP Switch2 rc.switch.log file and SDR equivalents . . . 225
43. SP Switch2 adapter Service Request Number failures and recovery actions . . . 227
44. SP Switch2 symptoms and recovery actions . . . 234
45. SP Switch2 worm return codes and analysis . . . 235
46. SP Switch2 Estart problems and analysis . . . 240
47. SP Security Services symptoms and recovery actions. . . . 272
48. AIX Error Log templates for Per Node Key Management. . . . 291
49. Per Node Key Management symptoms and recovery actions . . . 296
50. Remote commands - rsh and rcp dependencies . . . 303
51. Remote command (rsh/rcp) symptoms and recovery actions . . . 309
52. Remote to remote rcp . . . 314
53. secure remote command symptoms and recovery actions . . . 315

54. AIX Error Log templates for the System Monitor . . . 318
55. System Monitor symptoms . . . 328
56. AIX Error Log templates for the Logging daemon . . . 334
57. Logging daemon symptoms . . . 348
58. AIX Error Log templates for Topology Services . . . 354
59. Topology Services symptoms . . . 402
60. AIX Error Log templates for Group Services . . . 421
61. Group Services symptoms . . . 446
62. Arguments for Event Management daemon . . . 468
63. IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk Error Codes . . . 475
64. AIX Error Log templates for JSRT Services. . . . 493
65. Job Switch Resource Table (JSRT) Services symptoms . . . 494
66. JSRT Services return codes . . . 496
67. SWTBLSECMSGLEVEL environment variable values . . . 497
68. DCE security groups for JSRT Services . . . 499
69. User Access symptoms . . . 501
70. Automounter Related Commands . . . 501
71. Objects tested by SYSMAN_test on the control workstation only . . . 509
72. Objects Tested by SYSMAN_test on the control workstation and boot/install servers . . . 509
73. Objects tested by SYSMAN_test on the SP nodes . . . 509
74. Optional system management objects tested by SYSMAN_test . . . 510
75. Additional system management tests . . . 510
76. SP Perspectives symptom types. . . . 515
77. Launch pad and general Perspectives symptoms . . . 516
78. SP Hardware Perspective symptoms . . . 516
79. IBM Virtual Shared Disk Perspective symptoms . . . 516
80. SP Event Perspectives symptoms . . . 517
81. Perspectives - related file sets . . . 517
82. SP Perspectives resource variables . . . 519
83. File collection symptoms and recovery actions . . . 530
84. SP-specific LED/LCD values (Chronological order) . . . 535
85. SP-Specific LED/LCD values (Numerical order) . . . 537
86. SP-attached Server and Clustered Enterprise Server symptoms . . . 569
87. IBM @server pSeries 690 symptoms. . . . 588

About this book

This book contains information to help you diagnose and resolve problems for IBM RS/6000 SP systems and Parallel System Support Programs for AIX (PSSP). It does not contain the following:

- Information about diagnosing other SP products, such as Parallel Environment (PE) and LoadLeveler. These products have their own publications.
- Information about other SP system management issues. For this information, see *PSSP: Administration Guide*.

For a list of related books and information about accessing online information, see the Bibliography in the back of the book.

This book applies to PSSP Version 3 Release 4. To find out what version of PSSP is running on your control workstation (node 0), enter the following:

```
splst_versions -t -n0
```

In response, the system displays something similar to:

```
0 PSSP-3.4
```

If the response indicates **PSSP-3.4**, this book applies to the version of PSSP that is running on your system.

To find out what version of PSSP is running on the nodes of your system, enter the following from your control workstation:

```
splst_versions -t -G
```

In response, the system displays something similar to:

```
1 PSSP-3.4
2 PSSP-3.2
7 PSSP-3.1.1
8 PSSP-2.4
```

If the response for a particular node indicates **PSSP-3.4**, this book applies to the version of PSSP that is running on that node.

If you are running mixed levels of PSSP, be sure to maintain and refer to the appropriate documentation for whatever versions of PSSP you are running.

Who should use this book

This book is intended for system administrators, who are responsible for setting up and maintaining the SP system. This book can also be used by system operators and others, who are responsible for monitoring the status of the SP system and interacting with the hardware.

It is assumed that the reader has a working knowledge of AIX or UNIX and experience with network systems.

Typographic conventions

This book uses the following typographic conventions:

Typographic	Usage
Bold	• **Bold** words or characters represent system elements that you must use literally, such as commands, flags, and path names.
Italic	• *Italic* words or characters represent variable values that you must supply. • *Italics* are also used for book titles and for general emphasis in text.
`Constant width`	Examples and information that the system displays appear in `constant width` typeface.
[ ]	Brackets enclose optional items in format and syntax descriptions.
{ }	Braces enclose a list from which you must choose an item in format and syntax descriptions.
|	A vertical bar separates items in a list of choices. (In other words, it means “or.”)
< >	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, **<Enter>** refers to the key on your terminal or workstation that is labeled with the word Enter.
...	An ellipsis indicates that you can repeat the preceding item one or more times.
<Ctrl-*x*>	The notation **<Ctrl-*x*>** indicates a control character sequence. For example, **<Ctrl-c>** means that you hold down the control key while pressing **<c>**.
\	The continuation character is used in coding examples in this book for formatting purposes.

Part 1. Detecting and investigating PSSP problems

Chapter 1. Diagnosing SP problems overview

This chapter contains information to help you diagnose problems you may encounter installing or operating the SP system and PSSP. It helps you to identify whether a problem is related to the hardware or the software. It also shows you the procedure to follow if you require assistance from the IBM Support Center.

How to use this book

This section discusses which level of PSSP is supported by this manual, when to use this manual, and what other manuals are needed to diagnose PSSP problems.

SP systems and PSSP software supported by this book

The book applies only to PSSP Version 3 Release 2 (PSSP 3.2). This book does not supersede previous versions. To display the levels of PSSP installed on all nodes of your SP system, see “About this book” on page xix. If your SP system has mixed levels of PSSP, each version of PSSP has its own *PSSP: Diagnosis Guide*. Use the proper version of the manual to diagnose problems with a particular node. In order to diagnose problems on a node running a particular level of PSSP, you must use the manual that applies to that level.

When to use this book

Consult this manual for assistance during these administrative efforts:

- Before contacting the IBM Support Center to report a problem

  This manual lists the basic information that you should have available before contacting the IBM Support Center, and how to obtain that information. In addition to the basic information, specific PSSP software subsystems may require you to provide additional data that is specific to the failing subsystem or to the particular problem that you are experiencing. To understand what information is required and how to obtain it, consult both the basic instructions for preparing information for the IBM Support Center, and the diagnostic instructions for the failing PSSP subsystems in “Part 2. Diagnosing PSSP subsystems” on page 89..
- When you encounter problems while operating the SP hardware or the PSSP software

  This manual contains diagnostic procedures provided for PSSP software problems and some SP hardware problems. This manual also contains descriptions of available error information, how to retrieve it, and what to look for to analyze the problem. There are also solutions to some commonly encountered problems.
- When preparing to troubleshoot the SP hardware or PSSP software

  This manual provides instructions for running diagnostic procedures for PSSP software subsystems. These procedures are specific to the individual subsystem or SP hardware device being examined. Consult “Part 2. Diagnosing PSSP subsystems” on page 89 for these instructions.
- To become familiar with existing services that monitor system status

  PSSP provides graphical and command-line facilities to display the current status of system resources and to monitor changes in these resources. This manual introduces these tools and demonstrates how to use them to assess the current status of the SP hardware and software. Using these utilities, you can detect problems at an early stage and react to them before they propagate and magnify.

Special troubleshooting considerations

1. **Enhanced Security Option**

   PSSP 3.2 provides the option of running your RS/6000 SP system with an enhanced level of security. This function removes the dependency PSSP has to internally issue AIX **rsh** or **rcp** commands as a root user from a node. When this function is enabled, PSSP does not automatically grant authorization for a root user to issue AIX **rsh** or **rcp** commands from a node. If you enable this option, some procedures may not work as documented. For example, to run HACMP an administrator must grant the authorizations for a root user to issue AIX **rsh** or **rcp** commands, which PSSP would otherwise grant automatically.

2. **DCE Restriction**

   If you have DCE authentication enabled, you cannot run HACWS.

3. **Enablings secure remote command Option**

   PSSP 3.4 provides the ability to remove the dependency that PSSP has on the AIX **rsh** and **rcp** commands issued as root on the control workstation and nodes, by enabling the use of secure remote command and secure remote copy methods. When using the secure remote commands, the Restricted Root Remote commands is also enabled, limiting the use of remote commands to secure remote commands from the control workstation to the nodes. When this function is enabled, PSSP will use the secure remote command methods enabled for all remote command calls, no longer relying on the **rsh** and **rcp** commands.

   In addition, in PSSP 3.4 you have the ability to set Authorization for AIX Remote Commands to ″none″ on a system partition or on all system partitions, when secure remote commands are enabled. When this is set, PSSP code will not automatically grant authorization for the root user to issue the AIX **rsh** and **rcp** commands for a node or the control workstation to the system partition. Instead, all PSSP remote commands will be run using the secure remote command method enabled in that system partition.

   In order to set AIX Authorization for Remote Commands to ″none″ on any SP system partition, PSSP 3.4 must be installed on all nodes of that partition. If ″none″ is enabled, certain functions and procedures may not work as documented. See *PSSP: Administration Guide* for enabling secure remote commands and the ″none″ option.

Essential documentation - Other manuals to accompany this book

The “Bibliography” on page 609 lists manuals of general interest for the SP system and PSSP. This section lists those manuals specific to problem detection and problem solving for the SP system and for AIX. This section contains references to manuals not listed in the Bibliography.

1. For all levels of PSSP running on your SP system, *PSSP: Diagnosis and Messages Guide*

   These manuals are needed because SP systems allow different levels of PSSP to run on different nodes. Information for previous versions of PSSP and the hardware that they support may be present only in previous versions of this manual. This manual (PSSP 3.2 version) applies only to software currently in PSSP 3.2 and the hardware that PSSP 3.2 supports.

   When to use these books:

   - When a problem occurs on a node running a level of PSSP other than PSSP 3.2
   - When failures occur in hardware that was introduced on SP systems before the release of PSSP 3.2, and this hardware is not supported on PSSP 3.2

The information for that hardware will be included in *PSSP: Diagnosis and Messages Guide* for versions of PSSP that support the hardware.

- When a problem occurs in a PSSP subsystem that runs on several different nodes within the SP and those nodes are at different levels of PSSP, including PSSP 3.2

  In most cases, the subsystems that run on nodes with different levels of PSSP use a ″backward compatibility″ mode. This means that all nodes provide only that level of function available in the lowest version of PSSP on those nodes.

  For example, if Group Services (GS) is running on an SP system with four nodes, configured as follows:

  – two nodes running PSSP 3.2

  – one node running PSSP 2.4

  – one node running PSSP 2.3

  GS will essentially run as if PSSP 2.3 was installed on all nodes. Any additional function provided by GS for PSSP 2.4, 3.1, and 3.2 will not be available.

  This manual discusses diagnostic procedures for software provided in PSSP 3.2. Diagnostic procedures for software provided in PSSP 3.1 and PSSP 3.1.1 are documented in the appropriate version of *PSSP: Diagnosis Guide*. Diagnostic procedures for software provided in versions of PSSP before PSSP 3.1 are documented in the appropriate version of *PSSP: Diagnosis and Messages Guide*. These diagnostic procedures may still be valid when software is running in a ″backward compatibility″ mode on PSSP 3.2 nodes.

2. *PSSP: Messages Reference*

   This manual is the companion to this one (*PSSP: Diagnosis Guide*). The *PSSP: Messages Reference* provides a list of the specific error messages generated by PSSP 3.2, gives a detailed explanation of the error condition, and gives directions for responding to the error condition.

   When to use this manual:

   - When PSSP 3.2 software encounters problems or failures accompanied by error messages, to understand the nature of the failure and how to respond to it.
   - When you encounter a specific error message and wish to resolve the problem without having to diagnose the entire subsystem that issued the message.

3. *PSSP: Administration Guide*

   This manual describes the supported configurations of the SP system. This manual also describes how to configure and administer the SP system.

   This manual, *PSSP: Diagnosis Guide*, has many references to sections in *PSSP: Administration Guide* when describing how to analyze, circumvent, or repair problems.

   When to use this manual:

   - When installing, customizing, and configuring the SP system
   - When adding nodes and other computer resources to the SP system
   - When diagnosing potential SP hardware and PSSP software problems, to verify that the configuration and customization of your SP system is correct
   - When diagnosing potential SP hardware and PSSP software problems, to verify that the configuration and customization of your SP system is supported

- When responding to specific SP hardware or PSSP software error conditions, and you need more information on the commands and procedures you are instructed to use.

4. *RS/6000 SP: PSSP 2.2 Survival Guide (SG24-4928)*

   Although this manual is specific to PSSP 2.2, much of the information in this manual is still relevant to later releases of PSSP. This manual gives you insight to the SP hardware and PSSP structure, so that you can understand how problems in one component impact others.

   This manual provides specific error avoidance and recovery instructions for items that still exist or are supported in PSSP 3.2, such as:

   - Manual conditioning of a high node
   - Initial SP system setup
   - Node installation and Network Install Manager (NIM)
   - Switch topology and system partitioning
   - Network Time Protocol (NTP), time-of-day synchronization, and the problems that can happen with these components
   - Tips on using Problem Management for monitoring the SP and responding to error conditions
   - Tips on using the SP Error Log Management Facility

   When to use this manual:

   - When problems are encountered on nodes running PSSP 2.2
   - When nodes are encountered in distributed subsystems where one of the nodes involved is running PSSP 2.2
   - When looking for assistance in installing and customizing any version of PSSP
   - When trying to diagnose and recover from system partitioning problems

5. *RS/6000 SP: Problem Determination Guide (SG24-4778)*

   Although this manual is specific to PSSP 2.1, much of the information in this manual is still relevant to later releases of PSSP. This manual provides specific error detection and recovery instructions for items that are still used and supported in PSSP 3.2, such as:

   - Control workstation installation
   - Kerberos security subsystem
   - System partitioning
   - Changing IP addresses or hostnames of the nodes

   When to use this manual:

   - When troubleshooting Kerberos-related problems
   - When troubleshooting system partitioning problems
   - To use as a guide when changing node IP addresses or network hostnames

6. *RS/6000 SP Monitoring: Keeping It Alive (SG24-4873)*

   Although this manual is specific to PSSP 2.2, much of the information in this manual is still relevant to later releases of PSSP. This manual is needed because it describes the utilities built into the PSSP software that permit you to monitor what is going on in your SP system, how to be alerted when things go wrong, and how to automate the response to specific conditions. Topics include:

   - RS/6000 Cluster Technology (which is called ″HA Infrastructure″ in *RS/6000 SP Monitoring: Keeping It Alive*)

- SP Perspectives
- Problem Management

When to use this manual:

- To introduce yourself to the problem monitoring facilities available in PSSP, and how to use them.
- To understand what resources can be monitored.
- When learning how to automate your response to certain problems that can be handled without human interaction. An example of this is the process of expanding file system space when there is a shortage.

7. *AIX Version 4 Problem Solving Guide and Reference*

   This manual provides assistance in investigating and resolving AIX operating system problems. Consult this manual when you suspect a problem with the AIX operating system, or when you suspect that an AIX problem is contributing to an SP system problem.

8. *AIX Version 4 Messages Guide and Reference*

   This manual contains the list of three-digit LED/LCD display values for SP nodes. This manual is used when a node reports a three-digit LED/LCD display value, so that the reader can understand its meaning and appropriate action to take in response. SP hardware and PSSP software also issue LED/LCD display values, which are documented in "Chapter 32. SP-specific LED/LCD values" on page 535. If the LED/LCD cannot be located there, consult *AIX Version 4 Messages Guide and Reference*.

9. *IBM DCE for AIX, Version 3.1: Problem Determination Guide*

   This manual lists error messages and recovery actions along with administrative tips and general information. The purpose of this guide is to help programmers and administrators to interpret and to act on error messages and status codes when received.

Preparing for your first problem before it happens

This section explains how to obtain and record information about your SP system that you will need when you first problem occurs. You may not have the time or the means to obtain this information after a failure has occurred. The best strategy is to prepare this information before a failure occurs, and to have it handy before investigating possible problems.

Knowing your SP structure and setup

Problem investigation efforts are streamlined considerably by knowing the characteristics of the SP system at the time that a problem occurs. This includes what node types are being used, what software is installed on these nodes, what level of software is installed, what software service is installed, and so forth.

Create a log of your SP structure and setup

PSSP: Planning Volume 1 provides guidance in planning your physical site and selecting your hardware. *PSSP: Planning Volume 2* provides guidance for logically laying out the SP system structure and the SP administrative network, and selecting your software.

Examine your SP system, its structure and its software setup, and record this information in a log. Keep this log in a place where you will always have access to it, regardless of whatever failure occurs on your system. To avoid the possibility of losing this log to an online failure, it is best to keep this log in hardcopy format.

This list is the minimum amount of information to record in the log:

1. Your customer information:
   - Your access code, which is your customer number
   - Names and phone numbers of people whom the IBM Support Center should contact to assist you with problem resolution
2. Control workstation Information
   - The level of AIX installed
   - The PTF numbers for all fixes installed on AIX
   - The product number of PSSP on the control workstation. You need to use this to report a problem.
   - The PTF numbers for all fixes installed on PSSP
3. System Partitioning or Cluster information - what nodes are in the system partition or cluster. Also, what software is installed and active on that cluster.
4. Node information
   - The node’s number, its frame number and associated slot number, which can be found by issuing the **/usr/lpp/ssp/bin/splstdata -n** command (note: this is a system partition sensitive command.)
   - The node’s hostname and IP address, which can be found using the **/usr/lpp/ssp/bin/splstdata -a** command (note: this is a system partition sensitive command.)
   - The level of AIX installed on the node
   - The PTF numbers for all fixes installed to AIX
   - The product number of PSSP installed, which may be different from that installed on the control workstation
   - The PTF numbers for all fixes installed to PSSP
   - Optional software installed, including version numbers and fixes
   - Special hardware characteristics, such as: wide node, network attached node, twin-tailed DASD

Update the log whenever a failure occurs

Whenever an actual or suspected failure occurs on the SP system, make an update to this log. Record symptoms that are noticed at the time of the failure, and system conditions such as:

- The date and time that the problem was discovered
- The nature of the problem, such as a node halt, an abnormal program termination, request hang, poor response time, or hardware failure
- What nodes the problem was experienced on
- What software was running on the nodes at the time the problem was encountered
- What users were using the system at the time that the problem was encountered
- What actions that you or others took to repair or bypass the problem, and whether these actions were successful

Recording this information serves several purposes:

- It allows you to recognize recurring problems and to quickly find the steps needed to resolve or bypass the problem.
- It records details about the conditions that existed in the SP system at the time of the failure. This information is essential when contacting the IBM Support Center to report problems.

- It allows you to detect patterns in the occurrence of problems.

  Perhaps these problems occur on a regular basis, or whenever a specific program runs, or when a specific system resource is unavailable or reaching maximum limits. These patterns are difficult to detect unless historical data on past failures is available. Having this information available can assist you and the IBM Support Center in detecting patterns in these failure conditions.

Update the log whenever system conditions change

Using outdated or incomplete information when investigating a failure leads to wasted time. The wrong information is obtained and analyzed, the wrong diagnostic procedures are performed, and in some cases an incorrect solution is applied. This causes problem conditions to remain the same or become worse. It also may introduce additional problems. To avoid this wasted effort, be sure to update this log whenever the SP system structure or setup changes.

Update the log whenever the following occurs:

- New software is installed
- Upgrades are made to new levels of AIX or PSSP
- AIX or PSSP PTFs are installed
- Node hostname or IP addresses change
- Hardware changes, Switch adapter changes
- Problems occur

Making effective use of the IBM Support Center

There are several things you need to know in order to make effective use of the IBM Support Center. You need to know when to call IBM, how to contact IBM, and what information to collect before calling.

When to contact the IBM Support Center

Contact the IBM Support Center for the following situations:

- A repeated or persistent halt of an SP node
- A repeated or persistent hang of an SP node
- A repeated or persistent failure or hang of specific SP software

  These failures may not always occur on the same node, given the distributed nature of this software.
- A failure in mission-critical PSSP software

A single node or infrequent software failure that is not mission-critical may not be a cause to contact the IBM Support Center immediately. These problems may be caused by conditions that can be remedied through administrative techniques. Investigate these failures, using this manual as a guide for conducting the investigation. Follow these steps:

- Determine what was active on the system at the time.
- See who was using the system.
- Record the date and time of the failure.
- Determine what hardware was in use.
- Determine what specific services were being used at the time that the failure was detected
- Use the information in this manual and “Essential documentation - Other manuals to accompany this book” on page 4 to analyze and correct the problem.

Log information about these failures that you discover in the course of your investigations. This information can be used for your own future reference, and by the IBM Support Center if this failure becomes frequent enough or critical enough to require their assistance, as follows:

- The log information permits you to respond more quickly to similar failures in the future, and helps you to remember how to resolve the problem.
- The log information can be used for pattern analysis.

Problems and failures may appear to be unrelated at first, but they may have some relationship that is not immediately evident. Examine the conditions that were recorded for previous infrequent failures to see if there may be a pattern to them, even if the failure seem to be unrelated. Consider the following items when looking at the historical data on problems and failures:

- Do they happen when similar programs, procedures, or jobs are run?
- Do they happen when certain people or groups use the system?
- Do they happen at specific times, days, or shifts (peak or off-peak hours)?
- Does the failure occur when specific hardware is used?
- Are node reboots the only way to resolve the problem?

Contact the IBM Support Center when you discover any patterns in infrequent failures because:

- The system configuration may need repair.
- The IBM Support Center may have information about the problem you are experiencing.
- You may be experiencing a problem that no one else has ever encountered or reported.

Information to collect before contacting the IBM Support Center

ATTENTION - READ THIS FIRST

Read this section **in its entirety** and perform **ALL** of the instructions listed here before placing a call to IBM. Some of the required information must be captured **immediately** before system conditions change, the data is lost, or the data is overwritten.

1. Your Customer Information, which should be in the log discussed earlier. See "Create a log of your SP structure and setup" on page 7.
   - Your access code, which is your customer number
   - Names and phone numbers (external numbers, complete with area codes) where you can be reached by IBM service representatives
2. Your Product Information:
   - The PSSP product number for the level of PSSP running on the control workstation

     If PSSP 3.2 is running, the product number is **5765-D51**. If a level of PSSP other than PSSP 3.2 is running, the product number can be obtained from *PSSP: Diagnosis and Messages Guide* for that level.
   - The PTF numbers for all PSSP fixes installed on the control workstation
   - The version number of AIX running on the control workstation
   - The PTF numbers for all AIX fixes installed on the control workstation
   - For all nodes involved in the problem, obtain this information:

   - The version number of AIX running on the node
   - The PTF numbers for all AIX fixes installed on the node
   - The PSSP version number on the node
   - The PTF numbers for all PSSP fixes installed on the control workstation
3. Information about your problem. Different information is needed for different kinds of problems. Therefore, you cannot collect the same set of information for all problems. Here are some general rules:
   - For a single node halt or crash, you will need the following information:
     a. A system dump from the halted node. The system may already have created this dump for you, but you must verify this. Examine the halted node's LED indicator using the **spmon -Led** command. If the display shows a flashing **888**, a dump was started on the node. Use the **spmon -reset** command to step through the LED values until the flashing **888** appears again, recording all these LED values. Use Table 6 on page 83 to determine if the dump has completed and verify its contents.
     b. The **/unix** file from the halted node. This file will be obtained automatically by the service tools.
     c. The error log from the halted node. This log will also be created by the service tools.
     d. System Data Repository (SDR) information from the control workstation, that is used to find environment or configuration problems.
        1) Login the control workstation as **root**.
        2) Issue the following commands, redirecting the output to a file:
           a) **splstdata -e >** *filename*
           b) **splstdata -n -G >>** *filename*
           c) **splstdata -s -G >>** *filename*
        3) Write the file to the media, and label it as ″SDR information″.

     To obtain this information:
     a. Verify the contents of the dump. Use "Action 2. Verify the system dump" on page 83.
     b. Ensure that the **/tmp** file system has at least 8MB of free space.
     c. Make sure that the **/unix** file is the one that was used when the dump occurred (in the case where the **bosboot -k** command was used to select another **/unix**).
     d. Login as **root** and issue the **snap -r** command to clear the current contents of the **/tmp/ibmsupt** directory.
     e. Make sure that a tape drive is accessible to the node.
     f. Issue the **snap -Dgo** *tape device* command.
     g. Label the media with the node name, node number, its contents (dump, **/unix**, snap information), and the command used to create it: **snap -Dgo** *tape device*.
     h. Enable write-protect on the media and put in a safe place.
     i. For additional information on this process, see *RS/6000 SP: Problem Determination Guide* .
   - For multiple node halts or crashes, you will need the same information as in single node halts and crashes. Examine each system using the **spmon -Led** command to determine if a system dump was taken on each halted node.

     To obtain the necessary information from the halted nodes:

a. With the node still in its crashed or halted state, make sure that a system dump has been taken. If a dump does exist, do not re-create it. If a dump does not exist, one needs to be created. **DO NOT** create a dump if the node's LED shows flashing **888**. Use "Action 1. Produce a system dump" on page 81 to create the dump. Use the Primary Dump Device, unless this is impossible, because other tools will assume that the dump is located there.

b. Verify the contents of the dump. Use "Action 2. Verify the system dump" on page 83.

c. Once the node is rebooted and the dump verified, ensure that the **/tmp** file system has at least 8MB of free space available.

d. Make sure that the **/unix** file is the one that was used when the dump occurred (in the case where the **bosboot -k** command was used to select another **/unix**).

e. On the control workstation, ensure that the **/tmp** file system has at least 8MB of free space for each node that has halted. For example, if four nodes have halted, make sure that 32MB of space are available.

f. On the control workstation, build a file containing the hostnames of the nodes that have halted or crashed. The hostnames should be on one line, separated by commas with no intervening white space characters. For example: node1a,node1b,node5d. Save the name of this file for use in the next two steps.

g. On the control workstation, issue:

```
splm -a service -t filename -r
```

to clear the current contents of the **/tmp/ibmsupt** directories on these nodes. *filename* is the name of the file from Step 9f.

h. On the control workstation, issue:

```
splm -a service -t filename -c -p Dg
```

to start the **snap -Dg** command on these nodes. *filename* is the name of the file from Step 9f.

i. Ensure that a tape drive is available on the control workstation.

j. On the control workstation, issue:

```
splm -a gather -k service -t \
filename –l /tmp/servcol -o tape_device_name
```

to retrieve the service information. The command retrieves the information from the nodes listed in the file, writes this information temporarily to the **/tmp/servcol** file, then archives the data to the tape device in **tar** format. *filename* is the name of the file from Step 9f. *tape_device_name* is the name of the tape drive.

k. If the **/tmp/servcol** file remains on the control workstation, remove it.

l. Label the media with the names of the nodes involved, their node numbers, the contents of the tape (system dumps, **/unix** files, **snap** information), and the command used to create the tape from Step 9j.

m. Enable the write protection on the media and put it in a safe place.

- Node hangs or experiences response problems

a. A dump from the hung node is preferred. Go through the steps given previously for manually creating and verifying a dump of the hung nodes.

b. The **/unix** file is also needed, as in the previous procedure.

    c. Reboot the node and run the **snap** command given in the previous section to collect the data and create the media.

- Failures in specific PSSP software subsystems, including denial of service problems and performance problems
    a. Consult “Part 2. Diagnosing PSSP subsystems” on page 89 and find the diagnosis chapter for the failing PSSP subsystem. This chapter may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem.
    b. If error log information is needed from multiple nodes, the **splm** command can be used to consolidate these logs in one location.
    c. Do not generate a system dump unless the subsystem’s diagnosis chapter instructs you to do so.
    d. Write all information requested to the media and clearly label it.
- Failure in other SP software which is supplied by IBM
    a. Consult the diagnosis documentation for the failing product. This information may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem
    b. Do not generate a system dump unless the product’s diagnosis instructions instruct you to do so.
    c. Write all information requested to the media and clearly label it.
- Failure in non–IBM software
    a. Consult the diagnostic documentation for the failing product. This information may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem.
    b. Follow problem reporting procedures for that product.
- SP hardware failures
    a. Perform hardware diagnostic procedures associated with the hardware and record any information requested by these instructions. For details, see “RS/6000 SP hardware publications” on page 610.

How to contact the IBM Support Center

IBM Phone Numbers

In the United States:
The number for IBM software support is **1-800-237-5511**.
The number for IBM hardware support is **1-800-IBM-SERV**.
The number for the PC Help Center is **1–800-772-2227**.

Outside the United States, contact your local IBM Service Center.

Contact the IBM Support Center using the phone number above, for these problems:

- Node halt or crash not related to a hardware failure
- Node hang or response problems
- Failure in specific PSSP software subsystems
- Failure in other SP software supplied by IBM

The person with whom you speak will ask for the information from “Information to collect before contacting the IBM Support Center” on page 10 and give you a time period during which an IBM representative will return your call.

For failures in non-IBM software, follow the problem reporting procedures documented for that product.

For SP hardware failures, contact IBM Hardware Support at the number above.

For PC problems, contact the PC Help Center at the number above. Have your machine type and serial number ready.

For any problems reported to the IBM Support Center, a Problem Management Record (PMR) is created. A PMR is an online software record used to keep track of software problems reported by customers.

- The IBM Support Center representative will create the PMR and give you its number.
- Have the SDR information you collected earlier handy because it may be needed for inclusion in the PMR.
- Record the PMR number. **YOU WILL NEED IT** to send data to the IBM Support Center. YOU WILL ALSO NEED IT on subsequent phone calls to the IBM Support Center to discuss this problem.
- Write the PMR number on ALL media you created in the previous steps, even if you are not going to send this data to the IBM Support Center at this time. The Support Center may request the data at a later time, so you want to ensure that neither the media nor the PMR number corresponding to it is lost.
- To send the media to the IBM Support Center, use this address:

IBM Mailing Address

IBM RS/6000 Scalable POWERparallel Systems
Dept. 39KA, M/S P961, Bldg. 415
2455 South Road
Poughkeepsie, N.Y. 12601-5400

ATTN: APAR Processing

Note: If you are using multiple packages or envelopes to send the media, be sure to label them in a series, such as ″1 of 5″, ″2 of 5″, and so forth.

Be sure that the person you identified as your contact can be reached at the phone number you provided in the PMR.

Chapter 2. Detecting SP problems and keeping informed

The best way of streamlining your problem resolution is to prevent problems from occurring. To minimize the frequency and impact of problems, follow the configuration recommendations in *IBM RS/6000 SP: Planning, Volume 1, Hardware and Physical Environment* and *IBM RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, and use the tools documented in *PSSP: Administration Guide*. You should also follow the recommendations documented for any software you install.

However, problems may still occur. When they do, the best way to resolve these problems is to detect them as soon as they occur, and correct or bypass them before they impact the ability of other subsystems, causing secondary and tertiary failures. Several methods exist for detecting problems on the SP system.

The SP system provides the capability to detect problem situations in a runtime fashion when the system administrator is actively monitoring system conditions. The SP system also has asynchronous notification methods for use when the system administrator is not directly monitoring system conditions.

Runtime notification methods

PSSP provides tools to monitor system status and conditions in a runtime fashion, when the system administrator is actively monitoring the current status of the system. These tools are used when the system administrator wants to know immediately the current status of system resources, or to be notified immediately of problems and potential trouble situations.

Two sets of runtime tools are available. The choice of the tools depends on the capabilities of the system administrator's workstation and the system administrator's preferences. PSSP provides **graphical tools** for use on the control workstation or network-attached terminals. PSSP also provides **command-line tools** for those situations when only modem access or **s1term** access is available.

Graphical tools - SP Perspectives

PSSP provides graphical tools for system administration and monitoring through the SP Perspectives tool suite. Perspectives is engineered for ease-of-use for the system administrator. In order to be used effectively, SP Perspectives requires X11 graphics capable terminals or workstations and high-speed connections. Use Perspectives when monitoring the SP system from the control workstation or from a network-attached workstation.

The basic concepts of Perspectives and examples of its use are included in the SP Perspectives chapter of *PSSP: Administration Guide*. Perspectives also provides extensive online help information. To understand how to accomplish the tasks that are presented in this chapter, consult the SP Perspectives online help, using this section as a guide to the online help topics.

Individual Perspectives require that certain subsystems be operating and that the user is authorized to communicate with them. Such subsystems include the System Monitor, Event Management, System Data Repository, and Problem Management. For authorization required for each Perspective, see the discussion on using SP Perspectives in *PSSP: Administration Guide*.

The Perspectives launch pad is started using the **perspectives** command, which resides in the **/usr/lpp/ssp/bin** directory. Other Perspectives, such as the Event or Hardware Perspective, can be started from the launch pad. Before starting an SP Perspective, be sure that the **DISPLAY** environment variable is set to the machine that you want to display the SP Perspective. Also, be sure that you are permitted to display to that machine by running the **xhost** command on that machine.

Two Perspectives tools are useful for monitoring the system status and detecting problem situations:

The SP Event Perspective

This tool allows the user to specify system conditions that are of concern or importance, and to indicate what actions are to be taken when the condition exists. The Perspective interfaces with the Event Management software subsystem to monitor these conditions and alert the Perspective to the presence of the condition. To effectively use this Perspective, you must understand certain terminology.

Condition
The circumstances within the system that are of interest to the system administrator. Conditions can be created, viewed, and modified through the **Conditions** pane in the SP Event Perspective. To specify a condition, the system administrator must provide the necessary components to form the condition, including an event expression and, optionally, a rearm expression. Their definitions follow.

The rearm expression indicates when the SP Event Perspective should consider the event to have ″stopped″. For example, a file system is considered ″almost full″ when the available space is less than 10% of its capacity. The system administrator may want to consider the condition to exist until the available space reaches 13% of the file system’s capacity. The event expression would then be set to 10% and the rearm expression to 13%. As with the event expression, the system administrator can indicate an action to take when the rearm expression occurs, such as deactivating reserve resources that had been activated when the event occurred.

Event Expression
A relational expression that specifies the circumstances under which an event is generated.

Rearm Expression
A relational expression that specifies that the condition that triggered the event is no longer true. It is usually the inverse of the event expression.

Event Definition
An association made by the system administrator between a condition and a response to the presence of that condition.

Registration
The activation of an event definition. By registering an event definition, the system administrator instructs the Perspective to begin monitoring for the condition and to take the associated action if the condition should occur.

Once the user registers the event definition, the action will be run whenever the event or rearm expression occurs. This is independent of whether the Event Perspective is active at the time that the event or rearm expression occurs.

Event A change in the state of a system resource. For the purposes of this discussion, an event is more narrowly defined as the presence of the condition within the system.

To start the SP Event Perspective, double click on the **Event Perspective** icon in the SP Perspectives launch pad window.

Users can create conditions for situations that are important to them through the **Conditions** pane of the SP Event Perspective. A number of default conditions are provided through the SP Event Perspective. You may wish to add more or to tailor the predefined conditions to meet the specific needs of your particular SP installation. The Perspectives online help provides assistance on how to create conditions and how to modify existing conditions. To access this help, click on the **Help** button from the SP Event Perspective display, and select the **Tasks...** option. Assistance in handling conditions is available through the **Working with Conditions** topic.

Once a condition is defined through the SP Event Perspective, an action can then be associate with it. The action may be as simple as a visual notification that the event has occurred, or the action can be more sophisticated, including automatically invoking a command in response to the event. To associate the appropriate action with the presence of the condition (or to the absence of the condition), an event definition must be created. You can create these definitions and examine default definitions through the **Event Definitions** pane of the SP Event Perspective. The Perspective online help provides assistance on how to create event definitions and how to modify existing definitions. To access this help, click on the **Help** button from the SP Event Management Perspective display and select the **Tasks...** option. Assistance in handling event definitions is available through the **Working with Event Definitions** topic.

Only after both the condition and its associated event definition are defined to the Perspective, can you begin the monitoring of the condition. This is done by registering the event definition through the **Event Definitions** pane in the SP Event Perspective. To find how this is done, consult the Perspective's **Working with Event Definitions** online help topic.

Other basic SP Event Perspective tasks are described in the online help. To access this information, click on the **Help** button from the SP Event Management Perspective display, select the **Tasks...** option, and click on the **How Do I ...?** topic.

Depending on how the event definition was constructed, the SP Event Perspective reacts in one or more of the following ways when you register the event definition, and the condition that the event definition is based on occurs:

- The icon representing the event definition within the SP Event Perspective's Event Definitions pane changes to an envelope. This notification can be detected only if the SP Event Perspective is running. If the SP Event Perspective is shut down after you register for the condition, this visual notification is not presented.
- The action associated with the event definition is started. This action is specified when you create the event definition. Through this action, you can automate the response to the condition, such as sending e-mail to a system administrator, issuing a command to activate a pager, or issue an administrative command to allocate reserved resources to address the condition.

  Once you register the event definition, the action runs whenever the event occurs, whether or not the SP Event Perspective is active at the time the event occurs.

The actions performed when the event or the rearm expression occurs can be one of the following:

- A command - The command can perform controls, enable or disable resources, or notify you by other means (like mail or online messages).
- An SNMP Trap - This transmits a notification to the network using SNMP protocols, indicating that an event has occurred. This trap can be configured so that certain SNMP applications can receive the notification, or all can receive it. NetView is an example of such an application. Use an SNMP trap when you are using SNMP-based monitoring tools (such as NetView), and you want these tools to detect when events occur on the SP system.
- An entry in the AIX Error Log and the BSD System log - This is used to record a persistent record of the event, or the event's rearm condition. The AIX Error Log template **HA_PMAN_EVENT_ON** is used when the event condition occurs. The template **HA_PMAN_EVENT_OFF** is used when the rearm condition occurs. These templates are viewed by issuing the command:

  ```
  errpt -at -J HA_PMAN_EVENT_ON -J HA_PMAN_EVENT_OFF
  ```

  Notification can be sent to the system administrator whenever these templates are logged to the AIX Error Log. For instructions on setting up this notification, consult "Using the AIX Error Notification Facility" on page 70.

The SP Event Perspective is designed to be a multi-user tool. Multiple users can invoke the SP Event Perspective in parallel and monitor different conditions. Notifications are routed to those users that registered the associated event definition. The Perspective also stores event definitions created by each user in the user's **$HOME/.$USER:Events** file. By storing these definitions in different files, each user can tailor conditions and event definitions to best suit the user's needs. This also prevents users from accidentally modifying conditions or event definitions created or used by other SP Perspectives users.

The SP Hardware Perspective

This tool allows you to examine the current status of the SP system hardware. Through this tool, you can display a graphical representation of the system's overall structure, assess the current status of system hardware, and issue hardware control commands.

- **To examine the current status of a hardware device**, select the hardware device by single-clicking on the device's icon in the particular pane. Open the devices' notebook by single clicking the notebook icon (leftmost icon) in the toolbar at the top of the Perspective window. This displays a new window (notebook) that contains the device's current status, settings, and monitored conditions. This is useful for examining a node's LED values, its responsiveness to the network and the switch, its network configuration, and other information.

  For further assistance in using the notebook to view hardware status, consult the Perspective's online help. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Tasks...** option. Assistance in viewing hardware status is available in the **Viewing Hardware Attributes** topic.

  If you want to view the same hardware information from multiple entities, such as the responsiveness to the switch for a series of nodes, opening a notebook for each entity can be time-consuming. The SP Hardware Perspective offers an alternative method for displaying this information. Most information displayed in a notebook can also be displayed in the pane in a table format. The left column of the table contains the objects from the pane while the columns to the right contains the information you want displayed from the notebook.

  To switch from the icon to table view in a pane, select the icon on the right of the toolbar which shows a table and an icon. When you point at this icon, the descriptive text reads ″Show object in the table view or the icon view″. The first time you select this toolbar icon, the ″Set Table Attributes″ dialog is displayed.

This dialog lists the attributes from the objects notebook that you can display in the table. After selecting the desired items from the list, select ″OK″. The pane will be updated with the items you selected. The table entities that represent variable states of the hardware entity will be color coded to indicate ″good″, ″bad″, ″caution″ status as they did in the notebook.

For assistance on using the table view to examine hardware status, consult the Perspective online help. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Overview..** option. From the new window that appears, select the **Starting and Customizing SP Perspective** topic, then select the **Customizing SP Perspective** subtopic, and finally select the **Using Table View** item. The Help option from the selection list window provides a fast path to the help topic.

- **To monitor the status of hardware devices**, select the pane where the devices are contained. Looking at the top row of icons, a graph icon should be visible on the right. When you point at this icon, the descriptive text reads ″Set up and begin monitoring″. Click on this icon to bring up a window of items that can be monitored. Select the items to be monitored from this list. All objects in the pane will now be monitored for these conditions.

  When monitoring is active, the icons of the entities use a visual indication of the status. If all monitored conditions do not indicate a problem, the icons will be presented in a green color. If any of the monitored conditions indicate a problem, the icon will appear to have a red X drawn through it. Note that this will occur even if only one condition indicates a problem. For example, if five nodes are being monitored for five conditions, and one of these five conditions appears on node1a, the icon for node1a will appear with a red X through it, while the remaining nodes will be represented with green icons.

  To determine what condition may exist on a marked entity, select the object by single-clicking on its icon or its table entry in the pane. Then open the object’s notebook by clicking on the notebook icon in the upper left corner. When the notebook display comes up, page forward to the ″Monitored Conditions″ page. This page lists conditions being monitored for that object, along with the condition’s current state. Any state listed as ″Triggered″ indicates that the condition is present.

  If any object in a pane is presented in gray with a question mark (**?**) drawn over it for longer than a few seconds, a communication problem exists between the SP Hardware Perspective and the Event Management software subsystem. For assistance in resolving the problem, consult “Chapter 30. Diagnosing Perspectives problems on the SP System” on page 513.

  For further assistance in setting up and starting the hardware monitor, consult the Perspectives online help facility. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Tasks...** option. From the new window that appears, select the **Monitoring Hardware Objects** topic.

There are some characteristics of the SP Hardware Perspective that the user should keep in mind when using the tool. Unlike the SP Event Perspective, the SP Hardware Perspective does not permit the user to associate an action with the presence of a condition. Users that wish to automate a response to a specific system condition should use the SP Event Perspective. Also, the SP Hardware Perspective only monitors conditions while it is active. If the Perspective is shut down, any monitoring of hardware status is also shut down.

To be able to restart the Perspectives so that monitoring will automatically start, you will need to save the configuration to a profile. From the menu bar select **Options → Save Preferences...** The Save Preferences dialog will be displayed. For more information on using this dialog, select the Help button at the bottom. To start the Hardware Perspective with the saved profile:

- If you saved your preferences as a user profile, enter

```
sphardware -userProfile name_you_specified
```

- If you saved your preferences as a system profile, enter

```
sphardware -systemProfile name_you_specified
```

Previous versions of PSSP offered a graphical user interface as part of the System Monitor (**spmon**) command. PSSP Version 3.1 and later versions of PSSP, incorporate this hardware control capability into the SP Hardware Perspective. While the capability of the **spmon** command is available through the Perspective, the ″look and feel″ of the control is somewhat different. The SP Hardware Perspective offers a special online help facility to acclimate former **spmon** graphical interface users to the new controls. To access this help, click on the **Help** menu bar item in the SP Hardware Perspective, select the **Tasks...** option, and then select the **Transforming System Monitor Experience into Hardware Perspectives Skills** topic from the help menu.

Each Perspective provides its own unique capabilities. For the purposes of problem monitoring and determination, this manual recommends that the SP Event Perspective be used to monitor conditions of interest for the SP system. When the SP Event Perspective indicates that a hardware failure condition exists, the SP Hardware Perspective should be used to examine the current status of the system hardware and obtain more detailed information about the hardware problem.

Command line tools

PSSP provides command-oriented tools for system administration in addition to graphical tools for system administration and monitoring. These tools require no special workstation capability or high-speed connection, making them usable by almost any terminal type in any mode of access. Use these tools when examining system status through a modem connection or through a node’s S1 serial port. The tools discussed in this section are documented in greater detail in *PSSP: Command and Technical Reference*, *PSSP: Administration Guide*, and *AIX Version 4 Commands Reference*. These tools do not possess the same ease-of-use characteristics as their Perspectives based counterparts, although they do provide the same basic function.

Several commands are useful for monitoring the system status and detecting problem situations:

- spmon
- hmmon
- df
- dsh
- lsps
- lssrc

The **spmon** and **dsh** commands require the user to have specific authorizations. To learn how a user can acquire these authorizations, see ″Using the SP System Monitor″ chapter of *PSSP: Administration Guide*.

The **spmon** command permits the user to control and monitor SP hardware resources through a command-line interface without requiring a graphics-capable terminal or high-speed connection. The **spmon** command does not provide the capability to examine software status (such as paging space, file system space, or software subsystem activity). The **spmon** command provides access to more node-specific information than the **hmmon** command, which is introduced next. The **spmon** command provides a predefined system query to check the most basic problem conditions within the SP system.

The **hmmon** command provides hardware monitoring functions similar to the **spmon** command, and gives you access to more SP hardware information for frames and switches than the **spmon** command does. The **hmmon** command provides the capability to monitor frame and switch status as well as node status. The **hmmon** command is intended as a general-purpose SP hardware monitor. Although it has access to more SP information than the **spmon** command, it does not have access to some of the node-specific information that the **spmon** command does. The **hmmon** command does not provide a predefined system query, which the **spmon** command does.

The **df** command is an AIX command that examines the current status of file systems, such as current file system size and current available space within these file systems. While this command is designed to examine the AIX system on which it is issued, it can be invoked remotely with the **dsh** command to acquire this information for all nodes. Three file systems are of particular importance for all SP nodes:

- **/spdata**

  This directory contains configuration information for PSSP software and also contains copies of information from the SDR. By default, this directory resides in the **/** (or **root**) file system. Insufficient space in this file system can result in failures in PSSP software, especially those dependent on the SDR for proper operation. As a rule of thumb, ensure that this file system has at least 5% of its capacity available at any time.

  One method for avoiding space problems for the **/spdata** directory is to create a separate file system for this directory. Follow the instructions in *PSSP: Installation and Migration Guide* to create a separate volume group for this file system. Use the same rule of thumb for spotting potential trouble with this file system.

- **/var**

  This file system contains AIX system logs, such as the error log and user access logs. It also contains logs maintained by PSSP software for serviceability purposes. Some of these logs are never cleared except by explicit system administrator actions. If left unattended, they can grow to consume all available space.

  As a rule of thumb, ensure that 10MB of space is available within this file system at all times. If the file system reaches this threshold, consider either extending the file system's capacity with the **chfs** command, or examine the file system to determine where the space is being consumed and remove unneeded files.

  If **/var** is continually reaching the suggested threshold, this condition may indicate a chronic problem with some PSSP software or with specific hardware devices. Examine the logs listed in "Chapter 4. Error logging overview" on page 67 to determine if any show increased or extended activity, and perform any associated problem determination procedures if necessary.

- **/tmp**

  This file system is used by various user level applications, software products, and PSSP programs for temporary storage. Some legacy PSSP applications use this file system to store trace logs used for serviceability purposes. Some applications may inadvertently leave temporary files in the **/tmp** file system, or these applications may terminate before removing these files.

Insufficient space in **/tmp** can cause PSSP software to fail. As a rule of thumb, ensure that at least 8MB of space is available in this file system at any time. Eight MB is the amount of space a **snap** command will require if the system has to produce a dump to be sent to the IBM Support Center.

These space capacities can be verified using the **dsh** command to invoke the **df** command on all nodes in the SP system.

The **lsps** command provides an instant assessment of the currently available paging space for an AIX system. As with the **df** command, the **lsps** command provides information for the AIX system on which it runs. Using the **lsps** command with the **dsh** command or via a remote command, you can obtain the assessment for all nodes in the SP system.

Paging space availability by itself does not necessarily indicate a problem. Having only ten percent of 2 gigabytes of paging space available is not as significant a condition as having only ten percent of 100MB available. Also, one system 's critical situation may be a tolerable situation for another system. Because of this discrepancy, this manual will not suggest a default figure for a critical paging space situation. Use your knowledge of the system setup, system workload, and any past paging space problems to determine this value.

The **lssrc** command provides information for software services currently installed on an AIX system. Using **lssrc**, you can determine if a software service is active or inactive. Use this command in cases where a software service does not appear to be responding to requests for service on a specific node. To check software service status on multiple nodes, use this command through the **dsh** command.

The **dsh** command permits the user to issue a command on a remote node and to view the results on the local node. Using **dsh**, you can issue the commands listed previously on any SP node from a single location. This removes the need to login to each node individually. A user must have specific authorization to use the **dsh** command. To learn how a user can acquire this authorization, see ″Using the SP System Monitor″ chapter of *PSSP: Administration Guide*.

The following scenarios demonstrate how these tools are used to query and monitor the status of the SP system.

Assessing the current status of the SP system

This task is accomplished through the following series of steps:

1. **Preparing to Perform the System Check.** Prepare for this task by retrieving the log of the SP system structure. This log is discussed in “Create a log of your SP structure and setup” on page 7. This information is required to use the **hmmon** command effectively. The **hmmon** command obtains hardware information about nodes and switch devices using the frame number and slot number of the device, not the network name or IP address assigned to the device.

   This check should be performed by users authorized to invoke the **spmon** and **dsh** commands. To learn how a user can acquire this authorization, see the ″Using the SP System Monitor″ chapter in *PSSP: Administration Guide*.
2. **Perform a Preliminary Check of the SP System.** To perform a basic diagnostic check of the entire SP system, issue the following command from the control workstation:

   ```
   /usr/lpp/ssp/bin/spmon -G -d | more
   ```

This test verifies several items in the monitor program itself to make sure that it is running. Once the monitor verification completes, the **spmon** command checks the status of the SP frames and obtains information about the SP nodes. The **spmon** command performs these tests in a dependent order, so that if one of the early checks fails, subsequent checks are not performed. For example, if a frame cannot be queried, the frame and the nodes within that frame are not checked.

Example output from the **spmon -G -d** command:

```
1.  Checking server process
Process 10512 has accumulated 192 minutes and 53 seconds.
Check ok

2.  Opening connection to server
Connection opened
Check ok

3.  Querying frames (s)
1 frames (s)
Check ok

4.  Checking frames

       Controller  Slot 17  Switch    Switch    Power supplies
Frame   Responds   Switch   Power    Clocking   A   B   C   D
--------------------------------------------------------------
   1       yes       yes      on        0       on  on  on  on

5.  Checking nodes
------------------------------- Frame 1 ----------------------
Frame  Node  Node         Host/Switch   Key   Env   Front Panel   LCD/LED is
Slot  Number Type  Power   Responds    Switch Fail    LCD/LED     Flashing
  1     1    wide   on      yes  yes    normal  no  LEDs are blank   no
  3     3    thin   on      yes  yes    normal  no  LEDs are blank   no
  4     4    thin   on      yes  yes    normal  no  LEDs are blank   no
  5     5    thin   on      yes  yes    normal  no  LEDs are blank   no
  6     6    thin   on      yes  yes    normal  no  LEDs are blank   no
  7     7    wide   on      yes  yes    normal  no  LEDs are blank   no
  9     9    wide   on      yes  yes    normal  no  LEDs are blank   no
  11    11   wide   on      yes  yes    normal  no  LEDs are blank   no
  13    13   wide   on      yes  yes    N/A     no  LCDs are blank   no
```

Note that these tests are numbered. This makes it easy to detect if a test was omitted. The results of this command indicate potential problems if any of these conditions exist:

- The command does not run.
- The command does not perform all five verification checks.
- The fourth test indicates that the frame's controller is not responding, the switch power is not on, or any of the power supplies are listed as off.
- The fifth test indicates any abnormal conditions: a node's power is off, the **host responds** does not read **yes**, an environment failure is indicated, or the LCD or LED of the node is not blank (but not flashing).
- The fifth test indicates that the node's LCDs or LEDs are flashing.

  This indicates that a system dump was attempted.
- The fifth test indicates that the node is not responding to the switch device.

3. **Obtaining More Information**. If the **spmon** command mentioned previously indicates a potential problem situation, obtain more information in order to resolve the problem.

- If either the first or second test of **spmon -G -d** failed, consult “Chapter 21. Diagnosing System Monitor problems” on page 317.
- If the third or fourth test failed, use the **hmmon** command to detect if there are problems with the frame itself. Issue this command to obtain this information:

  ```
  hmmon -G -q -s -v frPowerOff*,controllerResponds,\
  controllerIDMismatch,nodefail* range_of_frame_nums:0
  ```

  Output is similar to:

  ```
  1 0 nodefail1            FALSE   0x8802  node 01 I2C not responding
  1 0 nodefail2            TRUE    0x8803  node 02 I2C not responding
  1 0 nodefail3            FALSE   0x8804  node 03 I2C not responding
  1 0 nodefail4            TRUE    0x8805  node 04 I2C not responding
  1 0 nodefail5            FALSE   0x8806  node 05 I2C not responding
  1 0 nodefail6            FALSE   0x8807  node 06 I2C not responding
  1 0 nodefail7            FALSE   0x8808  node 07 I2C not responding
  1 0 nodefail8            FALSE   0x8809  node 08 I2C not responding
  1 0 nodefail9            FALSE   0x880a  node 09 I2C not responding
  1 0 nodefail10           FALSE   0x880b  node 10 I2C not responding
  1 0 nodefail11           FALSE   0x880c  node 11 I2C not responding
  1 0 nodefail12           FALSE   0x880d  node 12 I2C not responding
  1 0 nodefail13           FALSE   0x880e  node 13 I2C not responding
  1 0 nodefail14           TRUE    0x880f  node 14 I2C not responding
  1 0 nodefail15           FALSE   0x8810  node 15 I2C not responding
  1 0 nodefail16           TRUE    0x8811  node 16 I2C not responding
  1 0 nodefail17           FALSE   0x8812  switch  I2C not responding
  1 0 frPowerOff           FALSE   0x8846  SEPBU   frame power off
  1 0 controllerIDMismatch FALSE   0x8871  frame ID mismatch
  1 0 controllerResponds   TRUE    0x88a8  frame responding to polls
  ```

  This command tests if any of the frame’s power supplies are off, if the frame controller is experiencing problems, or any of the node slot connections are bad. Keep in mind the warning made earlier, since wide and high nodes occupy more than one node slot in a frame, node failures will be detected for node slots that cannot be used because a wide or high node occupies that space.

  Such a situation is demonstrated in the example output listed previously. In this example, the nodes occupying slots 1 and 3 are wide nodes, as are the nodes occupying slots 13 and 15. Node slots 2, 4, 14, and 16 are therefore unusable, but the **hmmon** command indicates that nodes in these unavailable slots have failed. The log of the SP structure and setup is needed to understand which slots are ″supposed″ to indicate node failures, and which slots are not.

  Check for any of these conditions in the **hmmon** command output:

  – controllerResponds reads FALSE

  – controllerIDMismatch reads TRUE

  – nodefail17 reads TRUE (Indicating a failure in the SP switch)

  – Other nodefails show TRUE, and these node failure cannot be attributed to a wide or high node occupying that slot

  If a controller ID mismatch is shown, consult the Managing a HACWS Configuration chapter in *PSSP: Administration Guide*. For controller responsiveness problems, perform hardware diagnostics on the frame controller. For nodefail17 failures, perform hardware diagnostics on the switch device. For other node failures, perform hardware diagnostics on the node occupying that slot.

- If the fourth test of the **spmon -G -d** command indicated that the switch power was off, issue the following **hmmon** command to determine if this was caused by a hardware condition:

  ```
  hmmon -G -Q -s -v nodePower,powerLED,envLED,shutdownTemp frame_num:17
  ```

  Example output of the **hmmon** command, showing switch information for frame 1

  ```
  1  17  powerLED                  1  0x8c47  node/switch LED 1 (green)
  1  17  envLED                    0  0x8c48  node/switch LED 2 (yellow)
  1  17  nodePower            TRUE     0x8c4a  DC-DC power on
  1  17  shutdownTemp         FALSE    0x8c59  temperature shutdown
  ```

  This **hmmon** command will indicate if the switch has power, if power is available for the switch, if the switch's power was shut down automatically, and if the switch power was shut down due to high temperature. If the switch cannot obtain power, verify that the switch is correctly cabled to its power source. For other conditions, perform hardware diagnostics on the switch device.

- If the fifth test of the **spmon -G -d** command indicates that a node does not have power, and the node's power was not shut off manually, issue the following **hmmon** command to determine if the power was disabled because of a hardware condition:

  ```
  hmmon -G -Q -s -v nodePower,powerLED,envLED frame_num:node_num
  ```

  Example output of the hmmon command for a single node in a single frame:

  ```
  1   1  nodePower            TRUE     0x904a  DC-DC power on
  1   1  powerLED                  1  0x9047  node/switch LED 1 (green)
  1   1  envLED                    0  0x9048  node/switch LED 2 (yellow)
  ```

  This **hmmon** command will indicate if the node has power, if power is available for the node, and if the node's power was shut down automatically. If the node cannot get power, verify that the node is correctly cabled to its power source. For other conditions, perform hardware diagnostics on the node.

- If the fifth test of the **spmon -G d** command indicates that a node's LED/LCD display is not blank, a hardware or operating system error has occurred. The LED/LCD code contains important failure information. When this condition exists, examine the node's LED/LCD value and record the value displayed. Use the following command to examine this value:

  ```
  /usr/lpp/ssp/bin/spmon -L framenumber/nodenumber
  ```

To determine the explanation and action for the error, look up this code in “Chapter 32. SP-specific LED/LCD values” on page 535. If a three-digit LED/LCD code is not listed in this table, consult “Other LED/LCD codes” on page 539.

- If the fifth test of the **spmon -G -d** command indicated that the node’s LED/LCD value was flashing, and the **spmon -L** command in the previous bullet indicates that the LED/LCD value is **888**, a system dump was initiated on this node. The flashing **888** LED/LCD value indicates that a series of values are stored in the LED/LCD display.

  Step through this list of codes and record each value shown using the following sequence of steps:

  a. Issue the command

     ```
     /usr/lpp/ssp/bin/spmon -reset -t framenumber/nodenumber
     ```

     to step to the next stored LED/LCD value.

  b. Issue the command

     ```
     /usr/lpp/ssp/bin/spmon -L framenumber/nodenumber
     ```

     to retrieve the new LED/LCD value.

  c. Record this LED/LCD value

  Repeat these steps until the **spmon -L** command displays a value of **888** again. Retain this list of codes; they will be required by the IBM Support Center. To determine the explanation and action for these error codes, look up the codes in “Chapter 32. SP-specific LED/LCD values” on page 535. If a three-digit LED/LCD code is not listed in this table consult “Other LED/LCD codes” on page 539. Finally, save and verify the system dump, following the instructions provided in “Chapter 5. Producing a system dump” on page 81.

4. **Checking Basic Software Information.** Once hardware failures have been eliminated, it is time to perform some basic software verifications for the SP system. These checks will use the **dsh** command to invoke AIX commands on multiple nodes in parallel. To verify this, issue the following command from the control workstation:

   ```
   dsh -a -f32 hostname
   ```

   Example output of the **dsh -a -f32 hostname** command on a small SP system configuration:

   ```
   k21n01.ppd.pok.ibm.com: k21n01.ppd.pok.ibm.com
   k21n03.ppd.pok.ibm.com: k21n03.ppd.pok.ibm.com
   k21n04.ppd.pok.ibm.com: k21n04.ppd.pok.ibm.com
   k21n05.ppd.pok.ibm.com: k21n05.ppd.pok.ibm.com
   k21n06.ppd.pok.ibm.com: k21n06.ppd.pok.ibm.com
   k21n07.ppd.pok.ibm.com: k21n07.ppd.pok.ibm.com
   k21n09.ppd.pok.ibm.com: k21n09.ppd.pok.ibm.com
   k21n11.ppd.pok.ibm.com: k21n11.ppd.pok.ibm.com
   k21n13.ppd.pok.ibm.com: k21n13.ppd.pok.ibm.com
   ```

   This test will verify that the **dsh** command can reach the nodes within the SP system. Only nodes that were previously detected as being offline in the earlier tests should fail to respond to this command. If any other nodes within the SP system fail to respond, check for problems by referring to “Chapter 20. Diagnosing remote command problems on the SP System” on page 299.

   - Check the paging space that is in use on all nodes by using the **lsps** command on the control workstation:

```
dsh -av -f32 lsps -s | more
```

Example output of the **dsh -av -f32 lsps -s** command on a small SP system configuration:

```
k21n01.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n01.ppd.pok.ibm.com:       768MB               8%
k21n03.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n03.ppd.pok.ibm.com:       768MB              17%
k21n04.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n04.ppd.pok.ibm.com:       768MB               8%
k21n05.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n05.ppd.pok.ibm.com:       768MB              13%
k21n06.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n06.ppd.pok.ibm.com:       768MB              12%
k21n07.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n07.ppd.pok.ibm.com:       768MB              11%
k21n09.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n09.ppd.pok.ibm.com:       768MB               9%
k21n11.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n11.ppd.pok.ibm.com:       768MB               9%
k21n13.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n13.ppd.pok.ibm.com:       768MB              15%
```

Lack of available paging space can lead to thrashing conditions on a node. If these nodes are running parallel applications, the entire application will be slowed to the rate of the slowest responding node. The extent to which low paging space and thrashing can be tolerated differs from one customer environment to the next. As a general rule of thumb, investigate any nodes indicating that 80% or more of its paging space is currently in use.

- Check for file systems that are close to their capacity, concentrating on the file systems mentioned earlier in this section, by issuing the **dsh** command from the control workstation, to invoke the **df** command:

```
dsh -av -f32 df /spdata /var /tmp | more
```

Example output from the **dsh -av -f32 df /spdata /var /tmp** command on a small SP system configuration:

```
k21n01: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n01: /dev/hd4           32768       432   99%     1403    18%  /
k21n01: /dev/hd9var       147456     45480   70%      610     4%  /var
k21n01: /dev/hd3           98304     38632   61%       85     1%  /tmp
k21n03: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n03: /dev/hd4           32768     16960   49%     1431    18%  /
k21n03: /dev/hd9var       704512     99968   86%      595     1%  /var
k21n03: /dev/hd3           98304     50424   49%      278     3%  /tmp
k21n04: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n04: /dev/hd4           32768     16584   50%     1512    19%  /
k21n04: /dev/hd9var       147456    107312   28%      644     4%  /var
k21n04: /dev/hd3           98304     91232    8%       74     1%  /tmp
```

An obvious warning sign is if any of these file systems should appear to be more than 90% utilized. If any file systems appear over 90% utilized, examine the file systems for large files that can be removed or compressed, or consider extending the file system size. Attempt to keep 10MB available in the **/var** file system and 8MB available in the **/tmp** file system, to ensure that PSSP software and service software function correctly.

Keeping informed of status changes

Note: In order to successfully issue these commands, you must have ″monitor″ permission for kerberos (compat mode) or DCE mode, depending on the authentication method in use.

The previous discussion centered on obtaining the current status of SP system hardware and software. Such efforts are necessary if a problem is suspected and being actively investigated, but repeatedly issuing these commands periodically to examine the current status of the SP system can become tedious. To make the task of monitoring system status easier, PSSP provides monitoring capabilities within the **hmmon** and **spmon** commands as well. This avoids the necessity of reissuing the previously discussed commands over and over again to keep informed of the system status. This section describes some of the more common monitor commands.

To set up a monitor to check for frame hardware failures, issue the following background command:

```
hmmon -G -q -s -v frPowerOff*,controllerResponds,controllerIDMismatch,\
nodefail* range_of_frame_nums:0 &
```

Example initial output from the **hmmon** command:

```
1   0  nodefail1            FALSE    0x8802  node 01 I2C not responding
1   0  nodefail2            TRUE     0x8803  node 02 I2C not responding
1   0  nodefail3            FALSE    0x8804  node 03 I2C not responding
1   0  nodefail4            FALSE    0x8805  node 04 I2C not responding
1   0  nodefail5            FALSE    0x8806  node 05 I2C not responding
1   0  nodefail6            FALSE    0x8807  node 06 I2C not responding
1   0  nodefail7            FALSE    0x8808  node 07 I2C not responding
1   0  nodefail8            TRUE     0x8809  node 08 I2C not responding
1   0  nodefail9            FALSE    0x880a  node 09 I2C not responding
1   0  nodefail10           TRUE     0x880b  node 10 I2C not responding
1   0  nodefail11           FALSE    0x880c  node 11 I2C not responding
1   0  nodefail12           TRUE     0x880d  node 12 I2C not responding
1   0  nodefail13           FALSE    0x880e  node 13 I2C not responding
1   0  nodefail14           TRUE     0x880f  node 14 I2C not responding
1   0  nodefail15           TRUE     0x8810  node 15 I2C not responding
1   0  nodefail16           TRUE     0x8811  node 16 I2C not responding
1   0  nodefail17           FALSE    0x8812  switch I2C not responding
1   0  frPowerOff           FALSE    0x8846  SEPBU frame power off
1   0  controllerIDMismatch FALSE    0x8871  frame ID mismatch
1   0  controllerResponds   TRUE     0x88a8  frame responding to polls
```

This command is similar to the one presented previously, except that this version continually monitors the frame condition and generates a message to the terminal if any of the status should change. To stop monitoring this information, terminate the background process.

To set up a monitor to check for SP switch hardware status changes, issue the following background command:

```
hmmon -G -q -s -v nodePower,powerLED,envLED,\
shutdownTemp range_of_frame_nums:17 &
```

Example initial output from the **hmmon** command:

```
1  17  powerLED                  1  0x8c47  node/switch LED 1 (green)
1  17  envLED                    0  0x8c48  node/switch LED 2 (yellow)
1  17  nodePower            TRUE     0x8c4a  DC-DC power on
1  17  shutdownTemp         FALSE    0x8c59  temperature shutdown
```

This command is similar to one presented previously, except this version continually monitors the frame condition and generates a message to the terminal if any of the status should change. To stop monitoring this information, terminate the background process.

To set up a monitor to check for changes in a node’s LCD or LED status, issue the following background command:

```
hmmon -G -q -s -v LED7Seg* range_of_frame_nums:1-16 &
```

Example initial output from the **hmmon** command:

```
1   1  LED7SegA                255  0x909f  7 segment LED A
1   1  LED7SegB                255  0x90a0  7 segment LED B
1   1  LED7SegC                255  0x90a1  7 segment LED C
1   3  LED7SegA                255  0x949f  7 segment LED A
1   3  LED7SegB                255  0x94a0  7 segment LED B
1   3  LED7SegC                255  0x94a1  7 segment LED C
1   4  LED7SegA                255  0x949f  7 segment LED A
1   4  LED7SegB                255  0x94a0  7 segment LED B
1   4  LED7SegC                255  0x94a1  7 segment LED C
1   5  LED7SegA                255  0x949f  7 segment LED A
1   5  LED7SegB                255  0x94a0  7 segment LED B
1   5  LED7SegC                255  0x94a1  7 segment LED C
1   6  LED7SegA                255  0x949f  7 segment LED A
1   6  LED7SegB                255  0x94a0  7 segment LED B
1   6  LED7SegC                255  0x94a1  7 segment LED C
1   7  LED7SegA                255  0x909f  7 segment LED A
1   7  LED7SegB                255  0x90a0  7 segment LED B
1   7  LED7SegC                255  0x90a1  7 segment LED C
1   9  LED7SegA                255  0x909f  7 segment LED A
1   9  LED7SegB                255  0x90a0  7 segment LED B
1   9  LED7SegC                255  0x90a1  7 segment LED C
1  11  LED7SegA                255  0x909f  7 segment LED A
1  11  LED7SegB                255  0x90a0  7 segment LED B
1  11  LED7SegC                255  0x90a1  7 segment LED C
```

This command shows the initial status of these resources, and displays any status changes in these resources when they occur. All values should display a value of **255**, indicating that the associated readout element is blank. If any nodes indicate that a segment is not blank, issue the **spmon -L** command mentioned on 25 to obtain the current LCD or LED readout of the node.

To set up a monitor to check for nodes suddenly losing contact with the SP Switch, issue the following command:

```
spmon -q -M -l -t frame*/node*/switchResponds/value
```

Example initial output from the **spmon** command:

```
/SP/frame/frame1/node1/switchResponds/value/1
/SP/frame/frame1/node3/switchResponds/value/1
/SP/frame/frame1/node4/switchResponds/value/1
/SP/frame/frame1/node5/switchResponds/value/1
/SP/frame/frame1/node6/switchResponds/value/1
/SP/frame/frame1/node7/switchResponds/value/1
/SP/frame/frame1/node9/switchResponds/value/0
/SP/frame/frame1/node11/switchResponds/value/0
/SP/frame/frame1/node13/switchResponds/value/1
```

The **spmon** command also displays the current status, and a message to the terminal if any of these values change. All values should be 1. A value of 0 indicates that the node is not responding to the SP Switch. Note that this is the case with two of the nodes in this example, and these nodes should be investigated.

Other conditions can also be monitored using the **hmmon** and **spmon** commands; these suggestions offer the most basic of tests. To learn what other conditions can be monitored with these commands, and to tailor these commands to best suit your needs, refer to the **hmmon** and **spmon** sections of *PSSP: Command and Technical Reference*.

All commands can be issued from the same terminal session, but this can lead to confusing output when conditions change, or initial values can scroll off the terminal screen. To keep the monitoring manageable, consider issuing these commands from separate terminals, or from separate terminal windows from a XWindows capable terminal. Issue one monitoring command per terminal or terminal window. This will associate a terminal with each condition being monitored, and simplify the understanding of the monitor output.

Asynchronous (batch) notification methods

As system administrator, you cannot always devote your entire attention to monitoring the current status of a system, trying to detect problem conditions before they occur. For even moderately sized SP system configurations, this task can be time consuming and tedious. Other tasks require your attention, so actively monitoring the SP system for potential problem indications cannot become a task that consumes all your time and effort.

Fortunately, PSSP provides tools to monitor system status and conditions on your behalf, as well as the tools discussed previously to assess the current status of the system. Using these tools, you can indicate conditions of particular interest, request asynchronous notification of these events, and cause actions to be initiated when these conditions occur. In essence, the SP system monitors itself, takes action itself, and notifies you of the condition after it occurs. These monitoring tools can be used when you are not immediately available, such as during off-peak hours, or can be used to remove most of your monitoring burden.

Two sets of monitoring tools are available. As with the runtime notification tools mentioned previously, the choice of tool depends on the capabilities of your workstation and your preferences. PSSP provides **graphical monitoring tools** for use on the control workstation or a network attached terminals, and also provides **command-line monitoring tools** for those situations where only modem access or **s1term** access is available.

Graphics tools - SP Event Perspective

This tool was introduced in “The SP Event Perspective” on page 16. To use the SP Event Perspective effectively, you must understand certain terminology. These terms were introduced in “The SP Event Perspective” on page 16. Please refer to that section to become familiar with these terms.

Individual SP Perspectives require that certain subsystems be operating and that the user is authorized to communicate with them. Such subsystems include the System Monitor, Event Management, System Data Repository, and Problem Management. For the authorization required for each SP Perspective, see the discussion on using SP Perspectives in *PSSP: Administration Guide*.

Users of the SP Event Perspective can set up the Perspective to send a notification to the system administrator when conditions of interest exist on the system (or, to use the SP Perspectives terminology, when an event occurs). This is done by associating an action with the event in the event definition. This action can be any command or script that can be run from the AIX command line, including the creation of an electronic mail message, starting a process that can place a telephone call to the system administrator’s pager, send a message to a specific user at a specific terminal, or any other notification command. The action invoked when the event occurs is called ’the command’.

When creating or modifying the event definition, the user can specify a command to be issued when the condition exists. The following AIX command can be used to have the SP Event Perspective send an electronic mail message to a specific user when the event occurs:

```
/usr/bin/echo \
"event_condition has occurred `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_condition Notification" \
username@address
```

To understand the mechanics of setting up a command within an event definition, consult the Perspectives online help. Click on the **Help** menu button on the SP Event Perspective display, and select the **Tasks...** option. Assistance on specifying event definitions is available through the **Working with Event Definitions** topic.

Whenever the user registers for the event definition through the SP Event Perspective, ’the command’ will be issued if the condition exists in the system. The SP Event Perspective does not have to be currently active in order for ’the command’ to be issued; provided the user has registered for the event definition, the system will continue to monitor itself for this condition, and issue ’the command’ if the condition exists. In other words, the user can use the SP Event Perspective to set up event definitions, register for events, then shut down the SP Event Perspective, and the system will still issue the notification command when the condition occurs. The SP system continually monitors itself for the condition and issues the notification command until the user cancels the event registration.

’The command’ associated with an event definition can also be used to automate a response to this condition, instead of merely notifying the system administrator or another user of the condition. This topic will be discussed in “Automating your response to problems” on page 38.

Command line tools - Problem Management

Note: Problem management does not run in Restricted Root Access (RRA) mode. Problem management does not run when the system administrator has set Authorization for AIX Remote Commands to ″none″. See “Special troubleshooting considerations” on page 4 for more details.

Problem Management is a software subsystem used in command line and script oriented environments to specify conditions that should be monitored by the SP system, and to specify actions to take when these conditions exist on the SP system. This is the same software subsystem invoked internally by the SP Event Perspective discussed in the previous section.

For users attempting to connect to the SP nodes through low-speed modems or using non-graphical terminals, Problem Management provides a command line interface that can be used in place of the SP Event Perspective. As with other command line oriented tools, the Problem Management command line interface is not as intuitive to use or designed for ease of use as is its graphical counterpart.

1. **Prepare to monitor the system.** To become familiar with Problem Management, especially regarding the security requirements, see the chapter on Problem Management in *PSSP: Administration Guide*.
2. **Understand what you want to monitor.** Problem Management expects the user to know the conditions that are to be monitored. Unlike the SP Event Perspective, Problem Management does not provide an interactive method to query for the list of available conditions and a means to select from these conditions. The user must identify the conditions to be monitored, and provide them as a list to Problem Management. These conditions are identified by naming the associated resource variables, the internal mechanism that contains the current status of the associated resource.

   PSSP provides over 300 default resource variables. “Chapter 3. Conditions to monitor on the SP system” on page 43 provides a suggested list of resource variables to monitor, but specific SP systems may require that additional resources also be monitored. The full list of resource variables is maintained by the Event Management subsystem, and the list can be retrieved using the **haemqvar** command. This command generates large amounts of information, so it is best to start with a brief report from this command to identify those resources to be monitored:

   ```
   haemqvar -d | more
   ```

   This command provides the names of the available resource variables, and a short description of each resource variable:

```
IBM.PSSP.aixos.Proc.swpque   Average count of processes
                             waiting to be paged in.
IBM.PSSP.aixos.Proc.runque   Average count of processes that are
                             waiting for the cpu.
IBM.PSSP.aixos.pagsp.size   Size of paging space (4K pages).
IBM.PSSP.aixos.pagsp.%free   Free portion of this paging space (percent).
IBM.PSSP.aixos.PagSp.totalsize Total active paging space size (4K pages).
IBM.PSSP.aixos.PagSp.totalfree Total free disk paging space (4K pages).
IBM.PSSP.aixos.PagSp.%totalused Total used disk paging space (percent).
IBM.PSSP.aixos.PagSp.%totalfree   Total free disk space (percent).
IBM.PSSP.aixos.Mem.Virt.pgspgout 4K pages written to paging space by VMM.
IBM.PSSP.aixos.Mem.Virt.pgspgin 4K pages read from paging space by VMM.
IBM.PSSP.aixos.Mem.Virt.pagexct   Total page faults.
IBM.PSSP.aixos.Mem.Virt.pageout   4K pages written by VMM.
IBM.PSSP.aixos.Mem.Virt.pagein    4K pages read by VMM.
IBM.PSSP.aixos.Mem.Real.size   Size of physical memory (4K pages).
IBM.PSSP.aixos.Mem.Real.numfrb   Number of pages on free list.
IBM.PSSP.aixos.Mem.Real.%pinned   Percent memory which is pinned.
IBM.PSSP.aixos.Mem.Real.%free   Percent memory which is free.
                          :
                          :
```

The **haemqvar** command lists resources available only on the node where the command is issued. Keep in mind that resources may exist on some nodes and not on others. *PSSP: Command and Technical Reference* gives a detailed description of the **haemqvar** command, and how it can be used to locate any resource variable available within the SP system.

Once the resource variable to be monitored has been identified, the value type and locator for each resource variable must be identified. The locator informs Problem Management where to monitor the resource. For example, Problem Management needs to know the name of the file system and the node on which a file system resides, if it is to monitor that file system for the amount of space it has available. This information is conveyed to Problem Management through the locator value. To obtain the locator for a resource variable, issue the following **haemqvar** command:

```
haemqvar "" resource_variable_name "*"
```

This command provides details on the resource variable, including the locator keyword needed for Problem Management. The additional information can be helpful in constructing an effective Problem Management definition for the condition.

For example, to obtain the locator field for the **IBM.PSSP.CSS.ipackets_drop** variable, and to understand more about the variable, issue:

```
haemqvar "" IBM.PSSP.CSS.ipackets_drop "*"
```

which produces output similar to:

```
Variable Name:  IBM.PSSP.CSS.ipackets_drop
Value Type:     Quantity
Data Type:      long
Initial Value:  0
Class:          IBM.PSSP.CSS
Locator:        NodeNum
Variable Description:
    Number of packets not passed up.

    A message received by a node from the switch of the
    Communication SubSystem (CSS) is comprised of packets.
    IBM.PSSP.CSS.ipackets_drop is the count of the number of good
    incoming packets at the subject node's CSS interface which
    were dropped by the adapter microcode, since that interface
    was last initialized.

    If a node has too heavy a general workload, it may not service its
    CSS interface often enough, causing its messages to linger in the
    switch network.  If this is allowed to continue, the switch can
    become backed up causing other nodes to encounter poor switch
    performance; in fact, this condition can cause the entire
    switch to clog. Instead, the adapter microcode drops any "excess"
    packet -- a reliable protocol will eventually retry the message.

    For performance reasons, counts such as this are only updated
    approximately once every 2 minutes.

    This variable is supplied by the "IBM.PSSP.harmld" resource monitor.

    Example expression:

    To be notified when IBM.PSSP.CSS.ipackets_drop exceeds 100 on any
    node, register for the following event:

          Resource variable:  IBM.PSSP.CSS.ipackets_drop
          Resource ID:        NodeNum=*
          Expression:         X>100
          Re-arm expression:  X<100

    Resource ID wildcarding:

    The resource variable's resource ID is used to specify the number
    of the node (NodeNum) to be monitored. The NodeNum resource ID
    element value may be wildcarded in order to apply a query or event
    registration to all nodes in the domain.

    Related Resource Variables:

      IBM.PSSP.CSS.ibadpackets    Number of bad packets received
                                  by the adapter.
      IBM.PSSP.CSS.ipackets_lsw   Packets received on interface
                                  (lsw bits 30-0).
      IBM.PSSP.CSS.ipackets_msw   Packets received on interface
                                  (msw bits 61-31).
Resource ID:    NodeNum=int
  NodeNum: The number of the node for which the information applies.
```

The **Locator**: field indicates the keyword to be used with Problem Management to identify where the resource should be monitored. Note that the **haemqvar** command offers advice on how to use the locator field in the output.

3. **Identify the conditions of interest.** Problem Management is informed of the conditions to be monitored through the **pmandef** command. One **pmandef** command is needed for each condition to be monitored. This command is used

to subscribe to the event, which is similar in concept to the SP Event Perspective's registration of an event definition. To create the subscription, the following information is needed:

- The resource variable name, obtained in Step 2 on page 32.
- The resource variable locator, also obtained in Step 2 on page 32.
- The event expression, which indicates the condition of interest in this resource variable
- The rearm expression, which indicates when the condition is no longer of interest
- An event handle, which is a symbolic name that the system administrator will use to refer to this definition.

The event expression indicates the value that the resource variable will have when notice should be given. This value is assigned by the system administrator. The rearm expression indicates the value of the resource variable that indicates that the condition of interest is no longer present. How these expression are coded depends on the value type of the resource variable. The event handle is a name assigned by the system administrator, which should be descriptive of the condition being monitored.

For example, consider a case where the system administrator is interested in paging space on any SP node. If paging space reaches 90% capacity, the system administrator considers the node to be ″thrashing″ and wants to be notified. The system administrator considers the node to be ″thrashing″ once this threshold is reached, even if a little paging space frees up. The system administrator does not consider the ″thrashing″ problem to be resolved unless 40% of the paging space becomes available again. Using this scenario and the **haemqvar** commands from Step 2 on page 32, the system administrator identifies these conditions of interest:

- The resource variable name is **IBM.PSSP.aixos.PagSp.%totalused**, which contains the percentage of used paging space on a node.
- The resource variable locator is **NodeNum**, meaning that a node number is needed to indicate where the resource is to be monitored. The system administrator wants to monitor the condition on all nodes, so the locator expression is **NodeNum=***.
- The event expression must indicate the 90% capacity condition, so the expression **X>90** is used.
- The rearm expression must indicate the condition that ″turns off″ the event, which is when 40% of paging space becomes available. The expression **X< 60** is used.
- The system administrator assigns the name **Node_Thrash_Monitor** to this event definition.

Identify these conditions for all resource variables to be monitored. "Chapter 3. Conditions to monitor on the SP system" on page 43 lists some basic resource variables to monitor and the associated event expressions and rearm expressions.

4. **Decide how to notify the system administrator.** Problem Management associates an action with the event definition. When the condition exists within the system (or, to use the correct terminology, when the event occurs), Problem Management performs the action associated with the event. This action can be any command or script that can be issued from the AIX command line, including the creation of an electronic mail message, starting a process that can place a

telephone call to the system administrator’s pager, send a message to a specific user at a specific terminal, or any other notification command. This action is termed ’the command’.

The user can specify a command to be issued when the condition exists. The following AIX command can be used to have Problem Management send an electronic mail message to a specific user when the event occurs:

```
/usr/bin/echo \
"event_handle has occurred `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_handle Notification" \
username@address
```

An action can also be run when the condition no longer exists: when the rearm expression has been met. This action, called the ’rearm command’, can inform the system administrator that the condition no longer exists, so that the system administrator knows that the condition no longer needs attention. For example:

```
/usr/bin/echo \
"event_handle condition ended `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_handle Condition Ended"\
username@address
```

5. **Create an event definition file.** For every resource variable to be monitored, one **pmandef** command must be issued. If more than a handful of resources variables are to be monitored, this can result in a lot of typing. For convenience, create a file containing the **pmandef** commands to define these events to Problem Management. This will simplify the procedure for instructing Problem Management of the resources to monitor, and makes it easier to reissue these same commands at a later time.

   The **pmandef** command informs Problem Management of the conditions to be monitored by subscribing for events. This concept is almost exactly the same as the SP Event Perspective’s concept of registering event definitions. To subscribe for events on the chosen resource variables, create a file to contain **pmandef** commands in the following format:

   ```
   pmandef -s event_handle \
   -e 'resource_variable_name:locator:event_expression' \
   -r "rearm_expression" \
   -c event_command \
   -C rearm_command \
   -n 0
   ```

   Substitute the following information for the keywords in the previous command format:

   - *event_handle* is the event handle assigned by the system administrator in Step 3 on page 34.
   - *resource_variable_name* is the name of the resource variable, obtained in Step 2 on page 32.
   - *locator* is the locator expression, indicating where the resource is to be monitored, determined in Steps 2 on page 32 and 3 on page 34.
   - *event_expression* indicates the value the resource variable will have when the condition of interest exists, determined in Step 3 on page 34.
   - *rearm_expression* indicates the ″shut off″ value the resource variable will have when the condition no longer exists, determined in Step 3 on page 34.
   - *event_command* indicates the notification command (’the command’) to use for informing the system administrator that the condition exists, created in Step 4 on page 35.

- *rearm_command* indicates the notification command (’rearm command″) to use for informing the system administrator that the condition no longer exists, created in Step 4 on page 35.

Continuing with the previous example, the **pmandef** command to subscribe for the node thrashing condition would be:

```
pmandef -s Node_Thrash_Monitor \
-e 'IBM.PSSP.aixos.PagSp.%totalused:Nodenum=*:X>90' \
-r "X<60" \
-c '/usr/bin/echo "Node_Thrash_Monitor Alert `/usr/bin/date` \
   - Location Info: $PMAN_IVECTOR" | /usr/bin/mail\
-s "Node_Thrash_Monitor Alert" root@adminnode.ibm.com'
-C '/usr/bin/echo "Node_Thrash_Monitor Cancellation `/usr/bin/date` \
   - Location Info: $PMAN_IVECTOR" | /usr/bin/mail
-s "Node_Thrash_Monitor Cancellation" root@adminnode.ibm.com' \
-n 0
```

One **pmandef** command is required for each condition being monitored. Save this file and note its name for future reference.

6. **Subscribe to the Events through Problem Management.** To record these event definitions to Problem Management, issue the **pmandef** commands recorded in the file created in Step 5 on page 36 by issuing the **ksh** *filename* command, where *filename* is the name of the file created in Step 5 on page 36. Immediately after issuing the **ksh** command, issue the following command:

   ```
   pmandef -d all
   ```

   Problem Management not only subscribes to events with the **pmandef -s** command, but it also begins monitoring the resources as well. The **pmandef -d all** command disables the monitoring of these resources.

7. **Begin monitoring.** Begin monitoring these resources when you are ready. To begin monitoring the events that were subscribed in Step 6, issue this command:

   ```
   pmandef -a all
   ```

   This instructs Problem Management to begin monitoring all the conditions that were defined in Step 6. Should any of these events occur, Problem Management will issue ’the command’ associated with the event, to inform the system administrator of the event.

8. **Tailor the monitoring.** At times, certain conditions should not be checked on certain nodes. For example, Problem Management may be monitoring the space available in the **/tmp** file system on all nodes, but the system administrator expects **/tmp** to exceed that limit on a specific node (for example: node 5 on a 32-node SP system) for a certain period of time. If the monitoring is not tailored or modified to compensate for this expected event, the system administrator will be notified that the event occurred just as if it were an unexpected event.

   The system administrator can modify the subscribed event to Problem Management. To do this, the system administrator needs to know the following:

   - The event handle used for the condition, assigned in Step 3 on page 34.
   - The new locator that excludes the location where the condition is not to be monitored.

   Modification of the subscription is done in two steps:

   a. The event subscription is disabled, using the **pmandef** command:

```
pmandef -d event_handle
```

This deactivates the monitoring for this condition.

b. A new event expression is created, using a locator that excludes the location where the condition is not to be monitored. Using the example of **/tmp** monitoring, where node 5 is not to be monitored, the event expression would appear as:

```
IBM.PSSP.aixos.FS/%totused:NodeNum=1-4,6-32;VG=rootvg;
LV=hd3:X>90
```

c. Issue the **pmandef -s** command, using the structure provided in Step 5 on page 36 and the new event expression.

9. **Stop monitoring.** To stop monitoring of all events previously enabled in Step 7 on page 37, issue the following command:

```
pmandef -d all
```

These steps provide an overview of how Problem Management can be used to monitor system events and notify the system administrator when events occur. This is not a complete tutorial on the use of Problem Management. For greater detail on the capabilities and uses of Problem Management, especially with regard to security, consult the Problem Management chapter in *PSSP: Administration Guide*, the **pmandef** command, and the **haemqvar** command in *PSSP: Command and Technical Reference*.

Automating your response to problems

Detecting potential problem conditions before they become critical situations is the best way to resolve SP system problems. The condition is brought to the attention of the system administrator, allowing the system administrator to respond before the condition impacts other hardware and software components. But what if the procedure for correcting the situation is always the same? What if the system administrator will always run the same set of commands to address the condition whenever the condition occurs? Is it really necessary to require system administrator intervention, when the system administrator is going to perform the same action in all cases? The answer is **no**.

PSSP provides the capability to set an automated response to a system condition. This capability is provided as part of the SP Event Perspective, described in "The SP Event Perspective" on page 16 and the Problem Management subsystem, described in "Command line tools - Problem Management" on page 32. Using these tools, the system administrator can assign a specific action to be run automatically by these tools when the system condition exists, and also when the condition ″goes away″ (when the rearm event occurs). In the previous discussions of this chapter, this action was kept rather simple: the action caused a notification to be sent to the system administrator. The associated action does not need to be this simple: the action can be any AIX command or script. When the event occurs, these tools will run the command or script in response to the event.

Important - WHEN actions are performed

A response to a particular system condition may not always be the same, despite initial appearances. For example, when a file system is close to reaching its capacity, the appropriate response in most cases is to increase the file system's capacity using the **chfs** command. However, disk space is not a limitless resource, and eventually all disk space will be consumed if this approach is used whenever the file system reaches its capacity. Although this is a convenient solution, it is not always the correct solution. The file system should be checked for large obsolete files that can be deleted, users that exceed their quotas, directories that can be mounted on other file systems to save space in this file system, and other solutions.

When a system administrator associates an action with an event through Problem Management or the SP Event Perspective, this action is performed **each time the event occurs** (and Problem Management or Perspectives is monitoring the condition). Neither Problem Management nor the SP Event Perspective can decide that the action should not be performed for this specific occurrence of the event, but rather the system administrator needs to do some more analysis. This decision making process has to either be incorporated into the AIX command or script that will run in response to the event, or it has to be left to the system administrator's discretion.

Two strategies are offered:

- Specify an action for the event that performs two actions: the action notifies the system administrator of the event and the action also issues a command to address the event. This strategy allows the system to respond automatically to the condition and attempt to resolve the condition before it becomes a critical situation. The action also alerts the system administrator. The system administrator can then assess if the action should have been applied in this case.

  If the action was appropriate, the system administrator does not need to take action. However, if recent history indicates that the event has been recurring at an unusual rate, or history indicates that the action really does not resolve the condition and the event continues to occur, the system administrator still receives the notification and can respond to the condition.
- Build a response command or script that incorporates a decision making process within it. This response command can attempt to determine if a particular action is appropriate for the condition based on other information, such as other events or the recent history of actions taken in response to this event.

The second alternative can involve complicated logic, making it more difficult to implement. For this reason, the first strategy is recommended.

Important - WHERE actions are performed

Actions associated with events are performed **on the node that requested that the event be monitored**, which is **not necessarily the node where the condition exists**. For example, if a system administrator used Problem Management from the control workstation to monitor conditions on all nodes in the SP system and that condition suddenly exists on Node 42, the action is run on the control workstation, not Node 42. If the system administrator had associated a **chfs** command with the event, the **chfs** command would run on the control workstation and modify the control workstation's file system space, not the file system on Node 42.

When associating actions with events, keep in mind that the action will be performed by default on the node that asked for the event to be monitored. If action is to be taken on the node where the condition actually exists, the command invoked must determine where the condition occurred from the event information, and attempt to invoke remote processes on that node.

Both the SP Event Perspective and Problem Management make available several environmental values to the commands associated with the event. These variables are described in the chapter on using the Problem Management subsystem in *PSSP: Administration Guide*. Any command or script invoked by Problem Management or the SP Event Perspective has access to these variables. The variable **PMAN_IVECTOR** indicates where the condition exists. The command or script can parse the value of this environment variable, extract the node location information, and use that information to construct the appropriate remote command.

For example, consider the case where the **/var** file system is being monitored to ensure that it does not reach its capacity. When the file system does reach its capacity, the **chfs** command is to be invoked on the node where the condition exists to extend the size of the **/var** file system. To perform this action, a Korn Shell script is created. This script examines the contents of the **PMAN_IVECTOR** value, which has the following components to identify where the condition exists:

`VG=rootvg;LV=hd9var;NodeNum=`*`node_number_where_condition_exists`*

Once the node number has been found in the **PMAN_IVECTOR** value, the script will then find the host name for that node in the SDR. The script then uses the **dsh** command to issue the **chfs** command on the remote node to extend the size of the **/var** file system:

```
#!/bin/ksh
OLDIFS=$IFS
IFS=';'
set $PMAN_IVECTOR
for TOKEN in $*
do
if [[ $TOKEN = NodeNum* ]]
        then
              IFS='='
              print "$TOKEN" | read JUNK NODENUM
              HOST=$(SDRGetObjects -x Node node_number\=\=$NODENUM | \
                      awk '{print $11}')
        fi
done
IFS=$OLDIFS
if [[ "$HOST" != "" ]]
then
        dsh -w $HOST /usr/sbin/chfs -a size\=+1 /var
fi
```

This script is saved to an AIX file on the node where the monitoring request is made, and the execute permission is set on the file. The full path name of the file can then be provided to either Problem Management or the SP Event Perspective as a command to be issued when the event occurs.

Graphical tools - SP Event Perspective

The “Asynchronous (batch) notification methods” on page 30 introduced the concept of associating an action with a specific condition through the SP Event Perspective. This was done by providing a command within an event definition. In the previous section, this command was relatively simple: it issued an electronic mail message to a specific user to report the occurrence of the event.

To create an automatic response to the condition, provide an AIX command or script in the command field in addition to (or in place of) the notification command that was used in the previous discussion. Be sure to understand where the SP Event Perspective will attempt to run the command before assigning a command to this event, by referring to “Important - WHERE actions are performed” on page 39.

Command line tools - Problem Management

“Asynchronous (batch) notification methods” on page 30 introduced the concept of associating an action with a specific condition through the **pmandef** command of Problem Management. This association was done by specifying an argument to the

-c and **-C** options of the **pmandef** command. In the previous section, this argument was relatively simple: it issued an electronic mail message to a specific user to report the occurrence of the event.

To create an automatic response to the condition, provide an AIX command or script as an argument to **-c** or **-C** options of the **pmandef** command, in addition to (or in place of) the notification command that was used in the previous discussion. Be sure to understand where Problem Management will attempt to run the command before assigning a command to this event by referring to “Important - WHERE actions are performed” on page 39.

Chapter 3. Conditions to monitor on the SP system

Conditions to monitor using Perspectives or Problem Management

Whether you decide to monitor system condition using SP Perspectives or Problem Management, the sections that follow provide a list of the **minimum** hardware and software conditions for you to monitor. See “Descriptions of each condition” on page 44 for a detailed description of each condition.

Monitor these hardware conditions

For each frame, switch or node, monitor the following hardware conditions:

- For a frame:
  - Power
  - Controller responding
  - Controller ID mismatch (only applies to HACWS)
  - Temperature
  - Node slot failures
- For a switch:
  - Power and power availability
  - Environment light
  - Temperature
- For a node:
  - Power and power availability
  - Environment light
  - Temperature
  - Keymode switch
  - LED/LCD contents
  - LED/LCD flashing
  - Node responding
  - processorsOffline
- For the control workstation:

  Same as for a node

Monitor these software conditions

For each node, monitor the following software conditions:

- On each node:
  - Node can be reached by RSCT
  - **/tmp** becoming full
  - **/var** becoming full
  - **/** becoming full
  - Paging space low
  - Rising mbuf failures
  - Switch I/O errors
  - **inetd** daemon activity
  - **srcmstr** daemon activity
  - **ftpd** daemon activity

- **biod** daemon activity - applies only to NFS systems
- **portmap** daemon activity - used by RPC
- **xntpd** daemon activity (NTP time synch)
- **httpd** daemon activity (applies only to HTTP servers)
- **hatsd** daemon activity (cannot check using Event Management)
- **hadsd** daemon activity (cannot check using Event Management)
- **haemd** daemon activity (cannot check using Event Management)
- **cdsadv** daemon activity (DCE)
- **dced** daemon activity (DCE)

- On the control workstation:
  - All conditions to monitor on each node
  - **sdrd** daemon active
  - **kerberos** daemon activity
  - **secd** daemon activity (DCE)
  - **cdsd** daemon activity (DCE)

The DCE daemons that are listed (**cdsadv**, **dced**, **secd**, **cdsd**) are the minimum set that are required when using DCE on the SP system. You may have other DCE daemons running that you wish to monitor. For more information, consult *IBM DCE for AIX, Version 3.1: Administration Guide - Core Components*.

Descriptions of each condition

Table 1. Details about each condition to monitor

Condition	Details
Frame Power	**Description:** Whether the frame has its power on or off. When the frame power is off, nodes and switches in the frame cannot receive power. **Resource Variables:** • SP_HW.Frame.frPowerOff (overall power) • SP_HW.Frame.frPowerOff_A (power supply A) • SP_HW.Frame.frPowerOff_B (power supply B) • SP_HW.Frame.frPowerOff_C (power supply C) • SP_HW.Frame.frPowerOff_D (power supply D) • SP_HW.Frame.frACLED (AC power OK) • SP_HW.Frame.frDCLED (DC power OK) **Notes:** With the **frPowerOff_*** variables, only one needs to be **on** for the frame to receive power. If all of them are **off**, then the frame has no power. If power was not explicitly shut down by the system administrator, perform hardware diagnostics on the frame.

Table 1. Details about each condition to monitor (continued)

Condition	Details
Frame Controller Responding	**Description:** This indicates whether the frame controller is responding to command requests. The SP system can function when the frame controller is not responding, but it will not be possible to obtain certain node hardware status (such as key switch position and LED readouts) or issue certain hardware commands (such as resetting the node). When the controller fails, perform hardware diagnostics and replace the frame controller, if this is called for. Replacing the frame controller requires you to schedule down time for all nodes in that frame. **Resource Variable:** SP_HW.Frame.controllerResponds
Frame Controller ID Mismatch	**Description:** This indicates whether the ID of the frame controller agrees with the ID stored for it in the HACWS supervisor card. If the IDs do not match, this indicates that the HACWS supervisor card is not properly wired to the frame (possibly using the wrong ″tty″ line). Have the wiring between the control workstations (primary and backup) and the frame controller checked, and if that does not solve the problem, perform hardware diagnostics on both the control workstations and the frame controller. Monitor this condition only when HACWS is installed on the SP system. **Resource Variable:** SP_HW.Frame.controllerIDMismatch
Frame Temperature	**Description:** This indicates whether the frame's temperature is within the normal operational range. If the temperature becomes out of range, hardware within the frame may fail. Make sure that all fans are working properly. There are resource variables that you can check with the SP Event Perspective or the **hmmon** command to determine this. Make sure that the frame has proper ventilation. **Resource Variable:** SP_HW.Frame.tempRange

Table 1. Details about each condition to monitor (continued)

Condition	Details
Frame Node Slot Failures	**Description:** This indicates whether or not the frame supervisor can communicate with the node supervisor attached to the frame slot. It is possible to see a ″failure″ in this condition when no real failure exists. For example, since a wide node occupies two slots in the frame but only has one node supervisor, one of the slots associated with the wide node will always show a ″failure″. Any slots where nodes are not attached will show a ″failure″, but this is OK. This is why it is important to know the layout of the SP system. You should be concerned when the status changes to show a failure because it can indicate a failure in the node supervisor. The node may continue to function in this type of failure, but certain hardware status (LEDs, switch position) may not be available and commands (node reset) may not work. Run hardware diagnostics on the node connected to the frame slot showing a failure. **Resource Variables:** 1. SP_HW.Frame.nodefail1 2. SP_HW.Frame.nodefail2 3. SP_HW.Frame.nodefail3 4. SP_HW.Frame.nodefail4 5. SP_HW.Frame.nodefail5 6. SP_HW.Frame.nodefail6 7. SP_HW.Frame.nodefail7 8. SP_HW.Frame.nodefail8 9. SP_HW.Frame.nodefail9 10. SP_HW.Frame.nodefail10 11. SP_HW.Frame.nodefail11 12. SP_HW.Frame.nodefail12 13. SP_HW.Frame.nodefail13 14. SP_HW.Frame.nodefail14 15. SP_HW.Frame.nodefail15 16. SP_HW.Frame.nodefail16 17. SP_HW.Frame.nodefail17 **Note:** SP_HW.Frame.nodefail17 indicates a failure in the SP Switch supervisor. If the frame has no switch, this will always show a ″failure″.
Switch Power	**Description:** This indicates whether the switch power is on or off. If the frame has no power, the switch will not have power, so this should be checked first. If the frame has power but the switch does not, ensure that the switch was not manually shut down, and perform hardware diagnostics on the switch **Resource Variables:** • SP_HW.Switch.nodePower • SP_HW.Switch.powerLED • SP_HW.Switch.shutdownTemp **Notes:** SP_HW.Switch.nodePower indicates whether the power is on or off. SP_HW.Switch.powerLED indicates this as well, but also indicates whether the switch can receive power but is powered off. SP_HW.Switch.shutdownTemp indicates if the switch was powered off because of a high temperature condition.

Table 1. Details about each condition to monitor (continued)

Condition	Details
Switch Hardware Environment Indicator	**Description:** This indicates if the switch has detected any hardware anomalies that can cause or has caused a shut down of the switch. Such anomalies are: incorrect voltage, fan failure, temperature out of range, and internal hardware failure. This indicator shows whether all is well, whether a condition exists that should be investigated, or whether the switch was forced to shut down because of these errors. **Resource Variable:** SP_HW.Switch.envLED **Notes:** Any change in this indicator is worth investigating, even if the indicator shows that the problem is not yet critical. Check for fan failures in the SP Switch. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Perform hardware diagnostics on the switch. Schedule repair for any failing hardware components.
Switch Temperature	**Description:** This indicates if the temperature inside the switch hardware is out of the normal operational range. If the temperature becomes out of the normal range, the device may overheat and the hardware may fail. **Resource Variable:** SP_HW.Switch.tempRange **Notes:** Check for fan failures in the SP Switch. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command, and ensure that the frame has proper ventilation.
Node Power	**Description:** This indicates whether the node power is on or off. If the frame has no power, the node will not have power, so this should be checked first. If the frame has power but the node does not, ensure that the node was not manually shut down, and perform hardware diagnostics on the node. **Resource Variables:** SP_HW.Node.nodePower (should be used for SP-attached servers and clustered enterprise servers) SP_HW.Node.powerLED **Notes:** SP_HW.Node.nodePower indicates whether the power is on or off. SP_HW.Node.powerLED indicates this as well, but also indicates whether the node can receive power but is powered off.
Node Hardware Environment Indicator	**Description:** This indicates if the node has detected any hardware anomalies that can cause or have caused a shut down of the node. Such anomalies are: incorrect voltage, fan failure, temperature out of range, or internal hardware failure. This indicator shows whether all is well, whether a condition exists that should be investigated, or whether the node was forced to shut down because of these errors. **Resource Variable:** SP_HW.Node.envLED **Notes:** Any change in this indicator is worth investigating, even if the indicator shows that the problem is not yet critical. Check for fan failures in the node. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Perform hardware diagnostics on the node. Schedule repair for any failed hardware components.

Table 1. Details about each condition to monitor (continued)

Condition	**Details**
Node Temperature	**Description:** This indicates if the temperature inside the node hardware is out of the normal operational range. If the temperature becomes out of the normal range, the node may overheat and the hardware may fail. **Resource Variable:** SP_HW.Node.tempRange **Notes:** Check for fan failures in the node. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Ensure that the frame has proper ventilation, and check that all air paths within the frame are not clogged.
Node Key Mode Switch Position	**Description:** This shows the current setting of the node's mode switch. During node boot, the key switch position controls whether the operating system is loaded, and whether service controls are activated. During system operation, the position controls whether the node can be reset and whether system dumps can be initiated. For everyday operation, the key should be in the ″Normal″ position and should not change. A command must be issued to change the key position. If this does occur, locate the person changing this control and ensure that this action was taken for a proper reason. **Resource Variable:** SP_HW.Node.keyModeSwitch **Note:** Not all nodes have a key switch.
Node LED or LCD Readout	**Description:** Each node has an LCD or LED display. This display indicates the status of hardware and software testing during the node's boot process. This display is also used to display specific codes when node hardware or the operating system software fails. These codes indicate the nature of the failure, and whether any additional error data may be present. After a node has successfully booted, this display should be blank. If the display is not blank, use either the SP Hardware Perspective or the **spmon** command to determine what value is being displayed, and consult “Chapter 32. SP-specific LED/LCD values” on page 535, “Chapter 8. Network installation progress” on page 97, and “Other LED/LCD codes” on page 539 to determine what the LED or LCD value means. **Resource Variable:** SP_HW.Node.LCDhasMessage
Node Reachable by RSCT Group Services	**Description:** This indicates whether RSCT Group Services can reach the node through any of its network adapters and the switch. If this indicates that the node is not reachable, all of the node's network and switch adapters have either failed or been disabled. In this case, the only way to reach the node when it is powered on is through the node's serial link, using the **s1term** command. When this happens, check the network adapter status and issue the **/etc/ifconfig on** command to enable the adapter. Also, check the switch status and perform problem determination procedures for Group Services and Topology Services. **Resource Variable:** Membership.Node.state

Table 1. Details about each condition to monitor (continued)

Condition	Details
Node has a processor that is offline	**Description:** This indicates that when the node was booted, one or more processors did not respond. However, there is at least one active processor, so the node functions. **Resource Variable:** processorsOffline
/tmp file system becoming full	**Description:** Each node has its own locally available **/tmp** file system. This file system is used as temporary storage for many AIX and PSSP utilities. If this file system runs out of space, these utilities can fail, causing failures in those PSSP and LP utilities that depend on them. When this file system nears its storage capacity, it should be checked for large files that can be removed, or the file system size should be increased. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd3
/var file system becoming full	**Description:** Each node has its own locally available **/var** file system. This file system contains system logs, error logs, trace information, and other important node-specific files. If this file system runs out of space, log entries cannot be recorded, which can lead to loss of error information when critical errors occur, leaving you and IBM service personnel without an audit or debug trail. When the file system nears its storage capacity, it should be checked for old log information that can be removed or cleared, the file system size should be increased, or separate file systems should be made for subdirectories that consume large amounts of disk space. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd9var
/ file system becoming full	**Description:** Each node has its own locally available **root** file system. This file system contains important node boot and configuration information, as well as LP binaries and configuration files. If this file system runs out of space, it may not be possible to install products on the node, or update that node's configuration information (although the SMIT and AIX-based install procedures should attempt to acquire more space). When this file system nears its storage capacity, it should be checked for **core** files or any other large files that can be removed, or the file system's size should be increased. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd4

Table 1. Details about each condition to monitor (continued)

Condition	Details
Paging Space Low	**Description:** Each node has at least one locally available paging device. When all these paging devices near their capacity, the node begins to thrash, spending more time and resources to process paging requests than to process user requests. When operating as part of a parallel process, the thrashing node will delay all other parts of the parallel process that wait for this node to complete its processing. It can also cause timeouts for other network and distributed processes. A temporary fix is to terminate any non-critical processes that are using large amounts of memory. If this is a persistent problem, a more permanent fix is to restrict the node to specific processing only, or to add additional paging devices. **Resource Variable:** aixos.PagSp.%totalused
Kernel Memory Buffer Failures	**Description:** Kernel memory buffers, or ″mbufs″, are critical to network processing. These buffers are used by the kernel network protocol code to transfer network messages. If the kernel begins to encounter failures in acquiring these buffers, network information packets can be lost, and network applications will not run efficiently. An occasional failure can be tolerated, but numerous failures or a continuous stream of small failures indicates that not enough memory has been allocated to the kernel memory buffer pool. **Resource Variable:** aixos.Mem.Kmem.failures **Resource Identifier:** Type = mbuf
Switch Input and Output Errors	**Description:** The switch device driver tracks statistics on the device's use, including any errors detected by the driver. These errors are tracked as counters that are never reset, unless the node is rebooted. Consult "Chapter 15. Diagnosing SP Switch problems" on page 137 and "Chapter 16. Diagnosing SP Switch2 problems" on page 185 for assistance in diagnosing any reported errors for the SP Switch or SP Switch2. **Resource Variables:** • CSS.ibadpackets • CSS.ipackets_drop • CSS.ierrors • CSS.opackets_drop • CSS.oerrors • CSS.xmitque_ovf **Notes:** Any increment in the value of the CSS.ierrors or CSS.oerrors counters indicates that the switch adapter is about to go offline. Continual increments to the CSS.ibadpackets counter can indicate transmission problems or ″noise″ in the connection between the SP Switch adapter and the SP Switch, so the SP Switch cabling should be checked and hardware diagnostics performed. Continual increments to the CSS.ipackets_drop and CSS.opacktes_drop counters indicate that there is either too much input or too much output for the SP Switch device driver to handle, and packets are lost.

Table 1. Details about each condition to monitor (continued)

Condition	Details
inetd Daemon Activity	**Description:** The **inetd** master daemon is responsible for activating many AIX and PSSP service daemons when a client for that service connects to the node. If the daemon fails, these services cannot be started. Since many SP applications are network applications, this can cause widespread failure in all SP applications. If the daemon cannot be restarted manually, force a system dump of this node, collect information for the IBM Support Center, and restart the node. The reboot may temporarily resolve the problem. **Resource Variable:** Prog.xpcount **Resource Identifier:** ProgName=inetd UserName = root
srcmstr Daemon Activity	**Description:** The **srcmstr** daemon implements the System Resource Controller functions. If this daemon fails, services registered with the SRC cannot be controlled using SRC commands. If the daemon cannot be restarted manually, force a system dump of this node, collect information for the IBM Support Center, and restart the node. The reboot may temporarily repair the problem. **Resource Variable:** Prog.xpcount **Resource Identifier:** ProgName=srcmstr UserName = root
biod Daemon Activity	**Description:** The **biod** daemon handles block I/O requests for the NFS file system. In order for NFS to function on a node, at least one **biod** daemon must be active. For normal NFS activity, six to eight **biod** daemons are usually active on a node. For higher NFS activity, some nodes may have more. These daemons are started from node boot, run continuously, and should not shut down. If any daemons shut down, consult the NFS documentation for diagnostic procedures, and attempt to restart the daemon. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=biod UserName = root
portmap Daemon Activity	**Description:** This daemon knows all the registered ports on the node, and which programs are available on each of these ports. The daemon's task is to convert Remote Procedure Call (RPC) port numbers to Internet port numbers. RPC clients use this daemon to resolve their RPC port numbers. If the daemon fails, the daemon itself and all RPC servers on the node must be restarted. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=portmap UserName = root

Table 1. Details about each condition to monitor (continued)

Condition	Details
xntpd Daemon Activity	**Description:** This daemon is active on a node when the Network Time Protocol (NTP) time synchronization protocol is running. This daemon ensures that the node's time-of-day hardware is synchronized with the network's time server. A failure in the daemon does not necessarily mean that the time of day hardware on the node will no longer be synchronized with the network, although this danger does exist. A failure in the daemon does mean that time change updates from the network server will not be made on this node. Such problems can lead to failures in RSCT's Topology Services component, which may begin to see packets arriving out of chronological order, and may cause RSCT to falsely detect that one of its peer nodes has failed. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=xntpd UserName = root
kerberos Daemon Activity	**Description:** The **kerberos** daemon runs on the node where the Kerberos databases are stored. You need to know which node this is to properly check this condition. The daemon is responsible for accepting Kerberos V4 client requests for principal information, service tickets, and Kerberos V4 database maintenance. Failure in this daemon will cause failures in Kerberos V4 clients to acquire or validate credentials, which will lead to denial of service for users of the Kerberos V4 clients. If this daemon fails, consult "Chapter 18. Diagnosing SP Security Services problems" on page 251 for Kerberos V4 diagnostics, and attempt to restart the daemon. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=kerberos UserName = root
hatsd Daemon Activity	**Description:** This is the RSCT Topology Services daemon, which is responsible for maintaining an internal topology map of the SP system on this node. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, the node will be seen as ″down″ by all other nodes in this system partition. Other consequences of this daemon's failure to restart include the RSCT Group Services daemon on the node will fail and the RSCT Event Management daemon will fail. This daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing. To check this daemon's activity, you must use the **lssrc -g hats** command or the **ps -ef | grep hats** command.

Table 1. Details about each condition to monitor (continued)

Condition	Details
hagsd Daemon Activity	**Description:** This is the RSCT Group Services daemon, which is responsible for handling Group Services functions for all Group Services clients on this node. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, all Group Services clients on this node will appear to have failed, as far as the Group Services group members are concerned. Those groups will begin their member failure processing for the Group Services clients on this node. The daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing. To check this daemon's activity, you must use the **lssrc -g hags** command or the **ps -ef | grep hags** command.
haemd Daemon Activity	**Description:** This is the RSCT Event Management daemon, which is responsible for handling Event Management registrations on this node and communicating with other Event Management daemons in this system partition. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, none of the Event Management resource variables from this node will be available to Event Management applications for monitoring or event generation purposes. These affected applications include Problem Management and the SP Event Perspective. This daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing . To check this daemon's activity, you must use the **lssrc -g haem** command or the **ps -ef | grep haem** command.
sdrd Daemon Activity	**Description:** This daemon runs on the control workstation (and therefore must be checked only on that node), and services all requests made of the System Data Repository (SDR). Although a failure in this daemon may not have any immediate consequences, PSSP software services will not be able to access SDR information, and can fail at later times when this information is needed. Certain hardware monitoring capability can also be lost, and may result in widespread, falsely detected ″node not responding″ failures. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=sdrd UserName = root
dced Daemon Activity	**Description:** The DCE client for the **dcecp** command. This client runs on each host in the SP system when DCE authentication is used. Failures in this daemon can prevent you from administering principals, accounts, passwords and server keys and obtaining other necessary information regarding DCE users. For problem resolution, refer to *IBM DCE for AIX, Version 3.1: Problem Determination Guide*. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=dced UserName = root

Table 1. Details about each condition to monitor (continued)

Condition	**Details**
cdsadv Daemon Activity	**Description:** The DCE client for the Cell Directory Service. This client runs on each host in the SP system when DCE authentication is used. The **cdsd** and **cdsadv** together provide the cell directory service which is essentially a distributed information database. Failures in the client or server daemon may cause problems accessing information in the same way as a file system failure on a local machine. Items of interest that are kept in the Cell Directory Service are keytab objects and account objects, as well as others. For problem resolution, refer to *IBM DCE for AIX, Version 3.1: Problem Determination Guide*. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=cdsadv UserName = root
cdsd Daemon Activity	**Description:** The DCE server for the Cell Directory Service. This master server runs on the host you have designated to be the server host for DCE for your SP system. The **cdsd** and **cdsadv** together provide the cell directory service which is essentially a distributed information database. Failures in the client or server daemon may cause problems accessing information in the same way as a file system failure on a local machine. Items of interest that are kept in the Cell Directory Service are keytab objects and account objects, as well as others. For problem resolution, refer to *IBM DCE for AIX, Version 3.1: Problem Determination Guide*. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=cdsd UserName = root
secd Daemon Activity	**Description:** The DCE Security Server. This master server runs on the host you have designated to be the server host for DCE for your SP system. Failures in this daemon cause problems relating to authentication and authorization. A typical problem would be a login failure through **dce_login**. SP Services using DCE for authenticating trusted services may hang if the DCE Security Server is not running. The hang condition can also occur if **secd** is running but the registry has been disabled. You may see messages such as ″Cannot find KDC for requested realm″. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=secd UserName = root

Preparing to examine and monitor this information

This section describes how to define the conditions that are discussed previously, to the SP Event Perspective and how to define event definitions that are associated with these conditions.

SP Event Perspective - conditions that you can monitor using the default event definition

The following conditions have a default event definition which you can install using the SP Event Perspective.

Table 2. Conditions and default event definitions

Condition to monitor	Default event definition
Frame power	framePowerOff
Frame controller responding	frameControllerNotResponding
Switch power	switchPowerLED
Node power	nodePowerLED
Node power	nodePowerDown
Node responding	hostResponds
Node H/W environment	nodeEnvProblem
Node key switch	keyNotNormal
Node LED/LCD readout	LCDhasMessage
Node reachable by RSCT	nodeNotReachable
/tmp file system filling	tmpFull
/var file system filling	varFull
Page space low	pageSpaceLow
sdrd daemon activity	sdrDown
SP Switch responding	switchResponds
SP Switch2 responding	switchResponds0

Each of these conditions is to be monitored in all locations, with the exception of **sdrd** Daemon Activity, which is to be monitored only on the control workstation. Each default event definition contains a definition for the condition, so you need only load and register the event definition. Use the following procedure for each of the event definitions listed previously:

1. Bring up the SP Event Perspective.
2. Click in the Event Definitions pane.
3. Select Actions → Load Defaults from the menu bar.

   This opens the Load Default Event Definitions dialog box.
4. Select the default event definition of interest from the list.
5. If you wish to register the selected event definition, click **Register the selected event definitions**.
6. Click **OK** to load the selected default event definition and close the dialog box.

SP Event Perspective - conditions that you can monitor that you must define to the SP Event Perspective

This section contains instructions for using the SP Event Perspective to create the conditions that do not have default definitions.

To start:

1. Bring up the SP Event Perspective

2. If the Conditions pane is hidden, select View → Add Pane from the menu bar, to add the Conditions pane.
3. Click in the Conditions pane.
4. Select Actions → Create from the menu bar.
5. This displays the Create Conditions Notebook.

Now use the SP Event Perspective to create a condition for each of the following conditions:

- Frame Controller ID Mismatch - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameControllerIDMismatch.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Frame.controllerIDMismatch.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Frame Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameTempOutOfRange.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select tempRange.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Frame Node Slot Failure - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameSlotFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Frame.nodefail.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
  8. Repeat the previous steps for each IBM.PSSP.SP_HW.Frame.nodefail* resource variable.
- Switch Power Shutdown - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchShutdownTemp.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.

4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.switch.shutdownTemp.
5. In the **Event expression** field, enter X==1
6. Leave the **Rearm expression** field blank.
7. Press the **Create** button.

- Switch Hardware Environment Indicator - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchEnvLED.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Switch.envLED.
  5. In the **Event expression** field, enter X>0
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Switch Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchTemp.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Switch.tempRange.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Node Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, nodeTemp.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Node.tempRange.
  5. In the **Event expression** field, enter X>0
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- **/** File system Filling - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, rootFull.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.aixos.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.%totused.
  5. In the **Event expression** field, enter X>90

6. In the **Rearm expression** field, enter X<80
7. In the **Resource ID Elements to be Fixed for the Condition** field, enter LV=hd4;VG=rootvg.
8. Press the **Create** button.

- Kernel Memory Buffer Failures **-** create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, mbufFailures.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.aixos.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.Mem.Kmem.failures.
  5. In the **Event expression** field, enter X>X@P
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter Type=mbuf.
  8. Press the **Create** button.
- **srcmstr** Daemon Activity **-** create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, srcmstrFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.xpcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=srcmstr;UserName=root
  8. Press the **Create** button.
- **biod** Daemon Activity **-** create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, biodFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=biod;UserName=root
  8. Press the **Create** button.
- **portmap** Daemon Activity **-** create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, portmapFailure.
  2. Type a description of the condition.

3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
5. In the **Event expression** field, enter X@0<X@1
6. Leave the **Rearm expression** field blank.
7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=portmap;UserName=root
8. Press the **Create** button.

- **xntpd** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, xntpFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=xntpd;UserName=root
  8. Press the **Create** button.
- **kerberos** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, kerberosFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=kerberos;UserName=root
  8. Press the **Create** button.
- **dced** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, dcedFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=dced;UserName=root
  8. Press the **Create** button.

- **cdsadv** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, cdsadvFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=cdsadv;UserName=root
  8. Press the **Create** button.
- **cdsd** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, cdsdFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=cdsd;UserName=root
  8. Press the **Create** button.
- **secd** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, secdFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=secd;UserName=root
  8. Press the **Create** button.

SP Event Perspective - creating the event definitions

Create one event definition for each of the conditions that are listed previously. Most of these specify that the conditions be checked in all locations, while some are on specific nodes. Those that are checked in all locations are:

- Frame Controller ID Mismatch
- Frame Temperature
- Switch H/W Environment Indicator
- Switch Temperature

- Node Temperature
- **/** file system becoming full
- Kernel Memory Buffer Failures
- **srcmstr** Daemon Activity
- **portmap** Daemon Activity

Conditions to monitor only in specific locations

Conditions to monitor only in specific locations are:

- Frame Node Slot Failures - The slots to be monitored are those slots where nodes are directly connected in the frame.

  Those slots that cannot be used because wide or high nodes occupy previous slots are not to be monitored. You will need to identify which slots in which frames are to be monitored, and you can create one event per slot.

  When creating the event definition, select the proper ″nodefail″ resource name for the slot, and select the range of frame numbers where this slot is in use.
- **biod** Daemon Activity - This daemon is monitored only for nodes that use the NFS file system.

  When creating the event definition, select those nodes where NFS is used.
- **xntpd** Daemon Activity - This daemon is monitored only for nodes that use NTP time service.

  When creating the event definition, select only those nodes where NTP is used.
- **kerberos** Daemon Activity - This daemon is monitored on the control workstation if the Kerberos database resides on the control workstation.
- DCE - When creating the event definition, specify only the nodes where DCE is configured. Specify the control workstation if DCE is configured anywhere on the SP system.
  - **cdsd** Daemon Activity - This daemon is monitored on the control workstation if the DCE servers reside on the control workstation.
  - **secd** Daemon Activity - This daemon is monitored on the control workstation if the DCE servers reside on the control workstation.

Creating an event definition using the Create Event Definition notebook

On the Definition Page:

1. Give the event definition a descriptive name. For example, you can use rootFullEvent for the rootFull condition.
2. Click the down-arrow button for **Name** in the **Condition** box, and find the condition you want to use.
3. For those conditions that are to be observed in all locations, click the **Wild Card Element** selector under **Specify remaining resource ID elements** box.
4. For those that are to be observed in specific locations, select those locations under **Element Values**.

On the **Notifications** page, make sure the **Notification** (Get notified of events during the SP Event Perspective session) button is selected.

On the **Actions** page:

1. Click the **Take these actions when the event occurs** button.
2. In the **Run command** field, enter the following:

   ```
   'echo $PMAN_HANDLE Alert: Instance Vector $PMAN_IVECTOR |
   mail -s "$PMAN_HANDLE Alert" username@address'
   ```

Then click the **Create** button at the bottom of the page.

SP Hardware Perspective

The SP Hardware Perspective can be used to investigate further when the SP Event Perspective detects a hardware problem. This Perspective can show you the node's LED or LCD values (the SP Event Perspective cannot), and can be used to control the hardware or reset nodes when necessary.

Problem Management

Problem Management can be used as an alternative to the SP Event Perspective, to define event definitions that monitor conditions on your SP system.

Ensuring that the conditions you intend to monitor are known to Problem Management

The SP Event Perspective uses Problem Management whenever the event definition indicates that a command should be issued when the event occurs. If you used the SP Event Perspective to set up the event definitions and conditions, these definitions may already be known to Problem Management. Use the **pmanquery** command to check if the definition already exists. For example:

```
pmanquery -n varFullEvent
```

will check for the **varFullEvent** definition.

If the event definitions are not known, use the **pmandef** command to define them. Any event definitions made using the **pmandef** command are also usable by the SP Event Perspective later on. To create new event definitions, use the **pmandef** command. It is best to write a shell script with all the **pmandef** commands in it, and then invoke the shell script.

When the **pmandef** commands complete, the event definitions will be registered, which means that Problem Management will begin to check for these conditions immediately. If this is a problem, issue the **pmandef -d** command to disable the event until you are ready to monitor the condition. Then, use the **pmandef -a** command to activate it.

For more information about problem management, see the chapter on using the Problem Management subsystem in *PSSP: Administration Guide*. For more information about the **pmandef** command, see the man page for the command in *PSSP: Command and Technical Reference*.

The pmandef commands for specific conditions

Use these **pmandef** commands to set up event definitions for the following conditions. The field *username* is the user ID to receive the notification, and the *address* is a hostname.

- Frame power

```
pmandef -s framePowerEvent \
-e 'IBM.PSSP.SP_HW.Frame.frPowerOff:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Frame controller responding:

```
pmandef -s frameCtrlRespondingEvent \
-e 'IBM.PSSP.SP_HW.Frame.controllerResponds:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Frame controller ID mismatch:

```
pmandef -s frameCtrlIDMismatchEvent \
-e 'IBM.PSSP.SP_HW.Frame.controllerIDMismatch:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Frame temperature:

```
pmandef -s frameTempEvent \
-e 'IBM.PSSP.SP_HW.Frame.tempRange:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Frame node slot failure:

```
pmandef -s frameSlot1FailureEvent \
-e 'IBM.PSSP.SP_HW.Frame.nodefail1:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Frame Switch slot failure:

```
pmandef -s frameSwitchSlotFailEvent \
-e 'IBM.PSSP.SP_HW.Frame.nodefail17:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Switch power:

```
pmandef -s switchPowerEvent \
-e 'IBM.PSSP.SP_HW.Switch.powerLED:FrameNum=*:X!=1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Switch hardware environment indicator:

```
pmandef -s switchLEDEvent \
-e 'IBM.PSSP.SP_HW.Switch.envLED:FrameNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Switch shutdown due to extreme temperature:

```
pmandef -s switchTempShutdownEvent \
-e 'IBM.PSSP.SP_HW.Switch.shutdownTemp:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Switch temperature:

```
pmandef -s switchTempEvent \
-e 'IBM.PSSP.SP_HW.Switch.tempRange:FrameNum=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address'
```

- Node power:

```
pmandef -s nodePowerEvent \
-e 'IBM.PSSP.SP_HW.Node.powerLED:NodeNum=*:X!=1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node controller responding:

```
pmandef -s hostRespondsEvent \
-e 'IBM.PSSP.SP_HW.Node.hostResponds:NodeNum=*:X==0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node hardware environment indicator:

```
pmandef -s nodeLEDEvent \
-e 'IBM.PSSP.SP_HW.Node.envLED:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node temperature:

```
pmandef -s nodeTempEvent \
-e 'IBM.PSSP.SP_HW.Node.tempRange:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node key mode switch position:

```
pmandef -s nodeKeySwitchNotNormalEvent \
-e 'IBM.PSSP.SP_HW.Node.keyModeSwitch:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node LCD or LED readout:

```
pmandef -s nodeLCDMessageEvent \
-e 'IBM.PSSP.SP_HW.Node.LCDhasMessage:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node processorsOffline:

```
pmandef -s processorsOfflineEvent \
-e 'IBM.PSSP.SP_HW.Node.processorsOffline:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node reachable by RSCT Group Services:

```
pmandef -s nodeNotReachableEvent \
-e 'IBM.PSSP.Membership.Node.state:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/tmp** File system filling:

```
pmandef -s tmpFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;\
VG=rootvg;LV=hd3:X>90'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/var** File system filling:

```
pmandef -s varFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;\
VG=rootvg;LV=hd9var:X>90" -r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/** File system filling:

```
pmandef -s rootFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;VG=rootvg;LV=hd4:X>90'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Paging space low:

```
pmandef -s pageSpaceLowEvent \
-e 'IBM.PSSP.aixos.PagSp.%totalused:NodeNum=*:X>90' \
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Kernel memory buffer failures:

```
pmandef -s mbufFailureEvent \
-e 'IBM.PSSP.aixos.Mem.Kmem.failures:NodeNum=*;Type-mbuf:X>X@P'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch input errors:

```
pmandef -s switchInputErrEvent \
-e 'IBM.PSSP.CSS.ierrors:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch output errors:

```
pmandef -s switchOutputErrEvent \
-e 'IBM.PSSP.CSS.oerrors:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Transmit Queue overflows:

```
pmandef -s switchInputErrEvent \
-e 'IBM.PSSP.CSS.xmitque_ovf:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **inetd** daemon activity:

```
pmandef -s inetdFailureEvent \
-e 'IBM.PSSP.Prog.xpcount:NodeNum=*;\
ProgName=inetd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **srcmstr** daemon activity:

```
pmandef -s srcmstrFailureEvent \
-e 'IBM.PSSP.Prog.xpcount:NodeNum=*;\
ProgName=srcmstr;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **biod** daemon activity:

```
pmandef -s biodFailureEvent \
-e 'IBM.PSSP.Prog.pcount:NodeNum=*;\
ProgName=biod;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **portmap** daemon activity:

```
pmandef -s portmapFailureEvent \
-e 'IBM.PSSP.Prog.pcount:NodeNum=*;\
ProgName=portmap;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **xntpd** daemon activity:

```
pmandef -s xntpdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=range of nodes where xntpd daemon runs;\
ProgName=xntpd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **kerberos** daemon activity:

```
pmandef -s kerberosdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where kerberos daemon runs;\
ProgName=kerberos;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **dced** daemon activity:

```
pmandef -s dcedFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where dced daemon runs;\
ProgName=dced;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **cdsadv** daemon activity:

```
pmandef -s cdsadvFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where cdsadv daemon runs;\
ProgName=cdsadv;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **cdsd** daemon activity:

```
pmandef -s cdsdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where cdsd daemon runs;\
ProgName=cdsd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **secd** daemon activity:

```
pmandef -s secdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where secd daemon runs;\
ProgName=secd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

Chapter 4. Error logging overview

This section describes methods for managing the AIX Error Log on the SP system. Commands and SMIT interfaces for performing general log management, and for managing Syslog and the AIX Error Log, are also installed with the **sysman** option. Refer to the Error Log Management chapter in *PSSP: Administration Guide* for an overview.

Error logging is the writing of information to persistent storage to be used for debugging purposes. This type of logging is for subsystems that perform a service or function on behalf of an end user. The subsystem does not communicate directly with the end user and, therefore, needs to log events to some storage location. The events that are logged are primarily error events.

Error logging for the SP uses Berkley Software Distribution (BSD) syslog and AIX Error Log facilities to report events on a per node basis. The intent is to have the AIX Error Log be the starting point for diagnosing system problems. The AIX Error Log is named **/var/adm/ras/errlog**.

The BSD system log has a default name of **/var/adm/SPlogs/SPdaemon.log**. It is controlled by the **syslogd** daemon. This log must be created by issuing the **syslogd** command. For instructions on using **syslogd**, refer to the man page for the **syslogd** command, or *AIX 5L Version 5.1 Commands Reference*.

Error log entries include a “DETECTING MODULE” string that identifies the software component, module name, module level, and line of code or function that detected the event that was logged. The information is formatted depending on the logging facility the user is viewing. For example, the AIX Error Log facility information appears:

```
DETECTING MODULE
LP=LP name Fn=file name SID_level_of_the_file L#=Line number
```

The BSD syslog facility information appears:

```
timestamp, hostname, ID, PID
LP=LP name Fn=file name SID_level_of_the_file L#=Line number
```

Classifying Error Log events

The following table displays the mapping of SP Error Log label suffixes to syslog priorities and AIX Error Log error types.

Table 3. SP Error Log label suffixes mapped to syslog priorities and AIX Error Log types

Error label suffix	syslog priority field	syslog description	AIX Error Log error type	AIX Error Log description
EM	LOG_EMERG	Emergency, system unstable	PEND	The loss of availability of a device is imminent.
ER	LOG_ERR	Error condition	PERM	No recovery from this condition. A permanent error occurred.
ST	LOG_NOTICE	Normal, but significant condition	UNKN	It is not possible to determine the severity of the error.
TR	LOG_INFO	Informational message	UNKN	It is not possible to determine the severity of the error.

Table 3. SP Error Log label suffixes mapped to syslog priorities and AIX Error Log types (continued)

Error label suffix	syslog priority field	syslog description	AIX Error Log error type	AIX Error Log description
RE	LOG_DEBUG	Debug message	TEMP	Condition was recovered after several unsuccessful attempts.
DE	LOG_DEBUG	Debug message	UNKN	It is not possible to determine the severity of the error.

Effect of not having a battery on error logging

In a typical RS/6000, a battery is installed to maintain NVRAM. On an SP system there is no battery, and NVRAM may be lost when the node is powered off. AIX writes the last error log entry to NVRAM. During system startup the last entry is read from NVRAM and placed in the error log when the errdemon is started. This last error log entry may be important in diagnosis of a system failure.

On SP wide nodes the NVRAM does have power to it as long as the node is plugged into the frame and the frame is plugged into a working power source. On SP thin nodes, NVRAM is lost whenever the node is powered down. If the last error log entry is desired, the thin nodes should not be powered off. They should be re-IPLed in the “normal” key mode switch position if at all possible.

Managing and monitoring the error log

To manage and monitor the error log, you can do the following:
- View error log information in parallel.
- View SP switch error log reports.
- Use AIX error log notification.

Viewing error log information in parallel

It may be helpful when diagnosing a system problem to look at all of the error logs at once in parallel.

It is not a good idea to copy the **/var/adm/ras/errlog** files from the various nodes to a central place and then run **errpt** against the combined file. First, copying time is added to the sequential processing time of all the nodes and the total time required will be longer than viewing the logs in parallel. Second, error log analysis requires per node information from the ODM database (on each node).

Note: A user must have specific authorization to use the **dsh** command. To learn how a user can acquire this authorization, see ″Using the SP System Monitor″ chapter of *PSSP: Administration Guide*.

Use the **dsh** command with the **errpt** command and its options to view the error log. Perform the following steps:

1. View the summary information for all nodes to determine which ones are to be examined more closely. For example:

   ```
   dsh -a errpt -s 0930020094 |pg
   ```

   In this example, all error entries that occurred after September 30, 1994 at 2 a.m. for every node defined in the System Data Repository, are listed. The output is piped to **pg** in a one entry per line format.
2. Pick out the nodes that have error entries that require further examination.

3. View the selected nodes. For example:

   ```
   dsh -w host1,host2,host3 errpt -a -s 0930020094 > /tmp/930errors
   ```

   This example collects all the fully expanded error log reports after September 30, 1994 at 2 a.m. from nodes with a hostname of **host1, host2, host3**.

Summary log for SP Switch, SP Switch2, and switch adapter errors

For systems running PSSP 3.1 or higher, a centralized error log records information about SP Switch, SP Switch2, and switch adapter errors. Logging of switch and adapter errors in the AIX error log on nodes and on the control workstation causes the generation of a summary record in the summary log. This log has the name: **/var/adm/SPlogs/css/summlog** and is located on the control workstation. The summary log provides a centralized location for monitoring system-wide error activity. It also improves the usability of log output collected from individual nodes.

The summary log contains one summary entry for each CSS error log entry recorded on the failing node or control workstation. Entries in the log have the following fields, which are separated by blanks:

- **Timestamp** - A timestamp in the form: MMDDhhmmYYYY.
- **Node** - The *reliable_hostname* as stored in the SDR, for the originating node, with the domain portion removed.
- **Snap indicator** - A value that indicates whether a snap dump was taken:
  - **Y** indicates that a snap dump was taken.
  - **N** indicates that a snap dump was **NOT** taken.
- **Partition** - The name of the system partition to which the node belongs.

  For error log entries that do not pertain to a particular system partition, this field contains **global**.
- **Index** - The error log index for the entry being reported.
- **Label** - The error log entry label field for the entry being reported.

The summary log contains a record for each CSS error log entry produced on each node in the system. You can use this log to obtain a single image of error activity across the entire SP system. Using the log, you can identify situations involving multiple nodes and determine the nodes that are affected. You can use the timestamps to determine which node experienced a problem first, so that you can more easily identify the root cause of a problem.

Viewing SP Switch error log reports

Enter the following command to view all the SP switch adapter error reports in parallel:

```
dsh -a errpt -a -N css
```

It sends to stdout all the fully-formatted error log entries for all unusual status detected for the switch adapter device drivers that are contained in the error log. This may be for the past 90 days. AIX has a default **crontab** entry that removes all hardware error entries after 90 days and all software error entries after 30 days.

Enter the following command to view all the SP switch information in parallel:

```
dsh -a errpt -a -N Worm
```

It sends to stdout all the fully-formatted error log entries for the switch. This includes errors found during switch diagnostics.

Using the AIX Error Notification Facility

You can be notified of an SP error when it occurs by using the AIX Error Notification Facility.

IBM General Concepts and Procedures for RS/6000 (GC23-2202) explains how to use the AIX Error Notification Facility. *IBM RS/6000 Problem Solving Guide* (SC23-2204) explains the use of the AIX Error Log. This facility will perform an ODM method defined by the administrator when a particular error occurs or a particular process fails. The following classifications of errors can have notification objects defined by the administrator. Many of these messages will not occur often, so these notification objects can be defined even for large SP systems.

1. **PSSP AIX Error Log Labels that end in _EM.**

   The EM suffix signifies an emergency error and is usually used to tell the administrator information that would be needed to re-IPL a node. To find these messages, issue the command:

   ```
   errpt -t |grep "_EM "
   ```

2. **Any AIX Error Log entries that have an Error Type of PEND.**

   PEND signifies an impending loss of availability, and that action will soon be required of the administrator.

3. **Any AIX Error Log entries for the boot device of the node.**

   The boot device of the node usually has a resource name of **hdisk0**, but the name may vary if the installation has been customized.

4. **The AIX Error Label EPOW_SUS.**

   The EPOW_SUS error log entry is generated before power down when an unexpected loss of electrical power is encountered.

5. **The AIX Error Labels KERN_PANIC and DOUBLE_PANIC.**

   KERNEL_PANIC or DOUBLE_PANIC error log entries are generated when a kernel panic occurs.

The examples on the following pages may help the administrator in adding Error Notification Objects on the SP system. Adding a **dsh -a** command to the ODM commands will perform the action on all nodes of the SP system.

Example 1

Mail the error report to **root@controlworkstation** when a switch adapter fails online diagnostics.

- **Step 1**. Set up directories for the Error Notification objects and methods.

  ```
  mkdir /customerdefinedpath/errnotify/objects
  mkdir /customerdefinedpath/errnotify/methods
  ```

  Keep the methods scripts on each node so you can run them if distributed file system problems occur. File Collections is an excellent way to keep these scripts updated. The object files may be in a distributed file system since they are not used unless changes to the object are required.

- **Step 2**. Create the Error Notification Method scripts.

  Create a script or program that will be run when the error occurs. For example:

  ```
  #!/bin/ksh
  ######################################################################
  # Run errpt to get the fully expanded error report for the error
  # that was just written and redirect to a unique tempfile with the PID
  #of this script.
  ######################################################################
  errpt -a -l $1 > /tmp/tempfile.$$
  ```

```
####################################################################
# Mail the fully expanded error report to root@controlworkstation
# This could be anywhere in the network.
# root@controlworkstation is the user and hostname that the
# administrator wants to be notified at.
#####################################################################
mail root@controlworkstation < /tmp/tempfile.$$
```

- **Step 3**. Create the Error Notification Object

  Create a file that contains the Error Notification Object to catch the switch diagnostic failed error.

  ```
  errnotify:
            en_name = "tbx_diagerr.obj"
            en_persistenceflg = 1
            en_label = "SWT_DIAG_ERROR2_ER"
            en_method = "/customerdefinedpath/methods/errnot.
                        test.ksh$1"
  ```

  (The **en_name** value can be a maximum of 16 characters long.) Enter the **odmshow errnotify** command to view the Error Notification object.

  It is easy to modify an existing set of ODM errnotify stanzas. To do this, enter:

  ```
  odmget errnotify > file
  ```

  and edit the file. Include only attributes that have values.
- **Step 4.** Add the Error Notification Object to the errnotify class.

  ```
  odmadd  /customerdefinedpath/object/tbx_diagerr.obj
  ```

  (The file name is the name of the file with the Error Notification Object in it.)

  To delete this object, enter:

  ```
  odmdelete -o errnotify -q "en_name = tbx_diagerr.obj"
  ```

  To view this object in the ODM database, enter:

  ```
  odmget  -q "en_name = tbx_diagerr.obj" errnotify
  ```
- **Step 5.** The following mail will be sent to **root@controlworkstation** when an SP Switch MX adapter fails diagnostics:

  ```
  From root@sp2n5.kgn.ibm.com Mon Oct  3 11:25:59 1994
  Received: from sp2n5.kgn.ibm.com by ppsras.kgn.ibm.com
            (AIX 3.2/UCB 5.64/4.03)
            id AA24781; Mon, 7 May 1995 10:14:59 -0400
  Date: Mon, 3 Oct 1994 11:25:59 -0400
  From: root
  Message-Id: <9410031525.AA24781@sp2n5.kgn.ibm.com>
  To: root
  Status: RO

   ---------------------------------------------------------------------------
  ERROR LABEL: SWT_DIAG_ERROR2_ER
  ERROR ID: 323C48A0

  Date/Time:       Mon Oct  3 11:25:57
  Sequence Number: 18282
  Machine Id:      000004911800
  Node Id:         sp2n5
  Error Class:     H
  Error Type:      PERM
  Resource Name:   Worm
  Resource Class:  NONE
  Resource Type:   NONE
  Location:        NONE
  ```

```
Error Description
Switch adapter failed On-Line diagnostics

Probable Causes
Switch clock signal missing
Switch adapter failure

User Causes
Switch cable loose or disconnected

Recommended Actions
Run adapter diagnostics

Failure Causes
Switch adapter hardware

Recommended Actions
Run adapter diagnostics

Detail Data
DETECTING MODULE
LP=PSSP,Fn=dtb3mx,SID=1.35,L#=1303,
Service Request Number
763—942
```

Example 2

Error Notification when any Error Type of PEND occurs.

- **Steps 1 and 2** are the same as defined in the switch diagnostic failure example.
- **Step 3.** Create the Error Notification Object

  Create a file that contains the Error Notification Object to catch the pending availability problems. For example:

  ```
  errnotify:
          en_name = "errnot.PEND.obj"
          en_persistenceflg = 1
          en_type = "PEND"
          en_method = "/tmp/errnot.test.ksh $1"

  errnotify:
          en_name = "errnot.pend.obj"
          en_persistenceflg = 1
          en_type = "pend"
          en_method = "/tmp/errnot.test.ksh $1"

  errnotify:
          en_name = "errnot.Pend.obj"
          en_persistenceflg = 1
          en_type = "Pend"        en_method = "/tmp/errnot.test.ksh $1"
  ```

  (The variations of PEND are added because upper case is not strictly adhered to by all AIX LPs and vendors.)
- **Step 4.** Add the Error Notification Objects to the errnotify class. For example:

  ```
  odmadd  /customerdefinedpath/object/errnot.pend.obj
  ```

  (The file name is the name of the file with the Error Notification Object in it.)

  To delete these objects enter:

  ```
  odmdelete -o errnotify -q "en_name = errnot.PEND.obj"
  odmdelete -o errnotify -q "en_name = errnot.pend.obj"
  odmdelete -o errnotify -q "en_name = errnot.Pend.obj"
  ```

  To view this object in the ODM database, enter:

```
odmget  -q "en_name = errnot.PEND.obj" errnotify
odmget  -q "en_name = errnot.pend.obj" errnotify
odmget  -q "en_name = errnot.Pend.obj" errnotify
```

- **Step 5**. Mail is sent to the administrator when an error that has an Error Type of PEND occurs.

Example 3

Error Notification when any Error on the boot device of **hdisk0** occurs.

- **Step 1 and 2** are the same as defined in “Example 1” on page 70.
- **Step 3**. Create the Error Notification Object.

  Create a file that contains the Error Notification Object to catch the boot disk errors. Assume that **hdisk0** is the boot device.

  ```
  errnotify:
          en_name = "errnot.boot.obj"
          en_persistenceflg = 1
          en_resource = "hdisk0"
          en_method = "/tmp/errnot.test.ksh $1"
  ```

- **Step 4**. Add the Error Notification Object to the errnotify class.

  ```
  odmadd  /customerdefinedpath/object/errnot.boot.obj
  ```

  To delete this object, enter:

  ```
  odmdelete -o errnotify -q "en_name = errnot.boot.obj"
  ```

  To view this object in the ODM database, enter:

  ```
  odmget  -q "en_name = errnot.boot.obj" errnotify
  ```

- **Step 5**. Mail with the fully expanded error report will be sent to the administrator when an error on **hdisk0** occurs.

Example 4

Error Notification when unexpected power loss and kernel panics occur.

- **Steps 1 and 2** are the same as defined in “Example 1” on page 70.
- **Step 3**. Create the Error Notification Object

  Create a file that contains the Error Notification Object to catch the kernel panic and power loss Error Labels. For example:

  ```
  errnotify:
                 en_name = "power.obj"
                 en_persistenceflg = 1
                 en_label = "EPOW_SUS"
                 en_method = "/customerdefinedpath/methods/
                             errnot.test.ksh $1"

  errnotify:
                 en_name = "panic.obj"
                 en_persistenceflg = 1
                 en_label = "KERNEL_PANIC"
                 en_method = "/customerdefinedpath/methods/
                             errnot.test.ksh $1"

  errnotify:
                 en_name = "dbl_panic.obj"
                 en_persistenceflg = 1
                 en_label = "DOUBLE_PANIC"
                 en_method = "/customerdefinedpath/methods
                             /errnot.test.ksh $1"
  ```

- **Step 4.** Add the Error Notification Object to the errnotify class. For example:

  ```
  odmadd  /customerdefinedpath/object/power.panic.obj
  ```

The file name is the name of the file with the Error Notification Object in it.

- **Step 5**. Mail with the fully expanded error report will be sent to the administrator when any power loss or kernel panic occurs.

Using the SP logs

The SP System uses the standard logs provided by both AIX and the public domain software it includes, as well as SP-specific logs. Some logs reside on the control workstation only, and some reside only on the SP nodes. Others reside on both. Table 4 summarizes and shows the location of the logs to use when diagnosing SP problems. The abbreviation **CWS** stands for control workstation in this table.

Your IBM Support Center representative may ask you to provide information from these logs.

Table 4. SP log files

Type of message	Log file name	Location
Output of the **phoenix.snap** tool. See “phoenix.snap dump” on page 376.	**/tmp/phoenix.Snap/all.***timestamp***.tar.Z**	CWS, nodes
Output of the **SDR_test** command when run without **root** authority	**/tmp/SDR_test.log**	CWS, nodes
Standard AIX error log entries including the SP Switch	**/var/adm/ras/errlog**	Nodes
Error messages and verbose messages from security programs.	**/var/adm/SPlogs/auth_install/log**	CWS, nodes
Automounter messages	**/var/adm/SPlogs/auto/auto.log**	CWS, nodes
Messages from the **cstartup** command	**/var/adm/SPlogs/cs/cstart.***timestamp.pid*	CWS
Messages from the **cshutdown** command	**/var/adm/SPlogs/cs/cshut.***timestamp.pid*	CWS
Details of recovery actions for the IBM RVSD function	**/var/adm/SPlogs/csd/vsd.debuglog**	Nodes
Summary of recovery actions for the IBM RVSD function	**/var/adm/SPlogs/csd/vsd.log**	Nodes
Trace output of **pssp_script**, which performs the post install customization	**/var/adm/SPlogs/css/$nim_client_shr.config.log**	Nodes
Switch cable miswire information	**/var/adm/SPlogs/css/cable_miswire**	Primary node
Messages from the **css.snap** script	**/var/adm/SPlogs/css/css.snap.log**	CWS, nodes
Switch admin daemon messages	**/var/adm/SPlogs/css/cssadm.debug**	CWS
Switch admin daemon messages (stdout)	**/var/adm/SPlogs/css/cssadm.stdout**	CWS
Switch admin daemon messages (stderr)	**/var/adm/SPlogs/css/cssadm.stderr**	CWS
Fault service daemon messages for the SP Switch2	**/var/adm/SPlogs/css/daemon.log**	Nodes
Fault service daemon messages (stderr) for the SP Switch	**/var/adm/SPlogs/css/daemon.stderr**	Nodes

Table 4. SP log files (continued)

Type of message	Log file name	Location
Fault service daemon messages (stdout) for the SP Switch	**/var/adm/SPlogs/css/daemon.stdout**	Nodes
System error messages that occurred while distributing the topology file to the nodes.	**/var/adm/SPlogs/css/dist_topology.log**	Primary node
Trace of SP Switch adapter diagnostics failures	**/var/adm/SPlogs/css/dtbx_failed.trace**	Nodes
SP Switch adapter diagnostics trace information	**/var/adm/SPlogs/css/dtbx.trace**	Nodes
Messages from adapter diagnostics (stderr)	**/var/adm/SPlogs/css/dtbxworm.stderr**	Nodes
Log from the **Eclock** command	**/var/adm/SPlogs/css/Eclock.log**	CWS
Log from all **Ecommands** issued	**/var/adm/SPlogs/css/Ecommands.log**	CWS
Log from the **Emonitor** command	**/var/adm/SPlogs/css/Emonitor.log**	CWS
Result of **Estart** commands issued by the **Emonitor** daemon	**/var/adm/SPlogs/css/Emonitor.Estart.log**	CWS
Trace of last **Eunpartition** operation	**/var/adm/SPlogs/css/Eunpart.file**	Primary node
Trace file of fault service daemon messages	**/var/adm/SPlogs/css/fs_daemon_print.file**	Nodes
Switch fault information	**/var/adm/SPlogs/css/flt**	Nodes
stdout and stderr of the Event Management Resource Monitor and Methods	**/var/adm/SPlogs/css/logevnt.out**	CWS
SP Switch and SP Switch2 advanced diagnostics Messages Daemon log	**/var/adm/SPlogs/css/msdg.log**	CWS
Description of problems that arise while switch is initializing	**/var/adm/SPlogs/css/out.top**	Primary node
Initialization messages from the SP Switch support code	**/var/adm/SPlogs/css/rc.switch.log**	Nodes
Log from switch router generation	**/var/adm/SPlogs/css/router.log**	Nodes
Output log from switch router generation when it detects a failure	**/var/adm/SPlogs/css/router_failed.log**	Nodes
SP Switch advanced diagnostics tests and architecture components log	**/var/adm/SPlogs/css/spd.trace**	CWS, nodes
SP Switch2 advanced diagnostics tests and architecture components log	**/var/adm/SPlogs/cssX/pY/spd.trace**, where **X** and **Y** are variable. See “SP Switch2 log and temporary file hierarchy” on page 187.	CWS, nodes
SP Switch advanced diagnostics GUI log	**/var/adm/SPlogs/css/spd_gui.log**	CWS
Summary records for events logged to AIX error log on nodes.	**/var/adm/SPlogs/css/summlog**	CWS
stdout and stderr for the CSS logging daemon’s Event Management client	**/var/adm/SPlogs/css/summlog.out**	CWS

Table 4. SP log files (continued)

Type of message	Log file name	Location
Current state of the switch network, details about attached nodes and the topology file	**/var/adm/SPlogs/css/topology.data**	Primary node
Worm trace file from switch initialization	**/var/adm/SPlogs/css/worm.trace**	Primary node
SP Switch2 log files. See “SP Switch2 log and temporary file hierarchy” on page 187.	**/var/adm/SPlogs/css0/***	Primary node
SP Switch2 adapter diagnostics messages.	**/var/adm/SPlogs/css0/colad.trace** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
Trace information for the **cssadm2** daemon.	**/var/adm/SPlogs/css0/cssadm2.debug** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
Unexpected error messages received by the **cssadm2** daemon, while performing commands external to the daemon.	**/var/adm/SPlogs/css0/cssadm2.stderr** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
Unexpected informational messages received by the **cssadm2** daemon, while performing commands external to the daemon.	**/var/adm/SPlogs/css0/cssadm2.stdout** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
Errors logged by the **la_event_d** daemon for the SP Switch2 PCI Adapter.	**/var/adm/SPlogs/css0/la_error.log** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
la_event_d daemon verbose messages and raw dump data for the SP Switch2 PCI Adapter.	**/var/adm/SPlogs/css0/la_event_d.trace** See “SP Switch2 log and temporary file hierarchy” on page 187.	nodes
SP Switch2 log files. See “SP Switch2 log and temporary file hierarchy” on page 187.	**/var/adm/SPlogs/css1/***	Primary node
Results of the last CSS verification test	**/var/adm/SPlogs/CSS_test.log**	CWS
Output of the **supper** command	**/var/adm/SPlogs/filec/sup**_date.time_	Nodes
Actions **supper** performs when updating file collections	**/var/adm/SPlogs/filec/sup**_date.time_**r**	Nodes
Messages generated by the last **get_keyfiles** command issued	**/var/adm/SPlogs/get_keyfiles/get_keyfiles.log**	Nodes
Kerberos V4 authentication database administration daemon	**/var/adm/SPlogs/kerberos/admin_server.syslog**	Primary authentication server
Kerberos V4 primary authentication server log	**/var/adm/SPlogs/kerberos/kerberos.log**	Primary authentication server
Kerberos V4 secondary authentication server log	**/var/adm/SPlogs/kerberos/kerberos.slave_log**	Secondary authentication server
Kerberos V4 authentication database propagation daemon	**/var/adm/SPlogs/kerberos/kpropd.log**	Secondary authentication server

Table 4. SP log files (continued)

Type of message	**Log file name**	**Location**
Messages generated by transfer of srvtab files to nodes	**/var/adm/SPlogs/kfserver/kfserver.log.***PID*	CWS
Messages generated by registration process for the **kfserver** program	**/var/adm/SPlogs/kfserver/regserver.log**	CWS
Messages generated by the Problem Management daemon	**/var/adm/SPlogs/pman/pmand.log**	Nodes
Messages generated by the Problem Management daemon	**/var/adm/SPlogs/pman/pmand.***partition name***.log**	CWS
System Data Repository configuration messages	**/var/adm/SPlogs/sdr/SDR_config.log**	CWS
Output of the **SDR_test** command when run with **root** authority	**/var/adm/SPlogs/SDR_test.log**	CWS, nodes
System Data Repository error messages	**/var/adm/SPlogs/sdr/sdrdlog.***partition.pid*	CWS
Login control messages	**/var/adm/SPlogs/spacs/spacs.log**	Nodes
SP configuration Vital Product Data directory	**/var/adm/SPlogs/SPconfig/***	CWS, nodes
Vital Product Data output for the node	**/var/adm/SPlogs/SPconfig/***node number***.umcl**	CWS, nodes
lscfg -v command output	**/var/adm/SPlogs/SPconfig/***node number***.lscfg**	CWS, nodes
Messages generated by system daemons, including hardware errors	**/var/adm/SPlogs/SPdaemon.log**	CWS, nodes
Hardware Monitor hmc daemon log	**/var/adm/SPlogs/spmon/hmcd/ hmcd[***ipaddress***].log.** *julian_date*	CWS
SP extension node messages	**/var/adm/SPlogs/spmgr/spmgrd.log**	CWS
Hardware Monitor initialization and error messages	**/var/adm/SPlogs/spmon/hmlogfile.** *julian_date*	CWS
Node conditioning messages	**/var/adm/SPlogs/spmon/nc/nc.***frame.node*	CWS
Hardware Monitor **s70d** daemon error messages	**/var/adm/SPlogs/spmon/s70d/s70d**.*frame*.**log**.*julian_date*	CWS
Activity of the logging daemon (**splogd**).	**/var/adm/SPlogs/spmon/splogd.debug**	CWS
Contains the PID of the logging daemon, (**splogd**).	**/var/adm/SPlogs/spmon/splogd/splogd.pid**	CWS
SP Logging daemon state changes	**/var/adm/SPlogs/spmon/splogd.state_changes.** *timestamp* Note: this log is not shipped activated. To activate, see "Configuration test 1 - Check for state logging" on page 337.	CWS
Microcode download messages recorded when using smit (**smitty supervisor** command)	**/var/adm/SPlogs/spmon/ucode/ucode_log.***frame.node*	CWS
Output of the **spmon_ctest** command	**/var/adm/SPlogs/spmon_ctest.log**	CWS
Output of the **spmon_itest** command	**/var/adm/SPlogs/spmon_itest.log**	CWS

Table 4. SP log files (continued)

Type of message	Log file name	Location
Job Switch Resource Table Services information and error messages	**/var/adm/SPlogs/st/st_log**	Nodes
Sysctl server log messages	**/var/adm/SPlogs/sysctl/sysctld.log**	CWS, nodes
AIX error messages from mirroring a root volume group using SP volume group commands	**/var/adm/SPlogs/sysman/mirror.out**	Nodes
System Management configuration messages	**/var/adm/SPlogs/sysman/***node*.**config.log**.*pid*	Nodes
System Management first boot configuration messages	**/var/adm/SPlogs/sysman/***node*.**configfb.log**.*pid*	Nodes
System Management console messages	**/var/adm/SPlogs/sysman/***node*.**console.log**	CWS, nodes
System Management configuration messages	**/var/adm/SPlogs/sysman/spfbcheck.log**	Nodes
AIX error messages from unmirroring a root volume group using SP volume group commands	**/var/adm/SPlogs/sysman/unmirror.out**	Nodes
Informational and error messages from the **SYSMAN_test** command	**/var/adm/SPlogs/SYSMAN_test.log**	CWS, nodes
Event Management activity log	**/var/ha/log/em.default.***partition-name*	CWS, nodes
hags internal trace and log file	**/var/ha/log/hags***	CWS, nodes
GS service log - summary log	**/var/ha/log/hags***_node_incarnation***.***partition***.long** on PSSP nodes. **/var/ha/log/hags.***partition_node_incarnation***.***partition***.long** on the PSSP control workstation.	CWS, nodes
hagsglsm internal trace and log file	**/var/ha/log/hagsglsm***	CWS, nodes
Trace information for the topology services daemon	**/var/ha/log/hats.***dd.hhmmss.partition-name*	CWS, nodes
Information from the topology services startup script	**/var/ha/log/hats.***partition-name*	CWS, nodes
Trace information for Network Interface Modules, processes used by the Topology Services daemon to monitor network interfaces.	**/var/ha/log/nim.hats.***interface name***.***partition name***.00n** for PSSP nodes **/var/ha/log/nim.topsvcs***interface name***.***cluster name***.00n** for HACMP nodes.	CWS, nodes
Hardmon resource monitor messages	**/var/ha/run/haem.***hostname***/IBM.PSSP.hmrmd/ IBM.PSSP.hmrmd_log.***julian-date*	CWS
filec_config command log, no longer in use	**/var/sysman/logs/***	Nodes
SP SNMP Agent messages	**/var/tmp/SPlogs/spmgr/spgrd.log**	CWS, nodes

Log files are cleaned up on the nodes by the **cleanup.logs.nodes** command. Log files are cleaned up on the control workstation by the **cleanup.logs.ws** command. By default, continuously growing logs are trimmed to 400 lines every night, and non-growing files are deleted after seven days. The exceptions are:

- The **supper** message log, which is deleted after two days.
- The SDR log, which is deleted after seven days, only if it is **not** the current log.
- The **SPdaemon.log**, which is trimmed.
- The SP extension node SNMP Manager log file size is controlled by the user when the daemon is started.
- The Event Management log, which is trimmed when it reaches 256KB.
- The **kfserver.log.***PID*, which is deleted after 30 days.
- The **regserver.log**, which is deleted after 30 days.
- The **JSRT Services** log, which is trimmed when it reaches 100KB.
- The Group Services logs, **/var/ha/log/hags*** and **/var/ha/log/hagsglsm/***, which are trimmed according to the LOGSIZE specified when starting Group Services.
- The Topology Services logs - three instances of the daemon logs are kept. Current logs are trimmed to a given number of lines. This number of lines is a tunable parameter stored in the **Log_Length** attribute of the **TS_Config** SDR class. The default value is 5000 lines. The startup script log is not trimmed, and the seven latest instances are kept.
- The following logs are not trimmed:
  - The css (switch) logs
  - The **sysctld** logs
  - The logging daemon log
  - The SP logging daemon state changes log
  - The **s70d** log
  - The hardmon resource monitor log
  - The **SYSMAN_test** log
  - The **get_keyfiles** log

Chapter 5. Producing a system dump

Table 5. System dump information

Symptom	Recovery
Nodes do not respond or system crashes.	“Action 1. Produce a system dump” “Action 2. Verify the system dump” on page 83

Actions

Action 1. Produce a system dump

When your nodes do not respond or when your system crashes, a system dump may help you determine the cause of the problem. A system dump contains a copy of the kernel data on the system at the time of the crash. This section explains how to produce a dump, verify it, copy the dump to tape, and send the tape to IBM.

In some cases the system produces a dump automatically. If the system senses a fatal condition, it usually dumps automatically to the primary dump device and puts flashing **888** in the node’s three-digit display.

Attention

Do not initiate a system dump if the node’s three-digit display is **888**. If you initiate a dump, you will overwrite the dump that was taken at the time of the problem.

Instead, proceed to “Action 2. Verify the system dump” on page 83.

Dump methods

There are several ways you can produce a system dump. Some of the methods work with all configurations, and others do not. Each method explained here includes this configuration information.

Notes:

1. Graphical interface users can use the SP Hardware Perspective to operate the system controls. You can reset the node or put it in service mode either from the **Nodes Status** page of the **Node Notebook**, or the **Actions** menu.
2. Command interface users can use the **spmon** command to operate the system controls.

   A node can be reset using the command:

   ```
   /usr/lpp/ssp/bin/spmon -reset
   ```

   On systems that do not have a key mode switch, the **spmon -reset** command produces a dump.

   A node’s key mode switch can be altered using the command:

   ```
   /usr/lpp/ssp/bin/spmon -key state
   ```

   where *state* is either **normal**, **secure**, or **service**.

Dump to the primary dump device

Choose one of these methods to produce a dump on the primary dump device.

Method 1: This method works for all systems that have a key switch.

Set the key mode switch to the Service position and press the Reset button once.

Method 2: This method can only be done from a directly-attached keyboard. It cannot be done from a **tty** connection. This method works only on the control workstation.

Set the key mode switch to the Service position and, while holding the **<Ctrl>** and **<Alt>** keys, press the 1 on the numeric key pad.

Method 3: This method works for all system configurations, if the system is responding to commands.

Login as **root** and enter:

```
sysdumpstart -p
```

On the SP system, when **sysdumpstart** is issued on a PCI node, the SPLED display in SP Perspectives indicates **stby**. Following a stby, power off the node and then back on. **DO NOT RESET THE NODE**. A reset will lose the dump taken.

Method 4: This method works for nodes with virtual keys. It produces a dump to the default dump device, as defined by AIX. Issue the following commands from the control workstation, specifying the node's *frame* and *slot*.

1. `hmcmds service` *frame:slot*
2. `hmcmds reset` *frame:slot*
3. Wait for the **0c2** status code to change to **0c0**
4. `hmcmds normal` *frame:slot*
5. `hmcmds off` *frame:slot*

Dump to the secondary dump device

Choose one of these methods to produce a dump on the secondary dump device.

Note

If the secondary dump device is a removable media device, such as a tape or diskette drive, make sure that the medium is in the device.

Method 5: This method can only be done from a directly-attached keyboard. It cannot be done from a **tty** connection. This method works only on the control workstation.

Set the key mode switch to the Service position and, while holding down **<Ctrl>** and **<Alt>** keys, press the 2 on the numeric key pad.

Method 6: This method works for all system configurations, if the system is responding to commands.

Login as **root** and enter:

```
sysdumpstart -s
```

Action 2. Verify the system dump

You may have a system dump because you initiated it yourself or because the system produced one automatically. In either case, follow these steps to verify that the system dump was successful and that the information it contains is usable.

1. Record the three-digit codes.
   - If the system dumped automatically, the three-digit display will show flashing **888**. Press Reset repeatedly until **888** displays again and write down each three-digit code that is displayed. The last code before **888** displays again indicates if the dump was successful. Check the dump code status in the next table for more information.
   - If you initiated the dump yourself, the three-digit code that is displayed indicates if the dump was successful. Check the dump code status in the next table for more information.

Table 6. System dump status codes

Three-digit code	Meaning
0c0	The dump completed successfully.
0c1	An I/O error occurred while taking the dump.
0c2	A user-initiated dump is in progress.
0c4	The dump device was too small but the dump may still be usable. If zero bytes are written and 0c4 is displayed, it means the dump device was large enough but the system was hung and not able to initiate a dump.
0c5	An internal error occurred while taking the dump.
0c6	Prompts you to make the secondary dump device available.
0c7	The dump facility is waiting for a response from the NFS (Network File Server).
0c8	No dump device is defined.
0c9	A system-initiated dump is in progress.
0cc	The dump facility has switched to the secondary dump device.

2. On Micro Channel Nodes, change the key mode switch to “normal”, power off the node and power it back on. On PCI nodes, which do not have a key mode switch, power off the node and power it back on. This will allow the last error log entry stored in NVRAM to be placed in the error log. See “Effect of not having a battery on error logging” on page 68.

   Note: **DO NOT** hit the reset button, because this will cause the current dump information to be overwritten.
3. Log in as **root**.
4. Verify the dump device by entering:

   ```
   sysdumpdev
   ```

   This should return something like:

   ```
   primary          /dev/hd7
   secondary        /dev/sysdumpnull
   ```

Note the primary dump device name, and substitute it for **/dev/hd#** in the following steps.

5. Verify the dump by entering:

   ```
   sysdumpdev -L
   ```

   Output is similar to:

   ```
   0453-039

   Device name:         /dev/lv00
   Major device number: 10
   Minor device number: 10
   Size:                67108352 bytes
   Date/Time:           Wed Apr  5 14:52:35 EDT 2000
   Dump status:         -2
   dump device too small
   ```

   In this case, a 0c4 LCD was in the crash codes, and the dump device was too small. There maybe enough dump information, since 67108352 bytes of information were written to the dump device **/dev/lv00**. Continue to 6. If no bytes were written, the system was hung and no dump exists. **DO NOT** continue with these steps.

6. Verify the usability of the dump by entering:

   ```
      crash /dev/hd#
   ```

   This should return:

   ```
      Using /unix as the default namelist file.
      Reading in symbols.......................
   ```

   - If you get the message: “`ATTENTION: dumpfile does not appear to match namelist`”, either the dump did not take place or the **/unix** file does not match the dump that was in the dump device.

     The dump file is not useful. Enter **q** to quit the **crash** command. **DO NOT** continue with these steps and do not send the dump to IBM.
   - If messages are not displayed, proceed with the next step.
   - Enter **q** to quit the **crash** command.

7. Enter the **errdead** command to extract the error records from the **/dev/error** buffer and place them in the error log:

   ```
   /usr/lib/errdead /dev/hd#
   ```

8. Issue the **crash** command again:

   ```
   crash /dev/lvxx
   ```

   where **/dev/lvxx** is the dump device.

   When you see the **>** prompt, enter:

   ```
      stat
   ```

   Output is similar to:

   ```
      sysname:   AIX
      nodename:  journey
      release:   2
      version:   3
   ```

```
machine:   000052643100
time of crash:  Sun Jan 24 19:18:53 1993
age of system:  18 day, 1 hr., 29 min.
```

- If the **time of crash** in the output approximately matches the time the system crashed, the dump is sufficient for analysis. Continue with the next step.
- If the **time of crash** in the output does not approximately match the time of the system crash, Enter **q** to quit the **crash** command. The data is not useful. Do not continue with these steps and do not send the dump to IBM.

Now put the dump symptom string information into the error log. Issue the command:

```
symptom -e
```

This copies the symptom string into the error log, and can be used to search problem databases for duplicate problems.

9. Enter:

   ```
   trace
   ```

   Look for a trace report similar to this sample:

   ```
   STACK TRACE:
      .m_freem ()
      .soreceive
      ._recv
      .recv
   ```

   Enter **q** to quit the **crash** command.

Gather the dump and other **snap** or log information for the IBM Support Center. Contact your local service representative or call the IBM Support Center to open a Problem Management Record as explained in “How to contact the IBM Support Center” on page 13.

Chapter 6. Diagnosing hardware and software problems

This section provides troubleshooting information for the SP hardware and software. It contains tables to help you isolate the cause of SP problems and recover from them. The first table describes the high level symptoms and gives you a course of action or directs you to other tables to further analyze the problem. Note that some of the tables in the diagnosing sections for the various components list recovery actions. Such actions are further described within the body of the section.

High-Level SP symptoms

The following table lists the high-level symptoms you may experience and directs you to the corresponding chapter for each one.

Table 7. High-Level SP symptoms

Symptoms	Condition	Action
controllerResponds indicator on the Frame Notebook in the SP Hardware Perspective is red, **SPdaemon.log** reports this variable as 0, or the Frame Notebook in the SP Hardware Perspective displays blank nodes. The **hmmon -GQv controllerResponds** *frame*:*slot* command shows: `Frame responding to polls FALSE`.	Frame supervisor failure	Go to "Chapter 21. Diagnosing System Monitor problems" on page 317
You get error messages when you start the switch or applications that use the switch fail or hang.	Switch Failure	For the SP Switch, see "Chapter 15. Diagnosing SP Switch problems" on page 137. For the SP Switch2, see "Chapter 16. Diagnosing SP Switch2 problems" on page 185.
switchResponds indicator on the Nodes Noteboook in the SP Hardware Perspective is red, meaning that the node is not currently available to the switch network. When the indicator is yellow, the node is available to, but not part of, the network. The **spmon -G d** command shows **switchResponds** as **no**.	Switch or switch adapter problems	For the SP Switch, see "Chapter 15. Diagnosing SP Switch problems" on page 137. For the SP Switch2, see "Chapter 16. Diagnosing SP Switch2 problems" on page 185.
The **spmon**, **hmmon**, or **hmcmds** commands fail.	System Monitor problem	Go to "Chapter 21. Diagnosing System Monitor problems" on page 317
SP Perspectives is having trouble starting or running.	SP Perspectives problem	Go to "Chapter 30. Diagnosing Perspectives problems on the SP System" on page 513
Cannot monitor or operate hardware controls, cannot power frames, nodes, switch on or off.	Hardware problem or System Monitor problem	Go to "Chapter 21. Diagnosing System Monitor problems" on page 317
The configuration indicated by the SDR and Perspectives do not match your system configuration.	Frame Supervisor connectivity problem.	Go to "Chapter 21. Diagnosing System Monitor problems" on page 317.

Table 7. High-Level SP symptoms (continued)

Symptoms	**Condition**	**Action**
3DigitDisplay indicator in the SP Hardware Perspective displays three-digit codes. LED/LCD displays present for a given node.	Hardware or software problem	Check “Chapter 32. SP-specific LED/LCD values” on page 535 to see if the code is listed there. If not, see “Other LED/LCD codes” on page 539. Also, refer to *IBM RS/6000 Problem Solving Guide*, SC23-2204
Node hangs, cannot access system via **ping** or a remote command, or node crashes with **888** in the *3DigitDisplay* indicator.	Hardware or software problem	Go to “Chapter 5. Producing a system dump” on page 81
You have an orange icon at any level of the System Monitor topology display.	System connectivity problem	Go to “Chapter 12. Diagnosing system connectivity problems” on page 121
Cannot **ping** on the external network, remote commands or **telnet** and **rlogin** commands fail.	System connectivity problem	Go to “Chapter 12. Diagnosing system connectivity problems” on page 121
User access problems, cannot log in, password is not valid, cannot get to home directory.	Software problem	Go to “Chapter 28. Diagnosing User Access problems” on page 501
You get an error message in response to an SP command on the control workstation.	Hardware or software problem	Look up the message in *PSSP: Messages Reference* and follow the action suggested.
A number of SP Services fail at once on a node or the control workstation.	Software problem	See “Chapter 19. Diagnosing Per Node Key Management (PNKM) problems” on page 289 and “Chapter 18. Diagnosing SP Security Services problems” on page 251.

Part 2. Diagnosing PSSP subsystems

Chapter 7. Diagnosing NIM problems

This chapter discusses diagnostic procedures and failure responses for the Network Installation Management (NIM) service component of AIX, as it relates to PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 94. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 94.

Related documentation

The following publications provide information about the NIM component:

1. *PSSP: Installation and Migration Guide*
2. *PSSP: Administration Guide*
3. *PSSP: Messages Reference*
4. *IBM AIX 5L Version 5.1 Network Installation Management Guide and Reference*

Requisite function

This is a list of the software used by the NIM component. Problems within the requisite software may manifest themselves as error symptoms in the NIM. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the NIM component of PSSP, consider the following components as possible sources of the error.

1. Network connectivity - If there are problems in your network, NIM will experience failures. See “Chapter 12. Diagnosing system connectivity problems” on page 121 and “Chapter 13. Diagnosing IP routing problems” on page 123.
2. SP System Security Services

   Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

Obtain the following information relative to the failure:

1. NIM Configuration Information

   Issue this command on the boot/install server of the failing node:

   ```
   lsnim -l
   ```
2. SDR Configuration Information

   Issue these commands:

   - `splstdata -n`
   - `splstdata -v`
   - `splstdata -b`
3. Record all relevant error information, messages, and symptoms.
4. Node LED/LCD indicators are error status codes. They are located on the node’s LED/LCD if the node has one. If not, use the SP Hardware Perspective and go to the node notebook to see its LED/LCD code. For LED/LCD values not issued by PSSP, see “Other LED/LCD codes” on page 539.

   Many LED/LCD’s are displayed as status during the node installation. You should be concerned only when a LED/LCD value is displayed for an extended period of time. Refer to the rest of this section for common LED/LCD failures and to the hardware documentation for additional values. NIM LED/LCD values are all in the range of **600** through **699**. Status LED/LCD's have even-numbered values, error LED/LCD's have odd-numbered values.

5. Error messages from the **setup_server** command.

   These are errors encountered while configuring the NIM environment. They appear on the boot/install server. This is the stderr output of the **setup_server** command. Any messages that require user action will be stated as errors as opposed to the normal status messages.

Trace information

NIM debug SPOT

A NIM debug SPOT (Shared Product Object Tree) contains trace output from NIM invocation on the node. To diagnose a hang during installation (LED 611), it may be necessary to create and use a NIM debug SPOT. This section describes how to do this.

Note: To run diagnostics on a node, the node supervisor card must be at microcode version **1294** or later versions. To determine the microcode level of the card, issue this command on the control workstation, substituting # with the frame number and node number respectively.

```
/usr/lpp/ssp/bin/spmon –G –q –l frame#/node#/codeVersion/value
```

If your card is not at microcode version **1294** or later versions, debug installation may loop issuing this message: **032-001 You entered a command** *command_name* **that is not valid.**

1. Obtain the **lppsource** name and boot/install server for the failing node by issuing the command:

   ```
   splstdata -b -n node_number
   ```

2. On the control workstation, issue the **spbootins** command to set the boot response to **disk**. For example, for frame 1 node 15, issue the command:

   ```
   spbootins –r disk 1 15 1
   ```

   This will issue the necessary NIM commands to prepare for reallocation of the debug SPOT for frame 1 node 15.

3. From the boot/install server issue:

   ```
   nim –Fo check –a debug=yes spot_lppsource_name
   ```

4. When the previous command completes, issue the command:

   ```
   lsnim -l spot_lppsource_name
   ```

5. Look for lines that start with: **enter_dbg =**.

   Choose a line as follows:

   - If there is only one line, use that line.
   - If there is more than one line, determine which one to use as follows:
     a. If the node is a 332 MHz SMP wide node or 332 MHz SMP thin node, use the line that has **chrp.mp**.
     b. For an SMP node, use the line that has **rs6k.mp**.
     c. Otherwise, use the line with **rs6k.up**.

The line chosen contains an address, such as **0x0013afa0**. Omit the **0x** part and record the remainder of the address.

6. Issue the **spbootins** command to set the node's boot response to **install**.

   For example, for frame 1 node 15 issue the command:

   ```
   spbootins –r install 1 15 1
   ```

7. Condition the node. For example, for frame 1 node 15 issue the command:

   ```
   nodecond 1 15 &
   ```

8. Open a read-only tty. For frame 1 node 15, issue the command:

   ```
   s1term 1 15
   ```

   This may take a few minutes to complete. Do **not** enter anything until it finishes with:

   ```
   Trap instruction interrupt
   ```

   and a **0>** prompt is displayed. Type **<Ctrl-c>** to stop your **s1term**.

9. Now start a console log to capture debug output:

   ```
   script filename
   ```

   where *filename* is a name you choose for the file.

10. Open a write console to the node:

    ```
    s1term -w
    ```

11. You should see the **0>** prompt again. Issue:

    ```
    st hex_number 2
    ```

    (where *hex_number* is the address recorded in Step 5 on page 92, omitting the **0x**).

12. You should see the **0>** prompt, again. Issue:

    ```
    g
    ```

    The netboot will now be displayed as **live**. “Chapter 8. Network installation progress” on page 97 and “Chapter 32. SP-specific LED/LCD values” on page 535 give the meanings of the LED/LCD codes and help determine approximately where in the boot process your node is.

    As the node boots, it may hang with LED/LCD c46. This does not indicate a problem, but the debug netboot needs to be restarted by issuing **<Ctrl-q>**. If there is a hang at any other LED/LCD value, stop logging by going to Step 13.

13. When the node hangs, exit the tty by typing **<Ctrl-x>**. Then, stop the logging by sending a **kill** signal to the script process from Step 7, by issuing:

    ```
    kill pid
    ```

    To get the *pid*, issue:

    ```
    ps –ef | grep script
    ```

    If there are two scripts, killing the child process will stop both of them, or the **kill** command may be used on both processes.

    Now view the log file to determine what went wrong with the installation. If you contact the IBM Support Center, make sure that you have the log file available.

14. Finally, you will need to re-create a regular version of the SPOT. From the control workstation, issue this command:

    ```
    nim -Fo check spot_lppsource_name
    ```

where *lppsource_name* is the same name used in Step 3 on page 92.

NIM SPOT logs

NIM SPOT logs contain trace output from SPOT creation or update. The logs are located on the boot/install server. The trace is automatically activated when a SPOT is updated or created. These logs are located in: **/tmp/spot.out.***pid* and **/tmp/spot.updated.out.***pid*. Error messages give the exact file name.

NIM commands and their responses are recorded. Look for errors associated with command invocation, errors associated with file set installation, and errors associated with other NIM-related activities.

Information to collect before contacting the IBM Support Center

Obtain the following information before contacting IBM.

1. NIM Configuration Information. See “Error information” on page 91.
2. SDR Configuration Information. See “Error information” on page 91.
3. All relevant error information, messages, and symptoms.
4. Node LED/LCD indicators are error status codes.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the NIM component. Locate the symptom and perform the action described in the following table.

If you have a symptom that is not in the table, or the recovery action does not correct the problem, see “Information to collect before contacting the IBM Support Center” and contact the IBM Support Center.

Table 8. NIM symptoms

Symptom	Recovery
SPOT creation failure message from **setup_server**.	See “Action 1 - Check SPOT creation” on page 95.
lppsource creation failure message from **setup_server**.	See “Action 2 - Check lppsource creation” on page 95.
Node LED/LCD stops at **611**.	See “Action 3 - Verify exports” on page 95. See “Action 4 - Update SPOT” on page 95.
Node LED/LCD stops at **623**.	See “Action 4 - Update SPOT” on page 95.
setup_server, **allnimres**, **create_krb_files**, **delnimclient**, **delnimmastr**, **spgetdesc** or **unallnimres** fail with an authorization problem.	In “Chapter 10. Diagnosing boot problems” on page 113, see “Action 5 - Check for multiple Boot/Install servers in RRA mode, secure shell mode, or with AIX Authorization for Remote Commands set to ″none″” on page 115.

Actions

Action 1 - Check SPOT creation

The problem is that NIM failed to create a SPOT. Refer to the NIM log files listed in the error message. These files contain a trace of the SPOT creation and will contain any errors that occurred. This error is usually caused by one or more file sets that are required to create a SPOT being missing from the **lppsource**.

Verify that all file sets listed in *PSSP: Installation and Migration Guide* are present in the **lppsource**, and run **setup_server** again. Once the problem is corrected, **setup_server** should proceed normally.

Action 2 - Check lppsource creation

The problem is that NIM failed to create an **lppsource**. This error is usually caused by one or more missing file sets in the **lppsource** directory. Refer to *PSSP: Installation and Migration Guide* and ensure that all file sets listed are present. After correcting the problem, run **setup_server**. Once the problem is corrected, **setup_server** should proceed normally.

Action 3 - Verify exports

This problem is caused by an NFS mount failure. This problem may occur if there is an export or NFS problem on the boot/install server or the control workstation.

1. Verify that the client’s resources are exported on the boot/install server.

   Issue **exportfs** and look for lines for the **spot**, **pssplpp**, and **bosinst_data** directories. If any are not present, issue **setup_server** and look for errors.

   If there are any errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the export still does not exist, see “Information to collect before contacting the IBM Support Center” on page 94 and contact the IBM Support Center.
2. Verify that the **lppsource** is exported from the control workstation.

   If it is not present, issue **setup_server** and look for errors.

   If there are any errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the export still does not exist, see “Information to collect before contacting the IBM Support Center” on page 94 and contact the IBM Support Center.
3. If the exports appear in the **exportfs** output, you may be experiencing an NFS problem. Follow the AIX procedures for diagnosing NFS problems.

Once the problem is corrected, reinstall the node. It should complete the installation.

Action 4 - Update SPOT

This is a Bos install fatal error. This problem is usually caused by a SPOT, **mksysb**, and **lppsource** being out of sync. Use the command:

```
nim -o cust
```

to update the SPOT from the **lppsource** and reinstall the node.

Once the problem is corrected, the node should complete the installation.

Chapter 8. Network installation progress

When a network installation is in progress, the LED/LCD for the nodes involved show various values. These values indicate the installation stage.

Table 9 lists the sequence of LED/LCD values a node goes through. This table is only a sample list of LED/LCD values, and a node may not show all values listed. Depending on the adapters that are installed on the node, there may be additional values displayed. The table also lists the approximate time, after you start to install a node, when the LED/LCD values may be shown. Since the elapsed time to reach a specific LED/LCD value can vary, you should use these times as a gauge to determine the installation progress of the node. The factors that can affect the actual time it takes a node to reach a specific LED/LCD value include the following:

- The number of nodes being installed
- The size of the network installation image you are using
- The amount of traffic on the SP Ethernet
- The amount of work being conducted on the network installation node

Places in the node installation where a single LED/LCD value is displayed for an extended period of time are as follows:

1. c40
2. c54

These values are periods of high volumes of network traffic and are only a problem on a node when the value does not change for a very extended period of time. You can use Table 9 as a debugging tool when a node is stuck on a specific LED/LCD value.

Note: Some nodes have LED/LCD values that are of the form **uxx**, while others have LED/LCD values of the form **0axx**. They are equivalent. If your node has a value of **0axx**, look up the value in the following table as if it was **uxx**.

Table 9. Sample NIM installation trace

Time (min:sec)	LED/LCD value	Description
0:00		
	124	BIST started a CRC check on the OCS area of NVRAM.
	151	BIST started AIPGM test code.
	214	Power status register failed.
	219	Generating RAM POST bit map.
	291	Running standard I/O POST.
	200	Attempting IPL with key in secure position.
	102	BIST started following power-on reset.
	153	BIST started ACLST test code.
0:30		
	154	BIST started AST test code.
	100	BIST completed successfully; control was passed to IPL ROS.
	219	Generating RAM POST bit map.
	292	Running SCSI POST.
1:00		
	291	Running standard I/O POST.
	262	No keyboard connected to the system.
	260	Displaying information on the display console.
1:30		

Table 9. Sample NIM installation trace (continued)

Time (min:sec)	LED/LCD value	Description
	231	Attempting a normal mode IPL from Ethernet specified in IPL ROM.
2:30		
	Blank	
	606	Running the **if_config** command to bring up network interface.
	610	Attempting to NFS mount a remote file system.
	299	IPL ROM passed control to the loaded program code.
	608	Attempting to tftp the **.info** file from client's SPOT server.
	612	Accessing remote configuration files.
	622	Returning control to the **/sbin/rc.boot** program.
	520	Running bus configuration.
	890	SCSI-2 differential fast/wide adapter.
	620	Updating special device files.
3:00		
	890	SCSI-2 differential fast/wide adapter.
	570	Configuring virtual SCSI devices.
3:30		
	622	Returning control to the **/sbin/rc.boot** program
4:00		
	811	Identifying or configuring processor complex.
	Blank	
	570	Configuring virtual SCSI devices.
	727	Identifying or configuring unknown asynchronous device.
	538	The configuration manager is invoking a configuration method.
	622	Returning control to the **/sbin/rc.boot** program.
	c40	Restoring configuration files.
8:00		
	c42	Extracting data from diskette.
	c33	Selecting a tty terminal attached to serial ports S1 or S2.
	c44	Initializing installation database with target disk information.
8:30		
	c46	Normal installation processing.
	Blank	
9:00		
	c50	Creating root volume group on target disks.
	c46	Normal installation processing.
10:30		
	c54	Installing either BOS or additional packages.
19:00		
	c52	Changing from RAM environment to disk environment.
21:30		
	c46	Normal installation processing.
	570	Configuring virtual SCSI devices.
22:00		
	c46	Normal installation processing.
	731	Identifying or configuring PTY.
	539	The configuration method has terminated, returning to config.
	811	Identifying or configuring processor complex.
	538	The configuration manager is invoking a configuration method.

Table 9. Sample NIM installation trace (continued)

Time (min:sec)	LED/LCD value	Description
	570	Configuring virtual SCSI devices.
22:30		
	Blank	
23:00		
	u78	Running remote commands to complete node processing on boot/install server.
	u60	Creating **/etc/ssp/server_name** and updating **/etc/hosts**.
	u68	Copying kerberos realms file from boot/install server.
	u59	Running **config_node** to define adapters.
27:00		
	u65	SSP completing the **install/customize/maint**. Issuing shutdown.
27:30		
	c46	Normal installation processing.
28:00		
	292	Running SCSI POST.
	298	Attempting a software IPL.
	291	Running standard I/O POST.
	299	IPL ROM passed control to the loaded program code.
	Blank	
	890	SCSI-2 differential fast/wide adapter.
	820	
	538	The configuration manager is invoking a configuration method.
	570	Configuring Virtual SCSI devices.
28:30		
29:00		
	551	Running IPL varyon.
	517	Mounting client remote file system during network IPL.
29:30		
	553	IPL phase 1 is complete.
	570	Configuring virtual SCSI devices.
	538	The configuration manager is invoking a configuration method.
	c33	Selecting a tty terminal attached to serial ports S1 or S2.
30:00		
	Blank	
33:00		
	762	Running SSP configuration method for SP Switch Adapter
34:00		
	570	Configuring virtual SCSI devices.
	538	The configuration manager is invoking a configuration method.
34:30		
	Blank	

Chapter 9. Diagnosing node installation problems

This chapter discusses diagnostic procedures and failure responses for the node installation component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 102. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 102.

Related documentation

The following publications provide information about the node installation component of PSSP:

1. *PSSP: Installation and Migration Guide*
2. *PSSP: Messages Reference*
3. *IBM AIX 5L Version 5.1 Network Installation Management Guide and Reference*

Requisite function

This is a list of the software that is used by the node installation component of PSSP. Problems within the requisite software may manifest themselves as error symptoms in the node installation component of PSSP. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the node installation component of PSSP, consider the following components as possible sources of the error.

1. NIM component of AIX - If the error symptoms point to the AIX NIM component and the procedures here do not address the problem, see the AIX NIM documentation for assistance in diagnosing the problem.
2. SP System Security Services

   Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

- NIM Configuration Information

  Issue the **lsnim -l** command from the boot/install server of the failing node and record the output.
- SDR Configuration Information

  Issue the **splstdata -n**, **splstdata -v** and **splstdata -b** commands and record the output.
- Error messages, log files or failures.
- Node LED/LCD indicators

  These are error status codes available on each node. On some nodes, there is a hardware LED/LCD. If a node does not have one, use the SP Hardware Perspective to access the Node notebook to find the LED/LCD code.

  Many LED/LCD’s are displayed as status during the node installation. You should be concerned only when a LED/LCD value is displayed for an extended period of time. Refer to the rest of this section for common LED/LCD failures and to the hardware documentation for additional values. See “Other LED/LCD codes” on page 539.

- Error messages from the **setup_server** command.

  These are errors encountered while configuring the NIM environment. They appear on the boot/install server. This is the stderr output of the **setup_server** command. Any messages that require user action will be stated as errors as opposed to the normal status messages.

Trace information

Post-installation customization trace

This is trace output from **pssp_script** and **psspfb_script**. It is automatically activated by node installation. The output is located in: **/var/adm/SPlogs/sysman/***node name***.config.log.***pid* on the control workstation and **/var/adm/SPlogs/sysman/***node name***.configfb.log.***pid* on the installed node. It contains script debug output (**ksh -x** command).

Examine these files for error messages from commands that are issued by the install process.

Nodecond log

This is the trace output from the **nodecond** command, and is automatically activated when the **nodecond** command is issued. The trace is located in: **/var/adm/SPlogs/spmon/nc/nc.***frame.slot* on the control workstation.

Examine this file for error messages from the **nodecond** command.

Information to collect before contacting the IBM Support Center

Record these items before contacting IBM. For details, see “Error information” on page 101.

1. NIM Configuration Information.
2. SDR Configuration Information .
3. Node LED/LCD indicators.
4. Error messages, log files or other failures.
5. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the node installation component of PSSP. Locate the symptom and perform the action described in the following table.

If you have a symptom that is not in the table, or the recovery action does not correct the problem, see “Information to collect before contacting the IBM Support Center” and contact the IBM Support Center.

Table 10. Node Installation Symptoms

Symptom	Recovery
setup_server failure.	See “Action 1 - Correct NIM environment”.
setup_server fails with an authorization problem.	In “Chapter 10. Diagnosing boot problems” on page 113, see “Action 5 - Check for multiple Boot/Install servers in RRA mode, secure shell mode, or with AIX Authorization for Remote Commands set to ″none″” on page 115.
nodecond command failure.	See “Action 17 - Correct nodecond command” on page 110.
Node LED/LCD stops at **E105**.	See “Action 19 - Correct network problem” on page 111.
Node LED/LCD stops at **231/E1F7**.	See “Action 2 - Correct bootp failure”.
Node LED/LCD stops at **233**.	See “Action 20 - Perform reconfiguration of the node” on page 111.
Node LED/LCD stops at **260**.	See “Action 3 - Send boot image to node” on page 104.
Node LED/LCD stops at **c48**.	See “Action 18 - Correct bosinst_data file” on page 110.
Node LED/LCD stops at **u03**.	See “Action 4 - Send install_info file to node” on page 105.
Node LED/LCD stops at **u57**.	See “Action 5 - Send config_info file” on page 105.
Node LED/LCD stops at **u79**.	See “Action 6 - Send script.cust file” on page 105.
Node LED/LCD stops at **u50**.	See “Action 7 - Send tuning.cust file” on page 106.
Node LED/LCD stops at **u54**.	See “Action 8 - Send spfbcheck file” on page 106.
Node LED/LCD stops at **u56**.	See “Action 9 - Send psspfb_script file to the boot/install server” on page 107.
Node LED/LCD stops at **u58**.	See “Action 10 - Send psspfb_script file to the control workstation” on page 107.
Node LED/LCD stops at **u80**.	See “Action 16 - Send PSSP install images” on page 109.
Node LED/LCD stops at **u87**.	See “Action 15 - Check script.cust file” on page 109.
Node LED/LCD stops at **u62**.	See “Action 11 - Send spsec_overrides file” on page 108.
Node LED/LCD stops at **u67**.	See “Action 12 - Send krb.conf file” on page 108.
Node LED/LCD stops at **u68**.	See “Action 13 - Send krb.realms File” on page 108.
Node LED/LCD stops at **u69**.	See “Action 14 - Send krb-srvtab file” on page 109.
Error message indicating that **firstboot.cust** failed running copy_env_files.	See “Action 21 - Check installation with secure remote command option enabled” on page 111

Actions

Action 1 - Correct NIM environment

This is a problem configuring the NIM environment. Refer to *PSSP: Messages Reference* for the specific errors that are issued by the **setup_server** command. Follow the repair action described.

Once the repair is complete, run **setup_server** and the command should now complete.

Action 2 - Correct bootp failure

This is a node **bootp** failure. Perform the following steps:

1. Verify that the node’s **bootp** response is set to **install** by issuing:

```
splstdata -b -l node_number
```

If the node's **bootp** response is not set properly, reset it using the **spbootins** command and restart the installation.

2. On the node's boot/install server, run **setup_server** and verify that there are no errors. If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there.
3. Verify that there is an entry created in the **/etc/bootptab** file on the node's boot/install server. If there is no entry created after completing these steps, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
4. In a separate window, restart the bootp daemon in foreground mode.
   a. Edit **/etc/inetd.conf** and comment out the entry containing **bootpd** by inserting a ″#″ at the beginning of the line.
   b. Refresh **inetd** by issuing the command: **refresh -s inetd**.
   c. Start the bootp daemon in the foreground by issuing **bootpd -s -d -d -d**. The repeated **-d** flags set the level of trace output.
   d. The bootp daemon will now run and report status and errors to stdout.
5. In a separate window, open a read-only console to the node by using the **s1term** command.
6. Reissue **nodecond** to network boot the node.
7. Observe the output from the **bootpd** command.

Typical problems that may be encountered include:

- Hardware address mismatch.

  The **bootp** daemon reports that it is receiving a request from a hardware address that it does not recognize. This indicates a mismatch between the hardware address in the SDR and the adapter on the node.

  Delete the NIM client using the **delnimclient** command, reacquire the hardware address of the node using the **sphrdwrad** command, run **setup_server** on the boot/install server node, and attempt the installation again.
- No request received

  If the console output from the **s1term** command indicates that the node is sending **bootp** packets, but there is no output from the **bootp** daemon indicating that it is receiving them, you may have a network problem or an adapter failure. Use hardware diagnostics to isolate and repair the problem and perform the installation again.

Once the problem is corrected, reissue the **nodecond** command and the install should proceed normally.

Action 3 - Send boot image to node

The problem is that a **tftp** of boot image to node failed. Perform these steps:

1. On the node's boot/install server, run **setup_server** and verify that there are no errors. If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there.
2. Verify that the boot image exists in the **/tftpboot** directory on the boot/install server. This file is named with the node's long reliable hostname.
3. Verify that the permissions on the file are 644 or greater.
4. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.

5. Verify that the **/etc/tftpaccess.ctl** file on the boot/install server contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, reissue the **nodecond** command. The install should proceed normally.

Action 4 - Send install_info file to node

The problem is that the node failed to **tftp** the **install_info** file. Perform these steps:

1. Verify that the *node name***.install_info** file exists on the node's boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running the command **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 5 - Send config_info file

The problem is that the node failed to **tftp** the **config_info** file. Perform these steps:

1. Verify that the *node_name***.config_info** file exists on the node's boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 6 - Send script.cust file

The problem is that the node failed to **tftp** the **script.cust** file. Perform these steps:

1. Verify that the script.cust file exists on the node’s boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node’s **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

```
allow:/tftpboot
```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 7 - Send tuning.cust file

The problem is that the node failed to **tftp** the **tuning.cust** file. Perform these steps:

1. Verify that the **tuning.cust** file exists on the node’s boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node’s **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

```
allow:/tftpboot
```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 8 - Send spfbcheck file

The problem is that the node failed to **tftp** the **spfbcheck** file. Perform these steps:

1. Verify that the **spfbcheck** file exists on the node’s boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node’s **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.

3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 9 - Send psspfb_script file to the boot/install server

The problem is that the node failed to **tftp** the **psspfb_script** file. Perform these steps:

1. Verify that the **psspfb_script** file exists on the node's boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 10 - Send psspfb_script file to the control workstation

The problem is that the node failed to **tftp** the **psspfb_script** file. Perform these steps:

1. Verify that the **psspfb_script** file exists on the control workstation in the **/tftpboot** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the control workstation contains a line similar to

   ```
   allow:/tftpboot
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 11 - Send spsec_overrides file

The problem is that the node failed to **tftp** the **spsec_overrides** file. Perform these steps:

1. Verify that the **spsec_overrides** file exists on the control workstation in the **/spdata/sys1/spsec** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/spdata/sys1/spsec/
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 12 - Send krb.conf file

The problem is that the node failed to **tftp** the **krb.conf** file. Perform these steps:

1. Verify that the **krb.conf** file exists on the node's boot/install server in the **/etc** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see "Information to collect before contacting the IBM Support Center" on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

   ```
   allow:/etc/krb.conf
   ```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 13 - Send krb.realms File

The problem is that the node failed to **tftp** the **krb.realms** file. Perform these steps:

1. Verify that the **krb.realms** file exists on the node's boot/install server in the **/etc** directory.

   If it is not present, verify that the node's **bootp** response is set to **install**, and run **setup_server**.

If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.

2. Verify that the permissions on the file are 644 or greater.
3. Verify that the entry for the **tftp** daemon is not commented out in the **/etc/inetd.conf** file. If this is the problem, uncomment it and refresh **inetd** by running **refresh -s inetd**.
4. Verify that **/etc/tftpaccess.ctl** on the boot/install server contains a line similar to

```
allow:/etc/krb.realms
```

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 14 - Send krb-srvtab file

The problem is that the node failed to copy the **srvtab** file. Perform these steps:

1. Verify that the *node name***-new-srvtab** file exists on the control workstation in the **/spdata/sys1/k4srvtabs** directory. For nodes running levels of PSSP earlier than PSSP 3.2, this file is on the node’s boot/install server in the **/tftpboot** directory.

   If it is not present, verify that the node’s **bootp** response is set to **install**, and run **setup_server**.

   If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the file still does not exist, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.
2. Verify that the permissions on the file are 644 or greater.
3. Boot the node, and verify that the **/etc/krb-srvtab** file exists.

Once the problem is corrected, the node should move past the failing LED/LCD and continue the installation.

Action 15 - Check script.cust file

The problem is that running **script.cust** is causing a hang condition. Investigate **/tftpboot/script.cust**. This is a user-supplied script. Look for any problems that might cause the node to hang.

Once the problem is corrected, the node can be reinstalled, and the installation should succeed.

Action 16 - Send PSSP install images

The problem is that a PSSP directory failed to mount. Perform these steps to verify that directory is exported on the node’s boot/install server:

1. Obtain the node’s code version by issuing

   ```
   splstdata -b -l node_number
   ```
2. Verify that the directory **/spdata/sys1/install/pssplpp/***code version* exists on the boot/install server and that it contains the PSSP install images.

   If it is not present or is empty, create the directory and place the appropriate PSSP install images in it. Then run **setup_server** and reinstall the node.
3. Issue the **exportfs** command on the boot/install server. In the command output, look for a line similar to **/spdata/sys1/install/pssplpp**, and containing the failing node’s hostname.

If the line is not present, run **setup_server** on the node’s boot/install server.

If there are errors, refer to the entries for the message numbers given in *PSSP: Messages Reference* and follow the instructions there. If the export still does not succeed, see “Information to collect before contacting the IBM Support Center” on page 102 and contact the IBM Support Center.

4. If the export appears in the **exportfs** output, you may be experiencing an NFS problem. Follow the AIX procedures for diagnosing NFS problems.

Once the problem is corrected, reissue the **nodecond** command, and the install should proceed normally.

Action 17 - Correct nodecond command

The problem is that **nodecond** was unable to complete the network boot. Consult the log file listed in the error message to determine the exact cause of the error. Typical errors include:

- The node’s **bootp** response is not set to **install**. Use the **spbootins** command to set the node’s **bootp** response and reissue the **nodecond** command.
- A failure occurred attempting to connect to the node’s serial port. This may be caused by another user having an **s1term** open in write mode. Since there can be only one write mode **s1term** open at a time, **nodecond** fails.

  Close the write mode **s1term** and reissue the **nodecond** command. If **nodecond** is unable to obtain a serial port, and no other write mode **s1term** is open, refer to “Chapter 21. Diagnosing System Monitor problems” on page 317.
- A failure occurred attempting to power on or off the node. Refer to “Chapter 21. Diagnosing System Monitor problems” on page 317.
- A timeout has occurred. This usually indicates a hardware error. To verify this, open a read only **s1term** in another window and reissue the **nodecond** command. Depending on the error, there may be some diagnosis information in the console output. Perform hardware diagnostics to determine the nature of the problem.

After correcting the problem, reissue the **nodecond** command. The install should proceed normally.

Action 18 - Correct bosinst_data file

This is a noprompt installation failure. There is a mismatch in the data provided in the noprompt **bosinst_data** file being used by the node and the actual configuration of the node. The most common mismatch is the specification of a physical disk that does not exist on the node. To determine the nature of the failure, perform the following steps:

1. Open a write console to the node using the **s1term** command.
2. On the console there should be a panel displaying the current choices, and the error that the installation encountered.
3. If the physical disk specified in the SDR is not present on the system:
   a. Use the **spchvgobj** command to update the SDR.
   b. Reissue the **setup_server** command.
   c. Reinstall the node.
4. If the physical disk specified in the SDR is attached to the node, perform hardware diagnostics to determine why the device is not being configured. After correcting the problem, reinstall the node.

Once the problem is corrected, the node should proceed with the install.

Action 19 - Correct network problem

This generally indicates a network problem. It may be caused by a bad Ethernet card or by an Ethernet adapter being set to an incorrect duplex setting. You can verify the duplex setting on your nodes using the **lsattr** command. For example:

```
busio                        Bus I/O address                   False
businter                     Bus interrupt level               False
intr_priority 3              Interrupt priority                False
tx_que_size   64             TRANSMIT queue size               True
rx_que_size   32             RECEIVE queue size                True
full_duplex   no             Full duplex                       True
use_alt_addr  no             Enable ALTERNATE ETHERNET address True
alt_addr      0x000000000000 ALTERNATE ETHERNET address        True
```

Verify that the `full_duplex` setting is correct on all nodes for your particular network environment.

If all the adapters are correctly set, perform node diagnostics to determine if there is a bad adapter card in your system. If so, contact IBM Hardware Support to have it replaced. If none of these measures resolve the problem, record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 102, and contact the IBM Support Center for further assistance.

Action 20 - Perform reconfiguration of the node

This may be caused when a node is replaced with a different type of node without following the procedures documented in *PSSP: Installation and Migration Guide*. This causes an incorrect setting for the platform type in the NIM client object. If this is the case, follow the procedure for replacing a node with a different type of node in ″Reconfiguring the RS/6000 SP System″ of *PSSP: Installation and Migration Guide*. This will cause the NIM client object to be recreated properly.

Action 21 - Check installation with secure remote command option enabled

PSSP 3.4 provides the ability to replace **rsh** and **rcp** calls in the PSSP code with secure remote commands and secure remote copy calls.

The root user must be able to run secure remote commands from the control workstation to the nodes without password or passphrase prompts. This normally means that a root public key generated at the control workstation must have been installed on the nodes and the control workstation. In addition, if the boot/install server node for the node is not the control workstation, the boot/install server node's public key must have also been installed on the node.

If secrshell is enabled, and during installation the **dsh** command fails, you should check that root can issue the secure remote command and secure copy command from the control workstation to the nodes without being prompted for a password or passphrase.

A check should be made that the SDR **SP_Restricted** class, attributes **rcmd_pgm**, **dsh_remote_cmd**, and **remote_copy_cmd** are correct and consistent. The command **splstdata -e** displays the current values of these attributes.

If the system administrator is setting the **$RCMD_PGM**, **$DSH_REMOTE_CMD** and **$REMOTE_COPY_CMD** environment variables to override the setting in the SDR, check these variables to make sure that they are consistent. Use these commands:

```
echo $RCMD_PGM
echo $DSH_REMOTE_CMD
echo $REMOTE_COPY_CMD
```

to check that the remote shell command choice is accurate and consistent with the executable defined by the **$RCMD_PGM**, **$DSH_REMOTE_CMD** and **$REMOTE_COPY_CMD** environment variables.

If the **dsh_remote_cmd** or **remote_copy_command** attributes of the SDR **SP_Restricted** class are null, the remote command and remote copy methods used must be in or linked to the **bin** directory. For example, if **rcmd_pgm=rsh** and **dsh_remote_cmd** and **remote_copy_cmd** are null, then executables **/bin/rsh** and **/bin/rcp** must exist.

If root cannot issue a secure command to the node without being prompted for a password or passphrase, this will cause a secure remote command install of the nodes to fail. Check the following:

- If the root's public key on the control workstation is not installed properly on the node, check that **script.cust** was modified properly to **tftp** the root's public key to the node.
- If the boot/install server node's root public key is not installed properly on the boot/install server node, check that **script.cust** was modified to install the boot/install server node's public key on the node.
- If the secure remote command daemon (**sshd**) is not started, check to see if the daemon is enabled in the **/etc/inittab** file, from **script.cust**.

Chapter 10. Diagnosing boot problems

This chapter discusses diagnostic procedures and failure responses for the boot component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 114. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center”.

Related documentation

The following publications provide information about the boot component of PSSP:

1. *PSSP: Installation and Migration Guide*
2. *PSSP: Messages Reference*
3. *AIX 5L Version 5.1 User’s Guide*
4. Hardware diagnosis guides for individual nodes. See “Bibliography” on page 609.

Error information

Node LED/LCD indicators

The node’s LED/LCD indicator is an error status code. Some nodes have a hardware LED/LCD. In this case, the information can be obtained from the Node notebook of the SP Hardware Perspective.

Many LED/LCDs are displayed as status codes during the node boot process. You should be concerned only when an LED/LCD value is displayed for an extended period of time. Values that are specific to the RS/6000 SP System appear in “Chapter 32. SP-specific LED/LCD values” on page 535. The hardware manuals for your particular node, and the AIX messages contain the remaining LED/LCD values. See “Other LED/LCD codes” on page 539.

Trace information

Console log

This log is located on each node. It logs any error or output that is written to the AIX console. This logging is always active. The file is located in: **/var/adm/SPlogs/sysman/***node_name***.console.log**.

Examine this log for errors from commands issued by the boot process.

Information to collect before contacting the IBM Support Center

The following items are used to isolate problems in the boot component of PSSP. More detailed information about each item appears in “Error information”.

1. **/var/adm/SPlogs/sysman/***node_name***.console.log**.
2. Use the **errpt** command to obtain information on errors since the node was shutdown. For example, to obtain the information for errors that occurred since January 28, 2000 at 1:15 pm, issue:

   ```
   errpt -s 0128131500
   ```

3. The authentication method in use. Issue this command on the control workstation:

```
splstdata -p
```

The entry ″ts_auth_methods″ lists the authentication methods in use.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the boot component of PSSP. Locate the symptom and perform the action described in the following table.

If you have a symptom that is not in the table, or the recovery action does not correct the problem, see “Information to collect before contacting the IBM Support Center” on page 113 and contact the IBM Support Center.

Table 11. Boot symptoms

Symptom	Recovery
Node’s LED/LCD stops with a value.	See “Action 1 - Investigate LED/LCD”.
Error messages are in the **console.log** file.	See “Action 2 - Examine the log file”.
Node never becomes available.	See “Action 3 - Boot a node in maintenance mode” on page 115.
setup_server fails with an authorization problem.	See “Action 5 - Check for multiple Boot/Install servers in RRA mode, secure shell mode, or with AIX Authorization for Remote Commands set to ″none″” on page 115.
Problems with devices, including the hard disk	See “Action 4 - Boot a node in diagnostic mode” on page 115.

Actions

Action 1 - Investigate LED/LCD

This is most likely an AIX software or hardware problem. Refer to the entry for the LED/LCD in to determine the cause of the problem. Values that are specific to the RS/6000 SP System appear in “Chapter 32. SP-specific LED/LCD values” on page 535. The hardware manuals for your particular node, and the AIX messages contain the remaining LED/LCD values. See “Other LED/LCD codes” on page 539.

Follow any diagnostic or repair actions documented for the LED/LCD. Be sure to reinstall the node if the hard disk is replaced. If an Ethernet card or I/O planar is replaced, be sure to follow the procedure in the reconfigration chapter of *PSSP: Installation and Migration Guide* for this repair.

After this problem is resolved, the node’s boot process should proceed past the failing LED/LCD value and the boot should complete.

Action 2 - Examine the log file

This is an error with a command issued as a result of the boot process, or a subsystem invoked by the boot process. Examine the errors in the **console.log**. See “Console log” on page 113. Refer to the appropriate chapter in this book for the failing component. Since errors have a cascading effect, correct the earliest problem first, and then note its effect on later failures.

Once the problem is corrected, reboot the node.

Action 3 - Boot a node in maintenance mode

This is a fatal boot error. Check the **console.log**. See “Console log” on page 113. See the AIX error log for failures. If the node is not accessible via the **telnet** command or a remote command, try the **s1term** command to reach the node.

If **s1term** cannot reach the node, boot the node in maintenance mode by issuing these commands:

1. **spbootins -r maintenance -l** *node_number*
2. **nodecond** *frame_number slot_number*
3. A **s1term** window will be opened.
4. In the **s1term** window, choose these menu options:
   a. **start a limited function maintenance shell**
   b. **mount the root volume group and start a shell**
5. Examine the **console.log** and AIX error log, follow diagnostic procedures for the problem, and correct it.

Once the problem is corrected, reboot the node.

Action 4 - Boot a node in diagnostic mode

You can boot a node from Ethernet network in diagnostic mode when you want to run diagnostics on any device, including the hard disk, attached to that node. When a node is booted in diagnostics mode, it brings up the diagnostics menu, just as if the **diag** command was issued from AIX. But, because the hard disk of the node is not used as the boot device, you can format the hard disk, certify, diagnose, and download microcode.

Caution: Formatting destroys all data on that disk.

To boot a node in diagnostics mode, use the **spbootins** command. For example, to boot node 12 in diagnostics mode, issue:

```
spbootins -r diag -l 12
```

After the **spbootins** command has been issued, the next time the node is network booted, it will boot using the diagnostic image served over the network. The tty console will open on the display, and you will be able to select actions as in the AIX **diag** command. See *AIX 5L Version 5.1 Commands Reference* for a full description of the **diag** command.

A node in Diagnostic mode will NFS mount the Shared Product Object Tree (SPOT) from the boot/install server for use by the diagnostic image on the node. If device support is not present in the SPOT of the boot/install server, the device will not be supported by diagnostics on the node.

Action 5 - Check for multiple Boot/Install servers in RRA mode, secure shell mode, or with AIX Authorization for Remote Commands set to ″none″

Using multiple Boot/Install servers with Restricted Root Access is not recommended and is not automatically supported by PSSP. However, depending on the size of your system and network loads, it may not be possible to install your SP system without a single Boot/Install server.

Boot/Install servers are NIM masters and require remote command access to both the control workstation and the nodes that they serve. PSSP does not automatically create the correct entries in the authorization files to allow the remote commands to function.

To use multiple Boot/Install servers, follow this procedure to manually establish the correct authorizations on your system:

1. On the control workstation, change the authorization files, depending on the setting of the **auth_root_rcmd** attribute:

   standard
   An entry for the Boot/Install server node hostname in the **/.rhosts** file.

   k4 An entry for the Boot/Install server node remote command principal in the **/.klogin** file.

   DCE An entry for the **self-host** and the **spbgroot** principal for the Boot/Install server node.

2. On the Boot/Install server node, edit the **/etc/sysctl.conf** file and include these entries:

   ```
   /usr/lpp/ssp/sysctl/bin/install.cmds
   /usr/lpp/ssp/sysctl/bin/switch.cmds
   /usr/lpp/ssp/samples/sysctl/firstboot.cmds
   ```

   The last entry is included only if you are initiating a node instal customization.

In order for these changes to take effect, you must stop and restart **sysctld** on both the control workstation and all Boot/Install servers:

```
stopsrc -s sysctld
startsrc -s sysctld
```

Chapter 11. Diagnosing Root Volume Group problems

Table 12. Root Volume Group symptoms

Symptom	Recovery action
Creating the root volume group failed during installation because an incorrect disk was specified. The specified disk does not exist.	"Action 1 - Check disks"
Creating the root volume group failed during installation because the disk is in use by another volume group.	"Action 2 - Check disk allocation" on page 118
Mirroring failed because a disk specified in the physical volume list is already in use by another volume group. **Mirroring** is defined in "Root Volume Group terminology" on page 119.	"Action 2 - Check disk allocation" on page 118 "Action 3 - Force the Root Volume Group extension" on page 118
Mirroring failed because the root volume group is locked.	"Action 4 - Unlock the Root Volume Group" on page 118
Mirroring failed because there is insufficient space.	"Action 5 - Add space to physical volumes" on page 118
Mirroring failed because there is insufficient space for strictness. **Strictness** is defined in "Root Volume Group terminology" on page 119.	"Action 6 - Add physical volumes to the Root Volume Group" on page 118
Mirroring failed because an incorrect number of copies was specified.	"Action 7 - Verify the number of copies of AIX on the node for mirroring" on page 118
Unmirroring failed because the volume group is locked. **Unmirroring** is defined in "Root Volume Group terminology" on page 119.	"Action 4 - Unlock the Root Volume Group" on page 118
Unmirroring failed because an incorrect number of copies was specified.	"Action 8 - Verify the number of copies of AIX on the node for unmirroring" on page 119
Unmirroring failed because the **reducevg** command could not remove a physical volume from the root volume group.	"Action 9 - Check for user logical volumes on the physical volume" on page 119
Verification of mirroring or unmirroring is required.	"Action 10 - Verify mirroring or unmirroring" on page 119

Some root volume group terms are defined in "Root Volume Group terminology" on page 119. To understand more about root volume groups and mirroring concepts, see the appendix on Mirroring a Root Volume Group in *PSSP: Administration Guide*.

Actions

Action 1 - Check disks

Check to see which disks were used for installation, by issuing the **splstdata -b** command. If the disks are not valid, change the set of disks by issuing the **spchvgobj** command, and reinstall the node.

Action 2 - Check disk allocation

A physical volume can belong to only one volume group at a time. If you specified a list of disks for installation and one of the disks is in use by another volume group, you must remove the disk from the other volume group. Use the **reducevg** command to remove the disk, and then reinstall the node.

Action 3 - Force the Root Volume Group extension

A physical volume can belong to only one volume group at a time. If you attempt to extend the root volume group with a disk that is in use by another volume group, the **extendvg** command fails. If the disk is part of an inactive volume group, you can force the extension of the root volume group by specifying the **-f** (force) option. Use the **spmirrorvg -f** command.

Action 4 - Unlock the Root Volume Group

If a process terminated and left the volume group in a locked state, do the following:

1. Unlock the root volume group by issuing the **chvg -r** command.
2. Rerun the desired mirroring function, using the **spmirrorvg** command if mirroring is desired or the **spunmirrorvg** command if unmirroring is desired.

Action 5 - Add space to physical volumes

When mirroring, there must be enough total space on the additional physical volumes to contain all of AIX's logical volumes and still maintain strictness. Strictness is defined in "Root Volume Group terminology" on page 119. If there is not enough space in the additional physical volumes, add additional physical volumes to the root volume group by issuing the **spchvgobj** command. Then rerun mirroring using the **spmirrorvg** command.

Action 6 - Add physical volumes to the Root Volume Group

For each copy of AIX, you must have at least one physical volume in the root volume group. For example, if you have specified three copies of the root volume group (the original and two copies), you must have at least three physical volumes in the root volume group. If you have fewer physical volumes than copies, add additional disks to the physical volume list by issuing the **spchvgobj** command. Then, rerun mirroring by issuing the **spmirrorvg** command.

Action 7 - Verify the number of copies of AIX on the node for mirroring

To mirror successfully, you must specify more copies of AIX than are currently in effect on the node. For example, if there is one copy of the root volume group in effect on the node, you must specify two or three as the desired number of copies for mirroring. If there are two copies of the root volume group in effect, you must specify three copies for mirroring. If you specify the same number or fewer copies than are currently in effect on the node, mirroring has no effect.

Determine how many copies of the root volume group are currently in effect for the node and correct this number with these commands:

1. Use the **splstdata** command to find out how many copies of the root volume group are in effect for the node.
2. Use the **spchvgobj** command to change the number of desired copies.
3. Rerun the **spmirrorvg** command.

Action 8 - Verify the number of copies of AIX on the node for unmirroring

To unmirror successfully, you must specify fewer copies of AIX than are currently in effect on the node. For example, if there are three copies of the root volume group in effect on the node, you must specify one or two as the desired number of copies for unmirroring. If there are two copies of the root volume group in effect, you must specify one for unmirroring. If you specify the same number or more copies than are currently in effect on the node, unmirroring has no effect.

Determine how many copies of the root volume group are currently in effect for the node and correct this number with these commands:

1. Use the **splstdata -v** command to find out how many copies of the root volume group are in effect for the node.
2. Use the **spchvgobj** command to change the number of desired copies.
3. Rerun the **spunmirrorvg** command.

Action 9 - Check for user logical volumes on the physical volume

The root volume group may not be reduced by a physical volume unless all the logical volumes have been removed from the physical volume. During the unmirroring operation, all the AIX logical volumes are removed from the mirror's physical volumes. However, if additional (user) logical volumes were created on the physical volume, you cannot reduce the root volume group by the physical volume until all user logical volumes are moved or deleted.

Action 10 - Verify mirroring or unmirroring

Use this table to perform verification tasks for the root volume group.

Table 13. Verification of mirroring or unmirroring for Root Volume Groups

Verification task	Commands
Display the number of copies of the root volume group on the node.	Use the **lslv** command to show the number of copies of a logical volume in the root volume group. In AIX, mirroring is done on a logical volume basis. For example, **lslv hd2** will show the number of copies of the hd2 logical volume, which is used for the **/usr** file system.
Display the set of physical volumes in the root volume group.	Use the **lspv** command.
Display the boot list.	Use the AIX **diag** command, select the ″Task Selection″ option, and display the normal mode boot list.
Display the state of the quorum. **Quorum** is defined in "Root Volume Group terminology".	Use the **lsvg rootvg** command to list information about the root volume group. If the Quorum attribute is equal to 1, quorum is off. If the Quorum attribute is equal to anything else, quorum is on.

Root Volume Group terminology

Some root volume group terms are defined here:

mirroring. Mirroring provides redundant copies of AIX to prevent single points of failure. AIX provides for the original copy of AIX, and one or two additional copies. AIX places the data on the physical volumes in such a way that no two copies of the same data are ever on the same physical volume. This provides the redundancy necessary so that a single physical volume failure does not cause its volume group to fail.

quorum. A vote of the number of Volume Group Descriptor Areas (VGDAs) and Volume Group Status Areas (VGSAs) that are active. For a volume group of one disk, there are two VGDA/VGSAs. For a volume group of two disks, there are two VGDA/VGSAs on the first disk and one on the second disk. For a volume group of three or more disks, there is one VGDA/VGSA per disk.

A quorum ensures data integrity in the event of a disk failure. When a majority (51 %) of the VGDA/VGSAs in the volume group cannot be accessed, the group varies itself offline to prevent data loss or incorrect I/O operations. An error log entry is produced when this situation occurs.

strictness. A rule that AIX applies to the disks in a mirrored volume group. The strictness rule requires that enough physical volumes with sufficient space are part of a volume group, so that AIX can allocate the data according to a particular algorithm. This algorithm allocates data among the physical volumes in such a way that the loss of a single physical volume does not constitute the loss of the volume group. Refer to the definition of mirroring and unmirroring.

unmirroring. Unmirroring reduces the number of copies of a root volume group. For example, if there are currently three copies of the root volume group (the original and two copies), unmirroring would be used to reduce the number of copies from three to two or one. Optionally, physical volumes used for mirroring may be removed from the root volume group after unmirroring. When physical volumes are removed from the root volume group, they are made available for use by other volume groups.

Chapter 12. Diagnosing system connectivity problems

If a node becomes unresponsive or inaccessible, use the following table to diagnose the problem.

Table 14. System connectivity symptoms

Symptom	Recovery
Using the SP Hardware Perspective, bringing up **hostResponds** in a table view for nodes and multiple nodes shows: **Node Not Responding.** Using the SP Hardware Perspective, bringing up **switchResponds** in a table view for nodes and multiple nodes shows: **Adapter Not Configured.** The **spmon -G -d** command shows **hostResponds** as **no**. The **spmon -G -d** command shows **switchResponds** as **no**.	"Action 1 - Diagnose multiple nodes"
Using the SP Hardware Perspective, either the table view or nodes status page of the notebook shows: **Node Not Responding** for **hostResponds** or **Adapter Not Configured** for **switchResponds**. The **spmon -G -d** command shows **hostResponds** as **no** The **spmon -G -d** command shows **switchResponds** as **no**.	"Action 2 - Diagnose individual nodes"
Cannot access the node using a remote command or: **telnet, rlogin** or **ping**.	"Action 3 - Diagnose a network problem" on page 122
Cannot access the node using the **telnet or rlogin,** commands, but can access the node using the **ping** command.	The is a probable software error. See "Chapter 5. Producing a system dump" on page 81 to initiate a dump, record all relevant information and contact the IBM Support Center.
Can access the node using **telnet** or **ping**, but **hostResponds** still shows: **Node Not Responding.**	"Action 4 - Diagnose a Topology-related problem" on page 122

Actions

Action 1 - Diagnose multiple nodes

If several node icons in a frame report a failure, (either the nodes are not responding or several adapters are inactive) there may be a network problem.

If the failing nodes or communication adapters are on the same Local Area Network (LAN), verify the LAN hardware. If you determine that the hardware is functioning properly, call the IBM Support Center. Otherwise, follow local procedures for servicing your hardware.

If the nodes are not on the same LAN, diagnose the nodes individually as described in "Action 2 - Diagnose individual nodes".

Action 2 - Diagnose individual nodes

If an individual node icon in a frame reports a failure, use the SP Hardware Perspective to display the Nodes Status page in the Node notebook, for the failing node.

1. Check the node's LCD/LED indicator.

2. If a three-digit code is displayed, check “Chapter 32. SP-specific LED/LCD values” on page 535 to see if the code is described there. If the code is not described in this section, refer to *IBM RS/6000 Problem Solving Guide*.
3. Check the **hostResponds** indicator for a failure.
4. Check the node’s power indicator.

   If it shows that the node power is off, turn the node’s power on.

   If it shows that the node power is on or if the problem persists, call IBM hardware support.

Action 3 - Diagnose a network problem

If a node is not responding to a network command, you can access the node by using the tty. This can be done by using the SP Hardware Perspective, selecting the node and performing an **open tty** action on it. It can also be done by issuing the

```
s1term -w frame number slot number
```

command, where *frame number* is the frame number of the node and *slot number* is the slot number of the node.

Using either method, you can login to the node and check the hostname, network interfaces, network routes, and hostname resolution to determine why the node is not responding. The Appendix ″IP Address and Host Name Changes for SP Systems″ in *PSSP: Administration Guide* contains a procedure for changing hostnames and IP addresses.

If you are using IPv6 alias addresses, verify that each network (Ethernet or token ring) adapter on the affected system has a valid IPv4 address defined. To verify the adapter IP addresses on the control workstation and nodes, run the **SYSMAN_test** command. This command issues an error message if the node does not have valid IPv4 addresses for all Ethernet and token ring adapters that are used by the SP system. See “Chapter 29. Verifying System Management installation” on page 507.

Action 4 - Diagnose a Topology-related problem

If the **ping** and **telnet** commands are successful, but **hostResponds** still shows **Node Not Responding**, there may be something wrong with the Topology Services **(hats)** subsystem. Perform these steps:

1. Examine the en0 (Ethernet adapter) and css0 (switch adapter) addresses on all nodes to see if they match the addresses in **/var/ha/run/hats.***partition_name***/machines.lst**.
2. Verify that the netmask and broadcast addresses are consistent across all nodes. Use the **ifconfig en0** and **ifconfig css0** commands.
3. Examine the **hats** log file on the failing node. It is named: **/var/ha/log/hats.***dd.HHMMSS.partition_name*, where *dd.HHMMSS* is the day of the month and time of day when the Topology Services daemon was started, and *partition_name* is the name of the node’s system partition.
4. Examine the **hats** log file for the Group Leader nodes. Group Leader nodes are those that host the adapter whose address is listed below the line ″Group ID″ in the output of the **lssrc -ls hats** command. For more information, see “Chapter 23. Diagnosing Topology Services problems” on page 351, and the Topology Services chapter in *PSSP: Administration Guide*.

Chapter 13. Diagnosing IP routing problems

Ensure that the routing tables on the nodes accurately reflect the network. If they do, commands specified in **script.cust** should run correctly with all permissions and routes available.

Issue the following AIX commands to check the routing:

- To display network statistics and routing, enter:

  ```
  netstat -nr
  ```

- To see the routing to a specific point, enter:

  ```
  traceroute IP_address | host_name
  ```

IP source routing

Note that the **hats** script, which controls the operation of the Topology Services subsystem, issues the **no -o nonlocsrcroute=1** command. This command enables IP source routing. **DO NOT** change this setting, because the Topology Services subsystem requires this setting to work properly. If you change the setting, the Topology Services subsystems and a number of other subsystems that depend on it will no longer operate properly.

Chapter 14. Diagnosing SDR problems

This chapter discusses diagnostic procedures and failure responses for the SDR (System Data Repository) component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 131. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 128.

Related documentation

The following publications provide information about the SDR:

1. *PSSP: Command and Technical Reference*

   The commands that read and write to the SDR all have a prefix of **SDR**. Several other commands, such as **splstdata**, **splstnodes** and **spget_syspar** also read configuration data from the SDR.

   Information on the **SDR_dest_info** file is in the chapter ″Files and Other Technical Information″.
2. *PSSP: Administration Guide*

   The Appendix ″The System Data Repository″ gives an overview of the SDR and lists the classes and attributes that are stored in the SDR.
3. *PSSP: Messages Reference*

   SDR error messages are in the chapter ″0025 - System Data Repository Messages″.
4. *PSSP: Installation and Migration Guide*

   The SDR is installed in the process described in ″Task B. Install PSSP on the Control Workstation″. Step 22 describes how to run the SDR Verification Test. The SDR configuration data is then completed by scripts that run in the remaining installation steps.

Requisite function

This is a list of the software and operating system resources directly used by the SDR component of PSSP. Problems within the requisite software or resources may manifest themselves as error symptoms in the SDR. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the SDR component of PSSP, you should consider the following components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- AIX TCP/IP
- AIX catalog functions
- AIX SRC (system resource controller)
- **/spdata** file system
- **/var** file system
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

The SDR has many sources of error information:

1. From commands:

   The return codes from SDR commands correspond to the error message that the command displays. The SDR catalog is in **/usr/lib/nls/msg/***lang***/sdr.cat** where *lang* is the current AIX locale (language setting).

   To find the AIX locale for your node, issue the **locale** command on that node and find the line for LC_MESSAGES.

2. From the **sdrd** daemon:

   a. SDR Error Log

      There is one **sdrd** daemon for each system partition. Each **sdrd** has its own log file in the following location: **/var/adm/SPlogs/sdr/sdrdlog.***ipaddress.sdrdpid*, where *ipaddress* is the alias IP address of the system partition, and *sdrdpid* is the process id of that particular invocation of **sdrd**.

      Each time the **sdrd** is restarted, it creates a new log file. The **sdrd** logs are generally not large unless tracing is turned on, since tracing uses the same log file. There are often no errors logged, and only informational messages about the **sdrd** initialization are present.

      The oldest **sdrd** log is removed when **/usr/lpp/ssp/bin/cleanup.logs.ws** is run, if the log does not belong to a currently running **sdrd** and the log is older than one week old.

   b. SDR Error Messages in AIX Error Log

      The **sdrd** writes messages to the AIX error log with the error label **SDR_EMSG100_ER** under resource name **sdrd**. Refer to the Detail Data section of the logged error for information on the specific error that occurred. This information is in English. The types of problems that are recorded in the AIX error log are problems with: TCP/IP sockets, memory allocation, file system that are full, SP System Security Services initialization, and data inconsistencies.

   c. **SDR_config.log** file

      The **SDR_init** script is run to initialize the SDR's classes and attributes for the SP system. **SDR_config** creates the objects for the existing hardware. If they are invoked with the **-l** flag, both **SDR_init** and **SDR_config** write errors to a log in **/var/adm/SPlogs/sdr/SDR_config.log**.

      SDR_init is called when **install_cw** is run during PSSP installation. **SDR_config** is called by **splogd** when the hardware monitor reports a state change for the ″type″ hardware variable. For more information on **SDR_init** and **SDR_config**, see *PSSP: Command and Technical Reference*.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

SDR daemon trace

The **sdrd** daemon trace is intended for IBM Support Center personnel only. It is not intended for general customer use. When you work with IBM Service on a problem, you may be asked to turn this trace on to collect more information.

Trace information goes into the daemon error log in **/var/adm/SPlogs/sdr**. Since a large amount of information is logged, it may fill up the **/var** file system. Therefore, it is not recommended that the trace facility be left on for an extended period of time.

To turn on and off the **sdrd** trace, find the PID (process id) of the **sdrd** that you want to trace. Sending a hangup signal (**SIGHUP**) to the process will cause the daemon to toggle the trace setting. If the trace is on when the daemon receives the **SIGHUP**, the trace will be turned off. If the trace is off when the daemon receives the **SIGHUP**, the trace will be turned on.

To toggle the trace setting, perform these steps:

1. `lssrc -g sdr`

   This finds the PIDs of all of the **sdrd** daemons on the system.
2. `kill -HUP` *sdrdpid*, where *sdrdpid* is the process id of the **sdrd** that you are interested in tracing.

   Sending the **SIGHUP** signal toggles tracing on or off.

Before doing this, you must stop the currently running **sdrd** for that system partition:

```
stopsrc -s sdr.partition_name
```

where *partition_name* is the short hostname of the system partition of the **sdrd**.

If you want tracing to be on when the **sdrd** is started, use the **d** flag. Issue this command:

```
/usr/lpp/ssp/bin/sdrd partition_ip_address d
```

where *partition_ip_address* is the IP address being used for the system partition of the **sdrd**.

Note: The **d** flag here is **not** preceded by a **-** (minus sign).

Rather than start the **sdrd** in the foreground, you may use the **-a** flag on the **startsrc** command to pass the **d** flag. Issue this command:

```
startsrc -s sdr.partition_name -a d
```

Note: The **d** flag is not preceded by a **-** (minus sign).

Trace log entries consist of free form text messages.

Information to collect before contacting the IBM Support Center

Collect the following items if they are relevant to your problem:

1. If there is a core dump from an **sdrd**, it is located in the **/var/adm/SPlogs/sdr** directory, with a name of **core**. Search for and save this file.
2. Save the **sdrd** logs from **/var/adm/SPlogs/sdr**.
3. If there are **sdrd** messages in the AIX error log, capture them with this command:

   ```
   LANG=C errpt -aN sdrd > /tmp/sdrd.AIXerrlog
   ```
4. If the problem can be isolated to a specific command invocation, provide that information to the IBM Support Center.
5. The trusted services authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.
6. Information from SP Security Services. See “Information to collect before contacting the IBM Support Center” on page 258.

Diagnostic procedures

Use these procedures to determine the cause of an SDR failure.

Check current system partition

To find out which system partition you are currently in, issue the following command:

```
echo $SP_NAME
```

If **SP_NAME** is set, SDR commands are directed to the **sdrd** daemon for the hostname or IP address that **SP_NAME** represents. Make sure that your **SP_NAME** is not set to an unexpected partition. If **SP_NAME** is not set, SDR commands are directed to the primary partition that is defined in the **/etc/SDR_dest_info** file. See “Action 2 - Analyze system or network changes” on page 132 for the expected format of the **/etc/SDR_dest_info** file.

Query the state of the sdrd

The **sdrd** daemons are under the AIX SRC control, and therefore they can be queried using the **lssrc -g** command. There is one daemon for each system partition, and all of the daemons are in the **sdr** group. To query a single system partition, issue:

```
lssrc -s sdr.partition_name
```

where *partition_name* is the short hostname for the **sdrd** being queried.

Stop and start the sdrd

To stop or start all of the **sdrd** daemons, use the **-g** flag on the **stopsrc** or **startsrc** commands. This targets the whole **sdr** group. For example:

```
stopsrc -g sdr
```

Sometimes the **sdrd** does not stop right away. The **lssrc** command can be used to make sure that an **sdrd** daemon has stopped. If the **sdrd** daemon has not stopped, the **kill -9** command can be used to bring it down immediately.

The **sdr** command can also be used to start and stop the SDR. Issuing the command:

```
/usr/lpp/ssp/bin/sdr reset
```

stops and starts the SDR in the current partition. Using the **-spname** option, a specific partition's **sdrd** can be stopped, started, or reset.

Check the sdrd processes using the ps command

To see which **sdrd** daemons are running in the same system partition or an incorrect system partition name, issue the command:

```
ps -ef | grep sdrd
```

This can be used to find multiple **sdrd** daemons that are running in the same system partition. This situation is an error.

Check for sdrd memory leaks and CPU utilization

To see if there are **sdrd** memory leaks, issue these commands:

```
ps gvc | grep PID
ps gvc | grep sdrd
```

Issue these commands at regular intervals to check for a memory leak.

The `SIZE` and `RSS` fields indicate memory usage. They should not grow larger constantly over time. Be aware that the first time a class is accessed, it is cached in the **sdrd**, causing memory usage to increase. Also, a larger SP system has bigger **sdr** classes and therefore an **sdrd** that requires more memory.

The following sequence may be entered at the command line or put into a script to track **sdrd** CPU usage:

```
 ps gvc | grep PID > /tmp/sdrd.psgvc
 while true
  do
  ps gvc | grep sdrd >> /tmp/sdrd.psgvc
  sleep 60
  done
```

The `%CPU` field indicates how much CPU the **sdrd** is consuming. This will increase when clients are making many requests of the SDR. Typically, this peaks when nodes are booted and when IBM Virtual Shared Disks and GPFS are initialized.

Check for an sdrd hang

A command that connects to the **sdrd** of each system partition to retrieve system partition information is:

```
splstdata -p
```

If this command hangs when trying to connect to a particular system partition, the **sdrd** of that partition is most likely hung.

If this happens, use the **dbx** command to search for more information about the hang:

```
dbx -a sdrdpid
```

where *sdrdpid* is the process id of the **sdrd** that is hung.

Within **dbx**, the following subcommands are used to gather more information:

- **where** - to locate where (what routine name and memory location) the **sdrd** is stopped.
- **thread** - to display thread information
- If the **sdrd** created multiple threads, issue these subcommands:
  1. **thread** - to display the thread stack
  2. **thread current** *n* - to display the threads, for each thread number in the thread stack.
  3. **where** - to locate where (what routine name and memory location) the **sdrd** for the thread is stopped.

Check for an sdrd core dump

If there is a core dump from the **sdrd**, it is located in the **/var/adm/SPlogs/sdr** directory. To force a core dump, issue:

```
kill -ABRT sdrdpid
```

where *sdrdpid* is the process id of the **sdrd** that is to be dumped. Copy the **core** file to a safe place.

SDR verification test

The SDR verification test, **SDR_test**, runs various SDR commands to make sure that objects and files can be created, updated, queried and deleted. See *PSSP: Command and Technical Reference* for command details. SDR administrator access is required to run this command. To run the command, issue:

```
/usr/lpp/ssp/bin/SDR_test
```

To run the command via **SMIT**, do the following:

1. Issue the command: **smit SP_verify**
2. Select **System Data Repository**

Command results are logged in **/var/adm/SPlogs/SDR_test.log** as well as being sent to stdout and stderr, unless **quiet** mode is selected. If the command is issued by a user other than **root**, results are logged in **/tmp/SDR_test.log**.

Good results indicate that verification succeeded, as in this example:

```
SDR_test: Start SDR command line verification test
SDR_test: Verification succeeded
```

Error results are indicated by **SDR_test** error messages. One messages is issued for each failing test. There is also an error message from the specific SDR command that failed.

At the end of the test, there is a message with the number of tests that failed, such as:

```
SDR_test: 0037-213 Verification failed with 10 errors.
 See /var/adm/SPlogs/SDR_test.log
```

Look for MBCS data in the SDR

With National Language Support, two new fields have been added to the **SP** class in the SDR. These fields control whether MBCS (Multi-Byte Character Set) data may be written to the SDR. These fields are: **SDR_ASCII_only** and **admin_locale**.

- If **SDR_ASCII_only** is **true**, only data that is in the ASCII character range may be written to the SDR.
- If **SDR_ASCII_only** is **false**, data within the **admin_locale** language set as well as ASCII data may be written to the SDR.

The **SDRScan** routine look through the SDR classes and files and return a return code of 1, if any classes and files in the SDR have MBCS data. The file names and class and attribute names where the MBCS data is found are displayed by the command, unless the **quiet** option is chosen. MBCS data is only a problem if **SDR_ASCII_only** is **true**, or if the data is in a language other than the **admin_locale** language.

SDRScan can be run only by **root** on the control workstation. Issue this command:

```
SDRScan
```

You will receive messages for any non-ASCII data in the SDR. If there is no non-ASCII data, no messages are issued, and the return code will be zero. If you are running **ksh**, issue the command:

```
echo $?
```

to see the return code.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the SDR component of PSSP. Locate the symptom and perform the action described in the following table.

Table 15. System Data Repository (SDR) symptoms

Symptom	Recovery
Nonzero return code	See “Action 1 - Get the return code” on page 132.
Cannot connect to server	See “Action 2 - Analyze system or network changes” on page 132.
SDR class corrupted or missing	See “Action 3 - Analyze class situation” on page 134.
Cannot write to the SDR	See “Action 4 - Check authorization” on page 134.

Table 15. System Data Repository (SDR) symptoms (continued)

Symptom	Recovery
Error code 005 - Write authority required.	See “Action 4 - Check authorization” on page 134.
Error code 006 - Administrator authority required.	See “Action 4 - Check authorization” on page 134.

Actions

Action 1 - Get the return code

If you cannot run SDR commands, or a program that uses the SDR is failing when running SDR commands, get the return code or the message number from the failing SDR routine. The return codes from SDR routines are imbedded in the message numbers. The first four numbers in the SDR cataloged message are always **0025**, followed by a hyphen and a three-digit number. The three digit number is the return code. For example, the following SDR message is issued with a return code of 080 from any SDR routine that cannot connect to the SDR server:

```
0025-080 The SDR routine could not connect to server.
```

Some programs report the return code from an SDR routine, but not the message. Use **0025** and the return code to find the appropriate message in *PSSP: Messages Reference*. Follow the action for the particular error message to correct the error.

Once you have correct the problem, rerun the command that produced the error to verify that it is corrected.

Action 2 - Analyze system or network changes

System or network changes could affect the SDR. If an SDR command fails to connect to the server, do the following:

1. Issue the **spget_syspar** command on the node where SDR commands are failing.
2. If the **spget_syspar** command fails, check the **/etc/SDR_dest_info** file on the same node. It should have four records in it. These records are the **primary** and the **default** hostname and IP addresses. They should be similar to:

   ```
   default:default_syspar_ip_address
   primary:syspar_ip_address
   nameofdefault:default_hostname
   nameofprimary:syspar_hostnameprimary
   ```

   where

   - *default_syspar_ip_address* is the address of the default system partition.
   - *syspar_ip_address* is the address of the system partition that this node is in, or the default system partition if this file is on the control workstation.
   - *default_hostname* is the hostname of the default system partition.
   - *syspar_hostnameprimary* is the hostname of the partition that contains this node.

   Note: The default system partition may be the same as the primary system partition.

If this file is missing or does not have these four records, the node may not be properly installed, or the file may have been altered or corrupted. You can edit this file to correct it, or copy the file from a working node in the same system partition.

The **spget_syspar** command may also fail if:

- The value of the **SP_NAME** environment variable is a hostname (not an IP address)

  AND
- The system nameserver is not functioning properly.

3. If the **spget_syspar** command is successful, check to make sure that the address is also the address of a valid system partition. If it is, try to **ping** that address. Issue this command:

   ```
   ping -c 1 IP_address
   ```

   If the **ping** is successful, the output is similar to:

   ```
   PING 9.114.61.129: (9.114.61.129): 56 data bytes
   64 bytes from 9.114.61.129: icmp_seq=0 ttl=255 time=0 ms

   ----9.114.61.129 PING Statistics----
   1 packets transmitted, 1 packets received, 0% packet loss
   round-trip min/avg/max = 0/0/0 ms
   ```

   If the **ping** fails, output is similar to:

   ```
   PING 9.114.61.129: (9.114.61.129): 56 data bytes

   ----9.114.61.129 PING Statistics----
   1 packets transmitted, 0 packets received, 100% packet loss
   ```

   In this case, contact your system administrator to investigate a network problem.
4. If the value returned by the **spget_syspar** command is not the same as the address in the **primary** record of the **/etc/SDR_dest** information file, the **SP_NAME** environment variable is directing SDR requests to a different address. Make sure that this address (the value of the **SP_NAME** environment variable) is a valid system partition.
5. If the value of the **SP_NAME** environment variable is a hostname, try setting it to the equivalent dotted decimal IP address. If SDR commands now work, the system nameserver is not functioning.
6. If the address returned by **spget_syspar** is a valid system partition address and **pings** to that address are successful, check for the existence of the SDR server process (**sdrd**) on the control workstation with:

   ```
   ps -ae | grep sdrd
   ```

   If the process (**sdrd**) is not running, do the following:

   a. Check the **/var/adm/SPlogs/sdr** directory for a core dump. If one exists, see “Information to collect before contacting the IBM Support Center” on page 128 and contact the IBM Support Center.

   b. Check the SDR server logs in **/var/adm/SPlogs/sdr/sdrdlog.***ipaddr.pid*, where *ipaddr* is the IP address of the system partition and *pid* is a process ID.

   c. Issue the command:

```
/usr/bin/startsrc -g sdr
```

to start the SDR daemon. Start checks again at Step 5 on page 133. If the SDR daemon is now running and continues to run, check the **sdrd** entry in the file **/etc/inittab** on the control workstation. It should read:

```
sdrd:2:once:/usr/bin/startsrc -g sdr
```

Issue an SDR command again to see if it now connects to the server.

Action 3 - Analyze class situation

If an SDR command ends with RC=102 (internal data format inconsistency) or 026 (class does not exist), first make sure that the class name is spelled correctly and that the case is correct. See the table of classes and attributes in “The System Data Repository” appendix in *PSSP: Administration Guide*. Then, follow the steps in “SDR Shadow Files” in the System Data Repository appendix in the *PSSP: Administration Guide*.

This condition could be caused by the **/var** file system filling up. If this is the case, either define more space for **/var** or remove unnecessary files.

If the problem persists, contact the IBM Support Center.

Once you have corrected the problem, rerun the command that produced the error to verify that it is corrected.

Action 4 - Check authorization

The trusted services authentication methods for a system partition determine the rules used by that system partition’s **sdrd** to permit write and administrator access to the SDR. SDR administrator access is required for commands that change class definitions or create and delete files from the SDR. Write access is required for commands that add objects, change attributes and replace files in the SDR.

If the trusted services authentication methods are set to DCE only, appropriate credentials are needed to be able to issue write or administrator commands to the SDR. If the trusted services authentication methods are set to **dce:compat**, **compat**, or anything else, only the **root** user on the control workstation or an SP node in the **sdrd**’s system partition can issue write or administrator commands to that system partition’s SDR. For more information on authentication, see “The System Data Repository” appendix in *PSSP: Administration Guide*.

If an SDR command fails to write to the SDR, perform these steps:

1. Find out what trusted services authentication methods are in your system partition by issuing the command:

   ```
   lsauthpts
   ```
2. Make sure that you are in the system partition you expect, by issuing the command:

   ```
   spget_syspar -n
   ```

   If not, check the **SP_NAME** environment variable to see if it is set to connect to an unexpected system partition. If **SP_NAME** is not set, check the **/etc/SDR_dest_info** file for correctness. To correct the **SDR_dest_info** file, see “Action 2 - Analyze system or network changes” on page 132, Step 2 on page 132.

3. If you are connecting to the expected system partition's **sdrd**, perform the following actions based on the value of the trusted services authentication methods for the system partition:

   a. If the trusted services authentication methods are set to DCE only:

      Issue the **klist** command to see if you have DCE credentials. If so, you can see which SDR groups you belong to.

      If you do not belong to a group with **sdr** and **write** in the name, you cannot write to the SDR. If you do not belong to a group with **sdr** and **admin** in the name, you cannot issue SDR administrator access commands.

      If this is the case, ask your security administrator to add you to the appropriate **sdr** groups. As an alternative, **dce_login** to a principal that is in the appropriate **sdr** groups. If you have no credentials, **dce_login** to a principal in the appropriate sdr groups.

      Note: The group names may be overridden, however, in the **/spdata/sys1/spsec/spsec_overrides** file.

      Note that there are separate groups for access to system classes which are global to all system partitions, and for access to partition-sensitive classes. The objects of partition-sensitive classes may be written only from that system partition's **sdrd**. Groups for system classes have **system-class** in their name, unless the name was overridden in the **/spdata/sys1/spsec/spsec_overrides** file. Groups for partition-sensitive classes do not have **system-class** in their name.

      Partition-sensitive groups can also be partitioned. If the group has a **:p** appended to it in the **spdata/sys1/spsec/spsec_overrides** file, there will be a separate group for each partition, with its own access list.

      Note: The SDR administrator access authority includes write authority.

      To determine if there is a problem with DCE, see "Chapter 18. Diagnosing SP Security Services problems" on page 251.

   b. If the trusted services authentication methods are set to anything other than DCE only (The **lsauthpts** command returns anything other than DCE):

      The **root** user on the control workstation or **root** on a node in the SDR's system partition is allowed to perform write or administrator commands to the SDR.

      Issue the **whoami** command to make sure that you are running as **root**. If you are not **root**, and DCE is an option for your system partition, follow the actions in Step 3a.

      If you are **root**, perform these steps:

      1) To find out the hostname of the node you are running on, issue the command:

         ```
         hostname
         ```

      2) To find out the IP address for your hostname, issue the command:

         ```
         host hostname
         ```

         where *hostname* is the hostname found in the previous step.

      3) Issue the command:

```
SDRGetObjects Adapter netaddr==ipaddress
```

where *ipaddress* is the address found in the previous step.

If there is no object for the IP address you entered, the SDR will not recognize your node as being in its system partition. Also, the SDR will not allow **root** on the node to perform write or administrator SDR commands. Possible causes are: the adapters were not set up correctly during installation, and the network is not set up correctly on the node.

To see how routing is set up on a node, see “Chapter 13. Diagnosing IP routing problems” on page 123. For information on how to add adapters, see *PSSP: Installation and Migration Guide*.

Also, if there is an adapter on the node that cannot be defined in the **Adapter** class of the SDR, but commands from the node are routed across that adapter, the **sdrd** will not recognize the command as coming from one of its nodes.

In this case, a static route may be added from the node to the control workstation by using the **smit fastpath mkroute**. The destination address should be the IP address that represents the system partition of the node. The gateway IP address should be for an adapter that is defined in the SDR. Another possible workaround for this situation is to define the IP address of the unsupported adapter as a supported adapter type.

Once you have corrected the problem, check that the SDR can be written to by running the **SDR_test** command.

Chapter 15. Diagnosing SP Switch problems

This chapter discusses diagnostic procedures and failure responses for the SP Switch component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 177. All the recovery actions described require that the user have **root** access to the specified node. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 164.

Related documentation

The following publications provide information about the SP Switch:

1. *PSSP: Administration Guide*
2. *PSSP: Planning, Volume 2*
3. *PSSP: Installation and Migration Guide*
4. *PSSP: Messages Reference*

   These chapters contain messages related to the SP Switch:

   - 0028 - Switch Support Messages
   - 2510 - Switch Fault Service Daemon Messages
   - 2543 - Switch Admin Daemon (cssadm daemon) Messages
   - 2548 - SP Switch Advanced Diagnostic Messages
   - 2549 - SP Switch Advanced Diagnostic Messages
5. *PSSP: Command and Technical Reference*
6. *First Failure Data Capture Programming Guide and Reference*
7. *The RS/6000 SP Inside Out*, SG24-5374
8. *Understanding and Using the SP Switch*, SG24-5161

Requisite function

This is a list of the software and hardware directly used by the SP Switch component of PSSP. Problems within the requisite software and hardware may manifest themselves as error symptoms in the SP Switch. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the SP Switch component of PSSP, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

1. System Data Repository (SDR) component of PSSP
2. Ethernet component of AIX
3. Group Services **hags** and Topology Services **hats** components of PSSP
4. SP System Security Services

   Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
5. SP Switch hardware monitor and control: **hardmon** component of PSSP.

Internal SP Switch subsystem components

Software subsystems which the SP Switch service component contains are listed here. This section diagnoses problems that are handled by these subsystems.

1. **worm** - The fault-service daemon part that explores and initializes the SP Switch.
2. **Ecommands** - Commands such as **Estart**, **Efence**, **Eunfence**, and others.
3. **css** - The fault-service daemon part that handles switch recovery.
4. kernel extensions - Enabling IP, **Ecommands**, adapter error report, and switch recovery.
5. Hardware Abstraction Layer (**HAL**) - switch service.
6. CSS adapter device driver and adapter microcode - SP Switch communication.
7. CSS adapter device driver and adapter diagnostics - hardware diagnostics.

Error information

AIX Error Log information

In order to isolate an adapter or SP Switch error, first view the AIX error log.

- For switch related problem, login to the primary node. The **Eprimary** command lists the primary node's node number. In cases where the primary node failed, there will be no primary. In this case, login to the node listed under `oncoming primary`. In cases where the SP Switch continued working, it may have replaced the primary node several times. You cannot locate the primary node using the previous methods. In this case, the only way to locate the primary is to look inside the log files.
- For adapter related problems, you should login to the suspect node.
  - Once you are on the desired node, issue the AIX command:

    ```
    errpt | more
    ```
  - Output is similar to the following:

    ```
    ERROR_ID TIMESTAMP   T CL Res Name      ERROR_Description
    34FFBE83 0604140393T T H   Worm  Switch Fault-detected by switch chip
    C3189234 0604135793  T H   Worm  Switch Fault-not isolated
    ```

The Resource Name (`Res Name`) in the error log gives you an indication of what resource detected the failure.

Table 16. Resource Name failure indications - SP Switch

Resource name	Indication
Worm	The information was extracted from the SP Switch **worm** subsystem for SP Switch initialization.
css	Incorrect status was detected by the adapter (css device driver) for SP Switch recovery.
css0	css failed adapter diagnostics for SP Switch adapter recovery.

For a more detailed description, issue the AIX command:

```
errpt -a [-N resource_name] | more
```

where the optional *resource_name* is one of the entries in Table 16.

An error which was reported by the fault service daemon may be related to a more detailed message, or to a previous related error message saved in the stack file. View these linked related errors by issuing the FFDC **fcreport** command. The other trace files, and in particular the **flt** file, may give more information about the problem. Neighboring error log entries may give detailed information as well. For more information on FFDC services and tools, see *First Failure Data Capture Programming Guide and Reference*.

There are several subcomponents that write entries in the AIX error log:

1. TB3, TB3MX, or TB3PCI adapter errors, which are recognized by labels with a prefix of: **TB3_**.

Table 17. Possible causes of adapter failures - SP Switch

Label and Error ID	Error description and analysis
TB3_SLIH_ER BD3BFA74	**Explanation:** SP Switch adapter interrupt handler error. **Cause:** SP Switch adapter or SP Switch failure. **Action:** • Search for more logged information. • Run adapter diagnostics. • Call IBM Hardware service if the problem persists.
TB3_SVC_QUE_FULL_ER 9311420B	**Explanation:** SP Switch adapter service interface overrun. **Cause:** SP Switch adapter or SP Switch failure. **Action:** Call IBM Hardware service if the problem persists.
TB3_CONFIG1_ER A4BFBE61	**Explanation:** Failed to update the ODM during CSS configuration. **Cause:** Software error **Action:** Run configuration method with the **verbose** option. See “Adapter configuration error information” on page 152.
TB3_PIO_ER 4E93C6F0	**Explanation:** An I/O error was received on the SP Switch adapter device driver. **Cause:** SP Switch adapter hardware failure. **Action:** Run adapter diagnostics. Call IBM Hardware service if the problem persists.
TB3_HARDWARE_ER 17B8B70C	**Explanation:** SP Switch adapter hardware or microcode error. **Cause:** SP Switch adapter hardware failure or microcode error. **Action:** Run adapter diagnostics. Call IBM Hardware service if the problem persists.
TB3_MICROCODE_ER 0A67508C	**Explanation:** SP Switch adapter hardware or microcode error. **Cause:** SP Switch adapter or microcode failure. **Action:** Run adapter diagnostics. Call IBM Hardware service if the problem persists.
TB3_LINK_RE 3A1947F1	**Explanation:** SP Switch adapter link outage occurred. **Cause:** The node is fenced. **Action:** Run **Eunfence** to unfence the node. **Cause:** A cable is loose, disconnected, or faulty. **Action:** Run “Cable diagnostics” on page 173.

Table 17. Possible causes of adapter failures - SP Switch (continued)

Label and Error ID	Error description and analysis
TB3_BAD_PACKET_RE 1102E5B2	**Explanation:** Bad packet received. This entry always appears together with TB3_BAD_PKT_CONT_RE. **Cause:** An SP Switch cable failure. **Action:** Run "Cable diagnostics" on page 173. **Cause:** SP Switch adapter or SP Switch failure. **Action:** • Run adapter diagnostics. • Call IBM Hardware service if the problem persists.
TB3_BAD_PKT_CONT_RE 448DCF0D	**Explanation:** More detailed information about a bad packet received error. This entry completes the bad packet detailed information started in the previous TB3_BAD_PACKET_RE entry.
TB3_TRANSIENT_RE E94651FA	**Explanation:** SP Switch adapter transient error. **Cause:** A loose, disconnected, or faulty cable. **Action:** Run "Cable diagnostics" on page 173.
TB3_FAILURE_CONT_RE EBF464AF	**Explanation:** More detailed information about an SP Switch adapter transient, link, hardware or microcode error. This entry completes the detailed information started in the previous TB3_TRANSIENT_RE, TB3_LINK_RE, TB3_MICROCODE_ER or TB3_HARDWARE_ER entries.
TB3_THRESHOLD_ER 6E31CB46	**Explanation:** SP Switch adapter error threshold exceeded. The number of bad packets exceeded the threshold level. **Cause:** A loose, disconnected, or faulty cable. **Action:** Run "Cable diagnostics" on page 173.

2. TB3MX adapter errors, which are recognized by labels with a prefix of: TB3MX_.

Table 18. Possible causes of TB3MX specific adapter failures

Label and Error ID	Error Description and Analysis
TB3MX_ACCESS_ER 417A4BCB	**Explanation:** SP Switch adapter TB3MX user access error. A user process attempted to read or write a no access region of a switch adapter. **Cause:** User software error. **Action:** Review any user applications that have access to a an SP Switch adapter, such as those that issue LAPI, MPI or switch clock API calls.

3. Switch fault service daemon errors, which are recognized by labels with a prefix of: SP_SW_, TBS_ , SP_CSS_, SP_CLCK, or SP_MCLCK.

Table 19. Possible causes of fault service daemon failures - SP Switch

Label and Error ID	Error description and analysis
SP_SW_RECV_STATE_RE 2A095362	**Explanation:** SP Switch receiver state machine error. **Cause:** SP Switch adapter or SP Switch failure. **Action:** Call IBM Hardware Service if the problem persists.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_PE_ON_DATA_RE EBC5B66E	**Explanation:** SP Switch sender parity error on data. **Cause:** An SP Switch board failure. **Action:** Call IBM Hardware Service.
SP_SW_INVALD_RTE_RE BCCF8CCD	**Explanation:** SP Switch sender invalid route error. **Cause:** SP Switch adapter microcode or a switch daemon software error. **Action:** Call the IBM Support Center.
SP_SW_EDC_ERROR_RE 1A508382	**Explanation:** Receiver EDC class error. **Cause:** A transient error in data occurred during transmission over switch links. The EDC error may be one of the following: • A receiver EDC error. • A parity error on route. • An undefined control character was received. • Unsolicited data was received. • A receiver lost end-of-packet. • A token count mismatch. • A token sequence error. • A token count overflow. **Cause:** A loose, disconnected, or faulty cable **Action:** • See “Cable diagnostics” on page 173. • See **/var/adm/SPlogs/css/out.top** for cable information. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** See “Verify SP Switch node operation” on page 175. **Cause:** SP Switch adapter hardware failure. **Action:** • Run adapter diagnostics. • See “SP Switch device and link error information” on page 154 for additional information.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_SNDLOSTEOP_RE F9453ADD	**Explanation:** SP Switch sender lost EOP (end-of-packet) condition. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 173. • See **/var/adm/SPlogs/css/out.top** for cable information. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** See “Verify SP Switch node operation” on page 175. **Cause:** SP Switch adapter hardware failure. **Action:** • Run adapter diagnostics. • See “SP Switch device and link error information” on page 154.
SP_SW_RCVLNKSYNC_RE B4C98286	**Explanation:** SP Switch receiver link synchronization error or a switch chip lost clock synchronization on one of its receive ports. **Cause:** A loose, disconnected, or faulty cable. **Action:** • Check, reconnect, or replace the cable • See **/var/adm/SPlogs/css/out.top** for cable information. • See “Cable diagnostics” on page 173. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css/flt** for more information. • See “Verify SP Switch node operation” on page 175. **Cause:** SP Switch adapter hardware failure. **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css/flt** for more information. **Cause:** Remote switch adapter hardware failure. **Action:** • Run adapter diagnostics on remote node. • See **/var/adm/SPlogs/css/flt** to identify remote node.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_EDCTHRSHLD_RE E940D28E	**Explanation:** SP Switch receiver EDC errors exceeded the threshold level. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 173. • See **/var/adm/SPlogs/css/flt** for more information. • Call IBM Hardware Service if the problem persists.
SP_SW_SND_STATE_RE 6116CB9D	**Explanation:** SP Switch sender state machine error. **Cause:** SP Switch adapter or SP Switch failure. **Action:** Call IBM Hardware Service if the problem persists.
SP_SW_FIFOOVRFLW_RE D06B4832	**Explanation:** SP Switch receiver FIFO overflow error. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 173. • See **/var/adm/SPlogs/css/flt** for more information **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css/flt** for more information • See “Verify SP Switch node operation” on page 175. **Cause:** SP Switch adapter hardware failure **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css/flt** for more information. • See “SP Switch device and link error information” on page 154.
SP_SW_SNDTKNTHRS_RE C948155E	**Explanation:** SP Switch sender token errors exceeded the threshold level. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 173. • Call IBM Hardware Service if the problem persists. • See **/var/adm/SPlogs/css/flt** for more information.
SP_SW_PE_ON_NMLL_RE 9B7A16D3	**Explanation:** NMLL Switch central queue parity error. **Cause:** A switch board failure. **Action:** Call IBM Hardware Service.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_PE_ON_NCLL_RE 2E4C02D8	**Explanation:** NCLL Switch central queue parity error. **Cause:** A switch board failure. **Action:** Action: Call IBM Hardware Service.
SP_SW_NCLL_UNINT_RE 840FBC5A	**Explanation:** NCLL switch central queue was not initialized. **Cause:** A switch board failure. **Action:** Call IBM Hardware Service.
SP_SW_SNDLNKSYNC_RE 33164DD2	**Explanation:** SP Switch sender link synchronization error. Switch chip lost synchronization clock on one of its send ports. **Cause:** A loose, disconnected, or faulty cable. **Action:** • Check, reconnect, or replace the cable. • See **/var/adm/SPlogs/css/out.top** for cable information. • See “Cable diagnostics” on page 173. **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** • Replace the switch cable from the powered-off node with a wrap plug. • See **/var/adm/SPlogs/css/flt** for more information. • See “Verify SP Switch node operation” on page 175. **Cause:** A switch adapter hardware failure. **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css/flt** for more information. • See “SP Switch device and link error information” on page 154. **Cause:** A remote switch adapter hardware failure. **Action:** • Run adapter diagnostics on remote node. • See **/var/adm/SPlogs/css/flt** to identify remote node. • See “SP Switch device and link error information” on page 154.
SP_SW_SVC_PKTLEN_RE BE19105C	**Explanation:** SP Switch service logic encountered an incorrect packet length. **Cause:** SP Switch adapter microcode error or a switch daemon software error. **Action:** • See **/var/adm/SPlogs/css/flt** for more information. • Call the IBM Support Center if the problem persists.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_PE_INBFIFO_RE 8E1C91D6	**Explanation:** SP Switch service logic detected a bad parity in FIFO. **Cause:** A switch board failure. **Action:** Call IBM Hardware Service.
SP_SW_CRC_SVCPKT_RE 642E0890	**Explanation:** SP Switch service logic detected an incorrect CRC on a service packet. **Cause:** A transient error in the data occurred during transmission over the switch links. **Action:** See **/var/adm/SPlogs/css/flt** for more information. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See **/var/adm/SPlogs/css/flt** for more information. • See “Cable diagnostics” on page 173. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css/flt** for more information. • See “Verify SP Switch node operation” on page 175 for additional information. **Cause:** A switch adapter hardware failure. **Action:** • Run adapter diagnostics • See **/var/adm/SPlogs/css/flt** for more information. • See “SP Switch device and link error information” on page 154 for additional information.
SP_SW_PE_RTE_TBL_RE 5FEAC531	**Explanation:** SP Switch service logic determined bad parity in the route table. **Cause:** A switch board failure. **Action:** Call IBM Hardware Service.
SP_SW_LNK_ENABLE_RE 732582E8	**Explanation:** SP Switch service logic detected an invalid link enable value. **Cause:** A switch daemon software error. **Action:** Call the IBM Support Center.
SP_SW_SEND_TOD_RE 28F6B80F	**Explanation:** SP Switch service logic send TOD error. **Cause:** A switch daemon software error. **Action:** Call the IBM Support Center.
SP_SW_SVC_STATE_RE E51188C7	**Explanation:** SP Switch service logic state machine error. **Cause:** A switch board failure. **Action:** Call IBM Hardware Service if the problem persists.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_OFFLINE_RE DC85B51D	**Explanation:** Node fence request received. **Cause:** The operator ran the **Efence** command. **Action:** Run the **Eunfence** command to bring the node onto the switch.
SP_SW_PRI_TAKOVR_RE 80E1023C	**Explanation:** SP Switch primary node takeover. **Cause:** The switch primary node became inaccessible. **Action:** See the error log on the previous switch primary node.
SP_SW_BCKUP_TOVR_RE CE1A52F8	**Explanation:** SP Switch primary-backup node takeover. **Cause:** The primary backup node became inaccessible. **Action:** • See **/var/adm/SPlogs/css/out.top** for cable information. • See the error log on the previous switch primary backup node.
SP_SW_LST_BUP_CT_RE A5D521DE	**Explanation:** Primary-backup node not responding. **Cause:** The primary backup node became inaccessible. **Action:** See the error log on the current switch primary backup node.
SP_SW_UNINI_NODE_RE A4653F22	**Explanation:** SP Switch nodes that are listed were not initialized during **Estart** command processing. **Cause:** These nodes was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css/flt** file for more information. • See “Verify SP Switch node operation” on page 175 for additional information. **Cause:** Nodes fenced. **Action:** Run the **Eunfence** command. **Cause:** SP Switch adapter problem. **Action:** Run adapter diagnostics.
SP_SW_UNINI_LINK_RE DB5B6D5D	**Explanation:** SP Switch links were not initialized during **Estart** command processing. **Cause:** The switch cable is not wired correctly. **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly. **Cause:** Nodes fenced. **Action:** Run the **Eunfence** command. **Cause:** A loose, disconnected, or faulty cable **Action:** See “Cable diagnostics” on page 173.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_SIGTERM_ER 6E0AA114	**Explanation:** SP Switch fault service daemon received **SIGTERM**. **Cause:** Another process sent a **SIGTERM**. **Action:** Run the ′**rc.switch** command to restart switch daemon.
SP_SW_GET_SVCREQ_ER 1861B4F6	**Explanation:** Switch daemon could not get a service request. **Cause:** A switch kernel extension error. **Action:** Call the IBM Support Center.
SP_PROCESS_KILLD_RE DD37D0E5	**Explanation:** User process was killed due to a link outage. **Cause:** A switch adapter or switch failure. **Action:** See neighboring error log entries to determine the cause of the outage. **Cause:** The operator fenced this node. **Action: Eunfence** the node.
SP_SW_MISWIRE_ER C6A41256	**Explanation:** Switch cable miswired (not connected to the correct switch jack). **Cause:** The switch cable was not wired correctly. **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly.
SP_SW_LOGFAILURE_RE D35D236E	**Explanation:** Error writing to switch log files. **Cause:** The **/var** file system is full. **Action:** Obtain free space in the file system or expand the file system. **Cause:** There are too many files open in the system. **Action:** Reduce the number of open files in the system.
SP_SW_INIT_FAIL_ER 071B5E7A	**Explanation:** Switch fault service daemon initialization failed. **Cause:** The operating environment could not be established. **Action:** • See Detail Data in this error log entry, for the specific failure. • Correct the problem and restart the daemon, by issuing the **rc.switch** command. • Call the IBM Support Center if the problem persists.
SP_SW_SVC_Q_FULL_RE 42B75697	**Explanation:** Switch service send queue is full. **Cause:** There is a traffic backlog on the switch adapter. **Action:** Call the IBM Support Center if the problem persists.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	**Error description and analysis**
SP_SW_RSGN_PRIM_RE 66B1169C	**Explanation:** Resigning switch primary node responsibilities. **Cause:** Could not communicate over the switch. **Action:** See neighboring error log entries to determine the cause of the outage. **Cause:** Another node was selected as primary node. **Action:** None.
SP_SW_RSGN_BKUP_RE FDC35FFD	**Explanation:** Resigning as the switch primary-backup node. **Cause:** Could not communicate over the switch. **Action:** See neighboring error log entries to determine the cause of the outage. **Cause:** Another node was selected as the primary backup node. **Action:** None.
SP_SW_ACK_FAILED_RE A6524825	**Explanation:** Switch fault service daemon acknowledge of a service command failed. **Cause:** A switch communications failure. **Action:** Call the IBM Support Center if the problem persists. **Cause:** A traffic backlog on the switch adapter. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_SDR_FAIL_RE CD9CE5B3	**Explanation:** Switch fault service daemon failed to communicate with the SDR. **Cause:** An Ethernet overload. **Action:** Call the IBM Support Center if the problem persists. **Cause:** Excessive SDR traffic. **Action:** Call the IBM Support Center if the problem persists. **Cause:** The SDR daemon or the control workstation is down. **Action:** Check to see if the SDR daemon is up. **Cause:** A software error. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_SCAN_FAIL_ER 90AB4ED5	**Explanation:** Switch scan failed. **Cause:** Could not communicate over the switch. **Action:** Issue the **Estart** command if primary takeover does not occur. **Cause:** A switch adapter or switch failure. **Action:** Issue the **Estart** command if primary takeover does not occur.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_NODEMISW_RE F402BA0E	**Explanation:** SP Switch node miswired. **Cause:** A switch cable was not plugged into the correct node. **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly.
SP_SW_RTE_GEN_RE 7F223C8F	**Explanation:** Switch fault service daemon failed to generate routes. **Cause:** A software error. **Action:** Call the IBM Support Center.
SP_SW_FENCE_FAIL_RE 5515EE64	**Explanation:** Fence of a node failed. **Cause:** Could not communicate over the switch. **Action:** • See **/var/adm/SPlogs/css/flt** for more information. • See the error log on the failing node. • Issue the **Estart** command to initialize the switch network.
SP_SW_REOP_WIN_ER 5988A5BF	**Explanation:** Switch fault service daemon reopen adapter windows failed. **Cause:** A switch kernel extension error. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_ESTRT_FAIL_RE F9A6917F	**Explanation: Estart** failed - switch network could not be initialized. **Cause:** Could not initialize switch chips or nodes. **Action:** • See Detail Data in this error log entry, for specific failure. • Issue **Eclock -d** to reset the switch network and reestablish switch clocking. • Call the IBM Support Center if the problem persists.
SP_SW_CBCST_FAIL_RE 275E6902	**Explanation:** Switch fault service daemon command broadcast failed. **Cause:** Could not communicate over the switch. **Action:** Call the IBM Support Center if the problem persists. **Cause:** A traffic backlog on the switch adapter. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_UBCST_FAIL_RE 04839F9e	**Explanation:** Switch fault service daemon DBupdate broadcast failed. **Cause:** A switch communications failure. **Action:** Call the IBM Support Center if the problem persists. **Cause:** A traffic backlog on the switch adapter. **Action:** Call the IBM Support Center if the problem persists.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_SW_DNODE_FAIL_RE EAC0A694	**Explanation:** Switch fault service daemon failed to communicate with dependent nodes. **Cause:** A failure to communicate over the switch. **Action:** Call the IBM Support Center if the problem persists. **Cause:** A traffic backlog on the switch adapter. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_IP_RESET_ER F323DBDB	**Explanation:** Switch fault service daemon could not reset IP. **Cause:** A switch kernel extension error. **Action:** Call the IBM Support Center if the problem persists.
SP_SW_PORT_STUCK_RE A45FB6AB	**Explanation:** Switch port cannot be disabled. **Eunfence** failed. **Cause:** • switch-chip or adapter hardware error • cable failure **Action:** • See **/var/adm/SPlogs/css/worm.trace** on the primary node for more information. • Run adapter diagnostics on the node that failed to unfence.
SP_SW_CATALOG_OPEN_ST 7962AB9A	**Explanation:** Failed to open the switch catalog file. **Cause:** Too many files are opened. **Action:** Rerun the **rc.switch** script. **Cause:** Catalog file not found. **Action:** Check installation.
SP_SW_DIST_TOP_RE 4FFEADB6	**Explanation:** Switch topology file distribution failed to the listed nodes during autojoin. **Action:** See more information in file **/var/adm/SPlogs/css/dist_topology.log**. **Cause:** Ethernet connection lost. **Action:** Check Ethernet cable and configuration. **Cause: kerberos** security lost. **Action:** Get **kerberos** key. See **kerberos** command. **Cause:** Not enough space left in the node’s **/etc/SP** file system. **Action:** Free more space.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_CLCK_MISS_RE 0A137C5D	**Explanation:** Switch (non-master) lost clock. **Cause:** Switch clock signal is missing. **Action:** • See “Clock diagnostics” on page 170. • Call IBM Hardware service if problem persists.
SP_SW_FSD_TERM_ER E8817142	**Explanation:** Switch fault service daemon process was terminated. **Cause:** Faulty switch adapter or missing external clock source. **Action:** • See preceding error log entries for failure causes. • See “SP Switch adapter diagnostics” on page 172. **Cause:** Faulty system planar. **Action:** Run complete diagnostics on the node. If diagnostics fail to isolate the problem, contact IBM Hardware Service.
SP_MCLCK_MISS_RE 7A3D0758	**Explanation:** Switch (master oscillator) lost clock. **Cause:** • A node or switch board lost power. • A switch board failure. • A user incorrectly clocked the system. **Action:** • See “Clock diagnostics” on page 170. • Run the **Estart** command to initialize the switch network. • Call IBM Hardware service if the problem persists.
SP_SW_ECLOCK_RE 9EEBC3C4	**Explanation: Eclock** command issued by user. **Cause: Eclock** command run by the administrator. **Action:** Run the **Estart** command to initialize the switch network.
SP_SW_ADPT_BUS_ER 2ABBD2C0	**Explanation:** Switch adapter bus signal received. **Cause:** Bus error occurred during adapter access. Adapter configuration or hardware failure. **Action:** Call the IBM Support Center.
TBS_HARDWARE_ER F3448CD2	**Explanation:** Switch board hardware error. **Cause:** Switch board failure. **Action:** call IBM Hardware service.

Table 19. Possible causes of fault service daemon failures - SP Switch (continued)

Label and Error ID	Error description and analysis
SP_CSS_IF_FAIL_ER FCACBBD3	**Explanation:** Switch adapter interface system call failed. **Cause:** Could not communicate with the switch adapter. **Action:** • Check the switch adapter configuration. • Run adapter diagnosis, see “SP Switch adapter diagnostics” on page 172. • Call IBM Hardware service if the problem persists.

4. Adapter Diagnostic errors, which are recognized by labels with a prefix of: SWT_DIAG_

Table 20. Possible causes of adapter diagnostic failures - SP Switch

Label and Error ID	Error description and analysis
SWT_DIAG_ERROR1_ER 8998B96D	**Explanation:** SP Switch adapter failed post-diagnostics, see **diag** command. **Cause:** Faulty switch adapter or missing external clock source. **Action:** See “SP Switch adapter diagnostics” on page 172.
SWT_DIAG_ERROR2_ER 2FFF253A	**Explanation:** Switch adapter failed pre-diagnostics, failed power-on-self-test diagnostics. **Cause:** Faulty switch adapter or missing external clock source. **Action:** See “SP Switch adapter diagnostics” on page 172.

Adapter configuration error information

The following table lists the possible values of the **adapter_config_status** attribute of the **switch_responds** object of the SDR. Use the following command to determine its value:

```
SDRGetObjects switch_responds
```

Use the value of the **adapter_config_status** attribute for the node in question, to index into Table 21.

Note: The **adapter_config_status** table that follows uses the phrase ″adapter configuration command″. This refers to the SP Switch adapter (CSS adapter) configuration method. Use the following syntax to invoke it:

```
/usr/lpp/ssp/css/cfgtb3 -v -l css0 > output_file_name
```

Table 21. adapter_config_status values - SP Switch

adapter_config_status	Explanation and recovery
css_ready	Correctly configured CSS adapter.

Table 21. adapter_config_status values - SP Switch (continued)

adapter_config_status	Explanation and recovery
odm_fail genmajor_fail genminor_fail getslot_fail build_dds_fail	**Explanation:** An ODM failure has occurred while configuring the CSS adapter. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
lname_error	**Explanation:** The device logical name specified on the CSS adapter configuration command was incorrect. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
undefine_system_fail define_system_fail xilinx_system_fail	**Explanation:** The System Standard C Library Subroutine failed during CSS adapter configuration. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
define_fail	**Explanation:** The current instance of the CSS logical device could not be redefined. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
chkslot_fail	Verify that the CSS adapter is properly seated, then rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
busresolve_fail	**Explanation:** There are insufficient bus resources to configure the CSS adapter. **Action:** Contact the IBM Support Center.
xilinx_load_fail dd_load_fail fs_load_fail	**Action:** See “Verify software installation” on page 174. If software installation verification is successful and the problem persists, contact the IBM Support Center.
make_special_fail	**Explanation:** The CSS device special file could not be created during adapter configuration. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply them with the command output.
dd_config_fail fs_init_fail	**Explanation:** An internal device driver error occurred during CSS adapter configuration. **Action:** See “Information to collect before contacting the IBM Support Center” on page 164.
diag_fail	**Explanation:** SP Switch diagnostics failed. **Action:** See “SP Switch adapter diagnostics” on page 172.
not_configured	**Explanation:** The CSS adapter is missing or not configured.

SP Switch device and link error information

The device and link current status is gathered in the annotated switch topology file, **out.top**, that is created on the primary node. The file looks like the switch topology file except that for each device or link that differs from the operational default status, an additional comment is made.

These additional comments are appended to the file by the fault service daemon and reflect the current connectivity status of the link or device. No comment on a link or device line means that the link or device exists and is operational.

Not all the comments reflect an error. Some may be a result of the system configuration or current administration.

An example of a failing entry and description is in “out.top” on page 161. If the listed recovery actions fail to resolve your problem, contact the IBM Support Center.

The possible device status values for SP Switch systems, with their recovery actions, are listed in Table 22. The possible link status values for SP Switch systems, with their recovery actions, are listed in Table 23 on page 155.

Table 22. SP Switch device status and recovery actions

Device status number	Device status text	Explanation and recovery actions
-4	Device has been removed from the network - faulty.	**Explanation:** The device has been removed from the switch network, because of a fault on the device. **Cause:** A fault on the device. **Action:** If the device in question is a node, see “Verify SP Switch node operation” on page 175. Otherwise contact IBM Hardware Service.
-5	Device has been removed from the network by the system administrator.	**Explanation:** The device was placed offline by the system administrator (**Efence**). **Cause:** The switch administrator ran the **Efence** command. **Action: Eunfence** the device.
-6	Device has been removed from network - no AUTOJOIN.	**Explanation:** The device was removed and isolated from the switch network. **Cause:** The node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. **Action:** First attempt to **Eunfence** the device. If the node fails to rejoin the switch network, see “AIX Error Log information” on page 138. If the problem persists contact the IBM Support Center.
-7	Device has been removed from network for not responding.	**Explanation:** The device was removed from the switch network. **Cause:** An attempt was made to contact the device, but the device did not respond. **Action:** If the device in question is a node, see “Verify SP Switch node operation” on page 175. Otherwise contact the IBM Support Center.

Table 22. SP Switch device status and recovery actions (continued)

Device status number	Device status text	Explanation and recovery actions
-8	Device has been removed from network because of a miswire.	**Explanation:** The device is not cabled properly. **Cause:** Either the switch network is miswired, or the frame supervisor tty is not cabled properly. **Action:** First view the **/var/adm/SPlogs/css/ cable_miswire file**. Verify and correct all links listed in the file. Then issue the **Eclock -d** command and rerun the **Estart** command. If the problem persists, contact IBM Hardware Service.
-9	Destination not reachable.	**Explanation:** The device was not reachable through the switch network **Cause:** This is generally due to other errors in the switch network fabric. **Action:** Investigate and correct the other problem, then run the **Estart** command.

Table 23. SP Switch link status and recovery actions

Link status number	Link status text	Explanation and recovery actions
0	Link has been removed from network - no AUTOJOIN.	**Explanation:** The device was removed and isolated from the switch network. **Cause:** The node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. **Action:** First attempt to **Eunfence** the device. If the node fails to rejoin the switch network, see "AIX Error Log information" on page 138. If the problem persists contact the IBM Support Center.
-2	Wrap plug is installed.	**Explanation:** This link is connected to a wrap plug. **Cause:** The wrap plug is connected to the port in order to test the port. This is not normally a problem. **Action:** None.
-4	Link has been removed from network or miswired - faulty.	**Explanation:** The link is not operational and was removed from the network. **Cause:** Either the link is miswired or the link has failed. **Action:** First check the **/var/adm/SPlogs/css** directory for the existence of a **cable_miswire** file. If the file exists, verify and correct all links listed in the file. Then issue the **Eclock -d** command and rerun the **Estart** command. If the **cable_miswire** file does not exist, examine the **/var/adm/SPlogs/css/flt** file for entries relating to this link. If entries are found, verify that the cable is seated at both ends, then **rc.switch** the primary node and rerun the **Estart** command. If the problem persists, contact the IBM Support Center.

Table 23. SP Switch link status and recovery actions (continued)

Link status number	Link status text	Explanation and recovery actions
-7	Link has been removed from network - fenced.	**Explanation:** The device was placed offline by the Systems Administrator (**Efence**). **Action: Eunfence** the associated node.
-8	Link has been removed from network - probable miswire.	**Explanation:** The link is not cabled properly. **Action:** View the **/var/adm/SPlogs/css/cable_miswire** file. Verify and correct all links listed in the file, then **rc.switch** the primary node and rerun the **Estart** command.
-9	Link has been removed from network - not connected.	**Explanation:** The link cannot be reached by the primary node, therefore initialization of the link is not possible. **Cause:** This is generally caused by other problems in the switch network, such as a switch chip being disabled. **Action:** Investigate and correct the underlying problem, then run the **Estart** command.

Dump information

The dump files are created either along with the **css.snap** script, which automatically invokes the associated utility commands, or manually at the user's request. The **css.snap** script is issued by the user or by the switch support code (device driver, fault service daemon, kernel extension) whenever a serious error occurs. **css.snap** gathers dump and trace files into a snapshot compressed package. For more information, see "css.snap package" on page 164.

errpt.out

The AIX commands **errpt** and **errpt -a** are redirected into the dump file **errpt.out**, by the **css.snap** script. The file is gathered with the other dump and trace files, into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

css_dump.out

1. The command:

   ```
   /usr/lpp/ssp/css/css_dump -r
   ```

   dumps the adapter device driver trace buffer to stdout.

   When this command is run automatically from the **css.snap** script, the information is redirected into the file **css_dump.out**. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

   The command dumps status and error information for events such as **ioctl** calls to the device **css0** and adapter error interrupts.
2. This command dumps the adapter memory onto the standard output.

   ```
   /usr/lpp/ssp/css/tbXdump
   ```

X is automatically replaced in the **css.snap** script with 3, 3mx, or 3pci according to the adapter type: TB3, TB3MX, or TB3PCI respectively.

When this command is run automatically from the **css.snap** script, the information is redirected into the file **css_dump.out**, which is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

The command dumps adapter memory and state and is destructive, meaning that the adapter cannot be in use while issuing this command. Furthermore, after running the command the adapter should be restarted. Restarting the adapter is done by running the **rc.switch** script on the affected node:

```
/usr/lpp/ssp/css/rc.switch
```

fs_dump.out

The command:

```
/usr/lpp/ssp/css/fs_dump -r
```

dumps the fault-service kernel extension trace buffer to stdout.

When the command is run from the **css.snap** script, the output is redirected to the **fs_dump.out** file. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

The command dumps status and error information for events such as user-space system calls to the fault service kernel extension.

cssadm.debug

The file **cssadm.debug** is in the **/var/adm/SPlogs/css** directory. The file contains trace information of the actions of the **cssadm** daemon. This file contains entries for each event received and handled, as well as how the events were handled and the results. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

cssadm.stderr

The file **cssadm.stderr** is in the **/var/adm/SPlogs/css** directory. The file contains any unexpected error messages received by the **cssadm** daemon, while performing commands external to the daemon. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

cssadm.stdout

The file **cssadm.stdout** is in the **/var/adm/SPlogs/css** directory. The file contains any unexpected informational messages received by the **cssadm** daemon while performing commands external to the daemon. In general, this file should remain empty. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

regs.out

The command:

```
/usr/lpp/ssp/css/read_regs
```

dumps the switch adapter's Programmable Option Selection (POS) registers and TBIC registers in addition to the adapter clock status.

All information is dumped to stdout. When this command is run automatically from the **css.snap** script, the information is redirected into the file **regs.out**. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

The **read_regs** command reads all registers on the local SP switch adapter and adds information about the kernel extension and the clock source that are used by the adapter. The exact output differs for the TB3, TB3MX, or TB3PCI switch adapters.

You can check the TBIC_STATUS register from this information. Another way to check this register is by issuing the command:

```
/usr/lpp/ssp/css/diags/read_tbic -s
```

which is described in the relevant switch diagnostic section.

router.log

The file **router.log** is in the **/var/adm/SPlogs/css** directory. This file is used to record error and warning messages issued by the route table generator (RTG). It contains the RTG version and date, the date and time that the file was created, the node on which the file was created, information about the system topology and the routing algorithm type, and any messages generated during route generation. This file is created during service route generation on the primary node, and during processor route generation on all nodes.

If the file is found at the start of route generation, it is copied to **router.log.old**. If the file contains any messages at the end of route generation, the file is copied to **router_failed.log1**. Up to two failed router logs, **router_failed.log1** and **router_failed.log2** are maintained at any time.

This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 164.

CSS_test.log

The command:

```
/usr/lpp/ssp/bin/CSS_test
```

produces a log file **/var/adm/SPlogs/CSS_test.log**, which is then gathered into the snapshot compressed package. This file is present only if the **CSS_test** command was run on the node.

When running in Restricted Root Access (RRA) mode, **CSS_test** must be run from the control workstation. For more information, see “Verify software installation” on page 174, and the **CSS_test** entry of *PSSP: Command and Technical Reference*.

vdidl.out

These commands dump the switch kernel extension internal information when dealing with IP.

```
/usr/lpp/ssp/css/vdidlX -v -d
/usr/lpp/ssp/css/vdidlX -v -i
/usr/lpp/ssp/css/vdidlX -v -s
```

X is automatically replaced in the **css.snap** script with 3, 3mx, or 3pci, according to the adapter type: TB3, TB3MX, or TB3PCI respectively,

The exact output differs for the TB3, TB3MX, or TB3PCI switch adapters. All information is dumped to stdout. When this command is run automatically from the **css.snap** script, the information is redirected into the file **vdidl.out**. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 164.

spdata.out

The commands:

- **/usr/lpp/ssp/css/splstdata -f -G**
- **/usr/lpp/ssp/css/splstdata -s -G**
- **/usr/lpp/ssp/css/splstdata -n -G**
- **/usr/lpp/ssp/css/splstdata -b -G**

print network and switch current status information to stdout. The **css.snap** script dumps the command output to the file **spdata.out** and packages the file together with other snapshot files. See “Information to collect before contacting the IBM Support Center” on page 164. The **splstdata** command is described in *PSSP: Command and Technical Reference*.

netstat.out

The AIX commands:

- **netstat -I css0**
- **netstat -m**

print network related data to stdout. The **css.snap** script dumps the command output to the file **netstat.out** and packages the file together with the other snapshot files. See “Information to collect before contacting the IBM Support Center” on page 164. The **netstat** command is described in *AIX 5L Version 5.1 Command and Technical Reference*.

scan_out.log and scan_save.log

These files are in the **/adm/SPlogs/css** directory of each node that runs the adapter diagnostics and the TBIC test within this test (TBIC test within adapter diagnostics). The TBIC, which is a chip on the switch adapter, is scanned into this file and saved as part of the snapshot, when running the **css.snap** script. The **scan_save.log** is a previous TBIC scan.

This TBIC scan procedure is destructive and therefore taken only when the **css.snap** runs after a permanent failure. No other switch communication is available, and the switch adapter is reset after the scan is taken. See “SP Switch adapter diagnostics” on page 172.

Trace information

All the trace files use message catalogs to display the messages, as described in *PSSP: Messages Reference*.

flt

The **flt** file is in the **/var/adm/SPlogs/css** directory of each node. This file is used to log hardware error conditions about the switch, recovery actions taken by the fault service daemon, and other general operations that alter the switch configuration. The **flt** file is most important in the primary node where the switch initialization and servicing takes place.

Each entry in the **flt** file contains:

- severity:
  - **i** - informational
  - **n** - notification
  - **e** - error message
- time stamp: date and time that the message occurred
- catalog message number
- message text

Each entry may contain this global field:

- **device_id** is either a switch node number, or a switch chip id. Switch chip ids are numbers greater than 100000. The last digit represents the chip id within the switch board and is between 0 and 7. For example, a **device_id** of 100025 represents chip number 5 in switch board 2.

rc.switch.<PID>.<date>.<time>

This file is an FFDC stack file, and is in the **/var/adm/ffdc/stacks** directory of each node. This file can be read using FFDC tools such as **fcreport** and **fcstkrpt**. For more information on FFDC services and tools, see *First Failure Data Capture Programming Guide and Reference*. The file contains detailed information on errors, and provides a mechanism to link errors that occur in the same flow. A link is provided between an AIX error log message to a more detailed message, or to a flow related error in the stack file.

Each error entry is provided with the debug information:

- Time stamp
- Location within the code: function name, file name, line number
- Code version
- Catalog message id
- Message link information

The file provides most of the information that is already included in the **flt** file.

rc.switch.log

The **rc.switch.log** file is in the **/var/adm/SPlogs/css** directory of each node. This file contains log information from the **rc.switch** script. This is the script that is responsible for starting the fault service daemon. If a node fails or is unable to start the fault service daemon, this file may contain an explanation of the failure.

out.top

The **out.top** file is in the **/var/adm/SPlogs/css** directory of each node. The fault service daemon creates a concurrent topology file. The daemon uses the **expected.top** file from the SDR or from the primary node's **/etc/SP** directory. The **expected.top** is created by the **Eannotator** command, and is an ideal network topology file. The fault service daemon combines the current status of links and nodes (devices) to the ideal file. The result is then saved to the **out.top** file.

The **out.top** file has comments added only on non-operational links or devices. An example of such a comment is:

```
S  15 2 tb3 1 0          E01-S17-BH-J8 to E01-N10           -4 R:
     device has been removed from network - faulty
     (link has been removed from network - fenced)
```

This means that:

- Switch chip 15, port 2 is connected to switch node number 1.
- The switch is located in frame E01 slot 17.
- Its bulkhead connection to the node is jack 8.
- The switch node is in frame E01 (the same frame as the switch board), and it is node number 10.
- -4 is the device status of the right side device (tb0 node number1), which has the more severe device status of the two devices that are listed.

  The device status of the node is `device has been removed from network - faulty`.

SP Switch ports that have no connection are usually wrapped:

```
S 13 3   s 13 3         E01-S17-BH-J3 to E01-S17-BH-J3     2 L:
 initialized (wrap plug is installed)
```

The possible device and link status are listed with their relative recovery actions under "SP Switch device and link error information" on page 154.

dist_topology.log

The **dist_topology.log** file is in the **/var/adm/SPlogs/css** directory of the primary node. This file contains error messages that occur during the distribution of the topology file to the other nodes. The topology file distribution is done via the Ethernet and not through the SP Switch. Therefore, a failure to distribute the topology file is not directly related to the operation of the SP Switch.

worm.trace

The **worm.trace** file is in the **/var/adm/SPlogs/css** directory of each node. This file contains fault service daemon entries related to those in the **flt** file, but these entries are more detailed.

On a secondary node, the **worm.trace** file entries trace the initialization of the node connection to the SP Switch. On a primary node, the entries trace the whole switch initialization process: initializing each switch chip, calculating the routes to each node, establishing a primary-backup node, and setting the Time-Of-Day (TOD).

fs_daemon_print.file

The **fs_daemon_print.file** file is in the **/var/adm/SPlogs/css** directory of each node. This file contains the commands that the fault service daemon responded to. A time stamp along with a message is logged in the file whenever an event occurs. Important entries are the ones that refer to the adapter initialization, route table calculating, and downloading the table into the adapter.

daemon.stdout

The **daemon.stdout** file is in the **/var/adm/SPlogs/css** directory of each node. This file is a redirection of the stdout and contains the fault service daemon initial information taken before the establishment of any other log mechanisms.

logevnt.out

The **logevnt.out** file is in the **/var/adm/SPlogs/css** directory on the control workstation. The file contains records of errors which occurred in the components running on the node which experienced the error. These components are notified when switch-related error log entries are made, and report the summary data to Event Management for transmission to the control workstation. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 164 .

summlog.out

The **summlog.out** file is in the **/var/adm/SPlogs/css** directory on the control workstation. This file contains error information for the **css.summlog** daemon, which gathers summary log information and writes it to the summary log file. For information on the summary log file, see “Summary log for SP Switch, SP Switch2, and switch adapter errors” on page 69. **summlog.out** is a text file. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 164.

daemon.stderr

The **daemon.stderr** file is in the **/var/adm/SPlogs/css** directory of each node. This file is a redirection of the stderr and contains the error output from the fault service daemon during initialization.

dtbx.trace

The **dtbx.trace** file is in the **/var/adm/SPlogs/css** directory of each node. The file contains trace information of the last run of the css0 adapter diagnostics. The adapter diagnostics can run in two modes: Power-On-Self-Test (POST), and direct invocation, by issuing the command **diag**. The file creation time:

```
Midnight Dec 31 1969
```

indicates that the file was created during the POST, when the time had not yet been set. For more details, see “SP Switch adapter diagnostics” on page 172.

dtbx.failed.trace

The **dtbx.failed.trace** file is in the **/var/adm/SPlogs/css** directory of each node, if this file exists. This file contains trace information of the last failed run of the css0 adapter diagnostics. When the adapter diagnostics fail, they rename the **dtbx.trace** file to **dtbx.failed.trace**, in order to preserve the preliminary failure information. For more details, see "SP Switch adapter diagnostics" on page 172.

cable_miswire

The **cable_miswire** file is in the **/var/adm/SPlogs/css** directory of the primary node, if it exists. This file reports any miswire detected by the fault service daemon during SP Switch initialization (running the **worm** subsystem).

css.snap.log

The **css.snap.log** file is in the **/var/adm/SPlogs/css** directory of each node. This file is created every time that the **css.snap** command is run. This can be automatically on failure, or by direct invocation - when the user issues the **css.snap** command. This file contains information on what happened during the command invocation, and in particular:

- The time stamp of the snapshot.
- The node on which the snapshot was taken
- The list of files in the **/var/adm/SPlogs/css** directory before the command began running.
- The list of files which are assembled into the snapshot package.
- Information on all the running processes on the system at the snapshot.
- Information on the **ssp.css** software product.
- Information on the adapter and the microcode on the switch adapter.

Ecommands.log

The **Ecommands.log** file is in the **/var/adm/SPlogs/css** directory on the control workstation. This file contains trace information for the **Ecommands**.

spd.trace

The **spd.trace** file is in the **/var/adm/SPlogs/css** directory on the control workstation. This file contains messages about the SP Switch advanced diagnostics tests and architecture components. See "Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools" on page 245.

Missing error data warning

When the system is miswired, the fault service daemon cannot always detect the problem. The daemon updates the file **cable_miswire**, where it reports all the miswires it finds. When reporting a miswire problem, an entry may be missing from this file, or an entry may be incorrect.

If a node is powered on before the switch board receives power, a **diag_fail** condition may occur. See "diag_fail condition for an SP Switch adapter" on page 173.

Information to collect before contacting the IBM Support Center

The following items are used to isolate problems in the SP Switch component of PSSP. More detailed information about each item appears in “Error information” on page 138.

css.snap package

The **/usr/lpp/ssp/css/css.snap** script collects log, trace, and dump information that are created by SP Switch support code (device driver, worm, fault-service daemon, diagnostics) into a single compressed package.

The complete package output file is: **/var/adm/SPlogs/css/***hostname*.*yymmddhhmmss*.**css.snap.tar.Z** where *hostname* is the host name of the node where **css.snap** was issued, and *yymmddhhmmss* is the date and time that the **css.snap** information was collected.

The **css.snap** script creates a log file, **/var/adm/SPlogs/css/css.snap.log** where all the files gathered in the package are listed.

This script is called whenever a serious error is detected by the switch support code. To directly cause the system to create such a snapshot, login the desired node and manually issue the command:

```
/usr/lpp/ssp/css/css.snap [-c | -n | -s ]
```

where

- **-c** Flushes the adapter cache and prints the result. This is the default.
- **-n** Assumes that the device driver or daemon has flushed the cache.
- **-s** Takes a soft snap, which does not dump the adapter state. This excludes the **tb_dump.out**. This flag is used for temporary errors (TYPE=TEMP) where the integrity of the adapter is in doubt, or when it is not desirable to corrupt the adapter state by the use of diagnostic routines.

This table shows the error log entries that automatically take a snapshot, as well as the type of snap performed. The **soft** type enables a continuation of work with the switch. The **full** snap *might* corrupt the adapter, forcing an adapter reset and the node to be fenced off of the switch.

Table 24. AIX Error Log entries that invoke css.snap - SP Switch

Error Log entry	Snap type (full/soft)
SP_SW_FIFOOVRFLW_RE	soft
SP_SW_RECV_STATE_RE	soft
SP_SW_INVALID_RTE_RE	soft
SP_SW_NCLL_UNINT_RE	soft
SP_SW_PE_INBFIFO_RE	soft
SP_SW_PE_ON_DATA_RE	soft
SP_SW_PE_ON_NCLL_RE	soft
SP_SW_PE_RTE_TBL_RE	soft
SP_SW_SNDLOSTEOP_RE	soft
TB3_CONFIG1_ER	full
TB3_LINK_ER	full

Table 24. AIX Error Log entries that invoke css.snap - SP Switch (continued)

Error Log entry	Snap type (full/soft)
TB3_PIO_ER	soft
TB3_SVC_QUE_FULL_ER	full
TB3_THRESHOLD_RE	full

Collect the **css.snap** information from both the primary node and all nodes that are experiencing SP Switch problems. Do not reboot the nodes before running **css.snap**, because rebooting causes the loss of valuable diagnostic information.

The **css.snap** collects all the files which reside in the **/var/adm/SPlogs/css** directory, some additional files from the **/tmp** directory, and the last available stack file from the **/var/adm/ffdc/stacks** directory. Some of the files reside in each node, while others reside only on the primary nde or on the control workstation.Table 25 lists important files gathered by **css.snap**, and their location at **css.snap** time. Some of the files are created by the **css.snap** in order to gather concurrent information on the switch status.

Log files within css.snap package

The list of files collected by **css.snap** and their location is given in this table:

Table 25. SP Switch log files and their location

Number	Log file name	Contents	Location
1	cable_miswire	node-to-switch or switch-to-switch miswired connection information. For more information, see "cable_miswire" on page 163.	primary node
2	cable_miswire.old	previous miswire information. For more information, see "cable_miswire" on page 163.	primary node
3	core	fault service daemon dump file.	nodes
4	cssadm2.debug	Trace of **cssadm** daemon. For more information, see "cssadm.debug" on page 157.	control workstation
5	cssadm2.stderr	Unexpected error messages received by the **cssadm** daemon. For more information, see "cssadm.stderr" on page 157.	control workstation
6	cssadm2.stdout	Unexpected informational messages received by the **cssadm** daemon. For more information, see "cssadm.stdout" on page 157.	control workstation
7	css_dump.out	most recent **css_dump -r**. SP switch adapter device driver trace buffer dump file. For more information, see "css_dump.out" on page 156.	nodes

Table 25. SP Switch log files and their location (continued)

Number	Log file name	Contents	Location
8	css.snap.log	**css.snap** snapshot script information. A list of all files gathered in the last snapshot.	nodes
9	CSS_test.log	Present if **CSS_test** was run on the node. See “CSS_test.log” on page 158 and “Verify software installation” on page 174. For more information about the **CSS_test** command, see *PSSP: Command and Technical Reference*.	nodes
10	daemon.stderr	fault service daemon output file (stderr). For more information, see “daemon.stderr” on page 162.	nodes
11	daemon.stdout	fault service daemon output file (stdout). For more information, see “daemon.stdout” on page 162.	nodes
12	dist_topology.log	system error messages occurring during the distribution of the topology file to the nodes. For more information, see “dist_topology.log” on page 161.	primary node
13	dtbx.failed.trace	SP Switch adapter diagnostics last failed run information. For more information, see “dtbx.failed.trace” on page 163.	nodes
14	dtbx.trace	SP Switch adapter diagnostics messages. For more information, see “dtbx.trace” on page 162.	nodes
15	Eclock.log	**Eclock** related information	control workstation
16	Ecommands.log	log entries of all **Ecommands**. For more information, see “Ecommands.log” on page 163.	control workstation
17	errpt.out	most recent **errpt -a** and **errpt** results. See “errpt.out” on page 156. For more information about the **errpt** command, see *AIX Command and Technical Reference*.	nodes
18	flt	hardware error conditions found on the SP Switch, recovery action taken by the fault service daemon, and general operations that alter the SP Switch configuration. For more information, see “flt” on page 160.	nodes
19	fs_daemon_print.file	fault service daemon status information. For more information, see “fs_daemon_print.file” on page 162.	nodes

Table 25. SP Switch log files and their location (continued)

Number	Log file name	Contents	Location
20	fs_dump.out	most recent **fs_dump -r**. Fault service kernel extension trace buffer dump file. For more information, see “fs_dump.out” on page 157.	nodes
21	logevnt.out	Log error log events monitored by **ha**. For more information, see “logevnt.out” on page 162.	nodes
22	netstat.out	current **netstat -I css0** and **netstat -m** information. Network status information. See “netstat.out” on page 159. For more information about the **netstat** command, see*AIX Command and Technical Reference*.	nodes
23	out.top	SP switch link information. For more information, see “out.top” on page 161.	nodes
24	rc.switch.log	fault service daemon initialization information. For more information, see “rc.switch.log” on page 161.	nodes
25	regs.out	most recent **read_regs**. SP Switch adapter registers dump file. For more information, see “regs.out” on page 158.	nodes
26	router.log	SP Switch routing information For more information, see “router.log” on page 158.	nodes
27	scan_out.log	TBIC scan ring binary information. For more information, see “scan_out.log and scan_save.log” on page 159.	nodes
28	scan_save.log	previous TBIC scan ring binary information. For more information, see “scan_out.log and scan_save.log” on page 159.	nodes
29	stack file: fault_service_Worm_RTG_SP *.PID.date.time*	Customer Service related detailed error information with links between flow related errors. This is a binary format file.	nodes
30	spd.trace	Tracing of advanced switch diagnostics. See “spd.trace” on page 163.	control workstation
31	spdata.out	most recent **css.snap**’s **splstdata** dump file. SP data requests. For more information, see “spdata.out” on page 159.	
32	summlog.out	Error information from the **css.summlog** daemon. For more information, see “summlog.out” on page 162.	control workstation

Table 25. SP Switch log files and their location (continued)

Number	Log file name	Contents	Location
33	tb_dump.out	most recent ′**tb3dump**, **tb3mxdump**, or **tb3pcidump** according to the adapter card type: TB3, TB3MX or TB3PCI. Adapter memory dump file. For more information, see “css_dump.out” on page 156.	nodes
34	vdidl.out	most recent ′**vdidl3**, **vdidl3mx**, or **vdidl3pci** according to the adapter card type: TB3, TB3MX or TB3PCI. Fault service kernel extension IP services dump file. For more information, see “vdidl.out” on page 159.	nodes
35	worm.trace	switch initialization information. For more information, see “worm.trace” on page 161.	nodes

The files ending in **.out** are produced by running the appropriate command to dump internal (in memory) trace information or dump data to a file.

Disk space handling policy

The **css.snap** command avoids filling up the **/var** directory by following these rules:

1. If less than 10% of **/var** is free, **css.snap** exits.
2. If the css portion of **/var** is more than 30% of the total space in **/var**, **css.snap** erases old snap files until the css portion becomes less than 30%. If it is successful, the snap proceeds. If not, it exits.

The **css.snap** command is called automatically from the fault service daemon when certain serious errors are detected. It can also be issued from the command line when a switch or adapter related problem is suspected. See “css.snap package” on page 164.

Note: **css.snap** uses a number of undocumented utilities to collect information. Some of these can be destructive when used on a running system. After using **css.snap** to collect diagnostic information, it is advisable to run **/usr/lpp/ssp/css/rc.switch** in order to reset and reload the switch adapter and eliminate the residual effects of these utilities.

Diagnostic procedures

If your SP system or SP system partition shows signs of a switch failure, locate the symptom and perform the recovery action described. All the recovery actions described require that the user have **root** access on the specified SP Switch node.

Note: If your system is running in restricted root access mode, the following commands must be issued from the control workstation:

- **CSS_test**
- **Eclock**
- **Efence**
- **Eprimary**

- **Equiesce**
- **Estart**
- **Eunfence**
- **Eunpartition**
- **mult_senders_test**
- **switch_stress**
- **wrap_test**

SP Switch diagnostics

Verify the SP Switch topology configuration

The switch topology file is used to define the hardware configuration to the css support software. It should reflect the number of switches and nodes installed, as well as define how they are connected.

The topology file can reside in two places: in the SDR, or in the **expected.top** file in the **/etc/SP** directory of the primary node. Usually, the configuration in the SDR is used. If the configuration in **/etc/SP/expected.top** on the primary node exists, it overrides the configuration in the SDR. The **/etc/SP/expected.top** on the primary node is generally used for debugging proposes.

To verify that the topology in the SDR is correct, first read it out of the SDR using the command:

```
Etopology -read file_name
```

The **Etopology** command reads the switch topology from the SDR and places it in the specified file. For more information on this command, see *PSSP: Command and Technical Reference*.

Once the file is extracted, verify that the switch topology is an accurate representation of the installed hardware.

If changes to the switch topology file are required, remember to place them back into the SDR using the **Etopology** command:

```
Etopology file_name
```

A default set of topology configuration files is available in the **/etc/SP** directory. For more information, see *PSSP: Command and Technical Reference*.

The SP Switch uses an annotated topology file, produced by the **Eannotator** command. The system administrator is responsible to run the command and create the annotated topology file. When the file is not annotated, the fault service daemon will still work and the switch will function, but the switch jack numbers will not be correct. If you suspect that your topology file has a problem, you can verify it. Examine the **out.top** file in **/var/adm/SPlogs/css** of each node, and examine the topology file (using the **Etopology -read** *file_name* option as described previously).

Each link line in an annotated file is marked by E as in the example:

```
E01-S17-BH-J7 to E01-N1
```

Each link line in a file that is not annotated is marked by L as in the example:

```
L01-S17-BH-J7 to L01-N1
```

Verify the System Data Repository (SDR)

To verify that the SDR is installed and operating correctly, run the **SDR_test** command on the control workstation. It can be run either through **SMIT** panels or from the command line.

To verify SDR installation and operation from the **SMIT** panel:

1. Issue the command:

   ```
   smit SP_verify
   ```
2. The **RS/6000 SP Installation/Configuration Verification** menu appears.
3. Select: **System Data Repository**.
4. Press enter.
5. Review the output created.

To verify SDR installation and operation from the command line, enter:

```
/usr/lpp/ssp/bin/SDR_test
```

Review the output created.

Whenever the **SDR_test** command is run, a log file is created to enable the user to review the test results. The default log file created is: **/var/adm/SPlogs/SDR_test.log**. If the **SDR_test** command is run without **root** authority, the default log file created is: **/tmp/SDR_test.log**. Complete information on **SDR_test** can be found in *PSSP: Command and Technical Reference*. See also "SDR verification test" on page 130.

Next, login to the failing node and issue the command:

```
SDRGetObjects switch_responds
```

Examine the output that is returned. If the switch responds bits are returned, this indicates that the SDR is operating. You can also determine which nodes are operational on the switch by examining the value returned: A value of 1 indicates that the node is operational. A value of 0 indicates that the node is not operational.

Clock diagnostics

The following procedure should **not** be run on nodes that are operational on the switch. The utilities used for those verifications cannot coexist with normal switch operations on the node. If a clocking problem exists on all nodes in a rack, see "SP rack or system clock diagnostics" on page 171. Otherwise, see "SP Switch external clock diagnostics".

SP Switch external clock diagnostics

Perform these steps to determine if the external clock is operational at the node:

1. Login to the node in question.
2. Issue the following command:

   ```
   /usr/lpp/ssp/css/diags/read_tbic -s
   ```
3. The output is similar to:

```
TBIC status register   :78XXXXXX
```

4. Look at bit 3 and 4 (the bits are numbered from left to right, starting with 0):
   - If bits 3 and 4 are ON (equal 1), the external clock is operational at the node.
   - If either bit 3 or 4 are OFF (equal 0), the external clock is not operational at that node.

   In this example, bits 3 and 4 are both ON, indicating that the external clock is operational at the node.

Perform the following steps to restore the external clock at the node:

1. **rc.switch** the node with the following command:

   ```
   /usr/lpp/ssp/css/rc.switch
   ```
2. The output is similar to:

   ```
   adapter/mca/tb3
   ```
3. Determine if the clock is still not present using the same **read_tbic** command from the previous procedure.
4. If the clock is still not operational, try to **Eclock** the system.

   The **Eclock** command affects all switch boards in the system and requires exclusive use of the switch, and therefore the SP system partitions on the switch. For more information on the **Eclock** command see *PSSP: Command and Technical Reference*.

   To **Eclock** the system, issue the command:

   ```
   Eclock -d
   ```
5. Determine if the clock is present after the **Eclock**. Use the same **read_tbic** command from the previous procedure.
6. If the clock is still not present, run the “Cable diagnostics” on page 173 or contact IBM Hardware Service and request that they run the “Cable diagnostics” on page 173.

SP rack or system clock diagnostics

The following table list the possible clock loss problems on single racks and systems along with their recovery actions:

Table 26. Clock problems and recovery actions

Problem	Recovery action
All nodes in a single rack will not clock.	**Cause:** SP Switch is powered off. **Action:** Power the switch on, run **Eclock -d**, then **Estart**. **Cause:** The switch is not **Eclock**ed **Action:** Run **Eclock -d**, then **Estart**. **Cause:** The clock topology file used does not match the physical system topology or the switch board is defective. **Action:** Contact IBM Hardware Service.

Table 26. Clock problems and recovery actions (continued)

Problem	Recovery action
Some racks in the system will not clock.	**Cause:** SP Switches are powered off. **Action:** Power the switches on, run **Eclock -d**, then **Estart**. **Cause:** The system is not **Eclock**ed. **Action:** Run **Eclock -d**, then **Estart**. **Cause:** The clock topology file used does not match the physical system topology or the master clock switch is bad. **Action:** Contact IBM Hardware Service.

SP Switch adapter diagnostics

The adapter diagnostics have two modes of operation: the Power-On-Self-Test (POST) and by issuing the **diag** command from the command line.

For the automatic POST tests scenario, issue the command:

```
diag -c -d css0
```

For advanced diagnostics scenarios, issue the command:

```
diag -A -d css0
```

The advanced tests check the cable wrap. You will need the card and cable wrap plug to complete these tests.

Note: The complete set of adapter diagnostics needs the exclusive use of the css adapter on the current node that the diagnostics are run on. Any other processes that have the css device driver open must be closed (**kill**ed) before issuing the adapter diagnostics command. One of those processes is the fault service daemon: **fault_service_Worm_RTG_SP**. Processes such as the ″switch clock reader application″ make use of the fault-service daemon, and therefore they should be closed as well.

The diagnostics failures are reported to the AIX error log of the failing node. To view the adapter diagnostics errors:

1. Login to the failing node.
2. Issue the command:

   ```
   errpt -a |grep "Switch adapter failed POST diagnostics"
   ```

   to view the POST adapter diagnostics AIX error log entries.
3. In most cases each of the error entries will contain a Service Request Number (SRN).
4. Use the SRN to locate your error and its recovery actions in Table 27 on page 173.
5. Note that **x** may represent any value in Table 27 on page 173.

Table 27. SP Switch adapter Service Request Number failures and recovery actions

SRN	Recovery action
1xx	See “Verify software installation” on page 174. If the verification is successful and the problem persists, contact the IBM Support Center.
28x	See “Clock diagnostics” on page 170. If the verification is successful and the problem persists, contact the IBM Support Center.
Axx	See “Clock diagnostics” on page 170. If the verification is successful and the problem persists, contact the IBM Support Center.
All other SRNs	Contact IBM Hardware Service and arrange to have the adapter or cable replaced.

diag_fail condition for an SP Switch adapter

If a node is powered on before the switch boards receive power, the node’s switch adapter does not receive a clock signal. This is a situation that could occur during installation. This causes a **diag_fail** condition for the SP Switch adapter. If the adapter status is **diag_fail**, the **rc.switch** process terminates without starting the fault service daemon.

There are two ways to correct the problem:

1. Reboot the node.
2. Reconfigure the adapter. This requires stopping any process that could hold the css0 device, such as Topology Services (**hats**). Issue these command:
   a. stopsrc -s hats
   b. /usr/lpp/ssp/css/ucfgtb3 -l css0 -v
   c. /usr/lpp/ssp/css/cfgtb3 -l css0 -v
   d. startsrc -s hats

Note: If there are other processes, including application processes, holding the css0 device, these processes must be stopped and restarted as well.

Cable diagnostics

Switch to switch cable diagnostics

Visually inspect the cable in question:

1. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If everything looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them repair or replace the damaged components.
2. Repeat step 1 for the other end of the switch to switch cable.
3. Run the SP Switch Wrap Test and SP Switch Stress Test. See “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.
4. If everything checks out, contact IBM Hardware Service and have them replace the cable. If the problem persists, contact the IBM Support Center.

Node to switch cable diagnostics

Visually inspect the cable in question:

1. Remove the cable from the back of the node and examine the connectors (cable and back of the adapter) for bent pins or other visible damage. If everything looks OK, reconnect the cable to the adapter. If not, contact IBM Hardware Service and have them repair or replace the damaged components.
2. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If everything looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them repair or replace the damaged components.
3. Run the SP Switch Wrap Test and SP Switch Stress Test. See "Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools" on page 245.
4. If everything visually checks out, run advanced Adapter Diagnostics on the suspect adapter. The procedure is outlined in "SP Switch adapter diagnostics" on page 172. Follow the online instructions. If the diagnostics detect a failure, contact IBM Hardware Service and have them replace the failing components. If the adapter diagnostics pass and the problem persists, contact the IBM Support Center.
5. As a result of removing the cable, the node may be automatically fenced by the system. After reinstalling the cable, reboot the node or run the **rc.switch** command to reset the switch adapter. Only after this is complete, try to **Eunfence** the node.

SP Switch node diagnostics

Identify the failing node

Use this scenario if an application running on several nodes loses connectivity over the switch, or the **switch_responds** class indicates that several nodes are not on the switch. For more information on the **switch_responds** class, see the **SDRGetObjects** entry in *PSSP: Command and Technical Reference*.

1. View the summary log file, located on the control workstation.

   See "Summary log for SP Switch, SP Switch2, and switch adapter errors" on page 69.
2. Locate the first AIX error log entry that indicates a node or connectivity failure.
3. Examine other entries to see if the first failure is the cause of subsequent failures.
4. On the node that experienced the first failure, examine the AIX error log to see the complete version of the record described previously.
5. Use this as a starting point to debug the problem on this node.

Verify software installation

The software installation and verification are done using the **CSS_test** command on the control workstation. It can be run either through **SMIT** panels or the command line.

If you are using SP system partitions, **CSS_test** runs in the active SP system partition only. For more information on managing system partitions, see *PSSP: Administration Guide*.

If **CSS_test** is issued following a successful **Estart**, additional verification of the system is done to determine if each node in the system or system partition, can be **ping**ed.

To verify CSS installation from the **SMIT** panels:

1. Issue:

   ```
   smit SP_verify
   ```

2. The **RS/6000 SP Installation/Configuration Verification** menu appears.
3. Select: **Communications Subsystem**.
4. Press enter.
5. Review the output created.

Whenever the **CSS_test** command is run, a log file is created to enable the user to review the test results. The file is **/var/adm/SPlogs/CSS_test.log**. Complete information on **CSS_test** can be found in *PSSP: Command and Technical Reference*.

To verify CSS installation from the command line:

1. Issue the command:

   ```
   /usr/lpp/ssp/bin/CSS_test
   ```

2. Review the log file.

When running **CSS_test**, consider the following:

- The directory **/usr/lpp/ssp** on each of the nodes in the system partition should be accessible (execute and read permissions) to the user who performs the test.
- The script file **/etc/inittab** on each node should contain an entry for the script **rc.switch**.

Verify SP Switch node operation

Use this procedure to verify that a single SP Switch node is operating correctly. If the node you are attempting to verify is the primary node, start with Step 1. If it is a secondary node, start with Step 2.

1. Determine which node is the primary by issuing the **Eprimary** command on the control workstation. For complete information on the **Eprimary** command, see *PSSP: Command and Technical Reference*. For our purposes, consider this output:

   ```
   1 - primary
   2 - oncoming primary
   26 - primary backup
   26 - oncoming primary backup
   1 - autounfence
   ```

   If the command returns an oncoming primary value of none, reissue the **Eprimary** command, specifying the node you would like to have as the primary node. Following the completion of the **Eprimary** command (to change the oncoming primary) an **Estart** command is required to make the oncoming primary node the primary.

   If the command returns a primary value of none, an Estart is required to make the oncoming primary node the primary.

   The primary node on the SP Switch system can move to another node, if a primary node takeover is initiated by the backup. To determine if this has happened, look at the values of the primary and the oncoming primary backup. If they are the same value, then a takeover has occurred.

2. Ensure that the node is accessible from the control workstation. This is done by using the **dsh** command to issue the **date** command on the node as follows:

   ```
   /usr/lpp/ssp/bin/dsh -w problem_hostname date
   ```

   The output is similar to:

   ```
   TUE Jan 25 10:24:28 EDT 2000
   ```

   If the current date and time are not returned, refer to “Chapter 20. Diagnosing remote command problems on the SP System” on page 299.
3. Verify that the switch adapter (css0) is configured and is ready for operation on the node. This can be done by examining the **adapter_config_status** attribute in the **switch_responds** object of the SDR:

   ```
   SDRGetObjects switch_responds node_number==problem_node_number\
   node_number switch_responds autojoin isolated adapter_config_status
   ```

   The output is similar to:

   ```
   node_number switch_responds autojoin isolated adapter_config_status
    1              0               0       0        css_ready
   ```

   If the **adapter_config_status** object is anything other than `css_ready`, see “Adapter configuration error information” on page 152.

   To obtain the value to use for *problem_node_number*, issue an SDR query of the **node_number** attribute of the **Node** object, as follows:

   ```
   SDRGetObjects Node reliable_hostname==problem_hostname node_number
   ```

   The output is similar to the following:

   ```
   node_number
      1
   ```
4. Verify that the **fault_service_Worm_RTG_SP** daemon is running on the node. This can be accomplished by using the **dsh** command on the control workstation to issue a **ps** command to the problem node as follows:

   ```
   /usr/lpp/ssp/bin/dsh -w problem_hostname ps -e | grep Worm_RTG
   ```

   The output is similar to the following:

   ```
   18422  -0:00 fault_service_Worm_RTG_SP
   ```

   If the **fault_service_Worm_RTG_SP** daemon is running, SP Switch node verification is complete.

   If the **fault_service_Worm_RTG_SP** daemon is not running, see “AIX Error Log information” on page 138. The possible reasons why the **fault_service_Worm_RTG_SP** daemon is not running are:
   - The daemon exited due to an abnormal error condition.
   - A **SIGTERM**, **SIGBUS**, or **SIGDANGER** signal was processed by the daemon.

Node crash

A node crash is generally identified by the LED/LCD display on the node flashing **888**. Do **not** reboot the node. See “Chapter 5. Producing a system dump” on page 81.

SP Switch advanced diagnostics

To examine the SP Switch fabric in more detail, see “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.

Error symptoms, responses, and recoveries

If your system or system partition shows signs of a switch failure, locate the symptom and perform the recovery action described. All the recovery actions described require that the user have **root** access on the specified switch node. If any of the recovery actions fail to resolve your problem, contact IBM Support Center.

Note: If your system is running in restricted root access mode, the following commands must be issued from the control workstation:

- **CSS_test**
- **Eclock**
- **Efence**
- **Eprimary**
- **Equiesce**
- **Estart**
- **Eunfence**
- **Eunpartition**
- **mult_senders_test**
- **switch_stress**
- **wrap_test**

SP Switch symptoms and recovery actions

Table 28 lists the known symptoms of a failure in the SP Switch, and points the user to the location of the detailed diagnostics and recovery action. You may have a symptom that does not appear in the table. In this case, view the error entry in the AIX error log and see “AIX Error Log information” on page 138.

Table 28. SP Switch symptoms and recovery actions

Symptoms	Recovery actions
Estart failure: 1. System cannot find **Estart** command. 2. Primary node is not reachable. 3. **Estart** command times out or fails. 4. Expected number of nodes not initialized. 5. Some links do not initialize.	1. See “Verify software installation” on page 174. 2. See “Verify SP Switch node operation” on page 175. 3. See “Estart error recovery” on page 180. 4. See “SP Switch device and link error information” on page 154. 5. See “SP Switch device and link error information” on page 154.
Nodes drops off of the switch. (**switch_responds** is OFF for the node).	See “Verify SP Switch node operation” on page 175.

Table 28. SP Switch symptoms and recovery actions (continued)

Symptoms	**Recovery actions**
Nodes fails to communicate over the switch, but its **switch_responds** is ON (**ping** or **CSS_test** commands fail).	1. See "Verify SP Switch node operation" on page 175. 2. See "AIX Error Log information" on page 138.
Node crash.	See "Node crash" on page 177.
Node fails to **Eunfence**.	1. See "Unfence an SP Switch node" on page 181. 2. See "Eunfence error recovery" on page 182.
Oncoming primary node is fenced	1. See "Unfence an SP Switch node" on page 181. 2. See "Eunfence error recovery" on page 182.
Ecommand failure.	See "Ecommands error recovery" on page 180.
diag_fail condition for an SP Switch adapter	See "diag_fail condition for an SP Switch adapter" on page 173.
switch_responds is still ON after node panic.	See "switch_responds is still on after node panic" on page 183.

Recover an SP Switch node

You can restart the **fault_service_Worm_RTG_SP** daemon on the node by issuing:

```
/usr/lpp/ssp/css/rc.switch
```

Following the **rc.switch**, run this command to determine if the daemon is still running or has died:

```
ps -e | grep Worm
```

At this point you should be able to **Eunfence** the node by issuing:

```
Eunfence problem_node_number
```

The output is similar to the following:

```
All nodes successfully unfenced.
```

If you cannot resolve the problem, contact the IBM Support Center. You should also attempt to gather any log files that are associated with this failure. See "Information to collect before contacting the IBM Support Center" on page 164.

Worm error recovery

The following steps enable you to recover from switch initialization failures which impact the worm subsystem.

1. Login to the failing node. (see **Eprimary** command).
2. View the bottom of the file **/var/adm/SPlogs/css/worm.trace**. Look for a message similar to one of the following, where **xx** represents any number:
   - TBSworm_bfs_phase1() failed with rc=xx
   - TBSworm_bfs_phase2() failed with rc=xx

3. Use the rc (return code) number as entry in Table 29.
4. If the return code cannot be found in the table, or the actions taken did not correct the problem, contact the IBM Support Center.

Table 29. SP Switch worm return codes and analysis

Return code	Analysis
-3	**Explanation:** Local adapter receiver port is not enabled. **Cause:** The switch is not clocked. **Action:** From the Control Work Station (CWS) issue the command **Eclock -d** followed by the command **Estart**. **Cause:** Oncoming primary is fenced off the switch. **Action:** See "Unfence an SP Switch node" on page 181.
-4	**Explanation:** Unable to generate routes for the network. **Cause:** Corrupted topology file. **Action:** See "Verify the SP Switch topology configuration" on page 169.
-5	**Explanation:** Send packet from local node failed. **Cause:** Bad switch adapter. **Action:** Run the switch adapter diagnostics on the failing node. If diagnostics fails to isolate the problem, contact the IBM Support Center.
-6	**Explanation:** A switch miswire was detected. **Cause:** Switch network cabling does not match the switch topology file. **Action:** View the **/var/adm/SPlogs/css/cable_miswire** file to determine which cables are in question. Then disconnect and check the associated cables. If the problem persists, contact IBM Hardware Service.
-7	**Explanation:** A node miswire was detected. **Cause:** Switch network cabling does not match the switch topology file. **Action:** The device is not cabled properly. There are two possible causes for this condition: the switch network is miswired or the frame supervisor's tty is not cabled properly. First view the **/var/adm/SPlogs/css/cable_miswire** file. Verify and correct all links listed in the file. Then issue the command **Eclock -d** followed by **Estart**. If the problem persists, contact IBM Hardware Service.
-8	**Explanation:** Receive FIFO is full. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the failing node. If diagnostics fails to isolate the problem, contact the IBM Hardware Service. **Cause:** The switch is backed up from a node or a switch chip. **Action:** Contact the IBM Support Center.
-9	**Explanation:** Unable to initialize FIFOs. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the failing node. If diagnostics fails to isolate the problem, contact IBM Hardware Service.

Table 29. SP Switch worm return codes and analysis (continued)

Return code	Analysis
-27	**Explanation:** The TBIC was not initialized. **Cause:** The switch adapter is uninitialized. **Action:** Run the script **rc.switch** on the failing node, then issue the **Estart** command from the control workstation. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the failing node. If diagnostics fails to isolate the problem, contact IBM Hardware Service.
-36	**Explanation:** This node resigned as the primary node. **Cause:** The node determined it could no longer control and monitor the SP Switch. The primary backup node is now in control of the SP Switch. **Action:** No action required.
-43	**Explanation:** A read or write operation to the switch adapter failed. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the failing node. If diagnostics fails to isolate the problem, contact IBM Hardware Service.

Ecommands error recovery

Error isolation for any of the **Ecommands** (**Eclock**, **Eannotator**, and others) is as follows:

1. View the error output returned from the command. Note the error message number and text.
2. Find the message in *PSSP: Messages Reference*.
3. More information can be obtained from the **Ecommands.log** trace file, see “Ecommands.log” on page 163.
4. Perform the recommended recovery action.
5. If the **Ecommand** failed because it was unable to communicate with every node, see “Chapter 18. Diagnosing SP Security Services problems” on page 251.
6. If the **Ecommand** failed because it cannot access the SDR, or the SDR is set up incorrectly, see “Verify the System Data Repository (SDR)” on page 170.
7. After the recovery action is taken, if the problem persists, see “Information to collect before contacting the IBM Support Center” on page 164 and contact the IBM Support Center.

Estart error recovery

To isolate and recover from failures in the **Estart** command, follow these steps:

1. Login to the primary node.
2. View the bottom of the file **/var/adm/SPlogs/css/flt** file.
3. Use the failure message as an index to Table 30 on page 181.
4. If the failure message cannot be found in the table, or the actions taken did not correct the problem, contact the IBM Support Center.

Table 30. SP Switch Estart problems and analysis

Error	**Analysis**
Error in **buildDeviceDatabase()**	**Explanation:** Unable to build the device database. **Cause:** Missing or corrupted topology file. **Action:** See "Verify the SP Switch topology configuration" on page 169. **Cause:** malloc failures. **Action:** See "Information to collect before contacting the IBM Support Center" on page 164 and contact the IBM Support Center.
Error in **TBSswitchInit()**	**Explanation:** Unable to initialize the switch network. **Cause:** Switch initialization failed. **Action:** See "Worm error recovery" on page 178.
Error in **writeDeviceDatabase()**	**Explanation:** Unable to write **/var/adm/SPlogs/css/out.top**. **Cause:** Missing or corrupted topology file. **Action:** See "Verify the SP Switch topology configuration" on page 169. **Cause:** The **/var** file system is not large enough to accommodate the new **out.top** file. **Action:** Increase the size of **/var**.
No valid backup - SDR current Backup being changed to none	**Explanation:** Informational message. **Cause:** No node available as a backup. **Action:** No action required.
Cannot access SDR - SDR current Backup not changed	**Explanation:** SDR failure **Cause:** SDR not set up properly. **Action:** See "Verify the System Data Repository (SDR)" on page 170
Error in: • **fopen(act.top.***PID***)** • **fprintf(act.top.***PID***)** • **fclose(act.top.***PID***)** • **rename(act.top, act.top.***PID***)**	**Explanation:** An error occurred accessing file **/var/adm/SPlogs/css/act.top.***PID*. **Cause:** File access problems. **Action:** Evaluate the errno returned and take the appropriate action. If the problem persists contact the IBM Support Center.

Unfence an SP Switch node

The recovery action to take depends on the current status of the SP Switch, and the personality of the switch node to be unfenced. The SP Switch status is limited to whether it is operational or not. The personality of the switch node to unfence is whether or not the node is to become the primary node, primary backup node, or a secondary node. For more information on any of the commands used in this section, see *PSSP: Command and Technical Reference*.

To display the switch primary node and primary backup node, issue the command:

```
Eprimary
```

Example output is:

```
none - primary
2 - oncoming primary
none - primary backup
26 - oncoming primary backup
```

In this example, no primary node is available. Therefore, the SP Switch is not operational.

1. SP Switch is operational and the node is to be the secondary. The node to unfence is not listed as the primary or the oncoming primary, and there is a primary node.

   Use the command **Eunfence** to unfence the node.
2. SP Switch is operational and the node is to be the primary. The node to unfence is the oncoming primary and another node is currently the primary.
   - Use the command **Eunfence** to unfence the node.
   - Use the command **Estart** to set the node to its primary personality.
3. SP Switch is not operational and the node is to be the secondary: no node is listed as primary, and another node is listed as the oncoming primary.
   - Use the command **Estart** to operate the SP Switch.
   - Use the command **Eunfence** to unfence the node.
4. SP Switch is not operational and the node is to be the primary. No node is listed as primary, and the fenced node is listed as the oncoming primary.
   - Use the command **Eprimary** to set another node as oncoming primary.
   - Use the command **Estart** to operate the SP Switch.
   - Use the command **Eunfence** to unfence the node.
   - Use the command **Eprimary** to set the unfenced node to be the oncoming primary.
   - Use the command **Estart** to set the node as the switch primary.

Eunfence error recovery

This section is used to help you when you failed to unfence your node, following the unfence procedure described in “Unfence an SP Switch node” on page 181.

To isolate and correct most **Eunfence** problems, you should refer first to “Ecommands error recovery” on page 180.

The following list provides additional reasons for a particular node to fail to **Eunfence**:

1. The **Eunfence** of a node failed, but the SP Switch was not **Estart**ed. You cannot attempt to **Eunfence** any node on an SP Switch that is not started. Issue the **Estart** command.
2. The node can no longer be reached through the switch network. More information can be gathered from the **out.top** trace file, see “out.top” on page 161.
3. The SP Switch node failed to **Eunfence** because the switch topology could not be distributed. See “Chapter 18. Diagnosing SP Security Services problems” on page 251.
4. The node failed to respond when attempting to **Eunfence** it. See “SP Switch node diagnostics” on page 174 to isolate and correct the problem.
5. The user receives the message ″Cannot Unfence node xxx - timeout″, the most likely cause is that the fault service daemon (**fault_service_Worm_RTG_SP**) is

not running on the node. If the this is the case, issue the **/usr/lpp/ssp/css/rc.switch** command to start the daemon. If the daemon is still not running, refer to the **rc.switch.log** trace file. See “rc.switch.log” on page 161.

6. The user receives the message ″Cannot Unfence node xxx **-** timeout″, and you have replaced the switch cable. See “Cable diagnostics” on page 173. Even though the fault service daemon (**fault_service_Worm_RTG_SP**) is running, you must issue the **/usr/lpp/ssp/css/rc.switch** command to reload and reset the adapter before you can try again to **Eunfence** the node.
7. If any of the preceding procedures fail to resolve the problem, and the node is still fenced, gather the css logs of the primary node and the fenced node. This can be accomplished by logging into those nodes and issuing the **/usr/lpp/ssp/css/css.snap** command. See “Information to collect before contacting the IBM Support Center” on page 164.

switch_responds is still on after node panic

This section addresses the case where a node panics, **host_responds** becomes OFF and **switch_responds** is still ON. This is a valid condition when the SP Switch adapter, on the crashed node, has no outstanding requests to or from this node. The SP Switch is now in a state where it can become backlogged, since the link is still marked as up. This can cause problems on other parts of the SP Switch network.

Each node fault-service daemon is responsible for updating its **switch_responds**. The SP Switch primary node detects fallen links and turns off **switch_responds** of faulty ones. The **switch_responds** is turned on only during **Estart** or **Eunfence** command processing. A node panic with **switch_responds** ON is a legitimate occurrence. There are two cases:

1. With primary or backup SP Switch nodes running, the **switch_responds** is updated only after a packet is sent to the panicked node. Therefore, a user can change **switch_responds** by trying to **ping** the panicked node. Having HA (**hats** and **hags**) run on the nodes can remedy the situation, since they run IP packets between the nodes casually in order to check the links (LAN Adapter event).
2. Without primary or backup SP Switch node running, there is no switch control. In this case, **switch_responds** is not updated. Only a new **Estart** can correct this.

Chapter 16. Diagnosing SP Switch2 problems

This chapter discusses diagnostic procedures and failure responses for the SP Switch2 component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 233. All the recovery actions described require that the user have **root** access to the specified node. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 218.

Related documentation

The following publications provide information about the SP Switch2:

1. *PSSP: Planning, Volume 2*
2. *PSSP: Administration Guide*
3. *PSSP: Installation and Migration Guide*
4. *PSSP: Messages Reference*

   These chapters contain messages related to the SP Switch2:

   - 0028 - Switch Support Messages
   - 2510 - Switch Fault Service Daemon Messages
   - 2543 - Switch Admin Daemon (cssadm daemon) Messages
   - 2548 - SP Switch Advanced Diagnostic Messages
   - 2549 - SP Switch Advanced Diagnostic Messages
5. *PSSP: Command and Technical Reference*
6. *First Failure Data Capture Programming Guide and Reference*
7. *The RS/6000 SP Inside Out*, SG24-5374
8. *Understanding and Using the SP Switch*, SG24-5161

Requisite function

This is a list of the software and hardware directly used by the SP Switch2 component of PSSP. Problems with the requisite software and hardware may manifest themselves as error symptoms in the SP Switch2. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the SP Switch2, you should consider these software and hardware components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

1. System Data Repository (SDR) component of PSSP, running on the control workstation
2. Ethernet component of AIX - For SP Switch2 operation, there must be at least one connection between the control workstation and the primary node. SP Switch2 Time-Of-Day recovery depends on an Ethernet connection to all the nodes.
3. Group Services **hags**, Topology Services **hats**, and Hardware Monitor **hmmon** components of PSSP
4. SP System Security Services

   Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
5. SP Switch2 hardware monitor and control: **hardmon** component of PSSP.

6. Switch Fault service: the Fault Service daemon component of the SP Switch2 has to be operating on each node that is on the SP Switch2. When the daemon dies, the protocols (such as IP) are closed on this node. This is different than operation of the SP Switch. For the SP Switch, when the Fault Service daemon dies, the protocols continue, only SP Switch recovery is not available.
7. Switch Time Of Day (TOD): initiated and recovered by the **emasterd** daemon component of SP Switch2 software. This daemon runs on the control workstation and is responsible for selecting a suitable Master Switch Sequencer Node that maintain the Switch TOD. When the daemon dies, the Switch TOD recovery dies. The Switch TOD is maintained only to nodes that are on plane 0.

Internal SP Switch2 subsystem components

The SP Switch2 component contains the software subsystems listed here. This chapter has procedures to diagnose problems that are handled by these subsystems.

1. **worm** - The fault-service daemon part that explores and initialize the switch fabric.
2. **Ecommands** - Commands such as **Estart**, **Efence**, **Eunfence**, and others **Ecommands**.
3. **css** - The fault-service daemon part that handles switch recovery.
4. **Pseudo device driver** - Enables **Ecommands** to communicate with the fault-service daemon.
5. **Hardware Abstraction Layer (HAL)** - Enables the fault-service daemon to communicate with the adapters.
6. **Connectivity Matrix** - Enables IP and any other application to get information on the switch connectivity. IP uses it to get the node's route to a requested node. The fault-service daemon on each node maintains the concurrent switch connectivity information.
7. **CSS adapter device driver** - SP Switch2 connectivity
8. **CSS adapter microcode** - SP Switch2 connectivity
9. **CSS adapter Diagnostics** - Self test for the CSS adapter hardware. If it fails, the switch connectivity is lost.
10. **MSS Node Recovery Daemon: emasterd** - Running on the control workstation, this daemon initializes the Master Switch Sequencer (MSS) Node and monitors its function for Switch TOD recovery purposes.
11. **Switch Admin daemon: cssadm2** - Optional daemon running on the control workstation, this daemon monitors switch operation and will issue **Estart** when necessary.

Error information

SP Switch2 operation and recovery are initialized from the switch primary node. When the primary node fails, the primary backup node takes over and becomes the new switch primary node. For SP Switch2 related problems, the error messages that have been logged are found on the current primary node.

SP Switch2 Time Of Day (TOD) is maintained by the Master Switch Sequencer (MSS) node. The MSS node is selected and monitored from the control workstation by the **emasterd** daemon. For TOD related problems, the logged error messages are found on the control workstation.

If the Switch Admin daemon (**cssadm2**) is running on the control workstation, the logged error messages are found on the control workstation.

SP Switch2 log and temporary file hierarchy

All the SP Switch2 log and temporary files are organized in a directory hierarchy. Next to each directory the specified file level is given.

```
                              /var/adm/SPlogs/css     node
                               /                \
       /var/adm/SPlogs/css0 adapter   /var/adm/SPlogs/css1 adapter
                       /                              \
    /var/adm/SPlogs/css0/p0 port         /var/adm/SPlogs/css1/p0 port
```

Relevant files are found in the directories:

- **css** - for global node level log files.
- **css0** - for adapter level log files (the 0 in css0 is the adapter id).
- **css1** - for adapter level log files (the 1 in css1 is the adapter id).
- **css0/p0** - for port level log files for the css0 adapter (the 0 in p0 is the port number within the adapter).
- **css1/p0** - for port level log files for the css1 adapter (the 0 in p0 is the port number within the adapter).

In this chapter, whenever a temporary file is mentioned, the file level is given using this terminology:

1. **node** level - The file resides in the **/var/adm/SPlogs/css** directory. There will be only one file on the node with the same name. The file contains information relevant to all the adapters with all their ports.
2. **adapter** level - The file resides in the **/var/adm/SPlogs/css0** or **/var/adm/SPlogs/css1** directory, where 0 or 1 is the adapter ID. The file contains information relevant to a specific adapter, which affects all the ports of the adapter.
3. **port** level - The file resides in the **/var/adm/SPlogs/css0/p0** or **/var/adm/SPlogs/css1/p0** directory, where 0 or 1 is the adapter ID and 0 in the p0 is the port id. The file contains information relevant to the port of the adapter.

plane.info file

This file has a full path name of **/etc/plane.info**. It is created by the user who wishes to override the **SDR_config** switch to plane number calculations. This file is optional. This file consists of one line for each switch, and has the following format:

```
Frame#:Slot# Plane# Sequence#
```

The *Sequence#* is the switch number within the plane. For example, the first switch in plane 1 is sequence number 1, the second switch in plane 1 is sequence number 2, and the first switch in plane 2 is sequence number 1.

A sample file for a two-plane SP Switch2 system would be:

```
1:17 0 1
2:17 1 1
3:17 0 2
4:17 1 2
```

If something in this file is incorrect, the switch_plane and switch_plane_seq numbers in the **Switch** class of the SDR will reflect the errors. If one of your switches goes down or you have to disconnect it, the **SDR_config** command may try to renumber your switches if it does not see that switch. In this case, you can reserve the spot for that switch by creating an **/etc/plane.info** file consisting of what your system should look like when that broken switch is up and running.

The **/etc/plane.info** file should be deleted when no longer needed, or the **SDR_config** command will always use it to override its own calculations.

AIX Error Log information

In order to isolate an adapter or SP Switch2 error, first view the AIX error log.

- For switch related problems, login to the primary node. The **Eprimary** command lists the primary node by node number. In cases where the primary node failed, there will be no primary. In this case, login to the node listed under `oncoming primary`. In cases where the SP Switch2 continued working, it may have replaced the primary node several times. You cannot locate the primary node using the preceding methods. In this case, the only way to locate the primary is to look inside the log files.
- For adapter related problems, login to the suspect node.
- All the AIX log messages in this chapter contain five strings in the Detail Data section of each message:
  1. DETECTING MODULE - debug information
  2. ERROR ID - unprintable string with encoded information (FFDC)
  3. REFERENCE CODE - unprintable string with encoded information (FFDC)
  4. Adapter ID - the adapter the message refers to. The value **N** represents no specific adapter or port.
  5. Port Number - which of the adapter’s ports the message refers to

  The second and third strings enable error tracking. See *First Failure Data Capture Programming Guide and Reference*. The fourth and fifth strings give the level of this error (adapter or port), which can be used to point to more detailed log files.
- Once you are on the desired node, issue the AIX command:

  ```
  errpt | more
  ```

  Output is similar to the following:

  ```
  ERROR_ID TIMESTAMP   T CL Res Name      ERROR_Description
  34FFBE83 0604140393T T H   Worm  Switch Fault-detected by switch chip
  C3189234 0604135793  T H   Worm  Switch Fault-not isolated
  ```

The Resource Name (`Res Name`) in the error log gives you an indication of what resource detected the failure.

Table 31. Resource Name failure indications - SP Switch2

Resource name	**Indication**
Worm	The information was extracted from the SP Switch2 initialization.
css	Incorrect status was detected by the adapter (css device driver).
css0	Failed adapter diagnostics on adapter 0.
css1	Failed adapter diagnostics on adapter 1.

For a more detailed description, issue the AIX command:

```
errpt -a [-N resource_name] | more
```

where the optional *resource_name* is one of the entries in Table 31 on page 188.

There are several subcomponents that write entries in the AIX error log:

1. Adapter recovery errors, which are recognized by labels with prefixes: **CS_ADAPT_**, **CS_ATRANS_**, **CS_PTRANS_**, **CS_PTRANS**, **CS_PORT**, **CSS_DD**, **CSS_SLIH**, and **HACSSRMD**.

Table 32. Possible causes of adapter failures - SP Switch2

Label and Error ID	**Error description and analysis**
CS_PTRANS_HW_RE FECAAB29	**Explanation:** SP Switch2 adapter port - transient hardware error. **Cause:** SP Switch2 cable failure. **Action:** Check, reconnect, unfence, and if the problem persists, replace the cable. **Cause:** SP Switch2 adapter port hardware failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service. **Cause:** SP Switch2 failure. **Action:** Call IBM Hardware Service if the problem persists.
CS_PORT_HW_ER 984F7BA3	**Explanation:** SP Switch2 adapter port - permanent hardware error. **Cause:** SPSwitch2 cable failure. **Action:** Check, reconnect, unfence and if the problem persists, replace the cable. **Cause:** SP Switch2 adapter port hardware failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service. **Cause:** SP Switch2 failure. **Action:** Call IBM Hardware Service if the problem persists.
CS_ATRANS_HW_RE 506F0AF2	**Explanation:** SP Switch2 adapter - transient hardware error. **Cause:** SP Switch2 adapter hardware failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.

Table 32. Possible causes of adapter failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_ATRANS_MC_RE 6561F900	**Explanation:** SP Switch2 adapter - transient microcode error. **Cause:** SP Switch2 adapter microcode error. **Action:** If the problem persists: • Run adapter diagnostics. • Call the IBM Support Center
CS_ATRANS_SFW_RE 07E73229	**Explanation:** SP Switch2 adapter - transient software error. **Cause:** SP Switch2 adapter device driver error. **Action:** If the problem persists: • Run adapter diagnostics. • Call the IBM Support Center
CS_ATRANS_HW_MC_RE 218342C7	**Explanation:** SP Switch2 adapter - transient hardware or microcode error. **Cause:** SP Switch2 adapter failure. **Cause:** An adapter microcode failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.
CS_ATRANS_HW_SFW_RE 827413D0	**Explanation:** SP Switch2 adapter - transient hardware or software error. **Cause:** SP Switch2 adapter failure. **Cause:** An adapter device drive software failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.
CS_ADAPT_HW_RE CDECB780	**Explanation:** SP Switch2 adapter - critical hardware error. **Cause:** SP Switch2 adapter failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.
CS_ADAPT_MC_RE DF2AE96B	**Explanation:** SP Switch2 adapter - critical microcode error. **Cause:** An adapter microcode failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.

Table 32. Possible causes of adapter failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_ADAPT_SFW_RE 958E05ED	**Explanation:** SP Switch2 adapter - critical software error. **Cause:** SP Switch2 adapter device driver failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call the IBM Support Center
CS_ADAPT_HW_MC_RE D0999C1C	**Explanation:** SP Switch2 adapter - critical hardware or microcode error. **Cause:** SP Switch2 adapter failure. **Cause:** An adapter microcode failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call IBM Hardware Service.
CS_ADAPT_HW_SFW_RE DA7623E7	**Explanation:** SP Switch2 adapter - critical hardware or software error. **Cause:** SP Switch2 adapter failure. **Cause:** An adapter device drive software failure. **Action:** If the problem persists: • Run adapter diagnostics. • Call the IBM Support Center
CS_ADAPT_HW_ER DD4FECEA	**Explanation:** SP Switch2 adapter - permanent hardware error. **Cause:** SP Switch2 adapter failure. **Action:** • Run adapter diagnostics. • Call IBM Hardware Service.
CS_ADAPT_MC_ER 85421CF0	**Explanation:** SP Switch2 adapter - permanent microcode error. **Cause:** SP Switch2 adapter microcode failure. **Action:** • Run adapter diagnostics. • Call the IBM Support Center.

Table 32. Possible causes of adapter failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_ADAPT_SFW_ER F6E84D66	**Explanation:** SP Switch2 adapter - permanent software error. **Cause:** SP Switch2 adapter device driver failure. **Action:** • Run adapter diagnostics. • Call the IBM Support Center.
CS_ADAPT_HW_MC_ER 405FA51A	**Explanation:** SP Switch2 adapter - permanent hardware or microcode error. **Cause:** SP Switch2 adapter failure **Cause:** SP Switch2 adapter microcode failure. **Action:** • Run adapter diagnostics. • Call the IBM Support Center.
CS_ADAPT_HW_SFW_ER 035E8AD9	**Explanation:** SP Switch2 adapter - permanent hardware or software error. **Cause:** SP Switch2 adapter failure **Cause:** SP Switch2 adapter device driver failure. **Action:** • Run adapter diagnostics. • Call the IBM Support Center.
CSS_DD_DEBUG_ER 8ED77B7A	**Explanation:** SP Switch2 CSS device driver error. **Cause:** SP Switch2 device driver failure. **Action:** Call the IBM Support Center.
CSS_DD_CFG_ER 0195C376	**Explanation:** SP Switch2 CSS device driver configuration error. **Cause:** SP Switch2 device driver failure. **Action:** Run configuration method with verbose flag: **/usr/lpp/ssp/css/cfgcol -v -l css[0 | 1]** • Run the configuration method with the verbose option. • Call IBM Support Center.
CSS_SLIH_ER BAB33325	**Explanation:** SP Switch2 CSS device driver - an unexpected interrupt occurred. **Cause:** SP Switch2 adapter or SP Switch2 failure. **Action:** • See more logged information in the AIX error log. • Run adapter diagnostics. • If the problem persists, call IBM Hardware service.

Table 32. Possible causes of adapter failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
HACSSRMD_ERR C3E70E5D	**Explanation:** SP Switch2 CSS **hacssrmd** daemon terminated. **Cause:** Unknown failure. **Action:** If the problem persists, call the IBM Support Center.

2. SP Switch2 fault service daemon errors, which are recognized by labels with the prefix: **CS_SW_**.

Table 33. Possible causes of fault service daemon failures - SP Switch2

Label and Error ID	Error description and analysis
CS_SW_ADPT_TYPE_ER CC1DCEED	**Explanation:** The connected adapter type is not supported on SP Switch2. **Cause:** The user plugged an unsupported adapter or node into the SP Switch2 port. **Action:** Call the IBM Support Center.
CS_SW_SEND_HANG_RE 69AB5AEC	**Explanation:** A Sender Hang was detected. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_TKNCOUNTER_RE 3EF9DDC7	**Explanation:** A Token Counter Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_INIT_STATE_RE EB8CFA87	**Explanation:** Initialization State Machine error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_TOD_ECC_RE 35D40633	**Explanation:** Receiver TOD ECC Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CQ_PE_NCL_RE 066BD301	**Explanation:** Parity Error on Next Chunk Linked List occurred. **Cause:** SP Switch2 chip saw an error on a received package. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CQ_PE_FSL_RE E273ABC6	**Explanation:** Parity Error on Free Space Linked List occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CQ_SRM_EC_RE 5A19E266	**Explanation:** Source Routed Multicast ECC Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CQ_MCSRDT_RE 89316FCB	**Explanation:** Multicast Source Routed Decode Table Parity error occurred. **Cause:** A switch chip saw error on received package. **Action:** If the problem persists, call IBM Hardware Service.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_CQ_MCLRTD_RE A974CB87	**Explanation:** Multicast Lookup Table Route Decoder Parity Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CQ_RCA_PE_RE 27305E8F	**Explanation:** Repeat Count Array Parity Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_MULTICASTR_RE FADF4398	**Explanation:** Multicast Route Error occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_CHIP_ID_ER_RE 43D748CF	**Explanation:** Chip ID Error occurred. **Cause:** A switch chip configuration error. The hardware monitor daemon or the control workstation is down. **Action:** Check to see if the hardware monitor daemon is up. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_SVC_ARROVR_RE D32FD026	**Explanation:** Service Array Overflow Latch occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_PE_SVCARRI_RE BAEF6722	**Explanation:** Parity Error on input to Service Array occurred. **Cause:** SP Switch2 chip saw an error on a received package. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_PE_SVCARRO_RE 59D3D44A	**Explanation:** Parity Error on output to Service Array occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_INV_SVCCMD_RE D2833C50	**Explanation:** Invalid Service Command error occurred. **Cause:** SP Switch2 chip saw an error on a received package. **Action:** If the problem persists, call IBM Hardware Service.
CS_TOD_ERROR_RE A11A52E1	**Explanation:** Error occurred in TOD logic. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_CSS_IF_FAIL_ER B454D630	**Explanation:** SP Switch2 adapter service interface system call failed. **Cause:** Unable to communicate with the SP Switch2 adapter. **Action:** • Check the switch adapter configuration. • Run adapter diagnostics. See “SP Switch2 adapter diagnostics” on page 226. • If the problem persists, call IBM Hardware Service.
CS_SW_TKN_CNT_O_RE E5895205	**Explanation:** A Sender Token Count Overflow occurred. **Cause:** SP Switch2 chip failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_ACK_FAILED_RE D4E9B237	**Explanation:** SP Switch2 daemon failed to acknowledge a service command. **Cause:** SP Switch2 communication failure. **Cause:** A traffic backlog on SP Switch2 adapter. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_EDC_ERROR_RE E6E27F0C	**Explanation:** An EDC-class error was detected. **Cause:** A transient error in data occurred during transmission over switch links. The EDC error may be one of the following: • A receiver EDC error. • A parity error on route. • An undefined control character was received. • Unsolicited data was received. • A receiver lost end-of-packet. • A token count miscomparison. • A token sequence error. • A token count overflow. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 227. • See **/var/adm/SPlogs/css[0 | 1]/p0/out.top** for cable information. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** See “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter hardware failure. **Action:** For more information, see “SP Switch2 device and link error information” on page 206.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_RCVLNKSYNC_RE 37D7841C	**Explanation:** A Receiver Port Link Synch Failure occurred. **Cause:** A loose, disconnected, or faulty cable. **Action:** • Check, reconnect, or replace the cable. • See **/var/adm/SPlogs/css[0 | 1]/p0/out.top** for cable information. • See “Cable diagnostics” on page 227. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter hardware failure. **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. **Cause:** Remote SP Switch2 adapter hardware failure. **Action:** • Run adapter diagnostics on remote node. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** to identify remote node.
CS_SW_FIFOOVRFLW_RE 821465C1	**Explanation:** A Receiver FIFO Overflow error was detected. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 227. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter hardware failure. **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “SP Switch2 device and link error information” on page 206.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_EDCTHRSHLD_RE 39FCD5B9	**Explanation:** EDC Error Threshold condition occurred. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 227. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • If the problem persists, call IBM Hardware Service.
CS_SW_RECV_STATE_RE 255F1AA2	**Explanation:** SP Switch2 receiver state machine error. **Cause:** SP Switch2 adapter or switch failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_PE_ON_DATA_RE 1F59782A	**Explanation:** SP Switch2 sender parity error on data was detected. **Cause:** SP Switch2 board failure. **Action:** Call IBM Hardware Service.
CS_SW_INVALD_RTE_RE 02A63E85	**Explanation:** SP Switch2 sender invalid route error occurred. **Cause:** SP Switch2 adapter microcode or a switch daemon software error. **Action:** Call the IBM Support Center.
CS_SW_SNDLOSTEOP_RE F835CDED	**Explanation:** Sender Lost EOP (end-of-packet) condition occurred. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 227. • See **/var/adm/SPlogs/css[0 | 1]/p0/out.top** for cable information. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** See “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter hardware failure. **Action:** • Run adapter diagnostics. • See “SP Switch2 device and link error information” on page 206.
CS_SW_SNDTKNTHRS_RE 80CF3B5A	**Explanation:** A Token Error Threshold error occurred. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See “Cable diagnostics” on page 227. • If the problem persists, call IBM Hardware Service. • For more information, see **/var/adm/SPlogs/css[0 | 1]/p0/flt**.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_SND_STATE_RE 74CEAB0F	**Explanation:** A Sender State Machine Error occurred. **Cause:** SP Switch2 adapter or SP Switch2 failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_PE_ON_NMLL_RE 7F704673	**Explanation:** A Parity Error on the NMLL was detected. **Cause:** SP Switch2 board failure. **Action:** Call IBM Hardware Service.
CS_SW_CRC_SVCPKT_RE 8B091668	**Explanation:** SP Switch2 service logic detected an incorrect CRC on a Service Packet. **Cause:** A transient error in data occurred during transmission over SP Switch2 links. **Action:** See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. **Cause:** A loose, disconnected, or faulty cable. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “Cable diagnostics” on page 227. **Cause:** A node was shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter hardware failure. **Action:** • Run adapter diagnostics. • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See “SP Switch2 device and link error information” on page 206.
CS_SW_PE_RTE_TBL_RE E8F741CD	**Explanation:** SP Switch2 service logic detected an incorrect Parity Error in the Route Table. **Cause:** SP Switch2 board failure. **Action:** Call IBM Hardware Service.
CS_SW_SVC_STATE_RE CF66D3CC	**Explanation:** SP Switch2 service logic state machine error. **Cause:** SP Switch2 board failure. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_OFFLINE_RE 57959ED9	**Explanation:** Node received a fence (Offline) request. **Cause:** The operator ran the **Efence** command. **Action:** Run the **Eunfence** command to bring the node onto the SP Switch2.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_PRI_TAKOVR_RE A8978621	**Explanation:** SP Switch2 primary node takeover. **Cause:** The SP Switch2 primary node became inaccessible. **Action:** See the AIX error log on the previous SP Switch2 primary node.
CS_SW_BCKUP_TOVR_RE FD2D84AD	**Explanation:** SP Switch2 primary-backup node takeover. **Cause:** The SP Switch2 primary-backup node became inaccessible **Action:** See the AIX error log on the previous switch primary- backup node.
CS_SW_LST_BUP_CT_RE 2196D5B4	**Explanation:** SP Switch2 primary-backup node not responding. **Cause:** The SP Switch2 primary-backup node become inaccessible. **Action:** See the AIX error log on the current SP Switch2 primary-backup node.
CS_SW_UNINI_NODE_RE 96DD24B7	**Explanation:** SP Switch2 nodes not initialized during **Estart** command processing. **Cause:** The listed nodes were shutdown, reset, powered off, or disconnected. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • For more information, see “Verify SP Switch2 node operation” on page 229. **Cause:** SP Switch2 adapter problem. **Action:** Run adapter diagnostics on listed nodes.
CS_SW_UNINI_LINK_RE 362E5B7	**Explanation:** SP Switch2 links were not initialized during **Estart** command processing. **Cause:** The node was fenced. **Action:** Run the **Eunfence** command to unfence the node. **Cause:** The switch cable is not wired correctly. **Action:** See **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** to determine if cables were not wired correctly. **Cause:** A loose, disconnected, or faulty cable. **Action:** See “Cable diagnostics” on page 227.
CS_PROCESS_KILLD_RE D250F9DB	**Explanation:** User Process was killed due to link outage. **Cause:** SP Switch2 adapter failure or SP Switch2 failure. **Cause:** The operator fenced this node. **Action:** See neighboring error log entries to determine the cause of the outage.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_MISWIRE_ER 933B622E	**Explanation:** SP Switch2 cable miswired (not connected to the correct switch jack). **Cause:** The SP Switch2 cable was not wired correctly. **Action:** See **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** to determine if cables were not wired correctly.
CS_SW_HARDWARE_ER F96576C4	**Explanation:** Defective SP Switch2 board. **Cause:** SP Switch2 board configuration problem. **Action:** • Issue the command: **hmcmds -G setid** *frame*:*slot* to reconfigure the SP Switch2 board. • If the problem persists, call the IBM Support Center. **Cause:** Faulty SP Switch2 board. **Action:** If the problem persists, call IBM Hardware Service.
CS_SW_LOGFAILURE_RE 5ABE7E20	**Explanation:** Error writing SP Switch2 log files. **Cause:** The **/var** file system is full. **Action:** Obtain free space in the file system or expand the file system. **Cause:** There are too many files open in the system. **Action:** Reduce the number of open files in the system.
CS_SW_INIT_FAIL_ER 957E82AA	**Explanation:** Switch fault-service daemon initialization failed. **Cause:** The operating environment could not be established. **Action:** • See Detail Data of this entry for a specific failure. • Correct the problem and restart the daemon by running the **rc.switch** command. • If the problem persists, call the IBM Support Center.
CS_SW_SIGTERM_ER A98EF5D8	**Explanation:** SP Switch2 fault service daemon received SIGTERM. **Cause:** Another process sent a SIGTERM. **Action:** Run the **rc.switch** command to restart the daemon.
CS_SW_SVC_Q_FULL_RE 172826EF	**Explanation:** SP Switch2 service send queue is full. **Cause:** There is a traffic backlog on the SP Switch2 adapter. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_GET_SVCREQ_ER 4DFEC48	**Explanation:** SP Switch2 daemon could not get a service request. **Cause:** SP Switch2 device driver failure. **Action:** Call the IBM Support Center.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_RSGN_PRIM_RE 585D90B2	**Explanation:** The SP Switch2 Primary node resigned from the job as primary node. **Cause:** Could not communicate over the SP Switch2. **Action:** See neighboring AIX error log entries to determine the cause of the outage. **Cause:** Another node was selected as the primary node. **Action:** None.
CS_SW_RSGN_BKUP_RE C32FD9D3	**Explanation:** Resigning as the SP Switch2 primary-backup node. **Cause:** Could not communicate over the SP Switch2. **Action:** See neighboring error log entries to determine the cause of the outage. **Cause:** Another node was selected as the primary-backup node. **Action:** None.
CS_SW_SDR_FAIL_RE 3E6F3E2E	**Explanation:** Switch fault service daemon failed to communicate with SDR. **Cause:** An Ethernet overload. **Action:** If the problem persists, call the IBM Support Center. **Cause:** Excessive traffic to the SDR. **Action:** If the problem persists, call the IBM Support Center. **Cause:** The SDR daemon on the control workstation is down. **Action:** Check to see if the SDR daemon is up. **Cause:** A software error. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_SCAN_FAIL_ER 63589548	**Explanation:** SP Switch2 scan failed. **Cause:** Could not communicate over the SP Switch2. **Cause:** SP Switch2 adapter failure or a SP Switch2 failure. **Action:** Issue the **Estart** command if primary takeover does not occur.
CS_SW_PLANEMISW_ER 94E99A66	**Explanation:** SP Switch2 plane miswire. **Cause:** SP Switch2 cable is connected on one side to a switch-port or node-port belonging to a different SP Switch2 plane than the one that the other side of the cable is connected to. **Action:** See **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** to determine which cables were not wired correctly.
CS_SW_NODEMISW_RE A19DCA76	**Explanation:** SP Switch2 node miswired. **Cause:** SP Switch2 cable was not plugged into the correct node. **Action:** See **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** to determine if cables were not wired correctly.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_NODECONF_RE CEB4B5AF	**Explanation:** SP Switch2 node configuration error. **Cause:** A switch node was not configured properly to the system. **Cause:** An unknown node was plugged into the system - probable miswire. **Action:** • Use the switch node number and plane given to find the offending node. • See **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** to determine where the offending node is connected. • Reconfigure or disconnect the offending node.
CS_SW_RTE_GEN_RE 44D2A1B5	**Explanation:** SP Switch2 daemon failed to generate routes. **Cause:** A software error. **Action:** Call the IBM Support Center.
CS_SW_FENCE_FAIL_RE A6E635F9	**Explanation:** Fence of node off SP Switch2 failed. **Cause:** Could not communicate over the SP Switch2. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** for more information. • See the error log on the failing node. • Issue the **Estart** command to initialize the switch network.
CS_SW_REOP_WIN_ER 0C17D5C7	**Explanation:** Switch fault service daemon reopen adapter windows failed. **Cause:** SP Switch2 adapter or daemon recovered. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_ESTRT_FAIL_RE 4EE9669F	**Explanation: Estart** command failed - switch network could not be initialized. **Cause:** Could not initialize SP Switch2 chips or nodes. **Action:** • See Detail Data of this entry for the specific failure. • If the problem persists, call the IBM Support Center.
CS_SW_IP_RESET_ER A6BCABA3	**Explanation:** Switch fault service daemon could not reset IP. **Cause:** SP Switch2 device driver error. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_CBCST_FAIL_RE 31C01480	**Explanation:** Switch fault service daemon command broadcast failed. **Cause:** Could not communicate over the SP Switch2. **Cause:** A traffic backlog on the SP Switch2 adapter. **Action:** If the problem persists, call the IBM Support Center.

Table 33. Possible causes of fault service daemon failures - SP Switch2 (continued)

Label and Error ID	Error description and analysis
CS_SW_UBCST_FAIL_RE F7704403	**Explanation:** Switch fault service daemon database updates broadcast failed. **Cause:** SP Switch2 communication failure. **Cause:** A traffic backlog on the SP Switch2 adapter. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_DNODE_FAIL_RE 19337D09	**Explanation:** Switch daemon failed to communicate with dependent nodes. **Cause:** Failed to communicate over the SP Switch2. **Cause:** A traffic backlog on the SP Switch2 adapter. **Action:** If the problem persists, call the IBM Support Center.
CS_SW_PORT_STUCK_RE 889BE7C3	**Explanation:** SP Switch2 port cannot be disabled. **Eunfence** command failed. **Cause:** • SP Switch2 chip or adapter hardware error. • Cable failure. **Action:** • See **/var/adm/SPlogs/css[0 | 1]/p0/flt** on the primary node for more information. • Run adapter diagnostics on the node that failed to unfence.
CS_SW_FSD_TERM_ER 1C27CFCD	**Explanation:** Switch fault service daemon process was terminated. **Action:** See preceding error log entries to determine the cause of the failure. **Cause:** Faulty system planar. **Action:** Run complete diagnostics on the node. **Cause:** Not enough free space left in the node's **/var/adm/SPlogs** file system. **Action:** Obtain more space.

3. Adapter Diagnostic errors, which are recognized by labels with a prefix of: **SWT_DIAG_**

Table 34. Possible causes of adapter diagnostic failures - SP Switch2

Label and Error ID	Error description and analysis
SWT_DIAG_ERROR1_ER 8998B96D	**Explanation:** SP Switch2 adapter failed post-diagnostics, see the man page for the **diag** command. **Cause:** Faulty switch adapter. **Action:** See "SP Switch2 adapter diagnostics" on page 226.
SWT_DIAG_ERROR2_ER 2FFF253A	**Explanation:** SP Switch2 adapter failed diagnostics. **Cause:** Faulty switch adapter. **Action:** See "SP Switch2 adapter diagnostics" on page 226.

4. **emasterd** errors, which are recognized by labels with a prefix of: **CS_EMSTR_**

Table 35. Possible causes of SP Switch2 TOD management (emasterd) failures

Label and Error ID	**Error description and analysis**
CS_EMSTR_EXIT_ER FFDEEAA3	**Explanation:** SP Switch2 TOD management daemon (**emasterd**) exited on the control workstation. **Action:** Look for more information in the AIX error log on the control workstation. **Cause:** Insufficient free space in **/var/adm/SPlogs** file system. **Action:** Free space in the file system. **Cause:** The daemon cannot communicate with the SDR. **Action:** See that the SDR daemon is running, and restart the **emasterd**.
CS_EMSTR_RESIGN_ER 13D2008B	**Explanation:** SP Switch2 TOD management daemon (**emasterd**) failed to resign the current MSS (Master Switch Sequencer) node. **Cause:** No communication with the MSS node via the SP Switch2 or Ethernet. No new MSS will be assigned. **Action:** Resume communication with the node.
CS_EMSTR_MS_SRCH_ER EB6CBD01	**Explanation:** SP Switch2 TOD management daemon (**emasterd**) failed to find the current MSS (Master Switch Sequencer) node. The SDR information is incorrect. **Cause:** No communication with the MSS node via the SP Switch2 or Ethernet. No new MSS will be assigned. **Action:** Resume communication with the node.

5. SP Switch2 PCI Adapter errors, which are recognized by these error log entries.

Table 36. Possible causes of SP Switch2 PCI Adapter failures

Label and Error ID	**Error description and analysis**
CS_RCVY_START_RE 6317E423	**Explanation:** Critical error recovery is starting. **Action:** This is likely to be a switch adapter hardware or microcode failure. See neighboring error log entries for the cause of the outage.
CRS_FENCE_ER 1485995F	**Explanation:** CSS Adapter Port-Permanent Error(Fence) **Cause:** Switch cable or switch failure **Action:** Run adapter diagnostic. See **/var/adm/SPlogs/css/css[0 | 1]/la_error.log** and **/var/adm/SPlogs/css[0 | 1]/la_event_d.trace** for details. **Cause:** Switch cable loose or disconnected **Action:** Check, reconnect, or replace the cable. Unfence the node.
CRS_OFFLINE_ER 4E14B222	**Explanation:** CSS Adapter - Permanent Hardware or Software Error(Offline) **Cause:** Switch adapter hardware or software error. **Action:** Run adapter diagnostic. See **/var/adm/SPlogs/css/css[0 | 1]/la_error.log** and **/var/adm/SPlogs/css[0 | 1]/la_event_d.trace** for details. Record the above information anc contact the IBM Support Center.

Table 36. Possible causes of SP Switch2 PCI Adapter failures (continued)

Label and Error ID	Error description and analysis
CRS_RESTART_HWSW_ER 4007FD6A	**Explanation:** CSS Adapter - Critical Hardware or Software Error(Restart) **Cause:** Switch adapter hardware or software error. **Action:** See **/var/adm/SPlogs/css/css[0 | 1]/la_error.log** and **/var/adm/SPlogs/css[0 | 1]/la_event_d.trace** for details. If the problem persists, run adapter diagnostics and call hardware support.
CRS_LAMSG_SEND_FAIL 5144238C	**Explanation:** Event Handler failed to send La_event_d error action **Cause:** Switch adapter hardware or software error. **Action:** See **/var/adm/SPlogs/css/css[0 | 1]/la_error.log** and **/var/adm/SPlogs/css[0 | 1]/la_event_d.trace** for details.
CORSAIR_CONFIG1_ER FFCF4911	**Explanation:** CSS config failed. **Cause:** Software Error, ODM error. **Action:** Run config method with verbose option for more information.

Adapter configuration error information

The following table is based on the possible values of the **adapter_config_status** attribute of the **Adapter** object of the SDR. Use the following command to determine its value:

```
SDRGetObjects Adapter adapter_type==[css0 | css1] node_number adapter_config_status
```

Use the value of the **adapter_config_status** attribute for the node in question, to index into Table 37. The value of a correctly configured CSS adapter is **css_ready**.

Note: The **adapter_config_status** table that follows uses the phrase ″adapter configuration command″. This refers to the SP Switch2 adapter configuration method. Use this command to invoke it:

```
/usr/lpp/ssp/css/cfgcol -v -l [css0 | css1] > output_file_name
```

Table 37. adapter_config_status values - SP Switch2

adapter_config_status	Explanation and recovery
css_ready	Correctly configured CSS adapter.
odm_fail genmajor_fail genminor_fail getslot_fail build_dds_fail	**Explanation:** An ODM failure has occurred while configuring the CSS adapter. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply the command output.
getslot_fail	Verify that the CSS adapter is properly seated, then rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply them with the command output.

Table 37. adapter_config_status values - SP Switch2 (continued)

adapter_config_status	Explanation and recovery
busresolve_fail	**Explanation:** There are insufficient bus resources to configure the CSS adapter. **Action:** Contact the IBM Support Center.
dd_load_fail	See “Verify software installation” on page 228. If software installation verification is successful and the problem persists, contact the IBM Support Center.
make_special_fail	**Explanation:** The CSS device special file could not be created during adapter configuration. **Action:** Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply them with the command output.
dd_config_fail	**Explanation:** An internal device driver error occurred during CSS adapter configuration. **Action:** See “Information to collect before contacting the IBM Support Center” on page 218.
diag_fail	**Explanation:** SP Switch2 diagnostics failed. **Action:** See “SP Switch2 adapter diagnostics” on page 226.
not_configured	**Explanation:** The CSS adapter is missing or not configured.
pdd_init_fail load_khal_fail	See “Verify software installation” on page 228. If software installation verification is successful and the problem persists, contact the IBM Support Center.

SP Switch2 device and link error information

The device and link current status is gathered in the annotated switch topology file, **out.top**, that is created on each plane of each node that has its corresponding **switch_responds** set to 1. For plane 0, **switch_responds0** must be 1. For plane1, **switch_responds1** must be 1. The file looks like the switch topology file except that for each device or link that differs from the operational default status, an additional comment is made. For the directory that contains the **out.top** file, see “SP Switch2 log and temporary file hierarchy” on page 187.

These additional comments are appended to the file by the fault service daemon and reflect the current connectivity status of the link or device. No comment on a link or device line means that the link or device exists and is operational. The comment format is:

```
ideal-topology-line device-status-no which-device:device-status-string (link-status-string)
```

where:

- *ideal- topology-line* is the line copied from the topology file. No comment means that the device or link is up and running.
- *device-status-no* is the device error number. See Table 38 on page 207.
- *which-device* **L** for the left device or **R** for the right device.
- *device-status-string* is the device error description for this device-status number. See Table 38 on page 207.
- *link-status-string* is the link error description. See Table 39 on page 208.

Not all the comments reflect an error. Some may be a result of the system configuration or current system administration.

An example of a failing entry and description is in "out.top" on page 215. If the listed recovery actions fail to resolve your problem, contact the IBM Support Center.

The possible device status values for SP Switch systems, with their recovery actions, are listed in Table 38. The possible link status values for SP Switch systems, with their recovery actions, are listed in Table 39 on page 208. Additional miswire information can be found in "cable_miswire" on page 216.

Table 38. SP Switch2 device status and recovery actions

Device status number	Device status text	Explanation and recovery actions
2	Initialized	**Explanation:** Both devices are initialized. The port's link status is not operational. **Cause:** The link is faulty. **Action:** See Table 39 on page 208 for link status.
0	Uninitialized	**Explanation:** No device is connected to this port. **Cause:** There is no cable connected to this port. **Action:** If this is intentional, no action is needed. If not, connect a cable to the port.
-3	The device has been removed from network because of a bad signature	**Explanation:** The device was removed from the switch network - device configuration failure. **Cause:** A fault on the device. **Action:** Contact IBM Hardware Service.
-4	Device has been removed from network - faulty.	**Explanation:** The device has been removed from the switch network. **Cause:** A fault on the device. **Action:** If the device in question is a node, see "Verify SP Switch2 node operation" on page 229. Otherwise, contact IBM Hardware Service.
-5	Device has been removed from the network by the system administrator.	**Explanation:** The device was placed offline by the systems administrator (**Efence**). **Cause:** The switch administrator ran the **Efence** command. **Action: Eunfence** the device.
-6	Device has been removed from network - no AUTOJOIN.	**Explanation:** The device was removed and isolated from the switch network. **Cause:** The node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. **Action:** First attempt to **Eunfence** the device. If the node fails to rejoin the switch network, see "AIX Error Log information" on page 188. If the problem persists, contact the IBM Support Center.

Table 38. SP Switch2 device status and recovery actions (continued)

Device status number	Device status text	Explanation and recovery actions
-7	Device has been removed from the network for not responding.	**Explanation:** The device was removed from the switch network. **Cause:** An attempt was made to contact the device, but the device did not respond. **Action:** If the device in question is a node, see "Verify SP Switch2 node operation" on page 229. Otherwise, contact the IBM Support Center.
-8	Device has been removed from the network because of a miswire.	**Explanation:** The device is not cabled properly. **Cause:** Either the switch network is miswired, or the frame supervisor tty is not cabled properly. **Action:** First view the **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** file. Verify and correct all links listed in the file. Then issue the **Estart** command. If the problem persists, contact IBM Hardware Service.
-9	Destination not reachable.	**Explanation:** The device was not reachable through the switch network. **Cause:** This is generally due to other errors in the switch network fabric. **Action:** Investigate and correct the other problems, then run the **Estart** command.

Table 39. SP Switch2 link status and recovery actions

Link Status Number	Link Status Text	Explanation and Recovery Actions
0	Uninitialized	**Explanation:** The link is uninitialized. **Cause:** Switch Initialization was not complete. **Action:** Try to **Estart** the switch network again. If the problem persists, contact the IBM Support Center.
-1	The link is not operational - link re-timing	**Explanation:** The link is in the initialization stage. **Cause:** If the problem persists, the link may be faulty - Cable or interposer card faulty. **Action:** First attempt to **Estart** the switch again, If the link does not come up, try switching the cable or connecting a wrap plug to test the interposer card.
-2	Wrap plug is installed.	**Explanation:** This link is connected to a wrap plug. **Cause:** The wrap plug is connected to the port in order to test the port. This is not normally a problem. **Action:** None.

Table 39. SP Switch2 link status and recovery actions (continued)

Link Status Number	Link Status Text	Explanation and Recovery Actions
-3	The link is not operational - link failed to time.	**Explanation:** The link failed to initialize. **Cause:** If problem persists, the link maybe faulty - Cable or interposer card faulty. **Action:** First attempt to **Estart** again. If the link does not come up, try switching the cable or connecting a wrap plug to test the interposer card.
-4	Link has been removed from network or miswired - faulty.	**Explanation:** The link is not operational and was removed from the switch network. **Cause:** Either the link is miswired or the link has failed. **Action:** First check the **/var/adm/SPlogs/css[0 | 1]/p0** directory for the existence of a **cable_miswire** file. If the file exists, verify and correct all links listed in the file. Then issue the **Estart** command. If the **cable_miswire** file does not exist, examine the **/var/adm/SPlogs/css[0 | 1]/p0/flt** file for entries relating to this link. If entries are found, verify that the cable is seated at both ends, then run the **Estart** command. If the problem persists, contact the IBM Support Center.
-5	The link has been removed from network by the system administrator	**Explanation:** The link was removed (commented out) from the switch network by the switch administrator. This is not a problem.
-6	The link has been removed from network - no AUTOJOIN	**Explanation:** The device was removed and isolated from the switch network. **Cause:** The node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. **Action:** • First attempt to **Eunfence** the device. • If the node fails to rejoin the switch network, see “AIX Error Log information” on page 188. • If the problem persists, contact the IBM Support Center.
-7	Link has been removed from network - fenced.	**Explanation:** The device was placed offline by the Systems Administrator (**Efence**). **Action: Eunfence** the associated node.
-8	Link has been removed from network - probable miswire.	**Explanation:** The link is not cabled properly. **Action:** View the **/var/adm/SPlogs/css[0 | 1]/p0/cable_miswire** file. Verify and correct all links listed in the file, then run the **Estart** command.
-9	Link has been removed from network - not connected.	**Explanation:** The link cannot be reached by the primary node, therefore initialization of the link is not possible. **Cause:** This is generally caused by other problems in the switch network, such as a switch chip being disabled. **Action:** Investigate and correct the underlying problem, then run the **Estart** command.

Dump information

The dump files are created either along with the **css.snap** script, which automatically invokes the associated utility commands, or manually at the user's request. The **css.snap** script is issued by the user or by the switch support code (device driver, fault service daemon, HAL) whenever a serious error occurs. **css.snap** gathers dump and trace files into a snapshot compressed package. For more information, see "Information to collect before contacting the IBM Support Center" on page 218.

errpt.out

The AIX commands **errpt** and **errpt -a** are redirected into the file **errpt.out** on the port level, by the **css.snap** script. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is gathered with other dump and trace files into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 218.

cadd_dump.out

The command:

```
/usr/lpp/ssp/css/cadd_dump [-0 | -1] -r
```

dumps the adapter device driver css0 or css1 message buffer to stdout.

When this command is run automatically from the **css.snap** script, the information is redirected into the file **cadd_dump.out** on the port level. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 218.

The command dumps status and error information for events such as **ioctl** calls to a specific port on a specific adapter.

ifcl_dump.out

The command:

```
/usr/lpp/ssp/css/ifcl_dump -a -l -f
```

dumps the adapter device driver IP message buffer of css0 to stdout.

When this command is run automatically from the **css.snap** script, the information is redirected into the file **ifcl_dump.out** on the adapter level. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 218.

logevnt.out

The file contains records of errors which occurred in the components running on the node which experienced the error. These components are notified when switch-related error log entries are made, and report the summary data to Event Management for transmission to the control workstation. This file is on the adapter level of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

summlog.out

This file contains error information for the **css.summlog** daemon, which gathers summary log information and writes it to the summary log file. For information on the summary log file, see “Summary log for SP Switch, SP Switch2, and switch adapter errors” on page 69. **summlog.out** is a text file on the node level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

col_dump.out

The command:

```
/usr/lpp/ssp/css/col_dump [-0 | -1] -r
```

dumps the microcode trace buffer of css0 or css1 to stdout.

When the command is run from the **css.snap** script, the output is redirected to the **col_dump.out** file on the adapter level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

The command dumps status and error information for events such as user-space system calls to the Hardware Abstraction Layer (HAL).

The command dumps adapter memory and state, and may be destructive, meaning that the adapter cannot be in use while issuing this command. Furthermore, after running the command the adapter should be restarted. Restarting the adapter is done by running the **rc.switch** script on the affected node:

```
/usr/lpp/ssp/css/rc.switch -l [css0 | css1]
```

cssadm2.debug

The file contains trace information of the actions of the **cssadm2** daemon. This file contains entries for each event received and handled, as well as how the events were handled and the results. This file is on the node level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

cssadm2.stderr

The file contains any unexpected eror messages received by the **cssadm2** daemon, while performing commands external to the daemon. This file is on the node level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

cssadm2.stdout

The file contains any unexpected informational messages received by the **cssadm2** daemon while performing commands external to the daemon. In general, this file should remain empty. This file is on the node level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

regs.out

The command:

```
/usr/lpp/ssp/css/read_regs
```

dumps the SP Switch2 adapter’s Trail Blazer Interface Chip (TBIC3), MIC, and NBA registers.

All information is dumped to stdout. When this command is run automatically from the **css.snap** script, the information is redirected into the file **regs.out** on the adapter level. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 218.

You can check the TBIC_STATUS register from this information.

router.log

The file **router.log** is in the adapter level directory on each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is used to record error and warning messages issued by the route table generator (RTG). It contains the RTG version and date, the date and time that the file was created, the node on which the file was created, information about the system topology and the routing algorithm type, and any messages generated during route generation. This file is created during service route generation on the primary node, and during processor route generation on all nodes.

If the file is found at the start of route generation, it is copied to **router.log.old**. If the file contains any messages at the end of route generation, the file is copied to **router_failed.log1**. Up to two failed router logs, **router_failed.log1** and **router_failed.log2** are maintained at any time.

This file is then gathered into the snapshot compressed package. For more information on the snapshot, see “Information to collect before contacting the IBM Support Center” on page 164.

CSS_test.log

The command:

```
/usr/lpp/ssp/bin/CSS_test
```

produces a log file **CSS_test.log** at the node level directory, which is then gathered into the snapshot compressed package. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is present only if the **CSS_test** command was run on the node.

If your system is running in restricted root access mode, **CSS_test** must be run from the control workstation. For more information, see "Verify software installation" on page 228, and the **CSS_test** entry of *PSSP: Command and Technical Reference*.

odm.out

The commands:

```
/bin/odmget -q "attribute=adapter_status" PdAt
/bin/odmget -q "attribute=adapter_status" CuAt
```

dump the SP Switch2 adapter's configuration status as it was saved in the ODM, to stdout.

When this command is run automatically from the **css.snap** script, the information is redirected into the file **odm.out** on the adapter level. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is then gathered into the snapshot compressed package. For more information on the snapshot, see "Information to collect before contacting the IBM Support Center" on page 218.

spdata.out

The commands:

- **/usr/lpp/ssp/css/splstdata -f -G**
- **/usr/lpp/ssp/css/splstdata -s -G**
- **/usr/lpp/ssp/css/splstdata -n -G**
- **/usr/lpp/ssp/css/splstdata -b -G**

prints network and switch current status information to stdout.

When these commands are run automatically from the **css.snap** script, the information is redirected into the file **spdata.out** on the port level. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is then gathered into the snapshot compressed package. See "Information to collect before contacting the IBM Support Center" on page 218. The **splstdata** command is described in *PSSP: Command and Technical Reference*.

netstat.out

The AIX commands:

- **netstat -I css[0 | 1]**
- **netstat -m**

print network related data to stdout.

When these commands are run automatically from the **css.snap** script, the information is redirected into the file **netstat.out** on the adapter level. For the full path name, see "SP Switch2 log and temporary file hierarchy" on page 187. This file is then gathered into the snapshot compressed package. See "Information to collect before contacting the IBM Support Center" on page 218. The **netstat** command is described in *AIX 5L Version 5.1 Command and Technical Reference*.

scan_out.log and scan_save.log

These files are in the adapter level of each node that runs the adapter diagnostics and the TBIC test within this test (TBIC test within adapter diagnostics). For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. The TBIC, which is a chip on the SP Switch2 adapter, is scanned into this file and saved as part of the snapshot by the **css.snap** script. The **scan_save.log** is a previous TBIC scan. This TBIC scan procedure is destructive and therefore taken only when the **css.snap** runs after a permanent failure. No other switch communication is available, and the switch adapter is reset after the scan is taken. See “SP Switch2 adapter diagnostics” on page 226

DeviceDB.dump

The **DeviceDB.dump** file is in the port level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains the latest dump of the device data base from the fault service daemon. This dump is taken as part of the snapshot by the **css.snap** script.

Trace information

All the trace files use message catalogs to display the messages, as described in *PSSP: Messages Reference*.

rc.switch.log

The **rc.switch.log** file is in the node level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains log information created by the **rc.switch** script. It is responsible for starting the fault service daemon. If a node fails or is unable to start the fault service daemon, this file may contain an explanation of the failure.

daemon.stdout

The **daemon.stdout** file is in the node level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is a redirection of the stdout and contains the fault service daemon initial information, taken prior to the establishment of any other log mechanism.

daemon.log

The **daemon.log** file is in the node level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains the tracing of the fault service daemon’s node events.

adapter.log

The **adapter.log** file is in the port level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains the tracing of the fault service daemon adapter events. Trace information from adapter initialization and downloading the route table into the adapter is placed in this file.

flt

The **flt** file is in the port level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is used to log hardware error conditions on the switch, recovery actions taken by the fault service daemon, and other general operations that alter the switch configuration. The **flt** file is most important on the primary node, where the switch initialization and servicing take place.

Each entry in the **flt** file contains these items:

- severity:
  - **i** - informational
  - **n** - notification
  - **e** - error message
- time stamp: date and time that the message occurred
- catalog message number
- message text

Each entry may contain this global field:

- **device_id** is either a switch node number, or a switch chip id. Switch chip ids are numbers greater than 100000. The last digit represents the chip id within the switch board, and the digits before represent the switch board sequence number. For example, a **device_id** of 100025 represents chip number 5 in switch board sequence 2.

fs_daemon_print.file

The **fs_daemon_print.file** is in the port level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains the tracing of the fault service daemon port events. A time stamp along with a message is logged into the file whenever an event occurs. Important entries are the ones about the adapter initialization, route table calculation, and downloading the table into the adapter.

out.top

The **out.top** file is in the port level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. The fault service daemon creates a concurrent topology file. The daemon uses the **/etc/SP/expected.top** file, if found, or the topology file from the SDR, as the ideal switch network connectivity. The fault service daemon adds comments to any change from the ideal, and saves the result in the **out.top** file. The **out.top** file is updated whenever a change to the current status of links and nodes (devices) occurs.

The **out.top** file has comments added only on non-operational links or devices. An example of such a comment is:

```
S  15 2 col 1 0          E01-S17-BH-J8 to E01-N10           -4 R:
           device has been removed from network - faulty
           (link has been removed from network - fenced)
```

This means that:

- Switch chip 15 (board 1 chip 5), port 2 is connected to switch node number 1.
- The switch is located in frame 1 slot 17.
- Its bulkhead connection to the node is jack 8.
- The switch node is in frame 1 (the same frame as the switch board) slot 10.
- -4 is the device status of the right side device (col switch node number1), which has the more severe device status of the two devices listed.
- The device status text of this node is ″device has been removed from network - faulty″.
- The link status text ″link has been removed from network - fenced″.

SP Switch2 ports that have no connection are usually wrapped:

```
S 13 3   s 13 3         E01-S17-BH-J3 to E01-S17-BH-J3     2 L:
 initialized (wrap plug is installed)
```

When the fault service daemon on the primary node finds an unexpected device during it’s switch initialization (worm), the link will be marked as miswired. Miswire information will be put in the primary node’s **cable_miswire** file.

```
S 13 3   s 13 3        E01-S17-BH-J3 to E01-S17-BH-J3      2 L: initialized
      (link has been removed from network-probable miswire)
```

Device status codes and recovery actions are listed in Table 38 on page 207. Link status codes and recovery actions are listed in Table 39 on page 208.

topology.data

The **topology.data** file is in the port level directory of the primary node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains error messages that occur during the distribution of the topology file to the other nodes. The topology file distribution is done via the SP Switch2.

cable_miswire

The **cable_miswire** file is in the port level directory of the primary node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file reports any miswire detected by the fault service daemon during SP Switch2 initialization (running the **worm** subsystem).

css.snap.log

The **css.snap.log** file is in the node level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is created every time that the **css.snap** command is run, either automatically due to a failure, or by direct invocation when the user issues the **css.snap** command. This file contains information on what happened during the command invocation, and in particular:

- The time stamp of the snapshot.
- The node where the snapshot was taken
- The list of files in the **/var/adm/SPlogs/css** directory before the command began running.
- The list of files that are assembled into the snapshot package.
- Information on all the running processes on the system at the time of the snapshot.
- Information on the **ssp.css** software product.
- Information on the adapter and the microcode on the switch adapter.

colad.trace

The **colad.trace** file is in the adapter level directory of each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. The file contains trace information from css0 or css1 adapter diagnostics. The adapter diagnostics can run in two modes: Power-On-Self-Test (POST), and direct invocation, by issuing the command **diag**. A file creation time of

```
Midnight Dec 31 1969
```

indicates that the file was created during the POST, when the time had not yet been set. For more details, see “SP Switch2 adapter diagnostics” on page 226.

spd.trace

The **spd.trace** file is on port level directory of each adapter, on the control workstation. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. See “SP Switch2 log and temporary file hierarchy” on page 187. This file contains tracing of advanced switch diagnostics. See “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.

Ecommands.log

The **Ecommands.log** file is in the node level directory on the control workstation. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains trace information for the **Ecommands**.

emasterd.log

The **emasterd.log** file is in the node level directory on the control workstation. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains error and status information about the SP Switch2 TOD management.

emasterd.stdout

The **emasterd.stdout** file is in the node level directory on the control workstation. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains detailed trace information about the SP Switch2 TOD management.

chgcss.log

The **chgcss.log** file is in the node level directory on each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file is used to record all the RESERVE/RELEASE/QUERY events associated to the SP Switch2 adapter window attribute on the current SP node. This log file contains the event type, the value of the window attribute, and a timestamp that indicates when the **chgcss** command was invoked.

la_error.log

The **la_error.log** file is in the adapter level directory on each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. This file contains errors logged by the **la_event_d** daemon for the SP Switch2 PCI Adapter. When 1M of data is recorded, this log is copied with a suffix of **.old** and the existing **.old** file is overwritten.

la_event_d.trace

The **la_event_d.trace** file is in the adapter level directory on each node. For the full path name, see “SP Switch2 log and temporary file hierarchy” on page 187. ′This log contains **la_event_d** verbose messages and raw dump data, always in English, for the SP Switch2 PCI Adapter. When 1M of data is recorded, it is copied with a suffix of **.old**, and the existing **.old** file is overwritten. The level of information recorded can be increased with the **traceson -s la_event_d** command, and decreased with the **tracesoff -s la_event_d** command.

Missing error data warning

When the system is miswired, the fault service daemon cannot always detect the problem. The daemon updates the file **cable_miswire**, where it reports all the miswires it finds. When reporting a miswire, an entry may be missing in this file or an entry may be incorrect.

If a node is powered on before the switch board receives power, a **diag_fail** condition may occur. See “SP Switch2 adapter diagnostics” on page 226.

Information to collect before contacting the IBM Support Center

The following items are used to isolate problems in the SP Switch2 component of PSSP. More detailed information about each item appears in “Error information” on page 186. Before collecting any other information, check for the existance of the **/etc/plane.info** file, and make a copy of the file if it exists. See “plane.info file” on page 187.

css.snap package

The **/usr/lpp/ssp/css/css.snap** script collects log, trace, and dump information created by SP Switch2 support code (device driver, worm, fault-service daemon, diagnostics) into a single compressed package.

The complete package output file is in the directory: **/var/adm/SPlogs/css**. The file name varies according to the options:

- When the **-a** and **-p** options are not specified, the name is:

  hostname.yymmddhhmmss.**css.snap.tar.Z**
- When option **-a [css0 | css1]** is specified, but the **-p** option is not specified, the name is

  hostname.yymmddhhmmss.**adapter[0 | 1].css.snap.tar.Z**
- When option **-a [css0 | css1]** and **-p p0** are specified, the name is:

  hostname.yymmddhhmmss.**adapter[0 | 1].port0.css.snap.tar.Z**

where *hostname* is the hostname of the node where the **css.snap** command was issued, and *yymmddhhmmss* is the date and time that the **css.snap** information was collected.

The **css.snap** script creates a log file, **/var/adm/SPlogs/css/css.snap.log** where all the files gathered in the package are listed.

This script is called whenever a serious error is detected by the switch support code. To directly cause the system to create such a snapshot, login the desired node and manually issue the command:

```
/usr/lpp/ssp/css/css.snap [-c | -n] [-s] -a [css0 | css1] [-p p0]
```

where

- **-c** Flushes the adapter cache and prints the result. This is the default.
- **-n** Assumes that the device driver or daemon has flushed the cache.
- **-s** Takes a soft snap, which does not dump the adapter state. This excludes the **col_dump.out**. This flag is used for temporary errors (TYPE=TEMP) where the

integrity of the adapter is in doubt, or when it is not desirable to corrupt the adapter state by the use of diagnostic routines.

- **-a** Specifies a single adapter. When this flag is not used, all the node's configured adapters will be selected. This is the default.
- **-p** Specifies a single port on a specified adapter, hence a specific plane. This option must be used together with the **-a** option. When this flag is not specified, all the adapter's ports are selected. This is the default.

Note: For the SP Switch2 PCI Adapter, **css.snap** invokes the **css.snap.corsair** script. In this case, the output file has the name **css.snap.corsair** substituted where **css.snap** appears in the file naming conventions above.

Table 40 shows the error log entries that automatically take a snapshot, as well as the type of snap performed. The **soft** type enables a continuation of work with the switch. The **full** snap **might** corrupt the adapter, forcing an adapter reset and the node to be fenced off of the switch.

Table 40. AIX Error Log entries that invoke css.snap - SP Switch2

Error Log entry	Snap type (full/soft)
CRS_OFFLINE_ER	full
CRS_RESTART_HWSW_ER	full
CS_SW_FIFOOVRFLW_RE	soft
CS_SW_RECV_STATE_RE	soft
CS_RCVY_START_RE	full
CS_SW_INVALID_RTE_RE	soft
CS_SW_PE_ON_DATA_RE	soft
CS_SW_CQ_PE_NCL_RE	soft
CS_SW_PE_RTE_TBL_RE	soft
CSS_DD_CFG_ER	full
CS_SW_SVC_Q_FULL_RE	full

Collect the **css.snap** information from both the primary node and all nodes that are experiencing SP Switch2 problems. Do not reboot the nodes before running **css.snap**, because rebooting causes the loss of valuable diagnostic information.

The **css.snap** script collects all the files which reside in the **/var/adm/SPlogs/css**, **/var/adm/SPlogs/css0**, **/var/adm/SPlogs/css1**, **/var/adm/SPlogs/css1/p0**, and **/var/adm/SPlogs/css0/p0** directories, and some additional files from the **/tmp** directory. Some of the files reside on each node, while others reside only on the primary node or on the control workstation.

Table 41 on page 220 lists important files gathered by **css.snap** and their location at **css.snap** time. Some of the files are created by **css.snap** in order to gather concurrent information on the switch status. For the SP Switch2 PCI Adapter, additional files are collected by **css.snap.corsair** and they are noted after the table.

Log files within css.snap package

Table 41 on page 220 contains the list of files that are collected by **css.snap** and their location in the log directory hierarchy.

Table 41. SP Switch2 log files

Number	Log File name	Hierarchy	Contents	Location
1	adapter.log	adapter	Fault service daemon adapter status information. For more information, see “adapter.log” on page 214.	nodes
2	cable_miswire	port	Node-to-switch or switch-to-switch plane miswired connection information. For more information, see “cable_miswire” on page 216.	primary node
3	cadd_dump.out	adapter	Most recent **css.snap**’s **cadd_dump** command dump file. SP Switch2 adapter device driver trace buffer dump file. For more information, see “cadd_dump.out” on page 210.	nodes
4	chgcss.log	node	Log file of **chgcss**, which changes the adapter device driver’s attributes. For more information, see “chgcss.log” on page 217.	nodes
5	col_dump.out	adapter	The most recent **css.snap**’s **col_dump** command dump file. Microcode dump information. For more information, see “col_dump.out” on page 211.	nodes
6	colad.trace	adapter	SP Switch2 adapter diagnostics messages. For more information, see “colad.trace” on page 216.	nodes
7	core	node	Fault service daemon core dump file.	nodes
8	cssadm2.debug	node	Trace of **cssadm2** daemon. For more information, see “cssadm2.debug” on page 211.	control workstation
9	cssadm2.stderr	node	Unexpected error messages received by the **cssadm2** daemon. For more information, see “cssadm2.stderr” on page 211.	control workstation
10	cssadm2.stdout	node	Unexpected informational messages received by the **cssadm2** daemon. For more information, see “cssadm2.stdout” on page 212.	control workstation
11	css.snap.log	node	**css.snap** snapshot command log information - list of all files gathered in the last snapshot. For more information, see “css.snap.log” on page 216.	nodes
12	CSS_test.log	node	Present if the **CSS_test** command was run on the node. For more information, see “CSS_test.log” on page 212, “Verify software installation” on page 228 and the **CSS_test** command entry in *PSSP: Command and Technical Reference.*	nodes
13	daemon.log	node	Fault service daemon output file. For more information, see “daemon.log” on page 214.	nodes

Table 41. SP Switch2 log files (continued)

Number	Log File name	Hierarchy	Contents	Location
14	DeviceDB.dump	port	Latest dump of the device data base from the fault service daemon. See “DeviceDB.dump” on page 214.	nodes
15	Ecommands.log	node	Log entries of all **Ecommands**. For more information, see “Ecommands.log” on page 217.	control workstation
16	emasterd.log	node	TOD Management **emasterd** daemon - errors and notifications. For more information, see “emasterd.log” on page 217.	control workstation
17	emasterd.stdout	node	TOD Management **emasterd** daemon - more detailed trace file. For more information, see “emasterd.stdout” on page 217.	control workstation
18	errpt.out	node	Most recent **errpt -a** and **errpt** results. For more information, see “errpt.out” on page 210 and the **errpt** command entry in *AIX Command and Technical Reference*.	nodes
19	flt	port	Hardware error conditions found on the SP Switch2, recovery action taken by the fault-service daemon, and general operations that alter the SP Switch2 configuration. For more information, see “flt” on page 214.	nodes
20	fs_daemon_print.file	port	Fault service daemon port status information. For more information, see “fs_daemon_print.file” on page 215.	nodes
21	ifcl_dump.out	adapter	Most recent **css.snap**’s **ifcl_dump** command dump file. IP dump information. For more information, see “ifcl_dump.out” on page 210.	nodes
22	logevnt.out	node	Log error log events monitored by **ha**. For more information, see “logevnt.out” on page 210.	nodes
23	netstat.out	adapter	Most recent **css.snap**’s **netstat** command dump file. Network status information. For more information, see “netstat.out” on page 213, and the entry for the **netstat** command in *AIX Command and Technical Reference*.	nodes
24	odm.out	adapter	The node’s **adapter_status** configuration as saved in the ODM. For more information, see “odm.out” on page 213.	nodes
25	out.top	port	SP Switch2 plane link information. For more information, see “out.top” on page 215.	nodes
26	rc.switch.log	node	Fault service daemon initialization information. For more information, see “rc.switch.log” on page 214.	nodes

Table 41. SP Switch2 log files (continued)

Number	Log File name	Hierarchy	Contents	Location
27	rc.switch.log.previous	node	Node's previous fault service daemon initialization information. For more information, see "rc.switch.log" on page 214.	nodes
28	regs.out	adapter	Most recent **css.snap**'s **read_regs** command dump file. SP Switch2 adapter's registers dump file. For more information, see "regs.out" on page 212.	nodes
29	router.log	port	SP Switch2 routing information. For more information, see "router.log" on page 212.	nodes
30	scan_out.log	adapter	TBIC scan ring binary information. For more information, see "scan_out.log and scan_save.log" on page 213.	nodes
31	scan_save.log	adapter	Previous TBIC scan ring binary information. For more information, see "scan_out.log and scan_save.log" on page 213.	nodes
32	spd.trace	port	Tracing of advanced switch diagnostics. See "spd.trace" on page 217.	control workstation
33	spdata.out	port	Most recent **css.snap**'s **splstdata** command dump file. SP Switch2 data requests. For more information, see "spdata.out" on page 213.	primary node
34	summlog.out	node	Error information from the **css.summlog** daemon. For more information, see "summlog.out" on page 211.	control workstation
35	topology.data	port	System error messages from the distribution of the topology file to the secondary nodes. For more information, see "topology.data" on page 216.	primary node

If **css.snap** is invoked for an SP Switch2 PCI Adapter error, the **css.snap.corsair** script is invoked. The following log files are also included:

1. **sdr.out**, which is the result of issuing these commands:
   - **splstdata -f -G**
   - **splstdata -s -G**
   - **splstdata -a -G**
   - **splstdata -n -G**
   - **splstdata -b -G**
2. **khal_dump.out** which is the result of issuing the **khal_dump -Z** command.
3. **ifcr_dump.out** which is the result of issuing the **ifcr_dump -Z** command.
4. **var/adm/ffdc/stacks** which contains FFDC information.
5. **cssN/lsattr.out** which is the result of issuing the **lsattr -EL css[0 | 1]** command.

6. **css[0 | 1]/cordd_dump.ini** and **cordd_dump.fin** which are the results of issuing the **cordd_dump -Z -l css[0 | 1]** command. There is one file for the start and one for the end of the command.
7. **css[0 | 1]/ucode.out** which is the result of issuing the **cordd_dump -Z -a N** command.
8. **css[0 | 1]/read_regs.out** which is the result of issuing the **read_regs -l css[0 | 1]** command.

Note: The files ending in **.out** are produced by running the appropriate command to dump internal (in memory) trace information or dump data to a file.

Disk space handling policy

The **css.snap** command avoids filling up the **/var** directory by following these rules:

1. If less than 10% of **/var** is free, **css.snap** exits.
2. If the css portion of **/var** is more than 30% of the total space in **/var**, **css.snap** erases old snap files until the css portion becomes less than 30%. If it is successful, the snap proceeds. If not, it exits.

The **css.snap** command is called automatically from the fault service daemon when certain serious errors are detected. The **css.snap** command can also be issued from the command line when a switch or adapter related problem is suspected. See “css.snap package” on page 218.

Note: **css.snap** uses a number of undocumented utilities to collect information. Some of these can be destructive when used on a running system. After using **css.snap** with snap type **full** (see Table 40 on page 219), it is advisable to run **/usr/lpp/ssp/css/rc.switch**. This command resets and reloads the switch adapter, and eliminates the residual effects of **css.snap**.

Diagnostic procedures

If your SP system or SP system partition shows signs of a switch failure, locate the symptom and perform the recovery action described. All the recovery actions described require that the user have **root** access on the specified SP Switch2 node.

Note: If your system is running in Restricted Root Access mode, the following commands must be issued from the control workstation:

- **CSS_test**
- **Efence**
- **Eprimary**
- **Equiesce**
- **Estart**
- **Eunfence**
- **Eunpartition**
- **mult_senders_test**
- **switch_stress**
- **wrap_test**

SP Switch2 diagnostics

Verify the SP Switch2 topology configuration

The switch plane topology file is used to define the hardware configuration to the css support software. It should reflect the number of switches and nodes installed, as well as define how they are connected.

The topology file can reside in two places: in the SDR, or in the **expected.top** file in the **/etc/SP** directory of the primary node. Usually, the configuration in the SDR is used. If the configuration in **/etc/SP/expected.top** on the primary node exists, it overrides the configuration in the SDR. The **/etc/SP/expected.top** on the primary is generally used for debugging proposes.

To verify that the topology in the SDR is correct, first read it out of the SDR using the command:

```
Etopology -read [- p (0 | 1 | all)]file_name
```

The **-p** flag indicates the switch plane number.

The **Etopology** command reads the switch topology from the SDR and places it in the specified file. For more information on this command, see *PSSP: Command and Technical Reference*.

Once the file is extracted, verify that the switch topology is an accurate representation of the installed hardware.

If changes to the switch topology file are required, remember to place them back into the SDR by issuing the **Etopology** command:

```
Etopology [- p (0 | 1 | all)] file_name
```

A default set of topology configuration files is available in the **/etc/SP** directory. For more information, see *PSSP: Command and Technical Reference*.

The SP Switch2 uses an annotated topology file, produced by the **Eannotator** command. The system administrator is responsible for running the command and creating the annotated topology file. When the file is not annotated, the fault service daemon will still work and the switch will function, but the switch jack numbers will not be correct in the topology file, the **out.top** file, and the **cable_miswire** file. If you suspect your topology file was not annotated, verify it. Examine the **out.top** file in **/var/adm/SPlogs/css0/p0** and **/var/adm/SPlogs/css1/p0** of each node, and examine the topology file by using the **Etopology -read** command described earlier.

Each link line in an annotated file is marked by `E`, as in this example:

```
s 13 0    s 23 0    E01-S17-BH-J6 to E02-S17-BH-J6
```

Each link line in a file that is not annotated is marked by `L`, as in this example:

```
s 13 0    s 23 0   L01-S00-BH-J9  to L02-S00-BH-J9
```

Verify the System Data Repository (SDR)

To verify that the SDR is installed and operating correctly, run the **SDR_test** command on the control workstation. It can be run either through **SMIT** panels or from the command line.

To verify SDR installation and operation from the **SMIT** panels:

1. Issue the command:

   ```
   smit SP_verify
   ```
2. The **RS/6000 SP Installation/Configuration Verification** menu appears.
3. Select: **System Data Repository**.
4. Press enter.
5. Review the output created.

To verify SDR installation and operation from the command line, enter:

```
/usr/lpp/ssp/bin/SDR_test
```

and review the output created.

Whenever the **SDR_test** command is run, a log file is created to enable the user to review the test results. The default log file created is: **/var/adm/SPlogs/SDR_test.log**. If the **SDR_test** command is run without **root** authority, the default log file created is: **/tmp/SDR_test.log**. Complete information on **SDR_test** can be found in *PSSP: Command and Technical Reference*. See also "SDR verification test" on page 130.

Next, login to the failing node and issue the command:

```
SDRGetObjects switch_responds node_number switch_responds0
```

Examine the output that is returned. If the switch responds bits are returned, this indicates that the SDR is operating. You can also determine which nodes are operational on the switch by examining the value returned. A value of 1 indicates that the node is operational on the switch. A value of 0 indicates that the node is not operational on the switch.

Verify node configuration

Verify that the SDR and the node configuration match each other. This verification procedure can be done on a node that tried to run the fault service daemon. On the node, view the file **/var/adm/SPlogs/css/rc.switch.log**. This file lists the configuration information found by the **rc.switch** script just before it attempts to run the fault service daemon.

Table 42. SP Switch2 rc.switch.log file and SDR equivalents

Line in rc.switch.log	Value in file	Value in SDR
Line 1	**date** and **time** when the information was taken.	
Lines 2, 4, 8	node configuration: **reliable_hostname**, **node_number** and **switch_node_number**	**SDRGetObjects** Node frame_number==*tested_node_frame* slot_number==*tested_node_slot* reliable_hostname node_number switch_node_number
Line 5	**switch_type** should be equal to 132 for the SP Switch2.	
Line 6	**number_switch_planes** should be equal to 1.	

Table 42. SP Switch2 rc.switch.log file and SDR equivalents (continued)

Line in rc.switch.log	Value in file	Value in SDR
Line 7	**adapter_name** and **adapter_status** should be only one device - css0 with css_ready. If the status is other then css_ready, see Table 37 on page 205.	
Lines 9, 10, 11	IP configuration: **netaddr** and **netmask**	**SDRGetObjects** Adapter node_number==*tested_node* adapter_type==css0 netaddr netmask

If any of the above node configuration data do not match, correct the SDR configuration and re-configure the problem node. See *PSSP: Installation and Migration Guide* for more details on how to do this.

SP Switch2 adapter diagnostics

The adapter diagnostics have two modes of operation: the Power-On-Self-Test (POST) and by online issuing of the **diag** command.

For the automatic POST tests scenario, issue the command:

```
diag -c -d [css0 | css1]
```

For advanced diagnostics scenarios, issue the command:

```
diag -A -d [css0 | css1]
```

The advanced tests check the cable wrap. You will need the card and cable wrap plug to complete these tests.

Note: The complete set of adapter diagnostics needs the exclusive use of the css adapter on the current node that the diagnostics are run on. Any other processes that have the css device driver open must be closed (**kill**ed) before issuing the adapter diagnostics command. One of those processes is the fault service daemon: **fault_service_Worm_RTG_CS**. Processes such as ″switch clock (TOD) reader applications″ make use of the css device driver and therefore these processe should be closed as well.

The diagnostics failures are reported to the AIX error log of the failing node. To view the adapter diagnostics errors:

1. Login to the failing node.
2. Issue the command:

   ```
   errpt -a | grep "Switch adapter failed POST diagnostics"
   ```

   to view the POST adapter diagnostics AIX error log entries.
3. In most cases each of the error entries will contain a Service Request Number (SRN).
4. Use the SRN to locate your error and its recovery actions in Table 43 on page 227.
5. Note that **x** may represent any value in Table 43 on page 227.

Table 43. SP Switch2 adapter Service Request Number failures and recovery actions

SRN	Recovery action
765-x1xx	See “Verify software installation” on page 228. If the verification is successful and the problem persists, contact the IBM Support Center.
765-1xx6	Wrap test failed. This problem is caused by a bad switch cable. Contact IBM Hardware Service and arrange to have the switch cable replaced.
765-2xx1	RDRAM test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx2	RDRAM Controller test failed. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx3	SRAM test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx4	SRAM Controller test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx5	DMA test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx6	Wrap test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx7	Registers test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx8	740 access test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xx9	Reassembly test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xxA	Segmentation test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
765-2xxB	Interrupts test failed. This problem is caused by a faulty adapter. Contact IBM Hardware Service and arrange to have the adapter replaced.
All other SRNs	Contact IBM Hardware Service and arrange to have the adapter or cable replaced.

Cable diagnostics

Switch to switch Cable Diagnostics

Visually inspect the cable in question:

1. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If everything looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them repair or replace the damaged components.
2. Repeat step 1 for the other end of the switch to switch cable.
3. Run the SP Switch Wrap Test and SP Switch Stress Test. See “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.
4. If everything visually checks out, contact IBM Hardware Service and have them replace the cable. If the problem persists, contact the IBM Support Center.

Node to switch cable diagnostics

Visually inspect the cable in question:

1. Remove the cable from the back of the node and examine the connectors (cable and back of the adapter) for bent pins or other visible damage. If

everything looks OK, reconnect the cable to the adapter. If not, contact IBM Hardware Service and have them repair or replace the damaged components.

2. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If everything looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them repair or replace the damaged components.
3. Run the SP Switch Wrap Test and SP Switch Stress Test. See “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.
4. If everything visually checks out, run advanced Adapter Diagnostics on the suspect adapter. The procedure is outlined in “SP Switch2 adapter diagnostics” on page 226. Follow the online instructions. If the diagnostics detect a failure, contact IBM Hardware Service and have them replace the failing components. If the adapter diagnostics pass and the problem persists, contact the IBM Support Center.
5. As a result of removing the cable, the node may be automatically fenced by the system. After reinstalling the cable, reboot the node or run the **rc.switch** command to reset the switch adapter. Only after this is complete, try to **Eunfence** the node.

SP Switch2 node diagnostics

Identify the failing node

Use this scenario if an application running on several nodes loses connectivity over the switch, or the **switch_responds** class indicates that several nodes are not on the switch. For more information on the **switch_responds** class, see the **SDRGetObjects** entry in *PSSP: Command and Technical Reference*.

1. View the summary log file, located on the control workstation.

   See “Summary log for SP Switch, SP Switch2, and switch adapter errors” on page 69.
2. Locate the first error log entry that indicates a node or connectivity failure.
3. Examine other entries to see if the first failure is the cause of subsequent failures.
4. On the node that experienced the first failure, examine the error log to see the complete version of the error log record described previously.
5. Use this as a starting point to debug the problem on this node.

Verify software installation

The software installation and verification are done using the **CSS_test** command on the control workstation. **CSS_test** can be run either through **SMIT** panels or from the command line.

If **CSS_test** is issued following a successful **Estart**, additional verification of the system is done to determine if each node in the system or system partition, can be **ping**ed.

To verify CSS installation from the **SMIT** panels:

1. Issue:

   ```
   smit SP_verify
   ```
2. The **RS/6000 SP Installation/Configuration Verification** menu appears.
3. Select: **Communications Subsystem**.
4. Press enter.

5. Review the output created.

Whenever the **CSS_test** command is run, a log file is created to enable the user to review the test results. The log file is **/var/adm/SPlogs/CSS_test.log**. Complete information on **CSS_test** can be found in *PSSP: Command and Technical Reference*.

To verify CSS installation from the command line:

1. Issue the command:

   ```
   /usr/lpp/ssp/bin/CSS_test
   ```

2. Review the log file to determine the results.

When running **CSS_test**, consider the following:

- The directory **/usr/lpp/ssp** should be accessible.
- The script file **/etc/inittab** on each node should contain an entry for the script **rc.switch**.

Verify SP Switch2 node operation

Use this procedure to verify that a single SP Switch2 node is operating correctly. If the node you are attempting to verify is the primary node, start with Step 1. If it is a secondary node, start with Step 2 on page 230.

1. Determine which node is the primary by issuing the **Eprimary** command on the control workstation. For complete information on the **Eprimary** command, see *PSSP: Command and Technical Reference*. For our purposes, consider this output:

   ```
   plane 0:  1     - primary
   plane 0:  5     - oncoming primary
   plane 0:  49    - primary backup
   plane 0:  45    - oncoming primary backup
   plane 0:  1     - autounfence

   plane 1:  1     - primary
   plane 1:  5     - oncoming primary
   plane 1:  49    - primary backup
   plane 1:  45    - oncoming primary backup
   plane 1:  1     - autounfence
   ```

   If the command returns a primary value of none, an **Estart** is required to make the oncoming primary node the primary.

   If the command returns an oncoming primary value of none, reissue the **Eprimary** command specifying the node that you would like to have as the primary node. Following the completion of the **Eprimary** command (to change the oncoming primary) an **Estart** is required to make the oncoming primary node the primary.

   Note: Both the **Eprimary** and **Estart** commands have a flag (**-p**) which specifies the number of the switch plane that the command references. If the **-p** flag is omitted, the command applies to all planes.

   The primary node on the SP Switch2 system can move to another node, if a primary node takeover is initiated by the backup. To determine if this has happened, look at the values of the primary and the oncoming primary backup. If they are the same value, a takeover has occurred.

2. Ensure that the node is accessible from the control workstation. This is done by using the **dsh** command to issue the **date** command on the node as follows:

   ```
   /usr/lpp/ssp/bin/dsh -w problem_hostname date
   ```

   The output is similar to:

   ```
   TUE JAN 25 10:24:28 EST 2000
   ```

   If the current date and time are not returned, refer to “Chapter 20. Diagnosing remote command problems on the SP System” on page 299.
3. Verify that the switch adapter (css0 or css1) is configured and is ready for operation on the node. This can be done by examining the **adapter_config_status** attribute in the **switch_responds** object of the SDR:

   ```
   SDRGetObjects Adapter adapter_type==css0 node_number adapter_config_status
   ```

   or

   ```
   SDRGetObjects Adapter adapter_type==css1 node_number adapter_config_status
   ```

   The output is similar to:

   ```
   node_number adapter_config_status
      1            css_ready
   ```

   If the **adapter_config_status** object is anything other than `css_ready`, see “Adapter configuration error information” on page 205. More information on the error may be found in the AIX error log. Run the **errpt -a** command on this node and match the adapter error log to the error list found in “AIX Error Log information” on page 188.
4. Verify that the **fault_service_Worm_RTG_CS** daemon is running on the node. This can be accomplished by using the **dsh** command on the control workstation to issue a **ps** command to the problem node as follows:

   ```
   /usr/lpp/ssp/bin/dsh -w problem_hostname ps -e | grep Worm_RTG
   ```

   The output is similar to the following:

   ```
   18422  -0:00 fault_service_Worm_RTG_CS
   ```

   If the **fault_service_Worm_RTG_CS** daemon is running, SP Switch2 node verification is complete.

   If the **fault_service_Worm_RTG_CS** daemon is not running, see “AIX Error Log information” on page 188. The possible reasons why the **fault_service_Worm_RTG_CS** daemon is not running are:

   - The daemon exited due to an abnormal error condition.
   - A **SIGTERM**, **SIGBUS**, or **SIGDANGER** signal was processed by the daemon.
5. Verify that the adapter is running on the node. This can be accomplished by **tail**ing the end of the **/var/adm/SPlogs/css0/adapter.log** or **/var/adm/SPlogs/css1/adapter.log** file. If the any of the last several lines have `fsd_adapter_thread_exit`, the adapter is not running. This means that the adapter has a permanent adapter error, or that the adapter diagnostics are running.

Look in the node's AIX error log for local adapter errors and handle them if they are found. If no local adapter errors are found, and the adapter diagnostics are not running, and still the **adapter.log** is in exit state, run the **rc.switch -l [css0 | css1]** command to restart the fault service daemon. Then, see if the end of the **adapter.log** now contains the message **------Adapter Thread Started------**, signaling that the adapter has been restarted.

Node crash

A node crash is generally identified by the LED/LCD display on the node flashing **888**. Do **not** reboot the node. See "Chapter 5. Producing a system dump" on page 81.

SP Switch2 Time Of Day (TOD) diagnostics

Some applications use the Switch Time-Of-Day (Switch TOD). The Switch TOD is a value that is passed only to nodes that are on the switch (their corresponding switch_responds0 flag is 1). The Most-Significant-Bit (MSB) of this 64 bit value is called the 'valid bit'. When the 'valid bit' is 1, the Switch TOD is valid. This means that the value you see is synchronized with the switch TOD. When the 'valid bit' is 0, the Switch TOD is invalid. This means that the value you see is propagated by the node and not synchronized with the switch TOD. In this case, you are not assured to have this value within range of the Switch TOD.

When your node begins to get the Switch TOD, the value of your Switch TOD will change to the Switch TOD value (if necessary) and the 'valid bit' will turn to 1. The **Emaster** command will show you the node number that is responsible for the Switch TOD. The **emaster** daemon, which runs on the control workstation, is monitoring the switch and tries to recover the Switch TOD when necessary.

The sections and subsections that follow are ordered according to the more probable cause of the problem. After each item, check the Switch TOD again, and if the problem persists, continue to the next item.

SP Switch2 TOD on node is not valid

Some of the nodes in your system have Switch TOD 'valid bit' turned ON and some have it turned OFF.

1. Validate that your node is on the switch; the switch_responds0 for your node is 1. One way to see that is by issuing:

   ```
   SDRGetObjects switch_responds node_number==your_node_number switch_responds0
   ```

2. If your node is not on the switch, you have to unfence it, then check your node's Switch TOD value. See "Unfence an SP Switch2 node" on page 240.
3. If your node is on the switch and still only your node shows the Switch TOD 'valid bit' as OFF (all the other nodes that are on the switch have their Switch TOD 'valid bit' turned ON), contact IBM Hardware Service.

SP Switch2 TOD is not valid on all nodes

All the nodes in your system have Switch TOD 'valid bit' turned OFF.

1. Verify that the switch is up and running (the **Eprimary** command shows one node as the primary node,) and the switch_responds0 of the primary node is 1. If the switch is down, run the **Estart** command, then check again.
2. Verify that you do have a Master Switch Sequencer (MSS) node. To do this, issue the command **Emaster** on your control workstation. The result should look like

   ```
   49    - Master switch sequencing node
   ```

3. Verify that the MSS node is not fenced OFF the switch. If it is fenced, use the **Eunfence** command to bring the node back on the switch.
4. Verify that the **emaster** daemon is running on your control workstation. To do this, issue the command **lssrc -s emaster**. The result should look like:

```
Subsystem         Group            PID     Status
 emaster           swt             25380   active
```

5. If the Status is ″inoperative″, check the AIX error log on the control workstation for the resignation cause, and follow the recommended actions. Restart the **emaster** daemon by issuing: **startsrc -s emaster**. Check that the daemon stays in active status.
6. If the Status is still ″inoperative″, call the IBM Support Center.
7. If the Status is Active and you have one or more nodes on the switch, but there still is no MSS node assigned, call the IBM Support Center.

No Master Switch-Sequencer (MSS) node

The Emaster command shows no node number as the Master Switch Sequencer (MSS) node.

1. Verify that the **emaster** daemon is running on your control workstation. To do this, issue the command **lssrc -s emaster**. The result should look like:

```
Subsystem         Group            PID     Status
 emaster           swt             25380   active
```

2. If the Status is ″inoperative″, check the AIX error log on the control workstation for the resignation cause, and follow the recommended actions. Restart the **emaster** daemon by issuing: **startsrc -s emaster**. Check that the daemon stays in active status.
3. Verify that the switch is up and running (the **Eprimary** command shows a valid node number as primary node,) and the **switch_responds** of this node is 1. If the switch is down, issue the **Estart** command.
4. If you have successfully run the **Estart** command, and some of the nodes came up on the switch, and there still is no MSS, call the IBM Support Center.

No SP Switch2 TOD monitoring

There is no Switch TOD recovery on your system. An event happened on your system that should have caused a replacement of the MSS node, but failed to replace the MSS node.

1. Check the AIX error log for the reason for the failure. Follow the recovery action suggested in the error log entry.
2. Validate that the **emaster** daemon is running on your control workstation. To do this, issue the command **lssrc -s emaster**. The result should look like:

```
Subsystem         Group            PID     Status
 emaster           swt             25380   active
```

3. If the Status is ″inoperative″, check the AIX error log on the control workstation for the resignation cause, and follow the recommended actions. Restart the **emaster** daemon by issuing: **startsrc -s emaster**. Check that the daemon stays in active status.
4. If the Status is still ″inoperative″, call the IBM Support Center.
5. Verify that the Event Management, Group Services and Topology Services groups are running. Issue these commands on the control workstation:
   - lssrc -g hags
   - lssrc -g hats

- lssrc -g haem

Issue these commands on the control workstation, to verify the daemons for each node:

- dsh -av lssrc -g hags
- dsh -av lssrc -g haem
- dsh -av lssrc -g haem

If any of these have ″inoperative″ status, you may experience a problem with Switch TOD recovery. To restart these you should run:

- stopsrc -g haem
- stopsrc -g hats
- stopsrc -g hags

and then run:

- startsrc -g hags
- startsrc -g hats
- startsrc -g haem

6. If the Status is still ″inoperative″, call the IBM Support Center.
7. You can monitor events reaching the **emaster** daemon by looking in the **emaster.log** located in the node log directory level.

SP Switch2 advanced diagnostics

To examine the SP Switch2 fabric in more detail, see “Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools” on page 245.

Error symptoms, responses, and recoveries

If your system shows signs of a switch failure, locate the symptom and perform the recovery action described. All the recovery actions described require that the user have **root** access on the specified switch node. If any of the recovery actions fail to resolve your problem, contact IBM Support Center.

Note: If your system is running in restricted root access mode, the following commands must be issued from the control workstation:

- **CSS_test**
- **Efence**
- **Eprimary**
- **Equiesce**
- **Estart**
- **Eunfence**
- **Eunpartition**
- **mult_senders_test**
- **switch_stress**
- **wrap_test**

SP Switch2 symptoms and recovery actions

Table 44 on page 234 lists the known symptoms of a failure in the SP Switch2, and points the user to the location of the detailed diagnostics and recovery action. You may have a symptom that does not appear in the table. In this case, view the error entry in the AIX error log and see “AIX Error Log information” on page 188.

Table 44. SP Switch2 symptoms and recovery actions

Symptom	Recovery action
Estart failure: 1. System cannot find **Estart** command. 2. Primary node is not reachable. 3. **Estart** command times out or fails. 4. Expected number of nodes not initialized. 5. Some links do not initialize.	1. See “Verify software installation” on page 228. 2. See “Verify SP Switch2 node operation” on page 229. 3. See “Estart error recovery” on page 239. 4. See “SP Switch2 device and link error information” on page 206. 5. See “SP Switch2 device and link error information” on page 206.
Nodes drops off of the switch. (**switch_responds** is 0 for the node.).	See “Verify SP Switch2 node operation” on page 229.
Nodes fail to communicate over the switch, but its **switch_responds** is 1(**ping** or **CSS_test** commands fail).	1. See “Verify SP Switch2 node operation” on page 229. 2. See “AIX Error Log information” on page 188.
Node crash.	See “Node crash” on page 231.
Node fails to **Eunfence**.	1. See “Unfence an SP Switch2 node” on page 240. 2. See “Eunfence error recovery” on page 241.
Oncoming primary node is fenced	1. See “Unfence an SP Switch2 node” on page 240. 2. See “Eunfence error recovery” on page 241.
Ecommand failure.	See “Ecommands error recovery” on page 239.
diag_fail condition for an SP Switch2 adapter	See “SP Switch2 adapter diagnostics” on page 226.
switch_responds is still 1 after node panic.	See “switch_responds is still on after node panic” on page 242.
The switch_plane and switch_plane_seq numbers in the **Switch** class of the SDR are incorrect. The **SDR_config** command appears to number switches incorrectly.	See “plane.info file” on page 187

Recover an SP Switch2 node

You can restart the **fault_service_Worm_RTG_CS** daemon on the node by issuing:

```
/usr/lpp/ssp/css/rc.switch
```

Following the **rc.switch**, run this command to determine if the daemon is still running or has died:

```
ps -e | grep Worm
```

At this point you should be able to **Eunfence** the node by issuing:

```
Eunfence problem_node_number
```

The output should be similar to the following:

```
All nodes successfully unfenced.
```

If you cannot resolve the problem, contact the IBM Support Center. You should also attempt to gather all log files associated with this failure. See “Information to collect before contacting the IBM Support Center” on page 218.

Worm error recovery

The following steps enable you to recover from switch initialization failures that impact the **worm** subsystem.

1. Login to the switch primary node. See the **Eprimary** command to determine which node is the switch primary node.
2. Run the command **errpt -a |pg** and search for message `Estart failed` label **CS_SW_ESTRT_FAIL_RE**. The Detail Data will give the return code of the failure:

   ```
   switch initialization failed with xx
   ```

   where *xx* is the number to look up in Table 45.
3. If you did not find the error log entry, view the bottom of the file **/var/adm/SPlogs/css0/p0/flt** or **/var/adm/SPlogs/css1/p0/flt**. Look for a message similar to:
   - CSworm_bfs_phase1() failed with rc=*xx*
   - CSworm_bfs_phase2() failed with rc=*xx*

   where *xx* is the number to look up in Table 45.
4. Use the rc (return code) to retrieve the appropriate entry in Table 45.
5. If the return code cannot be found in the table, or the actions taken did not correct the problem, contact the IBM Support Center.

Table 45. SP Switch2 worm return codes and analysis

Return code	Analysis
-2	**Explanation:** Adapter sender port is not connected to the switch. (Phase 1 failure). **Cause:** Oncoming primary is fenced off the switch. **Action:** Pick another node as primary and run **Estart** again. See “Unfence an SP Switch2 node” on page 240.
-3	**Explanation:** Adapter receiver port is connected to switch **Cause:** Oncoming primary is fenced off the switch. (Phase 1 failure). **Action:** Pick another node as primary and Estart again. See “Unfence an SP Switch2 node” on page 240.
-4	**Explanation:** Unable to generate routes for the network. (Phase 1 failure). **Cause:** Corrupted topology file. **Action:** See “Verify the SP Switch2 topology configuration” on page 224.

Table 45. SP Switch2 worm return codes and analysis (continued)

Return code	Analysis
-5	**Explanation:** Send packet from local node failed. **Cause:** Bad switch adapter. **Action:** Run the switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem, contact the IBM Support Center.
-6	**Explanation:** A switch miswire was detected. **Cause:** Switch network cabling does not match the switch topology file. **Action:** View the **/var/adm/SPlogs/css0/p0/cable_miswire** or **/var/adm/SPlogs/css1/p0/cable_miswire** file to determine which cables are in question. Then disconnect and check the associated cables. If the problem persists, contact IBM Hardware Service.
-7	**Explanation:** Unable to generate routes for the network. (Phase 1 failure). **Cause:** Primary node is not connected in the right switch capsule. **Action:** View the **/var/adm/SPlogs/css0/p0/cable_miswire** or **/var/adm/SPlogs/css1/p0/cable_miswire** file to determine which cables are in question. See *PSSP: Planning, Volume 2* for wiring information. **Cause:** The frame supervisor's TTY is not cabled properly. **Action:** Reconnect the frame supervisor's TTY and try again. If the problem persist, contact IBM Hardware Service.
-8	**Explanation:** Receive FIFO is full. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fail to isolate the problem, contact IBM Hardware Service. **Cause:** The switch is backed up from a node or a switch chip. **Action:** Contact the IBM Support Center.
-9	**Explanation:** Unable to initialize FIFOs. (Phase 1 failure). **Cause:** Software problem. **Action:** Contact the IBM Support Center.
-10	**Explanation:** Node found in Switch Chip's FIFO. (Phase 1 failure). **Cause:** Software problem. **Action:** Contact the IBM Support Center.
-12	**Explanation:** The **worm** was unable to contact the oncoming primary backup node. New backup will be selected. **Cause:** This is not an error. **Action:** None.
-13	**Explanation:** Switch chip id mismatch from a previous connected switch chip. (Phase 1 failure). **Cause:** Hardware problem. **Action:** If the problem persists, contact IBM Hardware service.

Table 45. SP Switch2 worm return codes and analysis (continued)

Return code	Analysis
-23	**Explanation:** The switch chip that the oncoming primary node is connected to did not respond. The oncoming primary node failed to communicate with the switch. (Phase 1 failure). **Cause:** Oncoming primary node is fenced. **Action:** Pick another node as oncoming primary and **Estart** again. See “Unfence an SP Switch2 node” on page 240. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the oncoming primary node. If they fail, try to isolate the problem. Contact IBM Hardware Service. **Cause:** Bad switch board. **Action:** If problem persist, Contact IBM Hardware Service.
-27	**Explanation:** The TBIC was not initialized. **Cause:** The switch adapter is uninitialized. **Action:** Run the script **rc.switch** on the primary node, then issue the **Estart** command from the control workstation. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem, contact IBM Hardware Service.
-36	**Explanation:** This node resigned as the primary node. **Cause:** The node determined it could no longer control and monitor the switch. The primary backup node is now in control of the switch. **Action:** No action required.
-41	**Explanation: Worm** reached retry limit. **Cause:** System cables may have a problem. The system is not stable. **Action:** View **/var/adm/SPlogs/css0/p0/out.top** or **/var/adm/SPlogs/css1/p0/out.top** on the oncoming primary node. Check or comment out all links that are marked as fenced or faulty. Check or comment out all nodes that were not found by the **worm**. Run **Estart** again. If problem persists, contact IBM Hardware service.
-43	**Explanation:** A read or write operation to the switch adapter failed. **Cause:** Bad switch adapter. **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fail to isolate the problem, contact IBM Hardware Service.
-51	**Explanation:** Unexpected return. The software experienced an unexpected values. **Cause:** This can happen for these reasons: ID mismatch, lock handling failure, unexpected SDR access, query or setting failure, null pointer that should have a value, unexpected memory updates failure, unexpected value inside a packet. **Action:** The software automatically create a **css.snap** file. Call the IBM Support Center with this file.

Table 45. SP Switch2 worm return codes and analysis (continued)

Return code	Analysis
-52	**Explanation:** No response from the switch chip connected to the oncoming primary. (Phase 1 failure). **Cause:** Cable between the oncoming primary node and switch board is faulty. **Action:** Check or reconnect the oncoming primary cable and try again. **Cause:** Adapter error on the oncoming primary node. **Action:** Run switch adapter diagnostics on the oncoming primary node. If fails, isolate the problem. Contact IBM Hardware Service.
-54	**Explanation:** Unknown device id returned from the switch chip that the oncoming primary is connected to. (Phase 1 failure). **Cause:** The frame supervisor's TTY is not cabled properly. **Action:** Reconnect the frame supervisor's TTY and try again. If the problem persist, contact IBM Hardware Service. **Cause:** System not configured properly. **Action:** Check your configuration on the control workstation and try again. **Cause:** Switch board hardware failure. **Action:** If the problem persists, contact IBM Hardware Service.
-55	**Explanation:** Switch chip signature test failed or failed to reset switch chip errors. (Phase 1 failure). **Cause:** The frame supervisor's TTY is not cabled properly. **Action:** Reconnect the frame supervisor's TTY and try again. If the problem persist, contact IBM Hardware Service. **Cause:** Switch board hardware failure. **Action:** If the problem persists, contact IBM Hardware Service.
-56	**Explanation:** Switch chip connected to the oncoming primary reported the primary to be connected to internal switch port. (Phase 1 failure). **Cause:** The frame supervisor's TTY is not cabled properly. **Action:** Reconnect the frame supervisor's TTY and try again. If the problem persists, contact IBM Hardware Service. **Cause:** Switch board hardware failure. **Action:** If the problem persists, contact IBM Hardware Service.

Table 45. SP Switch2 worm return codes and analysis (continued)

Return code	Analysis
-57	**Explanation:** Oncoming primary connected to the wrong switch. (Phase 1 failure). **Cause:** The frame supervisor's TTY is not cabled properly. **Action:** Reconnect the frame supervisor's TTY and try again. If the problem persist, contact IBM Hardware Service. **Cause:** System not configured properly. **Action:** Check your configuration on the control workstation and try again. **Cause:** Switch board hardware failure. **Action:** If the problem persists, contact IBM Hardware Service.
-58	**Explanation:** Oncoming primary connected in the wrong place. (Phase 1 failure). **Cause:** Cable miswire. View the **/var/adm/SPlogs/css0/p0/cable_miswire** or **/var/adm/SPlogs/css1/p0/cable_miswire** file to determine which cables are in question. See *PSSP: Planning, Volume 2* for wiring information.
-61	**Explanation:** Failed to reset the oncoming primary's switch chip's error registers. (Phase 1 failure). **Cause:** Switch board hardware failure. **Action:** If the problem persists, contact IBM Hardware Service.

Ecommands error recovery

Error isolation for any of the **Ecommands** is as follows.

1. View the error output returned from the command. Note the error message number and text.
2. Find the message in *PSSP: Messages Reference*.
3. More information can be obtained from the **Ecommands.log** trace file, see "Ecommands.log" on page 217.
4. Perform the recommended recovery action.
5. If the **Ecommand** failed because it was unable to communicate with every node, see "Chapter 18. Diagnosing SP Security Services problems" on page 251.
6. If the **Ecommand** failed because it cannot access the SDR, or the SDR is set up incorrectly, see "Verify the System Data Repository (SDR)" on page 224.
7. After the recovery action is taken, if the problem persists, see "Information to collect before contacting the IBM Support Center" on page 218 and contact the IBM Support Center.

Estart error recovery

To isolate and recover from failures in the **Estart** command, follow these steps:

1. Login to the primary node.
2. View the bottom of the file **/var/adm/SPlogs/css0/p0/flt** or **/var/adm/SPlogs/css1/p0/flt**.
3. Use the failure message as an index to Table 46 on page 240.
4. If the failure message cannot be found in the table, or the actions taken did not correct the problem, contact the IBM Support Center.

Table 46. SP Switch2 Estart problems and analysis

Return Code	Analysis
Error in **buildDeviceDatabase()**	**Explanation:** Unable to build the device database. **Cause:** Missing or corrupted topology file. **Action:** See “Verify the SP Switch2 topology configuration” on page 224. **Cause:** malloc failures. **Action:** See “Information to collect before contacting the IBM Support Center” on page 218 and contact the IBM Support Center.
Error in **CSswitchInit()**	**Explanation:** Unable to initialize the switch network. **Cause:** Switch initialization failed. **Action:** See “Worm error recovery” on page 235.
Error in **writeDeviceDatabase()**	**Explanation:** Unable to write **/var/adm/SPlogs/css0/p0/out.top** or **/var/adm/SPlogs/css1/p0/out.top** file. **Cause:** Missing or corrupted topology file. **Action:** See “Verify the SP Switch2 topology configuration” on page 224. **Cause:** The **/var** file system is not large enough to accommodate the new **out.top** file. **Action:** Increase the size of **/var**.
No valid backup - SDR current Backup being changed to none	**Explanation:** Informational message. **Cause:** No node available as a backup. **Action:** No action required.
Cannot access the SDR - SDR current Backup not changed	**Explanation:** SDR failures **Cause:** SDR not set up properly. **Action:** See “Verify the System Data Repository (SDR)” on page 224.
Error in: • **fopen(act.top.*PID*)** • **fprintf(act.top.*PID*)** • **fclose(act.top.*PID*)** • **rename(act.top, act.top.*PID*)**	**Explanation:** An error occurred accessing file **/var/adm/SPlogs/css0/p0/act.top.***PID* or **/var/adm/SPlogs/css1/p0/act.top.***PID*. **Cause:** File access problems. **Action:** Evaluate the errno returned and take the appropriate action. If the problem persists contact the IBM Support Center.

Unfence an SP Switch2 node

The recovery action to take depends on the current status of the switch, and the personality of the switch node to be unfenced. The SP Switch2 status is limited to whether it is operational or not. The personality of the switch node to unfence is whether or not the node is to become the primary node, primary backup node, or a secondary node of the switch. For more information on any of the commands used in this section, see *PSSP: Command and Technical Reference*.

To display the switch primary node and primary backup node on all switch planes, issue the command:

```
Eprimary
```

Example output is:

```
plane 0:  none  - primary
plane 0:  5     - oncoming primary
plane 0:  none  - primary backup
plane 0:  45    - oncoming primary backup
plane 0:  1     - autounfence

plane 1:  none  - primary
plane 1:  5     - oncoming primary
plane 1:  none  - primary backup
plane 1:  45    - oncoming primary backup
plane 1:  1     - autounfence
```

In this example, no primary node is available. Therefore, the SP Switch2 is not operational. Remember that the **Efence** and **Eunfence** commands work only when the **Eprimary** command shows a valid node number as the primary node.

Note: In the following commands, the **-p** flag can be used to specify the number of the switch plane. If the **-p** flag is not used, the command is applied to all switch planes.

1. SP Switch2 is operational and the node is to be the secondary. The node to unfence is not listed as the primary or the oncoming primary, and there is a primary node.

   Use the command **Eunfence** to unfence the node.
2. SP Switch2 is operational and the node is to be the primary. The node to unfence is the oncoming primary, and another node is currently the primary.
   - Use the command **Eunfence** to unfence the node.
   - Use the command **Estart** to set the node to its primary personality.
3. SP Switch2 is not operational and the node is to be the secondary. No node is listed as primary, and another node is listed as the oncoming primary.
   - Use the command **Estart** to operate the SP Switch2.
   - Use the command **Eunfence** to unfence the node.
4. SP Switch2 is not operational and the node is to be the primary. No node is listed as primary and the fenced node is listed as the oncoming primary.
   - Use the command **Eprimary** to set another node as oncoming primary.

     Select a node that is not fenced as an oncoming primary, otherwise the **Estart** command with fail again.
   - Use the command **Estart** to operate the SP Switch2.
   - Use the command **Eunfence** to unfence the node.
   - Use the command **Eprimary** to set the unfenced node to be the oncoming primary.
   - Use the command **Estart** to set the node as the switch primary.

Eunfence error recovery

This section is used to help you when you failed to unfence your node, following the unfence procedure described in “Unfence an SP Switch2 node” on page 240.

To isolate and correct most **Eunfence** problems, you should refer first to “Ecommands error recovery” on page 239.

Note: In the following commands, the **-p** flag can be used to specify the number of the switch plane. If the **-p** flag is not used, the command is applied to all switch planes.

The following list provides additional reasons for a particular node to fail to **Eunfence**:

1. The **Eunfence** of a node failed, but the SP Switch2 was not **Estart**ed. You cannot attempt to **Eunfence** any node on an SP Switch2 that is not started. Issue the **Estart** command.
2. The node can no longer be reached through the switch network. More information can be gathered from the **out.top** trace file, see "out.top" on page 215.
3. The SP Switch2 node failed to **Eunfence** because the switch topology could not be distributed. See "Chapter 18. Diagnosing SP Security Services problems" on page 251.
4. The node failed to respond when attempting to **Eunfence** it. See "SP Switch2 node diagnostics" on page 228 to isolate and correct the problem.
5. The user receives the message ″Cannot Unfence node xxx - timeout″, the most likely cause is that the fault service daemon (**fault_service_Worm_RTG_CS**) is not running on the node. If the this is the case, issue the **/usr/lpp/ssp/css/rc.switch** command to start the daemon. If the daemon is still not running, refer to the **rc.switch.log** trace file. See "rc.switch.log" on page 214 .
6. The user receives a message similar to ″Cannot Unfence node xxx - timeout″, and you have replaced the switch cable. See "Cable diagnostics" on page 227. Even though the fault service daemon (**fault_service_Worm_RTG_CS**) is running, you must issue the **/usr/lpp/ssp/css/rc.switch** command to reload and reset the adapter before you can try to **Eunfence** the node.
7. If any of the preceding procedures fail to resolve the problem, and the node is still fenced, gather the css logs of the primary node and the fenced node. This can be accomplished by logging into those nodes and issuing the **/usr/lpp/ssp/css/css.snap** command. See "Information to collect before contacting the IBM Support Center" on page 218.

switch_responds is still on after node panic

This section addresses the case where a node panics, **host_responds** becomes 0 and **switch_responds0** or **switch_responds0** are still 1. This is a valid condition when the SP Switch2 adapter, on the crashed node, has no outstanding requests to or from this node. The SP Switch2 is now in a state where it can become backlogged, since the link is still marked as up. This can cause problems on other parts of the SP Switch2 network.

Each node fault-service daemon is responsible for updating its **switch_responds0** and **switch_responds1**. The SP Switch2 primary node detects fallen links and turns off the appropriate **switch_responds** The **switch_responds0** or **switch_responds1** is turned on only during **Estart** or **Eunfence** command processing. A node panic with **switch_responds** 1 is a legitimate occurrence. There are two cases:

1. With primary or backup SP Switch2 nodes running, the **switch_responds0** or **switch_responds1** is updated only after a packet is sent to the panicked node. Therefore, a user can change **switch_responds0** or **switch_responds1** by trying to **ping** the panicked node. Having HA (**hats** and **hags**) run on the nodes

can remedy the situation, since they run IP packets between the nodes casually in order to check the links (LAN Adapter event).

2. Without primary or backup SP Switch2 node running, there is no switch control. In this case, **switch_responds0** or **switch_responds1** is not updated. Only a new **Estart** command, specifying the **-p** flag with the correct switch plane number, can correct this.

Chapter 17. SP Switch and SP Switch2 advanced diagnostic tools

ATTENTION - READ THIS FIRST

Do **not** activate the SP Switch and SP Switch2 advanced diagnostic facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of the IBM Support Center, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

You can run advanced switch diagnostic tests when you suspect that a component in the switch network is not performing properly. These tests must be run from the control workstation, with the exception of the Adapter Error Log Analyzer (ELA), which may be run from any switch node. The reason for running a diagnostic test is usually an error reported in the system error log, or one of the CSS log files. In many cases, it is difficult to isolate the cause of these errors. This is because the same error may be caused by a hardware failure of different network components, and because the error may be caused by other reasons, such as the network being initialized by the switch management software.

If you detect such errors, you may decide to invoke an appropriate diagnostic test, as explained in the following sections. These tests help to identify the failing component, and determine whether your system has a ″real″ hardware problem or not. The ″positive″ result of a test is a message that indicates a hardware failure. In this case, contact IBM hardware support.

Detailed information about the commands used to invoke these tests can be found in *PSSP: Command and Technical Reference*.

There are cases in which the test reports software errors. This indicates either a temporary error condition (for example, the **Estart** command is invoked while the test is running, in which case the test should be restarted), or a permanent problem in the test itself, or in other PSSP software. For these cases, contact the IBM Support Center. The displayed message contains the number of the node reporting the problem.

Before calling IBM, save the following information: node number of the primary node and the reporting node; and the log file, **spd.trace** on the primary node and on the reporting node. For the SP Switch, the log file is in directory **/var/adm/SPlogs/css**. For the SP Switch2, the log file is located on the port level directory. It has a path name of **/var/adm/SPlogs/css0/p0** or **/var/adm/SPlogs/css1/p0**. Refer to “SP Switch2 log and temporary file hierarchy” on page 187.

During the test, different messages are displayed on the SPD GUI (graphical user interface for SP Switch and SP Switch2 diagnostics) or to your console. These messages include several fields: node ID of the reporting node, its personality (primary, backup or secondary), timestamp, type of the message, and the message itself. There are several types of messages: informational messages, warnings, and errors. Warnings and error messages are marked by ″!″ and ″E″ respectively (in the proper field on the GUI). You can ignore informational messages (which do not have such marks).

For detailed information about warning and error messages, see *PSSP: Messages Reference*.

Adapter Error Log Analyzer (ELA)

Adapter ELA is an extension of adapter diagnostics. It is invoked on switch nodes in order to diagnose problems that occurred on the node, but cannot be reproduced by the adapter diagnostics tests. The reason for running the adapter ELA and the commands to invoke it are the same as those for adapter diagnostics and AIX diagnostics in general. Adapter ELA does not have any impact on the system.

When to run the adapter ELA

Run adapter ELA in all cases where you suspect that the SP Switch or SP Switch2 adapter is not functioning properly. These are some examples:

- The system error log contains adapter-related errors.
- The node loses connectivity to the rest of the switch network.
- An SP Switch or SP Switch2 diagnostic test reports an adapter error.
- An SP Switch or SP Switch2 diagnostic test recommends that the adapter be checked on a specific node.

How to run the adapter ELA

There are several ways to run the adapter ELA using the AIX **diag** command:

1. To run the adapter ELA directly from the command line, issue this command:

   ```
   diag -d [css0 | css1] -e
   ```

   Wait until the tool gives you progress messages.
2. To run the adapter ELA from a GUI:
   a. Issue the AIX **diag** command without operands.
   b. When the first screen appears, press the enter key.
   c. Highlight the line ″Task Selection (Diagnostics, Advanced Diagnostics, Service Aids, etc...)″ and press the enter key.
   d. On the next screen, scroll down until you see ″Run Error Log Analysis″, highlight it and press the enter key.
   e. On the next screen, scroll down until you see **css0** or **css1**, highlight it and press the enter key.
   f. A plus sign **(+)** appears to the left of **css0** or **css1**.
   g. Press PF7 to commit. This starts the adapter ELA.
   h. Messages indicating the results of the adapter ELA appear on the top of the screen.

Interpreting the results of the adapter ELA

The adapter ELA first notifies you that it is testing the adapter, and asks you to stand by. Then, it displays an **ELA recommendations** screen, which either informs you that no problems were found, or gives a list of recommendations. These messages are self-explanatory. For example, the message may say that an SP Switch or SP Switch2 adapter problem was detected, and that you should contact IBM Software Support or IBM Hardware Support.

If the adapter ELA results are not displayed, or the adapter ELA reports that it failed to operate correctly (for example, wrong or missing files), the problem is probably due to incorrect PSSP installation. In this case, contact the IBM Support Center.

SP Switch or SP Switch2 stress test

This test verifies the functionality of a specific switch chip or switch port. The **switch_stress** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. The **-g** flag can be used to run the test with the SPD GUI, but you must specify the command operands and flags on the command line.

The **-n** flag is used to specify the switch plane number. **-n 0** specifies plane 0 and **-n 1** specifies plane 1. Both planes cannot be specified on one invocation of this command.

When to run the SP Switch or SP Switch2 stress test

Run this test when you suspect that a switch chip is faulty because the system error log on the primary node contains error status reports from the switch chip. This test checks the functionality of this switch chip and the attached links, in order to decide whether the switch chip should be replaced.

How to run the SP Switch or SP Switch2 stress test

First, decide which switch chip you want to test. Usually this is a switch chip that is reporting errors. Then, decide which nodes to use for the test. **These nodes cannot run parallel applications during this test.** By default, all nodes are allowed, so be careful to avoid disturbing applications that are running. The nodes that are not allowed will not be affected directly. However, since the test implies stress traffic in the switch network, the performance of applications running on **all** nodes may be affected.

Invoke the test, specifying the desired switch chip ID. Also specify the nodes that can participate in the test, or alternatively the nodes that are forbidden (because they are running critical applications).

The test does not require user intervention. It runs several iterations, each one using a different combination of switch chip ports. In each iteration, the test sends data through the ports under test. In the beginning of the iteration, the test notifies you as to which nodes are participating in the test. At the end of the iteration, it displays these statistics: number of packets that were sent and lost, and number of switch errors (reported from all switch chips used for sending data by the test iteration).

Interpreting the results of the SP Switch or SP Switch2 stress test

Each error reported by a switch chip is displayed. In a stable system, there should be few or no such reports. If the test succeeded in causing a critical fault on one of the switch chips, the test decides that its goal has been achieved. The faulty component is isolated. In this case, the test displays an appropriate message and terminates. Contact IBM hardware support to replace the faulty component.

Otherwise, the test just displays the statistics and continues to the next iteration. If the test did not cause critical faults, but did cause some failures (that were recovered), it does not necessarily mean that some hardware component should be replaced. This result gives an indication of a possible cause of problems. Contact IBM hardware support in this case also.

Multiple senders/single receiver test

This test detects nodes that are injecting corrupted packets into the switch network. The **mult_senders_test** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. The **-g** flag can be used to run the test with the SPD GUI, but you must specify the command operands and flags on the command line.

The **-n** flag is used to specify the switch plane number. **-n 0** specifies plane 0 and **-n 1** specifies plane 1. Both planes cannot be specified on one invocation of this command.

When to run the Multiple senders/single receiver test

Run this test if one or more of the nodes are reporting that they received ″bad packets.″ This may indicate a situation where there is a malfunctioning switch adapter in the system that is generating bad packets. You want to detect such ″bad sender″ nodes. The multiple senders/single receiver test finds the malfunctioning switch adapter among all of the nodes in the system.

How to run the multiple senders/single receiver test

Select a receiver node to be used by the test. The receiver node is usually one of the nodes that are reporting bad packet events. Then decide which nodes can be used as senders. **The receiver node and nodes that will be used as senders cannot run parallel applications during this test.** By default, all nodes are allowed, so be careful to avoid disturbing applications that are running. The nodes that are not allowed will not be affected directly. However, since the test implies stress traffic in the switch network, the performance of applications running on **all** nodes may be affected.

Invoke the test specifying the desired receiver node. Also specify the nodes that are allowed to be used as senders, or alternatively the nodes that are forbidden (because they are running critical applications).

The test does not require user intervention. In the beginning, the test notifies the user which sender nodes are participating. The selected senders send data to the receiver, and the test progress is monitored until it completes.

Interpreting the results of the multiple senders/single receiver test

The test monitors error reports from all switch network components during the iterations. If a critical fault occurs, the test displays a message about it and terminates. In this case, contact IBM hardware support to replace the faulty component. Otherwise, the test continues until all iterations are done. The test then either displays a message that contains the list of bad senders, or notifies you that no errors were found.

SP Switch or SP Switch2 wrap test

This test verifies the functionality of a link. The **wrap_test** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. This command must be used with the SPD GUI.

The **-n** flag is used to specify the switch plane number. **-n 0** specifies plane 0 and **-n 1** specifies plane 1. Both planes cannot be specified on one invocation of this command.

When to run the SP Switch or SP Switch2 wrap test

Run the SP Switch or SP Switch2 Wrap test if:

- The **Estart** command reports that it failed to initialize links or nodes.
- The error log on the primary node contains error reports about a link between two switches.
- The error log on the primary node contains error reports about a link between a node and a switch.

These conditions may indicate a hardware problem in the cable between two ports. Run this test to determine which component should be replaced. The SP Switch or SP Switch2 Wrap Test identifies the specific failing component.

How to run the SP Switch or SP Switch2 wrap test

First, decide which link to test, using the information in the error logs. If the link under test is a link connecting a switch to a node, fence the node before running the test. If the link under test is a link connecting two switches, be aware that during the test the link will be disabled.

The test guides you through several steps. In each step, a message box is displayed asking you to perform some operation, such as installing the wrap plug, according to the displayed information. You can either perform the requested operation and press OK, or press CANCEL. The test then checks the corresponding link component and displays the results (component passed or failed). If the user pressed CANCEL, or in a few other cases, the test cannot check the component and the test terminates. In these cases, it displays a message that explains the situation.

At the end of the test, you are requested to perform tasks to restore the system to the state it was in before the test.

Interpreting the results of the SP Switch or SP Switch2 wrap test

The test displays explicit diagnosis information for the link components.

Chapter 18. Diagnosing SP Security Services problems

This chapter discusses diagnostic procedures and failure responses for the Security Services component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 271. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 258.

Related documentation

The following publications provide information about SP Security Services.

- *PSSP: Planning Volume 2*

  This book describes configuration decisions that are made when setting up the SP Security Services on your system. You should be familiar with these choices and the resulting requirements. These requirements are for users of the system to be authenticated by logging into DCE or Kerberos V4, and for proper authorizations to allow appropriate access to system software components.
- *PSSP: Installation and Migration Guide*

  Several tasks that are performed when installing a new SP system or migrating from a prior release of PSSP determine the security capabilities of the software components. The understanding of, and correct performance of, these tasks are important for avoiding security-related problems when using SP system management facilities or network applications.
- *PSSP: Administration Guide*

  An important difference between this PSSP release and prior releases is the separation between roles of AIX super-user and system security administrator. When using DCE for security, the PSSP security infrastructure separates tasks that must be performed by the **root** user from those that must be performed by a DCE cell administrator. One side effect is that DCE has a larger number of discrete tasks to be performed for setup and administration, than Kerberos V4.
- *PSSP: Messages Reference*

  These chapters contain messages related to SP Security Services:
  - 2502 - Authentication Messages
  - 2503 - Kerberos Messages
  - 2504 - Kerberos Messages
  - 2545 - Authentication Installation and Configuration Messages
- *PSSP: Command and Technical Reference*

  Refer to the entries for the following commands and files for security considerations and restrictions on their use:
  1. Common PSSP security configuration
     - chauthpar
     - chauthpts
     - chauthts
     - get_keyfiles
     - kfserver
     - lsauthpar
     - lsauthpts
     - lsauthts

- SDRSetTsAuth
- setup_CWS
- spauthconfig
- spseccfg
- spsetauth

2. PSSP security configuration for DCE
   - config_spsec
   - create_dcehostname
   - create_keyfiles
   - setupdce
   - spsec_overrides
3. PSSP security administration for DCE
   - hmdceobj
   - spacl
   - spnkeymand
4. PSSP security use for DCE
   - dsrvtgt
   - spgrpname
   - sptgtprin
5. PSSP security configuration for Kerberos V4
   - add_principal
   - create_krb_files
   - kstash
   - setup_authent
6. PSSP security administration for Kerberos V4
   - chkp
   - ext_srvtab
   - hmacls
   - kadmin
   - kadmind
   - kdb_destroy
   - kdb_edit
   - kdb_init
   - kdb_util
   - kerberos
   - kprop
   - kpropd
   - krb.conf
   - krb.realms
   - ksrvutil
   - lskp
   - mkkp
   - rmkp
   - sysctl.acl
7. PSSP security use for Kerberos V4

  - k4destroy
  - k4init
  - k4list
  - kpasswd
  - ksrvtgt
  - rcmdtgt
- *IBM DCE for AIX, Version 3.1: Administration Commands Reference*

  Refer to entries for the following commands. This is the list of commands other than **dcecp**:

  1. For installation and configuration:
     - config.dce
     - unconfig.dce
     - kerberos.dce -type local
  2. For general DCE status: show.cfg
  3. For starting and stopping DCE:
     a. start.dce
     b. stop.dce
  4. For obtaining DCE credentials (for use with other DCE commands and with AIX secure remote commands):
     - dce_login -f
     - kinit -f
  5. For displaying the state of DCE credentials: klist -f
  6. For destroying DCE credentials: kdestroy

  These are all **dcecp -c** commands:

  - dcecp -c cell show
  - dcecp -c group catalog
  - dcecp -c group show -all
  - dcecp -c group list
  - dcecp -c group add
  - dcecp -c group remove
  - dcecp -c org catalog
  - dcecp -c org show -all
  - dcecp -c org list
  - dcecp -c org add
  - dcecp -c acl show
  - dcecp -c acl perm
  - dcecp -c acl modify
  - dcecp -c keytab show
  - dcecp -c keytab list
  - dcecp -c keytab catalog
  - dcecp -c keytab add
  - dcecp -c keytab remove
  - dcecp -c principal catalog
  - dcecp -c principal show -all
  - dcecp -c principal list

  - dcecp -c account catalog
  - dcecp -c account show -all
  - dcecp -c account list
  - dcecp -c registry catalog -master
  - dcecp -c registry show -policies
  - dcecp -c secval status
  - dcecp -c secval ping
- *IBM DCE for AIX, Version 3.1: Administration Guide - Introduction*
- *IBM DCE for AIX, Version 3.1: Administration Guide - Core Components*
- *IBM DCE for AIX, Version 3.1: DFS Administration Guide and Reference*
- *IBM DCE for AIX, Version 3.1: Application Development Guide - Introduction and Style Guide*
- *IBM DCE for AIX, Version 3.1: Application Development Guide - Core Components*
- *IBM DCE for AIX, Version 3.1: Application Development Guide - Directory Services*
- *IBM DCE for AIX, Version 3.1: Application Development Reference*
- *IBM DCE for AIX, Version 3.1: Problem Determination Guide*

  The following sections can help in diagnosing problems encountered when using DCE for SP security:

  - Message and Message ID Structure
  - File Systems Used by DCE
  - Keytab Files
  - Log Files on AIX
  - Checking the Security Servers
  - Checking User Accounts
  - Checking Access Permissions
  - Using DCE Debug and Trace Options
  - Common Problems and Their Resolution
  - Mapping DCE Daemon Core File Locations
- *IBM DCE for AIX, Version 3.1: Release Notes*

Requisite function

This is a list of the software and operating system resources directly used by the SP Security Services component of PSSP. Problems within the requisite software or resources may manifest themselves as error symptoms in SP Security Services. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with SP Security Services, you should consider the following components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

Distributed Computing Environment (DCE) Version 3.1 for AIX

> **DCE Restriction**
>
> If you have DCE authentication enabled, you cannot run HACWS.

Note: The DCE daemons will be affected if the **/var** file system becomes full. This will prevent users from logging into DCE. Refer to IBM DCE for AIX manuals for more information.

PSSP security services provides the option to use DCE to enhance the security of SP client/server programs known as SP trusted services, and of the AIX remote commands. This option is supported with DCE Version 3.1 or higher. If you choose to use this option, diagnostic information in messages and logs will be a combination of PSSP information and DCE information.

The various PSSP commands that you use to configure the system to use DCE for authentication and authorization invoke DCE commands that are described in the DCE product publications. Errors reported by those commands are logged with PSSP information in the log file described in “Log files” on page 256. SP trusted services receive error status from the security services library subroutines, which may contain DCE and DCE GSSAPI error codes. Messages and log entries created by the various services will contain message text obtained from DCE library subroutines.

If an error reported by DCE cannot be attributed to a readily corrected user problem, such as failure to login or failure to start DCE servers, consult the appropriate DCE publications.

System Data Repository (SDR)

Note: If you have migrated to PSSP 3.2 from a level of PSSP earlier than 3.1, there are extra security steps which are required. If these steps are omitted, the information in the SDR may be incorrect. See “Action 6 - Reset authentication values” on page 286.

The security configuration settings for each system partition are maintained in the **Syspar** object in the SDR. That information is set by various configuration commands. It is used by security administration commands, scripts that perform SP node initialization, and the security services runtime library, used by various SP trusted services.

For Kerberos V4 security, information is kept in the SP object about the type of Kerberos V4 authentication server that is being used. Also, the management of Kerberos V4 service principals used on SP nodes is dependent on node name and adapter information in the **Node** and **Adapter** objects.

The SDR contains information specific to each security setup. Common information includes:

- in the **Syspar** object for each system partition:
  - auth_install
  - auth_methods
  - auth_root_rcmd
  - ts_auth_methods

For DCE, this includes:

- in the **SP** object
  - sec_master
  - cds_server

  - cell_name
  - cw_dcehostname
- in the **Node** object
  - dcehostname

For Kerberos V4, this includes, in the **SP** object:

- authent_server

Error information

This section describes the information provided about error conditions detected by this component.

Log files

PSSP security components maintain several log files. One is used to record the administrative tasks that are performed when configuring the security services on the system. It contains text messages indicating both normal progress and error conditions encountered in performing those tasks. The other files are kept by the Kerberos V4 daemons. The content is also message text.

The **kerberos** and **kadmind** daemons record only their initial startup and any errors encountered. The **kpropd** daemon records the normal progress of Kerberos V4 database propagation as well as errors encountered.

SP Security Services configuration log

SP security configuration scripts that are used to set up the security services during installation, migration, and partitioning record their progress and completion status in the following file: **/var/adm/SPlogs/auth_install/log**. The configuration log file is trimmed automatically at midnight to its last 500 lines.

This is an example of the configuration log:

```
************************************************
*   Beginning of logging for -- spsetauth
*   Fri Dec  3 14:39:33 1999
************************************************

************************************************
*   Beginning of logging for -- updauthfiles
*   Fri Dec  3 14:39:35 1999
************************************************

updauthfiles: This command invoked as:  updauthfiles

updauthfiles: About to update local file: /.rhosts

updauthfiles: Updated local file: /.rhosts

updauthfiles: About to update local file: /.k5login

************************************************
***  End of logging for -- updauthfiles
***  Fri Dec  3 14:39:57 1999
```

Kerberos V4 daemon logs

Each Kerberos V4 daemon has a log file in which errors are recorded. The Kerberos V4 log files are not automatically trimmed.

- **kerberos**, running as the primary authentication server, uses: **/var/adm/SPlogs/kerberos/kerberos.log**.
- **kerberos**, running as a secondary authentication server, uses: **/var/adm/SPlogs/kerberos/kerberos_slave.log**.
- **kpropd**, the secondary authentication database server, uses: **/var/adm/SPlogs/kerberos/kpropd.log**.
- **kadmind**, the primary authentication database server, uses: **/var/adm/SPlogs/kerberos/admin_server.syslog**.

Sample DCE daemon log

For a full list of DCE daemon logs, refer to *IBM DCE for AIX, Version 3.1: Administration Guide - Core Components*. This is an example of the log file:

```
9:19:57: ------------------------------------------------------------
          show.cfg is starting at 1999-11-09-09:19:57.078-05:00I---.

09:19:57: The DCE configuration log file was successfully backed up to
                                /opt/dcelocal/etc/cfgdce.bck.
09:19:57: Parsing command line data.
09:19:57: show.cfg
09:19:57: Querying the supported protocols.
09:19:57: Using protseqs.rpc.
09:19:57: Querying IP information.
09:19:57: Querying currently configured components.
09:19:57: Reading the file, /opt/dcelocal/etc/cfgdce.dat...
09:19:57: Finished reading the file, /opt/dcelocal/etc/cfgdce.dat.
09:19:57: Determining the security server type.
09:19:57: Determining the Directory server type.
09:19:57: Waiting up to 30 seconds for "mget_dir_reptype_worker"
                                  to complete.
09:20:02: Authenticating as hosts/c166cw.ppd.pok.ibm.com/self.
09:20:03: Gathering component state information...
09:20:03: Querying the supported protocols.
09:20:03: Using protseqs.rpc.
09:20:03: Querying IP information.
09:20:03: Reading the file, /opt/dcelocal/etc/cfgdce.dat...
09:20:03: Finished reading the file, /opt/dcelocal/etc/cfgdce.dat.
09:20:03: Querying currently configured components.
09:20:03: Reading the file, /opt/dcelocal/etc/cfgdce.dat...
09:20:03: Finished reading the file, /opt/dcelocal/etc/cfgdce.dat.
09:20:03: Determining the security server type.
09:20:03: Determining the Directory server type.
09:20:03: Waiting up to 30 seconds for "mget_dir_reptype_worker"
                                             to complete.
09:20:03: Reading the file, /opt/dcelocal/etc/cfgdce.dat...
09:20:03: Finished reading the file, /opt/dcelocal/etc/cfgdce.dat.
09:20:03: Writing the file, /opt/dcelocal/etc/cfgdce.dat...
09:20:03: Finished writing the file, /opt/dcelocal/etc/cfgdce.dat.
09:20:04: Waiting up to 60 seconds for "check_secval_status_worker"
                                             to complete.
09:20:04: Querying the supported protocols.
09:20:04: Using protseqs.rpc.
09:20:04: Waiting up to 20 seconds for "isListeningWorker dced"
                                      to complete.
09:20:04: Querying the currently configured cell name.
09:20:04: Querying the supported protocols.
09:20:04: Using protseqs.rpc.
09:20:05: Waiting up to 20 seconds for "isListeningWorker rpc"
                                    to complete.
09:20:05: Querying the currently configured cell name.
09:20:05: Querying the supported protocols.
09:20:05: Using protseqs.rpc.
09:20:06: Waiting up to 20 seconds for "isListeningWorker sec_svr"
```

```
                                               to complete.
09:20:06: Querying the currently configured cell name.
09:20:08: Waiting up to 20 seconds for "isListeningWorker cds_svr"
                                               to complete.
09:20:08: Querying the currently configured cell name.
09:20:10: Waiting up to 20 seconds for "isListeningWorker cds_cl"
                                               to complete.
09:20:10: Querying the currently configured cell name.
09:20:10:
09:20:10:          Component Summary for Host:   c166cw.ppd.pok.ibm.com
09:20:10:          Component         Configuration State   Running State
09:20:10: Security Master server           Configured          Running
09:20:10: Security client                  Configured          Running
09:20:10: RPC                              Configured          Running
09:20:10: Initial Directory server         Configured          Running
09:20:10: Directory client                 Configured          Running
09:20:10:
09:20:10: The component summary is complete.
09:20:11: logout hosts/c166cw.ppd.pok.ibm.com/self
```

Dump information

Core files created by AIX for errors that occur in security services daemons are located as follows:

- All Kerberos V4 daemons use **/var/kerberos/database** as their current directory.
- For DCE, see *IBM DCE 3.1 Problem Determination Guide* Mapping DCE Daemon Core Locations section.

Information to collect before contacting the IBM Support Center

Before collecting the information listed here, make sure you have used the steps described in “Diagnostic procedures” on page 260 to verify the correct installation and configuration of security services on your system. Also, make sure that you have followed any other steps that apply to the specific error situation. You should always include the following information:

- The level of PSSP software on each affected system. See “Knowing your SP structure and setup” on page 7.
- The level of other installed products containing failing trusted services
- Record of tasks performed (**smit.log** or stdout and stderr from command-line session)
- The authentication method in use. Issue this command on the control workstation:

  ```
  splstdata -p
  ```

  The entry ″ts_auth_methods″ lists the authentication methods in use.

SP Security Services configuration errors

When errors occur while setting up the security services on your system, or error occur while initializing trusted services daemons, the following information may be required:

1. PSSP Security Services configuration for DCE:
   - Relevant parts of the **/var/adm/SPlogs/auth_install/log** file
   - Contents of the **spsec_overrides** file, if used
   - Output of the **dcecp -c account cat** command

- Output of the **dcecp -c keytab cat** command
- Output of the **ls -lR /spdata/sys1/keyfiles** command
- Output of the **splstdata -p** command
- Output of the **splstdata -e** command
- Output of the **splstdata -n** command
- Output of the **dcecp -c group list** *groupname* command for any relevant groups
- Output of the **lsauthpts -v -p** *sysparname* command for each affected partition
- Output of the **lsauthpar -v -p** *sysparname* command for each affected partition

2. Configuration for Kerberos V4 (Compatibility):
   - Contents of Kerberos V4 daemon log files. See “Kerberos V4 daemon logs” on page 256.
   - Contents of the **/etc/krb.conf** file
   - Contents of the **/etc/krb.realms** file
   - Output of the **splstdata -a -G** command
   - Output of the **netstat -in** command on the control workstation
   - Output of the **ksrvutil list** command
   - Output of the **splstdata -p** command
   - Output of the **splstdata -e** command

SP trusted services authentication errors

When errors indicating that the user could not be authenticated occur while running trusted services client programs, and you suspect a software problem, the following information may be required:

1. Using DCE:
   - Error messages displayed by failing commands or SP Perspectives panels
   - Output of the **klist** command for the client
   - Output of the **lsauthpts** command
   - Output of the **lsauthpar** command
   - Output of **dcecp** commands, showing principal, account, and keytab information. For command syntax, see *IBM DCE for AIX, Version 3.1: Administration Commands Reference*.
2. Using Kerberos V4 (Compatibility)
   - Error messages displayed by failing commands or SP Perspectives panels
   - Output of **k4list** command
   - Output of the **lsauthpts** command
   - Output of the **lsauthpar** command
   - Output of the **lskp** command, showing client principal and service principal
   - Contents of server log files. See “Kerberos V4 daemon logs” on page 256.

SP trusted services authorization errors

When errors indicating that the user is not authorized to perform a task occur while running trusted services client programs, but no authentication error is reported, check first with your security administrator to verify the user’s authorization. If a software problem is suspected, the following information may be required:

1. Using DCE (all trusted services)
   - Error messages displayed by failing commands or SP Perspectives panels
   - Output of the **klist** command for the client
   - Output of **dcecp** commands showing membership of relevant access groups
   - Output of **dcecp** commands showing ACL entries for relevant objects. For command syntax, see *IBM DCE for AIX, Version 3.1: Administration Commands Reference*.
   - Contents of server log files
2. Using Kerberos V4 (Sysctl, Hardware Monitor)
   - Error messages displayed by failing commands or SP Perspectives panels
   - Output of the **k4list** command for the client
   - Contents of hardmon ACL file, **spdata/sys1/spmon/hmacls**, if relevant
   - Contents of Sysctl ACL files, if relevant. See the chapter ″Sysctl″ in *PSSP: Administration Guide*
   - Contents of modified and added Sysctl configuration files, **etc/sysctl.conf**
   - Contents of server log files

Diagnostic procedures

This section describes how to check whether the security services component is operating as expected.

See 253 for a list of DCE commands that may be helpful. For more information on each command, see *IBM DCE for AIX, Version 3.1: Administration Commands Reference*.

Find out about your configuration

Perform these steps to determine your current configuration:

1. Issue **splstdata -p** to find out the number of partitions and the security methods enabled. See “Check enabled security” on page 265. SP Security Services will use DCE first if enabled and then Kerberos V4. Keep this in mind when performing problem determination.
2. Check if your SP system is configured for Kerberos V4 use by issuing the **lskp** command and look for principals similar to **hardmon.***partition_name*, **rcmd.***partition_name*, and **root.admin**.

   The output of **lskp** is similar to:

   ```
   hardmon.c166cw  tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   hardmon.c166s   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   hardmon.c166sp1 tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   hardmon.c166sp2 tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   hardmon.c186cw  tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   rcmd.c186sn04   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   rcmd.c186sn05   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   rcmd.c186sn06   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   rcmd.c186sp1    tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   root            tkt-life: 30d       key-vers: 2 expires: 2037-12-31 23:59
   root.SPbgAdm    tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
   root.admin      tkt-life: 30d       key-vers: 1 expires: 2037-12-31 23:59
   ```
3. Check if your SP system is configured for DCE use by issuing the **dcecp** command and looking for principals similar to: *cell_name***/ssp/***hostname***/***service*.

   Issue this command:

   ```
   dcecp -c principal cat | grep ssp
   ```

Output is similar to:

```
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/css
/.../c166dcecell/ssp/c168n01.ppd.pok.ibm.com/css
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/hardmon
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/pmand
/.../c166dcecell/ssp/c166s/sdr
/.../c166dcecell/ssp/c166sp1/sdr
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/sp_configd
/.../c166dcecell/ssp/c168n01.ppd.pok.ibm.com/sp_configd
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/spbgroot
/.../c166dcecell/ssp/c168n01.ppd.pok.ibm.com/spbgroot
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/spmgr
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/switchtbld
/.../c166dcecell/ssp/c168n01.ppd.pok.ibm.com/switchtbld
/.../c166dcecell/ssp/c166cw.ppd.pok.ibm.com/sysctl
/.../c166dcecell/ssp/c168n01.ppd.pok.ibm.com/sysctl

/.../c166dcecell/ssp/c166n01.ppd.pok.ibm.com/css
/.../c166dcecell/ssp/c166n13.ppd.pok.ibm.com/css
```

Verify DCE Security Services configuration

Verify the configuration for DCE or Kerberos V4, depending on which method you are using.

These procedures check installed file sets, daemons, principals, accounts, groups, rpc entries, keytabs, and key files. Note that names of DCE entities shown reflect the use of the default names, with any changes that may have been specified in the **spdata/sys1/spsec/spsec_overrides** file.

Check that required DCE file sets are installed

Systems using DCE for security services require the core client file sets. Refer to DCE product documentation for the list of expected file sets. Refer to PSSP release documentation for required PTF levels, if any. Compare the list of expected file sets, and those of any PTFs, with those on your system. Issue the command:

```
lslpp -l dce*
```

Output is similar to:

```
Fileset                      Level   State      Description
----------------------------------------------------------------
Path: /usr/lib/objrepos
dce.cds.rte                3.1.0.0 COMMITTED DCE Cell Directory Services
dce.client.core.rte        3.1.0.0 COMMITTED DCE Client Services
dce.client.core.rte.admin  3.1.0.0 COMMITTED DCE Client Administrative Tools
dce.client.core.rte.cds    3.1.0.0 COMMITTED DCE Client CDS Tools
dce.client.core.rte.config 3.1.0.0 COMMITTED DCE Client Configuration Tools
dce.client.core.rte.rpc    3.1.0.0 COMMITTED DCE Client RPC Tools
dce.client.core.rte.security
                           3.1.0.0 COMMITTED DCE Client Security Tools
dce.client.core.rte.time   3.1.0.0 COMMITTED DCE Client Time Tools
dce.client.core.rte.zones  3.1.0.0 COMMITTED DCE Client Time Zones
dce.compat.cds.smit
                        3.1.0.0 COMMITTED DCE SMIT Cell Directory Services
dce.compat.client.core.smit
                        3.1.0.0 COMMITTED DCE SMIT Client Tools
dce.compat.security.smit   3.1.0.0 COMMITTED DCE SMIT Security Services
dce.pthreads.rte
                        3.1.0.0 COMMITTED DCE Threads Compatibility Library
dce.security.rte           3.1.0.0 COMMITTED DCE Security Services
```

Check that the DCE client daemons are running

This example shows how to display the status of the DCE daemons on a node or workstation. This command should also be run on the workstation where the DCE servers are located. Issue the command:

```
lsdce -r
```

Output is similar to:

```
Gathering component state information...
                 Component Summary for Host: c58n03
           Component           Configuration State   Running State
Security client                     Configured           Running
RPC                                 Configured           Running
Directory client                    Configured           Running
```

Good results are indicated by the three components, `Security client`, `RPC`, and `Directory client` having a `Configuration State` of `Configured` and `Running State` of `Running`.

Error results are indicated if one or more of the three components are missing, or they have different values for `Configuration State` or `Running State`.

Check SP-unique principals, accounts, and keytabs

Each service defined in configuration file **/usr/lpp/ssp/config/spsec_defaults** should have an instance defined with a DCE principal, account, and keytab for each host configured to use DCE authentication. Exceptions are those services defined as having one instance for each system partition. These values may be overridden if you have made changes to the **spsec_overrides** file.

This example shows the subset of DCE registry information for PSSP and RSCT services on a control workstation. Here, the DCE hostname is **sp5cw**, and there are two system partitions named **sp5p1** and **sp5p2**. Issue the command:

```
dcecp -c acc cat -s | egrep "∧ssp/|∧rsct/" | egrep "sp5cw|sp5p1|sp5p2"
```

Output is similar to:

```
rsct/sp5cw/rsct
ssp/sp5cw/css
ssp/sp5cw/hardmon
ssp/sp5cw/pmandssp/sp5cw/sp_configd
ssp/sp5cw/spbgroot
ssp/sp5cw/spmgr
ssp/sp5cw/switchtbld
ssp/sp5cw/sysctl
rsct/sp5p1/hats
ssp/sp5p1/sdr
rsct/sp5p2/hats
ssp/sp5p2/sdr
```

Verify that there are keytab objects with the same names by issuing:

```
dcecp -c key cat -s | egrep "∧ssp/|∧rsct/" | egrep "sp5cw|sp5p1|sp5p2"
```

The key files for these services have the same names also. Find them by issuing the following:

```
( cd /spdata/sys1/keyfiles; find ssp -type f; find rsct -type f )
```

Check the existence and membership of SP access groups

Access to SP trusted services is generally granted through group membership, when using DCE for security. The groups are named as shown in the **/usr/lpp/ssp/config/spsec_defaults** file, unless you have specified alternate names in the **spsec_overrides** file. The following example shows how to check that the SP security administration group exists and has the cell administrator as a member. Issue the command:

```
dcecp -c group list $(spgrpname spsec-admin)
```

Output is similar to:

```
/.../cellname/cell_admin
```

Check the DCE CDS information required by SP ACL managers

The DCE ACL managers in the Hardware Monitor and Sysctl servers must have **rpcentry** objects defined in the CDS (Cell Directory Services), so that the ACL editor can locate the servers.

This example shows how to verify that the required entries exists and that their ACLs allow write access by the service principals. A control workstation must have entries for hardmon and sysctl. Other hosts have only sysctl entries. The following shell script is for a control workstation whose DCE hostname is **sp5cw**:

```
for rpcentry in $(dcecp -c dir list /.:/subsys/ssp/sp5cw)
do
dcecp -c acl show $rpcentry -entry | grep ${rpcentry##*/}
done{user ssp/sp5cw/hardmon rwdtc}
{user ssp/sp5cw/sysctl rwdtc}
```

Verify configuration of Kerberos V4

These procedures check the Kerberos V4 configuration files, Kerberos V4 authentication database, Kerberos V4 database administration ACL files, server key (srvtab) files, and Kerberos V4 daemons.

Check the Kerberos V4 configuration files

The configuration files are **/etc/krb.conf** and **/etc/krb.realms**. If you did not supply them, default files were created when you set up your primary authentication server. The default **krb.conf** file contains the realm name, derived from the domain part of the server's hostname. The default realms file is empty, unless hosts in the realm have network interfaces with different domain names.

When you create your own **krb.conf** file, you may have multiple entries per server, specifying different network interfaces. You would typically need to do this if your primary and secondary servers are connected by a different network than that which connects the primary server to the SP nodes. For example:

```
cat /etc/krb.conf
ABC.ORG
ABC.ORG sp5en0.abc.org admin server
ABC.ORG sp5en1.abc.org admin server
ABC.ORG bk5en1.abc.org
```

Check the Kerberos V4 authentication database

Use the **lskp** command to display information about Kerberos V4 principals. Specify the **-p** flag to inspect the principals predefined by Kerberos V4, including one named **default**, that supplies the default expiration date and maximum ticket lifetime values for new user principals:

```
lskp -p
krbtgt.ABC.ORG    tkt-life: 30d   key-vers: 1  expires: 2037-12-31 23:59
K.M               tkt-life: 30d   key-vers: 1  expires: 2037-12-31 23:59
changepw.kerberos tkt-life: 30d   key-vers: 1  expires: 2037-12-31 23:59
default           tkt-life: 30d   key-vers: 1  expires: 2037-12-31 23:59
```

Specify the **-s** flag to inspect the service principals used by SP servers:

```
lskp -s
hardmon.sp5cw  tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
hardmon.sp5en  tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.sp5cw     tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.sp5en     tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.node1en   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.node1sw   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.node1tr   tkt-life: Unlimited key-vers: 1 expires: 2037-12-31 23:59
rcmd.node2en   tkt-life: Unlimited key-vers: 2 expires: 2037-12-31 23:59
rcmd.node2sw   tkt-life: Unlimited key-vers: 2 expires: 2037-12-31 23:59
rcmd.node2tr   tkt-life: Unlimited key-vers: 2 expires: 2037-12-31 23:59
```

Check the Kerberos V4 administration ACL files

There are three ACL files that control access to the Kerberos V4 database through the **kadmin** command and through the Sysctl procedures **mkkp**, **chkp**, **lskp**, and **rmkp**. When you set up Kerberos V4, each ACL file was created with an entry for the administrators you defined. You may want to have different ACL entries, for example, authorizing the use of **lskp** and the **get** subcommand of **kadmin**:

```
cat /var/kerberos/database/admin_acl.getroot.admin
frank.admin
lucy.admin
```

Check the Kerberos V4 srvtab files

The **/etc/krb-srvtab** file should exist on each host, and contain keys that match those in the Kerberos V4 database. On the control workstation, file **/etc/spa-srvtab** should also exist, containing only the key for principal **root.SPbgAdm**. Use the **ksrvutil** command to show these files.

```
ksrvutil list
Version    Principal
1     rcmd.sp5cw@ABC.ORG
1     hardmon.sp5cw@ABC.ORG
1     rcmd.sp5en@ABC.ORG
1     hardmon.sp5en@ABC.ORG
1     root.SPbgAdm@ABC.ORG
```

Check that the Kerberos V4 daemons are running

This example shows how to display the status of the Kerberos V4 authentication daemon. If the status is shown as inoperative, check the daemon's log file for error information. See "Kerberos V4 daemon logs" on page 256. Normally the subsystem is started when **setup_authent** is run during PSSP installation, and automatically on reboot. An example of normal output is:

1. `lssrc -s kerberos`

   Output is similar to:

   ```
   Subsystem         Group            PID      Status
    kerberos                          16256    active
   ```
2. `tail /var/adm/SPlogs/kerberos/kerberos.log`

   Output is similar to:

```
 3-Jun-1999 14:16:58 Kerberos started, PID=20910
10-Jun-1999 13:56:02 Kerberos started, PID=18874
19-Jun-1999 21:04:56 Kerberos started, PID=16256
```

Check enabled security

You must have at least one security method in common enabled on the control workstation and on a node, for remote communication and for SP services using authentication to function properly.

These steps show a properly configured system:

1. Issue the command **splstdata -p**. Output is similar to:

```
List System Partition Information
System Partitions:
------------------
c166s
c166sp1
c166sp2

Syspar: c166s
-----------------------------------------------------------------------
syspar_name     c166s
ip_address      9.114.10.166
install_image   default
syspar_dir      /spdata/sys1/syspar_configs/3nsb0isb/config.4_4_40/layout.
 3syspar-coex/syspar.3.c166s
code_version    PSSP-2.4
haem_cdb_version 939568157,265012085,0

auth_install    dce:k4
auth_root_rcmd  dce:k4
ts_auth_methods dce:compat
auth_methods    k5:k4:std
Syspar: c166sp1
-------------------------------------------------------------------------
syspar_name     c166sp1
ip_address      9.114.10.220
install_image   default
syspar_dir      /spdata/sys1/syspar_configs/3nsb0isb/config.4_4_40/layout.
 3syspar-coex/syspar.1.c166sp1
code_version    PSSP-2.4
haem_cdb_version 939568026,670463251,0

auth_install    k4
auth_root_rcmd  k4
ts_auth_methods compat
auth_methods    k4:std

Syspar: c166sp2
-------------------------------------------------------------------------
syspar_name     c166sp2
ip_address      9.114.10.233
install_image   bos.obj.ssp.433
syspar_dir      /spdata/sys1/syspar_configs/3nsb0isb/config.4_4_40/layout.
 3syspar-coex/syspar.2.c166sp2
code_version    PSSP-3.2
haem_cdb_version 939568100,243454795,0

auth_install    dce
auth_root_rcmd  dce:std
ts_auth_methods dce
auth_methods    k5:std
```

2. Issue the command **export SP_NAME=c166s**.

3. Issue the command **dsh -w c166n01 lsauthent**. Output is similar to:

   ```
   c166n01: Kerberos 5
   c166n01: Kerberos 4
   c166n01: Standard Aix
   ```

4. Issue the command **dsh -w c166n01 lsauthts**. Output is similar to:

   ```
   c166n01: DCE
   c166n01: Compatibility
   ```

5. Issue the command **export SP_NAME=c166sp1**.
6. Issue the command **dsh -w c168n03 lsauthent**. Output is similar to:

   ```
   168n03: Kerberos 4
   c168n03: Standard Aix
   c168n03: kerberos: Couldn't get credentials for the server: Server
    not found in Kerberos database.
   ```

7. Issue the command **dsh -w c168n03 lsauthts**. Output is similar to:

   ```
   168n03: Compatibility
   c168n03: kerberos: Couldn't get credentials for the server: Server
    not found in Kerberos database.
   ```

8. Issue the command **export SP_NAME=c166sp2**.
9. Issue the command **dsh -w c166n03 lsauthent**. Output is similar to:

   ```
   166n03: Kerberos 5
   c166n03: Standard Aix
   ```

10. Issue the command **dsh -w c166n03 lsauthts**. Output is similar to:

    ```
    166n03: DCE
    ```

Check authentication

These sections verify that you have the proper credentials to perform the required functions.

Check DCE authentication and authorization

These steps obtain DCE authentication tickets, display them, use them to perform a trivial Sysctl task, destroy them, and verify their destruction.

1. `dce_login` *operator*

   ```
   Enter Password:
   DCE LOGIN SUCCESSFUL
   ```

2. `klist`

   Expected output is similar to:

   ```
   DCE Identity Information:
             Warning: Identity information is not certified
             Global Principal: /.../abc.org/operator
             Cell:   005741da-98d4-1eae-b63d-02608c2d0159 /.../abc.org
             Principal: 000017bc-2e20-2f5b-8a00-02608c2d0159 operator
             Group:     000403ef-db41-63cf-5802-00000018595c staff
             Local Groups:
                      000403ef-db41-63cf-5802-00000018595c staff
                      0000ac71-8883-21d0-a501-00000018595c hm-monitor
   Identity Info Expires: 1999/09/25:19:11:50
   Account Expires:       never
   Passwd Expires:        1999/10/22:10:41:12
   Kerberos Ticket Information:
   Ticket cache: /opt/dcelocal/var/security/creds/dcecred_67e4a7f7
   Default principal: operator@abc.org
   Server: krbtgt/abc.org@abc.org
   ```

```
      valid 1999/09/24:13:11:50 to 1999/09/25:19:11:50
Server: dce-rgy@abc.org
          valid 1999/09/24:13:11:51 to 1999/09/25:19:11:50
Server: dce-ptgt@abc.org
          valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
Client: dce-ptgt@abc.org      Server: krbtgt/abc.org@abc.org
          valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
Client: dce-ptgt@abc.org      Server: dce-rgy@abc.org
      valid 1999/09/24:13:11:52 to 1999/09/25:01:11:51
```

The `Local Groups` section shows which groups are checked for authorization access. The `Identity Info Expires` section shows when your credentials will expire.

3. `sysctl whoami -v`

   Expected output is similar to:

   ```
   DCE: /.../abc.org/operator
   K4:  operator@ABC.ORG
   AIX: joseph
   ```
4. `kdestroy`
5. `klist`

   Expected output is similar to:

   ```
   No DCE identity available: No currently established
      network identity for this context exists (dce / sec)

   Kerberos Ticket Information:

   klist: No credentials cache file found (dce / krb)
   (ticket cache /opt/dcelocal/var/security/creds/dcecred_67e4a7f7
   ```

Check DCE authentication for an SP service principal

These steps are used to check whether an SP service principal can login to DCE. If the login fails, it may indicate DCE server problems or key file problems for the SP service principal used. The following example shows results similar to what the command will return when there are no problems. Be sure to destroy the credentials obtained, as illustrated in the example.

1. `dsrvtgt`

   Output is similar to:

   ```
    dsrvtgt ssp/css
    FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600
   ```
2. `kdestroy -c FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600`

Check Kerberos V4 authentication

These steps obtain Kerberos V4 authentication tickets, use them to perform a trivial task using Sysctl, display them, destroy them, and verify their destruction.

1. `k4init operator`

   Output is similar to:

   ```
   Kerberos Initialization for "operator"
   Password:
   ```
2. `k4list`

   Output is similar to:

```
Ticket file:     /tmp/tkt1890
Principal:       operator@ABC.ORG
  Issued          Expires          Principal
Sep 22 13:55:32  Oct 22 13:55:32   krbtgt.ABC.ORG@ABC.ORG
```

3. `sysctl whoami -v`

   Output is similar to:

   ```
   CE:
   K4:  operator@ABC.ORG
   AIX: joseph
   ```

4. `k4destroy`

   Output is similar to:

   ```
   Tickets destroyed.
   ```

5. `k4list`

   Output is similar to:

   ```
   Ticket file:     /tmp/tkt1890
   k4list: 2504-076 Kerberos ticket file was not found
   ```

Check authorization

These sections verify that you have the proper authorization files to perform the required functions.

Check DCE ACL authorization

These steps check on authorization to perform a trivial Hardware Monitor task based on a DCE ACL granting access to a group.

1. `dce_login operator`

   Output is similar to:

   ```
   Enter Password:
   DCE LOGIN SUCCESSFUL
   ```

2. `klist`

   Output is similar to:

   ```
   DCE Identity Information:
      Warning: Identity information is not certified
      Global Principal: /.../abc.org/operator
      Cell:      005741da-98d4-1eae-b63d-02608c2d0159 /.../abc.org
      Principal: 000017bc-2e20-2f5b-8a00-02608c2d0159 operator
      Group:     000403ef-db41-63cf-5802-00000018595c staff
      Local Groups:
                 000403ef-db41-63cf-5802-00000018595c staff
                 0000ac71-8883-21d0-a501-00000018595c hm-monitor
   Identity Info Expires: 1999/09/25:19:11:50
   Account Expires:       never
   Passwd Expires:        1999/10/22:10:41:12
   Kerberos Ticket Information:
   Ticket cache: /opt/dcelocal/var/security/creds/dcecred_67e4a7f7
   Default principal: operator@abc.org
   Server: krbtgt/abc.org@abc.org
             valid 1999/09/24:13:11:50 to 1999/09/25:19:11:50
   Server: dce-rgy@abc.org
             valid 1999/09/24:13:11:51 to 1999/09/25:19:11:50
   Server: dce-ptgt@abc.org
             valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
   Client: dce-ptgt@abc.org      Server: krbtgt/abc.org@abc.org
   ```

```
          valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
Client: dce-ptgt@abc.org      Server: dce-rgy@abc.org
          valid 1999/09/24:13:11:52 to 1999/09/25:01:11:51
```

The `Local Groups` section shows which groups are checked for authorization access. The `Identity Info Expires` section shows when your credentials will expire.

3. `dcecp -c acl show /.:/subsys/ssp/sp5cw/hardmon/system`

   Output is similar to:

   ```
   {group spsec-admin -c-----}
   {group hm-control --vsmu-}
   {group hm-control-services --vsmu-}
   {group hm-monitor ----m--}
   {group hm-monitor-services ----m--}
   ```
4. `hmdceobj -q`

   Output is similar to:

   ```
   system
    hardmon
   ```
5. `kdestroy`

Check DCE group authorization

These steps allow you to check whether an id is located in a DCE group. Normally, you would add a group to a DCE ACL file and add DCE principals to the groups for access to SP services.

1. To check groups for your logged in DCE principal, issue this command:

   ```
   klist
   ```

   Output is similar to:

   ```
   DCE Identity Information:
      Warning: Identity information is not certified
      Global Principal: /.../abc.org/operator
      Cell:      005741da-98d4-1eae-b63d-02608c2d0159 /.../abc.org
      Principal: 000017bc-2e20-2f5b-8a00-02608c2d0159 operator
      Group:     000403ef-db41-63cf-5802-00000018595c staff
      Local Groups:
                 000403ef-db41-63cf-5802-00000018595c staff
                 0000ac71-8883-21d0-a501-00000018595c hm-monitor
   Identity Info Expires: 1999/09/25:19:11:50
   Account Expires:       never
   Passwd Expires:        1999/10/22:10:41:12
   Kerberos Ticket Information:
   Ticket cache: /opt/dcelocal/var/security/creds/dcecred_67e4a7f7
   Default principal: operator@abc.org
   Server: krbtgt/abc.org@abc.org
             valid 1999/09/24:13:11:50 to 1999/09/25:19:11:50
   Server: dce-rgy@abc.org
             valid 1999/09/24:13:11:51 to 1999/09/25:19:11:50
   Server: dce-ptgt@abc.org
             valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
   Client: dce-ptgt@abc.org      Server: krbtgt/abc.org@abc.org
             valid 1999/09/24:13:11:51 to 1999/09/25:01:11:51
   Client: dce-ptgt@abc.org      Server: dce-rgy@abc.org
             valid 1999/09/24:13:11:52 to 1999/09/25:01:11:51
   ```
2. To find out what groups are defined in DCE, issue this command:

   ```
   dcecp -c group cat
   ```

Output is similar to:

```
/.../c166dcecell/sdr-system-class-admin-services
/.../c166dcecell/sdr-system-class-write
/.../c166dcecell/sdr-system-class-write-services
/.../c166dcecell/sdr-write
/.../c166dcecell/sdr-write-services
/.../c166dcecell/switchtbld-clean
/.../c166dcecell/switchtbld-load
/.../c166dcecell/switchtbld-status
/.../c166dcecell/sysctl-cwsroot
/.../c166dcecell/sysctl-logmgt
/.../c166dcecell/sysctl-logmgt-services
/.../c166dcecell/sysctl-master
/.../c166dcecell/sysctl-mmcmd
/.../c166dcecell/sysctl-mmcmd-services
/.../c166dcecell/sysctl-pman
/.../c166dcecell/sysctl-vsd
/.../c166dcecell/sysctl-vsd-services
```

3. To find out members of a group, issue this command:

   ```
   dcecp -c group list group_name
   ```

   Output is similar to:

   ```
   /.../c166dcecell/joseph
   /.../c166dcecell/mary
   ```

Check Kerberos V4 ACL authorization

These steps check on authorization to perform trivial tasks, one based on Sysctl ACL permissions, the other based on hardmon ACL permissions.

1. `k4init operator`

   Output is similar to:

   ```
   Kerberos Initialization for "operator"
   Password:
   ```

2. `cat /etc/sysctl.acl`

   Output is similar to:

   ```
   acl
   _PRINCIPAL root.admin@ABC.ORG
   _PRINCIPAL operator@ABC.ORG
   ```

3. `sysctl puts {This is a test message}`

   Output is similar to:

   ```
   This is a test message
   ```

4. `cat /spdata/sys1/spmon/hmacls`

   Output is similar to:

   ```
   sp5cw.abc.org root.admin a
   sp5cw.abc.org root.SPbgAdm a
   1 root.admin vsm
   1 root.SPbgAdm vsm
   1 hardmon.sp5cw vsm
   1 operator m
   ```

5. `spmon -d`

   Output is similar to:

```
1.  Checking server process
    Process 71054 has accumulated 25 minutes and 14 seconds.
    Check successful
2.  Opening connection to server
    Connection opened
    Check successful
3.  Querying frame(s)
    1 frame
    Check successful
4.  Checking frames
    This step was skipped because the -G flag was omitted.
5.  Checking nodes
---------------------------------- Frame 1 ----------------------------
                     Host     Switch   Key    Env   Front Panel LCD/LED
Slot Node Type Power Responds Responds Switch Error LCD/LED     Flashes
---- ---- ---- ----- -------- -------- ------ ----- ----------- -------
  1    1 wide  on    yes      no      N/A    no   LCDs are blank  no
.
.
.
```

Error symptoms, responses, and recoveries

A user of SP trusted services interacts with SP security services in the following ways:

1. The user identifies himself to the underlying security mechanisms, DCE or Kerberos V4. This process obtains a ticket-granting ticket for the client principal based on either the user's password or a service's private key. The former is achieved interactively using the **dce_login** and **k4init** commands. Background processes running as **root** and invoking shell scripts use the **dsrvtgt** command for DCE and the **rcmdtgt** or **ksrvtgt** commands for Kerberos V4, to get tickets as service principals.
2. The user invokes a command that directly or indirectly invokes an SP trusted service. The client command uses security services to obtain a DCE authentication token to pass to a server. Clients that use Kerberos V4 in compatibility mode similarly obtain Kerberos V4 authentication tokens.
3. The user issues the **kdestroy** command for DCE or the **k4destroy** command for Kerberos V4 to remove the authentication tickets following their use.

Errors detected by SP Security Services are reported in three different ways:

1. Errors that occur while running security services commands (for example, **chauthpts**, **spgrpname**), are reported by error messages written to stderr. Most of the messages have the **2502**, **2503**, and **2504** SP Security Services prefix. Some of the messages have the **0016** installation and configuration prefix, or the **0025** SDR prefix.
2. Other errors are reported by either client or server daemon programs that comprise the SP trusted services. The security services runtime support used by those programs reports error conditions it detects, as well as errors detected by the DCE library subroutines it invokes.
3. The SP trusted services incorporate error message text describing those errors in their own output to stderr, or to daemon log files, or the AIX error log, as appropriate to the runtime environment.

Many of the errors, such as those caused by the failure of a user to login to DCE or Kerberos V4, or a user with expired credentials, are straightforward, and the recovery actions are clear. Other errors are reported by DCE-supplied error text, and the explanations and recovery actions for these error are located in *IBM DCE for AIX, Version 3.1: Problem Determination Guide*.

Other less common error conditions include those in Table 47.

Table 47. SP Security Services symptoms and recovery actions

Symptom	Recovery
The service principal for a target server is unknown. • DCE messages similar to ″registry object unknown″ • Kerberos V4 messages similar to ″Kerberos principal unknown″	See “Action 1 - Check the service principal”.
Authentication server is not available. • Users unable to login to DCE or Kerberos V4. • Servers do not initialize properly when using DCE. • Servers hang using DCE. • When DCE authentication fails due to an unavailable server, error messages similar to ″registry server unavailable″. • When a client cannot access the Kerberos V4 authentication server, error messages similar to ″Kerberos error: retry count exceeded″.	See “Action 2 - Check the authentication server” on page 273.
A user cannot login to Kerberos V4. The login fails with a message similar to: • ″Bad Kerberos name format’ • ″Kerberos principal unknown″ • ″Incorrect Kerberos password″	See “Action 3 - Propagate Kerberos V4 database” on page 274.
Authorization to use a trusted service is denied. Messages vary. A message similar to ″Cannot obtain service ticket for ...″ may appear.	See “Action 4 - Check for authorization or authentication problems” on page 275.
A background client cannot login as a service principal. Messages vary. A message similar to ″Incorrect Kerberos Password″ may appear.	See “Action 5 - Correct key files” on page 276.
PSSP configuration error: error message from **install_cw**, the SDR, and System Monitor indicate that all user requests are unauthorized. Corrupted or destroyed authentication method settings. Messages 2502–604, 2515–031, 0026–894, 0031–737, and 2539–498 or other similar messages.	See “Action 7 - Run the chauthts command” on page 286.
Problems with Sysctl using multiple authentication methods.	See “Action 8 - Diagnosing Sysctl security problems” on page 287.
Incorrect information in the SDR, after migrating to PSSP 3.2 from a level of PSSP earlier than PSSP 3.1.	See “Action 6 - Reset authentication values” on page 286.

Actions

Action 1 - Check the service principal

If a trusted services client program attempts to obtain a service ticket in order to pass an authentication token to a server, the service principal used by the server program as its identity for purposes of authentication must exist in the authentication database. When using DCE, this is the security registry. When using Kerberos V4, this is either the PSSP Kerberos V4 authentication database or an AFS authentication database.

This error is reported by DCE as ″registry object unknown″ error returned by the subroutine **gss_accept_sec_context()**. When the error occurs using Kerberos V4, the error is reported as ″Kerberos principal unknown″.

Use the diagnostic procedures in “Verify DCE Security Services configuration” on page 261 to verify that the proper SP Security Services configuration steps were performed. If necessary, refer to diagnostic procedures for SP Installation and Configuration. If those procedures indicate that the required service principal does not exist, identify and perform the required configuration task to create the entries in the authentication database. This may require that a DCE cell administrator run the **config_spsec** command, and the **root** user perform other configuration tasks as well. For a missing Kerberos V4 service principal, the Kerberos V4 administrator (usually **root.admin**) must run **setup_server** to customize the target server node.

Another possible cause of this error is incomplete propagation of an **spsec_overrides** file that contains an override name for one or more services to RS/6000 workstations. PSSP software automatically propagates the file from the control workstation to SP nodes, when they are installed. However, if you want to use security services client programs on another workstation, you must make sure the file is copied from the control workstation before running **config_spsec** on the workstation.

Action 2 - Check the authentication server

This error condition can manifest itself in several different ways and have many different causes. Users may be unable to login to DCE or Kerberos V4. Servers may not initialize properly when using DCE authentication. When DCE authentication cannot complete due to an unavailable server, the most common error message reported by DCE is ″registry server unavailable″.

When a client cannot get to the Kerberos V4 authentication server, the most common error message is ″Kerberos error: retry count exceeded″.

1. Follow the diagnostic procedures described previously to check server status, including the log files. See “Check that the Kerberos V4 daemons are running” on page 264.
2. Check the state of the server daemons.

   For DCE, issue the command **lsdce -r** on both application client and application server hosts, as well as on the DCE server hosts. If the servers are down, continue to the next step.
3. Check the DCE log files for server termination errors.

   If servers terminated, a full file system is a possible cause. If the server terminated, determine the reason from the log file, take corrective action, and restart the server.
4. Make sure that you have an appropriate number of backup (replica) servers for your system.
5. Also make sure that your **pe_site** file is kept current on all hosts in your DCE cell.

Errors may result from the inability to communicate with the authentication servers. If this is the case, check for network problems or interface and routing problems.

For Kerberos V4, the primary server system should have the **kerberos** and **kadmind** daemons running under the AIX SRC. A secondary server system should have the **kerberos** and **kpropd** daemons running (from **init**).

Check the Kerberos V4 daemon log - each Kerberos daemon program records errors and some status in a log file in the **/var/adm/SPlogs/kerberos** directory. Check these files if you suspect that one or more of the daemons has terminated. This is an example of the log file created by the **kerberos** daemon:

```
03-May-1999 14:06:45 Kerberos started, PID=22408
07-May-1999 07:53:07 Kerberos started, PID=8642
07-May-1999 07:53:07 kerberos: 2503-604 Cannot verify master key.
07-May-1999 07:55:11 Kerberos started, PID=8648
```

This is an example of the log file created by the **kadmind** daemon:

```
21-May-1999 16:40:12 The Kerberos administration server has started,
                                            PID=10188
21-May-1999 16:41:40 kadmind: 2503-101 Error: 2504-318 Could not verify
                                                         master key
21-May-1999 16:41:40 Shutting down the Kerberos administration server
```

This is an example of the log file created by the **kpropd** daemon:

```
13-Oct-1999 09:27:32 Established socket
13-Oct-1999 10:19:09 Connection from cwksta.xyz.abc.com, 129.49.100.41
13-Oct-1999 10:19:09 kpropd: Connection from rcmd.cwksta@XYZ.ABC.COM
13-Oct-1999 10:19:09 File received.
13-Oct-1999 10:19:09 Temp file renamed to /var/kerberos/database/slavedb
14-Oct-1999 07:36:41 Connection from cwksta.xyz.abc.com, 129.49.100.41
14-Oct-1999 07:36:41 kpropd: Connection from rcmd.cwksta@XYZ.ABC.COM
14-Oct-1999 07:36:41 File received.
14-Oct-1999 07:36:41 Temp file renamed to /var/kerberos/database/slavedb
```

Action 3 - Propagate Kerberos V4 database

Logging into Kerberos V4, these are the most common error messages encountered:

1. Bad Kerberos name format

   You probably entered **name.admin** in response to the prompt ″Kerberos name:″. You are not allowed to specify an instance in addition to the name. If you want to enter the name and instance together, enter them as the command-line argument when you invoke **k4init**. To request **k4init** to prompt you separately for the instance, invoke **k4init** with the **-i** flag.

2. Kerberos principal unknown

   You did not enter a principal name that is defined in the database. Perhaps you misspelled it, or the administrator did so when entering it into the database.

3. Incorrect Kerberos password

   You entered the wrong password, or your password was recently changed and has not yet been propagated to the secondary authentication server that you are using. If you incorrectly entered your password, just try again. Otherwise, if you suspect an out-of-date database, contact the administrator of your authentication service.

   This error can also occur when using a secondary authentication server, if the primary database has not been propagated since your principal was added. Check with the administrator responsible for maintaining the authentication service to determine if this is the case. The recovery action in this situation is to force the propagation of database changes without waiting for the normal cron process.

To correct the preceding problems 2 on page 274 and 3 on page 274, force the propagation of Kerberos V4 database:

1. Issue the **k4init** command, specifying as your principal name any user principal listed in the **root** user’s **.klogin** file on the primary server. The administrative principal name that was used to set up authentication on the primary server can be used. Any other user principal that the administrator has subsequently added to the file can be used.
2. Issue the following command to remotely perform the database propagation from the primary server to all secondary servers:

   ```
   dsh -w primary /usr/kerberos/etc/push-kprop
   ```

   where *primary* is the hostname of the primary Kerberos V4 authentication server.
3. Successful propagation is reported by a message for each secondary server hostname. If unsuccessful, review the **kpropd.log** file. See “Log files” on page 256.

Action 4 - Check for authorization or authentication problems

This is by far the most common problem encountered when using security services. Most security errors ultimately result in a failure to authorize a task that a client requested of a server. First, determine whether the problem is an authorization problem or an authentication problem. If the error resulted from an authentication problem, you should have at least one error message from the trusted service describing the type of failure, in addition to the authorization failure. When the trusted service uses DCE and Kerberos V4 for authentication (Sysctl and Hardware Monitor), the client could report authentication error messages regarding one authentication method or the other, or both.

A possible cause of authentication failures is the lack of a common authentication method on the client and server hosts. Client programs on a node in a system partition that is not configured to use DCE cannot access trusted services on nodes in another partition that is configured to use only DCE.

Similarly, a client host must use Kerberos V4 authentication, in order to run a Sysctl procedure that requires authentication on a target server that uses only Kerberos V4 for authentication.

Issue the commands from “SP trusted services authentication errors” on page 259, on the client and server hosts to look for error messages. If the authorization failure was accompanied by no message regarding authentication failure, consult the security administrator to determine whether the requisite permission had been granted. For most trusted services, the requisite permission requires membership in a DCE group, or an entry in a DCE ACL. For Sysctl and Hardware Monitor failures using Kerberos V4, the requisite permission usually requires an entry in an ACL file.

When using DCE authentication, an authorization failure may mean that the user failed to re-login to DCE after the administrator added his principal to one of the SP trusted services access groups.

When you use any of the authenticated client/server applications to administer or control the SP system, the error messages you receive on authentication failures will vary according to the application. For example, if you are using the command-line interface, you might see error messages such as the following, indicating that the System Monitor is unable to obtain Kerberos V4 credentials:

```
0026-706 Cannot obtain service ticket for hardmon.cwksta
Kerberos error code is 76, Kerberos error message is:
2504-076 Kerberos ticket file was not found.
spmon: 0026-001 Opening session failed.
```

The message states that you have no tickets, expired or unexpired, or your **KRBTKFILE** environment variable specifies a nonexistent file.

The following message states that you have tickets that have expired in the ticket cache file (specified by your **KRBTKFILE** environment variable or defaulted to **/tmp/tktuid**.

```
0026-706 Cannot obtain service ticket for hardmon.cwksta
Kerberos error code is 76, Kerberos error message is:
2504-032 Kerberos ticket expired.
spmon: 0026-001 Opening session failed.
```

If application error messages indicate probable authentication failure, use the **klist** or **k4list** command to check your authentication status. These commands always displays the current active ticket cache file.

Action 5 - Correct key files

SP Security Services include commands that can be used by a background process to obtain authenticated tickets, when no user is logged in to the client AIX system. SP installation and system management scripts use the **dsrvtgt**, **rcmdtgt**, or **ksrvtgt** command to login.

For DCE, the **dsrvtgt** command logs in as a service principal or a user, using a DCE key file. For a service principal that cannot be logged into, check the spelling of the service name with the default name contained in the **/usr/lpp/ssp/config/spsec_defaults** file. Use “Check SP-unique principals, accounts, and keytabs” on page 262 to verify that the required DCE registry information exists for the principal and that the key file exists.

If the configuration information and key file exist, check on the operation of the **spnkeymand** daemon that keeps the server keys refreshed. Check the password expiration policy for the spsec-services organization (or override organization name, if applicable). If a node or workstation is out of service for an extended period of time, it is possible for the keys to have not been refreshed before their expiration. See “Chapter 19. Diagnosing Per Node Key Management (PNKM) problems” on page 289.

If this is the case, you must use the recovery procedures provided by SP Installation and Configuration, using the **rm_spsec** and **config_spsec**. Distribute new key files to a node by customizing it or by running **create_keyfiles** on the node. If the error requires re-creation of key files for Topology Services, customize the nodes in the partition or run **get_keyfiles** on each node.

To recover from a single missing or corrupted DCE key file, see “Recover from a missing or corrupt DCE key file” on page 280.

Errors logging into Kerberos V4 could be reported as:

```
Incorrect Kerberos password
```

If this error occurs during installation, or when performing administrative tasks requiring remote invocation on SP nodes, it can indicate one of several error conditions:

- The process is not running as **root**. It cannot read the server key file, **/etc/krb-srvtab**, on the client system. You must, if authorized, login as **root** and retry the failing task.
- The server key file is out-of-date with respect to the authentication database.
- The server key file does not exist.

Compare server key versions. An administrator running as **root** can compare the versions of the server keys in the server key file and the database, when AFS is not being used. Use the **ksrvutil** command to show the version numbers of server keys in **/etc/krb-srvtab**. See “Check the Kerberos V4 srvtab files” on page 264. The actual value of the keys cannot be compared directly, because the copies in the database are encrypted in the master key. Compare the key versions.

When the key file, **/etc/krb-srvtab**, does not exist on the server system, or has the wrong key version, you must re-create the file. When the control workstation is an SP authentication server, customizing an SP node will automatically create a new server key file for the node. If customizing the node for this purpose is too disruptive, or if the system whose key file must be replaced is not an SP node, follow the following procedures.

Re-create Kerberos V4 server key files: You can use authentication administration commands to re-create an erroneous or missing server key file. Each system with an SP authentication service installed has its own unique server key file, containing the encrypted keys for the service instances that are available on the system.

On the control workstation and other IBM RS/6000 workstations that have client services installed and initialized, the file contains entries for services **rcmd** and **hardmon**. Separate instances of these principals are defined for each network interface on the system, where the instance-name is the short form of the network name.

There is also a special principal, **root.SPbgAdm**, that is used by **root** processes that run in the background and need access to services that use authentication. For example, on a client system with token ring and FDDI interfaces named **wksta5t.xyz.abc.com** and **wksta5f.def.abc.com**, the following principals’ keys are kept in the server key file:

- hardmon.wksta5f
- hardmon.wksta5t
- rcmd.wksta5f
- rcmd.wksta5t
- root.SPbgAdm

On the SP nodes, the hardmon entries are not included. The hardmon files are created by the **setup_server** command.

The server key files on these systems are created by the **setup_authent** command. During installation, they are kept in the **/spdata/sys1/k4srvtabs** directory on the control workstation, with file names of the form *hostname-new-srvtab*, where *hostname* is the short form of the hostname for each node.

When the control workstation is an SP authentication server, these files are retained there only until they are copied to the node during network boot. If the node boots from another node, these files are retained until they are copied to the node's boot/install server. A new server key file is generated any time the node is set up for a network boot.

When the control workstation is not configured as an authentication server, or when AFS authentication is used, the server key files for the SP nodes are **not** removed from the **/spdata/sys1/k4srvtabs** directory on the control workstation, once they are created.

If they are deleted or corrupted, or if you choose to change keys for any reason, follow the rest of the procedures to create new key files. In these procedures, *instance1* ... *instancen* are the network names (short form) of all the system's interfaces, and *hostname* is the short form of the system's hostname.

Replace a Kerberos V4 authentication server's key file: To re-create the file for a workstation (control workstation or other) that is configured as an authentication server, the **root** user follows these steps:

1. Create new key files in the **/tmp** directory for each instance:

   ```
   cd /tmp
   /usr/kerberos/etc/ext_srvtab -n instance1...instancen
   ```

2. Combine the key files into a single file:

   ```
   /bin/cat instance1-new-srvtab...instancen-new-srvtab\
                   /etc/spa-srvtab >/etc/krb-srvtab
   /bin/rm -f instance1-new-srvtab...instancen-new-srvtab
   ```

3. Make sure that the key file is readable by **root** only:

   ```
   /bin/chmod 400 /etc/krb-srvtab
   ```

Replace a client workstation's Kerberos V4 key file: When a workstation is not an authentication server, the **root** user can use the remote commands to perform the same function on a server system, and move the file to the local system. The principal name specified on the **k4init** command must be in the **root** user's **.klogin** file on the server:

1. Get a ticket-granting ticket to allow use of **rsh** and **rcp**.

   ```
   k4init principal
   ```

2. Create the new key files in **/tmp** on the server.

   ```
   rsh server cd /tmp\;
   /usr/kerberos/etc/ext_srvtab -n instance1...instancen SPbgAdm
   ```

3. Copy the files we created to the local **/tmp** directory.

   ```
   cd /tmp
   rcp server:/tmp/instance1-new-srvtab...instancen-new-srvtab\
          SPbgAdm-new-srvtab
   ```

4. Delete the files on the server.

```
rsh server /bin/rm -f /tmp/SPbgAdm-new-srvtab \
         /tmp/instance1-new-srvtab...instancen-new-srvtab
```

5. Combine the local files into a single file.

   ```
   /bin/cat instance1-new-srvtab...instancen-new-srvtab SPbgAdm-new-srvtab\
                >/etc/krb-srvtab
   /bin/rm -f instance1-new-srvtab...instancen-new-srvtab SPbgAdm-new-srvtab
   ```

6. Make sure that the key file is readable only by **root**.

   ```
   /bin/chmod 400 /etc/krb-srvtab
   ```

Replace an SP node's Kerberos V4 key file: The most straightforward way to replace a node's key file is to customize the node, using the **spbootins** command or **SMIT**. If you prefer, you can use procedures similar to the preceding example. The **root** user on the node can use the procedure if the control workstation is a Kerberos V4 authentication server. Specify the control workstation hostname as the server.

When the control workstation is not an authentication server, the server key file must be re-created at the server and then moved securely into the **/spdata/sys1/k4srvtabs** directory on the control workstation. For this, the **root** user should be logged into the control workstation.

When the authentication server is another workstation running the SP server or another MIT Kerberos Version 4 implementation, use the following procedure:

1. Get a ticket-granting ticket to allow use of **rsh** and **rcp**.

   ```
   k4init principal
   ```

2. Create the new key files in **/tmp** on the server.

   ```
   rsh server cd /tmp\;
   /usr/kerberos/etc/ext_srvtab -n instance1...instancen SPbgAdm
   ```

3. Copy the files we created to the local **/tmp** directory.

   ```
   cd /tmp
   rcp server:/tmp/instance1-new-srvtab...instancen-new-srvtab\
                SPbgAdm-new-srvtab
   ```

4. Delete the files on the server.

   ```
   rsh server /bin/rm -f /tmp/SPbgAdm-new-srvtab \
                 /tmp/instance1-new-srvtab...instancen-new-srvtab
   ```

5. Combine the local files into a single file.

   ```
   bin/cat instance1-new-srvtab...instancen-new-srvtab\
    SPbgAdm-new-srvtab >/etc/krb-srvtab
   /bin/rm -f instance1-new-srvtab...instancen-new-srvtab SPbgAdm-new-srvtab
   ```

6. Make sure that the key file is readable only by **root**.

   ```
   /bin/chmod 400 /etc/krb-srvtab
   ```

The new key file can then be installed on the node by either:

- Customizing the node, and then booting the node.
- Issue the **rcp** command to copy the new srvtab file from the control workstation **/spdata/sys1/k4srvtabs** directory to the **/etc/krb-srvtab** directory on the node.

Replace a server key file using AFS servers: When an AFS authentication server is being used, follow this procedure and be sure you are logged in to the control workstation as **root**.

1. Change the service password to a new, known value, but not one that is obvious.
2. Repeat this step for each instance (each short network name).

   ```
   kas setpassword -name rcmd.instance -new_password new-password
    -kvno 1 -admin_username afs-admin-name
    -password_for_admin afs-admin-passwd
   ```
3. Use the same principals and passwords to create a **srvtab** file.

   ```
   /usr/kerberos/bin/ksrvutil -afs
    -f /spdata/sys1/k4srvtabs/hostname-new-srvtab add
   ```
4. Use **ksrvutil**, an interactive program whose sequence of prompts and messages is as follows:

   ```
   Name: rcmd
   Instance: instance
   Realm: <Enter>
   Version number: 0
   New principal: rcmd.instance@realm Version 0
   Is this correct? y
   Password:
   Key successfully added.
   Would you like to add another key?
   (reply y until all instances have been entered)
   ```
5. Make sure that the key file is readable only by **root**.

   ```
   /bin/chmod 400 /spdata/sys1/k4srvtabs/hostname-new-srvtab
   ```

Recover from a missing or corrupt DCE key file: SP trusted services key files are under **/spdata/sys1/keyfiles** (main path). Under normal conditions, this directory exists and contains the following directories:

```
 drwxr-xr-x   3 root     system       512 Oct 26 13:04 LoadL/
 drwxr-xr-x   3 root     system       512 Oct 26 13:04 mmfs/
 drwxr-xr-x   3 root     system       512 Oct 26 13:04 ppe/
 drwxr-xr-x   4 root     system       512 Oct 26 13:06 rsct/
 drwxr-xr-x   3 root     system       512 Oct 26 13:04 ssp/
```

Under each of these directories is another directory. The name of this directory is the DCE hostname of the control workstation or node. For example, for node sp3en5:

```
 drwxr-xr-x   2 root     system       512 Nov 17 12:29 sp3en5
```

The DCE hostname directory contains the DCE key files for a given service or services:

```
 -rw-------   1 root     system       298 Oct 27 23:43 css
 -rw-------   1 root     system       382 Oct 27 23:43 pmand
 -rw-------   1 root     system       407 Oct 27 23:44 sp_configd
 -rw-------   1 root     system       397 Oct 27 23:44 spbgroot
 -rw-------   1 root     system       382 Oct 27 23:43 spmgr
 -rw-------   1 root     system       407 Oct 27 23:43 switchtbld
 -rw-------   1 root     system       152 Oct 26 12:22 sysctl
```

The key files on the control workstation and on the nodes differ in two ways:

1. On the control workstation, the **/ssp** directory contains at least two subdirectories. One subdirectory is named after the DCE hostname directory. Another subdirectory is named after the default partition of the system. If more than two subdirectories exist, their names correspond to the names of the remaining PSSP system partitions. This example relates to a system with three partitions, named **secsys1**, **secsys2**, and **secsys3**.

   ```
    ls -l
        total 32
        drwxr-xr-x   2 root     system       512 Nov 17 12:29 sp3en0/
        drwxr-xr-x   2 root     system       512 Oct 26 12:21 secsys1/
        drwxr-xr-x   2 root     system       512 Oct 26 12:22 secsys2/
        drwxr-xr-x   2 root     system       512 Oct 26 12:22 secsys3/
   ```

   The **ssp/***partition_name* subdirectories contain a DCE key file for each **sdrd**. The commands:

   ```
   cd secsys1
   ls -l
   ```

   produce output similar to:

   ```
   total 8
   -rw-------   1 root     system       112 Oct 26 12:22 sdr
   ```

2. The **ssp/***dce_hostname* directory has these files:

   ```
   ls -l
        total 64
        -rw-------   1 root     system       298 Oct 27 23:43 css
        -rw-------   1 root     system       154 Oct 26 12:21 hardmon
        -rw-------   1 root     system       382 Oct 27 23:43 pmand
        -rw-------   1 root     system       407 Oct 27 23:44 sp_configd
        -rw-------   1 root     system       397 Oct 27 23:44 spbgroot
        -rw-------   1 root     system       382 Oct 27 23:43 spmgr
        -rw-------   1 root     system       407 Oct 27 23:43 switchtbld
        -rw-------   1 root     system       152 Oct 26 12:22 sysctl
   ```

Note: Even if the SP system is not configured for LoadLeveler, PE, GPFS, or an SP switch, PSSP DCE key files are created for these services. This includes a css key file for use with PSSP switch services and commands. This is not a security exposure because the key files can be read and used only by a user with **root** authority.

In the event that a PSSP DCE key file does not exist in its expected location, determine if the key file was removed (deleted) from the host, or was not created during the installation and configuration process.

On the host where the DCE PSSP key file is missing, run the **create_keyfiles** command in **verbose** mode. If the command is run on the control workstation, ensure that the **-c** flag is specified on the **create_keyfiles** command. If the key file did exist, but was deleted, an attempt to create a new key file will fail, because the local DCE service will already contain a keytab object for the missing key file.

In this sequence, **sysctl** is the missing key file.

1. The key file catalog contains an entry for **sysctl**. Issue this command:

   ```
   dcecp -c keytab cat
   ```

   Output is similar to:

```
/.../sp_cell/hosts/sp3en0/config/keytab/self
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/GSmonitor
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Kbdd
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Master
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Negotiator
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Schedd
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Startd
/.../sp_cell/hosts/sp3en0/config/keytab/LoadL/sp3en0/Starter
/.../sp_cell/hosts/sp3en0/config/keytab/mmfs/sp3en0/mmfsd
/.../sp_cell/hosts/sp3en0/config/keytab/ppe/sp3en0/dpcl
/.../sp_cell/hosts/sp3en0/config/keytab/ppe/sp3en0/pmdv3
/.../sp_cell/hosts/sp3en0/config/keytab/rsct/sp3en0/rsct
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/css
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/pmand
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/sp_configd
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/spbgroot
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/switchtbld
/.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/sysctl
```

2. The physical **sysctl** key file is missing. Issue this command:

   ```
   ls -l
   ```

   Output is similar to:

   ```
        total 48
   -rw-------   1 root     system       223 Oct 27 23:38 css
   -rw-------   1 root     system       229 Oct 27 23:38 pmand
   -rw-------   1 root     system       244 Oct 27 23:38 sp_configd
   -rw-------   1 root     system       238 Oct 27 23:38 spbgroot
   -rw-------   1 root     system       244 Oct 27 23:38 switchtbld
   ```

3. Attempt to recreate the missing key file. Issue this command:

   ```
   create_keyfiles -v
   ```

   At the end of this output, you will see the expected error: `Cannot create object; already exists`. Output is similar to:

   ```
   Running "check_prereqs" subroutine ...
   Checking state of DCE ...
   Running "parse_defaults" subroutine ...
   Parsing spsec_defaults file ...
   Running "parse_overrides" subroutine ...
   Running "create_keys" subroutine ...
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/GSmonitor already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Kbdd already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Master already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Negotiator already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Schedd already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Startd already exists.
   Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Starter already exists.
   Keyfile /spdata/sys1/keyfiles/mmfs/sp3en0/mmfsd already exists.
   Keyfile /spdata/sys1/keyfiles/ppe/sp3en0/dpcl already exists.
   Keyfile /spdata/sys1/keyfiles/ppe/sp3en0/pmdv3 already exists.
   Keyfile /spdata/sys1/keyfiles/rsct/sp3en0/rsct already exists.
   Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/css already exists.
   Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/pmand already exists.
   Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/sp_configd already exists.
   Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/spbgroot already exists.
   Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/switchtbld already exists.
   Running "do_keytab_work" subroutine ...
   Creating keytab object, "ssp/sp3en0/sysctl" and randomizing keys.
   ***********************************************************
   /usr/lpp/ssp/bin/create_keyfiles: 0016-511 The dcecp command
   ```

```
                below returned non-zero return code.
/opt/dcelocal/bin/dcecp -c keytab create ssp/sp3en0/sysctl\
           -storage /spdata/sys1/keyfiles/ssp/sp3en0/sysctl -data
           { ssp/sp3en0/sysctl plain 1 "svc_pwd_was_here" }
Command output is:
Error: msgID=0x113DB0CE  Cannot create object; already exists
```

To re-create the missing key file with a random password, perform these steps:

1. As **root**, (which has access to the DCE self-host principal credentials), on the host with the missing key, remove the sysctl keytab object from the local catalog. Issue these commands:
   a. `dcecp -c keytab delete /.:/hosts/sp3en0/config/keytab/ssp/sp3en0/sysctl`
   b. `dcecp -c keytab cat | grep sysctl`
2. After the keytab object is deleted, delete and re-create the principal and account. On the control workstation, login as the cell administrator and delete the principal and account.

   ```
   dcecp -c principal delete ssp/sp3en0/sysctl
   ```

   Verify that the entries are gone:

   a. `dcecp -c principal show ssp/sp3en0/sysctl`

      The output is similar to:

      ```
      Error: msgID=0x1712207A  Registry object not found
      ```
   b. `dcecp -c account show ssp/sp3en0/sysctl`

      The output is similar to:

      ```
      Error: msgID=0x1712207A  Registry object not found
      ```

   Re-create the principal and account. If running on the control workstation, ensure that the **-c** flag is specified.

   ```
   config_spsec
   ```

   Answer the prompts that follow:

   ```
   This command requires cell administrator authority. Continue? (y/n) y


    Please enter cell administrator id to be added to ACL admin group: cell_admin


        Your cell administrator password is required to create accounts.
        Please enter your cell administrator password:
   ```

   Verify that the principal and account were re-created. Issue these commands:

   a. `dcecp -c prin show ssp/sp3en0/sysctl`

      Output is similar to:

      ```
      {fullname {}}
           {uid 1883}
           {uuid 0000075b-9dcd-21d3-9c00-0004ac493bac}
           {alias no}
           {quota unlimited}
           {groups spsec-services}
      ```

b. `dcecp -c account show ssp/sp3en0/sysctl -all`

   Output is similar to:

```
{acctvalid yes}
{client yes}
{created /.../sp_cell/cell_admin 1999-11-18-10:30:41.000-05:00I-----}
{description {}}
{dupkey no}
{expdate none}
{forwardabletkt yes}
{goodsince 1999-11-18-10:30:41.000-05:00I-----}
{group spsec-services}
{home /}
{lastchange /.../sp_cell/cell_admin 1999-11-18-10:30:41.000-05:00I-----}
{organization spsec-services}
{postdatedtkt no}
{proxiabletkt no}
{pwdvalid yes}
{renewabletkt yes}
{server yes}
{shell {}}
{stdtgtauth yes}
{usertouser no}
nopolicy
```

c. `kdestroy`

d. `exit`

3. On the host where the key file is missing, create a new key file. Run the PSSP script **create_keyfiles** in **verbose** mode, as the self-host principal. Messages about already existing key files are expected. If the **create_keyfiles** command is issued on the control workstation, ensure that the **-c** flag is specified. Issue this command:

   ```
   create_keyfiles -v
   ```

   Output is similar to:

```
Running "check_prereqs" subroutine ...
Checking state of DCE ...
Running "parse_defaults" subroutine ...
Parsing spsec_defaults file ...
Running "parse_overrides" subroutine ...
Running "create_keys" subroutine ...
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/GSmonitor already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Kbdd already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Master already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Negotiator already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Schedd already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Startd already exists.
Keyfile /spdata/sys1/keyfiles/LoadL/sp3en0/Starter already exists.
Keyfile /spdata/sys1/keyfiles/mmfs/sp3en0/mmfsd already exists.
Keyfile /spdata/sys1/keyfiles/ppe/sp3en0/dpcl already exists.
Keyfile /spdata/sys1/keyfiles/ppe/sp3en0/pmdv3 already exists.
Keyfile /spdata/sys1/keyfiles/rsct/sp3en0/rsct already exists.
Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/css already exists.
Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/pmand already exists.
Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/sp_configd already exists.
Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/spbgroot already exists.
Keyfile /spdata/sys1/keyfiles/ssp/sp3en0/switchtbld already exists.
Running "do_keytab_work" subroutine ...
Creating keytab object, "ssp/sp3en0/sysctl" and randomizing keys.
```

4. Verify that the key file physically exists and is in the keytab catalog. Issue these commands:

   a. `ls -ltr`

Output is similar to:

```
 total 48
     -rw-------   1 root     system       148 Oct 30 13:02 css
     -rw-------   1 root     system       152 Oct 30 13:02 pmand
     -rw-------   1 root     system       162 Oct 30 13:02 sp_configd
     -rw-------   1 root     system       158 Oct 30 13:02 spbgroot
     -rw-------   1 root     system       162 Oct 30 13:02 switchtbld
     -rw-------   1 root     system       154 Nov 18 11:25 sysctl
```

b. `dcecp -c keytab cat | grep sysctl /.../sp_cell/hosts/sp3en0/config/keytab/ssp/sp3en0/sysctl`

5. Verify that key file can be used to login to DCE. As **root**, issue these commands:

   a. `dsrvtgt ssp/sysctl`

   Output is similar to:

   ```
   FILE:/opt/dcelocal/var/security/creds/dcecred_68408200
   ```

   b. `export KRB5CCNAME=FILE:/opt/dcelocal/var/security/creds/dcecred_6840820`

   c. `klist | grep lob Global Principal: /.../sp_cell/ssp/sp3en0/sysctl`

   d. `kdestroy -c FILE:/opt/dcelocal/var/security/creds/dcecred_68408200`

   e. `klist`

   Output is similar to:

   ```
   No DCE identity available: No currently established network identity for this
                    context exists (dce / sec)
    Kerberos Ticket Information:
        klist: No credentials cache file found (dce / krb)
                      (ticket cache /opt/dcelocal/var/security/creds/dcecred_68408200)
   ```

   f. `unset KRB5CCNAME`

   g. `klist | grep lob Global Principal: /.../sp_cell/hosts/sp3en0/self`

   The **dsrvtgt** command can be used with other PSSP service/principal name pairs. See the entry for **dsrvtgt** in *PSSP: Command and Technical Reference*.

6. In this example, the **sysctl** key file was recreated. To ensure that the **sysctld** daemon can use the new key file, stop **sysctld** (if not already stopped), and then start **sysctld**. The **sysctld** daemon should not have any problem using the new key file, because the **dsrvtgt** validation in the previous step was successful.

   a. `startsrc -s sysctld -a '-d'`

   Output is similar to:

   ```
   0513-059 The sysctld Subsystem has been started. Subsystem PID is 21636.
   ```

   b. `lssrc -s sysctld`

   Output is similar to:

   ```
   Subsystem         Group            PID     Status
     sysctld                          21636   active
   ```

   Now verify that **sysctld** is working with the DCE method. In this example, the command was issued in a DCE-only environment.

   ```
   sysctl whoami -v
   ```

   The output is similar to:

```
DCE: /.../sp_cell/hosts/sp3en0/self
     K4:
     AIX: root
```

For other PSSP services whose key files were re-created, run appropriate client commands against those services. Review the output of the commands, and their corresponding server logs, for normal processing indications (responses and entries) in a DCE mode.

Action 6 - Reset authentication values

If you have migrated to PSSP 3.2 from a level of PSSP earlier than 3.1, there are extra steps for security that are required. If those steps are omitted, the information in the SDR will be incorrect.

If you have a situation where the SDR shows the following authentication states after a migration:

```
splstdata -p
.
.
. auth_install   K4
 auth_root_rcmd  std
 ts_auth_methods ''
 auth_methods    std
.
.
.
```

You must run the following steps on the control workstation in order to reset the authentication values for migration to complete properly.

1. `chauthent -k4 -std`
2. `spsetauth -p` *default partition name* `k4 std`

   If you have more than one system partition, issue this command for each of them.
3. `chauthpts -c compat`
4. `chauthpar -c k4 std`

   If you have more than one system partition, issue this command for each of them. Specify the partition with the **-p** flag.
5. `install_cw`

For more information, see the chapter ″Migrating to the Latest Level of PSSP″ of *PSSP: Installation and Migration Guide*. In this chapter, see the section ″Migrating the Control Workstation to PSSP 3.2″, steps ″Verify the Authentication Values in the SDR″ and ″Verify the Authentication Value for AIX Remote Commands″.

Action 7 - Run the chauths command

During configuration of the Control Workstation, the **chauthts** command must be run in order to inform the PSSP configuration code which authentication method will be used to complete the configuration. When this step is omitted, there are several symptoms.

- An error message is received from **install_cw**.
- The SDR and hardmon daemons declare all user requests to be unauthorized, even it the user is **root**, or has an appropriate DCE ticket.
- The SDR attribute **ts_auth_methods** has a value of **new**.

To recover from this situation:

1. Run the **chauthts** command to indicate your chosen authentication method.
2. Rerun **install_cw** or **SDR_init** to populate the SDR with appropriate values.
3. Continue with the configuration steps.

Action 8 - Diagnosing Sysctl security problems

Use this procedure to diagnose Sysctl problems that deal with multiple authentication methods.

1. If the node is running the current release of PSSP, on the client and server nodes, issue the **lsauthts** command to determine the authentication methods supported. If the node is running an earlier release, assume compat mode. Make sure that the client and server hosts have environments which will operate together. A Sysctl client on a host that uses neither DCE nor Kerberos V4 (compat mode) authentication is permitted to access Sysctl servers only as an unauthenticated user. A Sysctl client on a host using only DCE for authentication cannot access a Sysctl server on a host using only Kerberos V4. A Sysctl client on a host using only Kerberos V4 for authentication cannot access a Sysctl server on a host using only DCE.
2. Use the **getauth** Sysctl procedure to determine the authorization callback for an object you cannot access. If you are not authorized to use **getauth**, ask an administrator to do so. If the ACL callback is specified, use the **acllist** procedure to verify that you have permission to access the object, through an entry in the ACL file that protects the object. Make sure that the entry is appropriate for the authentication method that Sysctl is using. See Step 1.
3. If site security policy requires the user to be authenticated, use the DCE **klist** command to verify that the user has DCE credentials. If the system administrator has not configured AIX login to obtain DCE credentials, the user may need to login to DCE separately using the **dce_login** command with the user’s principal name. If the user just needs to refresh credentials, use the **dce_login** or DCE **kinit** command.
4. Use the **dcecp** command to verify that the client user has a DCE principal and account. If you are using ACLs to authorize access to the user, use **dcecp** also to list the ACLs that control access to the desired objects and make any necessary changes.
5. If previously obtained credentials appear to be no longer valid, check whether an administrator has recently changed the server’s keys. If so, you must use **dce_login** or **kinit** to obtain new credentials before retrying the client program.
6. Check the **sysctld** daemon log file for error information. The default name is **/var/adm/SPlogs/sysctl/sysctld.log**)

Chapter 19. Diagnosing Per Node Key Management (PNKM) problems

This chapter discusses diagnostic procedures and failure responses for the Per Node Key Management (PNKM) component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 296. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 292.

Per Node Key Management is a daemon that runs on the control workstation and each node of the SP system to update PSSP-related DCE server keys (passwords) in a DCE authentication environment.

Any principal or group names referred to in this chapter are the default names as shipped with the PSSP product. The principal or group names may be different if you have overridden them by updating the **/spdata/sys1/spsec/spsec_overrides** file.

Related documentation

The following publications provide information about Per Node Key Management:

1. SP system documentation
   - *PSSP: Planning, Volume 2*

     See ″Planning for Authentication″.
   - *PSSP: Administration Guide*
   - *PSSP: Installation and Migration Guide*

     ″Tasks to Install and Configure Authentication Methods″ section.
   - *PSSP: Command and Technical Reference*

     Entries for these commands and files:
     - spnkeyman_start
     - spsec_overrides file
   - *PSSP: Messages Reference*

     See ″Per Node Key Management Messages″.
   - *PSSP: Diagnosis Guide*

     The chapter on diagnosing authentication problems in this book, “Chapter 18. Diagnosing SP Security Services problems” on page 251.
2. AIX related documentation
   - *AIX 5L Version 5.1 Command and Technical Reference*

     The entry for the **errpt** command
3. DCE related documentation
   - *IBM Distributed Computing Environment 3.1 for AIX: Administration Guide and Reference*

     DCE organizations and password (key) expiration.
   - *IBM Distributed Computing Environment 3.1 for AIX: Command Reference*

     The entry for the **dcecp** command.

Requisite function

This is a list of the software and operating system resources directly used by the PNKM component of PSSP. Problems within the requisite software or resources may manifest themselves as error symptoms in PNKM. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the PNKM component of PSSP, you should consider the following components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

DCE Restriction

If you have DCE authentication enabled, you cannot run HACWS.

1. IBM DCE 3.1 for AIX

   DCE (Distributed Computing Environment) provides the functions that allow Per Node Key Management to obtain expiration times and update the PSSP related DCE server keys. DCE Version 3.1 or higher is required.
2. PSSP Installation and Configuration

   This software configures data in the DCE registry for PSSP services that will use DCE and whose keys will be managed by Per Node Key Management. PSSP Installation and Configuration also creates the key files that will be used by the PSSP services and managed by Per Node Key Management. To diagnose these functions, see “Chapter 7. Diagnosing NIM problems” on page 91 and “Chapter 9. Diagnosing node installation problems” on page 101.
3. SP System Security Services

   This software contains the routines necessary for Per Node Key Management to interact with DCE to obtain expiration times and update the PSSP related DCE server keys. To diagnose SP Security Services, see “Chapter 18. Diagnosing SP Security Services problems” on page 251.
4. AIX Version 4.3.3 or higher

   Per Node Key Management can run on stand-alone RS/6000 systems with the AIX operating system, when the prerequisite software is installed and properly configured for use.
5. **/var** file system

   The DCE daemons will be affected if the **/var** file system becomes full. This will prevent users from logging into DCE. Refer to IBM DCE for AIX manuals for more information.

Error information

AIX Error Log

Per Node Key Management logs errors to the AIX Error Log. Use the AIX **errpt** command to view the error logs. A helpful invocation of the **errpt** command is:

```
errpt -a -N spnkeyman
```

This shows a complete listing of all errors logged by **spnkeyman**.

An example of one error logged by **spnkeyman** is:

```
LABEL:                SPNKEYMAN_ERROR104
IDENTIFIER:           C0A5D912

Date/Time:            Tue Oct 26  10:55:49
Sequence Number:      3482
Machine Id:           000101774C00
Node Id:              c166n08
Class:                S
Type:                 PEND
Resource Name:        spnkeyman

Description
UNDETERMINED ERROR

Probable Causes
REVIEW EVENT DETAIL FOR PROBABLE CAUSE

Failure Causes
SOFTWARE SUBSYSTEM

         Recommended Actions
         REVIEW DETAILED DATA
         PERFORM PROBLEM DETERMINATION PROCEDURES
         CORRECT THEN RETRY OFFENDING SOFTWARE COMPONENT
         IF PROBLEM PERSISTS, CONTACT APPROPRIATE
                             SERVICE REPRESENTATIVE

 Detail Data
 DETECTING MODULE
 LP=PSSP, Fn- spnkeymand.c, SID=1.6, l#=582,

 DIAGNOSTIC EXPLANATION
 spsec_keyman_get_expiration error for ssp/sp_configd. err=44
 2502-618 The required keyfile was not found:
     /spdata/sys1/keyfiles/ssp/c166n08.ppd.pok.ibm.com
```

Underlying authentication mechanisms and libraries may have other NLS restrictions. Per Node Key Management displays messages received from these mechanisms without translating them. Consult necessary documentation for this authentication mechanisms for more information.

The AIX error log can wrap, and the error information can be overwritten. Periodically check the log for Per Node Key Management errors and, if any, save the information to a file.

This table lists the AIX error log templates for Per Node Key Management. All entries are of type PEND, meaning that loss of availability of a device or component is imminent.

Table 48. AIX Error Log templates for Per Node Key Management

Label	Error ID	Description
SPNKEYMAN_ERROR100	7A1D698B	**Explanation:** Generic initialization error. **Details:** Program initialization failed.
SPNKEYMAN_ERROR101	E06378AD	**Explanation:** Effective userid is not **root**. **Details:** The user starting the daemon was not using the **root** userid. The user is not authorized.
SPNKEYMAN_ERROR102	789B9E1E	**Explanation:** Thread creation failure. **Details:** The daemon could not create necessary threads.

Table 48. AIX Error Log templates for Per Node Key Management (continued)

Label	Error ID	Description
SPNKEYMAN_ERROR104	C0A5D912	**Explanation:** Generic SP security or DCE error. **Details:** An underlying library call or subsystem error. Messages may be passed through PNKM or issued by PNKM, if possible, concerning this error.

Debug information

Some messages can be truncated when received through the AIX error logging function. You may be able to generate some of the messages by issuing the following command:

```
spnkeymand -l
```

This function displays the SP services that Per Node Key Management controls, to stdout. If the expiration for any service cannot be retrieved, an error message is displayed.

Information to collect before contacting the IBM Support Center

Before calling IBM Service, make sure that you have verified that the environment where the error occurred is correct. Consult “Diagnostic procedures” on page 293. This verification includes:

- Verify that the key is not in an expired state.
- Verify that the key files exist in **/spdata/sys1/keyfiles/***product***/***host***/***service*. An example is **/spdata/sys1/keyfiles/ssp/c55s.ppd.pok.ibm.com/sdr**
- Verify that the data in the SP created DCE organization, **spsec_services**, using the **dcecp** command.
- Verify that the SP service is listed in the SP created DCE organization, **spsec-services**.
- Verify that the service has the proper principal and account in the DCE registry.
- Verify that the DCE servers are up and running and that network communication with the host where the DCE registry is located is OK.
- Verify that DCE is installed and configured on the host where Per Node Key Management is having problems.

If you are still having problems after verifying the environment, collect the following information for the source and target hosts to send to the IBM Support Center.

1. The PSSP level installed.
2. The current PSSP PTF level installed.
3. The current DCE level installed.
4. The current DCE PTF level installed.
5. The error information from the AIX error log, obtained by issuing the **errpt** command.
6. Gathered information on the SP system service:
   - DCE principal name of service that failed (if any).
   - DCE key file name of service that failed (if any).
   - Node on which **spnkeyman** was running.

   - Expiration data from the DCE organization, **spsec-services**. Use the **dcecp** command.
7. Any other diagnostic actions taken that you believe contributes to the solution of the problem or provides additional useful information for debugging the problem.
8. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.

Diagnostic procedures

Since Per Node Key Management depends on a variety of configuration events and customer configuration choices, it can be difficult to isolate a problem.

These are general steps that apply to Per Node Key Management. The steps are presented in decreasing order of likelihood that they will resolve the problem.

Has the spnkeyman daemon been started? Is it sleeping?

The **spnkeyman** daemon is started on the nodes on a reboot after a node has been configured with DCE. The **spnkeyman** daemon must be started manually on the control workstation. Verify whether the daemon was started, and if it is sleeping or running, using the AIX SRC command:

```
lssrc -s spnkeyman
```

Check the AIX error log for errors from **spnkeyman** that would prevent the daemon from running. If the daemon is not running, start it using either the **startsrc** command or the **spnkeyman_start** command.

Are key files for the SP System Services on the host where spnkeyman is running?

Key files for the SP services are located in **/spdata/sys1/keyfiles/***product***/***dcehostname***/***service*. These key files are used to login to DCE to obtain the expiration of the keys and to update the keys when necessary. The key files are created when DCE is configured for SP use. For information on the naming of DCE entities used on the SP system, refer to *PSSP: Planning, Volume 2*.

The **spnkeyman** daemon cannot use a key file if the key file is corrupted. This can occur if a key file has a zero length, contains corrupted data, or contains an entry for the service principal's account that does not match the DCE registry. If the key file contains the maximum number of entries, new entries cannot be added, and keys cannot be updated. You may need to delete entries if this situation occurs. For more information, see *IBM Distributed Computing Environment 3.1 for AIX: Administration Guide and Reference*.

The AIX error log has a limitation on the amount of data that can be issued, and some error messages may be truncated. If this happens for a specific error, you may be able to gather more information by attempting to add a key to the failing key file by issuing the **dcecp keytab add** command.

To correct the key files, see “Action 5 - Correct key files” on page 276.

Have the password expiration times been changed in the SP DCE organization, spsec-services, to indicate when keys should expire?

If no changes have been made to the password expiration in the SP DCE organization, **spsec-services**, then keys for SP services will never expire. If you have updated the password expiration for the SP services, it is recommended that only the password lifetime value be used. This value will trigger updates to the SP services keys through **spnkeyman**. If you have changed this value and wish to have **spnkeyman** check it immediately, stop and restart the **spnkeyman** daemon on each host. Otherwise, the value will be checked within 24 hours.

You can check using the DCE **dcecp** command. If no password expiration information is displayed, then the password expiration information has never been changed.

```
dcecp -c organization show spsec-services
```

Have the SP server keys expired?

Neither an SP service nor the **spnkeyman** daemon will be able to login if the service’s key has expired. You can check using the **dsrvtgt** command. If you receive output similar to the following, the key has not expired. Be sure to delete the credentials once verification is complete.

```
dsrvtgt ssp/css
FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600
kdestroy -c FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600
```

Otherwise, you may have received a ″keys have expired″ error message. Other messages are generated for other types of errors.

If the keys have expired, refer to the chapter ″Managing and Using SP Security Services″, heading ″Managing Server Keys for DCE″ in *PSSP: Administration Guide*. See also “Action 5 - Correct key files” on page 276. Per Node Key Management will generate an error message for each SP service, if the daemon cannot login as that service.

Was the password lifetime expiration time set to less than 24 hours?

The **spnkeyman** daemon checks the expiration time set in the SP DCE organization, **spsec-services**, every 24 hours. If you have set the expiration time to less than 24 hours, the daemon will not have checked, nor updated, the keys and the keys will expire.

Once the keys have expired, neither the SP services nor the **spnkeyman** daemon will be able to login to DCE, and they will fail. The key for each SP service must then be manually updated in the DCE registry and in the key file on the host using the DCE command, **dcecp**.

Check that the **spsec-services** file exists, and that its contents are correct, by using the DCE **dcecp** command. If no password expiration information is displayed, then the password expiration information has never been changed.

```
dcecp -c organization show spsec-services
```

Has the SP DCE organization, spsec-services, been created, and have the SP service principals been added to that organization?

The **spnkeyman** daemon uses the SP DCE organization, **spsec-services**, to obtain data on the key expiration. This organization is normally created when DCE is configured for SP system use. If the **spsec-services** file has not been created, refer to the DCE section in *PSSP: Installation and Migration Guide*.

You can check using the DCE **dcecp** command.

```
dcecp -c organization show spsec-services
```

Is the DCE registry populated with the principals and accounts for SP Services?

Before **spnkeyman** can manage the DCE keys for SP services, the DCE registry must be populated with the principals and accounts for the SP services. This is normally done when DCE is configured for SP system use. For more information, refer to *PSSP: Installation and Migration Guide*. You may also wish to verify that the principal name is correct if you caused the default to be overwritten by using the **spsec_overrides** file.

Use the DCE **dcecp** command to list the principals and accounts in the DCE registry. If there is a particular service you are interested in, you can use the **show** attribute rather than the **list** attribute.

```
dcecp -c principal cat
dcecp -c principal show /.../c55_cell/ssp/c55s/sdr
```

Is DCE installed on the host where spnkeyman is running?

DCE is required for **spnkeyman** to function. DCE must be installed and configured as well as configured for SP use before **spnkeyman** can provide its function.

Are the DCE daemons running on the host where spnkeyman is running?

The DCE client daemons must be running in order for the SP system to function in DCE mode for this host. The daemons must be running in order for **spnkeyman** to function regardless of which authentication is enabled.

The required DCE daemons are: **dced**, **cdsadv**, **secd**, and **cdsd**. See “Verify DCE Security Services configuration” on page 261.

Is There a route for the DCE client daemons to access the DCE server daemons?

The DCE client daemons must have access over the network to the DCE server daemons whether those server daemons are on the control workstation or elsewhere on your site. Name resolution must also be working properly.

Is the DCE secd daemon running on the DCE server host?

The **secd** daemon is the DCE security daemon. If the daemon is not running, PNKM and many other SP services daemons may hang. You may also see this situation if **secd** is running, but the registry has been disabled via a DCE registry disable command. You may see errors like:

```
2502-606 DCE error in sec_login_valid_from_keytable: Cannot find KDC for
requested realm (dce/krb) spsec_keyman_get_expiration error for LoadL/Schedd.
err=1"
```

Since the PNKM daemon restarts on error, you may need to forcibly stop the daemon until the situation is corrected. Do this using the **stopsrc** command with the **-c** flag or the **-f** flag. Issue the command with the **-c** flag first, and if this fails to stop the daemon, issue the command with the **-f** flag.

Error symptoms, responses, and recoveries

The recovery actions in Table 49 correspond to the questions listed in "Diagnostic procedures" on page 293.

Table 49. Per Node Key Management symptoms and recovery actions

Symptom	Recovery
spnkeymand or other SP services cannot login.	See 1. "Are key files for the SP System Services on the host where spnkeyman is running?" on page 293 2. "Have the password expiration times been changed in the SP DCE organization, spsec-services, to indicate when keys should expire?" on page 294 3. "Have the SP server keys expired?" on page 294
spnkeymand is not updating an SP server key.	See 1. "Are key files for the SP System Services on the host where spnkeyman is running?" on page 293 2. "Is the DCE registry populated with the principals and accounts for SP Services?" on page 295 3. "Has the SP DCE organization, spsec-services, been created, and have the SP service principals been added to that organization?" on page 295
spnkeymand is not updating any SP server keys.	See 1. "Are key files for the SP System Services on the host where spnkeyman is running?" on page 293 2. "Is the DCE registry populated with the principals and accounts for SP Services?" on page 295 3. "Has the SP DCE organization, spsec-services, been created, and have the SP service principals been added to that organization?" on page 295 4. "Have the password expiration times been changed in the SP DCE organization, spsec-services, to indicate when keys should expire?" on page 294 5. "Has the spnkeyman daemon been started? Is it sleeping?" on page 293
spnkeymand is not finding an SP service.	See 1. "Is the DCE registry populated with the principals and accounts for SP Services?" on page 295 2. "Are key files for the SP System Services on the host where spnkeyman is running?" on page 293 3. "Has the SP DCE organization, spsec-services, been created, and have the SP service principals been added to that organization?" on page 295

Table 49. Per Node Key Management symptoms and recovery actions (continued)

Symptom	**Recovery**
DCE failures.	See 1. “Is DCE installed on the host where spnkeyman is running?” on page 295 2. “Are the DCE daemons running on the host where spnkeyman is running?” on page 295 3. “Is There a route for the DCE client daemons to access the DCE server daemons?” on page 295
secd daemon not running.	See “Is the DCE secd daemon running on the DCE server host?” on page 295.

Chapter 20. Diagnosing remote command problems on the SP System

Enhanced Security Option

PSSP 3.4 provides the option of running your RS/6000 SP system with an enhanced level of security, called Restricted Root Access (RRA). This function removes the dependency PSSP has to internally issue **rsh** and **rcp** commands as a **root** user from a node. When this function is enabled, PSSP does not automatically grant authorization for a **root** user to issue **rsh** and **rcp** commands from a node. If you enable this option, some procedures may not work as documented. For example, to run HACMP an administrator must grant the authorizations for a **root** user to issue **rsh** and **rcp** commands that PSSP would otherwise grant automatically.

In AIX 4.3.1, the AIX Remote Command suite was enhanced to support Kerberos Version 5 authentication through DCE. These commands include **rsh**, **rcp**, **rlogintelnet**, and **ftp**. For SP migration purposes, the AIX remote commands, **rsh** and **rcp**, were enhanced to call an SP-supplied Kerberos Version 4 set of **rsh** and **rcp** routines. Therefore, the AIX commands **/usr/bin/rsh** and **/usr/bin/rcp** (also in **/bin/rsh** and **/bin/rcp**) on the SP system support the following authentication methods:

- Kerberos Version 5 (through DCE)
- Kerberos Version 4
- Standard AIX

The previously supplied remote commands are no longer shipped with PSSP. The **/usr/lpp/ssp/rcmd/bin/rsh** and **/usr/lpp/ssp/rcmd/bin/rcp** commands are now symbolic links to the AIX commands **/usr/bin/rsh** and **/usr/bin/rcp** respectively.

Things to be aware of when using Restricted Root Access (RRA)

When using Restricted Root Access, check this list of potential problems and restrictions:

1. RRA cannot be selectively applied to some nodes on an SP system. If RRA is activated, it takes effect for all nodes.
2. RRA requires that all nodes be a level PSSP 3.2 or higher.
3. After switching to RRA mode, it is advisable to manually verify all authorization files to ensure that no unwanted entries remain.
4. For the use of multiple Boot/Install servers, see “Action 5 - Check for multiple Boot/Install servers in RRA mode, secure shell mode, or with AIX Authorization for Remote Commands set to ″none″” on page 115.
5. HACMP/ES installation and configuration requires manual updates to authorization files when running in RRA mode
6. If running without RRA, and without a Kerberos authorization file, and then RRA is enabled, an empty **.klogin** file is created. This will prevent anyone from being able to **rlogin** or **telnet** to the node. This will effectively prevent the system administrator from accessing the SP system other than through the control workstation.
7. GPFS, IBM Virtual Shared Disk, The IBM Virtual Shared Disk perspective, and HACMP/ES will not start if RRA is enabled.

8. Do not distribute the **etc/sysctl.conf** file from the control workstation to the nodes when running in RRA mode.
9. Always run the **sysctld** daemon with the same port number (6680 recommended) across the entire SP system. When using RRA mode, critical PSSP functions rely on **sysctl** and failures will occur if there is a mismatch in the port number.
10. When running with HACWS, the **spsitenv** command cannot be used to enable the RRA mode from the backup control workstation. Once RRA has been enabled, the system administrator must manually copy the updated **.rhosts** and **.klogin** file from the control workstation to the backup control workstation

Using secure remote commands instead of AIX rsh and rcp commands

PSSP 3.4 provides the ability to remove the dependency that PSSP has on the AIX **rsh** and **rcp** commands issued as root, on the control workstation as well as on nodes, by enabling the use of a secure remote command method. It is the system administrator's responsibility to choose the secure remote command software and install it on the control workstation. This software must be installed and running, and the root user must have the ability to issue remote commands to the nodes and control workstation without being prompted for a password or passphrase, before the secure remote command facility is enabled for PSSP. All nodes must be at PSSP 3.2 or later releases before you can enable a secure remote command method.

When using the secure remote commands, the Restricted Root Access (RRA) must also be enabled, limiting the use of remote commands to secure remote commands from the control workstation to the nodes. When this function is enabled, PSSP will use the secure remote command methods enabled for all remote command calls, no longer relying on the AIX **rsh** and **rcp** commands.

A public key must be generated for the root ID on the contorl workstation and the boot/install server nodes, and installed on each node, along with the secure remote command software, to ensure that root can issue remote commands from the control workstation and any boot/install server nodes, to the other system nodes, without being prompted for a password or passphrase. Also, either StrictHostNameChecking must be disabled, or the system administrator must generate the **known_hosts** file such that the PSSP installation process can run without prompting from hostname checking.

To enable the secure remote command method, choose one of these options:

- Issue the **spsitenv rcmd_pgm=secrshell dsh_remote_cmd=/bin/ssh remote_copy_cmd=/bin/scp** command.
- Invoke **smit** and update the Site Information Menu to indicate that you want to run with secure remote commands, and specify where the executables are located.

See Step 28 in *PSSP: Installation and Migration Guide*. The PSSP 3.4 system defaults to using **rsh** and **rcp**, and the **bin/rsh** and **bin/rcp** executables for remote commands.

PSSP uses three environment variables that can be set by the user, to determine whether the AIX **rsh** and **rcp** commands, or a secure remote command method, are in effect. The user can use these environment variables to override the SDR settings for PSSP commands.

It is important to keep these environment variables consistent and pointing to the remote command method that you wish to use. If all three environment variables are null, the default is:

- **RCMD_PGM=rsh**
- **DSH_REMOTE_CMD=/bin/rsh**
- **REMOTE_COPY_CMD=/bin/rcp**

If **RCMD_PGM=secrshell** and both **DSH_REMOTE_CMD** and **REMOTE_COPY_CMD** are null, the default is:

- **RCMD_PGM=secrshell**
- **DSH_REMOTE_CMD=/bin/ssh**
- **REMOTE_COPY_CMD=/bin/scp**

In addition, in PSSP 3.4 you have the ability to set Authorization for AIX Remote Commands to ″none″ when secure remote commands are enabled. When this is set, PSSP code will not automatically grant authorization for the root user to issue the **rsh** and **rcp** commands for a node or the control workstation. Instead, all PSSP remote commands will be run using the secure remote command method enabled. In order to set AIX Authorization for Remote Commands to ″none″ on any SP system partition , PSSP 3.4 must be installed on all nodes of that partition.

If ″none″ is enabled, certain functions and procedures may not work as documented. See *PSSP: Administration Guide* for enabling secure remote commands and the ″none″ option. Also, see “Action 21 - Check installation with secure remote command option enabled” on page 111 for possible problems determination and resolution of secure remote command problems.

Related documentation

SP Related Documentation:

- *RS/6000 SP: Planning Volume 2*

  See the chapter ″Planning for Authentication″.
- *PSSP: Administration Guide*

  See the chapter ″Security Features of the SP System″, subheading ″Using the AIX Remote Commands″.
- *Installing and Configuring Distributed Authentication Services*
- *PSSP: Installation and Migration Guide*

  See the chapter ″Tasks to Install/Configure Authentication Methods″.
- *PSSP: Messages Reference*

  Remote command messages are documented in the section ″0041 - Remote Command Messages for Kerberos Version 4″.
- *PSSP: Command and Technical Reference*

  Entries for the following commands:
  - **dsh**
  - **rsh**
  - **rcp**
  - **chauthpar**
  - **config_spsec**
  - **create_keyfiles**
  - **create_dcehostname**

  - **k4list**
  - **k4init**
  - **lskp**
  - **setup_authent**
  - **setupdce**
  - **spauthconfig**
  - **spsetauth**
  - **updauthfiles**
- *PSSP: Diagnosis Guide*

  See “Chapter 18. Diagnosing SP Security Services problems” on page 251.

AIX Related Documentation

- *AIX Commands Reference*

  Entries for the following commands:
  - **chauthent**
  - **ftp**
  - **ftpd**
  - **inetd**
  - **krshd**
  - **krlogind**
  - **lsauthent**
  - **rcp**
  - **rlogin**
  - **rsh**
  - **telnet**
  - **telnetd**

DCE Related Documentation

- *IBM Distributed Computing Environment 3.1 for AIX: Administration Guide and Reference*
- *IBM Distributed Computing Environment 3.1 for AIX: Command Reference*
  - **dce_login**
  - **dcecp**

Requisite function

This is a list of the software and operating system resources directly used by the remote command component of PSSP. Problems within the requisite software or resources may manifest themselves as error symptoms in remote command processing. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with remote command processing, you should consider the following components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

Using AIX rsh and rcp commands

The **rsh** and **rcp** commands on the SP system depend on a number of varied configuration actions and user actions:

1. SP System Security Services

Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

2. The proper authentication methods must be installed and configured. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.
3. The authentication database must be configured for SP system use and have the proper identification for SP services and users. For Kerberos V5 (through DCE) or Kerberos V4, this equates to a principal name in the correct format in the database or registry.
4. The authentication method is enabled on the source and target hosts, on the hosts themselves, and in the partition if applicable. This means that one authentication method **must** be in common between the source and target host pair.
5. The SP service or user has obtained the proper tickets and credentials to pass authentication.
6. An authorization file is present, and principal is present in the file to allow access on the target system.

The **rsh** and **rcp** commands also depend on outside services such as **inetd**, proper network configuration, and reliable name serving and resolution.

Therefore, the SP remote commands have a dependency on the following for the source/target hosts depending on which methods you have installed and enabled.

Table 50. Remote commands - rsh and rcp dependencies

Mechanism	Required	Comments
Kerberos V5 through DCE	**lsauthent** shows Kerberos V5 **lsauthpar** shows Kerberos V5 **.k5login** authorization file	DCE clients installed only on nodes via SP Install/Config scripts.
Kerberos V4	**lsauthent** shows Kerberos V4 **lsauthpar** shows Kerberos V4 **.klogin** authorization file	
Standard AIX	**lsauthent** shows Standard AIX **lsauthpar** shows Standard AIX **.rhosts** authorization file	Not installed or configured by SP system. Authorization file created or updated and distributed if authorization method selected.

The setting of the authentication choices on the local hosts is done through SP configuration of your security selections. However, if those choices are changed via the **chauthent** command on one of the source/target host pairs, the remote commands may fail depending on the authentication methods in effect.

The Kerberos V4 support in the AIX remote commands is supplied through an SP system library which is called from the AIX remote commands. The library contains the ″client″ portion of the remote command and depends on the ″server″ portion as supplied by AIX. This server is named **krshd**.

There may be times when fixes for both the ″client″ side and ″server″ side are required. This means installing PTFs for both the SP system and AIX, to obtain a complete fix.

See “Using secure remote commands instead of AIX rsh and rcp commands” on page 300 for a discussion on secure remote commands, and the Authorization for AIX Remote Commands option of AIX.

Error information

On an SP system, it is now possible to have three different authentication methods in place and configured for different SP system partitions and users. Depending on how you have configured your authentication mechanisms and for what purpose, you may find error messages are displayed and the command finishes successfully.

The remote commands try the authentication methods in this order:

1. Kerberos V5 (DCE)
2. Kerberos V4
3. Standard AIX

Therefore, if you have Kerberos V5 and Kerberos V4 in place, and configured only Kerberos V4 with user ids, you will receive errors when the remote commands try Kerberos V5, before they fail and try Kerberos V4.

Note: If you are using Authorization for AIX Remote Commands=″none″. See “Using secure remote commands instead of AIX rsh and rcp commands” on page 300 .

AIX now supports an environment variable called **K5MUTE**. When set to 1, this variable allows you to mute error messages from the remote commands when you have more than one authentication mechanism enabled. You can set this variable on a system-wide or process basis. It is recommended, however, that this variable not be set when debugging remote command problems because it can hide important messages.

A useful AIX tool in debugging the **krshd** daemon is configuring the AIX syslog. You can have error messages from the **krshd** daemon sent to a log of your choice. You will see messages from other daemons as well. This action must be taken on the target system since that is where **krshd** runs. In general, Kerberos Version 5 messages (through DCE) have a prefix of ″Kerberos″.

To use syslog:

1. On the target system:
   a. Create your log file using the **touch** command. The file must exist before syslog will write to it.
   b. Edit the **/etc/syslog.conf** file and add the line:

      ```
      *.debug  file_name
      ```

      where *file_name* is your log file, with the full path name specified.
   c. Refresh the syslog subsystem to start logging, by issuing these command:
      1) `stopsrc -s syslogd`
      2) `startsrc -s syslogd`
2. On the source host, issue the command you are trying to debug.

3. On the target host, check your log file for **krshd** or **kerberos** errors.

Remember to unconfigure the **/etc/syslog.conf** file when you are done and to refresh the **syslogd** daemon.

Messages from the remote commands are translated to the language of the node on which the command is run. Ensure that the SP system is using either the English location or the SP administrative locale, so that the messages are readable.

Underlying authentication mechanisms may have other NLS restrictions. The remote commands display messages received from these mechanisms as is. Consult the necessary documentation for this authentication mechanisms for more information.

Errors from the remote commands are displayed to the user who issued the command. If you use the remote commands in a script, and do not capture error messages, they are lost. In this case, issue the remote command on the command line, in the same manner that the script issued the remote command.

To see all error messages, ensure that the AIX environment variable, **K5MUTE**, is set to 0.

If you wish to capture errors from the **krshd** daemon, you must configure syslog to do so.

Information to collect before contacting the IBM Support Center

Before calling IBM Service, it is your responsibility to verify that the environment where the error has occurred is correct. Perform these steps:

- Verify that the correct authentication mechanism is installed, configured, updated for SP system use, and enabled on the source and target hosts and in the system partition. See “SP Security Services configuration errors” on page 258.
- Verify that there is **one** authentication mechanism in common between the source/target hosts, both locally on each host and in the system partition where the hosts reside. See “Check enabled security” on page 265.
- Verify that the user has the proper principal in the proper database and registry. See “Verify configuration of Kerberos V4” on page 263.
- Verify that the user has done the proper **login** to obtain credentials before issuing the remote command. See “AIX remote command diagnostics” on page 306.
- Verify that the user is in the proper authorization file on the target host. See “AIX remote command diagnostics” on page 306.
- Verify that the user is using the proper flags if forwarding credentials (Kerberos V5) or when running the remote command in the background. See “AIX remote command diagnostics” on page 306.
- Verify that your secure remote command program is installed and running properly.

If you are still having problems after performing these steps, collect the following information for the source and target hosts, for review by the IBM Support Center.

1. The PSSP level installed
2. The current PSSP PTF level

3. The current PTF level for **bos.net.tcp.***xxxx*, where *xxxx* represents ″client″, ″server″ or ″smit″, followed by spaces and then version.release.modification.fix level. Issue this command:

   ```
   lslpp -L | grep tcp
   ```
4. The AIX Operating System Level, obtained by using the **oslevel** command.
5. The authentication and authorization settings, obtained by issuing the **lsauthent** and **lsauthpar** commands.
6. The exact syntax of the command that generated the error.
7. Any messages displayed with **K5MUTE** not set, or captured through syslog.
8. Any other diagnostic actions taken that you believe contribute to the solution of the problem, or provide additional information in debugging the problem.

AIX remote command diagnostics

To see all error messages, ensure that the AIX environment variable, **K5MUTE**, is set to 0.

Since the remote commands depend on a variety of configuration events, customer configuration choices and user actions, isolating a problem can be confusing. Knowing the authentication methods and states of enablement on the SP system is the first step to problem isolation. Also, user education on the proper actions for gaining tickets or credentials is crucial.

One event to be aware of is that the remote command may be working properly, but the user may not have the proper permissions. The user may need permissions to run a command issued through a remote command, or permission to write to a directory accessed through remote copy on the target machine. This can happen, for example, if the user tries to remote copy a file into an AFS file system to which permission has not been granted. The remote copy command simply returns the error message it received. The actual error is outside of the realm of the remote copy command.

In general, Kerberos V5 (through DCE) messages have a prefix of: **Kerberos** or sometimes **rshd**. Kerberos V4 messages usually have a prefix of **spk4rsh** or **spk4rcp**.

Another helpful problem isolation action is to try to duplicate the problem by running the remote command outside of the script. Run the command at the command line, with as much of the remote command's environment duplicated as possible.

This is a list of general steps that apply to the remote commands under each authentication mechanism. They are listed in decreasing order of probability that the step will find the source of the problem. Answer these questions and follow these steps:

1. For **rsh** and **rcp**, is there a link from **/usr/lpp/ssp/rcmd/bin/rsh** or **/usr/lpp/ssp/rcmd/bin/rcp** to **/usr/bin/rsh** or **/usr/bin/rcp** respectively?

   As of AIX 4.3.1, **/usr/bin/rsh** supports Kerberos V5 (through DCE), Kerberos V4 (through **libspk4rcmd.a**), and Standard AIX. The executables formerly in **/usr/lpp/ssp/rcmd/bin** are no longer shipped.
2. Is the user using the correct flags for the function being used?

   For example, the **-n** flag is necessary when running **rsh** in the background.
3. Is the user principal located in the authorization file on the target host?

- For Kerberos V5, this is **.k5login**. The principal format is:

  ```
  dce_principal_name@cell_name
  ```

  Using a DCE self_host principal in a cell named ″dcecell″ for example: `hosts/trw.pok.ibm.com/self@dcecell`.
- For Kerberos V4, this is **.klogin**. The principal format for an SP service principal is:

  ```
  rcmd.name@domain
  ```

  where *name* is the short hostname returned by the **hostname** command, and *domain* is the domain name. An example is:`rcmd.trw@POK.IBM.COM`.

4. Did the user obtain tickets or credentials?
   - For Kerberos V5, the user must issue the command: **dce_login**
   - For Kerberos V4, the command is **k4init** or **/usr/lpp/ssp/kerberos/bin/kinit**.
5. Does the user have a principal in the database or registry?

   To verify that the user has a valid principal:
   - For Kerberos V5, use the **dcecp** command under DCE.
   - For Kerberos V4, use the **lskp** command.
6. If the principal is a ″server″ principal and an SP system daemon or background function does not work, are the credentials valid? (that is, have the keys expired in Kerberos V5, or can the principal get a service ticket in Kerberos V4)?

   To verify that tickets or credentials are in place:
   - For Kerberos V5, use the **klist** command or **/usr/lpp/dce/bin/klist** command.
   - To verify if SP Services keys (passwords) are still valid under Kerberos V5, use the following **dsrvtgt** command. If you receive output similar to the following, then the keys have not expired. If you receive an error message, the keys have expired. Be sure to destroy the credentials once verification is complete. Use the SP Service principal you are checking in place of *ssp/css*.
     a. `dsrvtgt ssp/css`

        Output is similar to:

        ```
        FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600
        ```
     b. `kdestroy -c FILE:/opt/dcelocal/var/security/creds/dcecred_7aa04600`
   - For Kerberos V4, use the **k4list** command or **/usr/lpp/ssp/kerberos/bin/klist** command.
7. What are the authentication mechanisms enabled and should the user be authenticating via that method?

   If you have DCE (Kerberos V5) and Kerberos V4 enabled, but the user has a principal only for Kerberos V4, you will see messages coming from Kerberos V5. This is because Kerberos V5 tries and fails to authenticate the user before trying Kerberos V4. If you do not want these messages, set the AIX environment variable **K5MUTE**.
8. Are there errors in processing in the **krshd** server on the target host?

Initiate the logging through **syslogd** to verify that **krshd** is working or generating errors. **krshd** is the server for both Kerberos V5 and Kerberos V4 authentication mechanisms. If it generates error messages, save the log for analysis by the IBM Support Center.

9. Is the configuration correct to use **krshd**?

   Check the **/etc/services** and **/etc/inetd.conf** files for the **kshell** service. The **kshell** service in the **/etc/inetd.conf** file should point to the **/usr/sbin/krshd** file. See “Unknown service: kshell or TCP” on page 312.

10. Are the authentication methods you wish enabled for the remote commands actually enabled on the local source and target hosts and in the system partition if applicable?

    Check the local host settings using the **lsauthent** command. Check the partition settings using the **lsauthpar** command. The settings should have at least one authentication method enabled in common with the proper authorization files in place. Of course, the user should have obtained tickets or credentials for that particular method as well.

11. Are the required servers and clients for the authentication mechanism up and running properly?

    If servers are located off of the SP system, also be sure to check for network connections and proper routing for the particular authentication mechanism in use. If using Kerberos V5 under DCE, check whether permissions have been changed in ACL files for specific groups.

12. Has the SP system been configured to use that particular authentication mechanism?

    Enabling the method does not necessary mean the authentication mechanism has been properly installed and configured for SP system use. This can happen when methods are being enabled locally using the AIX **chauthent** command.

ATTENTION - READ THIS FIRST !

On the SP system, the **chauthent** command should be used with caution. The SP Installation and Configuration code will set the proper settings. The **chauthent** command has the unfortunate side effect that **all** of the current methods must be on the command line, along with the ones you are adding. It is easy to unintentionally turn off enablement of other methods using the **chauthent** command.

13. Has the proper service been applied to pick up any fixes for the SP system and AIX?

    Occasionally, fixes are required from both AIX and the SP system. AIX fixes usually pertain to the remote commands using Kerberos V5 or Standard AIX. SP system fixes usually pertain to **rsh** and **rsh** client routines using Kerberos V4 in the **libspk4rcmd.a** library. Note that the AIX **krshd** daemon processes requests from both Kerberos V5 and Kerberos V4 **rsh** and **rcp** commands.

Diagnostics specific to Kerberos V4

1. Is a file or link present in the **/usr/kerberos/bin** directory for **rcp**?

   A link should be present if the Kerberos V4 shipped with the SP system is installed and it should point to the file **/usr/lpp/ssp/rcmd/bin/rcp**. A file should be present only if an MIT version of Kerberos V4 is installed.

2. Has the system administrator run a **ksrvutil change** command?

All current tickets become no longer valid, and the **k4init** command must be run by each user to obtain a current ticket. It is likely that the **ksrvutil change** command will not be run very often.

Diagnostics specific to Kerberos V5

Has the user principal been set properly in the DCE Registry to allow credential forwarding? Has the user invoked **rsh** with the proper flag to indicate forwarding of credentials?

DCE requires that the forwarding-allowed attribute be set to true when creating the user account. Also, the user must use **dce_login -f** when logging into DCE. These actions will allow credential forwarding when the proper flag is present on the **rsh** command invocation. There are two flags associated with forwarding credentials through **rsh**. See the AIX **rsh** man page for more information. Also, consult *IBM Distributed Computing Environment 3.1 for AIX: Administration Guide and Reference* for information about principals, accounts, and forwarding of credentials.

Error symptoms, responses, and recoveries

Most of the error symptoms have obvious solutions such as ″add the user to the database″, or ″have the user issue the correct **login** command″. These types of errors will not have responses and recovery actions listed here.

There are some errors that are not normally seen, but can be generated and are seen through the remote commands. These types of errors are included here for reference.

Remote command (rsh/rcp) symptoms

Use this table to locate the problem symptom and corresponding recover action.

Table 51. Remote command (rsh/rcp) symptoms and recovery actions

Symptom	Recovery action
SIGINT error from **rsh**.	See “SIGINT error from rsh command” on page 310.
Protocol failure errors - Kerberos V4.	See “Protocol failure errors - Kerberos V4” on page 310.
Exec Format Error.	See “Protocol failure errors - Kerberos V4” on page 310.
SP services fail due to Kerberos V5 remote command authentication problems.	See “SP Services fail when using Kerberos V5” on page 311.
Non-SP system applications fail when using **rsh** or **rcp**.	See “Non-SP System applications fail when using rsh or rcp commands” on page 311.
kshell or **tcp**: Unknown service.	See “Unknown service: kshell or TCP” on page 312.
Unable to **rsh**, **telnet**, **rlogin**. **ping** shows host is up.	See “Unable to rsh, rcp, telnet, rlogin, but ping shows host is up” on page 312.

Table 51. Remote command (rsh/rcp) symptoms and recovery actions (continued)

Symptom	Recovery action
Error messages pointing to or from authentication subsystem.	See 1. "Protocol failure errors - Kerberos V4" 2. "SP Services fail when using Kerberos V5" on page 311 3. "Non-SP System applications fail when using rsh or rcp commands" on page 311 4. "Error messages pointing to or from an authentication subsystem" on page 312
Error messages indicating SP Security configuration problems.	See 1. "Protocol failure errors - Kerberos V4" 2. "SP Services fail when using Kerberos V5" on page 311 3. "Non-SP System applications fail when using rsh or rcp commands" on page 311 4. "Error messages pointing to or from an authentication subsystem" on page 312 5. "Error messages indicating SP Security Services configuration problems" on page 313
Kerberos **rsh** message `0041-010 Cannot import nflag or options variables.`	See "Kerberos rsh message 0041-010" on page 313.
Error messages pointing to connection problems.	See "Remote command connection problems" on page 313.
Error messages pointing to server configuration errors (Kerberos V5).	See "Remote command server configuration errors" on page 313.
Remote to remote **rcp** errors.	See "Remote to remote rcp errors" on page 314.
Cannot contact KDC for realm (Kerberos V5).	See "Cannot contact KDC (Kerberos V5)" on page 314.
Decrypt integrity check failed (Kerberos V5).	See "Decrypt integrity check failed (Kerberos V5)" on page 314.
Unexpected remote command authorization failures.	See "Remote command authorization failures" on page 315.

Remote commands (rsh/rcp) recovery actions

SIGINT error from rsh command

This error is received if **rsh** is issued at the command line and placed in the background without using the **-n** flag. The recovery action is to use the **-n** flag when issuing **rsh** in the background.

Protocol failure errors - Kerberos V4

This error message is usually received when the **rsh** client is passing a message through from either the **krshd** server, or one of the lower level Kerberos V4 libraries, where an error occurred. This message normally has the ″passed through″ message appended to it or on the next line.

If the message does not concern items like network errors or connection errors, the first recovery action is to verify that **krshd** is not having problems, by setting up syslog to capture any daemon error messages. See “Error information” on page 304. Remember to set syslog up on the target machine - not the machine where you issued the **rsh** command.

Any error messages logged from **krshd** should help pinpoint the problem and possible solution. The type of messages range from ″cannot locate servers″, or ″cannot reach server″, to errors obtained due to incorrect principal format for authentication or authorization.

While a complete list of ″passed through″ error messages cannot be specified, some general types of errors and recovery actions are helpful:

- **Principal not in database for an adapter that was just added.**

  Run the proper SP configuration scripts to update the database and authorization files.
- **Original hostname has been changed and the principal for that host was not updated or changed.**

  Run the proper SP configuration scripts to update the database and authorization files.
- **Realm or configuration problems may be indicated if servers cannot be located or reached.**

  For more information, see “Chapter 18. Diagnosing SP Security Services problems” on page 251.
- **Exec Format Error.**

  This is returned for incorrect handling of the return code for the Kerberos V4 or compatibility library call: **krb_recvauth**. This routine returns either an error number or message number. If an error number is returned and is treated like an error message number, this error will be returned. See “Information to collect before contacting the IBM Support Center” on page 305 and contact the IBM Support Center.

SP Services fail when using Kerberos V5

SP services can fail using the remote commands with the Kerberos V5 (through DCE) authentication method, due to expired credentials and the inability for services to obtain new credentials if the key used to login has expired.

An SP Service must login to DCE in order to obtain credentials for use with the remote commands. If the principal’s key (″password″) has expired, the service cannot log in to obtain credentials. Remote command failures may be the first indication of this situation, with other seemingly unrelated failures occurring. This type of error may affect services on only one node if that node has been down during the time an automatic refresh of keys had taken place.

For information on how to handle this situation, see “Chapter 18. Diagnosing SP Security Services problems” on page 251.

Non-SP System applications fail when using rsh or rcp commands

The AIX **/usr/bin/rsh** and **/usr/bin/rcp** commands now support Kerberos V5 (through DCE), Kerberos V4 (through an SP-supplied library) and Standard AIX. These commands expect users and applications to have the proper tickets or credentials for the authentication mechanism installed, and the method enabled on the SP system.

Non-SP system applications, such as database or workflow applications, may need to be updated to obtain tickets or credentials for the authentication method enabled. These applications may also need to be updated to place the proper principals, and create the proper accounts in the authentication database or DCE registry for that application's use.

The only bypass for applications that do not support the authentication methods installed, configured, and enabled on the SP system is to disable all authentication methods except for AIX Standard and to put in place the proper authorization files (**.rhosts**).

Unknown service: kshell or TCP

This message usually means that the proper **krshd** configuration has not been done. The primary cause is either a missing or commented out **kshell** line in the **/etc/services** file on the source host, target host or both.

Verify that the following line is valid in the **/etc/services** file on both the source and target host:

```
kshell          544/tcp         krcmd
```

Unable to rsh, rcp, telnet, rlogin, but ping shows host is up

This type of error can indicate one of two problems.

- It could be a configuration error in that the **kshell** line in **/etc/inetd.conf** is missing or commented out. These problems usually are located on the target host.

  Verify that the following line is valid in the **/etc/inetd.conf** file on the target host for **rsh** and **rcp** problems

  ```
  kshell  stream  tcp     nowait  root    /usr/sbin/krshd        krshd
  ```
- There could be a problem with the **inetd** subsystem itself. The **inetd** subsystem accepts connections for daemons and then starts the daemon and passes the socket connection to the daemon. If there are **inetd** problems, the **krshd**, **telnetd**, or **rlogind** daemons may never even get started. Therefore, they could not reply to the **rsh**, **rcp**, **telnet**, or **rlogin** request.

  You can get a variety of error messages, or remote commands may hang under these circumstances. You can also validate an **inetd** problem by setting up syslog to capture log errors from these daemons to see if the daemons start.

  It may help to stop and restart the **inetd** daemon with the AIX **stopsrc** and **startsrc** commands. Otherwise, consult the AIX manuals in debugging an **inetd** problem.

Error messages pointing to or from an authentication subsystem

When you receive messages indicating a problem with the underlying authentication mechanism, it is best to consult the proper manual for diagnosis and recovery information. Although the remote commands would not be the only function to see errors from the authentication mechanism, the remote commands may be the first indication of these errors.

Some authentication mechanism problems that can directly affect the remote commands are:

- Databases or registry unreachable, unreadable, or otherwise corrupted.
- Replica problems, such as:
  - Master servers down and replicas not in place.

  - There is no route to a replica.
  - The replica cannot be found by the client library or client daemon.
- Network problems preventing connection to Master servers or replicas.
- No replica designated to become writable if the Master server goes down.
- Kerberos V4 realm or DCE cell configuration problems.

Error messages indicating SP Security Services configuration problems

As indicated in the dependency section, the remote commands rely on the authentication method being installed and configured on the SP system, configured for SP system use and enabled. Depending on the authentication mechanism, the SP install and configure scripts perform a number of functions for the proper setup of the system.

Configuration and enablement of an authentication method depends on the choices you have made on the SP Security SMIT panels, and the resulting installation and configuration of the nodes after those choices have been made.

Some of the errors that directly affect the remote commands include:

- Errors in distribution of the authorization files.
- Errors in starting required daemons such as the DCE client.
- Errors in the distribution of necessary configuration files.
- Errors in populating the Kerberos V4 database or DCE registry with the proper SP principals

For more information, consult “Chapter 9. Diagnosing node installation problems” on page 101 in this book, *PSSP: Installation and Migration Guide*, and *RS/6000 SP: Planning Volume 2*.

Kerberos rsh message 0041-010

This message is received if you have not installed the AIX APAR IX85420 but have updated your SP system with **ssp.clients** fixes. Install the AIX APAR to obtain the companion fix to allow the Kerberos **rsh** routine to obtain the variables it needs from the AIX **rsh** client. The minimum AIX file set required is **bos.net.tcp.client.4.3.2.4**. If you have any questions, contact the IBM Support Center.

Remote command connection problems

The target host may be in the process of rebooting, or the **inetd** system may be down. Check whether the **krshd** daemon is listed in the file **/etc/inetd.conf** on the target system. If you had to add the **krshd** daemon, be sure to stop and restart the **inetd** subsystem.

For a ″Kerberos V5 Connection abort″ message, it means that DCE was deinstalled, but a DCE **unconfig_admin** was never done for the target.

For a ″Kerberos V5 Connection ended by software″ message, DCE configuration is incomplete. Client services are only partially available.

Remote command server configuration errors

The authentication services (servers and clients) may be inactive even though the services are still enabled. The authentication services may have been unconfigured. Use the **lsauthent** and **lsauthpar** commands to verify which authentication services are enabled, then check that they are running.

Remote to remote rcp errors

In order for remote to remote **rcp** to work, credentials must be present on the intermediate source host. Kerberos V5 (through DCE) supports forwarding of credentials. Kerberos V4 does not support forwarding of credentials. With a choice of authentication now available, this command may or may not work, depending on the authentication methods enabled on the systems involved.

For example, if you have three hosts named: **HostA**, **HostB**, and **HostC** and you issue the following command from **HostA**:

```
rcp HostB:/file HostC:/file
```

The following table shows what can happen and some of the necessary requirements.

Table 52. Remote to remote rcp

Authentication method	**Result**	**Comments**
Kerberos V5 (through DCE) on all hosts	Success	User must have forwardable credentials to allow use of those credentials on **HostB**.
Kerberos V4 on all hosts	Error	No credentials are forwarded, so no credentials exist on **HostB** for the **rcp** command to use from **HostB** to **HostC**. If you are **root**, you may use the **rsh** command to issue the **rcmdtgt** and **rcp** commands on **HostB**.
AIX standard on all hosts	Success	Based on IP addresses and not user authentication.
Kerberos V5 on **HostA** and **HostB** Kerberos V4 on **HostA**, **HostB** and **HostC**	Error	The **rcp** command from **HostA** to **HostB** works, but the **rcp** command started on **HostB** uses DCE first, and fails because DCE is not on **HostC**. The **rcp** command on **HostB** then tries Kerberos V4 and does not have the tickets necessary to continue.
Kerberos V5 on **HostA** and **HostB** Standard AIX on **HostA**, **HostB** and **HostC**	Success	The **rcp** command from **HostA** to **HostB** works, but the **rcp** command started on **HostB** uses DCE first, and fails because DCE is not on **HostC**. The **rcp** command on **HostB** then tries the next method (standard AIX), which is successful.
Kerberos V5 on **HostB** and **HostC** Kerberos V4 on **HostA**, **HostB** and **HostC**	Possible success	If the user has a DCE login session on **HostB**, the command works because the **rcp** command on **HostB** has the user’s credentials available. Otherwise, the command fails.

Cannot contact KDC (Kerberos V5)

The **/etc/krb5.conf** file contains an error in the **realms** stanza. The **kdc=** entry does not contain the correct address for the security server’s location. Usually, the problem is that the kdc entry address is on a different subnet than the current host.

Decrypt integrity check failed (Kerberos V5)

This usually indicates that a node’s DCE services are out of synch with the rest of the DCE cell. Perform these steps:

1. Stop the DCE servers and clients, by issuing the **stopsrc** command.
2. Issue the **clean_up.dce** command.
3. Issue the **clean_up.dce-core** command.

4. Issue the **rmxcreds** command.
5. Restart the DCE servers and clients, by issuing the **startsrc** command.

Remote command authorization failures

If you experience unexpected failure when issuing remote commands, check the following:

1. Check the SDR Site Environment settings. Run **splstdata -e** on the control workstation to make sure that the Restricted Root Access setting is correct. If not, issue the **spsitenv** command for the **restrict_root_rcmd** attribute. A value of `true` indicates that Restricted Root Access in to be used. A value of `false` indicates that Restricted Root Access is not to be used.
2. Check the remote command authorization files on the control workstation. If they are not correct, run **spsitenv** for the **restrict_root_rcmd** attribute to force all authorization files to be regenerated. If the **restrict_root_rcmd** attribute is `false` and your authentication is Kerberos Version 4, also run the **setup_authent** command.
3. Check the remote command authorization files on the nodes. If they are not correct, run **spsitenv** for the **restrict_root_rcmd** attribute to force all authorization files to be regenerated. If the **restrict_root_rcmd** attribute is `false` and your authentication is Kerberos Version 4, also run the **setup_authent** command.

Using secure remote commands - symptoms

Table 53. secure remote command symptoms and recovery actions

Symptom	Recovery action
Parallel command (**dsh**, **pcp**) hangs with secure shell enabled	See “Action 1”.
Parallel command (**dsh**, **pcp**) is using the wrong remote command method (**rsh** versus secure shell).	See “Action 2”.
Secure connection to **hostname** refused.	See “Action 3” on page 316.
After node install with secure shell enabled, failures in remote copy of files form the control workstation to the nodes from **firstboot.cust**.	See “Action 4” on page 316.

Using secure remote commands - recovery actions

Action 1

Perform these steps:

1. Check to see that the user public key is installed on the nodes to which the **dsh** command is being sent.
2. Check to see that either StrictHostName checking is disabled, or that the node’s name (long and short hostname) is in the **known_host** file.
3. Perform “Action 2”.

Action 2

Check your setting of the **RCMD_PGM**, **DSH_REMOTE_CMD** and **REMOTE_COPY_CMD** environment variables. These variables determine which remote command method is used by parallel commands.

Action 3

Check to see that the **sshd** daemon is running on the host that is listed in the error message received when the connection failed.

Action 4

Check to see that the secure remote command program was installed, and that the daemon started by **script.cust** on the node. Ensure that the **sshd** daemon was put in file **/etc/inittab** right after the **rctcpip** command.

Chapter 21. Diagnosing System Monitor problems

This chapter discusses diagnostic procedures and failure responses for the System Monitor component of PSSP. The list of known error symptoms and the associated responses are in the section "Error symptoms, responses, and recoveries" on page 328. A list of the information to collect before contacting the IBM Support Center is in the section "Information to collect before contacting the IBM Support Center" on page 321.

Related documentation

The following publications provide information about the System Monitor:

1. *PSSP: Command and Technical Reference*
   - hardmon daemon
   - hmadm
   - hmcmds
   - hmmon
   - spmon
   - s1term
   - spsvrmgr
   - hmckacls
   - hmgetacls
   - hmdceobj
   - hmreinit
   - spmon_ctest
   - spmon_itest
   - spapply_config
2. *PSSP: Messages Reference*

   The chapter ″0026 - System Monitor Messages″ contains messages issued by the System Monitor.
3. *PSSP: Administration Guide*
4. *PSSP: Diagnosis Guide*

   See "Chapter 6. Diagnosing hardware and software problems" on page 87.

Requisite function

This is a list of the software directly used by the System Monitor component of PSSP. Problems within the requisite software may manifest themselves as error symptoms in this software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the System Monitor component of PSSP, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

| - hmc daemon (**hmcd**)
- SP Supervisor hardware
  – Frame Supervisor cards
  – Node Supervisor cards
  – Switch Supervisor cards

  - I2C (I’squared’C) bus
- SP Supervisor microcode
- control workstation serial port
- S70 Daemon (**s70d**)
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
- System Data Repository
- TCP/IP Sockets
- AIX error logging (**errpt**) function
- **/var/adm/SPlogs/spmon** file system
- **/spdata/sys1/spmon/hmacls** file
- **/spdata/sys1/spmon/hmdceacls** file
- **/spdata/sys1/spmon/hmthresholds** file
- **/spdata/sys1/ucode** file system
- **/usr/lib/nls/msg/sysmon.cat** file system

Error information

AIX Error Logs and templates

The format of these entries is the format of the standard error log entries, which are viewed using the **errpt -a** command. These are hardware errors. An example is an entry that indicates: `Power module - DC power loss.`.

Entries are logged when the condition appears to be permanent. The error information can be cleared. **Do not** clear an entry if an error has occurred. Entries appear in the language specified by the current location setting (SP administrative locale) on the control workstation.

Table 54 shows the error log templates used by the system monitor. UNKN indicates an unknown error type. PERM indicates a permanent error type.

Table 54. AIX Error Log templates for the System Monitor

Label	Error ID	Type	Class	Description
HM001_TR	F55FBD1D	UNKN	S	**Explanation:** Informational System Monitor message. **Details:** These are informational messages that primarily report unrecognized debug flags specified by the **hmadm** command, or incorrect arguments to the System Monitor (hardmon) daemon.
HM002_TR	B463BBF0	UNKN	S	**Explanation:** Informational System Monitor message. **Details:** These are informational messages that report certain key events on the progress of the System Monitor daemon process. For example, initiating the main processing loop, or the opening or closing of a log file.

Table 54. AIX Error Log templates for the System Monitor (continued)

Label	Error ID	Type	Class	Description
HM003_ER	6CCA6600	PERM	S	**Explanation:** The System Monitor has encountered an internal error. **Details:** An internal error has occurred. Examples are system calls that have ended in error or internal program errors.
HM004_ER	3A814C25	PERM	S	**Explanation:** The System Monitor has encountered an internal error. **Details:** An internal error has occurred. Examples are system calls that have ended in error, or internal program errors.
HM005_ER	8D64B0C8	PERM	S	**Explanation:** The System Monitor has encountered an internal error. **Details:** An internal error has occurred. Examples are system calls that have ended in error or internal program errors. This can also be caused by an unexpected signal that killed an external hardware driver, or the external hardware driver's restart limit being exceeded.
HM006_ER	BB9D16B2	PERM	S	**Explanation:** The System Monitor has encountered a configuration error. **Details:** A configuration file, or the SDR, was found to contain incorrect data. Examples are: if the **hmthresholds** file contains an incorrect card type, or if there is a duplicate entry in the **hmacls** file.
HM007_ER	C70D5E9D	PERM	S	**Explanation:** A resource is unavailable to the System Monitor. **Details:** The primary reason for this error is that the System Monitor was unable to establish a session with the SDR.
HM008_ER	472C610A	PERM	S	**Explanation:** The System Monitor has encountered a data packet error. **Details:** There is some error involving the data packet that is sent from a frame supervisor to the System Monitor daemon. Examples are a missing data packet or an unexpected supervisor type in the data packet.
HM009_ER	60A89611	PERM	S	**Explanation:** The System Monitor has received a bad client command. **Details:** A command sent to the System Monitor from a client such as: hmmon, hmcmds, s1term is bad. For example, the command is incorrect or the command arguments are incorrect.
HM010_ER	A7DA9589	PERM	S	**Explanation:** The System Monitor could not open a file. **Details:** The System Monitor could not open a log file, a configuration file, or a tty.

System Monitor daemon log file

This file is located in: **/var/adm/SPlogs/spmon/hmlogfile.***ddd*, where *ddd* is the Julian day, on the control workstation. It is not trimmed. It is created automatically when the System Monitor (hardmon) is started.

A new system monitor daemon log file can be created by issuing the command:

```
hmadm clog
```

The previous log file is not deleted. A new one (with a different name) is created.

Entries in the file have this format:

```
hardmon:   message_number   message_text
```

Locate the message number in: *PSSP: Messages Reference*.

An example of some data in this file is :

```
 hardmon:  0026-801I  Hardware Monitor Daemon started at Wed Dec 22 17:46:06 1999
 hardmon:  0026-802I  Server port number is 8435, poll rate is 5.000000 seconds
 hardmon:  0026-805I  1 frames have been configured.
 hardmon:  0026-803I  Entered main processing loop
```

Dump information

A standard AIX core dump is located in: **/var/adm/SPlogs/spmon/hardmon/core**. Each time a core file is created, it overlays the previously generated core file. A dump is created on each occurrence of the following conditions:

1. Errors returned from system calls. For example:
   - Memory allocation errors.
   - File open errors.
   - TTY open errors.
   - Socket processing errors.
   - Child processing errors.
2. Configuration errors. For example:
   - **/spdata/sys1/spmon/hmacls** file errors.
   - **/spdata/sys1/spmon/hmdceobj** database errors.
   - **/spdata/sys1/spmon/hmthresholds** file errors.
   - System Data Repository (SDR) errors.
3. Various internal (consistency check) errors. For example:
   - About to index past an array boundary.
   - Control block chain is corrupted.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Hardmon provides a facility that allows the tracing of some internal functions, as well as its communication with the SP Frame Supervisor. It accomplishes this through the use of the **hmadm** command. Use this trace only when directed by the IBM Support Center. The entry for **hmadm** in *PSSP: Command and Technical Reference* gives details about running the trace and obtaining the output.

Information to collect before contacting the IBM Support Center

1. Any command output that seems to be helpful. For example, the output of any failing Diagnostic Procedures that were run.
2. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.
3. AIX error log. On the control workstation, issue this command:

   ```
   LANG=C errpt -a > /tmp/AIXerrlog
   ```
4. System Monitor (hardmon) daemon core dump, if one exists. See “Dump information” on page 320.
5. System Monitor (hardmon) daemon log file. See “System Monitor daemon log file” on page 320.

Diagnostic procedures

These procedures check the installation, configuration, and operation of the System Monitor.

Installation verification tests

Use these tests to check that the System Monitor is installed properly.

Installation test 1 – Check ssp.basic file set

This test verifies that the **ssp.basic** file set has been installed correctly. System Monitor function is included in the **ssp.basic** file set.

Issue this **lslpp** command on the control workstation:

```
lslpp -l ssp.basic
```

Good results are indicated by output similar to the following:

```
Path: /usr/lib/objrepos
 ssp.basic      3.1.0.8  COMMITTED  SP System Support Package

Path: /etc/objrepos
 ssp.basic      3.1.0.8  COMMITTED  SP System Support Package
```

In this case, proceed to “Installation test 2 - Check System Monitor files”.

Error results are indicated if entries for **ssp.basic** do not exist. In this case, try to determine why the file set was not installed, and either attempt to install it, or contact the IBM Support Center.

Installation test 2 - Check System Monitor files

This test verifies that the System Monitor daemon and associated commands and configuration files have been created in the proper directory on the control workstation. Issue these commands and verify that all these files exist:

```
        ls -l /usr/lpp/ssp/bin/hardmon
        ls -l /usr/lpp/ssp/bin/hmadm
        ls -l /usr/lpp/ssp/bin/hmcmds
        ls -l /usr/lpp/ssp/bin/hmmon
        ls -l /usr/lpp/ssp/bin/spmon
        ls -l /usr/lpp/ssp/bin/s1term
        ls -l /usr/lpp/ssp/bin/spsvrmgr
        ls -l /usr/lpp/ssp/bin/hmckacls
        ls -l /usr/lpp/ssp/bin/hmgetacls
        ls -l /usr/lpp/ssp/bin/hmdceobj
        ls -l /usr/lpp/ssp/bin/spmon_itest
        ls -l /usr/lpp/ssp/bin/spmon_ctest
        ls -l /usr/lpp/ssp/install/bin/hmreinit
        ls -l /spdata/sys1/spmon/hmacls
        ls -l /spdata/sys1/spmon/hmthresholds
        ls -l /spdata/sys1/spmon/hwevents
        ls -l /spdata/sys1/ucode
```

Notes:

1. The last entry, **ucode**, is a directory where one or more microcode files are located It is a part of the **ssp.ucode** file set, which is a prerequisite to the **ssp.basic** file set.
2. If the control workstation is in DCE only mode, the **hmacls** file is not used, and therefore does not need to exist.

Good results are indicated if all of the files exist and the files that are located in a **bin** directory are executable. Proceed to “Installation test 3 - Run the spmon_itest command”.

Error results are indicated if entries for one or more of these files does not exist. In this case, try reinstalling your SP system, or contact the IBM Support Center. If you have already reinstalled your SP system, resume diagnostics with “Installation test 1 – Check ssp.basic file set” on page 321.

Installation test 3 - Run the spmon_itest command

This test verifies that the system monitor is installed correctly. Issue this command on the control workstation:

```
spmon_itest
```

Good results are indicated by output similar to:

```
spmon_itest: Start spmon installation verification test
spmon_itest: Verification Succeeded
```

In this case, proceed to “Configuration test 1 - Check /etc/services file”.

Error results are indicated in all other cases. Check the file **/var/adm/SPlogs/spmon/spmon_itest.log** for error messages, and take appropriate action based on the messages. Repeat this test after taking corrective actions based on the messages from **spmon_itest**.

Configuration verification tests

Use these tests to check that the System Monitor is configured properly.

Configuration test 1 - Check /etc/services file

This test verifies that there is an entry for hardmon in the **/etc/services** file on the control workstation. Browse the file **/etc/services** and look for **hardmon** in the left column:

```
hardmon         8435/tcp
```

Good results are indicated if this entry is present. Proceed to “Configuration test 2 - Check SDR Frame class”.

Error results are indicated if this entry is not present. It should have been made during the installation of the SP system. Contact the IBM Support Center.

Configuration test 2 - Check SDR Frame class

This test verifies that the SDR **Frame** class exists, that all MACN values are the same for all frames, and that these values are the same as the output of the **vhostname** command.

Issue these commands on the control workstation:

1. `SDRGetObjects Frame`
2. `vhostname`

Good results are indicated if all of the MACN values, for each frame in the **Frame** object, are identical to the value returned by the **vhostname** command. Proceed to “Configuration test 3 - Check SDR SP_ports class” on page 324.

Error results are indicated if there are differences in the values returned by these two commands. In this case, perform one of these corrective actions:

1. Issue the command **hmreinit** to reinitialize the SDR **Frame** class.
2. Use the **SDRChangeAttrValues** command to correct one or more incorrect MACN values, in the respective **Frame** object. That is, change the value to be identical to what is returned by the **vhostname** command. Refer to the entries for the **hmreinit** and **SDRChangeAttrValues** commands in *PSSP: Command and Technical Reference*.
3. Change the system hostname to match the MACN attribute in the **Frame** class.

   Note: Whatever value the hostname has at this point, it must match the hostname attribute of the SDR **SP_ports** class.

Repeat this test after performing one of the corrective actions.

Configuration test 3 - Check SDR SP_ports class

This test verifies that the SDR **SP_ports** class exists, and that an object exists whose daemon attribute is **hardmon**. The test also verifies that the hostname and port attributes for this object are correct.

Issue these commands on the control workstation:

1. `SDRGetObjects SP_ports`

   The output is similar to:

   ```
   daemon      hostname              port
   hardmon     sup1.ppd.pok.ibm.com  8435
   haemd       ""                    10000
   ```

2. `vhostname`

Good results are indicated if all of these conditions are true:

1. There is an object in the **SP_ports** class whose daemon attribute is **hardmon** (as in the `SDRGetObjects SP_ports` output).
2. The hostname attribute of this object is identical to the value returned by the **vhostname** command.
3. The port attribute of this object is **8435**.

In this case, proceed to “Configuration test 4 - Check SDR Syspar class” .

Error results are indicated if one or more of these conditions is not true. In this case, perform one of these corrective actions:

1. If the hardmon entry does not exist, use the **SDRCreateObjects** command to create the missing **SP_ports** (hardmon entry) object.
2. If the hardmon entry exists, use the **SDRChangeAttrValues** command to modify the incorrect **SP_ports** (hardmon entry) object.
3. If the problem is that the hostname attribute does not match the value returned by the **vhostname** command, correct the problem by performing one of these steps:
   a. Change the system hostname so that they both match.
   b. Use the **SDRChangeAttrValues** command to change the hostname attribute so that they both match.

Note: Whatever value the hostname has at this point, it must match the hostname attribute of the SDR **SP_ports** class.

After performing corrective actions, proceed to “Configuration test 2 - Check SDR Frame class” on page 323.

Configuration test 4 - Check SDR Syspar class

This test verifies that the SDR **Syspar** class exists, and that one object exists for each system partition. Issue this command on the control workstation:

```
SDRGetObjects Syspar
```

Good results are indicated if the **Syspar** class exists, and there is an object for each system partition. In this case, proceed to “Configuration test 5 - Check SDR Switch class” on page 325.

Error results are indicated in all other cases. If the **Syspar** class does not exist or it is empty, contact the IBM Support Center. If the **Syspar** class exists and is not empty, but there is not an object for each system partition, perform one of these actions:

1. Try to reapply your partition configuration, by issuing the **spapply_config** command. For details of this command, refer to *PSSP: Command and Technical Reference*.
2. Contact the IBM Support Center.

Configuration test 5 - Check SDR Switch class

Do not perform this test unless your system has one or more switches.

This test verifies that the SDR **Switch** class exists, and that one object exists for each switch in the system. Issue the following command on the control workstation:

```
SDRGetObjects Switch
```

Good results are indicated if the **Switch** class exists, and there is an object for each switch in the system. In this case, proceed to “Configuration test 6 - Check SDR Syspar_map class”.

Error results are indicated in all other cases. If the **Switch** class does not exist, contact the IBM Support Center. If the **Switch** class exists, but it is empty, or it does not contain one object for each switch in the system, perform one of these actions:

1. Issue the **SDR_config** command. For details of this command, refer to *PSSP: Command and Technical Reference*.

   After running **SDR_config**, if there is still a problem with the SDR **Switch** class, it may be because the **SDR_config** command obtains information from hardmon, and hardmon is inactive or having problems. To determine if this is the case, see “Operational test 1 - Check that hardmon Is active” on page 327.
2. Contact the IBM Support Center

Configuration test 6 - Check SDR Syspar_map class

This test verifies that the SDR **Syspar_map** class exists, and that one or more objects exist. Issue this command on the control workstation:

```
SDRGetObjects Syspar_map
```

Good results are indicated if the **Syspar_map** class exists, and there is at least one object whose `used` attribute is 1. Proceed to “Configuration test 7 - Check SDR NodeExpansion class” on page 326.

Note: Nodes that do not physically exist may have entries, but their `used` attribute should be 0.

Error results are indicated in all other cases. If the **Syspar_map** class does not exist, contact the IBM Support Center. If the **Syspar_map** class exists, perform one of these actions:

1. Issue the **SDR_config** command. For details of this command, refer to *PSSP: Command and Technical Reference*. After running **SDR_config**, if there is still a problem with the SDR **Syspar_map** class, it may be because the **SDR_config** command obtains information from hardmon, and hardmon is inactive or having problems. To determine if this is the case, see “Operational test 1 - Check that hardmon Is active” on page 327.

2. Contact the IBM Support Center

Repeat this test after taking the suggested actions.

Configuration test 7 - Check SDR NodeExpansion class

Do not perform this test unless your system has one or more Expansion Nodes.

This test verifies that the SDR **NodeExpansion** class exists, and that one object exists for each Expansion Node in the system. Issue the following command on the control workstation:

```
SDRGetObjects NodeExpansion
```

Good results are indicated if the **NodeExpansion** class exists, and there is an object for each Expansion Node in the system. Proceed to “Configuration test 8 - Check frame, node, and switch supervisor cards”.

Error results are indicated in all other cases. If the **NodeExpansion** class does not exist, contact the IBM Support Center. If the **NodeExpansion** class exists, perform one of these actions:

1. Issue the **SDR_config** command. For details of this command, refer to *PSSP: Command and Technical Reference*. After running **SDR_config**, if there is still a problem with the **NodeExpansion** class, it may be because the **SDR_config** command obtains information from hardmon, and hardmon is inactive or having problems. To determine if this is the case, see “Operational test 1 - Check that hardmon Is active” on page 327.
2. Contact the IBM Support Center

Repeat this test after taking the suggested actions.

Configuration test 8 - Check frame, node, and switch supervisor cards

This test verifies that all frame, node, and switch supervisor cards that support microcode download, contain the latest level. Issue this command on the control workstation:

```
smit supervisor
```

Select either of the two buttons labeled `List Supervisors That Require Action`.

Good results are indicated by output similar to:

```
 spsvrmgr: All specified supervisor hardware is current and active.
           No further action is required at this time.
```

Proceed to “Configuration test 9 - Run the spmon_ctest command” on page 327.

Error results are indicated in all other cases. If this test fails, it means that one or more of the supervisor cards require some action. Either refer to the entry for the **spsvrmgr** command in *PSSP: Command and Technical Reference*, or use **smit** to take the action for all of the supervisor cards, or one at a time:

Issue this command:

```
smit supervisor
```

To take the required action on all supervisor cards, select the button labeled:

```
Update *ALL* Supervisors That Require Action
```

To take the required action on a single supervisor card, select the button labeled:

```
Update Selectable Supervisors That Require Action
```

Repeat this test after taking the suggested actions. If this test fails again after taking the required action on all supervisor cards that require action, contact the IBM Support Center.

Configuration test 9 - Run the spmon_ctest command

This test verifies that the System Monitor is configured correctly. Issue this command on the control workstation:

```
spmon_ctest
```

Good results are indicated if the output is similar to:

```
spmon_ctest: Start spmon configuration verification test
spmon_ctest: Verification Succeeded
```

Proceed to “Operational test 1 - Check that hardmon Is active”.

Error results are indicated in all other cases. Check the file **/var/adm/SPlogs/spmon_ctest.log** for error messages, and take appropriate action based on the messages. Repeat this test after taking corrective action.

Operational verification tests

Use these tests to check that the System Monitor is operating properly.

Operational test 1 - Check that hardmon Is active

This test verifies that the System Monitor (hardmon) is active and running correctly. Issue these commands:

1. `lssrc -s hardmon`
2. `ps -ef | grep hardmon`

Output from the **lssrc** command is similar to:

```
Subsystem         Group            PID     Status
 hardmon                           42532   active
```

Output from the **ps** command is similar to:

```
root 42532  5966  0  Sep 15   0  9:42 /usr/lpp/ssp/bin/hardmon -r 5
```

Good results are indicated if all of the following are true:

1. In the **lssrc** command output for the entry whose `Subsystem` is `hardmon`, the `Status` column is `active`.
2. In the **ps** command output, verify that the hardmon daemon uses the **-r** flag and that it has an argument of `5`.

   This means that the hardmon daemon polls each frame supervisor, including external hardware daemons, for state information, every five seconds. This is the default. If the hardmon daemon uses a value other than `5` for the **-r** flag argument, the daemon is not running according to IBM recommendations.

In this case, proceed to “Operational test 2 - Run query command”.

Error results are indicated in all other cases. To correct, see “Action 2 - Start the hardmon daemon” on page 329. Repeat this test after taking corrective actions. If the problem persists, contact the IBM Support Center.

Operational test 2 - Run query command

This test verifies that you can run a typical System Monitor query command. Issue this command:

```
hmmon -GQ 1:0-17
```

Good results are indicated if you get state output for slot 0 (the frame itself) in frame 1 (which must always exist), and any nodes in frame 1. An example is:

```
frame 001, slot 00:
  node 01 I2C not responding   FALSE
  node 02 I2C not responding   TRUE
  node 03 I2C not responding   FALSE
  node 04 I2C not responding   TRUE
  node 05 I2C not responding   FALSE
  node 06 I2C not responding   TRUE
  node 07 I2C not responding   TRUE
  node 08 I2C not responding   TRUE
  node 09 I2C not responding   FALSE
  node 10 I2C not responding   TRUE
  node 11 I2C not responding   FALSE
  node 12 I2C not responding   FALSE
  node 13 I2C not responding   FALSE
  node 14 I2C not responding   FALSE
  node 15 I2C not responding   TRUE
  node 16 I2C not responding   FALSE
  switch I2C not responding    FALSE
  controller tail is active    TRUE
  node 01 serial link open     FALSE
  node 02 serial link open     FALSE
           .
           .
           .
```

Error results are indicated in all other cases. This test can fail for several reasons. Read any error messages that are output by the **hmmon** command, and check the hardmon log **/var/adm/SPlogs/spmon/hmlogfile.***ddd*, where *ddd* is the Julian date of when the file was created, for any new messages. Refer to relevant actions in “Error symptoms, responses, and recoveries”.

Error symptoms, responses, and recoveries

Use this table to diagnose problems with the system monitor component of PSSP. Locate the symptom and perform the action described in the table.

Table 55. System Monitor symptoms

Symptom	Recovery
The **ssp.basic** file set is not installed.	See “Action 1 - Install ssp.basic” on page 329.
The hardmon daemon is not running.	See “Action 2 - Start the hardmon daemon” on page 329.
The hardmon daemon keeps terminating and then restarting.	See “Action 3 - Investigate the hardmon daemon” on page 330.

Table 55. System Monitor symptoms (continued)

Symptom	Recovery
A System Monitor command, for example **hmmon**, does not work correctly.	See “Action 4 - Investigate System Monitor command problems” on page 330.
Hardmon performance is poor.	See “Action 5 - Check paging space and CPU utilization” on page 332.

Actions

Action 1 - Install ssp.basic

Run “Installation test 1 – Check ssp.basic file set” on page 321 to verify that this is a problem. Install the **ssp.basic** file set, by issuing the **installp** command. Run “Installation test 1 – Check ssp.basic file set” on page 321 again to verify that the problem has been resolved.

Action 2 - Start the hardmon daemon

Run “Operational test 1 - Check that hardmon Is active” on page 327 to verify that this is a problem.

If the hardmon daemon is not running on the control workstation, you must start it. Do this by issuing the command:

```
startsrc -s hardmon
```

Run “Operational test 1 - Check that hardmon Is active” on page 327 to verify that hardmon was started successfully.

If the hardmon daemon uses an incorrect polling interval, it may cause problems. The polling interval is chosen when the hardmon daemon is started by the System Resource Controller. The value is in the **cmdargs** attribute in the hardmon ODM **SRCsubsys** object. Check the polling interval by issuing the ODM command:

```
odmget -q subsysname=hardmon SRCsubsys
```

Output is similar to the following, which is the default:

```
SRCsubsys:
        subsysname = "hardmon"
        synonym = ""
        cmdargs = "-r 5"
        path = "/usr/lpp/ssp/bin/hardmon"
        uid = 0
        auditid = 0
        standin = "/dev/console"
        standout = "/dev/console"
        standerr = "/dev/console"
        action = 1
        multi = 0
        contact = 2
        svrkey = 0
        svrmtype = 0
        priority = 20
        signorm = 15
        sigforce = 15
        display = 1
        waittime = 15
        grpname = ""
```

If the cmdargs attribute is not "-r 5", correct it issuing the command:

```
chssys -s hardmon -a "-r 5"
```

Then, reissue the **odmget** command to verify that the new cmdargs attribute is "-r 5" and run “Operational test 1 - Check that hardmon Is active” on page 327 to verify that the problem is resolved.

Action 3 - Investigate the hardmon daemon

Possible causes of this problem, and corrective actions are:

1. The **hmthresholds** file may not contain an entry for a type of hardware in the system, or there is a format error in the file.

   Repair the **/spdata/sys1/spmon/hmthresholds** file, referring to the comments located at the beginning of the file. If the file cannot be repaired, restore it from regular system backups or from the PSSP installation media.
2. Unable to open a log file.

   Other than a system error, the only reason that a log file would not be able to open is if the directory **/var/adm/SPlogs/spmon** does not exist. Verify that this directory exists, and create it if it does not.
3. SP Security Services has not been initialized properly by hardmon.

   Examine the file **/var/adm/SPlogs/spmon/hmlogfile.***ddd*, where *ddd* is the Julian date. Look for any error messages related to SP Security Services. Also, check the error log for error messages, by issuing the AIX **errpt** command. Take any action suggested to correct these errors.
4. There is an SDR configuration error.

   The correction is the same as for 3.
5. There is a system error. For example, a system call failed, a file descriptor was created that was larger than the allowed maximum size, or the system ran out of memory.

   The correction is the same as for 3.

Action 4 - Investigate System Monitor command problems

Possible causes of this problem, and corrective actions are:

1. The hardmon daemon is rejecting the command, because you are not authenticated to hardmon, or do not have the proper authorization for what the command is trying to do.

   If the problem is authentication or authorization, refer to “Chapter 18. Diagnosing SP Security Services problems” on page 251.
2. Some commands are partition sensitive. If the command takes slot numbers as a parameter, and one or more of the nodes are not in the current partition, and the ″global″ option (specified by the **-G** flag) is not used, the command will run as if that node does not exist. Also, for these commands, the ″global″ option must always be specified for frames and switches, because they do not reside in any system partition, including the current one.

   For example, the commands:

   ```
   hmmon -Q 1:0
   hmmon -Q 1:17
   ```

   produce an error message, since the **-G** flag was not specified, and 1:0 represents frame 1, and 1:17 represents the switch in frame 1.

Another example, if the current system partition is named PART1, and frame 3 node 8 is in the system partition named PART2, the command:

```
hmmon -Q 3:8
```

produces an error message, since the **-G** flag was not specified, and frame 3 node 8 is not in the current system partition.

In the case where the **-G** flag was not specified, refer to the entry for the particular command in *PSSP: Command and Technical Reference*.

Make sure that all objects in the SDR **Syspar_map** class reflect correct partitioning information. An error in one of these SDR objects can cause hardmon to be unable to locate nodes correctly.

3. If the command uses slot numbers as a parameter, and one or more of the nodes do not exist.

   Do not specify a target frame, node, or switch that does not exist in the system.
4. The rs232 tty cables to one or more frames are not connected correctly.

   Verify that the rs232 tty cables from the control workstation to the frame that is not responding are connected correctly. Note that the S70, S7A and S80 type server frames have two rs232 tty cables attached to the control workstation. All other frames have one rs232 tty cable attached to the control workstation.

   If the rs232 tty cables were not connected to the proper frames or servers, and you have already configured these frames using the **spframe** command, perform these steps:

   a. Issue the **spdelfram** command to delete the affected frames, **before** you re-cable the frames.
   b. Re-cable the rs232 tty cables to the proper frames.
   c. Add the frames that were deleted in step 4a. Issue the **spframe** command using the **-r yes** operand.
5. The hardmon daemon is rejecting the command because there is a frame ID mismatch, due to incorrect cabling. That is, the value of the **controllerIDMismatch** attribute for the frame is TRUE.

   Run the corrective action for cause 4. If the problem is not resolved, check for a frame ID mismatch by issuing the command:

   ```
   hmmon -GQv controllerIDMismatch F:0
   ```

   where *F* is the number of the frame on which the command is not working.

   If the value is TRUE, this means that the supervisor of this frame believes that it is attached to a frame other than the one it is physically attached to. To correct this, issue the command:

   ```
   hmcmds -G setid F:0
   ```

   where *F* is the number of the frame on which the command is not working.

   To verify the correction, wait 5 seconds, reissue the **hmmon** command and verify that the value of **controllerIDMismatch** is FALSE.
6. The **hmreinit** command was run, and it hangs.

   Use the **kill** command on the **hmreinit** process, and then reissue the **hmreinit** command.

Action 5 - Check paging space and CPU utilization

Possible causes of this problem, and corrective actions are:

1. Paging space is too low.

   Check that the paging space is adequate and adjust it if necessary.

2. Other processes are consuming CPU resources. For example, if you are using your control workstation as a boot server, the NFS daemons may be using most of the processor time.

   Use the **vmstat** command to check the overall CPU utilization.

   Check the CPU utilization of the hardmon and logging daemons. One method is to issue these commands:

   a. `ps gvc | grep hardmon`
   b. `ps gvc | grep splogd`

   If the CPU utilization rate is very high and this cannot be attributed to the hardmon or logging daemon, look for other processes which are consuming the CPU resources.

Chapter 22. Diagnosing SP Logging daemon Problems

This chapter discusses diagnostic procedures and failure responses for the Logging daemon component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 348. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 336.

Related documentation

The following publication provides information about the Logging daemon:

- *PSSP: Command and Technical Reference*

  Entries for these commands:

  – splogd Daemon

  – setup_logd

  – hwevents file

Requisite function

This is a list of the software directly used by the Logging daemon component of PSSP. Problems within the requisite software may manifest themselves as error symptoms in the Logging daemon. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the Logging daemon component of PSSP, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- The hardmon daemon must be running on the control workstation.
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
- AIX System Resource Controller (SRC).
- The **/var** file system is required on the control workstation.
- The **/spdata** file system is required on the control workstation.

Error information

AIX Error Log

This is the standard AIX error log located on the control workstation. Use **errpt -a** to display entries. These are hardware errors. An example would be an entry that indicates: ″Power module - DC power loss.″ Entries are logged when the condition appears to be permanent.

The error information can be cleared. Do **not** clear it if an error has occurred. The Detail Data in the error log entries comes from catalogs, and will appear in the language specified by the SP administrative locale on the control workstation. This table lists the Logging daemon error log entries.

Table 56. AIX Error Log templates for the Logging daemon

Label	Error ID	Type	Class	Description
SPLOGD01_ERR	E2ADA7BF	UNKN	S	**Explanation: splogd** input file error. **Details:** An error occurred opening or while processing the **splogd** input file.
SPMON_EMSG100_ER	4CEF5A08	PERM	H	**Explanation:** Link error. **Details:** The Logging daemon has detected that a frame is not responding to the System Monitor daemon (hardmon) .
SPMON_EMSG101_ER	A1843F1E	PERM	H	**Explanation:** Communication error. **Details:** The Logging daemon has detected that a frame is not responding, for a specific node, to the System Monitor daemon (hardmon).
SPMON_EMSG102_ER	001BB5DD	PERM	H	**Explanation:** Communications subsystem failure. **Details:** The Logging daemon has detected a problem with a frame supervisor bus, or a supervisor card may not be seated properly.
SPMON_EMSG103_EM	8D9F2E66	PEND	H	**Explanation:** Equipment malfunction. **Details:** The Logging daemon has detected a problem somewhere in the hardware. For example, this could be a failure of a fan, or a voltage or temperature going out of range.
SPMON_EMSG104_EM	76A4FAD9	PEND	H	**Explanation:** Impending workstation subsystem failure. **Details:** The Logging daemon has detected a hardware problem. If the problem persists, it will cause the supervisor subsystem to shut down the node.
SPMON_EMSG106_ER	E406336B	PERM	H	**Explanation:** Resources not active. **Details:** The Logging daemon has detected that a power module is not installed.
SPMON_EMSG107_ER	F708903E	PERM	H	**Explanation:** Loss of electrical power. **Details:** The Logging daemon has detected that a power module has failed.
SPMON_INFO100_TR	E720BFB5	UNKN	H	**Explanation:** Power Off detected. **Details:** The Logging daemon has detected that power has been turned off for a frame, node or switch.
SPMON_INFO101_TR	3E6F3CE7	UNKN	H	**Explanation:** Power On detected. **Details:** The Logging daemon has detected that power has been turned on for a frame, node or switch.

Table 56. AIX Error Log templates for the Logging daemon (continued)

Label	Error ID	Type	Class	Description
SPMON_INFO102_TR	93FA22BC	UNKN	H	**Explanation:** Mux value set. **Details:** The Logging daemon has detected that the mux value of a switch has been set.
SPMON_INFO103_TR	0D1620A8	UNKN	H	**Explanation:** Threshold has been exceeded. **Details:** The Logging daemon has detected that a threshold value has gone out of range. The value could represent a voltage, current, or temperature.
SPMON_INFO104_TR	E91A5929	TEMP	H	**Explanation:** An error condition has been resolved. **Details:** The Logging daemon has previously detected a problem. However, the problem has subsequently been cleared. For example, a voltage that has exceeded it's threshold value is now within it's threshold value.

hwevents file

Examine the **hwevents** file at the time of the failure. To determine where the **hwevents** file is located, issue the command:

```
odmget -q 'subsysname=splogd' SRCsubsys
```

If the cmdargs entry of the output contains the ″**-f**″ flag, the location of the **hwevents** file follows this flag. If the cmdargs entry of the output does not contain the ″**-f**″ flag, the **hwevents** file is in directory **/spdata/sys1/spmon/**. Make a copy of **hwevents** in case the original is subsequently modified.

splogd.debug file

This file is named **/var/adm/SPlogs/spmon/splogd.debug**. If the **debug** option was specified by the **splogd** command, the **splogd.debug** file contains useful information for analyzing problems. If the **debug** option was not specified, and the problem can be re-created, do the following in order to obtain the **splogd.debug** file:

1. Determine if the **debug** option was specified by **splogd**. Issue the command:

   ```
   odmget -q 'subsysname=splogd' SRCsubsys
   ```
2. If the cmdargs entry of the output contains the ″**-d**″ flag, the **splogd.debug** file should exist. Make a copy of **/var/adm/SPlogs/spmon/splogd.debug** in order not to lose the existing information in this file.
3. If the cmdargs entry of the output does not contain the ″**-d**″ flag, and you want to attempt to re-create the problem with debugging enabled, do the following:
   a. From the output of the previous **odmget** command, if the cmdargs entry of the output contains the ″**-f**″ flag, cut and paste the argument that follows the ″**-f**″ flag, and issue the command:

      ```
      chssys -s splogd -a "-d -f HWEVENTS_PATH"
      ```

where *HWEVENTS_PATH* is the argument that followed the ″**-f**″ flag.

For example, if what you cut and paste was **/spdata/sys1/spmon/hwevents**, issue the command:

```
 chssys -s splogd -a "-d -f /spdata/sys1/spmon/hwevents"
```

b. From the output of the previous **odmget** command, if the `cmdargs` entry of the output does not contain the ″**-f**″ flag, issue the command:

```
chssys -s splogd -a "-d"
```

4. Verify that the ″**-d**″ flag is now enabled, by again issuing the command:

```
odmget -q 'subsysname=splogd' SRCsubsys
```

5. Stop and start the Logging daemon, in an attempt to re-create the problem with the **debug** option enabled. Issue these commands:
   a. **stopsrc -s splogd**
   b. **startsrc -s splogd**

Now, the **splogd.debug** file should exist. If you are able to re-create the problem, then after the problem occurs make a copy of **/var/adm/SPlogs/spmon/splogd.debug** in order to preserve the information in this file. To remove the **debug** option:

1. Issue the last **chssys** command without the ″**-d**″ flag.
2. Issue the command: **stopsrc -s splogd**.
3. Issue the command: **startsrc -s splogd**.

Dump information

If a core dump exists, it is located in: **/var/adm/SPlogs/spmon/splogd/core**. This is a standard AIX core dump, containing the contents of the **splogd** process at time of the dump. The file **/var/adm/SPlogs/spmon/splogd/core** at the time of the failure should be saved in a safe place, where it will not be overwritten.

Information to collect before contacting the IBM Support Center

1. AIX error log. On the control workstation, issue the command:

```
LANG=C errpt -a > /tmp/AIXerrlog
```

2. **hwevents** file at the time of the failure. See “hwevents file” on page 335.
3. SP Logging daemon log file. See “splogd.debug file” on page 335.
4. SP Logging daemon core dump, if one exists. See “Dump information”.

Diagnostic procedures

Diagnostic procedures for the Logging daemon consist of installation verification tests, configuration verification tests and operational verification tests.

Installation verification tests

Use these tests to check that the Logging daemon has been properly installed.

Installation test 1 - Verify splogd and setup_logd files

This test verifies that the **splogd** and **setup_logd** files have been installed in directory **/usr/lpp/ssp/bin/** and that they are executable. On the control workstation, issue the commands:

1. `ls -l /usr/lpp/ssp/bin/splogd`
2. `ls -l /usr/lpp/ssp/bin/setup_logd`

Verify that these two files exist and are executable.

Good results are indicated if both files exist and are executable. Proceed to “Installation test 2 - Verify hwevents file”.

Error results are indicated in all other cases. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Installation test 2 - Verify hwevents file

Verify that the **hwevents** file is in the proper location. Refer to “hwevents file” on page 335, which describes how to determine the file’s location.

Good results are indicated if the file exists in the proper location. Proceed to “Configuration test 1 - Check for state logging”.

Error results are indicated in all other cases. Create an **hwevents** file in the proper location, as described in “hwevents file” on page 335. You can find a sample file in ″RS/6000 SP Files and Other Technical Information″ of *PSSP: Command and Technical Reference*. Modify the file as appropriate. Repeat this test. If the problem persists, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Configuration verification tests

Use these tests to check that the Logging daemon has been properly configured.

Configuration test 1 - Check for state logging

If you expected **splogd** to create file **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp*, but it was not created, verify that the **hwevents** file contains the proper line to indicate enabling of state logging. Otherwise, proceed to “Configuration test 2 - Check /etc/syslog.conf and /var/adm/SPlogs/SPdaemon.log files” on page 338.

Locate the **hwevents** file that is currently being used by **splogd**. Refer to “hwevents file” on page 335, which describes where the **hwevents** file is located. Verify that there exists a line in this file of the following format:

```
  *     *    *          =      *    c   SP_STATE_LOG
```

Note: The seven fields must be separated by white space. The line must **not** begin with a ’#’ character, because it will be ignored as a comment.

Good results are indicated if the **hwevents** file contains the uncommented line described. Proceed to “Configuration test 2 - Check /etc/syslog.conf and /var/adm/SPlogs/SPdaemon.log files” on page 338.

Error results are indicated in all other cases. Add the line to the **hwevents** file, and proceed to “Configuration test 2 - Check /etc/syslog.conf and /var/adm/SPlogs/SPdaemon.log files” on page 338.

Configuration test 2 - Check /etc/syslog.conf and /var/adm/SPlogs/SPdaemon.log files

If you expected **splogd** to write to the syslog, but it did not, verify the contents of file **/etc/syslog.conf**. Also, verify the existence of file **/var/adm/SPlogs/SPdaemon.log**. Otherwise, go to “Configuration test 3 - Check hwevents file”.

In file **/etc/syslog.conf** verify that there exists one of the following three lines:

```
daemon.notice   /var/adm/SPlogs/SPdaemon.log
daemon.debug    /var/adm/SPlogs/SPdaemon.log
daemon.info     /var/adm/SPlogs/SPdaemon.log
```

Verify that file **/var/adm/SPlogs/SPdaemon.log** exists, and that its permissions are set to 644.

Good results are indicated if the **/etc/syslog.conf** file contains one of the three lines described, and the **SPdaemon.log** file exists with permissions 644. Proceed to “Configuration test 3 - Check hwevents file”.

Error results are indicated in all other cases. Perform these steps:

1. Delete the following lines, if any of them exist, from the **/etc/syslog.conf** file. Do **not** simply comment these lines out. They **must** be deleted from the file for this step.

   ```
   daemon.notice        /var/adm/SPlogs/SPdaemon.log
   daemon.debug         /var/adm/SPlogs/SPdaemon.log
   daemon.info          /var/adm/SPlogs/SPdaemon.log
   ```
2. Issue the command:

   ```
   setup_logd
   ```
3. Repeat this test. If the problem persists, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Configuration test 3 - Check hwevents file

If you expected **splogd** to write error information to the AIX error log, but it did not, verify that the **hwevents** file contains the proper line to enable error logging. Otherwise, proceed to “Configuration test 5 - Check sser exits” on page 339.

Locate the **hwevents** file that is currently being used by **splogd**. Refer to “hwevents file” on page 335, which describes the location of the **hwevents** file. Verify that there exists a line in this file of the following format:

```
  *      *       *                  =      *     b    SP_ERROR_LOG
```

Note: The seven fields must be separated by white space. The line must **not** begin with a ’#’ character, because it will be ignored as a comment.

Good results are indicated if the **hwevents** file contains the uncommented line described here. Proceed to “Configuration test 4 - Check SPMON error record templates” on page 339.

Error results are indicated in all other cases. Add the line to the **hwevents** file and proceed to “Configuration test 4 - Check SPMON error record templates” on page 339.

Configuration test 4 - Check SPMON error record templates

This test verifies that the error record templates for SPMON have been defined. Issue the command:

```
errpt -t | grep SPMON_
```

The output is similar to:

```
001BB5DD SPMON_EMSG102_ER   PERM H  COMMUNICATIONS SUBSYSTEM FAILURE
0D1620A8 SPMON_INFO103_TR   UNKN H  THRESHOLD HAS BEEN EXCEEDED
3E6F3CE7 SPMON_INFO101_TR   UNKN H  ELECTRICAL POWER RESUMED
4CEF5A08 SPMON_EMSG100_ER   PERM H  LINK ERROR
6469E2D8 SPMON_EMSG108_ER   PERM S  UNABLE TO COMMUNICATE WITH REMOTE NODE
76A4FAD9 SPMON_EMSG104_EM   PEND H  IMPENDING WORKSTATION SUBSYSTEM FAILURE
8D9F2E66 SPMON_EMSG103_EM   PEND H  EQUIPMENT MALFUNCTION
93FA22BC SPMON_INFO102_TR   UNKN H  SOFTWARE
A1843F1E SPMON_EMSG101_ER   PERM H  UNABLE TO COMMUNICATE WITH REMOTE NODE
E406336B SPMON_EMSG106_ER   PERM H  RESOURCES NOT ACTIVE
E720BFB5 SPMON_INFO100_TR   UNKN H  POWER OFF DETECTED
E91A5929 SPMON_INFO104_TR   TEMP H  PROBLEM RESOLVED
F708903E SPMON_EMSG107_ER   PERM H  LOSS OF ELECTRICAL POWER
```

Good results are indicated if the output is similar to the example output. That is, there are about 13 lines where the data in the second column begins with `SPMON_`. You may not see the same number of these lines, depending on the software level. Proceed to “Configuration test 5 - Check sser exits”.

Error results are indicated in all other cases. Determine why error record templates for SPMON have not been defined. You may need to contact the IBM Support Center.

Configuration test 5 - Check sser exits

If you expected **splogd** to run user exits, but they are not running, verify that the **hwevents** file contains the proper line for the user exit. Otherwise, proceed to “Operational test 1 - Check System Monitor daemon” on page 340.

Locate the **hwevents** file that is currently being used by **splogd**. Refer to “hwevents file” on page 335, which describes the location of the **hwevents** file. Verify that there exists a line in this file, for the user exit you wish to run, as described in the **hwevents** entry of ″RS/6000 SP Files and Other Technical Information″ in *PSSP: Command and Technical Reference*.

Good results are indicated if the **hwevents** file contains the proper line. “Operational test 5 - Check user exit mechanism” on page 341 will verify that the mechanism for running user exits is working. Proceed to “Operational test 1 - Check System Monitor daemon” on page 340.

Error results are indicated in all other cases. Add the appropriate line to the **hwevents** file. Refer to the **hwevents** file entry of ″RS/6000 SP Files and Other Technical Information″ in *PSSP: Command and Technical Reference*, for information about the format of the user exit line.

Operational verification tests

Use these tests to check that the Logging daemon is operating properly.

Operational test 1 - Check System Monitor daemon

This test verifies if the hardware monitor daemon is active, that is, if the **hardmon** daemon is currently running. The Logging daemon will not operate properly if **hardmon** is not running. Issue the command:

```
 lssrc -s hardmon
```

Look for the entry whose `Subsystem` is `hardmon`.

Good results are indicated if the `Status` column is `active`. Proceed to “Operational test 2 – Check for frames responding”.

Error results are indicated in all other cases. Issue the command:

```
startsrc -s hardmon
```

to start the hardware monitor daemon. Repeat this test, after taking the suggested action. If this test still fails, refer to “Chapter 21. Diagnosing System Monitor problems” on page 317 to determine why **hardmon** will not run.

Operational test 2 – Check for frames responding

This test verifies if the frames are responding. The Logging daemon will not operate properly if the frames are not responding. To verify that a particular frame is responding, issue the command:

```
hmmon -GQv controllerResponds F:0
```

where *F* is the frame number you are checking. If there is more than one frame, you should check all frames. For example, if you have frames 1 through 5, you would issue the command:

```
hmmon -GQv controllerResponds 1-5:0
```

The output is similar to the following:

```
  frame F, slot 00:
     TRUE  frame responding to polls
```

Good results are indicated if the value is `TRUE` for each frame. Proceed to “Operational test 3 - Check Logging daemon status”.

Error results are indicated if the value is `FALSE`, or you do not receive any output. Refer to “Chapter 21. Diagnosing System Monitor problems” on page 317.

Operational test 3 - Check Logging daemon status

This test verifies if the Logging daemon is active, that is, if the **splogd** daemon is currently running. Issue the command:

```
lssrc -s splogd
```

Look for the entry whose `Subsystem` is **splogd**.

Good results are indicated If the `Status` column is `active`. Proceed to “Operational test 4 - Check error daemon status” on page 341.

Error results are indicated in all other cases. Issue the command:

```
startsrc -s splogd
```

Repeat this test. If the problem persists, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Operational test 4 - Check error daemon status

This test verifies that the error daemon is running. Perform this test if entries that you expect to receive, are not written to the error log. If you choose to skip this test, because you are not having a problem with error log entries, proceed to “Operational test 5 - Check user exit mechanism”.

Issue the command:

```
ps -ef | grep errdemon
```

The output is similar to:

```
root  3206     1   0   Jun 19      -  0:40 /usr/lib/errdemon
```

Good results are indicated by output showing that the **errdemon** is running. Proceed to “Operational test 5 - Check user exit mechanism”.

Error results are indicated in all other cases. Determine why the **errdemon** is not running. Start the **errdemon** by issuing the command:

```
/usr/lib/errdemon
```

Repeat this test. If the problem persists, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Operational test 5 - Check user exit mechanism

This test verifies that the mechanism used by **splogd**, to run user exits specified in the **hwevents** file, is working correctly. If you choose to skip this test because you are not having a problem with running user exits, proceed to “Operational test 6 - Check writing to syslog file” on page 343.

Notes:

1. This test cannot be run if any of the following are true:
   - **hardmon** is not active.
   - You cannot find a node with an attribute of `serialLinkOpen` with a value of `FALSE` in step 5 on page 342.
   - In step 9 on page 342, the `Status` of **splogd** is not active.
2. This test assumes that the **hwevents** file is in the **/spdata/sys1/spmon/** directory. If the file is in another directory, you may still follow these instructions, but substitute the actual path to the location of your **hwevents** file. Refer to “hwevents file” on page 335, which describes how to determine the location of the **hwevents** file.

Perform these steps:

1. Create a trivial program in C, Perl, or another language that simply creates a file named **/tmp/splogd_test_output**, and writes some text to it. Name the program **/tmp/splogd_test**. This assumes that these two files do not already exist. If they do, substitute your own file names in this procedure.

2. Verify that running **/tmp/splogd_test** creates file **/tmp/splogd_test_output**.
3. Remove file **splogd_test_output**.
4. Copy the **/spdata/sys1/spmon/hwevents** file to a backup copy, so that the original **hwevents** file can be restored at the end of this test.
5. Run the **hmmon** command to find a node that has the `serialLinkOpen` attribute and whose current value of this attribute is `FALSE`.

   For example, if the system contains frames 1 through 5, run the command:

   ```
   hmmon -GQs 1-5:
   ```

   to find one. Choose a node in a frame that fits this criteria.

   In subsequent steps, the frame number will be referred to as *F* and the node number as *N*. Always substitute the numeric frame number for the letter *F* and the numeric node number for the letter *N*.
6. Edit file **/spdata/sys1/spmon/hwevents** and place the '#' character in the leftmost column of every line in the file. This makes every line a comment. Add the following line to this file, in any location:

   ```
     F  N  serialLinkOpen  =  *  c  /tmp/splogd_test
   ```

   where *F* and *N* are the frame number and node number from step 5.

   For example, if frame number is 3 and node number is 12, the line is:

   ```
   3   12   serialLinkOpen  =  *  c  /tmp/splogd_test
   ```

   Save this file.
7. Issue this command to stop **splogd**:

   ```
   stopsrc -s splogd
   ```

   The output is a message similar to:

   ```
   The stop of the splogd Subsystem was completed successfully.
   ```
8. Issue this command to start **splogd**:

   ```
   startsrc -s splogd
   ```

   You should get a message similar to:

   ```
   The splogd Subsystem has been started. Subsystem PID is nnnnn.
   ```
9. Issue this command:

   ```
   lssrc -s splogd
   ```

   and verify that `Status` column indicates `active`.
10. Verify that the file **/tmp/splogd_test_output** has not yet been created.
11. Issue the following command to open up an s1term:

    ```
    s1term -G F N
    ```

    where *F* and *N* are the frame number and node number from step 5.
12. Wait 10 seconds and, in another window, issue the command:

```
hmmon -GQs F:N | grep serialLinkOpen
```

where *F* and *N* are the frame number and node number from step 5 on page 342.

The value of the `serialLinkOpen` attribute should now be `TRUE`.

13. Close the s1term (opened in step 11 on page 342), by typing the terminal interrupt key, which is usually **<Ctrl-c>**.
14. Verify that the file **/tmp/splogd_test_output** has been created.
15. Copy the backup copy of the **hwevents** file that you created in step 4 on page 342, to **/spdata/sys1/spmon/hwevents**, to restore this file.

Good results are indicated if the file **/tmp/splogd_test_output** did not exist when checked in step 10 on page 342, but existed when checked in step 14. Proceed to "Operational test 6 - Check writing to syslog file".

Error results are indicated in all other cases. Perform these steps:

1. Delete the following lines, if any exist, from the **/etc/syslog.conf file**. Do **not** simply comment these lines out. They **must** be deleted from the file for this step.

   ```
           daemon.notice        /var/adm/SPlogs/SPdaemon.log
           daemon.debug         /var/adm/SPlogs/SPdaemon.log
           daemon.info          /var/adm/SPlogs/SPdaemon.log
   ```
2. Issue the command:

   ```
   setup_logd
   ```
3. Repeat this test. If the problem persists, record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 336, and contact the IBM Support Center.

Operational test 6 - Check writing to syslog file

This test verifies that the mechanism used by **splogd**, to write to the syslog (if specified in the **hwevents** file) is working correctly. Perform this test only if there is a node that is currently powered on, for which it is safe to power the node off and then back on. If you choose to skip this test, because you are not having a problem with syslog entries, or you are not able to power off and on a node, proceed to "Operational test 7 - Check writing to the AIX Error Log" on page 345.

Notes:

1. This test cannot be run if any of the following are true:
   - **hardmon** is not active.
   - You cannot find a node that you can safely power off and on.
   - In step 4 on page 344, the `Status` of **splogd** is not `active`.
2. This test assumes that the **hwevents** file is in the **/spdata/sys1/spmon/** directory. If the file is in another directory, you may still follow these instructions, but substitute the actual path to the location of your **hwevents** file. Refer to "hwevents file" on page 335, which describes how to determine the location of the **hwevents** file.

Perform these steps:

1. Edit file **/spdata/sys1/spmon/hwevents** and add the following line, if it does not already exist. Make sure this line is not commented out (# in leftmost column).

   ```
       *    *    *    =    *    b    SP_ERROR_LOG
   ```

2. Copy the **/etc/syslog.conf** file to a backup copy, so that it can be restored at the end of this test.
3. Modify the original **/etc/syslog.conf** file so that all lines are commented out (# in leftmost column), and add the following line:

   ```
     daemon.info           /var/adm/SPlogs/SPdaemon.log
   ```

4. Issue this command to stop **splogd**:

   ```
   stopsrc -s splogd
   ```

   The output is a message similar to:

   ```
   The stop of the splogd Subsystem was completed successfully.
   ```

5. Issue this command to start **splogd**:

   ```
   startsrc -s splogd
   ```

   You should get a message similar to:

   ```
   The splogd Subsystem has been started. Subsystem PID is nnnnn.
   ```

6. Issue the following command:

   ```
   lssrc -s splogd
   ```

   and verify that `Status` column indicates `active` for Subsystem **splogd**.
7. Issue the following commands to stop and start the **syslogd** daemon, so that it rereads the **syslog.conf** file:

   ```
   stopsrc  -s syslogd
   startsrc -s syslogd
   ```

8. Choose a node *N*, in a frame *F*, that is currently powered on, which you can safely power off and then back on. Power off and then power on this node by issuing the commands:

   ```
   hmcmds -Gv off  F:N
   hmcmds -Gv on   F:N
   ```

   where *F* is the frame number, and *N* is the node number.
9. Copy the backup copy of the **syslog.conf** file that you created in step 2, to **/etc/syslog.conf**, to restore this file.
10. Issue the following commands to stop and start the **syslogd** daemon, so that it rereads the original **syslog.conf** file:

    ```
    stopsrc  -s syslogd
    startsrc -s syslogd
    ```

11. If you modified the original **hwevents** file in step 1, you may want to comment out the line that you had added or uncommented, and repeat steps 4 and 5.

Good results are indicated if the file **/var/adm/SPlogs/SPdaemon.log** contains a line, with the **current** date and time. Make sure that the date and time is current. The line is similar to:

```
Sep  2 12:42:05 sup1 sphwlog[19826]: LP=PSSP,Fn=splogd.c,SID=I,L#=1339,
Information; Node F:N; powerLED; Power is off.
```

Proceed to “Operational test 7 - Check writing to the AIX Error Log”.

Error results are indicated in all other cases. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336 and contact the IBM Support Center.

Operational test 7 - Check writing to the AIX Error Log

This test verifies that the mechanism used by **splogd**, to write to the error log (if specified in the **hwevents** file), is working correctly. Perform this test only if there is a node that is currently powered on, for which it is safe to power the node off and then back on. If you choose to skip this test, because you are not having a problem with error log entries, or you are not able to power off and on a node, proceed to “Operational test 8 - Check writing to splogd.state_changes file” on page 346.

Notes:

1. This test cannot be run if any of the following are true:
   - **hardmon** is not active.
   - You cannot find a node that you can safely power off and on.
   - In step 4, the `Status` of **splogd** is not `active`.
2. This test assumes that the **hwevents** file is in the **/spdata/sys1/spmon/** directory. If the file is in another directory, you may still follow these instructions, but substitute the actual path to the location of your **hwevents** file. Refer to “hwevents file” on page 335, which describes how to determine the location of the **hwevents** file.

Perform these steps:

1. Edit file **/spdata/sys1/spmon/hwevents** and add the following line, if it does not already exist. Make sure this line is not comments out (# in leftmost column).

   ```
     *    *    *    =    *    b    SP_ERROR_LOG
   ```
2. Issue this command to stop **splogd**:

   ```
   stopsrc -s splogd
   ```

   The output is a message similar to:

   ```
   The stop of the splogd Subsystem was completed successfully.
   ```
3. Issue this command to start **splogd**:

   ```
   startsrc -s splogd
   ```

   You should get a message similar to:

   ```
   The splogd Subsystem has been started. Subsystem PID is nnnnn.
   ```
4. Issue the following command:

   ```
   lssrc -s splogd
   ```

Verify that the `status` column has `active` for subsystem **splogd**.

5. Choose a node *N*, in a frame *F*, that is currently powered on and which you can safely power off and back on. Power off and then power on this node, by issuing the commands:

   ```
   hmcmds -Gv off F:N
   hmcmds -Gv on F:N
   ```

   where *F* is the frame number, and *N* is the node number.
6. If you modified the original **hwevents** file in 1 on page 345, you may want to comment out the line that you had added or uncommented, and repeat steps 2 on page 345 and 3 on page 345.

Good results are indicated if you see, in the error log, these two entries and they have just been made (meaning since you performed 5), and they contain timestamps:

```
LABEL:  SPMON_INFO100_TR
LABEL:  SPMON_INFO101_TR
```

The SPMON_INFO100_TR entry indicated that Power Off was detected. The entry SPMON_INFO101_TR indicated that Power On was detected. You can view the error log by running the command:

```
errpt -a
```

You should redirect the output to a file. Proceed to “Operational test 8 - Check writing to splogd.state_changes file”.

Error results are indicated in all other cases. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 336, and contact the IBM Support Center.

Operational test 8 - Check writing to splogd.state_changes file

This test verifies that the mechanism used by **splogd**, to write to the **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp* file, if specified by the **hwevents** file, is working correctly. If you choose to skip this test, because you are not having a problem with the writing to the **splogd.state_changes.***timestamp* file, you are finished with Logging daemon tests.

Notes:

1. This test cannot be run if any of the following are true:
   - **hardmon** is not active.
   - You cannot find a node with an attribute of `serialLinkOpen` with a value of `FALSE` in step 2 on page 347.
   - In step 7 on page 347, the `Status` of **splogd** is not `active`.
2. This test assumes that the **hwevents** file is in the **/spdata/sys1/spmon/** directory. If the file is in another directory, you may still follow these instructions, but substitute the actual path to the location of your **hwevents** file. Refer to “hwevents file” on page 335, which describes how to determine the location of the **hwevents** file.

Perform these steps:

1. Copy the **/spdata/sys1/spmon/hwevents** file to a backup copy, so that the **hwevents** file can be restored at the end of this test.

2. Run the **hmmon** command to find a node that has the `serialLinkOpen` attribute and whose current value of this attribute is `FALSE`.

   For example, if the system contains frames 1 through 5, issue the command **hmmon -GQs 1-5:** to find one. Choose a node in a frame that fits this criteria.

   In subsequent steps, the frame number will be referred to as *F* and the node number as *N*. Always substitute the numeric frame number for the letter *F* and the numeric node number for the letter *N*.
3. Edit file **/spdata/sys1/spmon/hwevents** and place the '#' character in the leftmost column of every line in the file. This makes every line a comment. Add the following line to this file in any location:

   ```
      *      *     *           =       *     c    SP_STATE_LOG
   ```

   Save this file.
4. Copy file **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp* to a backup file, and then delete **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp*. The *timestamp* part of this file name is numeric, related to the current date.
5. Issue the following command to stop **splogd**:

   ```
   stopsrc -s splogd
   ```

   You should receive a message similar to:

   ```
   The stop of the splogd Subsystem was completed successfully.
   ```
6. Issue the following command to start **splogd**:

   ```
   startsrc -s splogd
   ```

   You should receive a message similar to:

   ```
   The splogd Subsystem has been started. Subsystem PID is nnnnn.
   ```
7. Issue the following command:

   ```
   lssrc -s splogd
   ```

   Verify that the `Status` column indicates `active`.
8. Issue the following command to open up an s1term:

   ```
   s1term -G F N
   ```

   where *F* and *N* are the frame number and node number from step 2.
9. Wait 10 seconds and, in another window, issue the command:

   ```
   hmmon -GQs F:N | grep serialLinkOpen
   ```

   where *F* and *N* are the frame number and node number from step 2.

   The value of the `serialLinkOpen` attribute should now be `TRUE`.
10. Close the s1term (opened in step 8), by typing the terminal interrupt key, which is usually **<Ctrl-c>**.
11. Verify that the file **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp* has been created, and the state change is present in the file. The *timestamp*

part of the file name will actually appear as numbers, related to the current date. You should see lines similar to the following inside the file:

```
F N serialLinkOpen TRUE
F N serialLinkOpen FALSE
```

12. Copy the backup copy of the **hwevents** file that you created in step 1 on page 346, to **/spdata/sys1/spmon/hwevents** to restore this file.

Good results are indicated if the file **/var/adm/SPlogs/spmon/splogd.state_changes.***timestamp* was created, and contained the state change, when checked in step 11 on page 347. You are finished with SP Logging Daemon testing.

Error results are indicated in all other cases. Record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 336 and contact the IBM Support Center.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the Logging daemon component of PSSP. Locate the symptom and perform the action described in this table.

Table 57. Logging daemon symptoms

Symptom	Recovery
The Logging daemon dies.	See "Action 1 - Run the setup_logd command".
The Logging daemon is not running.	See "Action 1 - Run the setup_logd command".
The Logging daemon is not writing to the syslog.	See "Action 2 - Ensure that the Logging daemon writes to syslog file" on page 349.
The Logging daemon is not writing to the AIX Error Log.	See "Action 3 - Ensure that the Logging daemon writes to the AIX Error Log" on page 349.
The Logging daemon is not writing state changes to the **splogd.state_changes.***timestamp* file.	See "Action 4 - Ensure that the Logging daemon writes to splogd.state_changes file" on page 349.
The Logging daemon is not running the user exits.	See "Action 5 - Correct Logging daemon user exit mechanism" on page 349.

Actions

Action 1 - Run the setup_logd command

Perform test "Installation test 1 - Verify splogd and setup_logd files" on page 337.

On the control workstation, check directory **/var/adm/SPlogs/spmon/splogd** to see if a core dump was created. If so, save it, see "Information to collect before contacting the IBM Support Center" on page 336 and contact the IBM Support Center.

Whether or not there is a core file, attempt to restart the Logging daemon:

1. Delete the following lines, if any exist, from the **/etc/syslog.conf** file. Do **not** simply comment these lines out. They **must** be deleted from the file for this repair action.

```
daemon.notice        /var/adm/SPlogs/SPdaemon.log
daemon.debug         /var/adm/SPlogs/SPdaemon.log
daemon.info          /var/adm/SPlogs/SPdaemon.log
```

2. Issue the command:

```
setup_logd
```

To ensure that this action has succeeded, perform test “Operational test 3 - Check Logging daemon status” on page 340, which verifies that the Logging daemon is now running.

If you are attempting this action because the Logging daemon died, be aware that this action may only temporarily resolve the problem. If the original condition that caused the Logging daemon to die in the first place still exists, this condition may again cause the Logging daemon to die before or after performing test “Operational test 3 - Check Logging daemon status” on page 340.

Action 2 - Ensure that the Logging daemon writes to syslog file

Perform these tests:

1. “Installation test 2 - Verify hwevents file” on page 337
2. “Operational test 1 - Check System Monitor daemon” on page 340
3. “Operational test 2 – Check for frames responding” on page 340
4. “Configuration test 2 - Check /etc/syslog.conf and /var/adm/SPlogs/SPdaemon.log files” on page 338

To ensure that this action has succeeded, perform test “Operational test 6 - Check writing to syslog file” on page 343.

Action 3 - Ensure that the Logging daemon writes to the AIX Error Log

Perform these tests:

1. “Installation test 2 - Verify hwevents file” on page 337
2. “Operational test 1 - Check System Monitor daemon” on page 340
3. “Operational test 2 – Check for frames responding” on page 340
4. “Configuration test 3 - Check hwevents file” on page 338
5. “Configuration test 4 - Check SPMON error record templates” on page 339
6. “Operational test 4 - Check error daemon status” on page 341

To ensure that this action has succeeded, perform test “Operational test 7 - Check writing to the AIX Error Log” on page 345.

Action 4 - Ensure that the Logging daemon writes to splogd.state_changes file

Perform these tests:

1. “Installation test 2 - Verify hwevents file” on page 337
2. “Operational test 1 - Check System Monitor daemon” on page 340
3. “Operational test 2 – Check for frames responding” on page 340
4. “Configuration test 1 - Check for state logging” on page 337

To ensure that this action has succeeded, perform test “Operational test 8 - Check writing to splogd.state_changes file” on page 346.

Action 5 - Correct Logging daemon user exit mechanism

Perform these tests:

1. “Operational test 1 - Check System Monitor daemon” on page 340
2. “Operational test 2 – Check for frames responding” on page 340

3. “Installation test 2 - Verify hwevents file” on page 337
4. “Configuration test 5 - Check sser exits” on page 339

To ensure that this action has succeeded, perform test “Operational test 5 - Check user exit mechanism” on page 341.

Chapter 23. Diagnosing Topology Services problems

This chapter discusses diagnostic procedures and failure responses for the Topology Services component of RSCT. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 402. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 384.

Note that Topology Services is a subsystem of RS/6000 Cluster Technology (RSCT).

Related documentation

The following publications provide information about Topology Services:

1. *PSSP: Command and Technical Reference*

   This book provides detailed syntax and parameter information for the commands that are used to control the Topology Services subsystem. See entries for these commands:

   - hats
   - hatsctrl
   - hatsoptions
   - hatstune
2. *PSSP: Administration Guide*

   The chapter ″The Topology Services Subsystem″ provides information about the subsystem and its interaction with PSSP subsystems.
3. *PSSP: Messages Reference*

   These chapters contain messages related to Topology Services:

   - 2523 - Topology Services Messages
   - 2525 - RS/6000 Cluster Technology Common Messages
4. *HACMP Enhanced Scalability Handbook* SG24-5328
5. *HACMP Enhanced Scalability Installation and Administration Guide* SC23-4306-01.

   See the chapter ″Troubleshooting HACMP/ES Clusters″. This book provides information on HACMP/ES, including problem determination procedures. This book is useful in isolating problems that affect Topology Services when run in the HACMP/ES environment.
6. *RS/6000 High Availability Infrastructure* SG24-4838

   This book provides information on the RSCT infrastructure, including a presentation of how the internals of Topology Services work.
7. *AIX 5L Version 5.1 Differences Guide*, the chapter ″AIX Workload Manager″.

Requisite function

This is a list of the software directly used by the Topology Services component of RSCT. Problems within the requisite software may manifest themselves as error symptoms in Topology Services. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the Topology Services component of RSCT, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- AIX UDP/IP communication
- SP Switch or SP Switch2 (CSS subsystem) of PSSP
- System Data Repository (SDR) component of PSSP
- AIX UNIX Domain sockets
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
- AIX System Resource Controller (SRC)
- First Failure Data Capture (FFDC) library
- **/var/ha** directory

Error information

AIX Error Logs and templates

The error log file is stored in **/var/adm/ras/errlog** by default. One entry is logged for each occurrence of the condition. The condition is logged on every node where the event occurred.

The Error Log file may wrap, since the file has a limited size. Data is stored in a circular fashion. Also, the system is shipped with a crontab file to delete hardware errors more than 90 days old and software errors and operator messages more than 30 days old.

The command:

```
/usr/lib/errdemon -l
```

shows current settings for the error logging daemon.

The command:

```
/usr/lib/errdemon -s
```

is used to change the size of the error log file.

Both commands require **root** authority.

Unless otherwise noted, each entry refers to a particular instance of the Topology Services daemon on the local node. Unless otherwise noted, entries are created on each occurrence of the condition.

Table 58 on page 354 lists the error log templates used by Topology Services, sorted by **Error Label**. An **Explanation** and **Details** are given for each error.

The Topology Services subsystem creates AIX error log entries for the following conditions:

- TS_ASSERT_EM
- TS_AUTHMETH_ER

- TS_CMDFLAG_ER
- TS_CPU_USE_ER
- TS_CTIPDUP_ER
- TS_CTLOCAL_ER
- TS_CTNODEDUP_ER
- TS_CWSADDR_ER
- TS_DCECRED_ER
- TS_DEATH_TR
- TS_DMS_WARNING_ST
- TS_DUPNETNAME_ER
- TS_FD_INTFC_NAME_ST
- TS_FD_INVAL_ADDR_ST
- TS_HAIPDUP_ER
- TS_HALOCAL_ER
- TS_HANODEDUP_ER
- TS_IOCTL_ER
- TS_IPADDR_ER
- TS_KEYS_ER
- TS_LATEHB_PE
- TS_LIBERR_EM
- TS_LOC_DOWN_ST
- TS_LOGFILE_ER
- TS_LONGLINE_ER
- TS_LSOCK_ER
- TS_MACHLIST_ER
- TS_MIGRATE_ER
- TS_MISCFG_EM
- TS_NIM_DIED_ER
- TS_NIM_NETMON_ERROR_ER
- TS_NIM_OPEN_ERROR_ER
- TS_NODEDOWN_EM
- TS_NODENUM_ER
- TS_NODEUP_ST
- TS_OFF_LIMIT_ER
- TS_REFRESH_ER
- TS_RSOCK_ER
- TS_SDR_ER
- TS_SECMODE_ER
- TS_SECURITY_ST
- TS_SECURITY2_ST
- TS_SEMGET_ER
- TS_SERVICE_ER
- TS_SHMAT_ER
- TS_SHMEMKEY_ER
- TS_SHMGET_ER

- TS_SP_DIR_ER
- TS_SPIPDUP_ER
- TS_SPLOCAL_ER
- TS_SPNODEDUP_ER
- TS_START_ST
- TS_STOP_ST
- TS_SYSPAR_ER
- TS_THATTR_ER
- TS_THCREATE_ER
- TS_THREAD_STUCK_ER
- TS_UNS_SIN_TR

When you retrieve an error log entry, look for the Detail Data section near the bottom of the entry.

Table 58. AIX Error Log templates for Topology Services

Label and Error ID	Type	Description
TS_ASSERT_EM 82D819EF	PEND	**Explanation:** Topology Services daemon exited abnormally. **Details:** This entry indicates that the Topology Services daemon exited with an **assert** statement, resulting in a core dump being generated. Standard fields indicate that the Topology Services daemon exited abnormally. Detail Data fields contain the location of the **core** file. This is an internal error. Data needed for IBM Service to diagnose the problem is stored in the **core** file (whose location is given in the error log) and in the Topology Services daemon service log. See “Topology Services service log” on page 379. Since only six instances of the Topology Services daemon service log are kept, it should be copied to a safe place. Also, only three instances of the **core** file are kept. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_AUTHMETH_ER C1FDC4E7	PERM	**Explanation:** The Topology Services startup script cannot retrieve active authentication methods using command **/usr/sbin/rsct/bin/lsauthpts**. **Details:** This entry indicates that command **/usr/lpp/ssp/bin/lsauthpts**, run by the Topology Service startup script on the control workstation, was unable to retrieve the active authentication methods in a system partition. This error occurs when the startup script is running on the control workstation during initial startup or refresh. When this error occurs, all Topology Services daemons in the system partition will terminate their operations and exit. Diagnosing this problem requires collecting data only on the control workstation. Standard fields indicate that the startup script cannot retrieve active authentication methods in a system partition using command **lsauthpts**. The problem may be one of the following: • The system partition has an incorrect set of active partition methods. • The current system partition cannot be identified. Detail Data fields contain the return code of command **lsauthpts** and the location of the startup script log. The error message returned by command **lsauthpts** can be found in the startup script log. For more information about SP security services, see “Chapter 18. Diagnosing SP Security Services problems” on page 251.
TS_CMDFLAG_ER 979E20DB	PERM	**Explanation:** Topology Services cannot be started due to incorrect flags. **Details:** This entry indicates that the Topology Services daemon was unable to start because incorrect command line arguments were passed to it. This entry refers to a particular instance of Topology Services on the local node. Other nodes may have been affected by the same problem. Standard fields indicate that the daemon was unable to start because incorrect flags were passed to it. Detail Data fields show the pathname to the daemon user log, which contains more detail about the problem. This problem may be one of the following: • Topology Services was started manually in an incorrect way. • Incompatible versions of the daemon and startup script are being used. • The AIX SRC definition for the subsystem was manually set to an incorrect value. Information about the cause of the problem may not be available once the problem is cleared.
TS_CTIPDUP_ER	PERM	**Explanation:** See TS_HAIPDUP_ER.
TS_CTNODEDUP_ER	PERM	**Explanation:** See TS_HANODEDUP_ER.
TS_CTLOCAL_ER	PERM	**Explanation:** See TS_HALOCAL_ER.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	**Type**	**Description**
TS_CPU_USE_ER FD20FB81	PERM	**Explanation:** The Topology Services daemon is using too much CPU. The daemon will exit. **Details:** This entry indicates that the Topology Services daemon will exit because it has been using almost 100% of the CPU. Since Topology Services runs in a real time fixed priority, exiting in this case is necessary. Otherwise, all other applications in the node will be prevented from running, Also, it is likely that the daemon is not working properly if it is using all the CPU. A **core** dump is created to allow debugging the cause of the problem. This entry refers to a particular instance of Topology Services running on a node. The standard fields indicate that the Topology Services daemon is exiting because it is using too much of the CPU, and explains some of the possible causes. The detailed fields show the amount of CPU used by the daemon (in milliseconds) and the interval (in milliseconds) where the CPU usage occurred. Collect the data described in "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center. In particular, the daemon log file and the most recent core files should be collected.
TS_CWSADDR_ER A35F9C3B	PERM	**Explanation:** Topology Services cannot find the control workstation address. **Details:** This entry indicates that the Topology Services startup script was unable to choose a suitable Ethernet adapter on the control workstation to add to the **machines.lst** configuration file. This error occurs when the startup script is running on the control workstation The failure prevents the Topology Services subsystem from starting or refreshing on the control workstation and on the nodes. Diagnosing the problem requires collecting data only on the control workstation. Standard fields indicate that the startup script was unable to find the control workstation adapter to insert in the **machines.lst** file. The problem could be one of the following: • There is no Ethernet adapter on the control workstation that is on the same subnet as the en0 adapter on one of the nodes. • The netmask of the Ethernet adapters on the control workstation or the nodes is incorrect. • The Ethernet adapter on the control workstation which should belong to the (Topology Services) SP Ethernet adapter membership group is not configured correctly. Detailed information about the problem is stored in the startup script log file **/var/ha/log/hats.***partition_name* on PSSP. Only a limited number (currently seven) of copies of this log file is kept. Details about the problem are stored in the startup script log file. Messages in this log file are stored both in English and in the node's language. The following can be found in the startup script's log file: • The address and ″netmask″ for all Ethernet adapters on the control workstation. • The address and ″netmask″ of the en0 adapters in all the nodes. When examining this file, look for missing Ethernet adapters on the control workstation. Check also for the Ethernet adapter on the control workstation which is on the SP Ethernet, having a different ″netmask″ than the en0 on the nodes.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_DCECRED_ER 81132988	PERM	**Explanation:** Topology Services cannot obtain credentials to update the SDR. **Details:** This entry indicates that the Topology Services startup script was unable to obtain the DCE credentials it needs to write data into the SDR. The **hats** script failed to login as the service principal **ssp/spbgroot** using command **dsrvtgt**. This error affects the startup script on the control workstation. When this problem occurs, the daemon will not start on the control workstation. If the startup script is being run as part of a refresh operation, the refresh operation fails (does not take effect in any of the nodes). Standard fields indicate that the startup script was unable to obtain DCE credentials to write into the SDR, and present possible causes. Detailed fields contain the return code of the **/usr/lpp/ssp/bin/dsrvtgt ssp/spbgroot** command and the location of the startup script log, which contains more details about the problem. This error typically indicates problems in DCE or security services. For DCE configuration problems, see the configuration log file **/opt/dcelocal/etc/cfgdce.log**. For other DCE problems, see log files in the **/opt/dcelocal/var/svc** directory. For security services problems, see “Chapter 18. Diagnosing SP Security Services problems” on page 251. Specifically, the procedures for restoring key files is discussed in “Action 5 - Correct key files” on page 276.
TS_DEATH_TR A99EB4EA	UNKN	**Explanation:** Lost contact with a neighboring adapter. **Details:** This entry indicates that heartbeat messages are no longer being received from the neighboring adapter. This entry refers to a particular instance of the Topology Services daemon on the local node. The source of the problem could be either the local or remote node. Data from the remote node should also be obtained. Standard fields indicate that a local adapter is no longer receiving packets from the remote adapter. Detail Data fields contain the node number and IP address of the remote adapter. Data about the loss of connectivity may not be available after the problem is cleared. The local or remote adapter may have malfunctioned. Network connectivity to the remote adapter may have been lost. A remote node may have gone down. The Topology Services daemon on the remote node may have been blocked. If the problem is with the local adapter, an error log entry of type **TS_LOC_DOWN_ST** should follow in a few seconds. Information on the remote node should be collected to obtain a better picture of what failure has occurred.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_DMS_WARNING_ST 4F35BB80	INFO	**Explanation:** The Dead Man Switch timer is close to triggering. **Details:** This entry indicates that the Dead Man Switch has been reset with a small time-to-trigger value left on the timer. This means that the system is in a state where the Dead Man Switch timer is close to triggering. This condition affects the node where the error log entry appears. If steps are not taken to correct the problem, the node may be brought down by the Dead Man Switch timer. This entry is logged on each occurrence of the condition. Some possible causes are outlined. Detailed fields contain the amount of time remaining in the Dead Man Switch timer and also the interval to which the Dead Man Switch timer is being reset. Program **/usr/sbin/rsct/bin/hatsdmsinfo** displays the latest ″time-to-trigger″ values and the values of ″time-to-trigger″ that are smaller than a given threshold. Small ″time-to-trigger″ values indicate that the Dead Man Switch timer is close to triggering.
TS_DUPNETNAME_ER CE953608	PERM	**Explanation:** Duplicated network name in **machines.lst** file. **Details:** This entry indicates that a duplicate network name was found by the Topology Services daemon while reading the **machines.lst** configuration file. This entry refers to a particular instance of Topology Services on the local node. Other nodes may be affected by the same problem, since the **machines.lst** file is the same on all nodes. If this problem occurs at startup time, the daemon exits. Standard fields indicate that a duplicate network name was found in the **machines.lst** file. Detail Data fields show the name that was duplicated. In HACMP/ES, the command **/usr/es/sbin/cluster/utilities/cllsif** displays all the adapters in the HACMP configuration. Having adapters of different types belonging to the same network is the cause of this problem.
TS_FD_INVAL_ADDR_ST	PERM	**Explanation:** An adapter is not configured or has an address outside the cluster configuration. **Details:** This entry indicates that a given adapter in the cluster (PSSP, or HACMP) configuration is either not configured, or has an address which is outside the cluster configuration. This entry affects the local node, and causes the corresponding adapter to be considered down. Detailed data fields show the interface name, current address of the interface, and expected boot-time address. Probable causes for the problem are: • There is a mismatch between the cluster adapter configuration and the actual addresses configured on the local adapters. • The adapter is not correctly configured. Save the output of the command **netstat -in**. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center if the source of the problem cannot be found.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_FD_INTFC_NAME_ST	PERM	**Explanation:** An interface name is missing from the adapter configuration. **Details:** The Topology Services startup script reads information from the cluster repository, containing for each adapter its address, boot-time interface name, and node number. This error entry is created when the interface name information is missing. This usually points to a problem when generating the adapter configuration. The detailed data fields contain the address in the Topology Services configuration and the interface name which has been ″assigned″ to the adapter by the Topology Services daemon. In HACMP, the information is stored in the HACMP Global ODM. Commands **/usr/es/sbin/cluster/utilities/cllsif** and **/usr/es/sbin/cluster/utilities/clhandle** retrieve the adapter and node information used by Topology Services in HACMP. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center. This problem, in most of the cases, will not prevent Topology Services from correctly monitoring the adapter. However, internal problems may occur if a subsequent Topology Services refresh (which in HACMP is done via a Topology DARE) is attempted.
TS_HAIPDUP_ER 1BDC2F53	PERM	**Explanation:** IP address duplication in Topology Services configuration file. **Details:** This entry indicates that Topology Services was not able to start or refresh because the same IP address appeared twice in the configuration. This entry refers to a particular instance of Topology Services on the local node, but the problem may affect all the nodes. If this problem occurs at startup time, the daemon exits. To diagnose the problem in PSSP, retrieve data from the control workstation. To diagnose the problem in HACMP, retrieve data from any of the nodes that were affected by the problem. Standard fields indicate that the same IP address appeared twice in the Topology Services **machines.lst** configuration file. Detail Data fields show the node number of one of the nodes hosting the duplicated address and the duplicated IP address. Information about the cause of the problem may not be available once the problem is cleared. In the PSSP realm, the adapter configuration is stored in the **Adapter** class of the SDR. On the HACMP realm, information is stored in the HACMP Global ODM. Commands **/usr/es/sbin/cluster/utilities/clhandle** and **/usr/es/sbin/cluster/utilities/cllsif** retrieve the adapter and node information used by Topology Services in HACMP. If the problem is caused by an incorrect adapter address specification in PSSP, refer to the ″IP Address and Host Name Changes for SP Systems″ Appendix in *PSSP: Administration Guide* for instructions on how to change IP addresses.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_HALOCAL_ER F8949418	PERM	**Explanation:** Local node missing in Topology Services configuration file. **Details:** Standard fields indicate that the local node was not present in the **machines.lst** file. This is a problem with the configuration stored in the Registry. On the PSSP realm, information is stored in the **Adapter** class of the SDR. On the HACMP realm, information is stored in the HACMP Global ODM. Commands **/usr/es/sbin/cluster/utilities/clhandle** and **/usr/es/sbin/cluster/utilities/cllsif** retrieve the adapter and node information used by Topology Services in HACMP. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.
TS_HANODEDUP_ER BCE1B994	PERM	**Explanation:** Node number duplicated in Topology Services configuration file. **Details:** This entry indicates that Topology Services was not able to start or refresh because the same node appeared twice on the same network. This entry refers to a particular instance of Topology Services on the local node, but the problem should affect all the nodes. If this problem occurs at startup time, the daemon exits. To diagnose the problem in PSSP, retrieve data from the control workstation. To diagnose the problem in HACMP, retrieve data from any of the nodes that were affected by the problem. Standard fields indicate that the same node appeared twice in the same network in the Topology Services **machines.lst** configuration file. Detail Data fields show the interface name of one of the adapters and the node number that appears twice. Information about the cause of the problem may not be available once the problem is cleared. In the PSSP realm, the adapter configuration is stored in the **Adapter** class of the SDR. On the HACMP realm, information is stored in the HACMP Global ODM. Commands **/usr/es/sbin/cluster/utilities/clhandle** and **/usr/es/sbin/cluster/utilities/cllsif** retrieve the adapter and node information used by Topology Services in HACMP. If the problem is caused by an incorrect adapter address specification in PSSP, refer to the ″IP Address and Host Name Changes for SP Systems″ Appendix in *PSSP: Administration Guide* for instructions on how to change IP addresses.
TS_IOCTL_ER 7C090481	PERM	**Explanation:** An **ioctl** call failed. **Details:** This entry indicates that an **ioctl()** call used by the Topology Services daemon to obtain local adapter information failed. This is a possible AIX-related problem. The Topology Services daemon issued an **ioctl()** call to obtain information about the network adapters currently installed on the node. If this calls fails, there is a potential problem in AIX. The Topology Services daemon exits. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	**Type**	**Description**
TS_IPADDR_ER 209F6175	PERM	**Explanation:** Cannot convert IP address in dotted decimal notation to a number. **Details:** This entry indicates that an IP address listed in the **machines.lst** configuration file was incorrectly formatted and could not be converted by the Topology Services daemon. If this problem occurs at startup time, the daemon exits. Standard fields indicate that the daemon was unable to interpret an IP address listed in the **machines.lst** file. The Detail Data fields contain the given IP address in dotted decimal notation and the node number where the address was found. The problem may be that the file system where the **run** directory is located is corrupted, or information in the System Registry (SDR for PSSP, Global ODM for HACMP) is not correct. The **machines.lst** file is kept in the daemon ″run″ directory (**/var/ha/run/hats.***partition_name* for PSSP). The file is overwritten each time the subsystem is restarted. A copy of the file is kept in the startup script's log file, **/var/ha/log/hats.***partition_name*. A number of instances (currently 7) of this log file is kept, but the information is lost if many attempts are made to start the subsystem.
TS_KEYS_ER A45AC96A	PERM	**Explanation:** Topology Services startup script cannot retrieve key file information using command **/usr/sbin/rsct/bin/hats_keys**. **Details:** This entry indicates that command **/usr/sbin/rsct/bin/hats_keys**, run by the Topology Service startup script on the control workstation, was unable to retrieve the Topology Services key file information. This error occurs when the startup script is running on the control workstation during initial startup or refresh. When this error occurs, all Topology Services daemons in the system partition will terminate their operations and exit. Diagnosing this problem requires collecting data only on the control workstation. The pathname of Topology Services key file is **/spdata/sys1/keyfiles/rsct/***syspar_name***/hats**, where *syspar_name* is the name of the SP system partition. Standard fields indicate that the startup script was unable to retrieve the Topology Services key file information using command **hats_keys**, and present possible causes. Detail Data fields contain the return code of command **hats_keys** and the location of the startup script log. The error message returned by command **hats_keys** is in the startup script log. This error typically indicates problems in DCE or SP Security Services. For DCE configuration problems, see the configuration log file **/opt/dcelocal/etc/cfgdce.log**. For other DCE problems, see log files in the **/opt/dcelocal/var/svc** directory. For SP security services problems, see “Chapter 18. Diagnosing SP Security Services problems” on page 251. Specifically, the procedures for restoring key files is discussed in “Action 5 - Correct key files” on page 276.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_LATEHB_PE 7AD2CABA	PERF	**Explanation:** Late in sending heartbeat to neighbors. **Details:** This entry indicates that the Topology Services daemon was unable to run for a period of time. This entry refers to a particular instance of the Topology Services daemon on the local node. The node that is the Downstream Neighbor may perceive the local adapter as dead and issue a **TS_DEATH_TR** error log entry. A node's Downstream Neighbor is the node whose IP address is immediately lower than the address of the node where the problem was seen. The node with the lowest IP address has a Downstream Neighbor of the node with the highest IP address. Standard fields indicate that the Topology Services daemon was unable to send messages for a period of time. Detail Data fields show how many seconds late the daemon was in sending messages. This entry is created when the amount of time that the daemon was late in sending heartbeats is equal to or greater than the amount of time needed for the remote adapter to consider the local adapter as down. Data about the reason for the Topology Services daemon being blocked is not usually kept, unless system tracing is being run on the node. The Service log file keeps information about Topology Services events happening on the node at the time the daemon was blocked. See "Topology Services service log" on page 379. Look for error log entries **TS_NODEDOWN_EM** in other nodes that refer to this node. If these are present , it means that the daemon blockage needs to be corrected, or the Topology Services tuning parameters have to be changed. Refer to the ″Node appears to go down and then up a few/several seconds later″ symptom in "Error symptoms, responses, and recoveries" on page 402.
TS_LIBERR_EM 8C2164B7	PEND	**Explanation:** Topology Services client library error. **Details:** This entry indicates that the Topology Services library had an error. It refers to a particular instance of the Topology Services library on the local node. This problem will affect the client associated with the library (RSCT Event Manager or more likely RSCT Group Services). Standard fields indicate that the Topology Services library had an error. Detail Data fields contain the error code returned by the Topology Services API. Data needed for IBM Service to diagnose the problem is stored in the Topology Services daemon service log, located at **/var/ha/log/hats.*.**_partition_name_ for PSSP or **/var/ha/log/topsvcs.*.**_cluster_name_ for HACMP/ES. Since this file may wrap, it should be saved. The RSCT Group Services daemon (the probable client connected to the library) is likely to have exited with an assert and to have produced an error log entry with template **GS_TS_RETCODE_ER**. Refer to "Chapter 24. Diagnosing Group Services problems" on page 419 for a list of the information to save. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description

TS_LOC_DOWN_ST
A0B80A40
	INFO	**Explanation:** Local adapter down.
		Details: This entry indicates that one of the local adapters is down. This entry refers to a particular instance of the Topology Services daemon on the local node. If there are multiple Topology Services daemons running in the node (for example, the PSSP version and the HACMP version), each daemon creates its own error log entry.
		Standard fields indicate that a local adapter is down. Detail Data fields show the interface name, adapter offset (index of the network in the **machines.lst** file), and the adapter address according to Topology Services. This address may differ from the adapter's actual address if the adapter is incorrectly configured. Information about the source of the problem may be lost after the condition is cleared.
		Possible problems are:
		• The adapter may have malfunctioned.
		• The adapter may be incorrectly configured. See entry for **TS_UNS_SIN_TR**.
		• There is no other adapter functioning in the network.
		• Connectivity has been lost in the network.
		• The SP Switch or SP Switch2 adapter may be fenced.
		• A problem in Topology Services' adapter health logic.
		Perform these steps:
		1. Verify that the address of the adapter listed in the output of
		`ifconfig  interface_name`
		is the same as the one shown in this error log entry. If they are different, the adapter has been configured with an incorrect address.
		2. If the output of the **ifconfig** command does not show the **UP** flag, this means that the adapter has been forced down by the command:
		`ifconfig  interface_name down`
		If the adapter is an SP Switch or SP Switch2 adapter, it may be fenced.
		3. Issue the command **netstat -in** to verify whether the `receive` and `send` counters are being incremented for the given adapter. The counters are the numbers below the `Ipkts` (receive) and `Opkts` (send) columns. If both counters are increasing, the adapter is likely to be working and the problem may be in Topology Services.
		4. Issue the **ping** command to determine whether there is connectivity to any other adapter in the same network. If **ping** receives responses, the adapter is likely to be working and the problem may be in Topology Services.
		5. Refer to "Operational test 4 - Check address of local adapter" on page 394.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_LOGFILE_ER 42D96ED3	PERM	**Explanation:** The daemon failed to open the log file. **Details:** This entry indicates that the Topology Services daemon was unable to open its log file. Standard fields indicate that the daemon was unable to open its log file. Detail Data fields show the name of the log file. The situation that caused the problem may clear when the file system problem is corrected. The Topology Services daemon exits. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.
TS_LONGLINE_ER 4D0CE96E	PERM	**Explanation:** The Topology Services daemon cannot start because the **machines.lst** file has a line that is too long. **Details:** This entry indicates that the Topology Services daemon was unable to start because there is a line which is too long in the **machines.lst** configuration file. This entry refers to a particular instance of Topology Services on the local node. If this problem occurs at startup time, the daemon exits. The problem is likely to affect other nodes, since the **machines.lst** file should be the same at all nodes. Standard fields indicate that the daemon was unable to start because the **machines.lst** configuration file has a line longer than 80 characters. Detail Data fields show the pathname of the **machines.lst** configuration file. It is possible that the network name is too long, or there is a problem in the **/var/ha** file system.
TS_LSOCK_ER 4E358F5D	PERM	**Explanation:** The daemon failed to open a listening socket for connection requests. **Details:** This entry indicates that the Topology Services daemon was unable to open a socket connection to communicate with its clients. Standard fields indicate that the daemon was unable to open the socket. Detail Data fields show the operation being attempted at the socket (in English) and the system error value returned by the system call. The situation that caused the problem may clear with a reboot. The **netstat** command shows the sockets in use in the node. The Topology Services daemon exits. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.
TS_MACHLIST_ER EDFE80F3	PERM	**Explanation:** The Topology Services configuration file cannot be opened. **Details:** This entry indicates that the Topology Services daemon was unable to read its **machines.lst** configuration file. Standard fields indicate that the daemon was unable to read the **machines.lst** file. Detail Data fields show the pathname of the file. Information about the cause of the problem is not available after the condition is cleared. If this problem occurs at startup time, the daemon exits. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_MIGRATE_ER 8535A705	PERM	**Explanation:** Migration-refresh error. **Details:** This entry indicates that the Topology Services daemon has found a problem during a migration-refresh. The migration-refresh is a refresh operation issued at the end of an HACMP node by node migration, when the last node is moved to the newer release. The problem may be caused by the information placed on the Global ODM when the migration protocol is complete. This entry refers to a particular instance of the Topology Services daemon on the local node. It is likely that some of the other nodes have a similar problem. Standard fields indicate that the Topology Services daemon encountered problems during a migration-refresh. HACMP may have loaded incorrect information into the Global ODM. Data read by the Topology Services startup script is left on the Topology Services **run** directory and will be overwritten in the next refresh or startup operation. The data in the ″run″ directory should be saved. The Topology Services ″Service″ log file has a partial view of what was in the Global ODM at the time of the refresh operation. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.
TS_MISCFG_EM 6EA7FC9E	PEND	**Explanation:** Local adapter incorrectly configured. **Details:** This entry indicates that one local adapter is either missing or has an address that is different from the address that Topology Services expects. Standard fields indicate that a local adapter is incorrectly configured. Detail Data fields contain information about the adapter, such as the interface name, adapter offset (network index in the **machines.lst** file), and expected address. Possible sources of the problem are: • The adapter may have been configured with a different IP address. • The adapter is not configured. • Topology Services was started after a ″Force Down″ in HACMP. This entry is created on the **first occurrence** of the condition. No data is stored about the condition after the problem is cleared. Use the interface name in the error report to find the adapter that is incorrectly configured. Command: **ifconfig** *interface_name* displays information about the adapter.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_NIM_DIED_ER	PERM	**Explanation:** One of the NIM processes terminated abnormally. **Details:** This entry is created when one of the ″Network Interface Modules″ (NIM) -- processes used by Topology Services to monitor the state of each adapter, terminates abnormally. When a NIM terminates, the Topology Services daemon will restart another, but if the replacement NIM also terminates quickly then no other NIM will be started, and the adapter will be flagged as down. Detailed data fields show: • Process exit value, if not terminated with a signal (A value from 1 to 99), will be an ’errno’ value from invoking the NIM process. • Signal number (0: no signal). • Whether a core file was created (1: core file; 0: no core file). • Process id (PID). • Interface name being monitored by the NIM. • Pathname of NIM executable file. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.
TS_NIM_NETMON_ERROR_ER	PERM	**Explanation:** An error occurred in the netmon library, used by the NIM (″Network Interface Module″ -- processes used by Topology Services to monitor the state of each adapter) in determining whether the local adapter is alive. **Details:** This entry is created when there is an internal error in the netmon library. As a result, the local adapter will be flagged as down, even though the adapter may still be working properly. A possible cause for the problem (other than a problem in the library) is the presence of some non-supported adapter in the cluster configuration. Detailed data fields show: • Errno value. • Error code from netmon library. • Function name in library that presented a problem. • Interface name being monitored. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center. It is important to collect the information as soon as possible, since log information for the netmon library is kept in log files that may wrap within 30 minutes.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_NIM_OPEN_ERROR_ER	PERM	**Explanation:** NIM (″Network Interface Module″ -- processes used by Topology Services to monitor the state of each adapter) failed to connect to the local adapter that it is supposed to monitor. **Details:** This entry is created when the NIM is unable to connect to the local adapter that needs to be monitored. As a result, the adapter will be flagged as down, even though the adapter might still be working properly. Detailed data fields show: • Interface name. • Description 1: description of the problem. • Description 2: description of the problem. • Value 1 - used by the IBM Support Center. • Value 2 - used by the IBM Support Center. Some possible causes for the problem are: • NIM process was blocked while responding to NIM open command. • NIM failed to open non-IP device. • NIM received an unexpected error code from a system call. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.
TS_NODEDOWN_EM 4D9226A5	PEND	**Explanation:** Remote nodes were seen as down by Topology Services. **Details:** This is an indication that the Topology Services daemon detected one or more remote nodes as being down. This refers to a particular instance of the Topology Services daemon. Data should be collected on the remote nodes that were seen as dead. Standard fields indicate that remote nodes were seen as dead and present possible causes. Detailed fields contain the path name of a file containing the numbers of the affected nodes. The file with the node numbers may eventually be deleted by the system. The file is located in: **/var/adm/ffdc/dumps/hatsd.*** Verify that the nodes listed in the specified file actually went down and investigate why.
TS_NODENUM_ER 2033793C	PERM	**Explanation:** The local node number is not known to Topology Services. **Details:** This entry indicates that Topology Services was not able to find the local node number. Standard fields indicate that the daemon was unable to find its local node number. The Topology Services daemon exits. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description

TS_NODEUP_ST
95A9DAD0
	INFO	**Explanation:** Remote nodes that were previously down were seen as up by Topology Services. This is an indication that the Topology Services daemon detected one or more previously down nodes as being up. It refers to a particular instance of the Topology Services daemon.
		Details: In case the same nodes were seen as dead a short time before, data should be collected on the remote nodes. Standard fields indicate that remote nodes were seen as up and present possible causes. Detailed fields contain, in the section, a reference to the entry where the same nodes were seen as dead. If these nodes were seen as down before at different times, the reference code will be for one of these instances.
		The Detail Data also contains the path name of a file which stores the numbers of the nodes that were seen as up, along with the error id for the error log entry where each node was seen as dead previously. The file with the node numbers may eventually be deleted by the system. The file is located in: **/var/adm/ffdc/dumps/hatsd.***.
		If the same nodes were recently seen as dead (follow the REFERENCE CODE), examine the remote nodes for the reason why the nodes were temporarily seen as dead. This entry is logged when a remote node is seen as alive. The same node may have been seen as dead some time ago. If so, the **TS_NODEUP_ST** will have, as part of the Detail Data, a location of a file whose contents are similar to:
		`.ZOWYB/Z5Kzr.zBI14tVQ7....................`
		`     1`
		The file contains the ERROR ID of the error log entry of the corresponding TS_NODEDOWN_EM entry (when the same node was flagged as dead).

TS_OFF_LIMIT_ER	PERM	**Explanation:** Number of network offsets exceeds Topology Services limit. **Details:** This entry is created whenever the number of adapters and networks in the cluster configuration exceeds the Topology Services daemon's internal limit for maximum number of ″heartbeat rings″ of 16. Notice that a single cluster network may map to multiple ″heartbeat rings″. This will happen when a node has multiple adapters in the same network, since a heartbeat ring is limited to a single adapter per node. If this error occurs, a number of adapters and networks in the configuration may remain unmonitored by Topology Services. The detailed data fields contain the first network in the configuration to be ignored and the maximum number of networks allowed. When attempting to resolve the problem, initially focus on the nodes that have the most adapters in the configuration, and proceed to remove some adapters from the configuration.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description

TS_REFRESH_ER
5FB345F4
	PERM	**Explanation:** Topology Services refresh error.
		Details: This entry indicates that a problem occurred during a Topology Services refresh operation. A refresh operation can be a result of a **hatsctrl -r** command on PSSP systems, or a Topology DARE (Dynamic Automatic Reconfiguration Event) in HACMP. Topology DARE is an HACMP feature to change the configuration of the cluster dynamically, without having to shut the cluster down. Topology DARE is invoked by customers when a configuration change, such as adding a node, is made.
		This SMIT sequence performs the Topology DARE:
		`SMIT`
		`     Cluster Topology`
		`            Synchronize Cluster Topology`
		This entry refers to a particular instance of the Topology Services daemon on the local node. On HACMP, the problem may have occurred in other nodes as well. Standard fields indicate that a refresh error occurred.
		The **machines.lst** file has some incorrect information. The problem is probably created during a migration-refresh on an HACMP node by node migration. Data used to build the **machines.lst** file is stored in the daemon's ″run″ directory and may be lost if Topology Services is restarted or a new refresh is attempted.
		More details about the problem are in the User log file. See "Topology Services user log" on page 381. Additional details are stored in the Service log. See "Topology Services service log" on page 379. If this problem occurs at startup time, the Topology Services daemon may exit. See "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

TS_RSOCK_ER F11523E1	PERM	**Explanation:** The daemon failed to open socket for peer daemon communication. **Details:** This entry indicates that the Topology Services daemon was unable to open a UDP socket for communication with peer daemons in other nodes. Standard fields indicate that the daemon was unable to open the socket. Detail Data fields describe the operation being attempted at the socket (in English), the reason for the error, the system error value, and the port number. The port number may be in use by either another subsystem or by another instance of the Topology Services daemon. If the AIX SRC subsystem loses its connection to the Topology Services daemon, the AIX SRC may erroneously allow a second instance of the daemon to be started, leading to this error. The situation that caused the problem may clear with a node reboot. Follow the procedures described for the ″Nodes or adapters leave membership after refresh″ symptom in "Error symptoms, responses, and recoveries" on page 402 to find a possible Topology Services daemon running at the node and stop it. If no process is found that is using the peer socket, see "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center. Include also an AIX System Dump. See "Chapter 5. Producing a system dump" on page 81.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_SDR_ER 0BD8A620	PERM	**Explanation:** Cannot retrieve data from the SDR. **Details:** This entry indicates that the Topology Services startup script **hats** was unable to retrieve information from the System Data Repository (SDR). This entry refers to a particular instance of Topology Services on the local node. If the SDR itself is having problems, it is likely that other nodes are also affected. Standard fields indicate that data could not be retrieved from the SDR. This could be: • A problem with the SDR subsystem. • Too much contention for the SDR when a large number of nodes are trying to access the SDR simultaneously. • Too much traffic on the SP Ethernet. Information about the cause of the problem may not be available once the problem is cleared. Diagnose the SDR subsystem. See "Chapter 14. Diagnosing SDR problems" on page 125.
TS_SECMODE_ER D1BD179A	PERM	**Explanation:** Failed to determine local DCE security mode. **Details:** This entry indicates that errors have occurred while the Topology Services daemon is trying to obtain DCE security information in a partition where DCE is the sole authentication method. When this error occurs, the affected daemon will terminate its operation and exit. This error should not occur if ″Compatibility″ is one of the partition's authentication methods. The following are probable causes for the problem: • The Topology Services daemon failed to load the security services library (library **/usr/lib/libspsec.a** possibly not installed). • Calls to security services library functions failed to either initialize the library, obtain the authentication method, or read the key file. • The local node is not configured in DCE-only mode. • Topology Services on the control workstation could not determine the partition's security state, using the **/usr/lpp/ssp/bin/lsauthpts** command, or it could not access the Topology Services key file. The Detail Data fields contain the location of the Topology Services User log file, which includes more detailed information about the problem. For SP Security Services problems, see "Chapter 18. Diagnosing SP Security Services problems" on page 251.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_SECURITY_ST 78278638	INFO	**Explanation:** Authentication failure in Topology Services. **Details:** This entry indicates that the Topology Services daemon cannot authenticate a message from one of the peer daemons running in a remote node. This entry refers to a particular instance of the Topology Services daemon on the local node. The node which is sending these messages must also be examined. Standard fields indicate that a message cannot be authenticated. Detail Data fields show the source of the message. The possible problems are: • There is an attempt at a security breach. • The Time-Of-Day clocks in the nodes are not synchronized. • There are stale packets flowing through the network. • IP packets are being corrupted. • The security key file is not in sync across all nodes in the system partition. An entry is created the first time a message cannot be authenticated. After that, entries are created less frequently. Information about the network must be collected while the messages are still being received. The command **iptrace** should be used to examine the packets arriving at the node. Perform the following steps: 1. Examine the output of the **lssrc -ls hats** command on the local node and on the node sending the message. Look for field ″Key version″ in the output and check whether the numbers are the same on both nodes. 2. Check that the key file is the same in all the nodes in the partition.
TS_SECURITY2_ST E486FA26	INFO	**Explanation:** More authentication failures in Topology Services. **Details:** This entry indicates that there have been additional incoming messages that could not be authenticated. For the first such message, error log entry **TS_SECURITY_ST** is created. If additional messages cannot be authenticated, error log entries with label **TS_SECURITY2_ST** are created less and less frequently. The standard fields indicate that incoming messages cannot be authenticated. The detailed fields show an interval in seconds and the number of messages in that interval that could not be authenticated. For more details and diagnosis steps, see the entry for the **TS_SECURITY_ST** label.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_SEMGET_ER 68547A69	PERM	**Explanation:** Cannot get shared memory or semaphore segment. This indicates that the Topology Services daemon was unable to start because it could not obtain a shared memory or semaphore segment. This entry refers to a particular instance of the Topology Services daemon on the local node. The daemon exits **Details:** Standard fields indicate that the daemon could not start because it was unable to get a shared memory or a semaphore segment. The Detail Data fields contain the key value and the number of bytes requested for shared memory, or the system call error value for a semaphore. The reason why this error has occurred may not be determined if the subsystem is restarted and this error no longer occurs.
TS_SERVICE_ER F93756D2	PERM	**Explanation:** Unable to obtain port number from the **/etc/services** file. **Details:** This entry indicates that the Topology Services daemon was unable to obtain the port number for daemon peer communication from **/etc/services**. This entry refers to a particular instance of the Topology Services daemon on the local node. The daemon exits. Other nodes may be affected if their **/etc/services** have similar contents as that on the local node. Standard fields indicate that the daemon was unable to obtain the port number from **/etc/services**. Detail Data fields show the service name used as search key to query **/etc/services**.
TS_SHMAT_ER DA6A5149	PERM	**Explanation:** Cannot attach to shared memory segment. **Details:** This entry indicates that the Topology Services daemon was unable to start because it could not attach to a shared memory segment. Standard fields indicate that the daemon could not start because it was unable to attach to a shared memory segment. The daemon exits. The Detail Data fields contain the shared memory identifier and number of bytes requested. The reason why the error occurred may not be found if the subsystem is restarted and the same error does not occur.
TS_SHMEMKEY_ER 41E8D858	PERM	**Explanation:** Cannot get IPC key. **Details:** This indicates that the Topology Services daemon was unable to start because it could not obtain an IPC key. This refers to a particular instance of the Topology Services daemon on the local node. The daemon exits. Standard fields indicate that the daemon could not start because it was unable to obtain an IPC key. The Detail Data fields contain the pathname of the UNIX-domain socket used for daemon-client communication. This pathname is given to the **ftok()** subroutine in order to obtain an IPC key. This entry is created when the UNIX-domain socket file has been removed. The reason why this error has occurred may not be determined if the subsystem is restarted and this error no longer occurs.
TS_SHMGET_ER 42416EB1	PERM	See TS_SEMGET_ER

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_SP_DIR_ER 596A9ABD	PERM	**Explanation:** Cannot create directory. **Details:** This entry indicates that the Topology Services startup script **hats** was unable to create one of the directories it needs for processing. Standard fields indicate that a directory could not be created by the startup script **hats**. Detail Data fields show the directory that could not be created. Information about the cause of the problem may not be available once the problem is cleared.
TS_SPIPDUP_ER E68EB007	PERM	See TS_HAIPDUP_ER
TS_SPLOCAL_ER 74B0CCF7	PERM	See TS_HALOCAL_ER
TS_SPNODEDUP_ER C82AB176	PERM	See TS_HANODEDUP_ER
TS_START_ST 645637FC	INFO	**Explanation:** The Topology Services daemon has started. This is an indication that the Topology Services daemon has started. This entry refers to a particular instance of the Topology Services daemon on the local node, or particular partition on the control workstation. **Details:** Standard fields indicate that the daemon started. The Topology Services subsystem was started by a user or during system boot. Detail Data will be in the language where the **errpt** command is run. The Detail Data contains the location of the log and run directories and also which user or process started the daemon.
TS_STOP_ST A204A4EE	INFO	**Explanation:** The Topology Services daemon has stopped. This is an indication that the Topology Services daemon has stopped. This entry refers to a particular instance of the Topology Services daemon on the local node or particular partition on the control workstation. **Details:** The Topology Services subsystem shutdown was caused by a signal sent by a user or process. Standard fields indicate that the daemon stopped. The standard fields are self-explanatory. If stopping the daemon is not desired, you must quickly understand what caused this condition. If the daemon was stopped by the AIX SRC, the word ″SRC″ is present in the Detail Data . The REFERENCE CODE field in the Detail Data section refers to the error log entry for the start of Topology Services. Detail Data is in English. Detail Data fields point to the process (SRC) or signal that requested the daemon to stop.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_SYSPAR_ER 8C616BE5	PERM	**Explanation:** Cannot find system partition name. **Details:** This entry indicates that the Topology Services startup script **hats** was unable to obtain the partition name using the **/usr/lpp/ssp/bin/spget_syspar** command. Standard fields indicate that a problem occurred in **/usr/lpp/ssp/bin/spget_syspar**. Information about the cause of the problem in **spget_syspar** may be lost once the problem is cleared. Issue the commands: **/usr/lpp/ssp/bin/spget_syspar** and **/usr/lpp/ssp/bin/spget_syspar -n**. If either returns an error and a nonzero exit code, perform problem determination procedures on the SDR. See “Chapter 14. Diagnosing SDR problems” on page 125.
TS_THATTR_ER B705E4E5	PERM	**Explanation:** Cannot create or destroy a thread attributes object. **Details:** This entry indicates that Topology Services was unable to create or destroy a thread attributes object. Standard fields indicate that the daemon was unable to create or destroy a thread attributes object. Detail Data fields show which of the Topology Services threads was being handled. The Topology Services daemon exits. See “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.
TS_THCREATE_ER 5540C482	PERM	**Explanation:** Cannot create a thread. **Details:** This entry indicates that Topology Services was unable to create one of its threads. Standard fields indicate that the daemon was unable to create a thread. Detail Data fields show which of the Topology Services threads was being created.
TS_THREAD_STUCK_ER 47E4956B	PERM	**Explanation:** Main thread is blocked. Daemon will exit. **Details:** This entry indicates that the Topology Services daemon will exit because its main thread was blocked for longer than a pre-established time threshold. If the main thread remains blocked for too long, it is possible that the node is considered dead by the other nodes. The main thread needs to have timely access to the CPU, otherwise it would fail to send ″heartbeat″ messages, run adapter membership protocols, and notify Group Services about adapter and node events. If the main thread is blocked for too long, the daemon exits with a core dump, to allow debugging of the cause of the problem. This entry refers to a particular instance of Topology Services running on a node. The standard fields indicate that the Topology Services daemon will exit because the main thread was blocked for too long, and explains some of the possible causes. The detailed fields show the number of seconds that the main thread appeared to be blocked, the number of recent page faults involving I/O operations, and the interval in milliseconds where these page faults occurred. If the number of page faults is non-zero, the problem could be related to memory contention. For information about diagnosing and working around the problem in case its root cause is a resource shortage, see “Action 5 - Investigate hatsd problem” on page 406. If a resource shortage does not seem to be a factor, the cause could be a problem in the daemon or in a service invoked by it. Contact the IBM Support Center.

Table 58. AIX Error Log templates for Topology Services (continued)

Label and Error ID	Type	Description
TS_UNS_SIN_TR 029E523B	UNKN	**Explanation:** Local adapter in unstable singleton state. **Details:** This entry indicates that a local adapter is staying too long in a singleton unstable state. Though the adapter is able to receive some messages, there could be a problem with it, which may prevent outgoing messages from reaching their destinations. This entry refers to a particular instance of the Topology Services daemon on the local node. Examine the Service log on other nodes to determine if other nodes are receiving messages from this adapter. See “Topology Services service log” on page 379. Standard fields indicate that a local adapter is in an unstable singleton state. Detail Data fields show the interface name, adapter offset (index of the network in the **machines.lst** file), and the adapter address according to Topology Services, which may differ from the adapter’s actual address if the adapter is incorrectly configured. The adapter may be unable to send messages. The adapter may be receiving broadcast messages but not unicast messages. Information about the adapter must be collected while the adapter is still in this condition. Issue the commands: **ifconfig** *interface_name* and **netstat -in** and record the output. Perform these steps: 1. Check if the address displayed in the error report entry is the same as the actual adapter address, which can be obtained by issuing this command: **ifconfig** *interface_name*. If they are not the same, the adapter has been configured with the wrong address. 2. Issue command **ping** *address* from the local node for all the other addresses in the same network. If **ping** indicates that there is no reply (for example: `10 packets transmitted, 0 packets received, 100% packet loss`) for all the destinations, the adapter may be incorrectly configured. 3. Refer to “Operational test 6 - Check whether the adapter can communicate with other adapters in the network” on page 396.

Dump information

Topology services provides two dumps, a core dump which is created automatically when certain errors occur, and a **phoenix.snap** dump which is created manually.

Core dump

There is a core dump generated by the Topology Services daemon. It contains information normally saved by AIX in a core dump: user-space data segments for the Topology Services daemon. It refers to a particular instance of the Topology Services daemon on the local node. Other nodes may have a similar core dump. On PSSP systems, the dump is located in: **/var/ha/run/hats.***partition_name***/core**. On HACMP systems, the dump is located in: **/var/ha/run/topsvcs.***cluster_name***/core**. An approximate size for the core dump file is between 7 and 10MB.

The dump is created automatically when the daemon invokes an **assert()** statement, or when the daemon receives a segmentation violation signal for accessing its data incorrectly. Forcing **hatsd** to generate a dump is necessary, especially if the daemon is believed to be in a hung state. The dump is created manually by issuing the command:

```
kill -6 pid_of_daemon
```

The *pid_of_daemon* is obtained by issuing:

- **lssrc -s hats** on PSSP nodes
- **lssrc -s hats.** *partition_name* on the PSSP control workstation
- **lssrc -s topsvcs** on HACMP nodes

The dump remains valid as long as the executable file **/usr/sbin/rsct/bin/hatsd** is not replaced. Only the last three core file instances are kept. The core dumps and the executable should be copied to a safe place. To analyze the dump, issue the command:

```
dbx /usr/sbin/rsct/bin/hatsd core_file
```

Good results are similar to the following:

```
Type 'help' for help.
reading symbolic information ...
[using memory image in core]

IOT/Abort trap in evt._pthread_ksleep [/usr/lib/libpthreads.a] at
0xd02323e0 ($t6) 0xd02323e0 (_pthread_ksleep+0x9c) 80410014 lwz r2,0x14(r1)
```

Some of the **error results** are:

1. This means that the current executable file was not the one that created the core dump.

   ```
   Type 'help' for help.
   Core file program (hatsd) does not match current program (core ignored)
   reading symbolic information ...
   (dbx)
   ```

2. This means that the core file is incomplete due to lack of disk space.

   ```
   Type 'help' for help.
   warning: The core file is truncated.  You may need to increase the
   ulimit for file and coredump, or free some space on the filesystem.
   reading symbolic information ...
   [using memory image in core]

   IOT/Abort trap in evt._pthread_ksleep [/usr/lib/libpthreads.a]
    at 0xd02323e0 0xd02323e0 (_pthread_ksleep+0x9c) 80410014
    lwz   r2,0x14(r1) (dbx)
   ```

phoenix.snap dump

This dump contains diagnostic data used for RSCT problem determination. It is a collection of log files and other trace information used to obtain a global picture of the state of RSCT. The dump is specific to each node. It is located in the **/tmp/phoenix.snapOut** directory. The dump is created by this command:

```
/usr/sbin/rsct/bin/phoenix.snap
```

Note: The **phoenix.snap** tool is a service tool and not a PSSP command. The tool is shipped with PSSP 3.2 as is - without documentation. For assistance on using **phoenix.snap** in a manner other than what is described in this section, contact the IBM Support Center.

The **phoenix.snap** tool may be run from a node, where it collects data from the node only. The **phoenix.snap** tool may be run from the control workstation, where it collects data from a list of nodes if the **-l** flag is used. For small systems, the command should be run from the control workstation with the **-l** flag to collect data from all the nodes.

For larger systems, see “Information to collect before contacting the IBM Support Center” on page 384 for a list of nodes to collect data from. If the **-l** option is not used, data is collected only from the local node. When this command is run on the control workstation, data is also collected for the Topology Services Group Leader and Group Services Nameserver. See “Contents of the phoenix.snap file” on page 378.

Data about the state of the subsystems, such as output of the **lssrc -l** command, becomes partially obsolete as soon as the state of the subsystem changes. An example of this is an adapter or node event. Data in the log files collected by **phoenix.snap** remains valid even after the state of the subsystem changes, but the collected log files will contain data only up to the time when **phoenix.snap** was run.

Data in **/tmp/phoenix.snapOut** will not be overwritten when another instance of **phoenix.snap** is run on the same node. This is because file **phoenix.snap.***timestamp***.tar.Z** has a timestamp in the name.

These flags control **phoenix.snap**:

- The **-d** flag instructs **phoenix.snap** to save the collected information in a different directory.
- The **-l** flag instructs **phoenix.snap** to collect data from a list of nodes, specified by each node’s hostname, separated by commas. **-l ALL** collects data from all the nodes.

  The **-l** flag can be used only on the control workstation. When this flag is used, **phoenix.snap** produces a **tar** file that contains the collected output from all the requested nodes.

Good results from the **phoenix.snap** command are indicated by an output similar to the following:

```
phoenix.snap version 1.4, args:

Determining if there is enough space in /tmp/phoenix.snapOut

compressed file will be about 982303 bytes
phoenix.snap requires about 5893820 bytes
Think we have enough space.
####################################################################
Send file /tmp/phoenix.snapOut/phoenix.snap.node3.12211215.out.tar.Z
 to the RS6000/SP service team
```

Error results are indicated by messages such as:

```
There is not enough space in /tmp/phoenix.snapOut
```

which indicates that there is no space to store the resulting data file. In this case, rerun the command using the **-d** flag to specify another directory, or create additional space in the **/tmp** directory.

Contents of the phoenix.snap file

The dump is a collection of files archived with the **tar** command and compressed with the **compress** command. Some of the files are copies of daemon log files, and some are the output of certain commands.

This is a partial list of the data items collected:

- Output of these AIX commands:
  - LANG=C errpt -a
  - ps -edf
  - lslpp -L
  - lssrc -a
  - vmstat
  - vmstat -s
  - netstat (several options)
  - ifconfig (several adapters)
  - no -a
  - netstat *interface_name*
- Data obtained from the SDR:
  - **hats.machines.lst** SDR file
  - **hats.machines.inst** SDR file
  - **Syspar_ports** object
  - **TS_Config** object
  - **GS_config** object
  - **host_responds** object
  - **/etc/SDR_dest_info** file
- Log files and **run** directory for:
  - Topology Services
  - Group Services
- Output of component-specific commands, such as:
  - lssrc -l
  - hagsgr
  - hagsvote

The output of the **phoenix.snap** command is in English. Many of the logs collected by the program are in English. Some will be in the node’s language.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Consult these logs for debugging purposes. They all refer to a particular instance of the Topology Services daemon running on the local node.

Topology Services service log

This log contains trace information about the activities performed by the daemon. When a problem occurs, logs from multiple nodes will often be needed. These log files must be collected before they wrap or are removed.

The trace is located in:

- **/var/ha/log/hats.***DD.hhmmss.partition_name* for PSSP nodes.
- **/var/ha/log/topsvcs.***DD.hhmmss.cluster_name* for HACMP nodes.

where *DD* is the day of the month when the daemon was started, and *hhmmss* is the time when the daemon was started.

If obtaining logs from all nodes is not feasible, the following is a list of nodes from which logs should be collected:

1. The node where the problem was seen
2. The Group Leader node on each network

   The Group Leader is the node which has the highest IP address on a network.
3. The Downstream Neighbor on each network

   This is the node whose IP address is immediately lower than the address of the node where the problem was seen. The node with the lowest IP address has a Downstream Neighbor of the node with the highest IP address.
4. The control workstation

Service Log long tracing

The most detailed level of tracing is Service log long tracing. It is started with the command:

```
traceson -l -s subsystem_name
```

where *subsystem_name* is:

- **hats** on PSSP nodes
- **hats.***partition_name* on the PSSP control workstation
- **topsvcs** on HACMP nodes

The long trace is stopped with this command: **tracesoff -s** *subsystem_name*, which causes short tracing to be in effect. When the log file reaches the maximum line number, the current log is saved in a file with a suffix of **.bak**, and the original file is truncated. When the daemon is restarted, a new log file is created. Only the last five log files are kept.

With service log long tracing, trace records are generated under the following conditions:

- Each message sent or received
- Each adapter that is disabled or re-enabled
- Details of protocols being run
- Details of node reachability information
- Refresh
- Client requests and notifications
- Groups formed, elements added and removed

Data in the Service log is in English. Each Service log entry has this format:

```
date     daemon name     message
```

Adapters are identified by a pair:

```
(IP address:incarnation number)
```

Groups are identified by a pair:

```
(IP address of Group Leader:incarnation number of group)
```

Long tracing should be activated on request from IBM Service. It can be activated (just for about one minute, to avoid overwriting other data in the log file), when the error condition is still present.

Service Log normal tracing

Service log normal tracing is the default, and is always running. There is negligible impact if no node or adapter events occur on the system. An adapter death event may result in approximately 50 lines of log information for the Group Leader and ″mayor″ nodes, or up to 250 lines for the Group Leader and ″mayor″ nodes on systems of approximately 400 nodes. All other nodes will produce less than 20 lines. Log file sizes can be increased as described in “Changing the service log size” on page 381.

With normal tracing, trace records are generated for these conditions:

- Each adapter that is disabled or re-enabled
- Some protocol messages sent or received
- Refresh
- Client requests and notifications
- Groups formed, members added and removed

No entries are created when no adapter or node events are happening on the system.

With normal tracing, the log trimming rate depends heavily on the frequency of adapter or node events on the system. The location of the log file and format of the information is the same as that of the long tracing described previously.

If the Service log file, using normal tracing, keeps growing even when no events appear to be happening on the system, this may indicate a problem. Search for possible entries in the AIX error log or in the User log. See “Topology Services user log”.

Changing the service log size

The long trace generates approximately 10KB of data per minute of trace activity. By default, log files have a maximum of 5000 lines, which will be filled in 30 minutes or less if long tracing is requested. To change the log file size:

1. For PSSP, issue this command on the control workstation:

   ```
   hatstune -l new_max_lines -r
   ```

   The full path name of this command is: **/usr/sbin/rsct/bin/hatstune**.

   For example, **hatstune -l 10000 -r** changes the maximum number of lines in a log file to 10000. The **-r** flag causes the Topology Services subsystem to be refreshed in all the nodes.
2. For HACMP, use this sequence:

   ```
   smit hacmp
       Cluster Configuration
         Cluster Topology
           Configure Topology Services and Group Services
             Change / Show Topology and Group Services Configuration

         Cluster Topology
             Synchronize Cluster Topology
   ```

Topology Services user log

The Topology Services user log contains error and informational messages produced by the daemon. This trace is always running. It has negligible impact on the performance of the system, under normal circumstances.

The trace is located in: **/var/ha/log/hats.***DD.hhmmss.partition_name.lang* for PSSP nodes, and **/var/ha/log/topsvcs.***DD.hhmmss.cluster_name.lang* for HACMP nodes, where *DD* is the day of the month when the daemon was started, *hhmmss* is the time when the daemon was started, and *lang* is the language used by the daemon.

Data in the user log is in the language where the daemon is run, which is the node’s administrative language. Messages in the user log have a catalog message number, which can be used to obtain a translation of the message in the desired language.

The size of the log file is changed using the same commands that change the size of the service log. Truncation of the log, saving of log files, and other considerations are the same as for the service log.

Each user log entry has this format:

```
date     daemon name       message
```

Adapters are identified by a pair:

```
(IP address:incarnation number)
```

Groups are identified by a pair:

```
(IP address of Group Leader:incarnation number of group)
```

The main source for diagnostics is the AIX error log. Some of the error messages produced in the user log occur under normal circumstances, but if they occur repeatedly they indicate an error. Some error messages give additional detail for an entry in the error log. Therefore, this log file should be examined when an entry is created in the system error log.

hats or topsvcs script log

This is the Topology Services startup script log. It contains configuration data used to build the **machines.lst** configuration file. This log also contains error messages if the script was unable to produce a valid **machines.lst** file and start the daemon. The startup script is run at subsystem startup time and at refresh time.

This log refers to a particular instance of the Topology Services script running on the local node. In PSSP, the control workstation is responsible for building the **machines.lst** file from the adapter information in the SDR. Therefore, it is usually on the control workstation that the startup script encounters problems. In HACMP, the **machines.lst** file is built on every node.

The size of the file varies from 1KB to 50KB according to the size of the machine. The trace runs whenever the startup script runs. The trace is located in:

- **/var/ha/log/hats.***partition_name***.*** for PSSP nodes
- **/var/ha/log/topsvcs.default.*** for HACMP nodes

A new instance of the **hats** startup script log is created each time the script starts. A copy of the script log is made just before the script exits. Only the last seven instances of the log file are kept, and they are named *file***.1** through *file***.7**. Therefore, the contents of the log must be saved before the subsystem is restarted or refreshed many times.

The *file***.1** is an identical copy of the current startup script log. At each startup, *file***.1** is renamed to *file***.2**; *file***.2** is renamed to *file***.3**, and so on. Therefore, the previous *file***.7** is lost.

Entries in the startup script log are kept both in English and in the node's language (if different). Trace records are created for these conditions:

- The **machines.lst** file is retrieved from the SDR.
- The **machines.lst** file is built using information from the SDR.
- An error is encountered that prevents the **hats** script from making progress.

There is no fixed format for the records of the log. The following information is in the file:

- The date and time when the **hats** script started running
- The arguments passed to the script
- A copy of **machines.lst** file generated
- Topology Services tunable parameters
- The date and time when the **hats** script finished running
- The netmasks of the adapters in the configuration

- If the script fails to find an address on the control workstation to include in the configuration, the output of the **netstat -in** and **ifconfig** commands for each Ethernet adapter in the control workstation is included.
- If the script was called for a refresh operation, the output of the **refresh** command is included in the log file.

The following information is in the HACMP startup script log:

- The date when the **topsvcs** script finished running.
- The HACMP version.
- A copy of the **machines.lst** file generated.
- A copy of the output of the **cllsif** command, containing the HACMP adapter configuration.
- The contents of the **HACMPnim** and **HACMPtopsvcs** ODM classes.
- A copy of the output of the **clhandle -ac** command.

The main source for diagnostics is the AIX error log. The **hats** script log file should be used when the error log shows that the startup script was unable to complete its tasks and start the daemon.

For a PSSP system, **good results** are indicated when the message:

```
Exec /usr/sbin/rsct/bin/hatsd -n node_number
```

appears towards the beginning of the file.

For an HACMP system, **good results** are indicated by the absence of fatal error messages.

For a PSSP system, **error results** are indicated by one or more error messages. For example:

```
hats: 2523-605 Cannot find the address of the control workstation.
```

and the absence of the `Exec` message.

For an HACMP system, **error results** are indicated by the presence of the message:

```
topsvcs: 2523-600 Exit with return code: error code
```

and possibly other error messages.

All error messages have the message catalog number with them.

Network Interface Module (NIM) log

This log contains trace information about the activities of the Network Interface Modules (NIMs), which are processes used by the Topology Services daemon to monitor each network interface. These logs need to be collected before they wrap or are removed.

The trace is located in:

- **/var/ha/log/nim.hats.***interface name***.***partition name***.00n** for PSSP nodes.
- **/var/ha/log/nim.topsvcs***interface name***.***cluster name***.00n** for HACMP nodes.

Where **00n** is a sequence number of 001, 002, or 003. These three logs are always kept. Log file 003 is overwritten by 002, 002 is overwritten by 001, and 001 is overwritten by 003.

Trace records are generated under the following conditions:

1. A connection with a given adapter is established.
2. A connection with a given adapter is closed.
3. A daemon has sent a command to start or stop heartbeating.
4. A daemon has sent a command to start or stop monitoring heartbeats.
5. A local adapter goes up or down.
6. A message is sent or received.
7. A heartbeat from the remote adapter has been missed

Data in the NIM log is in English only. The format of each message is:

```
time-of-day    message
```

An instance of the NIM log file will wrap when the file reaches around 200kB. Normally, it takes around 10 minutes to fill an instance of the log file. Since 3 instances are kept, the NIM log files needs to be saved within 30 minutes of when the adapter-related problem occurred.

Information to collect before contacting the IBM Support Center

The following information needs to be collected from the node that presents the problem. For connectivity-related problems, the same information is needed from the other nodes. If collecting data from all the nodes is not feasible, data should be collected from at least the following nodes:

1. The node’s Downstream Neighbor on all networks. This is the node whose IP address is immediately lower than the address of the node where the problem was seen. The node with the lowest IP address has a Downstream Neighbor of the node with the highest IP address.
2. The Group Leader node, which is the node with the highest IP address in the network.
3. The control workstation

Collect the files listed in number 1, 3d on page 386, and 3c on page 386 (bullets 2 and 3). Then, issue **phoenix.snap** to collect the remaining information. See “phoenix.snap dump” on page 376.

1. FFDC dump files: **/var/adm/ffdc/dumps/***
2. Topology Services Files
   - Service and User log files for the daemon.
     – On PSSP, **/var/ha/log/hats.***dd.hhmmss.partition_name***.
     – On HACMP, **/var/ha/log/topsvcs.***dd.hhmmss.cluster_name***.

     Here, *dd* is the day of the month when the daemon was started, and *hhmmss* is the time when the daemon was started.
   - Startup Script log.
     – On PSSP, **/var/ha/log/hats.***partition_name***.
     – On HACMP, **/var/ha/log/topsvcs.default***.
   - Entire contents of the daemon **run** directory.
     – On PSSP, **/var/ha/run/hats.***partition_name***/***.

   - On HACMP, **/var/ha/run/topsvcs.***cluster_name***/***.

     This includes the **machines.lst** file and temporary files used by the startup script. It also includes the log files for the **/usr/sbin/rsct/bin/netmonAdapterHealth** program, which is used to determine adapter status. The daemon’s core files are also stored here.
 - Output of the **lssrc -ls** command.
   - On PSSP nodes, issue the command: **lssrc -ls hats**.
   - On the PSSP control workstation, issue the command: **lssrc -ls hats.***partition_name***.**
   - On HACMP, issue the command: **lssrc -ls topsvcs**.
 - Information used by the startup scripts:
   - On any PSSP node, issue these commands and record the output:
     a. **SDRGetObjects -x SP cw_ipaddrs**
     b. **SDRGetObjects -x TS_Config Frequency Sensitivity Run_FixPri FixPri_Value Log_Length Pinning**
     c. **SDRGetObjects -x Subnet**
     d. **SDRGetObjects -x Network**
     e. **SDRGetObjects -x Adapter**
     f. **SDRGetObjects -G -x Adapter**
     g. **SDRRetrieveFile hats.machines.lst** *local_file_name*
     h. **SDRGetObjects host_responds**
     i. **/usr/lpp/ssp/bin/lsauthpts -c**
     j. **/usr/bin/dcecp -c keytab show rsct/***partition_name***/hats -member**
     k. **/usr/lpp/ssp/bin/splstdata -p**
   - On HACMP nodes, issue these commands and record the output:
     a. **odmget HACMPnim**
     b. **odmget HACMPtopsvcs**
     c. **/usr/es/sbin/cluster/utilities/clhandle -ac**
     d. **/usr/es/sbin/cluster/utilities/clhandle -c**
     e. **/usr/es/sbin/cluster/utilities/cllsif -Sc**
     f. **/usr/es/sbin/cluster/utilities/cllsclstr -Sc**
     g. **/usr/es/sbin/cluster/utilities/clmixver** (output and return value)

3. System Data
   a. Installation data
      - Output of the command: **lslpp -L**
      - Contents of files:
        - **/usr/sbin/rsct/optlevel.rsct.basic.rte**
        - **/usr/sbin/rsct/optlevel.rsct.basic.hacmp**
        - **/usr/sbin/rsct/optlevel.rsct.basic.sp**

   b. The authentication method in use. Issue this command on the control workstation:

```
splstdata -p
```

   The entry ″ts_auth_methods″ lists the active authentication methods in use for the SP trusted services (such as Topology Services).

c. Network and adapter data

- Output of commands:
  1) **netstat -in**
  2) **ifconfig** (all interfaces)
  3) **netstat -s**
  4) **netstat -m**
  5) **netstat -rn**
  6) **netstat -D**
  7) **netstat en*** (for example, en0, en1 and so forth)
- If the problem is still occurring, output of **ping** to addresses in the **machines.lst** file, located in the daemon run directory.
- If the problem is still occurring, output of the **iptrace** command. Follow this sequence:
  1) **iptrace /tmp/iptrace.out**
  2) **ps -ef | grep iptrace**
  3) Wait one minute, and then issue: **kill** *pid*
  4) Save a copy of the file **/tmp/iptrace.out**. This is a binary file.

d. Memory - output of the commands:

1) **vmstat 5 5**
2) **vmstat -s**
3) **vmtune**

The full path name is: **/usr/samples/kernel**.

Diagnostic procedures

These test verify the installation, configuration and operation of Topology Services.

Installation verification test

This test determines whether Topology Services has been successfully installed.

1. Verify if RSCT has been installed. Issue the command:

```
lslpp -l | grep rsct
```

Good results are indicated by an output similar to:

```
  rsct.basic.hacmp    1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.basic.rte      1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.basic.sp       1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.clients.hacmp  1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.clients.rte    1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.clients.sp     1.2.0.0  COMMITTED  RS/6000 Cluster Technology
  rsct.core.utils     1.2.0.0  COMMITTED  RS/6000 Cluster Technology
```

Error results are indicated by no output from the command.

2. Issue the command:

```
lppchk -c "rsct*"
```

Good results are indicated by the absence of error messages and the return of a zero exit status from this command. The command produces no output if it succeeds.

Error results are indicated by a non-zero exit code and by error messages similar to these:

```
lppchk: 0504-206  File /usr/lib/nls/msg/en_US/hats.cat could not be located.
lppchk: 0504-206  File /usr/sbin/rsct/bin/hatsoptions could not be located.
lppchk: 0504-208  Size of /usr/sbin/rsct/bin/phoenix.snap is 29356,
                      expected value was 29355.
```

Some error messages may appear if an EFIX is applied to a file set. An EFIX is an emergency fix, supplied by IBM, to correct a specific problem.

If the test failed, verify the installation of RSCT. The following file sets need to be installed:

1. **rsct.basic.rte**
2. **rsct.core.utils**
3. **rsct.clients.rte**
4. **rsct.basic.sp**
5. **rsct.clients.sp**
6. **rsct.basic.hacmp**
7. **rsct.clients.hacmp**

If the test succeeds, proceed to “Configuration verification tests”. If the test fails, see if RSCT was installed, and install RSCT if it was not.

Configuration verification tests

These tests verify the configuration of Topology Services.

Configuration test 1 - Verify configuration data for PSSP

This test verifies that Topology Services in PSSP has the configuration data it needs. Proceed to “Configuration test 3 - Verify HACMP/ES configuration data” on page 389 if in the HACMP/ES environment.

Issue the following commands to display data from the SDR:

- `SDRGetObjects Syspar`
- `SDRGetObjects SP cw_ipaddrs`
- `SDRGetObjects TS_Config`
- `SDRGetObjects Adapter`
- `SDRGetObjects -G Adapter`

Good results are indicated by none of these commands giving an error message, and all commands giving non-null output. **SDRGetObjects Adapter** must show all the adapters in the current partition, and **SDRGetObjects -G Adapter** must show all the adapters in the machine.

Error results are indicated if these commands fail. The SDR could be experiencing problems. Diagnose the SDR subsystem. If the commands succeed but do not show the expected information, it is possible that a problem occurred in the installation of the nodes. Verify installation of the nodes.

If the test is successful, proceed to “Configuration test 2 - Check control workstation Ethernet adapter”.

Configuration test 2 - Check control workstation Ethernet adapter

This test determines whether the control workstation has an Ethernet adapter that can be included in the Topology Services configuration file. On the control workstation, issue the command **netstat -in**, followed by the command: **ifconfig en***n*, for each Ethernet adapter listed by **netstat**.

Verify that at least one of the ″en″ adapters on the control workstation is on the same subnet ID as the en0 adapter of at least one of the nodes. The subnet ID and subnet mask for the control workstation adapter can be derived from the **ifconfig** command output. Use this calculation:

```
 Subnet id = inet & netmask
 Subnet mask = netmask
```

where *inet* and *netmask* are given in the output of the previous **ifconfig** command , and ″&″ is the bitwise ″AND″ operator. Ignore **ifconfig** command output that begins with **inet6**. Those are addresses in **IPv6** format.

For example, if the command **ifconfig en0** produced this output:

```
en0:  flags=e080863<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,
    GROUPRT,64BIT> inet 9.114.61.125 netmask 0xffffffc0 broadcast 9.114.61.127
```

calculate the Subnet ID as follows:

1. Convert the inet and netmask to hexadecimal notation. Convert each octet separately, and remove the ″.″
2. In this example, the inet is **9.114.61.125**, which converts to 0x09723D7D
3. In this example, the netmask is **0xFFFFFFC0**, which is already in hexadecimal notation. The equivalent dotted decimal form is: 255.255.255.192.
4. Calculate the **subnet ID = inet & netmask**.

   In this example, 0x09723D7D & 0xFFFFFFC0 = **0x09723D40**.
5. Convert the result back to dotted decimal form: 0x09723D40 = **9.114.61.64**. This is the Subnet ID.

The information about the nodes’ adapters can be obtained by issuing the command: **SDRGetObjects -G Adapter**

```
 Subnet id = netaddr & netmask
 Subnet mask = netmask
```

where *netaddr* and *netmask* are given in the output of the **SDRGetObjects** command.

Good results are indicated by the existence of at least one node where the Subnet ID / Subnet mask pairs are the same as in one of the control workstation’s ″en″ adapters.

Error results are indicated by the absence of such a pair.

If the problem is in the control workstation’s or nodes’ netmask, the netmask problem must be corrected. Adapters that belong to the same subnet must have the same netmask. If the problem is due to a lack of an Ethernet adapter in the control workstation which is in the same subnet as one of the nodes, this adapter must be added.

If this test is a success, proceed to “Operational verification tests” on page 390.

Configuration test 3 - Verify HACMP/ES configuration data

This test verifies that Topology Services in HACMP/ES has the configuration data it needs. Configuration data is stored in the HACMP/ES Global ODM. Obtain the output of the following commands:

1. `/usr/es/sbin/cluster/utilities/cllsif`
2. `/usr/es/sbin/cluster/utilities/clhandle`
3. `/usr/es/sbin/cluster/utilities/clhandle -a`
4. `/usr/es/sbin/cluster/utilities/cllsclstr`
5. `odmget HACMPnim`
6. `odmget HACMPtopsvcs`

The output of **cllsif** is similar to the following:

```
 Adapter          Type     Network  NetType Attribute Node  IP Address Hardware Interface Global
  Name                                                                           Address  Name

c47n07             service  SP_ether  ether  public  c47n07  9.114.61.71  en0   glob_SP_ether
c47n07_hpsboot     boot     sw_net    hps    public  c47n07  1.1.1.1      css0
c47n15_hpsservice  service  sw_net    hps    public          1.1.1.12     css0
c47n13_hpsservice  service  sw_net    hps    public          1.1.1.10     css0
```

This output should contain all the adapters defined in the HACMP configuration. The adapter names and addresses (at least the boot and standby) should correspond to what is actually configured on the machine. Those can be obtained by issuing the **netstat -in** command.

The output of **clhandle -a** should show all the nodes configured in HACMP, while the output of **clhandle** should contain the local node name and number.

The output of **cllsclstr** should show the cluster name and id.

The output of **odmget HACMPtopsvcs** should be similar to the following:

```
HACMPtopsvcs:
                                   hbInterval = 1
                                   fibrillateCount = 4
                                   runFixedPri = 1
                                   fixedPriLevel = 38
                                   tsLogLength = 5000
                                   gsLogLength = 5000
                                   instanceNum = 2
```

The output of **odmget HACMPnim** should be similar to the following:

```
HACMPnim:
                                       name = "ether"
                                       desc = "Ethernet Protocol"
                                       addrtype = 0
                                       path = ""
                                       para = ""
                                       grace = 30
                                       hbrate = 500000
                                       cycle = 4

HACMPnim:
                                       name = "token"
                                       desc = "Token Ring Protocol"
                                       addrtype = 0
                                       path = ""
                                       para = ""
                                       grace = 90
                                       hbrate = 500000
                                       cycle = 4
```

Good results are indicated by the output of these commands reflecting the desired HACMP configuration, with respect to networks, network adapters, and tunable values. In this case, proceed to “Operational verification tests”.

Error results are indicated if there is any inconsistency between the displayed configuration data and the desired configuration data. In this case, the HACMP configuration has to be edited, and the Cluster Topology must be synchronized.

Operational verification tests

The following names apply to the operational verification tests in this section:

- Subsystem name:
  - On PSSP nodes **hats**
  - On the PSSP control workstation **hats.***partition_name*
  - On HACMP nodes **topsvcs**
- User log file:
  - On PSSP **/var/ha/log/hats.***dd.hhmmss.partition_name.lang*
  - On HACMP nodes **/var/ha/log/topsvcs.***dd.hhmmss.cluster_name.lang*
- Service log file:
  - On PSSP **/var/ha/log/hats.***dd.hhmmss.partition_name*
  - On HACMP nodes **/var/ha/log/topsvcs.***dd.hhmmss.cluster_name*
- **run** directory:
  - On PSSP **/var/ha/run/hats.***partition_name*
  - On HACMP nodes **/var/ha/run/topsvcs.***cluster_name*
- **machines.lst** file:
  - On PSSP **/var/ha/run/hats.***partition_name***/machines.lst**
  - On HACMP nodes **/var/ha/run/topsvcs.***cluster_name***/machines.***cluster_id***.lst**

Operational test 1 - Verify status and adapters

This test verifies whether Topology Services is working and that all the adapters are up. Issue the following command:

```
lssrc -ls subsystem_name
```

Good results are indicated by an output similar to the following:

```
        Subsystem         Group            PID     Status
         hats             hats             20494   active
        Network Name   Indx Defd Mbrs St Adapter ID      Group ID
        SPether        [ 0]   15   15  S 9.114.61.195    9.114.61.195
        SPether        [ 0] en0          0x3740dd5c      0x3740dd62
        HB Interval = 1 secs. Sensitivity = 4 missed beats
        SPswitch       [ 1]   14   14  S 9.114.61.139    9.114.61.139
        SPswitch       [ 1] css0         0x3740dd5d      0x3740dd62
```

```
HB Interval = 1 secs. Sensitivity = 4 missed beats
  Configuration Instance = 926566126
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  control workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size: 6358 KB. Number of outstanding malloc: 588
  Number of nodes up: 15. Number of nodes down: 0.
```

If the number under the `Mbrs` heading is the same as the number under `Defd`, all adapters defined in the configuration are part of the adapter membership group. The numbers under the `Group ID` heading should remain the same over subsequent invocations of **lssrc** several seconds apart. This is the expected behavior of the subsystem.

Error results are indicated by outputs similar to the following:

1. `0513-036 The request could not be passed to the hats subsystem. Start the subsystem and try your command again.`

   In this case, the subsystem is down. Issue the **errpt** command and look for an entry for the subsystem name. Proceed to “Operational test 2 - Determine why the Topology Services subsystem is inactive” on page 393.

2. `0513-085 The hats Subsystem is not on file.`

   The subsystem is not defined to the AIX SRC. In PSSP, the partition-sensitive subsystems may have been undefined by the **syspar_ctrl** command. The same command may be used to add the subsystems to the node. In HACMP/ES, HACMP may have not been installed on the node. Check the HACMP subsystem.

3. This output requires investigation because the number under `Mbrs` is smaller than the number under `Defd`.

   ```
   Subsystem         Group            PID     Status
    hats             hats             20494   active
   Network Name   Indx Defd Mbrs St Adapter ID      Group ID
   SPether        [ 0]   15    8  S 9.114.61.195    9.114.61.195
   SPether        [ 0] en0          0x3740dd5c      0x3740dd62
   HB Interval = 1 secs. Sensitivity = 4 missed beats
   SPswitch       [ 1]   14    7  S 9.114.61.139    9.114.61.139
   SPswitch       [ 1] css0         0x3740dd5d      0x3740dd62
   HB Interval = 1 secs. Sensitivity = 4 missed beats
     Configuration Instance = 926566126
     Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
     control workstation IP address = 9.114.61.125
     Daemon employs no security
     Data segment size: 6358 KB. Number of outstanding malloc: 588
     Number of nodes up: 8. Number of nodes down: 7.
     Nodes down: 17-29(2)
   ```

   Some remote adapters are not part of the local adapter’s group. Proceed to “Operational test 3 - Determine why remote adapters are not in the local adapter’s membership group” on page 393.

4. This output requires investigation because a local adapter is disabled.

   ```
   Subsystem         Group            PID     Status
    hats             hats             20494   active
   Network Name   Indx Defd Mbrs St Adapter ID      Group ID
   SPether        [ 0]   15   15  S 9.114.61.195    9.114.61.195
   SPether        [ 0] en0          0x3740dd5c      0x3740dd62
   HB Interval = 1 secs. Sensitivity = 4 missed beats
   SPswitch       [ 1]   14    0  D 9.114.61.139
   SPswitch       [ 1] css0
   ```

```
HB Interval = 1 secs. Sensitivity = 4 missed beats
  Configuration Instance = 926566126
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  control workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size: 6358 KB. Number of outstanding malloc: 588
  Number of nodes up: 15. Number of nodes down: 0.
```

A local adapter is disabled. Proceed to "Operational test 4 - Check address of local adapter" on page 394.

5. This output requires investigation because there is a **U** below the `St` heading.

```
Subsystem         Group            PID     Status
 hats             hats             20494   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
SPether        [ 0]   15    8  S 9.114.61.195    9.114.61.195
SPether        [ 0] en0          0x3740dd5c      0x3740dd62
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [ 1]   14    1  U 9.114.61.139    9.114.61.139
SPswitch       [ 1] css0         0x3740dd5d      0x3740dd5d
HB Interval = 1 secs. Sensitivity = 4 missed beats
  Configuration Instance = 926566126
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  control workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size: 6358 KB. Number of outstanding malloc: 588
  Number of nodes up: 8. Number of nodes down: 7.
  Nodes down: 17-29(2)
```

The last line of the output shows a list of nodes that are either up or down, whichever is smaller. The list of nodes that are down includes only the nodes that are configured and have at least one adapter that Topology Services monitors. Nodes are specified by a list of node ranges, as follows:

```
N1-N2(I1)  N3-N4(I2) ...
```

Here, there are two ranges, *N1*-*N2*(*I1*) and *N3*-*N4*(*I2*). They are interpreted as follows:

- *N1* is the first node in the first range
- *N2* is the last node in the first range
- *I1* is the increment for the first range
- *N3* is the first node in the second range
- *N4* is the last node in the second range
- *I2* is the increment for the second range

If the increment is 1, it is omitted. If the range has only one node, only that node's number is displayed. Examples are:

a. `Nodes down: 17-29(2)` means that nodes 17 through 29 are down. In other words, nodes 17, 19, 21, 23, 25, 27, and 29 are down.

b. `Nodes up: 5-9(2) 13` means that nodes 5, 7, 9, and 13 are up.

c. `Nodes up: 5-9 13-21(4)` means that nodes 5, 6, 7, 8, 9, 13, 17, and 21 are up.

An adapter stays in a singleton unstable membership group. This normally occurs for a few seconds after the daemon starts or after the adapter is re-enabled. If the situation persists for more than one minute, this may indicate a problem. This usually indicates that the local adapter is receiving some messages, but it is unable to obtain responses for its outgoing messages. Proceed to “Operational test 7 - Check for partial connectivity” on page 397.

6. An output similar to the expected output, or similar to output 3 on page 391, but where the numbers under the `Group ID` heading (either the address of the Group Leader adapter or the ″incarnation number″ of the group) change every few seconds without ever becoming stable.

   This kind of output indicates that there is some partial connectivity on the network. Some adapters may be able to communicate only with a subset of adapters. Some adapters may be able to send messages only or receive messages only. This output indicates that the adapter membership groups are constantly reforming, causing a substantial increase in the CPU and network resources used by the subsystem.

   A partial connectivity situation is preventing the adapter membership group from holding together. Proceed to “Operational test 10 - Check neighboring adapter connectivity” on page 400.

If this test is successful, proceed to “Operational test 11 - Verify node reachability information” on page 400.

Operational test 2 - Determine why the Topology Services subsystem is inactive

This test is to determine why the Topology Services subsystem is not active.

- For PSSP, issue command: **errpt -N ″hats*″ -a**
- For HACMP/ES, issue command: **errpt -N topsvcs -a**

The AIX error log entries produced by this command, together with their description in Table 58 on page 354, explain why the subsystem is inactive. If no entry that explains why the subsystem went down or could not start exists, it is possible that the daemon may have exited abnormally.

In this case, issue the **errpt -a** command and look for an error. Look for an error entry with a LABEL: of CORE_DUMP and PROGRAM NAME of **hatsd**. (Issue the command: **errpt -J CORE_DUMP -a**.) If such an entry is found, see “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.

Another possibility when there is no **TS_** error log entry, is that the Topology Services daemon could not be loaded. In this case a message similar to the following may be present in the Topology Services User startup log:

```
0509-036 Cannot load program hatsd because of the following errors:
0509-023 Symbol dms_debug_tag in hatsd is not defined.
0509-026 System error: Cannot run a file that does not have a valid format.
```

The message may refer to the Topology Services daemon, or to some other program invoked by the startup script **hats**. If such an error is found, contact the IBM Support Center.

For errors where the daemon did start up but exited during initialization, detailed information about the problem is in the Topology Services User error log.

Operational test 3 - Determine why remote adapters are not in the local adapter’s membership group

Issue the command:

```
lssrc -ls subsystem
```

on all the nodes and the PSSP control workstation. The command:

```
dsh -a "lssrc -ls subsystem"
```

issued from the control workstation can be used to issue **lssrc** on all the nodes.

If this test follows output 3 on page 391, at least one node will not have the same output as the node from where output 3 on page 391 was taken.

Some of the possibilities are:

1. The node is down or unreachable. Diagnose that node by using “Operational test 1 - Verify status and adapters” on page 390.
2. The output is similar to output of 3 on page 391, but with a different group id, such as in this output:

   ```
   Subsystem         Group            PID     Status
    hats             hats             20494   active
   Network Name   Indx Defd Mbrs St Adapter ID      Group ID
   SPether        [ 0]   15    7  S 9.114.61.199    9.114.61.201
   SPether        [ 0] en0          0x3740dd5c      0x3740dd72
   HB Interval = 1 secs. Sensitivity = 4 missed beats
   SPswitch       [ 1]   14    7  S 9.114.61.141    9.114.61.141
   SPswitch       [ 1] css0         0x3740dd5d      0x3740dd72
   HB Interval = 1 secs. Sensitivity = 4 missed beats
     Configuration Instance = 926566126
     Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
     Control Workstation IP address = 9.114.61.125
     Daemon employs no security
     Data segment size: 6358 KB. Number of outstanding malloc: 588
     Number of nodes up: 7. Number of nodes down: 8.
     Nodes up: 17-29(2)
   ```

   Compare this with the output from 3 on page 391. Proceed to “Operational test 8 - Check if configuration instance and security status are the same across all nodes” on page 398.
3. The output is similar to the outputs of 1 on page 391, 2 on page 391, 4 on page 391, or 5 on page 392. Return to “Operational test 1 - Verify status and adapters” on page 390, but this time focus on this new node.

Operational test 4 - Check address of local adapter

This test verifies whether a local adapter is configured with the correct address. Assuming that this test is being run because the output of the **lssrc** command indicates that the adapter is disabled, there should be an entry in the AIX error log that points to the problem.

Issue the command:

```
errpt -J TS_LOC_DOWN_ST,TS_MISCFG_EM -a | more
```

Examples of the error log entries that appear in the output are:

-
  ```
  LABEL:          TS_LOC_DOWN_ST
  IDENTIFIER:     D17E7B06

  Date/Time:       Mon May 17 23:29:34
  Sequence Number: 227
  Machine Id:      000032054C00
  Node Id:         c47n11
  Class:           S
  Type:            INFO
  ```

```
        Resource Name:   hats.c47s

        Description
        Possible malfunction on local adapter
```

-
  ```
          LABEL:          TS_MISCFG_EM
          IDENTIFIER:     6EA7FC9E

          Date/Time:       Mon May 17 16:28:45
          Sequence Number: 222
          Machine Id:      000032054C00
          Node Id:         c47n11
          Class:           U
          Type:            PEND
          Resource Name:   hats.c47s
          Resource Class:  NONE
          Resource Type:   NONE
          Location:        NONE
          VPD:

          Description
          Local adapter misconfiguration detected
  ```

Good results are indicated by the absence of the **TS_MISCFG_EM** error entry. To verify that the local adapter has the expected address, issue the command:

```
ifconfig interface_name
```

where *interface_name* is the interface name listed on the output of **lssrc**, such as:

```
        SPswitch       [ 1]   14    0  D 9.114.61.139
        SPswitch       [ 1] css0
```

For the **lssrc** command output, the output of **ifconfig css0** is similar to:

```
css0: flags=800847 <UP,BROADCAST,DEBUG,RUNNING,SIMPLEX>
        inet 9.114.61.139 netmask 0xffffffc0 broadcast 9.114.61.191
```

Error results are indicated by the **TS_MISCFG_EM** error entry and by the output of the **ifconfig** command not containing the address displayed in the **lssrc** command output.

Diagnose the reason why the adapter is configured with an incorrect address. For PSSP, the adapter may have been incorrectly configured in the SDR, or the adapter's address was incorrectly set manually. For HACMP, the cluster on the node may have been stopped with the ″Forced Down″ option. The adapters must be configured with their boot-time addresses before the cluster can be started on a node. This can be done by issuing command:

```
/etc/rc.net -boot
```

several times in a sequence. Issuing the command only once may not set all IP routes correctly.

If this test is a success, proceed to "Operational test 5 - Check if the adapter is enabled for IP".

Operational test 5 - Check if the adapter is enabled for IP

Issue the command:

```
ifconfig interface_name
```

The output is similar to the following:

```
 css0: flags=800847 <UP,BROADCAST,DEBUG,RUNNING,SIMPLEX>
         inet 9.114.61.139 netmask 0xffffffc0 broadcast 9.114.61.191
```

Good results are indicated by the presence of the UP string in the first line of the output. In this case, proceed to "Operational test 6 - Check whether the adapter can communicate with other adapters in the network".

Error results are indicated by the absence of the UP string in the first line of the output.

Issue the command:

```
ifconfig interface_name up
```

to re-enable the adapter to IP.

Operational test 6 - Check whether the adapter can communicate with other adapters in the network

Root authority is needed to access the contents of the **machines.lst** file. Display the contents of the **machines.lst** file. The output is similar to the following:

```
*InstanceNumber=925928580
*configId=1244520230
*!HaTsSeCStatus=off
*FileVersion=1
*!TS_realm=PSSP
TS_Frequency=1
TS_Sensitivity=4
TS_FixedPriority=38
TS_LogLength=5000
*!TS_PinText
Network Name SPether
Network Type ether
*
*Node Type Address
    0 en0 9.114.61.125
    1 en0  9.114.61.65
    3 en0  9.114.61.67
    11 en0  9.114.61.195
...
Network Name SPswitch
Network Type hps
*
*Node Type Address
    1 css0 9.114.61.129
    3 css0 9.114.61.131
    11 css0 9.114.61.139
```

Locate the network to which the adapter under investigation belongs. For example, the css0 adapter on node 11 belongs to network SPswitch. Issue the command:

```
ping -c 5 address
```

for the addresses listed in the **machines.lst** file.

Good results are indicated by outputs similar to the following.

```
PING 9.114.61.129: (9.114.61.129): 56 data bytes
64 bytes from 9.114.61.129: icmp_seq=0 ttl=255 time=0 ms
64 bytes from 9.114.61.129: icmp_seq=1 ttl=255 time=0 ms
64 bytes from 9.114.61.129: icmp_seq=2 ttl=255 time=0 ms
64 bytes from 9.114.61.129: icmp_seq=3 ttl=255 time=0 ms
64 bytes from 9.114.61.129: icmp_seq=4 ttl=255 time=0 ms

----9.114.61.129 PING Statistics----
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max = 0/0/0 ms
```

The number before `packets received` should be greater than 0.

Error results are indicated by outputs similar to the following:

```
PING 9.114.61.129: (9.114.61.129): 56 data bytes

----9.114.61.129 PING Statistics----
5 packets transmitted, 0 packets received, 100% packet loss
```

The command should be repeated with different addresses until it succeeds or until several different attempts are made. After that, pursue the problem as an adapter or IP-related problem. If the adapter is an SP Switch adapter, refer to “Chapter 15. Diagnosing SP Switch problems” on page 137. If the adapter is an SP Switch2 adapter, refer to “Chapter 16. Diagnosing SP Switch2 problems” on page 185.

If this test succeeds, but the adapter is still listed as disabled in the **lssrc** command output, collect the data listed in “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.

Operational test 7 - Check for partial connectivity

Adapters stay in a singleton unstable state when there is partial connectivity between two adapters. One reason for an adapter to stay in this state is that it keeps receiving PROCLAIM messages, to which it responds with a JOIN message, but no PTC message comes in response to the JOIN message.

Check in the User log file to see if a message similar to the following appears repeatedly:

```
2523-097 JOIN time has expired. PROCLAIM message was sent
                   by (10.50.190.98:0x473c6669)
```

If this message appears repeatedly in the User log, investigate IP connectivity between the local adapter and the adapter whose address is listed in the User log entry (10.50.190.98 in the example here). Issue command:

```
ping -c 5 address
```

address is 10.50.190.98 in this example.

See “Operational test 5 - Check if the adapter is enabled for IP” on page 395 for a description of **good results** for this command.

The local adapter cannot communicate with a Group Leader that is attempting to attract the local adapter into the adapter membership group. The problem may be with either the local adapter or the Group Leader adapter (″proclaimer″ adapter).

Pursue this as an IP connectivity problem. Focus on both the local adapter and the Group Leader adapter. See “Chapter 13. Diagnosing IP routing problems” on page 123.

If the **ping** command succeeds, but the local adapter still stays in the singleton unstable state, contact the IBM Support Center.

In an HACMP/ES environment, it is possible that there are two adapters in different nodes both having the same service address. This can be verified by issuing:

```
lssrc -ls subsystem_name
```

and looking for two different nodes that have the same IP address portion of `Adapter ID`. In this case, this problem should be pursued as an HACMP/ES problem. Contact the IBM Support Center.

If this test fails, proceed to “Operational test 4 - Check address of local adapter” on page 394, concentrating on the local and Group Leader adapters.

Operational test 8 - Check if configuration instance and security status are the same across all nodes

This test is used when there seem to be multiple partitioned adapter membership groups across the nodes, as in output 2 on page 394.

This test verifies whether all nodes are using the same configuration instance number and same security setting. The instance number changes each time the **machines.lst** file is generated by the startup script. In PSSP, the configuration instance always increases. In HACMP/ES, the configuration instance number normally increases, unless a snapshot of a previous configuration is applied.

Issue the command:

```
lssrc -ls subsystem_name
```

on all nodes. If this is not feasible, issue the command at least on nodes that produce an output that shows a different `Group ID`.

Compare the line `Configuration Instance = (number)` in the **lssrc** outputs. Also, compare the line `Daemon employs` in the **lssrc** command outputs.

Good results are indicated by the number after the `Configuration Instance` phrase being the same in all the **lssrc** outputs. This means that all nodes are working with the same version of the **machines.lst** file.

Error results are indicated by the configuration instance being different in the two ″node partitions″ (this is unrelated to the SP system partitions). In this case, the adapters in the two partitions cannot merge into a single group because the configuration instances are different across the node partitions. This situation is likely to be caused by a refresh-related problem. One of the node groups, probably that with the lower configuration instance, was unable to run a refresh. If a refresh operation was indeed attempted, consult the description of the ″Nodes or adapters leave membership after refresh″ problem in “Error symptoms, responses, and recoveries” on page 402.

The situation may be caused by a problem in the AIX SRC subsystem, which fails to notify the Topology Services daemon about the refresh. The description of the

″Nodes or adapters leave membership after refresh″ problem in “Error symptoms, responses, and recoveries” on page 402 explains how to detect the situation where the Topology Services daemon has lost its connection with the AIX SRC subsystem. In this case, contact the IBM Support Center.

If the security setting is not the same on all the nodes in a partition, some of the nodes may fail to authenticate each other’s messages. AIX error log entries with labels **TS_SECURITY_ST** and **TS_SECURITY2_ST** may appear on those nodes. For information about these error log entries, see **TS_SECURITY_ST** on page 371 and **TS_SECURITY2_ST** on page 371.

If this test is successful, proceed to “Operational test 9 - Check connectivity among multiple node partitions”.

Operational test 9 - Check connectivity among multiple node partitions

This test is used when adapters in the same Topology Services network form multiple adapter membership groups, rather than a single group encompassing all the adapters in the network.

Follow the instructions in “Operational test 8 - Check if configuration instance and security status are the same across all nodes” on page 398 to obtain **lssrc** outputs for each of the node partitions.

The IP address listed in the **lssrc** command output under the `Group ID` heading is the IP address of the Group Leader. If two node partitions are unable to merge into one, this is caused by the two Group Leaders being unable to communicate with each other. Note that even if some adapters in different partitions can communicate, the group merge will not occur unless the Group Leaders are able to exchange point-to-point messages. Use **ping** (as described in “Operational test 6 - Check whether the adapter can communicate with other adapters in the network” on page 396) to determine whether the Group Leaders can communicate with each other.

For example, assume on one node the output of the **lssrc -ls hats** command is:

```
Subsystem         Group            PID     Status
 hats             hats             15750   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
SPether        [0]   15    9   S 9.114.61.65     9.114.61.195
SPether        [0]              0x373897d2      0x3745968b
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [1]   14    14  S 9.114.61.129    9.114.61.153
SPswitch       [1]              0x37430634      0x374305f1
HB Interval = 1 secs. Sensitivity = 4 missed beats
```

and on another node it is:

```
Subsystem         Group            PID     Status
 hats             hats             13694   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
SPether        [0]  15    6  S 9.114.30.69     9.114.61.71
SPether        [0]              0x37441f24      0x37459754
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [1]   14   14  S 9.114.61.149    9.114.61.153
SPswitch       [1]             0x374306a4      0x374305f1
```

In this example, the partition is occurring in the SP Ethernet. The two Group Leaders are IP addresses 9.114.61.195 and 9.114.61.71. Login to the node that hosts one of the IP addresses and issue the **ping** test to the other address. In case the two adapters in question are in the same subnet, verify whether they have the same subnet mask. “Configuration test 2 - Check control workstation Ethernet adapter” on page 388 describes how to obtain the subnet id and subnet mask for an adapter.

Good results and **error results** for the **ping** test are described in “Operational test 6 - Check whether the adapter can communicate with other adapters in the network” on page 396. If the **ping** test is not successful, a network connectivity problem between the two Group Leader nodes is preventing the groups from merging. Diagnose the network connectivity problem. See “Chapter 12. Diagnosing system connectivity problems” on page 121.

Good results for the subnet mask test are indicated by the adapters that have the same subnet id also having the same subnet mask. If the subnet mask test fails, the subnet mask at one or more nodes must be corrected by issuing the command:

```
ifconfig interface_name address netmask netmask
```

All the adapters that belong to the same subnet must have the same subnet mask.

If the **ping** test is successful (the number of `packets received` is greater than 0), and the subnet masks match, there is some factor other than network connectivity preventing the two Group Leaders from contacting each other. The cause of the problem may be identified by entries in the Topology Services User log. If the problem persists, collect the data listed in “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center. Include information about the two Group Leader nodes.

Operational test 10 - Check neighboring adapter connectivity

This test checks neighboring adapter connectivity, in order to investigate partial connectivity situations. Issue the command **errpt -J TS_DEATH_TR | more** on all the nodes. Look for recent entries with label **TS_DEATH_TR**. This is the entry created by the subsystem when the local adapter stops receiving heartbeat messages from the neighboring adapter. For the adapter membership groups to be constantly reforming, such entries should be found in the error log.

Issue the **ping** test on the node where the **TS_DEATH_TR** entry exists. The target of the **ping** should be the adapter whose address is listed in the Detail Data of the AIX error log entry. “Operational test 6 - Check whether the adapter can communicate with other adapters in the network” on page 396 describes how to perform the **ping** test and interpret the results.

If the **ping** test fails, this means that the two neighboring adapters have connectivity problems, and the problem should be pursued as an IP connectivity problem.

If the **ping** test is successful, the problem is probably not due to lack of connectivity between the two neighboring adapters. The problem may be due to one of the two adapters not receiving the COMMIT message from the ″mayor adapter″ when the group is formed. The **ping** test should be used to probe the connectivity between the two adapters and all other adapters in the local subnet.

Operational test 11 - Verify node reachability information

Issue the following command:

```
lssrc -ls subsystem_name
```

and examine lines:

1. `Number of nodes up: # . Number of nodes down: #.`
2. `Nodes down: [...]` or `Nodes up: [...]`

in the command output.

Good results are indicated by the line `Number of Nodes down: 0`. For example,

```
Number of nodes up: 15    Number of nodes down: 0
```

However, such output can only be considered correct if indeed all nodes in the system are known to be up. If a given node is indicated as being up, but the node seems unresponsive, perform problem determination on the node. Proceed to “Operational test 12 - Verify the status of an unresponsive node that Is shown to be up by Topology Services”.

Error results are indicated by `Number of Nodes down:` being nonzero. The list of nodes that are flagged as being up or down is given in the next output line. An output such as `Nodes down: 17-23(2)` indicates that nodes 17, 19, 21, and 23 are considered down by Topology Services. If the nodes in the list are known to be down, this is the expected output. If, however, some of the nodes are thought to be up, it is possible that a problem exists with the Topology Services subsystem on these nodes. Proceed to “Operational test 1 - Verify status and adapters” on page 390, focusing on each of these nodes.

Operational test 12 - Verify the status of an unresponsive node that Is shown to be up by Topology Services

Examine the **machines.lst** configuration file and obtain the IP addresses for all the adapters in the given node that are in the Topology Services configuration. For example, for node 9, entries similar to the following may be found in the file:

```
 9 en0  9.114.61.193
 9 css0 9.114.61.137
```

Issue this command.

```
ping -c5 IP_address
```

If there is no response to the **ping** packets (the output of the command shows `100% packet loss`) for all the node’s adapters, the node is either down or unreachable. Pursue this as a node health problem. If Topology Services still indicates the node as being up, contact the IBM Support Center because this is probably a Topology Services problem. Collect long tracing information from the Topology Services logs. See “Topology Services service log” on page 379. Also obtain **iptrace** information from the node where the test is being run. See “Information to collect before contacting the IBM Support Center” on page 384.

If the output of the **ping** command shows some response (for example, `0% packet loss`), the node is still up and able to send and receive IP packets. The Topology Services daemon is likely to be running and able to send and receive heartbeat packets. This is why the node is still seen as being up. This problem should be pursued as an AIX-related problem.

If there is a response from the **ping** command, and the node is considered up by remote Topology Services daemons, but the node is unresponsive and no user application is apparently able to run, a system dump must be obtained to find the cause of the problem. See “Chapter 5. Producing a system dump” on page 81.

In a PSSP environment, make an attempt to connect to the node using the serial line interface. Issue this command:

```
spmon -o nodenode_number
```

If the connection is successful, the problem is likely to be lack of IP connectivity to the node. If the connection is not successful, a system dump is needed to diagnose the problem.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the Topology Services component of RSCT. Locate the symptom and perform the action described in the following table.

Table 59. Topology Services symptoms

Symptom	Recovery
Adapter membership groups do not include all the nodes in the configuration.	See “Operational test 1 - Verify status and adapters” on page 390.
Topology Services subsystem fails to start.	See “Action 1 - Investigate startup failure”.
The refresh operation fails or has no effect.	See “Action 2 - Investigate refresh failure” on page 403.
A local adapter is notified as being down by Topology Services.	See “Action 3 - Correct local adapter problem” on page 404.
Adapters appear to be going up and down continuously.	See “Action 4 - Investigate partial connectivity problem” on page 405.
A node appears to go down and then up a few seconds later.	See “Action 5 - Investigate hatsd problem” on page 406.
Adapter (or **host_responds**) appears to go down and then up a few seconds later.	See “Action 6 - Investigate IP communication problem” on page 413.
Group Services exits abnormally because of a Topology Services Library error. Error log entry with template **GS_TS_RETCODE_ER** is present.	See “Action 7 - Investigate Group Services failure” on page 413.
A node running HACMP/ES crashes and produces an AIX dump. System Dump analysis reveals a ″panic″ in function **haDMS_kex:dead_man_sw_handler**.	See “Action 8 - Investigate node crash” on page 414.
Nodes or adapters leave membership after a refresh.	See “Action 9 - Investigate problems after a refresh” on page 415.
If DCE is the only active authentication method, and it is suspected that the Topology Services authentication keys has been compromised.	See “Action 10 - Correct authentication keys” on page 418.

Actions

Action 1 - Investigate startup failure

Some of the possible causes are:

- SDR-related problems that prevent the startup script from obtaining configuration data from the SDR.

- Adapter configuration problems, such as duplicated IP addresses in the configuration or no control workstation Ethernet adapter in the same subnet as the nodes.
- AIX-related problems, such as a shortage of space in the **/var** directory or a port number already in use.
- SP Security Services problems that prevent Topology Services from obtaining credentials, determining the active authentication method, or determining the authentication keys to use.

See “Operational test 2 - Determine why the Topology Services subsystem is inactive” on page 393. To verify the correction, see “Operational test 1 - Verify status and adapters” on page 390.

Action 2 - Investigate refresh failure

The most probable cause is that an incorrect adapter or network configuration was passed to Topology Services. In PSSP, the following command:

```
/usr/sbin/rsct/bin/hatsctrl -r
```

may produce a message similar to:

```
2523-300 Refresh operation failed because of errors in machines.lst file.
```

The same message will be present in the startup script log. Another message may also appear:

```
hatsctrl: 2523-646 Refresh operation failed.
              Details are in the AIX Error Log and in the hats script log
              (/var/ha/log/hats.partition_name).
```

Also, configuration errors result in AIX Error Log entries being created. Some of the template labels that may appear are:

- TS_SPNODEDUP_ER
- TS_HANODEDUP_ER
- TS_SPIPDUP_ER
- TS_HAIPDUP_ER
- TS_CWSADDR_ER
- TS_SDR_ER

In addition, when DCE is the only active authentication method for SP trusted services, refresh may fail if the Topology Services startup script fails to get DCE credentials to update the SDR. When this happens, an AIX error log entry with the following label may appear: **TS_DCECRED_ER**.

The AIX error log entries should provide enough information to determine the cause of the problem. Detailed information about the configuration and the error can be found in the startup script log and the user log.

For the problems that result in the error log entries listed here, the solution involves changing the IP address of one or more adapters. The procedure is different depending on whether the problem occurs in PSSP or HACMP. After the adapter configuration problem is fixed, a new refresh operation can be attempted. On PSSP, the command to use is the **hatsctrl** command described previously.

On HACMP, the following sequence results in a Topology Services refresh:

```
smit hacmp
                   Cluster Configuration
                       Cluster Topology
                           Synchronize Cluster Topology
```

Good results are indicated by the lack of the error message 2523-300 or similar messages, and the lack of error AIX log entries listed earlier. **Good results** are also indicated by a change in the `Configuration Instance`, when checked using these steps:

1. Issue this command: **lssrc -ls** *subsystem_name* before the refresh.
2. Wait until one minute after the refresh completes.
3. Issue the command again: **lssrc -ls** *subsystem_name*.
4. Verify that the number near the `Configuration Instance` is different. The number remains unchanged when the refresh fails.

If the **lssrc** command is issued right after the refresh, text similar to the following may appear as part of the output:

```
Daemon is in a refresh quiescent period.
    Next Configuration Instance = 926456205
```

This message indicates that the refresh operation is in progress.

Error results are indicated by messages 2523-300 and 2523-646, and error log entries listed previously.

Action 3 - Correct local adapter problem

Probable causes of this problem are:

1. The adapter is not working.
2. The network may be down.
3. In the case of an SP Switch or SP Switch2 adapter, the switch may be down, or the local adapter may be fenced.
4. The adapter may have been configured with an incorrect IP address.
5. Topology Services is unable to get response packets back to the adapter.
6. There is a problem in the subsystem's ″adapter self-death″ procedures.

See "Operational test 4 - Check address of local adapter" on page 394 to analyze the problem. The repair action depends on the nature of the problem. For problems 1 through problem 4, the underlying cause for the adapter to be unable to communicate must be found and corrected.

For problem 5, Topology Services requires that at least one other adapter in the network exist, so that packets can be exchanged between the local and remote adapters. Without such an adapter, a local adapter would be unable to receive any packets. Therefore, there would be no way to confirm that the local adapter is working.

Note that for configurations having only two nodes, when a node or a node's switch adapter fails, the switch adapter on the other node will also be flagged as down. This is because the remaining adapter will have no other adapter to communicate with. The same is also true with two-node SP system partitions.

To verify the repair, issue the **lssrc** command as described in "Operational test 1 - Verify status and adapters" on page 390. If the problem is due to Topology Services being unable to obtain response packets back to the adapter (problem 5 on page 404), the problem can be circumvented by adding machine names to file **/usr/sbin/cluster/netmon.cf**.

These machines should be routers or any machines that are external to the configuration, but are in one of the networks being monitored by the subsystem. Any entry in this file is used as a target for a probing packet when Topology Services is attempting to determine the health of a local adapter. The format of the file is as follows:

```
machine name or IP address 1
machine name or IP address 2
..........
```

where the IP addresses are in dotted decimal format. The use of this file is explained in *HACMP: Planning Guide*. If the file does not exist, it should be created. To remove this recovery action, remove the entries added to the file, delete the file, or rename the file.

Action 4 - Investigate partial connectivity problem

The most probable cause is a partial connectivity scenario. This means that one adapter or a group of adapters can communicate with some, but not all, remote adapters. Stable groups in Topology Services require that all adapters in a group be able to communicate with each other.

Some possible sources of partial connectivity are:

1. Physical connectivity
2. Incorrect routing at one or more nodes
3. Adapter or network problems which result in packets larger than a certain size being lost
4. Incorrect ARP setting in large machine configurations

   The total number of entries in the ARP table must be a minimum of two times the number of nodes. The number of entries in the ARP table is calculated by multiplying the **arptab_bsiz** parameter by the **arptab_nb** parameter. The parameters **arptab_bsiz** and **arptab_nb** are tunable parameters controlled by the AIX **no** (**n**etwork **o**ptions) command.
5. High network traffic, which causes a significant portion of the packets to be lost.

To check whether there is partial connectivity on the network, run “Operational test 10 - Check neighboring adapter connectivity” on page 400. The underlying connectivity problem must be isolated and corrected. To verify the correction, issue the **lssrc** command from “Operational test 1 - Verify status and adapters” on page 390.

The problem can be bypassed if the connectivity test revealed that one or more nodes have only partial connectivity to the others. In this case, Topology Services can be stopped on these partial connectivity nodes. If the remaining adapters in the network have complete connectivity to each other, they should form a stable group.

Topology Services subsystem can be stopped on a node by issuing the command:

```
/usr/sbin/rsct/bin/hatsctrl -k
```

Note that the nodes where the subsystem was stopped will be marked as down by the others. Applications such as IBM Virtual Shared Disk and GPFS will be unable to use these nodes.

To test and verify this recovery, issue the **lssrc** command as described in "Operational test 1 - Verify status and adapters" on page 390. The `Group ID` information in the output should not change across two invocations approximately one minute apart.

Once this recovery action is no longer needed, restart Topology Services by issuing this command:

```
/usr/sbin/rsct/bin/hatsctrl -s
```

Action 5 - Investigate hatsd problem

Probable causes of this problem are:

1. The Topology Services daemon is temporarily blocked.
2. The Topology Services daemon exited on the node.
3. IP communication problem, such as mbuf shortage or excessive adapter traffic.

Probable cause 1 can be determined by the presence of an AIX error log entry with **TS_LATEHB_PE** template on the affected node. This entry indicates that the daemon was blocked and for how long. When the daemon is blocked, it cannot send messages to other adapters, and as a result other adapters may consider the adapter dead in each adapter group. This results in the node being considered dead.

The following are some of the reasons for the daemon to be blocked:

1. A memory shortage, which causes excessive paging and thrashing behavior; the daemon stays blocked, awaiting a page-in operation.
2. A memory shortage combined with excessive disk I/O traffic, which results in slow paging operations.
3. The presence of a fixed-priority process with higher priority than the Topology Services daemon, which prevents the daemon from running.
4. Excessive interrupt traffic, which prevents any process in the system from being run in a timely manner.

A memory shortage is usually detected by the **vmstat** command. Issue the command:

```
vmstat -s
```

to display several memory-related statistics. Large numbers for `paging space page ins` or `paging space page outs` (a significant percentage of the `page ins` counter) indicate excessive paging.

Issue the command: **vmstat 5 7** to display some virtual memory counters over a 30-second period or time. If the number of free pages (number under the `fre` heading) is close to 0 (less than 100 or so), this indicates excessive paging. A nonzero value under `po` (pages paged out to paging space) occurring consistently also indicates heavy paging activity.

In a system which appears to have enough memory, but is doing very heavy I/O operations, it is possible that the virtual memory manager may ″steal″ pages from processes (″computational pages″) and assign them to file I/O (″permanent pages″). In this case, to allow more computational pages to be kept in memory, the **vmtune** command can be used to change the proportion of computational pages and permanent pages.

The same command can also be used to increase the number of free pages in the node, below which the virtual memory manager starts stealing pages and adding them to the free list. Increasing this number should prevent the number of free pages from reaching zero, which would force page allocation requests to wait. This number is controlled by the **minfree** parameter of the **vmtune** command.

Command:

```
/usr/samples/kernel/vmtune -f 256 -F 264 -p 1 -P 2
```

can be used to increase **minfree** to 256 and give more preference to computational pages. More information is in the **minfree** parameter description of the Appendix ″Summary of Tunable AIX Parameters″, in *AIX Versions 3.2 and 4 Performance Tuning Guide*.

If the reason for the blockage cannot be readily identified, AIX tracing can be set up for when the problem recurs. The command:

```
/usr/bin/trace -a -l -L 16000000 -T 8000000 -o /tmp/trace_raw
```

should be run in all the nodes where the problem is occurring. Enough space for a 16MB file should be reserved on the file system where the trace file is stored (**/tmp** in this example).

The trace should be stopped with the command:

```
/usr/bin/trcstop
```

as soon as the **TS_LATEHB_PE** entry is seen in the AIX error log. The resulting trace file and the **/unix** file should be saved for use by the IBM Support Center.

The underlying problem that is causing the Topology Services daemon to be blocked must be understood and solved. Problems related to memory thrashing behavior are addressed by *AIX Versions 3.2 and 4 Performance Tuning Guide*. In most cases, obtaining the AIX trace for the period that includes the daemon blockage (as outlined previously) is essential to determine the source of the problem.

For problems related to memory thrashing, it has been observed that if the Topology Services daemon is unable to run in a timely manner, this indicates that the amount of paging is causing little useful activity to be accomplished on the node.

Memory contention problems in Topology Services can be reduced by using the AIX Workload Manager. See “Preventing memory contention problems with the AIX Workload Manager” on page 411.

For problems related to excessive disk I/O, these steps can be taken in AIX to reduce the I/O rate:

1. Set I/O pacing.

I/O pacing limits the number of pending write operations to file systems, thus reducing the disk I/O rate. AIX is installed with I/O pacing disabled. I/O pacing can be enabled with the command:

```
chdev -l sys0 -a maxpout='33' -a minpout='24'
```

This command sets the high-water and low-water marks for pending write-behind I/Os per file. The values can be tuned if needed.

2. Change the frequency of **syncd**.

   If this daemon is run more frequently, fewer number of pending I/O operations will need to be flushed to disk. Therefore, the invocation of **syncd** will cause less of a peak in I/O operations.

   To change the frequency of **syncd**, edit (as **root**) the **/sbin/rc.boot** file. Search for the following two lines:

   ```
   echo "Starting the sync daemon" | alog -t boot
   nohup /usr/sbin/syncd 60 > /dev/null 2>&1 &
   ```

   The period is set in seconds in the second line, immediately following the invocation of **/usr/sbin/syncd**. In this example, the interval is set to 60 seconds. A recommended value for the period is 10 seconds. A reboot is needed for the change to take effect.

If the problem is related to a process running with a fixed AIX priority which is higher (that is, smaller number) than that of the Topology Services daemon, the problem may be corrected by changing the daemon's priority. In PSSP, this can be done by issuing this command on the control workstation:

```
/usr/sbin/rsct/bin/hatstune -p new_value -r
```

Note: Command **hatstune** was introduced in PSSP 3.2 as a replacement for making direct changes to the **TS_Config** SDR class using command **SDRChangeAttrValues**.

Probable cause 2 on page 406 can be determined by the presence of an AIX error log entry that indicates that the daemon exited. See "AIX Error Logs and templates" on page 352 for the list of possible error templates used. Look also for an error entry with a LABEL of CORE_DUMP and PROGRAM NAME of **hatsd**. This indicates that the daemon exited abnormally, and a **core** file should exist in the daemon's **run** directory.

If the daemon produced one of the error log entries before exiting, the error log entry itself, together with the information from "AIX Error Logs and templates" on page 352, should provide enough information to diagnose the problem. If the CORE_DUMP entry was created, follow instructions in "Information to collect before contacting the IBM Support Center" on page 384 and contact the IBM Support Center.

Probable cause 3 on page 406 is the most difficult to analyze, since there may be multiple causes for packets to be lost. Some commands are useful in determining if packets are being lost or discarded at the node. Issue these commands:

1. `netstat -D`

   The `Idrops` and `Odrops` headings are the number of packets dropped in each interface or device.

2. `netstat -m`

The `failed` heading is the number of mbuf allocation failures.

3. `netstat -s`

   The `socket buffer overflows` text is the number of packets discarded due to lack of socket space.

   The `ipintrq overflows` text is the number of input packets discarded because of lack of space in the packet interrupt queue.
4. `netstat -v`

   This command shows several adapter statistics, including packets lost due to lack of space in the adapter transmit queue, and packets lost probably due to physical connectivity problems (″CRC Errors″).
5. `vmstat -i`

   This command shows the number of device interrupts for each device, and gives an idea of the incoming traffic.

There can be many causes for packets to be discarded or lost, and the problem needs to be pursued as an IP-related problem. Usually the problem is caused by one or more of the following:

1. Excessive IP traffic on the network or the node itself.
2. Inadequate IP or UDP tuning.
3. Physical problems in the adapter or network.

If causes 1 and 2 do not seem to be present, and cause 3 could not be determined, some of the commands listed previously should be issued in loop, so that enough IP-related information is kept in case the problem happens again.

The underlying problem that is causing packets to be lost must be understood and solved. The repair is considered effective if the node is no longer considered temporarily down under a similar workload.

In some environments (probable causes 1 on page 406 and 3 on page 406), the problem may be bypassed by relaxing the Topology Services tunable parameters, to allow a node not to be considered down when it cannot temporarily send network packets. Changing the tunable parameters, however, also means that it will take longer to detect a node or adapter as down.

Note: Before the tunable parameters are changed, record the current values, so that they can be restored to their original values if needed.

This solution can only be applied when:

1. There seems to be an upper bound on the amount of ″outage″ the daemon is experiencing.
2. The applications running on the system can withstand the longer adapter or node down detection time.

In PSSP, the Topology Services ″sensitivity″ factor can be changed by issuing this command on the control workstation:

```
/usr/sbin/rsct/bin/hatstune -s new_value -r
```

The adapter and node detection time is given by the formula:

```
2 * Sensitivity * Frequency
```

(two multiplied by the value of *Sensitivity* multiplied by the value of *Frequency*), where *Sensitivity* is the value returned by the command:

```
SDRGetObjects TS_Config Sensitivity
```

and *Frequency* is the value returned by the command:

```
SDRGetObjects TS_Config Frequency
```

Both values are also returned by the command:

```
/usr/sbin/rsct/bin/hatstune -v
```

In HACMP, the ″Sensitivity″ and ″Frequency″ tunable parameters are network-specific. The tunable parameters for each network may be changed with the sequence:

```
smit hacmp
                  Cluster Configuration
                    Cluster Topology
                      Configure Network Modules
                        Change/Show a Cluster Network Module
                          (select a network module)
                            Select the "Failure Detection Rate"
```

The topology must be synchronized for the tunable parameter changes to take effect. This is achieved with the sequence:

```
smit hacmp
                  Cluster Configuration
                    Cluster Topology
                      Synchronize Cluster Topology
```

To verify that the tuning changes have taken effect, issue the command:

```
lssrc -ls subsystem_name
```

approximately one minute after making the changes. The tunable parameters in use are shown in the output in a line similar to the following:

```
HB Interval = 1 secs. Sensitivity = 4 missed beats
```

For each network, `HB Interval` is the *Frequency* parameter, and `Sensitivity` is the *Sensitivity* parameter.

For examples of tuning parameters that can be used in different environments, consult the chapter ″The Topology Services Subsystem″ of *PSSP: Administration Guide*.

Good results are indicated by the tunable parameters being set to the desired values.

Error results are indicated by the parameters having their original values or incorrect values.

To verify whether the tuning changes were effective in masking the daemon outage, the system has to undergo a similar workload to that which caused the outage.

To remove the tuning changes, follow the same tuning changes outlined previously, but this time restore the previous values of the tunable parameters.

Preventing memory contention problems with the AIX Workload Manager: Memory contention has often caused the Topology Services daemon to be blocked for significant periods of time. This results in ″false node downs″, and in the triggering of the Dead Man Switch timer in HACMP/ES. An AIX error log entry with label **TS_LATEHB_PE** may appear when running RSCT 1.2 or higher. The message ″Late in sending Heartbeat by ...″ will appear in the daemon log file in any release of RSCT, indicating that the Topology Services daemon was blocked. Another error log entry that could be created is **TS_DMS_WARNING_ST**.

In many cases, such as when the system is undergoing very heavy disk I/O, it is possible for the Topology Services daemon to be blocked in paging operations, even though it looks like the system has enough memory. Two possible causes for this phenomenon are:

- In steady state, when there are no node and adapter events on the system, the Topology Services daemon uses a ″working set″ of pages that is substantially smaller than its entire addressing space. When node or adapter events happen, the daemon faces the situation where additional pages it needs to process the events are not present in memory.
- When heavy file I/O is taking place, the operating system may reserve a larger percentage of memory pages to files, making fewer pages available to processes.
- When heavy file I/O is taking place, paging I/O operations may be slowed down by contention for the disk.

The probability that the Topology Services daemon gets blocked for paging I/O may be reduced by making use of the AIX Workload Manager (WLM). WLM is an operating system feature introduced in AIX Version 4.3.3. It is designed to give the system administrator greater control over how the scheduler and Virtual Memory Manager (VMM) allocate CPU and physical memory resources to processes. WLM gives the system administrator the ability to create different classes of service, and specify attributes for those classes.

The following explains how WLM can be used to allow the Topology Services daemon to obtain favorable treatment from the VMM. There is no need to involve WLM in controlling the daemon’s CPU use, because the daemon is already configured to run at a real time fixed scheduling priority. WLM will not assign priority values smaller than 40 to any thread.

These instructions are given using **SMIT**, but it is also possible to use WLM or AIX commands to achieve the same goals. For versions of AIX before the 4330-02 Recommended Maintenance Level (which can be ordered using APAR IY06844), HACMP/ES should not be active on the machine when WLM is started. If HACMP/ES is active on the machine when WLM is started, it will not recognize the Topology Services daemon as being in the newly created class. For the same reason, in PSSP the Topology Services subsystem must be restarted after WLM is started. Starting with 4330-02, WLM is able to classify processes that started before WLM itself is started, so restarting Topology Services is not needed.

Initially, use the sequence:

```
smit wlm
   Add a Class
```

to add a TopologyServices class to WLM. Ensure that the class is at Tier 0 and has Minimum Memory of 20%. These values will cause processes in this class to receive favorable treatment from the VMM. Tier 0 means that the requirement from this class will be satisfied before the requirements from other classes with higher tiers. Minimum Memory should prevent the process's pages from being taken by other processes, while the process in this class is using less than 20% of the machine's memory.

Use the sequence:

```
   smit wlm
      Class Assignment Rules
         Create a new Rule
```

to create a rule for classifying the Topology Services daemon into the new class. In this screen, specify **1** as Order of the Rule, TopologyServices as Class, and **/usr/sbin/rsct/bin/hatsd** as Application.

To verify the rules that are defined, use the sequence:

```
    smit wlm
      Class Assignment Rules
          List all Rules
```

To start WLM, after the new class and rule are already in place, use the sequence:

```
    smit wlm
       Start/Stop/Update WLM
          Start Workload Management
```

To verify that the Topology Services daemon is indeed classified in the new class, use command:

```
ps -ef -o pid,class,args | grep hatsd | grep -v grep
```

One sample output of this command is:

```
15200     TopologyServices /usr/sbin/rsct/bin/hatsd -n 5
```

The TopologyServices text in this output indicates that the Topology Services daemon is a member of the TopologyServices class.

If WLM is already being used, the system administrator must ensure that the new class created for the Topology Services daemon does not conflict with other already defined classes. For example, the sum of all ″minimum values″ in a tier must be less than 100%. On the other hand, if WLM is already in use, the administrator must ensure that other applications in the system do not cause the Topology Services daemon to be deprived of memory. One way to prevent other applications from being more privileged than the Topology Services daemon in regard to memory allocation is to place other applications in tiers other than tier 0.

If WLM is already active on the system when the new classes and rules are added, WLM needs to be restarted in order to recognize the new classes and rules.

For more information on WLM, see the chapter ″AIX Workload Manager″ in *AIX 5L Version 5.1 Differences Guide*.

Action 6 - Investigate IP communication problem

Probable causes of this problem are:

1. The Topology Services daemon was temporarily blocked.
2. The Topology Services daemon exited on the node.
3. IP communication problem, such as mbuf shortage or excessive adapter traffic.

Probable cause 1 and probable cause 2 are usually only possible when all the monitored adapters in the node are affected. This is because these are conditions that affect the daemon as a whole, and not just one of the adapters in a node.

Probable cause 3, on the other hand, may result in a single adapter in a node being considered as down. Follow the procedures described to diagnose symptom ″Node appears to go down and then up″, “Action 5 - Investigate hatsd problem” on page 406. If probable cause 1 on page 406 or probable cause 2 on page 406 is identified as the source of the problem, follow the repair procedures described under the same symptom.

If these causes are ruled out, the problem is likely related to IP communication. The instructions in ″Node appears to go down and then up″, “Action 5 - Investigate hatsd problem” on page 406 describe what communication parameters to monitor in order to pinpoint the problem.

To identify the network that is affected by the problem, issue command **errpt -J TS_DEATH_TR | more**. This is the AIX error log entry created when the local adapter stopped receiving heartbeat messages from its neighbor adapter. The neighbor’s address, which is listed in the error log entry, indicates which network is affected.

A note on ***″local″ and ″remote″ adapter notifications:*** In PSSP, when there is a communication flicker, a node may temporarily appear down on **host_responds** without the same node ever failing or even considering its local adapter to be down. This happens because **host_responds** is actually the Ethernet adapter membership as seen from the control workstation. If the Ethernet adapter membership partitions because of a temporary communication problem, only the adapters on the same group as the control workstation will be considered up by **host_responds**. Consequently, if the control workstation is temporarily isolated, all the nodes will be considered down.

Action 7 - Investigate Group Services failure

This is most likely a problem in the Topology Services daemon, or a problem related to the communication between the daemon and the Topology Services library, which is used by the Group Services daemon. Usually the problem occurs in HACMP during IP address takeover, when multiple adapters in the network temporarily have the same address. The problem may also happen during a Topology DARE operation in HACMP or a Topology Services refresh in PSSP. See **TS_REFRESH_ER** on page 369.

When this problem occurs, the Group Services daemon exits and produces an error log entry with a LABEL of **GS_TS_RETCODE_ER**. This entry will have the Topology Services return code in the Detail Data field. Topology Services will produce an error log entry with a LABEL of **TS_LIBERR_EM**. Follow the instructions in “Information to collect before contacting the IBM Support Center” on page 384 and contact the IBM Support Center.

Action 8 - Investigate node crash

If a node crashes, perform AIX system dump analysis. Probable causes of this problem are:

1. The Dead Man Switch timer was triggered, probably because the Topology Services daemon was blocked.
2. An AIX-related problem.

When the node restarts, perform AIX system dump analysis. Initially, issue the following sequence in SMIT to obtain information about the dump:

```
smit
                  Problem Determination
                    System Dump
                      Show Information About the Previous System Dump
```

Then, issue this SMIT sequence to obtain information about the dump device. It is the device listed after the `primary` text in this sequence.

```
smit
                  Problem Determination
                    System Dump
                      Show Current Dump Devices
```

Issue the command:

```
/usr/sbin/crash dump_device/unix
```

The command will present a **>** prompt. Issue the **stat** subcommand. The output is similar to the following:

```
sysname: AIX
        nodename: bacchus
        release: 3
        version: 4
        machine: 000022241C00
        time of crash: Thu Mar 11 10:23:34 EST 1999
        age of system: 2 hr., 2 min.
        xmalloc debug: disabled
        abend code: 700
        csa: 0x322eb0
        exception struct:
                0x00000000 0x00000000 0x00000000 0x00000000 0x00000000
        panic:
```

Note the `panic` text. If this is absent, this dump was not caused by the ″Dead Man Switch″ timer trigger. If `panic` is present in the output, issue subcommand:

```
t -m
```

The output should be similar to the following:

```
Skipping first MST

MST STACK TRACE:
0x00322eb0 (excpt=00000000:00000000:00000000:00000000:00000000) (intpri=5)
        IAR:      .panic_trap+0 (000213a4):     teq   r1,r1
        LR:       .[haDMS_kex:dead_man_sw_handler]+54 (011572bc)
        00322dc8: .clock+f0 (0001307c)
        00322e28: .i_poll+6c (0006e58c)
        00322e78: ex_flih_rs1+e8 (000ce28c)
```

If the lines `panic_trap` and `haDMS_kex:dead_man_sw_handler` are present, this is a dump produced because of the Dead Man Switch timer trigger. Otherwise, there is another source for the problem. For problems unrelated to the ″Dead Man Switch″ timer, contact the IBM Support Center. Use the **quit** subcommand to exit the program. For more information about producing or saving System Dumps, see “Chapter 5. Producing a system dump” on page 81.

If the dump was produced by the Dead Man Switch timer, it is likely that the problem was caused by the Topology Services daemon being blocked. HACMP/ES uses this mechanism to protect data in multi-tailed disks. When the timer is triggered, other nodes are already in the process of taking over this node’s resources, since Topology Services is blocked in the node. If the node was allowed to continue functioning, both this node and the node taking over this node’s disk would be concurrently accessing the disk, possibly causing data corruption.

The Dead Man Switch (DMS) timer is periodically stopped and reset by the Topology Services daemon. If the daemon gets blocked and does not have a chance to reset the timer, the timer-handling function runs, causing the node to crash. Each time the daemon resets the timer, the remaining amount left in the previous timer is stored. The smaller the remaining time, the closer the system is to triggering the timer. These ″time-to-trigger″ values can be retrieved with command:

```
/usr/sbin/rsct/bin/hatsdmsinfo
```

The output of this command is similar to:

```
Information for Topology Services -- HACMP/ES
DMS Trigger time: 8.000 seconds.
Last DMS Resets                         Time to Trigger (seconds)
11/11/99 09:21:28.272                    7.500
11/11/99 09:21:28.772                    7.500
11/11/99 09:21:29.272                    7.500
11/11/99 09:21:29.772                    7.500
11/11/99 09:21:30.272                    7.500
11/11/99 09:21:30.782                    7.490

DMS Resets with small time-to-trigger    Time to Trigger (seconds)
Threshold value: 6.000 seconds.
11/11/99 09:18:44.316                    5.540
```

If small ″time-to-trigger″ values are seen, the HACMP tunables described in “Action 5 - Investigate hatsd problem” on page 406 need to be changed, and the root cause for the daemon being blocked needs to be investigated. Small time-to-trigger″ values also result in an AIX error log entry with template **TS_DMS_WARNING_ST**. Therefore, when this error log entry appears, it indicates that the system is getting close to triggering the Dead Man Switch timer. Actions should be taken to correct the system condition that leads to the timer trigger.

For additional diagnosis and repair procedures, follow the instructions for the symptom ″Node appears to go down and then up a few seconds later″ “Action 5 - Investigate hatsd problem” on page 406, which is also related to the Topology Services daemon being blocked.

Action 9 - Investigate problems after a refresh

Probable causes of this problem are:

1. A refresh operation fails on the node.
2. Adapters are configured with an incorrect address in the registry (SDR on PSSP nodes, Global ODM on HACMP nodes).

3. Topology Services startup script on the control workstation fails to determine the active authentication methods or fails to determine the authentication keys.
4. Topology Services daemon on the nodes fails to confirm the local active authentication methods or fails to obtain the authentication keys.

Verify whether all nodes were able to complete the refresh operation, by running “Operational test 8 - Check if configuration instance and security status are the same across all nodes” on page 398. If this test reveals that nodes are running with different Configuration Instances (from the **lssrc** output), at least one node was unable to complete the refresh operation successfully.

Issue the command **errpt -J ″TS_*″ | more** on all nodes. Entry **TS_SDR_ER** is one of the more likely candidates. It indicates a problem while trying to obtain a copy of the **machines.lst** file from the SDR. The startup script log provides more details about this problem.

Other error log entries that may be present are:

1. TS_REFRESH_ER
2. TS_MACHLIST_ER
3. TS_LONGLINE_ER
4. TS_SPNODEDUP_ER or TS_HANODEDUP_ER
5. TS_SPIPDUP_ER or TS_HAIPDUP_ER
6. TS_IPADDR_ER
7. TS_MIGRATE_ER
8. TS_AUTHMETH_ER
9. TS_KEY_ER
10. TS_SECMODE_ER

For information about each error log entry and how to correct the problem, see “Error information” on page 352.

If a node does not respond to the command: **lssrc -ls** *subsystem*, (the command hangs), this indicates a problem in the connection between Topology Services and the AIX SRC subsystem. Such problems will also cause in the Topology Services daemon to be unable to receive the refresh request.

If no **TS_** error log entry is present, and all nodes are responding to the **lssrc** command, and **lssrc** is returning different Configuration Instances for different nodes, contact the IBM Support Center.

If all nodes respond to the **lssrc** command, and the Configuration Instances are the same across all nodes, follow “Configuration verification tests” on page 387 to find a possible configuration problem. Error log entry **TS_MISCFG_EM** is present if the adapter configuration on the SDR (for PSSP) or ODM (for HACMP) does not match the actual address configured in the adapter.

For problems caused by loss of connection with the AIX SRC, the Topology Services subsystem may be restarted. For PSSP systems, issuing the command: **/usr/sbin/rsct/bin/hatsctrl -k WILL NOT WORK** because the connection with the AIX SRC subsystem was lost. To recover, perform these steps:

1. Issue the command:

   ```
   ps -ef | grep hats | grep -v grep
   ```

to find the daemon's *process_ID*:

The output of the command is similar to the following:

```
root 13446  8006  0  May 27   - 26:47 /usr/sbin/rsct/bin/hatsd -n 3
```

In this example, the *process_ID* is 13446.

Note: If HACMP is also running on the node, there will be two lines of output in the **ps** command. To find out which is the PSSP version of the daemon, the following methods can be used:

a. If the number displayed after the `-n` text in the output (the node number) is different on the two lines, the PSSP version of the daemon is the one where the number is the SP node number. The node's number is obtained by issuing the command:

   ```
   /usr/lpp/ssp/install/bin/node_number
   ```

b. If the node numbers are the same, issue the command:

   ```
   lssrc -s topsvcs
   ```

   to obtain the process id of the HACMP/ES version of the daemon. This process id should match that in one of the lines in the **ps** output. The PSSP version of the daemon is the other one.

2. Issue the command:

   ```
   kill process_ID
   ```

   This stops the Topology Services daemon.

3. If the AIX SRC subsystem does not restart the Topology Services subsystem automatically, issue this command:

   ```
   /usr/sbin/rsct/bin/hatsctrl -s
   ```

For HACMP, restarting the Topology Services daemon requires shutting down the HACMP cluster on the node, which can be done with the sequence:

```
smit hacmp
                       Cluster Services
                           Stop Cluster Services
```

After HACMP is stopped, follow the instructions for PSSP on page 416, to find the process id of the Topology Services daemon and stop it, using the command:

```
/usr/sbin/rsct/bin/topsvcsctrl
```

instead of the command:

```
/usr/sbin/rsct/bin/hatsctrl
```

Now restart HACMP on the node using this sequence:

```
smit hacmp
                       Cluster Services
                           Start Cluster Services
```

Follow the procedures in “Operational verification tests” on page 390 to ensure that the subsystem is behaving as expected across all nodes.

Note: In the HACMP/ES environment, **DO NOT STOP** the Topology Services daemon by issuing any of these commands.

- kill
- stopsrc
- topsvcsctrl -k

This is because stopping the Topology Services daemon while the cluster is up on the node results in the node being stopped by the HACMP cluster manager.

Action 10 - Correct authentication keys

The cause of this problem should be investigated thoroughly. The Topology Services key file can be re-created using the procedure described in “Action 5 - Correct key files” on page 276.

Alternatively, a new key can be added to the Topology Services key file, and Topology Services may be notified by a refresh command to use this new key. The procedure to add a new key is described in ″The Topology Services Subsystem,″ section ″Changing the Authentication Method and Key″ of *PSSP: Administration Guide*.

Chapter 24. Diagnosing Group Services problems

This chapter discusses diagnostic procedures and failure responses for the Group Services (GS) component of RSCT. The list of known error symptoms and the associated responses are in the section "Error symptoms, responses, and recoveries" on page 446. A list of the information to collect before contacting the IBM Support Center is in the section "Information to collect before contacting the IBM Support Center" on page 430.

Note that Group Services is a subsystem of RS/6000 Cluster Technology (RSCT).

Related documentation

The following publications provide information about the Group Services (GS) component of RSCT:

1. *RSCT: Group Services Programming Guide and Reference*

   This book provides detailed information for programmers, including the Group Services API (Application Programming Interfaces). It also provides sample programs.
2. *PSSP: Administration Guide*

   The chapter ″The Group Services Subsystem″ provides information about the GS subsystem and its interaction with PSSP subsystems.
3. *PSSP: Command and Technical Reference*

   This book provides detailed syntax and parameter information for the commands used to control the Group Services subsystem:
   - lssrc
   - nlssrc
   - hagsns
   - hagsvote
   - hagsctrl
4. *PSSP: Messages Reference*

   These chapters contain messages related to Group Services:
   - 2520 - Group Services Messages
   - 2525 - RS/6000 Cluster Technology Common Messages
5. *HACMP Enhanced Scalability Handbook*, SG24-5328-00

   This book provides information on HACMP/ES, including problem determination procedures. This book is useful in isolating problems that affect Group Services when run in the HACMP/ES environment.
6. *HACMP Enhanced Scalability Installation and Administration Guide* SC23-4306-01.

   See the chapter ″Troubleshooting HACMP/ES Clusters″. This book provides information on HACMP/ES, including problem determination procedures. This book is useful in isolating problems that affect Group Services when run in the HACMP/ES environment.
7. *RS/6000 High Availability Infrastructure*, SG24-4838-00

   This book provides information about the RSCT infrastructure, including a presentation of how Group Services works.

Requisite function

This is a list of the software directly used by the GS component of RSCT. Problems within the requisite software may manifest themselves as error symptoms in Group Services. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the GS component of RSCT, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- Topology Services subsystem of RSCT
- AIX System Resource Controller (SRC)
- **/var/ha** directory
- System Data Repository (SDR)
- SP Switch (CSS subsystem)
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
- DCE security
- FFDC library
- AIX: UDP communication
- AIX: Unix-Domain sockets

Error information

AIX Error Logs and templates

Table 60 on page 421 shows the AIX error log templates used by GS.

- GS_ASSERT_ER
- GS_AUTH_DENIED_ST
- GS_CLNT_SOCK_ER
- GS_DEACT_FAIL_ST
- GS_DOM_MERGE_ER
- GS_DOM_NOT_FORM_WA
- GS_ERROR_ER
- GS_GLSM_ERROR_ER
- GS_GLSM_START_ST
- GS_GLSM_STARTERR_ER
- GS_GLSM_STOP_ST
- GS_INVALID_MSG_ER
- GS_MESSAGE_ST
- GS_START_ST
- GS_STARTERR_ER
- GS_STOP_ST
- GS_TS_RETCODE_ER
- GS_XSTALE_PRCLM_ER

When you retrieve an error log entry, look for the Detail Data section near the bottom of the entry.

Each entry refers to a particular instance of the GS daemon on the local node. One entry is logged for each occurrence of the condition, unless otherwise noted in the Detail Data section. The condition is logged on every node where the event occurred.

The Detail Data section of these entries is not translated to other languages. This section is in English.

The error type is:

- A - Alert (failure in a GS client)
- E - Error (failure in GS)
- I - Informational (status information)

Table 60. AIX Error Log templates for Group Services

Label and Error ID	Type	Diagnostic explanation and details
GS_ASSERT_ER 5608839C	E	**Explanation:** The GS daemon produced a core dump. **Details:** The GS daemon encountered an irrecoverable assertion failure. This occurs only if the daemon core dumps due to a specific GS assertion failure. GS will be restarted automatically and the situation will be cleared. However, its state is not cleared and the system administrator must determine the cause of the failure. The REFERENCE CODE field in the Detail Data section may refer to the error log entry which caused this event. See "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.
GS_AUTH_DENIED_ST 23628CC2	A	**Explanation:** An unauthorized user tried to access GS. **Details:** An unauthorized user tried to connect to the GS daemon. Standard fields indicate that GS daemon detected an attempt to connect from an unauthorized user. Detailed fields explain the detail information. Possibilities are: the user is not a **root** user, or the user is not a member of the **hagsuser** group.
GS_CLNT_SOCK_ER E31202B3	E	**Explanation:** Warning or error on the Group Services client socket. **Details:** Group Services has an error on the client socket, or the AIX **hagsuser** group is not defined. Standard fields indicate that Group Services received an error or warning condition on the client socket. Detailed fields explain what error or warning caused this problem.
GS_DEACT_FAIL_ST 972737A5	I	**Explanation:** Failure of the deactivate script. **Details:** The GS daemon is unable to run the deactivate script. Standard fields indicate that the GS daemon is unable to run the script. Detailed fields give more information. The deactivate script may not exist, or system resources are not sufficient to run the deactivate script.

Table 60. AIX Error Log templates for Group Services (continued)

Label and Error ID	**Type**	**Diagnostic explanation and details**
GS_DOM_MERGE_ER 267369E6	A, E	**Explanation:** Two Group Services domains were merged. **Details:** Two disjoint Group Services domains are merged because Topology Services has merged two disjoint node groups into a single node group. There may be several nodes with the same entries. Detailed fields contains the merging node numbers. At the time of domain merge, GS daemons on the nodes that generate **GS_DOM_MERGE_ER** entries will exit and be restarted. After the restart, (by **GS_START_ST**) Group Services will clear this situation. The REFERENCE CODE field in the Detail Data section may refer to the error log entry that caused this event. See "Action 2 - Verify Status of Group Services Subsystem" on page 448. See "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.
GS_DOM_NOT_FORM_WA 83B3361F	I	**Explanation:** A Group Services domain was not formed. **Details:** The GS daemon writes this entry periodically until the GS domain is formed. There may be several nodes in the same situation at the same time. The GS domain cannot be formed because: • On some nodes, Topology Services may be running but GS is not. • Nameserver recovery protocol is not complete. This entry is written periodically until the domain is established. The entry is written as follows: every 5, 30, 60, 90 minutes, and then once every two hours as long as the domain is not established. The domain establishment is recorded by a **GS_MESSAGE_ST** template label. The REFERENCE CODE field in the Detail Data section may refer to the error log entry that caused this event.
GS_ERROR_ER 2C582BE2	A, E	**Explanation:** Group Services logic failure. **Details:** The GS daemon encountered an irrecoverable logic failure. Detailed fields describes what kind of error is encountered. The GS daemon exits due to the GS logic failure. Group Services will be restarted automatically and the situation will be cleared. However, if the state is not cleared, the administrator must determine what caused the GS daemon to terminate. The REFERENCE CODE field in the Detail Data section may refer to the error log entry that caused this event. See "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.

Table 60. AIX Error Log templates for Group Services *(continued)*

Label and Error ID	Type	Diagnostic explanation and details
GS_GLSM_ERROR_ER E0A52CFA	A, E	**Explanation:** Group Services GLSM daemon logic failure. **Details:** The Group Services GLSM daemon encountered an irrecoverable logic failure. Standard fields indicate that the daemon stopped. Detailed fields point to the error log entry created when the daemon started. The Group Services GLSM daemon exited due to the logic failure. The Group Services GLSM daemon will be restarted automatically and the situation will be cleared. However, if the state is not cleared, the administrator must determine what caused the problem. The standard fields are self-explanatory. The REFERENCE CODE field in the Detail Data section may refer to the error log entry that caused this event. See "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.
GS_GLSM_START_ST 3C447B26	I	**Explanation:** Group Services GLSM Daemon started (AIX error log entry). **Details:** The Group Services GLSM daemon has started. Standard fields indicate that the daemon started. Detailed fields contain the path name of the log file. The Group Services GLSM subsystem was started by a user or by a process. Issue this command: `lssrc -l -s` *`glsm_subsystem`* where *subsystem* is: • **hagsglsm.***partition* on the PSSP control workstation • **hagsglsm** on a PSSP node • **grpglsm** on a HACMP node If the daemon is started, the output will contain a status of ″operative″ for **hagsglsm**. Otherwise, the output will contain a status of ″inoperative″ for **hagsglsm**.
GS_GLSM_STARTERR_ER 2BF4FAC3	A, E	**Explanation:** Group Services GLSM daemon cannot be started. **Details:** The Group Services GLSM daemon encountered a problem during startup. Standard fields indicate that the daemon is stopped. Detailed fields point to the error log entry created when the daemon started. The GS daemon cannot be started because: 1. **Syspar** class in the SDR cannot be obtained. 2. **exec** to **hagsglsmd** has failed. The AIX log entry may be the only remaining information about the cause of the problem after it is cleared.

Table 60. AIX Error Log templates for Group Services (continued)

Label and Error ID	Type	Diagnostic explanation and details
GS_GLSM_STOP_ST 177CF680	I	**Explanation:** HAGSGLSM (HA Group Services GLobalized Switch Membership) daemon stopped. **Details:** The Group Services GLSM daemon was stopped by a user or by a process. Standard fields indicate that the daemon stopped. Detailed fields point to the error log entry created when the daemon started. If the daemon was stopped by the SRC, the word ″SRC″ will be present in the Detail Data. The REFERENCE CODE field in the Detail Data section may reference the error log entry that caused this event. Issue this command: `lssrc -l -s` *`glsm_subsystem`* where *subsystem* is: • **hagsglsm.***partition* on the PSSP control workstation • **hagsglsm** on a PSSP node • **grpglsm** on a HACMP node If the daemon is stopped, the output will contain a status of ″inoperative″ for **hagsglsm**. Otherwise, the output will contain a status of ″operative″ for **hagsglsm**.
GS_INVALID_MSG_ER 4242B470	A, E	**Explanation:** The GS daemon received an unknown message. **Details:** The GS daemon received an incorrect or unknown message from another daemon. The transmitted messages may be corrupted on the wire, or a daemon sent a corrupted message. The GS daemon will restart and clear the problem. See “Information to collect before contacting the IBM Support Center” on page 430 and contact the IBM Support Center.
GS_MESSAGE_ST F3AB2E54	I	**Explanation:** Group Services informational message **Details:** The GS daemon has an informational message about the Group Services activity, or condition. Detailed fields describes the information. It is one of the following: 1. The GS daemon is not connected to Topology Services. 2. The GS domain has not recovered or been established after a long time. 3. Any other message, which will be in the detailed field. The REFERENCE CODE field in the Detail Data section may refer to the error log entry that caused this event.
GS_START_ST B4D19120	I	**Explanation:** Group Services daemon started. **Details:** The GS subsystem is started by a user or by a process. Detailed fields contain the log file name.

Table 60. AIX Error Log templates for Group Services (continued)

Label and Error ID	Type	Diagnostic explanation and details
GS_STARTERR_ER 66E04502	A, E	**Explanation:** Group Services cannot be started. **Details:** The GS daemon encountered a problem during startup. **Information about the cause of this problem may not be available once the problem is cleared.** The GS daemon cannot start because one of the following conditions occurred: 1. **Syspar** class in the SDR cannot be obtained. 2. **exec** to **hagsd** failed. 3. The environment variables used by the startup scripts are not set properly. 4. Daemon initialization failed.
GS_STOP_ST 28EDAA82	I	**Explanation:** Group Services daemon stopped. **Details:** The GS daemon was stopped by a user or by a process. Detailed fields indicate how the daemon stops. If this was not intended, the system administrator must determine what caused the GS daemon to terminate. If the daemon was stopped by the AIX SRC, ″SRC″ will be present in the Detail Data.
GS_TS_RETCODE_ER EB38514E	A, E	**Explanation:** The Topology Services library detected an error condition. **Details:** The GS daemon received an incorrect or unknown message from another daemon. This entry refers to a particular instance of the Topology Services library on the local node. Standard fields indicate that Group Services received an error condition from Topology Services. Detailed fields contain the explanation and Topology Services library error number. The GS daemon will restart and clear the problem. The standard fields are self-explanatory. The REFERENCE CODE field in the Detail Data section may contain the Topology Services log entry that causes this event. See “Action 7 - Investigate Group Services failure” on page 413.
GS_XSTALE_PRCLM_ER 6D790718	A, E	**Explanation:** Non-stale proclaim message was received. This means that inconsistent domain join request messages were received. **Details:** The local node received a valid domain join request (proclaim) message from his Nameserver twice. This should not happen in a normal situation. Detailed fields point to the error log entry of a NodeUp event. Topology Services reports inconsistent node down and up events between nodes. The GS daemon will restart and clear the problem. The REFERENCE CODE field in the Detail Data may reference the error log entry that caused this event. For more information, see the symptom ″Non-stale proclaim message received″ in “Error symptoms, responses, and recoveries” on page 446. See “Information to collect before contacting the IBM Support Center” on page 430 and contact the IBM Support Center.

Dump information

Group Services creates a core dump automatically when certain errors occur, and also provides service information that can be obtained automatically by the **phoenix.snap** script.

Core dump

A core dump is generated by the Group Services daemon if it encounters an undefined condition. It contains normal information saved by AIX in a core dump: Group Services daemon's process user data and call stack information. The dump is specific to a particular instance of the GS daemon on the local node. Other nodes may have a similar core dump. Each core dump file is approximately 10MB in size.

For a PSSP node, the core dumps are located in: **/var/ha/run/hags.**_partition_**/core*** and **/var/ha/run/hagsglsm.**_partition_**/core***.

For a HACMP node, the core dumps are located in:
/var/ha/run/grpsvcs._cluster_**/core*** and **/var/ha/run/grpglsm.**_cluster_**/core***.

Core dumps are created automatically when:

- One of the GS daemons invokes an **assert()** statement if the daemon state is undefined or encounters an undefined condition by design.
- The daemon attempts an incorrect operation, such as division by zero.
- The daemon receives a segmentation violation signal for accessing its data incorrectly.

A core dump is created manually by issuing the command:

```
kill -6 pid_of_daemon
```

where _pid_of_daemon_ is obtained by issuing the command:

```
lssrc -s gs_subsystem
```

where _gs_subsystem_ is:

- **hags** on PSSP nodes
- **hags.**_partition_ on the PSSP control workstation
- **grpsvcs** on HACMP nodes

The core dump is valid as long as the executable file **/usr/sbin/rsct/bin/hagsd** is not replaced. Copy the core dumps and the executable file to a safe place.

To verify the core dump, issue this command:

```
dbx /usr/sbin/rsct/bin/hagsd core_file
```

where _core_file_ is one of the **core*** files described previously.

Good results are indicated by output similar to:

```
Type 'help' for help.
 reading symbolic information ...
```

```
[using memory image in core]
IOT/Abort trap in evt._pthread_ksleep [/usr/lib/libpthreads.a]
at 0xd02323e0 ($t6) 0xd02323e0 (_pthread_ksleep+0x9c) 80410014
 lwz    r2,0x14(r1)
```

Error results may look like one of the following:

1. This means that the current executable file was not the one that created the core dump.

   ```
   Type 'help' for help.
   Core file program (hagsd) does not match current program (core ignored)
   reading symbolic information ...
   (dbx)
   ```

2. This means that the dump is incomplete due to lack of disk space.

   ```
   Type 'help' for help.
   warning: The core file is truncated.  You may need to increase the ulimit
   for file and coredump, or free some space on the filesystem.
   reading symbolic information ...
   [using memory image in core]

   IOT/Abort trap in evt._pthread_ksleep [/usr/lib/libpthreads.a]
   at 0xd02323e0
   0xd02323e0 (_pthread_ksleep+0x9c) 80410014
          lwz   r2,0x14(r1)
   (dbx)
   ```

phoenix.snap dump

This dump contains diagnostic data used for RSCT problem determination. It is a collection of log files and the other trace information for the RSCT components. For more detailed information about **phoenix.snap**, see the Topology Services chapter, heading “phoenix.snap dump” on page 376.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

The log files, including the Group Services Trace logs and startup logs, are preserved as long as their total size does not exceed 5MB. If the total size is greater than 5MB, the oldest log file is removed at Group Services startup time.

GS service log trace

The GS service log contains a trace of the GS daemon. It is intended for IBM Support Center use only, and written in English. It refers to a particular instance of the GS daemon running on the local node. When a problem occurs, logs from multiple nodes are often needed.

If obtaining logs from all nodes is not feasible, collect logs from these nodes:

- The node where the problem was detected
- The Group Services Nameserver (NS) node. To find the NS node, see "How to find the GS nameserver (NS) node" on page 432.
- The PSSP control workstation
- If the problem is related to a particular GS group, the Group Leader node of the group that is experiencing the problem. To find a Group Leader node for a specific group, see "How to find the Group Leader (GL) node for a specific group" on page 433.

Service log short tracing is always in effect. Service log long tracing is activated by this command:

```
traceson -l -s subsystem_name
```

where *subsystem_name* is:

- **hags** for PSSP nodes
- **hags.***partition_name* for the PSSP control workstation
- **grpsvcs** for HACMP nodes

The trace is deactivated, (reverts to short tracing) by issuing this command:

```
tracesoff -s subsystem_name
```

where *subsystem_name* is the same as for trace activation.

The trace may produce 20MB or more of data, depending on GS activity level and length of time that the trace is running. Ensure that there is adequate space in the directory **/var/ha**.

The trace is located in:

- **/var/ha/log/hags_***node_incarnation***.***partition* on PSSP nodes
- **/var/ha/log/hags.***partition_node_incarnation***.***partition* on the PSSP control workstation
- **/var/ha/log/grpsvcs_***node_incarnation***.***domain* on HACMP nodes

where *incarnation* is an increasing integer set by the GS daemon. This value can be obtained from the **NodeId** field of the command:

```
hagsns -l -s gssubsys
```

The long trace contains this information:

1. Each Group Services protocol message sent or received
2. Each significant processing action as it is started or finished
3. Details of protocols being run

For many of the cases, log files from multiple nodes must be collected. The other nodes' log files must be collected before they wrap or are removed. By default, during the long tracing, log files will expand to a maximum of 5 times the configured log size value.

To change the configured value of the log size on a PSSP node, issue this command:

```
SDRChangeAttrValues TS_Config Log_Length=new_length
```

where *new_length* is the number of lines in the trace log file. Then, restart the GS daemon.

To change the configured value on a HACMP node, perform these steps:

1. Issue this command: **smit hacmp**.
2. Select **Cluster Configuration.**
3. Select **Cluster Topology**.
4. Select **Configure Topology Services and Group Services**.
5. Select **Change/Show Topology and Group Services Configuration**.
6. Select **Group Services log length** (number of lines).
7. Enter the number of lines for each Group Services log file.

When the log file reaches the line number limit, the current log is saved into a file with a suffix of **.bak**. The original file is then truncated. With the ″long″ trace option, the default of 5000 lines should be enough for only 30 minutes or less of tracing.

Each time the daemon is restarted, a new log file is created. Only the last 5 log files are kept.

Long tracing should be activated on request from IBM Service. It can be activated (for about one minute, to avoid overwriting other data in the log file) when the error condition is still present.

Each entry is in the format: *date message*.

The ″short″ form of the service log trace is always running. It contains this information:

1. Each Group Services protocol message sent or received.
2. Brief information for significant protocols being run.
3. Significant information for possible debugging.

GS service log trace - summary log

The GS service log - summary log contains a trace of the GS daemon, but records only important highlights of daemon activity. This log does not record as much information as the GS service log, and therefore it will not wrap as quickly as the GS service log. This log is more useful in diagnosing problems whose origin occurred a while ago. All information in this log is also recorded in the GS service log, provided that the log has not yet wrapped. The GS service log - summary log is intended for IBM Support Center use only, and written in English. It refers to a particular instance of the GS daemon running on the local node. When a problem occurs, both logs from multiple nodes are often needed.

The trace is located in:

- **/var/ha/log/hags_***node_incarnation***.***partition***.long** on PSSP nodes
- **/var/ha/log/hags.***partition_node_incarnation***.***partition***.long** on the PSSP control workstation
- **/var/ha/log/grpsvcs_***node_incarnation***.***domain***.long** on HACMP nodes

where *incarnation* is an increasing integer set by the GS daemon. This value can be obtained from the **NodeId** field of the command:

```
hagsns -l -s gssubsys
```

All other information about this log is identical to the GS service log. See “GS service log trace” on page 427.

Group Services startup script log

This log contains the GS daemon’s environment variables and error messages where the startup script cannot start the daemon. The trace refers to a particular instance of the GS startup script running on the local node. This trace is always running. One file is created each time the startup script runs. The size of the file varies from 5KB to 10KB.

On PSSP nodes, it is located in: **/var/ha/log/hags.default.***partition_name.node_incarnation*.

On HACMP nodes, it is located in: **/var/ha/log/grpsvcs.default.***domain.node_incarnation*.

The main source for diagnostic information is the AIX error log. The GS startup script log should be used when the error log shows that the startup script was unable to complete its tasks and could not start the daemon.

The data in this file is in English. This information is for use by the IBM Support Center. The format of the data is the same as that of the GS Service Log Trace, ″long″ option.

Information to collect before contacting the IBM Support Center

Collect information from these nodes:

1. Nodes that exhibit the problem
2. GS nameserver (NS) node. See “How to find the GS nameserver (NS) node” on page 432.
3. Group Leader (GL) node, if the problem is related to a particular group. See “How to find the Group Leader (GL) node for a specific group” on page 433.
4. The PSSP control workstation

Collect the files listed in number 1 here, and then issue **phoenix.snap** to collect the remaining information. See “phoenix.snap dump” on page 376.

1. FFDC (AIX) Error Logs, which consist of the following:
   a. AIX error log: **/var/adm/ras/errlog**
   b. Error log templates: **/var/adm/ras/errtmplt**
   c. FFDC dump files: **/var/adm/ffdc/dumps/***
   d. English message catalogs for Group Services and Topology Services:
      - **/usr/lib/nls/msg/en_US/ha_gs.cat**
      - **/usr/lib/nls/msg/en_US/hats.cat**

   This information can be obtained from a remote machine as follows:
   a. Copy the remote error log to the local node, giving it a name such as *errlogname*.
   b. Issue the command:

```
errpt -a errlogname -y errtmplt -z ha_gs.cat
```

c. Obtain the authentication method in use. Issue this command on the control workstation:

```
splstdata -p
```

The entry ″ts_auth_methods″ lists the authentication methods in use.

2. Group Services Files
   a. Service and User log files for the GS daemon.

      On a PSSP node, the directory is: **/var/ha/log/hags_***.

      On a HACMP node, the directory is: **/var/ha/log/grpsvcs_***.
   b. Startup Script log.

      On a PSSP node it is: **/var/ha/log/hags.default***.

      On a HACMP node it is: **/var/ha/log/grpsvcs.default***.
   c. Entire contents of the daemon ″run″ directory.

      On a PSSP node the directory is: **/var/ha/run/hags.***partition***/***.

      On a HACMP node, the directory is: **/var/ha/run/grpsvcs.***cluster name*.
   d. Output of the canonical form of the **lssrc** command.

      On a PSSP node, issue this command:

      ```
      /usr/sbin/rsct/bin/nlssrc -c -ls hags
      ```

      On the PSSP control workstation, issue this command:

      ```
       /usr/sbin/rsct/bin/nlssrc -c -ls hags.partition
      ```

      On a HACMP node, issue this command:

      ```
      lssrc -ls grpsvcs
      ```
   e. Output of the **hagsns** command.

      On a PSSP node, issue this command:

      ```
      /usr/sbin/rsct/bin/hagsns -c -s hags
      ```

      On the PSSP control workstation, issue this command:

      ```
      /usr/sbin/rsct/bin/hagsns -c -s hags.partition
      ```

      On a HACMP node, issue this command:

      ```
      /usr/sbin/rsct/bin/hagsns -c -s grpsvcs
      ```
   f. Output of the **hagsvote** command.

      On a PSSP node, issue this command:

      ```
      /usr/sbin/rsct/bin/hagsvote -c -ls hags
      ```

      On the PSSP control workstation, issue this command:

      ```
      /usr/sbin/rsct/bin/hagsvote -c -ls hags.partition
      ```

      On a HACMP node, issue this command:

```
/usr/sbin/rsct/bin/hagsvote -c -ls grpsvcs
```

g. On any PSSP nodes, collect the output of the following command: **SDRGetObjects GS_Config**.

3. System Data

   a. Installation data - output of the **lslpp -L** command.

   b. Contents of the following files:

      - **/usr/sbin/rsct/optlevel.rsct.basic.rte**
      - **/usr/sbin/rsct/optlevel.rsct.basic.hacmp**
      - **/usr/sbin/rsct/optlevel.rsct.basic.sp**

4. Topology Services information. See “Chapter 23. Diagnosing Topology Services problems” on page 351.

How to find the GS nameserver (NS) node

Perform these steps to find out which node is the GS nameserver node.

1. Issue one of these commands:

   - Issue this command on the PSSP control workstation:

     ```
     lssrc -ls hags.partition_name
     ```

   - Issue this command on a PSSP node:

     ```
     lssrc -ls hags
     ```

   - Issue this command on an HACMP node:

     ```
     lssrc -ls grpsvcs
     ```

   If the output is similar to:

   ```
    Subsystem         Group            PID     Status
     hags             hags            14460   active
   1 locally-connected clients. Their PIDs:
    10596 (hagsglsmd)
   HA Group Services domain information:
   Domain established by node 6
   Number of groups known locally: 1
                     Number of   Number of local
   Group name        providers   providers/subscribers
   cssMembership           9          1          0
   ```

   you can obtain the node number of the nameserver. In this case, it is node 6, from the line `Domain established by node 6`. Do not perform any of the remaining steps.

2. If the output indicates `Domain not established`, wait to see if the problem is resolved in a few minutes, and if not proceed to “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

3. There is another command that is designed for the NS status display. Issue the command:

   ```
   /usr/sbin/rsct/bin/hagsns -s subsystem
   ```

   Output is similar to:

```
HA GS NameServer Status
NodeId=1.16, pid=14460, domainId=6.14, NS established,
  CodeLevel=GSlevel(DRL=8)
NS state=kCertain, protocolInProgress=kNoProtocol,
  outstandingBroadcast=kNoBcast
Process started on Jun 19 18:34:20, (10d 20:19:22) ago.
 HB connection took (19:14:9).
Initial NS certainty on Jun 20 13:48:45, (10d 1:4:57) ago,
 taking (0:0:15).
Our current epoch of Jun 23 13:05:19 started on (7d 1:48:23),  ago.
Number of UP nodes: 12
List of UP nodes:  0 1 5 6 7 8 9 11 17 19 23 26
```

In the preceding example, `domainId=6.14` means that node 6 is the NS node. Note that the `domainId` consists of a node number and an incarnation number. The incarnation number is an integer, incremented whenever the GS daemon is started.

4. The **hagsns** command output on the NS also displays the list of groups:

```
We are: 6.14 pid: 10094 domaindId = 6.14 noNS = 0 inRecovery = 0
 CodeLevel=GSlevel(DRL=8)
NS state=kBecomeNS, protocolInProgress = kNoProtocol,
outstandingBroadcast = kNoBcast
Process started on Jun 19 18:35:55, (10d 20:22:39) ago.
HB connection took (0:0:0).
Initial NS certainty on Jun 19 18:36:12, (10d 20:22:22) ago,
  taking (0:0:16).
Our current epoch of certainty started on Jun 23 13:05:18,
  (7d 1:53:16) ago.
Number of UP nodes: 12
List of UP nodes: 0 1 5 6 7 8 9 11 17 19 23 26
List of known groups:
1.1 cssMembership: GL: 6 seqNum: 73
 theIPS: 6 1 26 17 0 8 7 9 5 11 lookupQ:
2.1 ha_em_peers: GL: 6 seqNum: 30
 theIPS: 6 0 8 7 5 11 lookupQ:
```

In this example, the groups are **cssMembership** and **ha_em_peers**.

How to find the Group Leader (GL) node for a specific group

There are two ways of finding the Group Leader node of a specific group:

1. The **hagsns** command on the NS displays the list of membership for groups, including their Group Leader nodes. To use this method:
   a. Find the NS node from “How to find the GS nameserver (NS) node” on page 432.
   b. Issue the following command on the NS node:

   ```
   /usr/sbin/rsct/bin/hagsns -s hags
   ```

   The output is similar to:

   ```
   HA GS NameServer Status
   NodeId=6.14, pid=10094, domainId=6.14, NS established,
    CodeLevel=GSlevel(DRL=8)
   NS state=kBecomeNS, protocolInProgress=kNoProtocol,
     outstandingBroadcast=kNoBcast
   Process started on Jun 19 18:35:55, (10d 20:22:39) ago.
     HB connection took (0:0:0).
   Initial NS certainty on Jun 19 18:36:12, (10d 20:22:22) ago,
   ```

```
  taking (0:0:16).
Our current epoch of certainty started on Jun 23 13:05:18,
  (7d 1:53:16) ago.
Number of UP nodes: 12
List of UP nodes:  0 1 5 6 7 8 9 11 17 19 23 26
List of known groups:
1.1 cssMembership: GL: 6 seqNum: 73
  theIPS: 6 1 26 17 0 8 7 9 5 11 lookupQ:
2.1 ha_em_peers: GL: 6 seqNum: 30 theIPS: 6 0 8 7 5 11 lookupQ:
```

The bottom few lines display the group membership information. For example, the GL node of the group **cssMembership** is node 6, and its participating nodes are ″6 1 26 17 0 8 7 9 5 11″.

2. If you need only the GL node of a specific group, the **hagsvote** command gives the answer. Issue the command:

   ```
   hagsvote -s hags
   ```

   The output is similar to:

   ```
   Number of groups: 3
   Group slot #[0] Group name [HostMembership] GL node [Unknown] voting data:
   No protocol is currently executing in the group.
   --------------------------------------------------

   Group slot #[1] Group name [cssRawMembership] GL node [Unknown] voting data:
   No protocol is currently executing in the group.
   --------------------------------------------------

   Group slot #[2] Group name [cssMembership] GL node [6] voting data:
   No protocol is currently executing in the group.
   ```

   In this output, node 6 is the GL node of the group **cssMembership**. If the GL node is `Unknown`, this indicates that no client applications tried to use the group on this node.

Diagnostic procedures

Installation verification test

This test determines whether RSCT has been successfully installed. Group Services is a part of RSCT. Perform the following steps:

1. Issue the command:

   ```
   lslpp -l | grep rsct
   ```

   Good results are indicated by output similar to:

   ```
    rsct.basic.hacmp     1.2.0.0  COMMITTED  RS/6000 Cluster Technology (HACMP domains)
    rsct.basic.rte       1.2.0.0  COMMITTED  RS/6000 Cluster Technology (all domains)
    rsct.basic.sp        1.2.0.0  COMMITTED  RS/6000 Cluster Technology (SP domains)
    rsct.clients.hacmp   1.2.0.0  COMMITTED  RS/6000 Cluster Technology (HACMP domains)
    rsct.clients.rte     1.2.0.0  COMMITTED  RS/6000 Cluster Technology (all domains)
    rsct.clients.sp      1.2.0.0  COMMITTED  RS/6000 Cluster Technology (SP domains)
    rsct.core.utils      1.2.0.0  COMMITTED  RS/6000 Cluster Technology (all domains)
   ```

   Error results are indicated by no output from the command.

2. Issue the command:

```
lppchk -c "rsct*"
```

Good results are indicated by the absence of error messages and the return of a zero exit status from this command. The command produces no output if it succeeds.

Error results are indicated by a non-zero exit code and by error messages similar to these:

```
lppchk: 0504-206  File /usr/lib/nls/msg/en_US/hats.cat could not be located.
lppchk: 0504-206  File /usr/sbin/rsct/bin/hatsoptions could not be located.
lppchk: 0504-208  Size of /usr/sbin/rsct/bin/phoenix.snap is 29356,
                        expected value was 29355.
```

Some error messages may appear if an EFIX is applied to a file set. An EFIX is an emergency fix, supplied by IBM, to correct a specific problem.

If the test fails, the following file sets need to be installed:

1. **rsct.basic.rte**
2. **rsct.core.utils**
3. **rsct.clients.rte**
4. **rsct.basic.sp**
5. **rsct.clients.sp**
6. **rsct.basic.hacmp**
7. **rsct.clients.hacmp**

If this test is successful, proceed to “Configuration verification test”.

Configuration verification test

This test verifies that Group Services on a PSSP node has the configuration data that it needs. Perform the following steps:

1. Perform the Topology Services Configuration verification diagnosis. See “Chapter 23. Diagnosing Topology Services problems” on page 351.
2. If it succeeds, issue the following commands to display data from the SDR and obtain the level of PSSP:
   a. `SDRGetObjects Syspar`
   b. `splst_versions -t`

   Good results are indicated by none of the preceding commands returning an error message, and all commands returning non-null output.

   Error results are indicated if the preceding commands fail. In this case, the SDR could be experiencing problems. Diagnose the SDR subsystem by referring to “Chapter 14. Diagnosing SDR problems” on page 125. If the commands succeed but do not show the expected information, it is possible that a problem occurred in the installation of the nodes. Verify installation of the nodes by consulting “Chapter 9. Diagnosing node installation problems” on page 101.

If this test is successful, proceed to “Operational verification tests” on page 436.

Operational verification tests

The following information applies to the diagnostic procedures that follow:

- Subsystem Name:
  - For PSSP on the control workstation: **hags.***partition_name*
  - For PSSP on nodes: **hags**
  - For HACMP/ES: **grpsvcs**
- Service and User log files:
  - For PSSP: **/var/ha/log/hags_***
  - For HACMP: **/var/ha/log/grpsvcs_***
- Startup Script log:
  - For PSSP: **/var/ha/log/hags.default***
  - For HACMP: **/var/ha/log/grpsvcs.default***

Operational test 1 - Verify that Group Services is working properly

Issue the following command:

```
lssrc -ls subsystem_name
```

Good results are indicated by an output similar to:

```
Subsystem         Group            PID     Status
 hags             hags             22962    active
2 locally-connected clients.  Their PIDs:
20898(hagsglsmd) 25028(haemd)
HA Group Services domain information:
Domain established by node 21
Number of groups known locally: 2
                   Number of   Number of local
Group name         providers   providers/subscribers
cssMembership          10          0           1
ha_em_peers            6           1           0
```

There **must** be an entry for **cssMembership**.

Error results are indicated by one of the following:

1. A message similar to:

   ```
   0513-036 The request could not be passed to the hags subsystem.
       Start the subsystem and try your command again.
   ```

   This means that the GS daemon is not running. The GS subsystem is down. Issue the **errpt** command and look for an entry for the subsystem name. Proceed to “Operational test 2 - Determine why the Group Services subsystem is not active” on page 438.
2. A message similar to:

   ```
   0513-085 The hags Subsystem is not on file.
   ```

   This means that the GS subsystem is not defined to the AIX SRC.

   In PSSP, the partition-sensitive subsystems may have been undefined by the **syspar_ctrl** command. Use **syspar_ctrl -a** to add the subsystems to the node.

In HACMP/ES, HACMP may have not been installed on the node. Check the HACMP subsystem.

3. Output similar to:

   ```
   Subsystem         Group            PID     Status
   hags.c47s         hags             7350    active
    Subsystem hags.c47s trying to connect to Topology Services.
   ```

   This means that Group Services is not connected to Topology Services. Check the Topology Services subsystem. See “Chapter 23. Diagnosing Topology Services problems” on page 351.

4. Output similar to:

   ```
   Subsystem         Group            PID     Status
   hags.c47s         hags             35746   active
    No locally-connected clients.
    HA Group Services domain information:
    Domain not established.
    Number of groups known locally: 0
   ```

   This means that the GS domain is not established. This is normal during the Group Services startup period. Retry this test after about three minutes. If this situation continues, perform “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

5. Output similar to:

   ```
   Subsystem         Group            PID     Status
   hags.c47s         hags             35746   active
    No locally-connected clients.
    HA Group Services domain information:
    Domain is recovering.
    Number of groups known locally: 0
   ```

   This means that the GS domain is recovering. It is normal during Group Services domain recovery. Retry this test after waiting three to five minutes. If this situation continues, perform “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

6. An output similar to the **Good results**, but no **cssMembership** group is shown on the control workstation or the PSSP nodes. Proceed to “Operational test 7 - Verify the HAGSGLSM (Group Services GLobalized Switch Membership) subsystem” on page 444.

7. Output similar to:

   ```
   Subsystem         Group            PID     Status
   hags              hags             25132   active
    No locally-connected clients.
    HA Group Services domain information:
    Domain established by node 1.
    Number of groups known locally: 0
   ```

   This means that no GS clients are connected, or no local groups are established. The GS daemon is working normally for a while at startup time, or one of the following conditions occurred:

   a. The **haem** subsystem is not running on the control workstation. Issue this command to start the **haem** subsystem.

      ```
      lssrc -s haem.partition_name
      ```

The output is similar to:

```
Subsystem         Group            PID     Status
haem.c47s         haem                    inoperative
```

b. The **haem** subsystem and the switch are not working on the nodes. Issue the command: **lssrc -ls hats**. The output is similar to:

```
Subsystem         Group            PID     Status
hats              hats             25074   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
SPether        [0]   15   12  S 9.114.61.65     9.114.61.195
SPether        [0] en0          0x376d296c      0x3779180b
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [1]   14    0  D 9.114.61.129
SPswitch       [1] css0
HB Interval = 1 secs. Sensitivity = 4 missed beats
  1 locally connected Client with PID:
hagsd( 14460)
  Configuration Instance = 925928580
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  Control Workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size 7052 KB
```

Look for `SPswitch`. The line

```
SPswitch       [1]   14    0  D 9.114.61.129    9.114.61.154
```

implies that the switch is not working or Topology Services thinks that the switch is down. For more information, see “Chapter 23. Diagnosing Topology Services problems” on page 351.

c. If the two preceding conditions do not apply, see “Operational test 5 - Verify whether the cssMembership or css1Membership groups are found on a node” on page 442.

Operational test 2 - Determine why the Group Services subsystem is not active

Issue the command:

```
errpt -N hags subsystem_name
```

where *subsystem_name* is:

- **hags** for PSSP nodes
- **hags.***partition_name* for the PSSP control workstation
- **grpsvcs** for HACMP nodes

and look for an entry for the *subsystem_name*. It appears under the `RESOURCE_NAME` heading.

If an entry is found, issue the command:

```
errpt -a -N hags subsystem_name
```

to get details about error log entries. The entries related to Group Services are those with `LABEL` beginning with **GS_**.

The error log entry, together with its description in “AIX Error Logs and templates” on page 420, explains why the subsystem is inactive.

If there is no **GS_** error log entry explaining why the subsystem went down or could not start, it is possible that the daemon may have exited abnormally. Look for an error entry with LABEL of CORE_DUMP and PROGRAM NAME of **hagsd**, by issuing the command:

```
errpt -J CORE_DUMP
```

If this entry is found, see "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.

Another possibility when there is no **GS_** error log entry is that the Group Services daemon could not be loaded. In this case, a message similar to the following may be present in the Group Services startup log:

```
0509-036 Cannot load program hagsd because of the following errors:
0509-026 System error: Cannot run a file that does not have a valid format.
```

The message may refer to the Group Services daemon, or to some other program invoked by the startup script **hags**. If this error is found, see "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.

For errors where the daemon did start up but then exited during initialization, detailed information about the problem is in the Group Services error log.

Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered

The **hagsns** command is used to determine the nameserver (NS) state and characteristics. Issue the command:

```
hagsns -s subsystem_name
```

The output is similar to:

```
HA GS NameServer Status
NodeId=0.32, pid=18256, domainId=0.Nil, NS not established,
  CodeLevel=GSlevel(DRL=8)
The death of the node is being simulated.
NS state=kUncertain, protocolInProgress=kNoProtocol,
  outstandingBroadcast=kNoBcast
Process started on Jun 21 10:33:08, (0:0:16) ago.
  HB connection took (0:0:0).
Our current epoch of uncertainty started on Jun 21 10:33:08,
  (0:0:16) ago.
Number of UP nodes: 1
List of UP nodes:  0
```

Error results are indicated by output of NS state is kUncertain, with the following considerations:

1. kUncertain is normal for a while after Group Services startup.
2. Group Services may have instructed Topology Services to simulate a node death. This is so that every other node will see the node down event for this local node. This simulating node death state will last approximately two or three minutes.

If this state does not change or takes longer than two or three minutes, proceed to check Topology Services. See "Chapter 23. Diagnosing Topology Services problems" on page 351.

If the Group Services daemon is not in kCertain or kBecomeNS state, and is waiting for the other nodes, the **hagsns** command output is similar to:

```
HA GS NameServer Status
NodeId=11.42, pid=21088, domainId=0.Nil, NS not established,
  CodeLevel=GSlevel(DRL=8)
NS state=kGrovel, protocolInProgress=kNoProtocol,
  outstandingBroadcast=kNoBcast
Process started on Jun 21 10:52:13, (0:0:22) ago.
  HB connection took (0:0:0).
Our current epoch of uncertainty started on Jun 21 10:52:13,
  (0:0:22) ago.
Number of UP nodes: 2
List of UP nodes:  0 11
Domain not established for (0:0:22).
  Currently waiting for node 0
```

In the preceding output, this node is waiting for an event or message from node 0 or for node 0. The expected event or message differs depending on the NS state which is shown in the second line of the **hagsns** command output.

Analyze the NSstate as follows:

1. kGrovel means that this node believes that the waiting node (node 0 in this example) will become his NS. This node is waiting for node 0 to acknowledge it (issue a Proclaim message).
2. kPendingInsert or kInserting means that the last line of the **hagsns** command output is similar to:

   ```
   Domain not established for (0:0:22).  Currently waiting for node 0.1
   ```

   This node received the acknowledge (Proclaim or InsertPhase1 message) and is waiting for the next message (InsertPhase1 or Commit message) from the NS (node 0).

   If this state does not change to kCertain in a two or three minutes, proceed to “Operational test 1 - Verify that Group Services is working properly” on page 436, for Topology Services and Group Services on the waiting node (node 0 in this example).
3. kAscend, kAscending, kRecoverAscend, or kRecoverAscending means that the last line of the **hagsns** command output is similar to:

   ```
   Domain not established for (0:0:22).  Waiting for 3 nodes: 1 7 6
   ```

   If there are many waiting nodes, the output is similar to:

   ```
   Domain not established for(0:0:22).Waiting for 43 nodes: 1 7 6 9 4 ....
   ```

   This node is trying to become a nameserver, and the node is waiting for responses from the nodes that are listed in the **hagsns** command output. If this state remains for between three and five minutes, proceed to “Operational test 1 - Verify that Group Services is working properly” on page 436, for Topology Services and Group Services on the nodes that are on the waiting list.
4. kKowtow or kTakeOver means that the last line of the **hagsns** command output is similar to:

   ```
   Domain not recovered for (0:0:22).  Currently waiting for node 0.1
   ```

After the current NS failure, this node is waiting for a candidate node that is becoming the NS. If this state stays too long, proceed to “Operational test 1 - Verify that Group Services is working properly” on page 436, for the Topology Services and Group Services on the node that is in the waiting list.

In this output, the value `0.1` means the following:

- The first number (″0″) indicates the node number that this local node is waiting for.
- The second number(″1″) is called the incarnation number, which is increased by one whenever the GS daemon starts.

Therefore, this local node is waiting for a response from the GS daemon of node 0, and the incarnation is 1.

Operational test 4 - Verify whether a specific group is found on a node

Issue the following command:

```
lssrc -ls subsystem_name
```

Error results are indicated by outputs similar to the **error results** of “Operational test 1 - Verify that Group Services is working properly” on page 436 through “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

Good results are indicated by an output similar to:

```
Subsystem         Group            PID     Status
 hags             hags             22962    active
2 locally-connected clients.  Their PIDs:
20898(hagsglsmd) 25028(haemd)
HA Group Services domain information:
Domain established by node 21
Number of groups known locally: 2
                   Number of   Number of local
Group name         providers   providers/subscribers
cssMembership          10          0          1
ha_em_peers            6           1          0
```

In this output, examine the `Group name` field to see whether the requested group name exists. For example, the group **ha_em_peers** has 1 local provider, 0 local subscribers, and 6 total providers.

For more information about the given group, issue the command:

```
hagsns -s subsystem_name
```

on the NS node. The output is similar to:

```
HA GS NameServer Status
NodeId=6.14, pid=10094, domainId=6.14, NS established,
  CodeLevel=GSlevel(DRL=8)
NS state=kBecomeNS, protocolInProgress=kNoProtocol,
 outstandingBroadcast=kNoBcast
Process started on Jun 19 18:35:55, (10d 20:22:39) ago.
 HB connection took (0:0:0).
Initial NS certainty on Jun 19 18:36:12, (10d 20:22:22) ago,
 taking (0:0:16).
Our current epoch of certainty started on Jun 23 13:05:18,
 (7d 1:53:16) ago.
Number of UP nodes: 12
```

```
List of UP nodes:  0 1 5 6 7 8 9 11 17 19 23 26
List of known groups:
1.1 cssMembership: GL: 6 seqNum: 73
 theIPS: 6 1 26 17 0 8 7 9 5 11 lookupQ:
2.1 ha_em_peers: GL: 6 seqNum: 30 theIPS: 6 0 8 7 5 11 lookupQ:
```

In the last line, the nodes that have the providers of the group **ha_em_peers** are 6 0 8 7 5 11.

Operational test 5 - Verify whether the cssMembership or css1Membership groups are found on a node

If “Operational test 1 - Verify that Group Services is working properly” on page 436 through “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439 succeeded, issue the following command:

```
lssrc -ls subsystem_name
```

The output is similar to:

```
Subsystem         Group            PID     Status
 hags             hags             22962    active
2 locally-connected clients.  Their PIDs:
20898(hagsglsmd) 25028(haemd)
HA Group Services domain information:
Domain established by node 21
Number of groups known locally: 2
                   Number of   Number of local
Group name         providers   providers/subscribers
cssMembership          10          1          0
ha_em_peers            6           1          0
```

In the preceding output, the **cssMembership** group has 1 local provider. Otherwise, the following conditions apply:

1. No **cssMembership** or **css1Membership** exists in the output.

   There are several possible causes:

   a. **/dev/css0** or **/dev/css1** devices are down.

      Perform switch diagnosis. See “Chapter 15. Diagnosing SP Switch problems” on page 137 or “Chapter 16. Diagnosing SP Switch2 problems” on page 185.

   b. Topology Services reports that the switch is not stable.

      Issue the following command:

      ```
      lssrc -ls hats_subsystem
      ```

      where *hats_subsystem* is:

      - **hats** on PSSP nodes
      - **hats.***partition_name* on the PSSP control workstation
      - **topsvcs** on HACMP nodes

      The output is similar to:

      ```
      Subsystem       Group           PID     Status
       hats           hats            17058   active
      Network Name   Indx Defd Mbrs St Adapter ID      Group ID
      SPether        [0]   15    2  S 9.114.61.65     9.114.61.125
      SPether        [0] en0          0x37821d69      0x3784f3a9
      ```

```
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [1]   14    0  D 9.114.61.129
SPswitch       [1] css0
HB Interval = 1 secs. Sensitivity = 4 missed beats
  1 locally connected Client with PID:
hagsd( 26366)
  Configuration Instance = 926456205
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  Control Workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size 7044 KB
```

Find the first `SPswitch` row in the `Network Name` column. Find the `St` (state) column in the output. At the intersection of the first `SPswitch` row and state column is a letter. If it is not **S**, wait for few minutes longer since the Topology Services SPswitch group is not stable. If the state stays too long as **D** or **U**, proceed to Topology Services diagnosis. See “Chapter 23. Diagnosing Topology Services problems” on page 351. If the state is **S**, proceed to Step 1c. In this example, the state is **D**.

The state has the following values:

- **S** - stable or working correctly
- **D** - dead, or not working
- **U** - unstable (not yet incorporated)

c. **HAGSGLSM** is not running or waiting for Group Services protocols.

Proceed to “Operational test 7 - Verify the HAGSGLSM (Group Services GLobalized Switch Membership) subsystem” on page 444.

2. **cssMembership** or **css1Membership** exist in the output, but the number of local providers is zero.

   Proceed to “Operational test 7 - Verify the HAGSGLSM (Group Services GLobalized Switch Membership) subsystem” on page 444.

Operational test 6 - Verify whether Group Services is running a protocol for a group

Issue the following command:

```
hagsvote -ls subsystem
```

Compare the output to this list of choices.

1. If no protocol is running, the output is similar to:

```
Number of groups: 3
Group slot #[0] Group name [HostMembership] GL node [Unknown]
 voting data: No protocol is currently executing in the group.
--------------------------------------------------------------

Group slot #[1] Group name [cssRawMembership] GL node [Unknown]
 voting data: No protocol is currently executing in the group.
--------------------------------------------------------------

Group slot #[2] Group name [theSourceGroup] GL  node [1]
 voting data: No protocol is currently executing in the group.
---------------------------------------------------------------
```

   In this output, no protocol is running for ″theSourceGroup″.

2. A protocol is running and waiting for a vote. For the group `theSourceGroup`, this node is soliciting votes and waiting for the local providers to vote. The output is similar to:

```
Group slot #[2] Group name [theSourceGroup] GL node [1]
 voting data: Not GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
Number of providers: 1
Number of providers not yet voted: 1 (vote not submitted).
Given vote:[No vote value] Default vote:[No vote value]
-------------------------------------------------------
```

The number of local providers is 1, and no voting is submitted. Its Group Leader (GL) node is 1. The output of the same command on the GL node (node 1) is similar to:

```
Group slot #[3] Group name [theSourceGroup] GL node [1] voting data:
GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
Number of providers: 1
Number of providers not yet voted: 0 (vote submitted).
Given vote:[Approve vote] Default vote:[No vote value]
Global voting data:
Number of providers not yet voted: 1
Given vote:[Approve vote] Default vote:[No vote value]
---------------------------------------------------
```

This indicates that a total of one provider has not voted.

Operational test 7 - Verify the HAGSGLSM (Group Services GLobalized Switch Membership) subsystem

Issue the following command:

```
lssrc -ls glsm_subsystem
```

where *glsm_subsystem* is:

- **hagsglsm** on PSSP nodes
- **hagsglsm.***partition-name* on the PSSP control workstation
- **grpglsm** on HACMP nodes

Good results are indicated by output similar to:

- On the control workstation,

```
Subsystem         Group            PID     Status
 hagsglsm.c47s   hags             22192   active
Status information for subsystem hagsglsm.c47s:
Connected to Group Services.
 Adapter  Group                  Mbrs   Joined  Subs'd  Aliases
 css0     (device does not exist)
          cssMembership         0       No      Yes     -
 css1     (device does not exist)
          css1Membership        0       No      Yes     -
 ml0      ml0Membership         -       No      -
Aggregate Adapter Configuration
 The current configuration id is 0x1482933.
 ml0[css0] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
 ml0[css1] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
```

- On other nodes,

```
Subsystem         Group            PID     Status
 hagsglsm         hags             16788   active
Status information for subsystem hagsglsm:
Connected to Group Services.
 Adapter  Group                  Mbrs   Joined  Subs'd  Aliases
 css0     cssRawMembership      16      -       Yes     1
          cssMembership         16      Yes     Yes     -
 css1     css1RawMembership     16      -       Yes     1
          css1Membership        16      Yes     Yes     -
 ml0      ml0Membership         16      Yes     -       cssMembership
Aggregate Adapter Configuration
 The current configuration id is 0x23784582.
 ml0[css0] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
 ml0[css1] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
```

Error results are indicated by one of the following outputs:

1. A message similar to:

   ```
   0513-036 The request could not be passed to the hags subsystem.
            Start the subsystem and try your command again.
   ```

   This means that the HAGSGLSM daemon is not running. The subsystem is down. Issue the **errpt** command and look for an entry for the subsystem name. Proceed to “Operational test 2 - Determine why the Group Services subsystem is not active” on page 438.
2. A message similar to:

   ```
   0513-085 The hagsglsm Subsystem is not on file.
   ```

   This means that the HAGSGLSM subsystem is not defined to the AIX SRC.

   For PSSP nodes, the partition-sensitive subsystems may have been undefined by the **syspar_ctrl** command. The same command may be used to add the subsystems to the node.

   In HACMP/ES, HACMP may have not been installed on the node. Check the HACMP subsystem.
3. Output similar to:

   ```
   Subsystem         Group            PID     Status
   hagsglsm.c47s     hags             26578   active
   Status information for subsystem hagsglsm.c47s:
   Not yet connected to Group Services after 4 connect tries
   ```

   HAGSGLSM is not connected to Group Services. The Group Services daemon is not running. If the state is **S**, proceed to “Operational test 1 - Verify that Group Services is working properly” on page 436 for Group Services subsystem verification.
4. Output similar to:

   ```
   Subsystem         Group            PID     Status
    bhagsglsm        bhags            16048   active
   Status information for subsystem bhagsglsm:
   Waiting for Group Services response.
   ```

   HAGSGLSM is being connected to Group Services. Wait for a few seconds. If this condition does not change after several seconds, proceed to “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439 or “Operational test 6 - Verify whether Group Services is running a protocol for a group” on page 443.
5. Output similar to:

   ```
   Subsystem         Group            PID     Status
    hagsglsm         hags             26788   active
   Status information for subsystem hagsglsm:
   ```

```
Connected to Group Services.
 Adapter  Group                    Mbrs   Joined  Subs'd  Aliases
 css0     cssRawMembership        -       -       No      -
          cssMembership           16       No     No      -
 css1     css1RawMembership       15      -       Yes     1
          css1Membership          15      Yes     Yes     -
 ml0      ml0Membership            -      -      -        -
Aggregate Adapter Configuration
 The current configuration id is 0x23784582.
 ml0[css0] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
 ml0[css1] Nodes: 1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
```

On nodes that have the switch, the line ″`cssRawMembership`″ or ″`css1RawMembership`″ have `No` in the `Subs'd` column.

Check Topology Services to see whether the switch is working. Issue the command:

```
lssrc -ls hats_subsystem
```

The output is similar to:

```
Subsystem         Group            PID     Status
 hats             hats             25074   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
SPether        [0]   15   11  S 9.114.61.65     9.114.61.193
SPether        [0] en0          0x376d296c      0x3784fdc5
HB Interval = 1 secs. Sensitivity = 4 missed beats
SPswitch       [1]   14    8  S 9.114.61.129    9.114.61.154
SPswitch       [1] css0         0x376d296d      0x3784fc48
HB Interval = 1 secs. Sensitivity = 4 missed beats
  1 locally connected Client with PID:
hagsd( 14460)
  Configuration Instance = 925928580
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
  Control Workstation IP address = 9.114.61.125
  Daemon employs no security
  Data segment size 7052 KB
```

Find the first row under `Network Name` with `SPswitch`. Find the column with heading `St` (state). Intersect this row and column. If the value at the intersection is not **S**, see **TS_LOC_DOWN_ST** on page 363 and proceed to “Action 3 - Correct local adapter problem” on page 404.

If the state is **S**, proceed to “Operational test 1 - Verify that Group Services is working properly” on page 436 to see whether the Group Services domain is established or not. If the Group Services domain is established, proceed to “Operational test 6 - Verify whether Group Services is running a protocol for a group” on page 443 for **cssMembership** protocol activity.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with Group Services. Locate the symptom and perform the action described in the following table:

Table 61. Group Services symptoms

Symptom	Error label	Recovery
GS daemon cannot start.	GS_STARTERR_ER	See “Action 1 - Start Group Services daemon” on page 447.

Table 61. Group Services symptoms (continued)

Symptom	Error label	Recovery
GS domains merged.	GS_DOM_MERGE_ER	See "Action 2 - Verify Status of Group Services Subsystem" on page 448.
GS clients cannot connect or join the GS daemon.	The following errors may be present: GS_AUTH_DENIED_ST GS_CLNT_SOCK_ER GS_DOM_NOT_FORM_WA	See "Action 3 - Correct Group Services access problem" on page 448.
GS daemon died unexpectedly.	The following errors may be present: GS_ERROR_ER GS_DOM_MERGE_ER GS_TS_RETCODE_ER GS_STOP_ST GS_XSTALE_PRCLM_ER	See "Action 4 - Correct Group Services daemon problem" on page 450.
GS domain cannot be established or recovered.	The following errors may be present: GS_STARTERR_ER GS_DOM_NOT_FORM_WA	See "Action 5 - Correct domain problem" on page 450.
GS protocol has not been completed for a long time.	None	See "Action 6 - Correct protocol problem" on page 451.
HAGSGLSM cannot start.	GS_GLSM_STARTERR_ER	See "Action 7 - Correct hagsglsm startup problem" on page 452.
HAGSGLSM has stopped.	GS_GLSM_ERROR_ER or None	See "Action 8 - hagsglsm daemon has stopped" on page 452.
Non-stale proclaim message received.	GS_XSTALE_PRCLM_ER	See "Action 9 - Investigate non-stale proclaim message" on page 452.

Actions

Action 1 - Start Group Services daemon

Some of the possible causes are:

- SDR-related problems that prevent the startup script from obtaining configuration data from the SDR.
- AIX-related problems such as a shortage of space in the **/var** directory or a port number already in use.
- SRC-related problems that prevent the daemon from setting the appropriate SRC environment.

Run the diagnostics in "Operational test 2 - Determine why the Group Services subsystem is not active" on page 438 to determine the cause of the problem.

Action 2 - Verify Status of Group Services Subsystem

The AIX error log has a **GS_DOM_MERGE_ER**, and the Group Services daemon has restarted. The most common cause of this situation is for Group Services daemon to receive a **NODE_UP** event from Topology Services after the Group Services daemon formed more than one domain.

If the Group Services daemon has been restarted and a domain has been formed, no action is needed. However, if the Group Services daemon is not restarted, perform “Operational test 1 - Verify that Group Services is working properly” on page 436 to verify the status of the GS subsystem.

Perform these steps:

1. Find a node with the **GS_DOM_MERGE_ER** AIX error log entry.
2. Find the **GS_START_ST** entry before the **GS_DOM_MERGE_ER** in the AIX error log.
3. If there is a **GS_START_ST** entry, issue the command:

   ```
   lssrc -l -s subsystem_name
   ```

   where *subsystem_name* is:

   - **hags.***partition_name* on the PSSP control workstation
   - **hags** on PSSP nodes
   - **grpsvcs** on HACMP nodes
4. The **lssrc** output contains the node number that established the GS domain.
5. Otherwise, proceed to “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

After the merge, the Group Services daemon must be restarted. See **TS_NODEUP_ST** on page 368. Check it with “Operational test 2 - Determine why the Group Services subsystem is not active” on page 438.

Action 3 - Correct Group Services access problem

For the nodes that cannot join, some of the possible causes are:

1. Group Services may not be running.
2. Group Services domain may not be established.
3. The clients may not have permission to connect to the Group Services daemon.
4. Group Services is currently doing a protocol for the group that is trying to join or subscribe.

Analyze and correct this problem as follows:

1. Issue the command:

   ```
   lssrc -s subsystem
   ```

   where *subsystem_name* is:

   - **hags.***partition_name* on the PSSP control workstation
   - **hags** on PSSP nodes
   - **grpsvcs** on HACMP nodes

   The output is similar to:

```
Subsystem         Group            PID     Status
 hags.c47s        hags             23482   active
```

If `Status` is not `active`, this indicates that the node cannot join the GS daemon. Perform “Operational test 2 - Determine why the Group Services subsystem is not active” on page 438. Start the Group Services subsystem by issuing this command:

```
/usr/sbin/rsct/bin/hagsctrl -s
```

If `Status` is `active`, proceed to Step 2.

2. Perform “Operational test 1 - Verify that Group Services is working properly” on page 436 to check whether the Group Services domain is established or not.
3. Issue the command:

   ```
   errpt -a -N subsystem_name | more
   ```

   where *subsystem_name* is:
   - **hags** on PSSP nodes
   - **hags.***partition name* on the PSSP control workstation
   - **grpsvsc** on HACMP nodes

   Check the AIX error log for this entry:

   ```
   ---------------------------------------------------
   LABEL:          GS_AUTH_DENIED_ST
   IDENTIFIER:     23628CC2

   Date/Time:       Tue Jul 13 13:29:52
   Sequence Number: 213946
   Machine Id:      000032124C00
   Node Id:         c47n09
   Class:           O
   Type:            INFO
   Resource Name:   hags

   Description
   User is not allowed to use Group Services daemon

   Probable Causes
   The user is not the root user
   The user is not a member of hagsuser group

   Failure Causes
   Group Services does not allow the user

           Recommended Actions
           Check whether the user is the root
   Check whether the user is a member of hagsuser group

   Detail Data
   DETECTING MODULE
   RSCT,SSuppConnSocket.C,           1.17, 421
   ERROR ID
   .0ncMX.ESrWr.Oin//rXQ7....................
   REFERENCE CODE

   DIAGNOSTIC EXPLANATION
   User myuser1 is not a supplementary user of group 111. Connection refused.
   ```

This explains that the user (myuser1) of the client program does not have correct permission to use Group Services.

The following users can access Group Services:

- The **root** user.
- A user who is a primary or supplementary member of the **hagsuser** group, which is defined in the **/etc/group** file.

Change the ownership of the client program to a user who can access Group Services.

4. Issue the command:

   ```
   hagsvote -ls subsystem
   ```

   to determine whether the group is busy, and to find the Group Leader node for the specific group.
5. Issue the same command on the Group Leader Node to determine the global status of the group. Resolve the problem by the client programs.

Action 4 - Correct Group Services daemon problem

Some of the possible causes are:

1. Domain merged.
2. Group Services daemon received a non-stale proclaim message from its NS.

   If the Topology Services daemon is alive when the current NS restarts and tries to become a NS, the newly started NS sends a proclaim message to the other nodes. These nodes consider the newly started node as their NS. The receiver nodes consider the proclaim message current (that is, ″non-stale″) but undefined by design. Therefore, the received Group Services daemon will be core dumped.
3. The Topology Services daemon has died.
4. The Group Services daemon has stopped.
5. Group Services has an internal error that caused a core dump.

Examine the AIX error log by issuing the command:

```
errpt -J GS_DOM_MERGE_ER,GS_XSTALE_PRCLM_ER,GS_ERROR_ER,GS_STOP_ST,\
GS_TS_RETCODE_ER | more
```

and search for **GS_** labels or a RESOURCE NAME of any of the GS subsystems. If an entry is found, the cause is explained in the DIAGNOSTIC EXPLANATION field.

If Group Services has taken a core dump, the AIX error log will have the **CORE_DUMP** label with RESOURCE NAME of any of the GS subsystems. In this case, the core file is in: **/var/ha/run/***gs_subsystem.partition*. Save this file. See “Action 7 - Investigate Group Services failure” on page 413.

Action 5 - Correct domain problem

Some of the possible causes are:

1. Topology Services is running, but the Group Services daemon is not running on some of the nodes.
2. Group Services internal NS protocol is currently running.

Proceed to “Operational test 3 - Determine why the Group Services domain is not established or why it is not recovered” on page 439.

Action 6 - Correct protocol problem

This is because the related client failed to vote for a specific protocol. Issue this command on any node that has target groups:

```
hagsvote -ls gs_subsystem
```

where *gs_subsystem* is:

- **hags** on PSSP nodes
- **hags.***partition* on the PSSP control workstation
- **grpsvcs** on HACMP nodes

If this node did not vote for the protocol, the output is similar to:

```
Group slot #[3] Group name [theSourceGroup] GL node [0] voting data:
Not GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
Number of providers: 1
Number of providers not yet voted: 1 (vote not submitted).
Given vote:[No vote value] Default vote:[No vote value]
ProviderId      Voted?  Failed? Conditional?
[101/11]        No      No      Yes
```

As the preceding text explains, one of local providers did not submit a vote. If this node has already voted but the overall protocol is still running, the output is similar to:

```
Group slot #[3] Group name [theSourceGroup] GL node [0] voting data:
Not GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
Number of providers: 1
Number of providers not yet voted: 0 (vote submitted).
Given vote:[Approve vote] Default vote:[No vote value]
ProviderId      Voted?  Failed? Conditional?
[101/11]        Yes     No      Yes
```

In this case, issue the same command on the Group Leader node. The output is similar to:

```
Group slot #[2] Group name [theSourceGroup] GL node [0] voting data:
GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
Number of providers: 1
Number of providers not yet voted: 1 (vote not submitted).
Given vote:[Approve vote] Default vote:[No vote value]
ProviderId      Voted?  Failed? Conditional?
[101/0] No      No      No

Global voting data:
Number of providers not yet voted: 1
Given vote:[Approve vote] Default vote:[No vote value]
Nodes that have voted: [11]
Nodes that have not voted: [0]
```

If there is no provider on the group leader node, the output of **hagsvote -ls** *subsystem_name* would be similar to:

```
Number of groups: 1
Group slot #[2] Group name [theSourceGroup] GL node [0] voting data:
GL in phase [1] of n-phase protocol of type [Join].
Local voting data:
No local providers to vote.  Dummy vote submitted.
```

```
Global voting data:
Number of providers not yet voted: 0
Given vote:[No vote value] Default vote:[No vote value]
Nodes that have voted: [0 ]
Nodes that have not voted: [2 ]
```

The GL's output contains the information about the nodes that did not vote. Investigate the reason for their failure to do so. Debug the GS client application.

Action 7 - Correct hagsglsm startup problem

Some of the possible causes are:

- SDR-related problems that prevent the startup script from obtaining configuration data from the SDR.
- AIX-related problems such as a shortage of space in the **/var** directory or a port number already in use.
- SRC-related problems that prevent the daemon from setting the appropriate SRC environment.

Proceed to "Operational test 7 - Verify the HAGSGLSM (Group Services GLobalized Switch Membership) subsystem" on page 444.

Action 8 - hagsglsm daemon has stopped

Issue this command:

```
lssrc -l -s subsystem_name
```

where *subsystem_name* is:

- **hags.***partition_name* on the PSSP control workstation
- **hags** on PSSP nodes
- **grpsvcs** on HACMP nodes

If the daemon is stopped, the output will contain a status of ″inoperative″ for **hagsglsm**. Otherwise, the output will contain a status of ″operative″ for **hagsglsm**. If stopping the daemon was not intended, see "Information to collect before contacting the IBM Support Center" on page 430 and contact the IBM Support Center.

Action 9 - Investigate non-stale proclaim message

The local Group Services daemon receives a valid domain join request (proclaim) message from its NameServer (NS) more than once. This typically happens when Topology Services notifies Group Services of inconsistent node events. This problem should be resolved automatically if a **GS_START_ST** AIX error log entry is seen after the problem occurs.

Perform these actions:

1. Find the **GS_START_ST** AIX error log entry after this one.
2. If there is a **GS_START_ST** entry, issue the command:

   ```
   lssrc -l -s subsystem_name
   ```

   where *subsystem_name* is:

   - **hags.***partition_name* on the PSSP control workstation
   - **hags** on PSSP nodes
   - **grpsvcs** on HACMP nodes
3. The **lssrc** output contains the node number that established the GS domain.

4. Otherwise, proceed to “Action 4 - Correct Group Services daemon problem” on page 450 .

If this problem persists, record all relevant information and contact the IBM Support Center.

Chapter 25. Diagnosing Event Management problems

This chapter discusses diagnostic procedures and failure responses for the Event Management (EM) component of RSCT. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 472. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 464.

Note that Event Management is a subsystem of RS/6000 Cluster Technology (RSCT).

Related documentation

The following publications provide information about the Event Management (EM) component of RSCT:

1. *PSSP: Administration Guide*

   ″The Event Management Subsystem″ contains information about the Event Management components, configuration, operation, and other components on which it depends.
2. *PSSP: Messages Reference*

   These chapters contain messages related to Event Management:
   - 2521 - Event Management Messages
   - 2522 - Resource Monitor Messages
   - 2525 - RS/6000 Cluster Technology Common Messages
3. *RSCT: Event Management Programming Guide and Reference*

   ″Understanding Event Management″ discusses the basic components of the Event Management subsystem. It is oriented to users of the Event Management API (EMAPI).

   ″EM Configuration Data Reference″ provides a discussion about the contents of the Event Management Configuration Database (EMCDB).
4. *The RS/6000 SP Inside Out*, SG24-5374

   ″RS/6000 Cluster Technology″ provides an overview of the Event Management subsystem, and in general of all RSCT subsystems.
5. *PSSP Version 3 Survival Guide*, SG24-5374

   ″RSCT Cluster Technology (RSCT)″ provides several hints and tips for dealing with problems with Event Management and the RSCT components in general.

Requisite function

This is a list of the software directly used by the EM component of RSCT. Problems within the requisite software may manifest themselves as error symptoms in Event Management. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the EM component of RSCT, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

1. Group Services (GS)

   Event Management uses the Group Services facility for peer coordination through the **ha_em_peers** group. Event Management also subscribes to the

HostMembership and **AdapterMembership** internal groups in Group Services. If Group Services is not running, Event Management hangs waiting for GS to resume.

2. Reliable Messages Passing Library (Topology Services)

   Event Management uses the Reliable Message Passing Library to send and receive messages between EM daemons. This library is based on the Network Connectivity Table (NCT) provided by Topology Services.

3. System Performance Measurement Interface (SPMI)

   | The RMAPI and the **haemaixos** subsystem (part of the EM subsystem) use the
   | SPMI library shipped with AIX as part of the **perfagent.tools** file set
   | (**/usr/lib/libSpmi.a**). The SPMI library is used to access AIX resource data.

4. System Data Repository (SDR)

   Event Management uses the SDR to store configuration information, such as the EMCDB (Event Management Configuration Database) file, the version number for the EMCDB, and the EM classes for resource variables and resource monitor definitions. It also uses the SDR for initialization, specifically to obtain the system partition name (domain). Event Management also uses the **switch_responds** class in the SDR to create the instances of the **IBM.PSSP.Response.Switch.state** resource variable.

5. SP System Security Services

   Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

6. **/var**

   Event Management uses the **/var** file system to store runtime and log information. Runtime information (along with any core dump) is located in **/var/ha/run/haem.***syspar* and log information is located in **/var/ha/log**. It also maintains UNIX domain sockets in the **/var/ha/soc** and **/var/ha/soc/haem** directories.

7. **/etc**

   Event Management stores configuration information in **/etc/ha/cfg**.

8. TCP/IP Sockets

   Event Management includes socket information in **/etc/services**. Event Management accepts two possible remote connections. The peers communication is done through a UDP port defined in the **Syspar_ports** SDR class. Event Management also supports remote client connections on the control workstation for inter-domain communication. It uses a TCP/IP port specified in the **SP_ports** SDR class.

Internal components of Event Management

Event Management uses the RMAPI (Resource Monitor API) to communicate with and control resource monitors (RMs).

Resource monitors

Resource monitors are of two types: external and internal.

External resource monitors

RSCT ships with seven external resource monitors. They supply data to the EM daemon. The seven external RMs are:

- **IBM.PSSP.hmrmd**

This monitor provides the state of the SP hardware. This information is obtained from the PSSP hardware monitoring subsystem (**hardmon**). The resource variables are of type state, and sent directly to the Event Management daemon as a message. The process name is **hmrmd**. It is started by the Event Management daemon. The executable is located in **/usr/sbin/rsct/bin/haemRM**.

- **IBM.PSSP.harmpd**

This monitor examines processes that are running a particular application. The resource variables can be used to determine whether or not a particular system daemon is running. The resource variables are of type state, and sent directly to the EM daemon as a message. The process name is **harmpd**. It is started by the Event Management daemon. The executable is located in **/usr/sbin/rsct/bin/haemRM**.

- **IBM.PSSP.harmld**

This monitor provides switch, IBM Virtual Shared Disk, LoadLeveler, processor on-line information, and internal variables. It uses shared memory, as it only reports on resource variables of type counter and quantity. The process name is **harmld**. It is started by the Event Management daemon. The executable is located in **/usr/sbin/rsct/bin/haemRM**.

- **IBM.PSSP.pmanrmd**

This monitor supplies the resource variables of the Problem Management subsystem (PMAN). On your behalf, **pmanrmd** can run a program, script, or command to report on various aspects of the system. The resource variables are of type state, and sent directly to the Event Management daemon as a message. The process name is **pmanrmd**. It is started by the AIX SRC. The subsystem name is **pmanrm.***syspar*. The Perl script is located in **/usr/lpp/ssp/bin/pmanrmd**.

- **aixos**

This monitor provides resource variables that represent AIX operating system resources. This is a daemon (**harmad**) with a connection type of server. The SPMI library provides the means to directly query AIX structures (such as the kernel and Logical Volume Manager) to supply data for a range of operating system resource variables. The aixos resource monitor calls the SPMI library (part of the **perfagent.tools** component). The process name is **harmad.** It is started by the AIX SRC. The subsystem name is **haemaixos.***syspar*. It is located in **/usr/sbin/rsct/bin/haemRM**.

- **IBM.PSSP.CSSLogMon**

This monitor supplies a resource variable that represents the state of CSS error log entries. This is a command-based resource monitor with a connection type of client.

- **IBM.PSSP.SDR**

This monitor provides a resource variable that represents the modification state of SDR classes. This is a command-based resource monitor with a connection type of client.

Internal resource monitors

The two internal RMs are:

- **Membership**

This monitor supplies the resource variables that correspond to node and network adapter state. The information is obtained directly from Group Services by subscribing to the system groups **hostMembership**, **enMembership**, and **cssMembership**.

- **Response**

This monitor supplies the **IBM.PSSP.Response** resource variables. The **IBM.PSSP.Response.Host.state** resource variable is updated based on information coming from the adapter membership information (**en0** adapter in particular) supplied by Group Services. The **IBM.PSSP.Response.Switch.state** resource variable is updated based on the **switch_responds** class in the SDR. This SDR class is updated by the switch daemon itself.

The Host_Responds daemon (**hrd**) obtains the Ethernet state information by subscribing to the **IBM.PSSP.LANAdapter.state** resource variable through the EMAPI. It receives events from the **Membership** monitor.

Service information

Service information is obtained automatically by the **phoenix.snap** tool, or it can be collected manually.

Automatic method - phoenix.snap

This tool collects data for reporting RSCT-related problems. In particular, it always collects data for Topology Services and Group Services, while other subsystems are optional. The full name is: **/usr/sbin/rsct/bin/phoenix.snap**.

The **phoenix.snap** tool can be run from the control workstation or any node. The script collects and generates a single file in **tar** format with the data from the control workstation and the selected nodes. For example, to collect data for Event Management, issue this command on the control workstation:

```
phoenix.snap -w HAEM -d /snap -l sp6n01,sp6n02
```

This tool collects information for Topology Services, Group Services, Event Management, and general configuration information for the control workstation and nodes 1 and 2 (listed by their hostname). It places the single ″snap″ file in **/snap** in **tar** format in a file named **all.***mmddhhmm***.tar**. The tool also generates a file containing errors from the script. This file is located in the same directory, and it is named **phoenix.snap_err.***mmddhhmm***.out**. For both files, *mmddhhmm* represents the date (month and day only) and time (hour and minute only) that the information was collected.

Note: The **phoenix.snap** tool is a service tool and not a PSSP command. The tool is shipped with PSSP 3.2 as is, without documentation. For assistance on using **phoenix.snap** in a manner other than what is described in this section, contact the IBM Support Center.

Manual method of data collection

This is a list of the information needed to diagnose a problem with Event Management.

SDR data:

- Output of the command: **splstdata -n**

- Output of the command: **SDRGetObjects Syspar_ports**
- Output of the command: **SDRGetObjects SP_ports**
- Output of the command: **splstdata -a**
- Output of the command: **SDRGetObjects TS_Config**
- Output of the command: **SDRGetObjects host_responds**
- File **/spdata/sys1/sdr/partitions/***syspar***/hats.machines_lst**
- File **/spdata/sys1/sdr/partitions/***syspar***/hats.machines.inst**

Files from failing nodes and the control workstation:

- **/etc/hosts**
- **/etc/resolv.conf**
- **/etc/netrvc.conf**
- **/etc/services**
- **/etc/SDR_dest_info**
- **/var/ha/run** directory (All entries except **pman**)
- **/var/ha/logs** (ten most recent files)
- **/var/ha/run/hats.***syspar***/hats.machines.inst**
- **/var/ha/run/hats.***syspar***/hats.machines.lst**

Output of these commands from failing nodes and the control workstation:

- **ypwhich**
- **ps -edf**
- **LANG=C errpt -a**
- **df -k**
- **hr query** (control workstation only)
- **echo $NSorder**
- **lslpp -L**
- **lssrc -a**

Output of these commands, which consists of network data from failing nodes and the control workstation:

- **echo $NSorder**
- **no -a**
- **netstat -m**
- **netstat -in**
- **netstat -rn**
- **netstat -D**
- **entstat en***
- **ifconfig** (on all adapters in the Topology Services groups)

Output of these commands, which consists of component-specific data from failing nodes and the control workstation:

- **lssrc -ls hats,hags,haem**
- **/usr/sbin/rsct/bin/hagsgr -s hags**
- **/usr/sbin/rsct/bin/hagsvote -l -a**
- **/usr/sbin/rsct/bin/hagspbs -s hags**
- **/usr/sbin/rsct/bin/hagsns -s hags**

- **/usr/sbin/rsct/bin/hagsmg -s hags**
- **/usr/sbin/rsct/bin/hagscl -ls hags**
- **ipcs -m**

Error information

The Event Management subsystem uses the AIX Error Log as the main repository for errors and informational messages. Besides the logging of informational messages and errors, the Event management daemon can log additional and detailed information if tracing is activated.

AIX Error Log for Event Management

Event Management does not use the AIX Error Log in the same way as Topology Services or Group Services. In PSSP 3.2, Event Management uses only two AIX error log templates. The important information from these entries consists of informational and error message placed in the Detail Data field. All error messages have numbers and are documented in *PSSP: Messages Reference*.

There are two types of error messages that Event Management logs in the AIX Error log:

1. HA001_TR - for informational messages
2. HA002_ER - for error messages

See the Detail Data field for each entry.

Error log files

The log file is located in the **/var/ha/log** directory. The file is named **em.default.***syspar*. It contains any error message from the Event Management daemon that cannot be written to the AIX error log. Normally, all the daemon error messages are written to the AIX error log. This log also contains error messages that result from repetitive operational errors. For example, errors that are logged when the Event Management daemon cannot connect to Group Services, and it retries every five seconds, or errors logged when it tries to join the **ha_em_peers** group every 15 seconds.

The size of the **em.default.***syspar* file is examined every two minutes. If the size exceeds 256KB, the file is renamed with a suffix of **.last**, and a new default file is created. No more than two copies of this file are kept.

If Event Management cannot start a resource monitor, it also records additional information in the **em.defaults.***syspar* log file. The error information includes the name of the resource monitor that could not be started.

Event Management daemon errors

The Event Management daemon errors are categorized here:

Initialization errors

- Not running as **root**.
- Cannot get group attribute for the **haemrm** group.
- Cannot set UID.
- **Malloc** failed.

Event Management Configuration Database (EMCDB) operations

- Cannot open, read, write or checksum the EMCDB file.

Signals

- Cannot ignore or set **SIGPIPE**, **SIGALRM**, **SIGCHLD**
- **sigthreadmask()** failed.

Sockets

- Cannot open, read, or write sockets (UDP and TCP/IP)

Register and unregister Events (EMAPI)

- Missing information, for example: **node_number**
- Syntax error, for example, and error in an instance vector or expression

Environment setting

- Cannot determine environment (SP or HACMP node)
- Incorrect or missing **node_number**

Group Services

- Connection to Group Services failed.
- Cannot join peers group.
- Error reading group state value.

Reliable Messages Library (PRM)

- Cannot initialize PRM services.
- Cannot set PRM limits.
- Cannot send or receive messages.

AIX SRC subsystem

- Event Management daemon not started by the AIX SRC

Event Management daemon

- Cannot create or access runtime directory.
- Cannot create lock file in runtime directory.

Resource Monitor operations

- Errors communicating with resource monitors.

SP System Security Services

- Cannot load security library
- Error from security routine (not equal to **SPSEC_SUCCESS**)

System Performance Measurement Interface (SPMI)

- Cannot get statistics from SPMI.

EMCDB problems

EMCDB Version is incorrect

This problem may happen when the system administrator re-creates the Event Management subsystem by using the **haemctrl** command on the control workstation. It can also happen if the system administrator creates a new EMCDB file by using the **haemcfg** command. Every time a new EMCDB is created (by using the **haemcfg** command), its version number is stored in the **Syspar** SDR class. Daemons use that version number and file only if the state value of the **ha_em_peers** group in Group Services contains the same value.

The file is transferred from the control workstation to the nodes using a remote copy command, and it is stored in **/etc/ha/cfg**. If the file does not exist, an error message is logged in **/etc/ha/log/em.default.***syspar* and in the AIX error log.

Resource Monitor problems

Problems with resource monitors are usually communication problems. One way of verifying that the RMs are connected to and communicating with the Event Management daemon is to issue the command:

```
lssrc -ls haem.syspar
```

and check the `Resource Monitor` section. The output is similar to:

```
Resource Monitor Information
        Name          Inst     Type      FD    SHMID    PID   Locked
IBM.PSSP.CSSLogMon       0        C       -1     -1      -2  00/00  No
IBM.PSSP.SDR             0        C       -1     -1      -2  00/00  No
IBM.PSSP.harmld          0        S       20     11   28954  01/01  No
IBM.PSSP.harmpd          0        S       19     -1   28684  01/01  No
IBM.PSSP.hmrmd           0        S       21     -1   21766  01/01  No
IBM.PSSP.pmanrmd         0        C       14     -1      -2  00/00  No
Membership               0        I       -1     -1      -2  00/00  No
Response                 0        I       -1     -1      -2  00/00  No
aixos                    0        S       12     10      -2  00/01  No
```

The connection type specifies how the resource monitor connects to Event Management:

- Type **server** (S) corresponds to external daemons, and their PID is in the PID column.
- Type **client** (C) are usually scripts or commands that run and send updates to the EM regarding resource variables.
- Type **internal** (I) corresponds to resource monitors internal to the Event Management daemon.

The last two columns of the output named `Locked`, represent counters for successful connections to the resource monitor. The Event Management daemon maintains two counters: one for start attempts and one for successful connections. If either of these counters reaches the start limit or connect limit respectively, the RM is locked.

The counters are cleared two hours after the first start or connect. For starts, the limit is three. For connects, the limits is the number of instances configured for the resource monitor (**rmNum_instances** in the **EM_Resource_Monitor** class) multiplied by three. For all resource monitors shipped with PSSP, **rmNUM_instances** is one.

Once the Event Management daemon has successfully connected to a resource monitor of type server, the daemon attempts to reconnect to the resource monitor if it should terminate. The reconnection is attempted at the rate of one per minute. However, reconnection attempts are limited under the following circumstances:

1. If it is necessary that the daemon start the resource monitor before each reconnection attempt. After three attempts within two hours, the resource monitor is locked, and no further attempts are made.

2. If the resource monitor cannot be started by the Event Management daemon after three unsuccessful reconnections within two hours, the resource monitor is locked. No further reconnection attempts are made.

The reason for locking the resource monitor is that if it cannot be started and remain running, or successful connections are frequently being lost, a problem exists with the resource monitor. Once you isolate and correct the problem, unlock and start the resource monitor by issuing the **haemunlkrm** command. This command resets the start and connect counters to zero and also resets the two hour window.

Note: Locking does not apply to **client** type resource monitors.

Dump information

Dump information from Event Management Resource Monitor daemons

If the Event Management daemon detects an error in the shared memory segment used by the daemon, or a resource monitor instance, it creates a dump file in the **/var/ha/run/haem.***syspar* directory. This dump contains the first 4KB of the shared memory segment. The dump is called **rzdump.RM***rmname.rminst.time*, where *rmname.rminst* is the resource monitor name and instance, and *time* is a timestamp.

The Event Management daemon (**haemd**) generates a full core dump in most of the error conditions listed in “Event Management daemon errors” on page 460.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Trace facility built into Event Management

The tracing function can be activated by using the **haemctrl** command. It supplies additional problem determination information when it is requested by the IBM Support Center. Normally, tracing should not be turned on, because it may degrade Event Management subsystem performance and can consume large amounts of disk space in the **/var** file system.

The trace files are located in the **/var/ha/log** directory. The following is a brief description of each file:

- **em.trace.***syspar* contains trace output from the Event Management daemon.

- **em.msgtrace.***syspar* contains message trace output from the Event Management daemon.

Information to collect before contacting the IBM Support Center

1. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry ″ts_auth_methods″ lists the authentication methods in use.
2. See “Manual method of data collection” on page 458 for the remaining list of item to collect.

Diagnostic instructions

Verify SP software installation

RSCT does not provide software verification. The following steps verify that the software and components are installed and defined:

1. Issue the command:

   ```
   lslpp -l rsct.*
   ```

   and verify that the RSCT components are installed. The output of this command is similar to:

   ```
     Fileset                   Level  State      Description
     ----------------------------------------------------------------------------
            Path: /usr/lib/objrepos
     rsct.basic.hacmp        1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    basic function (HACMP domains)
     rsct.basic.rte          1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    basic function (all domains)
     rsct.basic.sp           1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    basic function (SP domains)
     rsct.clients.hacmp      1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    client function (HACMP
                                                   domains)
     rsct.clients.perl5      1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    Perl5 Package
     rsct.clients.rte        1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                   client function (all domains)
     rsct.clients.sp         1.2.0.0  COMMITTED  RS/6000 Cluster Technology
                                                    client function (SP domains)
   ```
2. Issue the command:

   ```
   lssrc -s haem.syspar
   ```

   and verify that the Event Management subsystem is defined. If it is not defined, you may have to use the **syspar_ctrl** command to recreate the RSCT subsystems. See “Recover crashed node (Event Management Resource Monitor daemons)” on page 472.
3. Issue the command:

   ```
   lssrc -ls haem.syspar
   ```

and verify that the Event Management daemon is running. The output of this command is similar to:

```
Subsystem         Group            PID     Status
 haem.c166s       haem             30448   active

No trace flags are set

Configuration Data Base version from SDR:
        931790799,481121271,0

Daemon started on Tuesday 08/10/99 at 08:36:37
Daemon has been running 2 days, 7 hours, 9 minutes and 9 seconds
Daemon connected to group services: Yes
Daemon has joined peer group:       Yes
Daemon communications enabled:      Yes
Daemon security:                    Compatibility
Peer count:                         6

Peer group state:
        931790799,481121271,0
        NOSEC

Logical Connection Information for Local Clients
    LCID          FD           PID     Start Time
       0          11          17302    Tuesday 08/10/99 08:38:38
       1          13          23740    Tuesday 08/10/99 08:38:40
      10          16          33032    Tuesday 08/10/99 08:38:59
      11          17          33032    Tuesday 08/10/99 08:38:59

Logical Connection Information for Remote Clients
    LCID          FD           PID     Start Time

Logical Connection Information for Peers
    LCID         Node
       2            7
       4            3
       5            6
      20            8
      27            1
      31            5
Resource Monitor Information
         Name         Inst     Type      FD     SHMID    PID    Locked
IBM.PSSP.CSSLogMon      0        C        -1      -1      -2  00/00  No
IBM.PSSP.SDR            0        C        -1      -1      -2  00/00  No
IBM.PSSP.harmld         0        S        20      11   28954  01/01  No
IBM.PSSP.harmpd         0        S        19      -1   28684  01/01  No
IBM.PSSP.hmrmd          0        S        21      -1   21766  01/01  No
IBM.PSSP.pmanrmd        0        C        14      -1      -2  00/00  No
Membership              0        I        -1      -1      -2  00/00  No
Response                0        I        -1      -1      -2  00/00  No
aixos                   0        S        12      10      -2  00/01  No

Highest file descriptor in use is 21

Peer Daemon Status
   0 S S      1 I A      3 I A      5 I A      6 I A      7 I A
   8 I A     17 O A     21 O A


Internal Daemon Counters
    GS init attempts =         22  GS join attempts =          1
    GS resp callback =       1653  CCI conn rejects =          0
    RMC conn rejects =          0  HR conn rejects  =          0
    Retry req msg    =          0  Retry rsp msg    =          0
    Intervl usr util =          1  Total usr util   =       2107
    Intervl sys util =          2  Total sys util   =        2006
    Intervl time     =      12000  Total time       =    19847228
    lccb's created   =         33  lccb's freed     =          23
    Reg rcb's creatd =         41  Reg rcb's freed  =          30
    Qry rcb's creatd =        332  Qry rcb's freed  =         332
    vrr created      =         41  vrr freed        =          30
    vqr created      =      42147  vqr freed        =       42147
    var inst created =        939  var inst freed   =           0
    Events regstrd   =         41  Events unregstrd =          30
    Insts assigned   =         43  Insts unassigned =          19
    Smem vars obsrv  =       3306  State vars obsrv =      195596
    Preds evaluated  =      33132  Events generated =         232
    Smem lck intrvl  =          0  Smem lck total   =           0
    PRM msgs to all  =          8  PRM msgs to peer =           8
    PRM resp msgs    =         86  PRM msgs rcvd    =          32
    PRM_NODATA       =        126  PRM_BADMSG errs  =           0
    Sched q elements =         32  Free q elements  =          30
    xcb alloc'd      =       1303  xcb freed        =        1302
    xcb freed msgfp  =         16  xcb freed reqp   =           1
    xcb freed reqn   =          8  xcb freed rspc   =         678
    xcb freed rspp   =         86  xcb freed cmdrm  =         513
    xcb freed unkwn  =          0  Sec enable       =           0
    Sec disable      =          0  Sec authent      =           0
    Wake sec thread  =          0  Wake main thread =           0
    Missed sec rsps  =          0  Enq sec request  =           0
    Deq sec request  =          0  Enq sec response =           0
    Deq sec response =          0

Daemon Resource Utilization Last Interval
User:                 0.010 seconds    0.008%
System:               0.020 seconds    0.017%
User+System:          0.030 seconds    0.025%

Daemon Resource Utilization Total
User:                21.070 seconds    0.011%
System:              20.060 seconds    0.010%
User+System:         41.130 seconds    0.021%

Data segment size:  2132K
```

The first portion of the output gives a good deal of information. The first lines correspond to the output obtained without the **-l** flag. It tells you if the subsystem is active or inoperative. Also, there is the status of the trace flag, which is `off` in this case.

The next line shows the EMCDB version number stored in the SDR.

```
Configuration Data Base version from SDR:
        931790799,481121271,0
```

The line under:

```
Peer group state:         931790799,481121271,0         NOSEC
```

shows the version number used by the **ha_em_peers** group. In this case, both are `931790799,481121271,0`. Compare these values to see if the daemons are using an EMCDB version different that the one stored in the SDR.

This usually happens when you run the **haemcfg** command that creates a new EMCDB, but you have not restarted the Event Management daemons.

Remember that all the Event Management daemons in the system partition (domain) need to be stopped, and the **ha_em_peers** group dissolved, in order to use the new EMCDB.

The output also tells you how long the daemon has been running:

```
Daemon started on Tuesday 08/10/99 at 08:36:37
Daemon has been running 2 days, 7 hours, 9 minutes and 9 seconds
```

If the daemon was able to connect to Group Services:

```
Daemon connected to group services: Yes
```

If the daemon was able to join the **ha_em_peers** group:

```
Daemon has joined peer group:       Yes
```

The following line states if the daemon has enabled communication with clients:

```
Daemon communications enabled:      Yes
```

Since RSCT 1.2 supports DCE security, the next line states which security mode this daemon is working with:

```
Daemon security:                    Compatibility
```

The last line of this stanza gives you the number of providers in the **ha_em_peers** group, not counting this one. This is the number of Event Management daemons running. If all nodes are up and running, this number should be equal to the number of nodes, plus one for the control workstation, and one for the current node.

```
Peer count:                         6
```

Identify the failing node

Follow these steps to identify the failing node:

1. Issue the command:

   ```
   lssrc -ls haem.syspar
   ```

   on the control workstation. Check the number of `Peer count` explained in "Verify SP software installation" on page 464. If the number of `Peer count` is less than the number of nodes (including the control workstation) plus one, you have one or more failing nodes.

2. Issue the command:

   ```
   lssrc -ls hags.syspar
   ```

   on the control workstation and verify that the **ha_em_peers** group appears on the list of local groups, as shown here:

   ```
   lssrc -ls hags.c184s
   Subsystem         Group            PID     Status
   hags.c184s       hags             22446   active
   3 locally-connected clients.  Their PIDs:
   17804 27348 29440
   ```

```
HA Group Services domain information:
Domain established by node 0.
Number of groups known locally: 3
                   Number of   Number of local
Group name         providers   providers/subscribers
cssMembership            5           0           1
ha_em_peers              7           1           0
ha.vsd                   5           1           0
```

If the **ha_em_peers** group does not appear in the output, the Event Management daemon is not running on the control workstation, or it has not been able to join the **ha_em_peers** group. If the daemon is not running, follow the steps given in “Recover crashed node (Event Management Resource Monitor daemons)” on page 472.

If the daemon is not running on the control workstation (or the node where you are running the commands), the number of local providers for the **ha_em_peers** group will be zero.

3. Issue the following command to identify which node is failing:

```
/usr/sbin/rsct/bin/hagsgr -s hags.c184s -a ha_em_peers
```

The output is similar to the following:

```
  Number of: groups: 6
Group name[ha_em_peers] group state[Inserted |Idle |]
Providers[[1/0][1/65][1/5][1/9][1/13][1/1][1/17]]
Local subscribers[]
```

Note: This command is undocumented and not supported, but it is shipped with RSCT, and therefore available on any system running RSCT.

From the output of this command you can see the list of providers (members of the **ha_em_peers** group). They are listed in the order they joined the group, with the first in the list called the Group Owner. The list contains the members brackets [X/Y] where **X** is the instance number of the daemon, and **Y** is the node number.

From this list you can identify the failing node.

Verify event registration

There are three optional arguments that can be passed to the daemon by using the **haemtrcon** command. The syntax for the **haemtrcon** command is as follows:

```
haemtrcon [-h host] [-a argument] -g group_name
haemtrcon [-h host] [-a argument] -s subsystem_name
haemtrcon [-h host] [-a argument] -p subsystem_pid
```

The arguments accepted by the daemons for dumping information are shown in Table 62.

Table 62. Arguments for Event Management daemon

Argument	Description
regs	Dumps registered events.
dinsts	Dumps registered instances.
olists	Dumps observation lists.

To see all the events registered with Event Management, you could use this command:

```
haemtrcon -a regs -s haem.sp5en0
```

Output is similar to the following:

```
haemtrcon: the specified trace flags have been set (00000000)
```

This command sends a request to the Event Management daemon to dump all registered events. The daemon will then dump the requested information into the **em.trace** file in the **/var/ha/log** directory. The content of the file looks like this:

```
Trace Started at 11/30/99 14:38:01.509522432
Registered Events:

0 0x00000000 ( 0, 0) 0 IBM.PSSP.Membership.LANAdapter.state "X==0" "X==1"

AdapterNum=0 AdapterType=en NodeNum=0

AdapterNum=0 AdapterType=en NodeNum=5

AdapterNum=0 AdapterType=en NodeNum=9

AdapterNum=0 AdapterType=en NodeNum=13

AdapterNum=0 AdapterType=en NodeNum=1

1 0x00000000 ( 0, 0) 2 IBM.PSSP.Response.Host.state "X==1 && X@P==0" ""

NodeNum=5

NodeNum=9

NodeNum=13

NodeNum=1

2 0x00010000 (1,0) 2 IBM.PSSP.Membership.LANAdapter.state "X==0 && X@P==1" ""

AdapterNum=0 AdapterType=css NodeNum=13

AdapterNum=0 AdapterType=css NodeNum=9
AdapterNum=0 AdapterType=css NodeNum=5

AdapterNum=0 AdapterType=css NodeNum=1

3 0x00000000 ( 0, 0) 3 IBM.PSSP.CSSlog.errlog "X@1 != 0" ""

No instances currently assigned

4 0x00000000 ( 0, 0) 1 IBM.PSSP.pm.User_state1 "X@0!=X@P0" ""

No instances currently assigned

9 0x00000000 ( 0, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned

10 0x00010000 ( 1, 0) 18 IBM.PSSP.SDR.modification "" ""

Class=EM_Condition

14 0x00050000 ( 5, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned
```

```
15 0x00060000 ( 6, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned

16 0x00070000 ( 7, 0) 18 IBM.PSSP.SDR.modification "" ""

Class=Node

17 0x00080000 ( 8, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned
```

As you can see, there are several events registered with Event Management, but only a few of them have instances currently assigned. An event with no instances assigned is an event known to Event Management, but not currently active. The first three event are the only ones active. We can see that Event Management is monitoring the Ethernet (en) and SP Switch (css) adapters through the Membership resource variable. Also, it is monitoring the Response variable for host in all the nodes.

Let us activate a file system monitor and dump this information again. The result is as follows:

```
Trace Started at 11/30/99 14:51:02.012841216

 Registered Events:

0 0x00000000 ( 0, 0) 0 IBM.PSSP.Membership.LANAdapter.state "X==0" "X==1"

AdapterNum=0 AdapterType=en NodeNum=0

AdapterNum=0 AdapterType=en NodeNum=5

AdapterNum=0 AdapterType=en NodeNum=9

AdapterNum=0 AdapterType=en NodeNum=13

AdapterNum=0 AdapterType=en NodeNum=1

1 0x00000000 ( 0, 0) 2 IBM.PSSP.Response.Host.state "X==1 && X@P==0" ""

NodeNum=5

NodeNum=9

NodeNum=13

NodeNum=1
2 0x00010000 ( 1, 0) 2 IBM.PSSP.Membership.LANAdapter.state "X==0 && X@P==1" ""

AdapterNum=0 AdapterType=css NodeNum=13

AdapterNum=0 AdapterType=css NodeNum=9

AdapterNum=0 AdapterType=css NodeNum=5

AdapterNum=0 AdapterType=css NodeNum=1

3 0x00000000 ( 0, 0) 3 IBM.PSSP.CSSlog.errlog "X@1 != 0" ""

No instances currently assigned

4 0x00000000 ( 0, 0) 1 IBM.PSSP.pm.User_state1 "X@0!=X@P0" ""
```

```
No instances currently assigned

9 0x00000000 ( 0, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned

10 0x00010000 ( 1, 0) 18 IBM.PSSP.SDR.modification "" ""

Class=EM_Condition

14 0x00050000 ( 5, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned
15 0x00060000 ( 6, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned

16 0x00070000 ( 7, 0) 18 IBM.PSSP.SDR.modification "" ""

Class=Node

17 0x00080000 ( 8, 0) 18 IBM.PSSP.SDR.modification "" ""

No instances currently assigned

18 0x00090000 ( 9, 0) 18 IBM.PSSP.aixos.FS.%totused "X>90" "X<60"

VG=rootvg LV=hd3
```

As you can see from this new output, there is a new event (18) which is the one we have just activated. If you want more information about this event, you can use one of the other two arguments described in Table 62 on page 468. For example, the **olists** argument gives details on the events registered for this monitor, as follows:

```
Trace Started at 11/30/99 14:56:45.161388544



Obsv control = 0x200605c8, interval = 60.000000, flags = 0x0000,
                     last obsv = 943991746 874736

Obsv list = 0x2002b8e8, delay = 20.000000, number of ptr lists elements = 1

limit = 10000, inst count = 16

Normal list:


IBM.PSSP.aixos.FS.%totused

vector: VG=rootvg LV=hd3

API instance ID = 2, RM instance ID = 18465, RM instance number = 0

current value: 8.072917, raw: 8.072917

flags: 0003 qcnt: 0



Immediate list:
```

From this output, you can see that the sample interval for this variable is 60 seconds, and that the current value is a little bit over 8%. From the instance or vector, we can tell that this is a monitor of the **/tmp** file system (hd3). The previous output gave us the condition (X>90), and rearm condition (X<60). This tracing or dump facility from Event Management is helpful in situations where events registered through either the SP Event Perspective, Problem Management or the EMAPI directly, do not appear to be working.

Verify Resource Monitors

See “Resource Monitor problems” on page 462.

Error symptoms, responses, and recoveries

Recover crashed node (Event Management Resource Monitor daemons)

The Event Management subsystem, including resource monitors can be re-created by using the **syspar_ctrl** command as follows:

1. Remove the subsystem:

   ```
   syspar_ctrl -D haem
   ```
2. Re-create the subsystem:

   ```
   syspar_ctrl -A haem
   ```

   This command creates the Event Management subsystem and starts the daemons.

Recover EMCDB

The EMCDB can be recreated by using the **haemcfg** command on the control workstation. For the Event Management daemons to use this new EMCDB, they must be stopped. Use the following procedure:

1. Run the **haemcfg** command (no arguments)
2. Stop the Event Management daemon on the control workstation and on all the nodes within the partition. Issue these commands:
   - `stopsrc -s haem.`*syspar* on the control workstation.
   - `stopsrc -s haem` on the nodes.
3. Start the Event Management daemon on the control workstation and on all the nodes within the partition. Issue these commands:
   - `startsrc -s haem.`*syspar* on the control workstation.
   - `startsrc -s haem` on the nodes.

Security errors

The EMAPI returns errors if the client cannot be authenticated or authorized with the EM daemon. Error messages (with message numbers) are in the error return information. The daemon also logs errors (in the AIX error log) if it has problems accessing the security routines.

Chapter 26. Diagnosing IBM Virtual Shared Disk problems

This chapter discusses diagnostic procedures and failure responses for the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk components of PSSP. The list of known error symptoms and the associated responses are in the section "Error symptoms, responses, and recoveries" on page 487. A list of the information to collect before contacting the IBM Support Center is in the section "Information to collect before contacting the IBM Support Center" on page 479.

Enhanced Security Option

PSSP 3.4 provides the option of running your RS/6000 SP system with an enhanced level of security. This function removes the dependency PSSP has to internally issue **rsh** and **rcp** commands as a **root** user from a node. When this function, called Restricted Root Access (RRA), is enabled, PSSP does not automatically grant authorization for a **root** user to issue **rsh** and **rcp** commands from a node. If you enable this option, some PSSP components, such as IBM Virtual Shared Disk software, may not work as documented.

Related documentation

The following publications provide information about the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk components of PSSP:

1. *PSSP: Planning Volume 2*

   The chapter ″Planning for Virtual Shared Disks″ explains how to plan for the software aspects of the IBM Virtual Shared Disk facility.
2. *PSSP: Managing Shared Disks*

   This book explains how to install, manage and use the disk management facilities of PSSP. These facilities are: the IBM Virtual Shared Disk component, the Data Striping Device (HSD) component, and the IBM Recoverable Virtual Shared Disk component.
3. *PSSP: Command and Technical Reference*

   This books provides detailed syntax and parameter information for the commands used to monitor and control the IBM Virtual Shared Disk facility.
4. *PSSP: Messages Reference*

   These chapters contain messages related to IBM Virtual Shared Disk software:
   - 0021 - IBM Virtual Shared Disk Common Messages
   - 0034 - IBM Virtual Shared Disk Common Messages
   - 2500 - Hashed Shared Disk (HSD) Messages
   - 2506 - IBM Recoverable Virtual Shared Disk Messages

The IBM Virtual Shared Disk Perspective is a graphical user interface that lets you perform IBM Virtual Shared Disk tasks. It also lets you view and change IBM Recoverable Virtual Shared Disk options. This Perspective is started using the command:

```
/usr/lpp/ssp/bin/spvs
```

To see an online help system that explains the IBM Virtual Shared Disk Perspective, click on **Help** → **Tasks** at the top right-hand corner of the main panel.

Requisite function

This is a list of the software directly used by the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk functions of PSSP. Problems within the requisite software may manifest themselves as error symptoms in this software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk functions, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- AIX Logical Volume Manager
- Group Services component of PSSP
- SP Switch (CSS Subsystem) of PSSP
- Sysctl Subsystem of PSSP
- System Data Repository (SDR) component of PSSP
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Before you begin

Before starting to debug the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk subsystems, consider the following tasks that will make the job easier:

1. Set up a working collective of the virtual shared disk nodes. All of the diagnostic procedures that use the **dsh** command in the remainder of this chapter assume that you have done this. For instructions on how to set up the working collective, see the **dsh** command description in *PSSP: Command and Technical Reference*.
2. If you add the directory **/usr/lpp/csd/bin** to your path you will not have to use the fully qualified path name when issuing IBM Virtual Shared Disk commands. To do this, issue this command:

   ```
   export PATH=$PATH:/usr/lpp/csd/bin
   ```

   Note: The path must be fully qualified when using the **dsh** command, or the **DSHPATH** environment variable can be set.

Error information

Errors logged by the IBM Virtual Shared Disk device driver and IBM Recoverable Virtual Shared Disk subsystem

These errors identify which function has failed, along with the errno (error number). These errors fall into the following categories:

1. Configuration errors, which are generated by the client node and reported on the client node.
2. **read/write** errors:
   - Generated by the client node and reported on the client node.
   - Generated by the server node and reported on the server node.
   - Generated by the server node and reported on both the server and client nodes.

3. **ioctl** errors, which are generated by the client node and reported on the client node.

These errors are recorded in the AIX error log. The **Type** field identifies the error as either PERM (permanent), INFO (informational) or UNKN (unknown).

Table 63. IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk Error Codes

Label	Error ID	Type	Description
RVSDD_A_ER	E0CEAF98	PERM	**Explanation:** The RVSD daemon asserted. **Details:** The Detail Data field indicates the reason for the assert. The RVSD daemon will respawn.
RVSDD_ER	692D4985	PERM	**Explanation:** An RVSD subsystem failure detected - Exiting. **Details:** The Detail Data indicates the cause of the error.
VSD_ADAPTYPE_ER	8943D0AC	PERM	**Explanation:** The RVSD subsystem detected different types of communication adapters. **Details:** During startup, the RVSD subsystem detected that IBM Virtual Shared Disk nodes were defined with different types of communication adapters.
VSD_BKLEVELNODES_ST	2EC9F1DD	INFO	**Explanation:** The RVSD subsystem detected back-level nodes. **Details:** During startup, the RVSD subsystem detected nodes that had back-level software installed. See the **rvsdrestrict** command for more information.
VSD_COPYIN_ER	21DF8EA6	PERM	**Explanation:** An IBM Virtual Shared Disk ioctl **copyin()** failure was detected. **Details:** The **copyin()** system call failed. This typically indicates an internal device driver error.
VSD_COPYOUT_ER	E45D5194	PERM	**Explanation:** An IBM Virtual Shared Disk ioctl **copyout()** failure was detected. **Details:** The **copyout()** system call failed. This may indicate an internal device driver error.
VSD_EIO_ER	32DEFEB6	PERM	**Explanation:** An EIO error was detected. **Details:** The device driver received an EIO error on an I/O request.
VSD_EXT_ER	C0292121	PERM	**Explanation:** An external function call failed. **Details:** The device driver received an error while trying to process a command. This may indicate an internal device driver error.
VSD_FREEMEM_ER	673B210E	PERM	**Explanation:** A failure was detected while attempting to free memory. **Details:** The device driver detected an error while freeing memory. This may indicate an internal device driver error.

Table 63. IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk Error Codes (continued)

Label	Error ID	Type	Description
VSD_HC_A_ER	E3747DF7	PERM	**Explanation:** The HC daemon asserted. **Details:** The Detail Data indicates the reason for the assert.
VSD_HC_ER	F9EE499E	PERM	**Explanation:** An HC subsystem failure was detected. **Details:** The Detail Data indicates the reason for the error.
VSD_INT_ER	2C617B9F	PERM	**Explanation:** An internal error was detected. **Details:** The Detail Data indicates the reason for the error.
VSD_IPPROTO_ER	6DE5541A	PERM	**Explanation:** An IP protocol error was detected. **Details:** The Detail Data indicates the reason for the error.
VSD_IPPROTO_ST	CD11733B	INFO	**Explanation:** This is an IP protocol informational message. **Details:** The Detail Data indicates the message.
VSD_RESTRICTFUNC_ST	1A565093	INFO	**Explanation:** the **rvsdrestrict** command reduced function. **Details:** This messages is logged when the RVSD subsystem is being started and the **rvsdrestrict** command has explicitly set the function level at or below the version of the RVSD subsystem that is installed on this node.
VSD_RESTRICTHIGH_ST	F34DF5D1	INFO	**Explanation:** Back-level nodes were detected. **Details:** This messages is logged when the RVSD subsystem is being started and back-level nodes were detected and the **rvsdrestrict** command has not been set to reduce the function level. Back-level nodes are defined as: 1. Nodes discovered running levels of PSSP earlier than PSSP 3.2. 2. RVSD is not installed on all the IBM Virtual Shared Disk nodes. 3. The node has an earlier level of RVSD installed than what was specified using the **rvsdrestrict** command.
VSD_RESUME_ER	7005E98D	PERM	**Explanation:** A resume failure was detected. **Details:** A virtual shared disk was being internally resumed and detected a failure because the virtual shared disk is no longer in ACTIVE state.

Table 63. IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk Error Codes (continued)

Label	Error ID	Type	Description
VSD_STOPVSD_ST	1969E9BD	INFO	**Explanation:** EIO recovery is stopping IBM Virtual Shared Disks. **Details:** The device driver detected an EIO error and the RVSD subsystem was attempting to switch ownership of the volume group to the backup node, but experiencing one of these problems: 1. The backup node may be down. 2. There may not be a backup node. 3. EIO errors are occurring too quickly on both the backup and primary nodes. 4. EIO recovery has been turned off. In all these cases, all the virtual shared disks will be moved to STP state on the volume group that received the EIO error.
VSD_SUSPEND_ER	631B153D	UNKN	**Explanation:** A suspend failure was detected. **Details:** An error has been discovered in an internal counter. The counter will be reset to a good value.
VSD_TIMESTAMPS_ST	937F798C	INFO	**Explanation:** Volume group timestamps are out of sync. **Details:** The RVSD subsystem has discovered that changes have been made to a volume group on one node, that have not been communicated to the node that is taking over ownership of the volume group. An attempt will be made to import the volume group to learn of the changes.
VSD_VARYONVG_ER	D628A365	PERM	**Explanation: varyonvg** function failed - Exiting. **Details:** The RVSD subsystem is trying to vary on a volume group to this node, but the **varyonvg** command has failed.
VSD_XMALLOC_ER	207133A8	PERM	**Explanation:** Memory allocation failure - Exiting **Details:** The device driver is trying to allocate memory and has received a failure. This may indicate that the AIX kernel is out of pinned memory.

Dump information

The **vsd.snap** script collects all information necessary to report problems related to the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk components of PSSP.

Note: If the PSSP system has crashed, obtain a system dump and accompanying unix files. See “Chapter 5. Producing a system dump” on page 81.

The **vsd.snap** script collects general system environment information, IBM Virtual Shared Disk configuration information, and various trace logs. The **vsd.snap** script collects information from an individual node. The information may be placed in a location specified by the user. The default location is: **/tmp/vsd.snapOut**. This information is collected only when the user explicitly invokes the **vsd.snap** script.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Internal virtual shared disk device driver circular trace buffer

This trace is intended for IBM Service Personnel only. It is not intended for general user tracing. The virtual shared disk device driver maintains a circular trace buffer in the kernel. The default is to trace a minimum amount of information, but different levels of detail may be traced. This trace is local to each node. It is controlled by the **vsdtrace** command.

If full tracing is enabled, CPU utilization increases. The trace output may be viewed using the **vsddump** command. The output is displayed to stdout, or it may be sent to a file. If full tracing is enabled, the small circular buffer fills up and wraps quickly.

Trace records contain a timestamp, information about the source and target nodes, and the name of the routine that issued the trace information.

IBM Recoverable Virtual Shared Disk subsystem tracing to the console log

The IBM Recoverable Virtual Shared disk subsystem traces its activity to the console log. This trace shows rvsd responses to requests coming from: Group Services, the AIX System Resource controller, and the IBM Virtual Shared Disk device driver. This trace is local to each node.

A basic level of tracing is enabled by default, but more detail may be enabled or disabled with the **ha.vsd trace on|off** command. This trace has negligible impact on the system’s performance and other resources. The console log must be directed to a file in order to capture the output of this trace. This log may grow rapidly, consuming space in the **/var** directory.

Trace records contain a timestamp and the name of the routine that issued the trace information.

IBM Recoverable Virtual Shared Disk subsystem logging of recovery actions

The IBM Recoverable Virtual Shared Disk subsystem logs recovery information. This information is also logged to the console log, where you can see how recovery interacts with other system activities. Recovery actions are always logged.

This trace is local to each node. The log is located in: **/var/adm/csd/vsd.log** on each node. There is also a symbolic link to this file from: **/var/adm/SPlogs/csd/vsd.log**. This log is pre-allocated, and it is kept trimmed to the most recent 4000 lines. Trace records contain a timestamp and the name of the routine issuing the trace record.

Information to collect before contacting the IBM Support Center

The **vsd.snap** script collects all information necessary to report problems related to the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk components of PSSP.

Note: If the PSSP system has crashed, obtain a system dump and accompanying unix files. See “Chapter 5. Producing a system dump” on page 81.

The **vsd.snap** script collects general system environment information, IBM Virtual Shared Disk configuration information, and various trace logs. The **vsd.snap** script collects information from an individual node. The information may be placed in a location specified by the user. The default location is: **/tmp/vsd.snapOut**. This information is collected only when the user explicitly invokes the **vsd.snap** script.

Diagnostic procedures

These procedures check for errors in the installation, configuration and operation of the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk subsystems of PSSP.

Installation verification test

This test determines if you have successfully installed the IBM Virtual Shared Disk software. It also tests if you can successfully define, activate, read from, and write to an IBM Virtual Shared Disk. To run the test, follow these steps:

1. From the control workstation, use the **createvsd** command to create an IBM Virtual Shared Disk. For example:

   ```
   /usr/lpp/csd/bin/createvsd -n 1/:hdisk0/ -g rootvg -s 8 -c 1 -v junk
   ```
2. From the client node where you will run the test, the device driver and the virtual shared disk must be configured. For example:

   ```
   /usr/lpp/csd/bin/cfgvsd -a
   ```
3. From a client node, run the **vsdvts** command. For example,

   ```
   /usr/lpp/csd/bin/vsdvts vsd_name
   ```
4. If the virtual shared disk was created just for this test, remove it. From the control workstation, issue this command:

   ```
   /usr/lpp/csd/bin/removevsd -f -v vsd_name
   ```

The **vsdvts** function displays the commands it issues and whether they were successful or not. If **vsdvts** fails, see “Configuration test 1 - Check IBM Virtual Shared Disk nodes” on page 480.

Configuration verification tests

Use these tests to check that the IBM Virtual Shared Disk subsystem is installed properly.

Configuration test 1 - Check IBM Virtual Shared Disk nodes

This test checks that the IBM Virtual Shared Disk nodes have been designated appropriately. Obtain the IBM Virtual Shared Disk node information by issuing this command on the control workstation:

```
/usr/lpp/csd/bin/vsdatalst -n
```

The output is similar to the following:

```
VSD Node Information
                                   Initial Maximum     VSD      rw     Buddy Buffer
  node           VSD     IP packet  cache   cache request request minimum maximum size: #
number host_name adapter   size    buffers buffers   count   count    size    size maxbufs
------ --------- -------  -------- ------- ------- ------- ------- ------- ------- -------
     1 mynode1   css0        61440      64    4096     256      48    4096  262144     130
     3 mynode2   css0        61440      64    4096     256      48    4096  262144     130
     5 mynode3   css0        61440      64    4096     256      48    4096  262144     130
     7 mynode4   css0        61440      64    4096     256      48    4096  262144     130
```

Good results are indicated by:

1. The list contains all expected nodes.
2. All nodes must use the same type of communications adapter (virtual shared disk adapter).
3. If using the **css0** adapter, the IP packet size should be **61440**.

In this case, proceed to “Configuration test 2 - Check that all IBM Virtual Shared Disk nodes know about each other” on page 481.

Error results are indicated if any of the criteria listed are not met.

Actions to take if configuration test 1 detects an error: Some of these parameters require that the device driver is unloaded from the kernel and reloaded.

1. Stop the IBM Recoverable Virtual Shared Disk subsystem on the nodes by issuing this command on the control workstation:

   ```
   dsh /usr/lpp/csd/bin/ha.vsd stop
   ```

   This assumes that the WCOLL environment variable is set, and points to a file listing the nodes. See Step 1 on page 474 of “Before you begin” on page 474.
2. Unconfigure the virtual disks on the nodes by issuing this command on the control workstation:

   ```
   dsh /usr/lpp/csd/bin/ucfgvsd -a
   ```
3. Unload the device driver on the nodes by issuing this command on the control workstation:

   ```
   dsh /usr/lpp/csd/bin/ucfgvsd VSD0
   ```
4. Correct the node attributes. For example, on the control workstation issue:

   ```
   /usr/lpp/csd/bin/updatevsdnode -n ALL -a css0 -M 61440
   ```
5. Configure the device driver and the shared disks on the nodes and start the IBM Recoverable Virtual Shared Disk subsystem. Issue this command on the control workstation:

   ```
   dsh /usr/lpp/csd/bin/ha_vsd
   ```

Configuration test 2 - Check that all IBM Virtual Shared Disk nodes know about each other

Check that all IBM Virtual Shared Disk nodes know about each other. On the control workstation, issue this command:

```
dsh "/usr/lpp/csd/bin/lsvsd -i | wc -l"
```

Note that the quotes are important here.

The output is similar to the following:

```
node1:    16
node3:    16
node8:    16
```

Good results are indicated if all numbers in the right hand column are the same.

Error results are indicated if the numbers are not the same. This means that an IBM Virtual Shared Disk node has been added or deleted, and not all the nodes are aware of the change.

Actions to take if configuration test 2 detects an error: To correct the situation, perform one of the following steps:

1. If the rvsd subsystem has been started on all the nodes, issue this command from any node where rvsd is active:

   ```
   /usr/lpp/csd/bin/ha.vsd refresh
   ```
2. Follow the steps in “Actions to take if configuration test 1 detects an error” on page 480.

Then, issue this command from the control workstation:

```
dsh "/usr/lpp/csd/bin/lsvsd -i | wc -l"
```

Operational verification tests

These tests instruct the user to check for errors in the operation of the IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk subsystems. **Good results** and **Error results** paragraphs explain how to interpret the test results. Start with Test 1 and proceed to the next test or perform other actions as indicated in text of each test.

Operational Test 1 - Check the IBM Virtual Shared Disks states using SP Perspectives

Use the IBM Virtual Shared Disk Perspective table view. By putting the Nodes pane in table view, you can keep an eye on the states while continuing with other activities. To view virtual shared disk states (active, suspended or stopped) and the IBM Recoverable Virtual Shared Disk subsystem state in table view, do the following on the control workstation:

1. Start the IBM Virtual Shared Disk Perspective with this command: **/usr/lpp/ssp/bin/spvs**.
2. Click on the nodes pane
3. Click on **View** → **Show Objects in Table View** or click on the table icon in the tool bar.
4. Click on the **IBM VSD Node** tab in the **Set Table Attributes for Nodes** dialog.
5. Click on the attributes to view while pressing the **<Ctrl>** key.

6. Click on **OK**.

Good results are indicated if the IBM Recoverable Virtual Shared Disk subsystem is in an active state, and all IBM Virtual Shared Disks are in an active state.

Error results are indicated if IBM Virtual Shared Disks are in a suspended or stopped state. This indicates a potential problem.

In all cases, proceed to “Operational test 2 - Check the IBM Recoverable Virtual Shared Disk subsystem”.

Operational test 2 - Check the IBM Recoverable Virtual Shared Disk subsystem

This test checks that the IBM Recoverable Virtual Shared Disk subsystem is started, and that the IBM Virtual Shared Disks have been activated. To run this test, issue this command on the control workstation:

```
dsh /usr/lpp/csd/bin/ha.vsd query
```

Good results are indicated by output similar to the following. Note in particular that `active=1`.

```
Subsystem        Group            PID     Status
 rvsd            rvsd             19324   active
 rvsd(vsd): quorum= 8, active=1, state=idle, isolation=member,
            NoNodes=12, lastProtocol=nodes_joining,
            adapter_recovery=on, adapter_status=up,
            RefreshProtocol has never been issued from this node,
            Running function level 3.1.1.0.
```

Error results are indicated by output similar to one of these two examples:

```
Subsystem        Group            PID     Status
 rvsd            rvsd                     inoperative
```

OR

```
Subsystem        Group            PID     Status
 rvsd            rvsd             2570    active
 rvsd(vsd): quorum= 8, active=0, state=idle, isolation=isolated,
            NoNodes=0, lastProtocol=idle,
            adapter_recovery=on, adapter_status=down,
            RefreshProtocol has never been issued from this node,
            Running function level 3.1.1.0.
```

If the test indicates an error, follow these instructions:

1. If `active=0` from the output:
   - If `isolation=member`, compare the `quorum=` and `NoNodes=` fields. If `quorum` is greater than `NoNodes`, determine which nodes are not part of the `ha.vsd` group.

     Issue this command on the control workstation:

     ```
     dsh /usr/lpp/csd/bin/ha.vsd query | grep inoperative
     ```

  You could also check that the rvsd subsystem is not waiting for Group Services on some nodes, by issuing this command on the control workstation:

  ```
  dsh /usr/lpp/csd/bin/ha.vsd query |\
   grep "waiting for Group Services to connect"
  ```

- If `adapter_recovery=on` and `adapter_status=down` then there is a problem with the communications adapter, which must be corrected.
- If `adapter_recovery=on` and `adapter_status=unknown`, check that the **hats** (Topology Services) and **hags** (Group Services) subsystems have been started. Issue these commands:
  - `lssrc -ls hats`
  - `lssrc -ls hags`

2. On the affected nodes, issue the **vi `lscons`** command to view the console log. Use the log to determine why the IBM Recoverable Virtual Shared Disk daemon has exited, or to see if a recovery problem has been logged. An example is a failure varying a volume group online.
3. On the affected nodes, view the file **/var/adm/SPlogs/csd/vsd.log** to determine if the IBM Recoverable Virtual Shared Disk subsystem has logged a recovery problem such as a failure varying a volume group online.
4. On the affected nodes, view the system error log to check for IBM Virtual Shared Disk error entries, by issuing the command:

   ```
   errpt -a | pg
   ```

In all cases, proceed to “Operational test 3 - Check the IBM Virtual Shared Disk states using commands”.

Operational test 3 - Check the IBM Virtual Shared Disk states using commands

On the control workstation, see if the IBM Recoverable Virtual Shared Disk subsystem is active with this command:

```
dsh /usr/lpp/csd/bin/lsvsd -l | grep -E "SUS|STP" | dshbak -c
```

Good results are indicated if no shared disks are in the suspend SUS or stopped STP states. In this case, proceed to “Operational test 4 - Display the IBM Virtual Shared Disk device driver statistics” on page 484.

Error results are similar to the following example:

```
HOSTS -------------------------------------------------------------
c164n04.ppd.pok.ib
-------------------------------------------------------------------
  7      STP   -1       0       0      vsd2        nocache          8
  8      STP   -1       0       0      vsd1        nocache          8
307      SUS   -1       0       0      gpfsvsd7    nocache       2148
```

If the test indicates an error, determine why the shared disks are not active (ACT). Virtual shared disks are normally in suspend (SUS) state for only a short period of time, while recovery is taking place. Virtual shared disks are not normally in a stopped (STP) state unless they have been explicitly stopped. Perform these steps on the affected nodes:

1. View the console log to determine if the IBM Recoverable Virtual Shared Disk subsystem logged a recovery problem such as a failure varying a volume group online.
2. View the file **/var/adm/SPlogs/csd/vsd.log** to determine if the IBM Recoverable Virtual Shared Disk subsystem logged a recovery problem, such as a failure varying a volume group online.
3. View the system error log to check for IBM Virtual Shared Disk error entries by issuing the command:

```
errpt -a -N vsdd rvsd rvsdd | pg
```

If problems with one or more volume groups are detected, pursue this problem as a hardware problem or a potential AIX Logical Volume Manager problem.

Operational test 4 - Display the IBM Virtual Shared Disk device driver statistics

On the control workstation, display the device driver statistics by performing these steps from within the IBM Virtual Shared Disk Perspective:

1. Click on a virtual shared disk Node from the **Nodes** pane.
2. Click on the **Properties** notebook icon in the tool bar, or click the **Actions** → **View or Modify Properties...** to display the ″IBM Virtual Shared Disk Node″ notebook.
3. Click on the **Virtual Shared Disk Node Statistics** tab of the notebook.

Without using SP Perspectives, use the **statvsd** command on the nodes, to obtain the statistics.

Suspect an error if there are a large number of **requests queued** and **timeouts**.

Good results are indicated by **0 timeouts**, as in this example:

```
VSD driver (vsdd): IP/SMP interface:    PSSP Ver:3 Rel: 1.1

         9 vsd parallelism
     61440 vsd max IP message size
         0 requests queued waiting for a request block
         0 requests queued waiting for a pbuf
         0 requests queued waiting for a cache block
         0 requests queued waiting for a buddy buffer
       0.0 average buddy buffer wait_queue size
         0 rejected requests
         0 rejected responses
         0 rejected merge timeout.
         0 requests rework
       958 indirect I/O
         0 64byte unaligned reads.
         0 timeouts
retries: 0 0 0 0 0 0 0 0 0
         0 total retries
Non-zero Sequence numbers
 node#        expected        outgoing   outcast? Incarnation: 0
     1           12660               0
     2           51403               0
     4           19951              14

 8 Nodes Up with zero sequence numbers: 3 5 6 7 8 11 12 14
```

Error results are indicated by nonzero values in the **timeout** field, or large numbers (in the thousands or hundreds of thousands range) of **requests queued**.

If the test indicates an error, check the error log to determine the minor number of the IBM Virtual Shared Disks that timed out. Issue this command:

```
errpt -a -J VSD_INT_ER | pg
```

Use the minor number to determine the IBM Virtual Shared Disk name, by issuing this command:

```
lsvsd -l | grep minor_number
```

In all cases, proceed to “Operational test 5 - Check the IBM Virtual Shared Disk server”.

Operational test 5 - Check the IBM Virtual Shared Disk server

This test checks the IBM Virtual Shared Disk server to see if it is active and accessible. On the client node, issue:

```
lsvsd -l vsd_name
```

and refer to and refer to the `server_list` column to determine the node numbers of the servers. For the case of a Concurrent Virtual Shared Disk (CVSD), repeat the following steps for each server.

The output of the **lsvsd -l** command is similar to the following:

```
minor  state server lv_major lv_minor vsd-name  option  size(MB)  server_list
 5      ACT   13      42       1      Vsd1n13   nocache 128        13,14
```

Now issue the command:

```
vsdatalst -n
```

to correlate the server numbers to node names. The output is similar to the following:

```
       VSD Node Information
                                         Initial Maximum    VSD      rw      Buddy Buffer
  node                 VSD     IP packet   cache   cache request request minimum maximum size: #
number host_name       adapter    size   buffers buffers   count   count    size    size maxbufs
------ --------------- ------- --------- ------- ------- ------- ------- ------- ------- -------
     1 c164n01.ppd.pok css0        61440      64     256     256      48    4096  262144      66
     2 c164n02.ppd.pok css0        61440      64     256     256      48    4096  262144      66
    13 c164n13.ppd.pok css0        61440      64     256     256      48    4096  524288      18
    14 c164n14.ppd.pok css0        61440      64     256     256      48    4096  524288      18
    16 c164n16.ppd.pok css0        61440      64     256     256      48    4096  524288      18
```

Is the IBM Virtual Shared Disk local or remote? Issue the command:

```
ping host_name
```

where *host_name* is obtained from the previous step.

Good results are indicated if the **ping** is successful. Proceed to “Operational test 6 - Check network options” on page 486.

Error results are indicated if the **ping** fails. In this case, issue this command:

```
ifconfig VSD_adapter
```

to determine if the IBM Virtual Shared Disk adapter is active (UP). If the adapter is not active, pursue this as a switch or network problem.

Operational test 6 - Check network options

This test checks the network options that affect virtual shared disks. Issue this command:

```
/usr/sbin/no -a | grep -E "thewall|ipqmaxlen"
```

Good results are indicated by these recommended values:

- thewall = 65536
- ipqmaxlen = 1024

If the results are good, proceed to “Operational test 7 - Check ability to read from the IBM Virtual Shared Disk”.

Error results are indicated by values less than the recommended ones. This may adversely affect performance and cause timeouts. To change these network options, issue the commands:

```
/usr/sbin/no -o thewall=65536
```

and

```
/usr/sbin/no -o ipqmaxlen=1024
```

Operational test 7 - Check ability to read from the IBM Virtual Shared Disk

This test determines if we can read from the IBM Virtual Shared Disk, both locally and remotely. Issue this command both locally on the IBM Virtual Shared Disk server, and remotely on another node:

```
dd if=/dev/r{vsdname} of=/dev/null bs=4k count=1
```

Good results are indicated by the following output:

```
1+0 records in.
1+0 records out.
```

Error results are indicated if the command hangs, which may be for up to 15 minutes.

If the IBM Virtual Shared Disk is accessible locally but not remotely, there may be a virtual shared disk sequence number problem or a routing problem. Proceed to “Operational test 8 - Check that the client and server nodes have routes to each other” to check the device driver routing tables.

If the virtual shared disk is not locally accessible, issue the **dd** command from page 486 to the local logical volume. If the **dd** command fails, pursue the problem as a potential AIX Logical Volume Manager problem.

Operational test 8 - Check that the client and server nodes have routes to each other

On both the client node (where an IBM Virtual Shared Disk cannot be read) and on the server node for that disk, issue this command:

```
/usr/lpp/csd/bin/lsvsd -i
```

The output is similar to the following:

```
node             IP address
  1            9.114.68.129
  5            9.114.68.130
  12           [KLAPI 11]
```

Good results are indicated if the output is the same on both the client and server nodes.

Error results are indicated if the output differs between the client and server nodes. A typical problem is missing entries. In this case, issue the command:

```
ha.vsd refresh
```

This refreshes the node information on all nodes where the IBM Recoverable Virtual Shared Disk subsystem is active.

Error symptoms, responses, and recoveries

The recovery scenarios describe what a system administrator or operator might see when the IBM Recoverable Virtual Shared Disk component goes into action to recover from system problems.

Recognizing recovery

You know that the rvsd subsystem is performing recovery processing when you see IBM Virtual Shared Disks that are in the active or suspended state, and you did not put them there. Use the IBM Virtual Shared Disk Perspective to display the states of your IBM Virtual Shared Disk nodes. See the chapter on Managing and Monitoring Virtual Shared Disks, section on Monitoring Virtual Shared Disks, in: *PSSP: Managing Shared Disks*.

If you recognize that recovery is not taking place normally, check if the rvsd subsystem is active on all nodes. Use the IBM Virtual Shared Disk Perspective, or the **ha.vsd query** and **hc.vsd query** commands to see if the respective subsystems are active. If `active=0` or `state=idle` for an extended period of time, recovery is not taking place normally. The **ha.vsd query** command returns output in the following format:

```
Subsystem         Group            PID     Status
rvsd              rvsd             18320   active
rvsd(vsd): quorum= 7, active=0, state=idle, isolation=member,
           NoNodes=5, lastProtocol=nodes_failing,
           adapter_recovery=on, adapter_status=up,
           RefreshProtocol has never been issued from this node,
           Running function level 3.2.0.0.
```

The **hc.vsd query** command output looks like this:

```
Subsystem         Group            PID     Status
hc.hc              rvsd             20440   active
 hc(hc): active=0, state=waiting for client to connect
 PING_DELAY=600
 CLIENT_PATH=/tmp/serv.
 SCRIPT_PATH=/usr/lpp/csd/bin
```

Planning for recovery

You should have disabled or removed all user-provided scripts that issue change-of-state commands to IBM Virtual Shared Disks.

Monitor the activity of the rvsd subsystem using the IBM Virtual Shared Disk Perspective, to become aware of potential problems as soon as they arise. You can begin a monitoring session, leave it in a window on your workstation and go about doing other work while the monitoring activity continues.

Each of the recovery scenarios is organized as follows:
1. Symptoms
2. Detection
3. Affected components
4. Recovery steps
5. Restart

Virtual Shared Disk node failure

- **Symptoms**

  A node has either hung or has failed.
- **Detection**

  Node failure was detected by the Group Services program, and it has notified the rvsd subsystem. An operator may have seen the change in state displayed by the IBM Virtual Shared Disk Perspective.
- **Affected Components**

  The affected components can be everything accessing the IBM Virtual Shared Disks, including: the IBM Virtual Shared Disks themselves, software applications, and the rvsd and related subsystems running on the failed nodes.
- **Recovery Steps**

  The recovery steps are:
  1. The recovery services of surviving nodes suspend IBM Virtual Shared Disks served by the failing node. The IBM Virtual Shared Disk recovery process puts the suspended IBM Virtual Shared Disks into the active state on the secondary node.
  2. Optional subscribers to **hc** on surviving nodes are informed of the membership change.
- **Restarting the Failed Node**

  Some of the following actions must be done manually; others are done automatically by the rvsd subsystem. You should have procedures in place to instruct operators when and how to do the manual operations.
  1. An operator reboots the failed node.
  2. An operator may need to issue the **Estart** or **Eunfence** command to enable the rebooted node to access the switch. Nodes can be set up to reboot automatically, using the **Estart -M** command, which starts the monitor function.
  3. The rvsd subsystem is automatically brought up on reboot, once the communications adapter is available.
  4. The **hc.activate** script, if present, is invoked.
  5. The rebooted node rejoins the active group. If it was a primary node, it takes over again as a primary node.
  6. Optional subscribers to **hc** on surviving nodes are informed of the membership change.

Switch failure scenarios

The sequence of events triggered by a switch failure depends on whether adapter recovery is enabled.

When adapter recovery is enabled

- **Symptoms**

  The rvsd subsystem initiates recovery on failing nodes. Errors might be generated by applications running on these nodes. Error reports might be generated as well.
- **Detection**

  Node failure was detected by the rvsd subsystem. An operator may have seen the change in state displayed by the IBM Virtual Shared Disk Perspective.
- **Affected Components**

  The affected components can be everything accessing the IBM Virtual Shared Disks, including: the IBM Virtual Shared Disks themselves, software applications, and the rvsd and related subsystems running on the failed nodes.
- **Recovery Steps and Restart**

  Follow problem determination procedures for the switch failure. See “Chapter 15. Diagnosing SP Switch problems” on page 137 and “Chapter 16. Diagnosing SP Switch2 problems” on page 185.

When adapter recovery is disabled

- **Symptoms**

  Remote IBM Virtual Shared Disk I/O requests hang and then fail after about 15 minutes. The IBM Virtual Shared Disk clients see a time out error.
- **Detection**

  The rvsd subsystem neither detects nor handles switch failure when adapter recovery is disabled or when a non-supported adapter is used. An operator using the SP or IBM Virtual Shared Disk Perspective might recognize that **switch_responds** for the node is off.
- **Affected Components**

  Applications using IBM Virtual Shared Disks will hang and then fail after about 15 minutes.
- **Recovery Steps and Restart**

  An operator issues the **Estart** or **Eunfence** command to restart the switch.

  If **Estart** or **Eunfence** fails, use standard diagnostic methods for handling switch problems. See “Chapter 15. Diagnosing SP Switch problems” on page 137 and “Chapter 16. Diagnosing SP Switch2 problems” on page 185.

  The Problem Management interface could be used to run a script that would stop the rvsd subsystem on switch failures and restart it when the switch is active again.

Topology Services or recovery service daemon failure

- **Symptoms**
  - At the node where the failure occurred, all the IBM Virtual Shared Disks stop and restart, causing I/O errors to the application.
  - At other nodes, the IBM Virtual Shared Disks served by the problem node are switched to their secondary servers briefly, then returned to the primary server.
- **Detection**

  There is no automatic method to determine that the daemons are failing.
- **Recovery Steps**
  1. In the unlikely event that the recovery service daemons hang, no recovery is performed.

2. At the node where the failure occurred, the IBM Virtual Shared Disks remain in the active state. However, subsequent node failures or reboots can cause I/O requests to remote IBM Virtual Shared Disks to hang, and fail after 15 minutes. Issue the **ps** command at that node to check for the **rvsd** and **hc** daemons.
3. At the other nodes, some IBM Virtual Shared Disks remain indefinitely in the suspended state. All the **rvsd** daemons are present, and their presence can be verified by using the **ps** command.
4. When the daemons are available again, you can manually issue the **ha_vsd reset** command on the problem node. If this is insufficient, reboot the problem node.

For more information, see the book *RS/6000 Cluster Technology: Group Services Programming Guide and Reference*.

Disk EIO errors

- **Symptoms**

  I/O errors occur on some or all of the IBM Virtual Shared Disks served from a node.
- **Detection**

  There is no automatic method of determining that volume group failures are occurring. An entry is posted in the System Error Log and a hardware error (EIO) code is returned to the IBM Virtual Shared Disk device driver.
- **Affected Components**

  The IBM Virtual Shared Disk subsystem cannot access the data on the failed volume groups.
- **Recovery Steps**
  1. The volume group that contains the failed IBM Virtual Shared Disks is automatically put into the suspended state by the rvsd subsystem.
  2. If EIO recovery has not taken place on this volume group in approximately the last seven minutes, an attempt is made to switch the volume group over to the backup server. The rvsd subsystem switches the volume group over to the newly assigned primary server and retries the previously failed I/O request. This involves switching the primary and secondary fields in the VSD Global Volume Group Information that is stored in the SDR.

     If EIO recovery has taken place on this volume group in approximately the last seven minutes, the IBM Virtual Shared Disks in this volume group are placed in the stopped state.

     Note: You can tell if EIO recovery has occurred by using the **vsdatalst -g** command and looking at the `recovery` field in the results. If it contains a value other than zero, recovery has taken place at some point.
  3. Correct the condition that caused the error. You might need to issue the **vsdchgserver** command to switch the primary and backup servers back to their original settings. The command will swap the primary and backup server fields in the SDR, and initiate a failover of the specified volume groups to the new primary virtual shared disk server.

Note: If you mirror data from each IBM Virtual Shared Disk to a IBM Virtual Shared Disk on another adapter, you should not experience this type of error.

Chapter 27. Diagnosing Job Switch Resource Table Services problems

This chapter discusses diagnostic procedures and failure responses for the Job Switch Resource Table Services (JSRT Services) component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 494. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 493.

Related documentation

The following publications provide information about the Job Switch Resource Table Services (JSRT Services) components of PSSP:

1. *PSSP: Command and Technical Reference*

   This book provides detailed syntax and parameter information for the commands used by JSRT Services. It also contains a chapter on SP subroutines used for loading, unloading, querying and cleaning Job Switch Resource Tables.

   The commands used by JSRT Services are:

   - **st_status**
   - **st_clean_table**
   - **st_verify**

   The subroutines used by JSRT Services are:

   - **swtbl_load_table**
   - **swtbl_unload_table**
   - **swtbl_load_job**
   - **swtbl_unload_job**
   - **swtbl_clean_table**
   - **swtbl_status_node**
   - **swtbl_status**
   - **swtbl_query_adapter**
   - **swtbl_adapter_resources**
   - **swtbl_adapter_connectivity**
2. *PSSP: Messages Reference*

   The chapter ″2511 - Job Switch Resource Table Services Messages″ gives detailed information about JSRT Services messages.
3. Manpages for the PSSP commands listed in Item 1.
4. *The RS/6000 SP Inside Out*, SG24-5374
5. *Understanding and Using the SP Switch*, SG24-5161

Requisite function

This is a list of the software directly utilized by JSRT Services component of PSSP. Problems within the requisite software may manifest themselves as error symptoms in this software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with JSRT Services, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- SP Switch or SP Switch2 (CSS Subsystem) of PSSP
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.
- System Data Repository (SDR) component of PSSP
- SP Perspectives - nodegroup function

Error information

Job Switch Resource Table Services error and information log

Every node with the JSRT Services installed contains an error and information log. This log is located in **/var/adm/SPlogs/st/st_log**.

This log contains any error messages or general information messages that occur when the JSRT Services commands or APIs are issued. All entries contain a timestamp and the name of the command or function issuing the message. An entry is created for each occurrence of a condition that produces an error or general information message.

The JSRT Services log exists on both the client and the server nodes. For example, if you issue a **st_clean_table** command on node 1 to clean a window on node 2, then logs will exist on both nodes 1 and 2.

More detailed logging will occur when the environment variable **SWTBLAPIERRORMSGS** is set to **yes**. Setting this variable produces more messages in the **st_log**. This causes the log to be truncated and copied more frequently. To set the variable, issue this command:

```
export SWTBLAPIERRORMSGS=yes
```

This log is copied to **/var/adm/SPlogs/st/st_log.previous** when it reaches 100KB in size. Each subsequent copy overwrites the **/var/adm/SPlogs/st/st_log.previous** file. Therefore, only one copy of **st_log** and **st_log.previous** is retained.

Entries in this log are not translated to other languages. All entries are in English. For more information about **st_log**, see “Action 3 - Request more detailed log information” on page 497.

AIX Error Logs and templates for JSRT Services

The JSRT Services creates AIX error entries using the PSSP **ppslog** facility for the following cases:

- A load attempt fails because a window is already loaded.
- An unload attempt fails because a window is in use.
- The **LOAD_ST** ioctl call fails.
- The **UNLOAD_ST** ioctl call fails.
- The JSRT error log is copied and truncated.

To view the JSRT error events, issue the following command on the node where an error is suspected:

```
errpt -aN Switch_Table | more
```

When you retrieve an error log entry, look for the DIAGNOSTIC EXPLANATION section near the bottom of the entry.

One entry is logged for each occurrence of the condition. The condition is logged on every node where the event occurred.

The Detail Data section of these entries is not translated to other languages. This section is in English.

Table 64 shows the error log templates used by JSRT Services. **ST_TRUNCATE_ST** is used to indicate truncation of the log **/var/adm/SPlogs/st/st_log**. **ST_SWITCH_ERR** is used for all other JSRT Services records. UNKN indicates an unknown error type. PERM indicates a permanent error type.

Table 64. AIX Error Log templates for JSRT Services

Error Label and ID	Error type	Diagnostic explanation and action
ST_TRUNCATE_ST ABE9698F	UNKN	**Explanation:** check_size: Copied *logname* to *logname*.**previous** and truncated *logname.* **Cause:** The **st_log** exceeded 100KB. It was copied and truncated. **Action:** None required.
ST_SWITCH_ERR 918DF996	PERM	**Explanation:** An error occurred during the processing of a JSRT Service. **Cause:** The request for a load, unload, or clean failed. **Actions:** Perform these tasks: • Check the file **/var/adm/SPlogs/st/st_log** for further information. • Perform SP Switch or SP Switch2 diagnostics.

Information to collect before contacting the IBM Support Center

The following items are used to isolate problems in the JSRT Services component of PSSP.

1. JSRT log - **/var/adm/SPlogs/st/st_log** and **/var/adm/SPlogs/st/st_log.previous**.

   These logs are located on the node from which the JSRT Services command was issued, or the node from which the JSRT API was issued, and any nodes involved in the JSRT API activity. These logs should be collected as soon as an error occurs.

   More detailed information about this item appears in “Job Switch Resource Table Services error and information log” on page 492.
2. AIX error log

Any records pertaining to the JSRT Services from the **errpt** command. This log is located on the node from which the JSRT Services command was issued, or the node from which the JSRT API was issued, and any nodes involved in the JSRT API activity. The log states the nodes on which the error occurred.

More detailed information about this item appears in “AIX Error Logs and templates for JSRT Services” on page 492.

3. Output of the **st_verify** command.

   This command verifies the installation of the **ssp.st** file set. The log name can be specified by the user, or defaults to **/var/adm/SPlogs/st/st_verify.log**.

   More detailed information about this item appears in “Action 1 - Verify JSRT Services installation” on page 495.
4. The authentication method in use. Issue the **splstdata -p** command to obtain this information.

Diagnostic procedures - Installation verification

The **st_verify** command verifies the installation of the **ssp.st** file set. To run this command and interpret the results, see “Action 1 - Verify JSRT Services installation” on page 495.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the Job Switch Resource Table (JSRT) Services component of PSSP. Locate the symptom and perform the action described in the following table.

Table 65. Job Switch Resource Table (JSRT) Services symptoms

Symptom	Recovery
Cannot **load** or **unload** a Job Switch Resource Table (JSRT) on a node.	“Action 1 - Verify JSRT Services installation” on page 495 “Action 2 - Check the JSRT services log file” on page 496 “Action 3 - Request more detailed log information” on page 497 “Action 5 - Check the switch_node_number file” on page 498 “Action 6 - Check the current status of JSRT Services for a node” on page 498
Cannot obtain the status of a JSRT window.	“Action 1 - Verify JSRT Services installation” on page 495 “Action 2 - Check the JSRT services log file” on page 496 “Action 4 - Check the JSRT Services data files” on page 498
Cannot run the **switchtbld** daemon.	“Action 1 - Verify JSRT Services installation” on page 495 “Action 2 - Check the JSRT services log file” on page 496
st_status command hangs. No messages are displayed.	“Action 7 - Check communication paths for affected nodes” on page 499
Cannot load program **st_status**.	“Action 8 - Check the LIBPATH environment variable” on page 499
st_status returns **ST_SYSTEM_ERROR** or **ST_LOADED_BYOTHER** for a window.	“Action 9 - Issue the st_clean_table command” on page 499
st_status returns **ST_RESERVED** status for a window.	“Action 10 - Issue the chgcss command” on page 499 “Action 3 - Request more detailed log information” on page 497

Table 65. Job Switch Resource Table (JSRT) Services symptoms (continued)

Symptom	Recovery
The **st_status** command runs in DCE and compatibility mode and issues the message: swtbl_status: 2511-508 Time out while waiting for a response from host *hostname*.	"Action 13 - Ensure that the client and master DCE servers are running." on page 500
API returned **ST_NOT_AUTHEN**, **ST_NOT_AUTHOR** or **ST_SECURITY_ERROR**.	"Action 2 - Check the JSRT services log file" on page 496 "Action 3 - Request more detailed log information" on page 497 "Action 12 - Ensure that the caller is a member of the appropriate DCE security group" on page 499

Actions

Action 1 - Verify JSRT Services installation

When issued on the control workstation, the **st_verify** script checks the installation of the JSRT Services on every node that is defined in the current system partition. When issued on a single node, it verifies the installation of the JSRT Services on only that node.

The user must be logged in as **root** in order to perform this action. Run installation verification tests using SMIT, or from the command line, to ensure that installation is complete.

Using SMIT:

```
TYPE    smit SP_verify
        (the Installation/Configuration Menu appears)
SELECT  Job Switch Resource Table Services Installation
PRESS   Enter
```

Using the command line, enter: **/usr/lpp/ssp/bin/st_verify** .

The **st_verify** script checks that the correct files and directories were installed and that the necessary entries exist in the files. The files and directories are **/etc/services** , **/etc/inittab** , and **/etc/inetd.conf**.

Installation verification test output: The **st_verify** script produces an output log, located in **/var/adm/SPlogs/st/st_verify.log** (by default) or in a location that you specify. After completion, a message is written to stdout stating whether the verification passed or failed. If a failure occurred, examine the log for a list of the errors that were found.

Good results are indicated when a message similar to the following is written to stdout:

```
Verifying installation of the Job Switch Resource Table Services on node 0.
JSRT Services installation verification SUCCESSFUL on node 0.
Check /var/adm/SPlogs/st/st_verify.log file for further details.
```

Error results are indicated when a message similar to the following is written to stdout:

```
Verifying installation of the Job Switch Resource Table Services on node 6.
JSRT Services installation verification FAILED on node 6.
Total of 1 errors found on node 6.
```

If the test fails, check the **/var/adm/SPlogs/st/st_verify.log file** (or specified log) for further details about which files or directories are missing. Reinstall the **ssp.st** file set if necessary.

Action 2 - Check the JSRT services log file

The JSRT Services maintain a single log file, **st_log**, which is located in: **/var/adm/SPlogs/st**. This log is located on every node where the services are used. For example, if the **swtbl_load_job** API is used, entries are found on the local node where the API was invoked, and entries are also found on the nodes that were being loaded by the **swtbl_load_job** API.

Examine the logs and correct any obvious problems that have been identified.

The following table indicates return codes that may appear in the log. They are defined in **/usr/lpp/ssp/include/st_client.h**.

If a return code of 2 (ST_NOT_AUTHOR), 3 (ST_NOT_AUTHEN), or 20 (ST_SECURITY_ERROR) appears in the log, see “Chapter 18. Diagnosing SP Security Services problems” on page 251 and “Chapter 19. Diagnosing Per Node Key Management (PNKM) problems” on page 289.

Table 66. JSRT Services return codes

Return code	Name	Explanation
0	ST_SUCCESS	The service request was successful.
1	ST_INVALID_TASK_ID	An invalid task ID is specified as input.
2	ST_NOT_AUTHOR	The caller is not authorized to perform the service.
3	ST_NOT_AUTHEN	The caller is not authenticated to perform the service.
4	ST_SWITCH_IN_USE	The JSRT is already loaded or in use.
5	ST_SYSTEM_ERROR	A system error occurred.
6	ST_SDR_ERROR	An SDR error occurred.
7	ST_CANT_CONNECT	The **connect** system call failed.
8	ST_NO_SWITCH	No css device is installed.
9	ST_INVALID_PARAM	An invalid parameter was specified as input.
10	ST_INVALID_ADDR	The **inet_ntoa** command failed on the **st_addr** input value.
11	ST_SWITCH_NOT_LOADED	No JSRT is currently loaded.
12	ST_UNLOADED	A previously successful load was unloaded because of an error.
13	ST_NOT_UNLOADED	No unload request was issued.
14	ST_NO_STATUS	No status request was issued.
15	ST_DOWNON_SWITCH	The node is down on the switch.
16	ST_ALREADY_CONNECTED	The node has already been connected to, and a load request made by a previous **st_node_info** structure.
17	ST_LOADED_BYOTHER	The JSRT was loaded outside of the API.
18	ST_SWNODENUM_ERROR	An error occurred when processing the switch node number.
19	ST_SWITCH_DUMMY	For testing purposes.

Table 66. JSRT Services return codes (continued)

Return code	Name	Explanation
20	ST_SECURITY_ERROR	DCE security error.
21	ST_TCP_ERROR	Error using TCP/IP.
22	ST_CANT_ALLOC	Cannot allocate storage.
23	ST_OLD_SECURITY	Old security method used.
24	ST_NO_SECURITY	No security methods used.
25	ST_RESERVED	Window reserved outside of API.

Action 3 - Request more detailed log information

To facilitate debugging, you can set several environment variables before invoking a JSRT Service. The variables provide more detailed information in the **st_log** file.

The first environment variable is **SWTBLAPIERRORMSGS**, and it must be set to **yes** within the caller's environment.

For example, as a **ksh** user, enter:

```
export SWTBLAPIERRORMSGS=yes
```

This is an example of the more detailed log information for a call to **swtbl_load_table**:

```
Thu Jun 18 10:12:22 1998: swtbl_load_table: INPUT PARAMETERS: uid - 0
pid - 19118 job_key - 1 requestor_node - k10n11.ppd.pok.ibm.com
num_tasks - 1 job_desb - load client test
Thu Jun 18 10:12:22 1998: swtbl_load_table: INPUT PARAMETERS virtual
 task id=0 switch_node_number=10 window_id=1
```

The second environment variable is **SWTBLSECMSGLEVEL** and it can be set to a an integer value which represents a level of detail. See Table 67. If a value greater than 4 is used, the highest level of tracing (4) is assumed.

Table 67. SWTBLSECMSGLEVEL environment variable values

Value	Level of detail
0	No tracing is performed.
1	Trace error conditions.
2	Trace significant flow or data.
3	Trace calls and returns from function.
4	Trace flows and data - maximum detail.

For example, as a **ksh** user, enter:

```
export SWTBLSECMSGLEVEL=4
```

This is an example of the more detailed log information for a call to **swtbl_load_job**:

```
Wed Aug 25 09:43:25 1999: Ssobj: spsec_authenticate_client failed -
cl=c187n12.ppd.pok.ibm.com, gr=switchtbld-status, inf=2502-611
 An argument is missing or not valid.
```

```
Wed Aug 25 09:43:25 1999: Ssobj::addError(eC=7, s=3, r=0, e=0, Error 0,
 ex=0, er=0, sH = c187n12.ppd.pok.ibm.com
Wed Aug 25 09:43:25 1999: Ssobj::addError(g=switchtbld-status, sE=1)
Wed Aug 25 09:43:25 1999: Ssobj::addError(sA=2502-611 An argument is missing
 or not valid.)
Wed Aug 25 09:43:25 1999: Ssobj::addError() - Return 7, p=201322e8
Wed Aug 25 09:43:25 1999: Ssobj::returnErrors(s=3,
 host=c187n12.ppd.pok.ibm.com)
```

Action 4 - Check the JSRT Services data files

The JSRT Services maintains a set of data files that are located in the **/spdata/sys1/st** directory on every node. Verify that the directory exists and that the files have **root** access. Note that the files are not created until the load or unload services have been invoked. These files are also removed when a node is rebooted.

Action 5 - Check the switch_node_number file

The **/spdata/sys1/st/switch_node_number** file contains a single integer that represents the switch node number of the node. This file can be read using an AIX editor, or by issuing the **cat** command. The integer in the file should match the **switch_node_number** attribute in the SDR Node class for that node. Issue the **/usr/lpp/ssp/bin/st_set_switch_number** installation script on every node to create the **switch_node_number** file and set the correct value.

The following messages in the **st_log** indicate that there is a problem with the switch node number:

1. Error reading the **/spdata/sys1/st/switch_node_number** file. Failed to get the switch node number for this node. Issue the command:

   ```
   /usr/lpp/ssp/bin/st_set_switch_number
   ```

2. The switch node number read from the **/spdata/sys1/st/switch_node_number** file is invalid, switch_number=*number from file*.

Action 6 - Check the current status of JSRT Services for a node

The **st_status** command shows you the current status of the JSRT windows on all nodes or on the node specified. This tells you whether the JSRT windows are loaded, unloaded, reserved by another subsystem, or in error.

To show the status of JSRT windows on all nodes within the current system partition, issue **st_status**.

To show the status of all JSRT windows on node **k10n15**, issue: **st_status k10n15.** Output similar to the following appears:

```
*************************************************************
Status from node: k10n15  User: root
Load request from: k10n15 Pid: 12494 Uid: 0
Job Description: No_job_description_given
Time of request:Wed_Jan_24_13:38:21_EDT_1998
Adapter: /dev/css0 Memory Allocated: 10000
Window id: 0
*************************************************************
Node k10n15 adapter /dev/css0 window 1 returned ST_RESERVED.
Window 1 is RESERVED by VSD.
*************************************************************
Node k10n15 adapter /dev/css0 window 2 returned ST_SWITCH_NOT_LOADED
*************************************************************
Node k10n15 adapter /dev/css0 window 3 returned ST_SWITCH_NOT_LOADED
```

Action 7 - Check communication paths for affected nodes

If the **st_status** command hangs, it is trying to contact a node which has lost its TCP/IP communication path. If the command is issued without any arguments, it will try to contact every node in the current SP system partition. If one of the nodes is down on the Ethernet, the command will not return until the TCP/IP timeout value has been reached.

Issue the **SDRGetObjects** command to determine the nodes in the current SP system partition. For example,

```
SDRGetObjects Node reliable_hostname
```

Issue the AIX **ping** command using the names returned under **reliable_hostname** from the **SDRGetObjects** command to determine connectivity.

Now, reissue the **st_status** command with a specific list of node names.

Action 8 - Check the LIBPATH environment variable

The **st_status** command is compiled with a **LIBPATH** of **/usr/lpp/ssp/lib:/usr/lib:/lib**. If your **LIBPATH** conflicts with this, then issue the command as follows:

```
LIBPATH=/usr/lpp/ssp/lib st_status
```

Action 9 - Issue the st_clean_table command

The **st_clean_table** command forces the unload of the Job Switch Resource Table for a specified window on the specified node. See the **st_clean_table** command man page for more details.

Action 10 - Issue the chgcss command

The **chgcss** command can be used to view and manipulate reserved user space windows. See the **chgcss** command man page for more details.

Action 11 - Call the swtbl_adapter_resources API to obtain valid values

The **swtbl_adapter_resources** API returns the configured resources of the specified adapter on the node from which it is invoked. See the **swtbl_adapter_resources** API man page for more details.

Action 12 - Ensure that the caller is a member of the appropriate DCE security group

These are the DCE security groups defined for the JSRT Services:

Table 68. DCE security groups for JSRT Services

Function	Affected services	DCE group
clean window	swtbl_clean_table API st_clean_table command	ssp/switchtbld/clean
load or unload table	swtbl_load_job swtbl_unload_job	ssp/switchtbld/load
status	swtbl_status st_status	ssp/switchtbld/status

Action 13 - Ensure that the client and master DCE servers are running.

If the client DCE servers are running and the master DCE server is not, a timeout can occur even if DCE and compatibility mode is specified. Either stop the client DCE servers by issuing the **stopsrc** command, or start the master DCE server by issuing the **startsrc** command.

Chapter 28. Diagnosing User Access problems

If your users are having problems logging into the SP System or accessing their home directories, locate the symptom and perform the action described in the following table.

Table 69. User Access symptoms

Symptom	Recovery
No one can login to an SP node	Reboot the node in maintenance mode. If no one is able to login, including **root**, follow the steps in "Action 4 - Boot a node in diagnostic mode" on page 115 for rebooting a node in maintenance mode.
User unable to login to SP node	"Action 1 - Check the /etc/security/passwd file" "Action 2 - Check Login Control"
User unable to access directories managed by the automounter	"Action 3 - Verify that the automount daemon is running"

Actions

Action 1 - Check the /etc/security/passwd file

If a user is having problems logging in to nodes in the SP System, check the **login** and **rlogin** attributes for the user in the **/etc/security/passwd** file on the SP node.

Action 2 - Check Login Control

Check the Login Control facility to see whether the user's access to the node has been blocked.

The System Administrator should verify that the user is allowed access. The System Administrator may have blocked interactive access so that parallel jobs could run on a node.

Action 3 - Verify that the automount daemon is running

On AIX 4.3.1 and later systems, the AutoFS function replaces the automount function of AIX 4.3.0 and earlier systems. All automount functions are compatible with AutoFS. With AutoFS, file systems are mounted directly to the target directory instead of using an intermediate mount point and symbolic links.

Review the commands in the following table and issue the ones that are appropriate for diagnosing the problem.

Table 70. Automounter Related Commands

Command	Comments
ps -ef | grep automount	Verifies that the automount daemon is running on the system on which you are having problems accessing directories.
lssrc -g autofs	For AIX 4.3.1 and later systems, the automounter is controlled by the System Resource Controller (SRC). This command indicates whether the **automountd** daemon is active or not.

Table 70. Automounter Related Commands (continued)

Command	Comments
mount	For AIX 4.3.0 and earlier systems, provides the process id of the automount daemon if it is running, the names of the file systems controlled by the automount daemon, and the active mounts under the **/tmp_mnt** directory. For AIX 4.3.1 and later systems, provides the names of the file systems controlled by the automounter daemon and lists any currently active mounts under the **/tmp_mnt** directory.
view /var/adm/SPlogs/auto/auto.log	Contains error messages generated by PSSP.
view /var/adm/SPlogs/SPdaemon.log	Contains error messages generated by the automount daemon.
splstdata -e | grep amd_config	Informs whether SP automounter has been configured.
splstdata -e | grep usermgmt_config	Informs whether SP user management support has been configured.
splstdata -e | grep filecoll_config	Informs whether SP file collections have been configured.
view /etc/auto.master	Lists the file systems to be controlled by automount and their associated map files.
ls -l /etc/auto/maps	Lists of map files and whether they are readable. Specifically, the existence of **auto.u** map file.
view /etc/auto/maps/auto.u	Lists the user map entries for the **/u** file system.
ls -l /etc/auto/cust	Lists customization files and whether they are executable.
view /var/sysman/sup/lists/user.admin	Lists automounter files that are distributed through file collections.

It may be that the automounter daemon is not running. It is also possible that automount is running but that there is another problem. For AIX 4.3.0 or earlier systems, issue:

```
ps -ef | grep automount
```

For AIX 4.3.1 or later systems, issue:

```
lssrc -g autofs
```

1. Automount Is not running

If issuing the previous command did not show that the automount process was running, issue:

```
mount
```

to see if any automount points are still in use. If you see an entry similar to the following one, there is still an active automount mount point. This is for AIX 4.3.0 and earlier systems:

```
luna.pok.ibm.com (pid23450@/u) /u afs Nov 07 15:41 ro,noacl,ignore
```

For AIX 4.3.1 and later systems, the output is:

```
/etc/auto/maps/auto.u /u autofs Aug 07 11:16 ignore
```

Attempt to unmount the file system by issuing:

```
unmount /u
```

If the file system is busy, issue the following command to determine the processes that are accessing the file system. Stop all of these processes and attempt to unmount the file system again.

```
fuser /u
```

If the **mount** command does not show any active mounts for automount, issue the following command to start the automounter:

```
/etc/auto/startauto
```

Proceed as follows:

- If **startauto** succeeds

  If this command succeeds, issue the previous **ps** or **lssrc** command again to verify that the automount daemon is actually running. If so, verify that the user directories can be accessed or continue with “2. Automounter is running, but the user cannot access user files” on page 504.

  Note that the automount daemon should be started automatically during boot. Check to see if your SP system is configured for automounter support by issuing:

  ```
  splstdata -e | grep amd_config
  ```

  If the result is *true*, you have automounter support configured for the SP system in your Site Environment options.

  If the **startauto** command was successful but the automount daemon is still not running, check to see if the SP automounter function has been replaced by issuing:

  ```
  ls -l /etc/auto/cust
  ```

  If the result of this command contains an entry similar to:

  ```
  -rwx ----- 1 root system 0 Nov 12 13:20 startauto.cust
  ```

  the SP system function to start the automounter has been replaced. View this file to determine which automounter was started and follow local procedures for diagnosing problems for that automounter.

  If the result of the **ls** command does not show any executable user customization script, check both the automounter log file **/var/adm/SPlogs/auto/auto.log** and the daemon log file **/var/adm/SPlogs/SPdaemon.log** for error messages. Find the recorded error messages in *PSSP: Messages Reference* or in the AIX error message documentation and follow the recommended actions.

- If **startauto** fails

  If the **startauto** command fails, find the reported error messages in *PSSP: Messages Reference* and follow the recommended actions. Check the automounter log file **/var/adm/SPlogs/auto/auto.log** for additional messages. Also, check the daemon log file **/var/adm/SPlogs/SPdaemon.log** for messages that may have been written by the automounter daemon itself.

  If no error messages were recorded, the failure may be due to problems with the automount map files, master map file, or the **/u** directory. Check the following:

  – Verify that all entries in the automount master map file **/etc/auto.master** are correct and follow the format specified in the AIX publication *System Management Guide: Communications and Networks*. If there are no entries in this file, the automount daemon invocation will fail.

  – Verify that each map file the master map references is correct, and follows the format specified in the same AIX publication.

– Verify that each file system listed in the master map file is a local directory and is not a symbolic link to another directory. *PSSP: Administration Guide* contains a chapter on managing the Automount that contains useful information for understanding your SP automounter installation.

2. Automounter is running, but the user cannot access user files

For an AIX 4.3.0 or earlier system, if the result of issuing the **ps -ef | grep automount** command is similar to:

```
 root 21430    1   0  10:37:41    0:00 /usr/sbin/automount
 /etc/auto.master  -m -D HOST=k22n11
```

then the automount daemon is running.

For an AIX 4.3.1 or later system, if the result of issuing the **lssrc -g autofs** command is similar to:

```
Subsystem       Group      PID     Status
automountd      autofs     12126   active
```

then the automount daemon is running.

The problem may be that automount is waiting for a response from an NFS server that is not responding, or there is a problem with a map file.

Check the **/var/adm/SPlogs/SPdaemon.log** for information relating to NFS servers not responding. If a user’s files are mounted with NFS and the server is not responding, then automount may hang on the NFS mount request. Correct this NFS failure before continuing. After you resolve the NFS failure, you can restart the automount daemon.

If the problem does not appear to be related to an NFS failure, you will need to check your automount maps. Look at the **/ect/auto/maps/auto.u** map file to see if an entry for the user exists in this file. If the user’s name is test, and the command **cd /u/test** results in:

```
ksh:  /u/test:   not found
```

you can look at the **auto.u** map to see if there is an entry defined for the user by issuing:

```
grep test /etc/auto/maps/auto.u
```

The result may indicate that there is no entry for this user in the automounter map. This can happen if the user was recently added, and the maps have not been distributed in the file collections. To check if the file gets updated with the new map copied from the control workstation, issue the following command on the node experiencing the problems:

```
supper update -v user.admin
```

Note that the automount maps are automatically distributed to the nodes each hour by command in the cron. You can look at these commands with **crontab -l**.

After you updated the **auto.u** map with the version that contains the user information, reread the **auto.u** map by issuing:

```
grep test /etc/auto/maps/auto.u
```

If the result appears as follows:

```
test   luna:/home/luna:&
```

issue the following:

```
cd /u/test
```

If the **cd** command does not work, there may not be a route to the hostname specified in the host field of the user's automount map entry. This can happen on file servers where there are multiple interfaces, and the routing has not been defined for all of the interfaces, from the systems attempting to access the server. You can verify this by attempting to **ping** the server specified.

Another possible problem is that the server is exporting the file system to an interface that is not the interface from which the client is requesting the mount. This problem can be found by attempting to mount the file system manually on the system where the failure is occurring. For the map in the previous example, you could issue:

```
mount luna:/home/luna /mnt
```

to mount the file system on **/mnt**. Listing the contents would show the user's files. If the map file information is incorrect, use the **spchuser** command from the control workstation to update the map file entry for the user. For example, if the test user's home directory moves from *luna* to *starship*, you would issue:

```
spchuser home=starship:/home/starship/test test
```

This will update the automount map file. You must then wait for as long as five minutes with no access attempts to the **/u/test** directory. This will allow the automount daemon to time out any old access attempts to the previously mounted **luna:/home/luna** exported file system. Do **not** attempt to force the unmount by manually issuing the **unmount** command on the previously mounted file system. This will put the automounter daemon in an inconsistent state and you will need to stop and restart the daemon to recover access to the **/u/test** directory.

Stopping and restarting automount

If you have determined that you need to stop and restart the automount daemon, the cleanest and safest way is to reboot the system. However, this may not be desired due to other processes currently running on the system. If you cannot reboot the system, follow these steps:

1. Determine whether any users are already working on the directories mounted by the **autmountd** daemon. Issue: **mount**

   If automounter is controlling any file systems, you will see an entry similar to:

   ```
   /etc/auto/maps/auto.u /u  autofs Aug 07 11:16 ignore
   ```

   Also, if a user is working in a directory mounted by the atumounter, you will see an entry similar to:

   ```
   luna.pok.ibm.com luna.pok.ibm.com:/home/luna/test /u/test nfs Aug 10 10:37
   ```

   You can request the user to either logoff, or **cd** out of their home directory so that the directory will no longer be in use. If any directory managed by the automounter is accessed while the daemon is stopped, the file system may hang.

2. Stop the **automountd** daemon with this command:

   ```
   stopsrc -g autofs
   ```

   Note that it is important that you **DO NOT** stop the daemon with the **kill -kill** or **kill -9** command. This may cause file system hangs and force you to reboot your system to recover those file systems.

3. Restart the automounter

   ```
   /etc/auto/startauto
   ```

   You can verify that the daemon is running by issuing the previous **lssrc** command.

Chapter 29. Verifying System Management installation

PSSP includes verification tests you can run to check installation of this software. You must be the **root** user to invoke the verification test script. You can invoke the system management verification test on the control workstation from **SMIT** by selecting SSP System Management on the RS/6000 SP Installation/Configuration Verification menu.

Alternatively, you can invoke it from the command line by entering:

```
/usr/lpp/ssp/bin/SYSMAN_test
```

When you invoke the verification test from the control workstation, it uses the **dsh** command to run on all responding nodes. **SYSMAN_test** issues a message if one or more nodes are not tested. You can test a single node by issuing the script on that node directly by local login or using **dsh**.

Verification test output

SYSMAN_test has three output modes: normal, verbose, and quiet. In all cases the full log of test activity is stored in **/var/adm/SPlogs/SYSMAN_test.log** (or a user-specified alternative log file). Each node's log file is stored locally on the node. In normal mode, the test displays all error messages and a summary message reporting success or failure, grouped by node. In verbose mode, the output includes all information recorded in the log files. In quiet mode, only the success or failure message appears. You must examine the log files to see the errors that occurred.

The system management verification SMIT log that follows contains a sample log showing one failing node.

```
HOST: tserv11.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: 0037-031 The xntpd daemon is not running, but the ntp option was configured
SYSMAN_test: Verification failed. The number of errors found was 1

SYSMAN_test: Executing test on all active nodes

r05n11.hpssl.kgn.ibm.com: r05n11.hpssl.kgn.ibm.com: Connection timed out
dsh:  5025-509 r05n11.hpssl.kgn.ibm.com rsh had exit code 1
HOST: r05n01.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n03.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n05.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: 0037-005 Required ntp entry missing from /etc/services
SYSMAN_test: 0037-002 File /etc/rc.ntp does not exist
SYSMAN_test: 0037-003 Directory /etc/auto does not exist
SYSMAN_test: 0037-003 Directory /etc/amd does not exist
SYSMAN_test: 0037-028 /u should not be linked when using automounter
SYSMAN_test: 0037-031 The automount daemon is not running, but the Automount option was configured
SYSMAN_test: 0037-005 Required supfilesrv entry missing from /etc/services
SYSMAN_test: 0037-002 File /var/sysman/supper does not exist
SYSMAN_test: 0037-002 File /var/sysman/file.collections does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/sup does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/supfilesrv does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/supscan does not exist
```

```
SYSMAN_test: 0037-003 Directory /var/sysman/etc does not exist
SYSMAN_test: 0037-003 Directory /var/sysman/logs does not exist
SYSMAN_test: 0037-003 Directory /var/sysman/sup does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/nite does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/sum does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/fiscal does not exist
SYSMAN_test: 0037-002 File /var/adm/acct/nite/jobcharge does not exist
SYSMAN_test: 0037-005 Required acct/startup entry missing from /etc/rc
SYSMAN_test: 0037-004 Directory /var/adm/acct was not exported
SYSMAN_test: Verification failed. The number of errors found was 21

HOST: r05n07.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n09.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n10.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n12.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n13.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n14.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n15.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n16.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded
SYSMAN_test: The number of nodes that were not tested is 1
SYSMAN_test: The total number of errors found was 22
```

What system management verification checks

This test is intended to verify that the PSSP software on the control workstation and nodes and the configuration of the system management facilities completed normally. Since those activities, for the most part, do not result directly in the establishment of testable user interfaces, the verification test concentrates on verifying that various objects relating to node installation and configuration are in the expected state.

The following tables show the key objects that are tested by **SYSMAN_test**. Some objects need to be tested only on the control workstation, some on boot/install servers, some on **/usr** servers, some on **/usr** clients, and others on all SP nodes.

Objects tested by SYSMAN_test on the control workstation only

Table 71. Objects tested by SYSMAN_test on the control workstation only

Object	Verification
CMI SMIT stanzas	Defined in ODM
/etc/SP (directory)	Created
install images	Exported
/.klogin	Entry for each node
nfs (daemon)	Active

Objects Tested by SYSMAN_test on the control workstation and boot/install servers

Table 72. Objects Tested by SYSMAN_test on the control workstation and boot/install servers

Object	Verification
/etc/services	tftp, bootpc, bootps
/etc/inetd.conf	tftp, bootps, instsrv
netinst user	Defined, home directory and files
install image	Exported
bootp_response	Set to disk for clients
tftpd	Daemon running
install_info	File exists for each client
config_info	File exists for each client
net.image file	Link exists for each client
/.klogin	Entry exists for client
/etc/bootptab	Entry exists for client
/etc/exports	Client access to images
install image	Each client's image available

Objects tested by SYSMAN_test on all SP nodes (not the control workstation)

Table 73. Objects tested by SYSMAN_test on the SP nodes

Object	Verification
Node number	Exists in ODM
server_name file	Exists
/etc/hosts	Contains server_name
/.klogin	Contains control workstation and its boot/install server entries

Objects relating to optional system management services

These are objects relating to Optional System Management Services that are tested on all SP nodes and the control workstation.

Table 74. Optional system management objects tested by SYSMAN_test

Object	Verification
/etc/ntp.conf	Exists, content matches config option
xntpd daemon	Running or not
supfilesrv daemon	Running or not
crontab	Set for filec updates or not
/etc/inittab	Set up for filec or not
/etc/services	Set up for filec or not
supman	User defined for filec or not
Predefined file collections	Exist or not
Admin SMIT stanzas	Installed in ODM or not
Svcs.daemons	Contains entries for selected options
Automounter maps	In user.admin if automounter used
Automount	Daemon running or not
sp_passwd	Linked from /bin/passwd or not
pmswitch	Links from print commands or not
Print id	Defined and in password file or not
pmbec	Permissions set appropriately if exists
acct	Directories and files set up or not
crontab	Accounting entries added or not
Jobcharge	File matches SDR attribute if required
/etc/exports	Mount and root access for acct master if required

Additional tests

The following table shows the additional tests that are performed to verify the functionality of System Management services.

Table 75. Additional system management tests

Service	Verification
ntp	Issue the **xntpdc -l** command
Automount daemon	Issue the **ps -e | grep automount** command
filec	Issue a **supper where** command
Ethernet and token ring adapter addresses	Check for valid IPv4 address format

Interpreting the test results

If you find that all information was entered correctly, the tests that **SYSMAN_test** performs should never fail. If errors are reported in the log files, first review the log files created by the various installation steps and use the SP Hardware Perspective as recommended to review the hardware configuration information. Use the SMIT menus that display SP configuration and customization data.

If you find no erroneously entered information or incorrect procedures, you should report the error to the IBM Support Center. This may be as an error in either the installation and configuration process, or in the test itself (the error indication could be spurious). In some cases, for example a missing directory or file, you may want to try to manually correct the problem as a circumvention. In all cases, it should be reported to the IBM Support Center as a possible defect.

Chapter 30. Diagnosing Perspectives problems on the SP System

This chapter discusses diagnostic procedures and failure responses for the SP Perspectives component of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 515. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 514.

SP Perspectives is a set of graphical user interfaces that allow the user to monitor and control their SP system. The set of perspectives include the:

- SP Hardware Perspective (**sphardware**)
- SP Event Perspective (**spevent**)
- IBM Virtual Shared Disk Perspective (**spvsd**)
- System Partitioning Aid Perspective (**spsyspar**)
- Although not a perspective application, the **spled** application displays the SP Nodes LED/LCD values and is launched from some perspectives.

Related documentation

The following publication provides information about SP Perspectives:

- Perspectives online help

  This is available from each SP Perspective individually, or from the SP Perspectives launch pad. This online help discusses each perspective in detail, and is the primary source of information on the perspective.
- *PSSP: Managing Shared Disks*
- *PSSP: Administration Guide*

  This book contains a chapter on SP Perspectives. It also has chapters on dependent subsystems, as listed in “Requisite function” on page 514 (Event Management, Problem Management, SDR, **rsh**, **dsh**, IBM Virtual Shared Disk subsystem).
- *PSSP: Installation and Migration Guide*

  This book explains which file sets are to be installed for each of the SP Perspectives.
- *PSSP: Command and Technical Reference*

  This book contains entries for the following commands used to invoke the SP Perspectives:
  - **sphardware**
  - **spevent**
  - **spvsd**
  - **spsyspar**
  - **perspectives**
  - **spled**
- *SP Perspectives: A New View of Your SP System*, SG24-5180-00
- *RSCT: Event Management Programming Guide and Reference*
- *RS/6000 SP Monitoring: Keeping it Alive*, SG24-4873
- *The RS/6000 SP Inside Out*, SG24-5374

Requisite function

This is a list of the software and operating system resources directly used by the SP Perspectives component of PSSP. Problems within the requisite software or resources may manifest themselves as error symptoms in SP Perspectives. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with SP Perspectives, you should consider the following components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

SP Perspectives has dependencies on several subsystems and commands in the SP System:

- SDR
- System Monitor
- Event Management - including several resource monitors
- Problem Management
- IBM Virtual Shared Disk
- **rsh** and **dsh** commands
- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

Programming Error Message - A programming error message is displayed when a perspectives application catches an exception that was not expected. The error message lists the exception that is thrown, and what method it was thrown from. This message is displayed in a message dialog or in the window where the perspective was started.

Information to collect before contacting the IBM Support Center

Collect this information before calling the IBM Support Center:

1. The PSSP level installed.
2. The current PSSP PTF level
3. The authentication method in use. Issue the **splstdata -p** command to obtain this information.
4. A list of credentials that the user has for authorization of the respective component (**klist** or **k4list**).
5. A detailed description of what actions and events led to the error.
6. The circumstances under which the error occurred. Were you installing a new level, updating from a PTF, migrating the control workstation or a node? This will help to recreate and debug the problem. Supply the proper information if necessary. An example is the AIX levels you were migrating from and migrating to.
7. If the problem can be re-created, use the **-debug_script** flag when starting the perspectives application. This prints environment variables and other information that the perspectives application uses to stdout.
8. If the problem results in a core dump, supply the core file, along with the **dbx** trace information. To get the **dbx** trace information, issue the following commands in the directory where the core file is located:

```
/bin/dbx/usr/lpp/ssp/perspectives/bin/perspective_name
```

where *perspective_name* is the name of the perspective you were running at the time of the failure.

9. Set the environment variable **ICLUI_TRACE=NOPREFIX** and rerun the failing perspective. Gather the information that is displayed to stdout.
10. The profile being used for the perspective.
    - If using the default system profile, save the file **/usr/lpp/ssp/perspectives/profiles/***$LANG/.perspectives_name***Profile**, where *$LANG* is the value of the **LANG** environment variable and *perspectives_name* is the name of the perspective invoked.

      For example, if you are running the hardware perspective in the **en_US** locale, the file to save is: **/usr/lpp/ssp/perspectives/profiles/en_US/.sphardwareProfile**.
    - If you have a default user profile, which would have the same name as that of the default system profile, it is used if no other profile is specified. This profile is in the **$HOME** directory. For example, if you are **root**, the default user profile is: **/.sphardwareProfile**.
    - If you specify a profile on the command line, or through the perspectives launch pad, user profiles are found in the user's **$HOME** directory and the system profiles are found in **/usr/lpp/ssp/perspectives/profiles/***$LANG/profile_name*.

Diagnostics

Use these procedures to diagnose problems with SP Perspectives.

1. If you have problems running a perspective using a modified resource file or profile, such that the perspective is not displayed, then revert back to using the default system profile and resource file. To use the default profile, do not specify a profile name.
2. If you do not see the perspectives splash window displayed, this may indicate that there is a problem in running the splash program. Try running the perspective without displaying the splash window. For example:

   ```
   sphardware -nosplash
   ```
3. To eliminate problems related to a bad user or system profile, specify the **-noProfile** flag when starting the perspective. For example:

   ```
   sphardware -noProfile
   ```

Error symptoms, responses, and recoveries

Perspectives problems may arise in one of the following categories. See the accompanying table to diagnose these problems.

Table 76. SP Perspectives symptom types

Type of Problem	Table to Reference
General problems running the Launch Pad or any of the Perspectives	Table 77 on page 516
General problems running the SP Hardware Perspective	Table 78 on page 516
General problems running the IBM Virtual Shared Disk Perspective	Table 79 on page 516
General problems running the SP Event Perspective	Table 80 on page 517

Table 77. Launch pad and general Perspectives symptoms

Symptom	Recovery
Perspectives command not found (**perspectives**, **sphardware**, **spevent**, **spvsd**, or **spsyspar**).	See “Action 1 - Verify SP Perspectives installation” on page 517.
The Launch Pad or Perspectives fails to come up.	See “Action 2 - Export the DISPLAY environment variable” on page 518.
The Launch Pad or Perspectives terminates prematurely.	See “Information to collect before contacting the IBM Support Center” on page 514 and contact the IBM Support Center.
The Launch Pad or Perspectives hangs.	See “Action 8 - Check performance of the system” on page 520.
During startup, you receive a message stating that you cannot run this application directly.	See “Action 9 - Run Perspectives from /usr/lpp/ssp/bin” on page 520.
During startup, you receive a message indicating lack of access to dependent subsystems.	See “Action 3 - Obtain Access to dependent subsystems” on page 518.
When trying to open the SP Perspectives online help, you receive the following message: ″The requested online help is either not installed or not in the proper search path. The Help Volume is: **Help4Help**, Location ID: **QUICK-HELP**″.	See “Action 11 - Install the file set needed for Perspectives online help” on page 520.
When displaying SP Perspectives to a non-AIX X-server, you get a core dump. Examining the trace output, you determine that the default Perspective font is not found on the X-server you are displaying perspectives to.	See “Action 12 - Override the default fontList resource variable” on page 520.
Actions are not available for objects in SP Perspectives.	See “Action 3 - Obtain Access to dependent subsystems” on page 518.
During startup, you receive a message that you do not have access to one or more subsystems	See “Action 3 - Obtain Access to dependent subsystems” on page 518.
You receive a message that the connection to the Event Manager was lost. If you were monitoring, all icons now have question marks indicating an unknown state.	See “Action 5 - Check the Event Manager daemon” on page 518.

Table 78. SP Hardware Perspective symptoms

Symptom	Recovery
If the • Node Notebook Environment page is blank or the notebook status pages attributes such as Power, Controller responds, Node failure, Environment LED for nodes, switches, or frames have question marks. • Switch Notebook Status page is blank. • Frame Notebook Status page is blank.	See “Action 6 - Check the Resource Monitors” on page 519.
The Node notebook tabs: Node Status, Node Environment, Hardware Resource Variables, CSS Resource Variables, AIX OS Resource Variables and All Dynamic Resource Variables are grayed out (not selectable).	You do not have access to the Event Manager. See “Action 3 - Obtain Access to dependent subsystems” on page 518.

Table 79. IBM Virtual Shared Disk Perspective symptoms

Symptom	Recovery
You are not able to create IBM Virtual Shared Disks or IBM Hashed Shared Disks (HSDs).	See “Action 10 - Prepare disks for the createvsd or createhsd commands” on page 520.

Table 79. IBM Virtual Shared Disk Perspective symptoms (continued)

Symptom	Recovery
IBM Virtual Shared Disk information in notebooks is not being updated. For example, information in notebooks and the table view is not being update.	See “Action 6 - Check the Resource Monitors” on page 519.
SDR information is not being updated automatically. For example, information in notebooks and the table view is not being updated.	See “Action 6 - Check the Resource Monitors” on page 519.
Newly configured IBM Virtual Shared Disks and Hashed Shared Disks or IBM Virtual Shared Disk state changes are not being updated in the IBM Virtual Shared Disk Perspective.	See “Action 6 - Check the Resource Monitors” on page 519. Check the **IBM.PSSP.harmld** resource monitor.

Table 80. SP Event Perspectives symptoms

Symptom	Recovery
You receive a message in a dialog box that the connection to the Event Manager was lost. The Event Perspective closes when you press the OK button.	See “Action 5 - Check the Event Manager daemon” on page 518.
Buttons and fields of the Actions page of the Event Definition Notebook are not selectable.	You need access to the Problem Management subsystem. See “Action 3 - Obtain Access to dependent subsystems” on page 518.

For more detail on performing various actions, see the chapter on using the SP Perspectives in *PSSP: Administration Guide*.

Note: If the actions in these tables do not resolve your problem, determine if you can re-create the problem. If you can, see “Information to collect before contacting the IBM Support Center” on page 514 and call the IBM Support Center for further assistance.

Actions

Action 1 - Verify SP Perspectives installation

To ensure that SP Perspectives subcomponent was installed properly, perform these steps:

1. Ensure that the following file sets were installed as options to the **pssp.installp** image.

Table 81. Perspectives - related file sets

Command	Perspective	Required file sets
perspectives	Launch Pad	**ssp.gui**
sphardware	SP Hardware Perspective	**ssp.gui**
spevent	SP Event Perspective	**ssp.gui**
spvsd	IBM Virtual Shared Disk Perspective	**ssp.csd.gui** **ssp.gui**
spsyspar	System Partitioning Aid Perspective	**SSP.top** **ssp.gui**

For more information, see *PSSP: Installation and Migration Guide*.

2. Add **/usr/lpp/ssp/bin** to you **PATH** environment variable, or use the full path name with the command. The full path name of a perspectives command is: **/usr/lpp/ssp/bin/***command-name*.

Action 2 - Export the DISPLAY environment variable

In order to display Perspectives from the control workstation to another machine, do the following:

1. In your login window to the control workstation, export your **DISPLAY** variable to the X-server of the machine where you want SP Perspectives displayed.
2. On the machine where you want to display perspectives to, issue the command:

   ```
   xhost + control_workstation_name
   ```

   This enables windows to be displayed from the control workstation to that machine.

Action 3 - Obtain Access to dependent subsystems

If you do not have access to required subsystems, their functions may not be available in the SP Perspective. To find out what access is required for each action in each SP Perspective, see the perspectives online help Security section. The online help can be started from the Perspectives Launch Pad. You can also start the Perspectives online help by issuing the command:

```
/usr/dt/bin/dthelpview -helpVolume /usr/lpp/ssp/perspectives/help/$LANG/pgui.sdl
```

To find out how to obtain access to specific subsystems, see *PSSP: Administration Guide*.

Action 4 - Check SP Hardmon control access

Some SP Perspectives require monitor, serial, or control access to the hardmon subsystem. See the Security chapter of *PSSP: Administration Guide* to determine what security mechanism is in use on your SP system, and how to obtain the specific hardmon access for that security mechanism.

Action 5 - Check the Event Manager daemon

If you receive messages that SP Perspectives has lost its connection to the Event Manager, the problem could be that the Event Manager daemon, **haemd** has terminated. The problem could also be that the network connection to the Event Manager daemon was lost. Perform the following steps:

1. Check that the Event Manager daemon is up and running by issuing the command:

   ```
   lssrc -a | grep haem
   ```

   This lists the Event Management daemon. If the system is partitioned, the daemon will be listed for each system partition.
2. If any **haem** daemon is listed as inoperative, restart the daemon by issuing the command:

   ```
   startsrc -g haem
   ```
3. Issue the command:

   ```
   lssrc -a | grep haem
   ```

again to verify that the Event Manager daemon is now up and running.

Action 6 - Check the Resource Monitors

Resource monitors are software components that provide resource variables to the Event Manager daemon. Here are some examples of resource variables and the resource monitors that supply them to the Event Manager daemon:

Table 82. SP Perspectives resource variables

Resource variable name	Resource Monitor
IBM.PSSP.Response.Host.State	Response
IBM.PSSP.Response.Switch.State (Used for SP Systems with an SP Switch)	Response
IBM.PSSP.SwitchResponse.state (Used for SP Systems with an SP Switch2)	IBM.PSSP.Switch
IBM.PSSP.SP_HW.Node.powerLED	IBM.PSSP.hmrmd
IBM.PSSP.SP_HW.lcd1	IBM.PSSP.hmrmd
IBM.PSSP.aixos.FS.%totused	aixos

To check if any of the resource monitors are down or locked, issue the command:

```
lssrc -ls haem.syspar_name
```

for each system partition.

For example, if you have two system partitions named **k4s** and **k4sp1**, you would issue the command: **lssrc -ls haem.k4s** to check the resource monitors in the first system partition, and: **lssrc -ls haem.k4sp1** to check the resource monitors in the second system partition.

The listing of **haem** information is similar to the following:

```
Resource Monitor Information

Resource Monitor Name    Inst  Type   FD     SHMID      PID     Locked
IBM.PSSP.CSSLogMon        0      C    -1      -1         -2    00/00 No
IBM.PSSP.SDR              0      C    -1      -1         -2    00/00 No
IBM.PSSP.Switch           0      S    21      -1      20268    01/01 No
IBM.PSSP.harmld           0      S    23       7      25738    01/01 No
IBM.PSSP.harmpd           0      S    19      -1      20976    01/01 No
IBM.PSSP.hmrmd            0      S    22      -1      22136    01/01 No
IBM.PSSP.pmanrmd          0      C    15      -1         -2    00/00 No
Membership                0      I    -1      -1         -2    00/00 No
Response                  0      I    -1      -1         -2    00/00 No
aixos                     0      S    12       5         -2    00/01 No
```

If any resource monitors are locked, the entry in the `Locked` column will be `yes`. Issue the command:

```
haemunlkrm
```

to unlock the resource monitor.

For example, if the hardmon resource monitor (IBM.PSSP.hmrmd) is locked, issue the command:

```
/usr/sbin/rsct/bin/haemunlkrm -s haem -a IBM.PSSP.hmrmd
```

Also note that the **aixos** resource monitor provides variables which are retrieved through the Event Management daemons running on the node. For example, if you are getting question marks for the Node notebook attribute named CPUs online, this may mean that the **aixos** resource monitor is not running on the node.

Use the **telnet** command to access that node and issue the command:

```
lssrc -g haem
```

If either the **haem** or **haemaixos** subsystem is not active, then issue the command:

```
startsrc -g haem
```

to restart the daemons.

Action 7 - Check for a core dump

Check the directory you are running from for a file named **core** with a current timestamp. If the file exists, save it. See "Information to collect before contacting the IBM Support Center" on page 514 and contact the IBM Support Center.

Action 8 - Check performance of the system

Check the overall CPU utilization of the control workstation to see if any processes are consuming a large amount of time. See if any of the following are consuming a large amount of CPU time:

- The Perspectives processes (**perspectives**, **sphardware**, **spevent**, **spvsd**, or **spsyspar**).
- The underlying subsystems: **hardmon**, **s70d**, **hmcd**, **sdrd**, **haemd**, or **hagsd**

If any of these processes are continually consuming a large amount of CPU time, they may need to be brought down and restarted.

Action 9 - Run Perspectives from /usr/lpp/ssp/bin

You tried to run one of the Perspectives commands: **perspectives**, **sphardware**, **spevent**, **spvsd**, or **spsyspar** from **/usr/lpp/ssp/perspectives/bin**.

These executables need specific environment variables set in order to run correctly. These executables should not be run directly. Run the Perspectives command from the **/usr/lpp/ssp/bin** directory, or add **/usr/lpp/ssp/bin** to your **PATH** variable.

Action 10 - Prepare disks for the createvsd or createhsd commands

See *PSSP: Managing Shared Disks*.

Action 11 - Install the file set needed for Perspectives online help

In order for the SP Perspectives online help information help window to be available, the **X11.Dt.helpinfo** file set must be installed. For information on installing Perspectives, see *PSSP: Installation and Migration Guide*.

Action 12 - Override the default fontList resource variable

The SP Perspectives applications retrieve their font in the following priority:

1. SP Perspectives first searches for a font description specified in a profile. This profile can be specified on the command line using either one of the flags:
   - -systemProfile
   - -userProfile

The **-systemProfile** searches for the profile in the directory, **/usr/lpp/ssp/perspectives/profiles/$LANG**. The **-userProfile** searches for the profile in the **$HOME** directory.

If no profile is specified at startup, a default profile is used which is either a system profile or user profile. If the **userProfile** exists, it overrides the system default profile. Examples of the default user and system profile for the SP Hardware Perspective, **sphardware**, for the **root** user are:

- /.sphardwareProfile
- /usr/lpp/ssp/perspectives/profiles/en_US/.sphardwareProfile

Note: The default system profiles do not specify a font. However, if you have authority to write to the **/usr/lpp/ssp/perspectives/profil es/$LANG** directory, it is possible to overwrite the default system profile used at startup.

The font resources are saved in a profile using four specifications:

- font.family:
- font.pointSize:
- font.italic:
- font.bold:

This is an example of using a non-default SP Hardware Perspective user profile, example1, which is saved in the file, **$HOME/.sphardwareexample1**. Issue the following command:

```
spharwdare -userProfile example1
```

If this profile contains the four font resources listed previously, and the specification for that font exists on the X11 server you are displaying the application to, that font is used. If the four font resources do not match a font found on the X11 server, the default font for that X11 server is used.

2. If a font is not specified in a profile, the font specified in the application resource file is used. The SP Perspectives application resource files are found in the directory**/usr/lpp/ssp/perspectives/app-defaults/$LANG** and are specified as standard X11 resources. For example, the default font specification for the en_US locale for the SP Hardware Perspective is found in the file, **/usr/lpp/ssp/perspectives/app-defaults/en_US/Sphardware** and, is specified as:

```
Spharware*fontList: -misc-fixed-medium-r-normal--*-110-100-100-c-*-iso8859-1
```

The SP Perspectives applications force the use of the **XAPPLRESDIR** environment variable to specify where the application resources are found. You can override these values in your own **.Xdefaults** file with the same specification of **Sphardware*fontList**, but specifying your choice of a font found on their X11 server.

To determine which fonts are available on the X11 server that displays the SP Perspectives application, you can use the X11 program, **/usr/bin/X11/xlsfonts**, found on the control workstation. Make sure that the **DISPLAY** environment variable is set to the workstation that you want to display the SP Perspectives application.

3. If the font specified in a profile, and the font specified through X11 resources, are both not found on the X11 server that displays the SP Perspectives

application, the default font for your X11 server is used. If you still have problems obtaining a font for your X11 server, it is recommended that you use the **xlsfonts** program to find a font on your X11 server, and specify this font in your **.Xdefaults** file.

Chapter 31. Diagnosing file collections problems on the SP System

SP file collections is that subsystem of PSSP responsible for file distribution within the SP system. You must request it to be installed via the SP Site Environment SMIT panels. IBM ships predefined collections that handle the administrative files of File Collections, user administration if you have chosen to use SP User Management, and node administration. You can add files to these collections in order to have them distributed throughout the SP system.

Related documentation

The following publications provide information about file collections:

- *PSSP: Administration Guide*
- *PSSP: Messages Reference*

  Messages related to file collections are in ″0018 - supper Messages″..
- *PSSP: Command and Technical Reference*
  - **supfilesrv** daeamon - Daemon that serves the file collections on the SP system
  - **supper** - Manages the SP file collections.
  - **filec_host**
- *Inside the RS/6000 SP*, the chapter on File Collections
- *AIX 5L Version 5.1 Command Reference*
  - **cron** man page - runs commands automatically at specified dates and times
  - **crontab** man page - Submits, edits, lists, or remove cron jobs

Requisite function

This is a list of the software directly used by the file collections component of PSSP. Problems within the requisite software may manifest themselves as error symptoms in file collections software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the file collections component of PSSP, you should consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

1. File collections must be chosen to be installed by the customer.
2. In order for file collections to work properly, the subsystem depends on proper AIX and network configuration.
3. Hostname resolution must be functioning. A route must exist from the node to the file collection server. This is either the control workstation or the node’s boot/install server if you have chosen to configure a boot/install server. The network must be functioning properly. The **cron** subsystem must be functioning properly.
4. During SP installation, the necessary configuration is done for routes and hostnames by the Install and Configure scripts.
5. During File Collections configuration on each node, the **/usr/lpp/ssp/install/bin/filec_config** script adds an entry to the **root** crontab file to kick off automatic updates to the collections residing on that node. If the cron subsystem is taken off-line or problems occur, the collections on that node are not updated.

6. File Collections depends on the configuration script being run on the control workstation, any boot/install servers and nodes during installation and possibly customization. Should errors occur during SP install or customization, then you may experience File Collection errors or File Collections may not even be installed.
7. File Collections does not take advantage of the authentication mechanisms on the SP system other than its internal checking of AIX hostnames and its specific user id.

SP configuration

During configuration of File Collections, installed files are copied from **/usr/lpp/ssp/filec** to their proper location in directories **/var/sysman** and **/usr/lpp/ssp/bin**. If these files are updated during the application of a PTF, the install script ensures that the files in **/var/sysman** and **/usr/lpp/ssp/bin** are updated. However, if an error occurs during this processing, the files may not be updated properly.

IBM ships and installs the following files:

- **supfilesrvr**
- **sup**
- **supscan**
- **filec_config**
- **filec_host**
- **supper**
- **file.collections**
- **lists**
- Set of data files in each collection of **node.root**, **user.admin**, **sup.admin**, **power_system**:
  - **list**
  - **lock**
  - **prefix**
  - **refuse**
  - **supperlock**

Other files are created during configuration. If SP system configuration errors occur, these files may not be created.

- **supfilesrv.pid**
- Set of files in each collection of: **node.root**, **user.admin**, **sup.admin**, **power_system**:
  - **host**

This list is for the current level of PSSP. Some files may not exist for older levels of PSSP.

On a node, File Collections relies on the **/etc/mastername** command to supply the boot/install server hostname. The boot/install server could be the control workstation or a node which has been configured as a boot/install server. The **/etc/mastername** command, in turn, relies on the existence of the **/etc/ssp/server_name** file. This file is updated when necessary by the SP Installation and Configuration code. This is the method used by File Collections to obtain the File Collections server name when running on a node.

Normally, configuration errors can be corrected by running **/usr/lpp/ssp/bin/services_config**. The **services_config** script calls **filec_config** with the input parameters necessary to configure File Collections.

Error information

File Collections creates two log files. These files exist only on the nodes and, if configured, any boot/install servers. The files exist on each local host and are not consolidated in any way. These log files are:

- **/var/adm/SPlogs/filec/sup***date.time*. This log contains minimal status information of the last time the **supper** command was run via a cron job. The contents of this log are
  - Date and time when supper was last run
  - Hostname
  - Number of files that were updated during that run
- **/var/adm/SPlogs/filec/sup***date.time***r**. This log contains the client messages logged as the supper command attempted to update the collections. This is the file that can be scanned for error messages, if the client had problems getting a collection.

A directory, **/var/sysman/logs**, is created but no longer used.

File Collections creates a status file on the server hosts where the supfilesrv daemon runs. The status file contains the pid number of the supfilesrv daemon. The file is named: **/var/sysman/sup/supfilesrv.pid**.

Note: If you are using the **supper** command from the command line and are also logging the output, be aware that you can cause an overwrite of a previous log. This is done by issuing the **supper** command multiple times within one minute. You will see only the information produced by the last command.

Displayed and logged information is taken from catalogs and may be translated. In order for these logs to be readable by a system administrator, it is recommended that either the English locale or the SP administrative locale be used.

To determine the SP administrative locale, issue this command:

```
splstdata -e
```

and look for the **admin_locale** value.

Information to collect before contacting the IBM Support Center

Before calling the IBM Support Center, it is your responsibility to verify that the environment is correctly configured for File Collections. This includes verifying that any changes made to files specified by a collection were made on the master copy. The master is located on the control workstation for the SP File collections. Any changes made to files specified by a collection on a node will be overwritten on the next **supper** update.

If you have boot/install servers configured in your SP system, be aware that file distribution happens in two steps. First, the files on the boot/install server are updated from the control workstation. Then, the nodes are updated from their boot/install server. If you have changed a master copy of a file on the control workstation, and then issue **supper** on the node, the node will receive the old copy of the file from the boot/install server. Either wait for the normal update to occur, or be sure to issue the **supper** command on the boot/install server first, and then on the node, to obtain the updated copy of the master file on the control workstation.

Updates occur every hour through file collections based on an entry in cron. If you need updated copies of changed master files distributed before that time, you can issue a **dsh** command from the control workstation to the nodes to run the **supper** command. See the previous paragraph if you have defined boot/install servers on your SP system.

If you have changed the AIX hostnames of any host in your SP system, changed an IP address of a hostname, or changed the primary adapter of a host, consult *PSSP: Administration Guide* for instructions on updating the PSSP software to recognize the new AIX hostname, IP address, or adapter.

Run through the following diagnostic steps to further check the configuration of the environment for File Collections. If you are still having problems after this verification, collect the following information for the server (control workstation and boot/install servers if configured) and the client (usually a node) for use by the IBM Support Center:

1. The PSSP level installed.
2. The current PSSP PTF level.
3. A description of the error.
4. Diagnostic actions taken and the affect of those actions.
5. Whether any of the data files that describe a collection were modified to add customer files for distribution.
6. Whether you are using SP User Management.
7. Whether you have configured a boot/install server for the node where the error occurred.
8. The circumstances under which the error occurred. Were you installing a new level, updating from a PTF, migrating the control workstation or a node? Was File Collections running normally and then an error occurred? This will help to recreate and debug a problem.

Diagnostics

File Collections errors are divided into three types:

1. File Collections errors on the control workstation (the master server), which usually affect distribution of files across the entire SP system.
2. File Collections errors on a subset of nodes, which usually points to the boot/install server for those nodes as the source of the problem.

   A boot/install server is a File Collections client of the File Collections server running on the control workstation. A boot/install server is in turn the File Collections server for the File Collections client running on each node.
3. File Collections errors on a client (usually a node), which only affect updates of the collections, one collection, or a single file on that node.

One of the first diagnostic actions is to login to the node where a problem has occurred. Verify that a collection can be updated by running the **supper** command manually, either interactively or through the command line.

- **Interactive mode** is used by issuing the **supper** command without operands.

You will see the message `Supper Collection Maintenance -— type 'help' for a list of commands`. This is followed by the **supper** prompt line, `supper>`. By entering **help**, a list of supper options is displayed, along with a brief description of each option. For example, the **when** option shows the last update time of all resident collections. You remain in Supper Collection Maintenance mode until you enter the **quit** option.

Perform these steps to see if the collection update is taking place:

1. Touch or update a master file on the control workstation or boot/install server to ensure that an update will occur.
2. On the node that is being checked, issue the **supper** command.
3. Note that you are in interactive mode. Issue the **update** option.
4. The update message for each file collection is displayed. This is an example:

   ```
   Updating collection sup.admin from server
   c180cw.ppd.pok.ibm.com

   Files changes:  1 updated, 0 removed, 0 errors.
   ```

   This example shows that one master file on the control workstation was touched or changed, and the update was done on the node.

- **Command line entry** is done by issuing the **supper** command with the option you desire. These steps perform the same test as the preceding one, using a command line entry instead of interactive mode:
  1. Touch or update a master file on the control workstation or boot/install server to ensure that an update will occur.
  2. On the node being tested, issue the **supper update** command.
  3. The output is the same as for the preceding interactive command.

File collections errors on any SP host

Answer these questions and follow these steps to debug file collections errors on an SP host system:

1. Has File Collections been chosen for configuration on the SP system?

   To configure File Collection Management on the SP system, the Site Environment variable, **filecoll_config**, must be set to **true** through the **SMIT** SP Site Environment panel, or from the command line by issuing the command:

   ```
   spsitenv filecoll_config=true
   ```

   If you make this change after installation of your system, you are required to customize the nodes for this change to take effect.
2. Does the **/etc/services** file list the port number for the File Collections **supfilesrv** daemon?

   There should be an entry similar to: `supfilesrv 8431/tcp`.
3. Is the network up and running?

   Verify connectivity between the node and the node's boot/install server. Normally, File Collections uses the internal SP connection.
4. Has a new adapter been installed? Has the hostname changed? Has a different adapter been configured as the primary?

   Consult *PSSP: Administration Guide* for instructions on updating the PSSP software to recognize the new AIX hostname, IP address, or adapter.

5. Have you received configuration errors, either directly from File Collections or the Installation/Configuration SP software?

   Most File Collection configuration errors can be corrected by running **/usr/lpp/ssp/install/bin/services_config**. There is no easy way to run **/usr/lpp/ssp/install/bin/filec_config** directly, without specifying a number of environment variables that are already taken care of within **/usr/lpp/ssp/install/bin/services_config**. You should run this on each SP host with File Collections errors.

File collections configuration errors

Answer these questions and follow these steps to debug File Collections configuration errors:

1. Have you received installation errors or errors from applying a PTF?

   If errors occurred during installation or the application of a PTF, this could indicate that the files installed were not copied to the proper place. Depending on where the installation or PTF errors occurred, you may be able to correct the problem by running **/usr/lpp/ssp/install/bin/services_config** on the affected hosts.
2. Are the collections being updated from the wrong server?

   This can happen if a new boot/install server has been defined or an existing one undefined, and the affected nodes have not yet been rebooted. On a node, File Collections obtains the server hostname by use of the **/etc/mastername** command. This command relies on a file on the node. That file is updated when a node has been customized after a configuration change.

File collections server (control workstation or boot/install server) errors

Answer these questions and follow these steps to debug File Collections server errors:

1. Is the **supfilesrv** daemon running on the control workstation and any configured boot/install servers?

   This daemon runs on the control workstation and the boot/install servers. It responds to requests issued on the nodes to update the collections of files configured for File Collections. The **supfilesrv** daemon is under control of the AIX System Resource Controller (SRC). It is normally started as part of the SP initialization scripts. Under normal circumstances, the system administrator does not have to start or stop the daemon.
2. Is the File Collections ID, **supman**, listed in the **/etc/passwd** file?

   File Collections requires a unique, unused ID for **supman**, through which the File Collection daemon, **supfilesrv**, can communicate. The default installation configures the user ID, **supman_uid**, to 102 and the port, **supfilesrv_port**, to 8341. These values could be changed using **SMIT** or the **spsitenv** command. The default user name and uid attributes listed in the **/etc/passwd** file are: **supman** for the user name, and 102 for the uid.
3. Is the password ***** (asterisk) for the File Collections id, **supman**, in the **/etc/passwd** file?

   The **/etc/passwd** file is an ASCII file that contains basic user attributes. The entry for each user is unique. The second field of each entry is the Password field that contains a valid encrypted password, an ***** (asterisk) indicating an invalid password, or and ! (exclamation point) indicating that the password is in the **/etc/security/passwd** file. If the second field has an ***** (asterisk) and a password is required for authentication, the user cannot log in.

4. Is the File Collections id, **supman**, part of the AIX security group?

   The **/etc/group** file is an ASCII file that contains the basic group attributes. Each record is identified by a group name, and contains attributes separated by colons. The fields of interest to File Collections are:

   - First field - Name, which is the unique name of by which a group is known on the system.
   - Third field- ID, which is the ID of the group
   - Fourth field - Users, which list the members of the group

   The file collection daemon, **supfilesrv** requires read access permission to any files that you want managed by file collections. You must add the supman ID to the security group in the **/etc/group** file. This provides **supman** with read access to files that have security group permission and allows these files to be managed across the SP system by file collections. This is necessary to distribute the files in the **user.admin** collection. If you are seeing errors distributing only the **user.admin** collection, check whether the supman ID is part of the AIX security group.

   At the prompt on the control workstation, issue: **id supman** You should see output similar to:

   ```
   uid=102(supman) gid=7(security)
   ```

   If this is not the case, verify the configuration of SP User Management and run **/usr/lpp/ssp/install/ bin/services_config**.

5. Are all the AIX hostnames of the nodes being served by this server (control workstation or boot/install server) in the **host** file in each collection?

   During the configuration of file collections, a per-collection **host** file is created to limit access for the predefined collections to the nodes of the SP system. This **host** file resides in the directory **/var/sysman/***file_collection***/host**, where *file_collection* is the unique name of a file collection.

   The absence of the per-collection **host** file indicates that any hosts are allowed access. SP systems in earlier releases did not have this **host** file, allowing any client from any machine to access file collections.

   If you are using public code, you are responsible for updating and maintaining this **host** file. While some public code documentation exists on the per-collection **host** file, this information is not readily available in IBM documentation. IBM discourages the use of public code commands in favor of our supplied **supper** command and our IBM supplied collections. IBM creates and maintains the public code **host** file. This automatically grants access to the IBM supplied collections, but only to SP nodes by default. See “Notices” on page 597.

   IBM ships a new file, **collection.host.list** which lists the SP-supplied collections. Only the SP-supplied collections will be updated with a **host** file default. If you wish to have your own customer-defined collections in the **/var/sysman/** path updated with a **host** file, the collection names must be placed in this **collection.host.list** file. This file must always be distributed to be present and updated on both the control workstation and any defined boot/install servers.

6. Has the collection’s **list** file been altered on the server?

   When adding or removing a file or directory of files for distribution, it is very easy to produce errors in the **list** file for a collection. For example, if the ″**.**″

(period) is not the first character in the line of a list file, the file or directory may not be found since paths usually start at the **root** level.

You may also see unintended actions. Depending on the error in the **list** file, for example, you could distribute your entire **root** directory to a node.

File collections client errors

1. Is the **cron** daemon running, and is there a File Collections entry in the **root crontab** file?

   An entry is placed in the root crontab file during File Collections configuration. The default is to update the collections once each hour, usually ten minutes after the hour. During peak times when nodes are updating collections, you may see Server Busy messages if you are using the **supper** line command on a node interactively. This may also occur if you reboot an SP system with a large number of nodes all at once.
2. Is a **scan** file present in the collection on a node or the server?

   File Collections uses the **scan** file to prevent building new lists of files to check and update every time it runs. If you have changed a collection to add a file or remove a file from distribution, you need to remove the current **scan** file for each collection, or run the command:

   ```
   supper scan
   ```

   on the control workstation to build or update the **scan** file. It is important that you run the **supper scan** command when you make changes to a collection. This is because the existing **scan** file does not have the new changes, and the changes will not be included in the collection when a **supper update** command is issued from the node.
3. Has the master copy of a file been updated?

   If you do not see a file being updated, or if you see an old copy of a file replacing a new one, verify that the master copy of the file, located on the control workstation, has been updated.

   If a file was updated on a boot/install server and this same file is also distributed to the boot/install server from the control workstation, the updated file will be overwritten by the distributed file.

Error symptoms, responses, and recoveries

The error symptoms that are corrected using the diagnostic steps described previously are not repeated in this section. There are some errors that do not clearly appear to relate to a particular diagnostic action. These error symptoms are included here for reference, and make use of the steps from “Diagnostics” on page 526.

Table 83. File collection symptoms and recovery actions

Symptom	Actions
Permission Denied messages	See “File collections server (control workstation or boot/install server) errors” on page 528, Steps 2 on page 528, 3 on page 528, 4 on page 529 and 5 on page 529.
Server Busy messages	See “File collections configuration errors” on page 528, Step 2 on page 528, and “File collections client errors” Step 1.

Table 83. File collection symptoms and recovery actions (continued)

Symptom	Actions
Old copies of files are distributed	See “File collections server (control workstation or boot/install server) errors” on page 528, Step 6 on page 529, and “File collections client errors” on page 530 Steps 2 on page 530 and 3 on page 530.
Server IP address cannot be resolved	See “File collections errors on any SP host” on page 527, Steps 2 on page 527, 3 on page 527 and 4 on page 527, “File collections configuration errors” on page 528, Step 2 on page 528, and “File collections server (control workstation or boot/install server) errors” on page 528 Steps 1 on page 528 and 2 on page 528.
Specific files in a collection are not updated	See “File collections server (control workstation or boot/install server) errors” on page 528, and “File collections client errors” on page 530. Follow all steps in each section.
More than one or two types of Files Collection errors	See “File collections errors on any SP host” on page 527, Steps 4 on page 527 and 5 on page 528, and “File collections configuration errors” on page 528 Step 1 on page 528.

Part 3. Diagnosing SP node and network attached hardware

Chapter 32. SP-specific LED/LCD values

LED and LCD values generated by the SP system and Parallel System Support Programs are of two types: those that convey information and signal status, or those that indicate a problem. These are the PSSP-specific LED/LCD Values, generated during the NIM installation of the node. This list is displayed in chronological order, or the order in which the they occur during processing.

Note: Some nodes have LED/LCD values that are of the form **uxx**, while others have LED/LCD values of the form **0axx**. They are equivalent. If your node has a value of **0axx**, look it up in the following tables as if it was **uxx**.

For an LED/LCD not in this table, see “Other LED/LCD codes” on page 539.

Table 84. SP-specific LED/LCD values (Chronological order)

u20	Create log directory (enter function create_directories).
u21	Establish working environment (enter function setup_environment).
u03	Get the *node*.install_info file from the master.
u04	Expand *node*.install_info file.
u22	Configure node (enter function configure_node).
u57	Get the *node*.config_info file from the master.
u59	Get the cuat.sp template from the master.
u23	Create/update /etc/ssp files (enter function create_files).
u60	Create/update /etc/ssp files.
u24	Update /etc/hosts file (enter function update_etchosts).
u25	Get configuration files (enter function get_files).
u61	Get /etc/SDR_dest_info from boot/install server.
u79	Get script.cust from boot/install server.
u50	Get tuning.cust from boot/install server.
u54	Get spfbcheck from boot/install server.
u56	Get psspfb_script from boot/install server.
u58	Get psspfb_script from control workstation.
u27	Update /etc/inittab file (enter function update_etcinittab).
u28	Perform MP-specific functions (enter function upmp_work).
u52	Processor is ″MP″.
u51	Processor is ″UP″.

Table 84. SP-specific LED/LCD values (Chronological order) (continued)

u55	Fatal error in bosboot.
u29	Install prerequisite file sets (enter function install_prereqs).
u30	Install ssp.clients (enter function install_ssp_clients).
u80	Mount lppsource and install ssp.clients.
u31	Install ssp.basic (enter function install_ssp_basic).
u81	Install ssp.basic.
u32	Install ssp.ha (enter function install_ssp_ha).
u53	Install ssp.ha.
u33	Install ssp.sysctl (enter function install_ssp_sysctl).
u82	Install ssp.sysctl.
u34	Install ssp.pman (enter function install_ssp_pman).
u41	Configure switch (enter function config_switch).
u35	Install ssp.css (enter function install_ssp_css).
u84	Install ssp.css.
u36	Install ssp.jm (enter function install_ssp_jm).
u85	Install ssp.jm.
u37	Delete master .rhosts entry (enter function delete_master_rhosts).
u38	Create new dump logical volume (enter function create_dump_lv).
u86	Create new dump logical volume.
u39	Run customer’s tuning.cust (enter function run_tuning_cust).
u40	Run customer’s script.cust (enter function run_script_cust).
u87	Run customer’s script.cust script file.
u26	Get authentication files (enter function authent_stuff).
u67	Get /etc/krb.conf from boot/install server.
u68	Get /etc/krb.realms from boot/install server.
u69	Get krb-srvtab from boot/install server.
u42	Run psspfb_script (enter function run_psspfb_script).
u46	/tftpboot/tuning.cust is being run during node customization.

The following list contains the same PSSP LEDs/LCDs as in the previous list, but sorted numerically for reference use:

Table 85. SP-Specific LED/LCD values (Numerical order)

u03	Get the *node*.install_info file from the master.
u04	Expand *node*.install_info file.
u20	Create log directory (enter function create_directories).
u21	Establish working environment (enter function setup_environment).
u22	Configure node (enter function configure_node).
u23	Create/update /etc/ssp files (enter function create_files).
u24	Update /etc/hosts file (enter function update_etchosts).
u25	Get configuration files (enter function get_files).
u26	Get authentication files (enter function authent_stuff).
u27	Update /etc/inittab file (enter function update_etcinittab).
u28	Perform MP-specific functions (enter function upmp_work).
u29	Install prerequisite file sets (enter function install_prereqs).
u30	Install ssp.clients (enter function install_ssp_clients).
u31	Install ssp.basic (enter function install_ssp_basic).
u32	Install ssp.ha (enter function install_ssp_ha).
u33	Install ssp.sysctl (enter function install_ssp_sysctl).
u34	Install ssp.pman (enter function install_ssp_pman).
u35	Install ssp.css (enter function install_ssp_css).
u36	Install ssp.jm (enter function install_ssp_jm).
u37	Delete master .rhosts entry (enter function delete_master_rhosts).
u38	Create new dump logical volume (enter function create_dump_lv).
u39	Run customer’s tuning.cust (enter function run_tuning_cust).
u40	Run customer’s script.cust (enter function run_script_cust).
u41	Configure switch (enter function config_switch).
u42	Run psspfb_script (enter function run_psspfb_script).
u46	/tftpboot/tuning.cust is being run during node customization.
u50	Get tuning.cust from boot/install server.
u51	Processor is ″UP″.

Table 85. SP-Specific LED/LCD values (Numerical order) (continued)

u52	Processor is ″MP″.
u53	Install ssp.ha.
u54	Get spfbcheck from boot/install server.
u55	Fatal error in bosboot.
u56	Get psspfb_script from boot/install server.
u58	Get psspfb_script from control workstation.
u57	Get the *node*.config_info file from the master.
u59	Get the cuat.sp template from the master.
u60	Create/update /etc/ssp files.
u61	Get /etc/SDR_dest_info from boot/install server.
u67	Get /etc/krb.conf from boot/install server.
u68	Get /etc/krb.realms from boot/install server.
u69	Get krb-srvtab from boot/install server.
u79	Get script.cust from boot/install server.
u80	Mount lppsource and install ssp.clients.
u81	Install ssp.basic.
u82	Install ssp.sysctl.
u84	Install ssp.css.
u85	Install ssp.jm.
u86	Create new dump logical volume.
u87	Run customer's script.cust script file.

The following LEDs/LCDs are produced after NIM installation has occurred, and during the initial post-installation reboot of the node. This list is sorted chronologically, or in the order of which the LEDs/LCDs occur during processing.

u90	Setup working environment (enter function setup_environment).
u91	Unconfiguring any adapters left over from the mksysb.
u89	Cleans up and unconfigures DCE from mksysb and removes key files and sysctl acls.
u92	Configure adapters (enter function config_adapters).
u93	Configure inet0 (enter function config_inet0).

u94 Run cfgmgr (enter function run_cfgmgr).

u99 Set CLOCAL flag on tty0

u95 Run complete_node on boot/install server (enter function complete_node).

u78 Set the KRBTKFILE variable and get an rcmd ticket.

u96 Run customer’s firstboot.cust (enter function run_firstboot_cust).

The following list contains the same LEDs/LCDs as in the previous list, but sorted numerically for reference use:

u78 Set the KRBTKFILE variable and get an rcmd ticket.

u89 Cleans up and unconfigures DCE from mksysb and removes key files and sysctl acls.

u90 Setup working environment (enter function setup_environment).

u91 Unconfiguring any adapters left over from the mksysb.

u92 Configure adapters (enter function config_adapters).

u93 Configure inet0 (enter function config_inet0).

u94 Run cfgmgr (enter function run_cfgmgr).

u95 Run complete_node on boot/install server (enter function complete_node).

u96 Run customer’s firstboot.cust (enter function run_firstboot_cust).

u99 Set CLOCAL flag on tty0

The following LEDs/LCDs occur during a node IPL.

762 SP Switch Adapter configuring on node

763 SP Switch MX Adapter configuring on node

764 RS/6000 SP System Attachment Adapter configuring on node

Other LED/LCD codes

For LED/LCDs generated by the hardware, see the manual for your hardware type:

- *RS/6000 SP: 604 and 604e SMP High Node Service Guide*, GA22-7446
- *RS/6000 SP: POWER3 SMP High Node Service Guide*, GA22-7448
- *RS/6000 SP: SMP Thin and Wide Node Service Guide*, GA22-7447
- *RS/6000 SP: SP Switch Service Guide*, GA22-7443
- *RS/6000 SP: Uniprocessor Node Service Guide*, GA22-7445
- *RS/6000 Diagnostic Information for Multiple Bus Systems*, SA38-0509

- *SSA Adapters User’s Guide and Maintenance Information*, SA33-3272

The generic service guide also lists LED/LCD values: *RS/6000 SP: System Service Guide*, GA22-7442.

LED/LCD values issued by AIX are documented in: *AIX 5L Version 5.1 Messages Guide and Reference*.

Chapter 33. Diagnosing 604 and 604e High Node problems

This chapter provides information on:

- 604 and 604e High Node characteristics, including:
  - Addressing power and fan failures in these nodes
  - Rebooting the node after a system failure
- Error conditions and performance considerations
- Using SystemGuard and BUMP programs

604 and 604e High Node characteristics

The 604 and 604e High Node operation is different from other nodes in several areas:

- A power feature is available which adds a redundant internal power supply to the node. In this configuration, the node will continue to run in the event of a power supply failure. Error notification for a power supply failure is done through the AIX Error Log on the node.
- The cooling system on the node also has redundancy. In the event that one of the cooling fans fails, the node will continue to run. Error notification for a power supply failure is done through the AIX Error Log on the node.
- If a hardware related crash occurs on the node, SystemGuard will re-IPL the node using the **long** IPL option. During **long** IPL, some CPU or memory resources may be deconfigured by SystemGuard to allow the re-IPL to continue.

Error conditions and performance considerations

You need to be aware of the following conditions that pertain to the unique operation of this node:

- An error notification object should be set up on the node for the label **EPOW_SUS**. The **EPOW_SUS** label is used on AIX Error Log entries that may pertain to the loss of redundant power supplies or fans.
- If the node is experiencing performance degradation, use the **lscfg** command to verify that none of the CPU resources have been deconfigured by SystemGuard This happens if it has re-IPLed the node using the **long** IPL option.

Using SystemGuard and BUMP programs

SystemGuard is a collection of firmware programs that run on the bringup microprocessor (BUMP). SystemGuard and BUMP provide service processor capability. They enable the operator to manage power supplies, check system hardware status, update various configuration parameters, investigate problems, and perform tests.

The BUMP controls the system when the power is off or the AIX operating system is stopped. The BUMP releases control of the system to AIX after it is loaded. If AIX stops or is shut down, the BUMP again controls the system.

To activate SystemGuard, the key mode switch must be in the service position during the standby or init phases. The standby phase is any time the system power is off. The init phase is the time when the system is being initialized. The PSSP software utilizes SystemGuard IPL flags such as the **FAST IPL** default. The following example shows how to modify the **FAST IPL** flag.

1. Open two windows: one for invoking **hmcmds** and one for the **s1term** tty console for the 604 or 604e High Node.
2. Invoke **hmcmds** to power off the node and set the key switch to **service**:

   ```
   hmcmds -G off frame:node
   hmcmds -G service frame:node
   ```
3. Open the tty console in write mode for the target 604 or 604e High Node:

   ```
   s1term -G -w frame:node
   ```

   OR

   ```
   spmon -o node_number
   ```
4. Working with the tty console, press **<Enter>** and make sure that the BUMP prompt (>) appears in the window. Invoke **sbb** in the tty console to bring up the Stand-By Menu.
5. Select **1** to check the current flag settings on the tty console.
6. If the flag setting for **FAST IPL** is **disabled**, enter **5** to change the value from **disabled** to **enabled**
7. Press **x** a few times to exit out of Stand-By menu on tty console and close the console.
8. Invoke **hmcmds** to power on the 604 High Node and set the key switch to **Normal**:

   ```
   hmcmds -G on frame:node
   hmcmds -G normal frame:node
   ```
9. The 604 or 604e High Node will now initialize through the **BUMP** interface, and will IPL the node using the **FAST IPL**, bypassing the low-level BUMP diagnostics.

For more information about SystemGuard and BUMP, see the headings ″Using the 604 or 604e High Node BUMP menus″ and ″SystemGuard maintenance Menu access″ of Chapter 3, ″Service Procedures″ in *RS/6000 SP: 604 and 604e SMP High Node Service Guide*.

Chapter 34. Diagnosing POWER3 SMP Thin and Wide Node problems

The term ″POWER3 SMP Thin and Wide Node″ used in this chapter refers to the following nodes:

- 200 MHz POWER3 SMP Thin and Wide nodes
- 375 MHz POWER3 SMP Thin and Wide node

This section provides information on:

- POWER3 Symmetric MultiProcessing (SMP) node characteristics
- The boot sequence for the POWER3 SMP Thin and Wide Node
- Error conditions and performance considerations
- Service processor surveillance

POWER3 SMP Thin and Wide Node characteristics

The POWER3 SMP node is offered in either a thin or wide footprint. The POWER3 nodes have the same hardware monitor supervisor card types as the 332 MHz SMP Thin and Wide Nodes.

In order to provide an SMP in these more dense packages, the power and cooling design is more like that of earlier thin and wide uniprocessor nodes. N+1 power and cooling is not offered for these nodes.

The POWER3 SMP node has the following characteristics like the 332 MHz SMP node:

- RPA (RS/6000 Platform Architecture) architecture, formerly known as CHRP. Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```
- The RPA architecture dictates that there is no key switch on a POWER3 SMP node. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  sysdumpdev -K
  ```

  must be issued before a dump-on-demand condition arises.

  The **-K** (upper case K) flag is most likely the default for the node. The AIX command:

  ```
  sysdumpdev -k
  ```

  (lower case k) resets the setting.

Boot sequence for the POWER3 SMP Thin and Wide Node

The progress of a node's boot may be observed by watching its LCD change, either using the SP Hardware Perspective or via the command line using the command:

```
hmmon -Gq frame:slot
```

where *frame:slot* is the frame and slot number for the node.

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node’s Service Processor. This phase has completed once an LCD of the form **E000** through **E0FF** is displayed.

If the node boot stops during this phase, do the following:

1. Check the meaning of the final LCD code in ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide*.
2. Display the Service Processor logs by performing these steps:
   a. Power off the node from the control workstation.
   b. Open a TTY session to the node.
   c. Hit the return key to go to the Service Processor **Main Menu**.
   d. Select the **System Information Menu**.
   e. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with an LCD of **Exxx**, where **xxx** may be any value. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide*.

The final phase of node boot is the booting of the AIX operating system. This begins with a BootP request, which is denoted by the **E175** LCD code. Control is later handed over to AIX under the **E105** LCD code, after which AIX codes are seen, which all begin with **0xxx**. If the AIX boot has difficulties, record the LCD codes and consult ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error conditions and performance considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **4B2xxx01**, **4B2xxx02** or **4B2xxx10**, where **xxx** is:

- **654** for the 200 MHz POWER3 SMP Thin and Wide Node
- **768** for the 375 MHz POWER3 SMP Thin and Wide Nodes

Firmware routines are automatically invoked, which result in one of two new AIX error log entries:

- MACHINE_CHECK_CHRP - an immediate report of a hardware failure
- SCAN_ERROR_CHRP - a failure reported via periodic scan of NVRAM (non-volatile RAM)

Error data is encoded within these logs in accordance with the RPA architecture. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node’s performance. When this occurs, an AIX error log entry is written.

If the node chkstops, you may not be able to reboot the node and run ELA. In this case, refer to *RS/6000 SP: SMP Thin and Wide Node Service Guide*.

Service Processor surveillance

If surveillance is enabled, the Service Processor provides a ’processor-well’ checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung. The Service Processor posts an LCD code of **40A00000** or **40B00000**. If this occurs, reboot the node and check the AIX error log for hardware errors. If there are hardware errors, call IBM hardware service. Otherwise, collect information using a SNAP tool and call the IBM Support Center.

For more information about surveillance, see the heading ″Service processor system monitoring - surveillance″ of ″Chapter 3, Service procedures″ in *RS/6000 SP: SMP Thin and Wide Node Service Guide*.

Chapter 35. Diagnosing POWER3 SMP High Node problems

The term ″POWER3 SMP High Node″ used in this chapter refers to the following nodes:

- 222 MHz POWER3 SMP High Node
- 375 MHz POWER3 SMP High Node

This chapter provides information on:

- POWER3 SMP (Symmetric MultiProcessor) High Node characteristics
- The boot sequence for this node
- Error conditions and performance considerations
- Service processor surveillance
- The SP Expansion I/O Unit

POWER3 SMP High Node characteristics

The POWER3 SMP High Node is different from other nodes in several areas:

- The architecture of these nodes conforms to the RS/6000 Platform Architecture (RPA). Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```
- The RPA dictates that there is no key switch on these nodes. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  /usr/bin/sysdumpdev -K
  ```

  must be issued before a dump-on-demand condition arises.

  The -K (upper case K) flag is the default for the node. The AIX command:

  ```
  /usr/bin/sysdumpdev -k
  ```

  (lower case k) resets the setting.

Boot sequence for the POWER3 SMP High Node

The progress of a node’s boot may be observed by watching its LCD change, either using the SP Hardware Perspective or via the command line using the command:

```
/usr/lpp/ssp/bin/hmmon -Gq frame:slot
```

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node’s Service Processor. This phase has completed once an LCD of the form **E000** through **E0FF** is displayed.

If the node boot stops during this phase, do the following:

1. Check the meaning of the final LCD code in ″Appendix A, Messages and Codes ″ of *RS/6000 SP: POWER3 SMP High Node Service Guide*.
2. Display the Service Processor logs by performing these steps:

a. Power off the node from the control workstation.
b. Open a TTY session to the node.
c. Hit the return key to go to the Service Processor **Main Menu**.
d. Select the **System Information Menu**.
e. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with an LCD of **Exxx**, where **xxx** may be any value. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in ″Appendix A, Messages and Codes ″ of *RS/6000 SP: POWER3 SMP High Node Service Guide*.

The final phase of node boot is the booting of the AIX operating system. This begins with a bootp request, which is denoted by the **E175** LCD code. If the AIX boot has difficulties, record the LCD codes and consult ″Appendix A, Messages and Codes″ of *RS/6000 SP: POWER3 SMP High Node Service Guide* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error conditions and performance considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **45800001**. Firmware routines are automatically invoked, which result in one of two AIX error log entries:

- MACHINE_CHECK_CHRP
- SCAN_ERROR_CHRP

Error data is encoded with these logs in accordance with RPA. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node's performance. When this occurs, an AIX error log entry is written.

Service Processor Surveillance

If surveillance is enabled, the Service Processor provides a 'processor-well' checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung. The Service Processor posts an LCD code of **40B00000**. If this condition occurs, call IBM hardware service, or consult the hardware manual for a procedure to collect pertinent data from the Service Processor.

For more information about surveillance, see the heading ″Service processor system monitoring - surveillance″ of ″Chapter 3, Service procedures″ in *RS/6000 SP: POWER3 SMP High Node Service Guide*.

SP Expansion I/O Unit

The POWER3 SMP High Node can optionally have one or more SP Expansion I/O Units attached to the node. An SP Expansion I/O Unit is normally cabled to the node in a loop (with up to two expansion units per loop) which returns to the node. This creates a redundant data path from the expansion unit to the node.

An LCD value of **203w0xyz** indicates a cabling configuration between the node and the SP Expansion I/O Unit which does not result in a complete loop. In this LCD, the **w** is the expansion unit loop number, the **x** is always 0, the **y** indicates the expansion I/O port number on the node's system rack, and the **z** can have the following values:

z=0 Incorrectly cabled SP Expansion I/O Unit configuration

z=B A missing return link from the SP Expansion I/O Unit to the node's system rack

z=C A missing cable between two SP Expansion I/O Units

z=D A BIST (built-in self test) failure in the SP Expansion I/O Unit

z=E An SP Expansion I/O Unit was found connected to port 1, 3, or 5 with no return to the node's system rack, and no SP Expansion I/O unit was found connected to port 2, 4, or 6 (respectively).

In this case, the expansion unit connected to port 1, 3, or 5 is removed from the configuration since the cause of the error and the proper location of the expansion unit cannot be determined.

Chapter 36. Diagnosing 332 MHz SMP Thin and Wide Node problems

This chapter provides information on:
- 332 MHz Symmetric MultiProcessing (SMP) node characteristics
- The boot sequence for the 332 MHz SMP Node
- Error conditions and performance considerations
- Service processor surveillance

332 MHz SMP Node characteristics

The 332 MHz node was the first SP system SMP node offered in either a thin or wide footprint. Previous SMP nodes were available only as high nodes.

In order to provide an SMP in these denser packages, the power and cooling design is more like that of earlier thin and wide uniprocessor nodes. N+1 power and cooling is not offered for these nodes.

The 332 MHz SMP node is different from other nodes due to these characteristics:

- RPA (RS/6000 Platform Architecture), formerly known as CHRP. Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```
- The RPA architecture dictates that there is no key switch on a 332 MHz SMP node. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  sysdumpdev -K
  ```

  must be issued before a dump-on-demand condition arises.

  The **-K** (upper case K) flag is most likely the default for the node. The AIX command:

  ```
  sysdumpdev -k
  ```

  (lower case k) resets the setting.

Boot sequence for the 332 MHz SMP Node

The progress of a node’s boot may be observed by watching its LCD change, either using the SP Hardware Perspective or via the command line using the command:

```
hmmon -Gq frame:slot
```

where *frame:slot* is the frame and slot number for the node.

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node’s Service Processor. This phase has completed once an LCD of the form **E3xy** is displayed, where **x** and **y** may be any values.

If the node boot stops during this phase do the following:

1. Check the meaning of the final LCD code in ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide*.
2. Display the Service Processor logs by performing these steps:
   a. Power off the node from the control workstation.
   b. Open a TTY session to the node.
   c. Hit the return key to go to the Service Processor **Main Menu**.
   d. Select the **System Information Menu**.
   e. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with a memory test, denoted by an LCD of **E3xy**. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide*.

The final phase of node boot is the booting of the AIX operating system. This begins with a BootP request, which is denoted by the **E175** LCD code. Control is later handed over to AIX under the **E05** LCD code, after which AIX codes are seen, which all begin with **0xxx**. If the AIX boot has difficulties, record the LCD codes and consult ″Appendix A, Messages and Codes″ of *RS/6000 SP: SMP Thin and Wide Node Service Guide* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error conditions and performance considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **4B246101**, **4B246102**, **4B246110**, **4B24A101**, **4B24A102**, **4B24A110**. Firmware routines are automatically invoked, which result in one of two new AIX error log entries:

- MACHINE_CHECK_CHRP - an immediate report of a hardware failure
- SCAN_ERROR_CHRP - a failure reported via periodic scan of NVRAM (non-volatile RAM)

Error data is encoded within these logs in accordance with the RPA architecture. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node's performance. When this occurs, an AIX error log entry is written.

Service Processor surveillance

If surveillance is enabled, the Service Processor provides a 'processor-well' checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung. The Service Processor posts an LCD code of **40B0010p**, where *p* is the processor number of the hung processor. If this condition occurs, call IBM hardware service, or consult the hardware manual for a procedure to collect pertinent data from the Service Processor.

Chapter 37. Diagnosing dependent node configuration problems

The implementation of *dependent* nodes deals directly with RS/6000 SP Switch Routers. Therefore, configuration problems with a dependent node may be caused by either problems on the router node or on the SP. This chapter helps you diagnose these problems and determine how to correct configuration problems on the SP and on the SP Switch Router.

SP configuration diagnosis

The first step in diagnosing a dependent node problem is verifying that both the node and associated adapter have been properly configured and are operating correctly. To do this, perform steps 1, 2, and 3 here. You can optionally perform step 4.

1. Issue the following command to verify the definition of the extension node:

   ```
   splstnodes -t dependent node_number reliable_host_name\
    management_agent_hostname extension_node_identifier snmp_community_name
   ```

2. Issue the following command to verify the definition of the extension node adapter:

   ```
   splstadapters -t node_number netaddr netmask
   ```

3. Issue the following command to verify that the extension node is connected to the switch:

   ```
   SDRGetObjects switch_responds
   ```

   You should receive output similar to the following:

   ```
   node_number switch_responds autojoin  isolated     adapter_config_status
   1           0               0         0            css_ready
   3           0               0         0            css_ready
   4           0               0         0            css_ready
   5           0               0         0            css_ready
   6           0               0         0            css_ready
   7           0               0         0            css_ready
   ```

4. Use the SP Hardware Perspective to verify the definitions of the extension node and extension node adapter, and to verify that the extension node is connected to the switch.

   a. Issue **perspectives &** to bring up Perspectives as a background process.

   b.

      Double Click on:
      manage/control hardware

      Double Click on:
      IP Node

      Double Click on:
      Monitoring

      See the online help for an explanation of the notebook associated with the IP Node. Review the information associated with the node to verify that it is operating correctly.

Verify that these configuration parameters are correct. If you find errors here with any values other than the node number, use the **endefnode** and **endefadapter** commands to change those values. If the node number is incorrect, remove the node number and replace it with a correct node number. To remove it, use the **enrmnode** command or the **enrmadapter** command.

The next step is to determine what state your dependent node is in on the SP Switch. Use the **SDRGetObjects switch_responds** command to determine this. Here is an example of the output from this command, showing several dependent nodes in various states:

```
node_number  switch_responds autojoin  isolated  adapter_config_status
          1               0        0         0  not_configured
          2               0        0         0  css_ready
          3               0        0         0  micro_code_load_failed
          4               0        0         1  css_ready
          5               0        1         1  css_ready
          6               1        0         0  css_ready
```

An examination of each of these six nodes and their states will help to diagnose problems.

Dependent Node 1, in the previous example, is in the state of a newly-defined dependent node. This state means that the configuration is not complete, or there is no communication via SNMP to the node. See “SNMP configuration diagnosis” on page 556 for more information on diagnosing and correcting SNMP communication problems.

Dependent Node 2 is in the state of a dependent node after its dependent adapter has been defined and the node has been reconfigured.

Dependent Node 2 now shows the **adapter_config_status** is **css_ready**. This means that the configuration information has been delivered to and accepted by the node via SNMP. This node is ready to become active on the SP Switch network. If this reconfiguration was not successful, the **adapter_config_status** would either remain **not_configured** or become **micro_code_load_failed**, as seen with Dependent Node 3.

Dependent Node 3’s **micro_code_load_failed adapter_config_status** could also indicate that the node is still resetting the SP Switch Router Adapter in its chassis and has not finished. This reconfiguration can take some time, and should be checked again. If you suspect that your configuration information is not being properly transmitted via SNMP to the node, check that the SP Switch Router Adapter is in the same slot as defined in the adapter definition.

Dependent Nodes 4 and 5 are fenced and ready to be unfenced and brought onto the SP Switch network. The only difference is that Dependent Node 5 has **autojoin** turned on. This means that whenever a reconfigure of the SP Switch Router Adapter occurs in this node, it will automatically unfence. Dependent Node 4 will have to be unfenced manually by issuing the **Eunfence 4** command on the control workstation.

Dependent Node 6 indicates the active state of a properly configured dependent node. If you are unable to attain this state with a dependent node, and you cannot discover any SP–related configuration or SNMP network-related problems, then you will have to open an administrative telnet session to the node to diagnose problems there.

SP Switch Router configuration diagnosis

More information pertaining to the diagnosis of problems with the SP Switch Router and the SP Switch Router Adapter can be found in the documentation that ships along with the hardware. This documentation is listed in the Bibliography.

Some basic information about diagnosing problems with the SP Switch Router Adapter is included in the following section for ease of use, but you should also reference the documentation for the SP Switch Router.

Once you have exhausted configuration problems on the SP system, **telnet** to the SP Switch Router. To do this, you need the system userid (with **root** privileges) and password. You can open a **telnet** session from any network terminal screen or from SP Perspectives **only if remote logins have been enabled for the router node.** For more information about enabling remote logins, see *GRF Configuration Guide - Enable telnet access*. If you cannot remotely login, you can obtain local access to the router node from its RS-232 terminal, if the user has left it enabled. This terminal is only necessary for initial configuration.

After logging in with **root** privileges to the router node, there are several commands that are helpful in diagnosing problems. The first of these is the **grcard** command. Here is sample output from this command as run on a router node with 4 SP Switch Router Adapter cards installed:

```
# grcard -v
Slot    HWtype  State
----    ------  -----
0       DEV1_V1 held-reset
1       DEVI_V1 loading
2       DEV1_V1 dumping
3       DEV1_V1 running
```

The SP Switch Router Adapter is referred to as a DEV1_V1 media card to the router node, as you can see it listed under the HWtype column. Some of the common states of this media card include the four from the example, which are:

loading
: The media card is loading its configuration information. The SP Switch Router Adapter will remain in this state until configuration information from the SDR in the SP System is communicated to the router node via SNMP. If your card never leaves this state, you should read the section later in this chapter about “SNMP configuration diagnosis” on page 556.

dumping
: A media card that has been reset, or is failing, will dump its memory to a file before resetting if the MIB configuration field for this feature is turned on. By default, all SP Switch Router Adapter cards have this feature enabled. For more information on this feature see the GRF publications.

held-reset
: This is the state that the SP Switch Router Adapter will be in if the **grreset -h** (discussed later) command has been issued on it. This state is also present if the card has been put in a reset state from the SP System with the **enadmin** command. For more information about the **enadmin** command, refer to *PSSP: Command and Technical Reference*.

running
: This is the normal active state of the SP Switch Router Adapter.

Another useful command on the router node is the **grreset** command. Use this command to reset the SP Switch Router Adapter. When this is done, configuration information for the adapter is loaded from the control board on the router node. It is like rebooting the adapter. For more complete information about the **grreset** command, see *GRF Reference Guide.*

The **grconslog** command is useful in diagnosing problems on the router node. This command opens and displays the console log for the router node. Here you can see all configuration and network activity on the node. The most useful method for running this command is to run **grconslog -pdf** to open the console log, display information with port and date stamps, and keep the log open in flow mode.

Occasionally, the static file used for this log fills up. In this case, it then wraps or opens a new file depending on how the system is configured. When this happens, it will appear that the flow mode of your console log hangs. To recover, issue **<Ctrl-c>** out of the console log and start it up again. For more information on this command, see the GRF publications.

The primary use of the **grconslog -pdf** command is to open the console log and then issue commands from the SP System to see if they are getting to the router node and being run. You can find more information on this in the “SNMP configuration diagnosis” section in this chapter.

SNMP configuration diagnosis

The following section will aid you in diagnosing communication problems which may occur between the SNMP Agent administering the dependent node (residing on the router node) and the SP Manager residing on the control workstation. You should run with tracing enabled for the SPMGR subsystem during a dependent node configuration.

When you configure a dependent node within an extension node class, you create attribute values in the SDR DependentNode class which are used by the SP SNMP Manager to communicate with the SNMP Agent on the router node. These attributes are:

node_number
The node number for the dependent node

extension_node_identifier
The identifier assigned to the dependent node (this is the two-digit slot number of the dependent node adapter on the router node)

management_agent_hostname
The fully qualified hostname of the node on which the SNMP Agent administering the dependent node resides. This is used to communicate with the router node. It must resolve to an IP address.

snmp_community_name
The SNMP community name placed within SNMP messages passed between the SNMP Agent and the SNMP SP Manager for authentication. This value must match the community name value configured on the SNMP Agent host for communicating with the SP Manager on the control workstation.

If the **node_number** is specified in error, the configuration data may be sent to the SNMP Agent administering the dependent node successfully. However. problems will occur when attaching the switch adapter to the switch network.

When you have completed the definition of the dependent node on the control workstation, and have installed the SP Switch Router Adapter on the router node, check to see if the SDR *adapter_config_status* attribute value for the dependent node in the *switch_responds* class remains *configured.*. If so, then trap messages from the router node are not being processed successfully by the SNMP Manager on the control workstation. This can be caused by one of several problems:

1. If the **spmgr** subsystem trace file in the directory **/var/adm/SPlogs/spmgr** contains an entry indicating *init_io failed: udp port in use*, then the UDP port specified for service name **spmgrd-trap** in the **/etc/services** file on the control workstation is already in use. This error will also appear in an AIX error log entry written by the **spmgrd** daemon.

   Solution: Change the UDP port number for the **spmgrd-trap** service to an unused port number. The router node **snmpd** daemon configuration file, **/etc/snmpd.conf** on the router node, must also be updated to specify this same port number when sending trap messages to the control workstation. Both the **snmpd** daemon on the router node and the **spmgr** subsystem on the control workstation must be restarted after this change is made.
2. If the **lssrc -ls spmgr** command response contains zeros for both the number of switchInfoNeeded traps processed successfully, and the number processed unsuccessfully, then trap messages sent by the SNMP Agent on the router node are not being received by the SNMP Manager on the control workstation.

   Either the control workstation IP address or the UDP port number may have been specified in error in the **/etc/snmpd.conf** file on the router node. The UDP port number associated with the control workstation in file **/etc/snmpd.conf** on the router node must match the UDP port number specified for the **spmgrd-trap** service in the **/etc/services** file on the control workstation.

   Solution: Correct the erroneous value and restart the **spmgrd** daemon on the router node and the **spmgr** subsystem on the control workstation.
3. If the **lssrc -ls spmgr** command response contains zeros for the number of switchInfoNeeded traps processed successfully, and the number processed unsuccessfully is greater than zero, then trap messages sent by the SNMP Agent on the router node are being received by the SNMP Manger on the control workstation. However, they are not being successfully processed. This may be the result of one of the following errors:
   a. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: *'Dependent node <ext_id>* managed by the SNMP Agent on *<router_node_hostname>* is not configured in the SDR - switchInfoNeeded trap ignored', then either the *extension_node_identifier* or the *management_agent_hostname* attribute value for the corresponding extension node in the SDR DependentNode class is incorrect.

      Solution: Correct the attribute value.
   b. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: *'SDR attribute <attrname>* for *dependent node <ext_id>* in *class <classname>* has a null value for SNMP Agent on *host <router_node_hostname>'*, or an entry indicating: *'SDRGetAllObjects() DependentAdapter failed with return code 4'*, then required configuration values are missing from the indicated SDR class.

      Solution: Supply the missing attribute values.
   c. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: '*Dependent node <ext_id>* managed by the SNMP Agent on *host <router_node_hostname>* is configured with a bad community name - switchInfoNeeded trap ignored', then the *snmp_community_name* attribute value specified for the corresponding node

in the SDR DependentNode class does not match the community name specified for the control workstation in the **/etc/snmpd.conf** file on the router node.

Note that if the *snmp_community_name* attribute value is null, the community name to be specified in the router node is documented in the Ascend documentation.

Solution: Correct the community names in the **/etc/snmpd.conf** file on the router node and the *snmp_community_name* attribute for the corresponding SDR DependentNode class so that they match.

Some SNMP-related configuration problems occur when data is changed in the SDR after an initial configuration. Most of these problems are detected by the configuration-related commands, and messages are issued to the operator.

If you attempt to reconfigure a dependent node **after** doing one of the following:

- Issuing either the **endefnode** or **endefadapter** command with the **-r** operand.
- Selecting the reconfigure option from a SMIT extension node configuration panel.
- Issuing an **enadmin** command.

These problems could occur:

1. A time_out occurs on the **enadmin** command (invoked internally from the SMIT panels, **endefnode**, and **endefadpter** commands). This could be caused by one of the following configuration problems:
   a. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** or the AIX error log contains an entry indicating '2536-007 An authentication failure notification was received from an SNMP Agent running on host <router_node_hostname> which supports Dependent Nodes', then the SDR *snmp_community_name* attribute value in the DependentNode class for the extension node contains a name that does not match the community name specified for the control workstation in the **/etc/snmpd.conf** file on the router node.

      Solution: Correct the community names in the **/etc/snmpd.conf** file on the router node and the *snmp_community_name* attribute for the corresponding SDR DependentNode class so that they match.

   b. If no authentication error exists in either the trace file or the AIX error log, then the value specified for the SDR *management_agent_hostname* attribute in the DependentNode class for the extension node must not be the correct fully–qualified name for the router node.

      Solution: Correct the *management_agent_hostname* attribute value in the DependentNode class for the extension node.

Note: if the *extension_node_identifier* attribute value for an extension node is erroneously set to the ID of another existing extension node on the router node managed by another SP system, the results are unpredictable since two SNMP managers are trying to configure the same SP Switch Router Adapter.

Chapter 38. Diagnosing SP-attached Server and Clustered Enterprise Server problems

This chapter discusses diagnostic procedures and failure responses for the SP-attached server and clustered enterprise server software of PSSP which use the SAMI (Service and Manufacturing Interface) hardware protocol. The SP-attached servers and clustered enterprise servers which use the SAMI hardware protocol include the RS/6000 S70, S7A, and S80 servers and the IBM eServer pSeries 680. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 569. A list of the information to collect before contacting the IBM Support Center is in the section “Information to collect before contacting the IBM Support Center” on page 562.

SP-attached servers can be connected to SP systems that contain an SP Switch or those that have no switch. The S80 and the IBM eServer pSeries 680 servers can also connect to the SP Switch2. Clustered enterprise servers are clustered with an SP control workstation in a system that contains no SP node frames. Clustered enterprise server systems may contain an SP Switch or SP Switch2, in which case they are effectively treated as SP-attached server systems. In a clustered environment without a switch, or in a system with an SP Switch2, the system cannot be partitioned. The terms **the server** and **all servers** are used in this chapter in place of ’SP-attached server or clustered enterprise server which uses the SAMI hardware protocol’ and ’SP-attached servers or clustered enterprise servers which uses the SAMI hardware protocol’ respectively.

For information on diagnosing problems with SP-attached servers or clustered enterprise servers which use the CSP (Converged Service Processor) hardware protocol, see “Chapter 21. Diagnosing System Monitor problems” on page 317. The SP-attached servers and clustered enterprise servers which use the CSP hardware protocol include the RS/6000 H80 and M80 servers, and the IBM eServer pSeries 660 (6H0, 6H1, 6M1).

For information on diagnosing problems with SP-attached servers or clustered enterprise servers which use the HMC (Hardware Management Console) hardware protocol, see “Chapter 39. Diagnosing IBM eServer pSeries 690 problems” on page 577. The SP-attached servers and clustered enterprise servers which use the HMC hardware protocol include the IBM eServer pSeries 690.

Related documentation

The following publications provide information about the PSSP software that controls all servers:

1. *PSSP: Command and Technical Reference*

   Entries for these commands:
   - hmadm
   - hmcmds
   - hmmon
   - SDR_config
   - spdelfram
   - spframe
   - splstdata
   - s1term

2. *PSSP: Messages Reference*

   The chapter ″0026 - System Monitor Messages″ contains messages issued by all servers.

3. *PSSP: Administration Guide*

   See the Appendixes ″SP Daemons″ and ″The System Data Repository″.

Requisite function

This is a list of the software directly used by the PSSP software related to all servers. Problems within the requisite software may manifest themselves as error symptoms in this software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the server, consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- The System Monitor (hardmon) must be running on the control workstation.

  The external hardware daemons consist of two interfaces: the frame supervisor interface and the node supervisor interface.

  – The frame supervisor interface is responsible for keeping the state data in the frame’s packet current and formatting the frame packet for return to the hardmon daemon.

  – The node supervisor interface is responsible for keeping the state data in the node’s packet current. The node supervisor is also responsible for translating the commands received from the frame supervisor interface into SAMI (Service and Manufacturing Interface) protocol, before sending them to the service processor on all servers.

- The System Data Repository (SDR) subsystem

  The System Monitor requires the SDR subsystem. For more information, refer to “Chapter 14. Diagnosing SDR problems” on page 125.

- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

Logs

There are many logs that record error and status information.

Examine the following logs: daemon log, SDR configuration log, external hardware daemon communication log, S1 communication log, hardware monitor log, and SP hardware log.

Daemon log

This file contains the daemon’s error information. This file is located on the control workstation in **/var/adm/SPlogs/spmon/s70d/s70d.***frame***.log.***ddd* for the s70 daemon where *ddd* is the Julian date, and *frame* is the frame number of the server.

SDR configuration log

This log file contains information from the SDR configuration command **SDR_config**. The **SDR_config** command is automatically run under these conditions:

- Whenever the hardware monitor daemon is started

- When hardware changes are detected on the system
- In response to requests from the **spframe -r yes** command.
- In response to requests from the **spdelfram** command.

If **SDR_config** detects system configuration problems, error messages are written to the file located on the control workstation in: **/var/adm/SPlogs/sdr/SDR_config.log**.

External hardware daemon communication log

This file contains information on data sent to, and received from, the server's service processor. This file must be created manually on the control workstation as **/var/adm/SPlogs/spmon/s70d/s70d.***frame***.sami_dump** for the s70 daemon, where *frame* is the frame number of the server.

This log file contains informational, error and debug information. Examples of informational messages are timestamps for when the daemon started or stopped. An example of an error is the failure of the daemon to open a file. An example of debug information (enabled by the **hmadm** command), is a list of commands sent from the System Monitor (**hardmon**) to external hardware daemon.

For the s70 daemon, create this log whenever the s70 hardware appears to not be communicating with the s70 daemon. Issue this command:

```
hmadm -d sami setd
```

Issue this command to stop the log:

```
hmadm -d sami cleard
```

S1 communication log

This log file contains information on data sent to, and received from, the server's serial port. This file must be created manually on the control workstation. This file is located in: **/var/adm/SPlogs/spmon/s70d/s70d.***frame***.s1data_dump**, where *frame* is the frame number of the server.

Create this log whenever the s70 hardware serial interface appears to not be communicating with the s70 daemon. Issue this command:

```
hmadm -d s1data setd
```

Issue this command to stop the log:

```
hmadm -d s1data cleard
```

Hardware monitor log

This log file contains informational, error, and debug information. Examples of informational messages are timestamps for when daemons are started or stopped. An example of an error is a communication error between the System monitor (**hardmon**) an external hardware daemon. An example of debug information (enabled by the **hmadm** command), is a list of requests sent from the System monitor (**hardmon**) to an external hardware daemon. This file is located on the control workstation. The file is located in: **/var/adm/SPlogs/spmon/hmlogfile.***ddd*, where *ddd* is the Julian date.

SP hardware log

This log file contains information generated by system daemons, including hardware errors. SP hardware problems have a resource name of **sphwlog**. This file is located on the control workstation in: **/var/adm/SPlogs/SPdaemon.log**.

Dump information

The External hardware daemons create standard AIX core dumps when they fail. Any core file created should be saved in a safe place at the time of the failure, so that it is not overwritten if another failure occurs. The IBM Support Center can then examine the file at a later time. The dump is located on the control workstation in **/var/adm/SPlogs/spmon/s70d/core** for s70.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Each External hardware daemon provides facilities that allow the tracing of certain aspects of its internal function, as well as its communication with the SP Frame Supervisor. This is done using the **hmadm** command. Use this trace only when directed by the IBM Support Center. The *PSSP: Command and Technical Reference* gives details about running the trace and obtaining the output.

Information to collect before contacting the IBM Support Center

Collect this information before calling the IBM Support Center:

1. Any command output that seems to be helpful. For example, the output of any failing diagnostic procedures that were run.
2. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   The entry `ts_auth_methods` lists the authentication methods.
3. AIX error log. On the control workstation, issue the command:

   ```
   LANG=C errpt -a > /tmp/AIXerrlog
   ```
4. Daemon core dump, if it exists. See “Dump information”.
5. Daemon log. See “Daemon log” on page 560.
6. SDR Configuration Log. See “SDR configuration log” on page 560.

7. External hardware daemon communication log, if one has been created using the **hmadm** command. See “External hardware daemon communication log” on page 561.
8. S1 communication log, if one has been created using the **hmadm** command. See “S1 communication log” on page 561
9. Hardware monitor log. See “Hardware monitor log” on page 561.
10. The file **/var/adm/SPlogs/SPdaemon.log**.

Diagnostic procedures

These Diagnostic Procedures test the installation, configuration, and operation of all servers.

Installation verification tests

Use these tests to check that the server has been installed properly.

Installation test 1 - Verify the ssp.basic file set

This test verifies that the **ssp.basic** file set has been installed correctly. All server device drivers are included in the **ssp.basic** file set. Issue this **lslpp** command on the control workstation:

```
lslpp -l ssp.basic
```

The output is similar to the following:

```
Path: /usr/lib/objrepos
 ssp.basic      3.1.0.8  COMMITTED  SP System Support Package

Path: /etc/objrepos
 ssp.basic      3.1.0.8  COMMITTED  SP System Support Package
```

Good results are indicated if entries for **ssp.basic** exist. Proceed to “Installation test 2 - Check external hardware daemon”.

Error results are indicated in all other cases. Try to determine why the file set was not installed, and either install it, or contact the IBM Support Center.

Installation test 2 - Check external hardware daemon

This test verifies that the appropriate external hardware daemon and it’s directory have been created on the control workstation. Issue these commands on the control workstation:

1. `ls -l /usr/lpp/ssp/install/bin/s70d`

   The output is similar to the following:

   ```
   -r-x------  1 bin  bin  47166 Sep 15 13:00 /usr/lpp/ssp/install/bin/s70d
   ```
2. `ls -l /var/adm/SPlogs/spmon | grep s70d`

   The output is similar to the following:

   ```
    drwxr-xr-x  2 bin  bin   512 Sep 15 13:01 s70d
   ```

Good results are indicated if output for all the commands is similar to the examples provided. Proceed to “Configuration test 1 - Check SDR Frame object” on page 564.

Error results are indicated in all other cases. Record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 562 and contact the IBM Support Center.

Configuration verification tests

Use these tests to check that all servers have been configured properly.

Configuration test 1 - Check SDR Frame object

This test verifies that the SDR **Frame** object was created. The configuration data for all servers must reside in the SDR **Frame** object. On the control workstation, issue the command:

```
splstdata -f
```

For each server you should see a line of output similar to the following. The numbers may be different.

```
 frame#    tty      s1_tty    frame_type  hardware_protocol
 -----------------------------------------------------------
   3    /dev/tty2  /dev/tty3      ""          SAMI
```

Good results are indicated if all of the following are true:

1. One frame entry exists for every server installed on your SP system. If an entry does not exist for one of your servers, the **Frame** object has not been created.
2. The tty (serial port for SAMI communication) and s1_tty (serial port for S1 communication) values are correct. If either of these values is incorrect, the **Frame** object has not been created correctly.
3. The **hardware_protocol** is correct. The **hardware_protocol** value must be set to SAMI (Service and Manufacturing Interface) - communication protocol for an S70, S7A and S80 server.

   Note: If **hardware_protocol** is set to SP, the System Monitor will attempt to send SP frame supervisor commands to the server's service processor. The service processor does not understand this protocol and can become hung. If this happens, it becomes impossible to control the server, either from the SP system, or physically from the operator panel.

If all of these conditions are true, proceed to "Configuration test 2 - Check SDR Node object".

Error results are indicated if not all of these conditions are true. Attempt to fix the SDR data by issuing the **spframe** command with the appropriate parameters, or contact the IBM Support Center. For a description of the **spframe** command, refer to *PSSP: Command and Technical Reference*.

Repeat this test after issuing the **spframe** command. If the test still fails, record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 562 and contact the IBM Support Center.

Configuration test 2 - Check SDR Node object

This test verifies that the SDR **Node** object (for s70) was created. The configuration data for all servers must reside in the appropriate SDR object. For each S70, S7A and S80 server, issue this command on the control workstation:

```
splstdata -n -l N
```

where *N* is the node number of the server. If you do not know the node number, issue the command: **spmon -G -d** to determine node numbers.

For each S70, S7A and S80 server, you should see output similar to the following, which is output from **splstdata -n -l 33**:

```
node# frame# slot# slots  initial_hostname  reliable_hostname  dcehostname
   default_route   processor_type processors_installed description
--------------------------------------------------------------------------
  33      3     1     1  wild3n01.ppd.pok   wild3n01.ppd.pok  ""
     9.114.130.130              MP             4 7017-S70
```

Good results are indicated if an appropriate entry exists for each server in the system. Proceed to “Configuration test 3 - Check SDR Syspar_map object”.

Error results are indicated in all other cases. Attempt to fix the SDR data by issuing the **spframe** command with the appropriate parameters. If a particular entry exists, but contains incorrect information, you must first delete that **Frame** object by issuing the **spdelfram** command. For a description of these commands, refer to *PSSP: Command and Technical Reference*.

Repeat this test after issuing the **spframe** command. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Configuration test 3 - Check SDR Syspar_map object

This test verifies that the SDR **Syspar_map** object for all servers was created correctly. The switch port number for the server is stored in the **Syspar_map** object of the SDR. The switch port number for the server is either defined for the SP-attached server using the **spframe** command, or can be optionally automatically assigned by the SDR configuration command for clustered enterprise servers. On the control workstation, issue the command once for each server:

```
SDRGetObjects Syspar_map node_number==N  switch_node_number
```

where *N* is the node number of the server.

If you do not know the node number, issue the command: **spmon -G -d** to determine node numbers. For each command, you should see output similar to the following, which is output from **SDRGetObjects Syspar_map node_number==17 switch_node_number**:

```
  switch_node_number
          5
```

Good results are indicated if the switch port number that is returned matches the value requested when the server was originally defined. For systems with an SP Switch, this should be the switch port number associated with the port in which the server is cabled to the SP Switch. For systems without a switch, this should be any unused valid switch port number on the system. For clustered systems, this can be any unused value in the range 0 to 511. Proceed to “Operational test 1 - Check hardmon status” on page 566.

Error results are indicated if no entry exists for the node, or the returned value is incorrect. Attempt to fix the SDR data by issuing the **spframe** command with the appropriate parameters. If a **Frame** object already exists for this server, you must first delete that **Frame** object by issuing the **spdelfram** command. For a description of these commands, see *PSSP: Command and Technical Reference*. For information on assigning valid switch port numbers for all servers, see *PSSP: Planning Volume 2*.

Repeat this test after issuing the **spframe** command. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562, and contact the IBM Support Center.

Operational verification tests

Use these tests to check that all servers are operating properly.

Operational test 1 - Check hardmon status

This test verifies that the System Monitor (hardmon) is active and running correctly. External hardware daemons cannot run if hardmon is not running. Issue the commands:

1. `lssrc -s hardmon`

   The output is similar to the following:

   ```
    Subsystem         Group            PID     Status
     hardmon                           42532   active
   ```

2. `ps -ef | grep hardmon`

   The output is similar to the following:

   ```
    root 42532  5966  0  Sep 15   0  9:42 /usr/lpp/ssp/bin/hardmon -r 5
   ```

Good results are indicated if all of the following are true:

1. In the **lssrc** output, look for the entry whose `Subsystem` is `hardmon`. The `Status` column should be `active`.
2. In the **ps** output, verify that the hardmon daemon uses the **-r** flag and that the argument is `5`. This means that the hardmon daemon polls each frame supervisor, including external hardware daemons, for state information every five seconds. This is the default. If the hardmon daemon uses a value other than `5` for the argument to the **-r** flag, it is not running as IBM recommends.

If these conditions are met, proceed to “Operational Test 2 - Check external hardware daemons”.

Error results are indicated if these conditions are not met. To determine why hardmon is not running, or why the argument to the **-r** flag is not `5`, refer to “Chapter 21. Diagnosing System Monitor problems” on page 317.

Repeat this test, after taking any action suggested in “Chapter 21. Diagnosing System Monitor problems” on page 317. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Operational Test 2 - Check external hardware daemons

This test verifies that the **s70d** daemons are running. If you have one or more S70, S7A or S80 servers, whether they are SP-attached or clustered enterprise servers, issue the command:

```
ps -ef | grep s70d
```

For s70, S7A and S80 servers, a line of output for each server is similar to the following:

```
root 23384 42532 1 Sep 15  2 79:08 /usr/lpp/ssp/install/bin/s70d
                  -d 0 5 1 7 /dev/tty2 /dev/tty3
```

Good results are indicated if all of the following are true:

1. There is a line of output, as in this example, for each server.
2. The parameters of the command are correct. This is a description of each parameter from left to right:
   a. This parameter will always be **-d**.
   b. This parameter is the argument to the **-d** parameter. It is an integer in which each bit (of the binary representation) represents a particular debug option. The external hardware daemon, at the time it is created by the hardmon daemon, inherits this parameter from the current value of the hardmon daemon. In the example output, the value is 0 because no debug options were set in the hardmon daemon at the time the external hardware daemons were created.
   c. This parameter is the frame number of the server.
   d. This parameter is the node number of the server. It should always be 1.
   e. This parameter is the file descriptor of the hardmon side of the socket pair used for two-way communication. The value of this parameter will be whatever the next available file descriptor was at the time the external hardware daemon was created. Any integer value would be considered correct.
   f. This parameter is the SAMI communication port.
   g. This parameter is the S1 communication port. Note that for the s70 type, this tty must be **different than the tty** in the previous parameter.

To help verify the correctness of the last two parameters, issue the command:

```
splstdata -f
```

which produces output that shows what the communication ports are expected to be.

If you receive **good results**, proceed to “Operational test 3 - Check frame responsiveness”.

Error results are indicated if any of these conditions are not met. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Operational test 3 - Check frame responsiveness

This test verifies that the frames are responding. External hardware daemons cannot run properly if their frames are not responding. To verify that a particular frame is responding, issue the following command on the control workstation:

```
hmmon -GQv controllerResponds F:0
```

where *F* is the frame number of the server that you are checking. Repeat this test for each server in the system.

The output is similar to the following:

```
frame F, slot 00:
   TRUE  frame responding to polls
```

Good results are indicated if the value is TRUE for each server. Proceed to “Operational test 4 - Check SAMI communications”.

If the value is FALSE, or you do not get any output, the test may have encountered an error. During normal operation, this value may occasionally switch to FALSE, which may simply mean that the daemon happens to be busy and cannot respond to an individual System Monitor request in a timely manner. Therefore, if you get a FALSE value, repeat the **hmmon** command several more times, waiting at least five seconds between invocations. If the value is consistently FALSE after several attempts, assume this to be **error results**.

In the case of **error results**, perform these steps:

1. Verify that the tty cables are properly connected to the server. An S70, S7A and S80 type has two cables.
2. Verify that the server itself is operating properly.

Repeat this test after performing these steps. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Operational test 4 - Check SAMI communications

This test verifies that the SAMI communication is available. For each server, issue the following command twice on the control workstation, waiting at least five seconds between the two commands:

```
hmmon -GQs F:0,1
```

where *F* is the frame number of the server that you are checking. Repeat this test for each server in the system.

This is an example output for an s70 in frame 3:

```
  3   0  nodefail1           FALSE    0x8802  node 01 I2C not responding
  3   0  nodeLinkOpen1       FALSE    0x8813  node 01 serial link open
  3   0  diagByte                0    0x8823  diagnosis return code
  3   0  timeTicks           38701    0x8830  supervisor timer ticks
  3   0  type                    2    0x883a  supervisor type
  3   0  codeVersion           769    0x883b  supervisor code version
  3   0  daemonPollRate          5    0x8867  hardware monitor poll rate
  3   0  controllerResponds  TRUE     0x88a8  frame responding to polls
  3   1  nodePower           TRUE     0x944a  DC-DC power on
  3   1  serialLinkOpen      FALSE    0x949d  serial link is open
  3   1  DPOinProgress       FALSE    0x950b  delayed power off active
  3   1  SRChasMessage       FALSE    0x9512  SRC contains a message
  3   1  SPCNhasMessage      FALSE    0x9513  SPCN contains a message
  3   1  LCDhasMessage       FALSE    0x9506  LED/LCD contains a message
  3   1  src                 BLANK    0x9510  System Reference Code
  3   1  spcn                BLANK    0x9511  System Power Cntl Network
  3   1  hardwareStatus         72    0x94f3  hardware status byte
  3   1  diagByte               15    0x9423  diagnosis return code
  3   1  timeTicks           23850    0x9430  supervisor timer ticks
  3   1  type                   10    0x943a  supervisor type
  3   1  codeVersion           769    0x943b  supervisor code version
  3   1  lcd1                BLANK    0x94f4  LCD line 1
  3   1  lcd2                BLANK    0x94f5  LCD line 2
```

Good results are indicated if all of the following are true:

1. For each server, the value for `timeTicks`, for both slots 0 and 1, increases from the first invocation of the **hmmon** command to the second invocation of the **hmmon** command. The slot number is indicated by the second column of output.
2. For each server, the value for `nodefail1` is FALSE. Note that, during normal operation, this value may occasionally switch to TRUE. This may simply mean that the daemon happens to be busy and cannot respond to an individual System Monitor request in a timely manner. Therefore, if you get a TRUE value, repeat the **hmmon** command several more times, waiting at least five seconds in between invocations, before concluding that this test has failed.
3. For each server, information for slot 1 is displayed. If it is not, this indicates that the SAMI communication is not available. The slot number is indicated by the second column of output.

In this case, proceed to “Operational test 5 - Check S1 communications”.

Error results are indicated in all other cases. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Operational test 5 - Check S1 communications

This test verifies that the S1 communication is available. For each other server, issue the following command on the control workstation:

```
s1term -G F 1
```

where *F* is the frame number of the server that you are checking. Repeat this test for each server in the system.

Good results are indicated if the AIX login prompt is displayed. If you do not see the login prompt after issuing the **s1term** command, try typing the enter key a second time. If you still do not see the login prompt, consider this to be **error results**.

To correct the problem, perform these steps:

1. Verify that the s1 tty cable is properly connected to the server.
2. Verify that the server itself is operating properly.

Repeat this test after performing these steps. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 562 and contact the IBM Support Center.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the PSSP software for all servers. Locate the symptom and perform the action described in this table.

Table 86. SP-attached Server and Clustered Enterprise Server symptoms

Symptom	Recovery
The **ssp.basic** file set is not installed.	See “Action 1 - Install ssp.basic file set” on page 570.
An external hardware daemon or it’s directory have not been created on the control workstation.	See “Action 2 - Check permissions of ssp.basic” on page 570.

Table 86. SP-attached Server and Clustered Enterprise Server symptoms (continued)

Symptom	Recovery
The SDR **Frame** object was not created for a server.	See “Action 3 - Correct SDR Frame object”.
The SDR **Node** object was not created for a server.	See “Action 4 - Correct SDR Node object” on page 571.
The hardmon daemon is not running.	See “Action 5 - Correct System Monitor polling interval” on page 571.
The external hardware daemon is not running or not responding.	See “Action 6 - Investigate external hardware daemon failure” on page 572.
The SAMI communication is not available.	See “Action 7 - Restore SAMI communication” on page 573.
The S1 communication is not available.	See “Action 8 Restore S1 communication” on page 574.
The switch port number for the server is not set correctly.	See “Action 9 - Correct switch port number in SDR Syspar_map object” on page 574.

Actions

Action 1 - Install ssp.basic file set

Perform “Installation test 1 - Verify the ssp.basic file set” on page 563 to verify that this is a problem. Install the **ssp.basic** file set using the **installp** command. Repeat “Installation test 1 - Verify the ssp.basic file set” on page 563.

Action 2 - Check permissions of ssp.basic

Perform “Installation test 2 - Check external hardware daemon” on page 563 to verify that this is a problem. Install the **ssp.basic** file set using the **installp** command. If any of the permission or file attributes do not match what is shown in “Installation test 2 - Check external hardware daemon” on page 563, issue the **chmod** or **chown** commands, as appropriate, to correct the attributes. Repeat “Installation test 1 - Verify the ssp.basic file set” on page 563 and “Installation test 2 - Check external hardware daemon” on page 563.

Action 3 - Correct SDR Frame object

Perform test “Configuration test 1 - Check SDR Frame object” on page 564 to verify that this is a problem. Perform these steps:

1. If the **splstdata -f** command that was run in “Configuration test 1 - Check SDR Frame object” on page 564 returns an error message similar to:

   ```
   splstdata: 0022-001 The repository cannot be accessed. Return code was 80.
   ```

   refer to “Chapter 14. Diagnosing SDR problems” on page 125 to determine why the SDR cannot be accessed.
2. If you can successfully access the SDR, create a **Frame** object for the one that is missing by issuing the **spframe** command with the appropriate parameters. For a description of the **spframe** command, refer to *PSSP: Command and Technical Reference*.
3. If there were problems creating the SDR **Frame** object, investigate why the **SDR_config** command was unable to create the **Frame** object. Check the **SDR_config** log file, **/var/adm/SPlogs/sdr/SDR_config.log** for error messages. If there are error messages for the **Frame** object creation, refer to

“Configuration test 5 - Check SDR Switch class” on page 325 through “Configuration test 7 - Check SDR NodeExpansion class” on page 326.

4. If the **Frame** object was successfully created, but the server entry, output by the **splstdata -f** command, does not have correct information in it’s `tty` or `s1_tty` column, issue the **spframe** command with the correct values.
5. If the **Frame** object was successfully created, but the server entry, output by the **splstdata -f** command, does not have correct information in it’s `hardware_protocol` column, you must first issue the **spdelfram** command. Then, create a new definition by issuing the **spframe** command with the correct values.

   For an S70, S7A and S80 type server, the value must be SAMI.

After correcting the problem, repeat “Configuration test 1 - Check SDR Frame object” on page 564.

Action 4 - Correct SDR Node object

Perform “Configuration test 2 - Check SDR Node object” on page 564 to verify that this is a problem. Perform these steps:

1. If either of the two **splstdata** commands that were run in “Configuration test 2 - Check SDR Node object” on page 564 returns an error message similar to:

   ```
    splstdata: 0022-001 The repository cannot be accessed. Return code was 80.
   ```

   refer to “Chapter 14. Diagnosing SDR problems” on page 125 to determine why the SDR cannot be accessed.
2. If you can successfully access the SDR, first delete the incorrect frame definition with the **spdelfram** command, and invoke the **spframe** command with the appropriate parameters. Refer to *PSSP: Command and Technical Reference* for these commands. By issuing the **spframe** command, the hardmon and logging daemons together will create **Node** objects for S70, S7A and S80 servers.
3. If there were problems creating the **Node** or **ProcessorExtensionNode** objects, investigate why the **SDR_config** command was unable to create the objects. Check the **SDR_config** log **/var/adm/SPlogs/sdr/SDR_config.log** for error messages. If there are error messages for the object creation, refer to “Configuration test 5 - Check SDR Switch class” on page 325 through “Configuration test 7 - Check SDR NodeExpansion class” on page 326.

After correcting the problem, repeat “Configuration test 2 - Check SDR Node object” on page 564.

Action 5 - Correct System Monitor polling interval

Perform “Operational test 1 - Check hardmon status” on page 566 to verify that this is a problem. Perform these steps:

1. If the hardmon daemon is not running on the control workstation, you need to start it. Issue this command:

   ```
   startsrc -s hardmon
   ```

   Then, perform “Operational test 1 - Check hardmon status” on page 566 again to determine if hardmon was started successfully.
2. If the hardmon daemon uses an incorrect polling interval, it may cause problems. The polling interval is chosen when the hardmon daemon is started

by the System Resource Controller. The value is in the **cmdargs** attribute in the hardmon ODM **SRCsubsys** object. Check the polling interval by issuing the ODM command:

```
odmget -q subsysname=hardmon SRCsubsys
```

The output is similar to the following, which is the default:

```
SRCsubsys:
   subsysname = "hardmon"
    synonym = ""
    cmdargs = "-r 5"
    path = "/usr/lpp/ssp/bin/hardmon"
    uid = 0
    auditid = 0
    standin = "/dev/console"
    standout = "/dev/console"
    standerr = "/dev/console"
    action = 1
    multi = 0
    contact = 2
    svrkey = 0
    svrmtype = 0
    priority = 20
    signorm = 15
    sigforce = 15
    display = 1
    waittime = 15
    grpname = ""
```

If the `cmdargs` attribute is not ″**-r 5**″, correct this by issuing the following command:

```
chssys -s hardmon -a "-r 5"
```

Then reissue the **odmget** command to verify that the new **cmdargs** attribute is ″**-r 5**″.

After correcting the problem, repeat “Operational test 1 - Check hardmon status” on page 566.

Action 6 - Investigate external hardware daemon failure

Perform “Operational Test 2 - Check external hardware daemons” on page 566 to determine if the external hardware daemon is running. Perform “Operational test 3 - Check frame responsiveness” on page 567 to determine if the external hardware daemon is responding. If either test produces **error results**, perform these steps:

1. Several components of PSSP, involved with the operation of the servers, write data to log files. Check these log files and take appropriate action:
   - The daemon log file. Refer to “Daemon log” on page 560.
   - The hardware monitor log. Refer to “Hardware monitor log” on page 561.
   - The SP hardware log. Refer to “SP hardware log” on page 562.
   - The AIX Error log.

     The same messages that are in the SP hardware log are also found in the AIX Error log. To obtain full details of all SP hardware messages in this log, issue the command:

     ```
     errpt -aN sphwlog
     ```

You may want to redirect the output to a file, because there could be a large amount of output.

2. If one of the external hardware daemons is not running, but it should be, check to see if a core dump was created. Refer to “Dump information” on page 562.
3. If the System Monitor (hardmon) daemon is running, but an external hardware daemon is not running or not responding, issue the following command to start the external hardware daemon:

   ```
   hmcmds -G boot_supervisor F:0
   ```

   where *F* is the frame number of the server. This notifies the System Monitor that the external hardware daemon has stopped. The System Monitor then starts the daemon.
4. If you have attempted to start an external hardware daemon, and it still does not start, issue the following command to stop and restart the System Monitor daemon (hardmon):

   ```
   hmreinit
   ```

   The System Monitor daemon (hardmon) will be restarted by the System Resource Controller, and the daemon will then restart all of the external hardware daemons. It also causes the **SDR_config** command to run, updating the SDR as necessary.
5. If you have attempted to start an external hardware daemon by running the previous action, and it still does not start, issue the command:

   ```
   stopsrc -s hardmon
   ```

   to stop the System Monitor daemon (hardmon), and then issue the command:

   ```
   splstdata -f
   ```

   to see what ttys (tty and s1_tty) are needed by your external hardware daemons. Refer to “Configuration test 1 - Check SDR Frame object” on page 564 for typical output. For an S70, S7A and S80 servers, if one or more of it’s ttys has a corresponding entry in the **/etc/locks/** directory, delete these entries and repeat this step. A server may be prevented from starting if either of it’s two required ttys are locked.

After correcting the problem, repeat “Operational Test 2 - Check external hardware daemons” on page 566 and “Operational test 3 - Check frame responsiveness” on page 567.

Action 7 - Restore SAMI communication

Perform “Operational test 4 - Check SAMI communications” on page 568 to determine if the SAMI communication is available. If you receive an **error result**, perform these steps:

1. Verify that the SAMI (S70, S7A, S80) communication cable is not unplugged or loose, and that it is plugged into the correct tty socket of the control workstation.
2. Verify the tty definition on the control workstation. The **Enable LOGIN** characteristic must be set to `disable`. Use **smitty** as follows:

   ```
   TYPE   :   smitty
   SELECT :   devices
   SELECT :   TTY
   ```

```
SELECT :   Change / Show Characteristics of a TTY
select the TTY of interest and press ENTER
check the "Enable LOGIN" value
```

3. Verify that the serial port adapter on the control workstation does not have a hardware error, by checking the AIX Error Log.
4. Verify that the server is operating properly. For more information, refer to the manual for the specific server.

After correcting the problem, repeat “Operational test 4 - Check SAMI communications” on page 568.

Action 8 Restore S1 communication

Perform “Operational test 5 - Check S1 communications” on page 569 to verify that this is a problem. If you receive **error results**, perform these steps:

1. Verify that the S1 communication cable is not unplugged or loose, and that it is plugged into the correct tty socket of the control workstation.
2. Verify the S1 tty definition on both the control workstation and the server. On the control workstation, the **Enable LOGIN** characteristic must be set to `disable`. On the server itself, the **Enable LOGIN** characteristic must be set to `enable`. You can use **smitty** on the control workstation, and then **rsh** to the server. This is the **smitty** sequence:

```
TYPE   :   smitty
SELECT :   devices
SELECT :   TTY
SELECT :   Change / Show Characteristics of a TTY
select the TTY of interest and press ENTER
check the "Enable LOGIN" value
```

3. Verify that the serial port adapter on the control workstation and the server do not have a hardware error, by checking the AIX Error Log.
4. Verify that the server is operating properly. For more information, refer to the manual for the specific server.

After correcting the problem, repeat “Operational test 5 - Check S1 communications” on page 569.

Action 9 - Correct switch port number in SDR Syspar_map object

Perform “Configuration test 3 - Check SDR Syspar_map object” on page 565 to verify that the switch port number is not set correctly for the server. Perform these steps:

1. If the **SDRGetObjects Syspar_map** command that was run in “Configuration test 3 - Check SDR Syspar_map object” on page 565 returns an error message similar to:

   ```
   0025-080 The SDR routine could not connect to server.
   ```

   or some other message indicating a problem with the System Data Repository, refer to “Chapter 14. Diagnosing SDR problems” on page 125.
2. If you can successfully access the SDR, delete the **Frame** object for the server, if one exists, by issuing the **spdelfram** command. For a description of the **spdelfram** command, see *PSSP: Command and Technical Reference*.
3. Create a new **Frame** object for the server by issuing the **spframe** command, with the appropriate parameters. The **-n** option is used to specify the switch port number for the server. If this is a system of clustered enterprise servers (no SP frames or SP Switches), you do not need to specify the switch port number. In

this case, the SDR configuration command which is invoked during the **spframe -r yes** processing will automatically assign a valid value for you.

Refer to *PSSP: Planning Volume 2* for information on determining a valid switch port number, and for situations where you may not wish to have the SDR configuration command automatically assign one for you in a clustered enterprise server system. For a description of the **spframe** command, see *PSSP: Command and Technical Reference*.

4. If there were problems creating the **Frame** or **Syspar_Map** SDR objects, investigate why the **SDR_confg** command was unable to create the object by checking the **SDR_config** log file, **/var/adm/SPlogs/sdr/SDR_config.log**, for error messages. If there are error messages for the **Frame** or **Syspar_map** objects, refer to "Configuration test 5 - Check SDR Switch class" on page 325 through "Configuration test 7 - Check SDR NodeExpansion class" on page 326.

After correcting the problem, repeat "Configuration test 3 - Check SDR Syspar_map object" on page 565.

Chapter 39. Diagnosing IBM @server pSeries 690 problems

This chapter discusses diagnostic procedures and failure responses for the IBM @server pSeries 690 software of PSSP. The list of known error symptoms and the associated responses are in the section “Error symptoms, responses, and recoveries” on page 588. A list of the information to collect before contacting the IBM Support Center is in “Information to collect before contacting the IBM Support Center” on page 580.

IBM @server pSeries 690 servers can be connected to SP systems that contain an SP Switch, an SP Switch2, or those that have no switch. The terms **the server** and **all servers** are used in this chapter in place of **IBM @server pSeries 690**.

Related documentation

The following publications provide information about the PSSP software that controls all servers:

1. *PSSP: Command and Technical Reference*

   Entries for these commands:

   - hmadm
   - hmcmds
   - hmmon
   - sphmcid
   - spdelhmcid
   - SDR_config
   - spdelfram
   - spframe
   - splstdata
   - s1term
2. *PSSP: Messages Reference*

   The chapter ″0026 - System Monitor Messages″ contains messages issued by all servers.
3. *PSSP: Administration Guide*

   See the Appendixes ″SP Daemons″ and ″The System Data Repository″.

Requisite function

This is a list of the software directly used by the PSSP software related to all servers. Problems within the requisite software may manifest themselves as error symptoms in this software. If you perform all the diagnostic routines and error responses listed in this chapter, and still have problems with the server, consider these components as possible sources of the error. They are listed with the most likely candidate first, least likely candidate last.

- The System Monitor (**hardmon**) must be running on the control workstation.
- The external hardware daemons, which consist of two interfaces: the frame supervisor interface and the node supervisor interface.

  The frame supervisor interface is responsible for keeping the state data in the frame’s packet current and formatting the frame packet for return to the hardmon daemon.

The node supervisor interface is responsible for keeping the state data in the node's packet current.

- The System Data Repository (SDR) subsystem

  The System Monitor requires the SDR subsystem. For more information, refer to "Chapter 14. Diagnosing SDR problems" on page 125.

- SP System Security Services

  Principal and group names for DCE entities use the default SP chosen names. These may not be the actual names on the system if you have overridden them using the **spsec_overrides** file.

Error information

Logs

There are many logs that record error and status information. Examine the following logs: daemon log, SDR configuration log, external hardware daemon communication log, S1 communication log, hardware monitor log, and SP hardware log.

Daemon log

This file contains the daemon's error information. This file is located on the control workstation in **/var/adm/SPlogs/spmon/hcmd/hmcd[***ipaddress***].log.***ddd* for the **hmc** daemon where *ddd* is the Julian date, and *ipaddress* is the IP address of the HMC (Hardware Management Console) that the servers are connected to, as defined in the SDR.

SDR configuration log

This log file contains information from the SDR configuration command **SDR_config**. The **SDR_config** command is automatically run under these conditions:

- Whenever the hardware monitor daemon is started.
- When hardware changes are detected on the system.
- In response to requests from the **spframe -r yes** command.
- In response to requests from the **spdelfram** command.

If **SDR_config** detects system configuration problems, error messages are written to the file located on the control workstation in: **/var/adm/SPlogs/sdr/SDR_config.log**.

External hardware daemon communication log

This file contains information on data sent to, and received from, the HMC that the server is connect to. This file must be created manually on the control workstation as **/var/adm/SPlogs/spmon/hcmd/hmcd[***ipaddress***].java_trace** for the **hmc** daemon, where *ipaddress* is the IP address of the HMC (Hardware Management Console) that the servers are connected to.

Also contained in this log file is informational, error and debug information. Examples of informational messages are timestamps for when the daemon started or stopped. An example of an error is the failure of the daemon to connect to the HMC. An example of debug information (enabled by the **hmadm** command), is a list of requests sent from the System monitor **(hardmon**) to the external hardware daemon.

For the **hmc** daemon, create this log whenever the IBM @server pSeries 690 hardware appears to not be communicating with the **hmc** daemon. Issue this command:

```
hmadm -d java setd
```

To stop the log, issue this command:

```
hmadm -d java cleard
```

S1 communication log

This log file contains information on data sent to and data received from, the server's serial port. This file must be created manually on the control workstation. This file is located in: **/var/adm/SPlogs/spmon/hmc.***frame.node***.s1data_dump**, where *frame* and *node* are the frame and node numbers of the server, respectively. Create this log whenever the IBM @server pSeries 690 hardware serial interface appears to not be communicating with the **hmc** daemon. Issue this command:

```
hmadm -d s1data setd
```

To stop the log, issue this command:

```
hmadm -d s1data cleard
```

Hardware monitor log

This log file contains informational, error, and debug information. Examples of informational messages are timestamps for when daemons are started or stopped. An example of an error is a communication error between the System monitor (**hardmon**) an external hardware daemon. An example of debug information (enabled by the **hmadm** command), is a list of requests sent from the System monitor (**hardmon**) to an external hardware daemon. This file is located on the control workstation. The file is located in: **/var/adm/SPlogs/spmon/hmlogfile.***ddd*, where *ddd* is the Julian date.

SP hardware log

This log file contains information generated by system daemons, including hardware errors. SP hardware problems have a resource name of **sphwlog**. This file is located on the control workstation in: **/var/adm/SPlogs/SPdaemon.log**.

Dump information

The external hardware daemons create standard AIX core dumps when they fail. Any core file created should be saved in a safe place at the time of the failure, so that it is not overwritten if another failure occurs. The IBM Support Center can then examine the file at a later time. The dump is located on the control workstation in **/var/adm/SPlogs/spmon/hmcd/core** for the IBM @server pSeries 690.

Trace information

ATTENTION - READ THIS FIRST

Do **not** activate this trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

Each external hardware daemon provides facilities that allow the tracing of certain aspects of its internal function, as well as its communication with the SP Frame Supervisor. This is done using the **hmadm** command. Use this trace only when directed by the IBM Support Center. The *PSSP: Command and Technical Reference* gives details about running the trace and obtaining the output.

Information to collect before contacting the IBM Support Center

Collect this information before calling the IBM Support Center:

1. Any command output that seems to be helpful. For example, the output of any failing diagnostic procedures that were run.
2. The authentication method in use. Issue this command on the control workstation:

   ```
   splstdata -p
   ```

   . The entry **ts_auth_methods** lists the authentication methods.
3. AIX error log. On the control workstation, issue the command:

   ```
   LANG=C errpt -a > /tmp/AIXerrlog
   ```
4. Daemon core dump, if it exists. See “Dump information” on page 579.
5. Daemon log. See “Daemon log” on page 578.
6. SDR Configuration Log. See “SDR configuration log” on page 578.
7. External hardware daemon communication log, if one has been created using the **hmadm** command. See “External hardware daemon communication log” on page 578.
8. S1 communication log, if one has been created using the hmadm command. See “S1 communication log” on page 579.
9. Hardware monitor log. See “Hardware monitor log” on page 579.
10. The SP hardware log. See “SP hardware log” on page 579.

Diagnostic procedures

These Diagnostic Procedures test the installation, configuration, and operation of all servers.

Installation verification tests

Use these tests to check that the server has been installed properly.

Installation test 1 -Verify the ssp.basic file set

This test verifies that the **ssp.basic** file set has been installed correctly. All server device drivers are included in the **ssp.basic** file set. Issue this **lslpp** command on the control workstation:

```
lslpp -l ssp.basic
```

The output is similar to the following:

```
Path: /usr/lib/objrepos
ssp.basic 3.1.0.8 COMMITTED SP System Support Package

Path: /etc/objrepos
ssp.basic 3.1.0.8 COMMITTED SP System Support Package
```

Good results are indicated if entries for **ssp.basic** exist. Proceed to "Installation test 2 - Check external hardware daemon".

Error results are indicated in all other cases. Try to determine why the file set was not installed, and either install it, or contact the IBM Support Center.

Installation test 2 - Check external hardware daemon

This test verifies that the appropriate external hardware daemon and it's directory have been created on the control workstation. Issue these commands on the control workstation:

1. `ls -l /usr/lpp/ssp/install/bin/hmcd`

   The output is similar to the following:

   ```
   -r-x------ 1 bin bin 47166 Sep 15 13:00 hmcd
   ```
2. `ls -l /var/adm/SPlogs/spmon | grep hmcd`

   The output is similar to the following:

   ```
   drwxr-xr-x 2 bin bin 512 Sep 15 13:01 hmcd
   ```
3. `ls -l /usr/lpp/ssp/install/bin/HMCD.class`

   The output is similar to the following:

   ```
   -rw-r--r-- 1 bin bin 26927 Sep 11 01:03 HMCD.class
   ```
4. `ls -l /usr/lpp/ssp/lib/libHMCD.so`

   The output is similar to the following:

   ```
   -rwxr-x--x 1 bin bin 17680 Sep 15 13:00 libHMCD.so
   ```

Good results are indicated if output for all the commands is similar to the examples provided. Proceed to "Installation test 3 - Check external hardware daemon components".

Error results are indicated in all other cases. Record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 580 and contact the IBM Support Center.

Installation test 3 - Check external hardware daemon components

This test verifies that the appropriate external hardware daemon components and their directories have been created on the control workstation. Issue these commands on the control workstation:

1. `ls -l /usr/java130/jre/lib/ext/xerces.jar`

   The output is similar to the following:

```
-r--r--r-- 1 bin bin 1521373 Feb 12 13:20 xerces.jar
```

2. `ls -l /opt/freeware/cimom/org/snia/wbem/client/CIMClient.class`

   The output is similar to the following:

   ```
   -rwxr-xr-x 1 bin bin 8507 Jul 23 15:20 CIMClient.class
   ```

Good results are indicated if output for all the commands is similar to the examples provided. Proceed to "Configuration test 1 - Check SDR Frame object".

Error results are indicated in all other cases. Record all relevant information, see ″"Information to collect before contacting the IBM Support Center" on page 580 and contact the IBM Support Center.

Configuration verification tests

Use these tests to check that all servers have been configured properly.

Configuration test 1 - Check SDR Frame object

This test verifies that the SDR Frame object was created. The configuration data for all servers must reside in the SDR Frame object. On the control workstation, issue the command:

```
splstdata -f
```

For each server you should see a line of output similar to the following. The numbers may be different.

```
                                 List Frame Database Information
  frame#    tty          s1_tty   frame_type        hardware_protocol control_ipaddrs domain_name
 ---------- ------------ -------- ----------------- ----------------- --------------- ------------
 1          /dev/tty0    ""       switch            SP                ""              ""
 2          ""           ""       ""                HMC               9.114.62.123    huntley
```

Good results are indicated if all of the following are true:

1. One frame entry exists for every server installed on your SP system. If an entry does not exist for one of your servers, the Frame object has not been created.
2. The `hardware_protocol` is correct. The `hardware_protocol` value must be set to `HMC` ((Hardware Management Console) - communication protocol for an IBM @server pSeries 690 server.
3. The `control_ipaddrs` is correct. The `control_ipaddrs` represents the IP address of the HMC that the server is connect to.
4. The `domain_name` is correct. The `domain_name` represents the system name assigned to the IBM @server pSeries 690 server through the HMC Partition Management interface. You can verify the name by viewing the properties for the IBM @server pSeries 690 server directly using the HMC WebSM interface.

If all of these conditions are true, proceed to "Configuration test 2 - Check HMC password files" on page 583.

Error results are indicated if one or more of these conditions are not true. Attempt to fix the SDR data by issuing the **spframe** command with the appropriate parameters, or contact the IBM Support Center. For a description of the **spframe** command, refer to *PSSP: Command and Technical Reference*.

Repeat this test after issuing the **spframe** command. If the test still fails, record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 580 and contact the IBM Support Center.

Configuration test 2 - Check HMC password files

This test verifies that for each unique HMC IP address in your system, a corresponding password file has been created on the control workstation. On the control workstation, issue the command:

```
sphmcid
```

For each HMC in your system, you should see a line of output similar to the following. The numbers may be different:

```
9.114.58.22       hmcadmin
9.114.62.123      hmcadmin
```

Good results are indicated if for each HMC IP address displayed under the `control_ipaddrs` heading in “Configuration test 1 - Check SDR Frame object” on page 582, there is a corresponding HMC IP address displayed as a result of issuing the **sphmcid** command. Proceed to “Configuration test 3- Check SDR Syspar_map object”.

Error results are indicated if no entry exists for one or more HMC IP address. Attempt to create a password file on the control workstation for the missing HMC IP address. For a description of the **sphmcid** command , refer to *PSSP: Command and Technical Reference*.

Repeat the test after issuing the **sphmcid** command. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580, and contact the IBM Support Center.

Configuration test 3- Check SDR Syspar_map object

This test verifies that the SDR Syspar_map object for all servers was created correctly. The switch port number for the server is stored in the Syspar_map object of the SDR. If the server is attached to the SP Switch or an SP system without a switch, the switch port number for the server is defined for the SP-attached server using the **spframe** command. If the server is attached to the SP Switch2 or a clustered enterprise server system without a switch, the switch port number can be optionally automatically assigned by the SDR configuration command. On the control workstation, issue the command once for each server:

```
SDRGetObjects Syspar_map node_number==N switch_node_number
```

where *N* is the node number of the server.

If you do not know the node number, issue the command: **spmon -G -d** to determine node numbers. For example, the command:

```
SDRGetObjects Syspar_map node_number==17 switch_node_number
```

produces output similar to the following:

```
switch_node_number
5
```

Good results are indicated if the switch port number that is returned matches the value requested when the server was originally defined. For systems with an SP Switch, this should be the switch port number associated with the port in which the server is cabled to the SP Switch. For systems without a switch, this should be any unused valid switch port number on the system. For systems with an SP Switch 2 or for clustered systems, this can be any unused value in the range 0 to 511. Proceed to “Operational test 1 - Check hardmon status” on page 584.

Error results are indicated if no entry exists for the node, or the returned value is incorrect. Attempt to fix the SDR data by issuing the **spframe** command with the appropriate parameters. If a Frame object already exists for this server, you must first delete that Frame object by issuing the **spdelfram** command. For a description of these commands, see *PSSP: Command and Technical Reference*. For information on assigning valid switch port numbers for all servers, see *IBM RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*.

Repeat this test after issuing the **spframe** command. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580 and contact the IBM Support Center.

Operational verification tests

Use these tests to check that all servers are operating properly.

Operational test 1 - Check hardmon status

This test verifies that the System Monitor (hardmon) is active and running correctly. External hardware daemons cannot run if hardmon is not running. Issue the commands:

1.
   ```
   lssrc -s hardmon
   ```

   The output is similar to the following:

   ```
   Subsystem Group PID   Status
   hardmon         42532 active
   ```
2.
   ```
   ps -ef | grep hardmon
   ```

   The output is similar to the following:

   ```
   root 42532 5966 0 Sep 15 0 9:42 /usr/lpp/ssp/bin/hardmon -r 5
   ```

Good results are indicated if the following are both true:

1. There is a line of output, as in this example, for each server HMC that is actively controlling your IBM @server pSeries 690 servers.
2. In the **ps** output, the **hardmon** daemon uses the **-r** flag and the argument is `5`. This means that the **hardmon** daemon polls each frame supervisor, including external hardware daemons, for state information every five seconds. This is the default. If the **hardmon** daemon uses a value other than `5` for the argument to the **-r** flag, it is not running as IBM recommends.

If these conditions are met, proceed to “Operational test 2 - Check external hardware daemons”.

Error results are indicated if these conditions are not met. To determine why **hardmon** is not running, or why the argument to the **-r** flag is not `5`, refer to “Chapter 21. Diagnosing System Monitor problems” on page 317.

Repeat this test, after taking any action suggested in “Chapter 21. Diagnosing System Monitor problems” on page 317. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580 and contact the IBM Support Center.

Operational test 2 - Check external hardware daemons

This test verifies that the **hmc** daemons are running. If you have one or more IBM @server pSeries 690 servers, issue the command:

```
ps -ef | grep hmcd
```

For IBM @server pSeries 690 servers, a line of output for each server is similar to the following:

```
root 23384 42532 1 Sep 15 2 79:08 /usr/lpp/ssp/install/bin/hmcd
-d 0 9.114.58.22 1 minnow 1 5
```

Good results are indicated if all of the following are true:

1. There is a line of output, as in this example, for each server.
2. The parameters of the command are correct. This is a description of each parameter from left to right:
   - This parameter will always be **-d**.
   - This parameter is the argument to the **-d** flag. It is an integer in which each bit of the binary representation indicates a particular debug option. The external hardware daemon, at the time it is created by the **hardmon** daemon, inherits this parameter from the current value of the **hardmon** daemon. In the example output, the value is 0 because no debug options were set in the **hardmon** daemon at the time the external hardware daemons were created.

     These debug bit flags are for IBM service use. They are enabled with the **hmadm** command
   - This is the IP address of the HMC the servers are connected to. This value is obtained from the SDR control_ipaddrs attribute for the given frame.
   - This parameter is an integer that specifies the number of domain names, frame numbers, and file descriptor groups that follow. In the example output, the value 1 indicates that one group of these three parameters follows.
   - This parameter is the domain name of the server. This value is obtained from the SDR domain_name attribute for the given frame and should contain the system name for the server as defined through the HMC Partition Management interface.
   - This parameter is the frame number of the server.
   - This parameter is the file descriptor of the hardmon side of the socket pair used for two-way communication. The value of this parameter will be whatever the next available file descriptor was at the time the external hardware daemon was created. Any integer value is considered correct.

If you receive good results, proceed to “Operational test 3 - Check frame responsiveness”.

Error results are indicated if any of these conditions are not met. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580 and contact the IBM Support Center.

Operational test 3 - Check frame responsiveness

This test verifies that the frames are responding. External hardware daemons cannot run properly if their frames are not responding. To verify that a particular frame is responding, issue the following command on the control workstation:

```
hmmon -GQv controllerResponds F:0
```

where *F* is the frame number of the server that you are checking. Repeat this test for each server in the system.

The output is similar to the following:

```
frame F, slot 00:
TRUE frame responding to polls
```

Good results are indicated if the value is TRUE for each server. Proceed to "Operational test 4 - Check HMC communications".

If the value is FALSE, or you do not get any output, the test may have encountered an error. During normal operation, this value may occasionally switch to FALSE, which may simply mean that the daemon happens to be busy and cannot respond to an individual System Monitor request in a timely manner. Therefore, if you get a FALSE value, repeat the **hmmon** command several more times, waiting at least five seconds between invocations. If the value is consistently FALSE after several attempts, assume this to be **error results**.

In the case of **error results**, perform these steps:

1. Verify that the HMC IP address specified in the SDR is correct for the given server.
2. Verify that the server itself is operating properly.
3. Examine the daemon log. See "Daemon log" on page 578.

Repeat this test after performing these steps. If the test still fails, record all relevant information, see "Information to collect before contacting the IBM Support Center" on page 580. and contact the IBM Support Center.

Operational test 4 - Check HMC communications

This test verifies that the HMC communication is available. For each server, issue the following command twice on the control workstation, waiting at least five seconds between the two commands:

```
hmmon -GQs F:0
```

where *F* is the frame number of the server that you are checking. Repeat this test for each server in the system. This is an example output for an IBM @server pSeries 690 in frame 3:

```
3 0 nodefail1         FALSE 0x8802  node 01 I2C not responding
3 0 nodefail2         TRUE  0x8803  node 02 I2C not responding
3 0 nodefail3         TRUE  0x8804  node 03 I2C not responding
3 0 nodefail4         TRUE  0x8805  node 04 I2C not responding
3 0 nodefail5         TRUE  0x8806  node 05 I2C not responding
3 0 nodefail6         TRUE  0x8807  node 06 I2C not responding
3 0 nodefail7         TRUE  0x8808  node 07 I2C not responding
3 0 nodefail8         TRUE  0x8809  node 08 I2C not responding
3 0 nodefail9         TRUE  0x880a  node 09 I2C not responding
3 0 nodefail10        TRUE  0x880b  node 10 I2C not responding
3 0 nodefail11        TRUE  0x880c  node 11 I2C not responding
3 0 nodefail12        TRUE  0x880d  node 12 I2C not responding
3 0 nodefail13        TRUE  0x880e  node 13 I2C not responding
3 0 nodefail14        TRUE  0x880f  node 14 I2C not responding
3 0 nodefail15        TRUE  0x8810  node 15 I2C not responding
3 0 nodefail16        TRUE  0x8811  node 16 I2C not responding
3 0 nodeLinkOpen1     FALSE 0x8813  node 01 serial link open
3 0 nodeLinkOpen2     FALSE 0x8814  node 02 serial link open
3 0 nodeLinkOpen3     FALSE 0x8815  node 03 serial link open
3 0 nodeLinkOpen4     FALSE 0x8816  node 04 serial link open
3 0 nodeLinkOpen5     FALSE 0x8817  node 05 serial link open
3 0 nodeLinkOpen6     FALSE 0x8818  node 06 serial link open
3 0 nodeLinkOpen7     FALSE 0x8819  node 07 serial link open
3 0 nodeLinkOpen8     FALSE 0x881a  node 08 serial link open
3 0 nodeLinkOpen9     FALSE 0x881b  node 09 serial link open
3 0 nodeLinkOpen10    FALSE 0x881c  node 10 serial link open
3 0 nodeLinkOpen11    FALSE 0x881d  node 11 serial link open
3 0 nodeLinkOpen12    FALSE 0x881e  node 12 serial link open
3 0 nodeLinkOpen13    FALSE 0x881f  node 13 serial link open
3 0 nodeLinkOpen14    FALSE 0x8820  node 14 serial link open
```

```
3 0 nodeLinkOpen15     FALSE 0x8821  node 15 serial link open
3 0 nodeLinkOpen16     FALSE 0x8822  node 16 serial link open
3 0 CECUserName        huntley
                             0x8993  user defined CEC name
3 0 CECMode            0     0x8984  0=smp 1=partition
3 0 PowerOffPolicy     TRUE  0x8985  power off with last Lpar
3 0 CECCapability      1     0x8986  0=smp 1=lpar 2=numa
3 0 CECState           1     0x8987  0=off on init err inc con rec
3 0 diagByte           0     0x8823  diagnosis return code
3 0 timeTicks          38701 0x8830  supervisor timer ticks
3 0 type               5     0x883a  supervisor type
3 0 codeVersion        772   0x883b  supervisor code version
3 0 daemonPollRate     5     0x8867  hardware monitor poll rate
3 0 controllerResponds TRUE  0x88a8  frame responding to polls
```

Good results are indicated if all of the following are true:

1. For each server, the value for `timeTicks` increases from the first invocation of the **hmmon** command to the second invocation of the **hmmon** command. The slot number is indicated by the second column of output.
2. For each server, the value for `nodefail`*`N`* (where *N* is a configured node number) is `FALSE`. Note that, during normal operation, this value may occasionally switch to `TRUE`. This may simply mean that the daemon happens to be busy and cannot respond to an individual System Monitor request in a timely manner. Therefore, if you get a `TRUE` value, repeat the **hmmon** command several more times, waiting at least five seconds in between invocations, before concluding that this test has failed.

In this case, proceed to “Operational test 5 - Check S1 communications”.

Error results are indicated in all other cases. Record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580, and contact the IBM Support Center.

Operational test 5 - Check S1 communications

This test verifies that the S1 communication is available. For each server, issue the following command on the control workstation for one node chosen arbitrarily:

```
s1term -G F N
```

where *F* is the frame number and *N* is the node number for a logical partition in the server that you are checking. Repeat this test for each server in the system.

Good results are indicated if the AIX login prompt is displayed. If you do not see the login prompt after issuing the **s1term** command, try typing the enter key a second time. If you still do not see the login prompt, consider this to be error results. Note that the login prompt may take up to 90 seconds to appear.

To correct the problem, perform these steps:

1. Verify that the server itself is operating properly.
2. Examine the daemon log. See “Daemon log” on page 578.

Repeat this test after performing these steps. If the test still fails, record all relevant information, see “Information to collect before contacting the IBM Support Center” on page 580 and contact the IBM Support Center.

Error symptoms, responses, and recoveries

Use the following table to diagnose problems with the PSSP software for all servers.

Table 87. IBM @server pSeries 690 symptoms

Symptom	Recovery
The **ssp.basic** file set, Java 1.3, or CIMOM is not installed.	See “Action 1 - Install ssp.basic file set, Java 1.3, RPM, and CIMOM”.
An external hardware daemon or it’s directory have not been created on the control workstation.	See “Action 2 - Check permissions of ssp.basic” on page 589.
The SDR Frame object was not created for a server.	See “Action 3 - Correct SDR Frame object” on page 589.
The hardmon daemon is not running.	See “Action 4 - Correct System Monitor polling interval” on page 590.
The external hardware daemon is not running or not responding.	See “Action 5 - Investigate external hardware daemon failure” on page 590.

Actions

Action 1 - Install ssp.basic file set, Java 1.3, RPM, and CIMOM

Perform “Installation test 1 -Verify the ssp.basic file set” on page 581 to verify that this is a problem. Install the **ssp.basic** file set using the **installp** command. Repeat “Installation test 1 -Verify the ssp.basic file set” on page 581.

Perform “Installation test 3 - Check external hardware daemon components” on page 581 to verify that this is a problem. Perform the following installation procedures. Repeat “Installation test 3 - Check external hardware daemon components” on page 581.

Java 1.3 Installation:

```
Java130.rte.bin            1.3.0.8    C     F  Java Runtime Environment
                                               Executables
  - how to install
    smitty
  - where it gets installed
    /usr/java130/jre/bin/*

Java130.rte.lib            1.3.0.8    C     F  Java Runtime Environment
  - how to install
    smitty
  - where it gets installed
    /usr/java130/jre/lib/*
                                               Libraries
Java130.xml4j              1.3.0.0    C     F  XML Parser for Java
  - how to install
    smitty
  - where it gets installed
    /usr/java130/jre/lib/ext/xerces.jar
```

RPM Installation:

```
rpm.rte                   3.0.5.30    C     F  RPM Package Manager
AIX-rpm                   5.1.0.10    C     R  Virtual Package for libraries
                                               and shells installed on system
                                               (/bin/rpm)
  - how to install
    smitty
  - where it gets installed
    /bin/rpm
```

CIMOM Installation:

```
  openCIMOM                     0.61    C     R  The SNIA CIMOM (Common
                                                 Information Model Object
                                                 Manager)  (/bin/rpm)
    - name of the CD that cimom is delivered on:
      "AIX toolbox for Linux applications"
    - file name on disk
      openCIMOM-0.61-1.aix5.1.noarch.rpm
    - how to install
      rpm -i openCIMOM-0.61-1.aix5.1.noarch.rpm
    - where it gets installed
      /opt/freeware/cimom/org/snia/wbem/client/*
      /opt/freeware/cimom/org/snia/wbem/cim/*
    - requires rpm.rte & AIX-rpm
```

Action 2 - Check permissions of ssp.basic

Perform “Installation test 2 - Check external hardware daemon” on page 581 to verify that this is a problem. Install the **ssp.basic** file set using the **installp** command. If any of the permission or file attributes do not match what is shown in “Installation test 2 - Check external hardware daemon” on page 581, issue the **chmod** or **chown** commands, as appropriate, to correct the attributes.

Repeat “Installation test 1 -Verify the ssp.basic file set” on page 581 and “Installation test 2 - Check external hardware daemon” on page 581.

Action 3 - Correct SDR Frame object

Perform “Configuration test 1 - Check SDR Frame object” on page 582 to verify that this is a problem. Perform these steps:

1. If the **SDRGetObjects Frame** command that was run in “Configuration test 1 - Check SDR Frame object” on page 582 returns an error message similar to:

   ```
   splstdata: 0022-001 The repository cannot be accessed. Return code was 80.
   ```

   refer to “Chapter 14. Diagnosing SDR problems” on page 125 to determine why the SDR cannot be accessed.
2. If you can successfully access the SDR, create a Frame object for the one that is missing by issuing the **spframe** command with the appropriate parameters. For a description of the **spframe** command, refer to *PSSP: Command and Technical Reference*.
3. If there were problems creating the SDR Frame object, investigate why the **SDR_config** command was unable to create the Frame object. Check the **SDR_config** log file, **/var/adm/SPlogs/sdr/SDR_config.log** for error messages.
4. If the Frame object was successfully created, but the server entry, output by the **SDRGetObjects Frame** command, does not have correct information in it’s `control_ipaddrs` or `domain_name` fields, issue the **spframe** command with the correct options.
5. If the Frame object was successfully created, but the server entry, output by the **SDRGetObjects Frame** command, does not have correct information in it’s `hardware_protocol` column, you must first issue the **spdelfram** command. Then, create a new definition by issuing the **spframe** command with the correct options.

For an IBM @server pSeries 690 type server, the value of the Frame object must be HMC. After correcting the problem, repeat “Configuration test 1 - Check SDR Frame object” on page 582.

Action 4 - Correct System Monitor polling interval

Perform “Operational test 1 - Check hardmon status” on page 584 to verify that this is a problem. Perform these steps:

1. If the **hardmon** daemon is not running on the control workstation, you need to start it. Issue this command:

   ```
   startsrc -s hardmon
   ```

   Then, perform “Operational test 1 - Check hardmon status” on page 584 again to determine if **hardmon** was started successfully.

2. If the **hardmon** daemon uses an incorrect polling interval, it may cause problems. The polling interval is chosen when the **hardmon** daemon is started by the AIX System Resource Controller. The value is in the `cmdargs` attribute in the **hardmon** ODM SRCsubsys object. Check the polling interval by issuing the ODM command:

   ```
   odmget -q subsysname=hardmon SRCsubsys
   ```

   The output is similar to the following, which is the default:

   ```
   SRCsubsys:
      subsysname = "hardmon"
      synonym = ""
      cmdargs = "-r 5"
      path = "/usr/lpp/ssp/bin/hardmon"
      uid = 0
      auditid = 0
      standin = "/dev/console"
      standout = "/dev/console"
      standerr = "/dev/console"
      action = 1
      multi = 0
      contact = 2
      svrkey = 0
      svrmtype = 0
      priority = 20
      signorm = 15
      sigforce = 15
      display = 1
      waittime = 15
      grpname = ""
   ```

   If the `cmdargs` attribute is not ″**-r 5**″, correct this by issuing the following command:

   ```
   chssys -s hardmon -a "-r 5"
   ```

   Then reissue the **odmget** command to verify that the new cmdargs attribute is ″**-r 5**″.After correcting the problem, repeat “Operational test 1 - Check hardmon status” on page 584.

Action 5 - Investigate external hardware daemon failure

Perform “Operational test 2 - Check external hardware daemons” on page 584 to determine if the external hardware daemon is running. Perform “Operational test 3 - Check frame responsiveness” on page 585 to determine if the external hardware daemon is responding. If either test produces error results, perform these steps:

1. Several components of PSSP, involved with the operation of the servers, write data to log files. Check these log files and take appropriate action:
   - The daemon log file. Refer to “Daemon log” on page 578.
   - The hardware monitor log. Refer to “Hardware monitor log” on page 579.
   - The SP hardware log. Refer to “SP hardware log” on page 579.

- The AIX Error log.

  The same messages that are in the SP hardware log are also found in the AIX Error log. To obtain full details of all SP hardware messages in this log, issue the command:

  ```
  errpt -aN sphwlog
  ```

  You may want to redirect the output to a file, because there could be a large amount of output.

2. If one of the external hardware daemons is not running, but it should be, check to see if a core dump was created. Refer to “Dump information” on page 579.
3. If the System Monitor (**hardmon**) daemon is running, but an external hardware daemon is not running or not responding, issue the following command to start the external hardware daemon:

   ```
   hmcmds -G boot_supervisor F:0
   ```

   where *F* is the frame number of the server. This notifies the System Monitor that the external hardware daemon has stopped. The System Monitor then starts the daemon.
4. If you have attempted to start an external hardware daemon, and it still does not start, issue the following command to stop and restart the System Monitor daemon (**hardmon**):

   ```
   hmreinit
   ```

   The System Monitor daemon (**hardmon**) will be restarted by the AIX System Resource Controller, and the daemon will then restart all of the external hardware daemons. The **SDR_config** command will also run, updating the SDR as necessary.

Chapter 40. Diagnosing PSSP T/EC Event Adapter problems

If the Tivoli Enterprise Console (T/EC) Event Adapter fails to send events to the SP system, do the following:

1. Check your event subscription and test the event generation by forcing the event.
2. Verify that the **tecad_pssp** command is being run by issuing this command on the control workstation: **lssrc -ls pman.***your_partition_name*, where *your_partition_name* is the name of the system partition of the node for which you are subscribed.

   The output from this command shows whether the event is being properly triggered at this point. If not, check your subscription again.
3. Use the **wtdumprl** command in the T/EC side to see if you are getting any event notifications from the PSSP side.
   - If you are, the problem is not in the SP system. Check the event source, event group, and event filter definitions, as well as the event group assignments in the T/EC.
   - If you are not getting any event notifications from the PSSP side, then the problem may be on the SP system side.
4. If you suspect that the **tecad_pssp** command is being run, but nothing is being generated at the T/EC side, check to see if you have the proper configuration file **/usr/lpp/ssp/tecad/tecad_pssp.cfg** installed. Also, check that it points to the T/EC server.
5. Use the **/usr/lpp/ssp/tecad/test_agent** shell script to force the invocation of the **tecad_pssp** command. Check the results.
6. Check the network connectivity. See "Chapter 12. Diagnosing system connectivity problems" on page 121.

Part 4. Appendixes

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Department LJEB/P905
2455 South Road
Poughkeepsie, NY 12601-5400
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

| The following terms are trademarks of the International Business Machines
| Corporation in the United States, other countries, or both:

AFS
AIX
DFS
ESCON
IBM
IBMLink
LoadLeveler
Micro Channel
pSeries
Redbooks
RS/6000
Scalable POWERparallel Systems

SP
System/370
System/390
Tivoli Enterprise Console
TURBOWAYS

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, MS-DOS, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, and service names may be trademarks or service marks of others.

Publicly available software

PSSP includes software that is publicly available:

expect	Programmed dialogue with interactive programs
Perl	Practical Extraction and Report Language
SUP	Software Update Protocol
Tcl	Tool Command Language
TclX	Tool Command Language Extended
Tk	Tcl-based Tool Kit for X-windows

This book discusses the use of these products only as they apply specifically to the RS/6000 SP system. The distribution for these products includes the source code and associated documentation. **/usr/lpp/ssp/public** contains the compressed **tar** files of the publicly available software. (IBM has made minor modifications to the versions of Tcl and Tk used in the SP system to improve their security characteristics. Therefore, the IBM-supplied versions do not match exactly the versions you may build from the compressed **tar** files.) All copyright notices in the documentation must be respected. You can find version and distribution information for each of these products that are part of your selected install options in the **/usr/lpp/ssp/READMES/ssp.public.README** file.

Glossary of Terms and Abbreviations

A

ACL. Access Control List. A list that defines who has permission to access certain services; that is, for whom a server may perform certain tasks. This is usually a list of principals with the type of access assigned to each.

adapter. An adapter is a mechanism for attaching parts. For example, an adapter could be a part that electrically or physically connects a device to a computer or to another device. In the SP system, network connectivity is supplied by various adapters, some optional, that can provide connection to I/O devices, networks of workstations, and mainframe networks. Ethernet, FDDI, token-ring, HiPPI, SCSI, FCS, and ATM are examples of adapters that can be used as part of an SP system.

address. A character or group of characters that identifies a register, a device, a particular part of storage, or some other data source or destination.

AFS. A distributed file system that provides authentication services as part of its file system creation.

AIX. Abbreviation for Advanced Interactive Executive, IBM's licensed version of the UNIX operating system. AIX is particularly suited to support technical computing applications, including high function graphics and floating point computations.

API. Application Programming Interface. A set of programming functions and routines that provide access between the Application layer of the OSI seven-layer model and applications that want to use the network. It is a software interface.

application. The use to which a data processing system is put; for example, a payroll application, an airline reservation application.

application data. The data that is produced using an application program.

ARP. Address Resolution Protocol.

ATM. Asynchronous Transfer Mode. (See *TURBOWAYS 100 ATM Adapter.*)

authentication. The process of validating the identity of either a user of a service or the service itself. The process of a principal proving the authenticity of its identity.

authorization. The process of obtaining permission to access resources or perform tasks. In SP security services, authorization is based on the principal identifier. The granting of access rights to a principal.

authorization file. A type of ACL (access control list) used by the IBM AIX remote commands and the IBM PSSP Sysctl and Hardmon components.

B

batch processing. (1) The processing of data or the accomplishment of jobs accumulated in advance in such a manner that each accumulation thus formed is processed or accomplished in the same run. (2) The processing of data accumulating over a period of time. (3) Loosely, the execution of computer programs serially. (4) Computer programs executed in the background.

BOS. The AIX Base Operating System.

C

call home function. The ability of a system to call the IBM support center and open a PMR to have a repair scheduled.

CDE. Common Desktop Environment. A graphical user interface for UNIX.

charge feature. An optional feature for either software or hardware for which there is a charge.

CLI. Command Line Interface.

client. (1) A function that requests services from a server and makes them available to the user. (2) A term used in an environment to identify a machine that uses the resources of the network.

CMI. Centralized Management Interface provides a series of SMIT menus and dialogues used for defining and querying the SP system configuration.

Concurrent Virtual Shared Disk. A virtual shared disk that can be concurrently accessed by more than one server.

connectionless. A communication process that takes place without first establishing a connection.

connectionless network. A network in which the sending logical node must have the address of the receiving logical node before information interchange can begin. The packet is routed through nodes in the network based on the destination address in the packet. The originating source does not receive an acknowledgment that the packet was received at the destination.

control workstation. A single point of control allowing the administrator or operator to monitor and manage the SP system using the IBM AIX Parallel System Support Programs.

credentials. A protocol message, or part thereof, containing a ticket and an authenticator supplied by a client and used by a server to verify the client's identity.

css. Communication subsystem.

D

daemon. A process, not associated with a particular user, that performs system-wide functions such as administration and control of networks, execution of time-dependent activities, line printer spooling and so forth.

DASD. Direct Access Storage Device. Storage for input/output data.

DCE. Distributed Computing Environment.

DFS. distributed file system. A subset of the IBM Distributed Computing Environment.

DNS. Domain Name Service. A hierarchical name service which maps high level machine names to IP addresses.

E

Error Notification Object. An object in the SDR that is matched with an error log entry. When an error log entry occurs that matches the Notification Object, a user-specified action is taken.

ESCON. Enterprise Systems Connection. The ESCON channel connection allows the RS/6000 to communicate directly with a host System/390; the host operating system views the system unit as a control unit.

Ethernet. (1) Ethernet is the standard hardware for TCP/IP local area networks in the UNIX marketplace. It is a 10-megabit per second baseband type LAN that allows multiple stations to access the transmission medium at will without prior coordination, avoids contention by using carrier sense and deference, and resolves contention by collision detection (CSMA/CD). (2) A passive coaxial cable whose interconnections contain devices or components, or both, that are all active. It uses CSMA/CD technology to provide a best-effort delivery system.

Ethernet network. A baseband LAN with a bus topology in which messages are broadcast on a coaxial cabling using the carrier sense multiple access/collision detection (CSMA/CD) transmission method.

event. In Event Management, the notification that an expression evaluated to true. This evaluation occurs each time an instance of a resource variable is observed.

expect. Programmed dialogue with interactive programs.

expression. In Event Management, the relational expression between a resource variable and other elements (such as constants or the previous value of an instance of the variable) that, when true, generates an event. An example of an expression is `X < 10` where X represents the resource variable `IBM.PSSP.aixos.PagSp.%totalfree` (the percentage of total free paging space). When the expression is true, that is, when the total free paging space is observed to be less than 10%, the Event Management subsystem generates an event to notify the appropriate application.

F

failover. Also called fallover, the sequence of events when a primary or server machine fails and a secondary or backup machine assumes the primary workload. This is a disruptive failure with a short recovery time.

fall back. Also called fallback, the sequence of events when a primary or server machine takes back control of its workload from a secondary or backup machine.

FDDI. Fiber Distributed Data Interface.

FFDC. First Failure Data Capture.

Fiber Distributed Data Interface (FDDI). An American National Standards Institute (ANSI) standard for 100-megabit-per-second LAN using optical fiber cables. An FDDI local area network (LAN) can be up to 100 km (62 miles) and can include up to 500 system units. There can be up to 2 km (1.24 miles) between system units and concentrators.

file. A set of related records treated as a unit, for example, in stock control, a file could consist of a set of invoices.

file name. A CMS file identifier in the form of 'filename filetype filemode' (like: TEXT DATA A).

file server. A centrally located computer that acts as a storehouse of data and applications for numerous users of a local area network.

File Transfer Protocol (FTP). The Internet protocol (and program) used to transfer files between hosts. It is an application layer protocol in TCP/IP that uses TELNET and TCP protocols to transfer bulk-data files between machines or hosts.

First Failure Data Capture (FFDC). A set of utilities used for recording persistent records of failures and significant software incidents. It provides a means of associating failures to one another, thus allowing software to link effects of a failure to their causes and thereby facilitating discovery of the root cause of a failure.

foreign host. Any host on the network other than the local host.

FTP. File transfer protocol.

G

gateway. An intelligent electronic device interconnecting dissimilar networks and providing protocol conversion for network compatibility. A gateway provides transparent access to dissimilar networks for nodes on either network. It operates at the session presentation and application layers.

H

HACMP. High Availability Cluster Multi-Processing for AIX.

HACWS. High Availability Control Workstation function, based on HACMP, provides for a backup control workstation for the SP system.

Hardware Management Console (HMC). The *IBM Hardware Management Console for pSeries* is an installation and service support processor that runs only the HMC software. For an IBM @server pSeries 690 server to run the PSSP software, an HMC is required with a network connection to the PSSP control workstation . The HMC provides the following functions for the p690 server:

- Creating and maintaining a multiple partition environment
- Detecting, reporting, and storing changes in hardware conditions
- Acting as a focal point for service representatives to determine an appropriate service strategy

Hashed Shared Disk (HSD). The data striping device for the IBM Virtual Shared Disk. The device driver lets application programs stripe data across physical disks in multiple IBM Virtual Shared Disks, thus reducing I/O bottlenecks.

help key. In the SP graphical interface, the key that gives you access to the SP graphical interface help facility.

High Availability Cluster Multi-Processing. An IBM facility to cluster nodes or components to provide high availability by eliminating single points of failure.

HiPPI. High Performance Parallel Interface. RS/6000 units can attach to a HiPPI network as defined by the ANSI specifications. The HiPPI channel supports burst rates of 100 Mbps over dual simplex cables; connections can be up to 25 km in length as defined by the standard and can be extended using third-party HiPPI switches and fiber optic extenders.

home directory. The directory associated with an individual user.

host. A computer connected to a network, and providing an access method to that network. A host provides end-user services.

HMC. Hardware Management Console.

I

instance vector. Obsolete term for resource identifier.

Intermediate Switch Board. Switches mounted in the switch expansion frame.

Internet. A specific inter-network consisting of large national backbone networks such as APARANET, MILNET, and NSFnet, and a myriad of regional and campus networks all over the world. The network uses the TCP/IP protocol suite.

Internet Protocol (IP). (1) A protocol that routes data through a network or interconnected networks. IP acts as an interface between the higher logical layers and the physical network. This protocol, however, does not provide error recovery, flow control, or guarantee the reliability of the physical network. IP is a connectionless protocol. (2) A protocol used to route data from its source to it destination in an Internet environment.

IP address. A 32-bit address assigned to devices or hosts in an IP internet that maps to a physical address. The IP address is composed of a network and host portion.

ISB. Intermediate Switch Board.

K

Kerberos. A service for authenticating users in a network environment.

kernel. The core portion of the UNIX operating system which controls the resources of the CPU and allocates them to the users. The kernel is memory-resident, is said to run in “kernel mode” and is protected from user tampering by the hardware.

Kernel Low-Level Application Programming Interface (KLAPI). KLAPI provides transport service for communication using the SP Switch.

L

LAN. (1) Acronym for Local Area Network, a data network located on the user’s premises in which serial transmission is used for direct data communication among data stations. (2) Physical network technology that transfers data a high speed over short distances. (3) A network in which a set of devices is connected to another for communication and that can be connected to a larger network.

LAPI. Low-level Communication API.

local host. The computer to which a user's terminal is directly connected.

log database. A persistent storage location for the logged information.

log event. The recording of an event.

log event type. A particular kind of log event that has a hierarchy associated with it.

logging. The writing of information to persistent storage for subsequent analysis by humans or programs.

Low-level Communication API (LAPI). A low-level (low overhead) message passing protocol that uses a one-sided active message style interface to transfer messages between processes. LAPI is an IBM proprietary interface designed to exploit the SP switch adapters.

M

mask. To use a pattern of characters to control retention or elimination of portions of another pattern of characters.

menu. A display of a list of available functions for selection by the user.

Message Passing Interface (MPI). An industry standard message passing protocol that typically uses a two-sided send-receive model to transfer messages between processes.

Motif. The graphical user interface for OSF, incorporating the X Window System. Also called OSF/Motif.

MPI. Message Passing Interface.

MTBF. Mean time between failure. This is a measure of reliability.

MTTR. Mean time to repair. This is a measure of serviceability.

N

naive application. An application with no knowledge of a server that fails over to another server. Client to server retry methods are used to reconnect.

network. An interconnected group of nodes, lines, and terminals. A network provides the ability to transmit data to and receive data from other systems and users.

Network Interface Module (NIM). A process used by the Topology Services daemon to monitor each network interface.

NFS. Network File System. NFS allows different systems (UNIX or non-UNIX), different architectures, or vendors connected to the same network, to access remote files in a LAN environment as though they were local files.

NIM. (1) Network Installation Management is provided with AIX to install AIX on the nodes. (2) Network Interface Module is a process used by the Topology Services daemon to monitor each network interface.

NIM client. An AIX system installed and managed by a NIM master. NIM supports three types of clients:
- Standalone
- Diskless
- Dataless

NIM master. An AIX system that can install one or more NIM clients. An AIX system must be defined as a NIM master before defining any NIM clients on that system. A NIM master managers the configuration database containing the information for the NIM clients.

NIM object. A representation of information about the NIM environment. NIM stores this information as objects in the NIM database. The types of objects are:
- Network
- Machine
- Resource

NIS. Network Information System.

node. In a network, the point where one or more functional units interconnect transmission lines. A computer location defined in a network. The SP system can house several different types of nodes for both serial and parallel processing. These node types can include thin nodes, wide nodes, 604 high nodes, as well as other types of nodes both internal and external to the SP frame.

Node Switch Board. Switches mounted on frames that contain nodes.

NSB. Node Switch Board.

NTP. Network Time Protocol.

O

ODM. Object Data Manager. In AIX, a hierarchical object-oriented database for configuration data.

P

parallel environment. A system environment where message passing or SP resource manager services are used by the application.

Parallel Environment. A licensed IBM program used for message passing applications on the SP or RS/6000 platforms.

parallel processing. A multiprocessor architecture which allows processes to be allocated to tightly coupled multiple processors in a cooperative processing environment, allowing concurrent execution of tasks.

parameter. (1) A variable that is given a constant value for a specified application and that may denote the application. (2) An item in a menu for which the operator specifies a value or for which the system provides a value when the menu is interpreted. (3) A name in a procedure that is used to refer to an argument that is passed to the procedure. (4) A particular piece of information that a system or application program needs to process a request.

partition. See system partition.

Perl. Practical Extraction and Report Language.

perspective. The primary window for each SP Perspectives application, so called because it provides a unique view of an SP system.

pipe. A UNIX utility allowing the output of one command to be the input of another. Represented by the | symbol. It is also referred to as filtering output.

PMR. Problem Management Report.

POE. Formerly Parallel Operating Environment, now Parallel Environment for AIX.

port. (1) An end point for communication between devices, generally referring to physical connection. (2) A 16-bit number identifying a particular TCP or UDP resource within a given TCP/IP node.

predicate. Obsolete term for expression.

Primary node or machine. (1) A device that runs a workload and has a standby device ready to assume the primary workload if that primary node fails or is taken out of service. (2) A node on the switch that initializes, provides diagnosis and recovery services, and performs other operations to the switch network. (3) In IBM Virtual Shared Disk function, when physical disks are connected to two nodes (twin-tailed), one node is designated as the primary node for each disk and the other is designated the secondary, or backup, node. The primary node is the server node for IBM Virtual Shared Disks defined on the physical disks under normal conditions. The secondary node can become the server node for the disks if the primary node is unavailable (off-line or down).

Problem Management Report. The number in the IBM support mechanism that represents a service incident with a customer.

process. (1) A unique, finite course of events defined by its purpose or by its effect, achieved under defined conditions. (2) Any operation or combination of operations on data. (3) A function being performed or waiting to be performed. (4) A program in operation. For example, a daemon is a system process that is always running on the system.

protocol. A set of semantic and syntactic rules that defines the behavior of functional units in achieving communication.

R

RAID. Redundant array of independent disks.

rearm expression. In Event Management, an expression used to generate an event that alternates with an original event expression in the following way: the event expression is used until it is true, then the rearm expression is used until it is true, then the event expression is used, and so on. The rearm expression is commonly the inverse of the event expression (for example, a resource variable is on or off). It can also be used with the event expression to define an upper and lower boundary for a condition of interest.

rearm predicate. Obsolete term for rearm expression.

remote host. *See foreign host*.

resource. In Event Management, an entity in the system that provides a set of services. Examples of resources include hardware entities such as processors, disk drives, memory, and adapters, and software entities such as database applications, processes, and file systems. Each resource in the system has one or more attributes that define the state of the resource.

resource identifier. In Event Management, a set of elements, where each element is a name/value pair of the form `name=value`, whose values uniquely identify the copy of the resource (and by extension, the copy of the resource variable) in the system.

resource monitor. A program that supplies information about resources in the system. It can be a command, a daemon, or part of an application or subsystem that manages any type of system resource.

resource variable. In Event Management, the representation of an attribute of a resource. An example of a resource variable is `IBM.AIX.PagSp.%totalfree`, which represents the percentage of total free paging space. `IBM.AIX.PagSp` specifies the resource name and `%totalfree` specifies the resource attribute.

Restricted Root Access (RRA). Restricted root access (RRA) limits the uses of the **rsh** and **rcp** commands within PSSP software. When RRA is enabled, it restricts root **rsh** and **rcp** authorizations from the nodes to the control workstation, and from one node to another. However, control workstation to node **rsh** and **rcp** access is still permitted.

RISC. Reduced Instruction Set Computing (RISC), the technology for today's high performance personal computers and workstations, was invented in 1975. Uses a small simplified set of frequently used instructions for rapid execution.

rlogin (remote LOGIN). A service offered by Berkeley UNIX systems that allows authorized users of one machine to connect to other UNIX systems across a network and interact as if their terminals were connected directly. The rlogin software passes information about the user's environment (for example, terminal type) to the remote machine.

RPC. Acronym for Remote Procedure Call, a facility that a client uses to have a server execute a procedure call. This facility is composed of a library of procedures plus an XDR.

RRA. Restricted Root Access.

RSH. A variant of RLOGIN command that invokes a command interpreter on a remote UNIX machine and passes the command line arguments to the command interpreter, skipping the LOGIN step completely. See also *rlogin*.

S

SCSI. Small Computer System Interface.

Secondary node. In IBM Virtual Shared Disk function, when physical disks are connected to two nodes (twin-tailed), one node is designated as the primary node for each disk and the other is designated as the secondary, or backup, node. The secondary node acts as the server node for the IBM Virtual Shared disks defined on the physical disks if the primary node is unavailable (off-line or down).

server. (1) A function that provides services for users. A machine may run client and server processes at the same time. (2) A machine that provides resources to the network. It provides a network service, such as disk storage and file transfer, or a program that uses such a service. (3) A device, program, or code module on a network dedicated to providing a specific service to a network. (4) On a LAN, a data station that provides facilities to other data stations. Examples are file server, print server, and mail server.

shell. The shell is the primary user interface for the UNIX operating system. It serves as command language interpreter, programming language, and allows foreground and background processing. There are three different implementations of the shell concept: Bourne, C and Korn.

Small Computer System Interface (SCSI). An input and output bus that provides a standard interface for the attachment of various direct access storage devices (DASD) and tape drives to the RS/6000.

Small Computer Systems Interface Adapter (SCSI Adapter). An adapter that supports the attachment of various direct-access storage devices (DASD) and tape drives to the RS/6000.

SMIT. The System Management Interface Toolkit is a set of menu driven utilities for AIX that provides functions such as transaction login, shell script creation, automatic updates of object database, and so forth.

SNMP. Simple Network Management Protocol. (1) An IP network management protocol that is used to monitor attached networks and routers. (2) A TCP/IP-based protocol for exchanging network management information and outlining the structure for communications among network devices.

socket. (1) An abstraction used by Berkeley UNIX that allows an application to access TCP/IP protocol functions. (2) An IP address and port number pairing. (3) In TCP/IP, the Internet address of the host computer on which the application runs, and the port number it uses. A TCP/IP application is identified by its socket.

standby node or machine. A device that waits for a failure of a primary node in order to assume the identity of the primary node. The standby machine then runs the primary's workload until the primary is back in service.

subnet. Shortened form of subnetwork.

subnet mask. A bit template that identifies to the TCP/IP protocol code the bits of the host address that are to be used for routing for specific subnetworks.

subnetwork. Any group of nodes that have a set of common characteristics, such as the same network ID.

subsystem. A software component that is not usually associated with a user command. It is usually a daemon process. A subsystem will perform work or provide services on behalf of a user request or operating system request.

SUP. Software Update Protocol.

switch capsule. A group of SP frames consisting of a switched frame and its companion non-switched frames.

Sysctl. Secure System Command Execution Tool. An authenticated client/server system for running commands remotely and in parallel.

syslog. A BSD logging system used to collect and manage other subsystem's logging data.

System Administrator. The user who is responsible for setting up, modifying, and maintaining the SP system.

system partition. A group of nonoverlapping nodes on a switch chip boundary that act as a logical SP system.

T

tar. Tape ARchive, is a standard UNIX data archive utility for storing data on tape media.

Tcl. Tool Command Language.

TclX. Tool Command Language Extended.

TCP. Acronym for Transmission Control Protocol, a stream communication protocol that includes error recovery and flow control.

TCP/IP. Acronym for Transmission Control Protocol/Internet Protocol, a suite of protocols designed to allow communication between networks regardless of the technologies implemented in each network. TCP provides a reliable host-to-host protocol between hosts in packet-switched communications networks and in interconnected systems of such networks. It assumes that the underlying protocol is the Internet Protocol.

Telnet. Terminal Emulation Protocol, a TCP/IP application protocol that allows interactive access to foreign hosts.

ticket. An encrypted protocol message used to securely pass the identity of a user from a client to a server.

Tk. Tcl-based Tool Kit for X Windows.

TMPCP. Tape Management Program Control Point.

token-ring. (1) Network technology that controls media access by passing a token (special packet or frame) between media-attached machines. (2) A network with a ring topology that passes tokens from one attaching device (node) to another. (3) The IBM Token-Ring LAN connection allows the RS/6000 system unit to participate in a LAN adhering to the IEEE 802.5 Token-Passing Ring standard or the ECMA standard 89 for Token-Ring, baseband LANs.

transaction. An exchange between the user and the system. Each activity the system performs for the user is considered a transaction.

transceiver (transmitter-receiver). A physical device that connects a host interface to a local area network, such as Ethernet. Ethernet transceivers contain electronics that apply signals to the cable and sense collisions.

transfer. To send data from one place and to receive the data at another place. Synonymous with move.

transmission. The sending of data from one place for reception elsewhere.

TURBOWAYS 100 ATM Adapter. An IBM high-performance, high-function intelligent adapter that provides dedicated 100 Mbps ATM (asynchronous transfer mode) connection for high-performance servers and workstations.

U

UDP. User Datagram Protocol.

UNIX operating system. An operating system developed by Bell Laboratories that features multiprogramming in a multiuser environment. The UNIX operating system was originally developed for use on minicomputers, but has been adapted for mainframes and microcomputers. **Note:** The AIX operating system is IBM's implementation of the UNIX operating system.

user. Anyone who requires the services of a computing system.

User Datagram Protocol (UDP). (1) In TCP/IP, a packet-level protocol built directly on the Internet Protocol layer. UDP is used for application-to-application programs between TCP/IP host systems. (2) A transport protocol in the Internet suite of protocols that provides unreliable, connectionless datagram service. (3) The Internet Protocol that enables an application programmer on one machine or process to send a datagram to an application program on another machine or process.

user ID. A nonnegative integer, contained in an object of type *uid_t*, that is used to uniquely identify a system user.

V

Virtual Shared Disk, IBM. The function that allows application programs executing at different nodes of a system partition to access a raw logical volume as if it were local at each of the nodes. In actuality, the logical volume is local at only one of the nodes (the server node).

W

workstation. (1) A configuration of input/output equipment at which an operator works. (2) A terminal or microcomputer, usually one that is connected to a mainframe or to a network, at which a user can perform applications.

X

X Window System. A graphical user interface product.

Bibliography

This bibliography helps you find product documentation related to the RS/6000 SP hardware and software products.

You can find most of the IBM product information for RS/6000 SP products on the World Wide Web. Formats for both viewing and downloading are available.

PSSP documentation is shipped with the PSSP product in a variety of formats and can be installed on your system. The man pages for public code that PSSP includes are also available online.

Finally, this bibliography contains a list of non-IBM publications that discuss parallel computing and other topics related to the RS/6000 SP.

Information formats

Documentation supporting RS/6000 SP software licensed programs is no longer available from IBM in hardcopy format. However, you can view, search, and print documentation in the following ways:

- On the World Wide Web
- Online from the product media or the SP Resource Center

Finding documentation on the World Wide Web

Most of the RS/6000 SP hardware and software books are available from the IBM Web site at:

http://www.ibm.com/servers/eserver/pseries

You can view a book or download a Portable Document Format (PDF) version of it. At the time this manual was published, the Web address of the ″RS/6000 SP Hardware and Software Books″ page was:

http://www.rs6000.ibm.com/resource/aix_resource/sp_books

However, the structure of the RS/6000 Web site can change over time.

Accessing PSSP documentation online

On the same medium as the PSSP product code, IBM ships PSSP man pages, HTML files, and PDF files. In order to use these publications, you must first install the **ssp.docs** file set.

To view the PSSP HTML publications, you need access to an HTML document browser such as Netscape. The HTML files and an index that links to them are installed in the **/usr/lpp/ssp/html** directory. Once installed, you can also view the HTML files from the RS/6000 SP Resource Center.

If you have installed the SP Resource Center on your SP system, you can access it by entering the **/usr/lpp/ssp/bin/resource_center** command. If you have the SP Resource Center on CD-ROM, see the **readme.txt** file for information about how to run it.

To view the PSSP PDF publications, you need access to the Adobe Acrobat Reader. The Acrobat Reader is shipped with the AIX Bonus Pack and is also freely available for downloading from the Adobe Web site at:

http://www.adobe.com

To successfully print a large PDF file (approximately 300 or more pages) from the Adobe Acrobat reader, you may need to select the “Download Fonts Once” button on the Print window.

Manual pages for public code

The following manual pages for public code are available in this product:

SUP /usr/lpp/ssp/man/man1/sup.1

Perl (Version 4.036)
/usr/lpp/ssp/perl/man/perl.man

/usr/lpp/ssp/perl/man/h2ph.man

/usr/lpp/ssp/perl/man/s2p.man

/usr/lpp/ssp/perl/man/a2p.man

Manual pages and other documentation for **Tcl**, **TclX**, **Tk**, and **expect** can be found in the compressed **tar** files located in the **/usr/lpp/ssp/public** directory.

RS/6000 SP planning publications

This section lists the IBM product documentation for planning for the IBM RS/6000 SP hardware and software.

IBM RS/6000 SP:

- *Planning, Volume 1, Hardware and Physical Environment*, GA22-7280
- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281

RS/6000 SP hardware publications

This section lists the IBM product documentation for the IBM RS/6000 SP hardware.

IBM RS/6000 SP:

- *Planning, Volume 1, Hardware and Physical Environment*, GA22-7280
- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281
- *Installation and Relocation*, GA22-7441
- *System Service Guide*, GA22-7442
- *SP Switch Service Guide*, GA22-7443
- *SP Switch2 Service Guide*, GA22-7444
- *Uniprocessor Node Service Guide*, GA22-7445
- *604 and 604e SMP High Node Service Guide*, GA22-7446
- *SMP Thin and Wide Node Service Guide*, GA22-7447
- *POWER3 SMP High Node Service Guide*, GA22-7448

RS/6000 SP Switch Router publications

The RS/6000 SP Switch Router is based on the Ascend GRF switched IP router product from Lucent Technologies. You can order the SP Switch Router as the IBM 9077.

The following publications are shipped with the SP Switch Router. You can also order these publications from IBM using the order numbers shown.

- *Ascend GRF GateD Manual*, GA22-7327
- *Ascend GRF 400/1600 Getting Started*, GA22-7368
- *Ascend GRF Configuration and Management*, GA22-7366
- *Ascend GRF Reference Guide*, GA22-7367
- *SP Switch Router Adapter Guide*, GA22-7310

Related hardware publications

For publications on the latest IBM @server pSeries and RS/6000 hardware products, see the Web site:

http://www.ibm.com/servers/eserver/pseries/library/hardware_docs/

That site includes links to the following:

- General service documentation
- Guides by system (pSeries and RS/6000)
- Installable options
- IBM Hardware Management Console for pSeries guides

RS/6000 SP software publications

This section lists the IBM product documentation for software products related to the IBM RS/6000 SP. These products include:

- IBM Parallel System Support Programs for AIX (PSSP)
- IBM LoadLeveler for AIX 5L (LoadLeveler)
- IBM Parallel Environment for AIX (Parallel Environment)
- IBM General Parallel File System for AIX (GPFS)
- IBM Engineering and Scientific Subroutine Library (ESSL) for AIX
- IBM Parallel ESSL for AIX
- IBM High Availability Cluster Multi-Processing for AIX (HACMP)

PSSP Publications

IBM RS/6000 SP:

- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281

PSSP:

- *Installation and Migration Guide*, GA22-7347
- *Administration Guide*, SA22-7348
- *Managing Shared Disks*, SA22-7349
- *Diagnosis Guide*, GA22-7350
- *Command and Technical Reference*, SA22-7351

- *Messages Reference*, GA22-7352
- *Implementing a Firewalled RS/6000 SP System*, GA22-7874

RS/6000 Cluster Technology (RSCT):

- *Event Management Programming Guide and Reference*, SA22-7354
- *Group Services Programming Guide and Reference*, SA22-7355
- *First Failure Data Capture Programming Guide and Reference*, SA22-7454

LoadLeveler Publications

LoadLeveler:

- *Using and Administering*, SA22-7311
- *Diagnosis and Messages Guide*, GA22-7277

GPFS Publications

GPFS:

- *Problem Determination Guide*, GA22-7434
- *Administration and Programming Reference*, SA22-7452
- *Concepts, Planning, and Installation*, GA22-7453

Parallel Environment Publications

Parallel Environment:

- *Installation Guide*, GA22-7418
- *Messages*, GA22-7419
- *MPI Programming Guide*, SA22-7422
- *MPI Subroutine Reference*, SA22-7423
- *Hitchhiker's Guide*, SA22-7424
- *Operation and Use, Volume 1*, SA22-7425
- *Operation and Use, Volume 2*, SA22-7426

Parallel ESSL and ESSL Publications

- *ESSL Products: General Information*, GC23-0529
- *Parallel ESSL: Guide and Reference*, SA22-7273
- *ESSL: Guide and Reference*, SA22-7272

HACMP Publications

HACMP:

- *Concepts and Facilities*, SC23-4276
- *Planning Guide*, SC23-4277
- *Installation Guide*, SC23-4278
- *Administration Guide*, SC23-4279
- *Troubleshooting Guide*, SC23-4280
- *Programming Locking Applications*, SC23-4281
- *Programming Client Applications*, SC23-4282
- *Master Index and Glossary*, SC23-4285
- *HANFS for AIX Installation and Administration Guide*, SC23-4283

- *Enhanced Scalability Installation and Administration Guide, Volume 1*, SC23-4284
- *Enhanced Scalability Installation and Administration Guide, Volume 2*, SC23-4306

AIX publications

You can find links to the latest AIX publications on the Web site:

http://www.ibm.com/servers/aix/library/techpubs.html

DCE publications

The DCE library consists of the following books:

- *IBM DCE for AIX: Administration Commands Reference*
- *IBM DCE for AIX: Administration Guide—Introduction*
- *IBM DCE for AIX: Administration Guide—Core Components*
- *IBM DCE for AIX: DFS Administration Guide and Reference*
- *IBM DCE for AIX: Application Development Guide—Introduction and Style Guide*
- *IBM DCE for AIX: Application Development Guide—Core Components*
- *IBM DCE for AIX: Application Development Guide—Directory Services*
- *IBM DCE for AIX: Application Development Reference*
- *IBM DCE for AIX: Problem Determination Guide*
- *IBM DCE for AIX: Release Notes*

You can view a DCE book or download a Portable Document Format (PDF) version of it from the IBM DCE Web site at:

http://www.ibm.com/software/network/dce/library

Redbooks

IBM's International Technical Support Organization (ITSO) has published a number of redbooks related to the RS/6000 SP. For a current list, see the ITSO Web site at:

http://www.ibm.com/redbooks

Non-IBM publications

Here are some non-IBM publications that you might find helpful.

- Almasi, G., Gottlieb, A., *Highly Parallel Computing*, Benjamin-Cummings Publishing Company, Inc., 1989.
- Foster, I., *Designing and Building Parallel Programs*, Addison-Wesley, 1995.
- Gropp, W., Lusk, E., Skjellum, A., *Using MPI*, The MIT Press, 1994.
- Message Passing Interface Forum, *MPI: A Message-Passing Interface Standard, Version 1.1*, University of Tennessee, Knoxville, Tennessee, June 6, 1995.
- Message Passing Interface Forum, *MPI-2: Extensions to the Message-Passing Interface, Version 2.0*, University of Tennessee, Knoxville, Tennessee, July 18, 1997.
- Ousterhout, John K., *Tcl and the Tk Toolkit*, Addison-Wesley, Reading, MA, 1994, ISBN 0-201-63337-X.
- Pfister, Gregory, F., *In Search of Clusters*, Prentice Hall, 1998.

- Barrett, D., Silverman, R., *SSH The Secure Shell The Definitive Guide*, O’Reilly, 2001.

Index

Special Characters

/.klogin 509
.bak 429
/etc/bootptab 509
/etc/exports 509, 510
/etc/group 450, 529
/etc/ha/cfg 456, 462
/etc/hosts 509
/etc/inetd.conf 509
/etc/inittab 175, 510
  sdrd 134
/etc/locks 573
/etc/ntp.conf 510
/etc/passwd 528
/etc/rc.net 136
/etc/security/passwd 501
/etc/services 323, 372, 509, 510, 527
/etc/SP 509
/etc/syslog.conf 338
$HOME 515, 521
$HOME/.$USER:Events file 18
.info 101
.last 460
.out suffix 168, 222
ppslog 492
/spdata 21, 125, 333
/spmon 318, 330
/ssp 280, 281
/tmp 21, 37, 38, 43, 49, 64, 278, 279, 378, 471
/tmp_mnt 501
/tmp/SDR_test.log 130
/unix 84
/usr 508
/var 21, 40, 43, 49, 125, 127, 134, 168, 223, 254, 290, 333, 403, 447, 452, 456, 463, 478
/var/ha 352, 420, 428
.Xdefaults 521

Numerics

200 MHz POWER3 SMP Thin Node 543
200 MHz POWER3 SMP Wide Node 543
222 MHz POWER3 SMP High Node 547
3-DigitDisplay 87, 88
332 MHz SMP Node characteristics 551
332 MHz SMP Thin Node 551
332 MHz SMP Wide Node 551
375 MHz POWER3 SMP High Node 547
375 MHz POWER3 SMP Thin Node 543
375 MHz POWER3 SMP Wide Node 543
604 and 604e High Node characteristics 541
604 High Node 541
604e High Node 541

A

about this book xix
access groups 263
account 262
account object 53, 54
acct 510
ACL 260
ACL authorization 268, 270
ACL file 264
ACL Managers 263
ACL permission 270
act.top 181, 240
action associated to a condition 17
adapter
  css 138, 186
  css0 122, 246
    IP packet size 480
  en0 122, 137, 185
  SP Switch 138
  SP Switch2 186
  switch 246
  TB3 139
  TB3MX 139
  TB3PCI 139
adapter_config_status 205
  SP Switch 176
  SP Switch2 230
adapter_config_status values
  SP Switch 152
  SP Switch2 205
Adapter configuration error information
  SP Switch 152
  SP Switch2 205
adapter Error Log Analyzer (ELA) 246
adapter level 187
adapter.log
  SP Switch2 214
AdapterMembership 455
Additional system management tests 510
Additional tests
  SYSMAN_test 510
address
  IBM 14
admin_locale 131, 525
admin_server.syslog 257
Admin SMIT stanzas 510
administrator access
  SDR 130
advanced switch diagnostics 245
AFS authentication database 272
AFS authentication server 280
AIX 7, 8, 9, 10
  I/O pacing 407
AIX 5L Version 5.1 Differences Guide 351, 412
AIX default dump device 82

AIX Error Log 18, 67, 68, 74, 115, 245, 246, 247, 249, 345, 349, 352, 403, 460, 475, 483, 484, 485, 490, 541, 544, 548, 552, 558, 562, 572, 574, 580, 591
  /var/adm/ras/errlog 67
  daemon 341
  file name 67
  Group Services 420
  Job Switch Resource Table Services 492
  location 67
  Logging daemon 333, 336, 338
  missing entries 341
  name 67
  Per Node Key Management 290
  sample formatted output 67
  SDR 126, 128
  SP Switch 138, 139
  SP Switch2 188, 189
AIX Error Log entries that invoke css.snap
  SP Switch 164
  SP Switch2 219
AIX Error Log error type 67
AIX Error Log for Event Management
  Event Management 460
AIX Error Log notification 68
AIX Error Log templates for Group Services 421
AIX Error Log templates for JSRT Services 493
AIX Error Log templates for Per Node Key Management 291
AIX Error Log templates for the Logging daemon 334
AIX Error Log templates for the System Monitor 318
AIX Error Log templates for Topology Services 354
AIX Error Logs and templates
  System Monitor 318
  Topology Services 352
AIX Error Logs and Templates for JSRT Services 492
AIX Error Notification Facility 70
  boot device 73
  Error Type PEND 72
  hdisk0 73
  kernel panics 73
  power loss 73
AIX locale 126
AIX Logical Volume Manager 484, 486
AIX messages 7
AIX remote commands 4, 300
AIX SRC subsystem 461
AIX system logs 21
AIX Versions 3.2 and 4 Performance Tuning Guide 407
AIX Virtual Memory Manager 411
AIX Virtual Memory ManagerVirtual Memory Manager 411
AIX Workload Manager 407, 411, 412
aixos 457, 519, 520
annotated switch topology file 154, 169, 224
application resource files 521
Arguments for Event Management daemon 468
ARP table 405
arptab_bsiz 405
arptab_nb 405
Assessing the current status of the SP system 22
Asynchronous (batch) notification methods 30
audience of this book xix
auth_install/log 256, 258
authentication 28, 251, 265, 266, 267, 273, 277, 278, 286, 287, 289, 303, 306, 309, 311
authentication database 263, 272
authentication method 134, 272, 275, 303, 304, 305, 306, 307, 308, 311, 312, 313, 403, 416
  dce:compat 134
  DCE only 134, 135
  not DCE only 135
  SDR 128, 134
  SP Perspectives 514
authentication problem 275
authentication server 272, 273
authentication services 313
authentication ticket 271
authentication token 271, 272
authorization 22, 251, 266, 268, 270, 303, 306, 311
authorization file 313
authorization for AIX remote commands 4, 300
authorization problem 275
auto.log 503
auto.master 503
auto.u 502, 504
AutoFS 501
Automatic method - phoenix.snap
  Event Management 458
Automating your response to problems 38
automount 510
automount function 501
automountd 501, 502, 503, 504, 505
automounter 501
Automounter 74
automounter maps 510
Automounter Related Commands 501
available paging space 22

B
backup server 273
batch notification methods 30
battery, effect of not having on error logging 68
Before you begin
  IBM Virtual Shared Disk 474
Begin monitoring 37
Berkley Software Distribution (BSD) 67
bin 322
biod 44, 51, 58, 61, 65
BIST (built-in self test) failure
  POWER3 SMP High Node 549
boot errors 113
boot/install server 101, 115, 332, 508, 509, 523, 525, 526, 528, 529, 530
  hostname 524
Boot/Install server 115
boot log 113
boot problems 113
Boot sequence for the 332 MHz SMP Node 551
Boot sequence for the POWER3 SMP High Node 547

Boot sequence for the POWER3 SMP Thin and Wide Node 543
Boot symptoms 114
bootp 103
bootp_response 509
bootpd 104
bootptab 104
bosinst_data 110
bringup microprocessor (BUMP) 541
BSD syslog 18
BSD system log 18, 67
build_dds_fail 152, 205
buildDeviceDatabase 181, 240
BUMP program 541
busresolve_fail 153, 205

C

Cable diagnostics
  SP Switch 173
  SP Switch2 227
cable_miswire 74
  SP Switch 163
  SP Switch2 218
cable problems 249
cadd_dump.out
  SP Switch2 210
Cannot contact KDC (Kerberos V5) 314
cdsadv 44, 53, 54, 60, 66, 295
cdsd 44, 53, 54, 60, 61, 66, 295
Cell Directory Services (CDS) 263
centralized error log 69
cfg 456, 462
cfgdce.log 356, 361
Changing the service log size
  Topology Services 381
Check authentication
  SP Security Services 266
Check authorization
  SP Security Services 268
Check DCE ACL authorization 268
Check DCE authentication and authorization 266
Check DCE authentication for an SP service principal 267
Check DCE group authorization 269
Check enabled security 265
Check for an sdrd core dump 130
Check for an sdrd hang 130
Check for sdrd memory leaks and CPU utilization 129
Check Kerberos V4 ACL authorization 270
Check Kerberos V4 authentication 267
Check SP-unique principals, accounts, and keytabs 262
Check that required DCE file sets are installed 261
Check that the DCE client daemons are running 262
Check that the Kerberos V4 daemons are running 264
Check the DCE CDS information required by SP ACL managers 263
Check the existence and membership of SP access groups 263
Check the Kerberos V4 administration ACL files 264
Check the Kerberos V4 authentication database 263
Check the Kerberos V4 configuration files 263
Check the Kerberos V4 srvtab files 264
Check the sdrd processes using the ps command 129
chgcss.log
  SP Switch2 217
chkslot_fail 153
CHRP 551
class 131
  partition-sensitive 135
cleanup.logs.nodes 78
cleanup.logs.ws 126
client
  file collections 526
client node
  IBM Recoverable Virtual Shared Disk 474, 485, 486
  IBM Virtual Shared Disk 474, 479, 485, 486
client principal 259
client workstation 278
Clock Diagnostics
  SP Switch 170
Clock problems and recovery actions 171
cluster 8
clustered enterprise server 559
  command 559
  messages 560
  symptom table 569
clustered enterprise server symptoms 569
CMI SMIT stanzas 509
col_dump.out 218
  SP Switch2 211
colad.trace 76
  SP Switch2 216
collection.host.list 529
command
  bootp 104
  bootpd 104
  bosboot 11
  cadd_dump 210
  cat 270
  cd 505
  cfgcol 205
  cfgtb3 152, 173
  cfgvsd 479
  chauthent 286, 303, 308
  chauthpar 286
  chauthpts 271, 286
  chauthts 286, 287
  chdev 407
  chfs 21, 40
  chgcss 499
  chmod 570, 589
  chown 570, 589
  chssys 330, 335, 336, 572, 590
  chvg 118
  clean_up.dce 314
  clean_up.dce-core 314
  cleanup.logs.nodes 78
  clhandle 358, 359, 360, 383, 385, 389
  cllsclstr 385, 389
  cllsif 358, 359, 360, 383, 385, 389

command *(continued)*
- clmixver 385
- clustered enterprise server 559
- col_dump 211
- compress 378
- config_spsec 273, 276, 283
- crash 84, 414
  - trace subcommand 85
- create_keyfiles 276, 281, 282, 284
- createhsd 520
- createvsd 479, 520
- cron 523
- crontab 504, 523
- cshutdown 74
- css_dump 156
- CSS_test 76, 158, 166, 174, 175, 177, 212, 220, 228, 229, 234
- cstartup 74
- date 176, 230
- dbx 130, 376, 426
- dce_login 54, 135, 266, 268, 271, 287, 307, 309
- dcecp 53, 253, 258, 259, 260, 262, 263, 269, 281, 283, 285, 287, 292, 293, 294, 307, 385
- dd 486
- delnimclient 104
- df 20, 21, 22, 27
- diag 115, 119, 162, 172, 216, 226, 246
- dmpdt_chrp 543, 547, 551
- dsh 20, 21, 22, 26, 40, 68, 111, 176, 230, 233, 266, 275, 474, 481, 482, 483, 507, 514, 526
- dsrvtgt 267, 276, 285, 294, 307, 356
- Eannotator 161, 169, 180, 215, 224, 239
- echo 128, 131
- Eclock 75, 171, 179, 180
- Ecommands 75, 138, 178, 180, 186, 234, 239
- Efence 154
- Emaster 231, 232
- Emonitor 75
- enadmin 555, 558
- endefadapter 553, 558
- endefnode 553, 558
- enrmadapter 553
- enrmnode 553
- Eprimary 138, 175, 178, 181, 188, 229, 231, 232, 235, 241
- errdead 84
- errdemon 352
- errpt 18, 68, 113, 128, 138, 156, 172, 188, 210, 226, 290, 318, 333, 339, 346, 391, 393, 394, 400, 413, 416, 430, 436, 438, 445, 449, 450, 483, 484, 485, 492, 493, 562, 572, 580, 591
- Estart 75, 146, 155, 174, 175, 177, 178, 180, 182, 183, 186, 229, 231, 232, 234, 235, 236, 237, 238, 239, 242, 243, 245, 249, 488, 489
- Etopology 169, 224
- Eunfence 154, 174, 178, 182, 183, 228, 231, 234, 241, 242, 488, 489, 554
- exit 284
- export 265, 285, 492, 497
- exportfs 95, 109, 110
- extendvg 118

command *(continued)*
- fcreport 138, 160
- fcstkrpt 160
- filec_config 78
- filec_host 523
- fs_dump 157
- ftp 299
- fuser 503
- get_keyfiles 276
- getauth 287
- grcard 555
- grconslog 556
- grep 504
- grreset 556
- ha_vsd 480, 490
- ha.vsd 478, 480, 481, 482, 487
- haemcfg 461, 466, 472
- haemctrl 461, 463
- haemqvar 32, 33, 34, 35, 38
- haemtrcon 468, 469
- haemunlkrm 463
- hagsctrl 449
- hagsgr 468
- hagsns 428, 429, 431, 432, 433, 439, 440, 441
- hagsvote 431, 434, 443, 450, 451
- hats_keys 361
- hatsctrl 368, 403, 405, 416, 417
- hatsdmsinfo 357, 415
- hatstune 381, 408, 409
- hc.vsd 487
- hmadm 318, 319, 320, 321, 561, 562, 563, 578, 579, 580
- hmcmds 82, 87, 319, 331, 344, 346, 541, 542, 573, 591
- hmdceobj 269
- hmmon 20, 21, 22, 24, 25, 28, 29, 30, 45, 47, 48, 87, 319, 328, 330, 340, 342, 347, 543, 547, 551, 567, 568, 569, 585, 586
- hmreinit 323, 331, 573, 591
- host 135
- hostname 135, 307
- IBM @server@server pSeries 690 577
- ifcl_dump 210
- ifconfig 48, 122, 358, 363, 365, 375, 383, 386, 388, 395, 485
- install_cw 272, 286, 287
- installp 329, 570, 588
- iptrace 386, 401
- k4destroy 268
- k4estroy 271
- k4init 267, 270, 271, 274, 275, 278, 279, 307, 309
- k4list 259, 260, 267, 268, 276, 307, 514
- kadmin 264
- kdestroy 267, 269, 271, 284, 285, 307
- kill 93, 127, 129, 130, 331, 376, 417, 426
- kinit 287
- klist 135, 259, 260, 266, 267, 268, 276, 285, 307, 514
- ksh 37
- kshell 312
- ksrvtgt 271, 276

command *(continued)*
ksrvutil 259, 264, 277, 280, 308
list 524
locale 126
lock 524
login 305
lppchk 386, 434
ls 259, 284, 502, 503, 563, 581
lsattr 111
lsauthent 266, 303, 306, 308, 313
lsauthpar 259, 303, 306, 308, 313
lsauthpts 134, 135, 259, 354, 355, 370, 385
lsauthts 266, 287
lscfg 77, 541
lsdce 262, 273
lskp 259, 263, 264, 307
lslpp 261, 306, 321, 385, 386, 432, 434, 464, 563, 581
lslv 119
lsnim 91, 92, 101
lsps 20, 22, 26
lspv 119
lssrc 20, 22, 122, 129, 232, 264, 293, 376, 398, 405, 416, 423, 431, 462, 483, 501, 518, 557, 566, 584
lsvg 119
lsvsd 481, 483, 485, 486
mastername 524, 528
monitoring 28, 30
mount 501, 502, 505
mult_senders_test 248
netstat 123, 159, 213, 259, 358, 363, 364, 375, 383, 386, 388, 389, 408
nim 92, 93, 95
nlssrc 431
no 123
nonlocsrcroute option 123
no (network options) 405, 486
node_number 417
nodecond 93, 102, 103, 104, 105, 110, 115
odmadd 71, 72, 73
odmdelete 71, 72, 73
odmget 71, 72, 73, 213, 329, 330, 336, 385, 389, 571, 572, 590
cmdargs 335, 336
odmshow 71
oslevel 306
partition sensitive 330
perspectives 16, 513, 516, 517, 520, 553
phoenix.snap 376, 377, 378
ping 88, 121, 122, 133, 174, 177, 229, 234, 309, 312, 363, 375, 386, 396, 397, 399, 400, 401, 485, 499, 505
pmandef 34, 36, 37, 38, 40, 62
pmanquery 62
prefix 524
ps 93, 129, 133, 176, 230, 327, 332, 341, 412, 416, 490, 501, 502, 504, 510, 566, 584
rc.net 395
rc.switch 156, 161, 168, 171, 174, 175, 178, 182, 211, 214, 223, 225, 228, 234, 242

command *(continued)*
rcmdtgt 271, 276
rcp 4, 111, 278, 279, 299, 300, 302, 303, 308, 309, 310, 312, 314
read_regs 157, 212
read_tbic 157, 170, 171
reducevg 117, 118
refresh 104, 105, 106, 107, 108, 109, 383
refuse 524
removevsd 479
reset 129
rlogin 88, 121, 299, 309, 312
rm_spsec 276
rmxcreds 315
rsh 4, 111, 278, 279, 299, 300, 302, 303, 306, 308, 309, 310, 311, 312, 313, 514
rvsdrestrict 475, 476
s1term 48, 93, 104, 110, 115, 122, 319, 342, 347, 541, 542, 569, 587
sdr 129
SDR_confg 575
SDR_config 234, 325, 326, 570, 571, 573, 578, 589, 591
SDR_init 287
SDR prefix 125
SDR_test 74, 77, 130, 131, 136, 170, 224
non-root user 130
SDRChangeAttrValues 323, 324, 428
SDRCreateObjects 324
SDRGetObjects 135, 152, 170, 176, 205, 225, 230, 231, 323, 324, 325, 326, 385, 387, 388, 410, 435, 499, 553, 574, 583, 589
SDRScan 131
setup_authent 264, 277
setup_logd 338, 343, 349
setup_server 92, 94, 95, 102, 103, 104, 105, 106, 107, 108, 109, 110, 114, 273, 277
smit 67, 170, 174, 224, 300, 326
SDR 130
SMIT 495, 507, 511, 523, 527, 528
smitty 77, 573, 574
snap 11, 12, 21
SP-attached server 559
spapply_config 325
spbgroot 356
spbootins 92, 93, 104, 110, 115, 279
spchuser 505
spchvgobj 110, 117, 118, 119
spdelfram 331, 560, 565, 571, 574, 578, 589
spevent 513, 516, 517, 520
spframe 331, 560, 564, 565, 570, 571, 574, 578, 582, 589
spget_syspar 125, 132, 133, 134, 373
spgrpname 271
sphardware 19, 513, 516, 517, 520, 521
sphmcid 583
sphrdwrad 104
spled 513
splm 12
splogd
debug 335

command *(continued)*
- splst_versions 435
- splstadapters 553
- splstdata 8, 11, 91, 92, 101, 102, 103, 109, 114, 117, 118, 119, 125, 128, 130, 159, 213, 258, 259, 265, 286, 293, 303, 321, 385, 431, 464, 494, 502, 503, 514, 525, 562, 564, 567, 570, 571, 573, 580, 582, 589
- splstnodes 125, 553
- spmirrorvg 118
- spmon 11, 20, 21, 22, 23, 25, 26, 28, 29, 30, 48, 81, 87, 121, 270, 402, 542, 564, 565, 583
- spmon_ctest 77, 327
- spmon_itest 77, 322, 323
- spnkeyman_start 293
- spnkeymand 292
- spsetauth 286
- spsitenv 300, 527, 528
- spsvrmgr 326
- spsyspar 513, 516, 517, 520
- spunmirrorvg 119
- spvsd 513, 516, 517, 520
- st_clean_table 491, 492, 499
- st_set_switch_number 498
- st_status 491, 494, 498, 499
  - hang 499
- st_verify 491, 494, 495
- startauto 503, 506
- startsrc 128, 129, 133, 173, 232, 233, 285, 293, 304, 315, 329, 340, 342, 344, 345, 347, 472, 518, 520, 571, 590
  - splogd 336
- statvsd 484
- stopsrc 127, 129, 173, 233, 296, 304, 314, 342, 344, 345, 347, 472, 505, 573
  - splogd 336
- supfilesrv 523
- supper 76, 504, 510, 523, 525, 526, 527, 529, 530
- supperlock 524
- swtbl_adapter_connectivity 491
- swtbl_adapter_resources 491
- swtbl_clean_table 491
- swtbl_load_job 491
- swtbl_load_table 491
- swtbl_query_adapter 491
- swtbl_status 491
- swtbl_status_node 491
- swtbl_unload_job 491
- sysctl 267, 270, 272, 285, 287
- sysdumpdev 83, 84, 543, 547, 551
- sysdumpstart 82
- syslogd 67
- SYSMAN_test 78, 507, 508, 511
  - normal 507
  - quiet 507
  - verbose 507
- syspar_ctrl 391, 436, 445, 464, 472
- System Monitor 319, 328
- tar 378
- tbXdump 156
- tecad_pssp 593

command *(continued)*
- telnet 88, 115, 121, 122, 299, 309, 312, 520, 555
- tftp 104, 105, 106, 107, 108
- topsvcsctrl 417
- touch 304
- trace 407
- traceroute 123
- tracesoff 379, 428
- traceson 379, 428
- trcstop 407
- ucfgtb3 173
- ucfgvsd 480
- unmount 502, 505
- unset 285
- updatevsdnode 480
- varyonvg 477
- vdidlX 159
- vhostname 323, 324
- vi 483
- view 502
- vmstat 332, 386, 406, 409
- vmtune
  - minfree 406
- vsdatalst 480, 485, 490
- vsdchgserver 490
- vsddump 478
- vsdtrace 478
- vsdvts 479
- whoami 135
- worm.trace 178, 235
- wrap_test 249
- wtdumprl 593
- xhost 16, 518
- xntpdc 510

command line 32

command-line monitoring tools 30

Command line tools - Problem Management 32, 40

command line tools for system administration 20

Concurrent Virtual Shared Disk (CVSD) 485

condition 16, 17, 31, 32, 33, 36, 37, 38, 39, 40, 55, 56, 60, 61, 62, 471
- monitoring 19

Conditions and default event definitions 55

conditions that you can monitor that you must define to the SP Event Perspective 55

conditions that you can monitor using the default event definition 55

Conditions to monitor only in specific locations 61

Conditions to monitor using Perspectives or Problem Management 43

config_info 105, 509

config.log 102

config_node 99

configfb.log 102

configuration errors
- IBM Recoverable Virtual Shared Disk 474
- IBM Virtual Shared Disk 474

configuration file
- Tivoli Enterprise Console Event Adapter 593

Configuration verification test
- Group Services 435

Configuration verification tests
clustered enterprise server 564
IBM @server@server pSeries 690 582
IBM Recoverable Virtual Shared Disk 479
IBM Virtual Shared Disk 479
SP-attached Server 564
System Monitor 323
Topology Services 387
connection type 462
connectivity matrix
SP Switch2 186
console log
IBM Recoverable Virtual Shared Disk 478, 483, 484
IBM Virtual Shared Disk 478, 483, 484
console.log 113, 114, 115
Console log
boot 113
contacting IBM 13
contacting the IBM Support Center 3, 10, 13
Contents of the phoenix.snap file 378
control workstation 8, 10, 15, 39, 43, 44, 74, 82, 88, 92, 107, 163, 174, 217, 228, 245, 273, 318, 388, 507, 508, 509, 518, 523, 526, 528, 529, 544, 552, 557, 560, 561, 562, 563, 564, 569, 573, 574, 577, 578, 579, 580, 581, 587, 588
DCE-only mode 322
installation 6
rs232 tty cables 331
controllerIDmismatch 43
controllerIDMismatch 24
controllerResponds 24
FALSE 87
red 87
core dump
Group Services 421, 422, 423, 424, 425, 426
Logging daemon 336, 348
SDR 128, 130, 133
SP Perspectives 514, 516, 520
System Monitor 321
Topology Services 375
core file 49
IBM @server@server pSeries 690 579
IBM @server@server pSeries 690 problems 591
SP-attached server 562, 573
SP Security Services 258
System Monitor 320
CPU utilization 332, 520
SDR 129
crash
time of crash 85
trace subcommand 85
Create a log of your SP structure and setup 7
Create an event definition file 36
Creating an event definition using the Create Event Definition notebook 61
creating the event definitions 60
credentials 267
cron 530
crontab 510
crontab entry to delete hardware error entries 69
crontab entry to delete software error entries 69
css 138, 186
SP Switch 138
SP Switch2 188
css_dump.out
SP Switch 156
css_ready 152, 153, 205, 206
css.snap 74, 164, 165, 168, 183, 218, 219, 223, 242
SP Switch 156
SP Switch2 210
css.snap.log 164, 218
SP Switch 163
SP Switch2 216
css.snap package
SP Switch 164
css.snap Package
SP Switch2 218
CSS_test.log 175
SP Switch 158
SP Switch2 212
css0 246
SP Switch 138
SP Switch2 188
css0 adapter 173
css1
SP Switch2 188
cssadm.debug
SP Switch 157
cssadm.stderr
SP Switch 157
cssadm.stdout
SP Switch 157
cssadm2 186, 187
cssadm2.debug
SP Switch2 211
cssadm2.stderr
SP Switch2 211
cssadm2.stdout
SP Switch2 212
cssMembership 437, 442, 443, 446, 457
CSswitchInit 240
customer information 8, 10
customization files 502
CVSD 485
CWS 74

D

daemon
automountd 501, 502, 503, 504, 505
biod 44, 51, 58, 61, 65
bootpd 104
cdsadv 44, 53, 54, 60, 66, 295
cdsd 44, 53, 54, 60, 61, 66, 295
cron 530
cssadm2 186, 187
dced 44, 53, 59, 66, 295
emaster 232
emasterd 186
Emonitor 75
errdemon 341
external hardware 563, 579, 581

- daemon *(continued)*
  - fault service 74, 75, 138, 186
  - ftpd 43
  - hacssrmd 192
  - hadsd 44
  - haemd 44, 53, 456, 457, 460, 461, 462, 463, 464, 466, 467, 468, 469, 472, 518, 520
  - hagsd 52, 362, 421, 423, 426, 427, 428, 429, 430, 433, 436, 437, 439, 440, 446, 447, 448, 450, 452, 520
  - hagsglsm 422, 423, 424, 443, 444, 445, 447
  - hardmon 318, 319, 322, 327, 328, 329, 330, 332, 340, 520, 585
    - -r flag 327
  - harmad 457
  - harmld 457
  - harmpd 457
  - hatsd 44, 52, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 379, 381, 392, 393, 402, 406, 408, 411, 412, 413
  - hc 490
  - hmc 578
  - hmcd 520, 584
  - hmrmd 456
  - hrd 458
  - httpd 44
  - IBM Recoverable Virtual Shared Disk 475, 476, 477
  - id 529
  - inetd 43, 50, 65, 104, 105, 106, 107, 108, 109, 303, 312, 313
  - kadmind 256, 257, 274
  - kerberos 44, 52, 59, 61, 66, 256, 257, 274, 305
  - kpropd 256, 257, 274
  - krshd 303, 304, 305, 307, 308, 310, 311, 312, 313
  - la_event_d 217
  - logging 77, 332
  - mount 505
  - NFS 332
  - pmanrmd 457
  - portmap 44, 51, 58, 65
  - rlogind 312
  - rvsd 483, 490
  - s70 561, 562, 566
  - s70d 77, 318, 520
  - sdrd 44, 53, 126, 127, 128, 129, 130, 133, 134, 135, 136, 520
  - secd 44, 54, 60, 61, 66, 295, 297
  - snmpd 557
  - splogd 77, 340
  - spmgrd 557
  - spnkeyman 290, 293, 295, 296, 297
  - spnkeymand 276
  - srcmstr 43, 51, 58, 65
  - sshd 112
  - supfilesrv 510, 525, 527, 528, 529
  - switch admin 74
  - switchtbld 494
  - syncd 408
  - sysctld 116, 285, 287
  - syslogd 67, 304, 305, 307
- daemon *(continued)*
  - System Monitor 322
  - telnetd 312
  - tftp 104, 105, 106, 107, 108, 109, 509
  - xntpd 44, 51, 59, 61, 65, 510
- daemon log
  - IBM @server@server pSeries 690 578
  - SP-attached server 560
- daemon.log
  - SP Switch2 214
- daemon.stderr
  - SP Switch 162
- daemon.stdout
  - SP Switch 162
  - SP Switch2 214
- DARE 368
- data packet error
  - missing 319
  - System Monitor 319
- Data Striping Device 473
- DCE 4, 7, 259, 272, 289, 290, 296, 309, 356, 370, 402, 403
- DCE ACL file 275
- DCE authentication 267, 272
- DCE authentication ticket 266
- DCE authentication token 271
- DCE cell 273
- DCE cell administrator 251
- DCE client daemon 295
- DCE daemons 44, 262, 313
- DCE group authorization 269
- DCE GSSAPI 255
- DCE hostname 280
- DCE key file 261, 262, 276, 277, 280, 281, 284, 292
  - corrupted 293
  - inconsistent 293
- DCE login failure 273
- DCE messages 272, 306
- DCE principal 262
- DCE problem determination 7
- DCE publications 254, 255
- DCE registry 292, 293, 295, 311, 312, 313
- DCE restriction 4
- DCE security groups for JSRT Services 499
- DCE server 292
- DCE server daemon 295
- DCE server host 273
- DCE server key 289
- DCE server terminated 273
- dced 44, 53, 59, 66, 295
- dd_config_fail 153, 206
- dd_load_fail 153, 206
- Dead Man Switch Timer 357, 411, 414, 415
- debug
  - IBM @server@server pSeries 690 585
  - SP-attached server 567
- Debug information
  - Per Node Key Management 292
- debug_script
  - SP Perspectives 514
- debug spot 92

debug SPOT 115
debugging information
  error logging 67
Decide how to notify the system administrator. 35
Decrypt integrity check failed (Kerberos V5) 314
default condition 17
default event definition 55
default principal 263
default system partition 132
default system profile
  SP Perspectives 515
default user profile 515
define_fail 153
define_system_fail 153
dependent node 553
Descriptions of each condition 44
Details about each condition to monitor 44
Detecting SP problems and keeping informed 15
device icon 18
device notebook 18
device status
  SP Switch 154
  SP Switch2 207
DeviceDB.dump
  SP Switch2 214
devices, diagnosing 115
diag 246
diag_fail 153, 173, 178, 206, 234
  SP Switch 163
  SP Switch2 218
diag_fail condition for an SP Switch adapter 173
diagnosing
  200 MHz POWER3 SMP Thin Node problems 543
  200 MHz POWER3 SMP Wide Node problems 543
  222 MHz POWER3 SMP High Node problems 547
  332 MHz SMP Thin Node problems 551
  332 MHz SMP Wide Node problems 551
  375 MHz POWER3 SMP Thin Node problems 543
  375 MHz POWER3 SMP Wide Node problems 543
  375 MHzPOWER3 SMP High Node problems 547
  604 High Node problems 541
  604e High Node problems 541
  boot problems 113
  clustered enterprise server problems 559
  dependent node configuration problems 553
  Event Management 455
  extension node configuration problems 553
  Group Services problems 419
  hardmon 317
  IBM @server@server pSeries 690 problems 577
  IBM Recoverable Virtual Shared Disk problems 473
  IBM Virtual Shared Disk problems 473
  IP routing problems 123
  Job Switch Resource Table Services 491
  Logging daemon 333
  netinstall problems 101
  node install 101
  Per Node Key Management 289
  POWER3 SMP High Node problems 547
  POWER3 SMP Thin Node problems 543
  POWER3 SMP Wide Node problems 543
diagnosing *(continued)*
  rcp 299
  remote command problems 299
  Root Volume Group 117
  rsh 299
  SDR problems 125
  security services 251
  SP-attached server problems 559
  SP hardware problems 87
  SP Perspectives 513
  SP software problems 87
  SP Switch problems 137
  SP Switch2 problems 185
  splogd 333
  system connectivity problems 121
  System Monitor 317
  T/EC Event Adapter 593
  Tivoli Enterprise Console Event Adapter 593
  Topology Services 351
  user access problems 501
Diagnosing 332 MHz SMP Thin and Wide Node problems 551
Diagnosing 604 and 604e High Node problems 541
Diagnosing boot problems 113
Diagnosing dependent node configuration problems 553
Diagnosing Devices Including the Hard Disk 115
Diagnosing Event Management problems 455
Diagnosing file collections problems on the SP System 523
Diagnosing file collections problems on the SP SystemSP file collections 523
Diagnosing Group Services problems 419
Diagnosing hardware and software problems 87
Diagnosing IBM Virtual Shared Disk problems 473
Diagnosing IP routing problems 123
Diagnosing Job Switch Resource Table Services problems 491
Diagnosing NIM problems 91
Diagnosing node installation problems 101
Diagnosing Perspectives problems on the SP System 513
Diagnosing POWER3 SMP High Node problems 547
Diagnosing POWER3 SMP Thin and Wide Node problems 543
diagnosing PSSP problems 3
Diagnosing PSSP T/EC Event Adapter problems 593
Diagnosing remote command problems on the SP System 299
Diagnosing Root Volume Group problems 117
Diagnosing SP-attached Server and Clustered Enterprise Server problems 559
diagnosing SP problems 3
Diagnosing SP problems overview 3
Diagnosing SP Security Services problems 251
Diagnosing SP Switch problems 137
Diagnosing SP Switch2 problems 185
Diagnosing system connectivity problems 121
Diagnosing System Monitor Problems 317
Diagnosing Topology Services problems 351
Diagnosing User Access problems 501

diagnosis 9
Diagnostic instructions
  Event Management 464
diagnostic mode 115
diagnostic mode boot 115
Diagnostic procedure
  clustered enterprise server 563
  IBM @server@server pSeries 690 580
  SP-attached server 563
diagnostic procedures 3, 26
  Job Switch Resource Table Services 494
Diagnostic procedures
  Group Services 434
  IBM Recoverable Virtual Shared Disk 479
  IBM Virtual Shared Disk 479
  Logging daemon 336
  SP Security Services 260
  SP Switch 168
  System Monitor 321
  Topology Services 386
Diagnostic Procedures
  Per Node Key Management 293
  SDR 128
  SP Switch2 223
Diagnostic procedures - Installation verification
  Job Switch Resource Table Services 494
diagnostics
  switch 245
Diagnostics
  file collections 526
  SP Perspectives 515
Diagnostics specific to Kerberos V4
  remote commands 308
Diagnostics specific to Kerberos V5
  remote commands 309
directories for Error Notification objects 70
directory
  /etc/bootptab 509
  /etc/exports 509
  /etc/hosts 509
  /etc/inetd.conf 509
  /etc/locks 573
  /etc/services 509
  /etc/SP 509
  /spdata/sys1/k4srvtabs 109
  /spmon 330
  /ssp 280, 281
  /tmp_mnt 501
  /var 168, 223, 447, 452, 478
  /var/ha 352, 420, 428
  bin 112, 322
  etc 108
  lppsource 95
  perspectives/profiles 521
  splogd 348
  tftpboot 104, 105, 106, 107, 108
  ucode 322
Disk EIO errors
  IBM Virtual Shared Disk 490
Disk space handling policy
  SP Switch 168
Disk space handling policy *(continued)*
  SP Switch2 223
DISPLAY 16, 518, 521
dist_topology.log
  SP Switch 161
Distributed Computing Environment 254
Distributed Computing Environment (DCE) Version 3.1 for AIX 254
domain
  Group Services 422
domain merge
  Group Services 421
domain name
  IBM @server@server pSeries 690 585
DOUBLE_PANIC error label 70
Downstream Neighbor 361, 379, 384
dsh command to view error logs 68
DSHPATH 474
dtbx.failed.trace
  SP Switch 163
dtbx.trace
  SP Switch 162
dump 11, 12
  system 81
dump device
  default 82
  primary 81
  secondary 82
Dump information
  Event Management 463
  Group Services 426
  IBM @server@server pSeries 690 579
  IBM Recoverable Virtual Shared Disk 477
  IBM Virtual Shared Disk 477
  Logging daemon 336
  SP-attached server 562
  SP Security Services 258
  SP Switch 156
  SP Switch2 210
  System Monitor 320
  Topology Services 375
Dump information from Event Management Resource Monitor daemons 463
Dump Methods 81
Dump to the Primary Dump Device 82
Dump to the Secondary Dump Device 82
Dynamic Automatic Reconfiguration Event (DARE) 368

E

Ecommands error recovery
  SP Switch 180
  SP Switch2 239
Ecommands.log 180, 239
  SP Switch 163
  SP Switch2 217
Effect of not having a battery on error logging 68
ELA 246
ELA (Error Log Analysis) 544, 548, 552
ELA recommendations 246
em.default 460

em.msgtrace 464
EM_Resource_Monitor 462
EM suffix 70
em.trace 463, 469
EMAPI 455, 458, 461, 471, 472
emasterd 186
  MSS node error 204
  terminated 204
emasterd.log
  SP Switch2 217
emasterd.stdout
  SP Switch2 217
EMCDB 456, 461, 466
  version number 466
EMCDB problems 461
EMCDB Version is incorrect 461
emergency error 70
en_name 71
en0 adapter 356, 388, 458
Enable LOGIN 573, 574
enhanced security option 4
enMembership 457
Ensuring that the conditions you intend to monitor are known to Problem Management 62
Environment setting
  Event management 461
environment variable
  $DSH_REMOTE_CMD 111
  $HOME 515, 521
  $RCMD_PGM 111
  $REMOTE_COPY_CMD 111
  DISPLAY 16, 518, 521
  DSH_REMOTE_CM 301
  DSHPATH 474
  ICLUI_TRACE 515
  K5MUTE 304, 305, 306, 307
  KRB5CCNAME 285
  KRBTKFILE 276
  LANG 515
  LIBPATH 499
  PATH 474, 518, 520
  PMAN_IVECTOR 39, 40
  RCMD_PGM 301
  REMOTE_COPY_CMD 301
  SP_NAME 128, 133, 134, 265
  SWTBLAPIERRORMSGS 492, 497
  SWTBLSECMSGLEVEL 497
  WCOLL 474, 480
  XAPPLRESDIR 521
EPOW_SUS 541
EPOW_SUS AIX error label 70
errdemon 341
error avoidance 6
Error conditions and performance considerations
  332 MHz SMP Thin Node 552
  332 MHz SMP Wide Node 552
  604 High Node 541
  604e High Node 541
  POWER3 SMP High Node 548
  POWER3 SMP Thin Node 544
  POWER3 SMP Wide Node 544
error daemon 341
Error information
  boot 113
  Event Management 460
  file collections 525
  Group Services 420
  IBM @server@server pSeries 690 578
  IBM Recoverable Virtual Shared Disk 474
  IBM Virtual Shared Disk 474
  Job Switch Resource Table Services 492
  Logging daemon 333
  node install 101
  Per Node Key Management 290
  remote commands 304
  SDR 126
  SP-attached server 560
  SP Perspectives 514
  SP Security Services 256
  SP Switch 138
  SP Switch2 186
  System Monitor 318
  Topology Services 352
error label suffix 67
error log 6, 21, 67
error log entries
  deletion time 69
error log events
  classifying 67
Error log files
  Event Management 460
Error Log Management 67
error log management facility 6
error log template
  HA_PMAN_EVENT_OFF 18
  HA_PMAN_EVENT_ON 18
error logging 67
  effect of not having a battery 68
error logs
  SP 74
error message
  SP Perspectives 514
Error messages indicating SP Security Services configuration problem 313
Error messages pointing to or from an authentication subsystem 312
Error Notification method scripts 70
Error Notification object
  creation 71
Error Notification objects
  directories 70
error recovery 6
Error symptoms, responses, and recoveries
  boot 114
  clustered enterprise server 569
  Event Management 472
  file collections 530
  Group Services 446
  IBM @server@server pSeries 690 symptoms 588
  IBM Recoverable Virtual Shared Disk 487
  IBM Virtual Shared Disk 487
  Job Switch Resource Table Services 494

- Error symptoms, responses, and recoveries *(continued)*
  - Logging daemon 348
  - NIM 94
  - node install 102
  - Per Node Key Management 296
  - remote commands 309
  - SDR 131
  - secure remote commands 315
  - SP-attached server 569
  - SP Perspectives 515
  - SP Security Services 271
  - SP Switch 177, 233
  - System Monitor 328
  - Topology Services 402
- Errori Information
  - NIM 91
- Errors logged by the IBM Virtual Shared Disk device driver and IBM Recoverable Virtual Shared Disk subsystem 474
- errpt command to view error logs 68
- errpt.out
  - SP Switch 156
  - SP Switch2 210
- esource monitors
  - Event Management 456
- essential documentation 4
- Essential documentation - Other manuals to accompany this book 4
- Estart 245, 249
- Estart error recovery
  - SP Switch 180
  - SP Switch2 239
- etc 108
- etc/plane.info 187
- Ethernet adapter 114, 356, 388, 510
  - duplex setting 111
  - full_duplex 111
- Eunfence error recovery
  - SP Switch 182
  - SP Switch2 241
- Eunpartition 75
- event 16, 31, 40
- event command 36
- event definition 16, 17, 31, 35, 37, 40, 55, 60, 61, 62
- Event Definition Notebook 517
- event expression 16, 17, 35, 36, 38, 39
- event handle 35, 36, 37
- Event Management 16, 31, 32, 52, 53, 78, 455, 518
  - communication 456
  - core dump 463
  - esource monitors 456
  - events fail 471
  - External resource monitors 456
  - instance 470
  - Internal resource monitors 457
  - resource monitor 462
  - socket information 456
- Event Management Application Program Interface 461
- Event Management Configuration Database 456
- Event Management Configuration Database (EMCDB) operations 461
- Event Management Daemon 461
- Event Management daemon errors 460
- Event Management resource monitor
  - locked 462
  - reconnection 462
- Event Management Resource Monitor 75
- Event Manager 516, 517, 518, 519, 520
  - access 516
- event notification 593
- Event Perspective 16
- event registration 16, 31, 55, 471
- event subscription 34, 36, 37, 593
- expected.top 169, 224
  - SP Switch 161
  - SP Switch2 215
- extension node 553
- external clock
  - SP Switch 170, 171
- external hardware daemon 563, 566, 567, 569, 570, 577, 581, 584, 585, 588
- external Hardware Daemon 572, 573, 590, 591
- external hardware daemon communication log 561
  - IBM @server@server pSeries 690 578

F

- failure 10, 13
  - hardware 13, 14
  - kernel memory buffer 50
  - non-IBM hardware 14
  - non-IBM software 13
  - PC 14
  - PSSP software 13
  - software 13
- FAST IPL 541, 542
- fault service daemon 74, 75, 186
  - SP Switch 138, 154, 161, 162, 163, 168, 169, 172, 206
    - acknowledge service command failure 148
    - board hardware error 151
    - bus signal received 151
    - cable miswired 147
    - clock failure 150, 151
    - command broadcast failed 149
    - DBupdate broadcast failed 149
    - dependent node communication failure 149
    - Eclock 151
    - error writing to log files 147
    - Estart failure 146, 149
    - fence failure 149
    - generate routes failure 149
    - initialization failure 147
    - interface system call failed 151
    - IP reset failure 150
    - link outage 147
    - lost clock 150, 151
    - NCLL Switch central queue parity error 143
    - NMLL Switch central queue parity error 143
    - node fence request 145
    - node miswired 148
    - open switch catalog failure 150

fault service daemon *(continued)*
SP Switch *(continued)*
port cannot be disabled 150
primary-backup node not responding 146
primary-backup node takeover 146
primary node takeover 146
receiver EDC class error 141
receiver EDC error 142
receiver FIFO overflow error 143
receiver link synchronization error 142
receiver state machine error 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152
reopen adapter windows failure 149
resigning primary - backup node 148
resigning primary node 147
scan failed 148
SDR communication failure 148
sender invalid route error 141
sender link synchronization error 144
sender lost EOP 141
sender parity error 140
sender state machine error 143
sender token error 143
service command failure 148
service logic - bad parity in FIFO 144
service logic - bad parity in route table 145
service logic - incorrect CRC 145
service logic - incorrect packet length 144
service logic - invalid link enable value 145
service logic - send TOD error 145
service logic - state machine error 145
service request failure 147
service send queue is full 147
SIGTERM 146
terminated 151
topology file distribution failure 150
user process killed 147
SP Switch2 186, 188, 214, 215, 216, 218, 223, 224, 226
adapter service error 194
cable miswired 199
chip ID error 194
command broadcast failed 202
communication error 203
daemon error 195
database update error 202
defective switch board 200
EDC-class error 195
EDC error threshold 196
Estart error 199
fence error 202
fence request 198
incorrect CRC 198
incorrect parity error 198
initialization error 202
initialization failure 200
initialization state machine error 193
invalid service command 194
IP error 202
log error 200

fault service daemon *(continued)*
SP Switch2 *(continued)*
multicast lookup table route decoder parity error 193
multicast route error 194
multicast source routed decode table parity error 193
node configuration error 201
node miswire 201
parity error 193, 194, 198
plane miswire 201
port stuck 203
primary-backup node not responding 199
primary-backup node takeover 199
primary-backup resigned 201
primary node takeover 198
primary resigned 200
receiver FIFO overflow 196
receiver port link synch error 195
receiver stat machine error 197
receiver TOD ECC error 193
reopen error 202
repeat count array parity error 194
route error 202
scan error 201
SDR error 201
send queue is full 200
sender hang 193
sender invalid route error 197
sender lost EOP 197
sender parity error 197
sender state machine error 197
sender token count overflow 195
service array overflow latch 194
service logic state machine error 198
service request error 200
SIGTERM 200
source routed multicast ECC error 193
terminated 203
TOD logic error 194
token counter error 193
token error threshold 197
unsupported device 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203
user process error 199
fault_service_Worm_RTG_CS 226, 230, 234, 242
fault_service_Worm_RTG_SP 176, 178, 182, 183
fclose 181, 240
FFDC
Topology Services 384
file
.bak 429
/etc/group 450, 529
/etc/ha/cfg 456, 462
/etc/inittab 112, 175
/etc/krb-srvtab 109
/etc/ntp.conf 510
/etc/passwd 528
/etc/rc.net 136
/etc/security/passwd 501
/etc/services 323, 372, 527

file *(continued)*
/etc/syslog.conf 338
.last 460
.out suffix 168, 222
/tmp/SDR_test.log 130
.Xdefaults 521
act.top 181, 240
adapter.log 214
admin_server.syslog 257
application resource 521
auth_install/log 256, 258
auto.log 503
auto.master 503
auto.u 502, 504
bootptab 104
bosinst_data 110
cable_miswire 163, 216, 218
cadd_dump.out 210
cfg 456, 462
cfgdce.log 356, 361
chgcss.log 217
col_dump.out 211, 218
colad.trace 216
collection.host.list 529
config_info 105
config.log 102
configfb.log 102
console.log 113, 114, 115
css_dump.out 156
css.snap.log 163, 164, 216, 218
CSS_test.log 158, 175, 212
cssadm.debug 157
cssadm.stderr 157
cssadm.stdout 157
cssadm2.debug 211
cssadm2.stderr 211
cssadm2.stdout 212
daemon.log 214
daemon.stderr 162
daemon.stdout 162, 214
DeviceDB.dump 214
dist_topology.log 161
dtbx.failed.trace 163
dtbx.trace 162
Ecommands.log 163, 180, 217, 239
em.default 460
em.msgtrace 464
em.trace 463, 469
emaster.log 233
emasterd.log 217
emasterd.stdout 217
errpt.out 156, 210
etc/plane.info 187
expected.top 161, 169, 215, 224
file.collections 524
filec_config 524
filec_host 524
flt 138, 160, 161, 180, 214, 239
fs_daemon_print.file 162, 215
fs_dump.out 157
hmacls 260, 319, 322

file *(continued)*
hmlogfile 320, 321, 328, 330, 561, 579
hmthresholds 319, 330
host 524, 529
hwevents 336, 337, 338, 339, 341, 342, 343, 344, 345, 346, 347, 348
locating 335
SP_STATE_LOG 337
ifcl_dump.out 210
inetd.conf 104, 105, 106, 107, 108, 109, 312, 313
install_info 105
iptrace.out 386
k5login 303, 307
kerberos.log 257
kerberos_slave.log 257
klogin 116, 275, 278, 303, 307
known_hosts 300
kpropd.log 257, 275
krb.conf 108, 259, 263
krb.realms 108, 259, 263
krb-srvtab 264, 277
krb5.conf 314
la_error.log 217
la_event_d.trace 217
list 529
lists 524
logevnt.out 162, 210
machines.lst 122, 356, 358, 359, 360, 362, 364, 365, 374, 382, 383, 385, 386, 390, 396, 397, 398, 401, 416
net.image 509
netmon.cf 404
netstat.out 159, 213
odm.out 213
optlevel.rsct.basic.hacmp 385
optlevel.rsct.basic.rte 385
optlevel.rsct.basic.sp 385
out.top 154, 161, 169, 182, 206, 215, 224, 242
passwd 501
pe_site 273
phoenix.snap_err 458
phoenix.snapOut 376, 377
plane.info 187, 234
psspfb_script 107
rc.boot 408
rc.net 136
rc.switch 160
rc.switch.log 161, 182, 214, 225, 242
regs.out 157, 212
rhosts 116, 303, 312
router_failed.log1 158, 212
router_failed.log2 158, 212
router.log 158, 212
router.log.old 158, 212
rzdump.RM 463
s1data_dump 561
sami_dump 561
scan 530
scan_out.log 159, 213
scan_save.log 159, 213
script.cust 105, 109, 112

file *(continued)*
  SDR_config.log 126, 570, 571, 575, 589
  SDR_dest 133
  SDR_dest_info 125, 128, 132, 134
  SDR_test.log 130, 170, 225
  sdrdlog 126, 128, 133
  server_name 509, 524
  setup_logd 337
  smit.log 258
  snmpd.conf 557, 558
  spa-srvtab 264
  spd.trace 163, 217, 245
  SPdaemon.log 338, 345, 503, 504, 562, 579
    permissions 338
  spdata.out 159, 213
  spfbcheck 106
  splogd 337
  splogd.debug 336
    locating 335, 336
  splogd.state_changes 337, 346, 347, 348, 349
  spmon_ctest.log 327
  spmon_itest.log 323
  spot.out 94
  spot.updated.out 94
  spsec_defaults 262, 263, 276
  spsec_overrides 91, 101, 108, 125, 135, 137, 185, 258, 261, 262, 263, 273, 289, 302, 318, 333, 352, 420, 456, 474, 492, 514, 560, 578
  srvtab 76, 109, 280
  st_log 492, 493, 496, 497, 498
  st_log.previous 492, 493
  st_verify.log 494, 495, 496
  summlog.out 162, 211
  sup 524
  supfilesrv.pid 524, 525
  supfilesrvr 524
  supper 524
  supscan 524
  switch_node_number 498
  sysctl.conf 116, 260
  sysctld.log 287
  syslog 343, 349
  syslog.conf 304, 305, 338, 343, 344, 348
  SYSMAN_test.log 507
  System Monitor 322
  tb_dump.out 164
  tftpaccess.ctl 105, 106, 107, 108, 109
  topology.data 216
  tuning.cust 106
  vdidl.out 159
  vsd.log 479, 483, 484
  vsd.snapOut 477, 479
  worm.trace 161
File collection symptoms and recovery actions 530
file collections 502, 504, 510
  configuration 524
  diagnosing 523
  files not updated 531
  multiple errors 531
  old files distributed 530
  permission denied 530, 531
file collections *(continued)*
  problems 523
  server busy 530
  server errors 528
  varying errors 531
File collections client errors 530
file collections client log file 525
File collections configuration errors 528
File collections errors on any SP host 527
file collections log file 525
File collections server (control workstation or boot/install server) errors 528
file descriptor
  IBM @server@server pSeries 690 585
  SP-attached server 567
file distribution 525
file server 505
file set
  libspsec.a 370
  perfagent.tools 456
  rsct.basic.hacmp 387, 435
  rsct.basic.rte 387, 435
  rsct.basic.sp 387, 435
  rsct.clients.hacmp 387, 435
  rsct.clients.rte 387, 435
  rsct.clients.sp 387, 435
  rsct.core.utils 387, 435
  ssp.basic 321, 322, 328, 329, 563, 569, 570, 581, 588, 589
  ssp.clients 313
  ssp.csd.gui 517
  ssp.gui 517
  ssp.st 494, 496
  SSP.top 517
  ssp.ucode 322
  X11.Dt.helpinfo 520
file system 21, 27, 40
  /spdata 21, 125, 333
  /spmon 318
  /tmp 21, 37, 38, 43, 49, 64, 278, 279, 378, 471
  /unix 84
  /usr 508
  /var 21, 40, 43, 49, 125, 127, 134, 254, 290, 333, 403, 456, 463
  NFS 51
  root 21, 43, 49, 57, 64, 529
file system, shortage 27, 40
filec_config 523, 525, 528
filecoll_config 527
First Failure Data Capture Programming Guide and Reference 139
flt 180, 239
  SP Switch 160, 161
  SP Switch2 214
font
  SP Perspectives 520
font resources 521
fontList 520
fopen 181, 240
fprintf 181, 240

frame
  controller problems 24
  power supplies 24
  rs232 tty cables 331
frame controller 24, 45
frame controller ID mismatch 45, 56, 62
frame controller not responding 23, 24
frame controller problems 45
frame controller responding 44, 62
frame does not exist 331
frame node slot failure 56, 61, 63
frame node slot failures 45
frame not responding 340, 567, 585
Frame Notebook Status 516
frame number
  IBM @server@server pSeries 690 585
  SP-attached server 567
frame packet 560, 577
frame power 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 62
frame slot 45
frame status 21, 23
frame supervisor 45, 87, 319, 321
frame supervisor card 317, 326, 327
frame supervisor interface 560, 577
frame switch slot failure 63
frame temperature 45, 56, 63
fs_daemon_print.file
  SP Switch 162
  SP Switch2 215
fs_dump.out
  SP Switch 157
fs_init_fail 153
fs_load_fail 153
ftpd 43
full snap
  SP Switch 164
  SP Switch2 219

G

generic service guide 539
genmajor_fail 152, 205
genminor_fail 152, 205
getslot_fail 152, 205
GLSM daemon 423
graphical monitoring tools 30
Graphical tools - SP Event Perspective 40
Graphics tools - SP Event Perspective 31
GRF Configuration Guide - Enable telnet access 555
group
  partition-sensitive 135
Group Leader 122, 393
Group Leader node 428, 433, 434, 450, 451
Group Services 52, 78, 79, 137, 185, 419, 455, 457, 458, 461, 467, 483
  access 421
  AIX Error Log 420
  assert 426
  client 437, 447, 448
  client socket 421

Group Services *(continued)*
  core dump 421, 422, 423, 424, 425, 450
  daemon failure 447
  daemon not loaded 439
  daemon started 424
  daemon stopped 425
  deactivate script 421
  domain 437, 439, 446, 447, 448, 450, 452
  domain merge 421
  domain not formed 422
  error condition from Topology Services 425
  GLSM daemon started 423
  hagsglsm daemon logic failure 422
  hagsglsm start error 423
  hagsglsm stopped 423
  incorrect operation 426
  informational message 424
  internal error 450
  locating a group 441, 442
  log file name 424
  log size 428, 429
  logic failure 422
  long trace 428, 429, 430
  messages 419
  nodes to obtain data from 428
  NodeUp event 425
  proclaim message 425, 447, 450
  protocol 443, 447, 451
  segmentation violation signal 426
  Service log file 436
  short trace 428, 429
  start error 424
  started 424
  Startup Script log 436
  stopped 425
  Subsystem Name 436
  summary log 429
  symptom table 446
  undefined condition 426
  unknown message 424
  User log file 436
Group Services daemon 423, 426, 427, 428, 429, 430, 433, 436, 437, 439, 440, 446, 447, 448, 450, 452
Group Services nameserver 440, 450
Group Services Nameserver 377, 439
Group Services nameserver (NS) node 432
Group Services Nameserver (NS) node 428
Group Services service log trace 427
Group Services service log trace - summary log 429
Group Services startup script log 430
Group Services symptoms 446
GS_Config 378
GS service log trace 427
GS service log trace - summary log 429

H

ha_em_peers 455
ha_em_peers group 460, 461, 466, 467, 468
HACMP: Planning Guide 405
HACMP configuration 390

HACMP Enhanced Scalability Handbook 351, 419
HACMPnim ODM class 383
HACMPtopsvcs ODM class 383
hacssrmd 192
HACWS 4
HACWS supervisor card 45
hadsd 44
haem 437, 438, 519, 520
haemaixos 456, 520
haemd 44, 53, 456, 457, 460, 461, 462, 463, 464, 466, 467, 468, 469, 472, 518, 520
haemrm group 460
haemunlkrm 519
hags 78, 137, 439, 483
hagsd 52, 362, 421, 520
hagsglsm 78, 422, 423, 424, 443, 444, 445, 447
hagsuser group 421, 450
HAL 138
hard disk, diagnosing 115
hardmon 78, 317, 318, 319, 322, 327, 328, 329, 330, 332, 333, 340, 341, 343, 345, 346, 520, 585
  -r flag 327
  polling interval 566, 584
hardmon daemon log file 321
hardmon log file 320
hardmon problems 317
hardware
  control commands 18, 19
  monitoring 18, 19, 20
  status 18, 19, 20
Hardware Abstraction Layer (HAL) 138, 186
hardware error entries, automatic deletion 69
hardware monitor 31
Hardware Monitor 45, 77, 260
hardware monitor log
  IBM @server@server pSeries 690 579
  SP-attached server 561
Hardware Perspective 18
hardware_protocol 564
hardware support
  phone number 13
harmad 457
harmld 457
harmpd 457
hats 78, 122, 137, 483
hats log file 122
hats or topsvcs script log 382
hats script 356
hats script log 382, 384
hatsd 44, 52, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 379, 381, 392, 393, 402, 406, 408, 411, 412, 413
hc 490
hc.activate 488
HC daemon assert 475
hdisk0 70, 73
help 15
Help4Help 516
High-Level SP symptoms 87
High-Level SP Symptoms 87
hmacls 260, 319, 322
HMC 578, 585, 589
HMC communication 586
HMC password files 583
hmcd 520, 584
hmlogfile 320, 321, 328, 330, 561, 579
  example 320
  messages 320
hmrmd 456
hmthresholds 319, 330
host 524, 529
hostMembership 457
HostMembership 455
hostname 6, 8, 9, 122, 132, 133, 135, 505, 527, 529
hostname changed 311
hostResponds 23, 121, 122, 402, 413, 414
How to contact the IBM Support Center 13
How to find the Group Leader (GL) node for a specific group 433
How to Find the GS nameserver (NS) node 432
How to Run the adapter ELA 246
How to run the multiple senders/single receiver test 248
How to run the SP Switch or SP Switch2 stress test 247
How to run the SP Switch or SP Switch2 wrap test 249
How To Use This Book 3
hrd 458
HSD (Hashed Shared Disk) 473
httpd 44
hwevents 336, 337, 338, 339, 341, 342, 343, 344, 345, 346, 347, 348
  locating 335
  SP_STATE_LOG 337
HWEVENTS_PATH 335

I

I/O planar 114
IBM
  hardware support 13
  PC help center 13
  phone numbers 13
  software support 13
IBM @server@server pSeries 660 559
IBM @server@server pSeries 680 559
IBM @server@server pSeries 690 559, 577
  command 577
  daemon log 578
  debug 585
  domain name 585
  frame number 585
  messages 577
IBM @server@server pSeries 690 symptoms 588
IBM Concurrent Virtual Shared Disk (CVSD) 485
IBM Hashed Shared Disks (HSDs) 516, 517
IBM mailing address 14
IBM.PSSP.aixos.FS.%totused 519
IBM.PSSP.CSSLogMon 457
IBM.PSSP.harmld 457, 517

IBM.PSSP.harmpd 457
IBM.PSSP.hmrmd 456, 519
IBM.PSSP.LANAdapter.state 458
IBM.PSSP.pmanrmd 457
IBM.PSSP.Response 458
IBM.PSSP.Response.Host.state 458
IBM.PSSP.Response.Host.State 519
IBM.PSSP.Response.Switch.state 458
IBM.PSSP.Response.Switch.State 519
IBM.PSSP.SDR 457
IBM.PSSP.SP_HW.lcd1 519
IBM.PSSP.SP_HW.Node.powerLED 519
IBM.PSSP.Switch 519
IBM.PSSP.SwitchResponse.state 519
IBM Recoverable Virtual Shared Disk 74, 473
  back-level node 475, 476
  client node 474, 485, 486
  communication adapter mismatch 475
  configuration errors 474
  console log 478, 483, 484
  down-level software 476
  down–level software 475
  node failure 489
  read/write errors 474
  recovery problem 483, 484
  server node 474, 485, 486
  software 475, 476
  subsystem failure 475
  switch failure 489
  varyonvg failure 477
IBM Recoverable Virtual Shared Disk daemon 475, 476, 477
IBM Recoverable Virtual Shared Disk problems 473
IBM Recoverable Virtual Shared Disk subsystem 480, 482, 483
IBM Recoverable Virtual Shared Disk subsystem logging of recovery actions 478
IBM Recoverable Virtual Shared Disk subsystem tracing to the console log 478
IBM Support Center 3, 9, 10
  contacting 13
  mailing address 14
  phone numbers 13
  when to contact 9
IBM Virtual Shared Disk 129, 473, 516, 517
  application hang 489
  backup node down 476
  client node 474, 479, 485, 486
  client timeout 489
  configuration errors 474
  console log 478, 483, 484
  copyin failure 475
  copyout failure 475
  device driver 490
  disk I/O error 490
  EIO error 475, 476
  external function call failure 475
  free memory failure 475
  HC daemon assert 475
  HC subsystem failure 476
  I/O errors 489
IBM Virtual Shared Disk *(continued)*
  I/O requests hang 489
  internal device driver error 475
  internal error 476
  ioctl failure 475
  IP protocol error 476
  IP protocol informational message 476
  memory allocation failure 477
  minor number 485
  mirroring 490
  name 485
  node failure 488, 489
  node hang 488
  read/write errors 474
  recovery services 488
  requests queued 484
  resume failure 476
  routing problem 486
  secondary server 489
  server node 474, 485, 486
  stop and restart 489
  stopped 482, 483, 490
  suspend 490
  suspend failure 477
  suspended 482, 483
  switch failure 488, 489
  terminating 476
  timeouts 484
  timestamps out of sync 477
  virtual shared disk sequence number problem 486
  when adapter recovery is disabled 489
  when adapter recovery is enabled 489
  working collective 474, 480
IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk Error Codes 475
IBM Virtual Shared Disk Node 480
IBM Virtual Shared Disk Node Failure 488
IBM Virtual Shared Disk Perspective 473, 481, 484, 487, 488, 489, 515
IBM Virtual Shared Disk Perspective symptoms 516
IBM Virtual Shared Disk problems 473
IBM Virtual Shared Disk subsystem
  active 482
IBM Virtual Shared Disks 516, 517
ICLUI_TRACE 515
icon
  green 19
  notebook 19
  red X 19
  SP Perspectives 516
icon view
  SP Hardware Perspective 18
id 529
ideal network topology file 161, 215
Identify the conditions of interest. 34
Identify the failing node
  Event Management 467
  SP Switch 174
  SP Switch2 228
if_config 101

ifcl_dump.out
  SP Switch2 210
Important - WHEN actions are performed 38
Important - WHERE actions are performed 39
Iname_error 153
incarnation 428, 429, 432, 441
inetd 43, 50, 65, 104, 105, 106, 107, 108, 109, 303, 312, 313
inetd.conf 104, 105, 106, 107, 108, 109, 312, 313
Information to collect before contacting the IBM Support Center 10
  boot 113
  clustered enterprise server 562
  Event Management 464
  file collection 525
  Group Services 430
  IBM @server@server pSeries 690 580
  IBM Recoverable Virtual Shared Disk 479
  IBM Virtual Shared Disk 479
  JSRT Services 493
  Logging daemon 336
  NIM 94
  node install 102
  Per Node Key Management 292
  remote command 305
  SDR 128
  SP-attached server 562
  SP Perspectives 514
  SP Security Services 258
  SP Switch 164
  SP Switch2 218
  System Monitor 321
  Topology Services 384
informational messages
  System Monitor 318, 319
Initialization errors
  Event Management 460
inittab
  sdrd 134
install image 509
  PSSP 109
install_info 105, 509
install problems
  node 101
installation 95
  secure remote command 111
  time 95
installation trace, sample NIM 97
installation verification
  Job Switch Resource Table Services 494
Installation verification test
  Group Services 434
  IBM Recoverable Virtual Shared Disk 479
  IBM Virtual Shared Disk 479
  Topology Services 386
Installation verification tests
  clustered enterprise server 563
  IBM @server@server pSeries 690 580
  SP-attached server 563
  System Monitor 321
Installation Verification Tests
  Logging daemon 336
instance
  Event Management 470
Internal components of Event Management 456
Internal SP Switch subsystem components 137
Internal SP Switch2 subsystem component 186
Internal virtual shared disk device driver circular trace buffer 478
interpreting system management tests 511
Interpreting the results of the multiple senders/single receiver test 248
Interpreting the results of the SP Switch or SP Switch2 stress test 247
Interpreting the results of the SP Switch or SP Switch2 wrap test 249
Interpreting the test results
  SYSMAN_test 511
ioctl errors
  IBM Recoverable Virtual Shared Disk 475
  IBM Virtual Shared Disk 475
IP address 6, 9, 132, 133, 135, 136, 585
IP protocol informational message 476
IP routing problems 123
IP source routing 123
IPL 541
ipqmaxlen 486
iptrace.out 386
IPv4 address format 510

J

Job Switch Resource Table
  unload 499
Job Switch Resource Table (JSRT) Services symptoms 494
Job Switch Resource Table Services 77, 491
  AIX Error Log 492
  load failure 492, 494, 495
  messages 491, 492
  status 494
  unload failure 492, 494, 495
  window 492, 494, 498, 499
Job Switch Resource Table Services API 492
Job Switch Resource Table Services commands 492
Job Switch Resource Table Services error and information log 492
Job Switch Resource Table Services Error and Information Log 492
job switch resource table services problems
  symptom table 494
Job Switch Resource Table Serviceswindow
  Job Switch Resource Table Services 492
Jobcharge 510
JSRT Services 77, 79, 491
JSRT Services return codes 496

K

k5login 303, 307
K5MUTE 304, 305, 306, 307

kadmind 256, 257, 274
Keeping informed of status changes 28
kerberos 6, 44, 52, 59, 61, 66, 256, 257, 274, 305
kerberos.log 257
Kerberos rsh message 0041-010 313
kerberos_slave.log 257
Kerberos V4 52, 76, 271, 272
  login failure 272, 273
Kerberos V4 authentication database 263, 272
Kerberos V4 authentication token 271
Kerberos V4 configuration file 263
Kerberos V4 daemon logs 256
Kerberos V4 database 274
Kerberos V4 database administration ACL file 264
kerberos V4 key file 276
Kerberos V4 key file 278, 279
Kerberos V4 messages 272, 306
Kerberos V4 principal 263
KERN_PANIC error label 70
kernel memory buffer failures 50
key file 276
key mode switch 81, 82, 83, 541, 542, 543, 547, 551
keytab object 53, 54, 261, 262
kfserver 77
klogin 275, 278, 303, 307
Knowing your SP structure and setup 7
kpropd 256, 257, 274
kpropd.log 257, 275
krb.conf 108, 263
krb.realms 108, 259, 263
krb-srvtab 109, 264, 277
krb5.conf 314
KRB5CCNAME 285
KRBTKFILE 276
krshd 303, 304, 305, 307, 308, 310, 311, 312, 313
ksh 131, 497
kshell 308, 312

L

la_error.log
  SP Switch2 217
la_event_d.trace
  SP Switch2 217
LAN hardware 121
LANG 515
launch pad 513
Launch pad and general Perspectives symptoms 516
LC_MESSAGES 126
LCD 7
LED/LCD 7, 18, 48, 64, 83, 87, 91, 95, 101, 105, 106, 107, 108, 109, 113, 114, 121, 513, 543, 547, 551, 552
  0c0 82, 83
  0c1 83
  0c2 82, 83
  0c4 83, 84
  0c5 83
  0c6 83
  0c7 83
  0c8 83
  0c9 83
LED/LCD *(continued)*
  0cc 83
  0xxx 544, 552
  203w0xyz 548
  231 103
  233 103
  260 103
  40A00000 545
  40B00000 545, 548
  40B0010p 552
  45800001 548
  4B246101 552
  4B246102 552
  4B246110 552
  4B24A101 552
  4B24A102 552
  4B24A110 552
  4B265410 544
  4B276810 544
  4B2xxx01 544
  4B2xxx02 544
  4B2xxx10 544
  600 through 699 91
  611 92, 94
  623 94
  888 11, 12, 26, 81, 83, 88, 177, 231
  AIX 539
  c40 97
  c46 93
  c48 103
  c54 97
  chronological order 535
  E000 547
  E05 552
  E0FF 544, 547
  E105 103, 544
  E175 544, 548, 552
  E1F7 103
  E3xy 551, 552
  even number (status) 91
  Exxx 544, 548
  flashing 23, 26
  hardware 539
  installation 95
  monitoring 29
  NIM 91
  node boot 539
  not blank 23, 25
  numerical order 537
  odd number (error) 91
  SP-specific 535
  u03 103
  u50 103
  u54 103
  u56 103
  u57 103
  u58 103
  u62 103
  u67 103
  u68 103
  u69 103

- LED/LCD *(continued)*
  - u79 103
  - u80 103
  - u87 103
  - values 95, 535
- LED/LCD value 62
- LIBPATH 499
- libspsec.a 370
- limitations 4
- link problems 249
- link status
  - SP Switch 155
  - SP Switch2 208
- list 529
- load default event definition 55
- load_khal_fail 206
- LOAD_ST 492
- locale
  - AIX 126
- locator 34
- locked tty 573
- log 7, 9, 10
  - creating 7
  - Job Switch Resource Table Services error and information 492
  - maintaining 8, 9
  - SP setup 7, 8, 9, 22
  - SP structure 7, 8, 9, 22
  - updating 8, 9
- log file
  - hardmon 320, 321
  - System Monitor 321
- log file hierarchy
  - SP Switch2 187
- log files
  - cleaning up 78
  - not trimmed 79
  - SP Switch2 220
  - trimming 78
- Log files
  - SP Security Services 256
- Log files within css.snap package
  - SP Switch 165, 219
- logevnt.out
  - SP Switch 162
  - SP Switch2 210
- logging 332
- logging daemon 77
- Logging daemon 304, 333, 337, 340, 341, 348, 349
  - debug 335, 336
  - died 348, 349
  - not running 348, 349
  - not running user exit 348
  - not writing to AIX error log 348
  - not writing to AIX Error Log 349
  - not writing to splogd.state_changes 348, 349
  - not writing to syslog 348, 349
  - sample program 341
  - start 336
  - stop 336
  - stopped 348, 349
- Logging daemon *(continued)*
  - symptom table 348
  - terminated 348, 349
- Logging daemon commands 333
- Logging daemon problems 333
- Logging daemon symptoms 348
- logical volume 118, 119
- Logical Volume Manager 457, 484, 486
- login 77
  - fail 501
- Logs
  - IBM @server@server pSeries 690 578
  - SP-attached server 560
- long IPL
  - 604 High Node 541
  - 604e High Node 541
- Look for MBCS data in the SDR 131
- lppsource 94, 95
- lppsource name 92

M

- MACHINE_CHECK_CHRP 544, 548, 552
- machines.lst 122, 356, 358, 359, 360, 362, 364, 365, 374, 382, 383, 385, 386, 390, 396, 397, 398, 401, 416
- mailing address
  - IBM 14
- maintenance mode boot 115
- make_special_fail 153, 206
- Making effective use of the IBM Support Center 9
- Managing and Monitoring the Error Log 68
- Manual method of data collection
  - Event Management 458
- manual pages for public code 610
- manuals
  - AIX 7
  - problem determination 4
  - PSSP 4
  - SP system 6
- master file
  - file collections 530
- Master Switch Sequencer node 186, 231, 232
- MBCS (Multi-Byte Character Set) data 131
- mbuf failure 43, 50, 58, 64
- media card 555
- Membership
  - resource monitor 457
  - resource variable 470
- memory leak
  - SDR 129
- message dialog
  - SP Perspectives 514
- messages 5, 88
  - AIX 7
  - clustered enterprise server 560
  - file collections 523
  - Group Services 419
  - IBM @server@server pSeries 690 577
  - Job Switch Resource Table Services 491, 492
  - SP-attached server 560
  - supper 523

messages *(continued)*
  Topology Services 351
messaging
  required setting of IP source routing 123
mirror 118
mirroring 78, 117, 118, 119
Missing error data warning
  SP Switch 163
  SP Switch2 218
mksysb 95
Monitor these hardware conditions 43
Monitor these software conditions 43
monitoring commands 28
monitoring system status 3, 15
monitoring tools 15
mount 505
MSS node 186
mult_senders_test 248
multiple node problems 11
Multiple senders/single receiver test 248

N

nameserver 133
  Group Services 432
net.image 509
netinst user 509
netinstall problems 101
netmask 388
netmon.cf 404
netmonAdapterHealth 385
netstat.out
  SP Switch 159
  SP Switch2 213
NetView 18
Network Connectivity Table (NCT) 456
network installation management 91
Network installation progress 95
Network Interface Module log 383
Network Time Protocol (NTP) 51
NFS 51, 95, 110, 115, 332, 504, 509
NIM 6, 95, 101, 535, 538, 539
  symptom table 94
NIM (network installation management) 91
NIM client object 111
NIM debug SPOT 92
NIM Debug SPOT 94, 95
NIM environment 102, 103
NIM installation trace, sample 97
NIM log 383
NIM SPOT Logs 94
NIM symptoms 94
node 11, 12, 13
  604 High Node 541
  604e High Node 541
  bad packet 248
  bad sender 248
  boot failure 114
  crash 13
  falls off switch 246
  hang 12, 13, 121
node *(continued)*
  hostname 8
  installation 95, 535
  IP address 8
  key mode switch 81
  not accessible 121
  not responding 12, 121
  not responding to switch 23
  power off 23
  problem 11
  slot connectors 24
  switch_stress 247
node boot 48, 539
  diagnostic mode 115
  maintenance mode 115
node boot LED/LCD 539
  chronological order 538
  numerical order 539
node conditioning 6, 77
node connectivity 527
node controller responding 63
node crash 88
  SP Switch 178
  SP Switch2 234
Node crash
  SP Switch 177
  SP Switch2 231
node does not exist 331
node down 52, 53
node hang 81, 88
node hardware environment indicator 47, 63
node install 101
  symptom table 102
Node installation symptoms 103
node IPL 539
node key mode switch 48, 64
node LED 48
node LED/LCD 48, 64
Node LED/LCD indicators
  boot 113
node level 187
node not reachable 48
node not responding 52, 53
node notebook 91
Node notebook 101, 113
Node Notebook
  AIX OS Resource Variables 516
  All Dynamic Resource Variables 516
  CSS Resource Variables 516
  Environment page 516
  Hardware Resource Variables 516
  Node Environment 516
  Node Status 516
node number
  SP-attached server 567
Node number 509
node packet 560, 577
node panic 183, 242
  SP Switch 178
  SP Switch2 234
node power 25, 43, 47, 63, 110

node power indicator 122
node processorsOffiline 48
node processorsOffline 64
node reachable by Group Service 64
node reachable by RSCT Group Service 48, 64
node reset 48, 81
node.root 524
node serial port 110
node slot 24
node status 21, 23
node supervisor 45, 46
node supervisor card 92, 317, 326, 327
node supervisor interface 560, 578
node temperature 47, 57, 63
node thrashing 37
Node to switch cable diagnostics
  SP Switch 173
  SP Switch2 227
NODE_UP 448
Nodecond Log
  node install 102
nodefail 24
nodefail17 24, 25
Nodes pane 481, 484
Non-SP System applications fail when using rsh or rcp commands 311
nonlocsrcroute option
  no command 123
nonzero return code
  SDR 131, 132
noProfile 515
nosplash 515
not_configured 153, 206
notebook 19
notification 31, 39, 40, 61
NTP 6, 59, 61, 65, 510
NVRAM 68
NVRAM (non-volatile RAM) 552

O

Objects relating to optional system management services 510
Objects tested by SYSMAN_test on all SP nodes (not the control workstation) 509
Objects Tested by SYSMAN_test on the control workstation and boot/install servers 509
Objects tested by SYSMAN_test on the control workstation only 509
Objects tested by SYSMAN_test on the SP nodes 509
ODM
  SRCsubsys class
    cmdargs attribute 329, 330
  SRCsubsys object
    cmdargs attribute 571, 590
ODM errnotify stanzas 71
odm_fail 152, 205
odm.out
  SP Switch2 213
oncoming primary node 178, 182, 234, 241
online help 15
open tty 122
operational verification
  Topology Services 390
operational verification tests
  IBM Recoverable Virtual Shared Disk 481
  IBM Virtual Shared Disk 481
Operational verification tests
  clustered enterprise server 566
  Group Services 436
  IBM @server@server pSeries 690 584
  Logging daemon 339
  SP-attached server 566
Operational Verification Tests
  System Monitor 327
Optional system management objects tested by SYSMAN_test 510
optlevel.rsct.basic.hacmp 385
optlevel.rsct.basic.rte 385
optlevel.rsct.basic.sp 385
Other LED/LCD codes 539
other manuals 4
out.top 154, 169, 182, 206, 224, 242
  SP Switch 161
  SP Switch2 215

P

paging space 22, 26, 35, 37, 43, 49, 64, 332
paging space, shortage 27
partition-sensitive class 135
partition sensitive command
  -G flag 330
partition-sensitive group 135
passwd 501
password files 583
PATH 474, 518, 520
PC failures 14
PC help center
  phone number 13
PC problems 14
pdd_init_fail 206
pe_site 273
PEND error type 70, 72
Per Node Key Management
  symptom table 296
Per Node Key Management (PNKM) 289
Per Node Key Management symptoms and recovery actions 296
perfagent.tools 456
performance
  604 High Node 541
  604e High Node 541
  System Monitor 329
performance problems 13
perspectives 15
Perspectives - related file sets 517
perspectives/profiles 521
phoenix.snap 74, 75, 76, 77, 78, 426, 458
phoenix.snap dump 375, 376
phoenix.snap Dump 427
phoenix.snap_err 458

phoenix.snapOut 376, 377
phone numbers
  IBM 13
physical volume 117, 118, 119, 120
ping 133
plane.info 187, 234
Planning for recovery
  IBM Recoverable Virtual Shared Disk 487
  IBM Virtual Shared Disk 487
platform type 111
PMAN_IVECTOR 39, 40
pmandef commands for specific conditions 62
pmanrmd 457
pmbec 510
PMR 14, 85
pmswitch 510
PNKM 289
port level 187
port number 570, 574
portmap 44, 51, 58, 65
Possible causes of adapter diagnostic failures
  SP Switch 152
  SP Switch2 203
Possible causes of adapter failures
  SP Switch adapter 139
  SP Switch2 adapter 189
Possible causes of fault service daemon failures
  SP Switch 140
  SP Switch2 193
Possible causes of SP Switch2 PCI Adapter failures 204
Possible causes of SP Switch2 TOD management (emasterd) failures 204
Possible causes of TB3MX specific adapter failures
  SP Switch 140
Post-installation customization trace
  node install 102
power_system 524
POWER3 SMP High Node 547
  SP Expansion I/O Unit 548
POWER3 SMP High Node characteristics 547
POWER3 SMP Thin and Wide Node characteristics 543
POWER3 SMP Thin Node 543
  hang 545
  hardware failure 544
  performance 544
POWER3 SMP Wide Node 543
  hardware failure 544
  performance 544
  Service Processor 545
Preparing for your first problem before it happens 7
Preparing to examine and monitor this information 54
Preparing to Monitor the System 32
prerequisite knowledge for this book xix
Preventing memory contention problems with the AIX Workload Manager 411
primary backup node 45, 181, 182, 183, 240, 241, 242
primary dump device 81, 82, 84
primary node 45, 138, 146, 154, 175, 181, 182, 183, 206, 229, 240, 241, 242
primary server 274, 275
principal 262
principal missing 311
Print id 510
problem
  automatic response 38
problem determination 3, 9
  DCE 7
problem determination manuals 4
problem diagnosis 3, 67
Problem Management 6, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 53, 62, 77, 457, 471, 489, 517
Problem Management Record 14
problems 10
  PC 14
  PSSP 26
  reporting 3
  software 26
  SP system 26
processorsOffiline 48
processorsOffline 64
proclaim message 440, 447, 450
Producing a system dump 81
product information 10
program
  xlsfonts 521
protocol
  SAMI 560, 564
Protocol failure errors - Kerberos V4 310
PSSP: Administration Guide 5
PSSP: Diagnosis and Messages Guide 4
PSSP: Diagnosis Guide 3, 4
PSSP: Managing Shared Disks 473
PSSP: Messages Reference 5
pssp.installp 517
PSSP messages 5
pssp_script 74, 102
PSSP software 3
psspfb_script 102, 107
PTF 8, 9, 10

Q

Query the state of the sdrd 128
QUICK-HELP 516
quorum 119, 120

R

rc.boot 408
rc.net 136
rc.switch
  SP Switch 160
rc.switch.log 182, 214, 242
  SP Switch 161
  SP Switch2 214
rcp problems 299
rcp restriction 4
read_tbic 157
read/write errors
  IBM Recoverable Virtual Shared Disk 474

read/write errors *(continued)*
  IBM Virtual Shared Disk 474
rearm command 36, 37
rearm condition 471
rearm expression 16, 17, 35, 36, 38
receiver node 248
Recognizing recovery
  IBM Recoverable Virtual Shared Disk 487
  IBM Virtual Shared Disk 487
Recover an SP Switch node 178
Recover an SP Switch2 node 234
Recover crashed node (Event Management Resource Monitor daemons) 472
Recover EMCDB 472
Recover from a missing or corrupt DCE key file 280
Recoverable Virtual Shared Disk 473
recoveries
  boot 114
  clustered enterprise server 569
  Event Management 472
  file collections 530
  Group Services 446
  IBM @server@server pSeries 690 symptoms 588
  IBM Recoverable Virtual Shared Disk 487
  IBM Virtual Shared Disk 487
  Job Switch Resource Table Services 494
  Logging daemon 348
  NIM 94
  node install 102
  Per Node Key Management 296
  remote commands 309
  SDR 131
  secure remote commands 315
  SP-attached server 569
  SP Perspectives 515
  SP Security Services 271
  SP Switch 177
  SP Switch2 233
  System Monitor 328
  Topology Services 402
Register and unregister Events (EMAPI) 461
registration 16
regs.out
  SP Switch 157
  SP Switch2 212
Related documentation
  boot 113
  clustered enterprise server 559
  Event Management 455
  file collections 523
  Group Services 419
  hardmon 317
  IBM @server@server pSeries 690 577
  IBM Recoverable Virtual Shared Disk 473
  IBM Virtual Shared Disk 473
  Job Switch Resource Table Services 491
  Logging daemon 333
  NIM 91
  node install 101
  Per Node Key Management 289
  remote commands 301
Related documentation *(continued)*
  SDR 125
  SP-attached server 559
  SP Perspectives 513
  SP Security Services 251
  SP Switch 137
  SP Switch2 185
  System Monitor 317
  Topology Services 351
reliable_hostname 499
Reliable Messages Library (PRM) 461
Remote command (rsh/rcp) symptoms and recovery actions 309
Remote command connection problems 313
Remote command diagnostics 306
remote command problems 299
Remote command server configuration errors 313
remote commands 277
Remote commands - rsh and rcp dependencies 303
Remote commands recovery actions 310
Remote to remote rcp 314
Remote to remote rcp errors 314
rename 181, 240
replica server 273
reporting problems 3
Requisite function
  clustered enterprise server 560
  Event Management 455
  file collections 523
  Group Services 420
  hardmon 317
  IBM @server@server pSeries 690 577
  IBM Recoverable Virtual Shared Disk 474
  IBM Virtual Shared Disk 474
  Job Switch Resource Table Services 491
  Logging daemon 333
  NIM 91
  node install 101
  Per Node Key Management 290
  remote commands 302
  SDR 125
  SP-attached server 560
  SP Perspectives 514
  SP Security Services 254
  SP Switch 137
  SP Switch2 185
  System Monitor 317
  Topology Services 351
resource file 515
resource monitor
  dump 463
  haem 519
  haemunlkrm 519
  IBM.PSSP.harmld 517
  unlock 519
Resource Monitor API (RMAPI) 456
Resource Monitor Operations 461
Resource Monitor problems
  Event Management 462
resource name
  sphwlog 562, 579

- Resource Name failure indications
  - SP Switch 138
  - SP Switch2 188
- resource status 29
- resource variable 35, 36
  - fontList 520
  - IBM.PSSP.aixos.FS.%totused 519
  - IBM.PSSP.Response.Host.State 519
  - IBM.PSSP.Response.Switch.State 519
  - IBM.PSSP.SP_HW.lcd1 519
  - IBM.PSSP.SP_HW.Node.powerLED 519
  - IBM.PSSP.SwitchResponse.state 519
  - locator 35, 36, 37
  - name 35, 36
- resource variables 32, 33
  - locator 33, 34, 35
  - name 35
  - value type 33
- Resource Variables 44
- resources
  - monitoring 36
- Response
  - resource monitor 458
  - resource variable 470
- responses
  - boot 114
  - clustered enterprise server 569
  - Event Management 472
  - file collections 530
  - Group Services 446
  - IBM @server@server pSeries 690 symptoms 588
  - IBM Recoverable Virtual Shared Disk 487
  - IBM Virtual Shared Disk 487
  - Job Switch Resource Table Services 494
  - Logging daemon 348
  - NIM 94
  - node install 102
  - Per Node Key Management 296
  - remote commands 309
  - SDR 131
  - secure remote commands 315
  - SP-attached server 569
  - SP Perspectives 515
  - SP Security Services 271
  - SP Switch 177
  - SP Switch2 233
  - System Monitor 328
  - Topology Services 402
- restrictions 4
  - Topology Services
    - IP source routing setting 123
- rhosts 303, 312
- rlogin 501
- rlogind 312
- rmNUM_instances 462
- root
  - crontab file 530
- root access 498
- root authority 352
- root crontab file 523
- root file system 43, 49, 57, 64, 529
- root.SPbgAdm 264
- root user 77, 82, 83, 130, 131, 134, 135, 136, 137, 168, 177, 185, 223, 233, 251, 273, 275, 277, 278, 279, 280, 281, 283, 285, 286, 291, 299, 396, 408, 421, 450, 460, 495, 501, 507, 515, 521, 555
- root user restriction 4
- Root Volume Group 78
  - add space 118
  - add volumes 118
  - extending 118
  - failure 117
  - installation 117
  - locked 117
  - logical volumes 119
  - mirroring 118
  - unlock 118
  - unmirroring 119
  - verify 118
- Root Volume Group Symptoms 117
- Root Volume Groups terminology 119
- route table generato 158, 212
- router.log
  - SP Switch 158
  - SP Switch2 212
- routing
  - IP source 123
- RPA 551
- RPA (RS/6000 Platform Architecture) 543, 547, 551
- rpcentry object 263
- RS/6000 Cluster Technology 6
- RS/6000 High Availability Infrastructure 351, 419
- RS/6000 SP: 604 and 604e SMP High Node Service Guide 542
- RS/6000 SP: POWER3 SMP High Node Service Guide 547
- RS/6000 SP: Problem Determination Guide 6
- RS/6000 SP: PSSP 2.2 Survival Guide 6
- RS/6000 SP: SMP Thin and Wide Node Service Guide 544, 552
- RS/6000 SP: System Service Guide 539
- RS/6000 SP Monitoring: Keeping It Alive 6
- RS/6000 SP Switch Router 553
- RS232 tty 555
- rs232 tty cables 331
- RSCT: Event Management Programming Guide and Reference 455
- RSCT: Group Services Programming Guide and Reference 419
- rsct.basic.hacmp 387, 435
- rsct.basic.rte 387, 435
- rsct.basic.sp 387, 435
- rsct.clients.hacmp 387, 435
- rsct.clients.rte 387, 435
- rsct.clients.sp 387, 435
- rsct.core.utils 387, 435
- rsh problems 299
- rsh restriction 4
- run directory 360, 361, 364, 384
- Run Error Log Analysis 246
- Runtime notification methods 15
- rvsd 483, 490

- rzdump.RM 463

S

- S1 communication 569, 570, 574, 587
- S1 communication log
  - IBM @server@server pSeries 690 579
  - SP-attached server 561
- S1 communication port 567
- S1 tty cable 569
- s1data_dump 561
- s1term 347
- s70 561, 562, 564, 565, 566, 567, 568, 569, 571, 573, 574
- S70 318, 559
- s70 hardware
  - serial interface 561
- s70d 77, 318, 520
- S7A 559
- S80 559
- SAMI 571
- SAMI (Service and Manufacturing Interface) protocol 560, 564
- SAMI communication 568, 569, 570, 573
- SAMI communication port 567
- sami_dump 561
- Sample DCE daemon log 257
- sample formatted output
  - AIX Error Log 67
- Sample NIM Installation Trace 97
- sample program
  - Logging daemon 341
- Save Preferences dialog 19
- scan 530
- SCAN_ERROR_CHRP 544, 548, 552
- scan_out.log
  - SP Switch 159
  - SP Switch2 213
- scan_save.log
  - SP Switch 159
  - SP Switch2 213
- script, example 40
- script.cust 105, 109, 123
- SDR 11, 31, 53, 77, 87, 91, 101, 110, 137, 169, 180, 185, 224, 239, 272, 281, 318, 319, 330, 352, 356, 369, 373, 378, 387, 402, 420, 435, 447, 452, 456, 457, 458, 474, 492, 517, 560, 573, 578, 591
  - access 570, 571, 589
  - Adapter 387
    - adapter_config_status 205
  - Adapter class 136, 359, 360
  - administrator authority 132, 134
  - cannot connect to server 131, 132
  - class
    - corrupted 131, 134
    - missing 131, 134
  - CPU utilization 129
  - DependentNode class 556, 557, 558
  - diagnosing problems 125
  - Error code 005 131, 134
  - Error code 006 132, 134

SDR *(continued)*

  - Error code 026 134
  - Error code 080 132
  - Error code 102 134
  - Frame class 323
    - MACN attribute 323
    - MACN values 323
  - Frame object 564, 565, 569, 570, 571, 582, 589
  - hang 130
  - incorrect data 272, 286
  - information 21
  - memory leak 129
  - Node class
    - switch_node_number attribute 498
  - Node object 564, 570, 571, 588
    - dcehostname 256
    - node_number 176, 230
  - NodeExpansion class 326
  - nonzero return code 131, 132
  - not authorized 286
  - ProcessorExtensionNode 571
  - ProcessorExtensionNode object 564
  - return code 132
  - SP class
    - admin_locale 131
    - SDR_ASCII_only 131
  - SP object
    - authent_server 256
    - cds_server 255
    - cell_name 255
    - sec_master 255
    - ts_auth_methods 255
  - SP_ports class 323, 324, 456
    - daemon attribute 324
    - hostname 324
    - port 324
  - SP_Restricted 111, 112
  - SP Security Services 255
  - Switch class 325
  - switchResponds
    - adapter_config_status 152
  - symptom table 131
  - Syspar class 324, 325, 423, 425, 461
  - Syspar_map class 325, 331
  - Syspar_map object 565, 574, 583
  - Syspar object 255
  - Syspar_ports class 456
  - system partition 129
  - trace 127
  - ts_auth_methods attribute 286
  - TS_Config 410, 428
  - TS_Config class 408
  - VSD Global Volume Group 490
  - write authority 131, 134
  - write failure 131, 134
- sdr (name of SDR daemon error log) 127
- SDR administrator access 130
- SDR_ASCII_only 131
- SDR attributes 125
  - list 125, 134
- SDR catalog 126

SDR classes 125
  list 125, 134
sdr command
  reset 129
SDR command 134
SDR commands 125, 126, 132
SDR_config 126, 234
  script 126
SDR_config.log 126, 565, 570, 571, 575, 589
SDR configuration log
  IBM @server@server pSeries 690 578
  SP-attached server 560
SDR daemon error log 127
SDR daemon trace 127
SDR Daemon Trace 127
SDR_dest 133
SDR_dest_info 125, 128, 132, 134
SDR error log 126, 128
SDR error messages 125
sdr group 128, 129
SDR group 135
SDR_init 126
  script 126
SDR initialization 126
SDR installation 125
SDR messages 125, 126, 132
SDR problems 125
SDR return codes 126
SDR server log 133
SDR symptoms 131
SDR_test.log 130, 170, 225
SDR verification test 130
sdrd 44, 53, 126, 127, 128, 129, 130, 133, 134, 135, 136, 520
sdrdlog 126, 128, 133
secd 44, 54, 60, 61, 66, 295, 297
secondary dump device 82
secondary node 175, 181, 229, 240
secondary server 274, 275
secure remote command 4, 111, 300
secure remote command symptoms and recovery actions 315
secure remote copy 4
security 251
  diagnosing security problems 251
  enhanced 4
security breach 370
Security errors
  Event Management 472
security key file 370, 371
security option
  enhanced 4
security registry 272
sender node 248
sending mail to IBM 14
sending problem data to IBM 14
serial port 110
serial port adapter 574
server IP address 531
server key 290
server key file 263, 267, 276, 277, 278, 279, 280, 293, 294
server missing 311
server_name 509, 524
server node
  IBM Recoverable Virtual Shared Disk 474, 485, 486
  IBM Virtual Shared Disk 474, 485, 486
service guide
  generic 539
service information 12
Service information
  Event Management 458
Service Log long tracing
  Topology Services 379
Service Log normal tracing
  Topology Services 380
service principal 259, 263, 264, 271, 272, 273, 274, 276
service processor
  332 MHz SMP Thin Node 552
  332 MHz SMP Wide Node 552
  604 High Node 541
  604e High Node 541
  POWER3 SMP High Node 547, 548
  SP-attached server 561, 564
Service Processor
  332 MHz SMP Thin Node 551
  332 MHz SMP Wide Node 551
  POWER3 SMP Thin Node 544, 545
  POWER3 SMP Wide Node 544, 545
Service Processor log
  332 MHz SMP Thin Node 552
  332 MHz SMP Wide Node 552
  POWER3 SMP High Node 547
  POWER3 SMP Thin Node 544
  POWER3 SMP Wide Node 544
Service Processor Surveillance
  332 MHz SMP Thin Node problems 552
  332 MHz SMP Wide Node problems 552
  POWER3 SMP High Node 548
  POWER3 SMP Thin Node 545
  POWER3 SMP Wide Node 545
Service Request Number (SRN) 172, 173, 226, 227
service ticket 272
services_config 525, 528, 529
Set Table Attributes dialog 18
setup_logd 337
shared disks 473
Shared Product Object Tree 92
SIGBUS 176, 230
SIGDANGER 176, 230
SIGHUP 127
SIGINT 309, 310
SIGINT error from rsh command 310
Signals
  Event Management 461
SIGTERM 176, 230
SMIT 170, 174, 224, 228, 279, 313, 326, 368, 381, 403, 410, 411, 414, 417, 495, 507, 511, 523, 527, 528
  SDR 130
smit.log 258

snap 85
snapshot compressed package
  SP Switch 156
  SP Switch2 210
SNMP 78, 556
SNMP Agent 557, 558
SNMP communication problem 554
SNMP configuration diagnosis 556
SNMP Manager 557
SNMP trap 18
snmpd 557
snmpd.conf 557, 558
Sockets
  Event Management 461
soft snap
  SP Switch 164, 218
software error entries, automatic deletion 69
software support
  phone number 13
SP access groups 263
SP administrative locale 525
SP-attached server 572
  command 559
  daemon log 560
  debug 567
  frame number 567
  hang 564
  messages 560
  node number 567
  not responding 564
  s70 564, 565, 567, 568, 569, 571, 573, 574
  serial interface 561
  service processor 561, 564
SP-attached Server
  symptom table 569
SP-attached Server symptoms 569
SP configuration diagnosis 553
SP_ERROR_LOG 338, 344, 345
SP Error Log label suffixes mapped to syslog priorities and AIX Error Log types 67
SP error logs 74
SP Event Perspective 16, 20, 31, 32, 38, 39, 40, 45, 47, 48, 53, 55, 60, 62, 471, 515
  help 17, 31
  multiple user considerations 18
  with multiple users 18
SP Event Perspectives symptoms 517
SP Expansion I/O Unit
  POWER3 SMP High Node 548
SP file collections 502
SP hardware log
  IBM @server@server pSeries 690 579
  SP-attached server 562
SP Hardware Perspective 18, 20, 48, 62, 81, 87, 91, 101, 113, 121, 122, 511, 515, 521, 543, 547, 551, 553
  configuration profile 19
  help 18, 19, 20
  limitations 19
  monitoring hardware 19
  Node Notebook 121
  Node Status 121
SP Hardware Perspective *(continued)*
  system profile 19
  table format 18
  user profile 19
SP Hardware Perspective symptoms 516
SP hardware problems 87
SP log files 74
SP Logging daemon 333
SP_NAME 128, 133, 134, 265
sp_passwd 510
SP Perspectives 6, 15, 31, 43, 259, 260, 492
  access 516
  actions not selectable 516
  application resource files 521
  cannot access dependent subsystems 516
  cannot run directly 516
  Event 16
  failure 87
  font 520
  Hardware 18
  help 15, 513, 516, 518, 520, 553
  icon 516
  launch pad 513, 515, 516, 518
  launch pad failure 516
  message 516
  online help 15, 473
  question mark 516
  Security 518
  subsystem access 516
SP Perspectives resource variables 519
SP Perspectives symptom types 515
SP rack or system clock diagnostics 171
SP remote commands
  diagnosing problems 299
SP Security Services 251, 271, 330, 333, 403
  commands 251
  errors 271
  messages 271
SP Security Services configuration errors 258
SP Security Services configuration log 256
SP Security Services symptoms and recovery actions 272
SP Services fail when using Kerberos V5 311
SP software problems 87
SP-specific LED/LCD values 535
SP-specific LED/LCD values (Chronological order) 535
SP-Specific LED/LCD values (Numerical order) 537
SP_STATE_LOG 337, 347
SP Switch 137, 493, 554
  adapter 246
  adapter error 246, 248
  adapter Error Log Analyzer 246
  advanced diagnostics 245
  advanced diagnostics messages 245
  fault service daemon 138
  multiple senders test 248
  stress test 247
  summary error log 69
  symptom table 177
  wrap test 249

SP Switch adapter
  bad packet 139, 140
  device driver I/O error 139
  diag_fail 173
  error threshold exceeded 140
  hardware error 139
  interrupt handler error 139, 140
  link outage 139
  microcode error 139
  ODM update failure 139
  post-diagnostic failure 152
  power-on-self-test diagnostics 152
  pre-diagnostic failure 152
  service interface overrun 139
  summary error log 69
  transient error 140
  user access error 140
SP Switch adapter diagnostics 75, 172
SP Switch adapter diagnosticsfault_service_Worm_RTG_SP 172
SP Switch adapter Service Request Number failures and recovery actions 173
SP Switch advanced diagnostics 177
SP Switch and SP Switch2 advanced diagnostic tools 245
SP Switch device and link error information 154
SP Switch device status and recovery actions 154
SP Switch Diagnostics 169
SP Switch Estart problems and analysis 181
SP Switch external clock diagnostics 170
SP switch hardware
  monitoring 28
SP Switch link status and recovery actions 155
SP Switch node diagnostics 174
SP Switch or SP Switch2 stress test 247
SP Switch or SP Switch2 wrap test 249
SP Switch primary node 146
SP Switch problems 137
SP Switch Router 553
  bad community name 557
  configuration values missing 557
  console log 556
  log 556
  SDR values incorrect 557
SP Switch Router adapter 554
  port in use 557
SP Switch Router configuration diagnosis 555
SP Switch stress test 173, 174
SP Switch symptoms and recovery actions 177
SP Switch worm return codes and analysis 179
SP Switch wrap test 173, 174
SP Switch2 185, 493, 519
  adapter 246
  adapter error 246, 248
  adapter Error Log Analyzer 246
  advanced diagnostics 245
  advanced diagnostics messages 245
  fault service daemon 186
  multiple senders test 248
  stress test 247
  summary error log 69
SP Switch2 *(continued)*
  symptom table 233
  wrap test 249
SP Switch2 adapter
  critical error 191
  critical hardware error 190
  critical microcode error 190
  critical software error 190
  CSS device driver configuration error 192
  CSS device driver error 192
  CSS device driver unexpected interrupt 192
  CSS hacssrmd daemon terminated 192
  diagnostic failure 203
  permanent error 192
  permanent hardware error 191
  permanent microcode error 191
  permanent software error 191
  post-diagnostic failure 203
  summary error log 69
  transient error 190
  transient hardware error 189
  transient microcode error 189
  transient software error 190
SP Switch2 adapter diagnostics 216, 226
SP Switch2 adapter port
  permanent hardware error 189
  transient hardware error 189, 190, 191, 192, 193
SP Switch2 adapter Service Request Number failures and recovery actions 227
SP Switch2 advanced diagnostics 233
SP Switch2 device and link error information 206
SP Switch2 device status and recovery actions 207
SP Switch2 diagnostics 224
SP Switch2 Estart problems and analysis 240
SP Switch2 link status and recovery actions 208
SP Switch2 log and temporary file hierarchy 187
SP Switch2 log files 76, 220
SP Switch2 Node Diagnostics 228
SP Switch2 PCI Adapter
  critical error 204
  event handler 205
  permanent error 204
  recovery 204, 205
SP Switch2 problems 185
SP Switch2 rc.switch.log file and SDR equivalents 225
SP Switch2 stress Test 228
SP Switch2 Stress Test 227
SP Switch2 symptoms and recovery actions 233
SP Switch2 Time-Of-Day (Switch TOD) 231, 232
SP Switch2 Time Of Day (TOD) diagnostics 231
SP Switch2 worm return codes and analysis 235
SP Switch2 wrap Test 228
SP Switch2 Wrap Test 227
SP system
  characteristics 7
  configuration information 21
  monitoring 6, 22, 30
  setup 7
  status 22
  structure 7
SP system hardware 18, 19, 20

SP system hardware *(continued)*
control commands 18, 19
monitoring 18, 19
status 18, 19
SP system partition 324, 325, 330, 331
SP system problem
automatic response 38
SP system problems 3, 7, 15, 26, 67
SP System Security Service 461
SP System Security Services 91, 101, 125, 137, 185, 290, 302, 318, 352, 420, 456, 474, 492, 514, 560, 578
SP systems and PSSP software supported by this book 3
SP trusted services 255, 258, 259, 263, 271, 272, 275, 280, 403
errors 271
messages 271
SP trusted services authentication errors 259
SP trusted services authorization errors 259
spa-srvtab 264
SPD GUI 245, 247, 248
spd.trace 245
SP Switch 163
SP Switch2 217
SPdaemon.log 338, 345, 503, 504
IBM @server@server pSeries 690 579
permissions 338
SP-attached server 562
spdata.out
SP Switch 159
SP Switch2 213
Special troubleshooting considerations 4
spfbcheck 106
sphwlog
IBM @server@server pSeries 690 579
SP-attached server 562
splash program 515
splash window 515
splogd 77, 333, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348
splogd.debug 336
locating 335, 336
splogd.state_changes 337, 346, 347, 348
splogd_test 342
spmgr 557, 558
SPMGR subsystem 556
spmgrd 557
spmgrd-trap service 557
SPMI 456
spmon 22
SPMON
error templates 339
spmon, interpreting results 23
spmon_ctest.log 327
spmon_itest.log 323
spnkeyman 290, 295, 296, 297
spnkeymand 276
SPOT 115
SPOT (Shared Product Object Tree) 92
spot.out 94
spot.updated.out 94
spsec_defaults 262, 263, 276
spsec_overrides 91, 101, 108, 125, 135, 137, 185, 258, 261, 262, 263, 273, 289, 302, 318, 333, 352, 420, 456, 474, 492, 514, 560, 578
spsec-services 292, 293, 294, 295
srcmstr 43, 51, 58, 65
SRN 172
srvtab 280
ssp.basic 321, 322, 328, 329, 563, 569, 570, 581, 588, 589
ssp.clients 313
ssp.csd.gui 517
ssp.gui 517
ssp.st 494, 496
SSP.top 517
ssp.ucode 322
st_log 492, 493, 496, 497, 498
st_log.previous 492, 493
st_status
ST_LOADED_BYOTHER 494
ST_NOT_AUTHEN 495, 496
ST_NOT_AUTHOR 495, 496
ST_RESERVED 494
ST_SECURITY_ERROR 495, 496
ST_SYSTEM_ERROR 494
st_status program 494
st_verify.log 494, 495, 496
state data 560, 577
state logging 337
Stop and start the sdrd 129
Stop monitoring 38
Stopping and restarting automount 505
strictness 117, 118, 120
Subscribe to the Events through Problem Management. 37
subscription 34
subsystem
haem 520
haemaixos 520
summary error log 69, 75
summary error record 69
summary log file 174, 228
Summary log for SP Switch, SP Switch2, and switch adapter errors 69
summlog.out
SP Switch 162
SP Switch2 211
sup.admin 524
supervisor card 317, 326, 327
supfilesrv 510, 525, 527, 528, 529
supfilesrv.pid 524, 525
supfilesrv_port 528
supman 510, 528, 529
supman_uid 528
supper
help 526
update 526
when 526
surveillance
332 MHz SMP Thin Node 552
332 MHz SMP Wide Node 552

- surveillance *(continued)*
  - POWER3 SMP High Node 548
  - POWER3 SMP Thin Node 545
  - POWER3 SMP Wide Node 545
- Svcs.daemons 510
- switch
  - advanced diagnostics 245
  - bad packet 248
  - cable problems 249
  - link 249
  - link problems 249
  - wrap plug 249
- switch adapter error reports 69
- switch admin daemon 74
- Switch Admin daemon 186, 187
- switch advanced diagnostics 75
- switch chip 247
- switch chip error 247
- switch chip link 247
- switch does not exist 331
- switch error 247
- switch error log 69
- switch error reports 69
- switch failure 46, 47, 50, 87
- Switch failure scenarios
  - IBM Virtual Shared Disk 488
- switch fault 75
- switch hardware environment indicator 46, 57, 63
- switch I/O error 65
- switch I/O errors 43, 50
- switch information
  - viewing 69
- switch initialization 75
- switch input error 65
- switch network and topology log 75
- switch_node_number 498
- Switch Notebook Status 516
- switch output error 65
- switch port number 570, 574
- switch power 23, 25, 43, 46, 56, 63
- switch_responds 152, 174, 176, 177, 178, 183, 228, 230, 232, 234, 242, 243
- switch_responds is still on after node panic
  - SP Switch 183
  - SP Switch2 242
- switch_responds0 231
- switch router 75
- Switch Router 553
- switch shutdown 63
- switch status 21
- switch_stress 247
  - command 247
- switch supervisor card 317, 326, 327
- switch temperature 47, 57, 63
- Switch to switch cable diagnostics
  - SP Switch 173
  - SP Switch2 227
- switch topology 6
- switch topology file
  - SP Switch 169
  - SP Switch2 224
- switch transmit queue overflow 65
- switchInfoNeeded 557
- switchInfoNeeded trap 557
- switchResponds 29, 121, 152, 456, 458, 489, 553
  - no 87
  - red 87
- switchtbld 494
- swtbl_adapter_resources 499
- swtbl_clean_table 499
- swtbl_load_job 499
- swtbl_load_table 497
- swtbl_status 499
- swtbl_unload_job 499
- SWTBLAPIERRORMSGS 492, 497
- SWTBLSECMSGLEVEL environment variable values 497
- symptoms
  - boot 114
  - clustered enterprise server 569
  - Event Management 472
  - file collections 526
  - Group Services 446
  - hardmon 328
  - IBM @server@server pSeries 690 symptoms 588
  - IBM Recoverable Virtual Shared Disk 487
  - IBM Virtual Shared Disk 487
  - Job Switch Resource Table Services 494
  - Logging daemon 348
  - NIM 94
  - node install 102
  - Per Node Key Management 296
  - remote commands 309
  - SDR 131
  - secure remote commands 315
  - SP-attached server 569
  - SP Perspectives 515
  - SP Security Services 271
  - SP Switch 177
  - SP Switch2 233
  - System Monitor 328
  - Topology Services 402
- syncd 408
- Sysctl 78, 263, 264, 266, 267, 270, 275
- sysctl.conf 260
- Sysctl configuration file 260
- sysctld 285
- syslog 67, 338, 348, 349
- syslog.conf 304, 305, 338, 343, 344, 348
- syslogd 67, 304, 305, 307
- sysman 67
- SYSMAN_test.log 507
- system administration
  - command line tools 20
- system administrator 15, 16, 17, 18, 20, 21, 30, 31, 35, 36, 37, 38, 39
- system call 330
- system condition 16
- System connectivity symptoms 121
- system controls 81
- system crash 81
- System Data Repository 11

system dump 11, 12, 23, 26, 48, 81, 85
System Dump Information 81
System dump status codes 83
System Management configuration 78
System Monitor 47, 48, 78, 87, 137, 185, 260, 263, 268, 270, 272, 275, 277, 286, 317, 322, 327, 333, 340, 456, 518, 560, 561, 564, 566, 570, 571, 573, 577, 578, 584, 588, 590, 591
  AIX Error Log 318
  authentication 321, 330
  authorization 330
  client command 319
  configuration error 319
  configuration errors 320
  data packet error 319
  file problem 319
  incorrect client command 319
  informational messages 318
  internal error 318, 319, 320
  polling interval 566, 571, 584, 590
  resource not available 319
  system call failures 320
  topology 88
System Monitor command 317, 328
System Monitor daemon log file 320
System Monitor Daemon Log File 321
System Monitor messages 317
System Monitor performance 329
System Monitor problems 317
System Monitor symptoms 328
system monitoring 15
system nameserver 133
system partition 122, 133, 134, 135, 136, 324, 518
  default 132
system partitioning 6, 8
System Performance Measurement Interface 461
System Performance Measurement Interface (SPMI) 456
system profile 515
system resources 16
system status
  monitoring 3, 15
  monitoring tools 15
system verification test
  output 507
SystemGuard program 541, 542
systemProfile 520

T

T/EC Event Adapter 593
table attribute
  SP Hardware Perspective 18
table view
  SP Hardware Perspective 18
Tailor the monitoring 37
tb_dump.out 164
TBIC test 159, 213
TBSswitchInit 181
telephone numbers
  IBM 13
telnetd 312
temporary file hierarchy
  SP Switch2 187
tftp 104, 105, 106, 107, 108, 109, 509
tftpaccess.ctl 105, 106, 107, 108, 109
tftpboot 104, 105, 106, 107, 108, 109
thewall 486
ticket-granting ticket 271, 279
Time Of Day (TOD) Diagnostics
  SP Switch2 231
Tivoli Enterprise Console Event Adapter 593
Tivoli Enterprise Console Event Adapter configuration file 593
token ring adapter 510
tools 15
  monitoring 30
Topology DARE 368, 413
topology file 75, 169, 224
Topology Services 51, 52, 78, 79, 122, 123, 137, 173, 185, 276, 421, 422, 424, 425, 427, 432, 435, 437, 438, 442, 443, 446, 448, 450, 456, 458, 483
  adapter address 362
  adapter configuration problem 394, 402, 403
  adapter enabled for IP 395, 396
  adapter failed 404
  adapter membership group 399, 402
  adapter verification 390
  AIX-related problem 360
  authentication failure 370, 371
  authentication method 416
  broadcast message 374
  cannot create directory 372
  client library error 362
  configuration data 389
  configuration file 364
  configuration instance 398
  configuration problem 416
  connection request 364
  core file 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375
  CPU utilization 355, 393
  credentials 356
  daemon blocked 406, 414
  daemon failed 406
  daemon log file 363
  daemon started 373
  daemon stopped 373
  DCE security mode 370
  dead man switch timer 357
  Defd 390
  directory creation failure 372
  duplicate IP address 359
  duplicate network name 358
  duplicate node number 360
  excessive adapter traffic 413
  excessive disk I/O 406
  excessive interrupt traffic 406
  excessive paging 406
  failure 123
  heartbeat 361

Topology Services *(continued)*
- incorrect flags 355
- incorrect IP address 360, 404
- ioctl failure 360
- IP address 360, 361
- IP communication problem 406, 413
- IP connectivity 400, 401
- IP packets corrupted 370, 371
- IP packets received 396
- IPC key 372
- key file 418
- late heartbeat 361
- listening socket 364
- load failure 393
- local adapter 394, 404, 413, 414
- local adapter disabled 391
- local adapter down 362
- local adapter fenced 404
- local adapter incorrectly configured 365
- local node missing 359
- local node number unknown 367
- lost heartbeat 357
- machines.lst file 364
- Mbrs 390
- mbuf shortage 406, 413
- memory problems 407
- memory shortage 406
- migration-refresh error 364
- missing local node 359
- network configuration problems 403
- network connectivity 399
- network traffic 405
- node death 439
- node down 394, 400
- node not responding 401
- node number duplicated 360
- node reachability 400
- open socket error 369
- packet exchange 404
- partial connectivity 393, 397, 405
- peer communication 369
- peer daemon 370, 371
- port number 372
- problem 123
- refresh 398, 413, 414, 415
- refresh error 368
- refresh failure 402
- remote adapter 393, 404, 405, 413, 414
- remote nodes 367
- remote nodes down 367
- required setting of IP source routing 123
- run directory 390
- SDR error 369
- security breach 371
- security status 398
- semaphore segment 371
- sensitivity factor 409
- service log file 390
- shared memory segment 371, 372
- simulated node death 439
- singleton unstable membership group 392, 393

Topology Services *(continued)*
- singleton unstable state 397
- stale IP packets 370, 371
- startup script 354, 356, 361, 402
- state values 442
- status 390, 392
- subnet mask 399
- subsystem name 390
- switch down 404
- symptom table 402
- system partition name 373
- thread 374
- Time-Of-Day clock not synchronized 370, 371
- tuning parameters 361
- unicast message 374
- unstable singleton state 374
- user log file 390

Topology Services daemon 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375
- assert 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375
- exited 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375
- internal error 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375

Topology Services Group Leader 377, 379, 380, 384, 393, 397, 399

Topology Services or recovery service daemon failure
- IBM Virtual Shared Disk 489

Topology Services problems 351

Topology Services service log 379

Topology Services symptoms 402

Topology Services user log 381

topsvcs script log 382, 384

trace
- spmgr 557, 558

Trace facility built into Event Management 463

trace file
- clustered enterprise server 562
- IBM @server@server pSeries 690 580
- SP-attached server 562

Trace information
- boot 113
- clustered enterprise server 562
- Event Management 463
- Group Services 427
- IBM @server@server pSeries 690 580
- IBM Recoverable Virtual Shared Disk 478
- IBM Virtual Shared Disk 478
- SP-attached server 562
- SP Switch 160
- SP Switch2 214
- System Monitor 321
- Topology Services 379

Trace Information
- NIM 92
- node install 102

Trace Information *(continued)*
  SDR 127
trademarks 598
Transforming System Monitor Experience into Hardware Perspectives Skills 20
troubleshooting 3, 4
  messaging problems
    required setting of IP source routing 123
trusted services 255
trusted services authentication method 134
ts_auth_methods 128, 258, 293, 303, 321
TS_Config 378
tty
  locked 573
tuning.cust 106

U

ucode 322
Unable to rsh, rcp, telnet, rlogin, but ping shows host is up 312
unconfig_admin 313
undefine_system_fail 153
Understand What You Want to Monitor 32
Unfence an SP Switch node 181
Unfence an SP Switch2 node 240
Unknown service: kshell or TCP 312
UNLOAD_ST 492
unmirroring 78, 117, 118, 119, 120
Update the log whenever a failure occurs 8
Update the log whenever system conditions change 9
user access 88
user access logs 21
user access problems 501
  symptom table 501
User Access symptoms 501
user.admin 524, 529
user authentication 28
user authorization 22
user exit 341, 348
  Logging daemon 339
  splogd 339
user profile 515
userProfile 520
Using secure remote command 300
Using SystemGuard and BUMP Programs 541
Using the AIX Error Notification Facility 70

V

vdidl.out
  SP Switch 159
verification check of system management 508
Verification of Mirroring or Unmirroring for Root Volume Groups 119
verification test
  system management 507
Verification test output
  SYSMAN_test 507
Verify Configuration for DCE 261
Verify configuration of Kerberos V4 263
Verify event registration
  Event Management 468
Verify Resource Monitors
  Event Management 472
Verify software installation
  SP Switch 174
  SP Switch2 228
Verify SP Security Services configuration 261
Verify SP software installation
  Event Management 464
Verify SP Switch node operation 175
Verify SP Switch2 node operation 229
Verify the SP Switch topology configuration 169
Verify the SP Switch2 topology configuration 224
Verify the System Data Repository (SDR)
  SP Switch 170
  SP Switch2 224
Verify the System Dump 83
Verifying System Management installation 507
view 502
Viewing Error Log Information in Parallel 68
virtual keys 82
Virtual Shared Disk 473
Virtual Shared Disk node failure 488
Vital Product Data 77
VMM 411
Volume Group Descriptor Areas (VGDAs) 120
Volume Group Status Areas (VGSAs) 120
vsd.log 479, 483, 484
vsd.snap 477, 479
vsd.snapOut 477, 479

W

WCOLL 474, 480
What system management verification checks 508
WHEN actions are performed 38
When to contact the IBM Support Center 9
When to Run the adapter ELA 246
When to run the Multiple senders/single receiver test 248
When to run the SP Switch or SP Switch2 stress test 247
When to run the SP Switch or SP Switch2 wrap test 249
When to use this book 3
WHERE actions are performed 39
wide node 45
WLM 407
working collective
  IBM Virtual Shared Disk 474, 480
Workload Manager 407
worm
  SP Switch 138
  SP Switch2 186, 188
Worm error recovery
  SP Switch 178
  SP Switch2 235
worm.trace
  SP Switch 161
wrap plug 249

wrap_test 249
writeDeviceDatabase 181, 240
wrom trace 76

X

X-server 16, 516, 518, 521
X11.Dt.helpinfo 520
X11-server 521
XAPPLRESDIR 521
xhost 16
xilinx_load_fail 153
xilinx_system_fail 153
xlsfonts 521
xntpd 44, 51, 59, 61, 65, 510

Readers' Comments — We'd Like to Hear from You

Parallel System Support
Programs for AIX
Diagnosis Guide
Version 3 Release 4

Publication No. GA22-7350-03

Overall, how satisfied are you with the information in this book?

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied
Overall satisfaction	□	□	□	□	□

How satisfied are you that the information in this book is:

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied
Accurate	□	□	□	□	□
Complete	□	□	□	□	□
Easy to find	□	□	□	□	□
Easy to understand	□	□	□	□	□
Well organized	□	□	□	□	□
Applicable to your tasks	□	□	□	□	□

Please tell us how we can improve this book:

Thank you for your responses. May we contact you? □ Yes □ No

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

Name

Company or Organization

Phone No.

Address

Readers' Comments — We'd Like to Hear from You

GA22-7350-03

Cut or Fold
Along Line

Fold and Tape **Please do not staple** Fold and Tape

PLACE
POSTAGE
STAMP
HERE

IBM Corporation
Department 55JA, Mail Station P384
522 South Road
Poughkeepsie NY 12601-5400

Fold and Tape **Please do not staple** Fold and Tape

Cut or Fold
Along Line

Program Number: 5765-D51

GA22-7350-03