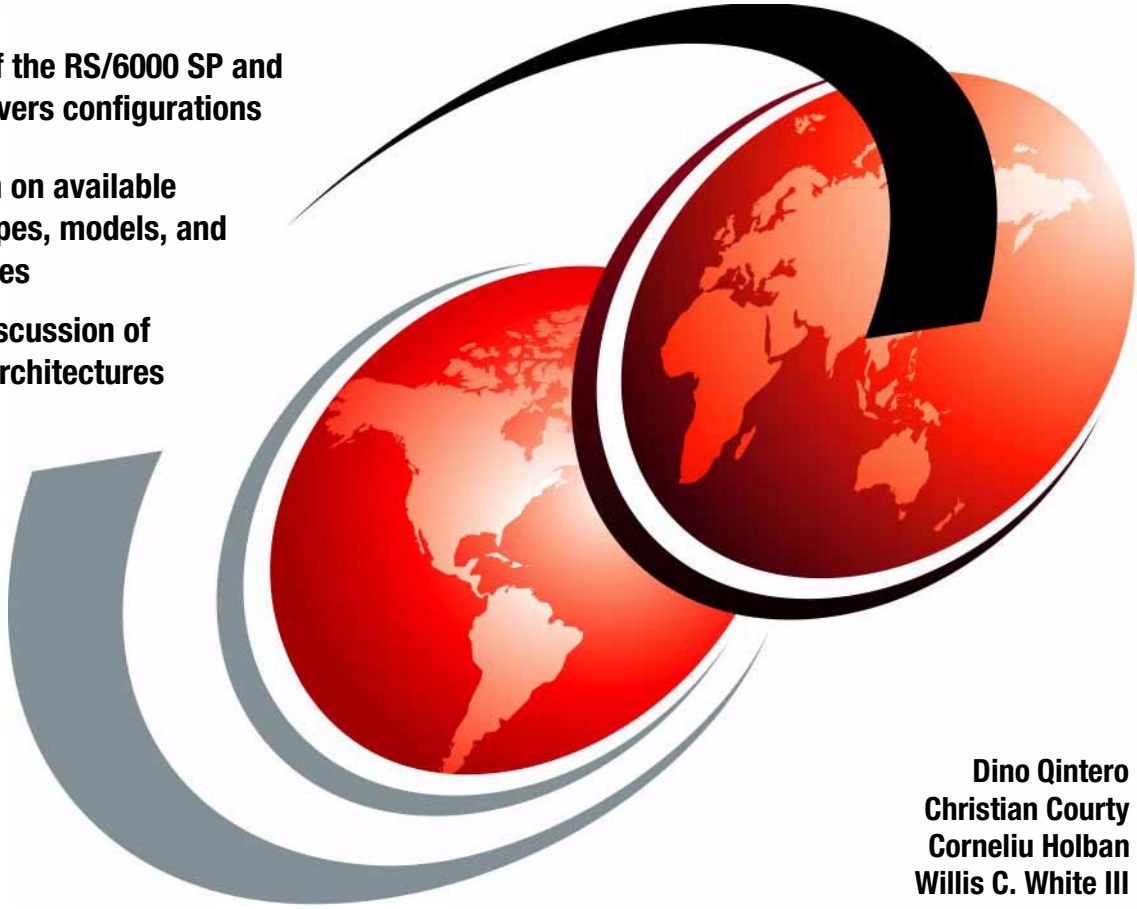
IBM

RS/6000 SP and Clustered IBM @server pSeries Systems Handbook

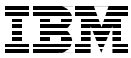
- Overview of the RS/6000 SP and pSeries servers configurations
- Information on available machine types, models, and feature codes
- In-depth discussion of hardware architectures



Dino Qintero
Christian Courty
Corneliu Holban
Willis C. White III

Redbooks

ibm.com/redbooks



International Technical Support Organization

RS/6000 SP and Clustered IBM @server pSeries Systems Handbook

September 2001

SG24-5596-02

Take Note! Before using this information and the product it supports, be sure to read the general information in “Special notices” on page 441.

Third Edition (September 2001)

This edition applies to Version 3, Release 2 of the IBM Parallel Support Programs for AIX (PSSP) Licensed Program (product number 5765-D51), and AIX Version 4 Release 3 Licensed Program (product number 5765-C34).

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. JN9B Building 003 Internal Zip 2834
11400 Burnet Road
Austin, Texas 78758-3493

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

Summary of changes

This section describes the technical changes made in this edition of the book and in previous editions. This edition may also include minor corrections and editorial changes that are not identified.

Third Edition, September 2001

This revision reflects the addition, deletion, or modification of new and changed information described below.

New information

- Blue Hammer clusters are discussed in Chapter 3, “SP-attached servers and Clustered Enterprise Servers” on page 77.
- Control workstation overview and requirements are discussed in Chapter 7, “Control workstations” on page 195.
- Clustering configurations are discussed in Chapter 10, “Clustering configurations” on page 285.

Changed information

- Overview of the RS/6000 SP and Clustered @server pSeries is discussed in Chapter 1, “Introduction” on page 1.
- SP cluster components are discussed in Chapter 2, “SP internal processor nodes” on page 25.
- SP-attached servers are discussed in Chapter 3, “SP-attached servers and Clustered Enterprise Servers” on page 77.
- SP Frame components are discussed in Chapter 4, “Frames” on page 139.
- The switch is discussed in Chapter 5, “The Switch” on page 159.
- SP Switch Routers are discussed in Chapter 6, “SP Switch Routers 9077” on page 181.
- PCI communication adapters are discussed in Chapter 8, “PCI communication adapters” on page 227.
- Software support is discussed in Chapter 9, “Software support” on page 261.

Contents

Summary of changes iii
Third Edition, September 2001 iii

Figures xv

Tables xvii

Preface xix
The team that wrote this redbook xix
Special notice xxi
IBM trademarks xxi
Comments welcome xxi

Chapter 1. Introduction 1
1.1 Overview on RS/6000 SP and Clustered @server pSeries 2
1.2 Features and benefits 4
1.3 Business solutions 7
1.4 Hardware components 9
1.4.1 Processor nodes 12
1.4.2 Frames 15
1.4.3 Extension nodes 16
1.4.4 Switches 17
1.4.5 Control workstations 17
1.5 RS/6000 SP and AIX clustering 18

Chapter 2. SP internal processor nodes 25
2.1 SP internal node 26
2.2 375 MHz POWER3 SMP nodes 27
2.2.1 375 MHz POWER3 processor 28
2.2.2 System memory 30
2.2.3 I/O subsystem 30
2.2.4 Service processor 31
2.2.5 System firmware and RTAS 32
2.2.6 Dynamic processor deallocation 32
2.2.7 375 MHz POWER3 SMP High Node 32
2.2.8 375 MHz POWER3 SMP Thin and Wide Nodes 35
2.3 375 MHz POWER3 SMP High Node (F/C 2058) 37
2.3.1 PCI bus description 37
2.3.2 Requirements 37

2.3.3 Options . . . 37
2.3.4 Processor requirements and options . . . 38
2.3.5 Memory requirements and options . . . 38
2.3.6 Disk requirements and options . . . 40
2.3.7 Switch and communication adapter requirements and options . . . 40
2.3.8 I/O adapters . . . 41
2.3.9 SP Expansion I/O unit (F/C 2055) . . . 41
2.3.10 Hard disk drive options . . . 42
2.4 375 MHz POWER3 SMP Wide Node (F/C 2057) . . . 43
2.4.1 Bus description . . . 43
2.4.2 Requirements . . . 43
2.4.3 Options . . . 44
2.4.4 Processor requirements and options . . . 44
2.4.5 Memory requirements and options . . . 45
2.4.6 Hard disk drive requirements and options . . . 46
2.4.7 Switch and communication adapter requirements . . . 47
2.5 375 MHz POWER3 SMP Thin Node (F/C 2056) . . . 48
2.5.1 Bus description . . . 48
2.5.2 Requirements . . . 48
2.5.3 Options . . . 49
2.5.4 Processor requirements and options . . . 49
2.5.5 Memory requirements and options . . . 50
2.5.6 Hard disk drive requirements and options . . . 51
2.5.7 Switch and communication adapter requirements . . . 51
2.6 POWER3 SMP nodes . . . 52
2.6.1 POWER3 SMP node system architecture . . . 53
2.6.2 POWER3 microprocessor . . . 53
2.6.3 6xx bus . . . 54
2.6.4 System memory . . . 54
2.6.5 I/O subsystem . . . 54
2.7 222 MHz POWER3 SMP High Node (F/C 2054) . . . 55
2.7.1 PCI bus description . . . 55
2.7.2 Requirements . . . 55
2.7.3 Options . . . 56
2.7.4 Processor requirements and options . . . 57
2.7.5 Memory options . . . 57
2.7.6 Disk requirements and options . . . 58
2.7.7 Switch and communication adapter requirements and options . . . 58
2.7.8 I/O adapters . . . 59
2.7.9 SP Expansion I/O unit (F/C 2055) . . . 59
2.7.10 Hard disk drive options . . . 60
2.8 332 MHz SMP nodes . . . 61
2.8.1 332 MHz SMP node system architecture . . . 61

2.8.2 SMP system bus 62
2.8.3 Processor/L2 cache controller. 63
2.8.4 Memory-I/O controller 64
2.8.5 I/O bridge bus 65
2.8.6 PCI I/O bridges 66
2.9 332 MHz SMP Wide Nodes (F/C 2051) 67
2.9.1 Bus description 67
2.9.2 Adapter placement restrictions 67
2.9.3 Requirements 68
2.9.4 Options 69
2.9.5 Processor requirements and options 69
2.9.6 Memory requirements and options 70
2.9.7 Disk requirements and options 70
2.9.8 Switch adapter requirements and options. 71
2.10 332 MHz SMP Thin Nodes (F/C 2050) 71
2.10.1 Bus description 72
2.10.2 Single SMP thin node configuration rules 72
2.10.3 Requirements 73
2.10.4 Options 73
2.10.5 Processor requirements and options 74
2.10.6 Memory requirements and options 74
2.10.7 Disk requirements and options 74
2.10.8 Switch adapter requirements and options. 75

Chapter 3. SP-attached servers and Clustered Enterprise Servers 77
3.1 Overview 78
3.2 SP-attached servers / CES - M/T 7017 supported models 79
3.2.1 @server pSeries 680 Model S85 (7017-S85) 79
3.2.2 RS/6000 Enterprise Server 7017 Model S80 (7017-S80) 85
3.2.3 RS/6000 Enterprise Server 7017 Model S7A (7017-S7A) 89
3.2.4 RS/6000 Enterprise Server 7017 Model S70 (7017-S70) 92
3.2.5 Summary features of the SP-attached servers/CES - M/T 7017 . . . 95
3.3 SP-attached servers/CES - M/T 7026 supported models 97
3.3.1 RS/6000 Enterprise Server 7026 Model M80 (7026-M80) 97
3.3.2 @server pSeries 660 Model 6H1 (7026-6H1) 103
3.3.3 @server pSeries 660 Model 6H0 (7026-6H0) 108
3.3.4 RS/6000 Enterprise Server 7026 Model H80 (7026-H80). 109
3.3.5 Summary features of the SP-attached servers/CES - M/T 7026 . . 114
3.3.6 Cable and PCI card for SP CWS attachment (F/C 3154) 116
3.4 SP-attached server environment 117
3.4.1 SP-attached server installation requirements 119
3.4.2 Planning the SP-attached server interfaces 122
3.4.3 SP-attached server configurator rules. 126

3.4.4 SP-attached servers connectivity diagram 128
3.4.5 Configuring Service Director on SP-attached servers. 128
3.5 Clustered Enterprise Server environment 129
3.5.1 Clustered Enterprise Server installation requirements 129
3.5.2 Planning the Clustered Enterprise Server interfaces 132
3.5.3 Clustered Enterprise Servers connectivity diagram 136
3.5.4 Configuring Service Director on CES systems 137

Chapter 4. Frames 139
4.1 Frames. 140
4.1.1 SP frame dimensions 142
4.1.2 Frame electrical power requirements 143
4.1.3 SP frame redundant power supply (F/C 3885, 3886) 144
4.2 Tall frames 145
4.3 Model 550 frames. 146
4.3.1 Model 550 non-switched configuration 146
4.3.2 Model 550 SP Switch-8 configuration 147
4.3.3 Model 550 single-stage SP Switch and SP Switch2 configuration . 147
4.3.4 Model 550 two-stage SP Switch and SP Switch2 configuration . . . 147
4.4 Model 555 standalone switch for clustered nodes. 148
4.5 Expansion frames (F/C 1550). 149
4.5.1 Using F/C 1550 non-switched expansion frames 151
4.5.2 Using F/C 1550 switched expansion frames. 152
4.6 SP Switch frames (F/C 2031, 2032) 153
4.7 Short frame 154
4.8 Model 500 frames. 155
4.8.1 Model 500 non-switched configuration 155
4.8.2 Model 500 switched configuration. 156
4.8.3 Expansion frames (F/C 1500) 156
4.8.4 Using F/C 1500 non-switched expansion frames 156
4.8.5 Configurations with processor nodes and a switch 157
4.8.6 Configurations with processor nodes only 157

Chapter 5. The Switch 159
5.1 SP Switch. 160
5.2 SP Switch components. 160
5.2.1 SP Switch link or data cable 160
5.2.2 SP Switch ports. 161
5.2.3 SP Switch chip 161
5.2.4 SP Switch board 163
5.2.5 SP Switch adapter. 163
5.2.6 SP Switch system 164
5.3 SP Switch, 16-port (F/C 4011) 166

5.4 SP Switch-8 (F/C 4008) 167
5.5 SP Switch adapter (F/C 4020, 4022, and 4023) 168
5.6 Special SP Switch adapter (F/C 4021 and F/C 8396) 169
5.6.1 SP Switch Router (Machine Type 9077) 169
5.7 Switch cabling 169
5.7.1 Switch-to-node cabling 170
5.7.2 Switch-to-switch cabling 171
5.8 SP Switch2 174
5.8.1 SP Switch2 (F/C 4012) 176
5.8.2 SP Switch2 adapter (F/C 4025) 178

Chapter 6. SP Switch Routers 9077 181
6.1 The 9077 SP Switch Routers 182
6.2 Overview 183
6.3 Installation requirements 186
6.3.1 System requirements 187
6.3.2 SP Switch Router adapter (F/C 4021) 187
6.3.3 Network media card requirements 188
6.3.4 Software requirements 189
6.4 Network interface 189
6.4.1 Connecting to the control workstation 190
6.4.2 Connecting to multiple SP systems 191
6.4.3 Connecting to the SP Switch 192

Chapter 7. Control workstations 195
7.1 Overview 196
7.2 Installation requirements 198
7.2.1 Supported pSeries and RS/6000 workstations and servers 198
7.2.2 System requirements 199
7.2.3 Interface adapter requirements 201
7.2.4 Software requirements 203
7.2.5 SP control workstation offerings 205
7.3 High Availability Control Workstation 220
7.3.1 Overview 221
7.3.2 Installation requirements 223
7.3.3 Limits and restrictions 224
7.4 Service Director on the CWS 225

Chapter 8. PCI communication adapters 227
8.1 PCI Communication adapters supported on the SP 228
8.2 Descriptions of adapter supported on the SP 231
8.2.1 IBM Network Terminal Accelerator adapter (F/C 2403) 231
8.2.2 SCSI-2 Differential Fast/Wide adapter/A (F/C 2416) 232
8.2.3 SCSI-2 Differential External I/O Controller (F/C 2420) 232

8.2.4 IBM Short Wave Serial HIPPI PCI adapter (FC 2732) 232
8.2.5 Long-Wave Serial HIPPI adapter (F/C 2733) 233
8.2.6 IBM Network Terminal Accelerator adapter (F/C 2402 & 2403) . . . 234
8.2.7 SCSI-2 Fast/Wide adapter/A (F/C 2416). 234
8.2.8 SCSI-2 Differential External I/O Controller (F/C 2420) 234
8.2.9 8-Port Async adapter - EIA-422A (F/C 2940) 234
8.2.10 SYSKONNECT SK-NET FDDI adapters (F/C 2741, 2742 & 2743)234
8.2.11 S/390 ESCON Channel PCI adapter (F/C 2751) 236
8.2.12 PCI Auto LANstreamer Token Ring adapter (F/C 2920). 237
8.2.13 EIA 232/RS-422 8-Port Asynchronous adapter (F/C 2943) 237
8.2.14 WAN RS232 128-port adapter (F/C 2944) 238
8.2.15 Turboways 622 Mbps PCI MMF ATM adapter (F/C 2646) 240
8.2.16 ARTIC960Hx 4-port Selectable adapter (F/C 2947) 240
8.2.17 2-port Multiprotocol X.25 adapter (F/C 2962) 241
8.2.18 ATM 155 TURBOWAYS UTP adapter (F/C 2963) 242
8.2.19 Ethernet 10/100 MB adapter (F/C 2968). 243
8.2.20 Gigabit Ethernet - SX adapter (F/C 2969). 244
8.2.21 10/100/1000 Base-T Ethernet PCI adapter (FC 2975) 245
8.2.22 TURBOWAYS 100 ATM adapter (F/C 2984) 245
8.2.23 Ethernet LAN PCI adapters (F/C) 2985 and 2987) 245
8.2.24 ATM 155 MMF adapter (F/C 2988). 246
8.2.25 Four-Port 10/100 BASE-TX Ethernet PCI adapter (F/C 4951) . . . 247
8.2.26 Crytographic coprocessor PCI adapter (F/C 4958). 248
8.2.27 High-Speed Token-Ring PCI adapter (F/C 4959) 249
8.2.28 SCSI-2 Ultra/Wide DE PCI adapter (F/C 6204) 250
8.2.29 Dual Channel Ultra2 SCSI PCI adapter (F/C 6205) 250
8.2.30 SCSI-2 Ultra/Wide Single Ended PCI adapter (F/C 6206) 251
8.2.31 Ultra SCSI Differential adapter (F/C 6207) 252
8.2.32 SCSI-2 F/W Single-Ended adapter (F/C 6208). 253
8.2.33 SCSI-2 F/W Differential adapter (F/C 6209) 254
8.2.34 SSA RAID 5 adapter (F/C 6215). 254
8.2.35 SSA Fast-Write Cache Module (F/C 6222). 255
8.2.36 Advanced SerialRAID Adapter (F/C 6225) 255
8.2.37 Gigabit Fibre Channel PCI adapter (F/C 6227). 255
8.2.38 Advanced SerialRAID Plus PCI adapter (F/C 6230). 256
8.2.39 ARTIC960RxD Quad Digital Trunk adapter (F/C 6310) 258
8.2.40 ARTIC960RxF Digital Trunk Resource adapter (F/C 6311) 259

Chapter 9. Software support 261
9.1 Parallel System Support Programs (PSSP) 262
9.1.1 Administration and operation 262
9.1.2 RS/6000 cluster technology (RSCT). 263
9.1.3 System monitoring and control 264

9.1.4 IBM Virtual Shared Disk (VSD) 264
9.1.5 Security 264
9.1.6 Performance Toolbox Parallel Extensions for AIX (PTPE) 265
9.1.7 Subsystem communications support 265
9.1.8 Network Time Protocol (NTP) 265
9.1.9 System availability 265
9.1.10 Other convenient services 266
9.1.11 New in PSSP 3.2 266
9.1.12 Software requirements 267
9.2 General Parallel File System (GPFS) 267
9.2.1 Administration and operation 268
9.2.2 Higher performance/scalability 268
9.2.3 Recoverability 268
9.2.4 Migration 269
9.2.5 New in GPFS 1.4 269
9.2.6 Software requirements 269
9.3 LoadLeveler 270
9.3.1 Administration 270
9.3.2 Capabilities 271
9.3.3 New in LoadLeveler 2.2 272
9.3.4 Software requirement 272
9.4 Mathematical Subroutine Libraries (PESSL) 272
9.4.1 Operations 274
9.4.2 New in ESSL 3.2 275
9.4.3 New in Parallel ESSL 2.2 275
9.4.4 New in OSLp 1.3 276
9.4.5 Software requirement 276
9.5 Parallel Environment (PE) 277
9.5.1 Operation 278
9.5.2 New in PE 3.1 278
9.5.3 Software requirements 280
9.6 IBM High Availability Cluster Multi-Processing for AIX (HACMP) . . . 281
9.6.1 Administration and operation 281
9.6.2 Features 282
9.6.3 New in HACMP 4.4 283
9.6.4 Software requirements 284

Chapter 10. Clustering configurations 285
10.1 Configurations rules 286
10.1.1 Placement rules for nodes, frames, and switches 287
10.1.2 Numbering rules for frames, slots, nodes, and switch ports 299
10.2 SP universal clustering capabilities 304
10.2.1 SP system in non-switched and switched environments 305

10.2.2 SP system with SP-attached servers 314
10.2.3 CES system in non-switched and switched environment 324
10.3 PSSP environment for SP/CES systems 337
10.4 Building SP/CES systems scenarios 344
10.4.1 Standard SP system 345
10.4.2 Adding SP-attached servers to an SP system 345
10.4.3 CES system 349
10.4.4 Migration an existing SP system to a CES system 350
10.4.5 Scaling up an existing CES system to an SP system 351
10.5 pSeries 680 Cluster Solutions 357
10.5.1 pSeries 680 Entry Cluster Server - S85/6H1 357
10.5.2 pSeries 680 Advanced Cluster Server - S85/M80 359

Appendix A. Naming convention 361
Definitions overview and naming convention 362
Naming convention 362

Appendix B. Hardware 365
SP Communication adapter requirements 366
PCI bus I/O adapter requirements for SMP nodes 366
MCA bus adapter requirements 381
Evaluating SP System power requirement 384
Legacy SP Nodes 389
SP Legacy Models by Frame and Switch Types 390
SP Legacy Models by Frame and Nodes Types 391
SP LAN topologies 392
SP Boot/Install server requirements 398
Disk Space required for PSSP file sets 404

Appendix C. SP system upgrades 407
Adding POWER3, and 332 MHz SMP Thin and Wide Nodes 408
Adding 375 MHz POWER3 and POWER3 SMP High Nodes 408
Adding SP Expansion I/O Units 408
Adding frames 408
Adding switches 409
Upgrading, replacing, and relocating nodes 409
Upgrading SMP-type nodes 410
Replacing withdrawn processor nodes 410
Relocating processor nodes 411
Upgrading SP frames 411

Appendix D. Software Appendix 413
PSSP and AIX 414
Supported levels 414

Migration. 415
Coexistence 416
High Availability Cluster Multi-Processing (HACMP) 417
Supported levels. 417
Migration considerations 418
Coexistence 418
VSD and RVSD. 419
Supported Levels 419
Migration considerations 420
Coexistence 420
General Parallel File System for AIX (GPFS) 421
Supported Levels 421
Coexistence 421
Parallel Environment (PE) 422
Migration considerations 422
ESSL and Parallel ESSL 423
LoadLeveler 423
Supported Levels 423
Migration considerations 424
Coexistence 424

Appendix E. SP/CES Tools 425
Network Installation Manager 426
Differences with SP/CES versus R/S6000 NIM 426
Service Director 432
Electronic Service Agent 434
What does Service Agent do. 435
How does Service Agent work 435
Electronic Service Agent in SP/CES environment 436
Obtaining Electronic Service Agent. 437
Inventory Scout (invscout) 437

Related publications 439
IBM Redbooks 439
Other resources 439
Referenced Web sites 439
How to get IBM Redbooks 440
IBM Redbooks collections. 440

Special notices 441

Abbreviations and acronyms 443

Index 445

Figures

1-1 RS/6000 SP system sample configuration. 10
1-2 RS/6000 SP system with SP-attached servers sample configuration . . 11
1-3 Clustered Enterprise Server System sample configuration 12
1-4 Load balancing cluster. 19
1-5 High availability cluster 19
1-6 High performance technical computing cluster 20
1-7 High performance commercial computing cluster 21
1-8 Common management domain 22
2-1 Mix of nodes in a frame 26
2-2 375 MHz POWER3 SMP architecture (High Node) 29
2-3 POWER3 High Node and five SP Expansion I/O Units connections . . . 31
2-4 375 MHz POWER3 High Node packaging. 34
2-5 375 Mhz POWER3 SMP Wide Node packaging 36
2-6 332 MHz SMP node system architecture block diagram 62
3-1 System architecture for 7017-S85 server 83
3-2 System architecture for 7026-M80 server 101
3-3 System architecture for 7026-6H1 server 106
3-4 SP-attached servers connectivity diagram. 128
3-5 Clustered Enterprise Servers connectivity diagram 137
4-1 SP frame model 550 141
4-2 Rear of the redundant power supply feature 144
4-3 Tall frame from front and rear views 145
4-4 9076 mod 555 with two SP-attached servers and a CWS 149
4-5 SP Switch frame front view 153
5-1 Front view of the SP Switch board. 160
5-2 Data flow between switch ports 161
5-3 SP Switch chip 162
5-4 Interconnections used on the switch board 163
5-5 logical diagram of an SP Switch adapter 164
5-6 SP Switch system 165
5-7 Single-stage switch cable paths. 172
5-8 Two-stage switch cable paths 173
5-9 SP Switch2 adapter hardware structure 179
6-1 IBM 9077 SP Switch Routers model 04S (left) and 16S (right) 182
6-2 SP Switch Router configuration 184
6-3 Rear view of the SP Switch Router model 04S 185
6-4 Rear view of SP Switch Router model 16S 186
6-5 SP Switch Router network interface 190

7-1 Control workstation interface 196
7-2 High Availability Control Workstation 222
10-1 Examples on node configuration 289
10-2 Example on frame configuration 290
10-3 Maximum short frame configurations with SP eight port Switch 292
10-4 Examples on SP Switch-8 Tall frame configurations 293
10-5 Examples on single SP Switch-16 configurations 294
10-6 Other examples on single SP Switch-16 configurations 295
10-7 Example on multiple SP-Switch-16 configurations 296
10-8 Example of an ISB configuration 297
10-9 Node numbering in an SP system with SP-attached server 301
10-10 Switch port numbering on an SP Switch 303
10-11 Switch port number assignment for SP-attached servers 304
10-12 Non-switched SP environment. 305
10-13 Switched SP environment 309
10-14 Switched SP environment with SP-attached servers. 315
10-15 Pure CES system environment 325
10-16 CES system in switched environment 331
10-18 Reconfigure the CES numbers to fit into an SP. 354
10-19 Changing a 15-node CES system to an SP system 355
10-21 pSeries 680 Entry Cluster Server 359
10-22 pSeries 680 Advanced Cluster Server. 360
B-1 Shared or switched 100BASE-TX network 393
B-2 Heterogeneous 10/100 Mbps network 394
B-3 Shared 10BASE2 network 395
B-4 Segmented 10BASE2 network 396
B-5 Switched 10BASE2 network 397
B-6 Ethernet topology with one adapter for a single-frame SP system . . . 399
B-7 Method 1 - Ethernet topology for multiple-frame SP system 400
B-8 Method 2 - Ethernet topology for multi-frame SP system 401
B-9 Method 3 - Ethernet topology for multi-frame SP system 402
B-10 Boot server frame approach 403

Tables

1-1 Servers supported as SP-attached servers 14
2-1 375 MHz POWER3 SMP High Node 32
2-2 375 MHz POWER3 SMP Thin and Wide Nodes 35
2-3 375 MHz POWER3 SMP High Node (F/C 2058) processor options . . . 38
2-4 375 MHz POWER3 SMP High Node (F/C 2058) memory features 39
2-5 375 MHz POWER3 SMP Wide Nodes processor option 44
2-6 375 MHz POWER3 SMP Wide Nodes memory features. 45
2-7 375 MHz POWER3 SMP Thin Nodes (F/C 2056) Processor option . . . 49
2-8 375 MHz POWER3 SMP Thin Nodes (F/C 2056) memory features . . . 50
2-9 POWER3 SMP High Node (F/C 2054) processor options 57
2-10 POWER3 SMP High Node (F/C 2054) memory features 57
2-11 332 MHz Wide Node processor options 69
2-12 32 MHz SMP Wide Node memory features 70
2-13 332 MHz SMP Thin Node processor option. 74
2-14 332 MHz SMP Thin Node memory features 74
3-1 Servers supported as SP-attached servers 78
3-2 Publications for the 7017-S85 server. 84
3-3 Publications for the 7017-S80 server. 88
3-4 Publications for the 7017-S7A server 92
3-5 Publications for the 7017-S70 server. 95
3-6 SP-attached servers / CES features - M/T 7017 95
3-7 Publications for the 7026-M80 server 103
3-8 Publications for the 7026-H80 server. 114
3-9 Features of the SP-attached servers and CES - M/T 7026 114
3-10 F/C 3154 characteristics 117
4-1 SP frame dimensions. 142
4-2 SP frame electrical power requirements 143
4-3 Short frame front view 154
5-1 AIX/PSSP/switches level matrix. 166
5-2 SP Switch adapter features 168
5-3 Quantities of Switch-to-Switch cables required per SP system 173
6-1 SP Switch Router network media cards and other options 188
7-1 Supported pSeries and RS/6000 workstations for SP/CES systems. . 198
7-2 Serial port adapters for PCI / MCA control workstations 202
7-3 Ethernet adapters for PCI/MCA control workstations 203
8-1 Supported PCI adapters 228
8-2 Cable information for 2-port multi protocol adapter 242
9-1 PSSP 3.2 software requirements. 267

9-2 Software requirements for GPFS 1.4 . . . 269
9-3 Software requirements for LoadLeveler 2.2 . . . 272
9-4 Requirements for ESSL 3.2, Parallel ESSL 2.2 and OSLp 1.3 . . . 276
9-5 Software requirements for PE 3.1 . . . 280
9-6 Software requirements for HACMP 4.4 . . . 284
10-1 Switch current compatibility with different types of nodes . . . 299
10-2 PSSP spframe command for an SPCES system . . . 338
10-3 PSSP spdelframe command for an SP/CES system . . . 340
10-4 PSSP Event Perspective - Event Definitions for CES . . . 342
10-5 Node number and IP address . . . 352
A-1 SP/CES logical names . . . 362
A-2 Machine names, M/T and Nicknames . . . 362
B-1 PCI bus description for SMP nodes . . . 367
B-2 Maximum quantities of PCI adapters for SMP thin and wide nodes . . 368
B-3 PCI adapters for high nodes and SP Expansion I/O Units . . . 370
B-4 PCI adapters weighting factors . . . 376
B-5 Micro Channel Adapter requirements . . . 381
B-6 SP system frame power computation worksheet . . . 385
B-7 Legacy SP nodes . . . 389
B-8 Legacy models by frames and switch types (table 1 of 2). . . . 390
B-9 Legacy models by frames and node types (table 2 of 2) . . . 391
10-6 Ethernet topology with two adapters for single-frame SP system 399
B-10 Disk space required for install images and to install PSSP file sets . . 404
C-1 Node upgrade features . . . 410
D-1 IBM LPPs per supported PSSP and AIX release . . . 414
D-2 AIX and PSSP support for different node types . . . 415
D-3 Supported migration paths for AIX and PSSP . . . 415
D-4 Levels of PSSP and AIX supported in a mixed system partition 416
D-5 HACMP supported levels . . . 417
D-6 VSD and RVSD supported levels . . . 419
D-7 Supported levels in relationship to AIX and PSSP supported levels . . 421
D-8 Supported PE LPP levels . . . 422
D-9 Available PE migration paths . . . 422
D-10 Supported levels in relationship to AIX and PSSP supported levels . . 423

Preface

This redbook is an update of the RS/6000 SP Systems Handbook published in 2000. This publication includes Blue Hammer and Blue Hammer Jr. clustering solution configurations.

This publication includes the following:

- Information on M80, H80, 6H1 SP-attached support as well as clustering solutions capabilities with PSSP 3.2.
- Overview on clustered @server pSeries systems, configurations, and hardware support.
- Updated information on the available software for RS/6000 SP and clustered systems (PSSP, GPFS, LoadLeveler, PESSL, PE and HACMP).
- Updated information on RS/6000 SP hardware support.
- Updated information on the supported control workstations for RS/6000 SP and clustered @server pSeries systems.
- Updated information on the supported communication adapters for the RS/6000 SP and clusted @server pSeries systems.
- Updated information on the SP Switch, SP Switch2 and the SP Switch Router.
- Updated information on available RS/6000 clustered hardware (machine types, models, and feature codes).

The team that wrote this redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization, Austin Center.

Dino Quintero is a project leader at the International Technical Support Organization (ITSO), Poughkeepsie Center. He has over nine years experience in the Information Technology field. He holds a BS in Computer Science and a MS degree in Computer Science from Marist College. Before joining the ITSO, he worked as a Performance Analyst for the Enterprise Systems Group and as a Disaster Recovery Architect for IBM Global Services. He has been with IBM

since 1996. His areas of expertise include enterprise backup and recovery, disaster recovery planning and implementation, and RS/6000. He is also a Microsoft Certified Systems Engineer. Currently, he focuses on RS/6000 Cluster Technology by writing redbooks and teaching IBM classes worldwide.

Christian Courty is a IBM IGS RS/6000 SP specialist in Montpellier (France). He joined IBM in 1970. He has been Mainframe Hardware CE and later MVS Software CE. He became a CATIA specialist and follows the migration of CATIA from mainframe to the workstation and AIX. He was involved on the first SP1 installations in Paris. He moved to Montpellier in 1994 to support the SP2 for EMEA in the SP support group.

Corneliu Holban is an IBM Advisory IT specialist in New York. He has 8 years of experience in system engineering, technical support, and sales on RS/6000 and RS/6000 SP systems. His areas of expertise include RS/6000 and RS/6000 SP – solutions sizing, capacity planning, high availability, and system performances. He has IBM Certification for RS/6000 Solutions Sales and IBM Qualification for RS/6000 SP Solutions Sales.

Willis C. White III is a software engineer in the Functional Verification Test department of the AIX development LAB in Poughkeepsie NY. A Poughkeepsie native with roots stretching back 100 years in the area, he is a third generation IBMer. He has two years of experience in the Information Technology field. He holds a BS in Electrical Engineering from The Rochester Institute of Technology and joined the cooperation right out of college. His areas of expertise include SP system administration, UNIX testing theory and problem determination, and resolution on PSSP systems, along with a host of other specialties relating to hardware and hardware engineering.

Thanks to the following people for their contribution to this project:

IBM Poughkeepsie
Brian Croswell, Margaret M. Cronin, Chester A. Peters, Amy M. Brown, Chris Algozzine, Bruno Bonetti, David Ayd, Hal Turner, David Delia, David Wong, Bob Leddy, Mike Coffey, Norman Nott, Pat Caffrey, Sarah Wong, Shujun Zhou, Joan McComb, Waiman Chan, Brian Herr

International Technical Support Organization, Poughkeepsie Center
Ella Buslovich, Al Schwab

International Technical Support Organization, Austin Center
Matthew Parente

Special notice

This publication is intended to help IBM customers, IBM business partners, IBM sales professionals, IBM I/T specialists, and the IBM technical support community when proposing RS/6000 cluster base solutions. The information in this publication is not intended as the specification of any programming interfaces that are provided by RS/6000 hardware, AIX software, or PSSP software. See the PUBLICATIONS section of the IBM Programming Announcement for RS/6000 for more information about what publications are considered to be product documentation.

IBM trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

AIX
ClusterProven
DirectTalk
e (logo)® 
HACMP/6000
LoadLeveler
Lotus Notes
MVS/ESA
NetView
OS/390
PowerPC Architecture
Portmaster
Redbooks Logo 
RS/6000
SP
SP2
System/390
VM/ESA
Wave
Chipkill
DB2
Domino
ESCON
IBM ®
Lotus
Micro Channel
Netfinity
Notes
PC 500
PowerPC 604
pSeries
Redbooks
Service Director
SP1
S/390
TURBOWAYS
VisualAge

Comments welcome

Your comments are important to us!

We want our IBM Redbooks to be as helpful as possible. Send us your comments about this or other Redbooks in one of the following ways:

- Use the online **Contact us** review redbook form found at:

ibm.com/redbooks

- Send your comments in an Internet note to:

  redbook@us.ibm.com

- Mail your comments to the address on page ii.

1

Introduction

The IBM RS/6000 SP high-performance system uses the power of parallel processing to expand your applications.

Designed and extended as a Clustered @server pSeries system for performance, scalability, reliability, availability, serviceability, and management, this system makes feasible the processing of applications characterized by large scale data handling and compute intensive applications.

The RS/6000 SP system is IBM's family of parallel computing solutions. It provides a state-of-the-art parallel computing system and industry-leading application enablers.

1.1 Overview on RS/6000 SP and Clustered @server pSeries

The RS/6000 SP system simultaneously brings dozens to hundreds of RISC processor nodes to a computing problem. Its parallel processing capability enhances computing performance and throughput many times in comparison to serial computing. In addition to helping improve the performance of existing applications, new applications, like complex data mining and modeling of the universe, are now possible.

The basic SP building block is the processor node. It consists of a POWER3 or PowerPC Symmetric Multiprocessors (SMP), memory, Peripheral Component Interconnect (PCI) expansion slots for Input/Output (I/O) and connectivity, and disk devices. The three types of nodes (thin, wide, and high) may be mixed in a system and are housed in a short (1.25 m) or tall (1.93 m) system frames. Depending on the type of nodes used, an SP tall frame can contain up to 16 nodes and an SP short frame can contain up to 8 nodes. These frames can be interconnected to form a system with up to 128 nodes (512 by special order). Each node contains its own copy of the AIX operating system.

The 222 MHz/375 MHz POWER3 SMP nodes, are powered by the same POWER3 processor technology introduced with the RS/6000 43P model 260/270 workstation. Hence, all applications that run on the 43P model 260/270 should run unchanged on a single node of the SP system. For example, Environmental Systems Research Institute, without having tested their applications specifically on the POWER3 SMP node, has stated that their applications set ARC/INFO, ArcView, SDE, and IMS works fine on this SP node.

The 332 MHz SMP nodes are powered by the PowerPC 604e processor. They represent the first general availability of the advanced technology used in the SP system IBM has delivered to Lawrence Livermore National Laboratory as part of the Department of Energy's Accelerated Strategic Computing Initiative (ASCI) project. This system (at the time of this writing, already upgraded to POWER3 nodes), performs the complex calculations required for the simulation to predict the performance, safety, reliability, and manufacturability of the U.S. nuclear stockpile.

The system is managed by the AIX operating system and the Parallel Systems Support Programs for AIX (PSSP) management software, from a single point of control, a pSeries or an RS/6000 system, called Control Workstation (CWS).

Effective parallel computing requires high-bandwidth, low-latency internode communications. The SP Switch, recommended for high-performance commercial computing, provides a one-way bandwidth of up to 150 MB per second between nodes (300 MB bidirectional) for interconnecting all node types. The SP Switch2, recommended for high-performance technical computing, provides a one-way bandwidth of up to 500 MB per second between nodes (1 GB bidirectional) for interconnecting POWER3 high nodes.

The SP Switch Router is a high-performance I/O gateway that provides the fastest available means of communication between the SP system and the outside world or among multiple SP systems. This SP gateway combines the Lucent GRF with the IBM SP Switch Router adapter to enable direct network attachment to the SP Switch. Other media cards connect to a variety of standard external networks. Each media card has its own hardware engine, enabling SP I/O to scale nearly one-to-one with the number of cards.

The SP system can also scale disk I/O nearly one-to-one with processors and memory making access to terabytes of data possible and expansions or upgrades easier to manage. If you outgrow your existing system, you can readily add increments of computing power.

The RS/6000 SP system, as a cluster of SP processor nodes, managed by the AIX operating system and PSSP management software from the CWS, is the first step to universal clustering.

Managed by the AIX and PSSP from the CWS, @server pSeries and RS/6000 Enterprise Servers, can function as SP-attached servers on an SP system, using the connectivity provided by the SP Switch or the LAN. The exceptional performance of these servers is especially impressive for online transaction processing applications. With their large, single node data capacity, they are also well-suited to the tasks associated with Enterprise Resource Planning (ERP). Plus, their excellent query capability is ideal for today's business intelligence applications. These characteristics combined make the SP-attached servers an excellent third-tier choice for data storing and for a three-tier e-business environments where POWER3 SMP nodes or 332 MHz SMP nodes are used as the middle tier.

Managed by the AIX and PSSP from the CWS, @server pSeries and RS/6000 Enterprise Servers, can be defined as Clustered Enterprise Servers (CES) on a Clustered @server pSeries system (CES system). The CES system contains pSeries servers and RS/6000 servers only. The connectivity is provided by the SP Switch or the LAN. Mission-critical enterprise applications such as ERP, Web serving, Supply Chain Management (SCM), Customer Relationship Management (CRM), e-business, Online Transaction Processing (OLTP), and business-intelligence may be managed by a CES system.

The SP universal cluster system (SP/CES system) delivers balanced performance with processor, memory, switch, and I/O scalability. Over time, the SP system has demonstrated leadership in standard industry benchmarks. An SP e-business server is recognized by the 1998 Guiness Book of Records for an Internet volume of 110,414 hits in one minute recorded at the Nagano Winter Olympic Games. The outstanding performance of the SP system is shown in the TOP500 supercomputer list at:

`http://www.top500.org`

1.2 Features and benefits

The SP/CES system is IBM's family of parallel computing solutions. The scalable architecture of the SP system, its high-performance communications, POWER3 and PowerPC processors give it the power to handle data-intensive, compute-intensive, and I/O-intensive jobs with ease.

You can execute both serial and parallel applications simultaneously, while managing your system from a single workstation. For scientific and technical applications, the SP/CES system delivers the power and connectivity for rapid turnaround; from structural analysis and seismic modeling to circuit simulation and molecular modeling. Multiple users can run complex queries against very large amounts of data and obtain results interactively. This makes the SP/CES system an ideal solution for database query, online transaction processing, business management, and batch processing applications.

The IBM software offerings for the SP/CES system provide an interlocking set of licensed programs designed to address a wide range of system and application needs. The open architecture, based on the AIX operating system (the IBM implementation of UNIX), is enabled to integrate the SP/CES system into existing environments. The software architecture is closely tuned to the SP/CES system hardware design for maximum usability.

The performance, scalability, reliability, availability, serviceability and management represent the main advantages offered by the SP/CES systems.

The SP system is a general-purpose scalable power parallel system based on share-nothing architecture. Generally available SP systems range from 1 to 128 processor nodes, and each processor node is functionally equivalent to a stand-alone RS/6000 workstation or server. Large SP systems with up to 512 nodes (special order) have been delivered and are successfully being used today.

Each processor node contains its own copy of the standard AIX operating system and other standard RS/6000 system software, providing access to thousands of available AIX applications. A set of new software products designed specifically for the SP (PSSP with Virtual Shared Disks (VSD), Recoverable Virtual Shared Disks (RVSD), Concurrent Virtual Shared Disks (CVSD), General Parallel File Systems (GPFS), and Parallel Environment (PE)) allows the parallel capabilities of the SP system to be effectively exploited.

The SP processor nodes, on SP Switch or SP Switch-less configurations, can be partitioned into pools of nodes. For example, four nodes can work as a Lotus Notes server, while eight others process a parallel database. The SP system partitioning is designed to divide the system into non-overlapping sets of nodes in order to make the system more efficient and more tailored to your needs.

On an SP/CES processor node, the resources may also be partitioned using the AIX WorkLoad Manager (WLM) facilities. For example a part of the processor and physical memory resources can be dedicated for a specific class of jobs - for example, interactive applications, and another part can be dedicated for other specific class of jobs - for example, batch applications. WLM is designed to give system administrators greater control over how the scheduler and Virtual Memory Manager (VMM) allocate CPU and physical memory resources to processes. This can be used to prevent different classes of jobs from interfering with each other and to allocate resources based on the requirements of different groups of users.

The RS/6000 SP system takes the advantages of the industrial leading RS/6000 workstation technology, the latest RS/6000 processors are repackaged for use as SP nodes, and the SP nodes are interconnected by a high-performance, multistage, packet-switched network for inter processor communication in order to perform the parallel functions.

The SP/CES system optimizes high availability through built-in redundancy, subsystem recovery, component error checking and correction, RAID5, external and internal disk mirroring, and hardware and software monitoring. High availability clusters of up to 32 SP nodes are supported by one of the industry's leading software products for critical application backup and availability, High Availability Cluster Multi-Processing (HACMP) for AIX. If an error such as a node failure occurs, the system can execute a recovery script that transfers the work to another node and prevents the application from going down. RS/6000 cluster technology, implemented also on the SP/CES systems, provides a collection of services (topology, group, event management) that define hardware and software resources, node relationships and coordinated actions to manage groups of nodes.

The SP/CES processor nodes incorporate a dynamic CPU de-allocation function, which enables the system to detect failing processors and take them offline without re-booting the system. The system service processor can record this action and notify the system administrator or service personnel of the condition. Computing operation continues with the processor deactivated, allowing repair to be scheduled at a time convenient to your processing schedule. Functions such as power-on and power-off, reading the service processor error logs and POST error logs, reading vital product data (VPD), and accessing NVRAM can be performed remotely.

Managing large systems is always a complex process. For the SP/CES system, a single graphical control workstation (CWS) that displays hardware, software, jobs, and user status makes system management easier. The system administrator uses the CWS and the PSSP software product (it is delivered with the SP standard nodes or it has to be ordered for the SP-attached servers and the Clustered Enterprise Servers) to perform management tasks including user and password management, job accounting (as well as system startup/shutdown), monitoring, and partitioning.

In addition, the SP/CES system offers a wide range of open system management software tools for operation and administration, availability, deployment, and security management. Included are the Tivoli Suite software for system management, Tivoli Storage Manager for backup and recovery, LoadLeveler for job scheduling and workload balancing, Performance Toolbox (PTX) for performance monitoring, Distributed Computing Environment (DCE) and Kerberos for security.

As a general-purpose parallel computer, the SP/CES systems are used productively in a wide range of application areas and environments in the high-end UNIX technical and commercial computing market. This broad-based success is attributable to the highly flexible and general-purpose nature of the system. RS/6000 SP system can be used for doing a variety of tasks including parallel, serial, batch, and interactive jobs.

The SP system is a member of the RS/6000 product family, and it is positioned in the top end of entire pSeries and RS/6000 product line. The following are the features and benefits of the SP system:

- Single point of administrative control

  Makes system management easier with less expertise and time required for most tasks.

- Comprehensive system management

  Provides the tools required to install, operate, and control an SP system and helps maintains consistency with other enterprise AIX systems management tools.

- Scalability

  Makes upgrading and expansion easier and allows for transparent application growth.

- AIX operating system

  Provides a wealth of multiuser communications and systems management technologies and complies with major industry standards. Additionally, it provides AIX binary compatibility, where most AIX 4 applications already running on other RS/6000 systems can run unmodified.

- Open system design

  Supports many communication protocols, adapters, and peripherals for a flexible system.

- Configuration flexibility

  Provides various node types that can be intermixed on the system and supports various PCI adapters.

- POWER family processor

  Delivers the processing power required for large, complex applications. It also allows the flexibility to configure the system for optimum commercial or technical computing application performance.

- Multistage packet switch

  Supports high-performance communications between processor nodes and maintains point-to-point communication time independent of node location. SP Switch Router provides fastest available communication between SP systems and external networks.

- System partitioning

  Isolates application subsystems and enables concurrent use of production and test AIX systems.

- High availability

  Helps avoid costly downtime due to system outages and provides an optional backup control workstation.

1.3 Business solutions

Installed in almost ten thousands customer locations worldwide, the SP/CES system delivers solutions for some of the most complex, large commercial and technical computing problems.

Commercial computing users implement the SP/CES mission-critical commercial computing solutions to address business intelligence applications, server consolidation, and collaborative computing comprised of Lotus Notes, Lotus Domino Server, internet, intranet, extranet, and groupware application.

Technical computing users, including corporations, universities, and research laboratories, use the SP/CES system for leading-edge applications, such as seismic analysis, computational fluid dynamics, engineering analysis, and computational chemistry.

The SP/CES solutions can be categorized into the following areas:

- Business intelligence

  Provides scalable database capacity with support of leading parallel databases including IBM DB2 UDB EEE, Oracle Parallel Server, and Informix Dynamic Server AD/XP. This also offers proven scalable performance with leadership TPC-H results and delivers mainframe inter operability for optimal data movement.

- e-business

  Scalable growth and a single management console virtually eliminate server proliferation issues associated with the addition of new servers to support the increasing number of Internet services and the complex dynamic workloads characteristic of network computing.

  Additionally, flexible node partitioning options permit multiple logical computing tiers for web business; logic and database servers are supported in a single physical system while system investment is preserved.

- Enterprise Resource Planning (ERP), Supply Chain Management (SCM), Customer Relationship Management (CRM)

  This SP/CES solution consolidates applications among multiple nodes within a single SP system, allowing ERP and supply chain planning applications from multiple vendors to take on a single-system appearance.

  Additionally, it provides LAN consolidation, allowing multiple systems (in a two- or three-tier client/server environment) to be managed as a single system. Also, these provide high availability computing, using the IBM industry-leading HACMP, to provide back up, recovery, and fault-tolerant computing for mission-critical applications.

- Technical computing

  This solution supports batch, interactive, serial, and parallel processing while providing outstanding floating-point performance. The SP/CES solutions lead the way by supporting industry initiatives, such as PVM, LAPI, MPI, and HPF.

- Server consolidation

Helps reduce the complexities and costs of systems management, lowering the total cost of ownership and allowing simplification of application service level management. Additionally, this solution leverages investment in hardware and software, allowing better sharing of resources and licenses and distributing idle cycles instead of hot spots. Also, it provides the infrastructure that supports improved availability, data sharing, and response time.

1.4 Hardware components

The basic hardware components of the SP/CES system are as follows:

- Processor nodes (include SP nodes, SP-attached servers, Clustered Enterprise Servers)
- Frames with integral power subsystem
- Extension nodes (include SP Switch Routers)
- SP switches
- Control Workstation (CWS)

These components connect to your existing computer network through a local area network (LAN) making the RS/6000 SP system accessible from any network-attached workstation.

The Machine Type (M/T) for the RS/6000 SP system is 9076.

Figure 1-1 on page 10, Figure 1-2 on page 11, and Figure 1-3 on page 12 illustrate sample configurations of RS/6000 SP system universal clustering: RS/6000 SP system, RS/6000 SP system with SP-attached servers, and CES system respectively. These figures provide you a rough idea on how the hardware components are connected.

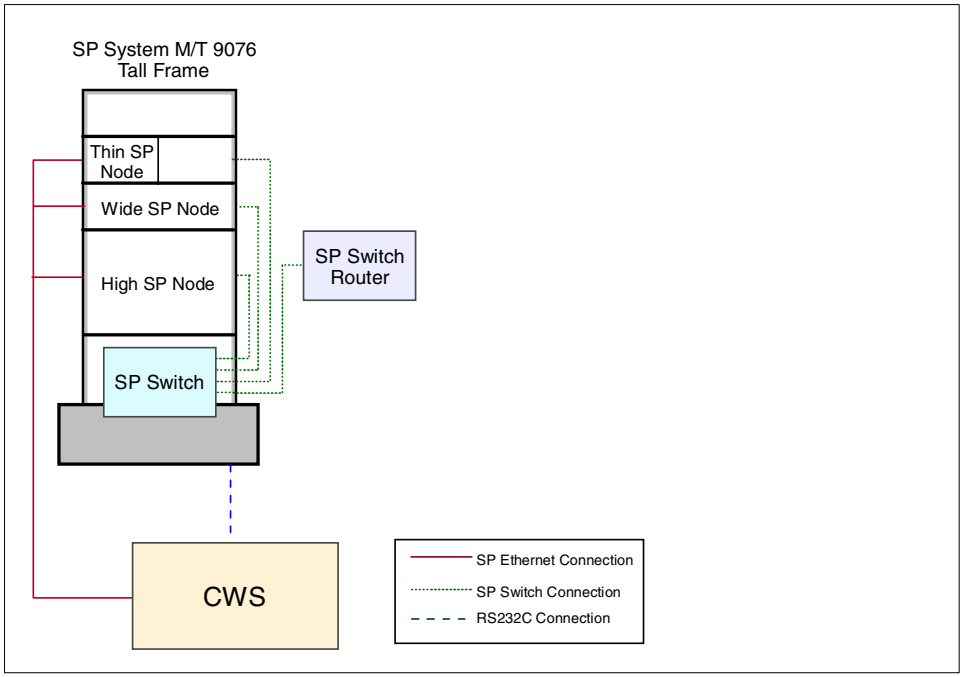


Figure 1-1 RS/6000 SP system sample configuration

As shown in Figure 1-1:

- The thin, wide, and high SP nodes, and the SP Switch are mounted in a tall frame.
- The SP nodes and SP Switch Router are connected to the SP Switch.
- The SP nodes and the SP Switch Router are connected to the SP Ethernet interface of the CWS.
- The frame with the SP nodes and the SP Switch are connected to the RS-232 interface of the CWS.

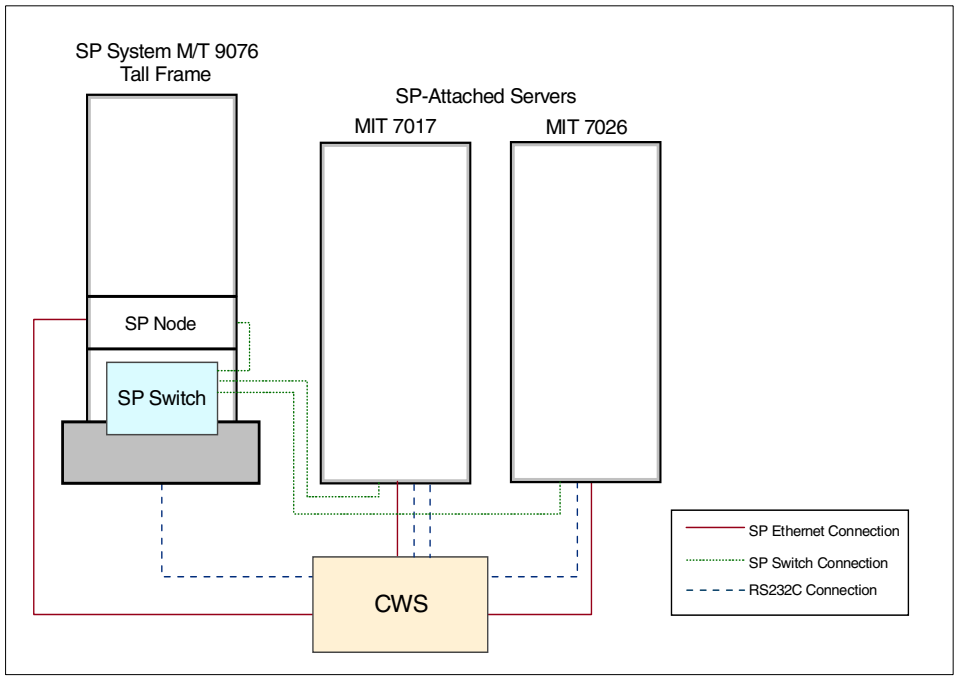


Figure 1-2 RS/6000 SP system with SP-attached servers sample configuration

Figure 1-2 shows an RS/6000 SP system with SP-attached servers sample configuration:

- The SP node and the SP Switch are mounted in a tall frame.
- The SP-attached servers (M/T 7017 and M/T 7026) are housed in their own frames.
- The SP node and the SP-attached servers are connected to the SP Switch.
- The SP node and the SP-attached servers are connected to the SP Ethernet interface of the CWS.
- The frame with the SP node and the SP Switch, and the SP-attached servers are connected to the RS-232 interface of the CWS.

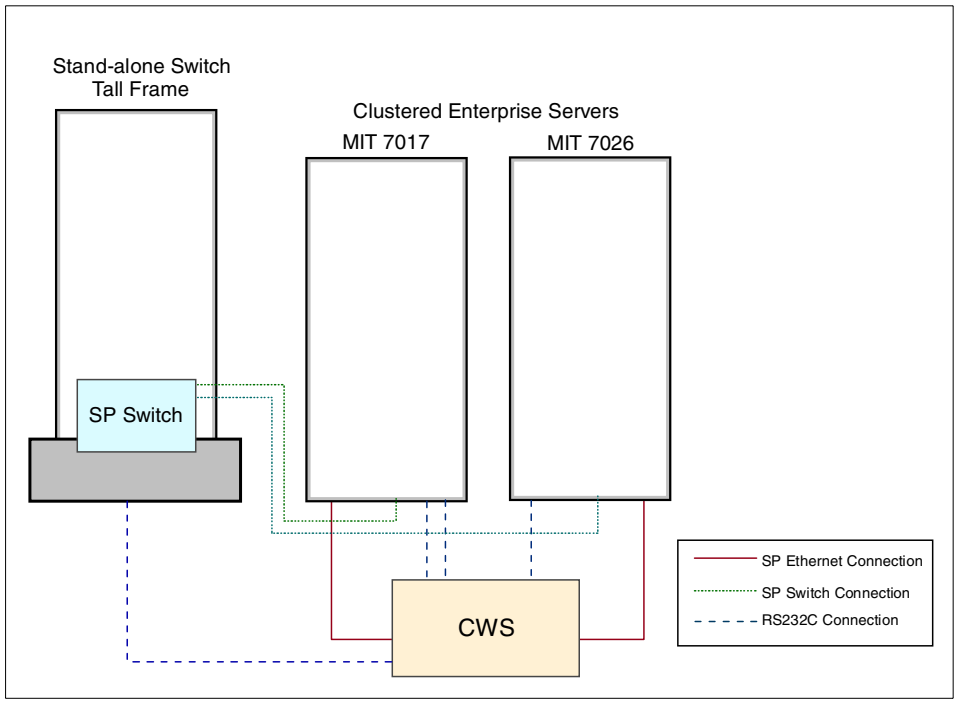


Figure 1-3 Clustered Enterprise Server System sample configuration

Figure 1-3 shows a pSeries CES system sample configuration:

- The Clustered Enterprise Servers (M/T 7017 and M/T 7026) are housed in their own frames.
- The SP Switch is mounted in the dedicated tall frame 9076-555 Stand-alone Switch for clustered nodes.
- The servers are connected to the SP Switch. The servers are connected to the SP Ethernet interface and to the RS-232 interface of the CWS.

1.4.1 Processor nodes

The IBM RS/6000 SP system is scalable from 1 to 128 processor nodes, which can be contained in multiple SP frames. Up to 16 processor nodes (SP nodes) can be mounted in a tall frame, while a short frame will hold up to eight processor nodes (SP nodes). SP systems consisting of more than 128 processor nodes are available (on a special order basis).

There are four types of RS/6000 SP processor nodes:

- High Nodes
- Wide Nodes
- Thin Nodes
- SP-attached servers and Clustered Enterprise Servers

All nodes are available in an SMP configuration and they use PCI architecture.

High nodes

The high nodes occupy two full drawer locations. These nodes require the 1.93 m tall frame, thus four nodes can be housed on this frame. The nodes can be placed in the first node slot of a frame without requiring additional nodes. The maximum number of high nodes (POWER3) supported in a 128-frame system is 128. The high nodes provide additional hard disk drives and adapter capacity by connecting to SP Expansion I/O units. These nodes are compatible with the SP Switch and the SP Switch2, but they are not compatible with the SP Switch-8 and the High Performance Switch.

The following high nodes are supported:

- 375 MHz POWER3 SMP nodes (Feature Code (F/C) 2058)
- 222 MHz POWER3 SMP nodes (F/C 2054)

The following high nodes are currently available:

- 375 MHz POWER3 SMP nodes (F/C 2058)

Wide nodes

The wide nodes occupy one full drawer, thus eight nodes can be housed in a tall frame. These nodes can be placed in the first node slot of a frame without requiring additional nodes. The wide nodes are compatible with the SP Switch and the SP Switch-8, but they are not compatible with the High Performance Switch.

The following wide nodes are supported:

- 375 MHz POWER3 SMP nodes (F/C 2057)
- 332 MHz SMP nodes (F/C 2051)

The following wide nodes are currently available:

- 375 MHz POWER3 SMP nodes (F/C 2057)

Thin nodes

The thin nodes occupy half of a drawer (one slot). Up to sixteen of these nodes can be housed in a tall frame. When installed singly within a drawer, these nodes must be placed in an odd-numbered node slot. The wide nodes are compatible with the SP Switch and the SP Switch-8, but they are not compatible with the High Performance Switch.

The following thin nodes are supported:

- 375 MHz POWER3 SMP nodes (F/C 2056)
- 332 MHz SMP nodes (F/C 2050)

The following thin nodes are currently available:

- 375 MHz POWER3 SMP nodes (F/C 2056)

SP-attached servers and Clustered Enterprise Servers

The SP-attached server is an @server pSeries or an RS/6000 Enterprise Server, supported to be configured to operate with an RS/6000 SP system.

The Table 1-1 shows the list with the @server pSeries and RS/6000 Enterprise Servers supported as an SP-attached server.

Table 1-1 Servers supported as SP-attached servers

Server name	Machine Type (M/T)	Model	Availability
@server pSeries 680 Model S85	7017	S85	Yes
RS/6000 Enterprise Server 7017 Model S80	7017	S80	Yes
RS/6000 Enterprise Server 7017 Model S7A	7017	S7A	No
RS/6000 Enterprise Server 7017 Model S70	7017	S70	No
RS/6000 Enterprise Server 7026 Model M80	7026	M80	Yes
@server pSeries 660 Model 6H1	7026	6H1	Yes
@server pSeries 660 Model 6H0	7026	6H0	Yes
RS/6000 Enterprise Server 7026 Model H80	7026	H80	Yes

Note: Up to 16 of the supported M/T 7017 Models and up to 32 of the supported M/T 7026 Models, up to 32 mixed servers in total, can be integrated on an SP system as SP-attached servers.

The Clustered Enterprise Server (CES) is an @server pSeries or an RS/6000 Enterprise Server, supported to be configured to operate with a Clustered @server pSeries system (CES system).

All the possible SP-attached servers to an SP system, can be defined as Clustered Enterprise Servers. The CES system contain pSeries systems and RS/6000 systems only, connected to the CWS.

Note: Up to 16 of the supported M/T 7017 Models and up to 32 of the supported M/T 7026 Models, and up to 32 mixed servers in total, can be integrated on a CES system.

Similar to an SP node, the SP-attached server and the Clustered Enterprise Server can perform most SP processing and administration functions. However, unlike a standard SP node, the SP-attached server and the Clustered Enterprise Server are housed in their own frame. Thus, the SP-attached server and the Clustered Enterprise Server have both node-like and frame-like characteristics.

Attention: These servers are compatible with the SP Switch.

For more information, refer to Chapter 2, “SP internal processor nodes” on page 25 and Chapter 3, “SP-attached servers and Clustered Enterprise Servers” on page 77.

1.4.2 Frames

The IBM RS/6000 SP system frames contain and provide power for processor nodes, switches, hard disk drives and other hardware. A frame feature code provides an empty frame with its integral power subsystem and AC power cable. The nodes and the other components have different specific feature codes.

The frames are offered in a list of five options:

- Tall (1.93 m) model frames (Model 550, Model 555)
- Tall expansion frames (F/C 1550)
- Short (1.25 m) model frames (Model 500)
- Short expansion frames (F/C 1500)
- SP switch frames (F/C 2031 for SP Switch and F/C 2032 for SP Switch2)

Frames have locations known as drawers into which the processor nodes are mounted. Tall frames have eight drawers and short frames have four. Each drawer location is further divided into two slots. A slot has the capacity of one thin node or SP Expansion I/O unit. A wide node occupies one full drawer, while a high node occupies two full drawers. The maximum number of SP frames supported in an SP system is 128.

Note: The Model 555 is a tall frame with an integral SEPBU power supply, and SP Switch (F/C 4011), but no processor nodes. This frame provides node interconnectivity to CES systems having configurations of 2 to 16 servers. You must add an SP Expansion Frame (F/C 1550) containing an SP Switch for configurations of 17 to 32 servers.

For more information, refer to Chapter 4, “Frames” on page 139.

1.4.3 Extension nodes

An extension node is a non-standard node that extends the SP system’s capabilities but that cannot be used in the same manner as a standard node. Note that SP systems with extension nodes require PSSP Version 2.3 or later.

One type of extension node is a dependent node. A dependent node depends on SP nodes for certain functions but implements much of the switch-related protocol that standard nodes use on the SP Switch. Typically, dependent nodes consist of four major components as follows:

1. A physical dependent node: The hardware device requiring SP processor node support.
2. A dependent node adapter: A communication card mounted in the physical dependent node. This card provides a mechanical interface for the cable connecting the physical dependent node to the SP system.
3. A logical dependent node: It is made up of a valid, unused node slot and the corresponding unused SP Switch port. The physical dependent node logically occupies the empty node slot by using the corresponding SP Switch port. The switch port provides a mechanical interface for the cable connecting the SP system to the physical dependent node.
4. A cable: It connects the dependent node adapter with the logical dependent node. It connects the extension node to the SP system.

A specific type of dependent node is the SP Switch Router (9077-04S and 9077-16S models). It is a licensed version of the Lucent GRF switched IP router that has been enhanced for direct connection to the SP Switch. The SP Switch router uses networking cards (Ethernet, FDDI, ATM, Sonet, HIPPI, HSSI) that fit

into slots in the SP Switch Router. In the same way that the SP Switch Router Adapter connects the SP Switch Router directly to the SP Switch, these networking cards enable the SP Switch Router to directly connect to an external network. Network connections through SP Switch Routers are typically faster and have better availability than network connections through the SP system nodes.

For more information, refer to Chapter 5, “The Switch” on page 159.

1.4.4 Switches

The SP switches provide a message-passing network that connects all processor nodes with a minimum of four paths between any pair of nodes. The SP series of switches can also be used to connect the SP system with optional external devices. A switch feature code provides you with a switch assembly and the cables to support node connections. The number of cables you receive depends on the type of switch you order.

The SP switches are available as either the 16-port SP Switch or the 8-port SP Switch-8:

- SP Switch2 (F/C 4012), 16-port switch (in SP systems exclusively with POWER3 high nodes only)
- SP Switch (F/C 4011), 16-port switch
- SP Switch-8 (F/C 4008), 8-port switch

For more information, refer to Chapter 5, “The Switch” on page 159.

1.4.5 Control workstations

You can view the SP/CES control workstation, as a server to the SP/CES system applications. The subsystems running on the CWS are the SP/CES server applications for the SP/CES nodes. The nodes are clients of the CWS server applications. The CWS server applications provide configuration data, security, hardware monitoring, diagnostics, a single point of control service, and optionally, job scheduling data and a time source. The CWS is an pSeries or an RS/6000 system. The available pSeries and RS/6000 systems which are supported as CWS are as follows:

- @server pSeries 660 Model 6H1 (7026-6H1)
- @server pSeries 620 Model 6F1 (7025-6F1)
- RS/6000 Enterprise Server Model H80 (7026-H80)
- RS/6000 Enterprise Server Model F80 (7026-F50)

- RS/6000 44P Model 170 (7044-170)

A High Availability Control Workstation (HACWS) option is also available on configurations with SP standard frames only.

For more information, refer to Chapter 7, “Control workstations” on page 195.

1.5 RS/6000 SP and AIX clustering

The new e-business environment creates the need for a computing infrastructure comprised of servers that optimize three classes of workloads: Data transaction servers to manage core business processes; web application servers to manage end-user experience; and appliance servers to manage specific network functions.

The environment must provide essential computing attributes that enable seamless integration and management of end-to-end web infrastructure. Computing resources in this diverse computing environment must be managed to effectively balance complex workloads as well as deliver the required Quality of Service (QoS), Capacity on Demand (CoD), security, scalability and manageability. This aggregation or clustered computing, include servers, operating systems, storage, networks, applications, racks and stacks.

The emergence of Web serving and e-business as mainstream application workloads has dramatically increased the value and relevance of clustered computing. Clustered computing now includes “tight clusters” (as RS/6000 SP) for high availability and single application scalability, “loose clusters” (as dense computing elements in racks) for appliances and web infrastructure, and “grid clusters” (heterogeneous, distributed collection of computing resources) for e-utilities and terascale applications. These diverse cluster solutions must deliver the needed attributes of QoS, CoD, security, scalability, and manageability.

On the IBM AIX operating system environment, we may consider the following different cluster computing models and their main characteristics:

- Sequential and Distributed Applications (little or no application alteration is required)
  - Load Balancing Cluster
    - Scale number of users or transactions throughput through application replication
    - Allocation of incoming transactions based on load-balancing algorithms

      Figure 1-4 illustrates a load balancing cluster.

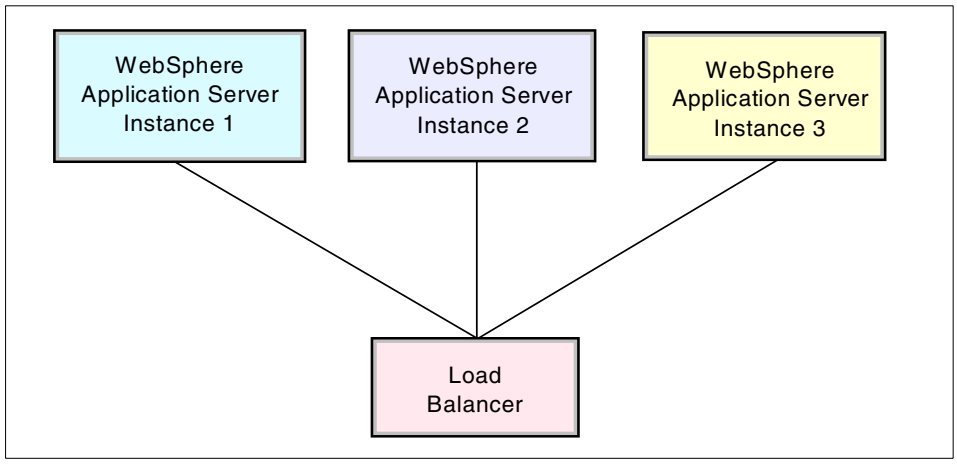


Figure 1-4 Load balancing cluster

– High Availability Cluster
  - Increases application or data availability
  - Two or more nodes (primary and backup)
  - Active/Passive and Active/Active configurations
  - “Fail-over” of resources to backup node (nodes) when primary node fails

  Figure 1-5 illustrates a high availability cluster.

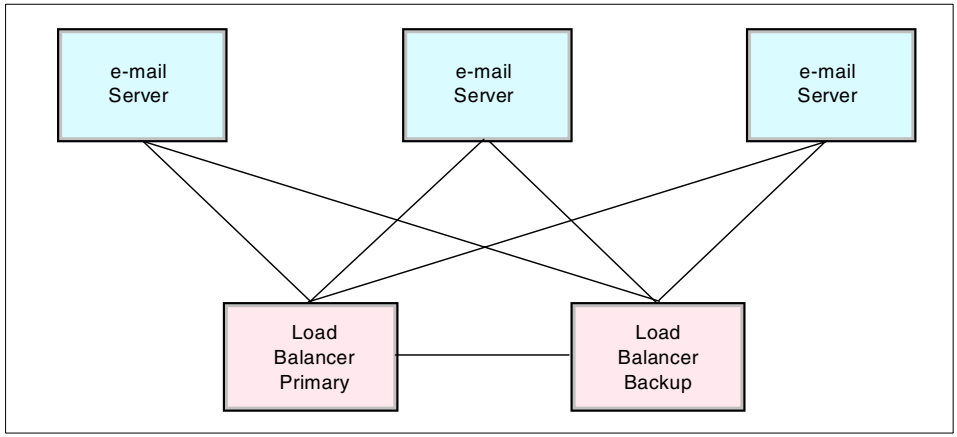


Figure 1-5 High availability cluster

- Parallel applications (applications are segmented to run in parallel and scale)
  - High Performance Technical Computing (Message Passing Cluster)
    - Extreme vertical scalability for a single large problem
    - Mostly based on MPI (Message Passing Interface) standard

      Figure 1-6 illustrates a high performance technical computing cluster.

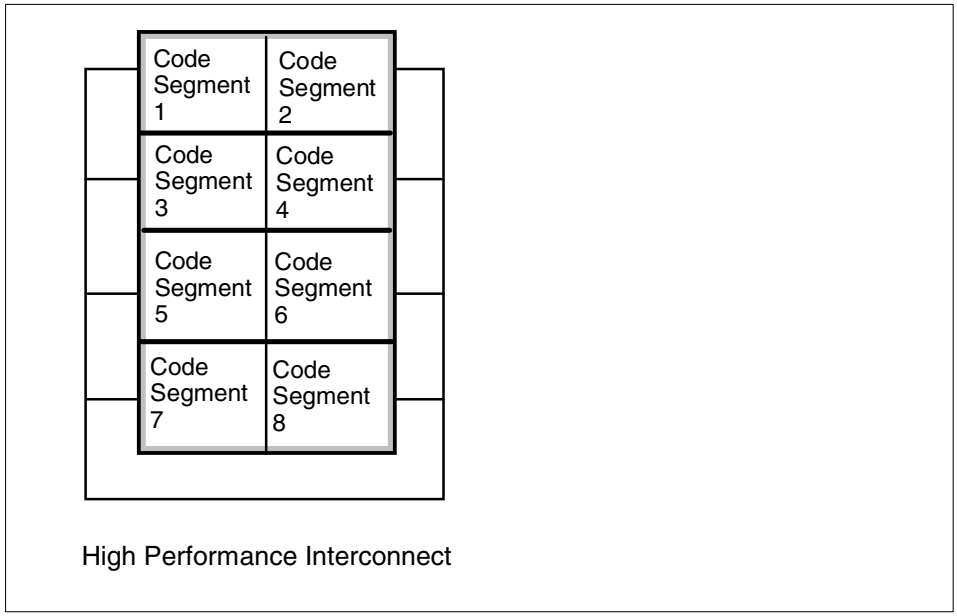


Figure 1-6 High performance technical computing cluster

  - High Performance Commercial Computing (Parallel Database Cluster)
    - Extreme database scalability
    - Shared nothing or shared disk
    - Commercial “Parallel” databases (DB2 EEE, Informix XPS, Oracle Parallel Server)

      Figure 1-7 on page 21 illustrates a high performance commercial computing cluster.

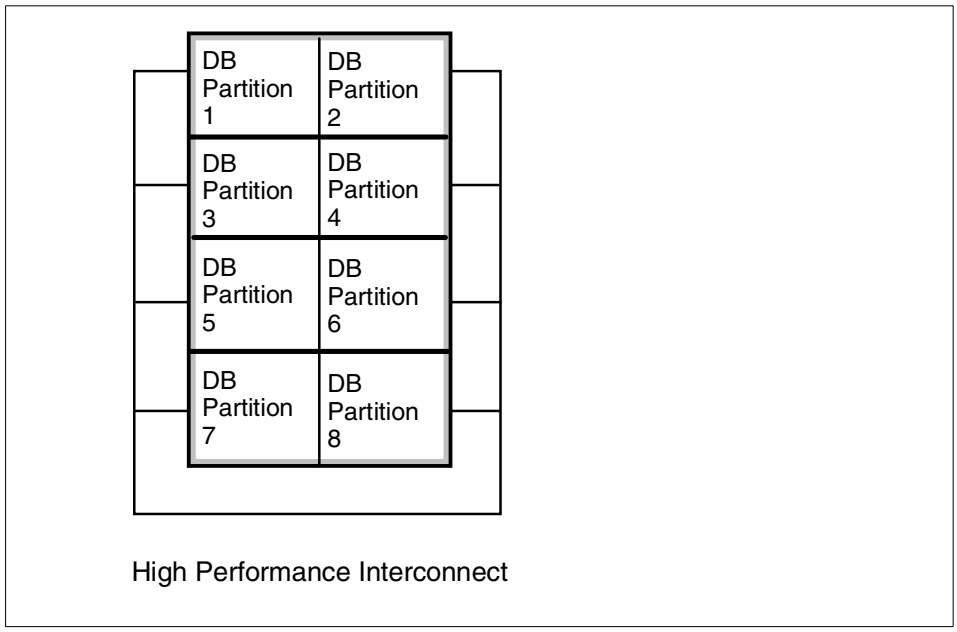


Figure 1-7 High performance commercial computing cluster

- Server Consolidation (applications are consolidated to reduce costs)
  - Common Management Domain
    - Consolidation of different workloads
    - Heterogeneous building blocks
    - Central management point-of control
    - Consolidated, simplified management interface for multiple servers

    Figure 1-8 on page 22 illustrates a common management domain.

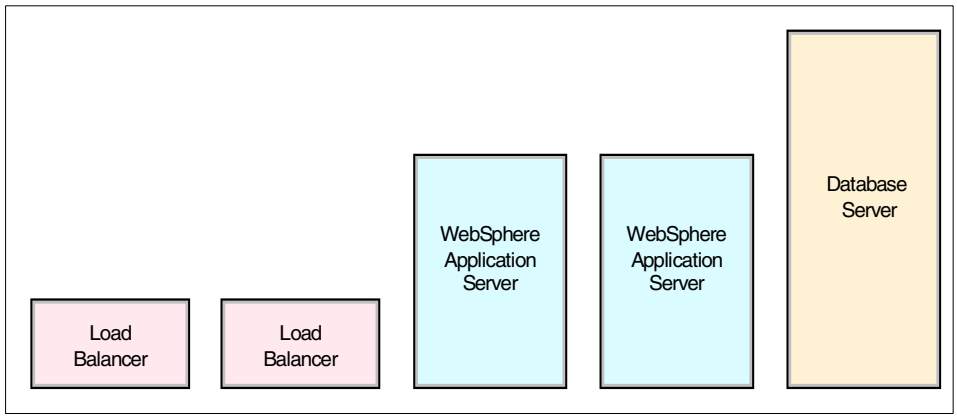


Figure 1-8 Common management domain

The following list shows benefits of AIX clustering:

- Physical isolation between partitions
- Can be used as fail-over for each other
- Easy accounting for user departments
- Economic solution for startup and growth (financial benefits of smaller systems)
- Upgrade hardware and software independently
- Easier to adopt newer technology
- Manage all systems from one console

The RS/6000 SP system, as an SP universal cluster, can implement any type of the mentioned cluster computing models separately or together.

Referring the RS/6000 SP systems and the AIX clusters evolution, we mention that:

- The original AIX cluster was based on the RS/6000 SP systems (M/T 9076), dated in 1993, and its code name is “Deep Blue.”
- The Clustered Enterprise Server system (M/T 7017 systems) is dated 2000, it included up to 16 of S85, S80, S7A and S70 systems and its code name is “Blue Hammer.”
- The Clustered Enterprise Server system (M/T 7026 systems) is dated 2001, it is based on up to 32 of M80, 6H1 and H80 systems and its code name is “Blue Hammer Jr.”

To meet the demands of e-business, the SP/CES systems, as clustered computing-based infrastructure solutions, deliver primary computing characteristics as:

- Capacity on Demand: The infrastructure may allocate and deallocate computing resources (servers, storage, processors) as workload demands grows and change; these conditions define the parameters around the delivery of a new form of horizontal scalability.
- Vertical scaling: The infrastructure is prepared to handle the largest and most complex data transaction workload through parallel execution; the users may view resources as entities which the system can partition and provision to achieve optimal execution.
- Manageable: Management of the aggregated computing facility do not require detailed actions per individual components, but instead present a system view of a unified structure that can be effectively used and coordinated as a unit.
- Quality of Service: High quality of service can be delivered in such a way that the values are perceived and recognized by a user (to provision an system with a high level of reliability, availability and serviceability).

The RS/6000 SP systems and the Clustered @server pSeries systems follow the IBM @server Clustered Computing Advantage strategies, as a set of customer values including Capacity Advantage, Availability Advantage, and Solution Advantage, enabling customers to manage their infrastructures efficiently and effectively.

The SP system is an IBM success story, not only because of Deep Blue, but because the SP, as a universal cluster system, fits any customer need.

2

SP internal processor nodes

SP processor nodes are independent RS/6000 systems mounted in SP frames. Each node has its own processors, memory, disks, I/O slots, and AIX operating system. They are the basic building blocks of the SP systems.

This chapter presents the supported nodes:

- In 375 MHz POWER3 architecture:
  - The 375 POWER3 SMP High Node (F/C 2058) or Nighthawk-II node.
  - The 375 POWER3 SMP Wide Node (F/C 2057) or Winterhawk-II Wide Node.
  - The 375 POWER3 SMP Thin Node (F/C 2056) or Winterhawk-II Thin Node.
- In POWER3 SMP architecture:
  - The POWER3 SMP High Node (F/C 2054) or Nighthawk node.
- In 332 MHz SMP architecture:
  - 332 MHz SMP Wide Node (F/C 2051) or Silver Wide Node.
  - 332 MHz SMP Thin Node (F/C 2050) or Silver Thin Node.

2.1 SP internal node

The basic RS/6000 SP building block is the servers or processor node. Each node has a processor or processors, memory, disk drives, and PCI expansion slots for I/O and connectivity such as any other server. There are three types of internal SP nodes: Thin, wide, and high. They may be mixed in a system and are housed in a frame. Figure 2-1 shows a mix of nodes in a frame.

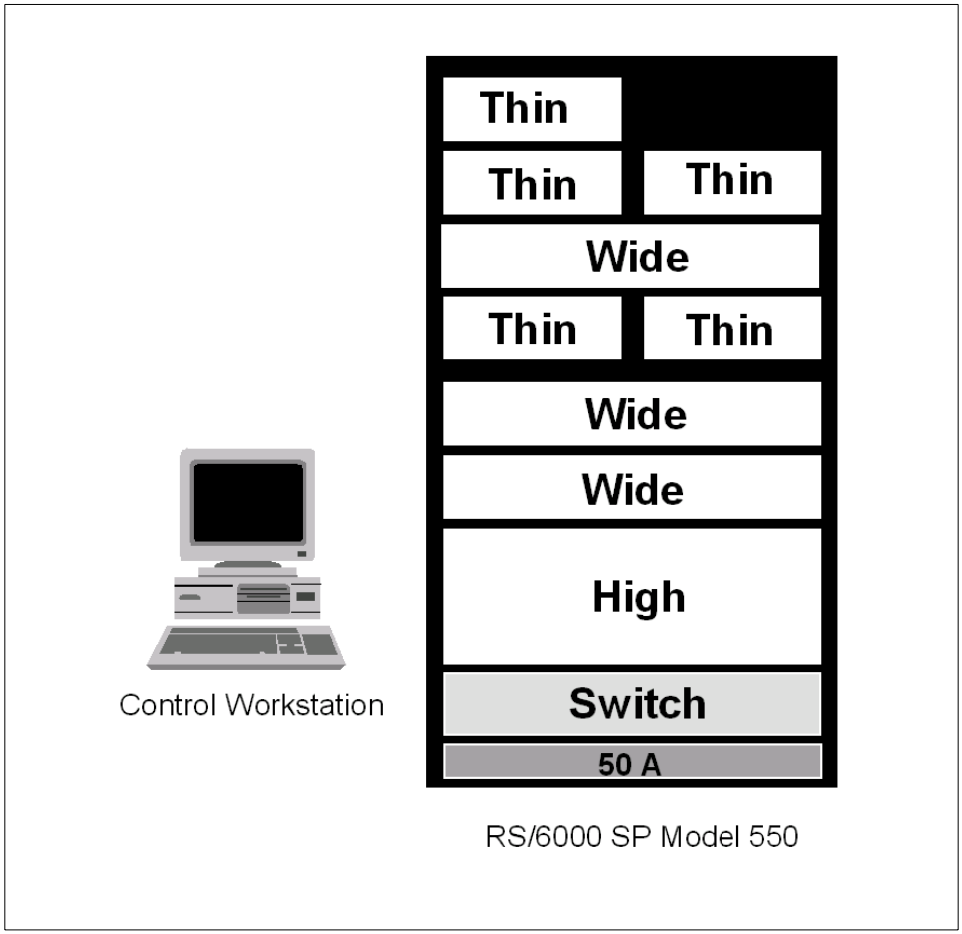


Figure 2-1 Mix of nodes in a frame

Current available nodes (can be purchased):

- 375 MHz POWER3 SMP High Node (F/C 2058)
- 375 MHz POWER3 SMP Wide Node (F/C 2057)

- 375 MHz POWER3 SMP Thin Node (F/C 2056)

Current supported nodes (cannot be ordered anymore):

- POWER3 SMP High Node (F/C 2054)
- 332 MHz SMP Wide Node (F/C 2051)
- 332 MHz SMP Thin Node (F/C 2050)

The previous nodes are referenced in Appendix B, “Hardware” on page 365.

2.2 375 MHz POWER3 SMP nodes

SP 375 MHz POWER3 SMP node system design is based on the IBM PowerPC Architecture and the RS/6000 Platform Architecture. The node is designed as a bus-based symmetrical multiprocessor (SMP) system, using a 64-bit address and a 128-bit data system bus running (2048-bit for High Node) at a 4:1 processor/clock ratio. Attached to the system bus (6xx bus) are from 2 to16 POWER3-II (630+) microprocessors, and up to two memory subsystem.

The memory-I/O controller is a general purpose chip set that controls memory and I/O for systems, such as the 375 MHz POWER3 SMP node, which implement the PowerPC MP System bus (6xx bus). This chip set consists of two semi-custom CMOS chips, one for address and control, and one for data flow. The memory-I/O controller chip set includes an independent, separately-clocked "mezzanine" bus (6xx-MX bus) to which PCI bridge chips and the SP Switch MX2 Adapter are attached. The 375 MHz POWER3 SMP system architecture partitions all the system logic into the high speed processor-memory portion and to the lower speed I/O portion. This design methodology removes electrical loading from the wide, high-speed processor-memory bus (6xx bus) allowing this bus to run much faster. The wide, high-speed 6xx bus reduces memory and intervention latency while the separate I/O bridge bus supports memory coherent I/O bridges on a narrower, more cost-effective bus.

Figure 2-2 on page 29 shows the 375 MHz POWER3 SMP High Node architecture.

2.2.1 375 MHz POWER3 processor

The 375 MHz POWER3 design contains a superscalar core, which is comprised of eight execution units, supported by a high bandwidth memory interface capable of performing four floating-point operations per clock cycle. The 375 MHz POWER3 design allows concurrent operation of fixed-point, load/store, branch, and floating-point instructions. There is a 32 KB instruction and 64 KB data level 1 cache integrated within a single chip in .22 um CMOS technology. Both instruction and data caches are parity protected.

The level 2 cache controller is integrated into the 375 MHz POWER3-II microprocessor with the data arrays and directory being implemented with external SRAM modules. The 375 MHz POWER3 microprocessor has a dedicated external interface (separate from 6xx bus interface) for the level 2 cache accesses. Access to the 6xx bus and the level 2 cache can occur simultaneously. The level 2 cache is a unified cache (that is, it holds both instruction and data), and is configured for four-way set associative configuration. The external interface to the 8 MB of level 2 cache has 256-bit width and operates at 250 MHz. This interface is ECC protected.

The 375 MHz POWER3 microprocessor is designed to provide high performance floating-point computation. There are two floating-point execution units, each supporting 3-cycle latency, 1-cycle throughput, and double/single precision Multiply-Add execution rate. Hence, the 375 MHz POWER3 microprocessor is capable of executing four floating-point operations per clock cycle, which results in a peak throughput of 1500 MFLOPS.

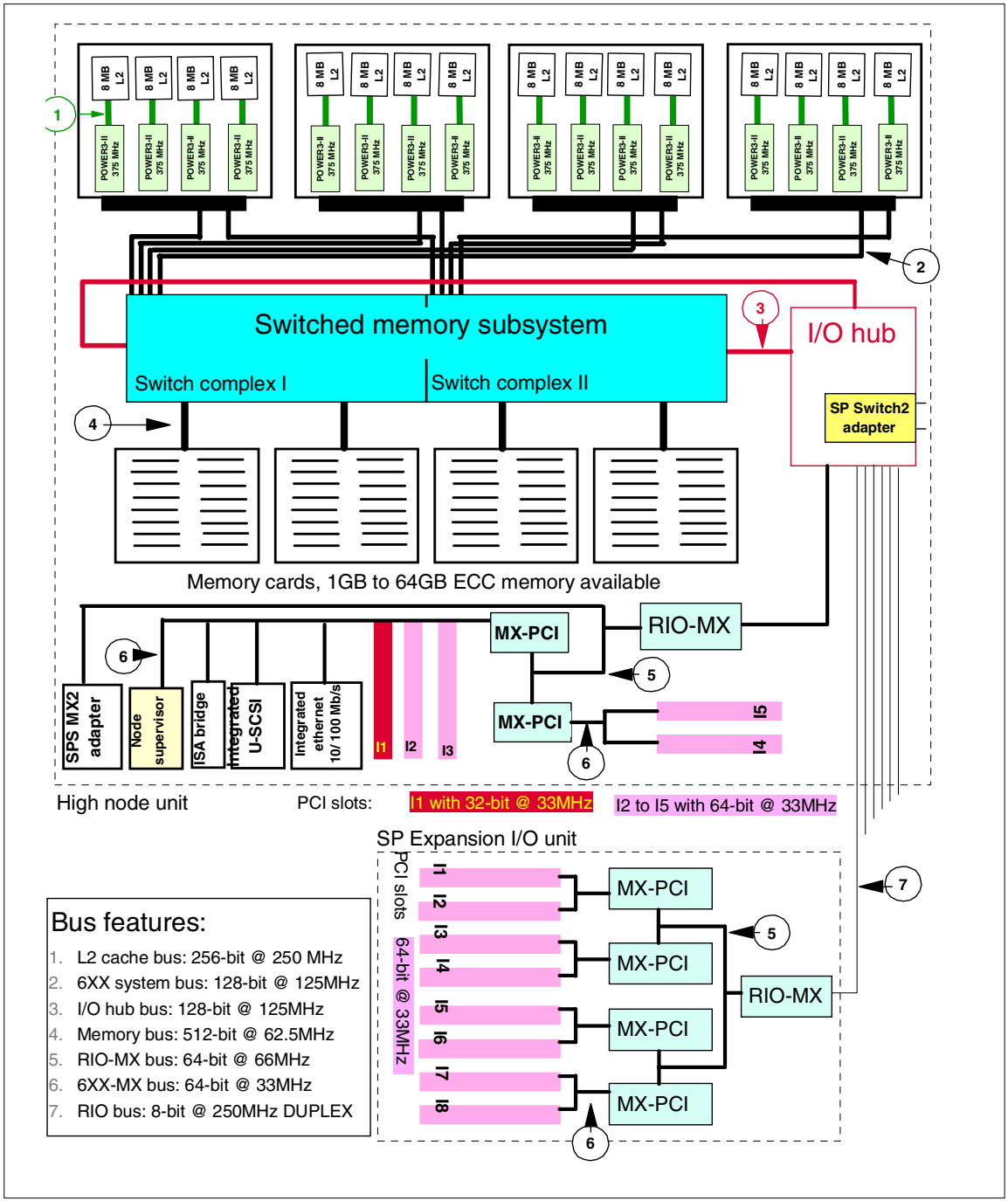


Figure 2-2 375 MHz POWER3 SMP architecture (High Node)

2.2.2 System memory

The SP 375 MHz POWER3 SMP system supports 256 MB to 64 GB of 10 ns SDRAM. System memory is controlled by the memory-I/O chip set through the memory bus. The memory bus consists of a 512-bit data bus and operates at 62.5 MHz clock cycle., this bus is separated from the System bus (6xx bus), which allows for concurrent operations on these two buses. For example, cache-to-cache transfers can occur while a DMA operation is in progress to an I/O device. There are memory card slots in the system. Each memory card contains DIMM slots; 512 MB, 256 MB and 128 MB memory DIMMs are supported. Memory DIMMs must be plugged in pairs and at least one memory card with a minimum of 256 MB of memory must be plugged in for the system to be operational. System memory is protected by Single Error Correction and Double Error Detection ECC code.

2.2.3 I/O subsystem

The I/O subsystem is different in Wide, Thin, or High Nodes:

In 375 MHz POWER3 Wide or Thin Node, the memory-I/O controller chip set implements a 64-bit plus parity, multiplexed address and data bus (6xx-MX bus) for attaching up to three PCI bridge chips and the SP Switch MX2 Adapter. The 6xx-MX bus runs at 60 MHz concurrently and independently from the 6xx and memory buses. At a 60 MHz clock cycle, the peak bandwidth of the 6xx-MX bus is 480 MBps. The three PCI bridge chips attached to 6xx-MX bus provides the interface up to 10 PCI slots: Two 32-bit PCI slots are in thin node and eight additional 64-bit PCI slots are in wide node.

One of the PCI bridge chips (Bridge Chip0) provides support for integrated Ultra2 SCSI and 10Base2, 100BaseT Ethernet functions. The Ultra2 SCSI interface supports up to four internal disks. An ISA bridge chip is also attached to PCI Bridge Chip0 for supporting two serial ports and other internally used functions in the 375 Mhz POWER3 SMP node.

The I/O subsystem in the 375 MHz POWER3 High Node has five PCI slots via the PCI bridges (MX-PCI), which have the RIO-MX and I/O hub as an additional interface.

The I/O hub can also drive up to six SP Expansion I/O units (RIO). Each RIO have eight PCI slots. Figure 2-3 on page 31 shows the rear view of POWER3 High Node and five SP Expansion I/O Units connections.

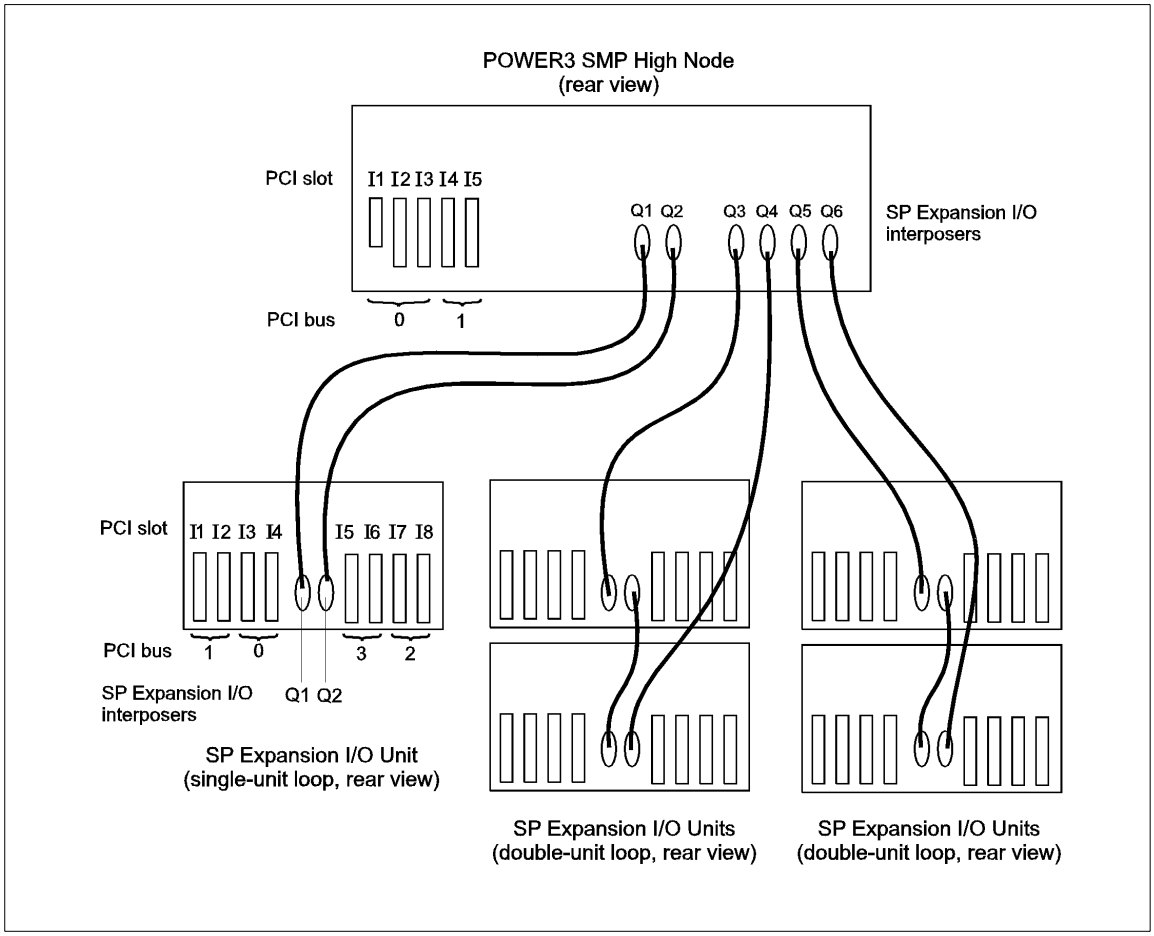


Figure 2-3 POWER3 High Node and five SP Expansion I/O Units connections

2.2.4 Service processor

The service processor function is integrated on the I/O planar board in the 375 MHz POWER3 SMP node. Service processor function is for initialization, system error recovery, and diagnostics. The service processor supports system diagnostics by saving the state of the system in a 128 KB non-volatile memory (NVRAM). The service processor code is stored in a 512 KB of flash memory and uses 512 KB of SRAM to execute. The service processor has access to latches and registers on the POWER3 microprocessors and has access to the memory-I/O controller chip set using the serial scan method.

2.2.5 System firmware and RTAS

The 375 MHz POWER3 SMP node system firmware flash memory is located on the I/O planar. System firmware contains code that is executed by the 375 MHz POWER3 microprocessor during the initial program load (IPL) phase of the system boot. It also supports various interactions between the AIX operating system and hardware. The extent and method of interaction is defined in the RS/6000 Platform Architecture (RPA). The Run Time Abstraction Software (RTAS) defined by RPA provides support for AIX and hardware for specific functions such as initialization, power management, time of day, I/O configuration, and capture and display of hardware indicators. RTAS and system IPL code are contained on 1 MB of flash memory.

2.2.6 Dynamic processor deallocation

Another high availability feature is processor deallocation. This feature allows the system to completely isolate a processor that has been determined to be operating below an acceptable level of reliability. To accomplish this, the system monitors processor status in real time to determine when the number of recoverable errors has exceeded a specified threshold. When the threshold is exceeded, the system sends a report to the operating environment that a processor is unreliable. The AIX operating environment then re-routes all interrupts and processes to other processors and disables the unreliable processor complex without interrupting end user services. The last available Firmware is recommended to be installed on the 375 MHz Node. The Web site where microcode/firmware can be downloaded from is located at:

http://www.austin.ibm.com/support/micro/download.html

2.2.7 375 MHz POWER3 SMP High Node

Table 2-1 outlines the 375 MHz POWER3 High Node and Figure 2-4 on page 34 shows the rear view of a 375 MHz POWER3 High Node packaging.

Table 2-1 375 MHz POWER3 SMP High Node

Node type	**375 MHz POWER3 SMP High Node**
Feature code	2058
Processor	64-bit POWER3-II 4-way, 8-way, 12-way, or 16-way
Clock	375 MHz
Data/instruction L1 cache	64 KB/32 KB
L2 cache/processor	8 MB

Node type	375 MHz POWER3 SMP High Node
Standard RAM	1 GB
Maximum RAM	64 GB
Disk/media bays	2/26 with SP Expansion I/O Unit
AIX storage	Internal/external boot disks Internal disks installed in pairs to support mirroring
Maximum internal storage	946.4 GB with SP Expansion I/O Unit
PCI bus speed	33 MHz
PCI slots	5 PCI (one 32-bit and four 64-bit) and 8 PCI in the Expansion Unit 53-max
Adapters	Integrated 10/100 Ethernet Integrated Ultra SCSI service processor
SP Switch adapter	MX2 adapter
SP Switch2 adapter	SP Switch2 adapter
ROLTP	81.7, 163.7, 242.3, 319.3

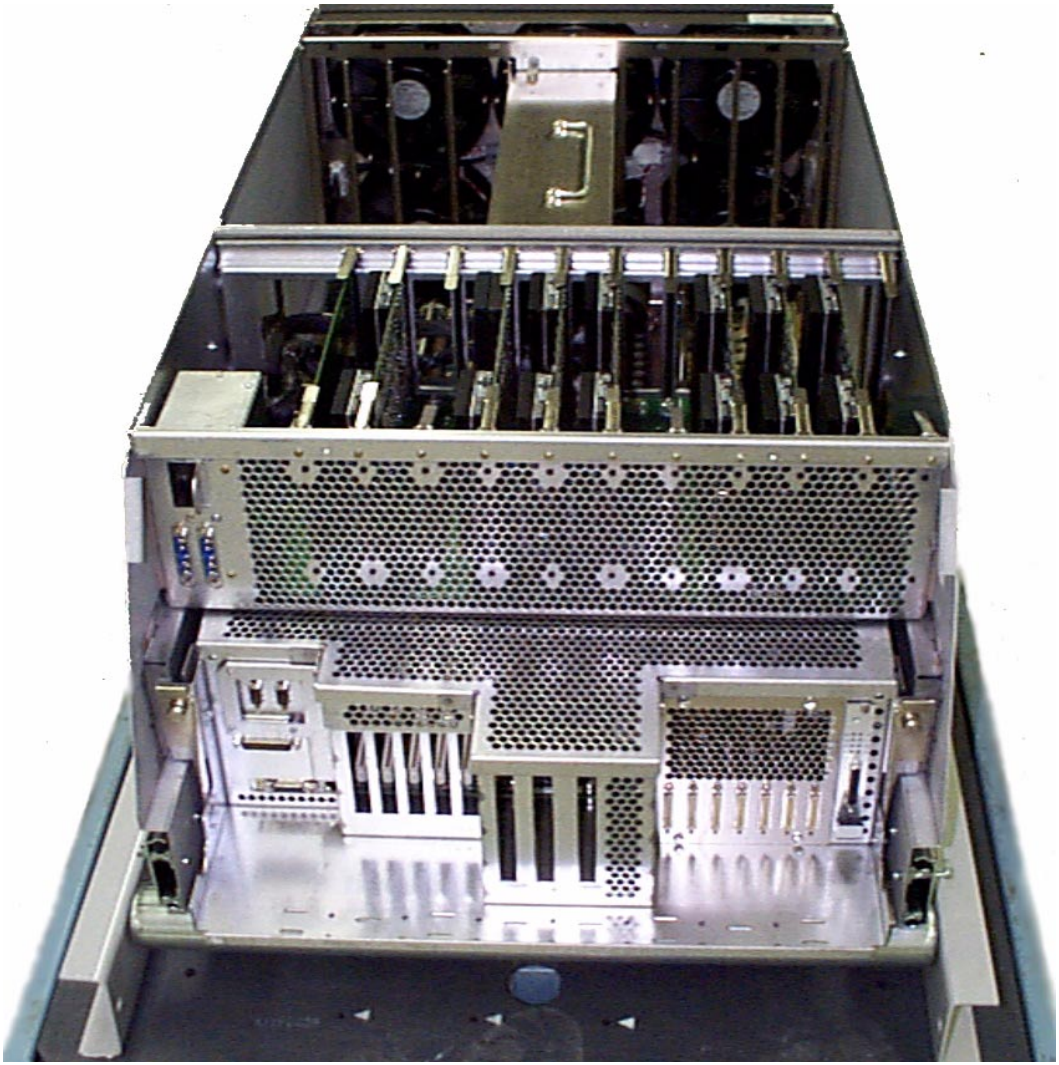


Figure 2-4 375 MHz POWER3 High Node packaging

2.2.8 375 MHz POWER3 SMP Thin and Wide Nodes

Table 2-2 outlines the 375 MHz POWER3 SMP Thin and Wide Nodes.

Figure 2-5 on page 36 shows the 375 MHz POWER3 SMP Wide Node packaging.

Table 2-2 375 MHz POWER3 SMP Thin and Wide Nodes

Node type	375 MHz POWER3 SMP Thin	375 MHz POWER3 SMP Wide
Feature code	2056	2057
Processor	64-bit POWER3-II 2-way or 4-way	64-bit POWER3-II 2-way or 4-way
Clock	375 MHz	375 MHz
Data/ Instruction (L1) cache	64 KB/32 KB	64 KB/32 KB
L2 cache/processor	8 MB	8 MB
Standard RAM	256 MB	256 MB
Maximum RAM	16 GB	16 GB
Disk/media bays	Two	Four
Storage	Internal/external boot disks Internal disks installed in pairs to support mirroring	Internal/external boot disks Internal disks installed in pairs to support mirroring
Max. internal storage	36.4 GB	109.2 GB
PCI bus speeds	33 MHz	33 MHz
PCI expansion slots	Two 32-bit	Eight 64-bit, two 32-bit
Integrated adapters	Integrated Ultra2 SCSI Ethernet (10/100 Mbps) High-speed serial port - 9 pin D-shell Service processor	Integrated Ultra2 SCSI Ethernet (10/100 Mbps) High-speed serial port - 9pin D-shell Service processor
SP Switch adapter	MX2 Switch is supported (# 4023)	MX2 Switch is supported (# 4023)
ROLTP	44.0, 67.7	44.0, 80.0

Figure 2-5 375 Mhz POWER3 SMP Wide Node packaging

2.3 375 MHz POWER3 SMP High Node (F/C 2058)

This section describes the 375 MHz POWER3 SMP High Nodes also called Nighthawk-II node. The 375 MHz POWER3 SMP High Nodes (F/C 2058) use PCI bus architecture and have four, eight, twelve, or sixteen 375 MHz 630FP 64-bit processors per node. Your IBM RS/6000 SP system must be operating at PSSP 3.1 APAR IY11093 (or later) to use these nodes.

The 375 MHz POWER3 SMP High Node occupies two full drawers locations, thus four nodes can be housed in a tall frame. These nodes require a 1.93 m (ft.)tall, deep frame (Model 550) or expansion frame (F/C 1550); they are not supported in the withdrawn 2.01 m (ft.) frame or in the 1.25 m (ft.) frame. These nodes can be placed in the first node slot of a frame without requiring additional nodes.

Note: 375 MHz POWER3 SMP High Nodes are not compatible with High Performance switches (F/C 4010 and F/C 4007).

2.3.1 PCI bus description

The 375 MHz POWER3 SMP High Node PCI bus contains one 32-bit and four 64-bit PCI slots for I/O adapters.

Additional PCI adapters can be attached to the bus by using up to six optional SP Expansion I/O Units. Each expansion unit has eight 64-bit PCI adapter slots.

2.3.2 Requirements

375 MHz POWER3 SMP High Nodes occupy two full node drawers. Up to four 375 MHz POWER3 SMP High Nodes may be installed in one tall, deep frame. The mandatory requirements are:

- PSSP 3.2 (or later) on the processor node, control workstation, and backup nodes
- Four processors (on one card, mounted in one slot)
- 1 GB of memory
- 9.1 GB mirrored Hard disk drive (with internal booting)

2.3.3 Options

Available options include the following:

- Four processors slots allowing a maximum of sixteen processors per node

- Four memory slots supporting up to 64 GB of memory
- Five PCI slots (four 64-bit and one 32-bit) for communication adapters
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter
- Integrated Ethernet with BNC and RJ45 ports (only one port can be used at a time):
  - 10Base2 Ethernet (BNC)
  - 10BaseT or 100BaseTX Ethernet (RJ45)
- Support for up to six SP Expansion I/O Units (F/C 2055)
- Two internal hard disk drive bays supporting up to 72.8 GB of storage (36.4 GB mirrored)
- Integrated Ultra SCSI network
- Two external, nine-pin, RS-232 connectors on the planar S2 and S3 ports. The S3 port is supported only for a HACMP serial heartbeat; the S2 port is not supported for this use. A 9- to 25-pin converter cable is included with the node:
  - Node-to-node HACMP cable (F/C 3124)
  - Frame-to-frame HACMP cable (F/C 3125)

2.3.4 Processor requirements and options

375 MHz POWER3 SMP High Nodes require a minimum of four 375 MHz processors mounted on one card. You can order up to three additional four-processor cards (F/C 4350) to configure the node with a total of sixteen CPUs. Table 2-3 shows the processor option for the 375 MHz POWER3 SMP High Node.

Table 2-3 375 MHz POWER3 SMP High Node (F/C 2058) processor options

Feature Code	Description	Minimum per node	Maximum per node
4350	One processor card with four CPUs	1	4

2.3.5 Memory requirements and options

375 MHz POWER3 SMP High Nodes have one to four memory cards, require a minimum of one GB of memory and support a maximum of 64 GB. Memory is supplied by 128 MB and 256 and 512 MB DIMMs which must be mounted in banks of eight DIMMs. Different capacity DIMMs (in banks of eight) can be mixed on the memory cards.

For the best memory-access bandwidth, memory DIMMs should be distributed evenly across four memory cards. As an example, you can realize better bandwidth by using four bank of 128 MB DIMMs (4 GB total) distributed evenly over four memory cards rather than by using one bank of 512 MB DIMMs (4 GB total) on one memory card. The following list illustrates this:

- 1 to 16 GB memory mounted on one card yields 16.8 - 24.3% of peak bandwidth
- 2 to 32 GB mounted on two cards yields 35.5 - 48.5% of peak
- 4 to 64 GB mounted on four cards yields 67 - 97% of peak

Table 2-4 shows the memory features for the 375 MHz POWER3 SMP High Node.

Table 2-4 375 MHz POWER3 SMP High Node (F/C 2058) memory features

Feature Code	Description	Minimum Node Requirement	Maximum Allowed Per Node
4880	Base memory card	1	4
4133	1 GB - (8 x 128 MB DIMMs)	1	16
4402	2 GB – (16 x 128 MB DIMMs)	0	8
4403	3 GB – (24 x 128 MB DIMMs)	0	4
4104	4 GB – (32 x 128MB DIMMs)	0	4
4134	2 GB – (8 x 256 MB DIMMs)	0	16
4412	4 GB – (16 x 256 MB DIMMs)	0	8
4413	6 GB – (24 x 256 MB DIMMs)	0	4
4414	8 GB – (32 x 256 MB DIMMs)	0	4
4421	4 GB – (8 x 512 MB DIMMs)	0	16
4422	8 GB – (16 x 512 MB DIMMs)	0	8
4423	12 GB – (24 x 512 MB DIMMs)	0	4
4424	16 GB - (32 x 512 MB DIMMs)	0	4

2.3.6 Disk requirements and options

375 MHz POWER3 SMP High Nodes can have one pair of internal hard disk drive attached through an integrated Ultra SCSI network. The node can have either no internal hard disk drive (with external booting) or from 9.1 GB to a maximum of 36.4 GB of mirrored, internal disk storage.

Additional hard disk drives can be attached to the 375 MHz POWER3 SMP High Node by connecting up to six SP Expansion I/O Units. Each expansion unit has four hard disk drive bays. For details, see Section 2.3.9, “SP Expansion I/O unit (F/C 2055)” on page 41.

Optional internal disk devices are available as follows:

- 9.1 GB Ultra SCSI disk pair (F/C 2909)
- 18.2 GB Ultra SCSI disk pair (F/C 2918)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3804)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3810)
- 36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3820)

External storage devices can be accessed through optional Ultra SCSI adapter (F/C 6207) and SSA adapter (F/C 6230).

2.3.7 Switch and communication adapter requirements and options

If you are planning to use the switch in your SP system, you need switch adapters to connect each RS/6000 SP node to the switch subsystem. SP switch adapters have several characteristics and restrictions, including the following:

- 375 MHz POWER3 SMP High Nodes are not supported with the SP Switch-8. You must use the SP Switch, 16-port (F/C 4011) or the SP Switch2, 16-port (F/C 4012).
- The 375 MHz POWER3 SMP High Node is not compatible with the older High Performance series of switches. If you install a POWER3 Wide Node into a switch-configured system, it must use only SP-type switches.

The switch adapter for 375 MHz POWER3 SMP High Nodes does not occupy a PCI slot; it is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

Important:

- For SP Switch systems, the nodes require the SP Switch MX2 Adapter (F/C 4023).
- For SP Switch2 systems, the nodes require the SP Switch2 Adapter (F/C 4025).

2.3.8 I/O adapters

The 375 MHz POWER3 SMP High Node has five PCI (Peripheral Component Interconnect) adapter slots. A full line of PCI adapters is offered for these nodes including:

- SCSI-2
- Ethernet
- Token Ring
- FDDI
- ATM
- Async
- Wide Area Network (WAN)
- SSA RAID5
- S/390 ESCON
- Serial HIPPI

Note: A 10BASE2 or 10BASE-T/100BASE-TX Ethernet adapter for the SP Ethernet is integrated into the POWER3 High Node and does not occupy a PCI slot.

2.3.9 SP Expansion I/O unit (F/C 2055)

This section describes the SP Expansion I/O Unit also call RIO. Each SP Expansion I/O Unit is an extension of the 375 MHz POWER3 SMP High Node, providing eight additional PCI adapter slots and four hard disk drive bays. PCI adapter hot-plug capability is supported for the SP Expansion I/O Unit with AIX 4.3.3 software loaded on the node.

Up to six expansion units can be connected to each processor node in one to three loops of one or two expansion units in each loop.

Each expansion unit (or pair of units) requires a mounting shelf (F/C 9935). This shelf occupies the space of one drawer in a frame. If only a single expansion unit is mounted in the shelf, a filler plate (F/C 9936) is required for the other side.

Expansion units can be mounted in the same frame as the node, using 2 m cables (F/C 3126), or in separate frames using 15 m cables (F/C 3127). These units require a tall, deep frame (Model 550 or F/C 1550); they are not supported in the withdrawn 2.01 m frame or in the 1.25 m frame. Figure 2-3 on page 31 shows a POWER3 High Node rear view and SP Expansion I/O Units connections

SP Expansion I/O Unit Placement: IBM suggests that SP Expansion I/O Units be mounted in separate frames, so as not to interfere with switch port utilization.

2.3.10 Hard disk drive options

Each SP Expansion I/O Unit has four hard disk drive bays, supporting one or two pairs of hard disk drives.

Attention: SCSI and SSA type hard disk drive cannot be mixed within an expansion unit.

Optional Hard disk drive pairs for SP Expansion I/O Units are available as follows:

- 9.1 GB Ultra SCSI disk pair (F/C 3800) - requires adapter (F/C 6206)
- 18.2 GB Ultra SCSI disk pair (F/C 3803) - requires adapter (F/C 6206)
- 9.1 GB SSA disk pair (F/C 3802) - requires adapter (F/C 6230)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3805) - requires adapter (F/C 6206)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3811) - requires adapter (F/C 6206)
- 18.2 GB SSA disk pair (F/C 3812) - requires adapter (F/C 6230)
- 36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3821) - requires adapter (F/C 6206) and an SP Expansion I/O Unit power upgrade (F/C 9955)
- 36.4 GB SSA 10K RPM disk pair (F/C 3822) - requires adapter (F/C 6230) and an SP Expansion I/O Unit power upgrade (F/C 9955)

Note: Empty, unused hard disk drive bay pairs require a filler plate (F/C 9612).

2.4 375 MHz POWER3 SMP Wide Node (F/C 2057)

This section describes the 375 MHz POWER3 SMP Wide Nodes, also called the Winterhawk-II Wide Node.

The 375 MHz POWER3 SMP Wide Nodes (F/C 2057) have PCI bus architecture and either two or four 375 MHz 64-bit processors per node. These nodes are functionally equivalent to an IBM RS/6000 7044-270 workstation. Your IBM RS/6000 SP system must be operating at PSSP 3.1.1 (or later) to use these nodes.

The node occupies one full drawer, thus eight nodes can be housed in a tall frame. These nodes can be placed in the first node slot of a frame without requiring additional nodes.

For electromagnetic compliance, these nodes are housed in an enclosure (F/C 9930).

If you plan to install a 375 MHz POWER3 SMP Wide Node into an early style 2.01 m or 1.25 m frame, a power system upgrade is required. Once the power system upgrade is done, these nodes are fully compatible with all existing SP system hardware except High Performance switches.

2.4.1 Bus description

The 375 MHz POWER3 SMP Wide Node PCI bus contains two 32-bit slots and eight 64-bit PCI slots divided into three logical groups. The first slot group (slot I2 and I3) is composed of the two 32-bit slots on the CPU side of the node. The second and third groups each contain four 64-bit PCI slots (slots I1-I4 and slots I5-I8) on the I/O side of the node. The I1 slot on the CPU side is reserved for the optional SP Switch MX2 Adapter.

2.4.2 Requirements

375 MHz POWER3 SMP Wide Nodes occupy one full node drawer. They can be asymmetrically configured for memory, hard disk drive, and adapters. Up to eight of these nodes can be installed in a tall frame and up to four in a short frame. The mandatory requirements are:

- PSSP 3.1.1 (or later) on the control workstation, backup nodes, and processor node.
- Two processors (mounted in one slot).
- 256 MB of memory.
- 4.5 GB of mirrored Hard disk drive (with internal booting).

- An upgraded power system on early-style frames.

2.4.3 Options

Available options include the following:

- Two processor slots allowing a maximum of four processors per node.
- Two memory slots supporting up to 8 GB of memory.
- Ten PCI slots (two 32-bit and eight 64-bit) for communication adapters.
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter.
- Integrated Ethernet with BNC and RJ45 ports (only one per port can be used at a time).
  - 10Base2 Ethernet (BNC).
  - 10BaseT Ethernet or 100BaseTX Ethernet (RJ45).
- Four hard disk drive bays supporting up to 109.2 GB of Storage (54.6 GB mirrored).
- Integrated Ultra SCSI.
- Standard service processor.
- External nine-pin RS232 on the planar S2 port (supported only for HACMP serial heartbeat); a 9- to 25-pin converter cable is included with the node.
  - Node-to-node HACMP cable (F/C 3124).
  - Frame-to-frame HACMP cable (F/C 3125).

2.4.4 Processor requirements and options

375 MHz POWER3 SMP Wide Nodes require a minimum of two processors mounted on one card. Optionally, you can order an additional processor card for a total of four CPUs. Table 2-5 shows the processor options for the 375 MHz POWER3 SMP Wide Nodes.

Table 2-5 375 MHz POWER3 SMP Wide Nodes processor option

Feature Code	Description	Minimum per node	Maximum per node
4444	One processor card with two CPUs	1	2

2.4.5 Memory requirements and options

375 MHz POWER3 SMP Wide Nodes require two memory cards and a minimum of 256 MB of memory. These nodes support a maximum of 16 GB of memory; each card has an 8 GB capacity. Memory is supplied by 128, 256, and 512 MB DIMMs which must be mounted in pairs. Different capacity DIMM-pairs can be mixed on the memory cards. Note that with the minimum memory (256 MB) installed, the second memory card contains no DIMMs. Table 2-6 shows the memory features for the 375 MHz POWER3 SMP Wide Nodes.

Table 2-6 375 MHz POWER3 SMP Wide Nodes memory features

Feature Code	Description	Minimum Node Requirement	Maximum Features Per Node
4098	Base Memory Card	2	2
4110	256 MB – (2 x 128 MB DIMMs)	1	16
4133	1 GB – (8 x 128 MB DIMMs)	0	4
4402	2 GB – (16 x 128 MB DIMMs)	0	2
4403	3 GB – (24 x 128 MB DIMMs)	0	1
4404	4 GB – (32 x 128 MB DIMMs)	0	1
4119	512 MB – (2 x 256 MB DIMMs)	0	16
4134	2 GB – (8 x 256 MB DIMMs)	0	4
4412	4 GB – (16 x 256 MB DIMMs)	0	2
4413	6 GB – (24 x 256 MB DIMMs)	0	1
4414	8 GB – (32 x 256 MB DIMMs)	0	1
4100	1 GB – (2 x 512 MB DIMMs)	0	16
4421	4 GB – (8 x 512 MB DIMMs)	0	4
4422	8 GB – (16 x 512 MB DIMMs)	0	2
4423	12 GB – (24 x 512 MB DIMMs)	0	1
4424	16 GB – (32 x 512 MB DIMMs)	0	1

2.4.6 Hard disk drive requirements and options

375 MHz POWER3 SMP Wide Nodes can have up to two pairs of internal hard disk drives attached through an integrated Ultra SCSI network. This node can have either no internal hard disk drive (with external booting) or from 4.5 GB to a maximum of 54.6 GB of mirrored internal disk storage.

Optional direct access storage devices are available as follows:

- 4.5 GB Ultra SCSI disk pair (F/C 2904)
- 9.1 GB Ultra SCSI disk pair (F/C 2909)
- 18.2 GB Ultra SCSI disk pair (F/C 2918)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3804)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3810)
- 36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3820) - available only for I/O side hard disk drive bays

Note: No special cables or adapters are required to mount these internal hard disk drive. However, this node has an option (F/C 1241) that provides an independent SCSI hookup with the following characteristics:

- Eliminates the hard disk drive controller as a single point of failure during mirroring
- Increases disk performance
- Balances disk loading

The F/C 1241 option requires a PCI type SCSI adapter F/C 6206.

External storage devices can be accessed through an optional Ultra SCSI adapter (F/C 6207), SCSI-2 adapter (F/C 6209), and SSA adapter (F/C 6230).

2.4.7 Switch and communication adapter requirements

If you are planning to use a switch in your SP system, you need switch adapters to connect each RS/6000 SP node to the switch subsystem. SP switch adapters have several characteristics and restrictions, including the following:

Switch adapters

The switch adapter for 375 MHz POWER3 SMP Wide Nodes does not occupy a PCI slot. The switch adapter for these nodes is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

In switch-configured systems, 375 MHz POWER3 SMP Wide Nodes require SP Switch MX2 adapter (F/C 4023).

Switch restrictions

375 MHz POWER3 SMP Wide Nodes are not compatible with the older High Performance series of switches. If you install this node into a switch-configured system, you must use an SP Switch or an SP Switch-8.

Switch adapters for these nodes are not interchangeable with either the switch adapters used on uniprocessor wide nodes or with the SP Switch MX adapter used on 332 MHz SMP nodes.

I/O adapters

The 375 MHz POWER3 SMP Wide Nodes have 10 PCI (Peripheral Component Interconnect) adapter slots. A full line of PCI adapters is offered for these nodes including:

- SCSI-2
- Ethernet
- Token Ring
- FDDI
- ATM
- Async
- Wide Area Network (WAN)
- SSA RAID5
- S/390 ESCON
- Serial HIPPI

Note: A 100BaseTX, 10BaseT, or 10Base2 adapter for the SP Ethernet is integrated into the node and does not occupy a PCI slot.

2.5 375 MHz POWER3 SMP Thin Node (F/C 2056)

This section describes the 375 MHz POWER3 SMP Thin Nodes, also called the Winterhawk-II Thin Node.

375 MHz POWER3 SMP Thin Nodes (F/C 2056) have PCI bus architecture and either two or four 375 MHz 64-bit processors per node. These nodes are functionally equivalent to an IBM RS/6000 7044-270 workstation. Your IBM RS/6000 SP system must be operating at PSSP 3.1.1 (or later) to use these nodes.

The node occupies half of a drawer (one slot). Up to 16 of these nodes can be housed in a tall frame. When installed singly within a drawer, these nodes must be placed in an odd-numbered node slot. For complete information on node/frame configurations, see *RS/6000 SP: Planning Volume 2, Control Workstation and Software Environment*, GA22-7281.

For electromagnetic compliance, these nodes are housed in an enclosure (F/C 9930). If you order a single node, a cover plate (F/C 9931) is included to fill the even-numbered slot opening.

If you plan to install a 375 MHz POWER3 SMP Thin Node into an early style 2.01 m or 1.25 m frame, a power system upgrade is required. Once the power system upgrade is done, these nodes are fully compatible with all existing SP system hardware except High Performance switches.

2.5.1 Bus description

The 375 MHz POWER3 SMP Thin Node PCI bus contains two 32-bit slots, I2 and I3. The I1 slot is reserved for the optional SP Switch MX2 Adapter.

2.5.2 Requirements

375 MHz POWER3 SMP Thin Nodes occupy one half node drawer. They can be asymmetrically configured for memory, hard disk drive, and adapters. Up to sixteen of these nodes can be installed in a tall frame and up to eight in a short frame. The mandatory requirements are:

- PSSP 3.1.1 (or later) on the control workstation, backup nodes, and processor node.

- Two processors (mounted in one slot).
- 256 MB of memory.
- 4.5 GB of mirrored Hard disk drive (with internal booting).
- An upgraded power system on early-style frames.

2.5.3 Options

Available options include the following:

- Four processors in two slots.
- Two memory slots supporting up to 16 GB of memory.
- Two (32-bit) PCI slots for communication adapters.
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter.
- Integrated Ethernet with BNC and RJ45 ports (only one per port can be used at a time).
  - 10Base2 Ethernet (BNC)
  - 10BaseT Ethernet or 100BaseTX Ethernet (RJ45)
- Two hard disk drive bays supporting up to 36.4 GB of Storage (18.2 GB mirrored).
- Integrated Ultra SCSI.
- Standard service processor.
- External 9-pin RS232 on the planar S2 port (supported only for HACMP serial heartbeat); a 9- to 25-pin converter cable is included with the node.
  - Node-to-node HACMP cable (F/C 3124)
  - Frame-to-frame HACMP cable (F/C 3125)

2.5.4 Processor requirements and options

375 MHz POWER3 SMP Thin Nodes require a minimum of two processors mounted on one card. Optionally, you can order an additional processor card for a total of four CPUs. Table 2-7 shows the processor options for the 375 MHz POWER3 SMP Thin Nodes.

Table 2-7 375 MHz POWER3 SMP Thin Nodes (F/C 2056) Processor option

Feature Code	Description	Minimum per node	Maximum per node
4444	One processor card with two CPUs	1	2

2.5.5 Memory requirements and options

375 MHz POWER3 Thin Nodes require two memory cards and a minimum of 256 MB of memory. These nodes support a maximum of 16 GB of memory; each card has an 8 GB capacity. Memory is supplied by 128, 256, and 512 MB DIMMs which must be mounted in pairs. Different capacity DIMM-pairs can be mixed on the memory cards. Note that with the minimum memory (256 MB) installed, the second memory card contains no DIMMs. Table 2-8 shows the memory features for the 375 MHz POWER3 SMP Thin Nodes.

Table 2-8 375 MHz POWER3 SMP Thin Nodes (F/C 2056) memory features

Feature Code	Description	Minimum Node Requirement	Maximum Features Per Node
4098	Base Memory Card	2	2
4110	256 MB – (2 x 128 MB DIMMs)	1	16
4133	1 GB – (8 x 128 MB DIMMs)	0	4
4402	2 GB – (16 x 128 MB DIMMs)	0	2
4403	3 GB – (24 x 128 MB DIMMs)	0	1
4404	4 GB – (32 x 128 MB DIMMs)	0	1
4119	512 MB – (2 x 256 MB DIMMs)	0	16
4134	2 GB – (8 x 256 MB DIMMs)	0	4
4412	4 GB – (16 x 256 MB DIMMs)	0	2
4413	6 GB – (24 x 256 MB DIMMs)	0	1
4414	8 GB – (32 x 256 MB DIMMs)	0	1
4100	1 GB – (2 x 512 MB DIMMs)	0	16
4421	4 GB – (8 x 512 MB DIMMs)	0	4
4422	8 GB – (16 x 512 MB DIMMs)	0	2
4423	12 GB – (24 x 512 MB DIMMs)	0	1
4424	16 GB – (32 x 512 MB DIMMs)	0	1

2.5.6 Hard disk drive requirements and options

375 MHz POWER3 SMP Thin Nodes can have one pair of internal hard disk drive attached through an integrated Ultra SCSI network. The node can have either no internal hard disk drive (with external booting) or from 4.5 GB to a maximum of 36.4 GB of mirrored internal disk storage.

Optional direct access storage devices are available as follows:

- 4.5 GB Ultra SCSI disk pair (F/C 2904)
- 9.1 GB Ultra SCSI disk pair (F/C 2909)
- 18.2 GB Ultra SCSI disk pair (F/C 2918)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3804)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3810)

Note: No special cables or adapters are required to mount these internal hard disk drives.

External storage devices can be accessed through an optional Ultra SCSI adapter (F/C 6207), SCSI-2 adapter (F/C 6209), and SSA adapter (F/C 6230).

2.5.7 Switch and communication adapter requirements

If you are planning to use a switch in your SP system, you need switch adapters to connect each RS/6000 SP node to the switch subsystem. SP switch adapters have several characteristics and restrictions, including the following:

Switch adapters

The switch adapter for 375 MHz POWER3 SMP Thin Nodes does not occupy a PCI slot. The switch adapter for these nodes is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

In switch-configured systems, the 375 MHz POWER3 SMP Thin Nodes require the SP Switch MX2 adapter (F/C 4023).

Switch restrictions

375 MHz POWER3 SMP Thin Nodes are not compatible with the older High Performance series of switches. If you install this node into a switch-configured system, you must use an SP Switch or an SP Switch-8.

Switch adapters for these nodes are not interchangeable with either the switch adapters used on uniprocessor wide nodes or with the SP Switch MX adapter (F/C 4022) used on 332 MHz SMP nodes.

I/O adapters

The 375 MHz POWER3 SMP Thin Nodes have two PCI (Peripheral Component Interconnect) adapter slots. A full line of PCI adapters is offered for these nodes including:

- SCSI-2
- Ethernet
- Token Ring
- FDDI
- ATM
- Async
- Wide Area Network (WAN)
- SSA RAID5
- S/390 ESCON

Note: A 100BaseTX, 10BaseT, or 10Base2 adapter for the SP Ethernet is integrated into the node and does not occupy a PCI slot.

2.6 POWER3 SMP nodes

The POWER3 SMP node is the first scalable processor node that utilizes the POWER3 64-bit microprocessor. The floating-point performance of the POWER3 microprocessor makes this node an excellent platform for compute-intensive analysis applications. The POWER3 microprocessor offers technical leadership for floating-point applications by integrating two floating-point, three fixed-point, and two load/store units in a single 64-bit PowerPC implementation. Since the node conforms to the RS/6000 Platform Architecture, compatibility is maintained for existing device drivers, other subsystems, and applications. The POWER3 SMP node supports AIX operating systems beginning with version 4.3.2.

2.6.1 POWER3 SMP node system architecture

POWER3 SMP node system design is based on the IBM PowerPC Architecture and the RS/6000 Platform Architecture. The node is designed as a bus-based symmetrical multiprocessor (SMP) system using a 64-bit address and a 128-bit data system bus running at a 2:1 processor clock ratio. The memory-I/O controller is a general purpose chip set that controls memory and I/O for systems, such as the POWER3 SMP node, which implement the PowerPC MP System bus (6xx bus). This chip set consists of two semi-custom CMOS chips, one for address and control, and one for data flow. The memory-I/O controller chip set includes an independent, separately-clocked mezzanine bus (6xx-MX bus) to which three PCI bridge chips and the SP Switch MX2 Adapter are attached. The POWER3 SMP system architecture partitions all the system logic into the high speed processor-memory portion and to the lower speed I/O portion. This design methodology removes electrical loading from the wide, high-speed processor-memory bus (6xx bus) allowing this bus to run much faster. The wide, high-speed 6xx bus reduces memory and intervention latency, while the separate I/O bridge bus supports memory coherent I/O bridges on a narrower, more cost-effective bus.

2.6.2 POWER3 microprocessor

The POWER3 design contains a superscalar core, which is comprised of eight execution units, supported by a high bandwidth memory interface capable of performing four floating-point operations per clock cycle. The POWER3 design allows concurrent operation of fixed-point, load/store, branch, and floating-point instructions. There is a 32 KB instruction and 64 KB data level 1 cache integrated within a single chip in .25 um CMOS technology. Both instruction and data caches are parity protected. The level 2 cache controller is integrated into the POWER3 microprocessor with the data arrays and directory being implemented with external SRAM modules. The POWER3 microprocessor has a dedicated external interface (separate from 6xx bus interface) for the level 2 cache accesses. Access to the 6xx bus and the level 2 cache can occur simultaneously. The level 2 cache is a unified cache (that is, it holds both instruction and data) and is configured for direct mapped configuration. The external interface to the 4 MB of level 2 cache has 256-bit width and operates at 200 MHz. This interface is ECC protected. The POWER3 microprocessor is designed to provide high performance floating-point computation. There are two floating-point execution units, each supporting 3-cycle latency, 1-cycle throughput, and double/single precision Multiply-Add execution rate. Hence, the POWER3 microprocessor is capable of executing four floating-point operations per clock cycle, which results in a peak throughput of 800 MFLOPS.

2.6.3 6xx bus

The 6xx bus or System bus, connects up to four processor cards to the memory-I/O controller chip set. This bus is optimized for high performance and multiprocessing applications. It provides 40 bits of real address and a separate 128-bit data bus. The address, data, and tag buses are fully parity checked, and each memory or cache request is range checked and positively acknowledged for error detection. Any error will cause a machine check condition and is logged in AIX error logs. The 6xx bus runs at a 100 MHz clock rate, and peak data throughput is 1.6 GB/second. Data and address buses operate independently in true split transaction mode and are pipelined so that new requests may be issued before previous requests are snooped or completed.

2.6.4 System memory

The SP POWER3 SMP system supports 256 MB to 16 GB of 10ns SDRAM. System memory is controlled by the memory-I/O chip set via the memory bus. The memory bus consists of a 128-bit data bus and operates at 100 MHz clock cycle. This bus is separated from the System bus (6xx bus), which allows for concurrent operations on these two buses. For example, cache to cache transfers can occur while a DMA operation is in progress to an I/O device.

There are two memory cards slots in the system. Each memory card contains 16 DIMM slots. Only 128 MB memory DIMMs are supported. Memory DIMMs must be plugged in pairs, and at least one memory card with minimum of 256 MB of memory must be plugged in for system to be operational. System memory is protected by Single Error Correction, Double Error Detection ECC code.

2.6.5 I/O subsystem

The I/O subsystem is similar to 375 MHz POWER3 describe in Section 2.2.3, “I/O subsystem” on page 30. The System firmware and RTAS are similar to 375 MHz POWER3 describe in Section 2.2.5, “System firmware and RTAS” on page 32.

Service Processor microcode and System firmware microcode are available in the following Web site:

`http://www.austin.ibm.com/support/micro/download.html`

2.7 222 MHz POWER3 SMP High Node (F/C 2054)

This section describes the 222 MHz POWER3 SMP High Node, also called Nighthawk Node. The 222 MHz POWER3 SMP High Nodes (F/C 2054) use PCI bus architecture and have either two, four, six, or eight 222 MHz 64-bit processors per node. Your IBM RS/6000 SP system must be operating at PSSP 3.1.1 (or later) to use these nodes.

The 222 MHz POWER3 High Node provides additional hard disk drive and PCI adapter capacity by connecting to SP Expansion I/O Units.

The 222 MHz POWER3 SMP High Node occupies two full drawer locations, thus four nodes can be housed in a tall (1.93 m) frame. POWER3 SMP High Nodes can be placed in the first node slot of a frame without requiring additional nodes.

The 222 MHz POWER3 SMP High Node requires a tall, deep frame (Model 550 or F/C 1550); they are supported in the withdrawn 2.01 m frame or in the 1.25 m frame.

Note:

- The 222 MHz POWER3 SMP High Node (F/C 2054) is supported but not available. The installed node can be upgrade to the 375 MHz POWER3 SMP High Node (F/C 2058) mainly by replacing the CPU cards and Firmware.
- The 222 MHz POWER3 SMP High Node is not compatible with High Performance switches (F/C 4010 and F/C 4007).

2.7.1 PCI bus description

The 222 MHz POWER3 SMP High Node PCI bus contains one 32-bit and four 64-bit PCI slots for I/O adapters.

Additional PCI adapters can be attached to the bus by using up to six optional SP Expansion I/O Units. Each expansion unit has eight 64-bit PCI adapter slots.

2.7.2 Requirements

222 MHz POWER3 SMP High Nodes occupy two full node drawers. Up to four POWER3 SMP High Nodes may be installed in one tall/deep frame. The mandatory requirements are:

- PSSP 3.1.1 (or later) on the processor node, control workstation, and backup nodes.

- Two Processors (on one card, mounted in one slot).
- 1 GB of memory.
- 9.1 GB mirrored Hard disk drive (with internal booting).

2.7.3 Options

Available options include the following:

- Four processors slots allowing a maximum of eight processors per node.
- Four memory slots supporting up to 16 GB of memory.
- Five PCI slots (four 64-bit and one 32-bit) for communication adapters.
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter.
- Integrated Ethernet with BNC and RJ45 ports (only one port can be used at a time):
  - 10Base2 Ethernet (BNC).
  - 10BaseT or 100BaseTX Ethernet (RJ45).
- Support for up to six SP Expansion I/O Units (F/C 2055).
- Two internal hard disk drive bays supporting up to 72.8 GB of storage (36.4 GB mirrored).
- Integrated Ultra SCSI network.
- Two external, nine-pin, RS-232 connectors on the planar S2 and S3 ports. The S3 port is supported only for HACMP serial heartbeat; the S2 port is not supported for this use. A 9- to 25-pin converter cable is included with the node for this connector.
  - Node-to-node HACMP cable (F/C 3124).
  - Frame-to-frame HACMP cable (F/C 3125).

2.7.4 Processor requirements and options

222 MHz POWER3 SMP High Nodes require a minimum of two 222 MHz, PowerPC processors mounted on one card. You can order up to three additional two-processor cards (F/C 4849) to configure the node with a total of sixteen CPUs. Table 2-9 shows the processor options for the POWER3 SMP High Nodes.

Table 2-9 POWER3 SMP High Node (F/C 2054) processor options

Feature Code	Description	Minimum per node	Maximum per node
4849	One processor card with two CPUs	1	4

2.7.5 Memory options

222 MHz POWER3 SMP High Nodes have one to four memory cards, require a minimum of one GB of memory, and support a maximum of 16 GB. Memory is supplied by 128 MB DIMMs which must be mounted in banks of eight DIMMs.

For the best memory-access bandwidth, memory DIMMs should be distributed evenly across four memory cards. The following list illustrates this:

- 1 to 4 GB memory mounted on one card yields 16.8 - 24.3% of peak bandwidth.
- 2 to 8 GB mounted on two cards yields 33.5 - 48.5% of peak.
- 4 to 16 GB mounted on four cards yields 67 - 97% of peak.

The configurator rules used for memory placement in these nodes are designed to yield the best memory performance. Any plans to increase the amount of memory in the future should always be taken into consideration when deciding what quantity of memory cards to order. Table 2-10 shows the memory features for the POWER3 SMP High Node.

Table 2-10 POWER3 SMP High Node (F/C 2054) memory features

Feature Code	Description	Minimum per node	Maximum per node
4880	Base memory card	1	4
4133	1 GB – (8 x 128 MB DIMMs)	1	16
4402	2 GB – (8 x 128 MB DIMMs)	0	8
4403	3 GB – (8 x 128 MB DIMMs)	0	4
4404	4 GB – (8 x 128 MB DIMMs)	0	4

2.7.6 Disk requirements and options

222 MHz POWER3 SMP High Nodes can have one pair of internal hard disk drive attached through an integrated Ultra SCSI network. The node can have either no internal hard disk drive (with external booting) or from 9.1 GB to a maximum of 36.4 GB of mirrored, internal disk storage.

Additional hard disk drive can be attached to the POWER3 High Node by using up to six SP Expansion I/O units. Each expansion unit has four hard disk drive bays.

Optional internal disk devices are available as follows:

- 9.1 GB Ultra SCSI disk pair (F/C 2909)
- 18.2 GB Ultra SCSI disk pair (F/C 2918)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3804)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3810)
- 36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3820)

External storage devices can be accessed through optional Ultra SCSI adapter (F/C 6207) and SSA adapter (F/C 6230).

2.7.7 Switch and communication adapter requirements and options

If you are planning to use a switch in your SP system, you need switch adapters to connect each RS/6000 SP node to the switch subsystem. SP switch adapters have several characteristics and restrictions, including the following.

Switch restrictions

POWER3 SMP High Nodes are not supported with the SP Switch-8. You must use the SP Switch, 16-port (F/C 4011) or the SP Switch2, 16-port (F/C 4012).

Note: The POWER3 SMP High Node is not compatible with the older high Performance series of switches. If you install a POWER3 Wide Node into a switch-configured system, it must use only SP-type switches.

Switch Adapters

The switch adapter for POWER3 SMP High Nodes does not occupy a PCI slot; it is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

Important:

- For SP Switch systems, these nodes require the SP Switch MX2 Adapter (F/C 4023).
- For SP Switch2 systems, these nodes require the SP Switch2 Adapter (F/C 4025).

2.7.8 I/O adapters

The POWER3 SMP High Node has five PCI (Peripheral Component Interconnect) adapter slots. A full line of PCI adapters is offered for these nodes including:

- SCSI-2
- Ethernet
- Token Ring
- FDDI
- ATM
- Async
- Wide Area Network (WAN)
- SSA RAID5
- S/390 ESCON
- Serial HIPPI

Note: A 10Base2 or 10BaseT/100BaseTX Ethernet adapter for the SP Ethernet is integrated into the POWER3 High Node and does not occupy a PCI slot.

2.7.9 SP Expansion I/O unit (F/C 2055)

Each SP Expansion I/O Unit is an extension of the POWER3 SMP High Node, providing eight additional PCI adapter slots and four hard disk drive bays. PCI adapter hot-plug capability is supported for the SP Expansion I/O Unit with AIX 4.3.3 software loaded on the node.

Up to six expansion units can be connected to each processor node in one to three loops of one or two expansion units in each loop. Figure 2-3 on page 31 shows a rear view of POWER3 High Node and five SP Expansion I/O Units connections

Each expansion unit (or pair of units) requires a mounting shelf (F/C 9935). This shelf occupies the space of one drawer in a frame. If only a single expansion unit is mounted in the shelf, a filler plate (F/C 9936) is required for the other side.

Expansion units can be mounted in the same frame as the node, using 2 m cables (F/C 3126), or in separate frames using 15 m cables (F/C 3127). These units require a tall, deep frame (Model 550 or F/C 1550); they are not supported in the withdrawn 2.01 m frame or in the 1.25 m frame. Figure 2-3 on page 31 shows a POWER3 High Node rear view and SP Expansion I/O Units connections.

SP Expansion I/O Unit Placement: IBM suggests that SP Expansion I/O Units be mounted in separate frames, so as not to interfere with switch port utilization.

2.7.10 Hard disk drive options

Each SP Expansion I/O Unit has four hard disk drive bays, supporting one or two pairs of hard disk drives.

Attention: SCSI and SSA type Hard disk drive cannot be mixed within an expansion unit.

Optional hard disk drive pairs for SP Expansion I/O Units are available as follows:

- 9.1 GB Ultra SCSI disk pair (F/C 3800) - requires adapter (F/C 6206)
- 18.2 GB Ultra SCSI disk pair (F/C 3803) - requires adapter (F/C 6206)
- 9.1 GB SSA disk pair (F/C 3802) - requires adapter (F/C 6230)
- 9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3805) - requires adapter (F/C 6206)
- 18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3811) - requires adapter (F/C 6206)
- 18.2 GB SSA disk pair (F/C 3812) - requires adapter (F/C 6230)
- 36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3821) - requires adapter (F/C 6206) and an SP Expansion I/O Unit power upgrade (F/C 9955)
- 36.4 GB SSA 10K RPM disk pair (F/C 3822) - requires adapter (F/C 6230) and an SP Expansion I/O Unit power upgrade (F/C 9955)

Note: Empty, unused hard disk drive bay pairs require a filler plate (F/C 9612).

2.8 332 MHz SMP nodes

On April 21, 1998, IBM first introduced the PCI-based nodes, the RS/6000 SP 332 MHz SMP thin nodes and wide nodes. They provide two or four way symmetric multiprocessing (SMP) utilizing PowerPC technology and extend the RS/6000 PCI I/O technology to the SP system.

2.8.1 332 MHz SMP node system architecture

The 332 MHz SMP node is designed as a bus-based symmetric multiprocessor (SMP) using a 64-bit address and a 128-bit data system bus running at a 2:1 processor clock ratio. Attached to the system bus are from one to four PowerPC 604e processors with dedicated, in-line L2 cache/bus converter chips and a two chip memory-I/O controller. Note that the 604e processor only uses a 32-bit address bus. The memory-I/O controller generates an independent, separately clocked mezzanine I/O bridge bus to which multiple chips can be attached to implement various architected I/O buses (for example, PCI) that also operate independently and are separately clocked.

This design partitions all the system logic into a high speed processor-memory portion and a lower speed I/O portion. This has the cost advantage of not having to design I/O bridges to run wide buses at high speeds and it also removes electrical loading on the SMP system bus, which allows that bus to run even faster. The wide, high speed processor-memory bus reduces memory and intervention latency, while the separate I/O bridge bus supports memory coherent I/O bridges on a narrower, more cost effective bus. The memory-I/O controller performs all coherency checking for the I/O on the SMP system bus but relieves the SMP system bus from all I/O data traffic.

Figure 2-6 on page 62 shows the 332 MHz SMP node system architecture block diagram.

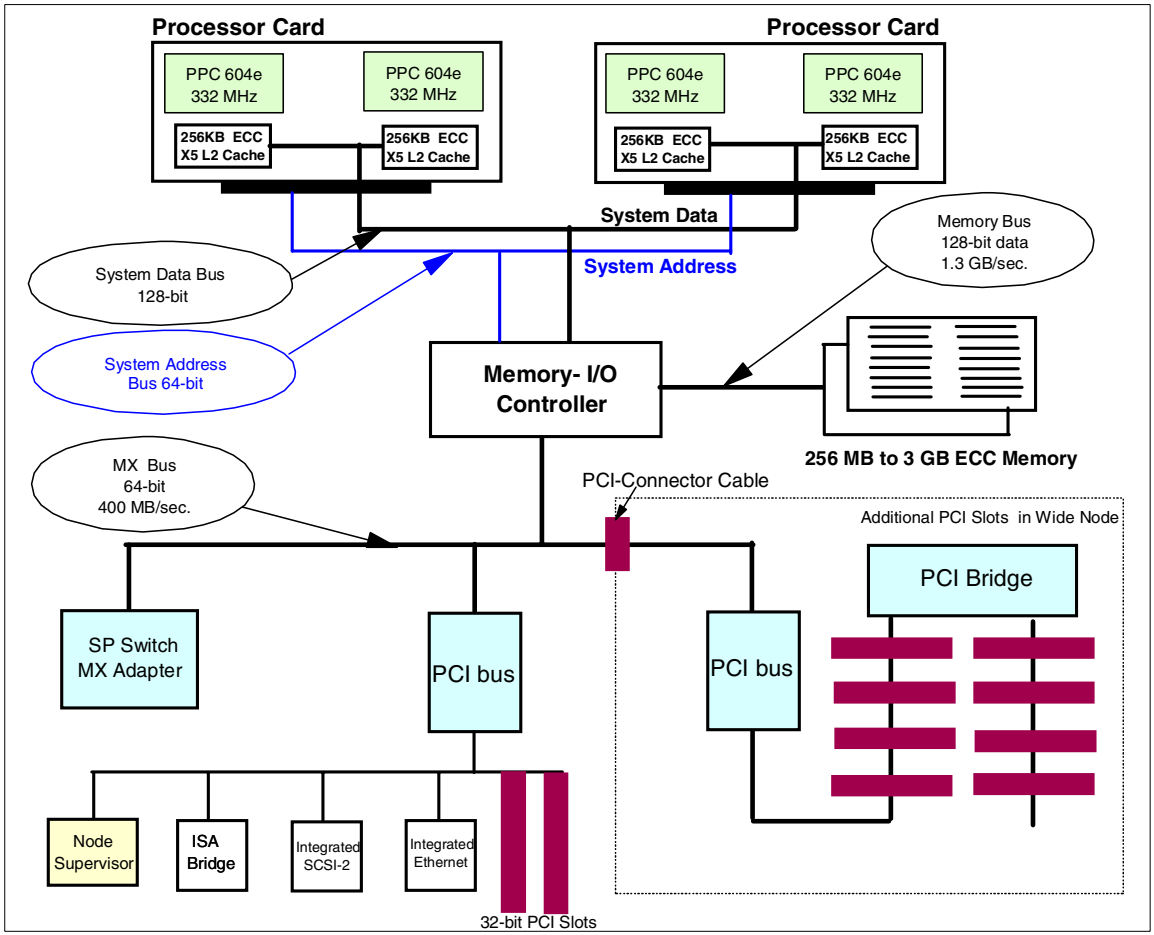


Figure 2-6 332 MHz SMP node system architecture block diagram

2.8.2 SMP system bus

The system bus of the 332 MHz SMP node is optimized for high performance and multiprocessing applications. It has a separate 64-bit + parity address bus and a 128-bit + parity data bus. These buses operate independently in true split transaction mode and are aggressively pipelined. For example, new requests may be issued before previous requests are snooped or completed. There is no sequential ordering requirement, and each operation is tagged with an 8-bit tag, which allows a maximum of up to 256 transactions to be in progress in the

system at any one time. The address bus includes status and coherency response buses for returning flow control, error reports, or coherency information for each request. It can support a new request every other bus cycle at a sustained rate of over 40 million/sec.

The coherency protocol used is an enhanced Modified Exclusive Shared and Invalid (MESI) protocol that allows for cache-to-cache transfers (intervention) on both modified and shared data found in the L2 or L1 caches. The coherency and data transfer burst size is 64 bytes. The bus was designed for glueless attachment using latch-to-latch protocols and fully buffered devices. This enables high system speeds, and the 332 MHZ SMP node can achieve a maximum data rate of 1.3 GB/s when the data bursts are brick-walled (for example, four quad words are transferred every 4 cycles).

The address, data, and tag buses are fully parity checked, and each memory or cache request is range checked and positively acknowledged for error detection. Any error will cause a machine check condition and be logged.

2.8.3 Processor/L2 cache controller

The 332 MHz SMP node X5 Level 2 (L2) cache controller incorporates several technological advancements in design providing greater performance over traditional cache designs. The cache controller is inline and totally contained in one chip.

Integrated on the same silicon die as the controller itself are dual tag directories and 256 KB of SRAM cache. The dual directories allow non-blocking access for processor requests and system bus snoops. These directory arrays are fully parity checked and, in the case of a parity error, the redundant array will be used. The data arrays employ ECC checking and correction and a single bit error can be detected and corrected with no increase in the cache latency. Multiple bit ECC errors will cause a machine check condition. The cache is configured as an 8-way set-associative, dual sectored, 64 byte line cache. Internal design trade-off performance studies have shown that for commercial workloads the miss rate is comparable to a 1MB direct mapped L2 cache with 32 byte lines.

The L2 cache has two independent buses: One 60x bus connecting with the PowerPC 604e microprocessor and one bus connecting to the SMP system bus. The 60x bus (64 data bits) operates at the same speed as the processor core (166 MHz) for a maximum data bandwidth of 1.2 GB/s. The L2 cache core logic runs 1:1 with the processor clock as well and can source L2 data to the processor in 5:1:1:1 processor clock cycles (5 for the critical double word, 1 additional clock for each additional double word). The SMP bus interface (128 data bits) operates at a 2:1 ratio (83 MHz) of the processor bus sustaining the 1.3 GB/s maximum data bandwidth.

Since the L2 cache maintains inclusion with the processor's L1 cache, it filters the majority of other processor snoop transactions, thus, reducing the contention for the L1 cache and increasing processor performance.

332 MHz SMP node X5 cache controller supports both the shared and modified intervention protocols of the SMP bus. If one L2 requests a cache line that another L2 already owns, the memory subsystem is bypassed, and the cache line is directly transferred from one cache to the other. The typical latency for a read on the system bus that hits in another L2 cache is 6:1:1:1 bus cycles compared to a best case memory latency of 14:1:1:1 bus cycles. This will measurably reduce the average memory latency when additional L2 caches are added to the system and accounts for the almost linear scalability for commercial workloads.

2.8.4 Memory-I/O controller

The 332 MHZ SMP node has a high performance memory-I/O controller capable of providing sustained memory bandwidth of over a 1.3 Gigabyte/sec and sustained maximum I/O bandwidth of 400 MBps with multiple active bridges.

The memory controller supports one or two memory cards with up to eight increments of synchronous dynamic RAM (SDRAM) memory on each card. Each increment is a pair of dual in-line memory modules (DIMMs). There are two types of DIMMs supported: 32 MB DIMMs that contain 16 Mb technology chips, or 128 MB DIMMs that contain 64 Mb technology chips. The DIMMs must be plugged in pairs because each DIMM provides 72 bits of data (64 data bits + 8 bits of ECC word), which, when added together, form the 144 bit memory interface. The memory DIMMs used in the 332 MHz SMP node are 100 MHz (10 ns), JEDEC standard non-buffered SDRAMs.

SDRAMs operate differently than regular DRAM or EDO DRAM. There is no interleaving of DIMMs required to achieve high data transfer rates. The SDRAM DIMM can supply data every 10 ns clock cycle once its access time is satisfied, which is 60 ns. However, as with a traditional DRAM, that memory bank is then busy (precharging) for a period before it can be accessed again. So, for maximum performance, another bank needs to be accessed in the meantime. The memory controller employs heavy queuing to be able to make intelligent decisions about which memory bank to select next. It will queue up to eight **`read`** commands and eight **`write`** commands and attempt to schedule all reads to non-busy banks and fill in with writes when no reads are pending. Writes will assume priority if the write queue is full or there is a new read address matching a write in the write queue.

The SDRAM DIMM itself contains multiple internal banks. The 32 MB DIMM has two internal banks, and the 128 MB DIMM has four internal banks. A different non-busy bank can be accessed to keep the memory fully utilized while previously accessed banks are precharging. In general, more banks of memory are better and allow the controller to schedule non-busy banks. So, 32 MB DIMMs with two banks internal will contain more banks than the same amount of memory in 128 MB DIMMs. For example, suppose you wanted 256 MB of memory total. That would be one pair of 128 MB DIMMs with a total of four internal banks or four pairs of 32 MB DIMMs with two internal banks, each for a total of eight banks. Once larger amounts of memory are installed (for example, more than 16 banks), it should not make much difference which DIMMs to use, for there would be plenty of non-busy banks to access. Note that there is no performance advantage to spreading DIMMs across two memory cards. The second card is needed only for expanding beyond a full first card.

Assuming the read requests can be satisfied to non-busy banks, the memory controller is capable of supplying a quad word of data every 83 MHz clock for an aggregate peak bandwidth of 1.32 GB/sec. While the memory bandwidth must be shared with any I/O, the SMP data bus bandwidth does not. This means that concurrent memory or cache-to-cache data can be sent on the SMP data bus while the memory controller is using the I/O bridge bus to transfer data or from I/O. There is also extensive queuing to support the I/O bus, seven command queues to hold outgoing processor requests to I/O, and six command queue's forwarding I/O requests for Direct Memory Access (DMA).

To provide memory coherency for I/O, but allowing the SMP bus to run at maximum speed, the memory-I/O controller acts as a snooping agent. This is accomplished by keeping track of the latest I/O reads in a chip directory. This directory is used whenever a processor request is issued on the SMP bus so that the request does not have to be forwarded to I/O to be snooped. When I/O issues a request, the controller places the request address on the SMP bus to be snooped by the processor caches before it is entered into the memory queues. A cache hit will retry the request and force the data to memory and then the request will be reissued by the memory controller on behalf of the I/O device. It is not rejected back to the device.

2.8.5 I/O bridge bus

The 332 MHz SMP node memory-I/O controller implements a 64-bit + parity, multiplexed address and data bus for attaching several PCI I/O bridges. This bus runs concurrent and independently from the SMP processor bus and the SDRAM memory buses. In the 332 MHz SMP node, this I/O bridge bus is clocked at 50 MHz, and there are three PCI bridges attached. The peak bandwidth of this bus

is 400 MBps, and it is also a split-transaction, fully tagged bus, like the SMP system bus. Multiple requests from each I/O bridge and the processors may be active at any one time, therefore, subject only to the amount of queuing and tags available in the bridges and the memory-I/O controller.

This bus is parity checked, and all addresses are range checked and positively acknowledged. In addition, the memory-I/O controller and the I/O bridges have fault isolation logic to record any detected error so that software can more easily determine the source of an error.

2.8.6 PCI I/O bridges

Each PCI hardware bridge chip attaches to the I/O bridge bus and is separately and concurrently accessible by any of the processors. There are no PCI bridge-bridge chips. All three of the 332 MHz SMP node's PCI buses have their own bridge chip designed to the PCI Version 2.0 Local bus Specification. These PCI bridges contain over 700 bytes of internal data buffering and will support either a 32-bit or 64-bit PCI data bus configuration. The internal data buffer structure of the bridges allows the I/O bridge bus interface functions and the PCI interface functions to operate independently. Processor initiated Load/Store operations may be up to 64 bytes long and are buffered separately from PCI adapter initiated DMA operations. Consequently, acceptance of processor Load/Store operations, which target the PCI bus, only depend upon buffer availability and not the availability of the target bus or completion of DMA operations. Likewise, the bridge will always accept PCI bus initiator DMA read or write operations that target system memory sends if there is an available data buffer and queue available and it does not have to wait for previous reads or writes to be completed. These PCI addresses are translated by the bridge chip using software initialized translation control elements (TCEs) fetched from memory. Each TCE will translate a 4K page of PCI address space.

The bridge design is optimized for large block sequential DMA read and DMA write operations by a PCI bus Initiator. Up to 128 bytes of DMA read data are continuously prefetched into one of four DMA read caches in order to sustain large block transfers. Two outstanding DMA read requests may be active at one time by each PCI bridge. When the transfer is terminated, the data and its TCE remain in the cache until the cache is reassigned or invalidated by a processor write to memory at that cached location. Each bridge also maintains an eight-entry TCE cache for PCI initiator DMA write operations and a 128 byte write-thru DMA buffer. The bridge PCI and system interfaces are parity checked and contain error detection and error reporting registers accessible by software for fault isolation. The combination of high performing, independent I/O bridges and the node split bus architecture yields outstanding I/O performance and scalability.

2.9 332 MHz SMP Wide Nodes (F/C 2051)

This section describes the 332 MHz POWER3 SMP High Nodes also called Silver Wide Node. The 332 MHz SMP wide nodes (F/C 2051) have PCI bus architecture and use either two or four 332 MHz PowerPC processors per node. These nodes are functionally equivalent to an IBM RS/6000 7025-F50 workstation. The SP system must be operating at PSSP 2.4 (or later) to use these nodes.

The 332 MHz SMP wide node occupies one full drawer; therefore, eight SMP wide nodes can be housed in a tall frame. SMP wide nodes can be placed in the first node slot of a frame without requiring additional nodes.

For electromagnetic compliance, these nodes are housed in an SMP enclosure. This enclosure (F/C 9930) is automatically included when you order a 332 MHz SMP wide node.

If you are going to mount a 332 MHz SMP wide node into an older 2.01 m or 1.25 m frame, a power system upgrade is necessary. However, once you have done the power system upgrade, these nodes are fully compatible with all existing SP hardware except for High Performance Switches.

2.9.1 Bus description

The 332 MHz SMP wide node PCI bus is divided into three logical groups of PCI slots. The first slot group (slots I2 and I3) is composed of the two 32-bit slots residing on the CPU side of the 332 MHz SMP wide node, and the second and third group reside on the I/O side of the node. Both the second and third group have four PCI slots each. The second group (slots I1 through I4) has three 64-bit slots and a single 32-bit slot. The third group (slots I5 through I8) is made up of the last four 32-bit slots on the I/O side of the node. The third group is a physical extension on the second group. The I1 slot on the CPU side of the node is reserved for the optional SP Switch MX2 Adapter. Previously installed 332 MHz SMP wide nodes may have a withdrawn SP Switch MX Adapter in the CPU side I1 slot.

2.9.2 Adapter placement restrictions

With few exceptions, the ten PCI slots in the 332 MHz SMP Wide Node can be used for any valid RS/6000 SP PCI system adapter. While most PCI adapters will function in any 332 MHz SMP Wide Node slot, the following adapters cannot be placed in any one of the third group of PCI slots:

- S/390 ESCON (F/C 2751)

- ARTIC960Hx 4-Port selectable (F/C 2947)
- 2-port Multiprotocol (F/C 2962)
- ATM 155 UTP (F/C 2963)
- Gigabit Ethernet - SX (F/C 2969)
- ATM 155 MMF (F/C 2988)
- Ultra SCSI SE (F/C 6206)
- Ultra SCSI DE (F/C 6207)
- SSA RAID5 (F/C 6215)
- ARTIC960RxD Quad Digital Trunk (F/C 6310)

To achieve the best performance with SSA RAID and Ultra SCSI disk subsystems, the following adapters for these devices should be distributed evenly across the two recommended PCI slot groups:

- SSA RAID5 (F/C 6215)
- Ultra SCSI SE (F/C 6206)
- Ultra SCSI DE (F/C 6207)

To avoid performance degradation, the following adapters should not be placed in slots I5, I6, I7, or I8 in 332 MHz SMP Wide Nodes:

- FDDI SK-NET LP SAS (F/C 2741)
- FDDI SK-NET LP DAS (F/C 2742)
- FDDI SK-NET UP SAS (F/C 2743)
- 10/100 MB Ethernet (F/C 2968)
- SCSI-2 F/W single-ended (F/C 6208)
- SCSI-2 F/W differential (F/C 6209)

For similar reasons, if two S/390 ESCON adapters (F/C 2751) are used in this node, one adapter must be placed in the CPU bus, and the other adapter must be placed in the first I/O bus.

2.9.3 Requirements

332 MHz SMP Wide Nodes occupy one full node drawer. These nodes are symmetrically configured for memory, disk, and adapters. Up to eight 332 MHz SMP Wide Nodes may be installed in one tall frame and up to four in a short frame.

The mandatory requirements are:

- PSSP 2.4 (or later) on the control workstation, backup nodes, and processor node
- Two processors (mounted in one slot)
- 256 MB of memory
- 4.5 GB of hard disk drive (with internal booting)
- An upgraded power system on older frames

2.9.4 Options

Each 332 MHz SMP wide node is functionally equivalent to an RS/6000 7025-F50 and has the following options:

- Two processor slots allowing a maximum of four processors per node.
- Two memory slots supporting up to 3 GB of memory.
- Ten PCI slots for communication adapters (seven 32-bit and 3 64-bit).
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter.
- Integrated 10BaseT/10Base2 Ethernet (only one port may be used at one time).
- Four disk bays supporting up to 72.8 GB of disk storage (36.4 GB mirrored).
- Integrated SCSI-2 Fast/Wide.
- Standard Service Processor.
- External 9-pin RS-232 on the planar S2 port (supported only for HACMP serial heartbeat); a 9- to 25-pin converter cable is included with the node:
  - Node-to-node HACMP cable (F/C 3124)
  - Frame-to-frame HACMP cable (F/C 3125)

2.9.5 Processor requirements and options

SMP wide nodes require a minimum of two 332 MHz PowerPC processors mounted on one card. However, you can order an additional processor card (F/C 4320) to configure the node with a total of four CPUs. Table 2-11 provides the processor options for the 332 MHz SMP Wide Nodes.

Table 2-11 332 MHz Wide Node processor options

Feature Code	Description	Minimum per node	Maximum per node
4320	One processor card with two CPUs	1	2

2.9.6 Memory requirements and options

332 MHz SMP wide nodes have two memory cards and require a minimum of 256 MB of memory. These nodes support a maximum of 3 GB of memory. Memory is supplied by 128 MB DIMMs that must be mounted in pairs (256 MB increments).

The memory cards are not required to be configured symmetrically. Each card has the capacity to mount 2 GB of DIMMs; however, only 3 GB are addressable per node. Memory cards and DIMMs are not interchangeable between SMP and non-SMP wide nodes. the following Table 2-12 provides the available memory features for the SMP Wide nodes.

Table 2-12 32 MHz SMP Wide Node memory features

F/C	Description	Minimum Node Requirement	Maximum Features Per Node
4093	Base Memory Card	2	2
4110	256 MB – (2 x 128 MB DIMMs)	1	12
4133	1 GB – (8 x 128 MB DIMMs)	0	3
4402	2 GB – (16 x 128 MB DIMMs)	0	1
4403	3 GB – (24 x 128 MB DIMMs)	0	1

2.9.7 Disk requirements and options

332 MHz SMP wide nodes can have up to four internal disks attached through an integrated SCSI-2 network. The 332 MHz SMP wide node can have either no internal disk (with external booting) or from 4.5 GB to a maximum of 72.8 GB of internal disk storage. External storage devices can be accessed through an optional Ultra SCSI Adapter (F/C 6207) or SCSI-2 adapter (F/C 6209).

Optional direct access storage devices are available as follows:

- 4.5 GB Ultra SCSI disk drive (F/C 2900)
- 4.5 GB Ultra SCSI disk drive pair (F/C 2904)
- 9.1 GB Ultra SCSI disk drive (F/C 2908)
- 9.1 GB Ultra SCSI disk drive pair (F/C 2909)
- 18.2 GB Ultra SCSI disk drive pair (F/C 2918)

This node does not require special cables or adapters to mount internal disks. However, the 332 MHz SMP wide node has an option (F/C 1241) that provides an independent SCSI hookup. It accomplishes the following:

- Eliminates the Hard disk drive controller as a single point of failure during mirroring
- Increases disk performance
- Balances disk loading

The (F/C 1241) option requires either an (F/C 6206) SCSI-2 Ultra/Wide Adapter PCI or an (F/C 6208) SCSI-2 Fast/Wide Adapter 4-A PCI as a PCI-type SCSI adapter.

2.9.8 Switch adapter requirements and options

The switch adapter for SMP wide nodes does not occupy a PCI slot. Instead, the switch adapter for these nodes is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

- In switch-configured systems, 332 MHz SMP wide nodes require the following switch adapter:
  - SP Switch MX2 Adapter (F/C 4023)

332 MHz SMP wide node switch restrictions

The 332 MHz SMP wide node is not compatible with the older High Performance series of switches. If an SMP wide node is going to be placed into an SP system configured with a switch, that switch must be either an SP Switch or an SP Switch-8.

2.10 332 MHz SMP Thin Nodes (F/C 2050)

This section describes the 332 MHz POWER3 SMP Thin Nodes, also called Silver Thin Node. The 332 MHz SMP thin nodes (F/C 2050) have PCI bus architecture and use either two or four 332 MHz PowerPC processors per node. These nodes are functionally equivalent to an IBM RS/6000 7025-F50 workstation. The SP system must be operating at PSSP 2.4 (or later) to use these nodes.

The 332 MHz SMP thin node occupies one half of a drawer and may be installed singly with systems operating at PSSP 3.1 or later. Therefore, up to 16 SMP thin nodes can be housed in a tall frame. When installed singly, 332 MHz SMP thin nodes must be placed in the odd numbered node slot. See Section 2.10.2, “Single SMP thin node configuration rules” on page 72 for details.

For electromagnetic compliance, these nodes are housed in an SMP Enclosure. This enclosure (F/C 9930) is automatically included when you order a 332 MHz SMP Thin Node. For installations using single SMP thin nodes, a cover plate (F/C 9931) is also included to cover the unused enclosure slot.

If you are going to mount a 332 MHz SMP thin node into an older 2.01 m or 1.25 m frame, a power system upgrade is necessary. However, once you have done the power system upgrade, these nodes are fully compatible with all existing SP hardware except for High Performance Switches.

2.10.1 Bus description

The 332 MHz SMP thin node PCI bus contains two 32-bit slots PCI slots (slots I2 and I3). The I1 slot is reserved for the optional SP Switch MX2 Adapter. Previously installed 332 MHz SMP thin nodes may have a withdrawn SP Switch MX Adapter in the I1 slot.

2.10.2 Single SMP thin node configuration rules

With PSSP 3.1, single POWER3 SMP thin nodes and single 332 MHz. SMP thin nodes are allowed in both tall and short frame configurations provided the following rules are observed:

- Single SMP thin nodes must be installed in the odd numbered node position. Single SMP thin nodes are not supported in the even numbered node position.
- Empty node drawers are allowed on tall frames if the frame is either a non-switched frame or configured with an SP Switch (16-port switch).
- Tall frames configured with the SP Switch-8 (8-port switch) must have all nodes placed in sequential order; no empty drawers are allowed. Therefore, the single SMP thin node in these frames is the last node in the configuration.
- Short frame configurations must have all nodes placed in sequential order; no empty drawers are allowed. Therefore, the single SMP thin node in these frames is the last node in the configuration.
- A single POWER3 SMP thin node and a single 332 MHz SMP thin node each occupy one half of a node drawer.

- Single POWER3 SMP thin nodes and single 332 MHz SMP thin nodes may be mixed in a thin node drawer.
- If a frame has more than six single SMP thin nodes installed, that frame will have an uneven weight distribution. You must be careful when moving these frames.

2.10.3 Requirements

332 MHz SMP thin nodes occupy one half of a node drawer. When two SMP thin nodes are placed in one drawer, the nodes may be asymmetrically configured for memory, disk, processor speed, and adapters. Up to sixteen 332 MHz SMP thin nodes may be installed in one tall frame and up to eight in a short frame.

The mandatory requirements are:

- PSSP 2.4 (or later) on the control workstation, backup nodes, and processor node
- Two processors (mounted in one slot)
- 256 MB of memory
- 4.5 GB of hard disk drive (with internal booting)
- An upgraded power system on older frames

2.10.4 Options

Each 332 MHz SMP Thin node is functionally equivalent to an RS/6000 7025-F50 and has the following options:

- Two processor slots allowing a maximum of four processors per node.
- Two memory slots supporting up to 3 GB of memory.
- Two disk bays supporting up to 36.4 GB of storage (18.2 GB mirrored).
- A dedicated Mezzanine bus (MX) slot for an optional switch adapter.
- Two 32-bit PCI slots for communication adapters.
- Integrated 10BaseT/10Base2 Ethernet (only one port may be used at one time).
- Integrated SCSI-2 Fast/Wide.
- Standard Service Processor.
- External nine-pin RS-232 on the planar S2 port (supported only for HACMP serial heartbeat); a 9 to 25-pin converter cable is included with the node
  - Node-to-node HACMP cable (F/C 3124)

– Frame-to-frame HACMP cable (F/C 3125)

2.10.5 Processor requirements and options

SMP thin nodes require a minimum of two 332 MHz PowerPC processors mounted on one card. However, you can order an additional processor card (F/C 4320) to configure the node with a total of four CPUs. Table 2-13 provides the processor options for the 332 MHz SMP Thin Nodes.

Table 2-13 332 MHz SMP Thin Node processor option

Feature Code	Description	Minimum per node	Maximum per node
4320	One processor card with two CPUs	1	2

2.10.6 Memory requirements and options

332 MHz SMP Thin Nodes require two memory cards, a minimum of 256 MB of memory, and support a maximum of 3 GB. Memory is supplied by 128 MB DIMMs, which must be mounted in pairs. The memory cards are not required to be configured symmetrically.

Each card has the capacity to mount 2 GB of DIMMs; however, only 3 GB are addressable per node. Table 2-14 provides the available memory features for the 332 MHz SMP Thin Nodes.

Table 2-14 332 MHz SMP Thin Node memory features

F/C	Description	Minimum Node Requirement	Maximum Features Per Node
4093	Base Memory Card	2	2
4110	256 MB – (2 x 128 MB DIMMs)	1	12
4133	1 GB – (8 x 128 MB DIMMs)	0	3
4402	2 GB – (16 x 128 MB DIMMs)	0	1
4403	3 GB – (24 x 128 MB DIMMs)	0	1

2.10.7 Disk requirements and options

332 MHz SMP thin nodes can have up to two internal disks attached through an integrated SCSI-2 network. The 332 MHz SMP thin node can have either no internal disk (with external booting) or from 4.5 GB to a maximum of 36.4 GB of internal disk storage (18.2 mirrored).

External storage devices can be accessed through an optional Ultra SCSI Adapter (F/C 6207) or SCSI-2 Adapter (F/C 6209) and SSA adapter (F/C 6230).

Optional direct access storage devices are available as follows:

- 4.5 GB Ultra SCSI disk drive (F/C 2900)
- 4.5 GB Ultra SCSI disk drive pair (F/C 2904)
- 9.1 GB Ultra SCSI disk drive (F/C 2908)
- 9.1 GB Ultra SCSI disk drive pair (F/C 2909)
- 18.2 GB Ultra SCSI disk drive pair (F/C 2918)

This node does not require special cables or adapters to mount an internal disk.

2.10.8 Switch adapter requirements and options

The switch adapter for SMP thin nodes does not occupy a PCI slot. Instead, the switch adapter for these nodes is installed into the Mezzanine bus (MX). The MX bus connects the I/O planar with the system planar. Placing the switch adapter in the MX bus enables switch traffic to proceed at higher bandwidths and lower latencies.

In switch configured systems, 332 MHz SMP thin nodes require the following switch adapter:

- SP Switch MX2 Adapter (F/C 4023)

332 MHz SMP thin node switch restrictions

- The 332 MHz SMP Thin Node is not compatible with the older High Performance series of switches.
- If a 332 MHz SMP Thin Node is to be placed into an SP system configured with a switch, that switch must be either an SP Switch or an SP Switch-8.
- Switch adapters for 332 MHz SMP Thin Nodes are not interchangeable with switch adapters used on uniprocessor wide nodes.

3

SP-attached servers and Clustered Enterprise Servers

The RS/6000 SP system, as a cluster of SP nodes, managed by the AIX operating system and PSSP management software from the Control Workstation (CWS), is the path to universal clustering.

Managed by the AIX and PSSP from the CWS, @server pSeries and RS/6000 Enterprise Servers, can function as SP-attached servers on an SP system.

Also managed by the AIX and PSSP from CWS, all the possible SP-attached servers to an SP system can be defined as Clustered Enterprise Servers (CES) on a Clustered @server pSeries system (CES system). The CES systems contain pSeries systems and RS/6000 systems only, connected to the CWS.

Any of these SP universal cluster systems (SP/CES systems) can function in switched or non-switched configurations.

3.1 Overview

The RS/6000 SP system, as a cluster of SP nodes and managed by the AIX operating system and PSSP management software from the Control Workstation, is the path to universal clustering.

We may consider SP systems with SP-attached servers. We may consider also Clustered Enterprise Server systems.

The SP-attached server is an @server pSeries or an RS/6000 Enterprise Server, supported to be configured to operate with an RS/6000 SP system.

The Table 3-1 shows the list with the @server pSeries and RS/6000 Enterprise Servers supported as SP-attached server.

Table 3-1 Servers supported as SP-attached servers

Server name	Machine Type (M/T)	Model	Availability
@server pSeries 680 Model S85	7017	S85	Yes
RS/6000 Enterprise Server 7017 Model S80	7017	S80	Yes
RS/6000 Enterprise Server 7017 Model S7A	7017	S7A	No
RS/6000 Enterprise Server 7017 Model S70	7017	S70	No
RS/6000 Enterprise Server 7026 Model M80	7026	M80	Yes
@server pSeries 660 Model 6M1	7026	6M1	Yes
@server pSeries 660 Model 6H1	7026	6H1	Yes
@server pSeries 660 Model 6H0	7026	6H0	Yes
RS/6000 Enterprise Server 7026 Model H80	7026	H80	Yes

Up to 16 of the supported M/T 7017 Models and up to 32 of the supported M/T 7026 Models, up to 32 mixed servers in total, can be integrated on an SP system as SP-attached servers.

The Clustered Enterprise Server (CES) is an @server pSeries or an RS/6000 Enterprise Server, supported to be configured to operate with a Clustered @server pSeries system (CES system).

All the possible SP-attached servers to an SP system, can be defined as Clustered Enterprise Servers on a CES system. The CES systems contain pSeries systems and RS/6000 systems only, connected to the CWS.

Up to 16 of the supported M/T 7017 Models and up to 32 of the supported M/T 7026 Models, up to 32 mixed servers in total, can be integrated on a CES system as Clustered Enterprise Servers.

Like an SP node, the SP-attached server and the Clustered Enterprise Server can perform most SP processing and administration functions. However, unlike an SP node, the SP-attached server and the Clustered Enterprise Server are housed in their own frame. Thus, the SP-attached server and the Clustered Enterprise Server have both node-like and frame-like characteristics.

These servers are compatible with SP Switch (F/C 4011). They need to be configured with the SP System Attachment Adapter (F/C 8396).

3.2 SP-attached servers / CES - M/T 7017 supported models

This section describes the M/T 7017 Models which are supported as SP-attached servers and Clustered Enterprise Servers.

3.2.1 @server pSeries 680 Model S85 (7017-S85)

The 7017-S85 server is the latest member of the popular S-family of 64-bit enterprise servers from IBM. The 7017-S85 server blends copper silicon-on-insulator technology and larger memory capacity with a proven system architecture to give you reliability you can count on. Scalability is enhanced by increasing the amount of memory to 96 GB by leveraging the performance enhancements of AIX Version 4.3.3 and the options of server clustering or SP attachment within your enterprise. With the 7017-S85 server, you can manage the evolution of your business into 64-bit computing while still supporting your existing 32-bit applications. In addition, the I/O subsystem supports 32-bit and 64-bit standard PCI adapters. Combined with an outstanding array of industry software and services, the 7017-S85 server delivers the power, capacity, reliability, and expandability to help move you into the mission-critical commercial computing.

The 7017-S85 server is further enhanced with the availability of Capacity Upgrade on Demand (CUoD) that allows you to have inactive processors installed in your system and can be made active quickly and easily as your business needs require. When more processing capacity is required, you simply issue an AIX command to increase the number of active processors in increments of two, four, or six. The processors become active at the next system re-boot. The 7017-S85 server periodically checks the number of active

processors, determines that the number of processors has been increased and notifies IBM. The customer is responsible for placing an order for the additionally activated processors with their IBM or IBM Business Partner sales representative.

The 7017-S85 server is packaged as a Central Electronics Complex (CEC) and an I/O rack. The 7017-S85 CEC entry configuration starts with a 4-way scalable SMP system that utilizes the 64-bit, 600 MHz, RS64 IV processor with 16 MB of Level 2 (L2) cache per processor. The 4-way SMP can be expanded to a 24-way SMP, and the system memory can be expanded to 96 GB. The I/O rack contains the first I/O drawer with up to three additional I/O drawers available.

The 7017-S85 is shipped and delivered with all the internal adapters and devices already installed and configured.

Highlights

The 7017-S85 delivers a cost-efficient growth path to the future through these attributes:

- 64-bit system scalability
  - 4-way, 6-way, 12-way, 18-way, or 24-way -- packaged on four processor cards (six RS64 IV processors per card)
  - 600 MHz system processors
  - 16 MB of L2 ECC cache for each 600 MHz processor
- Expandability and reliability
  - System memory expandable from 4 GB to 96 GB of ECC SDRAM
  - Multiple I/O racks can be attached
  - Up to four pSeries 680 I/O drawers per server
  - Each I/O drawer supports 14 of 33 MHz PCI bus slots (five 64-bit and nine 32-bit) and 12 disks that can be hot plugged
- Capacity Upgrade on Demand for processor growth
- Ultra SCSI, SSA, and FCAL interfaces supported
- 32-bit and 64-bit applications support
- AIX license included, can be preinstalled

Description

The 7017-S85 provides performance and supports the evolution of your enterprise to 64-bit computing. Its benefits include:

- Commercial performance in 4-, 6-, 12-, 18-, or 24-way enterprise servers using the RS64 IV microprocessor technology.
- Capacity Upgrade on Demand which allows additional processors to be easily added to the pSeries 680 as your business expands.
- Reliability, availability, and serviceability (RAS) providing:
  - Error checking and correcting (ECC) system memory and cache
  - I/O link failure recovery
  - Disk drives that can be hot-swapped
  - Environmental sensing
  - N+1 Power
  - N+1 Cooling units
  - Service processor for integrated system monitoring
  - Concurrent diagnostics
- Preconfigured tailored solutions that take advantage of the rack packaging to custom tailor your solution. Choose from multiple SCSI and SSA:
  - Storage
  - Media
  - Communication
  - Drawer options
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Enhanced processor speeds and memory growth. The minimum configuration, 4-way SMP processors can be expanded up to 24-way. Memory can be expanded up to 96 GB in 2 GB, 4 GB, 8 GB, 16 GB, or 32 GB increments. Industry-standard PCI slots can be expanded up to 56 total. System power supplies are designed for either 12-way or 24-way configurations.
- Industry recognized AIX operating system license with each system. New performance enhancements and a new Workload Management feature are also included with the latest release.
- Internal I/O system scalability and expandability that supports:
  - Up to four system I/O drawers with up to 56 PCI slots
  - Up to 873.6 GB internal disks
  - Eight media bays for tape and optical devices

- Industry-standard PCI subsystem supporting 32-bit and 64-bit adapters. Greater application throughput is achieved by the four independent PCI buses per rawer, providing up to 500 MB per second aggregate bandwidth per drawer.
- A high-performance ECC L2 cache of 16 MB per 600 MHz processor.
- RAID support SSA. The operating system provides RAID 0 and 1. The optional SSA RAID adapter provides RAID level 0, 1, and 5.

  Figure 3-1 illustrates the system architecture for 7017-S85 server.

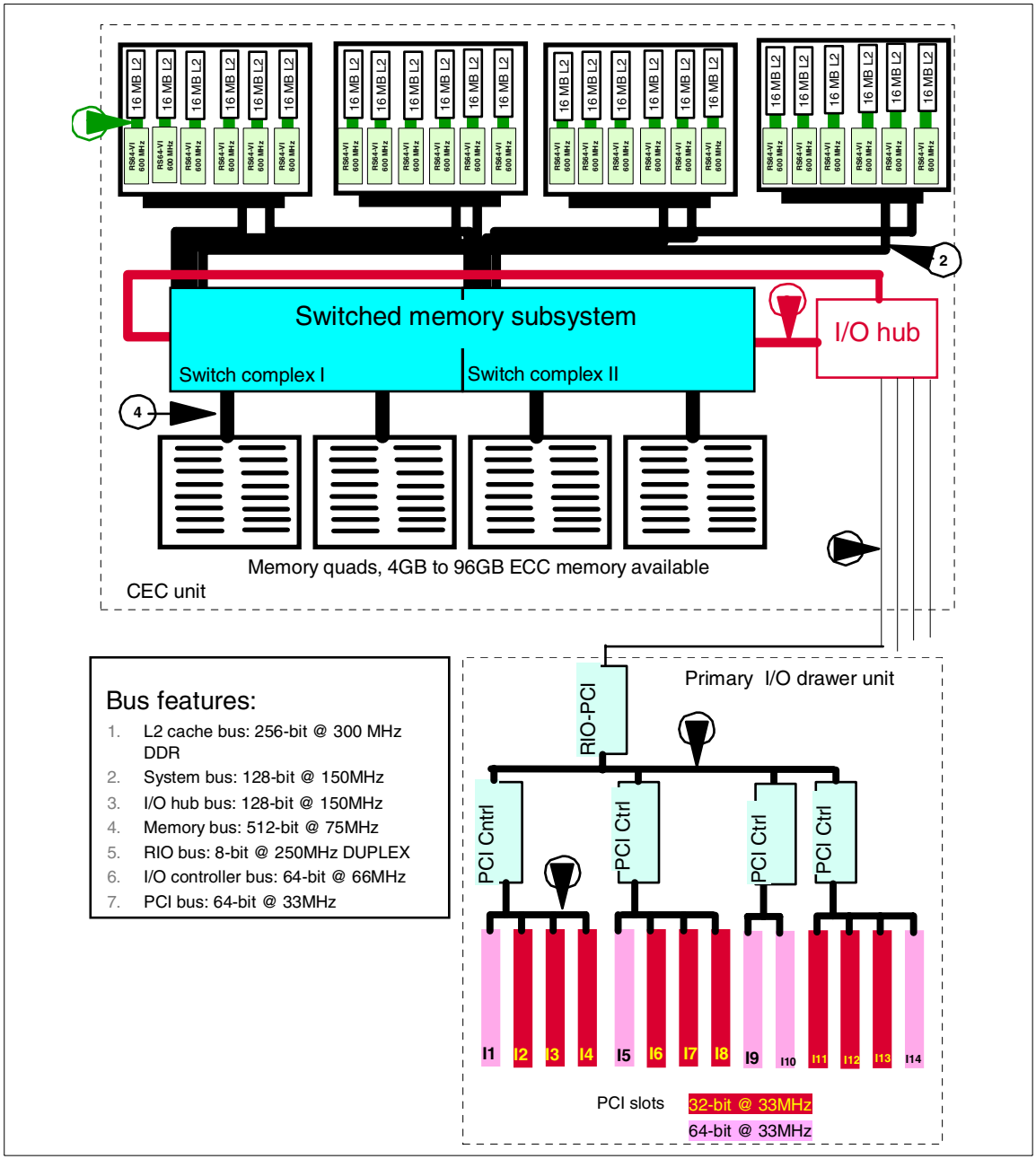


Figure 3-1 System architecture for 7017-S85 server

Product positioning

The 7017-S85 server has design and application performance that makes it a solution for mission-critical ERP/SCM, Data Warehouse/DataMart, OLTP, and e-business applications. It offers more performance, memory, and storage, and can run your current 32-bit applications alongside new 64-bit solutions.

It has the capacity to meet your current and future needs with:

- Choice of power from 4, 6, 12, 18, or 24 processors
- Extensive memory options up to 96 GB
- Up to 56 PCI expansion slots
- Up to 45 TB of disk storage

The 7017-S85 is designed from the ground up for mission-critical environments with a service processor, redundant cooling, and redundant power. In addition, you can cluster the highly reliable 7017-S85 with High Availability Cluster Multi-Processing (HACMP) for AIX (separately available) to achieve even greater availability and horizontal growth. The 7017-S85 performance, expansion, and reliability represent a combination of features for handling today's complex mission-critical applications.

Note: The IBM pSeries 680 Model S85 (7017-S85) is now available with 450 MHz RS64 III processors as well. Model S85 configurations equipped with 450 MHz processors are available in 6-, 12-, 18-, and 24-way configurations. If additional processing power is needed, customers can replace their 450 MHz processors with 600 MHz processors. Capacity Upgrade on Demand (CUoD) is available on Model S85 systems equipped with either the 450 MHz RS64 III or 600 MHz RS64 IV processors. CUoD allows you to have inactive processors installed in your system, which can be made active quickly and easily as your business needs require. AIX Version 4.3.3 with 4330-06 maintenance package, or later, or AIX 5L Version 5.1, or later.

For more information on this model, refer to the redbooks *RS/6000 Systems Handbook 2000 Edition*, SG24-5120 and *IBM eserver pSeries 680 Handbook including RS/6000 Model S80*, SG24-6023.

Table 3-2 provides the publications for the 7017-S85 server:

Table 3-2 Publications for the 7017-S85 server

Order Number	Title
SA38-0582	Enterprise Server S80 pSeries 680 Model S85 Installation Guide
SA38-0558	Enterprise Server S80 pSeries 680 Model S85 Service Guide

Order Number	Title
SA23-0557	RS/6000 Enterprise Server S80 pSeries S85 User's Guide
SA38-0538	PCI Adapter Placement Reference
SA38-0516	Adapters, Devices, and Cable Information for Multiple bus systems
SA38-0509	Diagnostic Information for Multiple bus systems

The publications are available for download from the Internet:

`http://www.ibm.com/servers/eserver/pseries/library/`

3.2.2 RS/6000 Enterprise Server 7017 Model S80 (7017-S80)

The 7017-S80 server is the member of the popular S-family of 64-bit enterprise servers from IBM. The 7017-S80 blends new copper technology and enhanced backplane speeds with a proven system architecture to provide reliability you can count on. Scalability is additionally enhanced by doubling the number of processors to 24, the amount of memory to 96 GB, and by leveraging the new performance enhancements of AIX Version 4.3.3. With the 7017-S80, you can manage the evolution of your business into 64-bit computing while still supporting your existing 32-bit applications. In addition, the I/O subsystem supports 32-bit and 64-bit standard PCI adapters. Combined with an outstanding array of industry software and services, the 7017-S80 provides the power, capacity, reliability, and expandability to help move you into the mission-critical commercial computing.

The 7017-S80 is packaged as a Central Electronics Complex and an I/O rack. The 7017-S80 Central Electronics Complex entry configuration starts with a 6-way scalable SMP system that utilizes the 64-bit, 450 MHz, RS64 III processor with 8 MB of Level 2 (L2) cache per processor. The 6-way SMP can be expanded to a 24-way SMP and the system memory can be expanded to 96 GB. Customers with installed 7017-S80 systems can increase their processing power by replacing their existing processors with 600 MHz RS64 IV processors.

The I/O rack contains the first I/O drawer with:

- Service processor
- High-performance disk drive
- Optical storage device (a CD-ROM readable device)
- 1.44 MB, 3.5-inch diskette drive
- Two PCI disk adapters

Up to three additional 7017-S80 I/O drawers can be added. Additional I/O racks can also be ordered with the7017-S80. Existing RS/6000 7015 Model R00 racks or the 7014 Model S00 rack can also be used for additional storage and communication drawers. This helps to protect your existing investment in SSA or SCSI DASD.

A fully configured system would be a 24-way processor with:

- 96 GB of system memory
- 56 PCI adapter slots (three of which are reserved for the service processor and the two required Ultra SCSI adapters)
- 48 disk bays that can be hot plugged
- Eight media bays (one of which is reserved for an optical storage device capable of reading CD-ROM discs).

The remaining space in the I/O racks can be used for various storage and communication subsystems.

The RS/6000 Enterprise Server 7017-S80 is shipped and delivered with all the internal adapters and devices already installed and configured. AIX Version 4.3.3 software is included with every 7017-S80 and can be preinstalled, if desired.

Highlights

The 7017-S80 delivers a cost-efficient growth path to the future through these attributes:

- 64-bit system scalability
  - 6-, 12-, 18-, or 24-way processors, which are packaged on four processor cards (six RS64 III processors per card)
  - 450 MHz system processors
  - 8 MB of L2 ECC cache for each 450 MHz processor
- Expandability and reliability
  - System memory expandable from 2 GB to 96 GB of ECC SDRAM
  - Multiple I/O racks can be attached
  - Up to four Model S80 I/O drawers per server
  - Each I/O drawer supports 14 33 MHz PCI bus slots (five 64-bit and nine 32-bit) and 12 disks that can be hot plugged
- Support for Ultra SCSI, SSA, and FCAL interfaces
- Support for 32-bit and 64-bit applications
- AIX license included, can be preinstalled

Description

The 7017-S80 provides performance and supports the evolution of your enterprise to 64-bit computing. Its outstanding benefits include:

- Commercial performance in 6-, 12-, 18-, or 24-way enterprise servers using the RS64 III microprocessor technology.
- Reliability, availability, and serviceability (RAS) providing:
  - Error checking and correcting (ECC) system memory and cache
  - I/O link failure recovery
  - Disk drives that can be hot-swapped
  - Environmental sensing
  - N+1 Power
  - N+1 Cooling units
  - Service processor for integrated system monitoring
  - Concurrent diagnostics
- Preconfigured tailored solutions that take advantage of the rack packaging to custom tailor your solution. Choose from multiple SCSI and SSA:
  - Storage
  - Media
  - Communication
  - Drawer options
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Processor and memory growth beyond any previous RS/6000 Enterprise Server. The base configuration, 6-way SMP processors can be expanded up to 24-way. Memory can be expanded up to 96 GB in 1 GB, 2 GB, 4 GB, 8 GB, 16 GB, or 32 GB increments. Industry-standard PCI slots can be expanded up to 56 total. System power supplies are designed for either 12-way or 24-way configurations.
- Industry recognized AIX operating system (one to two user) license included with each system. New performance enhancements and a new Workload Management feature are included with the latest release.
- Internal I/O system scalability and expandability that supports:
  - Up to four system I/O drawers with up to 56 PCI slots
  - Up to 873.6 GB internal disks
  - Eight media bays for tape and optical devices

- Industry-standard PCI subsystem supporting 32-bit and 64-bit adapters. Greater application throughput is achieved by the four independent PCI buses per drawer, providing up to 500 MB per second aggregate bandwidth per drawer.
- A high-performance ECC L2 cache of 8 MB per 450 MHz processor.
- RAID support SSA. The operating system provides RAID 0 and 1. The optional SSA RAID adapter provides RAID level 0, 1, and 5.

Product positioning

The 7017-S80 server has design and application performance that makes it a solution for mission-critical ERP/SCM, Data Warehouse/Data Mart, OLTP, and e-business applications. In addition, the 7017-S80 is the natural growth path for the Models S70 and S70 Advanced (S7A). It offers more performance, memory, and storage, and can run your current 32-bit applications alongside new 64-bit solutions.

It has the capacity to meet your current and future needs with:

- Choice of power from 6, 12, 18, or 24 processors
- Extensive memory options up to 96 GB
- Up to 56 PCI expansion slots
- Up to 45 TB of disk storage

The 7017-S80 server is designed from the ground up for mission-critical environments with a service processor, redundant cooling, and redundant power. In addition, you can cluster the highly reliable 7017-S80 with High Availability Cluster Multi-Processing (HACMP) for AIX (separately available) to achieve even greater availability and horizontal growth. The 7017-S80 performance, expansion, and reliability represent a combination of features for handling today's complex mission-critical applications.

For more information on this model, refer to the redbooks *RS/6000 Systems Handbook 2000 Edition*, SG24-5120 and *IBM eserver pSeries 680 Handbook including RS/6000 Model S80*, SG24-6023.

Table 3-3 provides the publications for the 7017-S80 server.

Table 3-3 Publications for the 7017-S80 server

Order Number	Title
SA38-0558	Enterprise Server S80 pSeries 680 Model S85 Installation Guide
SA23-0557	Enterprise Server S80 pSeries 680 Model S85 User's Guide
SA38-0538	PCI Adapter Placement Reference

Order Number	Title
SA38-0516	Adapters, Devices, and Cable Information for Multiple bus systems
SA38-0509	Diagnostic Information for Multiple bus systems

The publications are available for download from the Internet:

`http://www-1.ibm.com/servers/eserver/pseries/library/`

3.2.3 RS/6000 Enterprise Server 7017 Model S7A (7017-S7A)

The 7017-S7A (S70 Advanced) server is a member of the S- family - the generation of 64-bit, 4-way, 8-way, or 12-way, symmetric multiprocessing (SMP) enterprise servers from IBM. The 7017-S7A provides the power, capacity, reliability, and expandability to help move you into the mission-critical commercial computing. With the 7017-S7A, you can manage the evolution of your business into 64-bit computing while still supporting your existing 32-bit applications. The I/O subsystem will support 32-bit and 64-bit standard PCI adapters.

Highlights

The 7017-S7A delivers a cost-efficient growth path to the future through these attributes:

- 64-bit system scalability
  - 4-, 8-, or 12-way processors, which are packaged on three processor cards (four RS64 II processors per card)
  - 262 MHz system processors
  - 8 MB of L2 ECC cache for each 262 MHz processor
- Expandability and reliability
  - System memory expandable from 1 GB to 32 GB of ECC SDRAM
  - Multiple I/O racks can be attached
  - Up to four S70 Advanced I/O drawers per server
  - Each I/O drawer supports 14 33 MHz PCI bus slots (5 64-bit and 9 32-bit) and 12 hot-pluggable disks
- Support for Ultra SCSI and SSA interfaces
- Support for 32-bit and 64-bit applications
- AIX license included, can be preinstalled

Description

The 7017-S7A is packaged as a Central Electronics Complex and an I/O rack. The 7017-S7A Central Electronics Complex entry configuration starts with a 4-way scalable SMP system that utilizes the 64-bit, 262 MHz, RS64 II processor with 8 MB of Level 2 (L2) cache per processor. The 4-way SMP can be expanded to a 12-way SMP and the system memory can be expanded to 32 GB.

The I/O rack contains the first I/O drawer with:

- Service processor
- High-performance disk drive
- 32X Maximum speed CD-ROM
- 1.44 MB, 3.5-inch diskette drive
- Two PCI disk adapters

Up to three additional 7017-S7A drawers can be added. Additional I/O racks can also be ordered with the 7017-S7A. Existing RS/6000 7015 Model R00 racks or the 7014 Model S00 rack can also be used for additional storage and communication drawers. This helps to protect your existing investment in SSA or SCSI DASD.

A fully configured system would be a 12-way processor with:

- 32 GB of system memory
- 55 available PCI adapter slots
- 48 hot-pluggable disk bays
- Seven available media bays

The remaining space in the I/O racks can be used for various storage and communication subsystems.

The 7017-S7A server is shipped and delivered with all the internal adapters and devices already installed and configured. AIX Version 4.3.2 software is included with every 7017-S7A and can be preinstalled, if desired.

The 7017-S7A provides performance and supports the evolution of your enterprise to 64-bit computing. Its outstanding benefits include:

- Commercial performance in 4-, 8-, or 12-way enterprise servers, using the RS64 II microprocessor technology.
- Reliability, availability, and serviceability (RAS) providing:
  - Error checking and correcting (ECC) system memory and cache
  - I/O link failure recovery

  - Hot-swappable disk drives
  - Environmental sensing
  - N+1 Power
  - Service processor for integrated system monitoring
  - Concurrent diagnostics
- Preconfigured tailored solutions that take advantage of the rack packaging to custom tailor your solution. Choose from multiple SCSI and SSA:
  - Storage
  - Media
  - Communication
  - Drawer options
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Processor and memory growth beyond any previous RS/6000 Enterprise Server. The base configuration, 4-way SMP processors can be expanded up to 12-way. Memory can be expanded up to 32 GB in 512 MB, 1 GB, 2 GB, 4 GB, or 8 GB increments. System power supplies are designed for maximum configurations.
- Internal I/O system scalability and expandability that supports:
  - Up to four system I/O drawers with up to 56 PCI slots
  - Up to 436.8 GB internal disks
  - Eight media bays for tape and optical devices
- Robust industry-standard PCI subsystem supporting 32-bit and 64-bit adapters. Greater application throughput is achieved by the four independent PCI buses per drawer, providing up to 500 MB per second aggregate bandwidth per drawer.
- A high-performance ECC L2 cache of 8 MB per 262 MHz processor.

RAID support SSA. The operating system provides RAID 0 and 1. The optional SSA RAID adapter provides RAID level 0, 1, and 5.

Product positioning

The 7017-S7A is the natural growth path for the Models R30, R40, and R50. It offers more performance, memory, and storage, and has the added advantage of running your current 32-bit applications alongside new 64-bit solutions.

The 7017-S7A has the capacity to meet your current and future needs with:

- Choice of power from 4, 8, or 12 processors
- Extensive memory options up to 32 GB
- Up to 55 available expansion slots
- Up to 14 TB of disk storage

The 7017-S7A server is designed from the ground up for mission-critical environments with a service processor, redundant cooling, and redundant power. In addition, you can cluster the highly reliable 7017-S7A with High Availability Cluster Multi-Processing (HACMP) for AIX (separately available) to achieve even greater availability and horizontal growth. The 7017-S7A all adds up to the performance, expansion, and reliability to handle the most complex mission-critical, e-business applications.

For more information on this model, refer to the redbook *RS/6000 Systems Handbook 2000 Edition*, SG24-5120.

Table 3-4 provides the publications for the 7017-S7A server:

Table 3-4 Publications for the 7017-S7A server

Order Number	Title
SA38-0548	RS/6000 Enterprise Server Model S70 and S7A Installation Guide
SA23-0549	RS/6000 Enterprise Server S70 and S7A User's Guide
SA38-0538	PCI Adapter Placement Reference
SA38-0516	Adapters, Devices, and Cable Information for Multiple bus systems
SA38-0509	Diagnostic Information for Multiple bus systems

The publications are available for download from the Internet:

`http://www.ibm.com/servers/eserver/pseries/library/`

3.2.4 RS/6000 Enterprise Server 7017 Model S70 (7017-S70)

The 7017-S70 server is the first member of the S-family, which is the generation of 64-bit, 4-way to 12-way symmetric multiprocessing (SMP) enterprise servers from IBM. The 7017-S70 provides the power, capacity, reliability, and expandability to help move you into the mission critical commercial computing. With the 7017-S70, you can manage the evolution of your business into 64-bit computing while still supporting your existing 32-bit applications. The I/O subsystem will support both 32-bit and 64-bit standard PCI adapters.

Highlights

This system is the first 64-bit-ready member of the Enterprise Server Family. It delivers a cost-efficient growth path to the future through these attributes:

- 4-way to 12-way, 64-bit, 125 MHz system scalability
  - Packaged on three processor cards (four RS64 processors per card)
  - Each processor includes 4 MB of L2 ECC cache
- Support for up to 32 GB ECC SDRAM of system memory (512 MB base memory)
- Expandability and reliability:
  - Multiple I/O racks can be attached
  - Up to four S70 I/O drawers per server
  - Each I/O drawer supports fourteen 33 MHz PCI slots, five 64-bit and nine 32-bit 33 MHz PCI slots on its buses and 12 hot-pluggable disks
- AIX for Servers license included
- Support for SCSI and SSA interfaces
- Support for both 32-bit and 64-bit applications

Description

The 7017-S70 is packaged as a Central Electronics Complex and an I/O rack. The 7017-S70 Central Electronics Complex base configuration starts with a 4-way scalable symmetrical multiprocessing system (SMP) that utilizes the new 64-bit 125 MHz RS64 processor with 4 MB of Level 2 (L2) cache per processor. The standard 4-way SMP can be expanded to 12-way SMP; the system memory can be expanded to 32 GB. The I/O rack contains the base I/O drawer with the service processor, a 4.5 GB Ultra-SCSI-enabled, high-performance disk drive, a 20X Max speed CD-ROM, a 1.44 MB, 3.5-inch diskette drive, and two SCSI Fast-Wide PCI adapters. Up to three additional 7017-S70 I/O drawers may be added to the I/O racks. Additional I/O racks may also be ordered with the 7017-S70. Existing RS/6000 7015 Model R00 racks or the new 7014 Model S00 rack may also be used for additional storage and communication drawers. This helps to protect the customers existing investment in SSA or SCSI DASD.

A fully configured system would be a 12-way processor with 32 GB of system memory, 55 33MHz PCI adapter slots, 48 hot-pluggable SCSI disk bays, and up to 11 media bays. The remaining space in the I/O racks can be used for various storage and communication subsystems.

The 7017-S70 is shipped and delivered with all the internal adapters and devices already installed and configured. AIX Version 4.3 software is included with every 7017-S70 and may be pre- installed, if desired.

The 7017-S70 provides outstanding performance and supports the evolution of your enterprise to 64-bit computing. Its benefits include:

- Commercial performance in 4-way to 12-way enterprise servers that use the new RS64 Microprocessor Technology.
- Reliability, availability, and serviceability (RAS) provided by error checking and correcting (ECC) system memory and cache, I/O link failure recovery, hot-swappable disk drives, environmental sensing, N+1 Power, service processor for integrated system monitoring, and concurrent diagnostics.
- AIX, rated the best UNIX operating system by noted industry consultant D.H. Brown. Over 10,000 AIX applications are available that address many commercial and technical needs.
- Preconfigured tailored solutions. Take advantage of the rack packaging to custom-tailor your solution. Choose from multiple SCSI and SSA storage, media, communication, and drawer options.
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Processor and memory growth beyond any previous Enterprise Server. The base-configuration 4-way SMP processors can be expanded up to 12-way. Memory can be expanded up to 32 GB in 512 MB, 1 GB, 2 GB, 4 GB, or 8 GB increments. System power supplies are designed for maximum configurations.
- "Under the covers" I/O system scalability and expandability supporting up to four system I/O drawers, each system providing up to 56 PCI slots, up to 436.8 GB internal disks, and 12 media bays for tape and optical devices.
- Industry standard PCI subsystem, supporting both 32-bit and 64-bit adapters. Greater application throughput is achieved by the four independent PCI buses per drawer, providing up to 500 MB per second aggregate bandwidth per drawer.
- A high-performance ECC L2 cache of 4 MB per RS64 processor.
- RAID support is provided for SCSI or SSA. The operating system provides RAID 0 and 1. The optional SCSI RAID or SSA RAID adapters provide RAID level 0, 1, and 5.

Product positioning

The 7017-S70 is the natural growth path for the Models R30, R40, and R50. It offers more performance, memory, and storage, and the added advantage of running your current 32-bit applications alongside new 64-bit solutions.

It has the capacity to meet your current and future needs with a choice of power from four to 12 processors, extensive memory options to 32 GB, up to 54 expansion slots and 14 TB of disk storage.

The 7017-S70 is designed from the ground up for mission-critical environments with a service processor, redundant cooling, and optional redundant power supply. In addition, you can cluster the highly reliable 7017-S70 with High Availability Cluster Multi-Processing (HACMP) for AIX to achieve even greater availability and horizontal growth. The 7017-S70 all adds up to the performance, expansion, and reliability to handle the most complex mission-critical e-business applications.

Table 3-5 provides the publications for the 7017-S70 server:

Table 3-5 Publications for the 7017-S70 server

Order Number	Title
SA38-0548	RS/6000 Enterprise Server Model S70 and S7A Installation and Service Guide
SA23-0549	RS/6000 Enterprise Server S70 and S7A User's Guide
SA38-0538	PCI Adapter Placement Reference
SA38-0516	Adapters, Devices, and Cable Information for Multiple bus systems
SA38-0509	Diagnostic Information for Multiple bus systems

The publications are available for download from the Internet:

`http://www.ibm.com/servers/eserver/pseries/library`

3.2.5 Summary features of the SP-attached servers/CES - M/T 7017

A summary of the features of the SP-attached servers/Clustered Enterprise Servers - M/T 7017 is shown in Table 3-6.

Table 3-6 SP-attached servers / CES features - M/T 7017

Machine type and Model	7017-S85	7017-S80	7017-S7A	7017-S70
Processor type	RS64-IV	RS64-III/ RS64-IV	RS64-II	RS64
# of processors	4,6,12,18, 24	6,12,18,24	4,8,12	4,8,12
Clock rate MHz	600	450/600	262	125
Memory standard/maximum	4GB/96GB	2GB/96GB	1GB/32GB	512MB/32GB
Memory ECC	yes	yes	yes	yes

Machine type and Model	7017-S85	7017-S80	7017-S7A	7017-S70
Data/Instruction (L1) cache	128KB/128KB	128KB/128KB	64KB/64KB	64KB/64KB
L2 cache	16MB	8MB/16MB	8MB	4MB
L2 ECC	yes	yes	yes	yes
Service Processor	yes	yes	yes	yes
Hot-swappable disks	yes	yes	yes	yes
Dynamic CPU Deallocation	yes	yes	yes	-
Hot-plug slots	-	-	-	-
Redundant hot-plug power	yes	yes	yes	yes
Redundant hot/plug cooling	yes	yes	yes	yes
PCI slots available	53	53	53	53
PCI bus speed MHz	33	33	33	33
Disk/Media bays	48/8	48/8	48/8	48/8
Minimum/maximum internal disk storage GB	9.1/873.6	9.1/873.6	9.1/436.8	4.5/436.8
I/O Drawer	4	4	4	4
CUoD	yes	yes	-	-
System packaging	frame (+I/O rack)	frame (+I/O rack)	frame (+I/O rack)	frame (+I/O rack)
Servers per SP/CES	16	16	16	16
ROLTP performance	736.0	736.0	136.7	62.2

Note:

- Memory standard/maximum is shared memory.
- L1 cache and L2 cache are per processor.
- At least three PCI slots are already occupied by mandatory PCI adapters.

3.3 SP-attached servers/CES - M/T 7026 supported models

This section describes the M/T 7026 Models which are supported as SP-attached servers and Clustered Enterprise Servers.

3.3.1 RS/6000 Enterprise Server 7026 Model M80 (7026-M80)

The 7026-M80 server is a mid-range member of the 64-bit family of symmetric multiprocessing (SMP) enterprise servers from IBM. Positioned between the Model H80 and the powerful Model S80, the 7026-M80 server provides the power, capacity, and expandability required for e-business mission-critical computing. The 7026-M80 can assist you in managing the evolution of your business to incorporate the power of the Web and 64-bit computing into your environment while still supporting existing 32-bit applications.

Highlights

The 7026-M80 delivers a cost-efficient growth path to the future through such attributes as:

- 64-bit system scalability providing 2-, 4-, 6-, and 8-way configurations with combinations of the following processor options:
  - 2-way 500 MHz RS64 III with 4 MB L2 cache per processor
  - 4-way 500 MHz RS64 III with 4 MB L2 cache per processor
- Expandability and reliability
  - System memory expandable from 1 GB to 32 GB
  - Rack-mounted CEC and I/O drawers; CEC Drawer utilizes 8U (EIA Units) rack-space; I/O drawers utilize 5U rack space each
  - Up to four I/O drawers per server
  - Each I/O drawer supports 14 hot-plug PCI bus slots
- Support for SCSI, SSA, and Fibre Channel attached storage systems
- Support for 32-bit and 64-bit applications
- AIX Version 4.3.3 license included

Description

The 7026-M80 delivers a cost-efficient growth path to the future. It provides 64-bit scalability via the 64-bit RS64 III processor packaged as 2-way and 4-way cards. With its two processor positions, the 7026-M80 can be configured into 2-, 4-, 6-, or 8-way SMP configurations. The 7026-M80 also incorporates an I/O subsystem supporting 32-bit and 64-bit standard PCI adapters. The 7026-M80 has 2-and 4-way processor cards that operate at 500 MHz and incorporate 4 MB of L2 cache per processor.

The 7026-M80 is packaged as a rack-mounted Central Electronics Complex (CEC) drawer, cable attached to rack-mounted Remote I/O (RIO) drawers. The CEC and I/O drawers offer redundant power and redundant cooling. The CEC drawer incorporates the system processors, memory, and supporting systems logic. The primary (first) I/O drawer features:

- 14 available hot-plug PCI slots
- System support (service) processor
- An optical storage device capable of reading CD-ROM discs
- Diskette drive
- One additional media bay
- Integrated SCSI
- Ethernet port
- Keyboard port
- Mouse port
- Four serial ports and one parallel port

Up to three secondary I/O drawers can be added to provide up to a total of 56 PCI slots and eight media bays. System storage is provided via remotely attached SCSI, SSA, or Fibre Channel storage subsystems.

A fully configured 7026-M80 system would include:

- Two 4-way SMP card with 4 MB L2 cache per processor
- 32 GB of system memory
- 56 available hot-plug PCI slots
- 7 available media bays

The 7026-M80 CEC and I/O drawers are mounted in the IBM 7014 rack. Additional space in the 7014 rack may be utilized for various storage and communications subsystems.

The 7026-M80 server is shipped and delivered with all internal adapters and devices already installed and configured. AIX Version 4.3.3 software is included with every 7026-M80 and can be preinstalled.

The 7026-M80 provides performance and supports the evolution of your enterprise to Web integration and 64-bit computing. Its benefits include:

- Commercial performance in 2-, 4-, 6-, or 8-way enterprise servers using the RS64 III microprocessor technology.
- Reliability, availability, and serviceability (RAS) providing:
  - Error checking and correcting (ECC) system memory and cache
  - I/O link failure recovery
  - Environmental sensing
  - N+1 Power
  - N+1 Cooling units
  - Service processor for integrated system monitoring
  - Concurrent diagnostics
- Preconfigured tailored solutions that take advantage of the rack packaging to custom tailor your solution. Choose from multiple SCSI, SSA, and Fibre Channel attached:
  - Storage
  - Media
  - Communication
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Processor and memory growth. The minimum configuration 2-way processor can be increased to a 4-way, 6-way, or 8-way SMP. Memory can be expanded up to 32 GB in 1 GB, 2 GB, or 4 GB increments.
- Industry-standard hot-plug PCI slots can be expanded up to 56 total. These hot-plug slots allow most PCI adapters to be removed, added, or replaced without powering down the system. This function enhances system availability and serviceability.
- System power supplies are available in AC or DC and support the full 8-way SMP systems with 32 GB of memory. Redundant power supplies are incorporated in the M80 minimum configuration.
- Industry recognized AIX operating system license included with each system. New performance enhancements and a Workload Management feature are included with the latest release.

- Internal I/O system scalability and expandability that supports:
  - Up to four system I/O drawers with up to 56 PCI slots
  - Up to eight media bays for optical storage devices (a CD-ROM readable device) and tape devices
- Primary I/O drawer features:
  - Two integrated SCSI controllers (F/W and Ultra2)
  - An Ethernet 10/100 controller
  - Four serial ports
  - One parallel port
  - Keyboard and mouse ports

These are standard on the system and are made available without utilizing any of the expansion slots provided.

- Industry-standard PCI subsystem supporting 32-bit and 64-bit adapters. Greater application throughput is achieved via independent PCI buses for each of the 14 PCI slots in the I/O drawer. This design provides up to 1 GB per second aggregate bandwidth per drawer, depending on system configuration. The M80's independent PCI busses also are designed to enhance reliability by isolating adapters from other adapters within the drawer.
- A high-performance ECC L2 cache of 4 MB per processor on the 2-way and 4-way processor cards is provided.
- RAID support is available for SCSI and SSA. The operating system provides RAID 0 and 1. The optional SCSI and SSA RAID adapters provide RAID level 0, 1, and 5.
- Environmentally-induced (AC power) failure.

Figure 3-2 illustrates the system architecture for the 7017-M80 server.

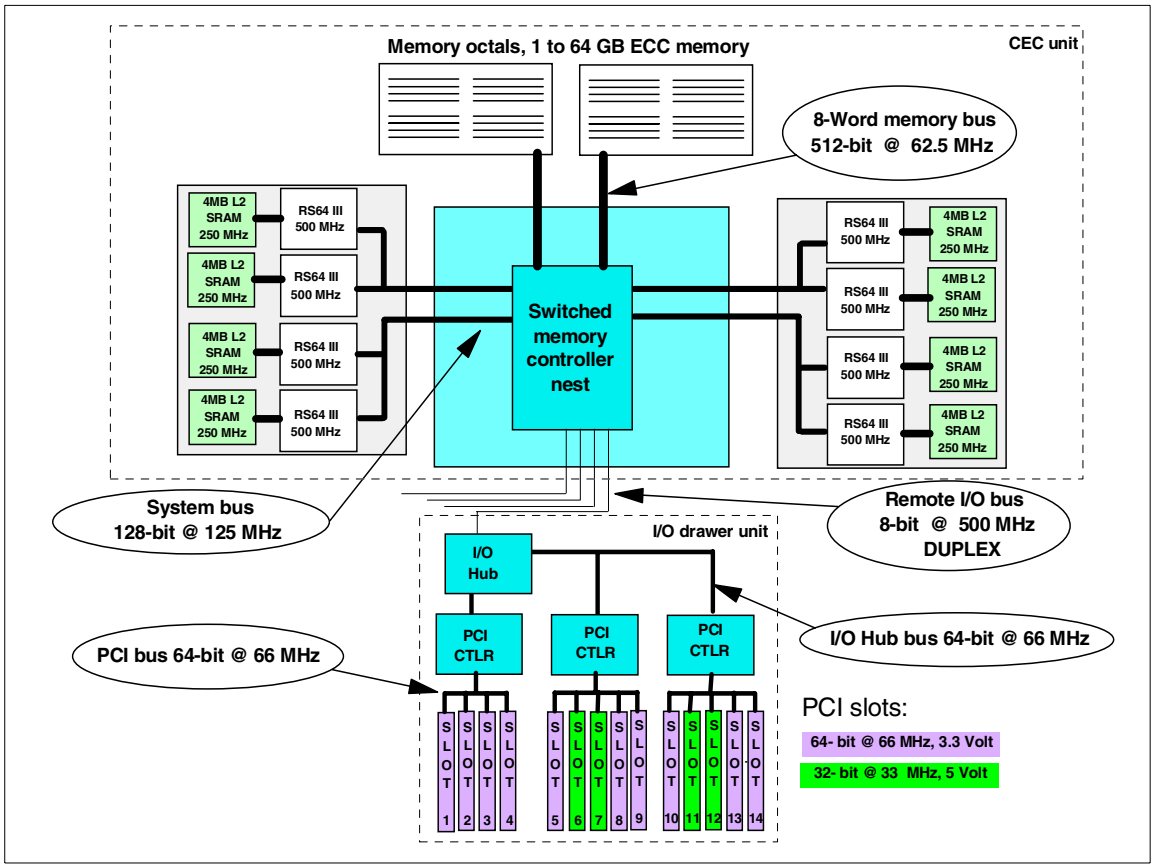


Figure 3-2 System architecture for 7026-M80 server

Product positioning

The 7026-M80 consists of a rack-mounted Central Electronics Complex (CEC) drawer containing the processors and memory connected to another rack-mounted rawer containing the media, hot-plug I/O slots, and optional boot bays. The CEC and I/O drawers implement redundant power and redundant cooling. The introduction of hot-plug PCI I/O slots brings new and increased levels of availability for this system. The need to power down and reboot when adding or replacing most PCI adapters is eliminated, thus reducing time delays while keeping network and application processing up and running. The 7026-M80 system can accommodate two 2-way or 4-way processor cards to create configurations of two, four, six, or eight processors as required, offering outstanding scalability as your business needs dictate.

The 7026-M80 is targeted at the key commercial processing segments of e-business, ERP, SCM, and Business Intelligence. In each segment, the 7026-M80 can meet critical requirements. For example:

- In ERP, the 7026-M80 is an application server with its powerful processors, memory capacity, and optional data storage capability.
- In e-business, the 7026-M80 can serve as a business- to-business Web server, going to other systems for business data or hosting the data storage itself.

The 7026-M80 rack-mounted drawer configuration offers flexibility regarding the number of CEC and I/O drawers that can be mounted in the rack, providing more compute and I/O power per square foot of valuable floor space. The 7026-M80 CEC requires 8U (EIA Units) of rack space, while each I/O drawer requires 5U of space. Thus, a minimum 7026-M80 configuration requires only 13U of rack space, while a maximum configuration fits into a 28U space. (The maximum configuration, which consists of one CEC drawer and four drawers, provides 56 available hot-plug PCI slots with up to eight processors and 32 GB of memory). Depending on the number of I/O drawers attached, up to two 7026-M80 systems can be installed in the 7014-S00 or the new 7014-T00 racks with room remaining to install external data storage drawers.

External disk storage is required for the 7026-M80 as data storage internal to the drawers is not provided; however, optional boot bays are available in the first I/O drawer. (The internal boot disk option requires the space of two PCI slots, leaving 12 available.) Memory in the CEC drawer can expand up to 32 GB, providing ample operating performance for demanding applications.

Its reliability improvements based on CPU de-allocation function enables the system to detect failing processors and take them off line without re-booting the system. The service processor can record this action and notify the systems administrator or service personnel of the condition. Computing operation continues with the processor deactivated, allowing repair to be scheduled at a time convenient to your processing schedule. Hot-plug functionality of the PCI I/O subsystem brings new levels of up-time to your system for the replacement or addition of I/O adapters. The design of this function shows a concern for the usability, convenience, and safety for the customer who performs this task. Selective disablement of an adapter can be made as necessary while system operations not dependent on this adapter continue to run without being affected. AIX assists the user with guidance through the process, helping to ensure success.

In addition, an HACMP packaged solution with dual 7026-M80s is offered to provide industry-leading high availability.The high availability solutions provided with the 7026-M80, can be ordered in a high-availability configuration with a choice of components. The high-availability configurations combined with

ClusterProven software give you a superior base for your high-availability requirements. The IBM high-availability solutions also offer a simple, but complete process for deciding which components are required and what software is best suited to a high- availability environment.

The 7026- M80 performs its role as the mid-range, rack-based platform of the "X80" family of SMP servers by complementing the deskside Model F80, the entry rack-based H80, and the powerful S80 in this broad server lineup. To assist in migrating to the "X80" family, a serial number protected model conversion is available for 7026-H80 customers who want to move up to the power of the 7026-M80. This model conversion replaces the 7026-H80 CEC drawer with an 7026-M80 CEC while allowing the 7026-M80 memory, I/O drawers, and adapters to be utilized with the new 7026-M80 system.

For more information on this model, refer to the redbook *RS/6000 Systems Handbook 2000 Edition*, SG24-5120.

Table 3-7 provides the publications for the 7026-M80 server.

Table 3-7 Publications for the 7026-M80 server

Order Number	Title
SA38-0576	RS/6000 Enterprise Server Model M80 Installation Guide
SA23-2652	System Unit Safety Information
SA38-0570	User's Guide
SA38-0571	Service Guide

The publications are available for download from the Internet:

`http://www.ibm.com/servers/eserver/pseries/library/`

3.3.2 @server pSeries 660 Model 6H1 (7026-6H1)

The 7026-6H1 server is a member of the 64-bit family of symmetric multiprocessing (SMP) servers from IBM. It offers a design that makes it a solution for mission-critical ERP/SCM, Data Warehouse/Data Mart, OLTP, and e-business applications. It blends copper silicon-on-insulator technology and larger memory capacity with a proven system architecture to give you reliability you can count on. Scalability is enhanced by increasing the amount of memory to 32 GB, by leveraging the enhancements of AIX Version 5.1, and the options of server clustering or SP attachment within your enterprise.

Combined with an array of industry software and services, the 7026-6H1 is a cost-efficient growth path to the future. It provides 64-bit scalability via the 64-bit RS64 III and IV processors packaged in 1-, 2-, 4-, and 6-way processor configurations and incorporates an I/O subsystem supporting 32-bit and 64-bit standard PCI adapters. The 7026-6H1 has 1-, 2-, and 4-way processors that operate at 450 MHz or 600 MHz, while the 6-way processor operates at 668 MHz.The L2 cache per processor is 2 MB for the 1-way processors, 4 MB for the 2-and 4-way processors, and 8 MB for the 6-way processor.

Highlights

The 7026-6H1 server delivers a cost-efficient growth path to the future through such attributes as:

- 64-bit system scalability with the following processor options:
  - 1-way 450 MHz RS64 III with 2 MB L2 cache
  - 1-way 600 MHz RS64 IV with 2 MB L2 cache
  - 2-way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 2-way 600 MHz RS64 IV with 4 MB L2 cache per processor
  - 4-way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 4-way 600 MHz RS64 IV with 4 MB L2 cache per processor
  - 6-way 668 MHz RS64 IV with 8 MB L2 cache per processor
- Expandability and reliability:
  - System memory expandable from 256 MB to 32 GB
  - Rack-mounted CEC and I/O drawers
  - Up to two I/O drawers per server
- CEC Drawer utilizes 5U (EIA Units) rack space
- I/O drawers utilize 5U rack space with each providing:
  - 14 available hot-plug PCI bus slots
  - Two media bays
- Support for SCSI, SSA, and Fibre Channel attached storage systems
- Support for 32-bit and 64-bit applications
- AIX Version 4.3.3 or 5.1 license included

Description

The 7026-6H1 is packaged as a rack-mounted Central Electronics Complex (CEC) drawer, cable attached to rack-mounted Remote I/O (RIO) drawers. The CEC and I/O drawers offer redundant power options and redundant cooling. The CEC drawer incorporates the system processors, memory, and supporting systems logic.

System storage is provided via remotely attached SCSI, SSA, or Fibre Channel storage subsystems. A fully configured 7026-6H1 server includes:

- One 6-way processor card with 8 MB L2 cache per processor
- 32 GB of system memory
- Two I/O drawers incorporating:
  - 28 available hot-plug PCI slots
  - Three available media bays

The 7026-6H1 CEC and I/O drawers are mounted in the IBM 7014 rack. Additional space in the 7014 rack may be utilized for various storage and communications subsystems.

AIX Version 4.3.3 or 5.1 software is included with every 7026-6H1 and can be preinstalled.

The 7026-6H1 brings new levels of availability features and functions to the Enterprise Server. The optional N+1 power subsystem provides redundancy in case of power supply failure. Hot-plugged power supplies in the CEC and I/O drawers allow concurrent repair when redundant power is installed. Optional uninterruptible power supply (UPS) systems are supported for this system.

The 7026-6H1's CEC and I/O drawer cooling subsystems have N+1 fans, which can be replaced concurrently if the optional redundant power is installed.

Systems equipped with optional redundant power have independent line cords for each power supply in the CEC and I/O drawers.

The 7026-6H1 provides Chipkill Memory with selected memory options. Using advanced technology derived from IBM's ultra-reliable System/390 enterprise servers, Chipkill Memory protects the server from any single memory chip failure and multi-bit errors from any portion of a single memory chip.

Memory chip failures can cause server system crashes that can result in the permanent loss of business data. Many file server applications lack the means to recover data after a catastrophic failure of this kind. Even when recovery is possible, it can take hours, resulting in significant losses in productivity or revenue.

Chipkill DIMMs for the 7026-6H1 provide the self-contained capability to correct real-time, multi-bit DRAM errors, including complete DRAM failures. This "RAID for memory" technology provides enhanced multi-bit error detection and correction that is transparent to the system.

Figure 3-3 illustrates the system architecture for 7026-6H1 server.

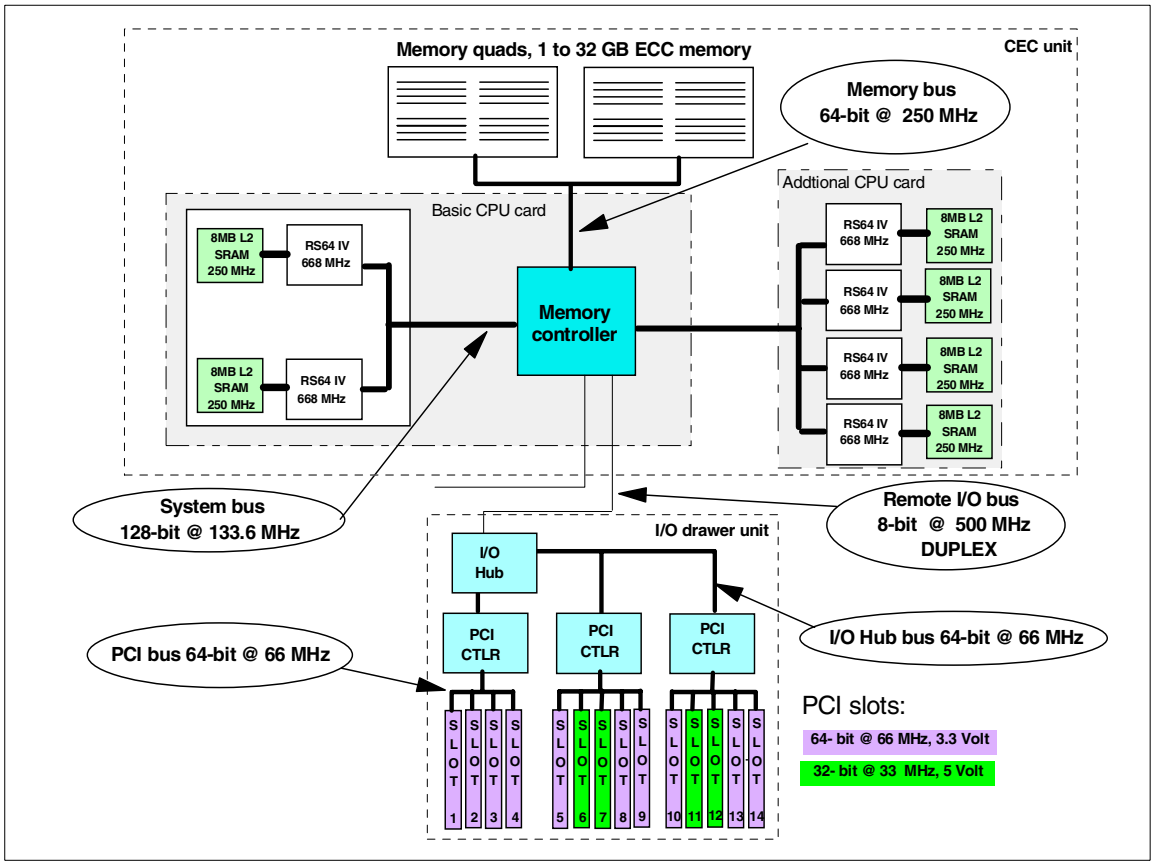


Figure 3-3 System architecture for 7026-6H1 server

Product positioning

The 7026-6H1 consists of a rack-mounted CEC drawer containing the processors and memory connected to another rack-mounted drawer containing the media, hot-plug I/O slots, and optional boot bays. Both drawers have redundant power options and redundant cooling. The incorporation of hot-plug PCI I/O slots brings new and increased levels of availability for this system. The need to power down and reboot when adding or replacing most PCI adapters is

eliminated, thus reducing time delays while keeping network and application processing up and running. The 7026-6H1 can accommodate one processor module in its smaller configuration and grow to two, four, or six processors as required, offering scalability as your business needs dictate.

The 7026-6H1 server is targeted at the key commercial processing segments of e-business, ERP, SCM, and Business Intelligence. In each segment, the 7026-6H1 can meet critical requirements. For example:

- In ERP, the 7026-6H1 is an application server with its powerful processors, memory capacity, and optional data storage capability.
- In e-business, the 7026-6H1 can serve as a business-to-business Web server, going to other systems for business data or hosting the data storage itself.

The 7026-6H1's rack-mounted drawer configuration offers flexibility regarding the number of CEC and I/O drawers that can be mounted in the rack, providing more compute and I/O power per square foot of valuable floor space. The 7026-6H1 CEC and I/O drawers each require 5U (EIA Units) of rack space. Thus, a 7026-6H1 configuration requires only 10U of rack space. This configuration, which consists of one CEC drawer and one I/O drawer, provides 14 available hot-plug PCI slots with up to six processors and 32 GB of memory.

External disk storage is required for the 7026-6H1 as data storage internal to the drawers is not provided; however, optional boot bays are available in the primary /O drawer. (The internal boot disk option requires the space of two PCI slots, leaving 12 available.) Memory in the CEC drawer can expand up to 32 GB, providing operating performance for demanding applications.

The 7025-6H1's reliability implementation, based on the CPU de-allocation function, enables the system to take a failing processor off line without re-booting the system. The service processor can record this action and notify the systems administrator, or service personnel, of the condition. Computing operation continues with the processor deactivated, allowing repair to be scheduled at a time convenient to your processing schedule.

Hot-plug functionality of the PCI I/O subsystem brings new levels of up-time to your system for the replacement or addition of I/O adapters. The design of this function shows a concern for the usability, convenience, and safety for the customer who performs this task. Selective disablement of an adapter can be made, as necessary, while system operations, not dependent on this adapter, continue to run without being affected. AIX assists the user by guiding him through the process, helping to ensure success.

In addition, an HACMP packaged solution with dual 7026-6H1 systems is offered to provide industry-leading high availability. The high-availability solutions provided with the 7026-6H1, can be ordered in a high-availability configuration with a choice of components. The high-availability configurations combined with ClusterProven software give you a superior base for your high-availability requirements. The IBM high-availability solutions also provide a complete and simple process for deciding which components are required and what software is best suited to a high-availability environment.

The 7026-6H1 server performs its role as the entry, rack-based platform of the pSeries family of SMP servers by complementing the deskside pSeries 620 Model 6F1, the mid-sized rack-based 7017-M80, and the pSeries 680 Model S85 in this broad server lineup.

3.3.3 @server pSeries 660 Model 6H0 (7026-6H0)

The 7026-6H0 server is a member of the 64-bit family of symmetric multiprocessing (SMP) servers from IBM. It delivers a cost-efficient growth path to the future through such attributes as:

- 64-bit system scalability with the following processor options:
  - 1-Way 450 MHz RS64 III with 2 MB L2 cache
  - 1-Way 600 MHz RS64 IV with 2 MB L2 cache
  - 2-Way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 2-Way 600 MHz RS64 IV with 4 MB L2 cache per processor
  - 4-Way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 4-Way 600 MHz RS64 IV with 4 MB L2 cache per processor
- Expandability and reliability
  - – System memory expandable from 256 MB to 32 GB
  - Rack-mounted CEC and I/O drawers
  - Up to two I/O drawers per server
- CEC Drawer utilizes 5U (EIA Units) rack space
- I/O drawers utilize 5U rack space with each providing:
  - 14 available hot-plug PCI bus slots
  - Two media bays
- Multiple power options include redundant 220 V AC or redundant -48 V DC
- NEBS level 3 compliant for the telecommunication carrier grade environment
- Support for SCSI, SSA, and Fibre Channel attached storage systems

- Support for 32-bit and 64-bit applications
- AIX Version 4.3.3 or 5.1 license included

3.3.4 RS/6000 Enterprise Server 7026 Model H80 (7026-H80)

The 7026-H80 server is a member of the 64-bit family of symmetric multiprocessing (SMP) enterprise servers from IBM. A follow-on to the popular Model H70, the 7026-H80 provides the power, capacity, and expandability required for e-business, mission-critical computing. The 7026-H80 can assist you in managing the evolution of your business to incorporate the power of the Web and 64-bit applications into your computing environment while still supporting existing 32-bit applications. The 7026-H80 is a cost-efficient growth path to the future. It provides 64-bit scalability via the 64-bit RS64 III processor packaged in 1-, 2-, 4-, and 6-way processor configurations and incorporates an I/O subsystem supporting 32-bit and 64-bit standard PCI adapters. The 7026-H80 has 1-, 2-, and 4-way processors that operate at 450 MHz, while the 6-way processor operates at 500 MHz. The 1-way processor provides 2 MB of L2 cache with the 2-, 4-, and 6-way processors incorporating 4 MB of L2 cache per processor.

Highlights

The 7026-H80 delivers a cost-efficient growth path to the future through such attributes as:

- 64-bit system scalability with the following processor options:
  - 1-way 450 MHz RS64 III with 2 MB L2 cache
  - 2-way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 4-way 450 MHz RS64 III with 4 MB L2 cache per processor
  - 6-way 500 MHz RS64 III with 4 MB L2 cache per processor
- Expandability and reliability:
  - System memory expandable from 256 MB to 16 GB
  - Rack-mounted CEC and I/O drawers
- CEC Drawer utilizes 5U (EIA Units) rack space
- I/O drawer utilizes 5U rack space:
  - I/O drawer provides 14 available hot-plug PCI bus slots
  - Integrated Ultra2 SCSI and 10/100 Ethernet without using adapter slots
  - An Optical Storage Device (a CD-ROM readable device) plus one available media bay
- Support for SCSI, SSA, and Fibre Channel attached storage systems

- Support for 32-bit and 64-bit applications
- AIX Version 4.3.3 license included

Description

The 7026-H80 is packaged as a rack-mounted Central Electronics Complex (CEC) drawer, cable attached to rack-mounted Remote I/O (RIO) drawers. The CEC and I/O drawers offer redundant power options and redundant cooling. The CEC drawer incorporates the system processors, memory, and supporting systems logic. The I/O drawer provides:

- Fourteen hot-plug PCI slots
- System support (service) processor
- An Optical Storage Device (a CD-ROM readable device)
- Diskette drive
- One additional media bay
- Integrated SCSI
- Ethernet port
- Keyboard port
- Mouse port
- Four serial ports and one parallel port

System storage is provided via remotely attached SCSI, SSA, or Fibre Channel storage subsystems.

A fully configured H80 system would include:

- One 6-way processor card with 4 MB L2 cache per processor
- 16 GB of system memory
- 14 available hot-plug PCI slots
- One available media bay

The 7026-H80 CEC and I/O drawers are mounted in the IBM 7014 rack. Additional space in the 7014 rack may be utilized for various storage and communications subsystems.

The 7026-H80 is shipped and delivered with all internal adapters and devices already installed and configured. AIX Version 4.3.3 software is included with every 7026-H80 and can be preinstalled.

The 7026-H80 server provides performance and supports the evolution of your enterprise to Web integration and 64-bit computing. Its benefits include:

- Commercial performance in 1-, 2-, 4-, or 6-way enterprise servers using the RS64 III microprocessor technology.
- Reliability, availability, and serviceability (RAS) providing:
  - Error checking and correcting (ECC) system memory and cache
  - I/O link failure recovery
  - Environmental sensing
  - N+1 Power
  - N+1 Cooling units
  - Service processor for integrated system monitoring
  - Concurrent diagnostics
- Preconfigured tailored solutions that take advantage of the rack packaging to custom tailor your solution. Choose from multiple SCSI, SSA, or Fibre Channel attached:
  - Storage
  - Media
  - Communication
- Open industry standards, such as Open Firmware based on IEEE P1275 standard.
- Processor and memory growth. The minimum configuration 1-way processor can be increased to a 2-way, 4-way, or 6-way SMP. Memory can be expanded up to 16 GB in 256 MB, 512 MB, or 1 GB increments.
- Industry-standard, hot-plug PCI slots allow most PCI adapters to be removed, added, or replaced without powering down the system. This function enhances system availability and serviceability.
- System power supplies are available in AC or DC and are designed to support up to 6-way SMP systems with 16 GB of memory. An optional modular AC power supply can be added to provide redundant power for AC systems. Redundant power is standard on systems with DC power.
- Industry recognized AIX operating system license included with each system. New performance enhancements and Workload Management features are included with the latest release.
- Primary I/O drawer that features:
  - Two integrated SCSI controllers (F/W and Ultra2)
  - An Ethernet 10/100 controller
  - Four serial ports
  - One parallel port

– Keyboard and mouse ports

These are standard on the system and are made available without utilizing any of the expansion slots provided.

- Industry-standard PCI subsystem supporting 32-bit and 64-bit adapters. Greater application throughput is achieved via independent PCI buses for each of the 14 PCI slots in the I/O drawer. This design provides up to 1 GBps aggregate bandwidth per drawer, depending on system configuration. The 7026-H80's independent PCI busses are designed to enhance reliability by isolating adapters from other adapters within the drawer.
- A high-performance ECC L2 cache of 2 MB on the 1-way processor and 4 MB per processor on the 2-, 4-, and 6-way processors is provided.
- RAID support available for SCSI and SSA. The operating system provides RAID 0 and 1. The optional SCSI and SSA RAID adapters provide RAID level 0, 1, and 5.

Product positioning

The 7026-H80 consists of a rack-mounted Central Electronics Complex (CEC) drawer containing the processors and memory connected to another rack-mounted drawer containing the media, hot-plug I/O slots, and optional boot bays. Both drawers have redundant power options and redundant cooling. The introduction of hot-plug PCI I/O slots brings new and increased levels of availability for this system. The need to power down and reboot when adding or replacing most PCI adapters is eliminated, thus reducing time delays while keeping network and application processing up and running. The 7026-H80 system can accommodate 1 processor module in its smallest configuration and grow to 2, 4, or 6 processors as required, offering scalability as your business needs dictate.

The 7026-H80 is targeted at the key commercial processing segments of e-business, ERP, SCM, and Business Intelligence. In each segment, the 7026-H80 can meet critical requirements. For example:

- In ERP, the 7026-H80 is an application server with its powerful processors, memory capacity, and optional data storage capability.
- In e-business, the 7026-H80 can serve as a business- to-business Web server, going to other systems for business data or hosting the data storage itself.

The 7026-H80's rack-mounted drawer configuration offers flexibility regarding the number of CEC and I/O drawers that can be mounted in the rack, providing more compute and I/O power per square foot of valuable floor space. The 7026-H80 CEC and I/O drawers each require 5U (EIA Units) of rack space. Thus, a 7026-H80 configuration requires only 10U of rack space. This configuration, which consists of one CEC drawer and one I/O drawer, provides 14 available hot-plug PCI slots with up to six processors and 16 GB of memory.

External disk storage is required for the 7026-H80 as data storage internal to the drawers is not provided; however, optional boot bays are available in the I/O drawer. (The internal boot disk option requires the space of two PCI slots, leaving 12 available.) Memory in the CEC drawer can expand up to 16 GB, providing ample operating performance for demanding applications.

The 7026-H80's reliability implementation, based on the CPU de-allocation function, enables the system to take a failing processor off line without re-booting the system. The service processor can record this action and notify the systems administrator, or service personnel, of the condition. Computing operation continues with the processor deactivated, allowing repair to be scheduled at a time convenient to your processing schedule.

Hot-plug functionality of the PCI I/O subsystem brings new levels of up-time to your system for the replacement or addition of I/O adapters. The design of this function shows a concern for the usability, convenience, and safety for the customer who performs this task. Selective disablement of an adapter can be made, as necessary, while system operations, not dependent on this adapter, continue to run without being affected. AIX assists the user by guiding him through the process, helping to ensure success.

In addition, an HACMP packaged solution with dual 7026-H80 systems is offered to provide industry-leading high availability. The high availability solutions provided with the H80, can be ordered in a high-availability configuration with a choice of components. The high-availability configurations combined with ClusterProven software give you a superior base for your high-availability requirements. The IBM high-availability solutions also provide a complete and simple process for deciding which components are required and what software is best suited to a high- availability environment.

The 7026-H80 performs its role as the entry, rack-based platform of the "X80" family of SMP servers by complementing the deskside Model F80, the mid-sized rack-based M80, and the powerful S80 in this broad server lineup. To assist in migrating to the "X80" family, a serial number protected model conversion is available for 7026-H70 customers who want to move to the 7026-H80 with memory and adapters being moved to the new 7026-H80 system. In addition,7026- H80 customers who require more system performance can obtain

a serial number protected model conversion to the 7026-M80 system. This model conversion replaces the 7026-H80 CEC drawer with an 7026-M80 CEC while allowing the 7026-H80 memory, I/O drawers, and adapters to be utilized with the new 7026-M80 system.

For more information on this model, refer to the redbook *RS/6000 Systems Handbook 2000 Edition*, SG24-5120-01.

Table 3-8 provides the publications for the 7026-H80 server.

Table 3-8 Publications for the 7026-H80 server

Order Number	Title
SA38-0575	RS/6000 Enterprise Server Model H80 Installation Guide
SA23-2652	System Unit Safety Information
SA38-0565	User's Guide
SA38-0566	Service Guide

The publications are available for download from the Internet:

`http://www-1.ibm.com/servers/eserver/pseries/library/`

3.3.5 Summary features of the SP-attached servers/CES - M/T 7026

A summary of the features of the SP-attached servers/Clustered Enterprise Servers - M/T 7026 is shown in Table 3-9.

Table 3-9 Features of the SP-attached servers and CES - M/T 7026

Machine type and Model	7026-M80	7026-6H1	7026-H80
Processor type	RS64-III	RS64-III/ RS64-IV	RS64-III/ RS64-IV
# of processors	2,4,6,8-way	1,2,4,6-way	1,2,4,6-way
Clock rate MHz	500	450 (1,2,4-way) 600 (1,2,4-way) 668 (6-way)	450 (1,2,4-way) 500 (6-way) 600 (1,2,4-way) 668 (6-way)
Memory standard/maximum	1GB/32GB	512MB/32GB	512MB/32GB
Memory ECC	yes	yes	yes
Data/Instruction (L1) cache	128KB/128KB	128KB/128KB	128KB/128KB

Machine type and Model	7026-M80	7026-6H1	7026-H80
L2 cache	4MB	2MB (1-way) 4MB (2,4-way) 8MB (6-way)	2MB (1-way) 4MB (2,4,6-way)
L2 ECC	yes	yes	yes
Service Processor	yes	yes	yes
Hot-swappable disks	External only	External only	External only
Dynamic CPU Deallocation	yes	yes	yes
Hot-plug slots	yes	yes	yes
Redundant hot-plug power	yes	Optionally	Optionally
Redundant hot/plug cooling	yes	yes	yes
PCI slots available	56	28	28
PCI bus speed MHz	33/66	33/66	33/66
Disk/Media bays	2/8	2/4	2/4
Minimum/maximum internal disk storage GB	0/36.4	0/36.4	0/36.4
I/O Drawer	4	2	1
CUoD	-	-	-
System packaging	rack drawer	rack drawer	rack drawer
Servers per SP/CES	32	32	32
ROLTP performance	222.5	187.1	111.9

Note:

- Memory standard/maximum is shared memory.
- L1 cache and L2 cache are per processor.
- One media bay is already occupied by the CD-ROM drive in the primary I/O drawer. The internal media bays are not capable of holding internal hard disks.
- Two internal disk drives can be attached to the primary I/O drawer via the dual boot bay option (F/C 6540), which eliminates the use of PCI slots 13 and 14. Three PCI slots are already occupied by mandatory PCI adapters.
- The 7026-6H0 Model, supported as SP-attached Server / CES, is not included in this table.

3.3.6 Cable and PCI card for SP CWS attachment (F/C 3154)

The SP CWS Attachment (F/C 3154) provides a cable and a PCI card necessary to attach an internal connector on the primary I/O drawer planar to the rear bulkhead of the I/O drawer. This connection is used with external cable F/C 3151 to attach to the SP CWS in support of SP-attached server and Clustered Enterprise Server environments.

This PCI card must be located in slot #7 of the primary I/O drawer. It combines into one F/C 3151 cable the functions carried by two cables (F/C 3150 and F/C 3151) on the SP-attached servers and Clustered Enterprise Servers - M/T 7017.

The hardware protocol of the SP-attached servers and Clustered Enterprise Servers - M/T 7026 within the frame is Convert Serial Processor (CSP).

Feature code 3154 (F/C 3154) uses no signal or power from the PCI bus. F/C 3154 has a ribbon cable, which attaches to the J12 port. The J12 port is a direct and dedicated pathway to the service processor, which gives greater scaling (up to 32 M/T 7026 servers). Remote monitoring, control, remote ASCII terminal capability (S1TERM), as well as power, are provided through the ribbon cable.

The device driver support is in the server base microcode.

F/C 3154 provides S1TERM, Hardmon, Sysmon and PSSP Perspectives capabilities.

Table 3-10 shows F/C 3154 characteristics.

Table 3-10 F/C 3154 characteristics

Feature Characteristics	F/C 3154 characteristics
Attributes provided	Internal port to bulkhead connection
Attributes required	One empty PCI slot (#7)
Minimum required per each SP-attached /CES	1
Maximum allowed	1
Software level required	AIX 4.3.3 with ML-08 or later PSSP 3.2 with APAR IY16350 or later

3.4 SP-attached server environment

The SP-attached servers use RS/6000 feature codes, not RS/6000 SP feature codes. The RS/6000 SP feature codes associated with the SP-attached server refers only to the system connections that attach the server to the RS/6000 SP system, they do not refer to hardware components. Since the SP-attached servers have cable attachment to the SP system, they have some attributes that appear node-like and others that appear frame-like to the SP system:

- F/C 9122 Node-like attachment of the SP-attached server to the SP system
- F/C 9123 Frame-like attachment of the SP-attached server to the SP system

The SP-attached servers M/T 7017 supported Models use F/C 9122 and F/C 9123 and appear nearly identical to nodes to the SP system.

The SP-attached servers M/T 7026 supported Models use F/C 9122, F/C 9123, and F./C 3154 to appear nearly identical to nodes to the SP system.

How the SP system views the SP-attached server

With a few hardware control exceptions, the SP-attached server performs the same functions that SP nodes perform. However, since the SP-attached server is mounted in its own frame and not in an SP frame, the SP system cannot view the SP-attached server as just another node. Instead, the SP system views the SP-attached server as an object with both frame and node characteristics. The node-like features of the SP-attached server are driven by F/C 9122 while the frame-like features of this device are driven by F/C 9123.

Because the SP-attached server has both frame and node characteristics, it must have both a frame number and a node number. However, since the SP-attached server does not have full SP frame characteristics, it cannot be considered a standard SP expansion frame. Thus, when you assign the frame number to the SP-attached server, you must observe the following rules:

- The SP-attached server cannot be inserted between a switch-configured frame and any non-switched expansion frame using that switch. As an example, consider that frames one and five of an SP system are switch configured. Frame two is a non-switched expansion frame attached to frame one. Frame six, seven, and eight are non-switched expansion frames attached to frame five. In this configuration, an SP-attached server could be given frame number three, but that would forbid any future attachment of non-switched expansion frames to the switch in frame one. If you assigned the SP-attached server to frame number nine, your system could still be scaled using other switch configured frames and non-switched expansion frames.
- The SP-attached server can be inserted between two switch-configured frames.

Note: Once the frame number has been assigned, the node number of the server (which is based on the frame number) is automatically generated. The following system defaults are used:

1. The SP-attached server is viewed by the system as a single frame containing a single node.
2. The system places the node-like features of the server in the slot one position.
3. Each SP-attached server installed in an SP system subtracts one node from the total node count allowed in the system; however, because the SP-attached server has frame-like features, it reserves sixteen node numbers that are used in determining the node number of nodes placed after the attached server.

Server attachment limits

When you add an SP-attached server to an SP system, the following limitations apply:

1. You must have at least one SP frame containing at least one node.
2. Each SP-attached server requires one valid, unused node slot in the SP system for switch port assignment. An assigned switch port is required in both switch-configured and switchless SP systems.

3. In some cases, the number of SP-attached servers you plan to install can exceed the number of available node slots in an SP frame. In this case, you can take advantage of any valid, unused node slots (and the associated switch ports) that exist in other SP frames in your system.

   As an example, consider a two-frame SP system. The first SP frame contains ten thin nodes and an SP Switch. The second SP frame contains five single SMP thin nodes and another SP Switch. You want to attach eight 7017 Enterprise Servers. In this example, you can attach six of the Enterprise Servers to the first frame and two Enterprise Servers to the second SP frame. As an alternative, all eight SP-attached servers can be connected to the second SP frame.

4. In some cases, the number of SP-attached servers you plan to install can exceed the number of available node slots in your SP system. In this case, you need to add an additional SP frame to your SP system. Only the first SP frame is required to have at least one node, additional SP frames can be empty.
5. Each SP-attached server counts as one node that must be subtracted from the total node count of 128 allowed in an SP system.
6. Each SP-attached server also counts as one frame that must be subtracted from the total frame count allowed in an SP system.

3.4.1 SP-attached server installation requirements

There are several requirements for hardware and software that you must meet before you can install an SP-attached server into your SP system as follows:

- System requirements
- Switch adapter requirements
- Network media card requirements
- Software requirements

SP-attached server system requirements

The following requirements must be met before you can place the SP-attached server into service:

1. Your SP system must be a tall frame system; short frames are not compatible with the SP-attached server.
2. If your SP system is switched, it must use 16-port SP Switches (F/C 4011); the SP Switch-8 and the SP Switch2 are not compatible with the SP-attached server.

3. You must make several cable connections. Since the SP-attached servers may require multiple RS-232 connections, you must use a multi-port asynchronous adapter inside the CWS. Do not use the integrated Ethernet for your SP LAN.
4. Since some of the cables used have limited lengths, you must keep those lengths and any required cable drops in mind when locating your SP-attached server in relation to the other SP system equipment.

SP-attached server switch adapter requirements

If you install an SP-attached server into a system that uses an SP Switch, you must install an SP Switch adapter in that server. The adapter you use to connect the SP-attached server to the SP Switch is the RS/6000 SP System Attachment adapter (F/C 8396). Because the RS/6000 SP System Attachment adapter is not an SP system adapter, you must order it using an RS/6000 feature code.

The single RS/6000 SP System Attachment adapter you place into each SP-attached server requires the following:

1. One valid, unused switch port on the SP Switch; corresponding to a legitimate node slot in your SP configuration. A legitimate node slot can be empty, the second half of a wide node, or one of the last three positions of a high node, provided that node slot satisfies the other rules for configuring nodes in an SP system. For example, if you have a frame with 16 thin nodes installed, you must not attach an RS/6000 SP System Attachment adapter to that frame until you remove a node and delete its configuration from the system image.
2. For 7017 models, three media card slots (slot 9-11) in the primary (first) I/O drawer of SP-attached server; For 7026 models, two media card slots (slot 6-7) in the primary I/O drawer.

If you attach an existing Enterprise Server to an SP system, you might find SCSI adapters installed in these slots. These SCSI adapters (if are supported on SP environment) must be relocated.

For example, if you attach an existing 7017 Enterprise Server to an SP system, you might find a SCSI adapter installed in slot #9 of the server. However, the SCSI adapter in slot 9 is typically connected to a boot device and requires special attention before it is removed.

M/T 7017 Boot device SCSI adapter relocation overview

1. Boot up the Enterprise Server you are attaching to the SP system.
2. Follow standard AIX procedures to change the boot device:
   a. Change device codes
   b. Change device address
3. Take the Enterprise Server down.
4. Move SCSI card from slot 9 to the new location (remember, slots 9 and 11 must be left open and the SP System Attachment adapter must be placed in slot 10).
   – Place SCSI adapter F/C 6206 and F/C 6208 either in slot 12 or slot 14
   – Place SCSI adapter F/C 6207 and F/C 6209 either in slot 12, slot 13, or slot 14
5. Reboot the server and continue with the SP attachment.

Note: Consult the appropriate documentation for specific installation procedures.

When you order an RS/6000 SP System Attachment adapter as server feature code, you must also order the SP Switch cable that connects the adapter to a valid switch port on the SP Switch. The SP Switch cable must be ordered with the RS/6000 SP system.

SP-attached server network media card requirements

Each network media card requires one media card slot in the SP-attached server. All network adapters in the SP-attached server use PCI architecture.

PCI adapter restrictions: Only SP system-supported PCI adapters can be used in an SP-attached server. Thus, if you attach an existing server to an SP system, you must remove any non-SP system supported PCI adapters.

If you plan to attach an existing 7017 or 7026 server to your system, you must place an SP LAN Ethernet adapter in the slot en0 position inside the server. Due to the fact that the Ethernet adapter in this slot must be configured for PSSP communications, any non-supported Ethernet adapter which is in the en0 slot must be removed. Additionally, if the Ethernet adapter in slot en0 is a supported Ethernet adapter, the adapter must be de-configured and then reconfigured as an SP-LAN Ethernet adapter.

For a complete listing of currently supported SP system PCI adapters, see Chapter 8, “PCI communication adapters” on page 227. With the exception of the RS/6000 SP System Attachment adapter (which is supported), if an adapter does not appear in the list, it is not supported for the SP-attached server. If you plan to use an existing Enterprise Server and any installed adapters do not appear in this list, they must be removed before it can be attached to the SP system.

SP-attached server software requirements

The SP-attached server M/T 7017 requires an SP system operating with the following software levels:

- PSSP 3.2 with APAR IY13026 (or later)
- AIX 4.3.3 with ML-06 (or later)

The SP-attached server M/T 7026 requires an SP system operating with the following software levels:

- PSSP 3.2 with APAR IY16350 (or later)
- AIX 4.3.3 with ML-08 (or later)

Notes:

1. Coexistence features provided by PSSP software permit other nodes in the system to operate at lower software levels. However, if you use an application that has coexistence requirements (such as PE), those requirements must be met.
2. If you attach an existing Enterprise Server, and that server is connected to an IPV6 network, you must remove the server from the network before making the SP attachement (some PSSP components tolerate IPV6 aliases for the IPV4 network but not with DCE, HACMP, HACWS, or an SP Switch).
3. AIX 5L operating system is not supported on SP-attached server environment.

Each SP-attached server also requires its own (chargeable) PSSP license.

3.4.2 Planning the SP-attached server interfaces

The SP-attached server requires different connections with your SP system.

Connections needed for SP-attached servers - M/T 7017

Each SP-attached server M/T 7017 must be attached to the control workstation via the following two serial cable features:

- A custom, IBM-supplied, 15 m (49 ft.) RS-232 cable connection from the control workstation to the SP-attached server SAMI port on the left side of the front operator panel
- A custom, IBM-supplied 15m (49 ft.) RS-232 cable connection from the control workstation to the SP-attached server S1 serial port on primary I/O drawer

Two 15 m (49 ft.) RS-232 cables are provided for each SP-attached server.

Note: Any server defined as SP-attached server on the SP system has to include the RS/6000 SP F/C 9122 Node-like attachement of the SP-attached server to the SP system, and F/C 9123 Frame-like attachement of the SP-attached server to the SP system. Based on this, the two RS-232 cables are provided for each SP-attached server.

Each SP-attached server M/T 7017 must be attached to the control workstation via an Ethernet LAN adapter. This adapter must be recognized by the clustered server as "en0" and must reside in slot #5 of the server primary I/O drawer.

The supported adapters for this system LAN are:

- For Twisted-pair cable connection:
  - 10/100 Ethernet 10BaseTX adapter (F/C 2968)
  - 10 MB AUI/RJ-45 Ethernet adapter (F/C 2987)

Note: Requires one F/C 9223 for each SP-attached server. It has to be ordered with the SP system (configures twisted-pair connection to server). The customer must supply all twisted-pair Ethernet cables.

- For BNC cable connection
  - 10 MB BNC/RJ-45 Ethernet adapter (F/C 2985)

Note: Requires one F/C 9222 for each SP-attached server. It has to be ordered with the SP system (configures BNC connection to server). Includes a 15 m (49 ft.) BNC Ethernet cable.

You must install a frame-to-frame, 10 m (33 ft.) electrical ground cable which is supplied with the SP system when you order F/C 9122 and F/C 9123. This cable is required in addition to the SP-attached server electrical ground. The frame-to-frame ground maintains the SP system and the SP-attached server at the same electrical potential.

The interconnection between the SP-attached server M/T 7017 and the SP system, can be accomplished utilizing the IBM RS/6000 SP System Attachment Adapter (F/C 8396) for the SP Switch connection.

If used, the adapter must be ordered as server feature code and it must be installed in slot #10 of the server primary I/O drawer and slots #9 and #11 must remain empty. Slot #9 must be left open to ensure the adapter has sufficient bandwidth. Slot #11 must be left open to provide clearance for the switch adapter heat sinks.

The SP Attachment Adapter uses an SP Switch cable to connect to an unused switch port on the current SP configuration.

Note: You must order one of the following cables which connect the adapter to a valid switch port on the SP Switch:

- 10 m SP Switch/Switch or Switch/Node cable (F/C 9310)
- 20 m SP Switch to Switch cable (F/C 9320)

The cable must be ordered as part of the SP system.

There are some SP-attached server placement limitations. Placement of the SP-attached server is limited by the length of its following supplied cables:

1. The 10 m (33 ft.) frame-to-frame ground cable and RS/6000 SP System Attachment adapter cable
2. The 15 m (49 ft.) RS-232 cables
3. The 15 m (49 ft.) BNC Ethernet cable

Approximately 3 m (10 ft.) of cable is needed for the vertical portion of the cable runs. Thus, the SP-attached server must be no more than 7 m (23 ft.) from the SP frame and no more than 12 m (40 ft.) from the control workstation.

Note: Refer to Chapter 10, “Clustering configurations” on page 285 for more details about the SP-attached server environment.

Connections needed for SP-attached servers - M/T 7026

Each SP-attached server M/T 7026 must be attached to the control workstation via a serial cable feature. It is a custom, IBM-supplied, 15 m (49 ft.) RS-232 cable connection from the control workstation to the SP-attached Server internal Attachment Adapter (F/C 3154) in slot #7 of the primary I/O drawer. Cable and PCI Card for SP Control Workstation Attachment (F/C 3154) provides internal connection from a connector on the planar of the primary I/O drawer to a PCI slot location on the rear bulkhead of the I/O drawer.

F/C 3154 must be ordered with the SP-attached server M/T 7026.

Two 15 m (49 ft.) RS-232 cables are provided for each SP-attached server. One of these cables may be used for this connection to F/C 3154.

Note: Any server defined as SP-attached server on the SP system, has to include the RS/6000 SP F/C 9122 Node-like attachment of the SP-attached server to the SP system and F/C 9123 Frame-like attachment of the SP-attached server to the SP system. Based on this, the two RS-232 cables are provided for each SP-attached server.

Each SP-attached server M/T 7026 must be attached to the control workstation via an Ethernet LAN adapter. This adapter must be recognized by the clustered server as "en0" and must reside in slot #1 of the server primary I/O drawer. The supported adapter for this system LAN is the 10/100 Ethernet 10BaseTX adapter (F/C 2968).

Note: Requires one F/C 9223 for each SP-attached server. It has to be ordered with the SP system (configures twisted-pair connection to server). The customer must supply all twisted-pair Ethernet cables.

You must install a frame-to-frame, 10 m (33 ft.) electrical ground cable which is supplied with the SP system when you order F/C 9122 and F/C 9123. This cable is required in addition to the SP-attached server electrical ground. The frame-to-frame ground maintains the SP system and the SP-attached server at the same electrical potential.

The interconnection between SP-attached server M/T 7026 and the SP system, can be accomplished utilizing the IBM RS/6000 SP System Attachment Adapter (F/C 8396) for the SP Switch connection.

If used, the adapter must be ordered as server feature code and it must be installed in slot #6 of the server primary I/O drawer. Slot #7 must only be used for the Internal Attachment Adapter (F/C 3154).

Note: You must remove the plastic divider between slots 6 and 7 due to the width of the SP-attachment adapter.

The SP Attachment Adapter uses an SP Switch cable to connect to an unused switch port on the current SP configuration.

Note: You must order one of the following cables which connect the adapter to a valid switch port.on the SP Switch:

- 10 m SP Switch/Switch or Switch/Node cable (F/C 9310).
- 20 m SP Switch to Switch cable (F/C 9320).

The cable must be ordered as part of the SP system.

There are some SP-attached server placement limitations. Placement of the SP-attached server is limited by the length of its following supplied cables:

1. The 10 m (33 ft.) frame-to-frame ground cable and RS/6000 SP System Attachment adapter cable.
2. The 15 m (49 ft.) RS-232 cables.

Approximately 3 m (10 ft.) of cable is needed for the vertical portion of the cable runs. Thus, the SP-attached server must be no more than 7 m (23 ft.) from the SP frame and no more than 12 m (40 ft.) from the control workstation.

Note: Refer to Chapter 10, “Clustering configurations” on page 285 for more details about the SP-attached server environment.

3.4.3 SP-attached server configurator rules

Ordering the SP-attached server requires the ordering of both SP features (cables and installation instructions, as well as tracking for virtual nodes and frames) and pSeries or RS/6000 features (SP LAN adapter, SP System Attachement adapter - where attachment to the SP Switch is desired, F/C 3154 for M/T 7026, PSSP code).

Configuring the system order for an SP-attached server to an existing SP system, consists of an MES order against the SP system to configure the SP-attached server as part of the SP system.

The RS/6000 SP system configurator session for all variations of SP-attached will include the following, when adding a SP node attachment:

- F/C 9122 Node Attachment Feature - An order for this feature will generate an order for two 15 Meter RS-232 cables for hardware control and connectivity between the CWS and the SP-attached server, and a 10 meter ground cable. It also traps some data so that the RS/6000 SP system configurator session can keep track of how many nodes (real and logical) are in the system.
- F/C 9123 Frame Attachment Feature - This feature keeps track of how many frames are currently in your RS/6000 SP system. Since the SP-attached server is both a logical node and a logical frame in the PSSP code logic, it is important to track this information to avoid exceeding allowable RS/6000 SP system limits for the number of frames.
- F/C 9222 Node Attachment Ethernet BNC Boot Feature (this feature tracks the choice to incorporate the SP-attached server as part of an SP system Ethernet BNC network and it provides BNC cable), or F/C 9223 Node Attachment Ethernet Twisted Pair Boot Feature (this feature tracks the choice to incorporate the SP-attached servers as part of an SP system Ethernet Twisted Pair network, but it provides no twisted pair cable)
- F/C 9310 or F/C 9320 Switch Connection Cable - This feature is required only if the SP-attached server is switch attached. It results in the ordering of one 10 or 20 meter switch connection cable.

Pre-Install required meeting - System Assurance Reviews

Before the order can be placed and the hardware installed, there will be a required Solutions Assurance Review (SAR) between Marketing, CEs, and the customer to ensure all participants understand prerequisite planning and activities for the SP-attach activities.

Required pre-SP-attachment activities

There are a number of potential system outage activities that must take place before the actual SP-attachment is made. These include both limited outage activities and system-wide outage activities. Limited system outage activities might include those activities that impact part of a system, such as possible movement or removal of adapters within the SP-attached Server or a planned failover between stand-alone systems to permit the installation of code upgrades on those stand-alones. System-wide outage activities might include the installation of new code levels on the Control Workstation and the subsequent required reboot. Planning for all of these activities will come out of the SAR meeting, and all should be conveyed to the customer as a necessary prerequisite to the merging of the SP-attached server into the SPs environment.

3.4.4 SP-attached servers connectivity diagram

Figure 3-4 on page 128 shows SP-attached servers connectivity diagram.

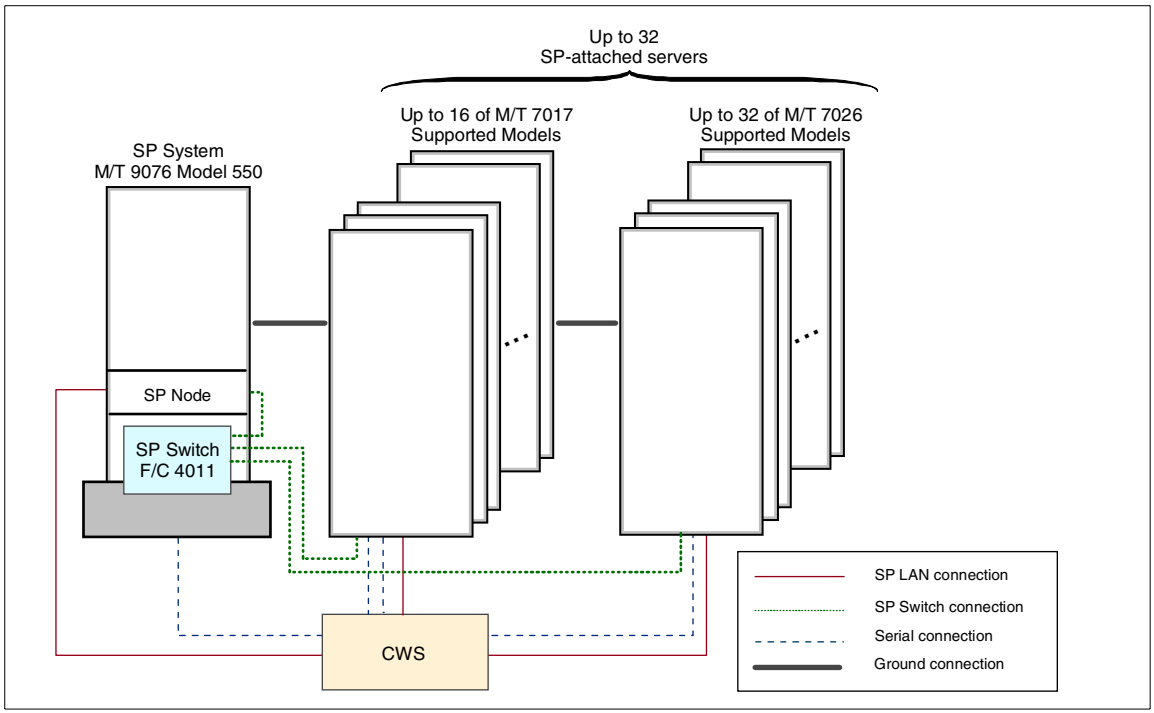


Figure 3-4 SP-attached servers connectivity diagram

3.4.5 Configuring Service Director on SP-attached servers

Service Director, as set of IBM software applications supplied with the Enterprise Servers M/T 7017 and M/T 7026, monitors the “health” of the system.

In a typical Enterprise Server installation, Service Director transmits reports through a modem supplied with that server. However, when the server is used as an SP-attached server, the modem supplied with the server is not used. In this installation, the SP-attached server acts like a Service Director client and forwards its Service Director messages to the local Service Director server. The local Service Director server can be the CWS of the SP which includes this SP-attached server, or any workstation connected to that LAN. When the local Service Director server receives messages from the SP-attached server, the messages are transmitted through the Service Director modem on the CWS.

To configure Service Director for the SP-attached server you must perform the following:

1. Configure the SP-attached server as a Machine Type 7017 or 7026 in Service Director. You must do this manually.
2. Configure Service Director on each SP-attached server to forward messages to the local Service Director server. The modem supplied with the server is not used.
3. Configure Service Director on the local Service Director server to forward messages received from the SP-attached server. The Service Director modem for the SP system is attached to the control workstation.

IBM intents to replace Service Director with Electronic Service Agent.

Note: Refer Appendix C, “SP system upgrades” on page 407 for more details about Service Director and Electronic Service Agent.

3.5 Clustered Enterprise Server environment

The Clustered Enterprise Servers use RS/6000 feature codes, not RS/6000 SP feature codes. The RS/6000 feature codes associated with the Clustered Enterprise Server refer to the cable connections that attach the servers to the control workstation.

3.5.1 Clustered Enterprise Server installation requirements

There are several requirements for hardware and software that you must meet before you can install a Clustered Enterprise Server into your CES system as follows:

- System requirements
- Switch adapter requirements
- Network media card requirements
- Software requirements

Clustered Enterprise Server system requirements

The following requirements must be met before you can place Clustered Enterprise Servers into service:

1. If your CES system is switched, it must use 16-port SP Switches (F/C 4011). The SP Switch-8 and the SP Switch2 are not compatible with the CES.

2. You must make several cable connections. Since the CES may require multiple RS-232 connections, you must use a multi-port asynchronous adapter inside the CWS. Do not use the integrated Ethernet for your SP LAN.
3. Some of the cables used with the CES have limited length. Those lengths, along with the vertical parts of the cable runs, limit the location of the servers in relation to the other system equipment.

Clustered Enterprise Server switch adapter requirements

If you install a Clustered Enterprise Server into a CES system that uses an SP Switch, you must install an SP Switch adapter in that server. The adapter you use to connect the Clustered Enterprise Server to the SP Switch is the RS/6000 SP System Attachment adapter (F/C 8396). Because the RS/6000 SP System Attachment adapter is not an SP system adapter, you must order it using an RS/6000 feature code.

The adapter uses an SP Switch cable to connect to an available switch port on a 9076-555 Stand-alone SP Switch for Clustered Nodes.

The model 555 is a tall frame with an integral SEPBU power supply and SP Switch (F/C 4011), but no processor nodes. This frame provides node interconnectivity in a CES system having configurations of 2 to 16 servers. By adding a second SP Switch (F/C 4011) on the 9076-555 Stand-alone SP Switch, the configuration will be extended with a Tall frame (F/C 1550), which will host the second SP Switch, and the interconnectivity in CES system will be increased up to 32 servers. If the 9076-555 is configured with two SP Switches (two F/C 4011), then the 16 switch-to-switch cables must be ordered with the 9076-555.

RS/6000 Server Attached Node (F/C 9122), which has to be ordered on the 9076-555 Stand-alone SP Switch, counts the number of RS/6000 Server Attached Nodes configured on the CES system (up to 16/32 Clustered Enterprise Servers). Every RS/6000 Server attached node (F/C 9122) selected generates an RS/6000 Server Frame Attachment (F/C 9123).

The RS/6000 SP System Attachment adapter requires, for 7017 models, three media card slots (slot 9-11) in the primary (first) I/O drawer of the SP-attached server. For 7026 models, two media card slots (slot 6-7) are required in the primary drawer.

If you attach an existing Enterprise Server to a CES system, you might find SCSI adapters installed in these slots. These SCSI adapter must be relocated (if supported on the SP environment).

For example, if you attach an existing 7017 Enterprise Server to a CES system, you might find a SCSI adapter installed in slot #9 of the server. However, the SCSI adapter in slot 9 is typically connected to a boot device and requires special attention before the SCSI adapter is removed.

M/T 7017 Boot device SCSI adapter relocation overview

1. Boot up the Enterprise Server you are attaching to the CES system.
2. Follow standard AIX procedures to change the boot device:
   a. Change device codes
   b. Change device address
3. Take the Enterprise Server down.
4. Move SCSI card from slot 9 to the new location (remember, slots 9 and 11 must be left open and the SP System Attachment adapter must be placed in slot 10).
   – Place SCSI adapter F/C 6206 and F/C 6208 either in slot 12 or slot 14
   – Place SCSI adapter F/C 6207 and F/C 6209 either in slot 12, slot 13, or slot 14
5. Reboot the server and continue with the Clustered Enterprise Server installation.

Note: Consult the appropriate documentation for specific installation procedures.

Clustered Enterprise Server network media card requirements

Each network media card requires one media card slot in the Enterprise Server. All network adapters in the Clustered Enterprise Server use PCI architecture.

PCI adapter restriction

Only SP system-supported PCI adapters can be used in a Clustered Enterprise Server. If you attach an existing server to a CES system, you must remove any non-SP system-supported PCI adapters.

If you plan to attach an existing 7017 or 7026 server to your system, you must place an SP LAN Ethernet adapter in the slot en0 position inside the server. Due to the fact that the Ethernet adapter in this slot must be configured for PSSP communications, any non-supported Ethernet adapter that is in the en0 slot must be removed. Additionally, if the Ethernet adapter in slot en0 is a supported Ethernet adapter, the adapter must be de-configured and then reconfigured as an SP-LAN Ethernet adapter.

For a complete listing of currently supported SP system PCI adapters, see Chapter 8, “PCI communication adapters” on page 227. With the exception of the RS/6000 SP System Attachment adapter (which is supported), if an adapter does not appear in the list, it is not supported for the CES. If you plan to use an existing Enterprise Server and any installed adapters do not appear in this list, they must be removed before it can be attached to the Clustered Enterprise Server.

Clustered Enterprise Server software requirements

The Clustered Enterprise Servers M/T 7017 require the following software levels:

- PSSP 3.2 with APAR IY13026 (or later)
- AIX 4.3.3 with ML-06 (or later)

The Clustered Enterprise Servers M/T 7026 require the following software levels:

- PSSP 3.2 with APAR IY16350 (or later)
- AIX 4.3.3 with ML-08 (or later)

Note:

1. AIX 5L operating system is not supported on Clustered Enterprise Server environment.
2. If you plan to attach an existing enterprise server in the CES system, and that server is connected to an IPV6 network, you must remove the server from that network before including it in the CES system (some PSSP components tolerate IPV6 aliases for the IPV4 network but not with DCE, HACMP, HACWS, or an SP Switch).
3. Each Clustered Enterprise Server requires its own (chargeable) PSSP license.

3.5.2 Planning the Clustered Enterprise Server interfaces

The Clustered Enterprise Server requires different connections with your CES system.

Connections needed for the CES - M/T 7017

Each Clustered Enterprise Server M/T 7017 must be attached to the control workstation via the following two serial cable features:

- Clustered Server Serial Port to Control Workstation Cable (F/C 3150) - server SAMI port on left side of the front operator panel.
- Clustered Server Control Panel to Control Workstation Cable (F/C 3151) - server S1 serial port on primary I/O drawer.

F/C 3150 and F/C 3151 cables must be ordered with the Clustered Enterprise Server M/T 7017 (or with the CWS as desired).

Each Clustered Enterprise Server M/T 7017 must be attached to the control workstation via an Ethernet LAN adapter. This adapter must be recognized by the clustered server as "en0" and must reside in slot #5 of the server's primary I/O drawer.

The supported adapters for this system's LAN are:

- For Twisted-pair cable connection:
  - 10/100 Ethernet 10BaseTX adapter (F/C 2968)
  - 10 MB AUI/RJ-45 Ethernet adapter (F/C 2987)
- For BNC cable connection
  - 10 MB BNC/RJ-45 Ethernet adapter (F/C 2985)

Note: The customer must supply all required Ethernet Twisted-pair or BNC cables.

The interconnection between servers in the CES system, can be accomplished utilizing the IBM RS/6000 SP System Attachment Adapter (F/C 8396) for the SP Switch connection.

If used, the adapter has to be ordered as a server feature code and it must be installed in slot #10 of the Clustered Enterprise Server primary I/O drawer; slots #9 and #11 must remain empty. Slot #9 must be left open to ensure the adapter has sufficient bandwidth. Slot #11 must be left open to provide clearance for the switch adapter heat sinks.

The RS/6000 Server Attached Node (F/C 9122), which has to be ordered on the 9076-555 Stand-alone SP Switch, counts the number of RS/6000 Server Attached Nodes configured on the CES system (up to 16/32 Clustered Enterprise Servers). Every RS/6000 Server attached node (F/C 9122) selected generates an RS/6000 Server Frame Attachement (F/C 9123).

Note: With F/C 9122 RS/6000 Server Attached Node and F/C 9123 RS/6000 Server Frame Attachement, two 15 m (49 ft.) RS/232 cables and a frame-to-frame electrical ground cable are provided. So, on a switched Clustered Enterprise Server environment, you may not need to order the F/C 3150 and F/C 3151 serial cables to connect the Clustered Enterprise Server to CWS. The frame-to-frame ground maintains the 9076-555 and the Clustered Enterprise Server at the same electrical potential.

The SP Attachment Adapter uses an SP Switch cable to connect to an available switch port on a 9076-555 Stand-alone SP Switch.

Note: You must order one of the following cables, which connect the adapter to a valid switch port on the SP Switch:

- 10 m SP Switch/Switch or Switch/Node cable (F/C 9310)
- 20 m SP Switch to Switch cable (F/C 9320)

The cable must be ordered as part of the 9076-555 Stand-alone SP Switch.

There are some Clustered Enterprise Server placement limitations. Placement of the Clustered Enterprise Server is limited by the length of its following supplied cables:

1. The 10 m (33 ft.) frame-to-frame ground cable and RS/6000 SP System Attachment adapter cable
2. The 15 m (49 ft.) RS-232 cables

Approximately 3 m (10 ft.) of cable is needed for the vertical portion of the cable runs. Thus, the Clustered Enterprise Server must be no more than 7 m (23 ft.) from the 9076-555 Stand-alone SP Switch (if it exists on your configuration) and no more than 12 m (40 ft.) from the control workstation.

Note: Refer to Chapter 10, “Clustering configurations” on page 285 for more details about the CES environment.

Connections needed for the CES - M/T 7026

Each Clustered Enterprise Server M/T 7026 must be attached to the control workstation via a serial cable feature. It is a custom 15 m (49 ft.) RS-232 cable – a Clustered Server Control Panel to CWS cable (F/C 3151) – to the CSP port connector on the Internal Attachment Adapter (F/C 3154) in slot #7 of the primary I/O drawer. Cable and PCI Card for SP Control Workstation Attachment (F/C 3154) provides internal connection from a connector on the planar of the primary I/O drawer to a PCI slot location on the rear bulkhead of the I/O drawer.

F/C 3151 and F/C 3154 must be ordered with the Clustered Enterprise Server M/T 7026.

Each Clustered Enterprise Server M/T 7026 must be attached to the control workstation via an Ethernet LAN adapter. This adapter must be recognized by the clustered server as "en0" and must reside in slot #1 of the server primary I/O drawer. The supported adapter for this system LAN is the 10/100 Ethernet 10BaseTX adapter (F/C 2968).

Note: The customer must supply all required Ethernet cables.

The interconnection between servers in the CES system, can be accomplished utilizing the IBM RS/6000 SP System Attachment Adapter (F/C 8396) for the SP Switch connection.

If used, the adapter has to be ordered as a server feature code and it must be installed in slot #6 of the server primary I/O drawer. Slot #7 must only be used for the Internal Attachment Adapter (F/C 3154).

Note: You must remove the plastic divider between slots 6 and 7 due to the width of the SP-attachment adapter.

RS/6000 Server Attached Node (F/C 9122), which has to be ordered on the 9076-555 Stand-alone SP Switch, counts the number of RS/6000 Server Attached Nodes configured on the CES system (up to 16/32 Clustered Enterprise Servers). Every RS/6000 Server attached node (F/C 9122) selected generates a RS/6000 Server Frame Attachement (F/C 9123).

Note: With F/C 9122 RS/6000 Server Attached Node and F/C 9123 RS/6000 Server Frame Attachement, two 15 m (49 ft.) RS/232 cables and a frame-to-frame electrical ground cable are provided. So, on a switched Clustered Enterprise Server environment, you may not need to order the F/C 3151 serial cable to connect the Clustered Enterprise Server to CWS. The frame-to-frame ground maintains the 9076-555 and the Clustered Enterprise Server at the same electrical potential.

The SP Attachment Adapter uses an SP Switch cable to connect to an available switch port on a 9076-555 Stand-alone SP Switch.

Note: You must order one of the following cables, which connect the adapter to a valid switch port.on the SP Switch:

- 10 m SP Switch/Switch or Switch/Node cable (F/C 9310)
- 20 m SP Switch to Switch cable (F/C 9320)

The cable must be ordered as part of the 9076-555 Stand-alone SP Switch.

There are some Clustered Enterprise Server placement limitations. Placement of the Clustered Enterprise Server is limited by the length of its supplied cables:

1. The 10 m (33 ft.) frame-to-frame ground cable and RS/6000 SP System Attachment adapter cable
2. The 15 m (49 ft.) RS-232 cables

Approximately 3 m (10 ft.) of cable is needed for the vertical portion of the cable runs. Thus, the Clustered Enterprise Server must be no more than 7 m (23 ft.) from the 9076-555 Stand-alone SP Switch (if it exists on your configuration) and no more than 12 m (40 ft.) from the control workstation.

Note: Refer to Chapter 10, “Clustering configurations” on page 285 for more details about the CES environment.

3.5.3 Clustered Enterprise Servers connectivity diagram

Figure 3-5 on page 137 shows Clustered Enterprise Servers connectivity diagram.

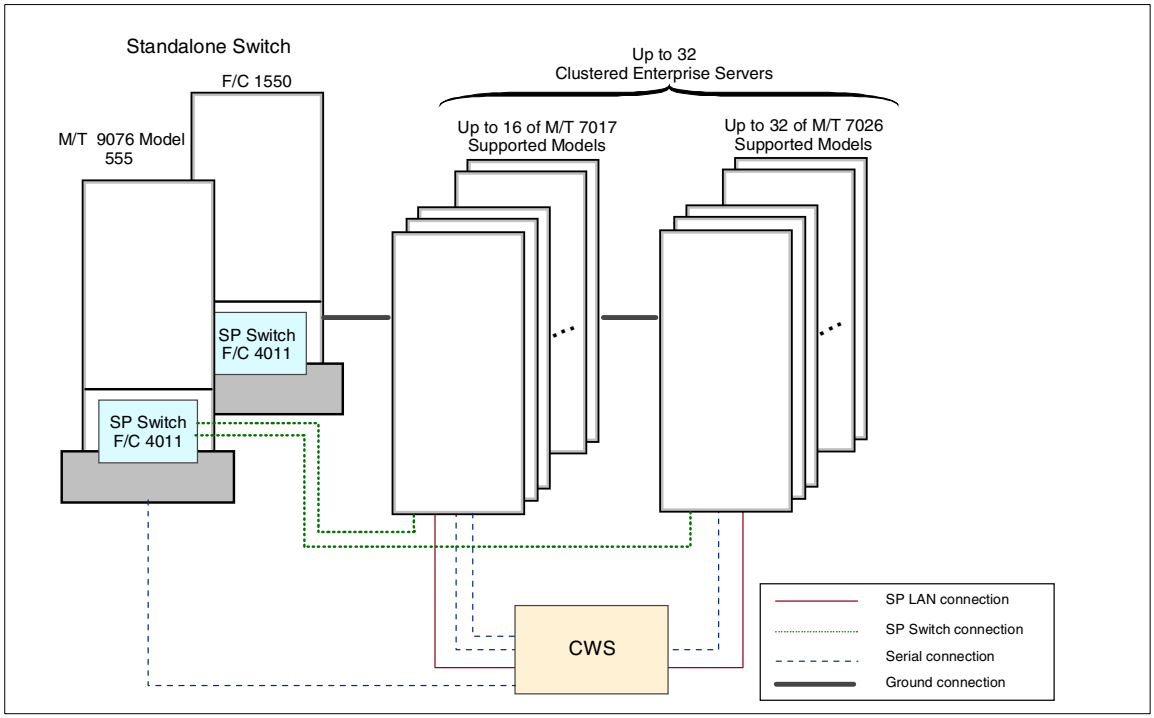


Figure 3-5 Clustered Enterprise Servers connectivity diagram

3.5.4 Configuring Service Director on CES systems

Service Director, a set of IBM software applications supplied with the Enterprise Servers M/T 7017 and M/T 7026, monitors the “health” of the system.

In a typical Enterprise Server installation, Service Director transmits reports through a modem supplied with that server. However, when the server is used as a Clustered Enterprise Server, the modem supplied with the server is not used. In this installation, the Clustered Enterprise Server acts like a Service Director client and forwards its Service Director messages to the local Service Director server. The local Service Director server can be the CWS of the CES system, or any workstation connected to that LAN. When the local Service Director server receives messages from the Clustered Enterprise Server, the messages are transmitted through the Service Director modem on the CWS.

To configure Service Director for the Clustered Enterprise Server you must perform the following:

1. Configure the Clustered Enterprise Server as a Machine Type 7017 or 7026 in Service Director. You must do this manually.

2. Configure Service Director on each Clustered Enterprise Server to forward messages to the local Service Director server. The modem supplied with the server is not used.
3. Configure Service Director on the local Service Director server to forward messages received from the Clustered Enterprise Server. The Service Director modem for the CES system is attached to the control workstation.

IBM intents to replace Service Director with Electronic Service Agent.

Note: Refer Appendix C, “SP system upgrades” on page 407 for more details about Service Director and Electronic Service Agent.

4

Frames

This chapter describe the different models of frames used in SP system and in CES.

The frames are used to hold nodes and switches in the SP system and to provide nodes interconnectivity in Clustered Enterprise Servers (CES) configuration.

There are two frames sizes; each size is designed to fit different system configurations:

- Tall frames
  - Tall model frames (Model 550, 555)
  - Tall expansion frames (F/C 1550)
  - SP Switch frames (F/C 2031, 2032)
- Short frames:
  - Short model frames (Model 500)
  - Short expansion frames (F/C 1500)

4.1 Frames

The SP processor nodes can be mounted in either a tall or short SP frame. The frame spaces that nodes fit into are called drawers. A tall frame has eight drawers, while a short frame has four drawers. Each drawer is further divided into two slots. One slot can hold one thin node. A wide node occupies one drawer (two slots), and a high node occupies two drawers (four slots).

An internal power supply (SEPBU) is included with each frame. A Frame Supervisor card handles the connection to the CWS via the RS232 to manage the installed node and Switch in the frame. Frames get equipped with the optional processor nodes and switches that you order. Strictly speaking there are three types of frames:

- Short frames
- Tall frames
- Switch frames

The tall and short frames are used to host nodes and they are usually just called frames. The switch frames are used to host switches or Intermediate Switch Boards (ISB). This special type of frame can host up to eight switch boards. After the first SP was made commercially available some years ago, there have been a number of model and frame configurations. Each configuration was based on the frame type and the kind of node installed in the first slot. This led to an increasing number of possible prepackaged configurations when more nodes became available. Figure 4-1 on page 141 shows the tall model frame (Model 550).

Figure 4-1 SP frame model 550

The product structure of the RS/6000 SP system was simplified with the introduction, on April 21, 1998, of a new frame. The new SP system was simplified with only two models (Model 500 and Model 550) and two expansion frame features (F/C 1500 and F/C 1550). The new frame replaces the old tall frame, which consisted of six models and six frame features.

See Appendix B, “Hardware” on page 365 for the old frames models; Table B-7 on page 389 shows the SP legacy models by node type; Table B-8 on page 390 and Table B-9 on page 391 show the SP legacy models by frame and switch types.

All new designs are completely compatible with all valid SP configurations using older equipment. Also, all new nodes can be installed in any existing SP frame provided that the required power supply upgrades have been implemented in that frame. The reason for this is that the SMP nodes have higher power consumption. Therefore, there is a higher power requirement for the frame.

4.1.1 SP frame dimensions

The most noticeable difference between the new and old tall frame is the reduction in height. Another physical difference is the footprint. Before this new frame offering, the frame and the first node in the frame were tied together, forming a model. Each new node that was made available was potentially installable in the first slot of a frame; so, a new model was born. With the new offering, IBM simplified the SP frame options by decoupling the imbedded node from the frame offering. Therefore, when you order a frame, all you receive is a frame with the power supply units and a power cord. All nodes, switches, and other auxiliary equipment are ordered separately. Table 4-1 details the SP frame dimensions, including the base.

Table 4-1 SP frame dimensions

Frame type	Height		Width		Depth		Weight	
	mm	in	mm	in	mm	in	kg	lbs
Tall Legacy models	2007	79	915	36	1118	44		
Tall 550 models and 555 models	1925	75.8	922	36.3	1295	51	441 to 984	971 to 2165
Short 500 models	1245	49	711	28	1015	40	232 to 414	510 to 910

The new tall frame has a height of 75.8 inches (1.9 m), which provides an easier fit for the entrance doors than the old 79 inches (2.01 m) tall frame. However, it requires more floor space because the depth of the tall frame is 51 inches (1.3 m), which is 7 inches (0.2 m) greater than the old tall frame. The new Short frame has the same dimensions as the old short frame.

4.1.2 Frame electrical power requirements

The electrical power specifications for the old and new power subsystems are summarized in Table 4-2.

Table 4-2 SP frame electrical power requirements

Power Requirements	Tall frame			Short frame	
	Legacy PDU	Legacy SEPBU	Model 550 555	Legacy	Model 500
SP Frame					
Power output (kW)	6.7	7	10.5	3.5	5.0
Utility loading (kvA)[1]	7.8	6.7	8.1	3.5	4.2
Phase	3	3	3	1	1
Current (Amps, low/high[2] V)	24/16	24/16	35/20	24/-	30/-
Base Power Regulators[3]	N/A	3	4	2	2
Computer room facility					
Conductor size (AWG, low/high[2] V)	10/10	10/10	8/8	10	6
Circuit breaker (Amps, low/high[2] V)	30/20 7	30/20 7	50/30 7	30/- 7	40/- 7
Connector (low voltage only[4])	460C9W 5	430C9W 5	460C9W 5	3933[6]	9C53U0[6]
Receptacle (low voltage only[4])	460R9W 5	430R9W 5	460R9W 5	3753[6]	9R53U0[6]

Notes:
[1] With thin nodes in every node position
[2] Low voltage: 200 - 240 V; High voltage: 380 - 415 V
[3] With n+1 redundant frame power
[4] High-voltage systems are wired to customer electrical service
[5] Any IEC309
[6] Russell-Stoll
[7] Use Circuit Breaker with motor start characteristics to allow inrush currents during the power-on of SP. The service you provide must be capable of an inrush current of 90 amperes maximum peak. This inrush current decays within five cycles.

4.1.3 SP frame redundant power supply (F/C 3885, 3886)

The redundant power supply feature permits feeding ac power to SP system frames from two separate sources to prevent system outages caused by the failure of a single ac power source. With this feature, the SP system operates normally with power applied to only one of the two power cables. Two separate power cables are attached to a special SEPBU enclosure having two side-by-side connectors and four standard power books. No special voltage or synchronization requirements of the two AC sources are required. The two AC sources are galvanically isolated within the SP system. When both power sources are operating within their normal range, the power cables share the system load.

Attention: This feature is offered only for tall frames.

Input branch circuit requirements

The Redundant Power Supply feature is available for three-phase, “open V” low-voltage (200-240 V ac nominal) and high-voltage (380-415 V ac nominal) sources (F/C 3885), and for single-phase low-voltage (200-240 V ac nominal) sources (F/C 3886). Figure 4-2 shows the rear of the redundant power supply feature.

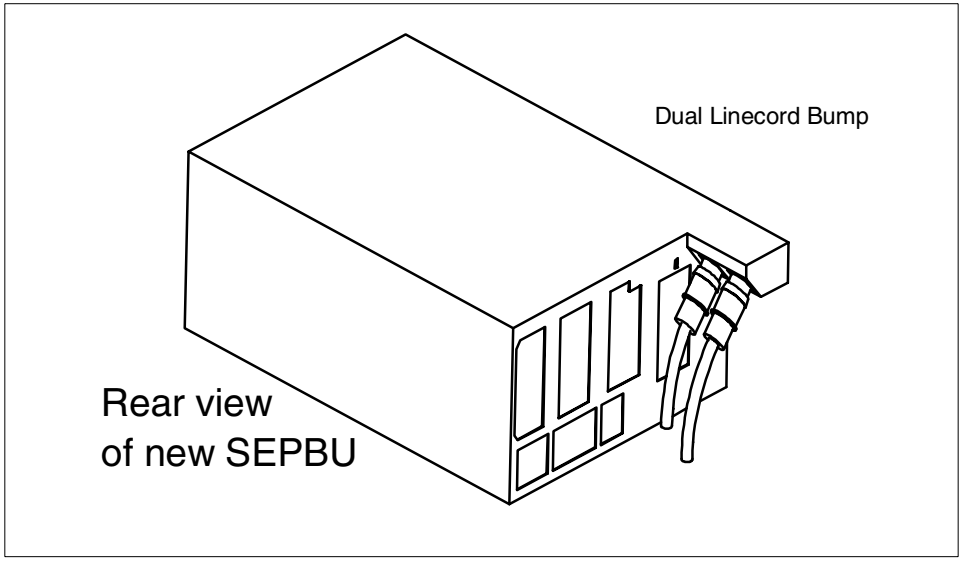


Figure 4-2 Rear of the redundant power supply feature

4.2 Tall frames

Tall frames are available in three variations:

- Tall model frames (Model 550, 555)
- Tall expansion frames (F/C 1550)
- SP switch frames (F/C 2031, 2032)

Figure 4-3 shows a tall frame's front and rear views.

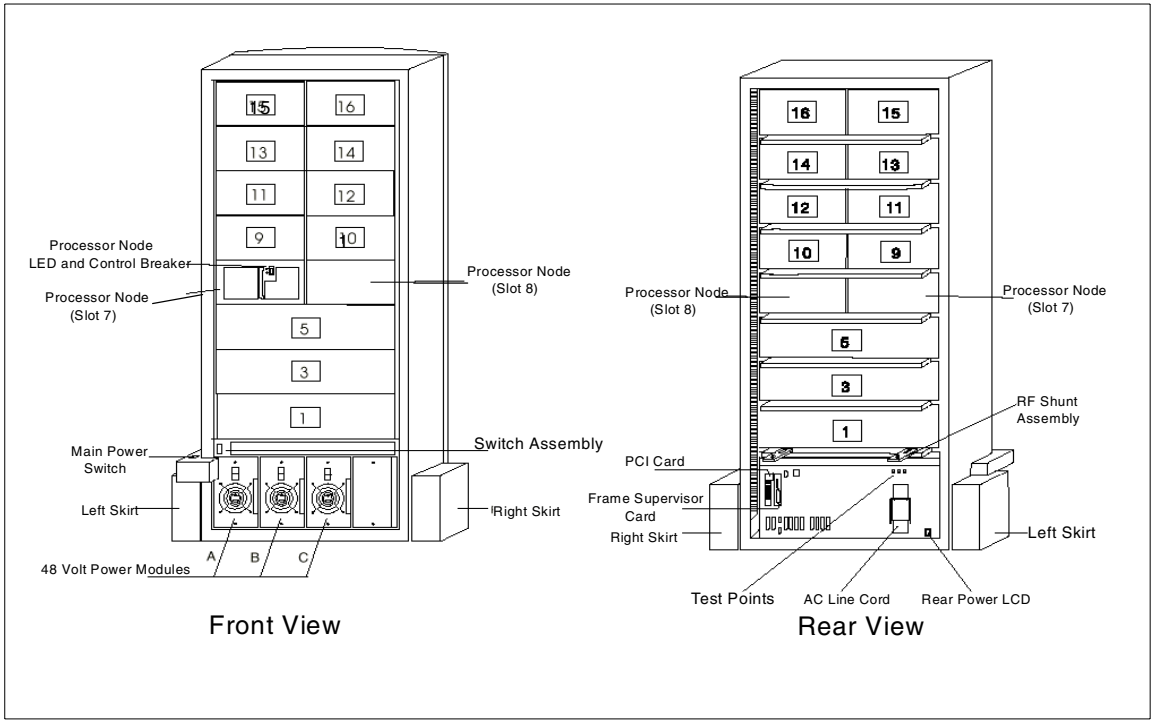


Figure 4-3 Tall frame from front and rear views

Tall frames have eight drawers and can house up to sixteen thin nodes, eight wide nodes, or four high nodes. All node types can be mixed in these frames. They also accommodate a switch. These frames are designed for concurrent maintenance; each processor node can be repaired without interrupting operations on other nodes.

Tall frames contain a SEPBU power subsystem to accommodate all SP processor nodes. The SEPBU is equipped with redundant (N+1) power supply modules; if one module fails, another takes over. The self-regulating SEPBUs are designed for concurrent maintenance; a failed power module can be removed and repaired without interrupting running processes on the nodes.

For improved system availability, you can use two separate ac power sources for your SP system using the SP frame redundant power supply (F/C 3885, 3886).

These 1.93 m frames are completely compatible with all existing SP systems. All existing nodes and SP switch types can be directly installed as long as configuration rules are not violated.

However, High Performance switches are withdrawn from production and are not compatible with most current SP system hardware.

When you order a frame feature, you receive the frame with its integral SEPBU and ac power cable. All nodes, switches, and other auxiliary equipment must be ordered separately.

Note: In order to maintain your entire SP system at the same electrical potential, you must attach a frame-to-frame ground between all frames in your SP system using IBM-supplied cables (P/N 46G5695).

4.3 Model 550 frames

The *model frame* is always the first frame in an SP system and it designates the type or model class of your SP system. The base level Model 550 SP system has a tall frame with eight empty node drawers and a 10.5 kW, three-phase SEPB power subsystem.

All processor nodes and optional switches must be purchased separately for these frames. One switch and up to sixteen thin nodes, eight wide nodes, or four high nodes, can be installed in these frames. Other frames that connect to the model frame are known as *expansion frames*.

Model 550 SP systems can be developed into several different configurations. Some of these configurations use switches and some do not.

4.3.1 Model 550 non-switched configuration

This configuration consists of 1-64 processor nodes mounted in one required Model 550 frame and in additional tall, non-switched expansion frames (F/C 1550).

4.3.2 Model 550 SP Switch-8 configuration

This configuration consists of one to eight processor nodes mounted in one required Model 550 frame equipped with an eight-port SP Switch-8 (F/C 4008). A non-switched expansion frame (F/C 1550) is supported in this configuration only if the model frame is filled before the total node count of eight is reached. In this configuration, each node requires either an SP Switch Adapter (F/C 4020), an SP Switch MX Adapter (F/C 4022), or an SP Switch MX2 Adapter (F/C 4023). Nodes in the non-switched expansion frames share unused switch ports in the model frame.

4.3.3 Model 550 single-stage SP Switch and SP Switch2 configuration

This configuration consists of 1 to 8 processor nodes mounted in one required Model 550 frame equipped with a sixteen-port switch. Depending on the number of nodes in your system, up to four switched expansion frames can be added to the system.

Single-stage system configurations can also utilize non-switched expansion frames. Nodes in the non-switched expansion frames share unused switch ports both in the model frame and in the switched expansion frames.

4.3.4 Model 550 two-stage SP Switch and SP Switch2 configuration

The standard two-stage switch configuration has 65-128 processor nodes. No more than 64 of the 128 nodes can be high nodes.

These nodes are mounted in one required Model 550 frame equipped with a sixteen-port switch and in switched expansion frames. The switches in these frames form the first switching layer.

This system configuration also requires a switch frame, which forms the second switch layer. The second stage switches in the switch frame are used for high-performance parallel communication between the switches mounted in the model frame and in switched expansion frames. Switch traffic is carried through concurrent data transmissions using the Internet Protocol (IP).

Two-stage switch system configurations can also utilize non-switched expansion frames. Nodes in the non-switched expansion frames share unused switch ports in the model frame and in switched expansion frames.

Note: Alternate two-stage switch configurations mounting fewer than 65 nodes or more than 128 nodes are available. Two-stage configurations using less than 65 nodes are simpler to scale up when you add more switches than are single-stage switch configurations.

4.4 Model 555 standalone switch for clustered nodes

The Model 555 is a tall frame with an integral SEPBU power supply, one SP Switch (F/C 4011), but no processor nodes. This frame provides node interconnectivity for Clustered Enterprise Server systems for configurations of 2 to 16 servers.

For configurations of 17 to 32 servers, add another SP Switch (F/C 4011) in the 9076 model 555. The configurator then adds an SP Expansion frame (F/C 1550), which contains an SP Switch. The SP Switch ports in these frames connect to the SP System Attachment adapters (RS/6000 F/C 8396) in the CES servers. Figure 4-4 on page 149 shows a 9076-555 model 555 with two SP-attached servers and a CWS.

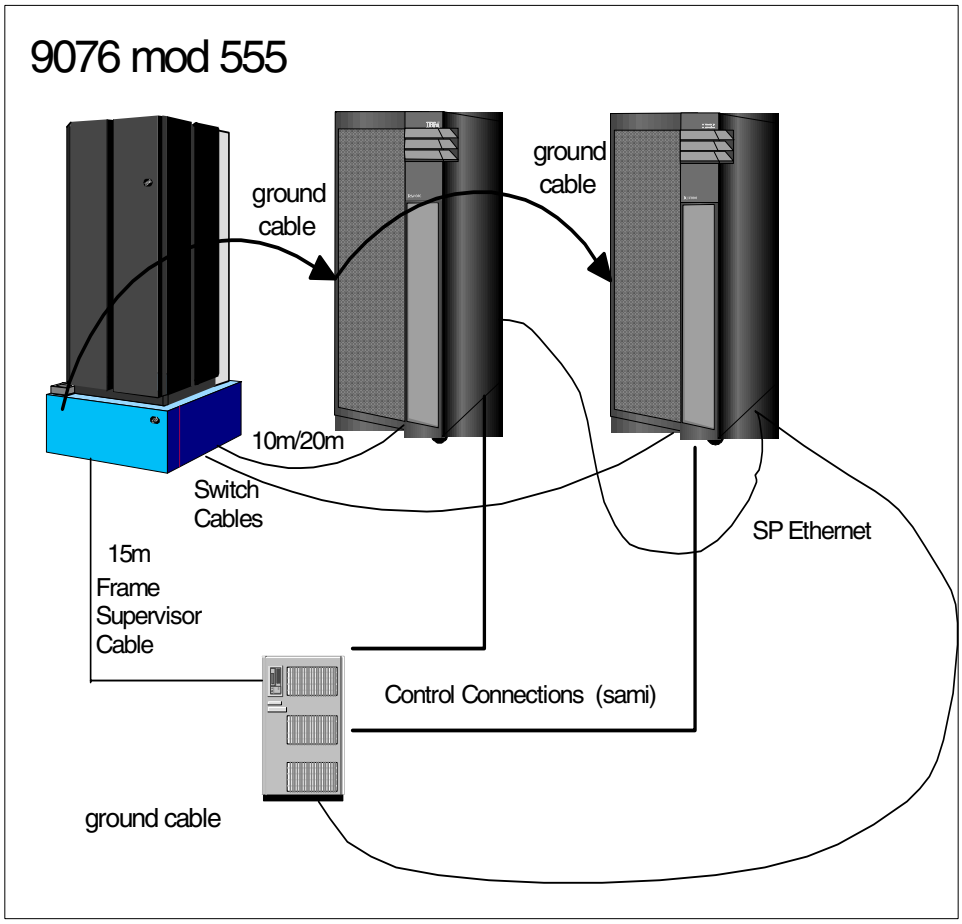


Figure 4-4 9076 mod 555 with two SP-attached servers and a CWS

4.5 Expansion frames (F/C 1550)

Used in Model 550, 3BX, 20X, 30X, 40X, and 55H systems, F/C 1550 is a tall frame with eight empty node drawers and a SEPBU power supply. All processor nodes and optional switches must be purchased separately for these frames. An SP Switch and up to 16 thin nodes, eight wide nodes, or four high nodes can be installed in these frames.

F/C 1550 is the base offering for the 1.93 m (75.8 inch) SP expansion frame. These frames are equipped with a 10.5 kW three-phase SEPBU self regulating power subsystem. All 1.93 m frames have eight empty node drawers for separately purchased nodes. Up to 16 thin nodes, eight wide nodes, or four high nodes can be installed in these frames.

You must populate each expansion frame with optional SP switches and nodes, as permitted by system configuration rules. These configuration rules impose limits on the number and location of each type of node and switch that can be included in each system and vary depending on how your SP model frame is configured.

There are two standard configurations for F/C 1550 expansion frames. These are:

1. An expansion frame configured with processor nodes only; known as a *non-switched expansion frame*.
2. An expansion frame configured with processor nodes and an SP switch; known as a *switched expansion frame*.

Non-Switched Expansion Frame Terminology:

Non-switched expansion frames were sometimes previously referred to as “logical switch expansion frames.”

Both of these terms indicate a shared-switch frame set. The frames of a shared-switch frame set include the first frame (containing the switch) and one or more additional frames, which do not contain a switch. The additional frames share the switch of the first frame in the set.

Uniprocessor and SMP Thin Node Pair Alert:

Frames that have uniprocessor thin nodes cannot be used as non-switched expansion frames. Similarly, if a frame has uniprocessor thin nodes and a switch with unused switch ports, it cannot have a non-switched expansion frame attached to the unused switch ports. Frames having uniprocessor thin nodes require an SP switch for expansion.

SMP thin nodes can, however, be used as single thin nodes. Frames having *single* SMP thin nodes *only* can be used for expansion and switch sharing. However, if a pair of SMP thin nodes is used, uniprocessor thin node rules apply to expansion and switch sharing

4.5.1 Using F/C 1550 non-switched expansion frames

A non-switched expansion frame is defined as a base offering expansion frame equipped with processor nodes only. Some Model 550 SP system configurations can be scaled into larger systems using these frames.

These SP system configurations are:

- Configurations using Model 550 frames equipped with processor nodes and a switch.
- Configurations using switch-configured Model 550 frames and F/C 1550 expansion frames equipped with processor nodes and a switch (switched expansion frames).
- Model 550 frames equipped with processor nodes only.

Frames containing processor nodes and a switch

Non-switched expansion frames are added to SP frames configured with processor nodes and a switch to take advantage of unused switch ports resulting from certain system configurations. These unused switch ports can be in the model frame or in switched expansion frames. In these cases, the switch, which can have ports to attach up to 16 nodes, is not fully utilized.

One example of an under-utilized switch would be a tall frame with eight wide nodes and an SP Switch. In this example, the frame is fully populated, yet only eight of the sixteen switch ports are used. In this case, you can add non-switched expansion frames to the switch-configured frame to take advantage of the eight unused switch ports. In these cases, node-to-node data transfers are completed through the switch.

Note: If the switch used in this configuration is an SP Switch-8, the nodes must be placed sequentially in this system configuration. Empty node drawers are not allowed. If the switch is a 16-port SP Switch, nodes can be placed in any order and empty drawers are permissible; however, node-to-node Ethernet BNC jumper cables might be too short if a drawer is left empty.

Frames containing processor nodes only

Non-switched expansion frames (configured with processor nodes only) can be added to the Model 550 SP frame to take advantage of unused node slots. In these cases, node-to-node data transfers are completed over the SP LAN.

4.5.2 Using F/C 1550 switched expansion frames

A switched expansion frame is defined as a base offering expansion frame equipped with processor nodes and a switch. These frames are added to SP systems with switch-configured Model 550 frames. Configuration rules permit you to attach up to four switched expansion frames to these model frames. In some system configurations, you can have unused switch ports in either the model frame or the switched expansion frames. Those unused switch ports can be used with non-switched expansion frames to complete your system.

- If your SP system uses single-stage switching, you can scale your SP system into a system containing up to 80 nodes.
- If your SP system uses two-stage switching, you can scale your SP system even larger. In these systems, 128 nodes (or more) are supported.

4.6 SP Switch frames (F/C 2031, 2032)

Switch frames are base, tall frames with integral SEPBU power supply equipped with switches but no processor nodes. A switch frame is required for systems having more than five switches; it interconnects the switches in the model and expansion frames. Figure 4-5 illustrates the SP Switch frame from the front view.

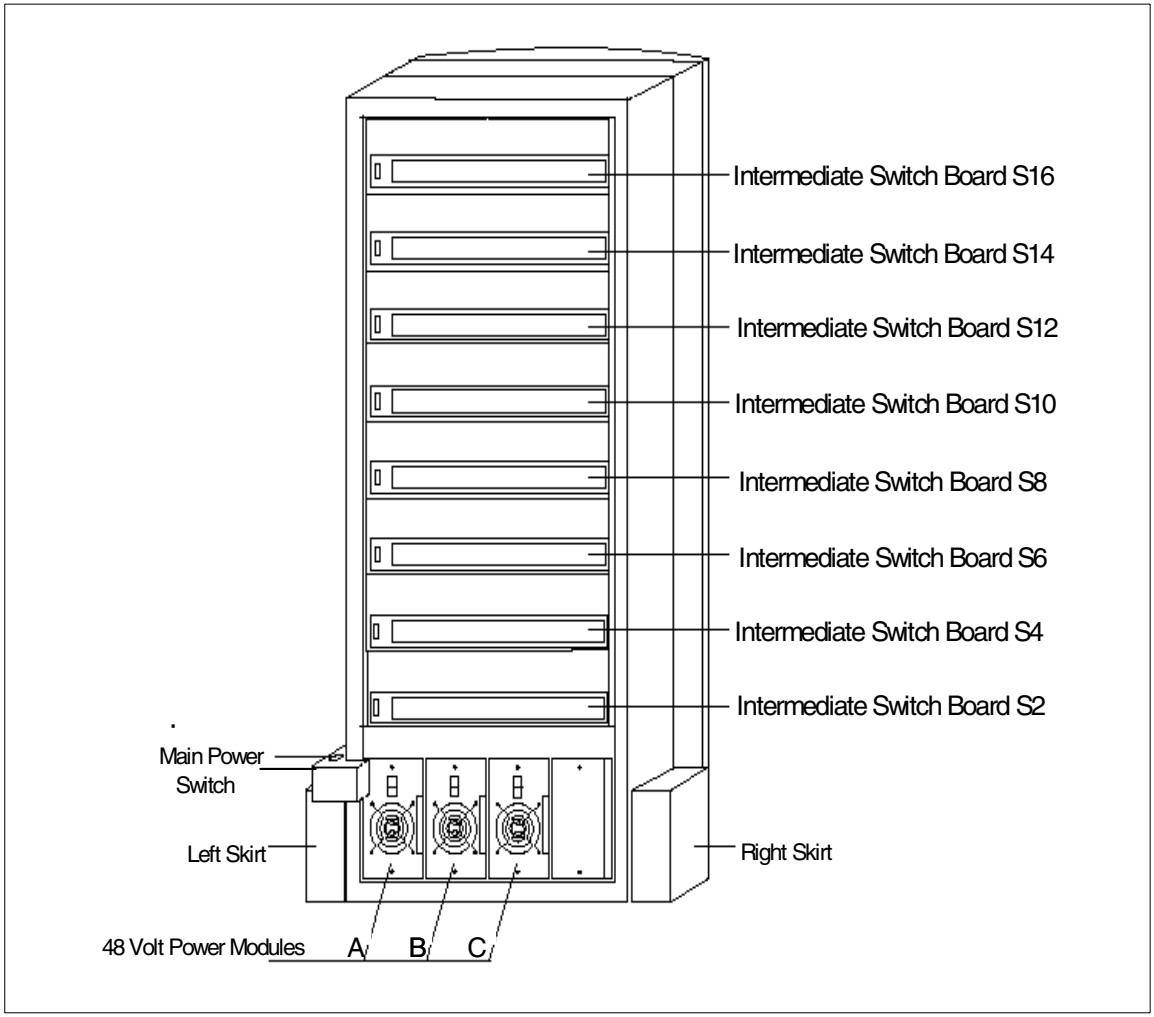


Figure 4-5 SP Switch frame front view

4.7 Short frame

Short frames are available in two variations:

- 1. Short *model* frames (Model 500)
- 2. Short *expansion* frames (F/C 1500)

Figure 4-3 illustrates a short frame from a front view.

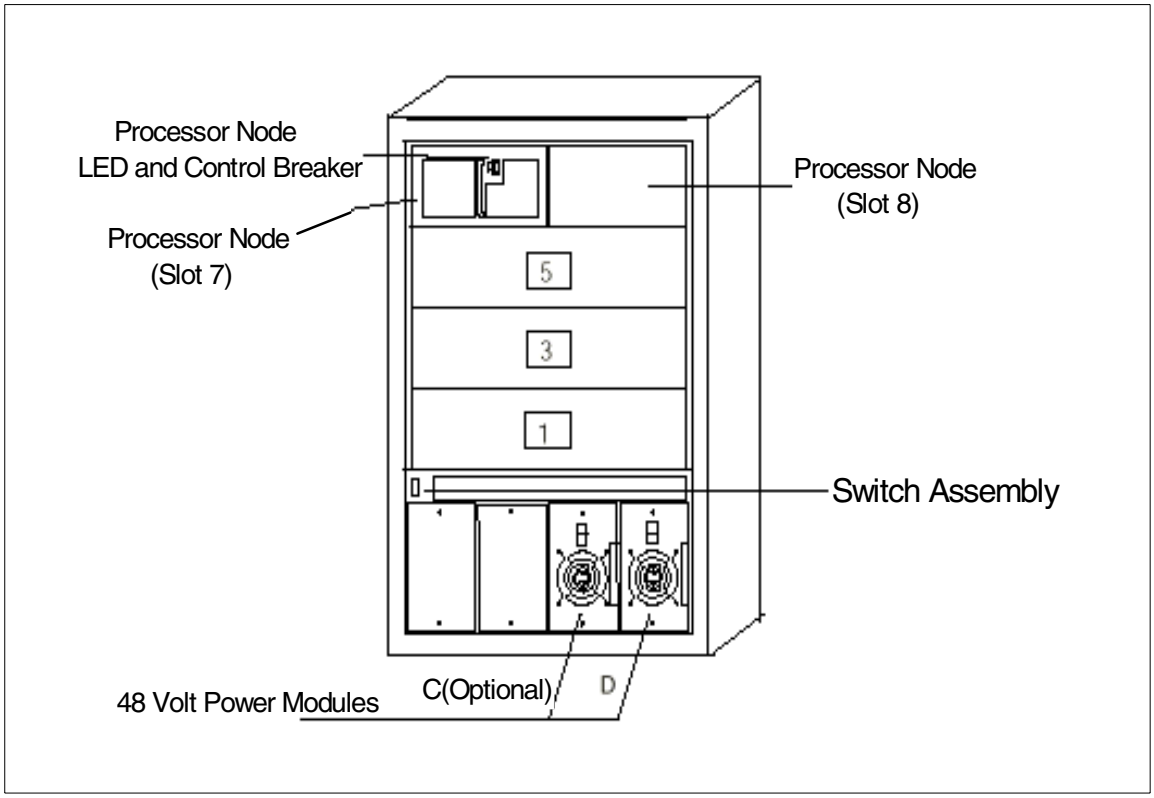


Table 4-3 Short frame front view

All short frames (including previous feature codes) are 1.25 m in height. However, the short frame Scalable Electric Power Base Unit (SEPBU) power subsystem has been redesigned to accommodate the latest SP processor nodes. This upgrade is required before you can use SMP wide nodes and SMP thin nodes in older frames.

Short frames have four drawers and can house up to eight thin nodes or four wide nodes. All node types can be directly installed as long as configuration rules are not violated. However, note that POWER3 SMP High Nodes cannot be installed in these frames due to a depth limitation. Short frames can also accommodate an eight-port switch. All frames are designed for concurrent maintenance; each processor node can be repaired without interrupting operations on other nodes.

Short frames contain a single-phase SEPBU power subsystem which provides 5.0 kW output power. Redundant power (F/C 1213) is an option with the short-frame SEPBU. With this option, if one power supply fails, another takes over. These self-regulating SEPBU units with the N+1 feature are also designed for concurrent maintenance; if a power supply module fails, it can be removed and repaired without interrupting running processes on the nodes.

When you order a frame feature, you receive the frame with its integral SEPBU and AC power cable. All nodes, switches, and other auxiliary equipment must be ordered separately.

Note: In order to maintain your entire SP system at the same electrical potential, you must attach a frame-to-frame ground between all frames in your SP system using IBM cables (P/N 46G5695).

4.8 Model 500 frames

The *model frame* is always the first frame in an SP system and it designates the type or model class of your SP system. The base level Model 500 SP system has a short frame with four empty node drawers and a 5.0 kW, single-phase SEPBU power subsystem.

All processor nodes and the optional switch must be purchased separately for these frames. One SP Switch-8 and up to eight thin nodes, four wide nodes, or two 200 MHz High Nodes (withdrawn from production) can be installed in the Model 500 frame. Other frames that connect to the model frame are known as an *expansion frame*. Model 500 SP systems can be developed into two different configurations, either non-switched or switched.

4.8.1 Model 500 non-switched configuration

This configuration consists of one to eight processor nodes mounted in one required Model 500 frame and up to three additional short, non-switched expansion frames (F/C 1500).

4.8.2 Model 500 switched configuration

This configuration consists of one to eight processor nodes connected through a single eight-port SP Switch-8 (F/C 4008). These nodes are mounted in one required Model 500 frame containing the SP Switch-8 and in up to three additional short, non-switched expansion frames (F/C 1500). In this configuration, each node requires either an SP Switch Adapter (F/C 4020), an SP Switch MX Adapter (F/C 4022), or an SP Switch MX2 Adapter (F/C 4023).

Nodes in the non-switched expansion frames (F/C 1500) share unused switch ports in the model frame. When short frames are used in a switched configuration, only the Model 500 frame can be equipped with a switch. SP switches cannot be mounted in the F/C 1500 frames.

4.8.3 Expansion frames (F/C 1500)

F/C 1500 is the base offering for the 1.25 m (49 in.) SP expansion frame. These frames are equipped with a 5.0 kW single-phase SEPBU self-regulating power subsystem. All 1.25 m frames have four empty node drawers for separately purchased nodes. Up to eight thin nodes, four wide nodes, or two high nodes can be installed in these frames.

Note: Switches cannot be mounted in F/C 1500 expansion frames. You must populate each expansion frame with optional nodes as permitted by system configuration rules. These configuration rules impose limits on the number and location of each type of node that can be included in each system and vary depending on how your SP model frame is configured.

- F/C 1500 expansion frames can only be configured with processor nodes.
- Expansion frames that are configured with processor nodes only are known as a *non-switched expansion frames*.

4.8.4 Using F/C 1500 non-switched expansion frames

Model 500, 2AX, 3AX, or 50H SP systems can be fully utilized using F/C 1500 expansion frames. Model 500 systems have a capacity for up to eight nodes. If you fill the model frame before installing all eight nodes, you can install additional nodes in the system by using F/C 1500 non-switched expansion frames. The model frame can be either of the following:

- Configured with processor nodes and a switch
- Configured with processor nodes only

Expansion Frame Terminology: Non-switched expansion frames were sometimes previously referred to as “logical switch expansion frames.”

Both of these terms indicate a shared-switch frame set. The frames of a shared-switch frame set include the first frame (containing the switch) and one or more additional frames which do not contain a switch. The additional frames share the switch of the first frame in the set.

4.8.5 Configurations with processor nodes and a switch

Non-switched expansion frames can be added to the Model 500 SP frame configured with processor nodes and an SPS-8 Switch to take advantage of unused switch ports.

One example of an under-utilized switch is a Model 500 frame with four wide nodes and an SP-8 Switch. In this case, the frame is fully populated yet only four of the eight switch ports are used. In this case, you can add non-switched expansion frames to the model frame to take advantage of the four unused switch ports. In these systems, node-to-node data transfers are completed through the switch.

Note: Nodes must be in sequence in this system configuration. Empty node drawers are not allowed.

4.8.6 Configurations with processor nodes only

Non-switched expansion frames can be added to the Model 500 SP frame configured with processor nodes only (no switch) to take advantage of unused node slots. Model 500 systems have a capacity for up to eight nodes. If you fill the model frame by placing two high nodes in that frame, you can install six additional nodes in the system by using F/C 1500 non-switched expansion frames. In these systems node-to-node data transfers are completed over the SP LAN.

5

The Switch

The Switch network is the most efficient network available in SP system and Clustered Enterprise Servers.

This chapter describe the two available switch series:

- The SP Switch or SPS (F/C 4011 and F/C 4008).
- The SP Switch2 or SPS2 (F/C 4012).

This chapter does not describe the previous High Performance Switch (HiPS) series.

5.1 SP Switch

The design objectives for the SP Switch network is low-latency, scalable high-bandwidth, and fault-tolerant communication to support the demands of parallel and client-server applications.

The SP Switch provides the message passing network that connects all nodes with a minimum of four paths between any pair of nodes.

The hardware component that supports this communication network consists of two basic components: The SP Switch adapter and the SP Switch board. There is one SP Switch adapter per processor node and generally one SP Switch board per frame. This setup provides connections to other processor nodes. Also, the SP system allows switch boards-only frames that provide switch-to-switch connections and greatly increase scalability. Figure 5-1 shows the front view of the SP Switch board.

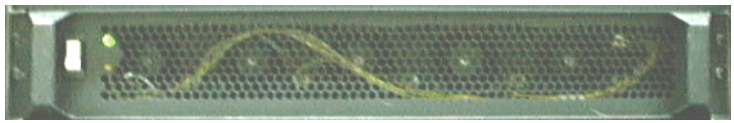


Figure 5-1 Front view of the SP Switch board

5.2 SP Switch components

This section discusses the hardware design that makes up the SP Switch network: The SP Switch link, the SP Switch port, the SP Switch chip, the SP Switch board, and the SP Switch adapter. The SP Switch link itself is the physical cable connecting two SP Switch ports. The SP Switch ports are hardware subcomponents that can reside on an SP Switch adapter that is installed in a node or on an SP Switch chip that is part of an SP Switch board.

5.2.1 SP Switch link or data cable

An SP Switch link connects two network devices. It contains two channels carrying packets in opposite directions. Each channel includes 8 bits Data, 1 bit Data Valid, and 1 bit Token signal.

5.2.2 SP Switch ports

An SP Switch port is part of a network device and is connected to other SP Switch ports through the SP Switch link. The SP Switch port includes two ports (input and output ports) for full duplex communication. The following diagram shows the relationship between the SP Switch link and the SP Switch port. Figure 5-2 shows the data flow between switch ports.

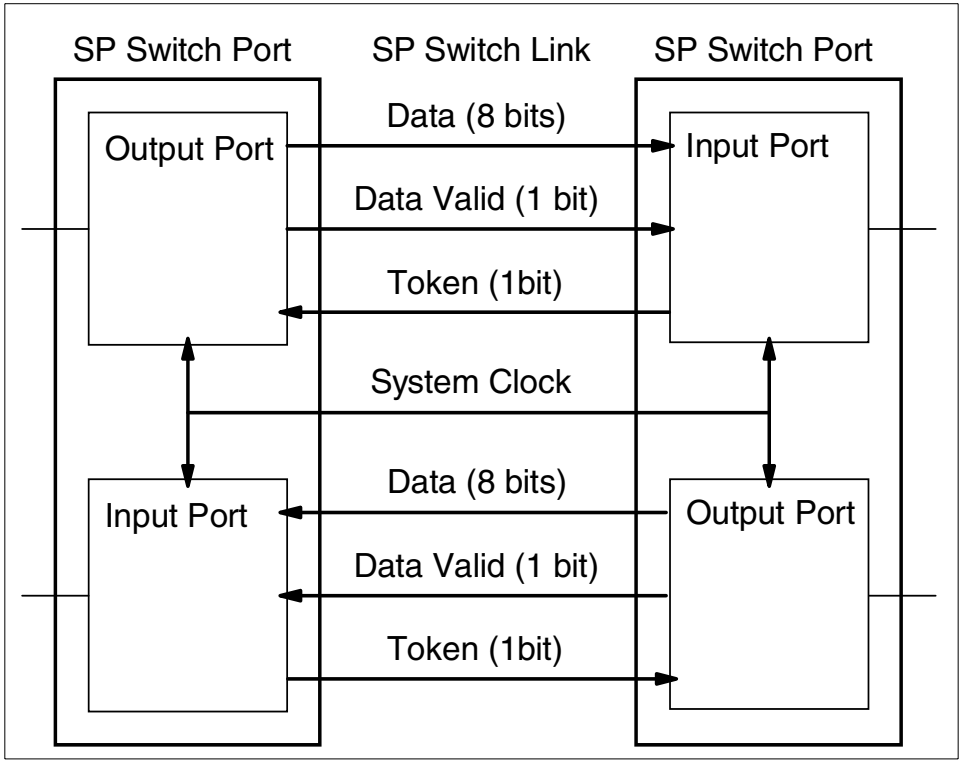


Figure 5-2 Data flow between switch ports

5.2.3 SP Switch chip

An SP Switch chip contains eight SP Switch ports, a central queue, and an unbuffered crossbar, which allows packets to pass directly from receiving ports to transmitting ports. These crossbar paths allow packets to pass through the SP Switch (directly from the receivers to the transmitters) with low latency whenever there is no contention for the output port. As soon as a receiver decodes the routing information carried by an incoming packet, it asserts a crossbar request to the appropriate transmitter. If the crossbar request is not granted, it is dropped (and, hence, the packet will go to the central queue). Each transmitter arbitrates

crossbar requests on a least recently served basis. A transmitter will honor no crossbar request if it is already transmitting a packet or if it has packet chunks stored in the central queue. Minimum latency is achieved for packets that use the crossbar.

Figure 5-3 on page 162 shows the SP Switch chip.

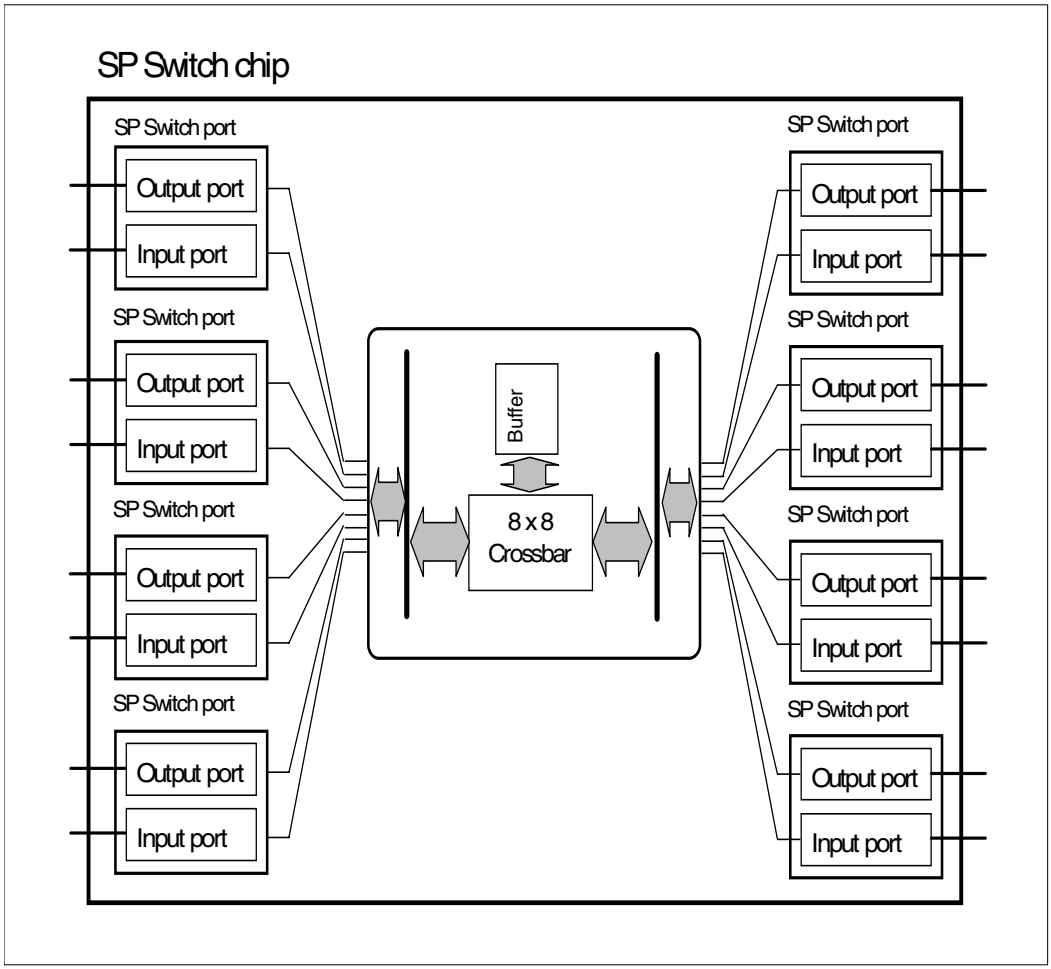


Figure 5-3 SP Switch chip

5.2.4 SP Switch board

An SP Switch board contains eight SP Switch chips that provide connection points for each of the nodes to the SP Switch network as well as for each of the SP Switch boards to the other SP Switch boards. There are 32 SP Switch ports in total. Of these, 16 could be connected to nodes and the other 16 to other SP Switch boards. The SP Switch board is mounted in the SP Frame. Figure 5-4 shows the interconnections used on the switch board.

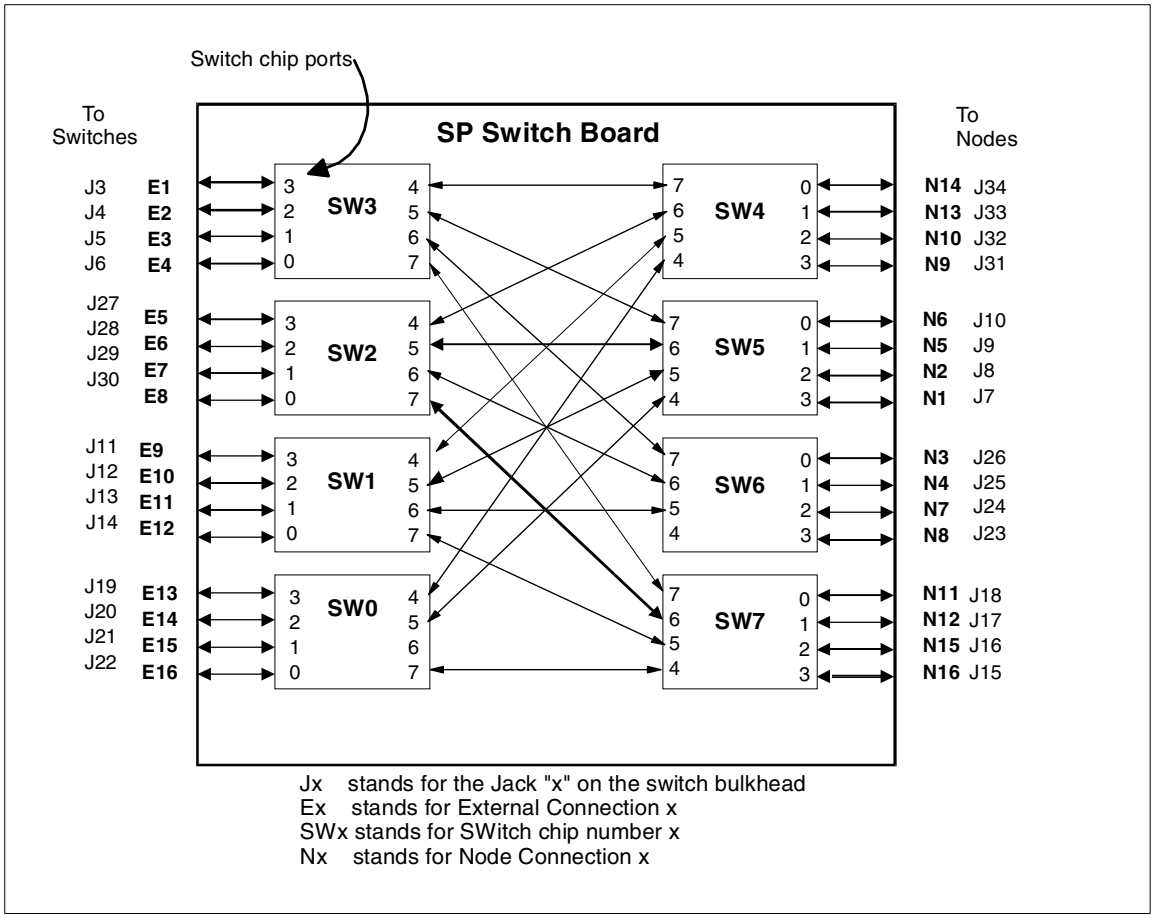


Figure 5-4 Interconnections used on the switch board

5.2.5 SP Switch adapter

An SP Switch adapter includes one SP Switch port that is connected to an SP Switch board. The SP Switch adapter is installed on an SP node.

An SP Switch adapter is designed for nodes that have both a PCI bus and an MX or MX2 bus.

Figure 5-5 on page 164 provides a logical diagram of an SP Switch adapter.

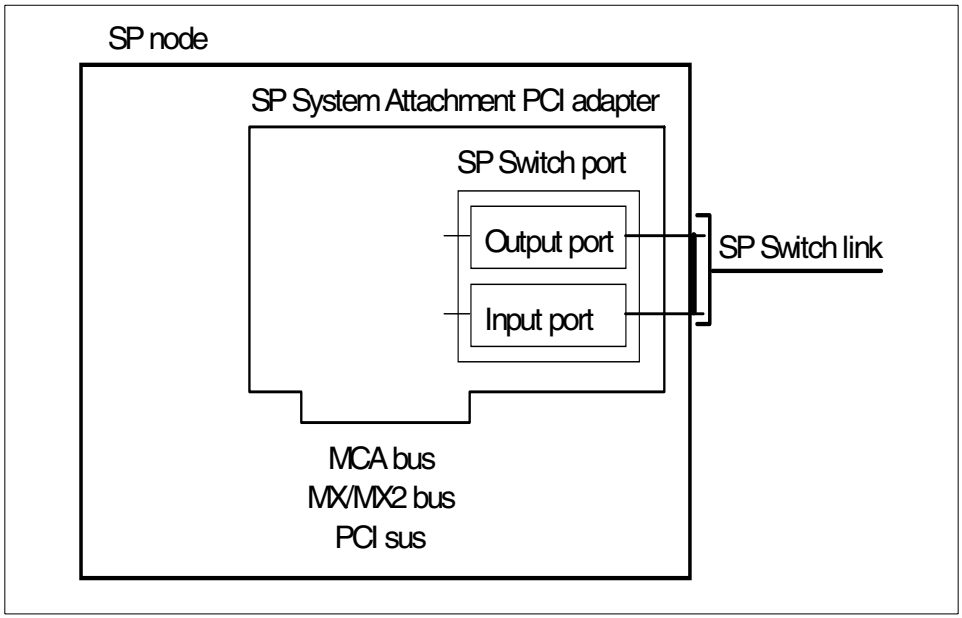


Figure 5-5 logical diagram of an SP Switch adapter

5.2.6 SP Switch system

Figure 5-6 on page 165 shows a SP Switch system with one SP frame, there are 16 nodes equipped with SP Switch adapters and one SP Switch board.

Sixteen SP Switch adapters are connected to 16 of 32 SP Switch ports in the SP Switch board. Another 16 SP Switch ports are available for other SP Switch boards; see Section 5.7.2, “Switch-to-switch cabling” on page 171 for the possible Switch to Switch connections.

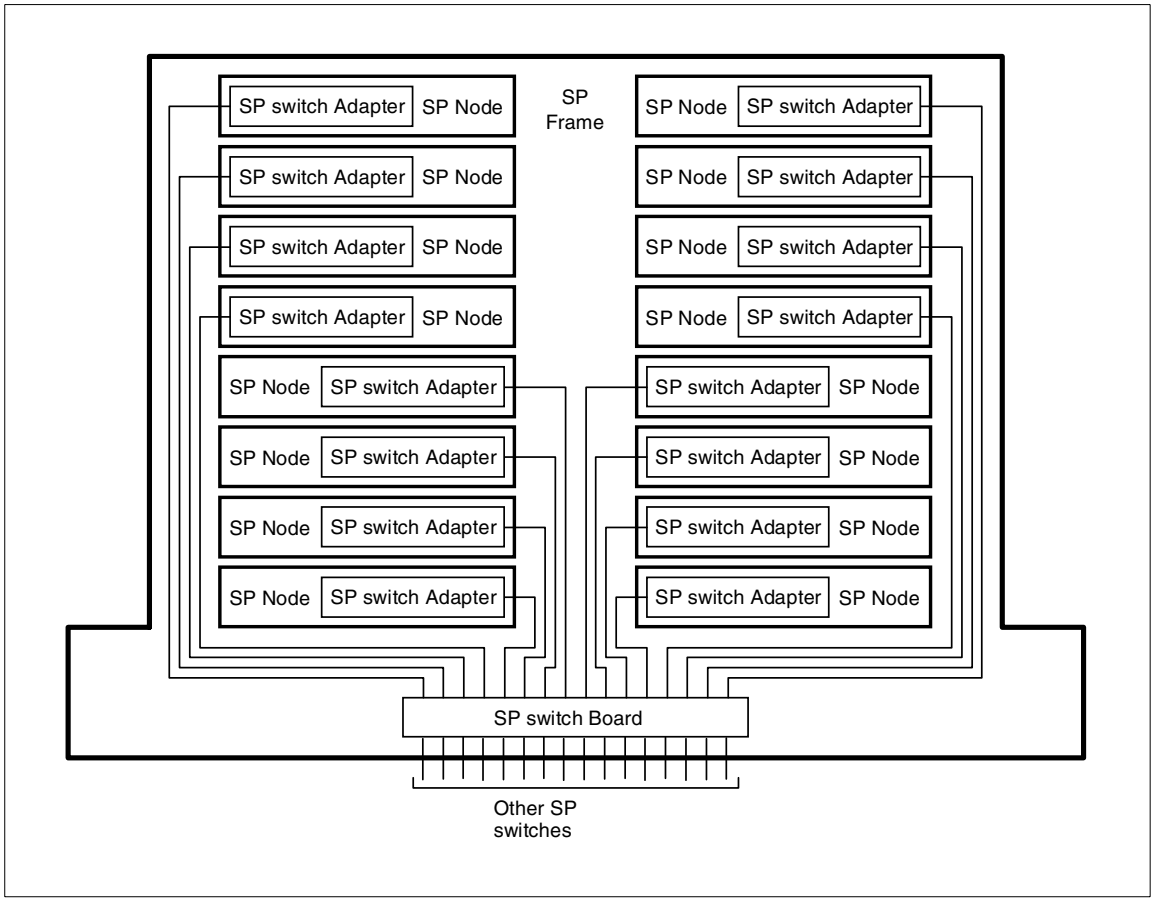


Figure 5-6 SP Switch system

Switch software considerations

Switch planning involves many issues; node placement, node addressing, and system partitioning are a few that you must consider when planning your switch layout. Make certain that you consult the *Planning Volume 2, Control Workstation and Software Environment*, GA22-7281 for software-related switch planning issues before you finalize your switch plans. Also refer to Chapter 10, “Clustering configurations” on page 285 for clustering configurations and nodes and adapter placement configuration rules.

Table 5-1 shows the AIX, PSSP, and the switches software level matrix relationship.

Table 5-1 AIX/PSSP/switches level matrix

	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3	PSSP 2.1	PSSP 2.2	PSSP 2.3	PSSP 2.4	PSSP 3.1.1	PSSP 3.2	HP Switch	SP Switch	SP Switch2
PSSP 2.1	X									X	X*	
PSSP 2.2	X	X		X						X	X	
PSSP 2.3		X	X	X	X					X	X	
PSSP 2.4		X	X	X	X	X				X	X	
PSSP 3.1.1			X		X	X	X				X	
PSSP 3.2			X								X	X

* Additional PTFs associated with APAR IX56314 are required for this configuration. Later modification levels of AIX and PSSP continue to support the latest levels of the other.

5.3 SP Switch, 16-port (F/C 4011)

There are two SP series of switch:

1. SP Switch, 16-port (F/C 4011)

2. SP Switch-8, 8-port (F/C 4008)

This section contains information on the SP Switch, 16-port (F/C 4011).

The SP Switch provides low-latency, high-bandwidth communication between nodes, supplying a minimum of four paths between any pair of nodes. The SP Switch can be used in conjunction with the SP Switch Router to dramatically increase the speed for TCP/IP, file transfers, remote procedure calls, and relational database functions. The required SP Switch Adapter (F/C 4020), SP Switch MX Adapter (F/C 4022), or SP Switch MX2 Adapter (F/C 4023) connects

each SP node to the SP Switch subsystem. One adapter of the required type must be ordered for each node in a switch-configured SP system. If you are using switch expansion frames, the SP Switch subsystem will allow you to scale your SP system up to 128 nodes.

When you order F/C 4011, you receive one 16-port SP Switch and all of the switch-to-node cables you need to connect the switch ports to up to sixteen nodes, both within the switch-equipped frame and in any non-switched expansion frames. You must specify the length of all switch-to-switch cables that make the connections between switch-equipped frames and to an SP Switch Frame.

Note:

- An SP Switch Router adapter is needed to connect the RS/6000 SP Switch Router to the SP Switch.
- An RS/6000 SP System Attachment adapter is needed to connect the SP-attached server and Clustered Enterprise Servers to the SP Switch.

5.4 SP Switch-8 (F/C 4008)

Eight port switches are a lower-cost alternative to the full-size 16 port switches. The 8-port SP Switch-8 (F/C 4008) provides switch functions for up to eight processor nodes in Model 500 and Model 550 systems. N+1 internal power for the SP Switch-8 is available using F/C 1212.

When you order F/C 4008, you receive one 8-port SP Switch and all of the cables you need to connect the switch ports to up to eight nodes, both within the switch-equipped frame and in any non-switched expansion frames.

SP Switch-8 can be configured in one of two ways:

- 1. In a Model 500 (1.25 m) frame with up to four F/C 1500 (1.25 m) non-switched expansion frames attached.
- 2. In a Model 550 (1.93 m) frame with F/C 1550 a non-switched expansion frame supporting up to a total of eight nodes.

The SP Switch-8 has two active switch chip entry points. Thus, your ability to create system partitions is restricted with this switch. With the maximum of eight nodes attached to the switch, you have two possible system configurations:

- A single partition with up to eight node system
- Two system partitions, with up to four nodes each

For upgrades to greater than eight node support, the SP Switch-8 is replaced by the 16 port SP Switch (F/C 4011). The SP Switch uses a similar network topology, proprietary protocol and communication physical layer as the SP Switch-8.

Note:
The SP-attached server, Clustered Enterpise Servers, and POWER3 High Nodes cannot be attached to the SP Switch-8.

5.5 SP Switch adapter (F/C 4020, 4022, and 4023)

If you plan to use a switch in your SP system, you will need a switch adapter to connect each RS/6000 SP node to the switch subsystem. SP Switches use either the SP Switch adapter (F/C 4020) for MCA-type nodes or the SP Switch MX Adapter (F/C 4022) or SP Switch MX2 Adapter (F/C 4023) for PCI-type nodes. One switch adapter is needed for each node in the SP system. Table 5-2 shows the SP Switch adapter features.

Attention:

- The SP Switch MX adapter (F/C 4022) has been withdrawn from production. This adapter has been replaced with the SP Switch MX2 adapter (F/C 4023).
- High Performance Switch adapters are not compatible with any of the SP Switch adapters. These adapters cannot coexist in the same system configuration.

Table 5-2 SP Switch adapter features

Adapter Feature Code	Description
4020	SP Switch adapter for installation as follows: ► Optional ► Order one adapter per MCA node
4022	SP Switch MX adapter (SP switch adapter for 332 MHz SMP nodes) for installation as follows: ► Optional ► Order one adapter per PCI node ► Withdrawn 10/98

Adapter Feature Code	Description
4023	SP Switch MX2 adapter (SP switch adapter for POWER3 SMP nodes and 332 MHz SMP nodes) for installation as follows: ► Optional ► Order one adapter per PCI node

5.6 Special SP Switch adapter (F/C 4021 and F/C 8396)

Some optional SP system equipment requires special adapters in order to connect to the SP Switch network. These devices are:

- SP Switch Router (Machine Type 9077)
- Clustered Enterprise Servers

5.6.1 SP Switch Router (Machine Type 9077)

The SP Switch Router has the following requirements:

- Requires one SP Switch Router adapter (F/C 4021) for each SP Switch connection. This adapter is ordered with the SP Switch Router.
- The SP Switch Router adapter is placed in the SP Switch Router. The included cable attaches to the SP Switch and uses up one valid node slot on the switch.

5.7 Switch cabling

The SP/CES systems in switched configurations require a number of cables that run between switches and processor nodes and between each switch and any switch frame in multi-switch systems.

Note: SP systems with up to five SP Switches (single-stage switch configuration) do not require a dedicated switch frame. Systems with more than five SP switches (two-stage configuration) do require an SP Switch frame (F/C 2031 or F/C 2032.).

For systems that require only single-stage switching, you might want to initially install an SP Switch Frame if you plan to scale up to a two-stage configuration at a later date. Switch cabling is much simpler in a two-stage switch configuration, due to the dedicated switch frame. Installing an SP Switch Frame initially greatly simplifies future expansion by eliminating the need to re-cable your switch fabric every time you add another switch.

5.7.1 Switch-to-node cabling

Switch-to-node cables are ordered and shipped with the system according to configurator rules.

The switch ports in a switch-equipped frame must be connected to all the nodes in that frame as well as to all the nodes in its non-switched expansion frames.

As an example, consider a Model 550 frame containing four high nodes and an SP Switch. The switch has 16 ports; the four nodes within that model frame are connected to four of the switch ports with four 2 meter cables. The rest of the ports can be used by up to twelve additional nodes mounted in non-switched expansion frames. In this example, twelve 10 meter cables are provided to connect the switch to the nodes in the non-switched expansion frames.

Use the following cables to connect switch ports to processor nodes:

- F/C 9302 – 2 meter switch-to-node cable (for nodes within the switch-equipped frame)
- F/C 9310 – 10 meter switch-to-node cable (for nodes in non-switched expansion frames)

The cables may be ordered when configure the processor nodes on SP system.

Use the following cables to connect switch ports to SP-attached servers and Clustered Enterprise Servers:

- F/C 9310 – 10 meter switch-to-node cable
- F/C 9320 – 20 meter switch-to-switch cable

The cables for the SP-attached servers may be ordered when you configure the SP-attached server as a RS/6000 Server attached frame (F/C 9123) and as a RS/6000 Server attached node (F/C 9122) on the SP system.

The cables for the Clustered Enterprise Servers may be ordered when you configure the Clustered Enterprise Server as a RS/6000 Server attached frame (F/C 9123) and as a RS/6000 Server attached node (F/C 9122) on the 9076-555 Standalone Switch for Clustered nodes.

5.7.2 Switch-to-switch cabling

The issues that you must consider when planning switch cable connections between switch-equipped frames in multi-switch SP systems include:

1. Determining the quantity of cables required
2. Determining the length of required cables
3. Minimizing cable lengths to reduce latency
4. Placing cables to reduce noise from other switch cables and ac sources
5. Making certain that your raised floor is high enough to contain all cables
6. Placing cables so that cooling air flow is unrestricted through raised floor space
7. Labeling and laying cables in an orderly manner, to allow for improved maintenance and reduced risk of mechanical failure
8. Placing cables and frames to allow for system growth

In SP systems with single-stage switching, each switch must be cable-connected to every other switch. You must determine the switch-cable lengths and quantities required for each of these connections.

Figure 5-7 illustrates single-stage configurations from two to five switches.

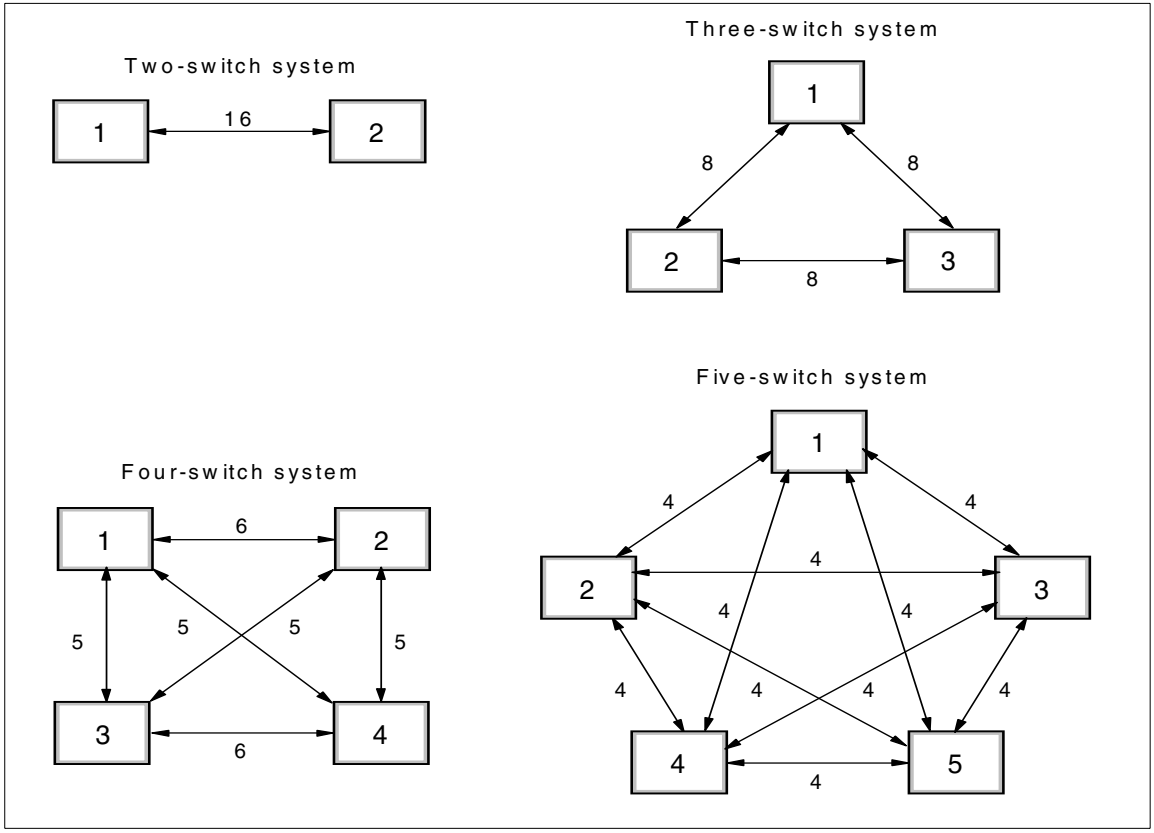


Figure 5-7 Single-stage switch cable paths

The numbered squares represent switch-equipped frames. The arrows are the cable sets with the quantity of individual cables in each set indicated. Switch-to-node cabling and non-switched expansion frames are not shown.

An SP Switch frame (F/C 2031 or F/C 2032) is required for systems using more than five switches (known as two-stage switching). Each switch in a two-stage system is cable-connected only to the SP Switch frame. Figure 5-8 illustrates an example of a two-stage switch configuration that has eight switch-equipped frames.

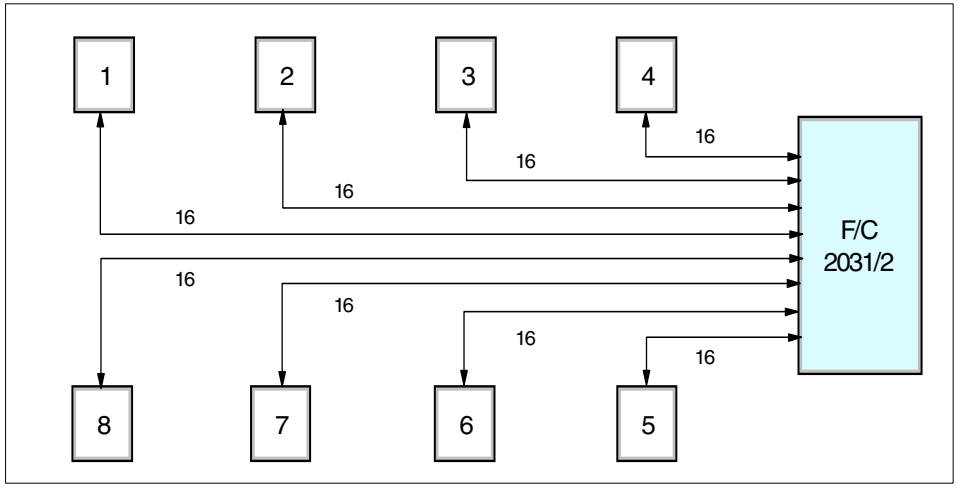


Figure 5-8 Two-stage switch cable paths

The quantity of switch cables depends on the quantity of switch-equipped frames in the system. Table 5-3 shows quantities of switch-to-switch cables required per SP system.

Table 5-3 Quantities of Switch-to-Switch cables required per SP system

Switch Configuration Type	Quantity of Switch-Equipped Frames (excluding SP Switch frame)	Quantity of Cables Between Any Two Switches	Total Quantity of Cables Required
Single-Stage	1	N/A	N/A
Single-Stage	2	16	16
Single-Stage	3	8	24
Single-Stage	4	5 / 6 (See Note 1)	32
Single-Stage	5	4	40
Two-Stage	6	16	96
Two-Stage	7	16	112
Two-Stage	8	16	128

Note:

1. Frames 1 to 2 and Frames 3 to 4 require six cables. All other paths require five.

2. Each cable is approximately 0.5 inches in diameter. We recommends that these cables be routed in a customer-supplied raceway or similar protective device to help prevent damage, both in raised and non-raised floor environments.

Use the switch cables listed here for switch-to-switch connections between switch-equipped frames. You can choose from the following lengths:

- F/C 9305: 5 meter (16 ft.)
- F/C 9310:10 meter (33 ft.)
- F/C 9315:15 meter (49 ft.)
- F/C 9320: 20 meter (66 ft.)

The cables for switch-to switch connections on the SP System may be ordered when you configure the frames with switches on the SP System.

The cables for switch-to-switch connections on the CES System may be ordered when you configure the 9076-555 Standalone Switch for Clustered Nodes with (2) F/C 4011.

Spending the time to calculate the ideal length for each cable enhances your SP/CES system operational efficiency by giving you optimal switch signal performance. Also, with less cable occupying the underlying floor area, there is an improved cooling air flow and better maintenance access.

5.8 SP Switch2

The SP Switch2 is the next step in the evolution of the SP interconnection fabric and offers enhanced performance and RAS (Reliability, Availability, and Serviceability) over the SP Switch. The performance gains are significant increases in bandwidth and reductions in latency. SP Switch2 provide a one-way bandwidth of up to 500 MB per second between nodes (1 GB bidirectional) for interconnecting POWER3 SMP high nodes.

Because the SP Switch2 design improvements are evolutionary steps on the SP and High Performance switches, the SP Switch2 is fully compatible with applications written for the older switches. The SP Switch2 software support continues to emphasize usage of MPI (Message Passing Interface), LAPI (Low level Application Interface), and IP (Internet Protocol), but the new switch subsystem enhances message handling so that the switch can keep pace with the increased performance of the POWER3 SMP High Node. Because of this, the SP Switch2 is currently only supported on those nodes.

Software enhancements

Several software component have been enhanced to support new functions available with the SP Switch2 and SP Switch2 adapter including:

- The Perspectives interface
- Switch management
- Software diagnostic

Perspectives

The Perspectives received the following updates:

System Notebook: Identification page shows the number of switch planes in use.

Switch Notebook:

- Configuration Page:
  - Switch name displays as SP_Switch2.
  - Switch supervisor type define as 132.
- Switch Notebook Status Page:
  - Shows the status of the four switch power supplies.
  - Shows the status of the four cooling fans.

Node Notebook

- Status Page: Need to monitor switchResponds0. The switchResponds attribute only applies to SP Switch (shows but does not apply to SP Switch2).
- Configuration Page:
  - Switch port number.
  - Switch chip.
  - Switch chip port are no longer displayed (moved to SP Adapter Page).
- SP Adapter Page:

- This is a new page listing for Ethernet, Token Ring, FDDI and CSS (Switch information) information.
- CSS attributes for SP Switch2 include: Adapter, Adapter port, Switch port number, Network address, Netmask, Subnet, Switch chip, Switch chip port, Switch number, Configuration status.

Switch management

Switch administration daemon provides node and global recovery. If the switch is started in a switch plane without a primary node, the switch administration daemon will select a new primary (or backup) and then run Estart.

- Switch Installation and configuration
  - The enhancement use “Discovery Node” to run switch initialization checkout: Calls out mis-wired switch to switch connections, Automatically builds the node portion of the switch topology (out.top) file.
- Logging, Event and error logging have been restructured.
  - Node level files: daemon.log, rc.switch.log, css.snap.log, Ecommands.log, logevnt.out, and daemon.log are in /var/adm/SPlogs/css directory.

In Single-Single configuration:

- Adapter level files: adapter.log, dtbx.trace, dtbx-failed.trace, scan-out.log, and scan-save.log are in /var/adm/SPlogs/css0 directory.
- Port level files: flt, fs-deamon-print.file, out.top, router.log, topology.data, and cable.miswire are in /var/adm/SPlogs/css0/p0 directory.

Software diagnostic

The SP Switch2 subsystem improves system Reliability, Availability, and Serviceability (RAS) by adding a number of new software diagnostic capabilities including:

- Node Location
- Differential line testing
- Mis-wire detection
- Build-in and functional self-tests
- Build-in port and cable wrap tests

5.8.1 SP Switch2 (F/C 4012)

Like the SP Switch, the SP Switch2 has 32 ports; 16 ports for switch-to-node connections and 16 ports for switch-to-switch connections. Management of the SP Switch2 network is the same as SP Switch, using service packets over the switch plane.

The N+1 feature of these switches allows for concurrent replacement of any failed power supplies or cooling fans. The supervisor can also be replaced while the switch is operating. PSSP level 3.2 software is required to use these switches.

This section describes some of the most important hardware characteristics of the SP Switch2 hardware.

TOD synchronization

The SP Switch2 is designed with multiple clock sources in the system. There is no master clock, unlike HPS or SP Switch. Therefore, instead of having clock selection logic, SP Switch2 switch has a separate oscillator for each switch chip. The TOD (time of day) logic on the SP Switch2 has also been significantly redesigned to simplify both coding and system requirements, as well as to improve the accuracy of the TOD across the system.

J-TAG interface

An additional interface, J-TAG interface, is added to the SP Switch2. This new interface will allow the supervisor to perform new functions such as writing initialization data to the switch chips and reading error status information back from the switch chip.

Adaptive routing

To support the double-port, a number of routing enhancements were added in SP Switch2, namely *adaptive routing* and *multicast packets*. Adaptive routing will allow the switch chip to determine which output port to route the packet to based on the route nibble. Multicast packets give the switch the ability to replicate and distribute packets to predefined groups of nodes.

When you order F/C 4012, you receive one SP Switch2 and, according to your SP system configuration, all of the switch-to-node cables you need to connect up to 16 POWER3 High Nodes. These nodes can be within the switch-equipped frame and in any non-switched expansion frames. You must specify the length of all switch-to-switch cables that make the connections between switch-equipped frames and to an SP Switch2 Frame (F/C 2032).

Each occupied switch port in the SP Switch2 contains an interposer card (F/C 4032). Interposer cards can be changed or added while the switch is operating.

Any unused switch ports must have blank interposer cards (F/C 9883) installed, which prevent contamination of the connector and ensure a proper cooling air flow.

Note:

- The SP Switch2 and its adapters are not compatible with the SP Switch or High Performance Switch series or their adapters; they cannot coexist in the same SP system.
- Only POWER3 and 375 MHz POWER3 High Nodes can be connected to the SP Switch2; no other node types can be included in the switch fabric.
- The SP Switch2 cannot be connected to an SP-attached server, Clustered Enterprise Server or to an SP Switch Router.

5.8.2 SP Switch2 adapter (F/C 4025)

Like the SP Switch Adapter, the SP Switch2 Adapter is controlled by node software and on-card microprocessor microcode. The adapter has two ports. Only one port (the upper port) is fully functional, the second is reserved. The configuration is called a *Single-Single* configuration.

The adapter has several design improvements including an on-board RAM memory component and a hardware driven datagram segmentation and reassemble engine to proceed 64KB IP datagram. To improve data flow through the adapter, the design include separate paths for microprocessor operations and data transfers and implement four 1000 MB/s uni-direction buses. Figure 5-9 on page 179 shows the SP Switch2 adapter hardware structure.

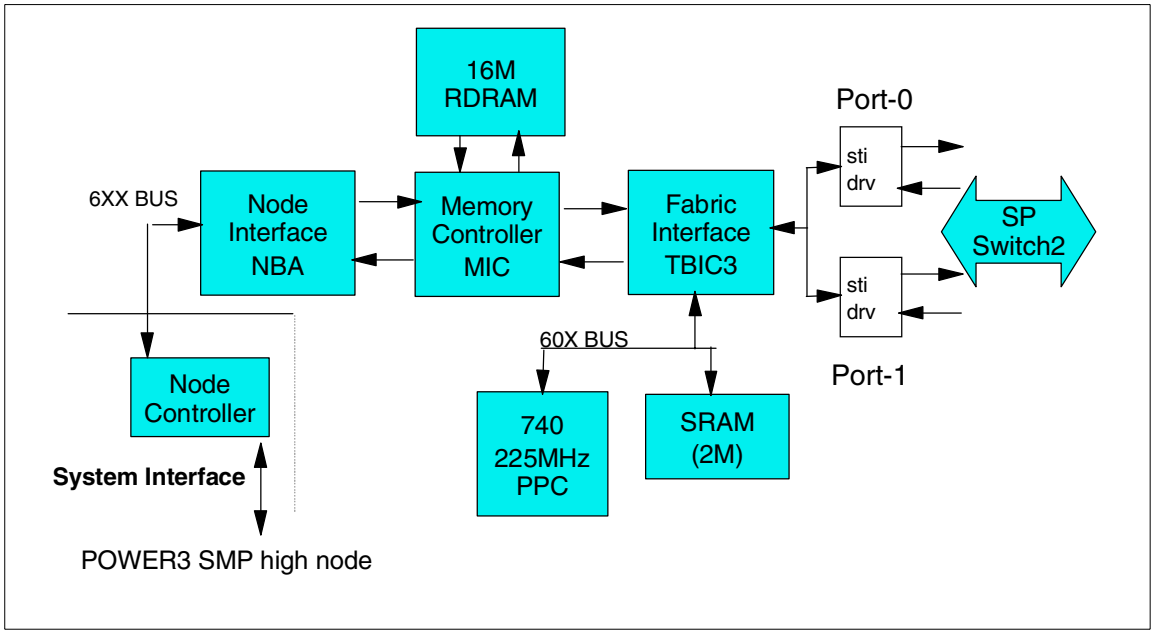


Figure 5-9 SP Switch2 adapter hardware structure

Node Bus Adapter (NBA)

The NBA chip moves data between the adapter and POWER3 SMP node's 6xx bus. The NBA chip's function is similar to that provided by the MBA, which was supplied on the TB3MX. The primary change is that it gives the adapter a memory type bus interface rather than the I/O style used in previous adapters.

Memory Interface Chip (MIC) and RDRam

The MIC chip is responsible for either passing information between the NBA and TBIC3, or for giving access to the adapter's data memory component. The adapter uses the 16 MB RDRam memory to buffer large blocks of data between the switch fabric and the main storage. This new function is not available on previous adapters.

TBIC3 Chip

The TBIC3 chip is an interface between the two STI interface chips, a 60X bus (to the 740 microprocessor), and the high speed interchip bus. Its function is very similar to those provided by TBIC on the TB3 adapter or the TBIC2 on the TB3MX adapter. The primary enhancements are its faster data transfer path (500 MB/s) and the addition of packet reassembly hardware in the chip.

The 740 microprocessor

The 740 microprocessor is responsible for passing control information between the adapter and the software on the POWER3 SMP node, for encoding or formatting packet headers, for building packet routing information, and for handling error conditions.

6

SP Switch Routers 9077

This chapter presents the IBM 9077 SP Switch Router within an SP environment only and with the SP Switch only.

6.1 The 9077 SP Switch Routers

The IBM 9077 SP Switch Router is a licensed version of the Ascend GRF switched IP router that has been enhanced for direct connection to the SP Switch. IBM remarkets models of the GRF that connect to the SP Switch as the SP Switch Router:

- IBM 9077 SP Switch Router model 04S (9077-04S) is based on Ascend GRF 400.
- IBM 9077 SP Switch Router model 16S (9077-16S) is based on Ascend GRF 1600.

To connect SP Switch Router to an SP system, the following adapter must be installed:

- SP Switch Router adapter (F/C 4021).

Figure 6-1 shows IBM 9077 SP Switch Routers.

Figure 6-1 IBM 9077 SP Switch Routers model 04S (left) and 16S (right)

Restrictions:

- The SP Switch Routers cannot be install in SP Switch2 environment.
- In CES environment with SP Switch, there is not obvious technical limitation to install SP Switch Routers, but this case has not been tested yet.

6.2 Overview

A physical dependent node, such as an RS/6000 SP Switch Router (Machine Type 9077), may have multiple logical dependent nodes, one for each dependent node adapter it contains. If the SP Switch Router contains more than one adapter, it can route data between SP systems. For the RS/6000 SP Switch Router, this card is called a Switch Router adapter (F/C 4021). Data transmission is accomplished by linking the dependent node adapters in the switch router with the logical dependent nodes located in different SP systems or system partitions.

In addition to the four major dependent node components (see Section 1.4.3, “Extension nodes” on page 16), the SP Switch Router (dependent node) has a fifth optional category of components. These components are networking cards that fit into slots in the SP Switch Router. In the same way that the SP Switch Router adapter connects the SP Switch Router directly to the SP Switch, these networking cards enable the SP Switch Router to be directly connected to an external network. The following networks can be connected to the RS/6000 SP Switch Router using available media cards:

- Ethernet 10/100 Base-T
- FDDI
- ATM OC-3c (single or multimode fiber)
- SONET OC-3c (single or multimode fiber)
- ATM OC-12c (single or multimode fiber)
- HiPPI
- HSSI

You can find a full list of these networking cards and other options in Table 6-1 on page 188.

Figure 6-2 on page 184 shows the SP Switch Router configuration. The SP Switch Router can be used for high-speed network connections or system scaling using HIPPI backbones or other communications subsystems, such as ATM or 10/100 Ethernet.

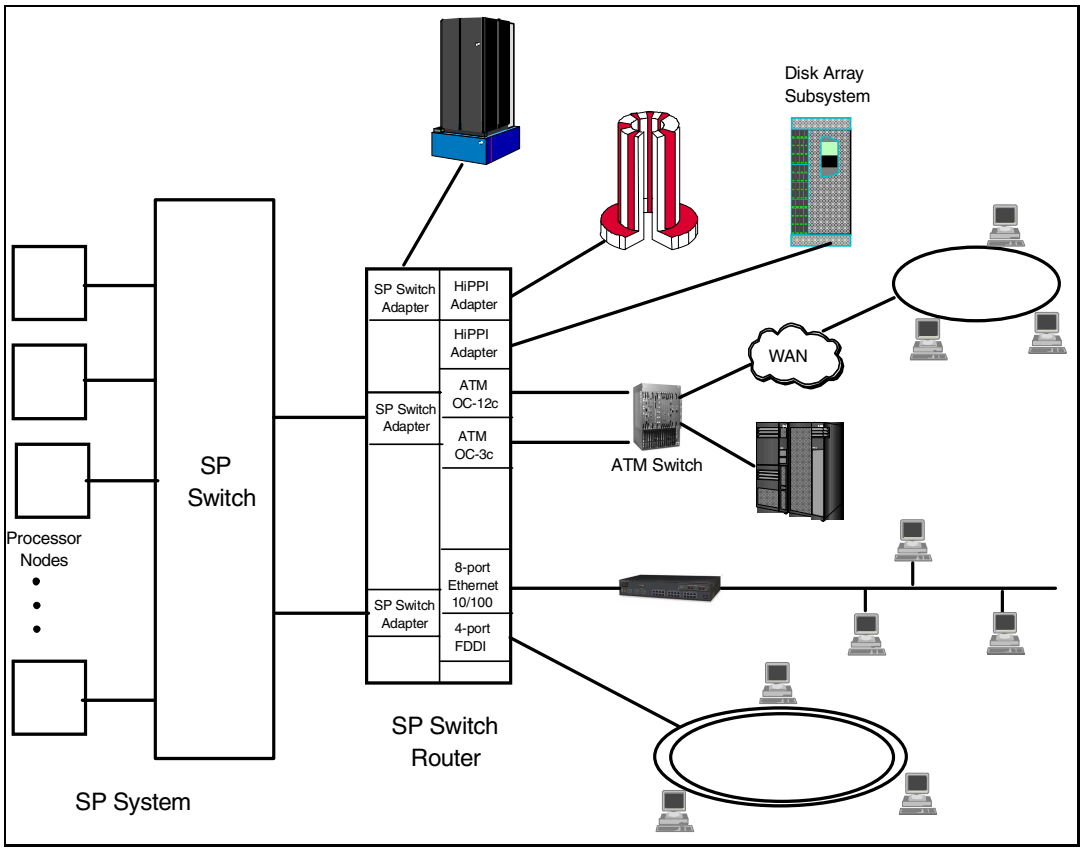


Figure 6-2 SP Switch Router configuration

Although you can equip an SP node with a variety of network adapters and use the node to make your network connections, the SP Switch Router with the Switch Router adapter and optional network media cards offers many advantages when connecting the SP to external networks.

- Each media card contains its own IP routing engine with separate memory containing a full route table of up to 150,000 routes. Direct access provides much faster lookup times compared to software driven lookups.
- Media cards route IP packets independently at rates of 60,000 to 130,000 IP packets per second. With independent routing available from each media card, the SP Switch Router gives your SP system excellent scalability characteristics.
- The SP Switch Router has dynamic network configuration to bypass failed network paths using standard IP protocols.

- Using multiple Switch Router adapters in the same SP Switch Router, you can provide high performance connections between system partitions in a single SP system or between multiple SP systems.
- A single SP system can also have more than one SP Switch Router attached to it, thus, further insuring network availability.
- Media cards are hot swappable for uninterrupted SP Switch Router operations.
- Each SP Switch Router has redundant (N+1) hot swappable power supplies.

Two versions of the RS/6000 SP Switch Router can be used with the SP Switch:

- The Model 04S offers four media card slots.
- The Model 16S offers sixteen media card slots.

Except for the additional traffic capacity of the Model 16S, both units offer similar performance and network availability.

Figure 6-3 illustrates, from a rear view, the SP Switch Router model 04S with four optional network media cards.

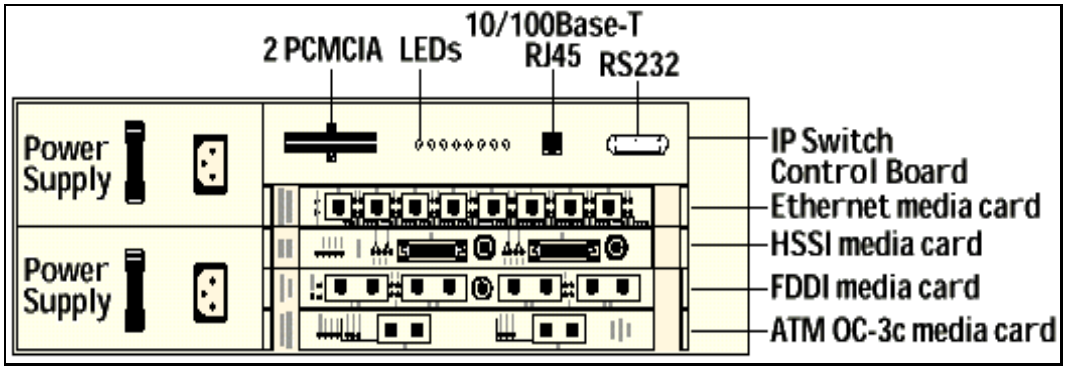


Figure 6-3 Rear view of the SP Switch Router model 04S

Figure 6-4 on page 186 illustrates, from the rear view, the SP Switch Router model 16S with nine optional network media cards.

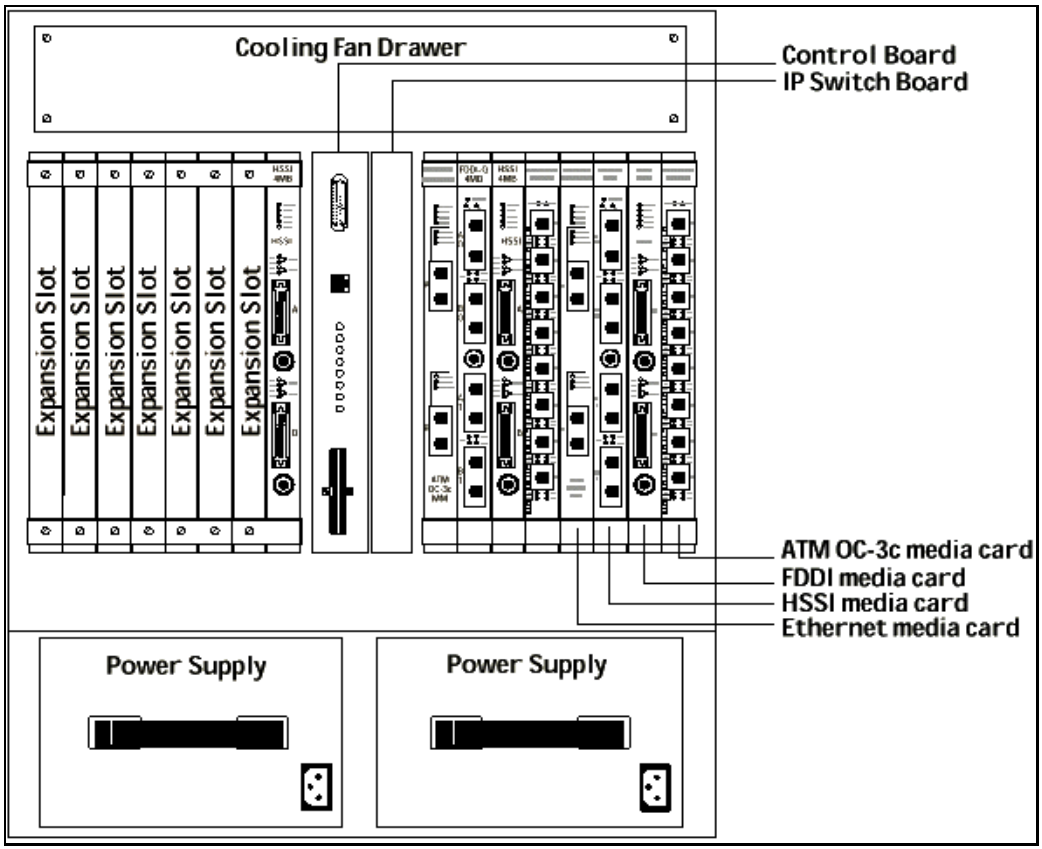


Figure 6-4 Rear view of SP Switch Router model 16S

6.3 Installation requirements

There are several requirements for hardware and software that must be met before you can place the RS/6000 SP Switch Router into service with your SP system. These requirements are in the following categories:

- System requirements
- Switch adapter requirements
- Network media card requirements
- Software requirements

6.3.1 System requirements

In addition to the SP Switch Router, the following requirements must be met before you can place the router into service:

- You must have at least one SP Switch Router adapter.
- You will need a VT100 compatible terminal with an RS-232 cable and null modem for initial configuration of the SP Switch Router.
- You will need a 10Base-T connection between your SP control workstation and the SP Switch Router. If your control workstation uses a 10Base-2 Ethernet, you must also supply a 10Base-T to 10Base-2 bridge.
- Your SP system must be equipped with either an SP Switch (F/C 4011) or an SP Switch-8 (F/C 4008).
- The SP Switch Router includes 128 MB of memory. This memory is used for storing routing information for 199,485 static and dynamic routes. F/C 1114 can be used to increase memory capacity on 64 MB increments up to a maximum of 256 MB, which allows up to 150,000 static route entries and 521,730 dynamic route prefixes per memory card.
- You must attach a frame-to-frame ground between the SP system and the SP Switch Router using the IBM supplied cable in order to maintain both systems at the same electrical potential.

6.3.2 SP Switch Router adapter (F/C 4021)

If you are placing an SP Switch Router into an SP system, you must install an SP Switch Router adapter (F/C 4021) in the SP Switch Router. The SP Switch Router adapter requires the following:

- One valid, unused switch port on the SP Switch corresponding to a valid node slot in your SP configuration.
- A valid node slot may be empty and may be the second half of a wide node, or one of the last three positions of a high node, provided that node slot satisfies the rules for configuring nodes in an SP system. For example, if you have a frame with 16 thin nodes installed, you cannot attach a Switch Router adapter to that frame until you remove a node and delete its configuration from the system image.
- One media card slot in the SP Switch Router. The RS/6000 SP Switch Router Model 04S has the capacity for a total of four SP Switch Router adapters and network media cards in any combination suiting the needs of your SP system. The RS/6000 SP Switch Router Model 16S has the capacity for a total of 16 SP Switch Router adapters and network media cards in any combination.

6.3.3 Network media card requirements

Each network media card requires one media card slot in the RS/6000 SP Switch Router. Remember, the network media cards use the same slots as the SP Switch Router adapters.

The following network media cards shown in Table 6-1 are available as options for the SP Switch Router.

Table 6-1 SP Switch Router network media cards and other options

SP Feature Code	Description
1101	ATM OC3, two port SM fiber
1102	ATM OC3, two port MM fiber
1103	SONET/IP OC3, one port MM fiber
1104	SONET/IP OC3, one port SM fiber
1105	ATM OC12, one port SM fiber
1106	FDDI, four port MM fiber
1107	Ethernet 10/100Base-T, eight port
1108	HIPPI, one port
1109	HSSI, two port
1112	Ethernet 10/100Base-T, four port
1113	Blank faceplate
1114	64 MB DRAM SIMM
1115	ATM OC12, one port MM fiber
1116	128 MB DRAM SIMM Memory option
1117	1000BASE-SX Gigabit Ethernet
1118	1000BASE-LX Gigabit Ethernet
4021	SP Switch Router adapter[1]
9310	SP Switch Router adapter cable, 10 meter option (includes 10 m frame-to-frame ground cable)
9320	SP Switch Router adapter cable, 20 meter option (includes 20 m frame-to-frame ground cable)
[1] Choice of either F/C 9310 or F/C 9320 is included with each F/C 4021.	

6.3.4 Software requirements

The SP Switch Router requires an SP system operating with PSSP 2.3 (or later) with the appropriate APAR level and AIX 4.2.1 (or later) on the primary and backup nodes for the SP Switch and on the control workstation.

If the SP Switch Router is used in an SP partition where there are nodes operating at lower than the required level of PSSP and AIX, you will have to apply service updates to the software operating on those nodes.

6.4 Network interface

The RS/6000 SP Switch Router (Machine Type 9077) requires a minimum of three connections with your SP system in order to establish a functional and safe network. These connections are:

1. A network connection with the control workstation.

   The SP Switch Router must be connected to the control workstation for system administration purposes. This connection may be either:

   - A direct Ethernet connection between the SP Switch Router and the control workstation.
   - An Ethernet connection from the SP Switch Router to an external network, which then connects to the control workstation. See Section 6.4.1, “Connecting to the control workstation” on page 190 for more information.

2. A connection between an SP Switch Router adapter and the SP Switch.

   The SP Switch Router transfers information into and out of the processor nodes of your SP system. The link between the SP Switch Router and the SP processor nodes is implemented by an SP Switch Router adapter (F/C 4021) and a switch cable connecting the Switch Router adapter to a valid switch port on the SP Switch. See Section 6.4.3, “Connecting to the SP Switch” on page 192 for more information.

3. A frame-to-frame electrical ground.

   The SP Switch Router frame must be connected to the SP frame with a grounding cable. This frame-to-frame ground is required in addition to the SP Switch Router electrical ground. The purpose of the frame-to-frame ground is to maintain the SP and SP Switch Router systems at the same electrical potential.

Both the SP Switch cable and the grounding cable are shipped with each SP Switch Router adapter. The recommended cable for connecting the SP Switch Router adapter to the SP Switch is 10 meters long (F/C 9310). An optional 20 meter cable (F/C 9320) is also available for the SP Switch connection. A frame-to-frame ground cable the same length as the SP Switch cable is included with both F/C 9310 and F/C 9320.

Network interface of the SP Switch Router is shown in Figure 6-5. VT100 terminal is optional and required for initial configuration time.

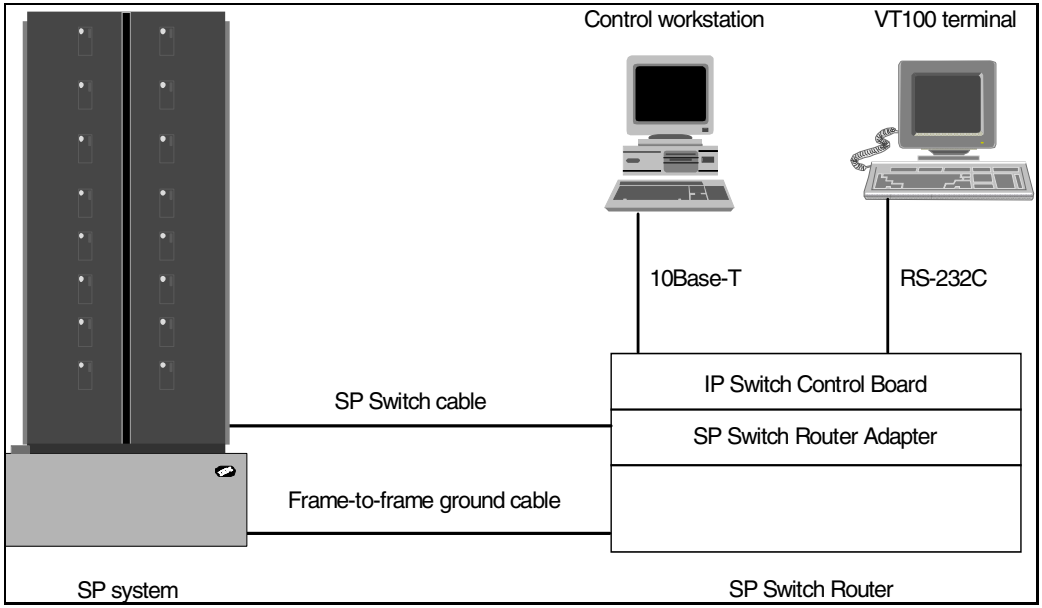


Figure 6-5 SP Switch Router network interface

The following sections describe how to connect the SP Switch Router to the control workstation and how to connect the SP Switch Router adapter to a valid SP Switch port.

6.4.1 Connecting to the control workstation

Although a dependent node, such as the SP Switch Router, does not function like a processor node, it must be administered by the SP system as if it were a processor node. Therefore, the SP Switch Router must be attached to the control workstation. All SP administrative connections are made to the SP Switch Router using the port on the Router's control board. From the SP Switch Router's control board, the connection to the control workstation is made using one of the following methods:

1. If the control workstation is connected to the SP system through a 10Base-2 (thin coax) network, the SP Switch Router may be connected to the network through a customer supplied 10Base-T to 10Base-2 hub (or bridge), such as the IBM 8222 Ethernet Workgroup Hub.

   All coax and twisted pair Ethernet cables must be supplied by the customer.

2. If the control workstation is connected to the SP system through a twisted pair (TP) Ethernet LAN, the SP Switch Router may be connected to an available port on the Ethernet hub (switch).

   All coax and twisted pair Ethernet cables must be supplied by the customer.

3. The RS/6000 SP Switch Router may also be connected to an additional 10Base-T adapter (such as F/C 2992) that has been installed directly in the control workstation for this purpose. If you decide to use this method, you must set up a separate Ethernet subnet for the SP Switch Router.

   When using separate 10Base-T adapters for the control workstation connection, in addition to the 10Base-T adapter, you must also supply a twisted pair Ethernet cable with a crossed connection appropriate for use between two network interface cards.

4. The SP Switch Router may also be indirectly connected to the control workstation using an external network. In this configuration, the Ethernet connection from the router's control board is attached to external network equipment. The external network connection to the control workstation may be:

   – A separate (non-SP Ethernet) Ethernet

   – ATM

   – FDDI

6.4.2 Connecting to multiple SP systems

If you are planning to connect one SP Switch Router to multiple, independent SP Systems, you will need:

- One SP Switch Router adapter for each SP system being connected.
- A Switch Router adapter cable to connect each of the adapters to an SP Switch located in each of the SP systems.
- An Ethernet connection from the SP Switch Router's control board to an external network. The Router's control board Ethernet connection is de0 and uses 10/100BaseT Ethernet.
- Connections from the external network must attach to the control workstations administering each SP system. The external networks may be:

  – Other Ethernets

  - ATM
  - FDDI
- Frame to frame grounds are required.

Note:

Other methods can be used to make the connection between the Router control board and all control workstations used in the SP systems. Any method providing the ability to ping SP control workstations from the Router control board will provide a valid path.

6.4.3 Connecting to the SP Switch

In addition to the control workstation Ethernet connection, the RS/6000 SP Switch Router requires a connection between the SP Switch and the SP Switch Router. To make this connection, your system will require the SP Switch Router adapter (F/C 4021). This adapter occupies one media card slot in the attached SP Switch Router.

Selecting a valid switch port

An SP Switch Router adapter in the SP Switch Router may be attached to an SP Switch to improve throughput of data coming into and going out of the RS/6000 SP system. Each SP Switch Router adapter in the RS/6000 SP Switch Router will require a valid unused switch port in the SP system. A valid unused switch port is a switch port that meets the rules for configuring frames and switches.

There are two basic sets of rules for choosing a valid switch port:

1. Rules for selecting a valid switch port associated with an empty node slot.
2. Rules for selecting a valid switch port associated with an unused node slot created by a wide or high node. These node slots are either the second half of a wide node or one of the last three positions of a high node.

Examples of using an empty node slot position

One example of using an empty node slot position is a single frame system with an SP Switch and 14 thin nodes located in slots 1 through 14. This system has two unused node slots in position 15 and 16. These two empty node slots have corresponding switch ports that provide valid connections for the SP Switch Router adapter.

Another example is a two-frame system with one switch. The first frame is fully populated with eight wide nodes. The second frame has three wide nodes in system node positions 17, 19, and 21. The only valid switch ports in this configuration would be those switch ports associated with system node numbers 23, 25, 27, 29, and 31 in the second frame.

In a four-frame system with an SP Switch and fourteen high nodes, there will be only two empty node positions. In this example, the first three frames are fully populated with four high nodes in each frame. The last frame has two high nodes and two empty high node slots. This means the system has two valid switch ports associated with system node numbers 57 and 61.

Using node slot positions within a wide or high node

The first example is a single frame with an SP Switch and eight wide nodes. These wide nodes occupy the odd numbered node slots. Therefore, all of the even number slots are said to be unoccupied and would have valid switch ports associated with them. These ports may be used for an SP Switch Router adapter.

A second example is a single frame system with an SP Switch, twelve thin nodes in slots 1 through 12, and a high node in slot 13. A high node occupies four slots but only uses one switch port. Therefore, the only valid switch ports in this configuration are created by the three unused node slots occupied by the high node. In other words, the switch ports are associated with node slots 14, 15, and 16.

7

Control workstations

The SP/CES systems require a customer-supplied pSeries or RS/6000 system known as a control workstation. The control workstation serves as a point of control for managing, monitoring, and maintaining the SP/CES system frames and individual processor nodes. A system administrator can perform these control tasks by logging into the control workstation from any other workstation on the network.

The control workstation also acts as a boot/install server for other servers in the SP/CES system. In addition, the control workstation can be set up as an authentication server using Kerberos. The control workstation can be the Kerberos primary server, with the master database and administration service, as well as the ticket-granting service. As an alternative, the control workstation can be set up as a Kerberos secondary server, with a backup database, to perform ticket-granting service.

Kerberos is no longer the only security method. The Distributed Computing Environment (DCE) can be used with Kerberos V4, or by itself, or nothing (AIX standard security).

A high availability solution may be implemented for the Control Workstation. The solution is named the High Availability Control Workstation.

7.1 Overview

You may consider the control workstation as a server to the SP/CES system applications. The subsystems running on CWS are the SP/CES server applications for the SP/CES system nodes. The nodes are clients of the CWS server applications. The CWS serves as a single point of control for these server applications that provide configuration data, security, hardware monitoring, diagnostics, and, optionally, job scheduling data and a time sourcing.

However, the CWS requires many hardware and software considerations. The SP/CES system requires a pSeries or an RS/6000 CWS with a color monitor. The CWS must be connected to each frame through RS-232 cable and SP Ethernet shown in Figure 7-1. Fifteen meter (49 ft.) cables are supplied with each SP standard frame for both RS-232 and BNC Ethernet only. It is the customer's responsibility to supply the cables for twisted-pair SP Ethernet network (if it exists). However, the control workstation and SP frames should be no more than 12 meters apart. This leaves 3 meters of cable for use in the vertical runs between the equipment. If you need longer vertical runs, or if there are floor obstructions, you must place the components closer together.

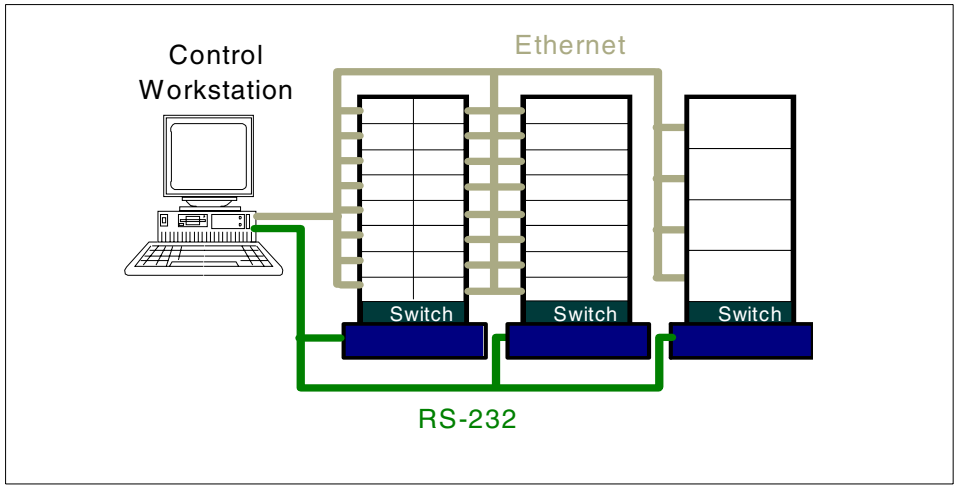


Figure 7-1 Control workstation interface

All SP-attached servers and Clustered Enterprise servers must also be connected to the control workstation. However, each SP-attached server M/T7017 or Clustered Enterprise Server M/T 7017 requires two RS-232 connections as well as the SP Ethernet connection. Each SP-attached server M/T 7017 or Clustered Enterprise Server M/T 7026 requires one RS-232 connection as well as the SP Ethernet connection. Fifteen meter (49 ft.) cables are supplied with each SP-attached server for both RS-232 and BNC Ethernet

only. You have to order the RS-232 cables for each Clustered Enterprise Server. No BNC Ethernet cables supplied for the Clustered Enterprise servers. On any of these cases, it is the customer's responsibility to supply the cables for twisted-pair SP Ethernet network (if it exists).

Note that most PCI control workstations provide either a 10BaseT or AUI connection for the SP Ethernet LAN. If you are attaching to nodes or SP-attached servers that require a BNC connection, make sure you have ordered the appropriate transceiver.

All SP Switch Routers must also be connected by the SP Ethernet to the control workstation. The SP Switch Routers require a 10BASE-T connection. All BNC or twisted-pair Ethernet cables must be supplied by the customer.

The CWS requires the AIX operating system and Parallel System Support Programs (PSSP) software. To take advantage of the latest software enhancements, your SP/CES system, along with its CWS, must be running the latest version of AIX and PSSP.

Note: You may have different AIX and PSSP releases installed in different nodes. The AIX and PSSP levels installed on the CWS must be equal to or greater than the levels of AIX and PSSP installed on the individual SP/CES nodes/servers. We suggests that you run only PSSP on your CWS to ensure optimum performance of your SP/CES system.

In general, the CWS offers a service processor as either a standard or optional feature (F/C 1001). The service processor is a standby processor that handles system start-up and some system monitoring functions. When this option is installed with a modem on the S1 serial port, you can perform remote system restarts on SP systems located in unmanned locations.

The CWS is the center of SP/CES system management. It acts as a control workstation - a focal point for systems administration. Any X-Windows capable system on a LAN can be used to log into the control workstation to perform system management, monitoring and control tasks. The administrator does not need to physically sit at the control workstation.

7.2 Installation requirements

The control workstation and some of its software are not part of the SP package and must be ordered separately. Make sure you have ordered them in time so they will arrive when the rest of your SP does. To coordinate delivery of the SP and control workstation, your IBM representative should link the SP and control workstation orders with a System Order Number. The installation requirements are shown in the following categories:

- Supported pSeries and RS/6000 workstations and servers
- System requirements
- Interface adapter requirements
- Software requirements

7.2.1 Supported pSeries and RS/6000 workstations and servers

The pSeries and RS/6000 workstations and servers shown in Table 7-1 are supported as control workstations for the SP/CES systems:

Table 7-1 Supported pSeries and RS/6000 workstations for SP/CES systems

Machine Type	Model	Currently Available
7026 (PCI) (Note 1,8,9)	6H1 (PSSP 3.2)	Yes
7025 (PCI) (1,9)	6F1 (PSSP 3.2)	Yes
7026 (PCI) (1,8)	H80 (PSSP 3.2)	Yes
7025 (PCI) (1)	F80 (PSSP 3.2), F50 (PSSP2.2 or later)	Yes
7044 (PCI) (1,2)	170 (PSSP 3.2)	Yes
7024 (PCI)	E20 (PSSP2.2 or later), E30 (PSSP 2.2 or later)	No
7025 (PCI) (1, 3)	F40 (PSSP 2.2 or later)	No
7025 (PCI) (4)	F30 (PSSP 2.2 or later)	No
7026 (PCI) (6)	H10 (PSSP 2.2 or later), H50 (PSSP 2.2 or later)	No
7043 (PCI) (1,2)	140 (PSSP 2.2 or later), 240 (PSSP 2.2 or later)	No
7012 (MCA) (5)	37T, 370, 375, 380, 39H, 390, 397, G30, G40	No
7013 (MCA) (5,6)	570, 58H, 580, 59H, 590, 591, 595, J30, J40, J50	No
7015 (MCA) (5,6,7)	97B, 970, 98B, 980, 990, R30, R40, R50	No
7030 (MCA) (5)	3AT, 3BT, 3CT	No

Notes:

1. The native RS-232 ports on the system planar can not be used as tty ports for the hardware controller interface. The 8-port asynchronous adapter EIA-232/ RS-422, PCI bus (F/C 2943) or the 128-port Asynchronous Controller (F/C 2944) are the only RS-232 adapters that are supported. These adapters require AIX 4.2.1 or AIX 4.3 on the control workstation.
2. The 7043 and 7044 models can only be used on SP systems with up to four frames. This limitation applies to the number of frames and not the number of nodes. This number includes expansion frames. The 7043 and 7044 models can not be used for SP/CES environments when the sum of SP frames and clustered servers exceeds four. The 7043 and 7044 models can not be used for SP/CES configured with 7017-S80 and 7017-S85 servers.(these servers use more CPU resources than the other nodes).
3. We strongly suggest using the service processor option (F/C 1001).
4. On systems introduced since PSSP 2.4, either the 8-port (F/C 2493) or 128-port (F/C 2944) PCI bus asynchronous adapter should be used for frame controller connections. We strongly suggest using the service processor option (F/C 1001). If you use this option, the frames must be connected to a serial port on an asynchronous adapter and not to the serial port on the control workstation planar board.
5. PSSP 3.1 and PSSP 3.2 support.
6. Requires a 7010 Model 150 X-Station and display. Other models and manufacturers that meet or exceed this model can be used. An ASCII terminal is required as the console.
7. Installed in either the 7015-99X or 7015-R00 Rack.
8. The Model is supported for twisted-pair LAN environment only.
9. The Model is member of pSeries family.

We strongly recommend using the 6H1, 6F1, H80, F80, and F50 models only, as CWS on SP systems with SP-attached servers and CES systems. The other CWS are considered underpowered due to the SP-attachment daemons put on the CWS.

7.2.2 System requirements

The minimum requirements for the CWS are:

- At least 128 MB of main memory. An extra 64 MB of memory should be added for each additional system partition. For SP systems with more than 80 nodes, 256 MB is required and 512 MB of memory is recommended.

- At least 9 GB of disk storage. If the SP is going to use an HACWS configuration, you can configure 9 GB of disk storage in the rootvg volume group and 9 GB for the spdata in an external volume group. Because the control workstation is used as a Network Installation Manager (NIM) server, the number of unique file sets required for all the nodes in the SP system might be larger than a normal single system. You should plan to reserve 6 GB of disk storage for the file sets and 2 GB for the operating system. This will allow adequate space for future maintenance, system mksysb images, and LPP growth. Keep in mind that if you have nodes at different levels of PSSP or AIX, each node requires its own LPP source, which will take up extra space. A good rule of thumb to use for disk planning for a production system is 4 GB for the rootvg to accommodate additional logging and /tmp space plus 4 GB for each AIX release and modification level for lppsource files. Additional disk space should be added for mksysb images for the nodes.If you plan on using rootvg mirroring, then for one mirror, double the number of physical disks you estimated so far. For two mirrors, triple the estimate.
- You might plan to use the control workstation to initially build your own customized mksysb AIX image. One reason might be that you want to enable the AIX trusted computing base option. If you plan to create a customized mksysb for any reason, you must have at least two physical disks in the control workstation, one for the alternate volume group (not rootvg) that the **mksysb** command will use for the output.
- Physically installed with the RS-232 cable to within 12 meters of each SP frame.
- Physically installed with two RS-232 cables to within 12 meters of each SP-attached server or Clustered Enterprise Server M/T 7017. Physically installed with one RS-232 cable to within 12 meters of each SP-attached server or Clustered Enterprise Server M/T 7026.
- Equipped with the following I/O devices and adapters:
  - A 3.5 inch diskette drive.
  - Four or eight millimeter (or equivalent) tape drive.
  - A SCSI CD-ROM device.
  - One RS-232 port for each SP frame.
  - Keyboard and mouse.
  - Color graphics adapter and color monitor. An X-station model 150 and display are required if an RS/6000 that does not support a color graphics adapter is used.
  - An appropriate network adapter for your external communication network. The adapter does not have to be on the control workstation. If it is not on the control workstation, the SP Ethernet must extend to another host that

is not part of the SP system. A backup control workstation does not satisfy this requirement. This additional connection is used to access the control workstation from the network when the SP nodes are down.

– SP Ethernet adapters for connection to the SP Ethernet The number of Ethernet adapters required depends completely on the Ethernet topology you use on your SP system. The following types of Ethernet adapters can be used:

  - Ethernet adapters with thin BNC. Each Ethernet adapter of this type can have only 30 network stations on a given Ethernet cable. The control workstation and any routers are included in the 30 stations.
  - Ethernet adapters with twisted pair (RJ45/AUI). A network hub or switch is required.
  - 10/100 Mbps Ethernet adapters. A network hub or switch is required.

► If your SP has more than two frames or you are considering a control workstation that has the PCI type of bus, you need an asynchronous adapter card to provide ports for the SP frames.

On SP systems with SP-attached servers and CES systems, an F50/F80 model configured with 256 MB of memory and 4 of 9 GB disk is the minimum configuration for a two frame system. It is recommended that an additional 256 MB is added for the next two frames, and then up to 1 GB for the larger systems.

7.2.3 Interface adapter requirements

Several different control workstations are available. Each model has different communications adapters offered as standard equipment. Depending on the model you choose, serial and Ethernet adapters may have to be added to the workstation to satisfy the needs of your SP/CES system.

Serial port adapters

All new PCI control workstations require a minimum of one additional asynchronous adapter. For additional PCI/MCA serial ports, select the equipment you need from the feature codes shown in Table 7-2 on page 202.

Note: In addition to the listed PCI bus adapters, the 7024-EXX and 7025-F30 control workstations also support the listed ISA bus adapters. All other PCI control workstations support only PCI bus adapters. PCI adapters offer performance advantages in all PCI control workstations and should be used whenever possible.

Table 7-2 Serial port adapters for PCI / MCA control workstations

F/C	Description	Currently Available
8-port PCI		
2931	8-port asynchronous adapter PCI BUS EIA-232 / RS-422	No
2932	8-port asynchronous adapter PCI BUS EIA-232 / RS-422A	No
2943	8-port asynchronous adapter PCI BUS EIA-232 / RS-422	Yes
128-port PCI		
2933	128-port asynchronous controller ISA bus	No
2934	Asynchronous terminal/printer cable, EIA-232 (2.4 MB/sec)	Yes
2944	128-port asynchronous controller ISA bus	Yes
3124	Serial port to serial port cable for drawer-to-drawer connections (2.4MB/sec)	Yes
3125	Serial port to serial port cable for rack-to-rack connections (2.4MB/sec)	Yes
8130	1.2 MB/sec remote asynchronous node (RAN) 16-port EIA-232 (US)	No
8131	128-port asynchronous controller cable, 4.5 m (1.2MB/sec)	Yes
8132	128-port asynchronous controller cable, 23 cm (1.2MB/sec)	Yes
8133	RJ-45 to DB-25 converter cable	Yes
8134	World Trade version of F/C 8130	No
8136	1.2 MB/sec rack-mountable asynchronous node (RAN) 16-port EIA-232	Yes
8137	2.4 MB/sec enhanced remote asynchronous node (RAN) 16-port EIA-232	Yes
8138	2.4 MB/sec enhanced remote asynchronous node (RAN) 16-port RS-422	Yes
8-port MCA		
2930	8-port asynchronous adapter	No
2995	multiport interface cable	No
16-port MCA		

F/C	Description	Currently Available
2955	16-port asynchronous adapter (not compatible with SP-attached servers and CES)	No
2996	multiport interface cable	No
128-port MCA		
8128	128-port asynchronous controller	No
8130	remote asynchronous node 16-port EIA-232	No
8134	World trade version of F/C 8130	No

Ethernet adapters

For additional PCI/MCA Ethernet ports, select the equipment you need from the feature codes shown in Table 7-3.

Table 7-3 Ethernet adapters for PCI/MCA control workstations

F/C	Description	Currently Available
PCI		
2968	IBM 10/100 Mbps Ethernet PCI adapter	Yes
2985	PCI Ethernet BNC/RJ-45 adapter	Yes
2987	PCI Ethernet AUI/RJ-45 adapter	Yes
4224	Ethernet 10BASE2 transceiver	Yes
MCA		
2980	Ethernet high performance LAN adapter	No
2992	Ethernet twisted pair (TP) adapter	No
2993	Ethernet BNC/AUI adapter	No
4224	Ethernet 10BASE2 transceiver	Yes

7.2.4 Software requirements

The control workstation and some of its software might not be part of the SP package and must be ordered separately.

The control workstation requires the following software:

- AIX 4.3.3 (5765-C34) (or later).

- PSSP 3.2 (5765-D51) (or later).
- C for AIX 3.6.6 (or later) or VisualAge C++ Professional 4.0 for AIX (or later).

Note:

1. PSSP does not support the incremental compiler and runtime libraries. It only supports the batch C and C++ 3.6.6 or later compilers and runtime libraries that are included in this VisualAge package.
2. AIX 5L is not supported on SP CWS environment.

An AIX license is required for the CWS. The PSSP ordered for the SP/CES system is entitled for use the entire SP/CES complex, which includes the CWS (and the backup CWS, if used).

Compilers are necessary for IBM service of PSSP. Also, without the compiler, dump diagnosis tools like *crash* cannot work effectively.

At least one concurrent user license is required for the SP system. Concurrent licensing is recommended so the one license can float across the SP nodes and the control workstation. You can order the license as part of the SP system. It is not specifically required on the control workstation if a license server for AIX for C and C++ exists some place in the network and the SP is included in the license server's cell.

AIX and PSSP are shipped from the factory preloaded on node disk storage. The SP system administrator must install AIX and PSSP on the CWS from the distribution medium, and perform system customization after the hardware is installed by the service representative. The proper levels of AIX and PSSP must be installed on the CWS before SP nodes can be installed.

Some PSSP components and related LPPs are optional but when you do choose to use them, some must be installed on the CWS while others can optionally be installed on the CWS. The following are optional but if you use them, they must be installed on your CWS:

- HACWS
- IBM Virtual Shared Disk and Recoverable Virtual Shared Disk
- PTPE
- AIX DCE (If you plan to use AIX DCE authentication methods as part of security on your SP, you must order and install the AIX DCE product)
- PSSP graphical user interfaces (These include the SP Perspectives and other optional interfaces such as:
  - Hardware Perspective

- Event Perspective
- Performance Monitor
- Virtual Shared Disk Perspective
- System Partitioning Aid
- SP TaskGuides
- SP Resource Center
- SP Perspectives Launch Pad
- Perspectives online help
- Netfinity Services Manager)

7.2.5 SP control workstation offerings

The latest CWS models are the RS/6000 7044-170 and 7025-F80, and the pSeries 7025-6F1 and 7026-6H1.

The RS/6000 7044-170 CWS is supported for low-end SP systems. If there is more than one SP frame or any CES planned in your environment, we strongly recommend the F80, 6H1, and 6F1 as CWS.

There are different SP Control Workstation offerings as prepackaged solutions for CWS:

7044-170 CWS default entry level configuration

This configuration is recommended for low-end SP systems. Example 7-1 shows a 7044-170 control workstation default entry level configuration.

Example 7-1 7044-170 control workstation default entry level configuration

```
*** HARDWARE ***

 Product           Description            Qty

 7044-170 RS/6000 44P Model 170             1
          1.44MB 3.5-in Diskette Drive      1
          Integrated Ultra SCSI Adapter     1
          Integrated External Ultra2 SCSI   1
          Port
          Integrated Ethernet Adapter       1
     2624 32x Speed CD-ROM Drive            1
     2830 POWER GXT130P Graphics Adapter    1
          (PCI)
     2909 18.2 GB 1" Ultra SCSI Hard Disk   2
          Drive
     2943 8-Port Async Adapter,             1
```

```
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet Adapter       1
     3628 P260 Color Monitor, Stealth        1
          Black
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     4223 Ethernet 10Base2 Transceiver       1
     4349 333MHz POWER3-II Processor Card    1
     5005 Preinstall                         1
     6159 12GB/24GB 4mm Tape Drive           1
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada (125V,      1
          15A)

*** SOFTWARE ***

 Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0903 CWS Diagnostic Pre-Load           1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1
```

7025-F80 control workstation small configuration

This configuration is recommended for SP systems that include no more than 30 nodes. Example 7-2 on page 207 shows a 7025-F80 control workstation small configuration.

Example 7-2 7025-F80 control workstation small configuration

```
*** HARDWARE ***

 Product           Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2943 8-Port Async Adapter,              1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI    1
     3023 18.2 GB 1" Ultra2 SCSI 16-bit      2
          Disk
     3109 SCSI External Port to SCSI         1
          Internal 6 Pack Cable Assembly
     3628 P260 Color Monitor, Stealth        1
          Black
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     5005 Preinstall                         1
     5201 1-Way RS64 III 450 MHz             1
          Processor Card, 2MB L2 Cache
     6159 12GB/24GB 4mm Tape Drive           1
     6553 SCSI Hot Swap 6-Pack               1
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada (125V,      1
          15A)

*** SOFTWARE ***

 Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0598 AIX Welcome Center                1
```

```
    0838 AIX 4.3 Update CD                1
    0857 AIX 4.3                          1
    0859 Preinstall AIX 4.3 Server        1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non         1
         Export.
    0903 CWS Diagnostic Pre-Load          1
    1004 CD-ROM Process Charge            1
    2924 English                          1
    3410 CD-ROM                           1
    5005 Preinstall                       1
    5924 AIX 4.3 PII - English            1
    9001 Asset Registration               1
```

7025-F80 control workstation medium configuration

This configuration is recommended for SP systems that include more than 30 nodes and for SP/CES systems that include no more than 16 enterprise servers.Example 7-3 shows a 7025-F80 control workstation medium configuration.

Example 7-3 7025-F80 control workstation medium configuration

```
*** HARDWARE ***

 Product          Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2943 8-Port Async Adapter,              1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI           3
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI    1
     3023 18.2 GB 1" Ultra2 SCSI 16-bit      4
          Disk
     3109 SCSI External Port to SCSI         1
          Internal 6 Pack Cable Assembly
     3628 P260 Color Monitor, Stealth        1
          Black
     4075 Memory Board, 16-position          1
     4110 256 MB (2x128MB) SDRAM DIMMs       2
     4223 Ethernet 10Base2 Transceiver       3
     5005 Preinstall                         1
```

```
    5202 2-Way RS64 III 450 MHz            1
         Processor Card, 4MB L2 Cache
    6159 12GB/24GB 4mm Tape Drive          1
    6553 SCSI Hot Swap 6-Pack              1
    8700 Quiet Touch Keyboard, Stealth     1
         Black - English (US)
    8741 3-Button Mouse - Stealth Black    1
    9300 Language - English (US)           1
    9800 Power Cord - US/Canada (125V,     1
         15A)

*** SOFTWARE ***

 Product          Description            Qty

 5765-C34 AIX Version 4                    1
    4061 AIX 4.3                           1
    5005 Preinstall                        1
    9001 Asset Registration                1

 5692-AIX System Software                  1
    0598 AIX Welcome Center                1
    0838 AIX 4.3 Update CD                 1
    0857 AIX 4.3                           1
    0859 Preinstall AIX 4.3 Server         1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non          1
         Export.
    0903 CWS Diagnostic Pre-Load           1
    1004 CD-ROM Process Charge             1
    2924 English                           1
    3410 CD-ROM                            1
    5005 Preinstall                        1
    5924 AIX 4.3 PII - English             1
    9001 Asset Registration                1
```

7025-F80 control workstation large configuration

This configuration is recommended for SP systems with a large number of nodes and for SP/CES systems that includes more than 16 enterprise servers. Example 7-4 shows a 7025-F80 control workstation large configuration.

Example 7-4 7025-F80 control workstation large configuration

```
*** HARDWARE ***

 Product          Description            Qty

 7025-F80 RS/6000 Enterprise Server Model   1
          F80
```

```
          1.44MB 3.5-in Diskette Drive      1
          Integrated Ultra2 SCSI Adapter    1
          Integrated Ethernet Adapter       1
     2432 Point to Point Cable, 1.1m        1
     2624 32x Speed CD-ROM                  1
     2830 POWER GXT130P Graphics Adapter    1
          (PCI)
     2944 128-Port Async Controller (PCI)   1
     2968 10/100 Mbps Ethernet PCI          3
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI   1
     3023 18.2 GB 1" Ultra2 SCSI 16-bit     6
          Disk
     3628 P260 Color Monitor, Stealth       1
          Black
     4075 Memory Board, 16-position         1
     4119 512 MB (2x256MB) SDRAM DIMMs      2
     4223 Ethernet 10Base2 Transceiver      3
     5005 Preinstall                        1
     5202 2-Way RS64 III 450 MHz            1
          Processor Card, 4MB L2 Cache
     6158 20GB/40GB 4mm Tape Drive          1
     6205 PCI Dual Channel Ultra2 SCSI      1
          Adapter
     6553 SCSI Hot Swap 6-Pack              1
     8131 4.5m Controller Cable             1
     8133 RJ-45 to DB-25 Converter Cables   1
          (4 cables per set)
     8137 Enhanced Async Node 16-Port       1
          EIA-232
     8700 Quiet Touch Keyboard, Stealth     1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black    1
     9300 Language - English (US)           1
     9800 Power Cord - US/Canada (125V,     1
          15A)

*** SOFTWARE ***

 Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
```

```
    0857 AIX 4.3                          1
    0859 Preinstall AIX 4.3 Server        1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non         1
         Export.
    0903 CWS Diagnostic Pre-Load          1
    1004 CD-ROM Process Charge            1
    2924 English                          1
    3410 CD-ROM                           1
    5005 Preinstall                       1
    5924 AIX 4.3 PII - English            1
    9001 Asset Registration               1
```

7026-6H1 control workstation small configuration

This configuration is recommended for SP systems that include no more than 30 nodes. Example 7-5 shows a 7026-6H1 control workstation small configuration.

Example 7-5 7026-6H1 control workstation small configuration

```
***  HARDWARE  ***

  Product           Description           Qty

 7026-6H1 pSeries 660                      1
          1.44MB 3.5-in Diskette Drive     1
          Integrated SCSI-2 F/W Adapter    1
          Integrated Ultra2 SCSI Adapter   1
          Integrated Ethernet Adapter      1
     2624 32x Speed CD-ROM                 1
     2830 POWER GXT130P Graphics Adapter   1
          (PCI)
     2943 8-Port Async Adapter,            1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI         1
          Adapter
     3102 18.2 GB 10K RPM Ultra SCSI Disk  2
          Drive
     3142 Remote I/O Cable - 3m            2
     3628 P260 Color Monitor, Stealth      1
          Black
     4110 256 MB (2x128MB) SDRAM DIMMs     1
     5005 Preinstall                       1
     5211 1-Way RS64 IV 600 MHz Processor  1
          Card, 2MB L2 Cache
     5992 System Control and               1
          Initialization Cable, 3m
     6132 CEC to Primary I/O Drawer Power  1
          Control Cable, 3m
     6159 12GB/24GB 4mm Tape Drive         1
```

```
     6324 Primary I/O Drawer, 5 EIA          1
     6540 IPL Disk Mounting Hardware,        1
          Cables, Terminator
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9172 AC Power Specify                   1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada             1

7014-T00 Enterprise Rack - 36 EIA            1
     0176 Content : FC 6324 (5 EIA)          1
     0188 Content : 7026-6H1 (5 EIA)         1
     6088 Front Door for 1.8m Rack, Black    1
     6098 Side Panel for 1.8 or 2.0m         2
          Rack, Black
     9171 Power Distribution Unit,           1
          Side-Mount, 1 Phase
     9300 Language - English (US)            1
     9800 Rack Power Cord - US/Canada        1

 7014-T00 Enterprise Rack - 36 EIA           1
     0183 Content : 9910-A30 (5 EIA)         1
     6088 Front Door for 1.8m Rack, Black    1
     6098 Side Panel for 1.8 or 2.0m         2
          Rack, Black
     9171 Power Distribution Unit,           1
          Side-Mount, 1 Phase
     9300 Language - English (US)            1
     9800 Rack Power Cord - US/Canada        1

 9910-A30 APC 5000VA Smart-UPS Rack-Mount    1
          SmartSlot 4x4 10BaseT SNMP         1
          Adapter (Standard)

***  SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                      1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1

 5692-AIX System Software                    1
     0598 AIX Welcome Center                 1
     0838 AIX 4.3 Update CD                  1
     0857 AIX 4.3                            1
     0859 Preinstall AIX 4.3 Server          1
```

```
     Bundle
0860 AIX 4.3 Bonus Pack - Non          1
     Export.
0903 CWS Diagnostic Pre-Load           1
1004 CD-ROM Process Charge             1
2924 English                           1
3410 CD-ROM                            1
5005 Preinstall                        1
5924 AIX 4.3 PII - English             1
9001 Asset Registration                1
```

7026-6H1 control workstation medium configuration

This configuration is recommended for SP systems that include more than 30 nodes and for SP/CES systems that include no more than 16 enterprise servers. Example 7-6 shows a 7026-6H1 control workstation medium configuration.

Example 7-6 7026-6H1 control workstation medium configuration

```
***  HARDWARE  ***

  Product           Description            Qty

 7026-6H1 pSeries 660                       1
          1.44MB 3.5-in Diskette Drive      1
          Integrated SCSI-2 F/W Adapter     1
          Integrated Ultra2 SCSI Adapter    1
          Integrated Ethernet Adapter       1
     2624 32x Speed CD-ROM                  1
     2830 POWER GXT130P Graphics Adapter    1
          (PCI)
     2943 8-Port Async Adapter,             1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI          3
          Adapter
     3102 18.2 GB 10K RPM Ultra SCSI Disk   2
          Drive
     3142 Remote I/O Cable - 3m             2
     3628 P260 Color Monitor, Stealth       1
          Black
     4075 Memory Board, 16-position         1
     4119 512 MB (2x256MB) SDRAM DIMMs      2
     5005 Preinstall                        1
     5212 2-Way RS64 IV 600 MHz Processor   1
          Card, 4MB L2 Cache
     5992 System Control and                1
          Initialization Cable, 3m
     6132 CEC to Primary I/O Drawer Power   1
          Control Cable, 3m
     6230 Advanced SerialRAID Plus          1
```

```
          Adapter
      6324 Primary I/O Drawer, 5 EIA          1
      6540 IPL Disk Mounting Hardware,        1
           Cables, Terminator
      8700 Quiet Touch Keyboard, Stealth      1
           Black - English (US)
      8741 3-Button Mouse - Stealth Black     1
      9172 AC Power Specify                   1
      9300 Language - English (US)            1
      9800 Power Cord - US/Canada             1

  7014-T00 Enterprise Rack - 36 EIA           1
      0176 Content : FC 6324 (5 EIA)          1
      0188 Content : 7026-6H1 (5 EIA)         1
      6088 Front Door for 1.8m Rack, Black    1
      6098 Side Panel for 1.8 or 2.0m         2
           Rack, Black
      9171 Power Distribution Unit,           1
           Side-Mount, 1 Phase
      9300 Language - English (US)            1
      9800 Rack Power Cord - US/Canada        1

 7014-T00 Enterprise Rack - 36 EIA            1
      0183 Content : 9910-A30 (5 EIA)         1
      6088 Front Door for 1.8m Rack, Black    1
      6098 Side Panel for 1.8 or 2.0m         2
           Rack, Black
      9171 Power Distribution Unit,           1
           Side-Mount, 1 Phase
      9300 Language - English (US)            1
      9800 Rack Power Cord - US/Canada        1

  7133-D40 Advanced SSA Disk Subsystem        1
           (Rack-Mounted)
           Single Mode                        1
      0550 Hungary Manufacturing Ship         1
           Direct to Customer
      8022 50/60Hz AC, 300 VDC Power          1
           Supplies
      8031 Raven Black Drawer Cover           1
      8518 One 10K/18.2GB Advanced Disk       4
           Drive Module
      8801 1m Advanced SSA Cable              1
      9300 Language - English (US)            1

  9910-A30 APC 5000VA Smart-UPS Rack-Mount    1
           SmartSlot 4x4 10BaseT SNMP         1
           Adapter (Standard)
```

```
***  SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0903 CWS Diagnostic Pre-Load           1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1
```

7025-6F1 control workstation small configuration

This configuration is recommended for SP systems that include no more than 30 nodes. Example 7-7 shows a 7025-6F1 control workstation small configuration.

Example 7-7 7025-6F1 control workstation small configuration

```
***  HARDWARE  ***

  Product           Description            Qty

 7025-6F1 pSeries 620                       1
          1.44MB 3.5-in Diskette Drive      1
          Integrated SCSI-2 F/W Adapter     1
          Integrated Ultra2 SCSI Adapter    1
          Integrated Ethernet Adapter       1
     2624 32x Speed CD-ROM                  1
     2830 POWER GXT130P Graphics Adapter    1
          (PCI)
     2943 8-Port Async Adapter,             1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI          1
          Adapter
     3109 SCSI External Port to SCSI        1
          Internal 6 Pack Cable Assembly
```

```
    3153 18.2 GB 10K RPM 1" Ultra3 SCSI     2
         16-bit Disk
    3628 P260 Color Monitor, Stealth        1
         Black
    3752 Service Package                    1
    4110 256 MB (2x128MB) SDRAM DIMMs       1
    5005 Preinstall                         1
    5211 1-Way RS64 IV 600 MHz Processor    1
         Card, 2MB L2 Cache
    6159 12GB/24GB 4mm Tape Drive           1
    6553 SCSI Hot Swap 6-Pack               1
    8700 Quiet Touch Keyboard, Stealth      1
         Black - English (US)
    8741 3-Button Mouse - Stealth Black     1
    9300 Language - English (US)            1
    9800 Power Cord - US/Canada (125V,      1
         15A)

9910-A13 APC Smart-UPS 1400VA, 103-132V,    1
         50/60Hz, 950W

***  SOFTWARE  ***

  Product           Description           Qty

 5765-C34 AIX Version 4                     1
    4061 AIX 4.3                            1
    5005 Preinstall                         1
    9001 Asset Registration                 1

 5765-E26 VisualAge C++ Professional V5     1
    0006 OTC                                1
    3483 Single Pallet Delivery             1
    5819 CD-ROM - English                   1
    9001 Asset Registration                 1

 5692-AIX System Software                   1
    0598 AIX Welcome Center                 1
    0838 AIX 4.3 Update CD                  1
    0857 AIX 4.3                            1
    0859 Preinstall AIX 4.3 Server          1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non           1
         Export.
    0903 CWS Diagnostic Pre-Load            1
    1004 CD-ROM Process Charge              1
    2924 English                            1
    3410 CD-ROM                             1
    5005 Preinstall                         1
```

```
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1
```

7025-6F1 control workstation medium configuration

This configuration is recommended for SP systems that include more than 30 nodes and for SP/CES systems that include no more than 16 enterprise servers. Example 7-8 shows a 7025-6F1 control workstation medium configuration.

Example 7-8 7025-6F1 control workstation medium configuration

```
***  HARDWARE  ***

  Product           Description            Qty

 7025-6F1 pSeries 620                       1
          1.44MB 3.5-in Diskette Drive      1
          Integrated SCSI-2 F/W Adapter     1
          Integrated Ultra2 SCSI Adapter    1
          Integrated Ethernet Adapter       1
     2624 32x Speed CD-ROM                  1
     2830 POWER GXT130P Graphics Adapter    1
          (PCI)
     2943 8-Port Async Adapter,             1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI          3
          Adapter
     3109 SCSI External Port to SCSI        1
          Internal 6 Pack Cable Assembly
     3153 18.2 GB 10K RPM 1" Ultra3 SCSI    4
          16-bit Disk
     3628 P260 Color Monitor, Stealth       1
          Black
     3752 Service Package                   1
     4075 Memory Board, 16-position         1
     4119 512 MB (2x256MB) SDRAM DIMMs      2
     5005 Preinstall                        1
     5202 2-Way RS64 III 450 MHz            1
          Processor Card, 4MB L2 Cache
     6159 12GB/24GB 4mm Tape Drive          1
     6553 SCSI Hot Swap 6-Pack              1
     8700 Quiet Touch Keyboard, Stealth     1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black    1
     9300 Language - English (US)           1
     9800 Power Cord - US/Canada (125V,     1
          15A)

9910-A13 APC Smart-UPS 1400VA, 103-132V,    1
          50/60Hz, 950W
```

```
***  SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5765-E26 VisualAge C++ Professional V5     1
     0006 OTC                               1
     3483 Single Pallet Delivery            1
     5819 CD-ROM - English                  1
     9001 Asset Registration                1

5692-AIX System Software                    1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0903 CWS Diagnostic Pre-Load           1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1
```

7025-6F1 control workstation large configuration

This configuration is recommended for SP systems with a large number of nodes and for SP/CES systems that includes more than 16 enterprise servers. Example 7-9 on page 219 shows a 7025-6F1 control workstation large configuration.

Example 7-9 7025-6F1 control workstation large configuration

```
***  HARDWARE  ***

  Product           Description            Qty

 7025-6F1 pSeries 620                        1
          1.44MB 3.5-in Diskette Drive       1
          Integrated SCSI-2 F/W Adapter      1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2944 128-Port Async Controller (PCI)    1
     2968 10/100 Mbps Ethernet PCI           3
          Adapter
     3109 SCSI External Port to SCSI         1
          Internal 6 Pack Cable Assembly
     3153 18.2 GB 10K RPM 1" Ultra3 SCSI     6
          16-bit Disk
     3628 P260 Color Monitor, Stealth        1
          Black
     3752 Service Package                    1
     4075 Memory Board, 16-position          1
     4119 512 MB (2x256MB) SDRAM DIMMs       2
     5005 Preinstall                         1
     5204 4-Way RS64 III 450 MHz             1
          Processor Card, 4MB L2 Cache
     6159 12GB/24GB 4mm Tape Drive           1
     6553 SCSI Hot Swap 6-Pack               1
     8131 4.5m Controller Cable              1
     8133 RJ-45 to DB-25 Converter Cables    1
          (4 cables per set)
     8137 Enhanced Async Node 16-Port        1
          EIA-232
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada (125V,      1
          15A)

 9910-A13 APC Smart-UPS 1400VA, 103-132V,    1
          50/60Hz, 950W

***  SOFTWARE  ***

  Product           Description            Qty
```

```
5765-C34 AIX Version 4                      1
    4061 AIX 4.3                            1
    5005 Preinstall                         1
    9001 Asset Registration                 1

5765-E26 VisualAge C++ Professional V5      1
    0006 OTC                                1
    3483 Single Pallet Delivery             1
    5819 CD-ROM - English                   1
    9001 Asset Registration                 1

5692-AIX System Software                    1
    0598 AIX Welcome Center                 1
    0838 AIX 4.3 Update CD                  1
    0857 AIX 4.3                            1
    0859 Preinstall AIX 4.3 Server          1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non           1
         Export.
    0903 CWS Diagnostic Pre-Load            1
    1004 CD-ROM Process Charge              1
    2924 English                            1
    3410 CD-ROM                             1
    5005 Preinstall                         1
    5924 AIX 4.3 PII - English              1
    9001 Asset Registration                 1
```

Note: For redundancy reasons, we recommend that you double the number of internal disks, add an additional hot swap SCSI 6-pack, and, if necessary, two Ultra2 SCI PCI adapters to each of the F80/6H1/6F1 configurations. This allows you to build a rootvg separately from a spdatavg which contains all the SP related data. Each half of the internal disks can be connected to a separate SCSI adapter; thus, the mirroring can be done over two separate SCSI interfaces.

7.3 High Availability Control Workstation

The High Availability Control Workstation (HACWS) is a component of the effort to reduce the possibility of single point of failure opportunities in the SP system. There are already redundant power supplies and replaceable nodes. However, there are also many elements of hardware and software that could fail on a

control workstation. With a HACWS, your SP system will have the added security of a backup control workstation. Also, HACWS allows your control workstation to be powered down for maintenance or updating without affecting the entire SP system.

The design of the HACWS is modeled on the High Availability Cluster Multi-Processing for the AIX (HACMP) licensed program product. HACWS utilizes HACMP running on two RS/6000 control workstations in a two-node rotating configuration. HACWS utilizes an external disk that is accessed non-concurrently between the two control workstations for storage of SP related data. There is also a dual RS-232 frame supervisor card with a connection from each control workstation to each SP frame in your configuration. This HACWS configuration provides automated detection, notification, and recovery of control workstation failures.

The HACWS option is available on configurations with SP standard frames only. The configurations with SP-attached servers and Clustered Enterprise Servers are not supported.

7.3.1 Overview

The SP system looks similar except that there are two control workstations connected to the SP Ethernet and TTY network.

.

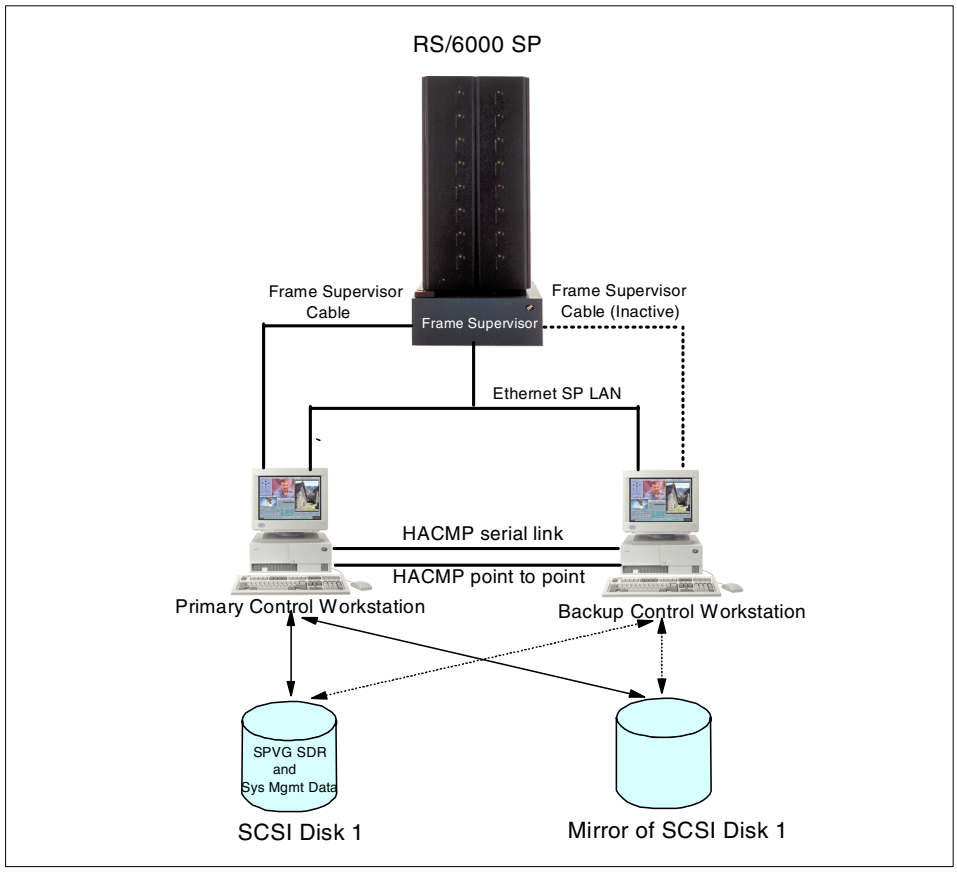


Figure 7-2 High Availability Control Workstation

The frame supervisor TTY network is modified to add a standby link. The second control workstation is the backup

Figure 7-2 shows a logical view of a High Availability Control Workstation. The figure shows disk mirroring, which is an important part of high availability planning.

If the primary control workstation fails, there is a disruptive failover that switches the external disk storage, performs IP and hardware address takeover, restarts the control workstation applications, remounts file systems, resumes hardware monitoring, and lets clients reconnect to obtain services or to update control workstation data. This means that there is only one active control workstation at any time.

The primary and backup control workstations are also connected on a private point-to-point network and a serial TTY link or target mode SCSI. The backup control workstation assumes the IP address, IP aliases, and hardware address of the primary control workstation. This lets client applications run without changes. The client application, however, must initiate reconnects when a network connection fails.

The SP data is stored in a separate volume group on the external disk storage.

The backup control workstation can run other unrelated applications if desired. However, if the application on the backup control workstation takes significant resources, that application may have to be stopped during failover and reintegration periods.

7.3.2 Installation requirements

There are a couple of requirements for hardware and software that must be met before you install the HACWS.

System requirements

In addition to the system requirements described in Section 7.2.2, “System requirements” on page 199, the following are required:

- Two supported pSeries or RS/6000 workstations. Each of these systems must have the same set of I/O required for control workstations as described in Section 7.2.1, “Supported pSeries and RS/6000 workstations and servers” on page 198. They can be different models, but the tty configuration must be exactly the same on each control workstation. The disks should be of the same type and configured the same way on both control workstations to allow the hdiskx numbers to be consistent between the two control workstations.
- External disk storage that is supported by HACMP and the control workstation being used. Two external disk controllers and mirrored disks are strongly recommended but not required. If a single external disk controller is used, the control workstation single point of failure has not been eliminated but moved to the disk subsystem.
- The HACWS connectivity feature (F/C 1245) on each SP frame.
- An additional RS232 connection for HACMP communication is needed if target mode SCSI is not being used for the HACMP communication.

Software requirements

The software requirements for the control workstation include:

- Two AIX server licenses (AIX Version 4.3.3 is required for PSSP Version 3.2).

- Two licenses for IBM C for AIX Version 3.6.6 or later or the batch C and C++ Version 3.6.6 or later compilers and runtime libraries of VisualAge C++ Professional Version 4.0 for AIX or later. If the compiler's license server is on the control workstation, the backup control workstation should also have a license server with at least one license. If there is no license server on the backup control workstation, an outage on the primary control workstation will not allow the SP system access to a compiler license.
- Two licenses and software sets for High Availability Cluster Multi-Processing for AIX (HACMP). HACMP Version 4.4 is required for AIX Version 4.3.3 and PSSP Version 3.2. Both the client and server options must be installed on both control workstations. You must purchase two licenses.
- PSSP 3.2 optional component HACWS. This is the customization software that is required for HACMP support of the control workstation. It comes with your order of PSSP 3.2 as an optionally installable component. Install a copy on both control workstations.

7.3.3 Limits and restrictions

The HACWS support has the following limitations and restrictions:

- You cannot split the load across a primary and backup control workstation either the primary or the backup provides all the functions at one time.
- The primary and backup control workstations must each be a pSeries or an RS/6000 system. You cannot use a node at your SP as a backup control workstation.
- The backup control workstation cannot be used as the control workstation for another SP system.
- The backup control workstation cannot be a shared backup of two primary control workstations.There is a one-to-one relationship of primary to backup control workstations; a single primary and backup control workstation combination can be used to control only one SP system.
- If your primary control workstation is a PSSP Kerberos V4 authentication server, the backup control workstation must be a secondary authentication server.
- If you plan to have DCE authentication enabled, you cannot use HACWS. If you already use HACWS, do not enable DCE authentication.
- HACWS does not tolerate IPV6 aliases for IPV4 addresses.

7.4 Service Director on the CWS

The Service Director software is provided with the SP/CES system. If you plan to use it, consider whether you want to install it on the CWS or somewhere else - any workstation connected to the LAN.

Service Director monitors the health of your SP/CES system. Service Director analyzes AIX error logs and runs diagnostics against those error logs, and if required, Service Director will notify the IBM Support Center. You can define which systems have the Service Director clients and servers and the level of error log forwarding or network access. In addition to notifying the IBM support center, you can also configure Service Director to send an automated E-mail message containing the fault information to your system administrator (requires mail to be active on each node). Upon receiving the fault notification, IBM will automatically dispatch a service engineer (with parts if needed) to correct the problem.

The local Service Director server must have an available serial port. However, if you are using a CWS which utilizes a service processor (F/C 1001), the service processor modem must be connected to the S1 serial port. Thus, you cannot use the S1 serial port for Service Director if your CWS has a service processor installed. The local Service Director server requires the serial port for a modem which transmits fault messages over local telephone lines. All new RS/6000 SP systems include a modem package as part of the ship group. The telecommunication cable plugs and ac power cable plugs vary according to the country code used for the SP system order. This package includes:

- An IBM compatible modem (minimum 9600 bps baud rate).
- A 9- to 25-pin serial cable.
- 15 meter, 25-pin extension cable.

The customer must supply the following:

- A dial-up, analog telephone line (public switched network) with 24-hour, 7-day availability (digital telephone lines cannot be used).
- A telephone extension cable (to connect the modem to the phone jack).

The local Service Director server, in general installed on CWS, and all nodes (SP nodes, SP-attached servers, Clustered Enterprise Servers) in your SP/CES system must have disk space available for installation of the Service Director software.

IBM intents to replace Service Director with Electronic Service Agent.

Note: Refer to Appendix C, “SP system upgrades” on page 407 for more details about the Service Director and the Electronic Service Agent.

8

PCI communication adapters

This chapter provides information for communication adapters and communication adapter extensions supported on the SP.

8.1 PCI Communication adapters supported on the SP

The following is a list of all supported communication adapters and most adapter extensions supported on the SP. Entries are ordered by feature code (F/C). Table 8-1 shows all currently supported PCI adapters. If an adapter in not in this table but it is listed in the descriptions, it is probably an adapter extension or enhancement.

Table 8-1 Supported PCI adapters

Feature code	Adapter ID	Description	Memory Address Space	Card size	AIX Level supported	PSSP level supported	Available
2732	9-W	Short Wave HPPI	8M	Long	4.2.1/4.3.1/ 4.3.3*	2.4*	Yes
2733	9-X	Long Wave HIPPI	8M	Long	4.2.1/4.3.1/ 4.3.3*	2.4*	Yes
2741		FDDI SK-Net LP SAS	2M	Short	4.2.1*	2.4*	Yes
2742		FDDI SK-Net LP DAS	2M	Short	4.2.1*	2.4*	Yes
2743		FDDI SK-Net UTP SAS	2M	Short	4.2.1*	2.4*	Yes
2751	5-5	s/390 ESCON	128M	Long	4.2.1/4.3.2 *	2.4*	Yes
2920	9-O	4/16 Auto LAN streamer	8M	Short	4.2.1*	2.4*	Yes
2943	3-B	8-Port Async EIA 232/RS-422	8M	Short	4.2.1*	2.4*	Yes
2944	3-C	128 Port Async	4M	Short	4.2.1*	2.4*	Yes
2946	A-B	Turboways 622Mbps PCI MMF ATM	8M	Short	4.2.1/4.3.2 *	2.4*	No

Feature code	Adapter ID	Description	Memory Address Space	Card size	AIX Level supported	PSSP level supported	Available
2947	9-R	IBM ARTIC960Hx 4-Port Selectable PCI	8M	Long	4.2.1/4.3.2*	2.4*	Yes
2962	9-L	2-Port Multiprotocol X.25	0	Short	4.2.1*	2.4*	Yes
2963	9-J	TURBOWAYS 155 UTP ATM	8M	Short	4.2.1*	2.4*	Yes
2968	9-P	10/100 Ethernet TX PCI	16M	Short	4.2.1*	2.4*	Yes
2969	9-U	Gigabit Ethernet SX	16M	Short	4.3.2*	2.4*	Yes
2975	A-A	IBM 10/100/1000 Base-T Ethernet PCI	16M	Short	?	?	No
2985	8-Y	10Base2 and 10BaseT (BNC/RJ-45) Ethernet LAN	2M	Short	4.2.1*	2.4*	Yes
2987	8-Z	10Base5 and 10BaseT (AUI/RJ-45) Ethernet LAN	2M	Short	4.2.1*	2.4*	Yes
2988	9-F	TURBOWAYS 155 MMF ATM	8M	Short	4.2.1*	2.4*	Yes

Feature code	Adapter ID	Description	Memory Address Space	Card size	AIX Level supported	PSSP level supported	Available
4951	9-Z	IBM 4-Port 10/100 Base-TX Ethernet PCI	16M	Short			No
4958	6-H	IBM PCI Cryptographic Coprocessor	128M	Long			No
4959	9-Y	4/16 Auto LAN streamer Full Duplex Token Ring	8M	Short	4.2.1/4.3.3*	2.4*	Yes
6204	4-U	IBM PCI Universal Differential Ultra SCSI	16M	Short	4.3.3*	3.1*	Yes
6205	4-R	DUAL CHANNEL ULTRA2 SCSI	32M	Long	4.2.1/4.3.3*	2.4*	Yes
6206	4-K	SCSI-2 Ultra/Wide	16M	Short	4.2.1*	2.4*	Yes
6207	4-L	SCSI-2 Differential Ultra/Wide	16M	Short	4.2.1*	2.4*	Yes
6208	4-A	SCSI-2 Fast/Wide	16M	Short	4.2.1*	2.4*	Yes
6209	4-B	SCSI-2 Differential Fast/Wide	16M	Short	4.2.1*	2.4*	Yes
6215	4-N	SSA RAID 5	2M	Long	4.2.1*	2.4*	Yes
6225	4-P	Advance Serial RAID	2M	Long	4.2.1/4.3.1/4.3.3*	2.4*	Yes

Feature code	Adapter ID	Description	Memory Address Space	Card size	AIX Level supported	PSSP level supported	Available
6227	4-S	Gigabit Fibre Channel	32M	Short	4.3.3*	3.1*	Yes
6230	4-P	Advanced Serial RAID Plus	2M	Long	4.2.1/4.3.2 *	2.4*	Yes
6210	6-E	IBM ARTIC960R xD Quad Digital Trunk PCI	4M	Long	4.2.1/4.3.2 *	2.4*	Yes
6311	6-G	IBM Artic960RxF Digital Trunk Resource	4M	Long			No

* This version or later

8.2 Descriptions of adapter supported on the SP

The following is a brief description of all PCI adapters supported on the SP. MCA (Micro Channel) and switch adapters are not discussed in this chapter. A list of MCA adapters supported on the SP is in **Appendix B, "Hardware" on page 365,** and the switch adapter lists can be found in Chapter 5, "The Switch" on page 159 and in Chapter 6, "SP Switch Routers 9077" on page 181.

8.2.1 IBM Network Terminal Accelerator adapter (F/C 2403)

The IBM Network Terminal Accelerator adapter (F/C 2403) is an Ethernet adapter that accelerates network performance by off-loading the telnet and rlogin daemons, TCP/IP protocol stack, and virtual terminal I/O management from the RS/6000 system. The adapter buffers the system from frequent CPU intensive packet interrupts, increases terminal I/O throughput and the number of concurrent online user sessions by up to three times, and reduces context switches, which dramatically reduces the CPU load.

Feature characteristics

This feature has the following characteristics:

- Network adapter software provides a pass-through capability for other Ethernet protocols, which can eliminate the need for a separate Ethernet adapter.
- The network adapter supports on board SNMP for network management.

8.2.2 SCSI-2 Differential Fast/Wide adapter/A (F/C 2416)

The SCSI-2 Differential Fast/Wide adapter (F/C 2416) is a dual-ported fast (10 MHz) and wide (two bytes) adapter. It provides synchronous SCSI bus rates up to 20 Mbps and attaches to SCSI fixed disks, CD-ROM devices, tape drives, R/W optical devices, and storage subsystems. The maximum data rate depends on the maximum rate of the attached device.

8.2.3 SCSI-2 Differential External I/O Controller (F/C 2420)

The SCSI-2 Differential External I/O Controller feature code (F/C 2420) allows you to attach external SCSI-2 differential devices. This adapter provides SCSI bus signal cable quality and a maximum SCSI bus length of up to 19 meters (62.3 feet).

8.2.4 IBM Short Wave Serial HIPPI PCI adapter (FC 2732)

The IBM Short-Wave Serial HIPPI PCI adapter (F/C 2732) provides high-speed connectivity via the ANSI serial HIPPI channel. It provides (via short-wave optics) the capability to participate in supercomputer environments, attach to disk-array subsystems, other SP systems, HIPPI switches, other vendor computers, and tape subsystems. It is a 32-bit, 33 MHz universal PCI serial HIPPI adapter and supports TCP/IP for communication. Data is sent and received over optical fiber at 1.2 Gbps using the HIPPI standard 20/24-bit encoding scheme. The effective maximum data rate of the HIPPI interface is 800 Mbps.

Feature characteristics

This feature has the following characteristics:

- Single-slot, full-size 32-bit PCI adapter
- PCI 2.1 specification compatible
- Intel 960 processor
- 2 MB DRAM program store
- 2 MB transmit and receive RAM
- Short-wave optics

Feature components

This feature order provides the following:

- Adapter card
- Diagnostic wrap plugs
- Installation instructions

Customer components

You must supply the following components for this feature:

- Short-wave cabling
- Dual SC connector 50/125 micron multi-mode fiber

8.2.5 Long-Wave Serial HIPPI adapter (F/C 2733)

The Long-Wave Serial HIPPI adapter (F/C 2733) provides high-speed connectivity via the ANSI serial HIPPI channel. It provides (via long-wave optics) the capability to participate in supercomputer environments, attach to disk-array subsystems, other SP systems, HIPPI switches, other vendor computers, and tape subsystems. It is a 32-bit, 33 MHz universal PCI serial HIPPI adapter and supports TCP/IP for communication. Data is sent and received over optical fiber at 1.2 Gbps using the HIPPI standard 20/24-bit encoding scheme. The effective maximum data rate of the HIPPI interface is 800 Mbps.

Feature characteristics

This feature has the following characteristics:

- Single-slot, full-size 32-bit PCI adapter
- PCI 2.1 specification compatible
- Intel 960 processor
- 2 MB DRAM program store
- 2 MB transmit and receive RAM
- Long-wave optics

Feature components

This feature order provides the following:

- Adapter card
- Diagnostic wrap plugs
- Installation instructions

8.2.6 IBM Network Terminal Accelerator adapter (F/C 2402 & 2403)

The IBM Network Terminal Accelerator adapter (F/C 2402) is an Ethernet adapter that accelerates network performance by off-loading the telnet and rlogin daemons, TCP/IP protocol stack and virtual terminal I/O management from the RS/6000 system. The network adapter buffers the system from frequent CPU intensive packet interrupts, increases terminal I/O throughput and the number of concurrent online user sessions, and reduces context switches, which dramatically reduces the CPU load.

The network adapter software provides a pass-through capability for other Ethernet protocols, which can eliminate the need for a separate Ethernet adapter. The network adapter supports on board simple network management protocol (SNMP) for network management.

8.2.7 SCSI-2 Fast/Wide adapter/A (F/C 2416)

The SCSI-2 Differential Fast/Wide adapter (F/C 2416) is a dual-ported fast (10 MHz) and wide (two bytes) adapter. It provides synchronous SCSI bus rates up to 20 Mbps and attaches to SCSI fixed disks, CD-ROM devices, tape drives, R/W optical devices, and storage subsystems. The maximum data rate depends on the maximum rate of the attached device.

8.2.8 SCSI-2 Differential External I/O Controller (F/C 2420)

The SCSI-2 Differential External I/O Controller feature (F/C 2420) allows you to attach external SCSI-2 differential devices. This adapter provides SCSI bus signal cable quality and a maximum SCSI bus length of up to 19 meters (62.3 feet).

8.2.9 8-Port Async adapter - EIA-422A (F/C 2940)

The 8-port Async feature (F/C 2940) provides the RS/6000 500 series system with up to eight EIA-422A asynchronous serial devices such as terminals and printers. The 8-port Async adapter contains all of the electronics required to support eight asynchronous ports and uses one I/O card slot.

8.2.10 SYSKONNECT SK-NET FDDI adapters (F/C 2741, 2742 & 2743)

There are three types of SYSKONNECT SK-NET FDDI adapters they include:

- LP Single Attach Station (SAS) feature code - 2741
- LP Dual Attach Station (DAS) feature code - 2742
- UP Dual Attach Station (DAS) feature code - 2743

They are all PCI fiber optic FDDI adapters that are compatible with the FDDI-ANSI X3T12 specifications and FDDI Standard Series. IBM no longer sells FDDI adapters, but it does support them. The following are some common characteristics of the adapters.

Feature characteristics

These adapters have the following characteristics:

- Supports single-ring FDDI attachment at 100 Mbps via a customer-supplied FDDI concentrator.
- Supports all TCP/IP protocols and ANSI Station Management (SMT) 7.3.

Feature components

This feature order provides the following:

- Adapter card
- Diagnostic wrap plug
- Diskette with adapter device driver
- Installation instructions

Customer supplied components

You must supply the following components for this feature:

- An FDDI concentrator, such as the IBM 8240 (or equivalent) concentrator, to connect to your FDDI local area network.
- Vice driver and FDDI common code (provided with adapter).

Unique features of FDDI-LP SAS

The adapter provides single attachment to an FDDI concentrator (or point-to-point) using fiber optic cabling (not supplied with the adapter).

Customer components

You must supply the following components for this feature:

- One 62.5/125 micron multimode fiber duplex cable with SC connectors.

Unique features of FDDI-LP DAS

The adapter provides either dual attachment to the main ring path or dual homing to one or two FDDI concentrators using fiber optic cabling (not supplied with the adapter).

Customer components

You must supply the following components for this feature:

- Two 62.5/125 micron multimode fiber duplex cables with SC connectors.

Unique features of FDDI-UP DAS

The adapter provides single attachment to a FDDI concentrator (or point to point) using Category 5 Unshielded Twisted Pair (UTP) cabling (not supplied with the adapter).

Customer supplied components

You must supply the following components for this feature:

- One UTP Category 5 cable

8.2.11 S/390 ESCON Channel PCI adapter (F/C 2751)

The PCI S/390 ESCON Channel adapter (F/C 2751) provides the SP system with an attachment to IBM Enterprise Systems Connection (ESCON) channels on System/390 mainframes. This direct ESCON channel connection provides a fiber optic link that can take advantage of ESCON Directors (fiber optic switches) permitting multiple channel connections. The adapter supports: VM/ESA, MVS/ESA, and OS/390.

Feature characteristics

This feature has the following characteristics:

- Full length PCI adapter
- Supports attachment to either 10 MB or 17 MB ESCON channels
- Supports VM, MVS, and OS/390
- Supports CLIO/S
- Supports ESCON multiple Image Facility (EMIF)
- Maximum distance supported, 43 Km using LED and XDF ESCON links
- S/390 TCP/IP for VM and MVS
- PCI 32-bit Bus Master adapter

Feature components

This feature order provides the following:

- One full length PCI adapter
- CD-ROM with device drivers
- Instruction manual
- Diagnostic wrap plug

Customer supplied components

The customer must supply the following components for this feature:

- ESCON cabling requires 62.5/125 multimode fiber cable with ESCON duplex connectors on both ends.
- AIX program feature, ESCON Control Unit LPP.

8.2.12 PCI Auto LANstreamer Token Ring adapter (F/C 2920)

The PCI Auto LANstreamer Token Ring adapter (F/C 2920) is a PCI 16/4 Token Ring adapter that is compatible with IEEE 802.5 specifications. The adapter has two external connections: RJ-45 to attach to UTP cabling and a 9-pin D-Shell to attach to STP cabling. This adapter is no longer available from IBM.

Feature characteristics

This feature has the following characteristics:

- Complies with IEEE 802.5 specifications
- Attaches to 4 MBps or 16 MBps token-ring area networks
- Supports both full and half duplex operations
- Has a PCI 32-bit Bus Master adapter

Feature components

This feature order provides the following:

- Adapter card
- Diskette with adapter device driver
- Installation instructions

Customer supplied components

The customer must supply the following components for this feature:

- Network equipment, such as a MAU and/or switching hub to connect the Token Ring network.
- UTP or STP cable to attach the adapter to the Token Ring network adapter device driver.

8.2.13 EIA 232/RS-422 8-Port Asynchronous adapter (F/C 2943)

The 8-port Async adapter (F/C 2943) provides the RS/6000 SP with up to eight EIA 232 or RS-422 asynchronous serial lines from a single PCI bus slot. This adapter adheres to the PCI Revision 2.1 standards for EIA 232 and RS-422. It features a low cost, high performance 32-bit card, 33 MHz bus speed, and a PCI bus transfer rate of 132 MBps.

This adapter provides a single DB-78 output that connects directly to the 8-port DB-25 connector box. All eight ports are software programmable to support either protocol at baud rates up to 230 K. The full set of modem control lines for asynchronous communication are provided for each port. Devices such as terminals, modems, processors, printers, and controllers may be attached.

Feature characteristics

This feature has the following characteristics:

- 8-port asynchronous device connections
- 32-bit Bus Master PCI bus (132 MB per second)
- Short-form factor PCI adapter
- EIA-232 maximum distance 31 m and 62 m dependent on baud rate and RAN
- RS-422 maximum distance 1200 m dependent on baud rate
- 230 K maximum baud rate
- Supports TxD, RxD, RTS, CTS, DSR, DCD, DTR, and RI on EIA 232
- Supports +TxD, -TxD, +RxD, and -RxD on RS-422

Feature components

This adapter order features the following:

- Adapter card
- 25-pin diagnostic wrap
- Diskette with adapter device driver
- Installation instructions
- External 3 meter DB78 cable to 8-port DB25 breakout box

Customer supplied components

A 3 m cable with an attached breakout box is supplied with each adapter. You must supply the following component for this feature:

- All cables needed to connect peripheral equipment to this adapter.

8.2.14 WAN RS232 128-port adapter (F/C 2944)

The 128-port Async adapter (F/C 2944) provides the SP with up to 128 EIA-232 asynchronous serial lines from a single PCI bus slot. This adapter adheres to the PCI standard. It features a low cost, high performance 32-bit card, 33 MHz bus speed, and a PCI bus transfer rate of 132 MBps.

Two 2.4 MBps synchronous channels link the adapter to a maximum of eight 16-port Remote Async Nodes (RANs). Each synchronous channel uses an HD-15 female connector to link up to four RANs. Each RAN supports either EIA-232 or RS-422 connections (16 per RAN), and up to eight RANs may be connected together yielding a total of 128 ports. The RAN utilizes an RJ-45 connector to provide interface signals at speeds up to 230 K baud at a limited number of ports.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master PCI bus
- Two synchronous channels to RAN
- EIA-232 maximum distance of 31 meter and 62 meter depending on baud rate and RAN
- RS-422 maximum distance 1200 meter dependent on baud rate

Customer supplied components

F/C 2944 utilizes the following optional RANs and device cables, which are available from IBM:

- 1.2 MBps RANs and cables:
  - F/C 8130: 1.2 MBps Remote Asynchronous Node, 16-port, EIA-232 (US)
  - F/C 8131: 128-port Asynchronous Controller Node Cable, 4.5 m
  - F/C 8132: 128-port Asynchronous Controller Cable 23 cm (9 in.)
  - F/C 8133: RJ-45 to DB-25 Converter Cable
  - F/C 8134:1.2 MBps Remote Asynchronous Node, 16-port, EIA-232 (world trade)
  - F/C 8136: 1.2 MBps Rack Mountable Remote Asynchronous Node, 16-port, EIA-232
- 2.4 MBps RANs and cables:
  - F/C 8137: 2.4 MBps Enhanced Remote Asynchronous Node, 16-port, EIA-232
  - F/C 8138: 2.4 MBps Enhanced Remote Asynchronous Node, 16-port, RS-422
  - F/C 2934: Asynchronous Terminal/Printer Cable, EIA-232
  - F/C 3124: Serial port to serial port cable for drawer-to-drawer connections
  - F/C 3125: Serial port to serial port cable for rack-to-rack connections

8.2.15 Turboways 622 Mbps PCI MMF ATM adapter (F/C 2646)

The IBM Turboways 622 Mbps PCI MMF ATM adapter (F/C 2646) is a 64-bit, Universal PCI Adapter. This adapter provides direct access to the ATM network at a dedicated 622 Mbps full-duplex connection. The Turboways 622 Mbps PCI MMF ATM Adapter is a short form-factor adapter that interfaces to the system via the PCI bus and connects to the 622 Mbps ATM network via dual-SC type, multi-mode fiber cables. The Turboways 622 Mbps PCI MMF ATM Adapter utilizes 16MB of SDRAM for control and 16MB of SDRAM for packet memory. This ATM adapter also provides a hardware assist for TCP checksum which can provide a performance improvement by minimizing the host CPU cycles.

Feature components

This feature order provides the following:

- Adapter Card
- One full-duplex connection to 622 ATM network

8.2.16 ARTIC960Hx 4-port Selectable adapter (F/C 2947)

The ARTIC960Hx 4-Port Selectable PCI adapter (F/C 2947) is a one-slot, standard-length, 32-bit PCI card. It provides 4-Ports of either EIA-232, EIA530, RS-449, X.21, or V.35. Only one standard-length adapter can be used at a time. Each port supports speeds up to 2.0 Mbps. Software support is provided by ARTIC960 Support for AIX, Developer's Kit, AIX versions 4.2.1 or 4.3.2 or later, that provide SDLC and Bisync support. The adapter can also be used for real-time device control, telephony signaling, and custom serial communication protocols.

This adapter is also equipped with a high-performance, eight-channel DMA controller. This DMA controller supports intelligent DMA operations, such as data buffer chaining and end-of-frame processing, to support high-performance communications protocols and high-throughput applications. The DMA controller is fully programmable for OEM and third-party device drivers.

Feature characteristics

This feature has the following characteristics:

- One 120-pin port
- Supports up to four connections of the same type
- Data transfer rates of up to 2 Mbps
- Supported interfaces are:
  - EIA-232

  - EIA-530
  - RS-449
  - X.21
  - V.35
- Support for SDLC and X.25 full-duplex, synchronous protocols

Featured components

This feature order provides the following:

- One ARTIC960Hx adapter (F/C 2947)
- A connecting cable (required). The following are available from IBM:
  - F/C 2861: ARTIC960Hx 4-port EIA-232 cable
  - F/C 2862: ARTIC960Hx 4-port RS-449 cable
  - F/C 2863: ARTIC960Hx 4-port X.21 cable
  - F/C 2864: ARTIC960Hx 4-port V.35 (DTE) cable
  - F/C 2865: ARTIC960Hx 4-port EIA-530 cable

8.2.17 2-port Multiprotocol X.25 adapter (F/C 2962)

The 2-Port Multiprotocol adapter (F/C 2962) provides the SP with high speed connections between stand alone system units on a wide area network (WAN). This adapter adheres to the Peripheral Component Interconnect (PCI) standard and also supports SDLC and X.25 protocols. The 2-port Multiprotocol adapter connects to WAN lines through externally attached data communication equipment including Channel Service Units (CSU), Data Service Units (DSU), and synchronous modems.

This adapter operates at speeds up to 2.048 Mbps and provides two ports that accommodate four selectable interfaces. These interfaces are:

- EIA 232D/V.24
- V.35
- V.36/EIA 449
- X.21

Interface configuration is selected by the type of cable attached. These cables are ordered separately, and you may configure with the 2-Port Multiprotocol adapter with two different cables.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master PCI 2.1 adapter
- Provides two, 36-pin high density (male) ports
- Provides four interface types, EIA 232D/V.24, V.35, V.36/EIA 449, and X.21
- Simultaneously supports two different interfaces
- Supports SDLC and X.25 full duplex synchronous protocols

Customer supplied components

If you plan to operate this adapter using X.25 protocols, then you must separately order the IBM AIXLINK/X.25 LPP (5696-926). This package provides a V.24, V.35, or X.21 port connection to X.25 packet switched networks.

The system interface is determined by the cable connected to this adapter. See Table 8-2 for a list of available cables and the interface supported by each cable.

The 2-port Multi protocol adapter can be configured with different cable types on each port.

Table 8-2 Cable information for 2-port multi protocol adapter

F/C	Interface Configuration	Cable Terminations (Length)
2951	EIA 232D/V.24 cable	36-pin to male DB25 (3 m)
2952	V.35 cable	36-pin to 34-pin male (3 m)
2953	V.36/EIA 449 cable	36-pin to 37-pin male (3 m)
2954	X.21 cable	36-pin to male DB15 (3 m)

8.2.18 ATM 155 TURBOWAYS UTP adapter (F/C 2963)

The TURBOWAYS 155 UTP ATM adapter (F/C 2963) enable TCP/IP applications to work in an asynchronous transfer mode (ATM) environment. This adapter provides dedicated 155 MB per second, full-duplex connection to ATM networks using either Permanent Virtual Circuits (PVC) or ATM Forum compliant Switched Virtual Circuits (SVC) UNI 3.1 signalling. The adapter supports AAL-5 adaptation layer interface and communication with devices located on an ATM network, bridged token ring, Ethernet, or other LAN. LAN Emulation (LANE) is provided by the AIX operating system.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master PCI 2.1 adapter
- External RJ45 connector
- Provides signaling channel setup
- Provides virtual connection setup and tear down
- Supports point-to-point and point-to-multipoint switching
- Supports virtual circuits (maximum 1024)
- Supports classical IP and ATRP over ATM (RFC 1577)
- Supports Ethernet LAN Emulation and token ring
- Supports ATM SNMP
- Best effort service

Customer supplied components

You must supply the following components with this feature:

- Category 5 High Speed UTP cables (or shielded) with RJ45 connectors (100 m maximum length).
- If you plan to use multipoint connections, you must provide an ATM switch.

8.2.19 Ethernet 10/100 MB adapter (F/C 2968)

The IBM 10/100 Ethernet TX PCI adapter (F/C 2968) is a 10/100 PCI Ethernet adapter that is compatible with IEEE 802.3 and 802.3u specifications. The adapter has one RJ-45 connection that supports connections to 100BaseTX and 10BaseT networks.

Feature characteristics

This feature has the following characteristics and requirements:

- Compatible with IEEE 802.3 Standards
- 32-bit Bus Master PCI Bus 132 MBps
- Supports auto-negotiation of media speed and duplex operation
- Supports both full and half duplex operation over 10BaseT networks via the RJ-45 connector

Feature components

This feature order provides the following:

- Adapter card
- Diskette with adapter device driver

- Installation instructions

Customer supplied components

You must supply the following components for this feature:

- Network equipment, such as a hub or switch, required to attach to 10BaseT Ethernet LANs.
- All Ethernet cables.
- For 100BASE-TX connections, Unshielded Twisted Pair (UTP) Category 5 cabling is required.

8.2.20 Gigabit Ethernet - SX adapter (F/C 2969)

The PCI Gigabit Ethernet - SX adapter (F/C 2969) is a 1000 Mbps PCI Ethernet adapter that is compatible with IEEE 802.3z specifications. The adapter has one external fiber connection that attaches to 1000BaseSX networks via 50 and 62.5 micron multimode cables with SC connectors.

Feature characteristics

This feature has the following characteristics and requirements:

- Compatible with IEEE 802.3z Standards
- Supports full duplex operation over 1000BaseSX networks
- Supports jumbo frames with AIX 4.3.2 device driver

Feature components

This feature order provides the following:

- Adapter card
- Fiber wrap plug
- Installation instructions

Customer supplied components

You must supply the following components for this feature:

- Network equipment such as a switch or router is required to attach to 1000BASESX networks.
- All Ethernet cables.The maximum operating distances for the fiber cables are:
  - 260 meters with 62.5 micron multimode fiber
  - 440 meters with 50 micron multimode fiber

8.2.21 10/100/1000 Base-T Ethernet PCI adapter (FC 2975)

10/100/1000 Base-T Ethernet PCI adapter (F/C 2975) is a Full Duplex Gigabit Ethernet adapter designed with highly integrated components to optimize cost and performance. The adapter interfaces to the system via the PCI bus and connects to the network using a 4-pair CAT-5 Unshielded Twisted Pair (UTP) cable for distances of up to 100m. The 10/100/1000 Base-T Ethernet PCI Adapter supports jumbo frames (up to 9000 bytes) for full duplex Fast & Gigabit Ethernet.

Feature components

This feature order provides the following:

- Adapter card
- One full-duplex 1000Base-T UTP connection to a Gigabit Ethernet LAN

Limitations:

- AIX's Network Install Manager (NIM) boot is not supported with this adapter.
- The 1000 Mbps speed is not supported in Half Duplex (HDX) mode.

8.2.22 TURBOWAYS 100 ATM adapter (F/C 2984)

The TURBOWAYS 100 ATM adapter (F/C 2984) enables TCP/IP applications to work in an asynchronous transfer mode (ATM) environment. One virtual connection is dedicated to each IP address, and a transformation of each IP address to the corresponding virtual connection is performed.The initial release supports AAL-5 adaptation layer interface and supports 1024 active virtual connections.

8.2.23 Ethernet LAN PCI adapters (F/C) 2985 and 2987)

The PCI Ethernet BNC/RJ-45 and AUI/RJ-45 adapters (F/C 2985 and 2987 respectively) are 10 Mbps PCI Ethernet adapters that is compatible with IEEE 802.3 specifications. The adapters has two external connections: BNC to attach to 10BASE2 or 10BASE5 networks and RJ-45 to attach to 10BASE-T networks.

Feature characteristics

This feature has the following characteristics and requirements:

- 10 Mbps Ethernet compatible with IEEE 802.3 Standards
- 32-bit Bus Master PCI Bus 132 Mbps

- Supports half duplex operations over 10BASE2 networks via the BNC connector
- Supports both full and half duplex operation over 10BASE-T networks via the RJ-45 connector

Feature components

This feature order provides the following:

- Adapter card
- RJ-45 and BNC diagnostic wrap plugs
- Installation instructions

Customer supplied components

You must supply the following components for this feature:

- Network equipment such as a hub or switch required to attach to 10BASE-T Ethernet LANs.
- All Ethernet cables.

  For 10BASE-T connections, Unshielded Twisted Pair (UTP) Category 3, 4, or 5 cabling is required. UTP Category 5 cabling is strongly suggested to facilitate upgrades to 100 Mbps Ethernet LAN without cabling changes.

8.2.24 ATM 155 MMF adapter (F/C 2988)

The TURBOWAYS 155 ATM adapter (F/C 2988) enables TCP/IP applications to work in an asynchronous transfer mode (ATM) environment. This adapter provides dedicated 155 MB per second, full-duplex connection to ATM networks using either Permanent Virtual Circuits (PVC) or ATM Forum compliant Switched Virtual Circuits (SVC) UNI 3.1 signalling. The adapter supports AAL-5 adaptation layer interface and communication with devices located on an ATM network, bridged token ring, Ethernet, or other LAN. LAN Emulation (LANE) is provided by the AIX operating system.

Feature characteristics

This feature has the following characteristics and requirements:

- Provides signaling channel setup
- Provides virtual connection set up and tear down
- Supports point-to-point and point-to-multipoint switching
- Supports virtual circuits (maximum 1024)
- Supports classical IP and ATRP over ATM (RFC 1577)

- Supports Ethernet LAN Emulation and Token Ring
- Supports ATM SNMP

Customer supplied components

You must supply the following components with this feature:

- Plenum rated 62.5/125 multimode fiber cables terminated with an SC connector
- An ATM switch

8.2.25 Four-Port 10/100 BASE-TX Ethernet PCI adapter (F/C 4951)

The IBM Four-Port 10/100 BASE-TX Ethernet PCI adapter (F/C 4951) makes available four Ethernet ports using a single PCI slot. It is a 32/64-bit, long PCI adapter, supporting four industry-standard Ethernet 10BASE-T or 100BASE-TX interfaces supporting 10 or 100 Mbps data rates, either half or full duplex on each of four separate ports. Each port has an RJ-45 connector for attachment to standard CAT-3/5 unshielded, twisted-pair cable. The adapter is IEEE 802.3u compatible and has full auto-negotiation for detecting speed and duplex capability across each port. It provides network boot and Network Install Manager (NIM) capability.

Feature characteristics

This feature has the following characteristics and requirements:

- Supports UTP-5 (RJ45) cable connections
- Capable of communicating at 10 or 100 Mbps
- Has auto-negotiation to determine speed and duplex capability
- Supports 32/64 bit PCI data width
- Supports both full and half duplex operation
- Includes status LEDs for speed and data activity
- Supports NIM install
- Meets PCI 2.1 specifications
- Operates at PCI bus speed of 33 MHz
- Fits in full-size PCI slots

Featured components

This feature order provides the following:

- Adapter card

- Diskette with adapter device driver
- Installation Instructions.

Customer supplied components

You must supply the following components for this feature:

- Network equipment such as a hub or switch.
- All Ethernet cables. For 100BASE-TX connections, Unshielded Twisted Pair (UTP) Category 5 cabling is required.

8.2.26 Crytographic coprocessor PCI adapter (F/C 4958)

The PCI Cryptographic Coprocessor adapter (F/C 4958) is a 2/3 length PCI adapter combining hardware and software to provide high performance, secure hardware engines for secure internet transactions such as secure data exchange, verifying electronic signatures, bulk data encryption and decrypting. Cryptographic processes are performed within a tamper-proof enclosure on the adapter that is designed to meet FIPS PUB 140-1 standard for commercial cryptographic devices at security Level 3.

Feature characteristics

This feature has the following characteristics:

- PCI 2.1 compatible, universal 3.3/5 V adapter
- 486 DX2 Internal processor at 99 MHz
- 8 MB of RAM
- 4 MB of FLASH ROM
- 32 KB of battery-backed RAM
- Supports standard interrupts, DMA, controls, real-time clock
- PCI 32-bit Bus Master

Featured components

This feature provides the following:

- Adapter card
- CD-ROM with adapter device driver
- Installation instructions

Limitations:

- The PCI Cryptographic Coprocessor Adapter is a field-installed only device in order to meet restrictive shipping requirements.

8.2.27 High-Speed Token-Ring PCI adapter (F/C 4959)

The High-Speed Token-Ring PCI adapter (F/C 4959) is a PCI 16/4 Token Ring Adapter that is compatible with IEEE 802.5 specifications. The adapter has two external connections: RJ-45 to attach to UTP cabling and a 9-pin D-Shell to attach to STP cabling.

Feature characteristics

This feature has the following characteristics:

- Supports full duplex operation at all three speeds
- Supports both UTP-5 (RJ-45) and STP (9-Pin D-shell)
- PCI bus specification 2.1:
  - Fits in PCI half-size slots
  - Supports both 5.0 and 3.3 volt signaling
  - Supports PCI data streaming
  - Operates in 64-bit slots as a 32-bit device
  - Operational at PCI bus speeds from 16 MHz to 33 MHz
- Consumes less than 2 watts of power
- Includes adapter and ring-status LEDs
- Supports field update of on-card microcode, via sectored FLASH EPROM
- On-card diagnostics implemented in microcode
- FCC Class B and CISPR Class B certified
- Supports NIM functions

Feature components

This feature order provides the following:

- Adapter card
- CD-ROM with adapter device driver
- Installation instructions

Customer supplied components

You must supply the following components for this feature:

- Network equipment such as a MAU and/or switching hub to connect the token-ring network
- UTP or STP cable to attach to the token-ring network

8.2.28 SCSI-2 Ultra/Wide DE PCI adapter (F/C 6204)

The PCI SCSI-2 Ultra/Wide Differential adapter (F/C 6204) provides a differential SCSI-2 Ultra/Wide interface that can burst data between devices on the SCSI bus at 40 MBps. F/C 6204 supports Ultra and Fast/Wide synchronous data transfers and it supports external devices (no internal connections) up to 25 m away. This adapter conforms to SCSI-2 standards and the Fast-20 (Ultra) documentation. Industry standard SCSI P (68-pin) connectors are incorporated on the adapter.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master Adapter
- Supports attachment of external 8-bit or 16-bit SCSI devices on the J2 port using a 68 pin SCSI-3 standard connector

Customer components

You must supply the following component for this feature:

- Cabling (available through IBM)

Limitations: Data transfer rates with F/C 6204 are limited to the speed of the slowest device on the SCSI bus.

8.2.29 Dual Channel Ultra2 SCSI PCI adapter (F/C 6205)

The Dual Channel Ultra2 SCSI PCI adapter (F/C 6205) is an ideal solution for applications requiring large block data transfers (more than 64K block size) in a multi-disk-drive environment utilizing Ultra/Ultra2 SCSI protocol. It provides up to 160 MBps aggregate SCSI throughput and is able to support single-ended Fast/Ultra devices or LVD Ultra/Ultra2 devices. The dual channels offer increased connectivity without utilizing an additional PCI slot. One or both channels can be dedicated to LVD devices or as an alternative, one channel can be used to support mixed performance single-ended devices. Industry standard VHDCI connectors are available for external connection to each channel.

Feature characteristics

This feature has the following characteristics:

- Two Ultra2/LVD SCSI buses
- PCI bus specification 2.1
  - Fits in PCI full-size slots
  - Supports both 5.0 and 3.3 volt signaling
  - Supports PCI data streaming
  - Two independent DMA Channels
  - 64-bit PCI Bus Master adapter also operates in a 32-bit PCI slot
  - Operational at PCI bus speeds from 16 MHz to 33 MHz
- Supports 16 bit single ended or LVD connections
- Uses Ultra2 SCSI standard external VLHDCI (Very High Density Cable Interconnect or 8mm) SCSI connectors per channel
- Ultra2 SCSI provides increased connectivity (cable length and number of SCSI devices supported) over Ultra SCSI
- NIM Boot
- Native Boot support on AIX 4.3.3

Feature components

This feature order provides the following:

- Adapter card
- CD-ROM with adapter device driver
- Installation instructions

Customer components

You must supply the following component for this feature:

- Cabling (available through IBM)

8.2.30 SCSI-2 Ultra/Wide Single Ended PCI adapter (F/C 6206)

The PCI SCSI-2 Ultra/Wide Single Ended adapter (F/C 6206) provides a single ended SCSI-2 Ultra/Wide interface that can burst data between devices on the SCSI bus at 40 MBps (twice the fast/wide rate) using block sizes greater than 64 K. It conforms to SCSI-2 standards and Fast-20 (Ultra) documentation. F/C 6206 supports both internal and external devices connected to the same SCSI bus. Industry standard SCSI P (68-pin) connectors are incorporated on the adapter.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master PCI 2.1 adapter
- Supports attachment of internal and external single ended 8-bit and 16-bit SCSI or Ultra SCSI devices
  - External connections on J2 with 68 pin SCSI-3 standard P connector
  - Internal connections on J3 with 68 pin high density SCSI connector for 16-bit attachments
  - Internal connections on J4 with 50 pin (2x25) SCSI connector for 8-bit attachments

Customer components

You must supply the following components for this feature:

- If you are using F/C 6206 to configure independent internal disk in an 332 MHz SMP wide node, you must also order F/C 1241.

Adapter limitations:

- Data transfer rates are limited to the speed of the slowest attached device. For example, if you connect an Ultra drive and a fast/wide drive, the adapter will limit data transfers to fast/wide rates.
- If a cable is attached to the external J2 connector, data transfer rates will be limited to fast/wide rates.
- Ultra data transfer rates can only be achieved using the internal connections with cable lengths of 1.5 m or less.
- External cable lengths are limited to 3 m for fast/wide data transfer rates.
- The internal J3 and J4 connectors cannot be used at the same time.

8.2.31 Ultra SCSI Differential adapter (F/C 6207)

The PCI SCSI-2 Ultra/Wide Differential adapter (F/C 6207) provides a differential SCSI-2 Ultra/Wide interface that can burst data between devices on the SCSI bus at 40 MBps. F/C 6207 supports Ultra and Fast/Wide synchronous data transfers, and it supports external devices (no internal connections) up to 25 m away. This adapter conforms to SCSI-2 standards and the Fast-20 (Ultra) documentation. Industry standard SCSI P (68-pin) connectors are incorporated on the adapter.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master adapter.

- Supports attachment of external 8-bit or 16-bit SCSI devices on the J2 port using a 68 pin SCSI-3 standard connector.

Customer components

You must supply the following component for this feature:

- Cabling (available through IBM)

Adapter limitations:

Data transfer rates with F/C 6207 are limited to the speed of the slowest device on the SCSI bus.

8.2.32 SCSI-2 F/W Single-Ended adapter (F/C 6208)

The PCI SCSI-2 Fast/Wide Single Ended adapter (F/C 6208) provides a single ended SCSI-2 Fast/Wide interface that can burst data between devices on the SCSI bus at 20 MBps. It conforms to SCSI-2 standards and supports Fast/Wide synchronous data rates of up to 10 MHz. F/C 6208 supports both internal and external devices connected to the same SCSI bus.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master adapter.
- Supports attachment of internal and external single ended 8-bit and 16-bit SCSI devices.
- External connections on J2 with 68 pin SCSI-3 standard P connector.
- Internal connections on J3 with 68 pin high density SCSI connector for 16-bit attachments and on J4 with 50 pin SCSI connector for 8-bit attachments.
- The J3 and J4 connectors cannot be used at the same time.

Customer components

You must supply the following components for this feature:

- If you are using F/C 6208 to connect an internal DASD in a 332 MHz SMP wide node, you must also order F/C 1241.

8.2.33 SCSI-2 F/W Differential adapter (F/C 6209)

The PCI SCSI-2 Fast/Wide Differential adapter (F/C 6209) provides a differential SCSI-2 Fast/Wide interface that can burst data between devices on the SCSI bus at 20 MBps. It conforms to SCSI-2 standards and supports Fast/Wide synchronous data rates of up to 10 MHz. F/C 6209 supports external devices connected to the same SCSI bus.

Feature characteristics

This feature has the following characteristics:

- 32-bit Bus Master adapter
- Supports attachment of external 16-bit SCSI devices on the J2 port using a 68 pin SCSI-3 standard P connector

Customer components

No customer components required.

8.2.34 SSA RAID 5 adapter (F/C 6215)

The PCI SSA RAID 5 adapter (F/C 6215) supports RAID 5 SSA disk arrays and can be used to access non-RAID disks between multiple hosts. It has the capability to improve the write response time in the single initiator mode for both RAID and non-RAID disks by the addition of the Fast-Write Cache Option (F/C 6222). For more details on the F/C 6222 option, refer to 8.2.35, “SSA Fast-Write Cache Module (F/C 6222)” on page 255.

Feature characteristics

This feature has the following characteristics and requirements:

- 32-bit PCI bus.
- Support for floating hot spares on the same loop.
- RAID 5 arrays from (2+P) up to (15+P).
- Up to 6 (15+P) or 32 (2+P) RAID 5 Array Groups per adapter.
- All members of a RAID 5 array must be on the same SSA loop.

8.2.35 SSA Fast-Write Cache Module (F/C 6222)

The SSA Fast-Write Cache Module (F/C 6222) is an optional 4 MB fast-write module that plugs into the PCI SSA RAID 5 adapter (F/C 6215). The F/C 6222 cache option uses non-volatile RAM, which has over seven years of memory retention. Non-volatile memory allows you to transfer the cache module from a failing Multi-Initiator adapter to a new adapter during the unlikely event of an adapter failure. This helps insure data integrity and operational reliability.

Feature characteristics

This adapter features the following characteristics and requirements:

- Only one F/C 6222 is supported on each PCI SSA RAID 5 adapter (F/C 6215).

8.2.36 Advanced SerialRAID Adapter (F/C 6225)

The Advanced SerialRAID Adapter (F/C 6225) is a 4-port (2 loop) Serial Storage Architecture (SSA) adapter providing an instantaneous data transfer rate of up to 160 MB per second per loop. It also provides eight-initiator, non-RAID capability, or two-initiator RAID 5 capability. The adapter accepts a 32 MB SSA Fast-Write Cache Option Card (F/C 6235) that can improve write performance in both RAID 5 and non-RAID one-initiator applications. The Advanced SerialRAID Adapter when operated in a RAID 5 configuration will support (2+P) to (15+P) arrays and up to 6 (15+P) arrays. The adapter also supports Hot Spares in RAID 5 mode. The Advanced SerialRAID Adapter also supports connectivity to external disk enclosures and internal RS/6000 SSA configurations. Optional SSA Fiber-Optic Extender is also supported (Refer to IBM 7133 Sales literature for additional information).

Any supported RS/6000 system can be set up to boot from an Advanced SerialRAID Adapter (F/C 6225), provided a non-RAID SSA disk is included as part of the configuration. Other disks associated with the adapter can be RAID but at least one disk must be a non-RAID SSA disk.

8.2.37 Gigabit Fibre Channel PCI adapter (F/C 6227)

The Gigabit Fibre Channel PCI adapter (F/C 6227) provides single initiator capability through an optical fiber link running up to 100MB/s. With the use of optical fiber cabling, this adapter provides the capability for a network of high-speed local and remote located storage. F/C 6227 is a 32-bit PCI to NL-Port Fibre Channel node port connected to an arbitrated loop) host bus adapter. The

single speed supported is 1.0625 Gbps (wire-speed) which corresponds to approximately 100 MBps. The supported media includes 50 and 62.5 micron multimode fiber. The supported transmitter is short-wave laser with open fiber control.

Feature characteristics

This feature has the following characteristics:

- Single Fibre Channel loop per adapter
- Single initiatior support (one adapter for PCI bus per loop)
- Optical fiber link cabling supporting a distance of up to 500 m (1640 ft.)
- Fiber optic interface data rates of up to 100 MBps
- PCI 2.1 compatible, universal 3.3/5 volt adapter

Feature components

This feature order provides the following:

- Adapter card
- CD-ROM with adapter device driver
- Installation instructions

Customer supplied components

You must supply the following components for this feature:

Cables:

- Multimode 50/125 micron fiber with SC connectors:
  - 1062.5 MBps
  - Two 500 m cables
- Multimode 62.5/125 micron fiber with SC connectors:
  - 1062.5 MBps
  - Two 175 m cables

8.2.38 Advanced SerialRAID Plus PCI adapter (F/C 6230)

The Advanced SerialRAID Plus adapter (F/C 6230) is a functional enhancement to the withdrawn Advanced SerialRAID Adapter (F/C 6225)and supports 2-initiator FWC and 2-way RAID 10. The Advanced SerialRAID Adapter (F/C 6225), which has been withdrawn from marketing, has a data transfer rate of up to 160 MBps per loop. This high-performance multi-initiator Serial Storage Architecture (SSA) adapter provides eight-initiator non-RAID capability, two-initiator RAID-5 capability, and one-initiator RAID-0 capability.

The adapter utilizes the SSA Enhanced Loop (EL) architecture, which offers a choice for SSA HA data protection, disk mirroring for the best performance, or multi-initiator RAID for the lowest total system cost. SSA EL adapter architecture enables RS/6000 SP PCI systems to share SSA storage in a multi-host SSA environment (cluster/SP).

The Advanced SerialRAID Adapter with up to 160 MBps data transfer rate per loop and optional 32 MB Fast-Write Cache increases the RS/6000 SP storage performance in single-initiator and multi-initiator/multi-host environments. Boot support is only via the AIX Network Install Manager (NIM). The adapter accepts a 32 MB Fast-Write Cache Option Card (F/C 6235) in either a one-initiator RAID or one-initiator non-RAID application.

Feature characteristics

This feature has the following characteristics:

- Up to 160 MBps data transfer rates per loop
- Support for the following options:
  - 8-way JBOD
  - 2-way RAID 5
  - 2-way RAID 10
  - 1-way RAID 0
  - 2-initiator Fast Write Cache with fail over (F/C 6235)
  - 128 MB DIMM (F/C 6231) available for full capacity FWC with two initiatiors
  - TMSSA and fencing (used primarily by HACMP)
  - Two adapters in same host on same loop - automatic adapter takeover on
  - failure
  - Hot spares on same loop - automatic rebuild on disk failure (RAID 5 and RAID 10 configurations)
- Interoperates with PCI SSA Multi-Initiator/RAID EL Adapter (F/C 6215), MCA SSA
- Multi-Initiator/RAID EL Adapter (F/C 6219), and Advanced SerialRAID Adapter (F/C 6225) in JBOD and RAID 5
- PCI 2.1 compatible, universal 3.3/5 volt adapter

Native and NIM boot/install from JBOD disks:

  - Boot after NIM AIX install on AIX 4.2.1 and 4.3.2/n
  - Boot after Native AIX install on AIX 4.3.n

- Supports 48 drives per loop, 96 drives per adapter
- Operates at 40 MBps on SSA link when attached to another 40 MBps capable port

Feature components

This feature order provides the following:

- Adapter card
- CD-ROM with adapter device driver
- installation instructions

Customer supplied components

No customer components required.

8.2.39 ARTIC960RxD Quad Digital Trunk adapter (F/C 6310)

The ARTIC960RxD Quad Digital Trunk adapter (F/C 6310) provides voice processing for up to four T1 or E1 digital trunk lines, providing connectivity for 96 (T1) or 120 (E1) voice channels in a single PCI slot. The voice processing function is provided by Direct Talk for AIX, Version 2.1 LPP. The adapter provides high-function control of I/O operations and serves to off-load I/O tasks from the system microprocessor.

Feature characteristics

This feature has the following characteristics:

- 32-bit PCI 2.1 adapter
- One 36-pin, high-density port
- Support for up to four T1 or E1 trunk lines
- Supports voice processing using DirectTalk for AIX

Feature components

This feature order provides the following:

- One ARTIC960RxD adapter (F/C 6310).
- A connecting cable (required). The following are available from IBM:
  - F/C 2709: ARTIC960Hx 4-port T1 RJ45 cable
  - F/C 2710: ARTIC960Hx 4-port E1 RJ45 cable
  - F/C 2871: ARTIC960RxD Quad DTA, T1, 100 ohm, 3 m 4-port cable
  - F/C 2872: ARTIC960RxD Quad DTA, T1, 100 ohm, 15 m extension cable
  - F/C 2873: ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 3 m 4-port cable
  - F/C 2874: ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 7.5 m extension cable
  - F/C 2875: ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-grounded, 1.8 m 4-port cable

- F/C 2876: ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-ungrounded, 1.8 m 4-port cable
- F/C 2877: ARTIC960RxD Quad DTA, H.100, 4-drop cable

8.2.40 ARTIC960RxF Digital Trunk Resource adapter (F/C 6311)

The Artic960RxF Digital Trunk Resource adapter provides H.100/SCBus support for TDM-based functions such as conferences, DSP speech recognition and text-to-speech plus one-call FAX. It is intended for use with the Corepoint DT/6000 Voice Recognition product. It has no external connections.

Feature characteristics

This feature has the following characteristics:

- Must be used in conjunction with DTXA.
- Intel 80960 RD processor with 4 MB DRAM.
- Universal PCI adapter.
- Replaces Dialogic Antaries ISA adapter.
- 32-bit PCI interface.
- A PLX 9080 local bus provides access to the SC4000s and DSPs from the i960.
- Two TI C67x floating point DSPs; each with 512 KB external SBSRAM, 4MB of external SDRAM, operating at a clock cycle of 167 MHz.
- Two SC4000 chips provide H.100/SCBus support for TDM-based functions such as conferencing, DSP speech recognition, and text-to-speech plus one-call FAX.
- A PMC connector to allow attachment of an optional daughter card.

Feature components

This feature order provides the following:

- One ARTIC960RxF adapter (F/C 6311)
- ARTIC RTE CD kit
- Diagnostics disk
- A connecting cable (required); the following cables are available from IBM:
  - (F/C 2877): ARTIC960RxD Quad DTA, H.100, 4-drop cable
  - (F/C 4353): ARTIC960RxD Quad DTA, H.100, 8-drop cable

9

Software support

This chapter describes the software available for SP systems. It covers the following software:

- Parallel System Support Programs (PSSP)
- General Parallel File System (GPFS)
- LoadLeveler
- Scientific Subroutine Library (ESSL, PESSL & OSL)
- Parallel Environment (PE)
- IBM High Availability Cluster Multi-Processing for AIX (HACMP)

9.1 Parallel System Support Programs (PSSP)

Program number- 5765 - D51

The PSSP software is a comprehensive suite of applications to manage an SP systems as a full-function parallel processing system. It provides a single point of control for administrative tasks and helps increase productivity by letting administrators view, monitor, and control how the system operates. The PSSP software is discussed in terms of functional entities called components of PSSP. Most functions are base components of PSSP while others are optional; they come with PSSP, but you choose whether to install and use them.

PSSP offers the following advantages:

- A full suite of system management applications with the unique functions required to manage the RS/6000 SP system.
- Simplified installation, operation, and maintenance of all nodes in an RS/6000 SP system. You can operate from a single control workstation.
- Parallel system management tools for allocating SP resources across the enterprise provided.
- Advanced performance monitoring for consolidated analysis and reporting.
- Error detection and recovery features that reduce the impact and occurrence of unplanned outages.
- Coexistence is allowed for several releases within an SP partition allowing for easier software migration.

9.1.1 Administration and operation

The PSSP system administration component packages make operating and administering an SP system easy and efficient. This is accomplished through tools and capabilities that PSSP offers.

Administrative features of PSSP:

- PSSP allows the system administrator/operator to perform all local and remote administrative functions from the control workstation (CWS). This makes the CWS a single point of control.
- It offers parallel system management functions across multiple SP nodes.
- PSSP contains the System Data Repository (SDR) for storing management data that can be retrieved across the control workstation, file servers, and SP nodes.
- It includes File collections for managing duplicated files and directories on multiple nodes.

- PSSP comes with login control for blocking unauthorized user or group access to a specific SP node or a set of nodes.
- It offers consolidated accounting for centralizing records at the node level (for tracking use by wall clock time rather than processor time) and for gathering statistics on parallel jobs.

Administrative tools of PSSP:

SP Task Guides

Online assistance designed to walk you through complex or infrequently performed tasks. Task Guides do not simply list the required steps; each Guide actually performs the steps for you, automating the steps to the highest degree possible and prompting you for input only when absolutely necessary. You might recognize them as *wizard*s.

SP Perspectives

A set of graphical user interfaces (GUI's) which can simplify administrative work.

Centralized Management Interface

Menu-based interface for system management commands, as well as command line equivalents.

9.1.2 RS/6000 cluster technology (RSCT)

RS/6000 cluster technology gives the SP systems high availability, online monitoring and automatic recovery actions on clustered nodes. It does this by the use of three distributed subsystems.

RSCT's subsystems:

Topology Services Maintains availability information regarding nodes and network adapters. Therefore seamless takeover of functions in the event of a node failure is possible.

Group Services Provides services for synchronization and coordination of recovery actions of other subsystems.

Event Management Monitors hardware and software resources in the SP and notifies an interested application or subsystem when the state of the resource changes. Resources on any node in the SP can be monitored from any other node.

9.1.3 System monitoring and control

PSSP provides system management tools that enable the system administrator to manage the SP system. PSSP allows authorized users to monitor and manipulate SP hardware variables at node or frame levels. It also enables the system administrator to gracefully shut down rather than rebooting the nodes or the complete system. In addition, it provides for consolidation of error and status logs for expediting problem determination.

9.1.4 IBM Virtual Shared Disk (VSD)

The IBM Virtual Shared Disk (IBM VSD) allows multiple nodes to access a disk as if the disk were attached locally to each node. This feature can enhance the performance of applications that provide concurrent control for data integrity, such as Oracle databases or the GPFS file system.

IBM Recoverable Virtual Shared Disk (RVSD)

The IBM Recoverable Virtual Shared Disk (IBM RVSD) allows transparent recovery of IBM VSD. The IBM RVSD function provides recovery from failures of IBM VSD server nodes and takes advantage of the availability services provided by PSSP to determine which nodes are up and operational.

Hashed Shared Disk

This component that works with the IBM Virtual Shared Disk component to offer data striping for your virtual shared disks.

9.1.5 Security

PSSP 3.2 offers three security settings, which can be used in various combinations: You can use all three of them at the same time, or you can use a combination of any two, or you could choose just one setting to use.

Security configurations:

- Standard authentication
- Kerberos Version 4
- Distributed Computing Environment (DCE)

9.1.6 Performance Toolbox Parallel Extensions for AIX (PTPE)

As its name implies, PTPE extends the capabilities of Performance Toolbox (PTX) or AIX by adding monitoring functions and statistics specific to the SP. PTPE provides SP performance statistics from software, such as LoadLeveler and IBM Virtual Shared Disk (VSD), and hardware components. These statistics are available in live (current) and archived (historical) forms for use in supporting both active monitoring and comparative analysis over time.

9.1.7 Subsystem communications support

The Communication Subsystems Support component contains Switch software support. It contains adapter diagnostics, switch initialization and fault-handling software, device driver and configuration methods, plus parallel communications application programing interfaces (APIs).

Each node can be:

- A communication path through the switch to other nodes.
- A monitor of the switch hardware.
- In control of the switch hardware for startup and, in case of error, execution of any appropriate recovery action.

9.1.8 Network Time Protocol (NTP)

Handles synchronizing time-of-day clocks on your control workstation and processor nodes. In PSSP 3.2, the public domain NTP 3.3 has been replaced with AIX NTP 3.4.

9.1.9 System availability

To significantly improve system availability, PSSP also contains functions and interfaces to other products that can help reduce unplanned outages and minimize the impact of outages that do occur. These include:

- System partitioning, which makes it possible to create a separate logical system partition. It allows testing of different levels and PTFs on the operating system, system software, or hardware. It also allows different production environments to be concurrently executed for workload isolation.
- Coexistence can reduce scheduled maintenance time. PSSP can coexist with several releases within an SP partition, thus, allowing easier migration to new software levels.
- Node isolation, which removes an SP node from active duty and enables it to be reintegrated without causing an SP Switch fault or disrupting switch traffic.

This isolation is useful for correcting an error condition or installing new hardware and software without impacting production.

- High Availability Control Workstation (HACWS) connects two RS/6000 workstations (with HACMP installed) to an SP system to provide a backup control workstation in the event the primary one becomes unavailable (only one is active at any time). A twin-tailed disk configuration, along with the IP address takeover afforded by HACMP, enables rapid switch over to the backup control workstation with little or no impact on operational access to the SP system or System Data Repository.
- The SP Resource Center provides one simple interface for all soft copy SP documentation and information resources. It consists of HTML, Java, and Java script files and works with a Web browser. The Resource Center provides access to a variety of information including publications, READMEs, Redbooks, white papers, product information, as well as up-to-date service information.
- PSSP supports a multi-threaded, standards-compliant Message Passing Interface (MPI) through an IBM Parallel Environment for AIX (PE) as well as maintaining its single-threaded MPI support. In addition, PSSP includes a Low-level Application Programming Interface (LAPI) with a flexible, active message style, communications programming model on the SP Switch.

9.1.10 Other convenient services

PSSP offers others convenient services:

- PSSP offers Perl programming language for developing system-wide shell scripts.
- PSSP comes with Tool command language (Tcl) for controlling and extending applications.

9.1.11 New in PSSP 3.2

PSSP 3.2 was announced to the public on July 28, 2000 and includes the following enhancements:

- The PSSP software is NLS-enabled. This means that the software has been made translation-ready and can be translated to any language supported by AIX.
- In PSSP 3.2, the public domain NTP 3.3 has been replaced with AIX NTP 3.4.
- SP and cluster software serviceability, the First Failure Data Capture (FFDC) facility is introduced in this version of PSSP. Among other things it provides interfaces to SP and cluster internal software components and applications to record sufficient information about failures.

9.1.12 Software requirements

Table 9-1 shows the software requirements for PSSP 3.2.

Table 9-1 PSSP 3.2 software requirements

AIX version	other software
4.3.3	► C for AIX, V4.4, or later. ► C for AIX, V5.0, or later. ► VisualAge C++ Professional for AIX V4.0

9.2 General Parallel File System (GPFS)

Program number - 5765 -B95

IBM General Parallel File System (GPFS) is a standards-based, parallel file system that delivers high performance, high availability, and high scalability while preserving the application interfaces used in standard file systems. GPFS allows access to files within an SP system from any GPFS node in the system and can be exploited by parallel jobs running on multiple SP nodes as well as by serial applications that are scheduled to nodes based on processor availability.

Advantages:

- GPFS is scalable therefore incremental improvements may be made to the file system by adding additional hardware of the same, or even lesser, capability. GPFS file systems may be multiple Terabytes in capacity and provide throughput of multiple GB/sec to and from one file. It was designed expressly to deliver scalable performance across multiple file system nodes.
- It is general-purpose file system in that it is suitable for many kinds of workloads including commercial and technical tasks.
- It provides global access (or uniform access) to files. That is, it is possible to mount a GPFS file system from every node on an SP system.
- GPFS is a portable solution based on the POSIX standard; no GPFS-specific system calls are needed or provided. This means that applications using standard POSIX functions will run on GPFS without modification or re-compilation.
- GPFS is easy to use. Since it is based on POSIX standards and POSIX is well known, initial use is easier for most developers.

9.2.1 Administration and operation

GPFS provides functions that simplify multinode administration and can be performed from any node in the SP configuration. These functions are based on, and are in addition to, the AIX administrative commands that continue to operate. A single GPFS multinode command can perform a file system function across the entire SP system. In addition, most existing UNIX utilities will also run unchanged. All of these capabilities allow GPFS to be used as a replacement for existing UNIX file systems where parallel optimization is desired.

All file systems provide a number of data management services, such as the ability to duplicate files, remove files, rename files, and so forth. As a POSIX file system, GPFS appears the same as any other POSIX file system but with exceptional capacity and read (or write) performance. The user performs ordinary file system operations, such as copying a file, and so forth, on GPFS in the same manner as they would on other standard file system.

9.2.2 Higher performance/scalability

By delivering file performance across multiple nodes and disks, GPFS scales beyond single-server (node) performance limits. Unlike NFS and JFS, GPFS file performance scales as additional file server nodes and disks are added to the SP system.

In a parallel environment GPFS can outperforms:

- PIOFS
- Network File System (NFS)
- Distributed File System (DFS)
- Journalled File System (JFS

9.2.3 Recoverability

GPFS can survive many system and I/O failures. Through its use of the RS/6000 Cluster Technology capabilities of PSSP in combination with IBM RVSD, GPFS is able to automatically recover from node, disk connection and disk adapter failures. GPFS will transparently failover locked servers and other GPFS central services. Through its use of IBM RVSD, GPFS continues to operate in the event of disk connection failures. GPFS allows data replication to further reduce the chances of losing data if storage media fail. Unlike PIOFS, GPFS is a logging file system that allows the recreation of consistent structures for quicker recovery after node failures. GPFS also provides the capability to mount multiple file systems, each of which can have its own recovery scope in the event of component failures.

9.2.4 Migration

Upgrading to a new release of GPFS can be tested on a system currently running GPFS. This eases migration by allowing the testing of a new level of code without inhibiting the production GPFS application.

9.2.5 New in GPFS 1.4

GPFS 1.4 was announced to the public on December 22, 2000 and includes the following enhancements:

- It supports concurrent data sharing in a cluster environment.
- Nodes may be added dynamically to a node set when using TCP/IP as the communications protocol.
- GPFS 1.4 supports a single-node quorum in a two-node set.
- It gives you the capability to operate in an SP environment with directly attached disks.

9.2.6 Software requirements

Table 9-2 shows the software requirements for GPFS 1.4.

Table 9-2 Software requirements for GPFS 1.4

AIX version	PSSP version	other software
4.3.3 or higher	3.2 or higher	► IBM Virtual Shared Disk image ► Recoverable Virtual Shared Disk image

9.3 LoadLeveler

Program number - 5765-D61

IBM LoadLeveler is a job management system that allows users to run more jobs in less time by matching the jobs' processing needs with the available resources. LoadLeveler schedules jobs, and provides functions for building, submitting, and processing jobs quickly and efficiently in a dynamic environment. It balances your workload, efficiently manages job flow in the network, and distributes work across all LoadLeveler-defined hardware resources. LoadLeveler is an application that runs as a set of daemons on each node in the network.

Advantages

- It offers a a single point of control for administration, job management and workload scheduling.
- LoadLeveler give users full scalability across processors and jobs.
- It comes with APIs to enable alternate scheduling algorithms.
- It supports thousands of jobs across hundreds of SP nodes and IBM workstations.
- It provides automatic recovery of central scheduler and can be configured with HACMP for node and network failover.

9.3.1 Administration

From a system management perspective, LoadLeveler allows a system administrator to control all jobs running in a cluster including the SP system. Thus, providing administrators with the information needed to make adjustments to job classes and changes to LoadLeveler-controlled resources.

LoadLeveler has a command line interface and a Motif-based graphical user interface (GUI) making it easy to submit and cancel jobs, monitor job status, and set and change job priorities. The system can be configured at the network and node levels, where workstations or SP nodes may be identified as job submitters, compute servers, or both.

LoadLeveler offers the option of using its own scheduler or an alternative scheduler, such as EASY from the Cornell Theory Center. The product also provides a user- or system-initiated checkpoint/restart capability for certain types of FORTRAN, C, or C++ jobs linked to the LoadLeveler libraries.

9.3.2 Capabilities

Parallel processing

For parallel jobs, LoadLeveler interfaces with parallel programming software to obtain the multiple SP nodes required for the job's parallel tasks.

Individual control

At the node level, users can specify to LoadLeveler when their processing nodes are available and how they are to be used. For example, some users might let their workstations accept any job during the night, but only certain jobs during the day when they most need their resources. Other users might simply tell LoadLeveler to monitor their keyboard activity and make their workstations available whenever they have been idle for a sufficient time.

Central control

LoadLeveler gives administrators a view of all jobs running on the system and the resources available to those jobs. Administrators can change the availability of theses resources to best fit their needs.

Scalability

As nodes are added, LoadLeveler automatically scales upward so that the additional resources are transparent to the user.

Automated job tracking

By tracking processes spawned off by a job and cancelling any processes left behind when a job is terminated, LoadLeveler ensures that resources assigned to cancelled jobs become available for future use.

Multiple user space tasks

LoadLeveler supports multiple user space tasks on both SP Switch and SP Switch2. With the use of Parallel Environment V3.1 and the new SP Switch2, LoadLeveler V2.2 can handle up to 4096 user-space tasks.

Additional scheduling algorithm

Includes the Backfill scheduler, which runs both serial and parallel jobs.

National language support (NLS)

LoadLeveler is enabled for NLS. Error messages are externalized in both message catalogues and the LoadLeveler Diagnosis and Messages Guide.

9.3.3 New in LoadLeveler 2.2

LoadLeveler 2.2 was announced to the public on July 18, 2000 and includes the following enhancements:

- Improved security: By exploiting the PSSP Security Services within the Distributed Computing Environment (DCE), administrators can authenticate user identity and delegate credentials, so only authorized users can schedule jobs on specified resources.
- Better resource control: Version 2.2 incorporates a process tracking function that locates processes leftover from terminated jobs. LoadLeveler can then cancel them and free up valuable resources for other tasks.
- More task assignment options: More job command file options allow various task geometry for a parallel job.
- Improved performance: LoadLeveler daemons are now running multi threaded providing improved command response time.

9.3.4 Software requirement

Table 9-3 shows the software requirements for LoadLeveler 2.2.

Table 9-3 Software requirements for LoadLeveler 2.2

AIX version	PSSP version	other software
4.3.2 or higher	3.1 or higher	► If you plan to run POE jobs in user space mode, sp.st 3.1 or later, compatible releases are required. ► If you plan to configure LoadLeveler to support DCE security services: fileset ssp.clients 3.2 is required.

9.4 Mathematical Subroutine Libraries (PESSL)

There are three types of Mathematical Subroutine Libraries; they include:

- Program number for ESSL - 5765-C42

- Program number for PESSL - 5765-C41
- Program number for OSLp - 5765-392

The Engineering and Scientific Subroutine Library (ESSL) family of products is a state-of-the-art collection of mathematical subroutines that provide a wide range of high-performance mathematical functions for many different scientific and engineering applications.

Parallel Engineering and Scientific Subroutine Library (PESSL) is built on top of ESSL and is specifically tuned to exploit the full power of the SP hardware with scalability across the range of system configurations.

ESSL and PESSL can be used to develop and enable many different types of scientific and engineering applications. New applications can be designed to take advantage of all the capabilities of the ESSL family. Existing applications can be easily enabled by replacing comparable routines and in-line code with calls to ESSL subroutines.

Parallel Optimization Subroutine Library (OSLp) is a collection of 60 high performance mathematical subroutines used by application programmers to solve large optimization problems. It includes all of the functions of the AIX OSL/6000 product but solves linear and mixed-integer programming problems in parallel on the SP processor achieving a significant performance improvement.

OSLp is a set of over 60 subroutines callable from your application program in order to find the optimal solution to several types of problems using linear programming (LP), mixed-integer programming (MIP), and quadratic programming (QP) mathematical techniques. Some of the solutions use serial algorithms; that is, all computations are performed in sequence on an RS/6000 or single node of an SP system.

Advantages of ESSL and PESSL:

- They are callable from XL FORTRAN, C, and C++ applications.
- Both are designed for high mathematical computational performance.
- They support many scientific and engineering applications used by multiple industries.
- They are tuned to the characteristics of SP hardware.
- ESSL and PESSL can be used with existing programs via relinking rather than recompiling.
- They support easy development of parallel applications with Single Program Multiple Data model and/or Shared Memory Parallel Processing model for SMPs.

9.4.1 Operations

ESSL's mathematical functions:

- Basic Linear Algebra Subroutines (BLAS)
- Linear Algebraic Equations
- Eigensystem Analysis
- Fourier Transforms

The ESSL products are compatible with public domain subroutine libraries, such as Basic Linear Algebra Subprograms (BLAS), Scalable Linear Algebra Package (ScaLAPACK), and Parallel Basic Linear Algebra Subprograms (PBLAS), Basic Linear Algebra Communication subprograms (BLACS), and MPL/MPI making it easy to migrate applications that utilize these libraries to use ESSL and/or Parallel ESSL.

OSL capabilities include:

- Linear programming
- Network programming
- Mixed integer programming
- Quadratic programming

Benchmarks have shown that Parallel OSL achieves significant speed increases on LP and MIP problems with its parallelized solution algorithms. LPs solved with the interior point algorithm are typically able to achieve a sub-linear speed increases of 40 to 50 percent processor utilization. That is, running such a problem on an SP with eight processors is three times as fast as running OSL/6000 on a single node.

Examples of applications that use these types of mathematical subroutines are:

- Structural analysis
- Time series analysis
- Computational chemistry
- Computational techniques
- Fluid dynamics analysis
- Mathematical analysis
- Seismic analysis
- Dynamic systems simulation
- Reservoir modeling

- Nuclear engineering
- Quantitative analysis
- Electronic circuit design

For more detailed information on ESSL, PESSL and OSLp refer to *RS/6000 Scalable POWER parallel System: Scientific and Technical Computing Overview*, SG24-4541-01.

9.4.2 New in ESSL 3.2

The ESSL Libraries are tuned for the RS/6000 POWER3-II processors. New enhancements have been made to ESSL's Dense Linear Algebraic Equations Subroutines, which now include:

- Symmetric Indefinite Matrix Factorization and Multiple Right-Hand Side Solve
- Symmetric Indefinite Matrix Factorization
- Symmetric Indefinite Matrix Multiple Right-Hand Side Solve

The Linear Least Squares subroutines now include:

- General Matrix QR Factorization

9.4.3 New in Parallel ESSL 2.2

The Parallel ESSL Libraries are tuned for the RS/6000 POWER3-II processors and the SP Switch and SP Switch2.

The Dense Linear Algebraic Equations Subroutines now include:

- Inverse of a real or complex general matrix.
- Reciprocal of the condition number of a real or complex general matrix.
- General matrix QR factorization.
- Least Squares solutions to linear systems of equations for real general matrices.

The Eigensystems Analysis Subroutines now include:

- Selected Eigenvalues and optionally the Eigenvectors of a complex Hermitian matrix
- Selected Eigenvalues and optionally the Eigenvectors of a real symmetric positive definite generalized Eigenproblem
- Reduce a complex Hermitian matrix to tridiagonal form

- Reduce a real symmetric positive definite generalized Eigenproblem to standard form

9.4.4 New in OSLp 1.3

The following are new modules and routines found in OSLp 1.3:

- A pair of new routines provided for transforming a general integer problem into an equivalent binary integer problem, and transforming the solution of the binary integer problem back into the solution of the original general integer problem.
- A new module provided for identifying relationships between the variables of a mixed integer programming problems.
- A new module used to analyze infeasibility in an linear programming problems.
- This new routine enables the user to custom order the LP matrix before symbolic factorization in interior point methods to reduce fill-in of the Cholesky factor.

9.4.5 Software requirement

Table 9-4 shows the software requirements for ESSL 3.2, Parallel ESSL 2.2 and OSLp 1.3

Table 9-4 Requirements for ESSL 3.2, Parallel ESSL 2.2 and OSLp 1.3

Libraries	AIX version	PSSP version	other software
ESSL 3.2	4.3.3.10 or higher		► XL FORTRAN for AIX, Version 7.1 ► IBM VisualAge C++Professional for AIX Version 5.0 ► C for AIX, Version 5.0.1 ► PL/1 Set for AIX, Version 1.1 or higher ► C libraries (Included in the AIX Version 4 Application Development Toolkit.)

Libraries	AIX version	PSSP version	other software
Parallel ESSL 2.2	4.3.3.10 or higher	3.2 or higher	► XL FORTRAN for AIX, Version 7.1 ► IBM VisualAge C++ Professional for AIX Version 5.0 ► C for AIX, Version 5.0.1 ► ESSL V3.2 ► Parallel Environment for AIX V3.1 ► C libraries (Included in the AIX Version 4 Application Development Toolkit.)
OSL 1.3	4.1 or higher		► AIX XL FORTRAN Run-Time Environment/6000 Version 3.1 ► XL C Compiler operating system 1.3 included with the AIX ► AIX XL C++/6000 One of the following must be present: ► IBM Parallel Environments (PE) ► IBM PVMe (PVMe) ► Parallel Virtual Machine (PVM)

9.5 Parallel Environment (PE)

Program number 5765-D93

IBM Parallel Environment for AIX is a high-function development and execution environment for parallel applications using the SP system. It is a complete solution for enterprises that need to develop, debug, analyze, tune, and execute parallel programs on the AIX platform.

The advantages of Parallel Environment include:

- It exploits threads and thread-safe MPI message passing on all nodes, including symmetric multiprocessors (SMPs).
- It supports Low-level Application Programming Interface (LAPI) programs.
- PE makes for easier parallel application development.
- PE gives enhanced XProfiler graphical performance tool.
- It enables easy application portability to networked RS/6000 or RS/6000 SP systems.

9.5.1 Operation

PE application development solution consists of:

1. Parallel Message Passing APIs for full implementation of the MPI 1.2 standard plus full implementation of the MPI-I/O and MPI-1 sided communication chapters of the MPI-2 standards. Also, continued support for the IBM Message Passing Library for communications between executing tasks in a Fortran, C, or C++ parallel program.
2. A Parallel Operating Environment (POE) for managing the development and execution of parallel applications.
3. Parallel debuggers offering both command-line and Motif-based interfaces. These debuggers extend traditional AIX capabilities and provide features for parallel application task debugging.
4. Xprofiler graphical performance tool.
5. Dynamic Probe Class Library (DPCL) parallel tools development API.

9.5.2 New in PE 3.1

PE version 3.1 was announced on July 18, 2000 and includes the following enhancements:

- New application program interfaces (API).
- FORTRAN 95 compiler scripts.
- Support for Distributed Computing Environment (DCE) security through the PSSP Security Services.
- Support for MPI one-sided communication allowing one process to specify all communication parameters for the sending operation as well as the receiving operation.

- Support for MPI shared memory message passing on symmetric multiprocessor (SMP) nodes, for the Internet Protocol (IP) library and for the User Space (US) library.
- Instrumentation probes that the analysis tool defines and inserts into the target application to collect data.
- An asynchronous communication callback facility that connects the class library with the daemons. It is this callback facility that enables an analysis tool to respond to data collected and sent by its probe.

9.5.3 Software requirements

Table 9-5 shows the software requirements of PE 3.1

Table 9-5 Software requirements for PE 3.1

AIX version	PSSP version	other software
4.3.3.10 or higher	3.2 or higher	► XL FORTRAN, V7, or higher, if FORTRAN programs are to be run ► If User Space batch or interactive jobs are to be submitted under Parallel Environment, then LoadLeveler V2.2 is required. At least one concurrent use license of C for AIX compiler or C++ compiler currently available versions include: ► C for AIX, V4.4, or higher; or C for AIX V5.1, or higher. ► VisualAge C++ Professional for AIX V4.1 ► VisualAge C++ Professional for AIX V5.0, or higher.

9.6 IBM High Availability Cluster Multi-Processing for AIX (HACMP)

Program number - 5765-E54

HACMP aims to eliminate any single point of failure, be it disks, network adapters, power supply, processors, or other components. HACMP provides a highly-available environment by identifying the set of cluster resources that are essential to processing and defining the takeover relationships among the cluster nodes that access these resources.

Detecting the failure of a system or an adapter is achieved by sending heart beats (small messages that indicate the sender is functioning correctly) between systems. Loss of heart beats for an extended period of time can be reasonably interpreted as meaning that the sending system has failed.

HACMP offers the following advantages:

- It minimize down time by quickly restoring essential services.
- It provides system-wide, shared resources that cooperate to provide essential services.
- HACMP makes for ease of use with task guides.
- It gives extensive capability to customize to individual requirements.
- It supports emulation of events for AIX error notification.
- It provides monitoring of applications.

9.6.1 Administration and operation

The HACMP Administration Guide, SC23-4279-02, documents the common administration tasks and tools that exist to help carry out these tasks. Management and maintenance of the cluster, once it is up and running, is one of the aspect of the HACMP software that has been significantly developed to make it easier to set up and manage a complex HA environment.

To be made highly-available by HACMP, each resource must be included in a resource group. Resource groups allow related resources to be combined into a single logical entity for easier configuration and management.

Cluster resources:

- Disks
- Volume groups
- File systems

- Network addresses
- Application servers

The takeover relationships among cluster nodes determine which cluster nodes control a resource group and which cluster nodes take over control of the resource group when the original node relinquishes control. Takeover relationships are defined by assigning on of the following

HACMP takeover relationships:

Cascading	The resource is always owned by the active node having the highest priority for that resource.
Rotating	The resource rotates among all the nodes defined in the chain. The node with the highest priority in the resource chain for the resource takes over for a failed node.
Concurrent	The resource is owned and accessed simultaneously by all owning nodes.

9.6.2 Features

HACMP offers the following features:

- Recovery flexibility

  HACMP clusters can be configured to meet complicated application availability and recovery needs. HACMP configurations can include mutual takeover or standby takeover recovery processes. Mutual takeover configurations allow all servers to be active information processors and perform as a backup server for a failed server.

- Workload distribution

  With an HACMP mutual takeover configuration, applications and their workloads are assigned to specific servers. For example, in a two-node HACMP cluster, one server is designated to be the application server and the second server is designated to be a database server. Segmenting the application processing to one server and the database processing to the second server utilizes both servers in a productive manner, and maximizes application throughput and investments in hardware and software.

- Scalable growth

  With HACMP, two to thirty two servers may be configured for high availability. HACMP offers built- in growth and helps protect investments in cluster hardware, software, services, and training.

- Easy upgrade/migration

The cluster configuration can be expanded or modified without service disruption through Dynamic Reconfiguration options. Version Compatibility and Dynamic Reconfiguration eliminate the most common causes of scheduled cluster down time.

- Concurrent processing

  With the HACMP Concurrent Resource Manager (CRM) feature option, data can be processed concurrently when using Oracle Parallel Server (OPS) Version 7.3 or 8. The combination of CRM and OPS offers an additional level of high availability. With CRM and OPS, all servers in the cluster share the data and system resources; if one or more servers should fail, the remaining servers continue processing with minimal disruption, if any, since no fail over is required.

For more detailed information see *RS/6000 Systems Handbook*, SG24-5120 or *Exploiting HACMP 4.4: Enhancing the Capabilities of Multi-Processing*, SG24-5979.

9.6.3 New in HACMP 4.4

HACMP 4.4 was announced to the public on June 20, 2000 and includes the following enhancements:

- Application Monitoring (ES and ESCRM only) provides Process Application monitoring and User-defined Application monitoring to determine the state of an application and to restart the application or fall the resource group over to another node.
- Tivoli Cluster Monitoring allows users to monitor the state of an HACMP cluster and its components on a Tivoli Desktop window.
- Cascading without Fallback permits specifying that the resource group not return to the original node when that node rejoins the cluster.
- Cluster Verification is enhanced to detect additional startup of failover problems.
- New Performance Tuning Parameters provide easier and more granular control.
- New documentation provides guidance for setting up, monitoring, and managing 7 x 24 clusters.
- Enhanced LVM Task Guide now provides a display of the physical location of each available disk and will create automatically a JFS log file.
- NFS Migration (HAS only) provides for migration between HACMP 4.3.1 HANFS feature and HACMP 4.4 HAS feature. HANFS is no longer included as a feature of HACMP.

- HAS and CRM support clusters of up to eight systems.
- ES supports clusters of up to 32 SP nodes or RS/6000 systems.
- ESCRM supports clusters of up to 8 SP nodes of RS/6000 systems.

9.6.4 Software requirements

Table 9-6 on page 284shows the software requirements for HACMP 4.4

Table 9-6 Software requirements for HACMP 4.4

AIX version	PSSP version	other software
4.3.3.10 or later	3.2 or later	► HAView facility requires the installation of Tivoli NetView for AIX ► Data Path Optimizer (DPO) Version 1.1 is only supported in the CRM and ESCRM

10

Clustering configurations

In this chapter you find information on how to build your complex cluster. Section 10.1, “Configurations rules” on page 286 shows all the rules you need to know before you design your cluster system; frame, switch, and node numbering rules are discussed.

Section 10.2, “SP universal clustering capabilities” on page 304 shows different SP universal clustering capabilities. The sample scenarios provide some possible clustering scenarios from the traditional one frame cluster to a mix of all possible clusterable servers.

Section 10.3, “PSSP environment for SP/CES systems” on page 337 shows the PSSP enhancements made to the SDR; commands and sample monitoring definitions are provided.

Section 10.4, “Building SP/CES systems scenarios” on page 344 provides different scenarios for building your SP/CES system.

10.1 Configurations rules

The basic components that comprise an SP/CES system are nodes/servers, frames, and switches. The SP architecture can generally support up to 128 nodes (512 with special request). Within these 128 nodes, you can have any combination of thin/wide/high SP nodes (up to 128), SP-attached servers M/T 7017 supported Models (up to 16), and SP-attached servers M/T 7026 supported Models (up to 32).

The CES architecture can support up to 32 servers. Within these 32 servers, you can have any combination of CES M/T 7017 supported Models (up to 16), and CES M/T 7026 supported Models (up to 32).

The SP architecture can generally support up to 128 frames. The model frame (i.e., 550 or 500) is always the first frame in an SP system and it designates the type or model class of your SP system. The other frames that connect to the model frame are the expansion frames (i.e., F/C 1550 - Tall frame or F/C 1500 - Short frame) and they can be:

- Non-switched expansion frames (configured with SP nodes only)
- Switched expansion frames (configured with SP nodes and an SP Switch)

There are three types of SP switches:

- High Performance Switch (HiP Switch) - 8 ports (F/C 4007) and 16 ports (F/C 4010)
- Scalable POWER Parallel Switch (SP Switch) - 8 ports (F/C 4008) and 16 ports (F/C 4011)
- Scalable POWER Parallel Switch2 (SP Switch2) - 16 ports (F/C 4012)

Eight port Switches have limited scalability and flexibility that can support an eight-node configuration only. For 16 port switches, the system can scale up to 128 nodes, with switch frames support.

Short frames can be developed into non-switched and SP Switch-8 configurations only. Tall frames provide more flexibility than Short frames in configurations. At this time, there are four main types of Tall frame configurations:

- Non-switched configuration
- SP-Switch (eight ports) configuration
- SP Switch (16 ports) and SP Switch2 (16 ports) configuration
- SP Switch (16 ports) and SP Switch2 (16 ports) with Intermediate Switch Board (ISB) configuration

Any SP-attached server is considered a node (F/C 9122) and a frame (F/C 9123) in an SP system. Any Clustered Enterprise Server is considered a node and a frame in a CES system.

Any SP-attached server and any Clustered Enterprise Server can be developed into non-switched and SP Switch (F/C 4011) configurations only. The switched CES system requires a Standalone Switch for Clustered Nodes (M/T 9076 Model 555).

However, there are still some configuration restrictions or rules, on the combination of these components (nodes, frames and switches) that need to be considered.

10.1.1 Placement rules for nodes, frames, and switches

This section describes the placement rules for node, frames, and switches.

Placement rules for nodes

Node configuration rule 1:

Attention: PCI Thin/Wide nodes can be installed in model 550 Tall frames (and F/C 1550 Tall Expansion frames) or model 500 Short frames (and F/C 1500 Short Expansion frames). Old SP Tall/Short frames with the power supply upgrade are also supported.

For the PCI Thin/Wide nodes, it is required to install them in the 550 and 500 models and in the afferent expansion frames (F/C 1550 and F/C 1500). These nodes can also be hosted in the old SP frames with the frame power supply upgraded, as the PCI nodes require new power requirements. This configuration allows customers with old SP systems installed to add new PCI nodes with higher performance. A precaution is that if the HiP Switch is used in the old SP frame, you have to upgrade the HiP Switch to SP Switch in order to switch the new PCI nodes and the old nodes together. In addition, you need to upgrade the HiP Switch adapters to SP Switch adapters in the SP nodes.

The PCI Thin/Wide nodes are supported in switchless, SP Switch (F/C 4011) or SP-Switch-8 (F/C 4008) configurations only.

Node configuration rule 2:

Attention: PCI High nodes can be installed in model 550 Tall frames only. An Expansion I/O unit can not coexist with another PCI Thin node in a same drawer.

The PCI High nodes are supported on model 550 Tall frames and Expansion frames (F/C 1550) only. In addition, they are supported in switchless, SP-Switch (F/C 4011) or SP-Switch2 (F/C 4012) configurations only.

The Expansion I/O units for PCI high nodes (F/C 2055) occupy one slot in a node drawer and can be installed in a single unit. However, they can not be mixed with other PCI thin nodes in the same node drawer. We suggest that Expansion I/O units (F/C 2055) be mounted in separate frames, so as not to interfere with switch port utilization.

Node configuration rule 3:

Attention: Single PCI thin node must be installed in the odd slot position of a drawer (the left side of the drawer).

You are allowed to place only a single PCI thin node drawer, while with MCA thin nodes, you need to fill the drawer with two thin nodes. The condition is that it must be installed in the odd slot position (left side) of the drawer. This is because the lower slot number is what counts when a drawer is not fully populated. Moreover, different types of PCI thin nodes can be mixed in the same drawer, so one can install a POWER3 Thin node in the left side of a drawer and a 332 MHz Thin node in the right side of the same drawer. Additionally, for Short frames, a single PCI thin node (if it exists) must be the last node in the frame.

Figure 10-1 on page 289 shows examples on this node configuration rule.

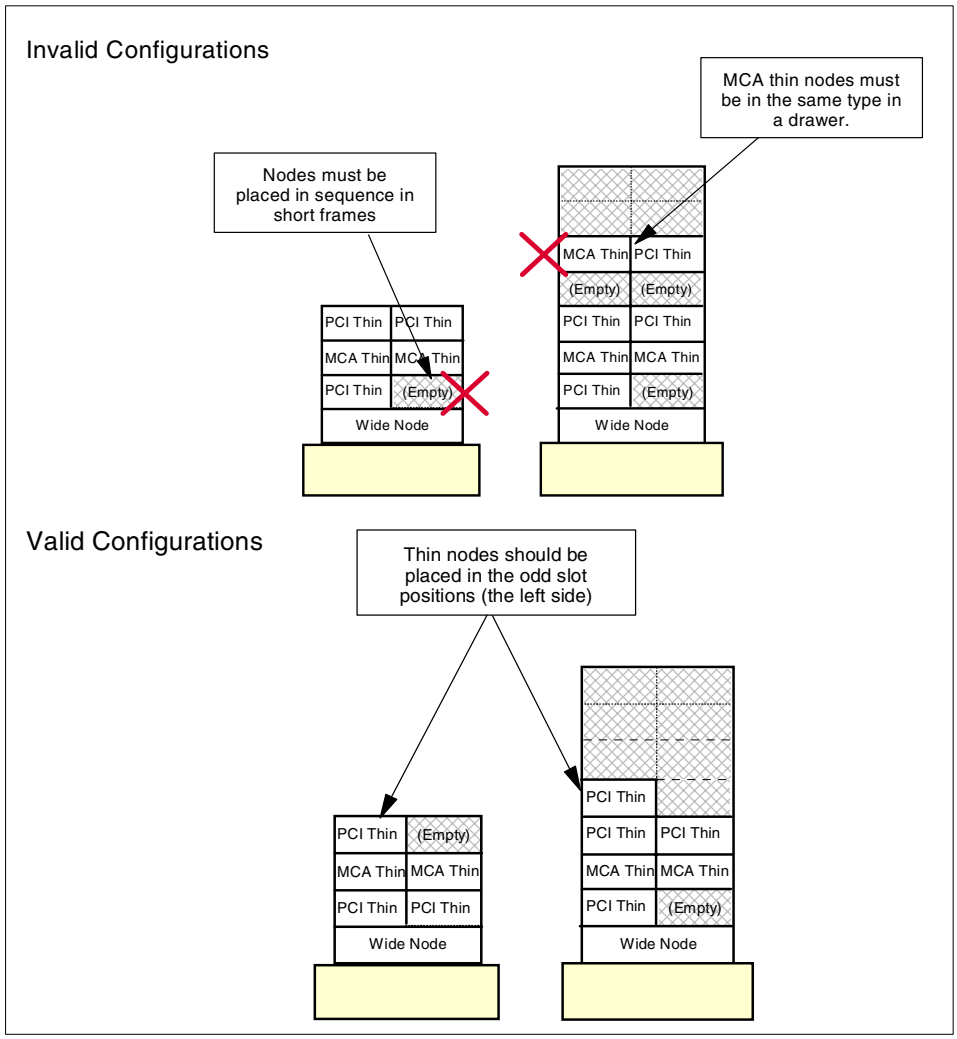


Figure 10-1 Examples on node configuration

Note that for MCA nodes, a pair of nodes is still required for each SP node drawer.

Node configuration rule 4:

SP-attached servers are supported on Model 550 only.

SP-attached servers are supported on SP Model 550 only.

Placement rules for frames

Frame configuration rule 1:

> Tall frames and Short frames can not be mixed in an SP system.

All the frames in an SP system must be either Tall frames or Short frames; a mixture of both is not supported. The SP Switch frame is classified as a Tall frame and therefore can be used with the Tall frames.

Frame configuration rule 2:

> A Short frame must be completely populated before adding another Short expansion frame. No imbedded empty drawer is allowed.

The Short frames must be completely full before the Short expansion frame can mount nodes, as shown in Figure 10-2.

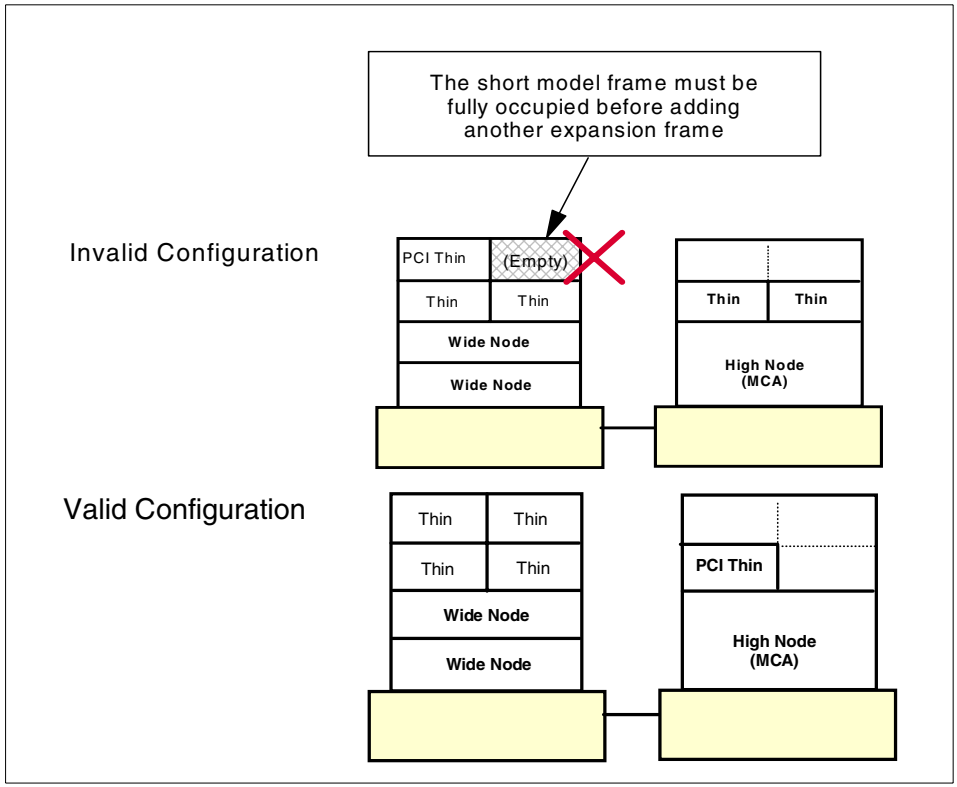


Figure 10-2 Example on frame configuration

Frame configuration rule 3:

A Short frame supports one SP Switch-8 board only.

The Short frame supports only the SP Switch-8 (F/C 4008) and only one Switch can be configured in an SP system composed of Short frames. This configuration mounts one to eight nodes through a single SP eight port Switch and the nodes can be hosted in a single Short frame, or with non-switched Short expansion frames after the first Short frame is fully populated. Nodes in the non-switched Short expansion frames share unused switch ports in the switched Short model frame with nodes in that frame. Figure 10-3 on page 292 shows examples of maximum SP eight port Switch Short frame configurations.

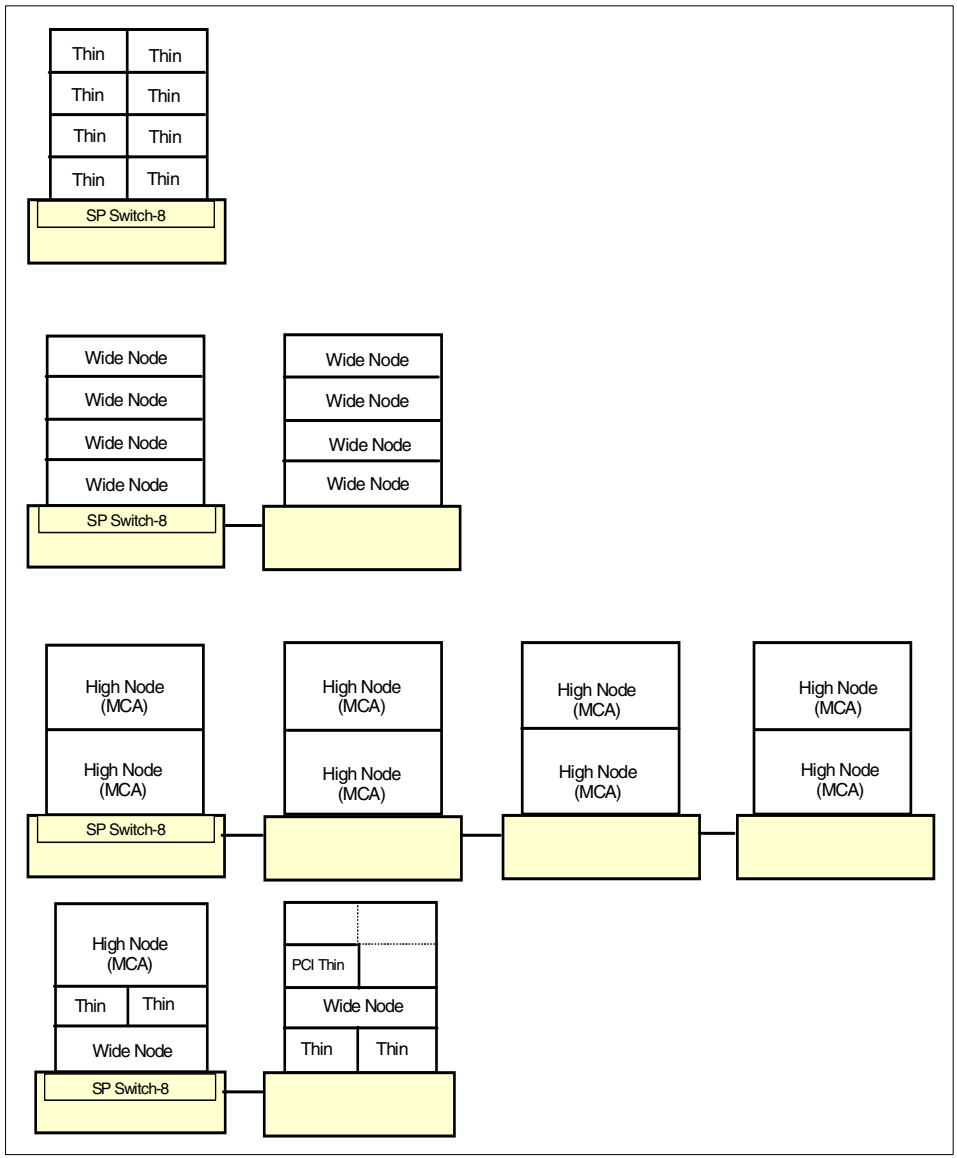


Figure 10-3 Maximum short frame configurations with SP eight port Switch

Frame configuration - rule 4:

When a Tall frame is installed with an eight port Switch, a maximum of eight nodes and one frame is supported. No expansion frame can be added. No imbedded empty drawer is allowed.

When using the SP Switch-8 (F/C 4008), the tall frame can only mount one to eight nodes. A minimum configuration is formed by one Tall model frame equipped with an SP Switch-8 and a single PCI thin node, one wide node, one MCA high node (a PCI high node is not supported), or one pair of MCA thin nodes.

Similar to Short frames, no imbedded empty drawer is allowed and, if there is a single PCI thin node in a drawer, it must be placed as the last node in a frame. Figure 10-4 on page 293 shows examples of SP Switch-8 Tall frame configurations.

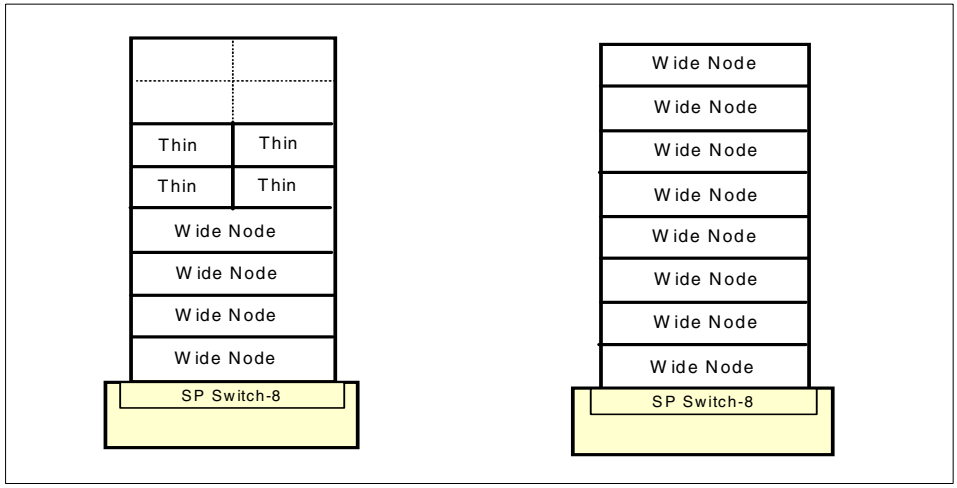


Figure 10-4 Examples on SP Switch-8 Tall frame configurations

Frame configuration - rule 5:

The SP Switch (16 ports) allows a maximum of four switched Tall expansion frames to attach to the Tall model frame, without an Intermediate Switch Board (ISB).

The SP Switch (16 ports) (F/C 4011 and F/C 4012) allows a maximum of four switched Tall expansion frames to attach to the Tall model frame, without ISB. Therefore, 1 to 80 processor nodes are supported in this configuration. If your SP system has no more than 16 nodes, a single SP Switch is enough to fulfill the system requirements. In this circumstance, non-switched expansion frames may be added, depending on the number of nodes and node locations. Figure 10-5 on page 294 and Figure 10-6 on page 295 show several examples of single stage SP Switch (16 ports) configurations with no more than 16 nodes.

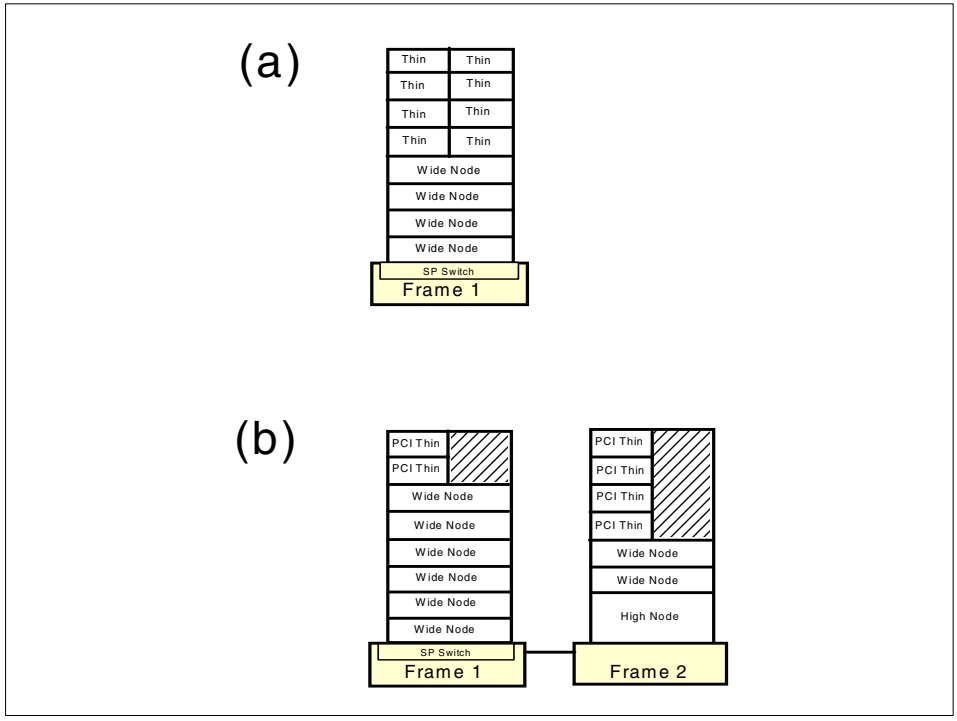


Figure 10-5 Examples on single SP Switch-16 configurations

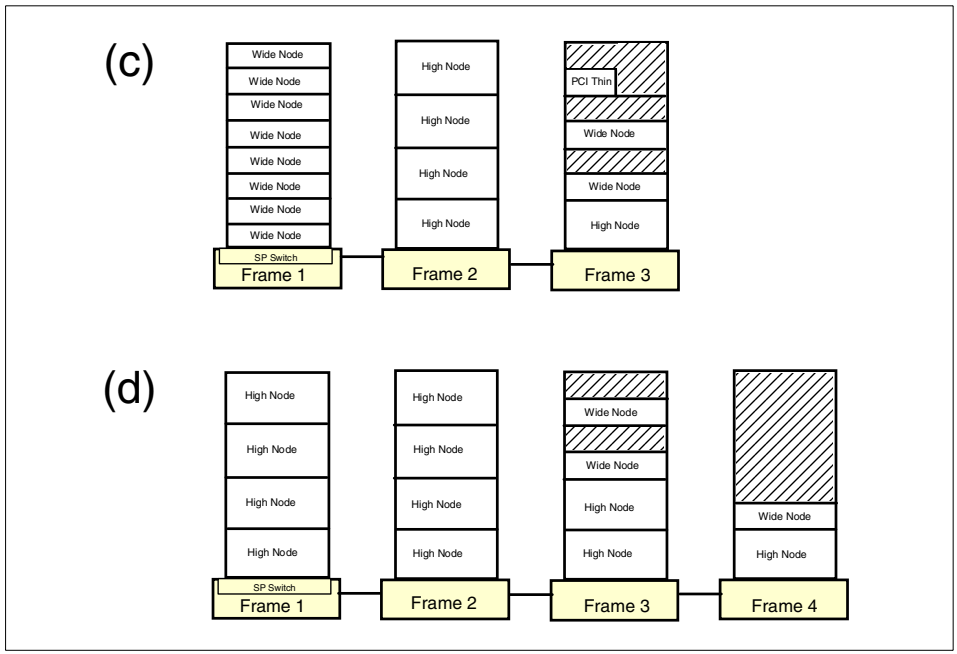


Figure 10-6 Other examples on single SP Switch-16 configurations

Frame configuration - rule 6:

> If a switched model frame or switched expansion frame has Thin nodes on the right side, it cannot support expansion frames.

In configuration (a) of Figure 10-5 on page 294, four wide nodes and eight thin nodes are mounted in a Tall model frame equipped with an SP Switch. There are four available switch ports that you can use to attach SP-attached servers or SP Switch routers. If an expansion frame is added, the four thin nodes on the right size of the frame are not longer in valid positions, since these slots do not have available switch port numbers.

In configuration (b) in Figure 10-5 on page 294, six wide nodes and two PCI thin nodes are are mounted in a Tall model frame equipped with an SP Switch. There also is a High node, two Wide nodes, and four PCI Thin nodes mounted in a non-switched expansion frame. Note that all PCI Thin nodes on the model frame must be placed on the left side to comply with frame configuration rule. All Thin nodes on an expansion frame are also placed on the left side to comply with the switch port numbering rule. There is one available switch port that you can use to attach SP-Attached servers or SP Switch routers.

In configuration (c) of Figure 10-6 on page 295, there are eight Wide nodes mounted in a tall model frame equipped with an SP Switch and four High nodes mounted in a non-switched expansion frame (Frame 2). The second non-switched expansion frame (Frame 3) is housed in a High node, two Wide nodes, and one PCI Thin node. This configuration occupies all 16 switch ports in the model frame. Note that Wide nodes and PCI Thin nodes in Frame 3 have to be placed in the location that is similar to a high node.

Frame configuration rule 7:

A maximum of three non-switched Tall expansion frames can be attached to each Tall model frame or switched Tall expansion frame.

Each Tall frame can attach a maximum of three non-switched Tall expansion frames. Configuration (d) in Figure 10-6 on page 295 shows an example.

If your SP system has 17 to 80 nodes, switched expansion frames are required. You can add switched expansion frames and non-switched expansion frames. Nodes in the non-switched expansion frame share unused switch ports that may exist in the model frame and in the switched expansion frames. Figure 10-7 shows an example of a Single Stage SP Switch with both switched and non-switched expansion frame configurations. There are four SP Switches and each can support up to 16 processor nodes, a total of 64-nodes configuration.

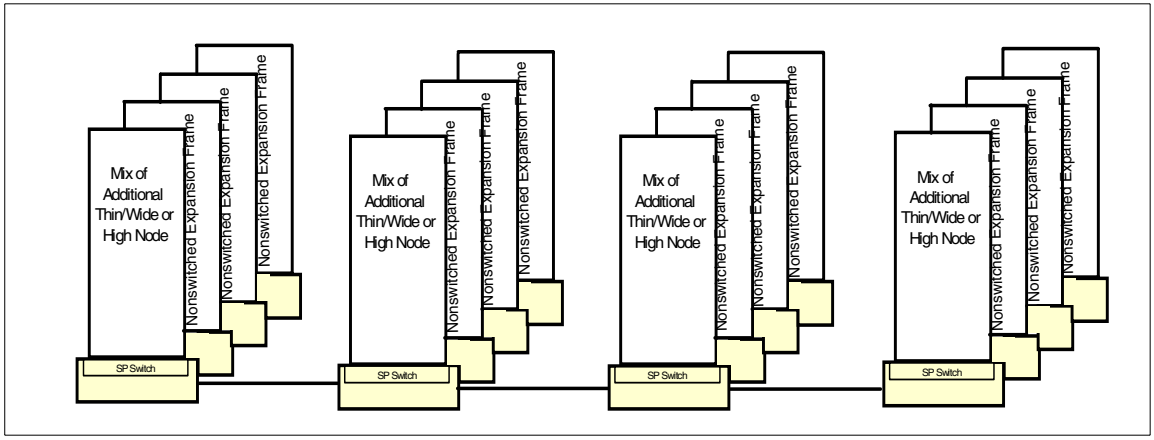


Figure 10-7 Example on multiple SP-Switch-16 configurations

Frame configuration rule 8:

For a more than five SP Switches configuration, an intermediate switch frame is required, which can expand up to 128 nodes.

An SP intermediate switch frame is a base, tall frame equipped with four SP Switches but no processor nodes. The switch frame is required for systems using more than five SP Switches to interconnect all the switches in the system. A switch frame supports systems with 65 to 128 nodes. However, it can also be configured into systems with fewer than 65 nodes to greatly simplify future expansion as more switches are added.

An example of using an SP Switch frame that forms the second switching layer is shown in Figure 10-8. The nodes are mounted in Tall model frames equipped with an SP Switch (16 ports). The SP Switch in these frames forms the first switching layer. This configuration can utilize both switched and non-switched expansion frames as well. Nodes in the non-switched expansion frame share unused switch ports that may exist in the model frame.

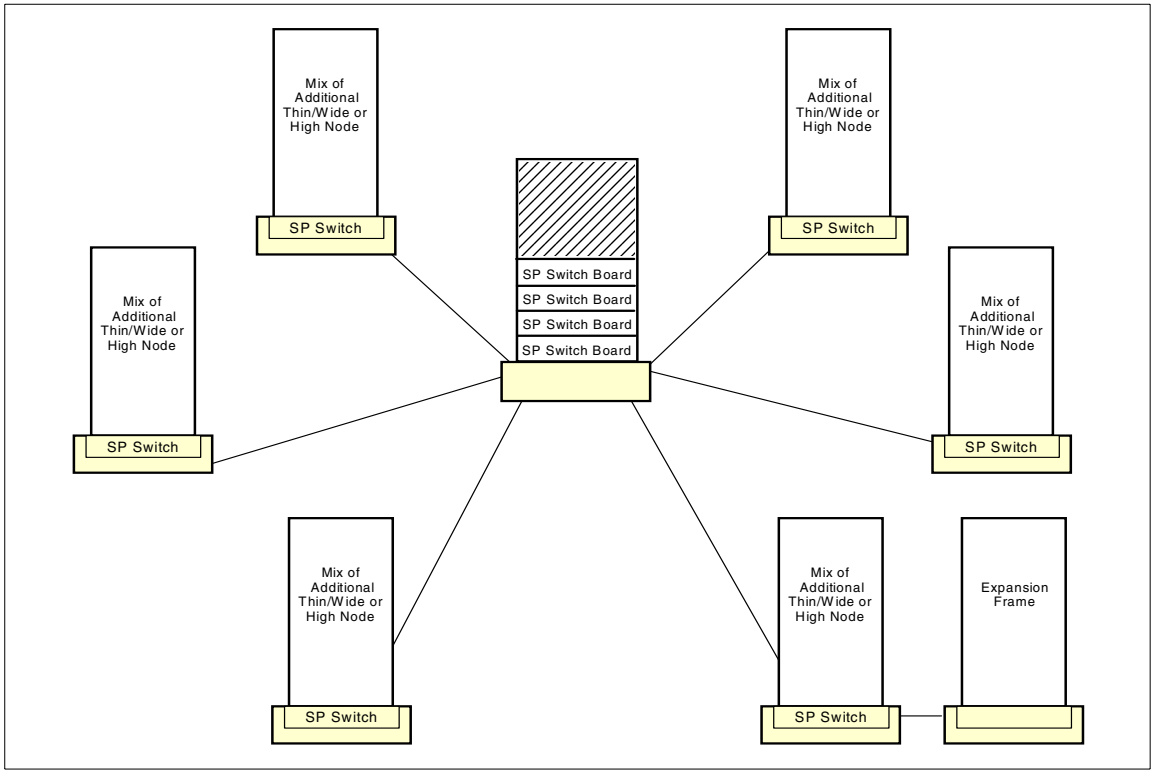


Figure 10-8 Example of an ISB configuration

Placement rules for switches

Switch configuration rule 1:

> An eight port switch does not support PCI high nodes, I/O Expansion units, or SP-attached servers.

All eight port switches do not support PCI high nodes, I/O expansion units, or SP-attached servers. As for the limited scalability of the eight port switches, they support up to a maximum of eight nodes.

Switch configuration rule 2:

> HiP Switch is compatible with MCA nodes only.

HP Switch requires HP Switch adapter-2 (F/C 4018) in an SP node, which is only available for for MCA, so HP Switch is only compatible with MCA nodes. Furthermore, HiP Switch is not supported in PSSP Version 3.1 or later.

Switch configuration rule 3:

> SP Switch is compatible with both MCA and PCI nodes.

The SP Switch (F/C 4011) can be used with both MCA and PCI nodes. A mixed environment of MCA nodes and PCI nodes is supported by SP Switch. However, as PSSP V3.1 is required for PCI POWER3 nodes, the coexistence of PSSP versions should be considered carefully when upgrading the SP system.

Switch configuration rule 4:

> The SP Switch (F/C 4011) only can be used with the SP-attached Servers and Clustered Enterprise Servers.

The SP Switch (F/C 4011) only can be used with the SP-attached Servers and Clustered Enterprise Servers. For the Clustered Enterprise Servers you have to configure 9076-555 Standalone Switch for Clustered Nodes with one up to two F/C 4011.

Switch configuration rule 5:

> SP Switch2 is compatible with PCI high nodes only (requires PSSP Version 3.2).

SP Switch2 (F/C 4012) provides a higher throughput and performance than SP Switch (F/C 4011) and it supports PCI high nodes and model 550 Tall frames only.

Table 10-1 shows the current compatibility of the switches and nodes; note that the switches in an SP/CES system cannot be mixed.

Table 10-1 Switch current compatibility with different types of nodes

	MCA nodes	PCI nodes (thin/wide)	PCI nodes (high)	SP-attached/ CES
HP Switch - 8	Yes	No	No	No
HP Switch - 16	Yes	No	No	No
SP Switch - 8	Yes	Yes	No	No
SP Switch - 16	Yes	Yes	Yes	Yes
SP Switch2	No	No	Yes	No

10.1.2 Numbering rules for frames, slots, nodes, and switch ports

There are some numbering rules in an SP/CES system that define the ordering of the hardware components:

- The frame numbering rule
- The slot numbering rule
- The node numbering rule
- The SP Switch port numbering rule

Frame numbering

Frame numbers define the order of frames in the SP/CES system. It describes the frame sequence to be examined during the configuration process. Each frame is referenced by the tty port to which the frame supervisor is attached and is assigned a numeric identifier. The frame numbers are used to assign global identifiers to the switch ports and nodes. This is also the order used to determine which frames share a switch. Normally, the frame numbers are assigned at the beginning of the system installation.

There are some general rules for frame numbering:

- If you have an SP Switch frame, you must configure it as the last frame in your SP system. Assign a high frame number to an SP Switch frame to allow for future expansion.

- Each SP-attached server and Clustered Enterprise Server occupies one frame number.
- If using PSSP Version 3.1.1 or before, the SP-attached server cannot be the first frame. It also cannot be inserted between a switched frame and any non-switched expansion frame using that switch.

Slot numbering

A Tall frame contains eight drawers and each drawer has two slots; there are 16 slots, in total. A Short frame has only four drawers; eight slots, in total, are available. A slot number describes the slot position of a node in an SP system. This is sensed by the software and can be viewed through software tools. When viewing a frame from the front, the slots are numbered sequentially from bottom left to top right.

For slot numbering, the following rules are applied:

- A thin node occupies a single slot in a drawer and its slot number is the corresponding slot.
- A wide node occupies two slots and its slots number is the odd-numbered slot.
- A high node occupies four consecutive slots in a frame. Its slots number is the first (lowest number) of these slots.
- An SP-attached server and a Clustered Enterprise Server is the only node in its frame. Its slot number is always 1, but uses the full 16 slot numbers. Therefore, 16 is added to the node number of the SP-attached server or Clustered Enterprise Server to get the node number of the next node.

Other special slot numbers are:

- Slot number zero (0) is used for the frame supervisor card.
- Slot number 17 is used for the SP Switch board supervisor card.

Node numbering

A node number is a global ID assigned to a node and it is automatically generated in the SP/CES system. Each node in the system, are assigned a node number, regardless of node or frame type, by the following formula (where slot_number is the lowest slot number occupied by the node):

```
node_number =((frame_number - 1) x 16) + slot_number
```

Each type (size) of node occupies one slot or a consecutive sequence of slots. For each node, there is an integer n such that a thin node (or Expansion I/O unit) occupies slot n, a wide node occupies slots n and n+1, and a high node occupies n, n+1, n+2 and n+3. For single thin nodes (not in a pair), wide nodes, and high nodes, n must be odd. For an SP-attached server and a Clustered Enterprise Server, n is 1. Use n in place of slot_number in the formula.

Node numbers are assigned independent of whether the frame is fully populated. Figure 10-9 demonstrates node numbering. Frame 4 represents an SP-attached server in a position where it does not interrupt the switched frame and companion non-switched expansion frame configuration. It can use a switch port on Frame 2, which is left available by the high nodes in frame 3. Frame 4's node number is determined by using the previous formula.

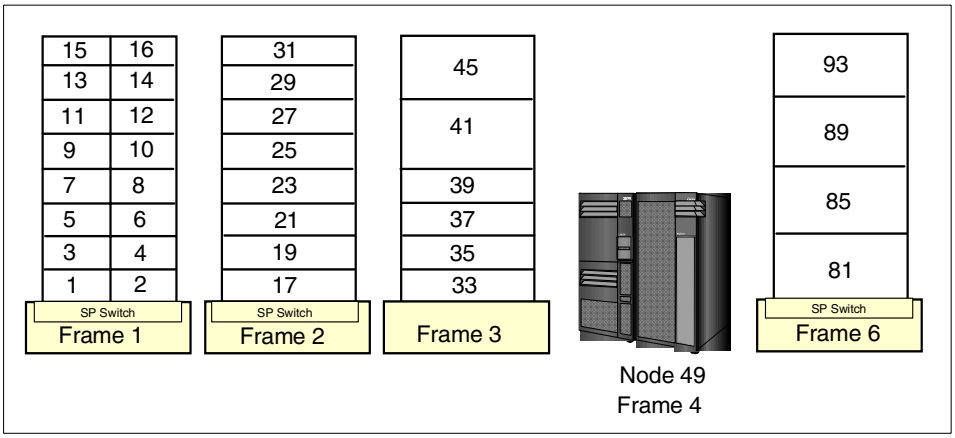


Figure 10-9 Node numbering in an SP system with SP-attached server

Switch port numbering

In a switched system, the switch boards are attached to each other to form a larger communication fabric. Each switch provides some number of ports to which a node can connect (16 ports for an SP Switch and 8 ports for the SP Switch-8). In larger systems, additional switch boards (intermediate switch boards) in the SP Switch frame are provided for switch board connectivity; such boards do not provide node switch ports.

Switch port numbers are used internally in PSSP software as a direct index into the switch topology and to determine routes between switch nodes.

If an SP System or a CES System has an SP Switch, Switch port number (also known as switch node number) is calculated by switch number and port number:

```
switch_port_number = ((switch_board_ number - 1) x 16) + port_number
```

Switch boards are numbered sequentially starting with 1 from the frame with the lowest frame number to that with the highest frame number.

Each full switch board contains a range of 16 port numbers that can be assigned. These ranges are also in sequential order with their switch board number. For example, switch board 1 contains switch port numbers 0 through 15.

Figure 10-10 on page 303 shows some possible examples of the switch port number assignment.

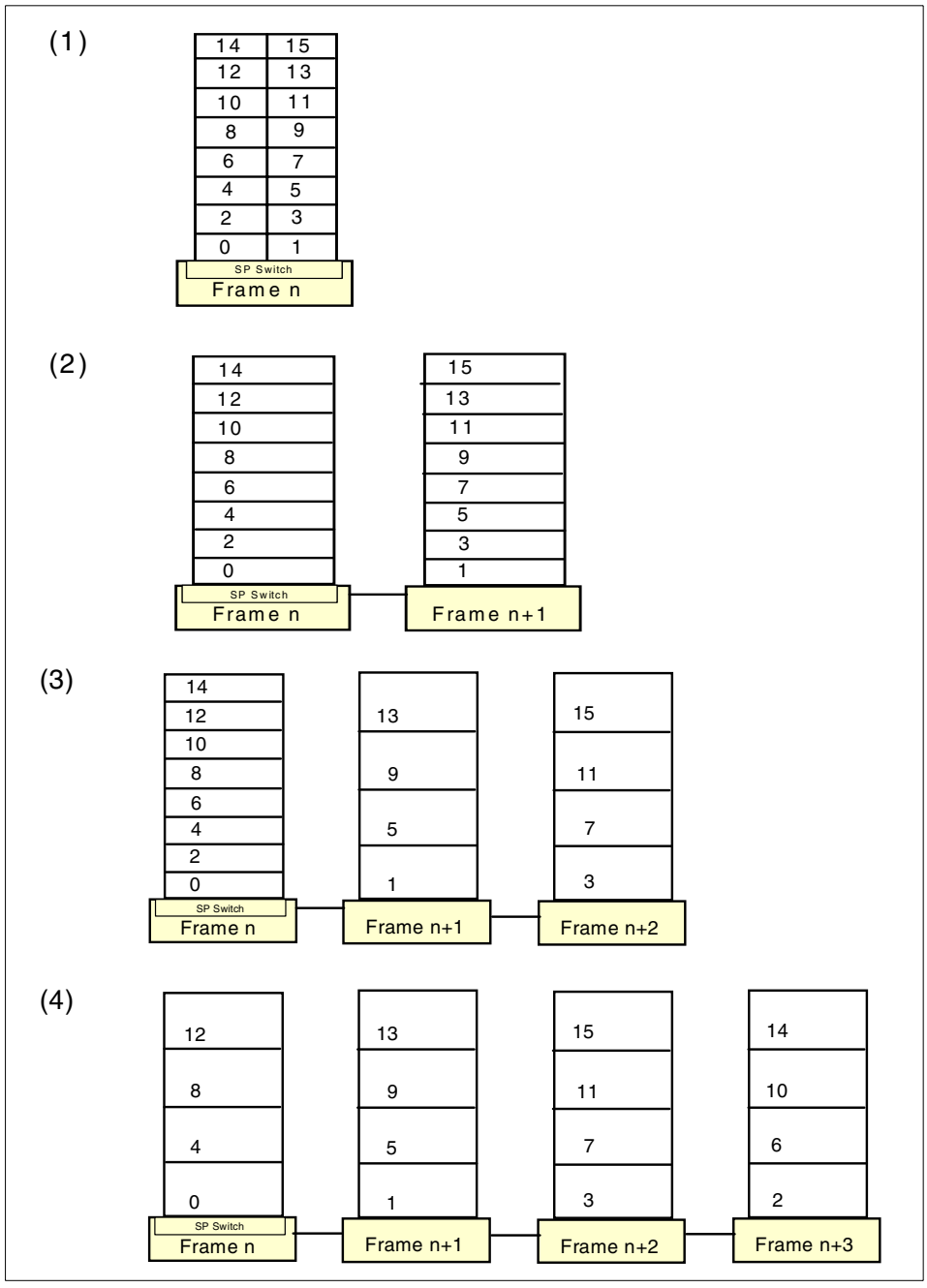


Figure 10-10 Switch port numbering on an SP Switch

SP-attached servers can connect to any available switch port of a switched SP frame, regardless of the frame number assigned to the SP-attached server. For example, in Figure 10-11 on page 304, an SP-attached server can be assigned with switch port number 1, 2, 3,17,18, or 19, no matter if it is physically connected to the Switch of Frame 1 or Frame 2.

Clustered Enterprise Servers can connect to any available switch port of a 9076-555 Standalone Switch for Clustered Nodes, regardless of the frame number assigned to the Clustered Enterprise Server.

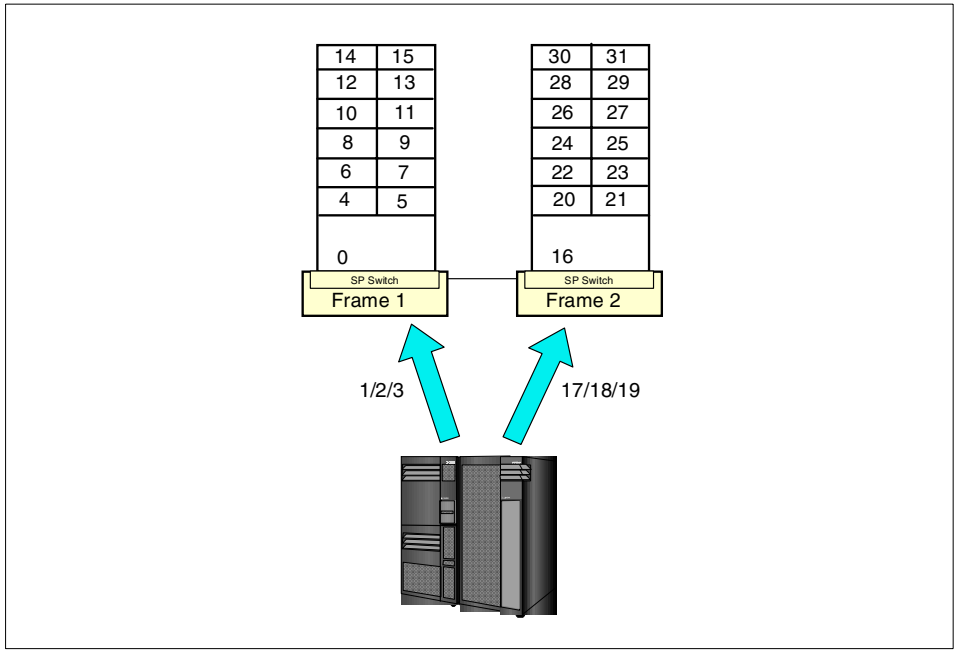


Figure 10-11 Switch port number assignment for SP-attached servers

If an SP/CES system does not have an SP Switch, the PSSP software still assigns switch port numbers. These switch port numbers are evaluated using frame numbers and slot numbers with the following formula:

```
switch_port_number = ((frame_number -1) x 16) + slot_number - 1
```

10.2 SP universal clustering capabilities

This section shows different SP universal clustering capabilities:

- SP system in non-switched and switched environments
- SP system with SP-attached servers in switched environment

- CES system in non-switched and switched environments

10.2.1 SP system in non-switched and switched environments

We may consider an SP system in non-switched or switched environments.

SP system in non-switched environment

Figure 10-12 on page 305 shows an example of a non-switched SP system.

The SP system has 16 thin nodes in an SP tall frame and a 7044-170 RS/6000 system as CWS.

Each SP node and the CWS are connected to the SP LAN administrative Ethernet. The SP frame has a serial line connection to CWS.

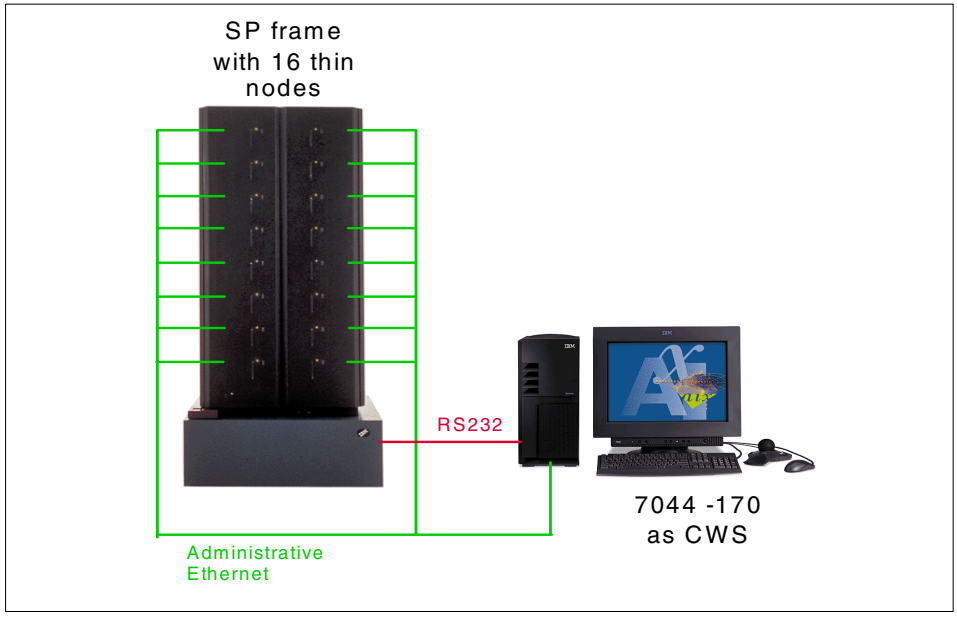


Figure 10-12 Non-switched SP environment

Example 10-1 shows a sample feature codes configuration for the non-switched SP environment considered.

Example 10-1 Sample configuration for non-switched SP environment

```
SP

***  SP HARDWARE  ***
```

```
  Product           Description            Qty

 9076-550 SP Tall Frame Model                1
     9123 RS/6000 Server Frame Attachment    1
     2056 375MHz Thin Node                  16
     2909 9.1 GB Ultra-SCSI Disk Drive      16
          Pair
     4098 Memory Expansion Feature          32
     4110 256 MB (2x128MB) SDRAM DIMMs      16
     4444 375 MHz POWER3-II Processor,      16
          2-way
     9222 Node Attachment - BNC Ethernet     1
     9432 AIX 4.3 with PSSP V3.2            16
     9920 Low Voltage Power (200-240V,       1
          50/60Hz)
     9930 PCI Node Enclosure                 8

Below is a description of your proposed system.  This includes the nodes used,
placement, and their devices.

---------------------------------------------------------------------
Frame: 1
Feature: 9076-550
Drawer        Left                Right

8 2056        thin                thin

7 2056        thin                thin

6 2056        thin                thin

5 2056        thin                thin

4 2056        thin                thin

3 2056        thin                thin

2 2056        thin                thin

1 2056        thin                thin

---------------------------------------------------------------------
thin: 375MHz Thin Node (2056)
Quantity placed: 16
  Product           Description            Qty
     2909 9.1 GB Ultra-SCSI Disk Drive       1
          Pair
     4098 Memory Expansion Feature           2
     4110 256 MB (2x128MB) SDRAM DIMMs       1
```

```
     4444 375 MHz POWER3-II Processor,       1
          2-way
     9432 AIX 4.3 with PSSP V3.2             1

***  SP SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                      1
     4061 AIX 4.3, per Node                 16
     9001 Asset Registration                 1

 5765-D51 Parallel System Support Program    1
          V3.2
     4011 SP Nodes                          16
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
     9001 Asset Registration                 1

 5765-D61 LoadLeveler V2.2                   1
     2005 SP Nodes                          16
     5000 CD-ROM                             1
     9001 Asset Registration                 1

 5765-E32 IBM C for AIX V5                   1
     0016 Base License                       1
     5809 CD-ROM                             1
     9001 Asset Registration                 1

 5765-E68 Performance AIDE V3.0              1
     0001 Installs                          1
     9001 Asset Registration                 1

 5765-E74 Performance Toolbox V3.0           1
     0001 Server Installs                    1
     9001 Asset Registration                 1

 5692-AIX System Software                    1
     0838 AIX 4.3 Update CD                  1
     0857 AIX 4.3                            1
     0860 AIX 4.3 Bonus Pack - Non           1
          Export.
     0917 Performance AIDE V3.0              1
     0922 Performance Toolbox V3.0           1
     1004 CD-ROM Process Charge              1
     2924 English                            1
     3410 CD-ROM                             1
     9001 Asset Registration                 1
```

```
170 CWS Default Entry

***  170 CWS Default Entry HARDWARE  ***

  Product           Description            Qty

 7044-170 RS/6000 44P Model 170              1
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra SCSI Adapter      1
          Integrated External Ultra2 SCSI    1
          Port
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM Drive             1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2943 8-Port Async Adapter,              1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet Adapter       1
     3102 18.2 GB 10K RPM Ultra SCSI Hard    2
          Disk Drive
     3628 P260 Color Monitor, Stealth        1
          Black
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     4223 Ethernet 10Base2 Transceiver       1
     4349 333MHz POWER3-II Processor Card    1
     5005 Preinstall                         1
     6159 12GB/24GB 4mm Tape Drive           1
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada (125V,      1
          15A)

***  170 CWS Default Entry SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                      1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1

 5692-AIX System Software                    1
     0598 AIX Welcome Center                 1
     0838 AIX 4.3 Update CD                  1
```

```
0857 AIX 4.3                          1
0859 Preinstall AIX 4.3 Server        1
     Bundle
0860 AIX 4.3 Bonus Pack - Non         1
     Export.
0903 CWS Diagnostic Pre-Load          1
1004 CD-ROM Process Charge            1
2924 English                          1
3410 CD-ROM                           1
5005 Preinstall                       1
5924 AIX 4.3 PII - English            1
9001 Asset Registration               1
```

SP system in switched environment

Figure 10-13 shows an example of a switched SP system using the SP Switch.

The SP cluster has two SP tall frames with a total of 16 wide nodes and a 7025-F80 RS/6000 system as CWS.

Each SP node and the CWS are connected to the SP LAN administrative Ethernet. Each SP frame has a serial line connection to CWS.

Each SP node is connected to the SP Switch.

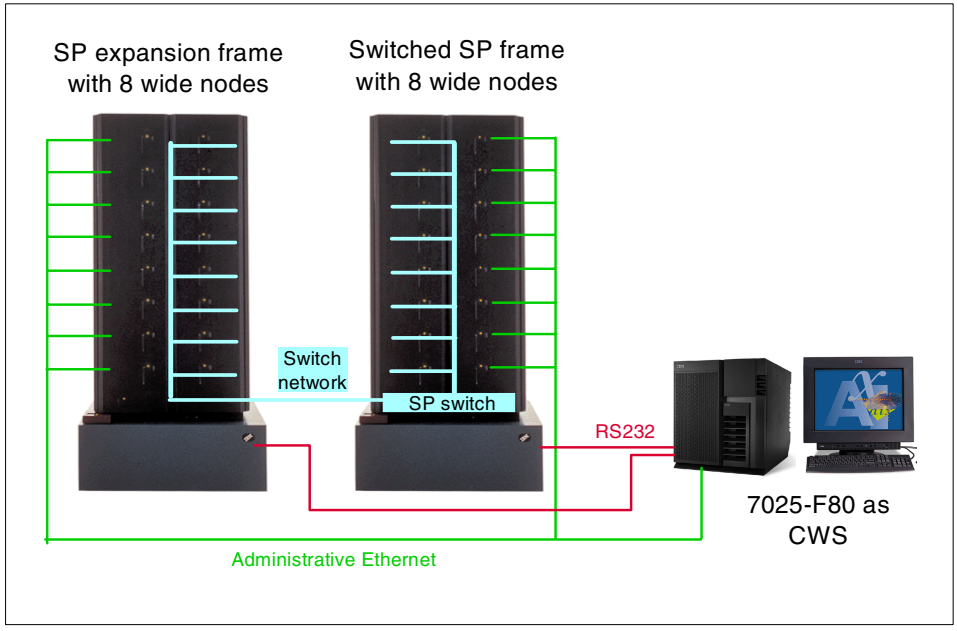


Figure 10-13 Switched SP environment

Example 10-2 shows a sample feature codes configuration for the switched SP environment considered.

Example 10-2 Sample configuration for switched SP environment

```
SP

***  SP HARDWARE  ***

  Product           Description            Qty

 9076-550 SP Tall Frame Model               1
     1550 Tall Expansion Frame, 10.5kW      1
          Three-phase SEPBU
     9123 RS/6000 Server Frame Attachment   1
     2057 375MHz Wide Node                 16
     2909 9.1 GB Ultra-SCSI Disk Drive     16
          Pair
     4011 SPS Switch                        1
     4023 SP Switch MX2 Adapter            16
     4098 Memory Expansion Feature         32
     4110 256 MB (2x128MB) SDRAM DIMMs     16
     4444 375 MHz POWER3-II Processor,     16
          2-way
     9201 Logical Frame Designation         1
     9222 Node Attachment - BNC Ethernet    1
     9302 2 meter SP Switch to Node cable   8
     9310 10 meter SP Switch/Switch or      8
          Switch/Node cable
     9432 AIX 4.3 with PSSP V3.2           16
     9920 Low Voltage Power (200-240V,      1
          50/60Hz)
     9930 PCI Node Enclosure               16

Below is a description of your proposed system.  This includes the nodes used,
placement, and their devices.

-----------------------------------------------------------------------
Frame: 1
Feature: 9076-550
Drawer        Left                Right

8 2057        wide

7 2057        wide

6 2057        wide

5 2057        wide
```

```
4 2057        wide

3 2057        wide

2 2057        wide

1 2057        wide

------------------------------------------------------------------------
Frame: 2
Feature:
Drawer        Left                Right

8 2057        wide

7 2057        wide

6 2057        wide

5 2057        wide

4 2057        wide

3 2057        wide

2 2057        wide

1 2057        wide

------------------------------------------------------------------------
wide: 375MHz Wide Node (2057)
Quantity placed: 16
  Product           Description           Qty
     2909 9.1 GB Ultra-SCSI Disk Drive      1
          Pair
     4023 SP Switch MX2 Adapter             1
     4098 Memory Expansion Feature          2
     4110 256 MB (2x128MB) SDRAM DIMMs      1
     4444 375 MHz POWER3-II Processor,      1
          2-way
     9432 AIX 4.3 with PSSP V3.2            1

***  SP SOFTWARE  ***

  Product           Description           Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3, per Node                16
     9001 Asset Registration                1
```

```
 5765-D51 Parallel System Support Program    1
          V3.2
     4011 SP Nodes                          16
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
     9001 Asset Registration                 1

 5765-D61 LoadLeveler V2.2                   1
     2005 SP Nodes                          16
     5000 CD-ROM                             1
     9001 Asset Registration                 1

 5765-E32 IBM C for AIX V5                   1
     0016 Base License                       1
     5809 CD-ROM                             1
     9001 Asset Registration                 1

 5765-E68 Performance AIDE V3.0              1
     0001 Installs                          1
     9001 Asset Registration                 1

 5765-E74 Performance Toolbox V3.0           1
     0001 Server Installs                    1
     9001 Asset Registration                 1

 5692-AIX System Software                    1
     0838 AIX 4.3 Update CD                  1
     0857 AIX 4.3                            1
     0860 AIX 4.3 Bonus Pack - Non           1
          Export.
     0917 Performance AIDE V3.0              1
     0922 Performance Toolbox V3.0           1
     1004 CD-ROM Process Charge              1
     2924 English                            1
     3410 CD-ROM                             1
     9001 Asset Registration                 1



F80 CWS Default Small

***  F80 CWS Default Small HARDWARE  ***

  Product           Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
```

```
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2943 8-Port Async Adapter,              1
          EIA-232/422 (PCI)
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI    1
     3109 SCSI External Port to SCSI         1
          Internal 6 Pack Cable Assembly
     3153 18.2 GB 10K RPM 1" Ultra3 SCSI     2
          16-bit Disk
     3628 P260 Color Monitor, Stealth        1
          Black
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     5005 Preinstall                         1
     5201 1-Way RS64 III 450 MHz             1
          Processor Card, 2MB L2 Cache
     6159 12GB/24GB 4mm Tape Drive           1
     6553 SCSI Hot Swap 6-Pack               1
     8700 Quiet Touch Keyboard, Stealth      1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black     1
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada (125V,      1
          15A)

***  F80 CWS Default Small SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0903 CWS Diagnostic Pre-Load           1
     1004 CD-ROM Process Charge             1
```

```
2924 English                          1
3410 CD-ROM                           1
5005 Preinstall                       1
5924 AIX 4.3 PII - English            1
9001 Asset Registration               1
```

10.2.2 SP system with SP-attached servers

Figure 10-14 shows an example of an SP system with SP-attached servers. It is a switched SP environment (the SP-attached servers can be connected to a non-switched SP environment as well). The SP system consists of an SP tall frame with eight wide nodes. Two enterprise servers M/T 7017 and three enterprise servers M/T 7026 are included on the SP environment as SP-attached servers. The CWS is a 7025-F80 RS/6000 system which uses a 128-port Asynchronous adapter and the necessary Random Asynchronous Node (RAN) as the serial connection interface.

Each 7017 server has two serial line connection and each 7026 server has one serial connection to the CWS. The SP frame has a serial connection to CWS.

Each SP node and each SP-attached server is connected by the SP administrative Ethernet LAN to the CWS.

Each SP node and each SP-attached server is connected to the SP Switch, which is hosted on the SP standard tall frame.

For a Twisted Pair network, a hub must be provided for twisted pair connections.

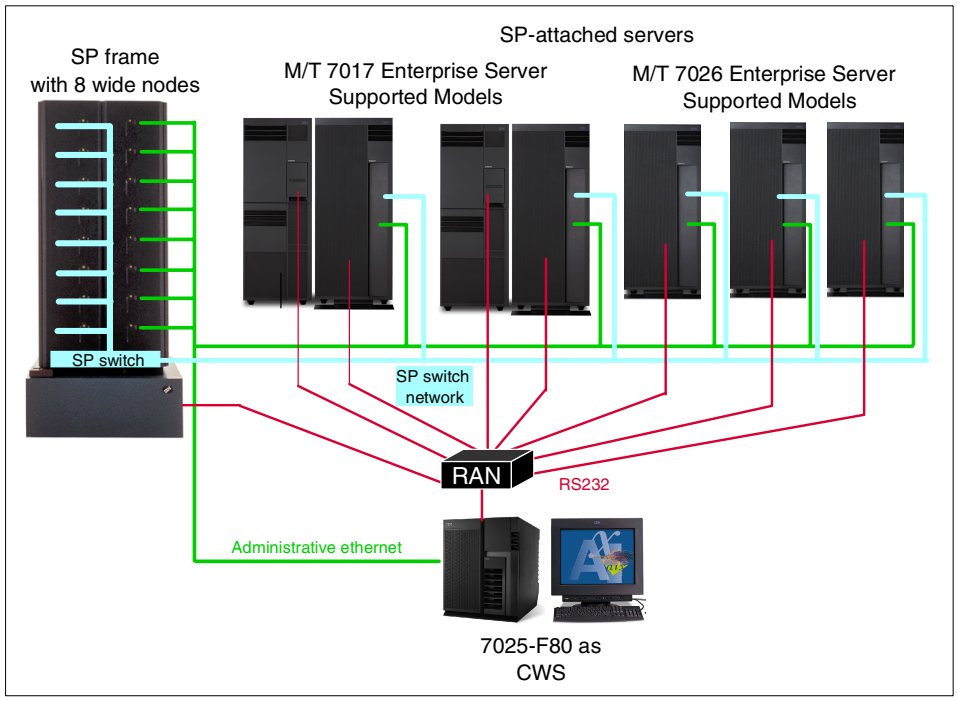


Figure 10-14 Switched SP environment with SP-attached servers

Example 10-3 shows a sample feature codes configuration for the switched SP environment with SP-attached servers considered.

Example 10-3 Sample configuration for switched SP environment with SP-attached servers

```
SP

***  SP HARDWARE  ***

  Product           Description            Qty

 9076-550 SP Tall Frame Model               1
     9123 RS/6000 Server Frame Attachment    5
     2057 375MHz Wide Node                   8
     9122 RS/6000 Server Attached Node       5
     2909 9.1 GB Ultra-SCSI Disk Drive       8
          Pair
     4011 SPS Switch                         1
     4023 SP Switch MX2 Adapter              8
     4098 Memory Expansion Feature          16
     4110 256 MB (2x128MB) SDRAM DIMMs       8
     4444 375 MHz POWER3-II Processor,       8
```

```
          2-way
     9223 Node Attachment - T/P Ethernet     5
     9302 2 meter SP Switch to Node cable    8
     9310 10 meter SP Switch/Switch or       5
          Switch/Node cable
     9432 AIX 4.3 with PSSP V3.2             8
     9920 Low Voltage Power (200-240V,       1
          50/60Hz)
     9930 PCI Node Enclosure                 8

Below is a description of your proposed system.  This includes the nodes used,
placement, and their devices.

-----------------------------------------------------------------------
Frame: 1
Feature: 9076-550
Drawer        Left                Right

8 2057        wide

7 2057        wide

6 2057        wide

5 2057        wide

4 2057        wide

3 2057        wide

2 2057        wide

1 2057        wide

-----------------------------------------------------------------------
Frame: 2
Feature: 9123
Drawer        Left                Right

8

7

6

5

4
```

```
3

2 9122

1 Node        RS/6000 Server Attac

------------------------------------------------------------------------
Frame: 3
Feature: 9123
Drawer        Left                Right

8

7

6

5

4

3

2 9122

1 Node        RS/6000 Server Attac

------------------------------------------------------------------------
Frame: 4
Feature: 9123
Drawer        Left                Right

8

7

6

5

4

3

2 9122

1 Node        RS/6000 Server Attac

------------------------------------------------------------------------
```

```
Frame: 5
Feature: 9123
Drawer        Left                Right

8

7

6

5

4

3

2 9122

1 Node        RS/6000 Server Attac

------------------------------------------------------------------------
Frame: 6
Feature: 9123
Drawer        Left                Right

8

7

6

5

4

3

2 9122

1 Node        RS/6000 Server Attac

------------------------------------------------------------------------
RS/6000 Server Attached Node: RS/6000 Server Attached Node
Quantity placed: 5
------------------------------------------------------------------------
wide: 375MHz Wide Node (2057)
Quantity placed: 8
  Product           Description           Qty
     2909 9.1 GB Ultra-SCSI Disk Drive      1
```

```
          Pair
     4023 SP Switch MX2 Adapter              1
     4098 Memory Expansion Feature           2
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     4444 375 MHz POWER3-II Processor,       1
          2-way
     9432 AIX 4.3 with PSSP V3.2             1

***  SP SOFTWARE  ***

  Product           Description           Qty

 5765-C34 AIX Version 4                     1
     4061 AIX 4.3, per Node                 8
     9001 Asset Registration                1

 5765-D51 Parallel System Support Program   1
          V3.2
     4011 SP Nodes                          8
     5812 RS/6000 SP Resource Center        1
          (CD-ROM)
     5813 CD-ROM                            1
     9001 Asset Registration                1

 5765-D61 LoadLeveler V2.2                  1
     2005 SP Nodes                          8
     5000 CD-ROM                            1
     9001 Asset Registration                1

 5765-E32 IBM C for AIX V5                  1
     0016 Base License                      1
     5809 CD-ROM                            1
     9001 Asset Registration                1

5765-E74 Performance Toolbox V3.0           1
     0001 Server Installs                   1
     9001 Asset Registration                1

 5692-AIX System Software                   1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0917 Performance AIDE V3.0             1
     0922 Performance Toolbox V3.0          1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     9001 Asset Registration                1
```

```
M/T 7017 (Quantity: 2)

***  M/T 7017 HARDWARE  ***

  Product           Description            Qty

 7017-S85 pSeries 680                        1
          1.44MB 3.5-in Diskette Drive       1
          This machine will function as      1
          an SP node
     2447 16-Bit PCI SCSI SE Adapter to 6    1
          SCSI Bays Cable
     2624 32x Speed CD-ROM Drive             1
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     3002 9.1 GB 10K RPM 1" Ultra SCSI       1
          Hot Swap Disk Drive
     3143 Remote I/O Cable - 6m              2
     4191 2048 MB Memory (4x512MB Cards)     1
     5005 Preinstall                         1
     5320 6-way 600 MHz RS64 IV Processor    1
          Card, 16MB L2 Cache
     6000 System Control and                 1
          Initialization Cable
     6008 Processor Complex to Rack Power    2
          Control Cable
     6088 Front Door for 1.8m Rack, Black    1
          with Copper
     6098 Side Panel for 1.8 or 2.0m         2
          Rack, Black
     6206 Ultra SCSI PCI-Bus Adapter         2
     6320 SCSI I/O Drawer, 10 EIA            1
     6321 Primary I/O Drawer Group           1
     6326 Support Processor Group            1
     6503 Remote I/O Hub, Dual Loop          1
     6547 SCSI 6-pack Hot Swap Back Plane    1
     7036 36U I/O Rack                       1
     8396 SP System Attachment Adapter       1
     9171 Power Distribution Unit,           1
          Side-Mount , 1 Phase
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada             1

***  M/T 7017 SOFTWARE  ***

  Product           Description            Qty
```

```
 5765-C34 AIX Version 4                      1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1

 5765-D51 Parallel System Support Program    1
          V3.2
     4012 External Attached Servers (Sxx     1
          models)
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
     9001 Asset Registration                 1

 5765-E68 Performance AIDE V3.0              1
     0001 Installs                           1
     9001 Asset Registration                 1

 5692-AIX System Software                    1
     0598 AIX Welcome Center                 1
     0838 AIX 4.3 Update CD                  1
     0857 AIX 4.3                            1
     0859 Preinstall AIX 4.3 Server          1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non           1
          Export.
     0917 Performance AIDE V3.0              1
     1004 CD-ROM Process Charge              1
     2924 English                            1
     3410 CD-ROM                             1
     5005 Preinstall                         1
     5924 AIX 4.3 PII - English              1
     9001 Asset Registration                 1


M/T 7026 (Quantity: 3)

***  M/T 7026 HARDWARE  ***

  Product           Description            Qty

 7026-6H1 pSeries 660                        1
          1.44MB 3.5-in Diskette Drive       1
          Integrated SCSI-2 F/W Adapter      1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
          This machine will function as      1
          an SP node
```

```
     2624 32x Speed CD-ROM                   1
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     3027 9.1 GB 10K RPM Ultra SCSI Disk     1
          Drive
     3142 Remote I/O Cable - 3m              2
     3154 Cable and PCI Card for SP          1
          Control Workstation Attachment
     4110 256 MB (2x128MB) SDRAM DIMMs       1
     5005 Preinstall                         1
     5211 1-Way RS64 IV 600 MHz Processor    1
          Card, 2MB L2 Cache
     5992 System Control and                 1
          Initialization Cable
     6132 CEC to Primary I/O Drawer Power    1
          Control Cable, 3m
     6324 Primary I/O Drawer, 5 EIA          1
     6540 IPL Disk Mounting Hardware,        1
          Cables, Terminator
     8396 SP System Attachment Adapter       1
     9172 AC Power Specify                   1
     9300 Language - English (US)            1

 7014-T00 Enterprise Rack - 36 EIA           1
     0176 Content : FC 6324 (5 EIA)          1
     0188 Content : 7026-6H1 (5 EIA)         1
     6088 Front Door for 1.8m Rack, Black    1
     6098 Side Panel for 1.8 or 2.0m         2
          Rack, Black
     9171 Power Distribution Unit,           1
          Side-Mount, 1 Phase
     9300 Language - English (US)            1
     9800 Rack Power Cord - US/Canada        1

***  M/T 7026 SOFTWARE  ***

  Product           Description            Qty

5765-C34 AIX Version 4                       1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1

 5765-D51 Parallel System Support Program    1
          V3.2
     4014 H80 Servers                        1
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
```

```
     9001 Asset Registration                1

5692-AIX System Software                    1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0917 Performance AIDE V3.0             1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1


F80 CWS Default Med

***  F80 CWS Default Med HARDWARE  ***

  Product           Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2944 128-Port Async Controller (PCI)    1
     2968 10/100 Mbps Ethernet PCI           3
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI    1
     3109 SCSI External Port to SCSI         1
          Internal 6 Pack Cable Assembly
     3153 18.2 GB 10K RPM 1" Ultra3 SCSI     2
          16-bit Disk
     3628 P260 Color Monitor, Stealth        1
          Black
     4075 Memory Board, 16-position          1
     4110 256 MB (2x128MB) SDRAM DIMMs       2
     4223 Ethernet 10Base2 Transceiver       3
     5005 Preinstall                         1
     5202 2-Way RS64 III 450 MHz             1
          Processor Card, 4MB L2 Cache
```

```
     6159 12GB/24GB 4mm Tape Drive          1
     6553 SCSI Hot Swap 6-Pack              1
     8131 4.5m Controller Cable             1
     8137 Enhanced Async Node 16-Port       1
          EIA-232
     8700 Quiet Touch Keyboard, Stealth     1
          Black - English (US)
     8741 3-Button Mouse - Stealth Black    1
     9300 Language - English (US)           1
     9800 Power Cord - US/Canada (125V,     1
          15A)

***  F80 CWS Default Med SOFTWARE  ***

  Product           Description           Qty

 5765-C34 AIX Version 4                    1
     4061 AIX 4.3                          1
     5005 Preinstall                       1
     9001 Asset Registration               1

 5692-AIX System Software                  1
     0598 AIX Welcome Center               1
     0838 AIX 4.3 Update CD                1
     0857 AIX 4.3                          1
     0859 Preinstall AIX 4.3 Server        1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non         1
          Export.
     0903 CWS Diagnostic Pre-Load          1
     1004 CD-ROM Process Charge            1
     2924 English                          1
     3410 CD-ROM                           1
     5005 Preinstall                       1
     5924 AIX 4.3 PII - English            1
     9001 Asset Registration               1
```

10.2.3 CES system in non-switched and switched environment

We may consider a CES system in non-switched or switched environments.

CES system in non-switched environment

Figure 10-15 on page 325 shows an example of a CES system in non-switched environment (pure CES system).

It is composed of seven Enterprise Servers M/T 7017 and 12 Enterprise Servers M/T 7026. The CWS is an 7025-F80 RS/6000 system with one 128-port Asynchronous adapter and a chain of two necessary RANs.

Each 7017 server has two serial line connection and each 7026 server has one serial connection to the CWS. Each server is connected by the SP administrative Ethernet LAN to the CWS.

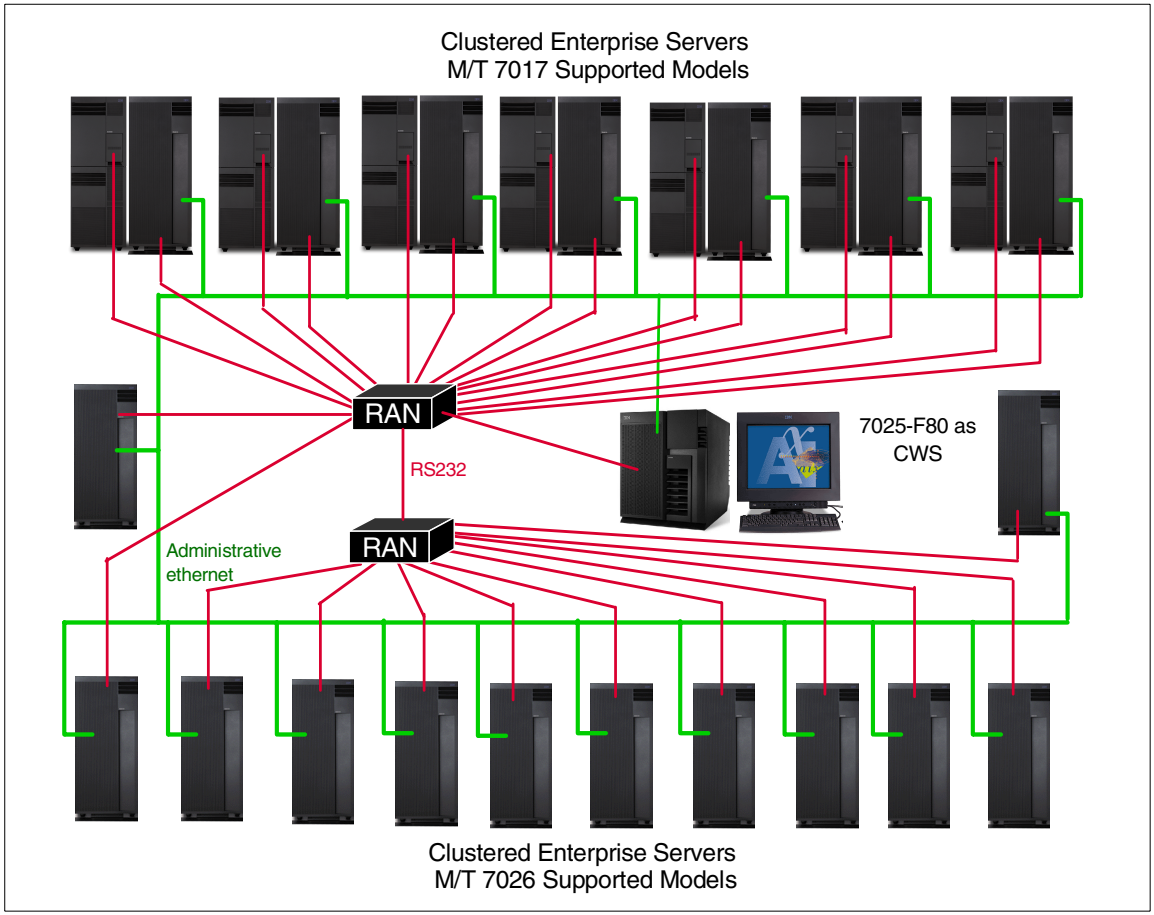


Figure 10-15 Pure CES system environment

Example 10-4 shows a sample feature codes configuration for the Pure CES system environment considered.

Example 10-4 Sample configuration for a pure CES system environment

```
M/T 7017 (Quantity: 7)
```

```
***  M/T 7017 HARDWARE  ***

  Product           Description            Qty

 7017-S80 RS/6000 Enterprise Server S80      1
          1.44MB 3.5-in Diskette Drive       1
          This machine will be part of a     1
          cluster
     2447 16-Bit PCI SCSI SE Adapter to 6    1
          SCSI Bays Cable
     2624 32x Speed CD-ROM Drive             1
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     3002 9.1 GB 10K RPM 1" Ultra SCSI       1
          Hot Swap Disk Drive
     3143 Remote I/O Cable - 6m              2
     3150 Cable, Clustered Server Serial     1
          Port to Control Workstation
     3151 Cable, Clustered Server Control    1
          Panel to Control Workstation
     4190 1024 MB Memory (4x256MB Cards)     2
     5005 Preinstall                         1
     5318 6-way 450MHz RS64 III Processor    1
          Card, 8MB L2 Cache
     6000 System Control and                 1
          Initialization Cable
     6008 Processor Complex to Rack Power    2
          Control Cable
     6206 Ultra SCSI PCI-Bus Adapter         2
     6320 SCSI I/O Drawer, 10 EIA            1
     6321 Primary I/O Drawer Group           1
     6326 Support Processor Group            1
     6503 Remote I/O Hub, Dual Loop          1
     6547 SCSI 6-pack Hot Swap Back Plane    1
     7000 I/O Rack                           1
     9171 Power Distribution Unit,           1
          Side-Mount , 1 Phase
     9300 Language - English (US)            1
     9800 Power Cord - US/Canada             1

***  M/T 7017 SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                      1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1
```

```
 5765-D51 Parallel System Support Program    1
          V3.2
     4012 External Attached Servers (Sxx     1
          models)
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
     9001 Asset Registration                 1

5692-AIX System Software                     1
     0598 AIX Welcome Center                 1
     0838 AIX 4.3 Update CD                  1
     0857 AIX 4.3                            1
     0859 Preinstall AIX 4.3 Server          1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non           1
          Export.
     0917 Performance AIDE V3.0              1
     1004 CD-ROM Process Charge              1
     2924 English                            1
     3410 CD-ROM                             1
     5005 Preinstall                         1
     5924 AIX 4.3 PII - English              1
     9001 Asset Registration                 1


M/T 7026 (Quantity: 12)

***  M/T 7026 HARDWARE  ***

  Product           Description            Qty

 7026-M80 RS/6000 Enterprise Server Model    1
          M80
          1.44MB 3.5-in Diskette Drive       1
          Integrated SCSI-2 F/W Adapter      1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
          This machine will be part of a     1
          cluster
     2624 32x Speed CD-ROM                   1
     2968 10/100 Mbps Ethernet PCI           1
          Adapter
     3027 9.1 GB 10K RPM Ultra SCSI Disk     1
          Drive
     3142 Remote I/O Cable - 3m              2
     3151 Cable, Clustered Server Control    1
          Panel to Control Workstation
     3154 Cable and PCI Card for SP          1
```

```
          Control Workstation Attachment
     4073 Memory Board, 32-position          1
     4133 1024 MB Memory (8x128MB DIMMs)     1
     5005 Preinstall                         1
     5200 2-Way RS64 III 500MHz Processor    1
          Card, 4MB L2 Cache
     5992 System Control and                 1
          Initialization Cable
     6132 CEC to Primary I/O Drawer Power    1
          Control Cable, 3m
     6196 76A Processor Power Regulator      1
     6283 Redundant AC Power Supply, I/O     1
          Drawer, 595W
     6284 Redundant AC Power Supply, CEC,    1
          1100W
     6324 Primary I/O Drawer, 5 EIA          1
     6540 IPL Disk Mounting Hardware,        1
          Cables, Terminator
     9172 AC Power Specify                   1
     9300 Language - English (US)            1

7014-T00 Enterprise Rack - 36 EIA            1
     0129 Content : 7026-M80 (8 EIA)         1
     0176 Content : FC 6324 (5 EIA)          1
     6088 Front Door for 1.8m Rack, Black    1
     6098 Side Panel for 1.8 or 2.0m         2
          Rack, Black
     9171 Power Distribution Unit,           1
          Side-Mount, 1 Phase
     9300 Language - English (US)            1
     9800 Rack Power Cord - US/Canada        1

***  M/T 7026 SOFTWARE  ***

  Product           Description            Qty

 5765-C34 AIX Version 4                      1
     4061 AIX 4.3                            1
     5005 Preinstall                         1
     9001 Asset Registration                 1

 5765-D51 Parallel System Support Program    1
          V3.2
     4013 M80 Servers                        1
     5812 RS/6000 SP Resource Center         1
          (CD-ROM)
     5813 CD-ROM                             1
     9001 Asset Registration                 1
```

```
5692-AIX System Software                    1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0917 Performance AIDE V3.0             1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1


F80 CWS Default Lg

***  F80 CWS Default Lg HARDWARE  ***

  Product           Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2432 Point to Point Cable, 1.1m         1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2944 128-Port Async Controller (PCI)    1
     2968 10/100 Mbps Ethernet PCI           3
          Adapter
     2985 Ethernet BNC/RJ-45 Adapter, PCI    1
     3153 18.2 GB 10K RPM 1" Ultra3 SCSI     6
          16-bit Disk
     3628 P260 Color Monitor, Stealth        1
          Black
     4075 Memory Board, 16-position          1
     4119 512 MB (2x256MB) SDRAM DIMMs       2
     4223 Ethernet 10Base2 Transceiver       3
     5005 Preinstall                         1
     5202 2-Way RS64 III 450 MHz             1
          Processor Card, 4MB L2 Cache
     6158 20GB/40GB 4mm Tape Drive           1
     6205 PCI Dual Channel Ultra2 SCSI       1
          Adapter
```

```
    6553 SCSI Hot Swap 6-Pack              1
    8131 4.5m Controller Cable             2
    8133 RJ-45 to DB-25 Converter Cables   1
         (4 cables per set)
    8137 Enhanced Async Node 16-Port       2
         EIA-232
    8700 Quiet Touch Keyboard, Stealth     1
         Black - English (US)
    8741 3-Button Mouse - Stealth Black    1
    9300 Language - English (US)           1
    9800 Power Cord - US/Canada (125V,     1
         15A)

***  F80 CWS Default Lg SOFTWARE  ***

  Product          Description           Qty

 5765-C34 AIX Version 4                    1
    4061 AIX 4.3                           1
    5005 Preinstall                        1
    9001 Asset Registration                1

 5765-E32 IBM C for AIX V5                 1
    0016 Base License                      1
    3483 Single Pallet Delivery            1
    5809 CD-ROM                            1
    9001 Asset Registration                1

 5765-E68 Performance AIDE V3.0            1
    0001 Installs                          1
    9001 Asset Registration                1

 5765-E74 Performance Toolbox V3.0         1
    0001 Server Installs                   1
    9001 Asset Registration                1

 5692-AIX System Software                  1
    0598 AIX Welcome Center                1
    0838 AIX 4.3 Update CD                 1
    0857 AIX 4.3                           1
    0859 Preinstall AIX 4.3 Server         1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non          1
         Export.
    0903 CWS Diagnostic Pre-Load           1
    0917 Performance AIDE V3.0             1
    0922 Performance Toolbox V3.0          1
    1004 CD-ROM Process Charge             1
    2924 English                           1
```

```
3410 CD-ROM                          1
5005 Preinstall                      1
5924 AIX 4.3 PII - English           1
9001 Asset Registration              1
```

CES system in switched environment

Figure 10-16 on page 331 shows an example of a CES system in switched environment. It consists of two Enterprise Servers M/T 7017 and three Enterprise Servers M/T 7026. The CWS is a 7025-F80 RS/6000 system which uses a 128-port Asynchronous adapter and the necessary Random Asynchronous Node (RAN) as the serial connection interface.

Each 7017 server has two serial line connection and each 7026 server has one serial connection to the CWS. Each server is connected by the SP administrative Ethernet LAN to the CWS and is connected to the SP Switch, which is hosted in the 9076-555 Standalone Switch.

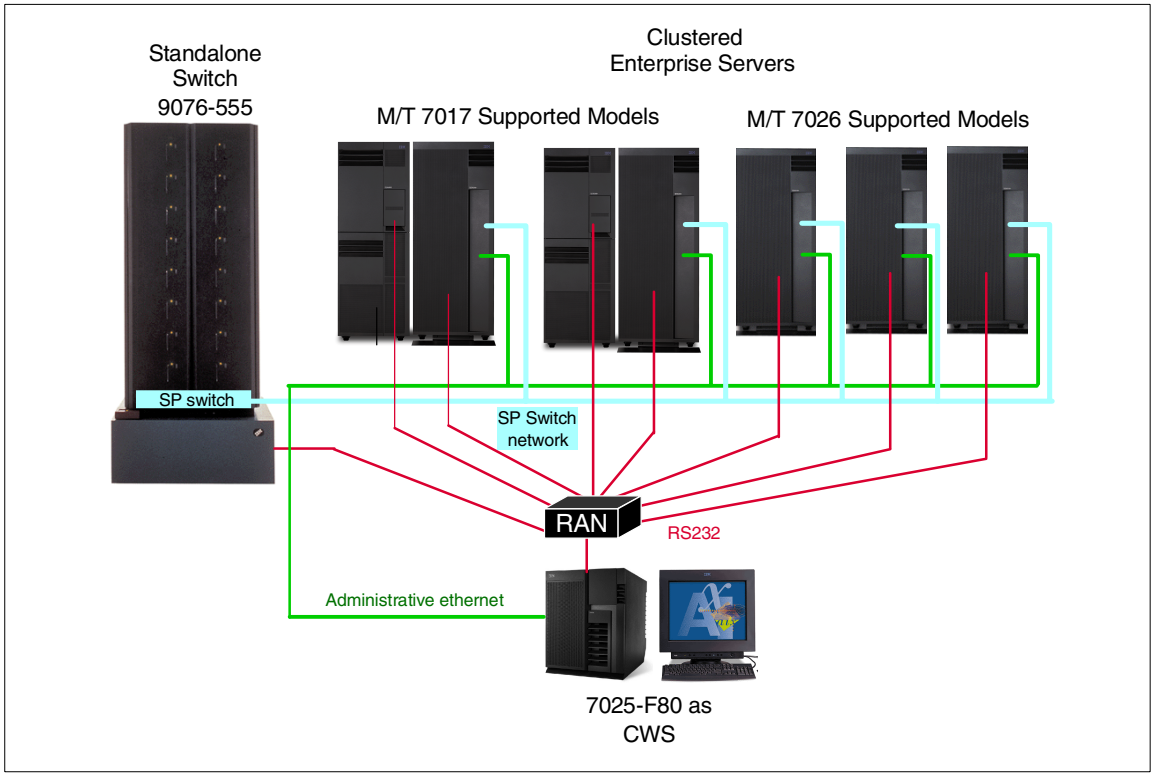


Figure 10-16 CES system in switched environment

Example 10-5 shows a sample feature codes configuration for the CES system in switched environment considered.

Example 10-5 CES system in switched environment

```
M/T 7017 (Quantity: 2)

***  M/T 7017 HARDWARE  ***

  Product           Description            Qty

 7017-S85 pSeries 680                       1
          1.44MB 3.5-in Diskette Drive      1
          This machine will function as     1
          an SP node
     2447 16-Bit PCI SCSI SE Adapter to 6   1
          SCSI Bays Cable
     2624 32x Speed CD-ROM Drive            1
     2968 10/100 Mbps Ethernet PCI          1
          Adapter
     3002 9.1 GB 10K RPM 1" Ultra SCSI      1
          Hot Swap Disk Drive
     3143 Remote I/O Cable - 6m             2
     4191 2048 MB Memory (4x512MB Cards)    1
     5005 Preinstall                        1
     5320 6-way 600 MHz RS64 IV Processor   1
          Card, 16MB L2 Cache
     6000 System Control and                1
          Initialization Cable
     6008 Processor Complex to Rack Power   2
          Control Cable
     6088 Front Door for 1.8m Rack, Black   1
          with Copper
     6098 Side Panel for 1.8 or 2.0m        2
          Rack, Black
     6206 Ultra SCSI PCI-Bus Adapter        2
     6320 SCSI I/O Drawer, 10 EIA           1
     6321 Primary I/O Drawer Group          1
     6326 Support Processor Group           1
     6503 Remote I/O Hub, Dual Loop         1
     6547 SCSI 6-pack Hot Swap Back Plane   1
     7036 36U I/O Rack                      1
     8396 SP System Attachment Adapter      1
     9171 Power Distribution Unit,          1
          Side-Mount , 1 Phase
     9300 Language - English (US)           1
     9800 Power Cord - US/Canada            1

***  M/T 7017 SOFTWARE  ***
```

```
  Product           Description            Qty

5765-C34 AIX Version 4                      1
     4061 AIX 4.3                           1
     5005 Preinstall                        1
     9001 Asset Registration                1

 5765-D51 Parallel System Support Program   1
          V3.2
     4012 External Attached Servers (Sxx    1
          models)
     5812 RS/6000 SP Resource Center        1
          (CD-ROM)
     5813 CD-ROM                            1
     9001 Asset Registration                1

5692-AIX System Software                    1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0917 Performance AIDE V3.0             1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1


M/T 7026 (Quantity: 3)

***  M/T 7026 HARDWARE  ***

  Product           Description            Qty

 7026-6H1 pSeries 660                       1
          1.44MB 3.5-in Diskette Drive      1
          Integrated SCSI-2 F/W Adapter     1
          Integrated Ultra2 SCSI Adapter    1
          Integrated Ethernet Adapter       1
          This machine will function as     1
          an SP node
     2624 32x Speed CD-ROM                  1
     2968 10/100 Mbps Ethernet PCI          1
          Adapter
```

```
    3027 9.1 GB 10K RPM Ultra SCSI Disk     1
         Drive
    3142 Remote I/O Cable - 3m              2
    3154 Cable and PCI Card for SP          1
         Control Workstation Attachment
    4110 256 MB (2x128MB) SDRAM DIMMs       1
    5005 Preinstall                         1
    5211 1-Way RS64 IV 600 MHz Processor    1
         Card, 2MB L2 Cache
    5992 System Control and                 1
         Initialization Cable
    6132 CEC to Primary I/O Drawer Power    1
         Control Cable, 3m
    6324 Primary I/O Drawer, 5 EIA          1
    6540 IPL Disk Mounting Hardware,        1
         Cables, Terminator
    8396 SP System Attachment Adapter       1
    9172 AC Power Specify                   1
    9300 Language - English (US)            1

 7014-T00 Enterprise Rack - 36 EIA          1
    0176 Content : FC 6324 (5 EIA)          1
    0188 Content : 7026-6H1 (5 EIA)         1
    6088 Front Door for 1.8m Rack, Black    1
    6098 Side Panel for 1.8 or 2.0m         2
         Rack, Black
    9171 Power Distribution Unit,           1
         Side-Mount, 1 Phase
    9300 Language - English (US)            1
    9800 Rack Power Cord - US/Canada        1

***  M/T 7026 SOFTWARE  ***

  Product           Description           Qty


 5765-C34 AIX Version 4                     1
    4061 AIX 4.3                            1
    5005 Preinstall                         1
    9001 Asset Registration                 1

 5765-D51 Parallel System Support Program   1
         V3.2
    4014 H80 Servers                        1
    5812 RS/6000 SP Resource Center         1
         (CD-ROM)
    5813 CD-ROM                             1
    9001 Asset Registration                 1
```

```
5692-AIX System Software                    1
     0598 AIX Welcome Center                1
     0838 AIX 4.3 Update CD                 1
     0857 AIX 4.3                           1
     0859 Preinstall AIX 4.3 Server         1
          Bundle
     0860 AIX 4.3 Bonus Pack - Non          1
          Export.
     0917 Performance AIDE V3.0             1
     1004 CD-ROM Process Charge             1
     2924 English                           1
     3410 CD-ROM                            1
     5005 Preinstall                        1
     5924 AIX 4.3 PII - English             1
     9001 Asset Registration                1


9076-555

***  9076-555 HARDWARE  ***

  Product           Description            Qty

 9076-555 Standalone Switch for Clustered    1
          Nodes
     4011 SPS Switch                        1
     9122 RS/6000 Server Attached Node      5
     9123 RS/6000 Server Frame Attachment   5
     9310 10 meter SP Switch/Switch or      5
          Switch/Node cable
     9920 Low Voltage Power (200-240V,      1
          50/60Hz)


F80 CWS Default Med

***  F80 CWS Default Med HARDWARE  ***

  Product           Description            Qty

 7025-F80 RS/6000 Enterprise Server Model    1
          F80
          1.44MB 3.5-in Diskette Drive       1
          Integrated Ultra2 SCSI Adapter     1
          Integrated Ethernet Adapter        1
     2624 32x Speed CD-ROM                   1
     2830 POWER GXT130P Graphics Adapter     1
          (PCI)
     2944 128-Port Async Controller (PCI)    1
```

```
    2968 10/100 Mbps Ethernet PCI          3
         Adapter
    2985 Ethernet BNC/RJ-45 Adapter, PCI   1
    3109 SCSI External Port to SCSI        1
         Internal 6 Pack Cable Assembly
    3153 18.2 GB 10K RPM 1" Ultra3 SCSI    2
         16-bit Disk
    3628 P260 Color Monitor, Stealth       1
         Black
    4075 Memory Board, 16-position         1
    4110 256 MB (2x128MB) SDRAM DIMMs      2
    4223 Ethernet 10Base2 Transceiver      3
    5005 Preinstall                        1
    5202 2-Way RS64 III 450 MHz            1
         Processor Card, 4MB L2 Cache
    6159 12GB/24GB 4mm Tape Drive          1
    6553 SCSI Hot Swap 6-Pack              1
    8131 4.5m Controller Cable             1
    8137 Enhanced Async Node 16-Port       1
         EIA-232
    8700 Quiet Touch Keyboard, Stealth     1
         Black - English (US)
    8741 3-Button Mouse - Stealth Black    1
    9300 Language - English (US)           1
    9800 Power Cord - US/Canada (125V,     1
         15A)

***  F80 CWS Default Med SOFTWARE  ***

  Product          Description           Qty

 5765-C34 AIX Version 4                    1
    4061 AIX 4.3                           1
    5005 Preinstall                        1
    9001 Asset Registration                1

 5765-E32 IBM C for AIX V5                 1
    0016 Base License                      1
    3483 Single Pallet Delivery            1
    5809 CD-ROM                            1
    9001 Asset Registration                1

 5765-E68 Performance AIDE V3.0            1
    0001 Installs                          1
    9001 Asset Registration                1

 5765-E74 Performance Toolbox V3.0         1
    0001 Server Installs                   1
    9001 Asset Registration                1
```

```
5692-AIX System Software                    1
    0598 AIX Welcome Center                 1
    0838 AIX 4.3 Update CD                  1
    0857 AIX 4.3                            1
    0859 Preinstall AIX 4.3 Server          1
         Bundle
    0860 AIX 4.3 Bonus Pack - Non           1
         Export.
    0903 CWS Diagnostic Pre-Load            1
    0917 Performance AIDE V3.0              1
    0922 Performance Toolbox V3.0           1
    1004 CD-ROM Process Charge              1
    2924 English                            1
    3410 CD-ROM                             1
    5005 Preinstall                         1
    5924 AIX 4.3 PII - English              1
    9001 Asset Registration                 1
```

10.3 PSSP environment for SP/CES systems

To support CES systems, PSSP 3.2 has made some enhancements to the SDR, various commands, and SP Perspectives.

SDR enhancements

To provide support for CES systems, a new system-wide boolean attribute is created in the SP class of the System Data Repository (SDR) called IsPartitionable. This attribute is set to the value of “false” for SP Switch2 systems and CES systems. It is set to the value of “true” for all other SP systems that have SP frames, whether or not these frames are switched or switchless.

The IsPartitionable attribute is set by the PSSP software based on the system configuration. There is no interface provided to the user to set this value.

Command enhancements

There are two commands enhanced due to support CES systems. They are:

- The **spframe** command
- The **spdelfram** command

The **spframe** command has been enhanced for the administrator who has no knowledge of a switch port number in SP systems. The **spdelfram** command has been enhanced to support in migration from SP systems to CES systems.

For details on these and other commands, see *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

The spframe command

The **spframe** command is used to enter configuration data for a frame or a series of frames and, optionally, to set up the initial System Data Repository (SDR).

The *-n* flag for the **spframe** command has been made optional for CES systems. If you are planning to have an SP-frame in the future, you have to reserve the switch port numbers for the future nodes. If you are not planning to have an SP-frame, then it is not necessary to assign switch port numbers. If the flag is not used, SDR_config dynamically assigns switch port numbers for these Enterprise Servers.

The *-p* flag for the **spframe** command has been updated to support two hardware protocol of the node within the frame:

- Service and Manufacturing Interface (SAMI) for M/T 7017 models.
- Converge Service Processor (CSP) for M/T 7026 models.

Table 10-2 is a summary of the **spframe** command for a CES system:

Syntax:

```
spframe -p {SAMI|CSP} [-n starting_switch_port] [-s {s1tty}] [-r yes | no]
start_frame frame_count starting_tty_port
```

Table 10-2 PSSP spframe command for an SPCES system

FLAGS	
-p {SAMI | CSP}	Indicates the hardware protocol of the node within the frame.This flag must be SAMI for M/T 7017 SP-attached server or Clustered Enterprise Server and, respectively, CSP for M/T 7026 SP-attached server or Clustered Enterprise Server.
-n starting_switch_port	Indicates the switch_ port_ number that is assigned to the Enterprise Server. This flag is required for SP-attached servers in systems with no switch or with an SP Switch. This flag is optional on CES systems only.
-s s1tty	Indicates the s1 tty port for the single Enterprise Server. This flag is optional, but if it is used, the s1tty must be specified as a fully qualified special device file name (such as /dev/tty2). The default value of the s1tty port is one plus the starting_tty_port operand.

FLAGS	
-r no | yes	Indicates wether you want to initialize the SDR. If this is the last or only time you are invoking this command during installation, specify -r yes. If -r yes is specified, the /spdata/sys1/spmon/hmacls file has the default entries created. The default is -r no.
OPERANDS	
start_frame	Specifies the frame number of the Enterprise Server to be used in this operation. Specify a value between 1 and 128 inclusive.
frame_count	Specifies the number of Enterprise Servers being added with this operation. The tty and s1tty port values are assigned sequentially beginning with the starting_tty_port operand. If the -s flag is specified, this operand must be one (1).
starting_tty_port	Specifies the device special file name of the tty port to be assigned to the Enterprise Server on this operation. Specify the full path of a valid tty special device file name (such as /dev/tty1).

Note: When using the **spframe** command with the -s flag to specify the s1 tty port, you can only add one frame at a time.

The following command example is allowed:

```
spframe -p SAMI -s /dev/tty10 1 1 /dev/tty0
```

However, the following command example is not allowed:

```
spframe -p SAMI -s /dev/tty10 1 2 /dev/tty0
```

The spdelfram command

The **spdelfram** command is used to remove configuration data for a frame or a series of frames from the System Data Repository (SDR).

```
Syntax
spdelfram [-c] {start_frame frame_count | -l frame_list}
```

The command has been enhanced to provide a new *-l* flag to specify a list of frame numbers to be deleted instead of using the start_frame and frame_count values. This list option can be used to delete all SP frames in a single operation to reconfigure an SP system directly to a CES system. Also, a second new option, the *-c* flag, has been added. This flag allows you to indicate that the system is being reconfigured to a CES system and that switch port number checking should be bypassed.

The following Table 10-3 is a summary of the **spdelfram** command for a CES system:

Table 10-3 PSSP spdelframe command for an SP/CES system

FLAGS	
-c	Specifies that switch port number verification should be bypassed. Use this flag when reconfiguring your existing SP system to a CES system.
-l frame_list	Specifies a list of frames to be used for this operation. frame_list is a comma-delimited list of frame numbers. If you use this flag, do not use the start_frame and frame_count operands
OPERANDS	
start_frame	Specifies the frame number of first node to be deleted in this operation. Specify a value between 1 and 128, inclusive. If you are using the -l flag, do not use this operand.
frame_count	Specifies the number of frames to be deleted in this operation. Specify a value between 1 and 128 inclusive. If you are using the -l flag, do not use this operand.

SP Perspectives enhancements

SP Perspectives is a suite of five SP Perspective applications:

- The Hardware Perspective
- The Event Perspective
- The VSD Perspective
- The System Partitioning Aid Perspective
- The Performance Monitor Perspective

SP Perspectives are enabled to perform system management tasks for your CES through direct manipulation of system objects represented by icons.

You simply select one or more Enterprise Server system objects (managed CES system resources, such as a frames or nodes) by clicking on them with a mouse and then selecting an action to perform on those system objects from the menu or tool bar. You can use this pattern of selecting one or more system objects and then selecting an action to perform on these system objects to accomplish numerous system management tasks with the SP Perspectives.

SP Perspectives provides function for hardware control and system object monitoring. For example, you can use SP Perspectives to power on and off or reset Enterprise Servers. Some configuration functions are also provided by SP Perspectives (for example, the ability to group Enterprise Servers into node groups).

You can utilize the following SP Perspective applications to perform the system management tasks for your CES system:

1. Hardware Perspective: Monitor and control Enterprise Servers

   For the case of CES, Hardware Perspective has an icon named “Hardware: Manage SP-attached S-Family Servers.”

   Each Enterprise Server has both frame and node personalities; therefore, it shows both the Nodes pane (Nodes:1) and Frames and Switches pane (Frames and Switches:1). The Nodes pane is displayed in the icon view, and the Frames and Switches pane is displayed in the tree view.

   To check for frame related Enterprise Server configuration information, click on one of the Frame icons in the Frames and Switches pane. Then click on the Notebook icon. This opens the “View or Modify Properties Frame” notebook. The frame status page can provide information as hardware protocol, CWS hostname, serial ports used on CWS.

   To check for node related Enterprise Server information or to control the Enterprise Server hardware, click on one of the Node icons in the Nodes pane. Then click on the Notebook icon. This opens the “View or Modify Properties Node” notebook. The Node Status page can provide information as if the power is on, if the node is up and running, how many CPUs are online. In addition to providing the information, this page allows you to control the Enterprise Server. You can perform the following operations for this Enterprise Server: Power it off, open its tty, and network boot it. From the Configuration page, you can find information as hostname, IP address, PSSP version and so on.

   The “View or Modify Properties Node” notebook has more pages than the “View or Modify Properties Frame” notebook. To see the specific page, click on the “page” tab on the notebook.

In general, the “View or Modify Properties Frame” notebook provides information on relationships between the Enterprise Server and the CWS. On the other hand, the “View or Modify Properties Node” notebook provides information on the Enterprise Server itself.

Other than checking configuration information and controlling hardware, the Hardware Perspective provides you with the ability to monitor system conditions on the Enterprise Servers.

You can monitor multiple conditions in one pane and see the aggregate monitoring state. The pane shows you a visual indication of what you are monitoring. To monitor system conditions, click on one of the Node icons in the Nodes pane. Then click on the Monitor icon. You can monitor as many conditions as you want. To monitor the AIX Error Log, the /tmp file system, and the /var file system, for example, select these three conditions. Then click on the Apply button.

2. Event Perspective: Create and monitor system events.

   Using the Event Perspective, you can create event definitions that let you know automatically when resource changes that are important to you have occurred in your CES system. An event is the result of a change in the state of a resource on your CES system. Examples of resources include Enterprise Servers, disk drives, memory, software applications, and file systems.

   Using the Event Perspective, you can be notified when an Enterprise Server goes down or becomes unreachable, or when the operating system is close to running out of paging space, or when there is something else wrong. More importantly, you can know these things have occurred without having the Event Perspective running.

   There are 20 default event definitions provided. They are shown on Table 10-4.

Table 10-4 PSSP Event Perspective - Event Definitions for CES

EVENT	DEFINITION
LCDhasMessage	The Enterprise Server’s LED or LCD contains a message.
errLog	A permanent error entry has been added to the AIX Error Log.
fileSystems	One of the file systems is over 90 percent full.
frameControllerNotResponding	The frame controller is not responding.
framePowerOff	The power to the frame has been turned off.
hostResponds	The node is not responding.

EVENT	DEFINITION
keyNotNormal	The key mode switch on a node was taken out of the Normal position.
nodeEnvProblem	The environment indicator LED on the node is illuminated. A hardware problem has been detected.
nodeNotReachable	Group Services has found no way to communicate with the node. The node is presumed to be down.
nodePowerDown	The power to the node is off.
nodePowerLED	Node power is off (power LED != 1).
nodeSerialLinkOpen	The serial link to the node (TTY) is open.
pageSpaceLow	The paging space utilized on the node exceeds 85 percent.
processorOffline	One or more processors may have been taken off-line.
sdrDown	The sdr daemon on the control workstation has died.
switchNotReachable	The switch adaptor on the node is not responding or the node is isolated.
switchPowerLED	Switch power is off.
switchResponds0	The switch adapter, CSS0-port 0, on the node is not responding, or the node is isolated.
tmpFull	The file system for LV=hd3 and VG=rootvg is over 90 percent full.
varFull	The file system for LV=hd9var and VG=rootvg is over 90 percent full.

Because of hardware limitations of CES systems, the following default definitions are not useful:

- framePowerOff
- keyNotNormal
- nodeEnvProblem
- nodePowerLED
- switchNotReachable
- switchPowerLED
- switchResponds0

If you are interested in some of these event definitions, you can use them by registering them. For example, if you are interested in checking that all the Enterprise Servers in your CES system are up and running, you can use the hostResponds event definition.

To register the hostResponds event definition, click the hostResponds event definition icon in the Event Definitions pane. Then click the Register icon on the tool bar. The hostResponds event definition icon changes into one with four colors.

When the event occurs, the event definition icon changes into an envelope icon. At the same time, the Global View of Event Notification Log window, will pop up.

When you do not need to manage the event definition that you selected, you can unregister the event definition. Click the hostResponds event definition icon in the Event Definitions pane. Then click the Unregister icon on the tool bar. The icon changes back to grey. You will not be notified if the event occurs.

3. Performance Monitor Perspective

   Performance Monitor Perspective may set up performance and monitoring hierarchies and archiving on your CES.

10.4 Building SP/CES systems scenarios

We may consider different scenarios of building your SP/CES system. Some of them are:

- Standard SP system: This is the case when you want to develop an SP system using the standard thin, wide and/or high SP nodes.
- Adding SP-attached servers to an SP system: This is the case when you have an SP system and you want to extend it using Enterprise Servers. If the Enterprise Servers are new servers, you install your SP-attached servers from scratch. If the Enterprise Servers are existing servers, you migrate them to SP-attached servers.
- CES system: This is the case when you want to manage Enterprise Servers from a single point of control. Therefore, you add a CWS to your Enterprise Servers and install PSSP on all of them. If the Enterprise Servers are new servers, you install your CES system from scratch. If the Enterprise Servers are existing servers, you migrate them to CES.
- Migrating an existing SP system to a CES system: This is the case when you have an SP system with SP-attached servers. You want use all of your SP frames for other SP systems. Therefore, you remove all of your SP frames from your SP system and manage it as a CES system.

- Scaling up an existing CES system to an SP system: This is the case when you have a CES system and you want to make it a larger scale SP system. Therefore your Enterprise Servers will become SP-attached servers and they will beneficiate of the advantages of the servers attached to the SP system.

For details on installation and migration on SP/CES systems, see *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347.

10.4.1 Standard SP system

If you want to develop an SP system using the standard thin, wide and/or high SP nodes, use the steps described in Chapter 2, “Installing and Configuring a New RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347 to install this type of system.

10.4.2 Adding SP-attached servers to an SP system

The PSSP software is manages the SP-attached server as though it is a node in a frame of its own. However, the SP-attached server is treated differently than an SP frame and has the following characteristics:

- It is the only node in its frame. the SP-attached server begins in slot number 1 and occupies the full 16 slots in the frame, using slot number 1 to calculate its node number. Therefore, 16 is added to that node number to derive the next available node number. However, the next node number assigned is dependent on the frame number of the next node.
- It connects to an available switch port number of an existing SP frame. An available switch port number is defined as a switch port number not used by any other node. Note that the switch port number associated with a slot may still be available if the slot is occupied by a node. For example, a high node occupies four slots but only uses the switch port number associated with the lowest slot it occupies. Any switch port number associated with the other three slots is available.
- It cannot be within a switch capsule (between a switched frame and any non-switched expansion frame within a capsule). Give it a frame number that fits before, between, or after switch capsules.
- The SP-attached server can be configured with the SP Switch (it is not supported with the SP Switch2).

Assume an SP system environment including a SP Switch capsule composed of frame A (which hosts the SP Switch), frame A+1 and frame A+2), and a second SP Switch capsule composed of frame B (which hosts the SP Switch) and frame B+1.

If you want to add one SP-attached server (as frame C) to the SP system, there are three possible positions that you can place the SP-attached server in. These positions are the followings:

- Before frame A (1 =< C < A)
- Between frame A+2 and frame B (A+2 < C < B)
- After frame B+1 (B+1 < C =< 128)

The SP-attached server that you add can connect to any available switch port of an existing SP frame. In this example, the SP-attached server can connect to an SP Switch board in either frame A or frame B. It is not dependent on where you have placed the SP-attached server.

The SP-attached servers can be new Enterprise Servers or existing Enterprise Servers.

SP-attached servers as new Enterprise Servers

If the SP-attached servers are new Enterprise Servers, you install your SP-attached servers from scratch. Your Enterprise Servers are either newly obtained or you do not care about the software previously installed on them. Use the steps described in Chapter 2, “Installing and Configuring a New RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347 to install this type of system.

For example, we consider to attach two M/T 7017 Enterprise Servers (frames 5 and 6) and one M/T 7026 Enterprise Server (frame 7).

The first server has the following characteristics:

- Frame Number: 5
- tty port for operator panel connection: /dev/tty4
- tty port for serial terminal connection: /dev/tty5
- switch port number: 14

The second server has the following characteristics:

- Frame Number: 6
- tty port for operator panel connection: /dev/tty6
- tty port for serial terminal connection: /dev/tty7
- switch port number:15

The third server has the following characteristics:

- Frame Number: 7

- tty port: /dev/tty8
- switch port number:16

In order to initiate “Step 29: Enter Frame Information and Reinitialize the SDR,” described in Chapter 2, “Installing and Configuring a New RS/6000 SP System” of the IBM Parallel System Support Programs for AIX: Installation and Migration Guide, GA22-7347, you would use the **spframe** command as follows:

```
spframe -r no -p SAMI -n 14 -s /dev/tty5 5 1 /dev/tty4
spframe -r no -p SAMI -n 15 -s /dev/tty7 6 1 /dev/tty6
spframe -r yes -p CSP -n 16 7 1 /dev/tty8
```

In our example, we specified the *-n* option for each series of contiguous SP-attached servers; the *-r* yes option when running the command for the final series of frames; included the starting frame number, the number of frames, the starting tty port value, and the starting switch port number, for each invocation of the command.

SP-attached servers as existing Enterprise Servers

If the SP-attached servers are existing Enterprise Servers, you have to migrate these existing servers to SP-attached servers. So you have to preserve the current content of these servers that are currently used for your business.

Performing the steps described in Chapter 6, “Reconfiguring the RS/6000 SP system” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347, you can attach your Enterprise Servers to an SP system without reinstalling their software. These steps preserve your existing software environment.

The following steps are excerpted from the IBM Parallel System Support Programs for AIX: Installation and Migration Guide, GA22-7347 (section “Adding Nodes” in Chapter 6). These are the steps that you have to pay special attention to:

- Step 9: Set up nodes to be installed: In this step, you must set the node to “customize” instead of “install.” For example:

  ```
  spbootins -r customize -l 33
  ```

- Step 18: Network Boot Optional Boot/Install Servers: In this step, you must proceed with the following operations:
  - Step 18.1: Upgrade AIX: If your Enterprise Server is not at AIX 4.3.3, you must first upgrade to that level of AIX before proceeding.
  - Step 18.2: Set up name resolution of the enterprise server: In order to execute a PSSP customization, the following must be resolvable on the Enterprise Server:

- The CWS host name.
- The name of the boot/install server's interface that is attached to the Enterprise Server's en0 interface (SP administrative LAN interface).

– Step 18.3: Set up routing to the CWS host name: If you have a default route set up on the Enterprise Server, you will have to delete it. If you do not remove the route, customization will fail when it tries to set up the default route defined in the SDR. In order for customization to occur, you must define a static route to the CWS's host name. For example, if the CWS's host name is its token ring address, such as 9.114.73.76, and your gateway is 9.114.73.254, execute the following:

```
route add -host 9.114.73.76 9.114.73.254
```

– Step 18.4: FTP the SDR_dest_info file: During customization, certain information will be read from the SDR. In order to get to the SDR, you must FTP the /etc/SDR_dest_info file from the CWS to the /etc/SDR_dest_info file on the Enterprise Server and check the mode and ownership of the file.

– Step 18.5: Verify perfagent: Ensure that perfagent.tools 2.2.32.x is installed on the Enterprise Server.

– Step 18.6: Mount the pssplpp directory: Mount the /spdata/sys1/install/pssplpp directory on the boot/install server from the Enterprise Server. For example, issue:

```
mount k3n1:/spdata/sys1/install/pssplpp /mnt
```

– Step 18.7: Install ssp.basic: Install ssp.basic and its prerequisites onto the Enterprise Server. For example, issue:

```
installp -aXgd/mnt/PSSP-3.2 ssp.basic 2>&1 | tee /tmp/install.log
```

– Step 18.8: Unmount the pssplpp directory: Unmount the /spdata/sys1/install/pssplpp directory on the boot/install server from the Enterprise Server. For example, issue:

```
umount /mnt
```

– Step 18.9: Run pssp_script: Run the `pssp_script` by issuing:

```
/usr/lpp/ssp/install/bin/pssp_script
```

– Step 18.10: Reboot: Perform a reboot. For example:

```
shutdown -Fr
```

10.4.3 CES system

If you want to manage Enterprise Servers from a single point of control, you have to integrate these servers into a CES system, adding a CWS to your servers and install PSSP on all of them. Within the CES system, the Enterprise Servers can be configured to use the SP Switch (the SP Switch2 is not supported).

The Enterprise Servers can be new Enterprise Servers or existing Enterprise Servers.

CES system with new Enterprise Servers

In this case, you install your CES system from scratch. Your Enterprise Servers are either newly obtained or you do not care about the software previously installed on them. Use the steps described in Chapter 2, “Installing and Configuring a New RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347 to install this type of system.

For example, we consider the installation of a CES system with one CWS and two M/T 7017 Enterprise Servers (SAMI port: /dev/tty0 and S1 port: /dev/tty1 on frame 1, respectively SAMI port: /dev/tty2 and S1 port: /dev/tty3 on frame 2).

In order to initate “Step 29: Enter Frame Information and Reinitialize the SDR” described in Chapter 2, “Installing and Configuring a New RS/6000 SP System” of the *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347, you would use the **spframe** command for our example as follows:

```
# spframe -p SAMI -r yes 1 2 /dev/tty0
```

With this command, you added two Enterprise Servers as frames 1 and 2 and initialized the SDR.

You can assign any frame number to a frame as long as the number is between 1 and 128, inclusively. It is not necessary to start from frame 1. Also, the sequence of frame number assignment does not matter.

For example, if you have three Enterprise Servers, you can add them to your CES system one by one, in any order. You could add frame 28 first, frame 9 second, and then frame 57 third.

Instead of using the **spframe** command, you can use SMIT menus for these operations.

If you are considering migrating your CES system to an SP system in the future, you have to use the -n flag on the **spframe** command and consider switch port numbers for that environment.

CES system with existing Enterprise Servers

In this case, you have to migrate these existing servers to CES. So, you have to preserve the current content of these servers that are currently used for your business.

Performing the steps described in Chapter 6, “Reconfiguring the RS/6000 SP system” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347, you can migrate your Enterprise Servers to a CES system without reinstalling their software. These steps preserve your existing software environment.

You have to pay attention to Step 9 and Step 18.

Planning consideration: By migrating your Enterprise Servers to a CES system, you can manage them effectively and easily. However, you need to consider other factors that will influence your migration plans:

- A CWS must run AIX 4.3.3 or later and PSSP 3.2 or later. We recommend: AIX 4.3.3 with ML-08 (or later) and PSSP 3.2. with APAR IY16350 (or later).
- All the Enterprise Servers in a CES system must be installed with one of the combinations of AIX and PSSP that are supported for coexistence. We recommend AIX 4.3.3. with ML-06 (or later) and PSSP 3.2. with APAR IY13026 (or later) for M/T 7017 systems, respectively AIX 4.3.3 with ML-08 (or later) and PSSP 3.2. with APAR IY16350 (or later) for M/T 7026 systems.

If you are considering migrating your CES system to an SP system in the future, you have to use the *-n* flag on spframe command and to consider switch port numbers for that environment.

10.4.4 Migration an existing SP system to a CES system

This is the case when you have an SP system with SP-attached servers and you want to use all of your SP frames for other SP systems. Therefore, you remove all of your SP frames from your SP system and manage the remaining SP-attached servers as a CES system.

Assume your SP system has three SP frames (frames 1, 4, and 6) and three SP-attached servers (frames 2, 3, and 5). You would like to remove all of the SP frames from your SP system and create a CES system with all of the SP-attached servers.

An important point to know here is that you have to remove all the SP frames at the same time. You are not allowed to remove them one by one. To remove all the SP frames from your SP system at the same time, use the **spdelfram** command with the *-l* flag as follows:

```
# spdelfram -c -l 1,4,6
```

In this example, you removed the SP frames (frames 1, 4, and 6) at the same time.

Planning consideration

By migrating your SP system to a CES system, you can manage your Enterprise Servers effectively and easily. However, you need to consider other factors that will influence your migration plans:

- A CWS must run AIX 4.3.3 or later and PSSP 3.2 or later. Before you do this migration. We recommend: AIX 4.3.3 with ML-08 (or later) and PSSP 3.2. with APAR IY16350 (or later).
- Each of the Enterprise Servers in the CES system must have one of the combinations of AIX and PSSP that are supported for coexistence installed on them. We recommend AIX 4.3.3. with ML-06 (or later) and PSSP 3.2. with APAR IY13026 (or later) for M/T 7017 systems, respectively AIX 4.3.3 with ML-08 (or later) and PSSP 3.2. with APAR IY16350 (or later) for M/T 7026 systems.
- You need to consider the dependencies and limitations that exist between applications. Some specific SP software are not supported with CES systems, so any applications that are dependent on this specific SP software must be removed.

10.4.5 Scaling up an existing CES system to an SP system

Note that, while planning your CES system, it can become a larger scale SP system in the future. However, with future growth, you might eventually add SP frames to your CES system to make it an SP system. Your system will then be subject to all the rules of an SP system, and your Enterprise Servers will become SP-attached servers.The SP Switch only is supported on CES systems and on SP-attached Servers. If you intend to scale your existing CES system to an SP system, plan your CES system with appropriate frame numbers and switch port numbers, as if you are planning for SP-attached servers with SP frames, so that you can migrate to an upscaled SP system in the future without having to totally reconfigure your existing CES system.

Adding SP frames to a CES system

This section discusses upgrading an existing CES system to an SP system. There are similar conditions when you add SP frames to a CES system. Remember, when you add SP frames to a CES system, the system is no longer a CES system. It is an SP system.

Assume your CES system has two Enterprise Servers (frame A and frame B), and you want to add one SP frame with an SP Switch (frame C) and one non-switched expansion frame (frame C+1) that is in the same switch capsule. In this case, there are three possible positions that you can place the SP frames (one switch capsule) in:

- Before frame A (1 =< C, C+1 < A)
- Between frame A and frame B (A < C, C+1 < B)
- After frame B (B < C, C+1 =< 128)

The SP-attached servers can each connect to any available switch port of an existing SP frame. In this case, SP-attached servers can connect to the SP Switch board in frame C. It does not matter where you have placed the SP frames.

When adding SP Switch adapters using the **spadaptrs** command with the *-n* yes flag (this is the default), the starting IP address may not necessarily be the lowest IP address in the system due to the new configurations allowed in an SP system from the addition of CES support to PSSP. Be sure to use the IP address of the lowest numbered node in the system as opposed to the lowest IP address in the system.

Assume your SP system has two SP-attached servers (frame 1 and 2) and one SP frame with an SP Switch (frame 3). The two SP-attached servers use switch port numbers 11 and 15. One high node is placed in slot 1 on the SP frame and uses switch port number 0.

Now, you want to assign IP addresses to the SP Switch adapters as shown in Table 10-5. Note that the IP addresses assigned won't match those given in Table 10-5.

Table 10-5 Node number and IP address

Node number	IP address	Netmask
1	129.33.34.11	255.255.255.0
17	129.33.34.12	255.255.255.0
33 (high node)	129.33.34.13	255.255.255.0

In this case, the lowest IP address is node 1. However, you cannot assign IP addresses to SP Switch adapters as follows:

```
spadaptrs 3 1 3 css0 129.33.34.13 255.255.255.0
```

Instead, use the IP address of the node with the lowest node number as follows:

```
spadaptrs 1 1 3 css0 129.33.34.11 255.255.255.0
```

The potential problem of changing from a CES system to an SP-attached system is that the frame numbers configured in the CES system might duplicate those used in the SP system. In such a case, you have to re-configure the frame numbers and switch port numbers of all the SP and SP-attached nodes such that the SP environment can be rebuilt. To avoid this problem, it is recommended to reserve some frame numbers and switch port numbers for possible future SP attachment.

For example, in Figure 10-17 on page 353, you have an CES system with frame 1 and frame 2 assigned with switch port numbers 0 and 1 respectively. If you add a *switchless* SP frame in the future, you have to assign the frame number 3 to the SP, and then you have to switch port numbers 32-47 in the SP frame according to the switch port numbering methodology.

You might find that the enterprise servers cannot use the switch port number 0 and 1 anymore since they are changed to SP-attached servers and must use the switch port numbers in the SP system. However, switch port number 0 and 1 do not exist.

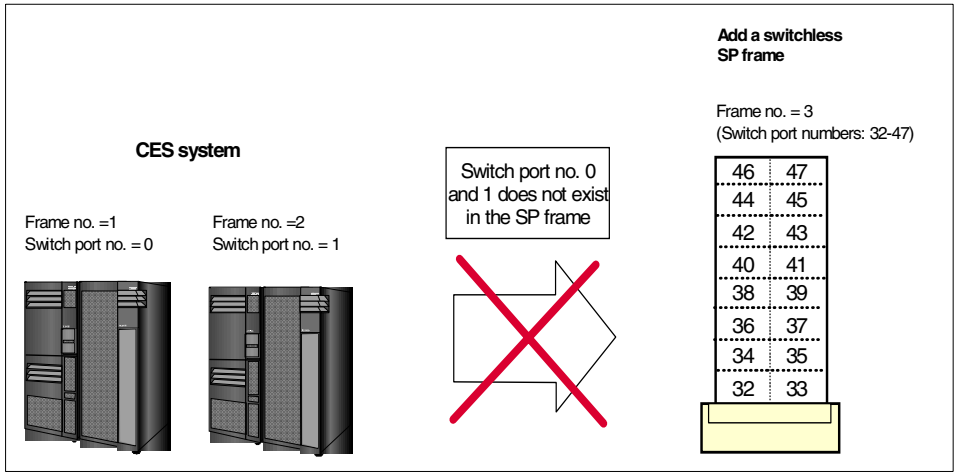


Figure 10-17 Problem for CES to SP-attached migration

To make the migration possible, you have to reconfigure the frame numbers of the enterprise servers to a larger number, for example 2 and 3, such that the SP

frame can be assigned the frame number 1. In this case, the enterprise servers can be SP-attached to the SP with switch port number 0 and 1, as long as slot 0 and 1 are available. However, in this way, reconfiguration of the CES numbering is required and if in a large scale CES (up to 32 nodes), the reconfiguration may require lots of effort. The reconfiguration is shown in Figure 10-18.

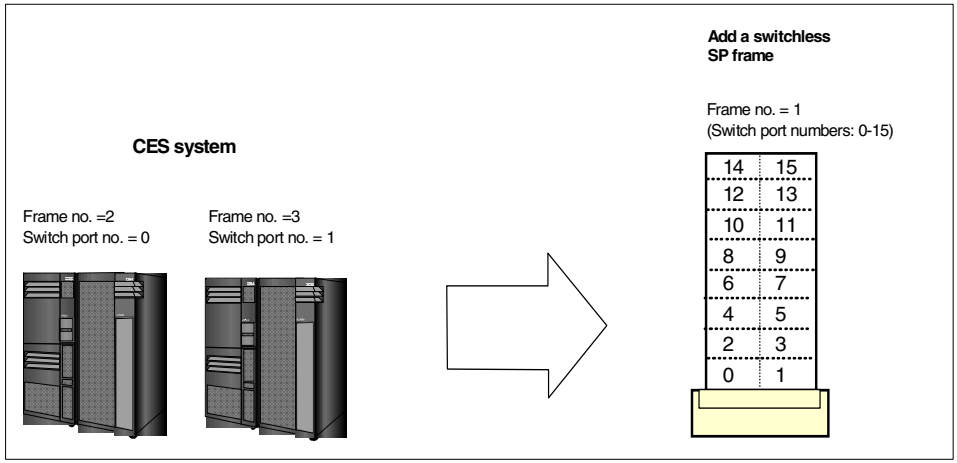


Figure 10-18 Reconfigure the CES numbers to fit into an SP

Another way to solve the problem is to add an SP switch. If you have an SP switch, you can add the enterprise servers to use any available switch port numbers available in the SP switch regardless to the frame numbers.

Therefore, you can see an appropriate numbering in a CES can prevent potential problems in future planning.

For CES numbering, there are also some recommendations provided here to provide greater flexibility:

1. For a CES system with 16 or less enterprise servers, start the frame number larger than 1. For a CES system with 17-32 nodes, start the frame number larger than 2.

   As described previously, if you have a 16-node CES system, you should start the frame number with at least 2. This helps avoid the switch port numbering problem (described above) when you add an switchless SP frame in the future. Since a switch frame can provide up to 16 switch port numbers, you have to install a second switch frame if you have more than 16 enterprise servers. In this case, it may be better to skip the frame number 2 also when assigning frame numbers to the CES from the beginning. Sometimes, if SP nodes are installed in the SP frame, the starting frame number for CES may be larger than 2.

2. If the CES system has less than 32 enterprise servers and add SP frames, leave space for flexible SP node addition.

   For example when a CES with 15 enterprise servers is redefined to be SP-attached servers in the SP system, the servers occupy at least 15 switch port numbers. If you reserve numbers 0-14 to the enterprise servers, only 15 is left. If in future you want to add an SP node, you can use this number only. With this switch port number 15, only one thin node is allowed. Therefore, to provide more flexibility for the addition SP node, it is a good practice to skip switch port number 12 which can allow a thin node, a wide node, or a high node to be added to the SP system. One of these nodes is possible to be hosted in node slot number 13, as shown in Figure 10-19.

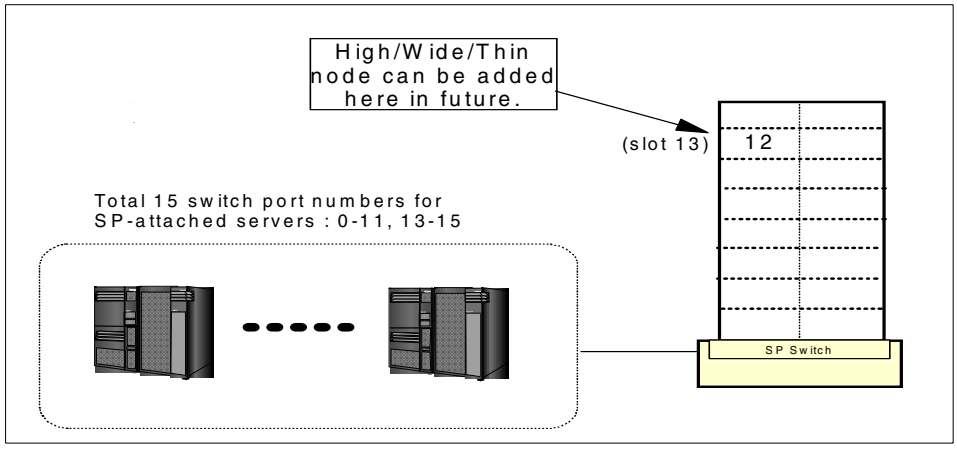


Figure 10-19 Changing a 15-node CES system to an SP system

The same practice can be applied to a CES with more than 16 and less than 32 nodes. Another example is shown in Figure 10-20 on page 356. To change a CES to an SP-attached system with 31 enterprise servers, a minimum of two SP frames are required, since each SP frame can provide up to 16 switch port numbers only. In this example, you may also skip the switch port number 12 or 28 such that a high, wide or thin node can be added in future.

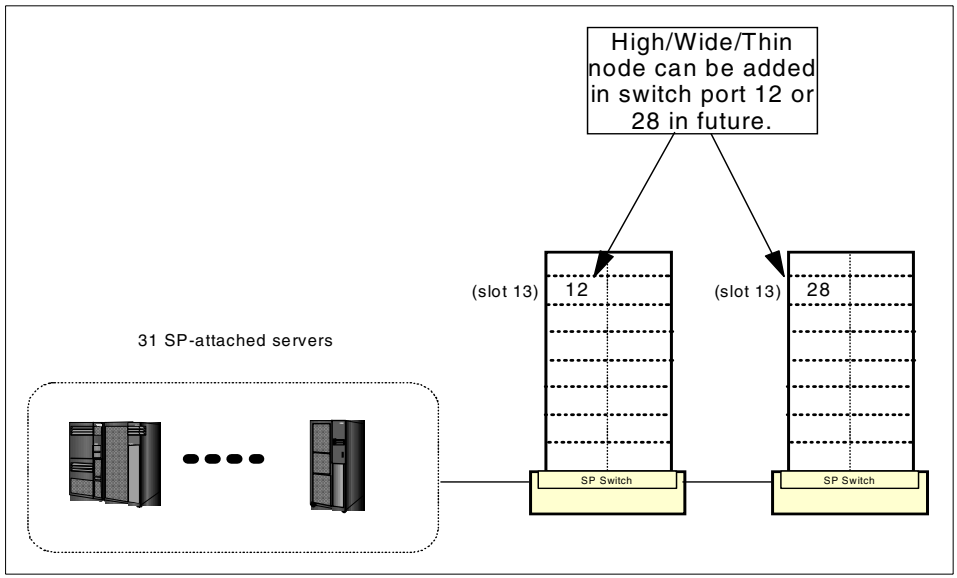


Figure 10-20 Changing a 31-node CES system to an SP system

Recommended expansion plan

If you are thinking of scaling up your CES system to an SP system in the future, you have to plan out appropriate frame numbers and switch port numbers for your Enterprise Servers so that they will still be valid when you add SP frames at a later time.

Attention: You have to plan the appropriate frame and switch port numbers that will work in an SP system with SP frames when you first install your CES system. This way you can migrate to an SP system without having to reconfigure your Enterprise Servers that are already installed.

There can be many expansion plans. An expansion plan could be considered when you plan to upscale from a CES system to a switchless SP system. Another expansion plan could be considered when you plan to upscale from a CES system to a switched SP system

Both expansion plans take in consideration numbers such as:

- Size of CES
- Frame/Node Type
- Frame Number
- Node Number

- Switch Port Number

For details on recommended expansion plans on CES systems to SP systems, see *IBM RS/6000 Clustered Enterprise Servers Systems Handbook*, SG24-5978.

10.5 pSeries 680 Cluster Solutions

Entry and advanced cluster ready solution packages are available and they provide enhanced scalability for the pSeries 680. These solution packages combine one pSeries 680 with two entry or mid-range servers and the cabling necessary to incorporate them into a new or existing cluster. The clusters operate under the control of IBM Parallel Systems Support Programs (PSSP) for AIX.

The servers in the cluster ready solutions can be used in a non-switched cluster environment or in an SP switch attached environment. When used in an SP switch attached environment, each server must be equipped with an SP Switch Attachment Adapter feature F/C 8396.

The entry or mid-range servers are mounted in the first I/O rack of the pSeries 680 at the Customer Solutions Center in Rochester, MN. World wide orders for the entry and advanced Cluster Server Solutions will be fulfilled from Rochester.

AIX 4.3.3 with maintenance package 4330-08 or later is required for each clustered server. AIX will be pre-loaded on each server in the cluster ready solution at the factory.

IBM PSSP for AIX version 3.2 with APAR IY16350 or later is required for each server in the cluster. PSSP will be shipped separately and installed by the customer.

Each server in the cluster is controlled via a customer supplied control workstation and requires a customer supplied ethernet LAN connection. Additional information regarding server clustering and control workstations is available at the following IBM PSSP Web site:

`http://www.rs6000.ibm.com/resource/aix_resource/sp_books/`

10.5.1 pSeries 680 Entry Cluster Server - S85/6H1

Each pSeries 680 Entry Cluster Server Solution order MUST INCLUDE a MINIMUM of the following hardware and software that must be ordered together, as specified, to qualify for the solution package:

- One - Complete pSeries 680 Server (7017-S85), incorporating the following:
  - One - Software Preinstall (F/C 5005)
  - One - Entry Cluster Solution Indicator (F/C 0701)
  - Two - SCSI Boot Disks
  - One - ASYNCH adapter for Control Workstation Attachment
  - One - Clustered Server Serial Port to CWS cable (F/C 3150)
  - One - Clustered Server Control Panel to CWS cable (F/C 3151)
  - One - IBM 10/100 Mbps Ethernet PCI Adapter (F/C 2968)
  - One - PSSP Version 3.2 or later license
  - One - AIX 4.3.3 or later license
  - Two - Rack Content Specify: 7026-6H1, 5U (F/C 0188)
  - Two - Rack Content Specify: FC 6324, 5EIA (F/C 0176)
  - One - Customer Service Specify - Rochester (F/C 5001)

  Note: Additional optional pSeries 680 features may be added as desired.
- Two - Complete pSeries 660 Model 6H1 (7026-6H1 Servers), each incorporating the following:
  - One - Entry Cluster Solution Indicator (F/C 0701)
  - Two - SCSI boot disks
  - One - Clustered Server Control Panel to CWS cable (F/C 3151)
  - One - Cable and PCI Card for SP CWS Attachment (F/C 3154)
  - One - IBM 10/100 Mbps Ethernet PCI Adapter (F/C 2968)
  - One - PSSP Version 3.2 or later license
  - One - AIX 4.3.3 or later license
  - One - Customer Service Specify - Rochester (F/C 5001)
  - One - Software Preinstall (F/C 5005)

Note: Additional optional 7026-6H1 features may be added as desired.

A pSeries 680 Entry Cluster Server is illustrated in Figure 10-21.

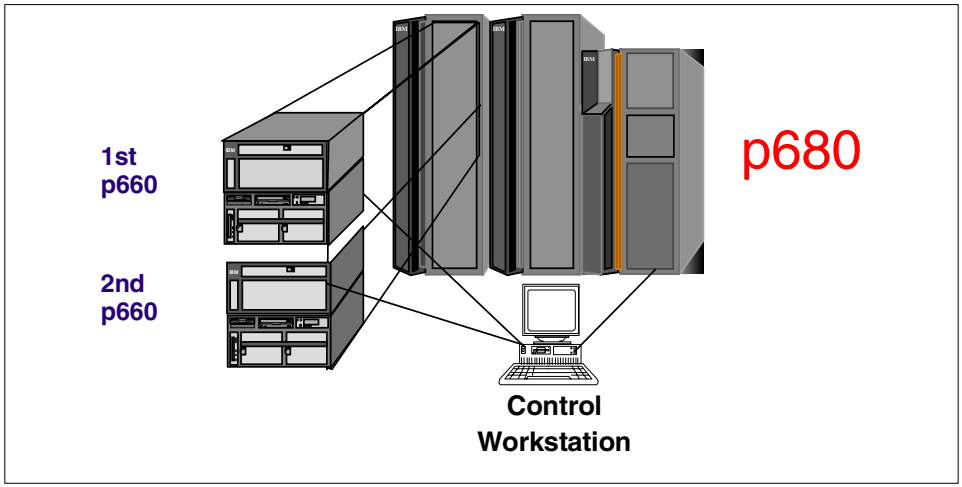


Figure 10-21 pSeries 680 Entry Cluster Server

Note: A similar configuration is offered based on a 7017-S85 Server with two 7026-6H0 servers.

10.5.2 pSeries 680 Advanced Cluster Server - S85/M80

Each pSeries 680 Advanced Cluster Server Solution order MUST INCLUDE a MINIMUM of the following hardware and software that must be ordered together as specified to qualify for the solution package:

- One - Complete pSeries 680 (7017-S85) Server, incorporating the following:
  - One - Advanced Cluster Solution Indicator (F/C 0702)
  - Two - SCSI Boot Disks
  - One - ASYNCH adapter for Control Workstation Attachment
  - One - Clustered Server Serial Port to CWS cable (F/C 3150)
  - One - Clustered Server Control Panel to CWS cable (F/C 3151)
  - One - IBM 10/100 Mbps Ethernet PCI Adapter (F/C 2968)
  - One - PSSP Version 3.2 or later license
  - One - AIX 4.3.3 or later license
  - Two - Rack Content Specify: 7026-M80, 8U (F/C 0129)
  - Two - Rack Content Specify: FC 6324, 5EIA (F/C 0176)
  - One - Customer Service Specify - Rochester (F/C 5001)

– One - Software Preinstall (F/C 5005)

Note: Additional optional pSeries 680 features may be added as desired.

- Two - Complete 7026-M80 Servers, each incorporating the following:
  - One - Advanced Cluster Solution Indicator (F/C 0702)
  - Two - SCSI boot disks
  - One - Clustered Server Control Panel to CWS cable (F/C 3151)
  - One - Cable and PCI Card for SP CWS Attachment (F/C 3154)
  - One - IBM 10/100 Mbps Ethernet PCI Adapter (F/C 2968)
  - One - PSSP Version 3.2 or later license
  - One - AIX 4.3.3 or later license
  - One - Customer Service Specify - Rochester (F/C 5001)
  - One - Software Preinstall (F/C 5005)

Note: Additional optional 7026-M80 features may be added as desired.

A pSeries 680 Advanced Cluster Server is illustrated in Figure 10-22.

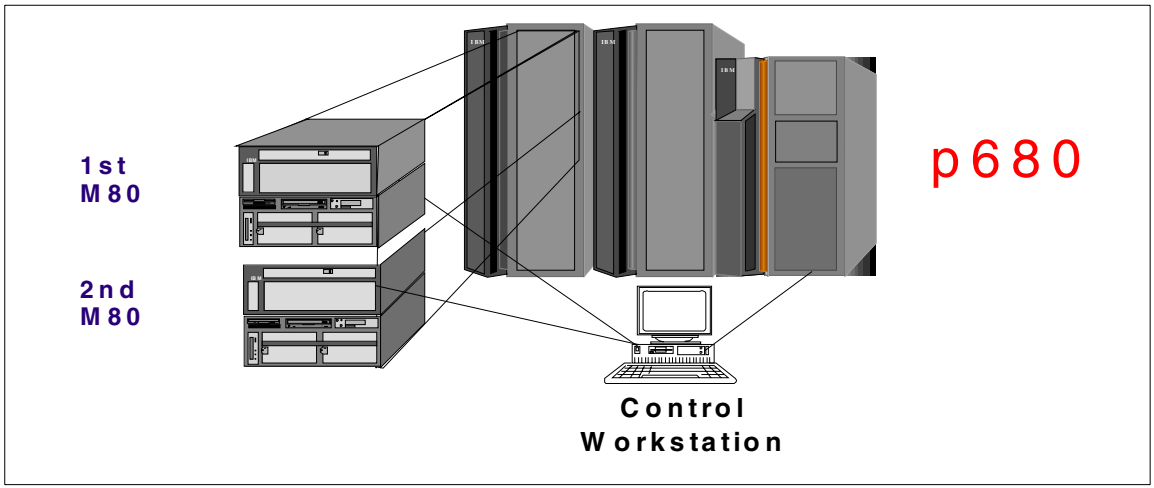


Figure 10-22 pSeries 680 Advanced Cluster Server

A

Naming convention

The naming convention appendix provides naming conventions used in the contents of the redbook. This appendix is formatted in the following manner:

- Overview
  - Naming convention

Definitions overview and naming convention

IBM announced changes in product names with the arrival of the new IBM @server pSeries Servers. It is important to bear in mind that this redbook uses the new naming convention.

Naming convention

Table A-1 shows the different combinations of SP/CES names.

Table A-1 SP/CES logical names

Solutions :	**SP**	**SP-aAttached server**	**Clustered Entreprise Servers, or CES**
Component : CWS	CWS	CWS	CWS
Frame	Frame	Server frame	Server Frame
Node	Internal Node	External Node *	Server or Node*
GRF 9077 Router	SP Switch	SP Switch	Not available
Login	Into a SP system	Into a SP system	Into a CES system
System naming configurations	SP systems	SP systems	CESs

(*) The external node or Server can be one of the following machine or associated machine type or nickname described in Table A-2.

Table A-2 Machine names, M/T and Nicknames

Node name	**Machine type (M/T)**	**Nickname used here**
IBM @server pSeries 680 Model S85	7017-S85	S85
IBM @serverpSeries 660 Model 6H1	7026-6H1	6H1
IBM RS/6000 7026 model H80	7026-H80	H80
IBM RS/6000 7026 model M80	7026-M80	M80
IBM RS/6000 7017 model S80	7017-S80	S80
IBM RS/6000 7017 Model S70 Advanced	7017-S7A	S7A

Node name	Machine type (M/T)	Nickname used here
IBM RS/6000 7017 Model S70	7017-S70	S70

B

Hardware

This appendix includes information regarding:

1. SP Communication adapter requirements
2. Evaluating SP System power requirement
3. Legacy SP Node
4. Legacy Models by Frame and Switch Types
5. Legacy Model by Frame and Nodes Types
6. SP LAN Topologies
7. Boot/install Server requirements
8. Disk space required for PSSP file sets

SP Communication adapter requirements

The SP Communication adapter requirements Annex contains information on the quantities of adapters of each type that can be installed in specific SP node types, suggested quantities to use, and adapter placement rules and restrictions.

PCI bus I/O adapter requirements for SMP nodes

If you install the maximum permissible quantity of adapters in the SP nodes, you can achieve maximum connectivity for the nodes in your system. Use Table B-2 on page 368 and Table B-3 on page 370 to determine the maximum quantity of adapters you can use and to determine the resulting power load.

Improved I/O performance, or optimum operation, can usually be realized if you install fewer than the maximum quantity of adapters. For guidelines to help you achieve this improved performance, see “Suggested PCI adapter quantities for optimum operation” on page 376.

To effectively use the information in this chapter, you need to know the PCI bus and slot naming conventions; for these, see “PCI bus group and slot descriptions” on page 366.

For important installation restrictions for specific adapters, see PCI adapter plugging rules and restrictions.

PCI bus group and slot descriptions

This section contains details on PCI bus group and slot naming and labeling conventions for the RS/6000 SP nodes.

PCI bus slot labeling

In both POWER3 and 332 MHz wide nodes, two slots are labeled I2 and two are labeled I3. They are differentiated in this book as ***I2 and I3 (CPU side)*** and as ***I2 and I3 (I/O side***).

The CPU side is on the left and the I/O side is on the right, seen as you face the side of the node having the slot labels.

Descriptions of PCI bus group numbering, bus type, slot width and slot speed for POWER3 and 332 MHz SMP nodes are shown in Table B-1.

Table B-1 PCI bus description for SMP nodes

Node type	Bus group	Bus type	Slot numbers	Slot width	Slot speed
375 MHz POWER3 and POWER3 High	0 0 1	Primary Primary Primary	I1 I2 and I3 I4 and I5	32-bit 64-bit 64-bit	33 MHz 33 MHz 33 Mhz
SP Expansion I/O	0-3	Primary	I1 to I8	64-bit	33 MHz
375 MHz POWER3 and POWER3 Wide	1 2 3	Primary Primary Primary	I2 and I3 (CPU side) I1 to I4 (I/O side) I5 to I8 (I/O side)	32-bit 64-bit 64-bit	33 MHz 33 MHz 33 Mhz
375 MHz POWER3 and POWER3 Thin	1	Primary	I2 and I3	32-bit	33 MHz
332 MHz Wide	1 2 3	Primary Primary Secondary	I2 and I3 (CPU side) I1 to I4 (I/O side) I5 to I8 (I/O side)	32-bit See Note 1 32-bit	33 Mhz 33 Mhz 33 MHz
332 MHz Thin	1	Primary	I2 and I3	32-bit	33 MHz

Notes:

1. Slot I1 through I3 (I/O side) are 64-bit slots; slot I4 is a 32-bit slot.

PCI adapter maximum quantities

The maximum quantities of adapters for 375 MHZ POWER3, POWER3, and 332 MHz SMP Thin and Wide nodes are shown in Table B-2.

Table B-2 Maximum quantities of PCI adapters for SMP thin and wide nodes

F/C	PCI Adapter Name	Number of PCI slots Required	Max. Number Allowed Per Node Wide/Thin	Notes
2732	Serial HIPPI Channel	1	2/1	
2733	Serial HIPPI Channel	1	2/1	
2741	FDDI SK-NET LP SAS	1	4/2	1
2742	FDDI SK-NET LP DAS	1	4/2	1
2743	FDDI SK-NET UP SAS	1	4/2	1
2751	S/390 ESCON	1	2/1	
2920	Token Ring Auto Lanstreamer	1	8/2	
2943	RS-422 8-port Async	1	6/2	
2944	WAN RS-232 128-port	1	7/2	
2947	ARTIC960Hx 4-port selectable	1	8/2	
2962	2-port Multiprotocol Adapter	1	8/2	
2963	ATM Turboways 155 UTP	1	4/2	1
2968	Ethernet 10/100 Mb	1	6/2	1
2969	Gigabit Ethernet-SX	1	3/1	2
2985	Ethernet 10 Mb BNC	1	8/2	
2987	Ethernet 10 Mb AUI	1	8/2	
2988	ATM 155 MMF	1	4/2	1

F/C	PCI Adapter Name	Number of PCI slots Required	Max. Number Allowed Per Node Wide/Thin	Notes
4951	Four-port 10/100 LAN	1	3/1	2
4958	Crypto	1	4/0	
4959	Token Ring	1	6/2	
6204	Ultra SCSI DE	1	6/2	3
6205	Dual Channel Ultra3 SCSI	1	3/1	2
6206	Ultra SCSI SE	1	6/2	
6207	Ultra SCSI DE	1	6/2	
6208	SCSI-2 F/W SE	1	8/2	
6209	SCSI-2 F/W DE	1	8/2	
6215	SSA RAID5	1	6/2	
6222	F/W Cache Option	0	(mounts on 6215)	
6225	SSA RAID EL	1	6/2	
6227	FC-AL	1	3/1	2
6228	FC-AL 64-bit	1	3/1	2
6230	SSA RAID EL	1	6/2	
6231	128 MB DIMM Option	0	(mounts on 6225 and 6230)	
6235	Fast-Write Cache Option for F/C 6225	0	(mounts on 6225 and 6230)	
6310	ARTIC960RxD Quad Digital Trunk	1	4/2	
6311	ARTIC960RxF Digital Trunk Resource	1	3/1	3

Notes:

1. Group restrictions apply.
2. Quantity of two (2) for POWER3 Wide Nodes.
3. Used only in 375 MHz POWER3 SMP Thin and Wide nodes.

The maximum quantities for 375 MHz POWER3 High Nodes, POWER3 High Nodes, SP Expansion I/O Units, and node/expansion unit combinations are shown in Table B-3.

Table B-3 PCI adapters for high nodes and SP Expansion I/O Units

F/C	PCI Adapter Name	Max. Quantity per High node	Max. Quantity per SP Expansion I/O Unit	Combi n-ation (Note 1)	Combi n-ation (Note 2)
2732	Serial HIPPI SW	1	3	8	8
2733	Serial HIPPI LW	1	3	8	8
2741	FDDI SK-NET LP SAS	5	8	24	24
2742	FDDI SK-NET LP DAS	5	8	24	24
2743	FDDI SK-NET UP SAS	5	8	24	24
2751	S/390 ESCON	0	4	10	10
2920	Token Ring Auto Lanstreamer	5	8	24	24
2943	RS-422 8-port Async	5	0	5	5
2944	WAN RS-232 128-port	5	0	5	5
2947	ARTIC960Hx 4-port selectable	4	8	24	24
2962	2-port Multiprotocol Adapter	5	8	24	24
2963	ATM Turboways 155 UTP	3	8	16	24
2968	Ethernet 10/100 Mb	3	8	24	24
2969	Gigabit Ethernet-SX	1	2	8	12

F/C	PCI Adapter Name	Max. Quantity per High node	Max. Quantity per SP Expansion I/O Unit	Combi n-ation (Note 1)	Combi n-ation (Note 2)
2985	Ethernet 10 Mb BNC	5	8	24	24
2987	Ethernet 10 Mb AUI	5	8	24	24
2988	ATM 155 MMF	3	8	24	24
4951	Four-port 10/100 LAN	1	3	6	12
4958	Crypto	2	4	4	4
4959	Token Ring	3	8	24	24
6204	Ultra SCSI DE	3	8	24	24
6205	Dual Channel Ultra3 SCSI	1	2	6	8
6206	Ultra SCSI SE	3	8	24	24
6207	Ultra SCSI DE	3	8	24	24
6215	SSA RAID5	2	4	24	24
6222	F/W Cache Option	(mounts on 6215)	(mounts on 6215)		
6225	SSA RAID EL	2	2	12	16
6227	FC-AL	2	3	10	14
6228	FC-AL 64-bit	2	3	10	14
6230	SSA RAID EL	2	2	12	16
6231	128 MB DIMM Option	(mounts on 6225 and 6230)	(mounts on 6225 and 6230)		
6235	Fast-Write Cache Option for F/C 6225	(mounts on 6225 and 6230)	(mounts on 6225 and 6230)		
6310	ARTIC960RxD Quad Digital Trunk	3	0	3	3

F/C	PCI Adapter Name	Max. Quantity per High node	Max. Quantity per SP Expansion I/O Unit	Combi n-ation (Note 1)	Combi n-ation (Note 2)
6311	ARTIC960RxF Digital Trunk Resource	3	0	3	3

Notes:

1. The maximum quantity of adapters supported for a specific POWER3 SMP High Node and six (6) SP Expansion I/O Units connected to that node.
2. The maximum quantity of adapters supported for a specific 375 MHz POWER3 SMP High Node and six (6) SP Expansion I/O Units connected to that node.

PCI adapter plugging rules and restrictions

This section lists specific PCI adapter installation rules and restrictions by node type.

375 MHz POWER3 SMP Wide Node:

- If one F/C 2969 is installed, no more than four (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If two F/C 2969 are installed, no more than two (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If three F/C 2969 are installed, none of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If one F/C 2969 is installed, place it in slot I1 through I8 (I/O side).
- If two F/C 2969 are installed, place one in slot I1 through I4 (I/O side) and the other in slot I5 through I8.
- F/C 2751 must be placed in either slot I3 (CPU side) or in slot I4.
- F/C 2963 and 2988 cannot be placed in slot I5.
- F/C 2732, 2733, 2969, 6205, 6225, 6227, 6228, 6230 cannot occupy the same PCI bus.
- The combined quantity of F/C 6225 and 6230 cannot exceed two per PCI bus.
- F/C 6205 cannot be placed in slot I5 to I8.

- If two F/C 2732 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- If two F/C 2733 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- Sum of F/C 6310 and 6311 cannot be more than four.
- F/C 4958 cannot be placed in slot 1 or 2 (CPU side).

375 MHz POWER3 SMP Thin Node:

- No more than four (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If one F/C 2969 is installed, no more than two (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If two F/C 2969 are installed, none of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- F/C 2751 must be placed in slot I3. (If only one of any other adapter is installed, placing the other adapter in slot I2 initially avoids moving it if you install F/C 2751 later.)
- F/C 2732, 2733, 2969, 6205, 6225, 6227, 6228, 6230 cannot occupy the same PCI bus.
- The combined quantity of F/C 6225 and 6230 cannot exceed two per PCI bus.
- Sum of F/C 6310 and 6311 cannot be more than two.

375 MHz POWER3 and POWER3 High Node:

- F/C 6204, 6206, 6207 cannot be placed in slot 1 or 2.
- F/C 6215, 6225, 6227, 6228, 6230 must be placed in slot 3 or 5.
- F/C 2947, 6310, 6311 cannot be placed in slot 1.
- F/C 2732, 2733, 2751, 2969, 6205 must be placed in slot 5.
- F/C 2963 and 2988 cannot be placed in the adjacent slot to the right of F/C 2751
- Sum of F/C 6310 and 6311 cannot be more than four.
- Only one of 2732, 2751, 2969, 6205, 6225, 6227, 6228, 6230 can be installed per PCI bus. The other slot in that bus should be empty.
- F/C 4958 cannot be placed in slot 1, 2, or 3.

SP Expansion I/O Unit:

- F/C 2732, 2733, 6227, 6228 must be placed in slot 1, 3, or 5.

- F/C 6215 must be placed in slot 1, 3, 5, or 7.
- F/C 2969, 6205, 6225, 6230 must be placed in slot 1 or 3.
- F/C 2963 and 2988 cannot be placed in the adjacent slot to the right of F/C 2751
- F/C 2751 must be placed in slot 1, 3, 5, or 7 in unit 3 or 4, or in slot 5 or 7 in unit 5.

POWER3 Wide Node:

- No more than four (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed
- If one F/C 2969 is installed, no more than two (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If two F/C 2969 are installed, none of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- F/C 2751 must be placed in either slot I3 (CPU side) or in slot I4.
- If one F/C 2969 is installed, place it in slot I1 through I8 (I/O side).
- If two F/C 2969 are installed, place one in slot I1 through I4 (I/O side) and the other in slot I5 through I8.
- F/C 2963 and 2988 cannot be placed in slot I5.
- Maximum of one F/C 2732, 2733, 2969, 6205 per PCI bus.
- Maximum of two F/C 6206, 6207, 6225, 6230 per PCI bus.
- The combined quantity of F/C 6225 and 6230 cannot exceed two per PCI bus.
- F/C 6205 cannot be placed in slot I5 to I8.
- If two F/C 2732 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- If two F/C 2733 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- F/C 4958 cannot be placed in slot 1 or 2 (CPU side).

332 MHz Wide Node:

- No more than four (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If one F/C 2969 is installed, no more than two (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If two F/C 2969 are installed, none of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- F/C 2751 must be placed in either slot I3 (CPU side) or in slot I4, except: If F/C 2969 is placed in slot I4, F/C 2751 must be placed in slot I3 (I/O side).
- If one F/C 2969 is installed, it must be placed in slot I2 or I3 (CPU side) or in slot I4.
- If two F/C 2969 are installed, place one in slot I2 or I3 (CPU side) and the other in slot I4.
- F/C 2732, 2733, 2751, 2947, 2962, 2963, 2968 (when operating in 100 Mbps mode), 2969, 2988, 6205, 6206, 6207, 6208 (when operating in SCSI–2 Fast/Wide mode), 6209 (when operating in Fast/Wide mode or to tape), 6215, 6225, 6230, 6227, 6228, 6310, 6311 cannot be placed in slot I5 through I8.
- Maximum of one F/C 2732, 2733, 2969, 6205, 6225, 6230 per PCI bus.
- If two F/C 6225 or 6230 are installed, each must be placed in a separate PCI bus in slot I2 or I3 (CPU side) or slot I1 to I4 (I/O side).
- If two F/C 2732 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- If two F/C 2733 are installed, one must be placed in slot I2 or I3 (CPU side) and the other must be placed in slot I1 to I4 (I/O side).
- Sum of F/C 6310 and 6311 cannot be more than four.
- F/C 4958 cannot be placed in slot 1 or 2 (CPU side).

POWER3 and 332 MHz Thin Node:

- No more than four (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If one F/C 2969 is installed, no more than two (in any combination) of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- If two F/C 2969 are installed, none of F/C 2741, 2742, 2743, 2963, 2968, 2988 can be installed.
- F/C 2751 must be placed in slot I3. (If only one of any other adapter is installed, placing the other adapter in slot I2 initially avoids moving it if you install F/C 2751 later.)

- Maximum of one F/C 2732, 2733, 2969, 6205 per thin node.
- The combined quantity of F/C 6225 and 6230 cannot exceed two per PCI bus.

Suggested PCI adapter quantities for optimum operation

The guidelines in this section help you configure your PCI nodes for optimum operation with good throughput for each adapter. To achieve this, each adapter type is assigned a *weighting factor*, which provides the following:

1. Estimates of the quantity of adapters to use concurrently.
2. The bus locations for the different adapter types.

The weighting factors are based on the node and PCI bus architecture, and processor and memory utilization for larger I/O read and write operations. (Larger and smaller I/O reads and writes refers to the basic I/O payload.) Nodes are assumed to have the maximum quantity of CPUs along with sufficient memory as required by a particular application.

Smaller I/O reads and writes increase the required node resources and decrease the quantity of adapters from that which is suggested for optimum operation.

The weighting factor for each PCI adapter type is shown in Table B-4.

Table B-4 PCI adapters weighting factors

F/C	Description	Type	Weighting Factor	Notes
2732	Serial HIPPI SW	Comm	High	
2733	Serial HIPPI HW	Comm	High	
2741	FDDI SK-NET LP SAS	Comm	Low	
2742	FDDI SK-NET LP DAS	Comm	Low	
2743	FDDI SK-NET UP SAS	Comm	Low	
2751	ESCON CU Emulation	Comm	Medium	
2920	Token Ring Auto Lanstreamer	Comm	Low	
2943	RS-422 8-port Async	Comm	Low	

F/C	Description	Type	Weighting Factor	Notes
2944	WAN RS-232 128-port	Comm	Low	
2947	ARTIC960Hx 4-port selectable	Comm	Low	
2962	2-port Multiprotocol Adapter	Comm	Low	
2963	ATM Turboways 155 UTP	Comm	Medium	
2968	Ethernet 10/100 Mb	Comm	Low (10) Medium (100)*	2
2969	Gigabit Ethernet-SX	Comm	High	
2985	Ethernet 10 Mb BNC	Comm	Low	
2987	Ethernet 10 Mb AUI	Comm	Low	
2988	ATM 155 MMF	Comm	Medium	
4951	Four-port 10/100 LAN	Comm	Low (no 100) Med (1-2 100) High (3-4 100)	7
4958	Crypto	Crypto	Low	
4959	Token Ring	Comm	Low	
6204	Ultra SCSI DE	Storage	High	
6205	Dual Channel Ultra3 SCSI	Storage	High	5
6206	Ultra SCSI SE	Storage	Low (F) Medium (F/W)* High (Ultra)	3

F/C	Description	Type	Weighting Factor	Notes
6207	Ultra SCSI DE	Storage	Low (F) Medium (F/W) High (Ultra)*	4
6208	SCSI-2 F/W SE	Storage	Medium	
6209	SCSI-2 F/W DE	Storage	Medium	
6215	SSA RAID5	Storage	Medium (RAID) High (Dual Loop)	
6225	SSA RAID EL	Storage	High	6
6227	FC-AL	Storage	High	
6228	FC-AL 64-bit	Storage	High	
6230	SSA RAID EL	Storage	High	
6310	ARTIC960RxD WAN DT Quad	Comm	Low	
6311	ARTIC960RxF Dig.Trunk	Comm	Low	

Notes:

1. Asterisk (*) indicates the default weighting factor used by the configurator
2. Low for 10 Mbps operation, Medium* for 100
3. Low for Fast SCSI operation, Medium* for Fast/Wide, High for Ultra
4. Low for Fast SCSI operation, Medium for Fast/Wide, High* for Ultra
5. Medium for single channel in Fast SCSI operation, High for all other single or dual channel Fast SCSI, Fast/Wide, Ultra or Dual Channel Ultra2 operation
6. Medium* for RAID single loop operation, High for non-RAID
7. Low for no ports at 100 Mb, Medium with 1 or 2 ports at 100 Mb, High for 3 or 4 ports at 100 Mb

Following are lists of weighting factor guidelines and suggested quantities of adapters (by node type, PCI bus, and weighting factor) to help you achieve optimum operation on PCI buses:

Adapter guidelines for using weighting factors

1. Distribute adapters across the PCI buses to equalize the weighting factor on each bus (application usage might require a different distribution).
2. Place adapters with high and medium weighting factors in slots on a ***primary*** PCI bus.
3. Except as suggested otherwise in this section, adapters with a low weighting factor can be placed in slots on a secondary PCI bus where allowed by "PCI adapter plugging rules and restrictions" on page 109. (Adapters on a primary PCI bus usually give better performance than on a secondary bus.)
4. F/C 2969 is a 64-bit PCI adapter, capable of operating in a 32-bit slot at lower throughput.
5. The total quantity of adapters used on a node can actually be fewer than the sum suggested below, due to system resource limitations.

POWER3 High Node

Select *one* of the following options for PCI Bus 0:

- 1 High
- 2 Medium
- 1 Medium and 2 Low
- 3 Low

Select *one* of the following options for PCI Bus 1:

- 1 High
- 2 Medium
- 1 Medium and 1 Low
- 2 Low

SP Expansion I/O Unit

Select *one* of the following options for *each* PCI Bus (0, 1, 2, and 3):

- 1 High
- 2 Medium
- 1 Medium and 1 Low

- 2 Low

POWER3 Thin and Wide Node

Select *one* of the following options for PCI Bus 1:

- 1 High
- 2 Medium
- 1 Medium and 1 Low
- 2 Low

POWER3 Wide Node

Select ***one*** of the following options for ***both*** PCI Bus 2 and PCI Bus 3:

- 1 High
- 2 Medium and 2 Low
- 1 Medium and 3 Low
- 4 Low

332 MHz Thin and Wide Node

Select **one** of the following options for PCI Bus 1:

- 1 High
- 2 Medium
- 1 Medium and 1 Low
- 2 Low

332 Mhz Wide Node

Select **one** of the following options for PCI Bus 2 and PCI Bus 3 combined:

- 1 High (place on Bus 2)
- 2 Medium (place on Bus 2) and 2 Low (should be on Bus 2)
- 1 Medium (place on Bus 2) and 4 Low (3 should be on Bus 2 with precedence given to comm-type adapters, and 1 on Bus 3)
- 6 Low (4 on Bus 2, with precedence given to comm-type adapters, and 2 should be on Bus 3)

MCA bus adapter requirements

Table B-5 contains reference information on MCA communication adapters.

Table B-5 Micro Channel Adapter requirements

F/C	MCA Adapter Name	Number of MCA slots Required	Max. Number Allowed Per Node Wide/Thin/High	Notes
1902	Optics Daughter Card	0	2/1/NA	9
1904	Fibre Channel 1 GB	1	2/2/NA	1,9
1906	Fibre Channel 266	1	2/2/NA	1,9
2402	IBM Network Terminal Accelerator 256	1	7/4/4	9
2403	IBM Network Terminal Accelerator 2048	1	7/4/4	9
2410	SCSI-2 External I/O Controller	1	7/4/NA	
2412	Enhanced SCSI-2 Differential Fast/Wide Adapter/A	1	7/4/14	
2415	SCSI-2 Fast/Wide Adapter/A	1	7/4/14	
2416	SCSI-2 Differential Fast/Wide Adapter/A	1	7/4/14	9
2420	SCSI-2 Differential External I/O Controller	1	7/2/NA	9
2700	4-Port Multiprotocol Communications Controller	1	7/3/8	
2723	FDDI Dual Ring	1	3/2/4	3
2724	FDDI SAS Single Ring	1	6/2/8	
2735	HIPPI	5	1/NA/2	4

F/C	MCA Adapter Name	Number of MCA slots Required	Max. Number Allowed Per Node Wide/Thin/High	Notes
2754	S/390 ESCON Channel Emulator Adapter	2	2/1/4	
2755	BMCA	1	2/2/2	5,6
2756	ESCON Control Unit Adapter	2	2/1/4	
2930	RS-232 8-port Adapter	1	7/4/14	
2940	8-port Async Adapter	1	7/4/14	
2960	X.25 Interface Co-Processor/2	1	7/4/8	
2970	Token-Ring	1	7/4/12	
2972	Auto Token-Ring LANstreamer MC 32	1	7/3/12	
2980	Ethernet	1	7/3/12	2
2984	ATM TURBOWAYS 100	1	2/2/2	9
2989	ATM TURBOWAYS 155	1	4/2/4	9
2992	Ethernet/FDX 10 Mbps TP/AUI MC Adapter	1	7/3/12	2
2993	Ethernet/FDX 10 Mbps BNC MC Adapter	1	7/3/12	2
2994	10/100 Ethernet Twisted Pair MC Adapter	1	2/2/4	
4224	Ethernet 10BASE-T Transceiver	0	8/4/15	
6212	9333 HPSA	1	6/2/8	7

F/C	MCA Adapter Name	Number of MCA slots Required	Max. Number Allowed Per Node Wide/Thin/High	Notes
6214	SSA 4-Port Adapter	1	4/2/8	
6216	SSA 4-Port Adapter	1	4/2/8	
6217	SSA 4-Port RAID Adapter	1	4/2/8	
6219	Micro Channel SSA Multi-Initiator/RAID EL Adapter	1	4/2/8	8
6222	SSA Fast-Write Cache Option	0	Mounts on 6219	
6305	Digital Trunk Dual	1	6/3/2	
7006	Realtime Interface Co-Processor Portmaster	1	7/4/8	
8128	128-Port Async Controller	1	7/4/7	

Notes:

1. Not supported in 120 MHz thin nodes or 135 MHz wide nodes.
2. High nodes and wide nodes have a minimum of one if used.
3. FDDI dual ring adapters **(F/C 272**3) have a mandatory prerequisite of the FDDI-SAS **(F/C 272**4).
4. The HIPPI feature **(F/C 273**5) uses three physical MCA slots and requires a total of five slots to satisfy power and thermal requirements.
5. The BMCA node-to-node **(F/C 275**2) cables are required on any node having the BMCA adapter and not having an external host connection.
6. BMCA adapters cannot be installed in two adjacent MCA slots due do physical cabling constraints. There is a limit of 32 adapters per frame, and a limit of 16 cables exiting the frame.
7. A maximum of 128 9333 High Performance Subsystem Adapter cables is allowed per frame.
8. Will not operate in SSA loops containing F/C 6214 or F/C 6216.
9. Withdrawn from production.

Evaluating SP System power requirement

The total power requirement of your RS/6000 SP system is the sum of the power required by all of the following individual components:

- Each base frame with integral power subsystem.
- Each processor node.
- All options and adapters in each node.
- Each switch.
- All auxiliary equipment, such as SP-attached servers, routers, RAIDs, extension nodes, printers, etc.

For the all auxiliary equipment, see the appropriate equipment documentation.

After calculating frame power requirements, add the total of all auxiliary equipment power to the total power usage of all the frames in your system.

Make one copy of Table B-6 for each frame in your SP system. Fill in the worksheet copies using your quantities of all applicable features. Complete the calculations, using the power factors given for each feature, to determine the total power requirement of each frame. After calculating the power requirements of each frame, you can use the results to calculate the thermal output of your system.

Table B-6 SP system frame power computation worksheet

Component	Feature	Qty x Factor	Watts
Frame	Empty frame with power subsystem (500, 550, 555, 1500, 1550)	(1) x 41	41
Nodes	375 MHz POWER3 SMP High Node (F/C 2058)	(___) x 650	
	375 MHz POWER3 SMP Wide Node (F/C 2057)	(___) x 104	
	375 MHz POWER3 SMP Thin Node (F/C 2056)	(___) x 59	
	POWER3 SMP High Node (F/C 2054)	(___) x 575	
	SP Expansion I/O Unit (F/C 2055)	(___) x 52	
	POWER3 SMP Wide Node (F/C 2053)	(___) x 104	
	POWER3 SMP Thin Node (F/C 2052)	(___) x 59	
	332 MHz SMP Wide Node (F/C 2051)	(___) x 154	
	332 MHz SMP Thin Node (F/C 2050)	(___) x 108	
	160 MHz Uniprocessor Thin Node (F/C 2022)	(___) x 135	
CPUs	375 MHz POWER3 SMP High Node CPU card (F/C 4350)	(___) x 168	
	(POWER3 SMP High Node CPU card (F/C 4849)	(___) x 132	
	375 MHz POWER3 SMP Thin or Wide Node CPU card (F/C 4444)	(___) x 160	
	POWER3 SMP Thin or Wide Node CPU card (F/C 4342)	(___) x 113	
	332 MHz SMP Thin or Wide Node CPU card (F/C 4320)	(___) x 76.5	

Component	Feature	Qty x Factor	Watts
Memory	POWER3 SMP High Node memory card (F/C 4880)	(___) x 66	
	375 MHz POWER3 SMP Thin and Wide Node memory card (F/C 4098)	(___) x 65	
	POWER3 SMP Thin and Wide Node memory card (F/C 4098)	(___) x 65	
	332 MHz SMP Thin and Wide Node memory card (F/C 4093)	(___) x 39.8	
	160 MHz Uniprocessor Thin Node memory card (F/C 4086, 7, 8, 9)	(___) x 11.5	
Switches and switch adapters	SP Switch (F/C 4011, 4008)	(___) x 167	
	SP Switch adapters (F/C 4020, 4022, 4023)	(___) x 26	
	SP Switch2 (F/C 4012)	(___) x 618	
	SP Switch2 adapter (F/C 4025)	(___) x 78	
Hard disk drives	4.5 GB Ultra SCSI (F/C 2900)	(___) x 8.9	
	4.5 GB Ultra SCSI disk pair (F/C 2904)	(___) x 17.8	
	9.1 GB Ultra SCSI (F/C 2908)	(___) x 22.1	
	9.1 GB Ultra SCSI disk pair (F/C 2909)	(___) x 44.2	
	18.2 GB Ultra SCSI disk pair (F/C 2918)	(___) x 57.6	
	9.1 GB Ultra SCSI disk pair (F/C 3800)	(___) x 44.2	
	9.1 SSA disk pair (F/C 3802)	(___) x 54.2	
	18.2 GB Ultra SCSI disk pair (F/C 3803)	(___) x 57.6	
	9.1 GB Ultra SCSI 10K RPM disk pair (F/C 3804)	(___) x 41.7	
	18.2 GB Ultra SCSI 10K RPM disk pair (F/C 3810)	(___) x 51.3	
	36.4 GB Ultra SCSI 10K RPM disk pair (F/C 3820)	(___) x 74.1	

Component	Feature	Qty x Factor	Watts
PCI adapters	Serial HIPPI SW (F/C 2732)	(___) x 39.7	
	Serial HIPPI LW (F/C 2733)	(___) x 39.7	
	FDDI SK-NET LP SAS (F/C 2741)	(___) x 9.9	
	FDDI SK-NET LP DAS (F/C 2742)	(___) x 15.2	
	FDDI SK-NET UP DAS (F/C 2743)	(___) x 9.9	
	ESCON Channel (F/C 2751)	(___) x 22.8	
	Token Ring (F/C 2920)	(___) x 2.3	
	Eight Port Async (F/C 2943)	(___) x 19.1	
	WAN 128-Port Async (F/C 2944)	(___) x 9.0	
	ARTIC960Hx 4-Port Selectable (F/C 2947)	(___) x 26.9	
	2-Port Multiprotocol X.25 (F/C 2962)	(___) x 10.8	
	TURBOWAYS 155 UTP ATM (F/C 2963)	(___) x 10.5	
	10/100 Ethernet 10BASE-TX (F/C 2968)	(___) x 9.1	
	Gigabit Ethernet-SX (F/C 2969)	(___) x 27.4	
	10BASE2/10BASE-T Ethernet LAN (F/C 2985)	(___) x 3.0	
	10BASE5/10BASE-T Ethernet LAN (F/C 2987)	(___) x 5.0	
	TURBOWAYS 155 ATM (F/C 2988)	(___) x 10.2	
	Token Ring (F/C 4959)	(___) x 2.7	
	SCSI-2 Ultra/Wide DE (F/C 6204, 6207)	(___) x 11.9	
	Ultra2 SCSI (F/C 6205)	(___) x 24.5	
	SCSI-2 Ultra/Wide SE (F/C 6206)	(___) x 11.9	
	SCSI-2 Fast/Wide SE (F/C 6208)	(___) x 11.9	
	SCSI-2 Fast/Wide DE (F/C 6209)	(___) x 11.9	
	SSA RAID 5 (F/C 6225)	(___) x 36.5	

Component	Feature	Qty x Factor	Watts
	Fibre Channel (F/C 6227)	(___) x 26.6	
	SSA RAID 5 (F/C 6230)	(___) x 36.5	
	F/W Cache Module (F/C 6235)	(___) x 5.9	
	ARTIC960RxD Quad Digital Trunk (F/C 6310)	(___) x 30.4	
MCA adapters	MCA slots occupied in uniprocessor thin nodes	(___) x 46	
	MCA slots occupied in uniprocessor high and wide nodes	(___) x 35.6	

Notes:

1. 1 Watt = 3.412 BTU/Hour
2. The SEPBU has a power factor of 1, thus W = VA
3. The PDU power supply ***does not*** have a power factor of 1, thus:
4. For frames with a PDU, multiply W x 1.16 for VA
5. For a PDU ***and*** High-voltage transformer, multiply W x 1.19 for VA

Legacy SP Nodes

Table B-7 shows Legacy SP Nodes.

Table B-7 Legacy SP nodes

Feature Code	Name	clock	Architecture
2001	THIN	62Mhz	Power1
2002	THIN	66Mhz	Power2
2003	WIDE	66Mhz	Power2
2004	THIN	66Mhz	Power2
2005	WIDE	77Mhz	Power2
2006	HIGH	112Mhz	PowerPC604
2007	WIDE	135Mhz	P2SC
2008	THIN	120Mhz	P2SC
2009	HIGH	200Mhz	PowerPC604
2022	THIN	160Mhz	P2SC
2050	332 SMP Thin	332Mhz	604 PowerPC
2051	332 SMP Wide	332Mhz	604 PowerPC
2052	POWER3 SMP Thin	200Mhz	630 PowerPC
2053	POWER3 SMP Wide	200Mhz	630 PowerPC
2054	POWER3 SMP High	225Mhz	630 PowerPC

SP Legacy Models by Frame and Switch Types

Table B-8 shows the first two digits defining the Frame Models in Legacy frames by Switch Types.

Table B-8 Legacy models by frames and switch types (table 1 of 2).

Model	Frame	Switch	Switch Frame	Nodes	Electrical Power
2Ax	Short	None	N/A	1 to 8	1 phase
3Ax	Short	SPS-8	No	1 to 8	1 phase
3Bx	Tall	SPS-8	No	1 to 8	3 phase
20x	Tall	None	N/A	1 to 64	3 phase

Note:

- x - Designates node type included with frame
- Short - 124.5 cm (49 in)
- Tall - 200.7 cm (79 in)
- SPS - SP Switch
- SPS-8 - SP Switch 8-port model
- Switch Frame - Not available for < 64 nodes
  - Optional for 65-80 nodes
  - Required for > 81 nodes

SP Legacy Models by Frame and Nodes Types

Table B-9 shows the third digits defining the Frame Models in Legacy frames by Node Types.

Table B-9 Legacy models by frames and node types (table 2 of 2)

Model	MHz	Form	Processor Type	I/O Bus	Available
xx1	62	Thin	Uni	MCA	No
xx2	66	Thin	Uni	MCA	No
xx3	66	Wide	Uni	MCA	No
xx4	66	Thin-2	Uni	MCA	No
xx5	77	Wide	Uni	MCA	No
xx6	112	High	SMP	MCA	No
xx7	135	Wide	Uni	MCA	No
xx8	120	Thin	Uni	MCA	No
xx9	200	High	SMP	MCA	No
xxA	160	Thin	Uni	MCA	No

SP LAN topologies

The network topology on SP/CES systems depends mainly on the size of the system. If the applications needed perform significant communication among nodes, additional network connectivity should be provided by SP Switch, Ethernet, Token Ring, FDDI, or ATM networks. To avoid overloading with application traffic, the SP LAN should be used only for SP node installations and system management tasks.

Note: IBM supplies an Ethernet LAN and one 15 meter (50 foot) Ethernet cable for each RS/6000 SP frame and each SP-attached server, as well as the required node-to-node cables. The Ethernet cable provided are BNC-type cables. IBM does not supply twisted-pair Ethernet cables or hubs. The customer must provide all other cables, network hubs, LAN switches, and routers.

The supported Ethernet adapters for SP LAN Ethernet communication on the 7017-S85, 7017-S80, 7017-S7A, and 7017-S70 servers - M/T 7017, as SP-attached servers and Clustered Enterprise Servers are the follows:

- Twisted-pair cable connection:
  - 10/100 Ethernet 10BASE-TX adapter (F/C 2968)
  - 10 MB AUI/RJ-45 Ethernet adapter (F/C 2987)
- BNC cable connection – 10 MB BNC/RJ-45 Ethernet adapter (F/C 2985)

The only supported Ethernet adapter for SP LAN Ethernet communication on the 7026-M80, 7026-6H1, and 7026-H80 servers - M/T 7026, as SP-attached servers and Clustered Enterprise Servers is 10/100 Ethernet 10BASE-TX adapter (F/C 2968) - Twisted-pair cable connection.

Shared or switched 100BASE-TX network

SP LANs, used in systems populated exclusively with POWER3 SMP nodes, can run at 100 Mbps including network installation. They require using unshielded twisted-pair (UTP) cabling.

A possible setup is to use a repeater capable of sustaining 100 Mbps with a fast Ethernet adapter in the CWS. An Ethernet switch, which supports full-duplex at 100 Mbps, provides improved performance over a repeater. Many node-to-node and node-to-CWS connections can be processed by the Ethernet switch simultaneously rather than with shared access through a repeater. The limiting factor for the quantity of simultaneous network installations of nodes is most likely the processing power of the CWS and not the network bandwidth.

However, switches with the required switching capabilities are available. An illustration of a 100BASE-TX network is shown in Figure B-1.

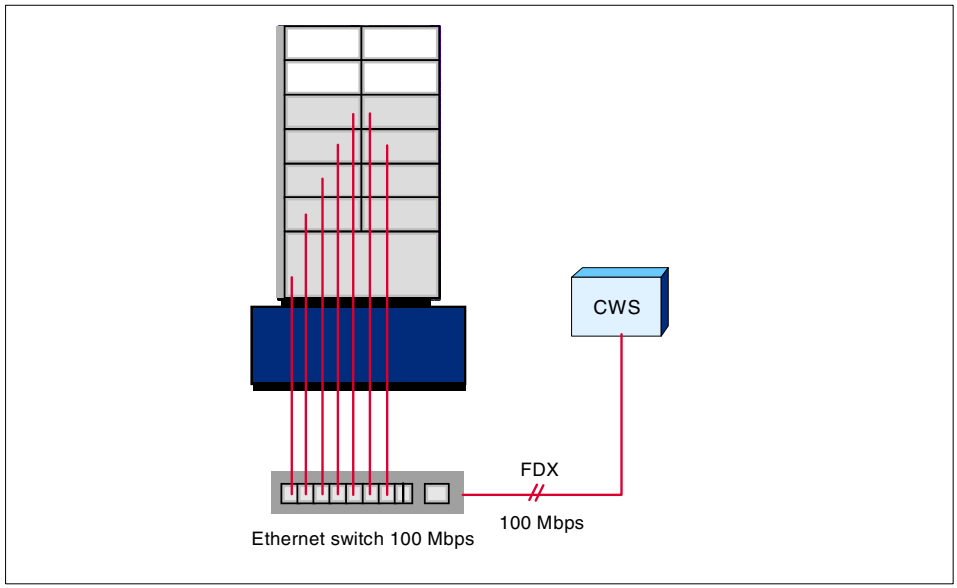


Figure B-1 Shared or switched 100BASE-TX network

Heterogeneous 10/100 Mbps network

SP systems can be upgraded with new nodes having fast Ethernet connections, while older or more lightly loaded nodes can continue to run with 10 Mbps SP LAN connections. In this case, you can use an Ethernet switch to provide a single LAN, which connects to the CWS at 100 Mbps FDX. As illustrated in Figure B-2 on page 394, one frame has new nodes with a 100 Mbps Ethernet, which are individually cabled by 100BASE-TX UTP to ports of the Ethernet switch and operate in full-duplex mode. Two frames with older nodes having 10BASE2 cabling are connected to ports of the same Ethernet switch using BNC to UTP media converters.

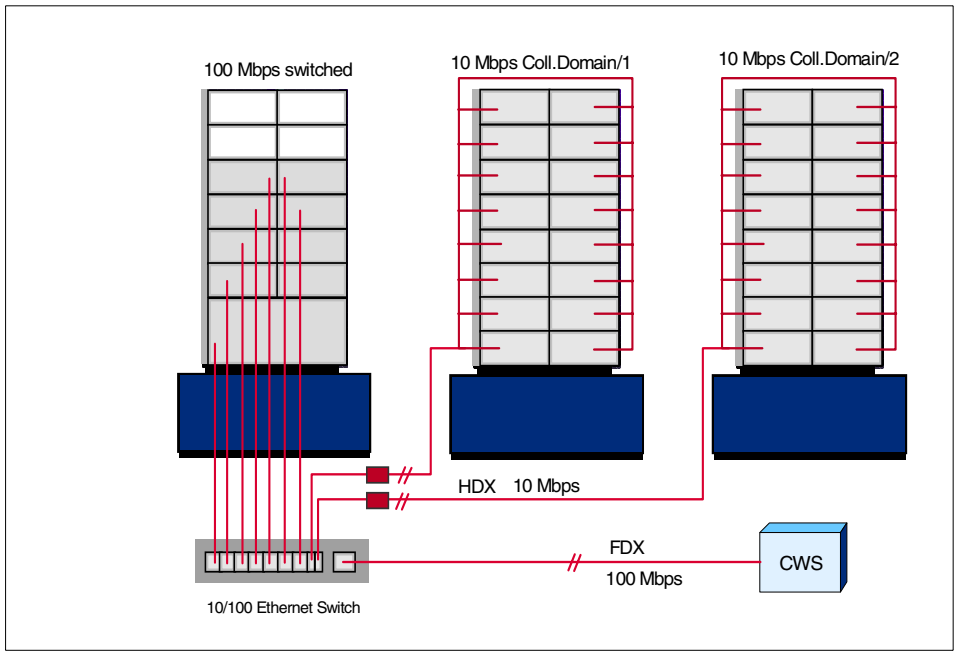


Figure B-2 Heterogeneous 10/100 Mbps network

In a similar case, you can use this same setup if a particular frame contains a mix of older nodes with 10 Mbps BNC and newer nodes with 100 Mbps UTP connections.

Shared 10BASE2 network

In small SP systems such as a single frame, the control workstation and the nodes typically share a single thin-wire Ethernet as illustrated in Figure B-3 on page 395.

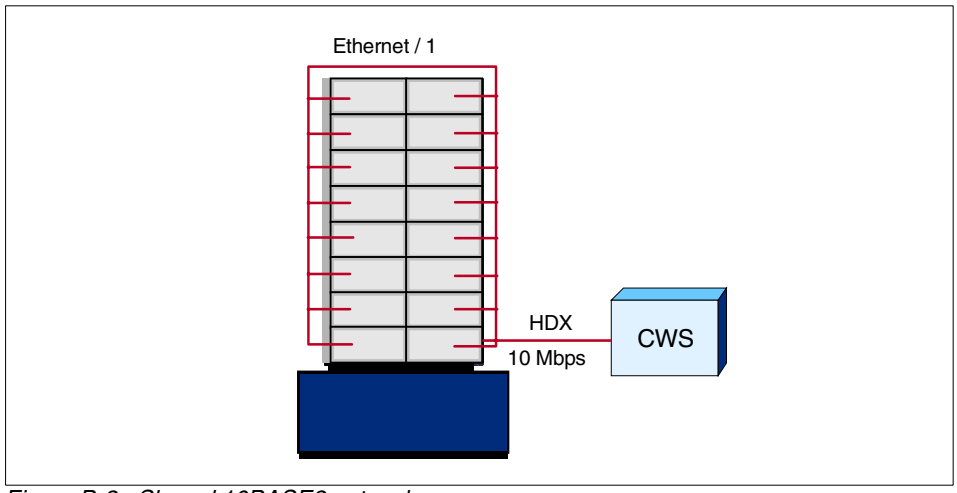


Figure B-3 Shared 10BASE2 network

This network has the following properties:

- No routing is required since the CWS and all nodes share one subnet.
- The entire SP LAN is a single broadcast domain as well as a single collision domain.
- Performance is limited to one 10 Mbps HDX connection at a time.

Even if this performance limitation is acceptable, this setup is limited by a maximum quantity of 30 stations on a 10BASE2 segment. In practice, no more than 16 to 24 stations should be connected to a single 10BASE2 Ethernet segment.

Segmented 10BASE2 network

Ethernet segmentation is used to overcome the limitations of a single shared Ethernet. The CWS is equipped with additional Ethernet adapters and each one is connected to a different shared 10BASE2 subnet as illustrated in Figure B-4 on page 396.

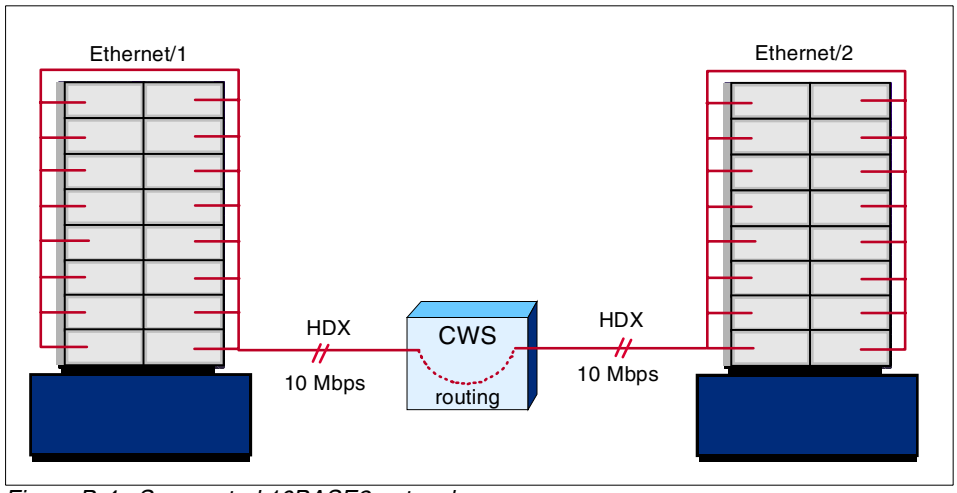


Figure B-4 Segmented 10BASE2 network

A network with N separate subnets and N Ethernet cards in the CWS has the following properties:

- Nodes in one subnet need static routes to the (N-1) other subnets through the CWS and routing or IP forwarding must be enabled on the CWS.
- The SP LAN is split into N broadcast domains.
- Aggregate performance is limited to a maximum of N times 10 Mbps HDX. This is only achievable if the CWS communicates with one node in each of the subnets simultaneously.

Note: This approach is limited by the number of available adapter slots in the CWS and the ability of the CWS to simultaneously handle the traffic among these subnets or to serve 6N to 8N simultaneous network installations. In practice, no more than four subnets should be used.

Switched 10BASE2 network

A network that overcomes performance limitations in shared or segmented Ethernet networks is Ethernet switching. Even with an added router, this setup is usually preferable to a segmented network with boot/install servers, from the viewpoints of performance, management, and complexity. This network is illustrated in Figure B-5 on page 397.

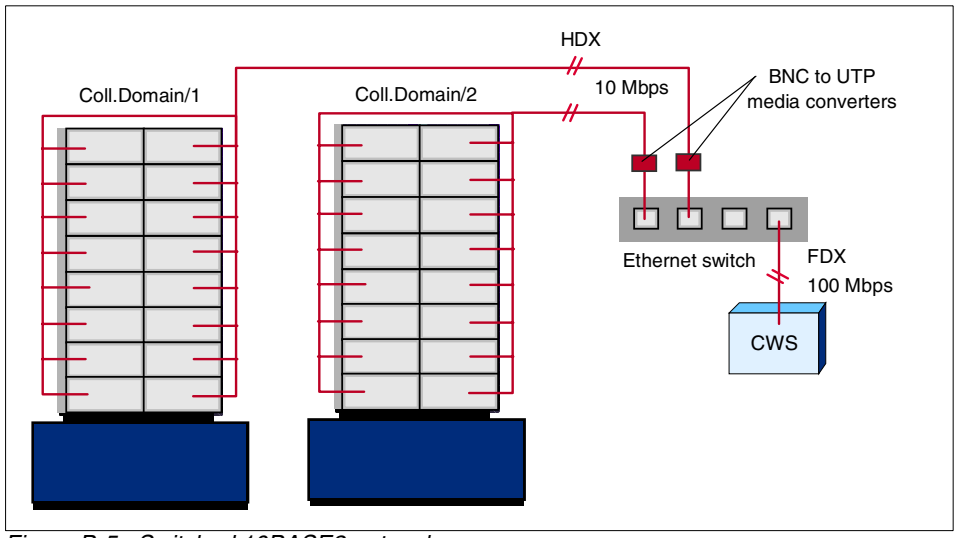


Figure B-5 Switched 10BASE2 network

This network setup has the following properties:

- No routing is required. All Ethernet segments are transparently combined into one large LAN by the Ethernet switch.
- Node-to-node connections within a single Ethernet segment still have to share that 10BASE2 medium in half-duplex mode. However, many communications between different ports can be switched simultaneously by the Ethernet switch. The uplink to the CWS can be run at 100 Mbps in full-duplex mode.

  This setup eliminates the routing overhead for communications between nodes or a node and the CWS. With a 100Mbps, full-duplex uplink to the CWS there should be no bottleneck in the connection to the CWS if the quantity of 10BASE2 segments is not much larger than ten.

To avoid problems with broadcast traffic, no more than 128 nodes should be connected to a single switched Ethernet subnet. Larger systems should be set up with a suitable quantity of switched subnets.

SP Boot/Install server requirements

For the SP Ethernet topology, you should consider your network install server requirements. The network install process uses the SP Ethernet for transferring the install image from the install server to the SP nodes. Running lots of concurrent network installs can exceed the capacity of the SP Ethernet. The following are suggested guidelines for designing the SP Ethernet topology for efficient network installs. Many of the configuration options will require additional network hardware beyond the minimal node and control workstation requirements.There are also network addressing issues to consider.

The following requirements exist for all configurations:

- Each boot-install server's en0 Ethernet adapter must be directly connected to each of the control workstations' Ethernet adapters.
- The Ethernet adapter configured as en0 must always be in the SP node's lowest hardware slot of all Ethernets.
- The NIM clients that are served by boot-install servers must be on the same subnet as the boot-install server's Ethernet adapter.
- NIM clients must have a route to the control workstation over the SP Ethernet.
- The control workstation must have a route to the NIM clients over the SP Ethernet.

Single frame systems

For small systems, you can use the control workstation as the network install server. This means that the SP Ethernet is a single network connecting all nodes to the control workstation. When installing the nodes, you should limit yourself to installing eight nodes at a time because this is the limit of acceptable throughput on the Ethernet. Figure B-6 on page 399 shows an Ethernet topology for a single-frame system.

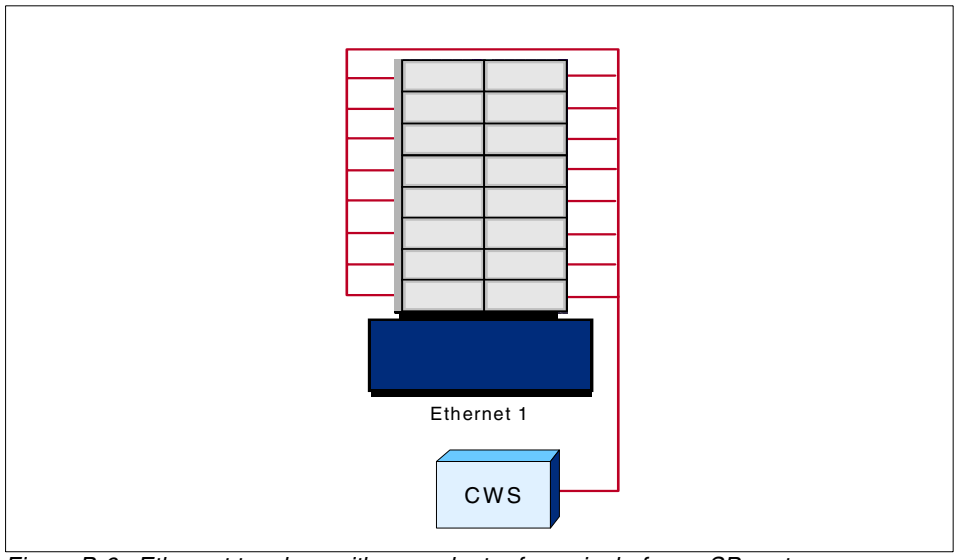


Figure B-6 Ethernet topology with one adapter for a single-frame SP system

An alternate way to configure your system is to install a second Ethernet adapter in your control workstation, if you have an available I/O slot, and use two Ethernet segments to the SP nodes. Each network should be connected to half of the SP nodes. During network install, you can install all 16 nodes at the same time. Figure 10-6 shows this alternate Ethernet topology for a single-frame system.

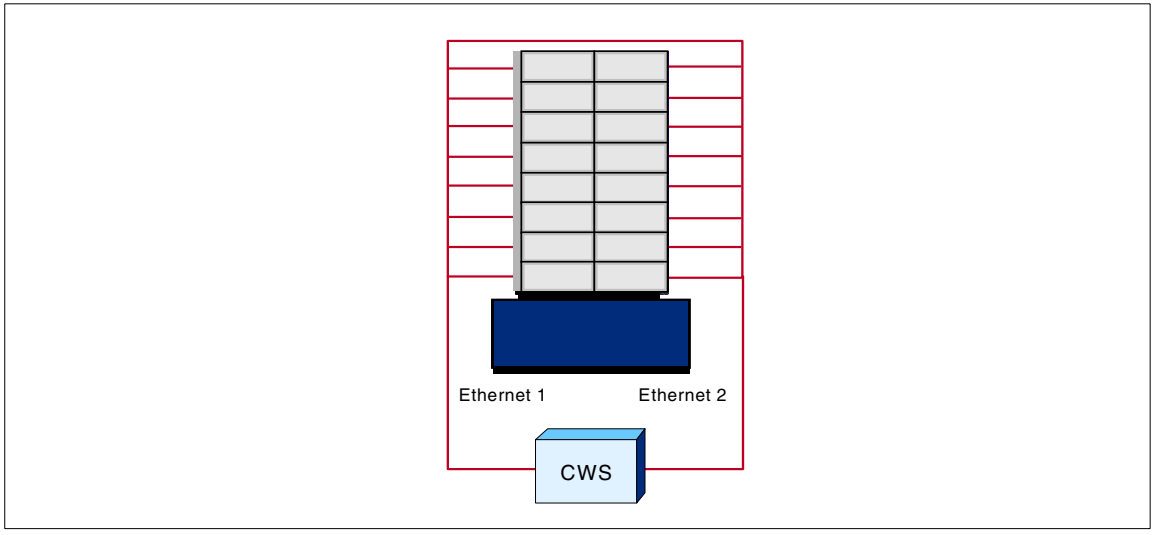


Table 10-6 Ethernet topology with two adapters for single-frame SP system

You have to set up your SP Ethernet routing so nodes on one Ethernet can communicate to nodes on the other network. You also need to set up your network mask so that each SP Ethernet is its own subnet within a larger network address.

Multiple frame systems

For multiple frame systems, you want to spread the network traffic over multiple Ethernets, and keep the maximum number of simultaneous installs per network to eight. You can use the control workstation to network install specific SP nodes, which will be the network install servers for the rest of nodes.

There are three ways to accomplish this.

1. The first method uses a control workstation with one Ethernet adapter for each frame of the system, and one associated SP Ethernet per frame. So, if you have a system with four frames as in Figure B-7 on page 400, the control workstation must have enough I/O slots for four Ethernet adapters, and each adapter connects one of the four SP frame Ethernet segments to the control workstation. Using this method, you install the first eight nodes on a frame at a time, or up to 32 nodes if you use all four Ethernet segments simultaneously. Running two installs will install up to 64 nodes. Figure B-7 shows an Ethernet topology for this multi-frame system.

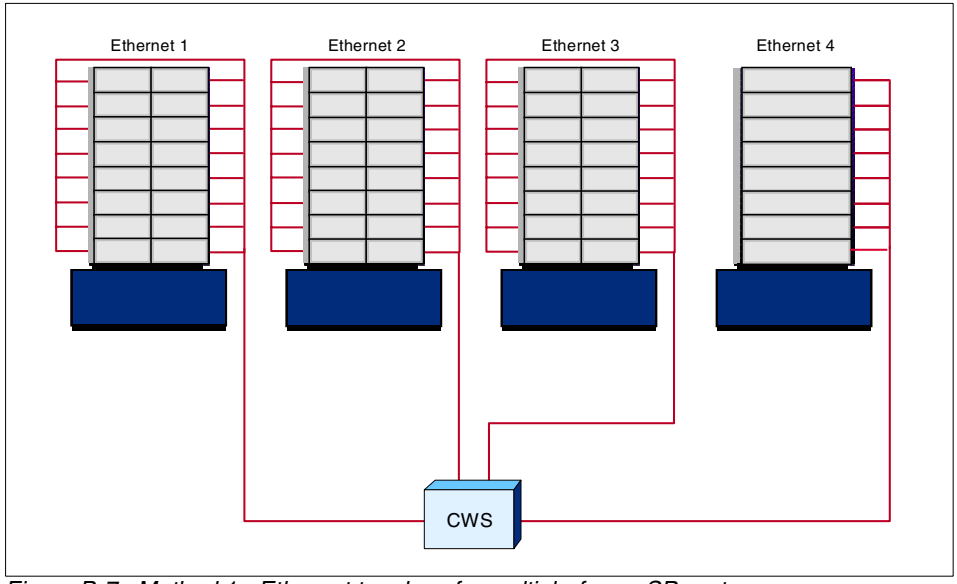


Figure B-7 Method 1 - Ethernet topology for multiple-frame SP system

You will have to set up your SP Ethernet routing so nodes on one Ethernet can communicate with nodes on another. You also need to set up your network mask so that each SP Ethernet is its own subnet within a larger network address. This method is applicable up to the number of slots your control workstation has available.

2. A second approach designates the first node in each frame as a network install server, and then the remaining nodes of that frame are set to be installed by that node. This means that, from the control workstation, you will have an SP Ethernet segment connected to one node on each frame. Then the network install node in each frame has a second Ethernet card installed, which is connected to an Ethernet card in the rest of the nodes in the frame. Figure B-8 shows an Ethernet topology for this multi-frame system.

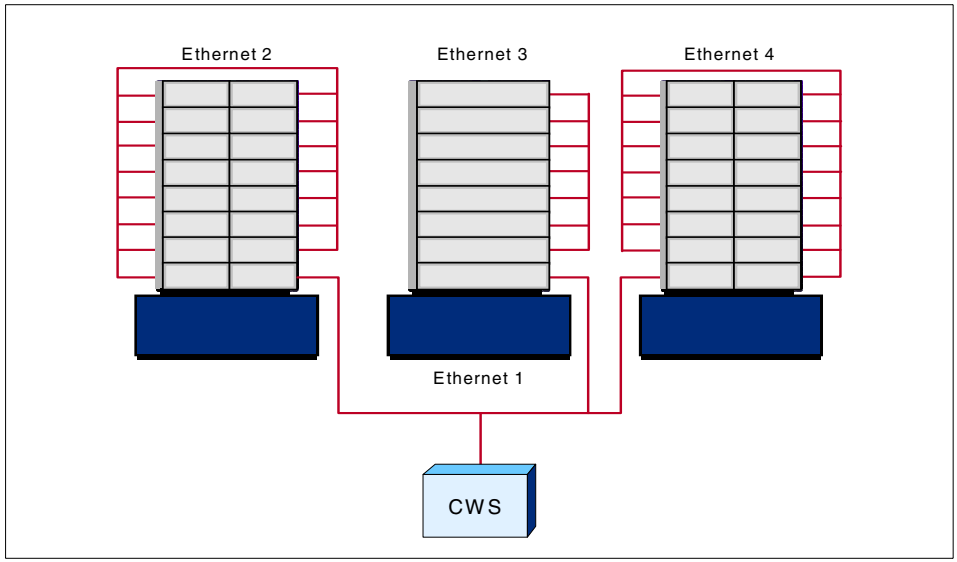


Figure B-8 Method 2 - Ethernet topology for multi-frame SP system

When using this method, installing the nodes requires that you first install the network install node in each frame. The second set of installs will install up to eight additional nodes on the frame. The last install, if needed, installs the rest of the nodes in each frame.

3. A third method adds an external router to the topology of the previous approach. This router is made part of each of the frame Ethernets, so that traffic to the outside need not go through the control workstation. You can do this only if the control workstation can also be attached externally, providing another route between nodes and the control workstation. Figure B-9 shows this Ethernet topology for such a multi-frame system.

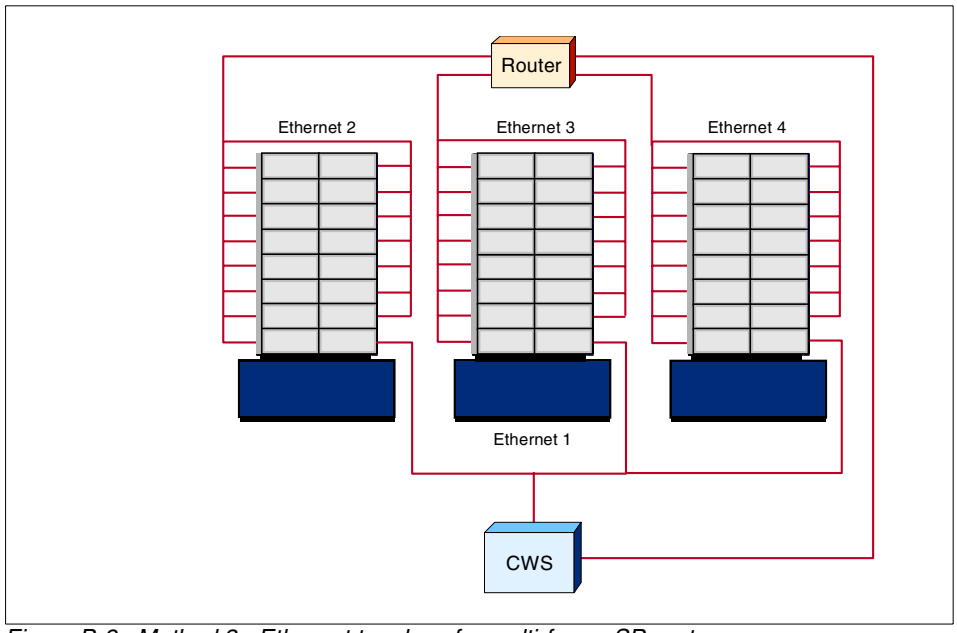


Figure B-9 Method 3 - Ethernet topology for multi-frame SP system

An alternative to the router in this configuration is an Ethernet switch, which could have a high-speed network connection to the control workstation.

Future expansion considerations

If your configuration will grow over time to a large configuration, you might want to dedicate your network install nodes in a different manner.

For very large configurations, you might want to dedicate a frame of nodes as designated network install nodes, as shown in Figure B-10 on page 403. In this configuration, each SP Ethernet from the control workstation is connected to up to eight network install nodes in a frame. These network install nodes are, in turn, connected to additional frames.

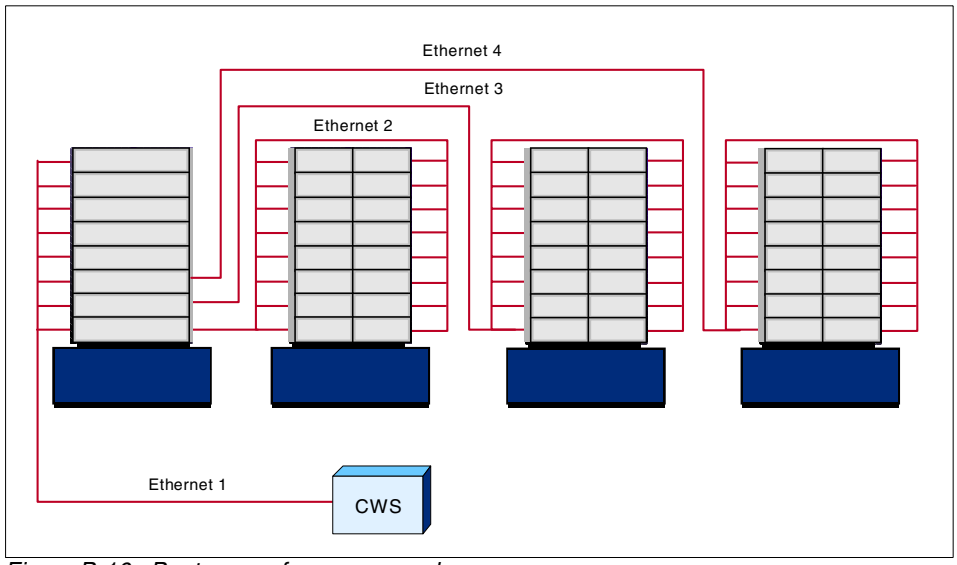


Figure B-10 Boot server frame approach

The advantage of this is that when you add an additional frame to your SP configuration, all you need to do is connect the new frame to one of the network install nodes, and reconfigure the system.

The network install procedure for this system is the same as for multiple frame systems. You first install the network install servers at a rate of eight per SP Ethernet segment. The network install servers then install eight other nodes until all nodes are installed.

The network address usually used for the SP Ethernet is a class C internet address. This address has a limit of 256 individual addresses before you need to add additional network addresses for the SP Ethernet. If your system is expected to grow beyond this number of nodes, you should plan with your Network Administrator additional network addresses for future SP Ethernet expansion. This will save you from having to re-assign the SP Ethernet addresses when you reach the address limit.

Disk Space required for PSSP file sets

Table B-10 shows the approximate sizes of the install image and individual fileset shipped with the PSSP software.

Table B-10 Disk space required for install images and to install PSSP file sets

Image or Fileset Name	Image Size	File Size
PSSP minimal AIX image: spimg		
spimg.433	91MB	
PSSP image: rsct.basic	16MB	
rsct.basic.hacmp		280KB
rsct.basic.rte		20MB
rsct.basic.sp		1.6MB
PSSP image: rsct.clients	500KB	
rsct.clients.hacmp		18KB
rsct.clients.rte		533KB
rsct.clients.sp		27KB
PSSP image:rsct.core	900KB	
rsct.core.utils		1.4MB
PSSP image: ssp	150MB	
ssp.authent		660KB
ssp.basic		6.8MB
ssp.cediag		433KB
ssp.clients		10.5MB
ssp.css		79.2MB
ssp.docs		30MB
ssp.gui		31.2MB
ssp.ha_topsvcs.compat		1KB
ssp.jm		550KB
ssp.perlpkg		1.6MB

Image or Fileset Name	Image Size	File Size
ssp.pman		1.2MB
ssp.public		12MB
ssp.spmgr		634KB
ssp.st		765KB
ssp.sysctl		1.5MB
ssp.sysman		1.3MB
ssp.tecad		203KB
ssp.tguides		1.1MB
ssp.top		1.4MB
ssp.top.gui		1.2MB
ssp.ucode		907KB
PSSP image: ssp.hacws	160KB	
ssp.hacws		141KB
PSSP image: ptpe	7MB	
ptpe.docs		3.1MB
ptpe.program		6.2MB
PSSP image: vsd	21MB	
vsd.cmi		173KB
vsd.hsd		180KB
vsd.rvsd.hc		374KB
vsd.rvsd.rvsdd		330KB
vsd.rvsd.scripts		249KB
vsd.sysctl		391KB
vsd.vsdd		652KB
PSSP images for other graphical user interface		
ssp.ptpegui	1.7MB	
ssp.vsdgui	3MB	

Image or Fileset Name	Image Size	File Size
PSSP image ssp.resctr for the Resource Center		
ssp.resctr.rte	4MB	

Note: The total storage can cross multiple file systems.

C

SP system upgrades

This section contains information on available RS6000 SP upgrades, conversions, and feature additions. Examples of typical upgrades include adding frames to your SP system, installing new higher-performance nodes or switches, and converting MCA-type nodes to PCI nodes.

If you are planning an upgrade for your SP system, you might need the following identifying information for the ordering process:

- Machine Type (M/T 9076)
- Model number
- SP System serial number

Adding POWER3, and 332 MHz SMP Thin and Wide Nodes

SMP Thin and Wide Nodes can be added to earlier SP systems. Adding these nodes to earlier withdrawn frames requires a power system upgrade. Both PDU and SEPBU power subsystems must be upgraded to 10.5 kW level SEPBUs for tall frames or 5.0 kW for short frames.

These nodes can be added to any Model 500, 550, 20X, 30X, or 40X, and F/C 1500 or 1550 expansion frame that has available drawer space and an upgraded PDU.

Adding 375 MHz POWER3 and POWER3 SMP High Nodes

POWER3 High Nodes can be added to any Model 550 or F/C 1550 expansion frame that has available drawer space. POWER3 High Nodes can also be added to Models 20X, 30X, or 40X only with the addition of a tall, deep expansion frame (F/C 1550), since the POWER3 High Node does not physically fit into the old-style frame. According to configurator rules, if there are no existing POWER3 High Nodes in your system, the MES will include a tall, deep expansion frame. If you determine that a new frame is not required, use RPQ8P2091 to delete the frame from the MES. The POWER3 High Node Lift Tool is required and included to install these nodes. The tool is returned after use.

Adding SP Expansion I/O Units

SP Expansion I/O Units for POWER3 High Nodes can be installed only in Model 550, Model T70, or F/C 1550 expansion frames. If more than eight SP Expansion I/O Units are installed in a single frame, an EMC cover gasket set (ECA065) must be installed in that frame.

Adding frames

You can add a tall expansion frame (F/C 1550) to a Model 550, 20X, 30X, or 40X system allowing additional nodes to be installed. You can also add a short expansion frame (F/C 1500) to a Model 500 (eight nodes maximum) system.

Adding switches

SP switches can be added in the following configurations:

1. Switchless to SP Switch (F/C 4011), SP Switch-8 (F/C 4008), or SP Switch2 (F/C 4012)
2. SP Switch (F/C 4011) to SP Switch2 (F/C 4012)
3. High Performance Switch (F/C 4010) to SP Switch (F/C 4011)
4. High Performance Switch LC8 (F/C 4007) to SP Switch-8 (F/C 4008)

Note:

1. The SP Switch2 requires the three-book, 10.5 kW SEPBU, which has a unique power connector.
2. Since the SP Switch2 is only supported with POWER3 and 375 MHz POWER3 High Nodes, you might need to split an existing system to remove any nodes which are not supported with the SP Switch2. You will then have two SP systems, each with its own serial number. You can order this “split system” feature using RPQ 8P2009.

Upgrading, replacing, and relocating nodes

Unique requirements for high nodes:

1. POWER3 High Nodes (F/C 2054, 2058) require a tall, deep frame (Model 550 or F/C 1550).
2. POWER3 High Nodes can be installed into Models 20X, 30X, or 40X only with the addition of a tall, deep expansion frame (F/C 1550), since they do not physically fit into the old-style frame. According to configurator rules, if there are no existing POWER3 High Nodes in your system, the MES will include a tall, deep expansion frame. If you determine that a new frame is not required, use RPQ8P2091 to delete the frame from the MES.
3. The POWER3 High Node Lift Tool is required and included with the MES to install these nodes. The tool is returned after use.

Upgrading SMP-type nodes

Available node upgrade features are shown in Table C-1.

Table C-1 Node upgrade features

Original node type	Can be upgraded to
332 MHz Thin (F/C 2050)	375 MHz POWER3 Thin (F/C 2056) 375 MHz POWER3 Wide (F/C 2057) 375 MHz POWER3 High (F/C 2058)
332 MHz Wide (F/C 2051)	375 MHz POWER3 Wide (F/C 2057) 375 MHz POWER3 High (F/C 2058)
POWER3 Thin (F/C 2052)	375 MHz POWER3 Thin (F/C 2056) 375 MHz POWER3 Wide (F/C 2057) 375 MHz POWER3 High (F/C 2058)
POWER3 Wide (F/C 2053)	375 MHz POWER3 Wide (F/C 2057) 375 MHz POWER3 High (F/C 2058)
POWER3 High (F/C 2054)	375 MHz POWER3 High (F/C 2058)
375 MHz POWER3 Thin (F/C 2056)	375 MHz POWER3 Wide (F/C 2057) 375 MHz POWER3 High (F/C 2058)
375 MHz POWER3 Wide (F/C 2057)	375 MHz POWER3 High (F/C 2058)

Note: These are one-for-one upgrades since some of the parts from the original nodes are used in the replacement nodes. Thus for example, you cannot upgrade four thin nodes to one high node. If the node being upgraded contains multiple processor cards, at least one of the processor cards must be installed into the new replacement node. Any additional processor cards can also be installed in the new node or can be used in other supported nodes in the system.

Replacing withdrawn processor nodes

Replacement features are available to replace withdrawn early processor nodes with 332 MHz, POWER3, and 375 MHz POWER3 SMP Thin, Wide, and High nodes.

Relocating processor nodes

If you plan to relocate any nodes or switches to another frame, you must order one of the following no-charge features:

- Relocating a Thin node (F/C 9250)
- Relocating a Wide node (F/C 9251)
- Relocating a High node (F/C 9252)
- Relocating an SP Switch (F/C 9253)

If you are relocating thin or wide nodes to a frame that does not have the 48 V dc power and Ethernet cables already installed, you can order the following features for each node you are moving:

1. F/C 9303 - 48 V dc cable
2. F/C 9304 - BNC Ethernet cable

Upgrading SP frames

The following SP frame upgrades can be performed:

Upgrading Model 500 and F/C 1500 to Model 550 and F/C 1550

This conversion is available if you want to upgrade your Model 500 short-frame SP system with a maximum of eight nodes to a Model 550 tall-frame system with a maximum capacity of 128 nodes. You receive a new tall frame with its integral SEPBU power subsystem and linecord and then your nodes are transferred from the old frame to the new. In this case, the system retains its original serial number.

Upgrading Model T70 to Model 550

This conversion transfers the POWER3 High Node and any SP Expansion I/O Units in a Model T70 Workgroup Server to a Model 550 tall frame.

D

Software Appendix

This appendix contains detailed information about the relationships between supported software packages.

In this appendix, the following software packages are described:

1. PSSP and AIX
2. High Availability Cluster Multi-Processing (HACMP)
3. Virtual shared disk and Recoverable virtual shared disk (VSD and (RVSD)
4. General Parallel File System for AIX (GPFS)
5. Parallel Environment (PE)
6. Engineering and Scientific Subroutine Library (ESSL) and Parallel ESSL (PESSL)
7. LoadLeveler

PSSP and AIX

Since PSSP and AIX must work together on the SP to allow other applications to run, they will be considered as one unit in this section.

Supported levels

Each LPP is not supported in every level on PSSP and AIX. Table D-1 shows what levels of PSSP and AIX support which LPP levels.

Table D-1 IBM LPPs per supported PSSP and AIX release

PSSP and AIX	IBM LPPs
PSSP3.2 AIX 4.3.3	► LoadLeveler 2.2, 2.1 ► PE 3.1 ► ESSL 3.1* ► PESSL 2.1* ► GPFS 1.3, 1.2, 1.1 ► HACMP/ES and HACMP 4.3 ► HACMP with HAGEO or GeoRM 2.1
PSSP 3.1.1. AIX 4.3.3	► LoadLeveler 2.2, 2.1 ► PE 2.4 ► ESSL 3.1* ► PESSL 2.1* ► GPFS 1.2, 1.1 ► HACMP/ES and HACMP 4.3
PSSP 2.4 AIX 4.2.1 or 4.3.3	► LoadLeveler 1.3 ► PE 2.3 ► PESSL 2.1* ► GPFS 1.1 ► RVSD 2.1.1 ► HACMP/ES and HACMP 4.2

Note: Before PSSP 3.1, IBM Recoverable Virtual Shared Disk was a separate LPP. The High Availability Control WorkStation and the Performance Toolbox Parallel Extensions components were priced features. They are now optional components which you receive with PSSP.

Table D-2 shows AIX and PSSP support for different node types:

Table D-2 AIX and PSSP support for different node types

AIX/PSSP	120 MHZ Thin	135 MHZ Wide	604 MHZ High	160 MHZ Thin	332 MHZ Wide	POWER3 Thin	POWER3 Wide	POWER3 High
AIX 4.2.1 PSSP 2.4	S	S	S	S	S	N	N	N
AIX 4.3.1 PSSP 2.4	S	S	S	S	S	N	N	N
AIX 4.3.2 PSSP 3.1	A	A	A	A	A	A	A	N
AIX 4.3.3 PSSP 3.1.1	A	A	A	A	A	A	A	A
AIX 4.3.3 PSSP 3.2	A	A	A	A	A	A	A	A

A = Available; factory and field orderable

S = Supported

N = Not supported

Note:

- AIX V4.2.1 with PSSP V2.4 or later is required for low cost models with High Nodes. The low cost models include tall and short frames with 8-port switch, and short frames without switch.
- A single 332 MHz Thin Node in a drawer is supported only by AIX V4.3.2 with PSSP V3.1 or later.
- 375 MHZ POWER3 SMP Nodes (features 2056 and 2057) are not supported on AIX 4.3.2 with PSSP 3.1

Migration

Only certain migration paths are possible. Table D-3 list them.

Table D-3 Supported migration paths for AIX and PSSP

From	To
PSSP 2.2 and AIX 4.1.5 or 4.2.1	PSSP 3.2 and AIX 4.3.3
PSSP 2.3 and AIX 4.2.1 or 4.3.3	PSSP 3.2 and AIX 4.3.3

From	To
PSSP 2.4 and AIX 4.2.1 or 4.3.3	PSSP 3.2 and AIX 4.3.3
PSSP 3.1.1and AIX 4.3.3	PSSP 3.2 and AIX 4.3.3

Coexistence

Coexistence is supported in the same system partition or a single default system partition (the entire SP system) for nodes running any combination of AIX and PSSP as shown in Table D-4.

Table D-4 Levels of PSSP and AIX supported in a mixed system partition

Product	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3
PSSP 3.2	N	N	S
PSSP 3.1.1	N	N	S
PSSP 2.4	N	S	S
PSSP 2.3	N	S	S
PSSP 2.2	S	S	N

S = Supported

N = Not supported

High Availability Cluster Multi-Processing (HACMP)

This section discussed HACMP supported levels, migration paths and coexistence options.

Supported levels

While PSSP has no direct requirement for HACMP, if you install hardware that requires PSSP 3.2, regardless of system partitioning, and you already use or are planning to use HACMP, you must also install and use HACMP 4.3. HACMP 4.3 can be run on the SP nodes and SP-attached servers. HACMP/ES 4.3 can run on the control workstation but prior levels can only run on nodes, not on SP-attached servers and not on the control workstation.

HACMP has a dependency on the RSCT Group Services. The RSCT component of PSSP is included in the HACMP install stack. The dependencies of HACMP on AIX and PSSP are shown in Table D-5.

Table D-5 HACMP supported levels

HACMP	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3	PSSP 2.4	PSSP 3.1.1	PSSP 3.2
4.4.0	N	N	S	N	N	S
4.3.1	N	N	S	N	N	S
4.3.0	N	N	S	N	S	S
4.2.2 4.2.1	N	S	S	S	S	N
4.2.0 4.1	S	S	S	S	S	N

S = Supported

N = Not supported

Migration considerations

HACMP 4.3.1 on the SP requires AIX 4.3.3 (or later) and PSSP 3.2. You have the following migration options:

- Migrating from HACMP/6000 1.2, 2.1, or 3.1 to HACMP 4.3.1 involves taking configuration snapshots, bringing down the cluster, reinstalling HACMP on all nodes in the cluster, and bringing it back up again.
- Migrating from HACMP 4.1, 4.2, or 4.3.0 to HACMP 4.3.1 also involves reinstalling HACMP on all nodes in the cluster; however the version compatibility function allows you to upgrade the cluster one node at a time without taking the entire cluster offline.
- Due to HACMP dependencies on levels of AIX, migrating one node at a time might require you to upgrade AIX, PSSP, and HACMP on the node all during the same service window.

Coexistence

HACMP 4.3 is not compatible with any of its lower level versions. While there is a version compatibility function to allow HACMP 4.3 to coexist in a cluster with HACMP 4.2 or HACMP 4.1, this function is intended as a migration aid only. Once the migration is completed, each processor in the HACMP cluster must be at the same AIX and HACMP release levels, including all PTFs.

VSD and RVSD

In order to exploit the new functions available in the Recoverable Virtual Shared Disk (RVSD) option of PSSP 3.2, all nodes in the system must be running PSSP 3.2 with the RVSD component. RVSD requires the Virtual Shared Disk (VSD) optional components to be running also.

After the last node is migrated to PSSP 3.2 with the RVSD option, all nodes in the system partition must have the RVSD subsystem reset. This requires stopping and starting applications such as Oracle, which can occur during a service window of less than four hours (approximately ten minutes). The **`rvsdrestrict`** command allows the administrator to select the level at which the RVSD subsystem is to run in a mixed system partition. If any node in the system partition has a level earlier than that set by the command, the RVSD subsystem will not be activated on that node. See Table D-6 for supported VSD and RVSD levels.

Attention: VSD and RVSD are not supported in a system comprised of clustered enterprise servers.

Supported Levels

Table D-6 VSD and RVSD supported levels

VSV & RVSD	PSSP 2.2	PSSP 2.3	PSSP 2.4	PSSP 3.1.1	PSSP 3.2	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3
1.2	S	S	S	S	S	S	S	S
2.1	N	S	S	S	S	N	S	S
2.1.1	N	N	S	S	S	N	S	S
3.1	N	N	N	S	S	N	N	S
3.2	N	N	N	N	S	N	N	S

S = Supported

N = Not supported

Migration considerations

The following is a list of points that should be considered before starting a migration.

1. When migrating from IBM Recoverable Virtual Shared Disk 1.2 and PSSP 2.2 you can migrate one node at a time but you must migrate both IBM Recoverable Virtual Shared Disk and PSSP at the same time.
2. You can migrate later levels of PSSP one node at a time without having to migrate IBM Recoverable Virtual Shared Disk at the same time. This is due to the fencing support that was added in PSSP 2.3.
3. In order to exploit the new functions available in the Recoverable Virtual Shared Disk option of PSSP 3.2, all nodes in a system partition must be running PSSP 3.2 with the Recoverable Virtual Shared Disk component. Recoverable Virtual Shared Disk requires the Virtual Shared Disk optional component be running also.

Coexistence

The Recoverable Virtual Shared Disk component of PSSP can coexist with IBM Recoverable Virtual Shared Disk LPPs in a mixed system partition with any combination of the following groupings:

- Recoverable Virtual Shared Disk component of PSSP 3.2
- Recoverable Virtual Shared Disk component of PSSP 3.1.1
- IBM Recoverable Virtual Shared Disk 2.1.1 with PSSP 2.4, or 3.1
- IBM Recoverable Virtual Shared Disk 2.1 with PSSP 2.3, 2.4, or 3.1
- IBM Recoverable Virtual Shared Disk 1.2 with PSSP 2.2, 2.3, 2.4, or 3.1

General Parallel File System for AIX (GPFS)

A file system managed by the GPFS LPP can only be accessed from within the system partition. GPFS 1.3 changes the locking semantics which control access to data and as a result requires that all nodes be at the same level. The task of migrating one node at a time is not supported. GPFS is supported in systems with the SP Switch or SP Switch2. To see what GPFS levels are supported by what PSSP and AIX levels see Table D-7.

Attention: GPFS is not supported in a system comprised of clustered enterprise servers.

Supported Levels

Table D-7 Supported levels in relationship to AIX and PSSP supported levels

GPFS	PSSP 2.4	PSSP 3.1.1	PSSP 3.2	AIX 4.2.1	AIX 4.3.3
1.4	N	N	S	N	S
1.3	N	N	S	N	S
1.2	N	S	S	N	S
1.1	S	S	S	S	N

S = Supported

N = Not supported

Coexistence

GPFS 1.3 does not coexist in a system partition with nodes using earlier levels of GPFS. It cannot coexist in a mixed system partition–it works only with PSSP 3.2 and AIX 4.3.3. However, all applications which execute on GPFS 1.1 will execute on GPFS 1.2 and all file systems created with GPFS 1.1 can be used with GPFS 1.2 and can be upgraded to a GPFS 1.2 file system.

In GPFS Release 3, all nodes in a GPFS nodeset must be in the same partition. A GPFS file system may only be accessed from within a single SP partition. File systems created under older releases of GPFS may continue to be used under GPFS 1.4.

Parallel Environment (PE)

Parallel Environment is comprised of:

1. Parallel Operating Environment (POE)
2. Message Passing Libraries (MPI, MPL)
3. Parallel Utilities which facilitate file manipulations (MPI sample programs)

IBM Parallel Environment for AIX, is not supported in a mixed system partition. See Table D-8 for AIX and PSSP support levels.

Table D-8 Supported PE LPP levels

PE	PSSP 2.4	PSSP 3.1.1	PSSP 3.2	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3
3.1	N	N	S	N	N	S
2.4	N	S	N	N	N	S
2.3	S	N	N	N	S	S
2.2	N	N	N	S	S	N

S = Supported

N = Not supported

Migration considerations

Parallel Environment does not support node by node migration. Migration paths are shown in Table D-9. There are some migration limitations as follows:

1. All nodes in the system partition need to be migrated to a new level of PE within the same service window.
2. You can run a particular level of PE with plus or minus one level of AIX or PSSP so you can migrate to a new level of AIX or PSSP without having to change to a new level of PE.
3. Applications using threads can migrate from PSSP 2.4 or PSSP 3.1.1 to PSSP 3.2 through binary compatibility without recompiling but they might not pick up the D7 libpthreads.a (shr.o) library.
4. You should migrate AIX, then PSSP, then PE.

Table D-9 Available PE migration paths

From	To
AIX 4.2.1, PSSP 2.4, PE 2.3 + PTF	AIX 4.3.2, PSSP 2.4, PE 2.3 + PTF

From	To
AIX 4.3.2, PSSP 2.4, PE 2.3 + PTF	AIX 4.3.2, PSSP 3.1.1, PE 2.4
AIX 4.3.2, PSSP 2.4, PE 2.3 + PTF	AIX 4.3.3, PSSP 3.2, PE 2.4
AIX 4.3.2, PSSP 3.1.1, PE 2.4	AIX 4.3.3, PSSP 3.2, PE 3.1

ESSL and Parallel ESSL

Engineering and Scientific Subroutine Library (ESSL) is not directly dependent on a level of PSSP. ESSL coexistence and migration is however, the same as for Parallel Environment because it is dependent on the Parallel Environment LPP. Parallel ESSL 2.2 supports PE 3.1. These packages are discussed in more detail in Chapter 9, “Software support” on page 261.

LoadLeveler

LoadLeveler is the PSSP-related LPP that supports scheduling and load balancing of parallel jobs on the SP system. The supported levels are shown in Table D-10.

Supported Levels

Table D-10 Supported levels in relationship to AIX and PSSP supported levels

LoadLeveler	PSSP 2.4	PSSP 3.1.1	PSSP 3.2	AIX 4.1.5	AIX 4.2.1	AIX 4.3.3
2.2	N	S	S	N	N	S
2.1	S	S	S without PE	N	N	S
1.3	S	N	N	S	S	S

S = Supported

N = Not supported

Migration considerations

LoadLeveler 2.2 supports node by node migration from LoadLeveler 2.1 (if a PTF is applied to 2.1). LoadLeveler provides other mechanisms for migration from LoadLeveler 1.3 to 2.1, including the use of separate LoadLeveler clusters.

Migration limitations:

1. If LoadLeveler and PE are used with PSSP 3.2, the migration to PSSP 3.2, LoadLeveler 2.2 and PE 3.1 must be completed simultaneously.
2. LoadLeveler 1.3.0 with PSSP 2.4 and AIX 4.3.2 can be migrated to LoadLeveler 2.1 without changing PSSP and AIX levels.
3. LoadLeveler 2.1 permits the migration of PSSP 2.4 or 3.1 and AIX 4.3.2 to PSSP 3.2.

Coexistence

LoadLeveler 2.2 can coexist with LoadLeveler 2.1 within a cluster (PTF on 2.1 is required for this). LoadLeveler 2.2 does coexist with PSSP 3.1.1 and 3.2 and AIX 4.3.3. LoadLeveler 1.3.0 coexists with PSSP 2.2, 2.3 and 2.4 and the corresponding levels of AIX. LoadLeveler 2.1 coexists with PSSP 2.4 and 3.1 and AIX 4.3.3. LoadLeveler 2.1 coexists with PSSP-3.2 and AIX 4.3.3, without PE.

Coexistence of LoadLeveler with Parallel Environment: When LoadLeveler and Parallel Environment exist on the same node, they must be at one of the following combinations:

- LoadLeveler 2.2 with Parallel Environment 3.1
- LoadLeveler 2.2 with Parallel Environment 2.4
- LoadLeveler 2.1 with Parallel Environment 2.4
- LoadLeveler 1.3 with Parallel Environment 2.3
- LoadLeveler 1.3 with Parallel Environment 2.2

E

SP/CES Tools

This chapter presents several tools that are used in SP/CES systems management. The tools are:

- Network Installation Manager (for installation and migration)
- Service Director
- Electronic Service Agent (to alert user and IBM about a problem)
- Inventory Scout (to check if the last available microcode is needed and gather the VPD)

Network Installation Manager

Network Installation Manager, or NIM, is an efficient tool to manage an installation or migration.

This appendix will highlight some characteristics of the way NIM operates in an SP environment including CES environment. Although the concepts behind NIM remain the same as for RS/6000, the way in which we configure and manage it is fundamentally different.

Customers may already be familiar with the classic NIM for RS/6000. However, if a server is attached to an SP or a CES, the SP NIM method must be used. The appendix assumes the user is familiar with both the RS/6000 NIM and SP NIM. For more detailed information on NIM, reference the redbook *NIM: From A to Z in AIX 4.3*, SG24-5524.

Creating a CES SP system as a clustered node system uses NIM intensively; adding the Enterprise Server to an SP system, or creating a cluster from several servers, still uses the SP NIM documentation and specification. As soon as a Enterprise Server is part of an SP or Clustered Enterprise Server (CES) it is seen as a SP node from the NIM point of view and, consequently, be referenced here as node or SP node.

The SP has some rules about what is and what is not supported. Each node is defined as a stand-alone machine; diskless or dataless nodes are not supported. Nodes must also be installed over the SP's internal Ethernet network, although it would be possible to boot a node over, for example, a token ring and install it; however, this is not supported.

The resource objects remain the same, though there are several scripts that are automatically allocated as resources when a node is installed, which would normally be an option on a classic RS/6000 environment. Finally, the group resource is not used within the SP environment.

Differences with SP/CES versus R/S6000 NIM

The main differences to be considered are:

1. setup_server
2. SDR
3. Installation overview
4. Boot/install server requirements
5. Netbooting
6. Node customization

7. Wrappers

setup_server

setup_server is a crucial function since it is called on every node and BIS (Boot/Install Server) on bootup and is central to configuring and maintaining the SP NIM environment. For more detailed information on boot/install servers, refer to PSSP: Administration Guide, SA22-7348.

The **setup_server** Perl source code configures and controls a huge portion of the SPs environment but is smaller than expected with less than 1000 lines of code.

This is because **setup_server** (since PSSP 2.2) is modular. It can be broken down into stand-alone pieces of code (wrappers), such as delnimmast, allnimres, mkconfig, and so forth. Each one of these wrappers is called as part of **setup_server** but can also be called manually from the command line.

Exactly how **setup_server** runs these wrappers depends upon how the environment is currently configured and the changes that have to be made, such as allocating some resources or building an extra boot install server.

setup_server only has a single allowable flag with which it can be called: -h. However, all this does is display the help text.

setup_server runs on every boot of a node, boot/install server, or the CWS (called from /etc/rc.sp). However, setup_server will also need to run manually whenever we are changing an element of a node install.

For example, if the node is set to install and we set it back to disk, we need to run **setup_server** so that it will deallocate the NIM resources, remove the nodes entry in /etc/tftpboot, unexport the exported file systems, and so on.

System Data Repository

The System Data Repository (SDR) is something unique to the SP. It contains SP-specific information in a central repository that resides on the CWS. It is held in plain ASCII text format under the /spdata/sys1/sdr/ directory structure.

All nodes within the SP system may need to query data or change an attribute of the SDR at some point; so, on the CWS, the sdrd daemon runs to handle requests from the nodes and itself (the CWS) and listens on TCP port 5712.

The activity of this daemon is also written to a log in /var/adm/SPlogs/sdr/sdrlog.syspar_ip_addr.pid where syspar_ip_addr is the IP address of the system partition, and pid is the process identifier (PID) of the SDR daemon.

SDR is similar in concept to the AIX Object Data Manager (ODM) and is a database containing Classes, Attributes, and Objects. It is crucial to the installation procedure. In order to install our SP system, we must initialize and populate the SDR with configuration information about the environmental details of the SP, the frame(s), and the nodes. When we then run the setup_server script for the first time, it queries the SDR in order to retrieve the relevant information to make and define the NIM objects necessary to go ahead with the installation.

To provide support for CES, a new system-wide boolean attribute is created in SP class of the SDR called IsPartitionable (set to false for SP Switch2 and CES).

Installation process

This is a very brief overview of the SP installation process concentrating on the steps linked with creating the NIM resources; for detailed information on the complete installation process, refer to the following guides:

- *IBM RS/6000 SP: Planning Volume 2, Control Workstation and Software Environment*, GA22-7281

- *PSSP: Installation and Migration Guide*, GA22-7347

The first part of the installation procedure deals with setting up the CWS and making sure it has the necessary prerequisites to actually act in this capacity.

Connect the frames to the CWS, then configure and tune the network adapters (the PSSP installation and migration guide has some specific tuning recommendations).

The second part of the installation process deals with getting PSSP installed on the CWS.

Copy the PSSP installation images over to disk. This step has some relevance to NIM because the script, pssp_script, which setup_server defines as a script object, remotely mounts the directory in which the PSSP images reside and installs PSSP on the nodes as part of the customization.

Next, copy a basic AIX (mksysb) image to disk. Again, this step has particular relevance to NIM because the mksysb image will later be defined as a NIM mksysb object. You can now install PSSP on the CWS. Then, initialize SP Authentication Services.

Run install_cw. This completes the configuration of the CWS, and this final step, among other things, installs the PSSP smitty panels, configures the SDR, and starts the SP daemons.

- At this step the installation of the CWS is complete, we can go ahead and enter our configuration details for the frame(s) and nodes into the SDR.
  - Enter Site Environment Details
  - Enter Frame Information and Reinitialize the SDR.
  - Verify System Monitor Installation by running smitty SP_verify
  - Enter the required Node Information.

There are many additional steps that can be followed to further customize the nodes, most significantly setting up the switch (if your SP system has one). However, at this point of the installation, we now have all the node and frame information in the SDR.

Boot/Install server requirements

By default, a single frame SP system will configure the CWS as the only Boot/Install server (BIS) - the NIM Master.

In a multiple frame SP system, by default, the CWS will act as the BIS for the first frame and then the first node in each additional frame will be configured to act as a BIS for the rest of the nodes in its own frame.

In a large CES configuration a BIS can always be selected to improve the installation performance.

Netbooting

To netboot an RS/6000, we would turn the keyswitch and go into SMS or use a ROM IPL disk, depending on the model. On the SP, this is all handled for us, thanks to our connection with the frame supervisor card and serial connection. In order to netboot a node, we simply need to click on netboot on the perspectives panel or, from the command line, use the **nodecond** command. It is possible to manually netboot the node if you wish to use the **spmon** command to change the power or keyswitch setting and a writable s1term in order to enter the boot address or, in the case of a PCI node, enter SMS mode.

Using the **nodecond** command, or perspectives to netboot a node, also provides a log file to check for errors, although you can also watch the progress of a netboot by using a read-only s1term. Do not use the open a console option from the perspectives panel if you want to simply watch the netboot progress of a node because this will open a writable console and stop automatic node conditioning.

Node customization

The customization of the SP nodes is performed by a Korn shell script called pssp_script. This is used by NIM after a migration or installation of a node.

setup_server defines this as a NIM resource of the type script; therefore, it is run on the node before NIM reboots it.

pssp_script is also run on bootup if a node is set to customize in the SDR. In order to detect this, the script, /etc/rc.sp, which is run from inittab, checks the bootp response of the node in the SDR, and, if it is set to customize, spawns off the pssp_script process.

During the customization phase, pssp_script configures the node's environment based on the data in the two files in the /tftpboot directory:

```
<node>.config_info and <node>.install_info
```

where <node> is the hostname of the node in question.

These two files are created by the setup_server wrappers, mkconfig and mkinstall.

Although the main function of pssp_script is to install the PSSP software and configure it. The script has several more notable functions. Among its extra tasks are configuring a separate dump logical volume, updating the /etc/hosts file, and starting or stopping volume group mirroring.

The final part of node customization is performed after the node is rebooted. As part of an installation or migration operation, pssp_script places an entry for a script called spfbcheck in /etc/inittab. pssp_script copies this script, along with another called psspfb_script from the directory /usr/lpp/ssp/install/bin on the BIS to the local /tftpboot directory on the node.

On reboot, the /tftpboot/spfbcheck script is run, which renames the /tftpboot/psspfb_script so that it is not run again accidently, and executes it. It then removes its own entry from /etc/inittab to stop itself from being run on subsequent boots.

The main job of psspfb_script is to configure the network adapters that have previously been defined in the SDR. Once this has been done, the final stage is for the script to set the bootp response field of the node back to disk in the SDR.

Wrappers

In an SP environment, we never configure NIM manually. Instead, the PSSP software package has scripts called *wrappers* for configuring the NIM environment. Wrappers are functions written in Perl and each one is an independent script.

These wrappers are called by the **setup_server** command for configuring the NIM environment in the CWS and Boot/Install servers (BIS). These scripts can also be called from the command line if you know the sequence in which they have to be executed. These scripts use the information from the SDR to initiate the appropriate NIM commands with the required options.The NIM wrappers are part of the ssp.basic filesets. They are:

- mknimmast: This wrapper initializes the NIM master.
- mknimint: This wrapper creates network objects on the NIM master to serve the NIM clients.
- mknimclient: This wrapper takes input from the SDR to create the clients on the CWS and the BIS that are configured as NIM masters.
- mkconfig: This wrapper creates the /tftpboot/<reliable_hostname>.config_info file for every node.
- mkinstall: This wrapper creates the /tftpboot/<reliable_hostname>.install_info file for every node in the SDR whose bootp_response is not set to *dis*k.This file is used during the network installation of the nodes.
- mknimres: This wrapper creates all the NIM resources for installation, diagnostics, migration and customization. The resources created will be used by the allnimres wrapper for allocation to the clients, depending on the bootp_response field.
- delnimmast: This wrapper is used to delete the NIM master definitions and all the NIM objects on the CWS or boot/install server.
- delnimclient: This wrapper is used to deallocate the resources and to delete the NIM clientdefinitions from the NIM master.
- allnimres: This wrapper is used to allocate the NIM resources to the clients depending on the value of the bootp_response attribute defined in the SDR. If the bootp_response value is set to install, migration, maintenance or diagnostic, the appropriate resources will be allocated. In case of a disk or customize value, all the resources will be deallocated.

Refer to *PSSP: Command and Technical Reference*, SA22-7351.

Service Director

Note:

- Service Director for SP was widely distributed and used, it is now replaced by Electronic Service Agent.
- Service Director is still supported and available.

Service Director overview

Service Director for RS/6000 is an application program that operates on a customer's IBM RS/6000. It is invoked by the proper installation of the program and successful registration, set up, customizing, and distribution of entitlement data and application code.

Service Director for RS/6000 is offered as a part of the IBM Warranty or IBM Maintenance Service package for no additional charge. Due to entitlement checking at the IBM server, machines not on IBM Warranty or Service cannot be entitled to this service.

Advantages of Service Director for RS/6000

The Service Director application can automatically report hardware related problems to IBM for service by the modem on the local server. The Service Director application can also automatically do problem analysis on those problems before calling for service. This level of Service Director supports all classic RS/6000 machine types and the 9076. Classic RS/6000 refers to machines that have concurrent diagnostics installed. Some of the PCI-based PowerPC machines had diagnostics on CD-ROM, and were not concurrent.

Service Director aids IBM Service Representatives in problem source identification, and it may be used to automatically place service calls to IBM for most hardware errors. System errors are dynamically monitored and analyzed; no customer intervention is required. Service Director for RS/6000 further simplifies error analysis for some errors once the Service Representative is onsite by using the Service Director event log. Customers' hardware error logs can now be reduced, because errors are being maintained within Service Director event

In summary, the advantages of Service Director for RS/6000 include:

- Automatic problem analysis
- Manual creation of problem reports
- Automatic customer notification
- Automatic problem reporting

- Service calls placed to IBM without operator intervention
- Vital Product Data (VPD) reported to IBM
- Common focal point for service interface
- Problem-definable threshold levels for call placement
- Reduced hardware error logs
- High-availability cluster multiprocessing (HACMP) support for full fallback; includes high-availability
- Cluster workstation (HACWS) for 9076
- Simple and complex environment support with minimum analog lines

Service Director components

The Service Director (SD) contains three components:

Product Support Application (PSA)

The PSA determines appropriate error disposition, then captures and passes information required to resolve any problem identified on a particular product or option to the analysis engine.

Analysis Routine

The analysis routine within Service Director schedules the execution of the PSAs. They may be configured to run constantly or on a specific time schedule. When the analysis routine runs, it monitors errors or events identified by a PSA. Errors or events are logged and depending on customer configured options, the analysis routine may notify a person or persons within the customer account and automatically transmit hardware errors and associated problem information to an IBM support center for remote analysis and action. If necessary, an IBM Service Representative is dispatched to the customer site with the parts needed to correct the problem reported.

Display Routine

The display function is the user's interface to Service Director for RS/6000. It provides a structured view of problem management information, such as: Status of recent hardware events logged by Service Director History of hardware events Remote viewing of client event logs Configuration and run options Lock and reset functions.

Obtaining Service Director code

- Service Director Code can be obtain by the following ways;
- SD can is delivered with any new RS/6000 SP machine,
- SSR can get it from CORE

- SD may be obtain using Internet from:
- `ftp://ftp.software.ibm.com/aix/servdir_client_code/`

Electronic Service Agent

The Electronic Service Agent (SA) is a software that resides on your system to monitor events and transmit data to IBM.

Electronic Service Agent has Product ID 5765-50400. It is replacing the Service Director. This is a no-charge function from IBM. Service Agent monitors system hardware error logs and automatically reports identified problems to IBM while the machine is under a service agreement or within the warranty period. The Service Agent collects system inventory and performance information to enable service improvements and to enable the provision of electronic services offerings.

Knowing about malfunctions early allows for proactive service delivery, which assists you in maintaining higher availability and performance. It allows IBM Service to respond faster and with the correct repair parts. In some cases, IBM may notify a customer of an error and provide a resolution prior to an outage.

Information collected through Service Agent will be made available to IBM service support representatives when they are helping to answer questions or diagnosing problems.

Terminology

The following terms that are used in SA:

IBM Service Agent Server (SAS)

The Service Agent Server (SAS) is located at the IBM company. It is the machine to which your computer(s) sends information that is stored, analyzed, and acted upon by IBM.

Gateway machine

The Gateway machine is the host system where Service Agent is first installed using the svcagent.installp. The Gateway machine contains the central database and the processes for controlling the Service Agent system processes and communications to IBM.

Forwarder

A machine or system interface defined to the Gateway database as a client or interface that forwards requests to or from other client machines.

What does Service Agent do

Here are some of the key things you can accomplish using Electronic Service Agent for pSeries and RS/6000:

- Automatic problem analysis.
- Problem-definable threshold levels for error reporting.
- Automatic problem reporting; service calls placed to IBM without intervention.
- Automatic customer notification.
- Commonly viewed hardware errors; you can view hardware event logs for any monitored machine on the network from any Service Agent host user interface.
- High-availability cluster multiprocessing (HACMP) support for full fallback; includes high-availability cluster workstation (HACWS) for 9076.
- Network environment support with minimum telephone lines for modems.

How does Service Agent work

Machines are defined and Service Agent installed by using the Service Agent user interface. After machines are defined, they are registered with the IBM Service Agent Server (SAS).

During the registration process, an electronic key is created that becomes part of your resident. Service Agent program. This key is used each time Service Agent places a call for service. The IBM Service Agent Server checks the current customer service status from the IBM entitlement database; if this reveals that you are not on Warranty or MA, then the service call is refused and posted back via e-mail notification.

Service Agent is not designed to arbitrarily pick up any general information without it having been programed to do so. However, there is some data that Service Agent does bring to IBM to help with problem resolution. This information consists of the problem or error information itself and Vital Product Data (VPD) or Inventory data.

Usually, the only data, besides error information, being sent to IBM is Vital Product Data (VPD), which is generated by either the **`lscfg`** command or the new invscout program. You can run one of these commands on a system and determine if this information is of a sensitive nature or not.

Prerequisites:

- Ensure your pSeries or RS/6000 is at Version 4.1 or above with concurrent diagnostics installed. 9076 Parallel System Support Program (PSSP) code must be at Release 1.2 or above.
- Java is required on all monitored machines. Java for AIX 4.3.3 is on the system disk.You will have to obtain Java from another source for machines with AIX versions earlier than 4.3.3. Java versions supported are 1.1.6 to 1.3.
- Ensure your Gateway server has an available serial port for the local modem if used. A TTY (0 - 16) device must be available and configured on the gateway.
- A modem is required to call the IBM Service Agent Server (SAS). For security, only outbound calls are required by Service Agent, so the auto answer capability of the modem should be disabled. An asynchronous modem with a minimum communications speed of 9600 baud and error correction (in the United States) is required. Please refer to local procedures in your country to see what the modem requirements are for Service Agent.
  - IBM ships the following modem types for use with Service Agent on some products: 7852 Model 400, 7857-017 or 7858-336
- On your Gateway server, ensure Point-to-Point Protocol (PPP) is installed and configured. PPP is only required if a modem is going to be used for error reporting to IBM.

Electronic Service Agent in SP/CES environment

Electronic Service Agent in SP/CES environment differ from basic R/6000 and pSerie SA in two ways,

1. In SP or in SP Attached Server environment, the CWS (or SA Gateway) must be set to Machine Type (M/T) 9076.
   - The nodes are added using the Add SP Nodes function.
2. In CES environment the servers are added using the Add machines function.
   - in the model type enter C80. This information is used to create a proper RETAIN record at IBM. With the C80 information the RETAIN call is forwarded to the correct queue monitored by CES skilled people.

Note: For the performance reason, sometime it is not recommended to install the SA Gateway in the CWS.

Obtaining Electronic Service Agent

Electronic Service Agent for pSeries and RS/6000 User's Guide, SC38-7105 is available at:

`ftp://ftp.software.ibm.com/aix/service_agent_code`

Inventory Scout (invscout)

Inventory Scout is a tool which contains Microcode Discovery Service and, in SP and CES, Vital Product Data (VPD) Capture Service. These two services are made possible through a new AIX command **`invscout`**, and a new daemon invscoutd.

What does invscout do

1. Scouts the user's system for microcode levels and compares them with an IBM database of latest levels. When used with web-based, Microcode Discovery Service, a html report is generated for the user that includes links to the latest levels of microcode for the system, should it find the currently installed level is downlevel.

   In the case of the SP and CES systems, the customer will not be given a link to the latest microcode, but will be directed to "Contact CE" to obtain the new level of microcode.

   In the html report for all RS/6000 systems, a link will be provided to the README of the latest level of microcode.

2. Gathers Vital Product Data (VPD) from the user's machine and when used with web-based VPD Capture Service, uploads it to IBM's MRPD database. This VPD will be helpful in determining the correct components to ship when a Miscellaneous Equipment Specification (MES) upgrade is ordered. Initially, only the 9076 SP VPD will be stored in the MRPD. All other RS/6000 systems are being targeted to work with MRPD mid 2001.

   Inventory Scout runs on AIX V4.1.5 or higher, an can be invoked by Java applets or run from the command line. The last level of Inventory Scout is 1.2.0.1

Where can I locate these services

Invscout and invscoutd were included in the 10/00 AIX 4.3.3 update and AIX 5.0 releases. You do not need to be at one of these levels to use this command/daemon and Services. These can also be obtained from the following Internet Service links:

Microcode Discovery Service

http://techsupport.services.ibm.com/rs6k/mds.html

VPD Capture Service

Http://techsupport.services.ibm.com/rs6k/vcs.html

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information on ordering these publications, see “How to get IBM Redbooks” on page 440.

- *IBM RS/6000 Clustered Enterprise Servers Systems Handbook,* SG24-5120
- *RS/6000 Systems Handbook 2000 Edition*, SG24-5120
- *Exploiting HACMP 4.4: Enhancing the Capabilities of Mutil-Processing, SG24-5979*
- *IBM @server pSeries 680 Handbook including RS/6000 Model S80*, SG24-6023

Other resources

These publications are also relevant as further information sources:

- *IBM RS/6000 SP: Planning Volume 2, Control Workstation and Software Environment*, GA22-7281
- *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference, SA22-7351*
- *HACMP Administration Guide*, SC23-4279

Referenced Web sites

These Web sites are also relevant as further information sources:

- `ftp://ftp.software.ibm.com/aix/service_agent_code`

  Site to get the code and the documentation for Electronic Service Agent
- `http://techsupport.services.ibm.com/rs6k/mds.html`

  Microcode Discovery Service

- `http://techsupport.services.ibm.com/rs6k/vcs.html`

  VPD Capture Service
- `http://www.ibm.com/servers/eserver/pseries/library/`

  pSeries servers publications website
- `http://www.austin.ibm.com/support/micro/download.html`

  Service Processor and system firmware microcode download website

How to get IBM Redbooks

Search for additional Redbooks or redpieces, view, download, or order hardcopy from the Redbooks Web site:

ibm.com/redbooks

Also download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

Redpieces are Redbooks in progress; not all Redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

IBM Redbooks collections

Redbooks are also available on CD-ROMs. Click the CD-ROMs button on the Redbooks Web site for information about all the CD-ROMs offered, as well as updates and formats.

Special notices

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

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, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere.,The Power To Manage., Anything. Anywhere.,TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States and/or other countries.

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

UNIX is a registered trademark in the United States and other countries licensed exclusively through The Open Group.

SET, SET Secure Electronic Transaction, and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.

Other company, product, and service names may be trademarks or service marks of others

Abbreviations and acronyms

API	Application Programming Interfaces
ATM	Asynchronous Transfer Mode
BLAS	Basic Linear Algebra Subroutines
CSP	Converged Service Processor
CSU	Channel Service Units
CRM	Concurrent Resource Manager
CWS	Control workstation
DAS	LP Dual Attach Station
DAS	UP Dual Attach Station
DCE	Distributed Computing Environment
DFS	Distributed File System
DPCL	Dynamic Probe Class Library
DSU	Data Service Units
EMIF	ESCON Multiple Image Facility
ESSL	Engineering and Scientific Subroutine Library
F/C	Feature Code
FFDC	First Failure Data Capture
GPFS	General Parallel File System
GUI	Graphical User Interface
HACMP	IBM High Availability Cluster Multi-Processing for AIX
HACWS	High Availability Control Workstation
IBM	International Business Machines Corporation
IP	Internet Protocol
ITSO	International Technical Support Organization
JFS	Journaled File System
LANE	LAN Emulation
LAPI	Low-level Application Programming Interface
LP	Linear Programming
MCA	Micro Channel
MIP	Mixed-integer programming
MPL	Message Passing Libraries
NFS	Network File System
NTP	Network Time Protocol
OPS	Oracle Parallel Server
OSL	Optimization Subroutine Library
OSLp	Parallel Optimization Subroutine Library
PCI	Peripheral Component Interconnect
PE	Parallel Environment
PESSL	Parallel Engineering and Scientific Subroutine Library
POE	Parallel Operating Environment
PSSP	Parallel System Support Programs
PVC	Permanent Virtual Circuits
QP	Quadratic Programming
RAN	Remote Async Nodes
RSCT	RS/6000 Cluster Technology
RVSD	Recoverable Virtual Shared Disk
SAS	LP Single Attach Station
SDR	System Data Repository
SMP	Symmetric Multiprocessors
SVC	Switch Virtual Circuits

Tcl	Tool command Language
US	User Space
UTP	Unshielded Twisted Pair
VSD	Virtual Shared Disk
WAN	Wide Area Network

Index

Symbols

/etc/hosts 430
/etc/inittab 430
/etc/rc.sp 430
/spdata/sys1/install/pssplpp 348
/spdata/sys1/sdr 427
/tftpboot 430
/tftpbooy/pssfb_script 430
/usr/lpp/ssp/install/bin 430
/vad/adm/SPlogs/css 176
/var/adm/SPlogs/css0 176

Numerics

1.2 MB/sec Rack Mountable Remote Asynchronous Node, 16-port, EIA-232 239
1.2 MB/sec Remote Asynchronous Node, 16-port, EIA-232 (US) 239
1.2 MB/sec Remote Asynchronous Node, 16-port, EIA-232 (world trade) 239
128-port Asynchronous Controller Cable 23 cm (9 in.) 239
128-port Asynchronous Controller Node Cable, 4.5 m 239
2.4 MB/sec Enhanced Remote Asynchronous Node, 16-port, EIA-232 239
2.4 MB/sec Enhanced Remote Asynchronous Node, 16-port, RS-422 239
332 MHz SMP architecture 25
332 MHz SMP node 61
  architecture 61
  I/O bridge 66
  I/O bridge bus 65
  L2 cache 63
  memory-I/O controller 64
  microprocessor 63
  system bus 62
332 MHz SMP thin node
  bus 72
  configuration rule 72
  disk option 74
  disk requirement 74
  memory option 74
  memory requirement 74
  option 73
  processor option 74
  processor requirement 74
  requirement 73
  switch adapter option 75
  switch adapter requirement 75
  switch restriction 75
332 MHz SMP thin nodes 71
332 MHz SMP wide node
  adapter placement restriction 67
  bus 67
  disk option 70
  disk requirement 70
  memory option 70
  memory requirement 70
  option 69
  processor option 69
  processor requirement 69
  requirement 69
  switch adapter option 71
  switch adapter requirement 71
  switch restriction 71
375 MHz POWER3 architecture 25
375 MHz POWER3 SMP Thin Node 48
375 MHz POWER3 SMP Wide Node 43
7025-6F1 17
7026-6H1 17
7026-F50 17
7026-H80 17
7044-170 18
8-Port Async Adapter - EIA-422A 234
9076-555 12

A

accounting 263
advantages of an SP/CES 4
ARTIC960Hx 4-port E1 RJ45 cable 258
ARTIC960Hx 4-port EIA-232 cable 241
ARTIC960Hx 4-port EIA-530 cable 241
ARTIC960Hx 4-port RS-449 cable 241
ARTIC960Hx 4-port T1 RJ45 cable 258
ARTIC960Hx 4-port V.35 (DTE) cable 241
ARTIC960Hx 4-port X.21 cable 241

ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 3 m 4-port cable 258
ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 7.5 m extension cable 258
ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-grounded, 1.8 m 4-port cable 258
ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-ungrounded, 1.8 m 4-port cable 259
ARTIC960RxD Quad DTA, H.100, 4-drop cable 259
ARTIC960RxD Quad DTA, T1, 100 ohm, 15 m extension cable 258
ARTIC960RxD Quad DTA, T1, 100 ohm, 3 m 4-port cable 258
Ascend 182
Ascend GRF 1600 182
Ascend GRF 400 182
Asynchronous Terminal/Printer Cable, EIA-232 239
available nodes 26

B

base power regulators for SP frames 143
Basic Linear Algebra Subprograms 274
BLAS 274
business intelligence 8

C

Capacity on Demand 23
Capacity Upgrade on Demand 80
CEC (Central Electronics Complex) 80
CES (Clustered Enterprise Servers) 139
Chipkill DIMMs 106
Chipkill Memory 105
circuit breaker for SP computer room 143
Clustered Enterprise Server 15
Clustered Enterprise Servers 13
CoD (Capacity on Demand) 18
Commands
  lscfg 435
  mksysb 200
  nodecond 429
  rvsdrestrict 419
  setup_server 431
  spadaptrs 352
  spdelfram 337
  spframe 337, 338, 347, 349
  spmon 429
Communication Subsystems Support 265
conductor size for SP computer room 143
control workstation 195
  interface adapter requirement 201
  software requirement 203
  supported IBM RS/6000 workstation 198
  system requirement 199
Cornell Theory Center 270
CRM (Customer Relationship Management) 8
CSP (Convert Serial Processor) 338
CUoD (Capacity Upgrade on Demand) 79
CVSD (Concurrent Virtual Shared Disks) 5
CWS (Control Workstation) 2

D

daemons
  invscoutd 437
DCE (Distributed Computing Environment) 6, 264
Deep Blue 23
DFS 268
Directories
  /spdata/sys1/install/pssplpp 348
  /spdata/sys1/sdr 427
  /usr/lpp/ssp/install/bin 430
  /var/adm/SPlogs/css 176
  /var/adm/SPlogs/css0 176
  /var/adm/SPlogs/css0/p0 176
Distributed File System 268
diverse cluster solutions 18
drawer 140
Driectories
  /tftpboot 430
dynamic CPU de-allocation 6

E

EASY 270
e-business 8
ECC (Error Checking and Correcting) 81
EIA 232D/V.24 cable 242
electrical current for SP power supplies 143
Electronic Service Agent 425
ERP (Enterprise Resource Planning) 3, 8
ESSL (Scientific Subroutine Library) 261
Event Perspective 340
expansion frame 155
expansion frames 146

F

feature code
1241 Independent SCSI Hookup 71
1500 15
1550 15, 37
18.2 GB Ultra SCSI 10K RPM disk pair 40
2031 15
2032 15
2050 14
2050 332 MHz SMP thin node 71
2051 13
2051 332 MHz SMP wide node 67
2054 13
2054 POWER3 SMP High Node 55
2056 14
2056 375 MHz POWER3 SMP Thin Node 48
2057 13
2057 375 MHz POWER3 SMP Wide Node 43
2058 13
2058 375 MHz POWER3 SMP High Node 37
2402 IBM Network Terminal Accelerator - 256 Session 234
2403 IBM Network Terminal Accelerator - 2048 Session 231
2415 SCSI-2 Fast/Wide Adapter/A 234
2416 SCSI-2 Differential Fast/Wide Adapter/A 232
2420 SCSI-2 Differential External I/O Controller 232, 234
2709 ARTIC960Hx 4-port T1 RJ45 cable 258
2710 ARTIC960Hx 4-port E1 RJ45 cable 258
2861 ARTIC960Hx 4-port EIA-232 cable 241
2862 ARTIC960Hx 4-port RS-449 cable 241
2863 ARTIC960Hx 4-port X.21 cable 241
2864 ARTIC960Hx 4-port V.35 (DTE) cable 241
2865 ARTIC960Hx 4-port EIA-530 cable 241
2871 ARTIC960RxD Quad DTA, T1, 100 ohm, 3 m 4-port cable 258
2872 ARTIC960RxD Quad DTA, T1, 100 ohm, 15 m extension cable 258
2873 ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 3 m 4-port cable 258
2874 ARTIC960RxD Quad DTA, E1, 120 ohm balanced, 7.5 m extension cable 258
2875 ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-grounded, 1.8 m 4-port cable 258
2876 ARTIC960RxD Quad DTA, E1, 75 ohm unbalanced-ungrounded, 1.8 m 4-port cable 259
2877 ARTIC960RxD Quad DTA, H.100, 4-drop cable 259
2900 4.5 GB Ultra SCSI disk drive 70, 75
2904 4.5 GB Ultra SCSI disk drive pair 70, 75
2908 9.1 GB Ultra SCSI disk drive 70, 75
2909 9.1 GB Ultra SCSI disk drive pair 70, 75
2918 18.2 GB Ultra SCSI disk drive pair 70, 75
2920 Token Ring Auto LANstreamer 237
2934 Asynchronous Terminal/Printer Cable, EIA-232 239
2940 8-Port Async Adapter - EIA-422A 234
2943 EIA 232/RS-422 8-Port Asynchronous Adapter 237
2944 WAN RS232 128-port 238
2947 ARTIC960Hx 4-port Selectable Adapter 240
2951 EIA 232D/V.24 cable 242
2952 V.35 cable 242
2953 V.36/EIA 449 cable 242
2954 X.21 cable 242
2962 2-port Multiprotocol X.25 Adapter 241
2963 2-port Multiprotocol X.25 Adapter 242
2968 Ethernet 10/100 MB 243
2969 Gigabit Ethernet - SX 244
2984 TURBOWAYS 100 ATM Adapter 245
2988 ATM 155 MMF 246
3124 38
3124 Serial port to serial port cable for drawer-to-drawer connections 239
3125 38
3125 Serial port to serial port cable for rack-to-rack connections 239
4008 17
4011 16
4012 17
4020 168
4020 SP Switch Adapter 168
4021 SP Switch Router Adapter 169, 187
4022 168
4022 SP Switch MX Adapter (withdrawn) 168
4023 169
4023 SP Switch MX2 Adapter 168
4025 SP Switch2 Adapter 178
4093 Base Memory Card 70, 74
4110 One Pair of 128 MB DIMMs 70, 74
4350 38
4958 IBM PCI Cryptographic Coprocessor 230
6206 Ultra SCSI Single Ended 251
6207 Ultra SCSI Differential 252
6208 SCSI-2 F/W Single-Ended 253

6209 SCSI-2 F/W Differential 254
6215 SSA RAID 5 254
6222 SSA Fast-Write Cache Module 255
6310 ARTIC960RxD Quad Digital Trunk Adapter 258
8130 1.2 MB/sec Remote Asynchronous Node, 16-port, EIA-232 (US) 239
8131 128-port Asynchronous Controller Node Cable, 4.5 m 239
8132 128-port Asynchronous Controller Cable 23 cm (9 in.) 239
8133 RJ-45 to DB-25 Converter Cable 239
8134 1.2 MB/sec Remote Asynchronous Node, 16-port, EIA-232 (world trade) 239
8136 1.2 MB/sec Rack Mountable Remote Asynchronous Node, 16-port, EIA-232 239
8137 2.4 MB/sec Enhanced Remote Asynchronous Node, 16-port, EIA-232 239
8138 2.4 MB/sec Enhanced Remote Asynchronous Node, 16-port, RS-422 239
feature code 2904
4.5 GB Ultra SCSI disk pair 46, 51
feature code 2909
9.1 GB Ultra SCSI disk pair 40, 46, 51
feature code 2918
18.2 GB Ultra SCSI disk pair 40, 46, 51
feature code 3800
9.1 GB Ultra SCSI disk pair 42
feature code 3802
9.1 GB SSA disk pair 42
feature code 3803
18.2 GB Ultra SCSI disk pair 42
feature code 3804
9.1 GB Ultra SCSI 10K RPM disk pair 40, 46, 51
feature code 3805
9.1 GB Ultra SCSI 10K RPM disk pair 42
feature code 3810
18.2 GB Ultra SCSI 10K RPM disk pair 46, 51
feature code 3811
18.2 GB Ultra SCSI 10K RPM disk pair 42
feature code 3812
18.2 GB SSA disk pair 42
feature code 3820
36.4 GB Ultra SCSI 10K RPM disk pair 40, 46, 58
feature code 3821
36.4 GB Ultra SCSI 10K RPM disk pair 42
feature code 3822
36.4 GB SSA 10K RPM disk pair 42
feature code 4008
SP Switch-8 293
feature code 4011 40
feature code 4012 40
feature code 4023
SP Switch MX2 Adapter 71
feature code 6207
Ultra SCSI adapter 40
feature code 6230
SSA adapter 40
feature code 8396
SP System Attachment Adapter 79
file collections 262
Files
/etc/hosts 430
/etc/inittab 430
/etc/SDR_dest_info 348
adapter.log 176
cable.miswire 176
css.snap.log 176
daemon.log 176
deamon.log 176
dtbx.trace 176
dtbx-failed.trace 176
Ecommands.log 176
flt 176
fs-deamon-print.file 176
logevnt.out 176
out.top 176
rc.switch.log 176
router.log 176
scan-out.log 176
scan-save.log 176
topology.data 176
floating-point computation 28
frame dimensions for the SP 142
frame numbering rule 299

G

GPFS (General Parallel File System) 261
GPFS (General Parallel File Systems) 5
GRF 182
grid clusters 18

H

HACMP 221
HACMP (High Availability Cluster Multi-Processing) 261

HACMP (High Availability Cluster Multi-Processing) 5
HACWS 220
HACWS (High Availability Control Workstation) 18
Hardware Perspective 340
high availability 5
High Availability Cluster 19
High Availability Cluster Multi-Processing for AIX 221
High Availability Control Workstation 220
  limitation and restriction 224
  software requirement 223
  system requirement 223
High Nodes 13
high nodes 13

I

IBM 9077 SP Switch Router model 04S 182
IBM 9077 SP Switch Router model 16S 182
IBM DB2 UDB EEE 8
IBM Network Terminal Accelerator - 2048 Session 231
IBM Network Terminal Accelerator - 256 Session 234
IBM Recoverable Virtual Shared Disk 264
IBM Virtual Shared Disk 264
Informix Dynamic Server AD/XP 8
installation and migration coexistence 265
Intermediate Switch Board 140
internal SP nodes 26
Inventory Scout 425
invscout 437
ISB 140

J

JFS 268
job flow 270
job priority 270
job status 270
job submitter 270
Journalled File System 268

L

LAPI 266
Load Balancing Cluster 18
logical dependent node 16
login control 263
Logs
  /var/adm/SPlogs/sdr/sdrlog.syspar_ip_addr.pid 427
  adapter.log 176
  cable.miswire 176
  css.snap.log 176
  daemon.log 176
  dtbx.trace 176
  dtbx-failed.trace 176
  Ecommands.log 176
  fs-deamon-print.file 176
  logevnt.out 176
  out.top 176
  rc.switch.log 176
  router.log 176
  scan-out.log 176
  scan-save.log 176
  topology.data 176
loose clusters 18
Low-level Application Programming Interface 266
lscfg 435

M

M/T 7017 11, 14, 22
M/T 7026 11, 14, 22
M/T 9076 9, 22
machine type
  9077 RS/6000 SP Switch Router 183
Manageable 23
mathematical function 274
  Basic Linear Algebra Subroutines 274
  Eigensystem Analysis 274
  Fourier Transforms 274
  Linear Algebraic Equations 274
  quadratic programming 274
mathematical programming
  linear programming 274
  mixed integer programming 274
  network programming 274
mathematical subroutine 274
  computational chemistry 274
  computational techniques 274
  dynamic systems simulation 274
  electronic circuit design 275
  fluid dynamics analysis 274
  mathematical analysis 274
  nuclear engineering 275
  quantitative analysis 275

reservoir modeling 274
seismic analysis 274
structural analysis 274
time series analysis 274
Memory-I/O 27
Message Passing Cluster 20
Message Passing Interface 266
mksysb 200
Model 500 15
Model 550 15, 37
Model 555 15
model frame 146
Motif 270
MPI 266
MX (Mezzanine bus) 40

N

Network File System 268
Network Installation Manager 265, 425
NFS 268
NIM 265
NLS 271
node numbering rule 299
nodecond 429
non-switched expansion frame 150

O

Oracle Enterprise Edition 8

P

Parallel applications 20
Parallel Basic Linear Algebra Subprograms 274
Parallel Database Cluster 20
Parallel Operating Environment 278
parallel processing 2
partition 5
PBLAS 274
PCI (Peripheral Component Interconnect) 2
PCI adapters 41
PE (Parallel Environment) 5, 261
Performance Monitor Perspective 340
Perl 266
physical dependent node 16
POE 278
power output for SP frames 143
power supply phases for SP frames 143
POWER3 2
POWER3 SMP architecture 25
POWER3 SMP High Node 37, 55
POWER3 SMP node 52
6xx bus 54
POWER3 microprocessor 53
service processor 31
system architecture 53
system memory 54
POWER3-II 28
PowerPC 2
processor deallocation 32
processor node 25
PSSP (Parallel System Support Programs) 261
PSSP (Parallel Systems Support Programs) 2
PTX (Performance Toolbox) 6, 265
PVM 271

Q

QoS (Quality of Service) 18
Quality of Service 23

R

RAN (Random Asynchronous Node) 314
receptacle for SP computer room 143
Redbooks Web site 440
Contact us xxi
redundancy 5
RJ-45 to DB-25 Converter Cable 239
RS/6000 SP building block 26
Russell-Stol 143
RVSD (Recoverable Virtual Shared Disks) 5
rvsdrestrict 419

S

SAMI (Service and Manufacturing Interface) 338
Scalable Electric Power Base Unit 154
Scalable Linear Algebra Package 274
ScaLAPACK 274
SCM (Supply Chain Management) 8
Scripts
/etc/rc.sp 430
/tftpboot/psspfb_script 430
pssp_script 348, 428
psspfb_script 430
SCSI-2 Differential External I/O Controller 232, 234
SCSI-2 Differential Fast/Wide Adapter/A 232
SCSI-2 Fast/Wide Adapter/A 234

SDR 262
SEPBU 140, 154
Serial port to serial port cable for drawer-to-drawer connections 239
Serial port to serial port cable for rack-to-rack connections 239
Server Consolidation 21
Server consolidation 8
server consolidation 8
Service Director 425
setup_server 427, 431
Shared Memory Parallel Processing 273
short frame 142
Single Program Multiple Data 273
Single-stage system 147
slot 140
slot numbering rule 299
SMP (Symmetric Multiprocessors) 2
SP
  375 MHz High node 32
  375 MHz Nodes 35
  frame dimensions 142
  POWER3 SMP High 375 MHz 32
  POWER3-II SMP Thin node at a glance 35
  POWER3-II SMP Wide node at a glance 35
  Switch Adapter 163
  Switch Chip 161
  Switch Link 160
  Switch Ports 161
  Switch system 164
SP benefits 6
SP features 6
SP frame 140
SP Model 500 142
SP Model 550 142
SP POWER3-II SMP Thin node
  at a glance 35
SP POWER3-II SMP Wide node
  at a glance 35
SP Switch 3
SP Switch (F/C 4011) 16
SP Switch 16-port (F/C 4011) 40
SP Switch Adapter 168
SP Switch chip 161
SP Switch MX2 Adapter 168
SP Switch MX2 Adapter (F/C 4023) 41
SP Switch port numbering rule 299
SP Switch Router 3, 16, 182
  connection to the control workstation 190
  connection to the SP Switch 192
  network interface 189
  network media card requirement 188
  software requirement 189
  system requirement 187
SP Switch Router Adapter 169
SP Switch Router model 04S 185
SP Switch Router model 16S 185
SP Switch2 3
  Hardware
    740 microprocessor 180
    MIC chip 179
    NBA chip 179
    TBIC3 chip 179
  J-TAG 177
SP Switch2 (F/C 4012) 17
SP Switch2 16-port (F/C 4012) 40
SP Switch2 Adapter (F/C 4025) 41
SP Switch-8 (F/C 4008) 17
SP switches 286
spadaptrs 352
SP-attached server 14, 15
SP-attached servers 11, 13
spdelfram 337
spframe 337, 338, 347, 349
spmon 429
SSA Fast-Write Cache Module 255
SSA RAID 5 254
supported nodes 27
Switch Adapter for the SP 163
Switch Chip for the SP 161
Switch Link for the SP 160
Switch Ports for the SP 161
Switch system for the SP 164
switched expansion frame 150
System Data Repository 262
System memory 30
system partitioning 265
System Partitioning Aid Perspective 340

T

tall frame 142
Tcl 266
technical computing 8
Thin Nodes 13
thin nodes 14
tight clusters 18
Tool command language 266

TPC-D 8
TURBOWAYS 100 ATM Adapter 245
Twisted Pair network 314
Two-stage switch system 147

U

utility loading for SP frames 143

V

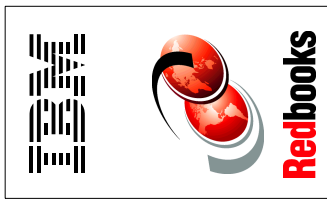
V.35 cable 242
V.36/EIA 449 cable 242
Vertical scaling 23
VMM (Virtual Memory Manager) 5
VPD (Vital Product Data) 6
VSD (Virtual Shared Disk) 265
VSD (Virtual Shared Disks) 5
VSD Perspective 340

W

Wide Nodes 13
wide nodes 13
WLM (AIX WorkLoad Manager) 5

X

X.21 cable 242
XL FORTRAN 273



RS/6000 SP and Clustered IBM @server pSeries Systems Handbook

(0.1”spine)
0.1”<->0.169”
53<->89 pages

RS/6000 SP and Clustered IBM @server pSeries Systems Handbook

Overview of the RS/6000 SP and pSeries servers configurations

Information on available machine types, models, and feature codes

In-depth discussion of hardware architectures

This redbook is an update of the RS/6000 SP Systems Handbook published in 2000. This publication includes Blue Hammer and Blue Hammer Jr. clustering solutions configurations.

This publication includes the following:

- Information on M80, H80, and 6H1 SP-attached support as well as clustering solutions capabilities with PSSP 3.2.
- Overview on clustered @server pSeries systems, configurations, and hardware support.
- Updated information on the available software for RS/6000 SP and clustered systems (PSSP, GPFS, LoadLeveler, PESSL, PE, and HACMP).
- Updated information on RS/6000 SP hardware support.
- Updated information on the supported control workstations for RS/6000 SP and clustered @server pSeries systems.
- Updated information on the supported communication adapters for the RS/6000 SP and clustered @server pSeries systems.
- Updated information on the SP Switch, SP Switch2 and the SP Switch Router.
- Updated information on available RS/6000 clustered hardware (machine types, models, and feature codes).

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-5596-02 ISBN 0738422991