

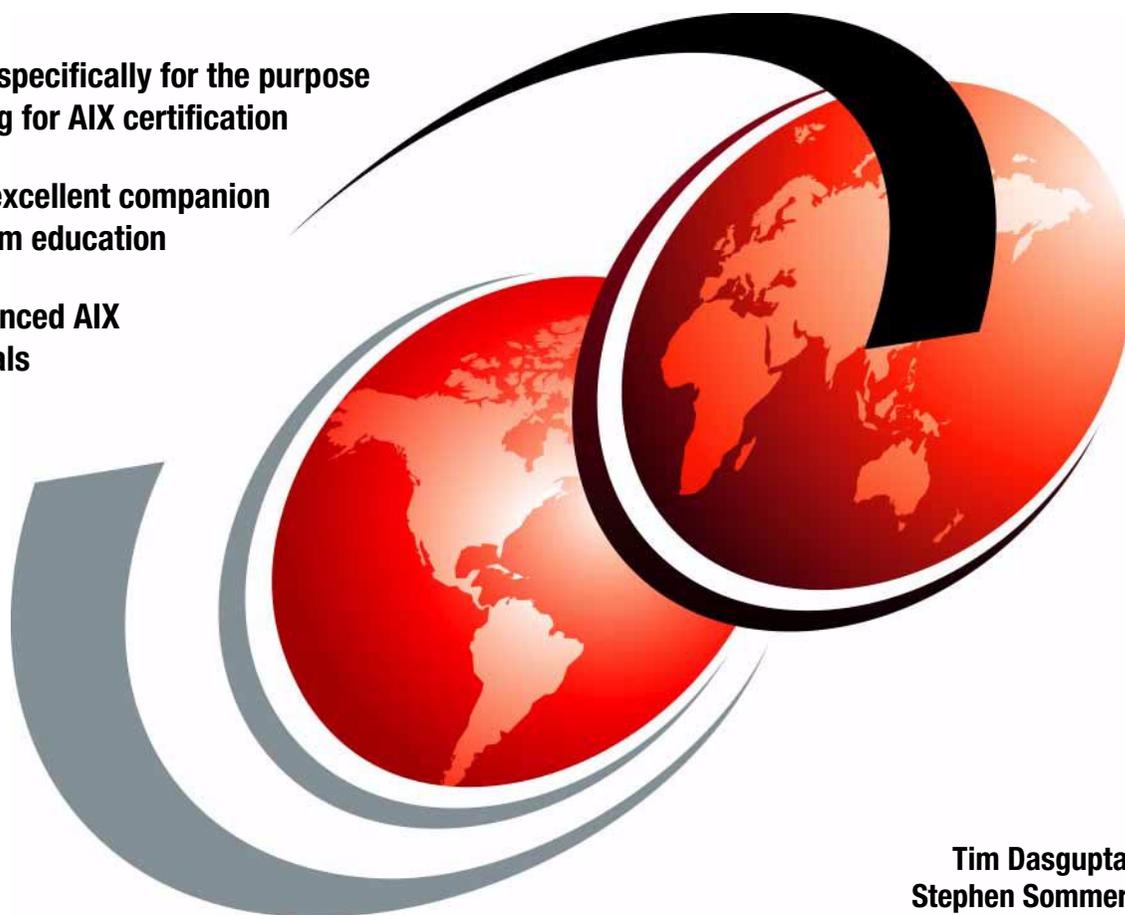


IBM **e**server Certification Study Guide - AIX 5L Installation and System Recovery

Developed specifically for the purpose
of preparing for AIX certification

Makes an excellent companion
to classroom education

For experienced AIX
professionals



Tim Dasgupta
Stephen Sommer



International Technical Support Organization

**IBM @server Certification Study Guide -
AIX 5L Installation and System Recovery**

December 2002

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

Second Edition (December 2002)

This edition applies to AIX 5L Version 5.1 (5765-E61) and subsequent releases running on an IBM @server pSeries or RS/6000 server.

© Copyright International Business Machines Corporation 2000, 2002. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Figures	ix
Tables	xi
Notices	xiii
Trademarks	xiv
Preface	xv
The team that wrote this redbook	xvi
Become a published author	xviii
Comments welcome	xviii
Chapter 1. Certification overview	1
1.1 Certification requirements	2
1.1.1 Required prerequisite	2
1.1.2 Recommended prerequisite	2
1.1.3 Information and registration for the certification exam	2
1.1.4 Core requirements	2
1.2 Certification education courses	5
Chapter 2. System installation and installation options	7
2.1 Installation methods	8
2.1.1 Installation startup	8
2.1.2 New and Complete Overwrite Installation	10
2.1.3 Migration Installation	10
2.1.4 Preservation Installation	13
2.2 Configuration Assistant Taskguide	14
2.3 Determining maintenance levels	15
2.4 Alternate Disk Installation	16
2.4.1 Cloning a running rootvg	18
2.5 Cloning your system	19
2.6 Installing additional hardware	20
2.6.1 Post-installation problems	22
2.7 Installing online documentation	23
2.8 Command summary	24
2.8.1 The blvset command	24
2.8.2 The cfgmgr command	24
2.8.3 The alt_disk_install command	25
2.9 Quiz	26

2.9.1 Answers	28
2.10 Exercises	28
Chapter 3. Additional software installation	29
3.1 What can be installed	30
3.1.1 Installation options	30
3.1.2 Installation methods	32
3.1.3 Installation reports	35
3.2 How to install using different scenarios	37
3.2.1 Bundles	37
3.2.2 Filesets	38
3.3 Uninstall software	46
3.4 Summary	47
3.5 Command summary	47
3.5.1 The installp command	48
3.5.2 The instfix command	49
3.5.3 The lspp command	50
3.6 Quiz	51
3.6.1 Answers	52
3.7 Exercises	52
Chapter 4. System boot	53
4.1 General overview of the boot process	54
4.2 BIST - POST	55
4.2.1 MCA systems	56
4.2.2 PCI systems	60
4.3 Boot phase 1	62
4.4 Boot phase 2	64
4.4.1 LED 551, 555, or 557	66
4.4.2 LED 552, 554, and 556	66
4.4.3 LED 518	68
4.4.4 The alog command	68
4.5 Boot phase 3	69
4.5.1 The /etc/inittab file	70
4.5.2 LED 553	72
4.5.3 LED c31	72
4.6 pSeries servers	73
4.6.1 Phase 1: Service processor initialization	73
4.6.2 Phase 2: Hardware initialization by the service processor	73
4.6.3 Phase 3: System firmware initialization	73
4.6.4 Phase 4: AIX boot	74
4.7 Summary	74
4.7.1 Boot phases	74

4.7.2 LED codes	74
4.8 Quiz	75
4.8.1 Answers	79
4.9 Exercises	79
Chapter 5. Disks and file systems	81
5.1 Volume groups	83
5.1.1 Creating a volume group	83
5.1.2 Migrating data to another disk in the same volume group	88
5.1.3 Importing a volume group from another system	89
5.1.4 Replacing a disk with another	92
5.1.5 Reducing a volume group	93
5.1.6 Reorganizing a volume group	94
5.2 Logical volumes	95
5.2.1 Creating a logical volume and setting the attributes	96
5.2.2 Different LVs associated with a JFS	102
5.2.3 Moving a logical volume from one physical volume to another	103
5.2.4 Mirroring logical volumes	107
5.3 Journalized file systems	110
5.3.1 Metadata	110
5.3.2 Adding a file system	110
5.3.3 Mounting and unmounting a file system	111
5.3.4 Reducing the size of a file system	114
5.3.5 Removing a file system	120
5.3.6 Creating a new JFS log	120
5.4 Managing a paging space	122
5.4.1 How to create a new paging space	123
5.4.2 How to expand an existing paging space	125
5.4.3 How to remove an existing paging space	126
5.4.4 How to shrink an existing paging space size	127
5.5 Working with physical volumes	129
5.5.1 Removing a physical disk drive from a system	129
5.5.2 Recovering an incorrectly removed disk	130
5.5.3 General tips on boot problems	132
5.6 Command summary	132
5.6.1 The <code>cfgmgr</code> command	132
5.6.2 The <code>bosboot</code> command	133
5.6.3 The <code>bootlist</code> command	134
5.6.4 The <code>chps</code> command	135
5.6.5 The <code>cplv</code> command	135
5.6.6 The <code>fsck</code> command	136
5.6.7 The <code>fuser</code> command	139
5.6.8 The <code>migratepv</code> command	140

5.6.9	The mkboot command	140
5.6.10	The mklv command	140
5.6.11	The mkps command	144
5.6.12	The replacepv command	145
5.6.13	The restvg command	145
5.6.14	The sysdumpdev command	147
5.7	Quiz	149
5.7.1	Answers	152
5.8	Exercises	152
Chapter 6.	System backup and recovery	153
6.1	Backup	154
6.1.1	Backup of the rootvg volume group	154
6.1.2	The data layout of a mksysb tape	154
6.1.3	Back up non-rootvg volume group	156
6.1.4	Back up file systems	156
6.2	Recovery	157
6.2.1	Recovery of rootvg	157
6.2.2	Restore a non-rootvg volume group	158
6.2.3	Restore a file system	158
6.2.4	Extracting data from mksysb tape	158
6.2.5	A customized mksysb restore	159
6.3	Other UNIX backup commands	161
6.4	Sysback	162
6.4.1	Sysback installation	163
6.4.2	Virtual devices	163
6.4.3	Other features	163
6.4.4	Optional features	164
6.5	Tivoli Storage Manager	164
6.6	Troubleshooting	165
6.6.1	Remote backup	165
6.6.2	Sparse files	166
6.6.3	Block size determination	166
6.6.4	Tape block size determination scenario	167
6.7	Common backup and restore errors	167
6.8	Command summary	168
6.8.1	The mksysb command	168
6.8.2	The savevg command	169
6.8.3	The restvg command	169
6.8.4	The backup command	170
6.8.5	The restore command	170
6.8.6	The tar command	171
6.9	Quiz	172

6.9.1 Answers	174
6.10 Exercises	174
Chapter 7. Network installation management	175
7.1 Debugging NIM boot images	177
7.1.1 Establishing network communication	177
7.1.2 Obtaining the boot image from the server	177
7.1.3 Running the boot image on the client	179
7.2 Error 611: Network error	180
7.3 The nim command	181
7.4 The showmount command	187
7.5 Producing debug output for NIM BOS installations	189
7.5.1 Producing debug output from a network boot image	190
7.5.2 Producing debug output from the BOS install program	191
7.6 Common NIM errors	192
7.7 Quiz	193
7.7.1 Answers	194
7.8 Exercises	194
Abbreviations and acronyms	195
Related publications	205
IBM Redbooks	205
Other resources	206
Referenced Web sites	206
How to get IBM Redbooks	207
IBM Redbooks collections	207
Index	209

Figures

2-1	Installations and Settings screen	9
2-2	Configuration Assistant Taskguide	15
3-1	List Software on Media and Related Information menu	30
3-2	SMIT installation menus.	32
3-3	Installing user-defined bundle	38
3-4	Preview option in SMIT	40
3-5	Saving previous version when updating fileset	42
3-6	Software maintenance and utilities	43
3-7	SMIT menu for rejecting an applied fileset.	43
3-8	Install newer version of software	44
3-9	Removing software through SMIT	46
3-10	Removing installed software and dependent software through SMIT	47
4-1	General boot order.	54
4-2	Function selection menu in diag	57
4-3	Task selection menu in diag	58
4-4	Display/alter bootlist menu in diag	58
4-5	SMS main menu	61
4-6	Boot phase 1	63
4-7	Boot phase 2, part one.	64
4-8	Boot phase 2, part two.	65
4-9	Boot phase 3	69
4-10	Example of rc.boot 3 in /etc/inittab.	71
5-1	The Volume Groups screen	85
5-2	The Add a Volume Group selection screen	85
5-3	The Add a Volume Group warning screen.	86
5-4	The Add a Volume Group Command Status screen	87
5-5	The SMIT Logical Volumes screen	97
5-6	The Add a Logical Volume screen	98
5-7	The Add a Logical Volume menu 1 of 2.	98
5-8	Five regions of a disk.	100
5-9	The Add a Logical Volume menu 2 of 2.	101
5-10	The Command Status screen for the logical volume	101
5-11	The SMIT crfs screen.	111
5-12	The SMIT restvg screen.	119
5-13	The SMIT Add Another Paging Space dialog	124
5-14	The SMIT Change/Show Characteristics of a Paging Space dialog	125
6-1	Contents of mksysb image.	155

Tables

2-1	Commonly used flags of the blvset command	24
2-2	Commonly used flags of the cfgmgr command	24
2-3	Commonly used flags of the alt_disk_install command	25
3-1	Flags used at installation with the installp command	33
3-2	Flags for installp used by SMIT	39
3-3	Commit or apply?	41
3-4	Choice of options when reinstalling software	44
3-5	Commonly used flags of the installp command	49
3-6	Commonly used flags of the instfix command	49
3-7	Commonly used flags of the lspp command	50
4-1	Common MCA LED codes	59
4-2	The telinit command examples	72
4-3	MCA POST LED	74
4-4	Boot phase 2 LED codes	75
4-5	Boot phase 3 LED codes	75
5-1	Commonly used flags of the cfgmgr command	133
5-2	Commonly used flags of the bosboot command	134
5-3	Commonly used flags of the bootlist command	134
5-4	Commonly used flags of the chps command	135
5-5	Commonly used flags of the cplv command	136
5-6	Commonly used flags of the fsck command	137
5-7	Commonly used flags of the fuser command	139
5-8	Commonly used flags of the migratepv command	140
5-9	Commonly used flags of the mkboot command	140
5-10	Commonly used flags of the mklv command	141
5-11	Commonly used flags of the mkps command	144
5-12	Commonly used flags of the replacepv command	145
5-13	Commonly used flags of the restvg command	146
5-14	Commonly used flags of the sysdumpdev command	147
6-1	Common backup and restore errors	168
6-2	Commonly used flags of the mksysb command	169
6-3	Commonly used flags of the savevg command	169
6-4	Commonly used flags of the restvg command	170
6-5	Commonly used flags of the backup command	170
6-6	Commonly used flags of the restore command	171
6-7	Commonly used flags of the tar command	171
7-1	Types of machines managed by NIM	176
7-2	Commonly used flags of the nim command	181

7-3	Commonly used flags of the showmount command	187
7-4	Common NIM codes	192

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

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

AFP™	Micro Channel®	Redbooks(logo)™ 
AFS®	Perform™	RS/6000®
AIX®	PowerPC®	Sequent®
AIX 5L™	PowerPC Reference Platform®	SP™
DFS™	pSeries™	Tivoli®
@server	PTX®	Versatile Storage Server™
IBM®	QMF™	
LoadLeveler®	Redbooks™	

The following terms are trademarks of International Business Machines Corporation and Lotus Development Corporation in the United States, other countries, or both:

Lotus®	Domino™
--------	---------

The following terms are trademarks of other companies:

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

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.

Preface

The AIX and IBM @server pSeries Certifications offered through the Professional Certification Program from IBM are designed to validate the skills required of technical professionals who work in the powerful and often complex environments of AIX and IBM @server pSeries. A complete set of professional certifications is available. It includes:

- ▶ IBM Certified AIX User
- ▶ IBM Certified Specialist - Business Intelligence for RS/6000
- ▶ IBM Certified Specialist - Domino for RS/6000
- ▶ IBM @server Certified Specialist - p690 Solutions Sales
- ▶ IBM @server Certified Specialist - p690 Technical Support
- ▶ IBM @server Certified Specialist - pSeries Sales
- ▶ IBM @server Certified Specialist - pSeries AIX System Administration
- ▶ IBM @server Certified Specialist - pSeries AIX System Support
- ▶ IBM @server Certified Specialist - pSeries Solution Sales
- ▶ IBM Certified Specialist - RS/6000 SP and PSSP Version 3
- ▶ IBM Certified Specialist - Web Server for RS/6000
- ▶ IBM @server Certified Specialist - pSeries HACMP for AIX
- ▶ IBM @server Certified Advanced Technical Expert - pSeries and AIX 5L

Each certification is developed by following a thorough and rigorous process to ensure the exam is applicable to the job role and is a meaningful and appropriate assessment of skill. Subject matter experts who successfully perform the job participate throughout the entire development process. They bring a wealth of experience into the development process, making the exams much more meaningful than the typical test that only captures classroom knowledge and ensuring the exams are relevant to the *real world*. Thanks to their effort, the test content is both useful and valid. The result of this certification is the value of appropriate measurements of the skills required to perform the job role.

This IBM Redbook is designed as a study guide for professionals wishing to prepare for the AIX 5L Installation and System Recovery certification exam as a selected course of study in order to achieve the IBM @server Certified Advanced Technical Expert - pSeries and AIX 5L certification.

This publication is designed to provide a combination of theory and practical experience needed for a general understanding of the subject matter. It also provides sample questions that will help in the evaluation of personal progress

and provide familiarity with the types of questions that will be encountered on the exam.

This publication does not replace practical experience, nor is it designed to be a stand-alone guide for any subject. Instead, it is an effective tool that, when combined with education activities and experience, can be a very useful preparation guide for the exam.

For additional information about certification and instructions on how to register for an exam, visit our Web site at:

<http://www.ibm.com/certify>

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.

Tim Dasgupta is an IBM Certified AIX Advanced Technical Expert (CATE). He works as a Senior Systems Architect at IBM Global Services in Canada. He has over eight years of experience in the areas of AIX, RS/6000, and pSeries. He is currently the Team Leader of Midrange Architecture Group in Montreal, Canada.

Stephen Sommer is an IBM Certified AIX Advanced Technical Expert (CATE), AIX Version 4.3.3 and 5.1. He works as a Senior IT Specialist at Faritec Services, an IBM Business Partner in Johannesburg, South Africa. He has eight years of experience in Midrange Support for AIX, RS/6000, and pSeries, both in South Africa and the United Kingdom.

The authors of the first edition are:

IBM Denmark

Thomas Herlin

IBM South Africa

André de Klerk

IBM Sweden

Thomas C. Cederlöf

Prokom Software SA in Poland

Tomasz Ostaszewski

The project that produced this publication was managed by:

International Technical Support Organization, Austin Center

Scott Vetter

Special thanks to:

IBM Atlanta

Shannan L DeBrule

Program Manager, AIX Certification

Darin Hartman

Thanks to the following people for their invaluable contributions to this project:

IBM Austin

Jesse Alcantar, Greg Althaus, Erwin Behnen, Karl Borman, Larry Brenner, Greg Flaig, Shawn Mullen, Brian Nicholls, Lindsey Thomson

IBM Atlanta

Michelle Page-Rivera

IBM Australia

Yun-Wai Kong

IBM Boston

Edward Geraghty

IBM Japan

Keigo Matsubara

IBM Netherlands

Ronald Van der Knaap

IBM Pakistan

Adnan Ikram

IBM Raleigh

Christopher Snell

IBM Toronto

Darrin Woodard

IBM U.K.

Peter Mayes

ILS Sweden

Malin Cederberg and Robert Olsson

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our 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:

IBM Corporation, International Technical Support Organization
Dept. JN9B Building 003 Internal Zip 2834
11400 Burnet Road
Austin, Texas 78758-3493



Certification overview

This chapter provides an overview of the skill requirements needed to obtain an IBM Advanced Technical Expert certification. The following chapters are designed to provide a comprehensive review of specific topics that are essential for obtaining the certification IBM @server Certified Advanced Technical Expert - pSeries and AIX 5L.

This level certifies an advanced level of pSeries and AIX knowledge and understanding, both in breadth and depth. It verifies the ability to perform in-depth analysis, apply complex AIX concepts, and provide resolutions to critical problems, all in a variety of areas within AIX, including the hardware the supports it.

1.1 Certification requirements

To attain the IBM @server Certified Advanced Technical Expert - pSeries and AIX 5L certification, you must pass four tests.

One test is the prerequisite in either pSeries AIX System Administration or pSeries AIX System Support. The other three tests are selected from a variety of pSeries and AIX topics. These requirements are explained in greater detail in the sections that follow.

1.1.1 Required prerequisite

Prior to attaining the IBM @server Certified Advanced Technical Expert - pSeries and AIX 5L certification, you must be certified as either an:

- ▶ IBM @server Certified Specialist - pSeries AIX System Administration
- or
- ▶ IBM @server Certified Specialist - pSeries AIX System Support

1.1.2 Recommended prerequisite

A minimum of six to twelve months of experience in performing in-depth analysis and applying complex AIX concepts in a variety of areas within AIX is a recommended prerequisite.

1.1.3 Information and registration for the certification exam

For the latest certification information, see the following Web site:

<http://www.ibm.com/certify>

1.1.4 Core requirements

Select three of the following exams. You will receive a Certificate of Proficiency for tests when passed.

AIX 5L Installation and System Recovery

The following objectives were used as a basis when the certification test 233 was developed. Some of these topics have been regrouped to provide better organization when discussed in this publication.

Preparation for this exam is the topic of this publication.

Section I - Installation

To install:

1. Develop installation strategy.
 - a. Determine whether migration or new install is necessary (for example, upgrade, overwrite, maintenance level).
 - b. Determine whether network or local install is required.
2. Perform local install.
 - a. Install base operating system (boot off CD).
 - b. Install additional licensed products.
 - c. Install necessary updates to operating system and licensed products.
3. Verify installation.
 - a. Reboot system to verify proper system startup and verify that subsystems have started properly.
 - b. Verify installation of desired licensed products.
 - c. Verify that applicable updates have been installed (**lppchk**, **lslpp**, **instfix**).
 - d. Verify that applications and/or databases are running properly.
4. Customize the system environment.
 - a. Set up system environment variables (date, time, and others).
 - b. Set up networking.
 - c. Set up NIM if necessary.
 - d. Set up devices and printers.
 - e. Prepare files system and logical volumes.
 - f. Set up additional resources.
 - g. Install applications/databases (prepare data with mirroring, additional volume group).

Section II - Installation problem determination

To determine installation problems:

1. Diagnose and identify boot problem.
 - a. Determine best path based on symptoms of failure.
 - b. Record boot problem symptoms (for example, LED hang, system crash, hang with text on screen).
 - c. Consult problem determination documentation.

2. Follow procedure for system recovery.
 - a. Boot system in maintenance mode.
 - b. Perform file system recovery.
 - c. Perform boot image/bootlist recovery with **bosboot** and **bootlist**.
 - d. Fix hangs due to /etc/inittab problems.
3. Verify problem is resolved.

Reboot system to verify proper operation.

Section III - Backup

To back up:

1. Determine strategy for backup (type, media, time).
 - a. Determine rootvg (OS) (for example, **mksysb**).
 - b. Determine data and applications strategy (what AIX tools, for example).
2. Perform backup (tools).

Determine backup utility required (command line tools).
3. Verify backup success.
 - a. Boot off of **mksysb** backup to verify it is bootable.
 - b. Troubleshoot using AIX backup methods.
 - c. List table of contents from **mksysb**.

Section IV - Recovery

To recover:

1. Determine strategy for recovery.

Determine if **mksysb** restore is necessary.
2. Perform recovery.
 - a. Recover LVM.
 - b. Restore data.
3. Verify Recovery Success.

Verify proper functionality of system applications after recovery.
4. Software maintenance.
 - a. Perform de-installation if necessary.
 - b. Perform operating system updates.

AIX 5L Performance and System Tuning

Test 234 was developed for this certification.

Preparation for this exam is the topic of *IBM @server Certification Study Guide - AIX 5L Performance and System Tuning*, SG24-6184.

AIX 5L Problem Determination Tools and Techniques

Test 235 was developed for this certification.

Preparation for this exam is the topic of *IBM @server Certification Study Guide - AIX 5L Problem Determination Tools and Techniques*, SG24-6185.

AIX 5L Communications

Test 236 was developed for this certification.

Preparation for this exam is the topic of *IBM @server Certification Study Guide - AIX 5L Communications*, SG24-6186.

pSeries HACMP for AIX

Test 187 was developed was developed for this certification.

Preparation for this exam is the topic of *IBM @server Certification Study Guide - pSeries HACMP for AIX*, SG24-6187.

RS/6000 SP and PSSP V3.1

Test 188 was developed for this certification.

Preparation for this exam is the topic of *IBM Certification Study Guide - RS/6000 SP*, SG24-5348.

p690 Technical Support

Test 195 was developed for this certification.

An IBM Redbook is planned for first quarter 2003 on this subject.

1.2 Certification education courses

Courses are offered to help you prepare for the certification tests. For a current list, visit the following Web site, locate your test number, and select the education resources available:

<http://www.ibm.com/certify/tests/info.shtml>



System installation and installation options

This chapter covers the following topics:

- ▶ An overview of the most common AIX installation methods
- ▶ Installation of additional hardware
- ▶ Highlights of the typical post-installation problems
- ▶ Installation of online documentation

System installation is an important part of system administration of RS/6000 or pSeries systems, as well as older systems that need to be upgraded.

2.1 Installation methods

AIX provides three types of installation methods:

- ▶ **New and Complete Overwrite Install**
Use this installation method on a new machine or if you want to completely overwrite an existing version of the base operating system (BOS) that is installed on your system.
- ▶ **Migration Install**
Use this installation method to upgrade AIX between versions and releases while preserving the existing root volume group.
- ▶ **Preservation Install**
Use this installation method when a version of BOS is installed on your system and you want to preserve the user data in the root volume group.

The following sections describe the installation methods in more detail.

2.1.1 Installation startup

In this section we discuss installation startup.

Note: Before you start installing AIX, make sure that your system has enough disk space. If not enough disk space is available, then the installation will fail.

To start the installation of AIX, your system needs to first boot from the CD-ROM. In order to force a boot from the media, you have to perform the following sequence. If you have an older RS/6000 system (non-PCI) with a system key present, turn the key to the Service position:

1. Power the system on.
2. Immediately insert the installation CD-ROM into the CD-ROM device.
3. When the keyboard indicator is shown on the screen (the word keyboard on an ASCII terminal or the keyboard icon on a graphical display), press F5 on the directly attached keyboard or 5 on the ASCII terminal. This will temporarily alter the bootlist. You can ignore this step if you have an older system (non-PCI) with a system key.
4. At this stage, the system boots up using the media. If the system does not boot up but enters SMS (Figure 4-5 on page 61), there is a problem with the media. Check if it is correctly inserted and undamaged. Reinsert media or replace media and restart at step 1.

5. Select the system console by pressing F1 and Enter on the attached keyboard or 1 and Enter on the ASCII terminal.
6. Select the installation language. The default is English.
7. On the Welcome to Base Operating System Installation and Maintenance screen press 2 to select **Change/Show Installation Settings and Install** to change the installation method. A screen similar to the one shown in Figure 2-1 will present the different installation options.

```

                                Installation and Settings

Either type 0 and press Enter to install with current settings, or type the
number of the setting you want to change and press Enter.

  1 System Settings:
    Method of Installation.....New and Complete Overwrite
    Disk Where You Want to Install....hdisk0

  2 Primary Language Environment Settings (AFTER Install):
    Cultural Convention.....English (United States)
    Language .....English (United States)
    Keyboard .....English (United States)
    Keyboard Type.....Default

  3 Install Trusted Computing Base..... No

>>> 0 Install AIX with the current settings listed above.

-----+-----
  88 Help ?      | WARNING: Base Operating System Installation will
  99 Previous Menu | destroy or impair recovery of ALL data on the
                   | destination disk hdisk0.
>>> Choice [0]:

```

Figure 2-1 Installations and Settings screen

8. At this point, you can select the method of your installation, the disk for rootvg, the set up language environments, and the installation of Trusted Computing Base (TCB), or other options on different software releases. Depending on the current system, the default installation method varies. In the following sections, the different installation methods are individually described.

The primary language environment can be set up for cultural convention, language, and keyboard by choosing a language that fits the actual requirements. Different sets of languages are supported (see AIX version-specific installation guides for additional information).

TCB provides an additional level of security by installing the trusted path, the trusted shell, and system integrity checking. Note that you cannot enable the TCB at a later time; it must be at the installation time.
9. Select 0 once you are satisfied with the settings, and the installation will start.

10. On older non-PCI models with a system key, you must, at this point, switch the key back to the Normal position.

2.1.2 New and Complete Overwrite Installation

The New and Complete Overwrite Installation method installs the BOS by overwriting all data on the selected destination disk.

In the following scenarios, a New and Complete Overwrite Installation should be used:

- ▶ A new machine without AIX installed.
- ▶ An existing system must be reinstalled with the BOS because:
 - rootvg has become corrupted, and a backup does not exist.
 - rootvg is spread over several hard disks, and you want to reassign it to fewer hard disks to make your rootvg smaller.
- ▶ The pre-installed system is not TCB enabled. If you need the features provided by TCB, you must re-install the system using this method.

After the installation is complete, the system will reboot and start the Installation Assistant for post-installation tasks. See 2.2, “Configuration Assistant Taskguide” on page 14.

2.1.3 Migration Installation

The Migration Installation method is used to upgrade AIX to a different version or release while preserving the existing root volume group. Migration is the default installation method to move from AIX 3.2 or any version of AIX Version 4 or the operating system to AIX 5L Version 5.1.

Migration Installation preserves all logical volumes and file systems on rootvg except /tmp. Other volume groups are not affected.

During a Migration Installation, the installation process determines which optional software products must be installed.

Previously installed software products remain on the system.

When migrating from AIX Version 3.2, all files in /usr/lib/drivers, /usr/lib/microcode, /usr/lib/methods, and /dev are removed from the system, so software support for non-IBM device drivers must be reinstalled. Furthermore, a set of software products is removed.

Note: Before performing any Migration Installation, make sure that you have a complete backup of your system, as well as your data.

Steps when performing a Migration Installation

At the beginning of a Migration Installation, the system verifies that there will be enough space to attempt the migration. If there is not, a message is displayed explaining how much extra space is needed.

At this point, you must reboot the server from the disk and make some space available in the rootvg volume group to do the migration. This can be done by either adding another disk to the rootvg or moving user data from rootvg to another volume group.

If the additional space could not be found to match the requirement, you might consider performing a Preservation Installation (see 2.1.4, “Preservation Installation” on page 13).

When enough free space is found, reboot from your installation media and try the Migration Installation again. To complete the Migration Installation, you must have at least 8 MB of free disk space.

If there is insufficient space to complete the Migration Installation during the BOS installation process, a message similar to the following is displayed at the end of the installation:

```
An error occurred while migrating packages.  
Some packages have not been installed.  
Please see /var/adm/ras/devinst.log for details or perform an overwrite or  
preservation install.
```

When you are migrating to AIX Version 4.3 or higher, a confirmation screen will appear and enable you to do the following:

- ▶ List the software on your system that is incompatible with the installed AIX version and that will be installed at the new level.
- ▶ List the saved base system configuration files that will not be merged into the system. These files are saved in /tmp/bos.
- ▶ List the filesets that will be removed and not replaced.
- ▶ Reboot the system without completing the migration. This selection cancels the installation.
- ▶ Continue with the installation.

If you, at this point, press 0 and Enter, the Migration Installation will begin.

If you are migrating AIX Version 3.2 programs, a BOS Migration Installation only migrates those optional software products that can run on later AIX versions. If you are migrating AIX Version 4.1 or AIX Version 4.2 programs and there is a newer level of a previously installed fileset on the installation media, a BOS Migration Installation will try to install the newer level. Any previously installed products that can run on newer BOS will not be altered during the Migration Installation.

After the Migration Installation is finished, check the migration by listing the filesets in your system and compare them with the previous ones. As an example, look for committed filesets of the previous version. If you were migrating from AIX Version 4.3 to AIX Version 5.1 use:

```
# ls1pp -l | grep 4.3
```

Any AIX filesets that are listed in the committed state should be removed. An exception to this might be third-party software products.

You might also check previous configuration files saved by the Migration Installation with the new ones created and resolve the possible differences.

Missing Migration Installation option

If the Migration Installation option is not selectable as you would expect when migrating from one AIX version to another (for example, 4.2 to 4.3), there is an incorrect pad string in the boot logical volume (BLV). This wrong system information can happen, for example, if a migration attempt was canceled or did not succeed.

A way to overcome this is to use the **blvset** command, which resets the pad string. Use the **-g** flag to get the pad string and the **-p** flag to write the pad string to the volume group.

Use **blvset** in the following way:

```
# /usr/lpp/bosinst/blvset -d /dev/hdisk0 -p level
```

The **blvset** command is expecting a version string from the standard input. Enter the version number (for example, 4.3), and press Enter.

This example illustrates a pad string reset from a broken 4.2 to 4.3 migration, where the rootvg is located on hdisk0.

After you have updated the pad strings, you can retry the Migration Installation.

When Migration Installation fails

There are times when your Migration Installation fails and you need to recover the system. You should have a system backup available and if you do, restore

your system from it. If your system backup does not exist, or is corrupted, the best way to recover is to perform a Preservation Installation. This will disrupt some information, such as network configuration, but it is much better than overwriting your entire system including user data.

When migrating to a later version of AIX, the server must be rebooted to ensure the updated filesets are being used. The reboot should occur as part of the migration process but may be interrupted for various reasons. The following example shows the output of the **lppchk** command on a server that has not been rebooted after migration:

```
# lppchk -v
bos.rte V => 4R is>4
bos.rte V>5
bos.rte not installed...5.1.0
```

Although when using the **lslpp** command on the bos.rte fileset, it shows that it has been installed at the required level, as in the following example:

```
# lslpp -l |grep bos.rte
bos.rte 5.1.0.25 committed Base Operating System Runtime
```

To correct this scenario the server should be rebooted and the fileset level checked with the **lslpp** command. This will complete the migration process.

2.1.4 Preservation Installation

The Preservation Installation method is chosen when a version of BOS is installed on a system where you want to preserve the user data in the root volume group. It is the default installation setting for a machine already installed with AIX Version 4.3 or higher.

Note: Preservation Installation overwrites the /usr, /tmp, /var, and /(root) file systems by default, so any user data in these directories is lost. These file systems are removed and recreated, so any other LPPs or filesets that you installed on the system will also be lost.

A Preservation Installation will only save all the *non-system* logical volumes and file systems on rootvg (for example, /home), or any user-created logical volumes.

In order to preserve data, the Preservation Installation uses the /etc/preserve.list file to list the system files to be copied and saved during the installation. The /etc/filesystems file is listed by default. Add the full path names of any additional files that you want to save during the Preservation Installation to the preserve.list file. This file does not exist on systems prior to AIX Version 4.1, so in this case you must explicitly create it.

You can modify the `preserve.list` file and copy it to a diskette to be used during BOS installation. Information on how to create this diskette is found in the AIX installation guides. If a `preserve.list` file exists on diskette, the installation program uses this information instead of the default `/etc/preserve.list` file.

Ensure that you have sufficient disk space in the `/tmp` file system to store all the files listed in the `/etc/preserve.list` file.

A system configuration must be redone after performing a Preservation Installation. Additionally, you must reinstall any applications that were installed in your rootvg after the Preservation Install has completed.

After a Preservation Installation, you must also reconfigure devices, as well as recreate users and groups.

2.2 Configuration Assistant Taskguide

After the installation is completed, the system reboots and the Configuration Assistant Taskguide program, `/usr/sbin/install_assist`, is automatically started. When using an ASCII terminal for install, the Configuration Assistant Taskguide is referred to as the Installation Assistant.

The Configuration Assistant Taskguide program allows you to set up basic configuration, such as the following:

- ▶ Setting date, time, and time zone
- ▶ Setting the root password
- ▶ Adjusting system storage and paging space
- ▶ Configuring TCP/IP
- ▶ Updating installed software after Migration Installation
- ▶ Configuring Web-based System Management (websm)
- ▶ Configuring online documentation and search

On a Version 4.3.1 system with a graphical display attached, the Installation Assistant will be similar to the screen shown in Figure 2-2 on page 15.

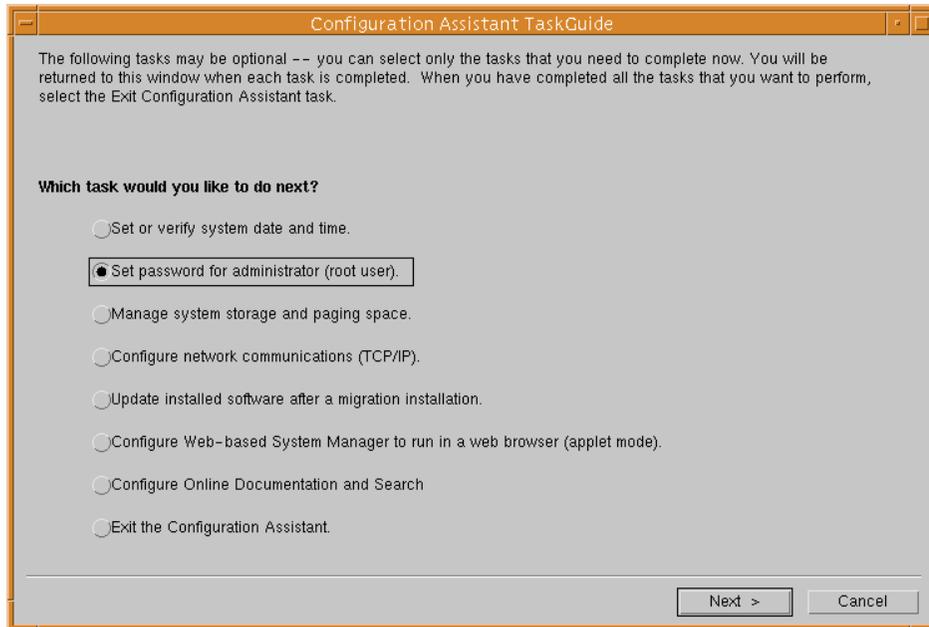


Figure 2-2 Configuration Assistant TaskGuide

On a system installed from an ASCII terminal, a similar task list will appear in alphanumeric format.

After you completed these tasks using the Configuration Assistance Taskguide, then you should disable the automatic restart of the Configuration Assistance Taskguide. Otherwise it might appear on the system console invoked by the init process on every reboot. To disable the automatic restart, you can select from the following menu item.

- ▶ Exit the Configuration Assistant.
- ▶ Finish now, and do not start Configuration Assistant when restarting AIX.

2.3 Determining maintenance levels

The `oslevel` command reports the level of the operating system using a subset of all filesets installed on your system. The `oslevel` command also prints information about maintenance levels, including which filesets are not at a specified maintenance level.

If your system, after an update, is not showing the new maintenance level, use the `-l` flag to determine what filesets need to be updated to match the new BOS level.

Note: Never use the newly installed AIX system for production use without proper update filesets installed from the AIX Update CD. AIX media shipments are accompanied by AIX Update CD-ROMS, which provide necessary fileset updates for the target hardware system support. To install or update the AIX filesets, refer to Chapter 3, “Additional software installation” on page 41.

An AIX operating system software component is divided into four levels. The first three of these levels are referred as VRM (each correspond to version, revision, and modification level number). For example, AIX 5L Version 5.1.1 means, Version 5 Release 1 modification level 1. These component levels are examined by the `oslevel` command.

The fourth component level is the *preventive maintenance level* delivered as a *pack* of the filesets update (PTF) from the IBM software support. The AIX preventive maintenance levels are usually distributed on the Internet (using the `fixdist` client on AIX), or by the AIX Update CD-ROM media.

To confirm the AIX preventive maintenance level on your system, you can use the `instfix` command as in the following example.

```
# instfix -i | grep ML
All filesets for 5.0.0.0_AIX_ML were found.
All filesets for 5.1.0.0_AIX_ML were found.
All filesets for 5.1.0.0_AIX_ML were found.
All filesets for 5100-01_AIX_ML were found.
All filesets for 5100-02_AIX_ML were found.
```

In this case, the last AIX preventive maintenance level 5100-02_AIX_ML is fully installed on this system. To check for additional maintenance level requirements, you can use the following method.

```
# instfix -ivk 5100-02_AIX_ML | grep not | grep ':'
```

Additional filesets should be updated from the Update CD or AIX BOS, Bonus Pack CDs. If you cannot find the necessary filesets in there, you should ask it for IBM software service or download it from the `fixdist` service.

2.4 Alternate Disk Installation

On systems where the downtime is critical, AIX Version 4.3 or later offers an additional way to install a system. The Alternate Disk Installation allows installing

a system while it is still up and running, thus decreasing the installation or upgrade down time considerably.

Alternate Disk Installation is used for:

- ▶ Installation of a **mksysb** image on another disk. For example, a **mksysb** from a machine with similar hardware that already has been upgraded.
- ▶ Cloning an existing rootvg on another disk. Optionally, the cloning offers the possibility of installing updates and new filesets to the cloned rootvg.

Using the Alternate Disk Installation method allows larger installations to prepare all the machines that need to be upgraded and perform the switch to the new version at the same time.

The command used for Alternate Disk Installation is **alt_disk_install**.

The command creates an **altinst_rootvg** volume group on the destination disk and prepares the same logical volume groups as in the rootvg, except the names are prepended with **alt_** (for example, **alt_hd1**). Similar are the file systems renamed to **/alt_inst/filesystemname**, and the original data (**mksysb** or rootvg) is copied.

After this first phase, a second phase begins where an optional configuration action can be performed, either a custom script or update of software, when cloning rootvg.

The third phase unmounts the **/alt_inst/** file systems and renames the file systems and logical volumes by removing the **alt** names. When this is done, the **altinst_rootvg** is varied off, and the bootlist is altered to boot from the new disk.

After the system is rebooted, the original rootvg is renamed to **old_rootvg**.

Note: The Alternate Disk Installation requires the filesets **bos.alt_disk_install.boot_images** and **bos.alt_disk_install.rte** to be installed.

Instead of using the **alt_disk_install** command directly, the SMIT menu's **smitty alt_mksysb** and **smitty alt_clone** are provided for **mksysb** installation and cloning, respectively.

2.4.1 Cloning a running rootvg

The following example shows the use of the `alt_disk_install` command performing a cloning of a running rootvg on hdisk0 to an unused hdisk1:

```
# lspv
hdisk0          000fa17debb9c93a    rootvg
hdisk1          000fa17dc9bf376f    None

# alt_disk_install -C hdisk1
Calling mkoszfile to create new /image.data file.
Checking disk sizes.
Creating cloned rootvg volume group and associated logical volumes.
Creating logical volume alt_hd5
Creating logical volume alt_hd6
Creating logical volume alt_hd8
Creating logical volume alt_hd4
Creating logical volume alt_hd2
Creating logical volume alt_hd9var
Creating logical volume alt_hd3
Creating logical volume alt_hd1
Creating /alt_inst/ file system.
Creating /alt_inst/home file system.
Creating /alt_inst/tmp file system.
Creating /alt_inst/usr file system.
Creating /alt_inst/var file system.
Generating a list of files
for backup and restore into the alternate file system...
Backing-up the rootvg files and restoring them to the
alternate file system...
Modifying ODM on cloned disk.
Building boot image on cloned disk.
forced unmount of /alt_inst/var
forced unmount of /alt_inst/usr
forced unmount of /alt_inst/tmp
forced unmount of /alt_inst/home
forced unmount of /alt_inst
forced unmount of /alt_inst
Changing logical volume names in volume group descriptor area.
Fixing LV control blocks...
Fixing file system superblocks...
Bootlist is set to the boot disk: hdisk1

# lspv
hdisk0          000fa17debb9c93a    rootvg
hdisk1          000fa17dc9bf376f    altinst_rootvg
#
```

Notice the difference in the `lspv` command before and after performing the `alt_disk_install`. The cloning example shown previously does not perform any update or customizing.

In addition to the creation of the Alternate Disk Installation, the `alt_disk_install` program can also be used for operations on the `altinst_rootvg`.

To access the `altinst_rootvg`, use the `-W` flag of `alt_disk_install` to *wake up* the alternate volume group. This is useful if you have to synchronize data on both `rootvg` and `altinst_rootvg`. Respectively, the `altinst_rootvg` can be put to sleep using the `-S` flag of `alt_disk_install`.

For further information on Alternate Disk Installation, refer to the man pages for `alt_disk_install`.

2.5 Cloning your system

A `mksysb` image enables you to clone one system image onto multiple target systems. The target systems might not contain the same hardware devices or adapters, require the same kernel (uniprocessor or microprocessor), or be the same hardware platform (`rs6k`, `rspc`, or `chrp`) as the source system. Please note that although `mksysb` cloning works, it is not a supported method of installation.

Use this procedure to install a `mksysb` backup on a target system it was not created on. Be sure to boot from the product media appropriate for your system and at the same maintenance level of BOS as the installed source system that the `mksysb` backup was made on. For example, you can use BOS Version 4.2.1 product media with a `mksysb` backup from a BOS Version 4.2.1 system. This procedure is to be used when installing a backup tape to a different system.

After booting from product media, complete the following steps when the Welcome to the Base Operating System Installation and Maintenance screen is displayed.

1. Select the **Start Maintenance Mode for System Recovery** option.
2. Select the **Install from a System Backup** option.
3. Select the drive containing the backup tape and insert the media for that device. The system reads the media and begins the installation.
4. You will be prompted again for the BOS install language, and the Welcome screen is displayed. Continue with the Prompted Installation process, as cloning is not supported for Nonprompted Installations.

Note: Cloning from a **mksysb**:

- ▶ Booting from tape product media is not supported on some rspc platform systems. When a backup tape is created on one of these systems, the **mksysb** command will display a message indicating that the system does not support tape boot. To determine what your platform system is, enter the following command:

```
bootinfo -p
```

or

```
lscfg -vp | grep Arch
```

- ▶ If you are cloning from the product tape to restore a backup tape, create a diskette that contains a `./bosinst.data` file with `SWITCH_TO_PRODUCT_TAPE=yes` in the `control_flow` stanza if this was not set prior to making the **mksysb**.
- ▶ If `SWITCH_TO_PRODUCT_TAPE` is set to `yes`, the system will prompt you to remove the **mksysb** media and insert the product media after the **mksysb** backup has been restored.

After the **mksysb** backup installation completes, the installation program automatically installs additional devices and the kernel (uniprocessor or microprocessor) on your system using the original product media you booted from. Information is saved in BOS installation log files. To view BOS installation log files, enter `cd /var/adm/ras` and view the `devinst.log` file in this directory.

If the source system does not have the correct passwords and network information, you may make modifications on the target system now. Also, some products ship device-specific files such as `graPHIGS`. If your graphics adapter is different on the target system, verify that the device-specific filesets for graphics-related LPPs are installed.

If the system you have cloned is using OpenGL or `graPHIGS`, there may be some device filesets from these LPPs that must be installed after a clone.

2.6 Installing additional hardware

When the system is installed, only the drivers and software related to the current hardware and adapters are installed. When you add additional hardware (for example, a new adapter), additional software, such as drivers and software related to the adapter, might be needed.

When the system boots after you have installed the new hardware, `cfgmgr` is automatically running trying to configure the new hardware. This will typically succeed if similar hardware is already present. For example, if you added an additional SCSI device, such as a tape drive.

If the new hardware is not listed using the `lscdev` command (for example, `lscdev -Cc adapter`), you can determine the missing software by running `cfgmgr` from a command window. The `cfgmgr` command will display a warning and indicate the missing driver filesets. For example:

```
# cfgmgr
cfgmgr: 0514-621 WARNING: The following device packages are required for
device support but are not currently installed.
devices.pci.token-ring:devices.pci.14101800:devices.pci.IBM.42H0658:devices.pci
.
IBM.25H3037:devices.pci.IBM.38H5818
```

Install the missing fileset from the AIX CDs using standard fileset installation procedures. Rerun the `cfgmgr` command, and if no warnings are displayed, then the new hardware is installed. Alternatively, you can use the `-i` flag of `cfgmgr` to install the missing driver software directly. For example, to install device drivers automatically during configuration (with the software contained in a directory), enter:

```
# cfgmgr -i /usr/sys/inst.images
```

Depending on the type of hardware, continue with the setup of the hardware following the provided instructions. For third-party hardware, the necessary drivers are normally provided with the product.

Some hardware might require additional software in order to be used. For example, when adding a graphics card on a server that originally has been installed without a graphic adapter, graphic software, such as the X11/CDE software, is not installed (at least not by default). Running the `xinit` command without properly installed filesets will produce the following errors:

```
# xinit
Could not load /usr/lpp/gai/
1356-800 xinit: Unable to start the X server.
1356-805 xinit: Giving up.
1356-811 xinit: Unable to connect to X server.
1356-803 xinit: Server error.
1356-804 xinit: Client error.
```

Install this software by using the Graphics-Startup installation bundle. A bundle is a collection of packages, products, or individual filesets.

To install this, use the SMIT command:

```
# smitty install_bundle
```

Then select the Graphics-Startup bundle. The list of system bundles is located in `/usr/sys/inst.data/sys_bundles`.

When finished installing the new hardware, remember to reconnect all parts (keyboard, mouse, and so on) you detached from the system. Some RS/6000 systems are critical about missing components and will stop on the SMS menu when you boot the system.

2.6.1 Post-installation problems

After installation, the Installation Assistant allows you to configure the basic system settings; however, a common set of problems can occur.

TCP/IP problems

Communication problems with TCP/IP are, in themselves, a large topic, and detailed coverage is beyond the scope of this publication. However, the following set of problems is quite common with systems that have just been installed.

- ▶ If either the DNS is missing or it is not configured correctly, you will not be able to ping other computers by host name, provided that you do not want to use the `/etc/hosts` file. Check the DNS setting by either looking in `/etc/resolv.conf` or checking the DNS setting using the SMIT `mktcpip` fastpath.
- ▶ A similar problem occurs when the system hangs if you log in using the CDE. As the host name is used in the `DISPLAY` environment, a connection cannot be established to the X-Server. This is especially critical for logins over the network from X-Terminals.

Common problems with CDE

Typical for new installations is that user accounts are established and customized with environment settings in the `.profile` read by the korn shell. When a user logs into the system using the common desktop environment (CDE), these environment settings are *not* performed; however, if the user performs a telnet log in, it works. This is due to the fact that, as per default, the CDE session does not source the `.profile`. In the `.dtprofile`, you have to uncomment the setting of the `DTSOURCEPROFILE` flag. The following is extracted from the `.dtprofile` file:

```
#
# If $HOME/.profile (.login) has been edited as described above, uncomment
# the following line.
#
DTSOURCEPROFILE=true
```

The CDE is known for name resolution problems, since it uses the INET domain socket instead of the UNIX domain socket to set up the necessary connection. For example, if you configure the DNS client configuration on the system, and pull off the Ethernet cable to move the system physically, it may take some time before the login panel on the CDE appears.

If you are using the DNS for a name resolution mechanism, you should consider applying the following method to avoid this situation.

1. Create the local host file (/etc/hosts) containing your host name and local host as follows.

```
# /etc/hosts -- local host table
127.0.0.1 localhost.yourdomain.com localhost loopback.yourdomain.com
loopback
XX.YY.ZZ.AA yourhost.yourdomain.com yourhost
```

2. Create the /etc/netsvc.conf file, which alters the system default name resolution order.

```
hosts = local4,bind4
```

3. Confirm the file owner and permission (should be set to permission 0644, and owned by user root, group system).

```
# ls -l /etc/hosts
-rw-r--r--  1 root    system      807 Jun 20 18:13 /etc/hosts
# ls -l /etc/netsvc.conf
-rw-r--r--  1 root    system      21 Jun 20 18:13 /etc/netsvc.conf
```

For further reading on post-installation problem determination, see *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496.

2.7 Installing online documentation

The AIX documentation is *not* automatically installed when you install the BOS. The documentation is contained on two documentation CD-ROMs.

For information about installing the AIX online documentation and setting up the Documentation Library Service, see the AIX version-specific installation guide.

Note: Note that the traditional man pages presented with the `man` command in AIX Version 4.3 or higher are pages converted from the HTML pages (filesets bos.html.en_US.cmds.*). These are part of the commands' reference on the base documentation CD.

2.8 Command summary

The following section provides a list of the key commands discussed in this chapter.

2.8.1 The blvset command

The **blvset** command resets the pad string in the boot logical volume (BLV). The command has the following syntax:

```
/usr/lpp/bosinst/blvset -d Device -p [level | menu ]  
/usr/lpp/bosinst/blvset -d Device -g [level | menu ]
```

The commonly used flags of the **blvset** command are provided in Table 2-1.

Table 2-1 Commonly used flags of the blvset command

Flag	Description
-d <i>device</i>	Device location of the rootvg, for example, /dev/hdisk0.
-p	Puts pad string to the BLV. Uses level option for AIX Version 4.x or 5.x. Uses menu option for AIX Version 3.2.
-g	Gets pad string from the BLV. Uses level option for AIX Version 4.x or 5.x. Uses menu option for AIX Version 3.2.

2.8.2 The cfgmgr command

The **cfgmgr** command configures devices and optionally installs device software. The command has the following syntax:

```
cfgmgr [ -f | -s | -p Phase ] [ -i Device ] [ -l Name ] [ -v ]
```

The commonly used flags of the **cfgmgr** command are provided in Table 2-2.

Table 2-2 Commonly used flags of the cfgmgr command

Flag	Description
-i <i>device</i>	This flag enables the installation of additional software needed by the hardware identified by cfgmgr .

2.8.3 The `alt_disk_install` command

The `alt_disk_install` command installs an alternate disk with a `mksysb` install image or clones the currently running system to an alternate disk. The command has the following syntax:

► Creation of Alternate Disk Installation:

```
alt_disk_install { -d device | -C } [ -i image.data ] [ -s script ] [ -R  
resolv_conf ] [ -D ] [ -B ] [ -V ] [ -r ] [ -p platform ] [ -L mksysb_level  
]  
[ -b bundle_name ] [ -I installp_flags ] [ -l images_location ]  
[ -f fix_bundle ] [ -F fixes ] [ -e exclude_list ] [ -w filesets ] [ -n ]  
[ -P phase_option ] target_disks
```

► Operation on existing Alternate Disk Installation:

```
alt_disk_install { -X [ volume group ] | -q disk | -S | -W disk | -v  
new_volume_group_name disk }
```

The commonly used flags of the `alt_disk_install` command are provided in Table 2-3.

Table 2-3 Commonly used flags of the `alt_disk_install` command

Flag	Description
-C	Clones the rootvg.
-d <i>device</i>	Installs <code>mksysb</code> from device, for example, <code>/dev/rmt0</code> .
-W <i>disk</i>	Wakes up a volume group for data access.
-S	Puts the alternate volume group to sleep.
-X	Removes the <code>altinst_rootvg</code> volume group.

2.9 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

1. By default, which of the following installation methods will retain existing user data in the root volume group (excluding user data in /usr, /tmp, /var, and / (root) file systems)?
 - A. Migration Installation
 - B. Preservation Installation
 - C. User data is not retained by default
 - D. Both Migration Installation and Preservation Installation
2. A graphics card has been added to a system and the system is rebooted. When running the `lsdev -Cc adapter` command, the graphics card is not listed in the output. Which of the following actions should be performed?
 - A. Reboot system using `run mkdev -l`.
 - B. Install device support using the `cfgmgr` command.
 - C. Use `smit console` to redirect the console.
 - D. Use `smit devices` and choose the correct adapter.
3. A graphics card has been added to a system. After rebooting, the system hangs on login with a blue screen indicating Please wait... and Opening the AIXWindows Desktop. Which of the following procedures should be performed next?
 - A. Reinstall CDE software.
 - B. Check name resolution.
 - C. Uncomment the last line in the .dtprofile file.
 - D. Boot in maintenance mode and perform file system checks.
4. The AIX level was Version 4.2.1. After upgrading the operating system to Version 4.3.3, the `man` pages are missing. Which of the following procedures should be performed to access the `man` pages?
 - A. Reinstall Info Explorer.
 - B. Relink /usr/lpp/info/data/ispaths to ispaths.full.
 - C. Reinstall the `man` fileset from the third AIX Version 4.3.3 CD-ROM.
 - D. Reinstall the commands reference from the base documentation CD-ROM.

5. Following an AIX upgrade from AIX Version 4.2.1 to Version 4.3.3, and booting from the CD-ROM, the Migration option is not offered as an option for the installation type. Which of the following procedures should be performed to activate the migration as an option?
 - A. Continue using the Preservation Installation option.
 - B. Run `installp` to clean up after a failed installation.
 - C. Run `sync1vodm` to synchronize the ODM information.
 - D. Run `/usr/lpp/bosinst/b1vset` to reset the *pad string* in the boot logical volume.

6. A system administrator is attempting to recover a system from a failed migration. However, there is no current backup of the system. Which of the following procedures should be performed to recover the system with minimum loss of data?
 - A. Install the system by using the migration option.
 - B. Install the system by using the preservation option.
 - C. Install the system by using the New and complete overwrite option.
 - D. Boot the system in maintenance mode and copy the ODM from ramfs to the ODM disk.

2.9.1 Answers

The following are the preferred answers to the questions provided in this section.

1. B
2. B
3. B
4. D
5. D
6. B

2.10 Exercises

The following exercises provide sample topics for self study. They will help ensure comprehension of this chapter.

1. Perform a New and Complete Overwrite BOS installation on a dedicated test system. If no available systems exists, try to add an new hard disk and perform the BOS installation on this new disk. If you do, remember to change the bootlist to point to the new disk.
2. If an old AIX Version 3.1, AIX Version 4.1, or AIX Version 4.2 spare system is available, try a Migration Installation.
After the migration, check the system and have a look at the LPPs installed.
3. Install the filesets bos.html.en_US.cmds.cmds1 to bos.html.en_US.cmds.cmds6 and afterwards try the **man** command.
4. *Advanced Exercise:* On a 4.3.3 test system, install to a new unallocated (additional) disk using the Alternate Disk Installation method and the corresponding **alt_disk_install** command.



Additional software installation

The following topics are discussed in this chapter:

- ▶ What can be installed
- ▶ How software can be installed
- ▶ How you can remove or fix filesets after a failed installation

After installing the base operating system, you will normally install additional software. Such additional software can include anything between a database or a single freeware command. This chapter will explain the differences in the installation methods.

3.1 What can be installed

AIX provides a variety of software installation options. The challenge is to select the right option for the right purpose.

A preview of the software is always a good place to start. SMIT provides a panel to do this, as shown in Figure 3-1.

```

                                List Software on Media and Related Information

Move cursor to desired item and press Enter.

List Filesets in a Bundle
List Software on Installation Media
List Software Fixes (APARs) on Installation Media
List Supplemental Fileset Information on Installation Media
Show License Agreements on Installation Media

F1=Help          F2=Refresh      F3=Cancel      F8=Image
F9=Shell         F10=Exit       Enter=Do
```

Figure 3-1 List Software on Media and Related Information menu

An overview of the installation options is provided in the following sections.

3.1.1 Installation options

Understanding what you are to install, and what is meant by the word *install*, will help you to choose the right command. In order to better understand this, a brief description of the available installation packaging options is provided in the following sections.

Files

When working with files, remember that an ordinary file cannot be installed through the **smitty installp** dialog, which uses the **installp** or the **instfix** command. The most common scenario when you want to install a single file is probably restoring a file from a backup tape. For example, if the RS/6000 LED code indicates a syntax error in `/etc/inittab` (LED code 553) you could choose to

fix the problem by reinstalling `/etc/inittab` from a **mksysb** tape or any other available backup media that has the file. The way to do this depends on the backup procedure used. If SMIT was used for generating the tape, the **backup** command has been used. In this case, the corresponding command to restore a file is **restore**.

Fileset

Fileset installation is probably the most common installation once a system installation is completed. A fileset is the smallest installable base unit for the AIX operating system. A fileset includes all files that constitute either a complete product, such as `bos.net.uucp`, or a separately installable part of a product, such as `bos.net.nfs.client`. More information on how to work with filesets can be found in 3.2.2, “Filesets” on page 38.

Bundle

A bundle is a list of filesets that are suited for a particular use. The bundle is only used as a template; you still have to define the media where the install image files are placed. The templates are placed in `/usr/sys/inst.data/sys_bundles` and in `/usr/sys/inst.data/user_bundles`. The templates are command text files, so this gives you an opportunity to manipulate the content of the system-defined bundles as well as create your own bundles, as shown in 3.2.1, “Bundles” on page 37. A sample of the system-defined bundles is:

- ▶ App-Dev
- ▶ Client
- ▶ Server
- ▶ Media-Defined
- ▶ Pers-Prod
- ▶ Graphics-Startup

Licensed Program Products

Licensed Program Products (LPPs) are complete software products, including all packages associated with the licensed programs. A good example of an LPP is the base operating system (BOS).

PTFs and emergency fixes (e-fixes)

PTF is the acronym for program temporary fix. A PTF is an updated, or fixed, fileset. The e-fix, or emergency fix, is a temporary solution to a unique problem. Each fix has an Authorized Program Analysis Report (APAR) number. The **instfix** command and the **smitty update_by_fix** allows you to install a fix or set of fixes without knowing any information other than the APAR number or other unique keywords that identify the fix. Emergency fixes will eventually, after testing and verification, become PTFs. The **installp** command can also be used

because PTFs are just regular filesets. More on how to work with APARs in discussed in “The instfix command” on page 34.

3.1.2 Installation methods

As discussed earlier in this chapter, you need to have an understanding of what you are to install. This will decide what command to use. The two commands mentioned so far are the **installp** command, which is to be used for fileset, LPP, and bundle installation, and the **instfix** command, which is to be used when AIX Support asks you to install a fix.

When either the **installp** or **instfix** command is used, AIX usually expands the necessary file system (for example, /usr) space automatically. This ability is based upon the information provided by the installable package. In the rare case that you receive an error message such as insufficient disk space, then you can increase the target file system explicitly to avoid failure of the installation.

SMIT installation menus

SMIT installation menus provide all the tools you need to install software. Under the fastpath, **smitty installp**, you will find menus that either use the **installp** command or the **instfix** command. All menus, except the highlighted one shown in Figure 3-2, use the **installp** command. The highlighted one uses the **instfix** command.

```
Install and Update Software

Move cursor to desired item and press Enter.

Install and Update from LATEST Available Software
Update Installed Software to Latest Level (Update All)
Install and Update Software by Package Name (includes devices and printers)
Install Software Bundle (Easy Install)
Update Software by Fix (APAR)
Install and Update from ALL Available Software

F1=Help          F2=Refresh      F3=Cancel      F8=Image
F9=Shell         F10=Exit       Enter=Do
```

Figure 3-2 SMIT installation menus

The installp command

The **installp** command is used for various installation tasks. One of the most useful features is the possibility to choose the state of the installed fileset, applied or committed. This is defined by the **-a** flag for apply and **-ac** for commit. After you have installed a fileset in applied state, you can choose to commit the fileset by using the **-c** flag.

Another useful function the **installp** command provides is the possibility to look at the media for installable fileset. This is done with the **-l** flag.

Table 3-1 lists some commonly used **installp** flags and their descriptions.

Table 3-1 *Flags used at installation with the installp command*

Flag	Description
<code>installp -ad /dev/cd0 fileset</code>	Installing fileset in applied state
<code>installp -acd /dev/cd0 fileset</code>	Installing of fileset in committed state
<code>installp -c fileset</code>	Committing an applied fileset
<code>installp -ld /dev/cd0</code>	Listing of installable filesets on media

Finally, the **installp** command can also be used to clean up or reject an applied fileset and to uninstall a product, as discussed in “How to reject applied software” on page 42 through “How to clean up after a failed or interrupted installation” on page 45.

To examine the current fileset level or installed fileset status, use the **ls1pp** command.

To display the installation and update history information for the `bos.sysmgmt.trace` fileset, use:

```
# ls1pp -h bos.sysmgmt.trace
Fileset      Level      Action      Status      Date        Time
-----
Path: /usr/lib/objrepos
bos.sysmgmt.trace
              4.3.3.0    COMMIT      COMPLETE    06/15/00    09:57:28
              4.3.3.11    COMMIT      COMPLETE    06/16/00    11:19:13

Path: /etc/objrepos
bos.sysmgmt.trace
              4.3.3.0    COMMIT      COMPLETE    06/15/00    09:57:33
              4.3.3.11    COMMIT      COMPLETE    06/16/00    11:19:14
```

As shown, the fileset bos.sysmgt.trace was once updated. It is now in the committed state at the fix level 4.3.3.11.

The **instfix** command

As mentioned in “PTFs and emergency fixes (e-fixes)” on page 31, the **instfix** command provides the function discussed in this section.

- ▶ Specifies whether APARs are already installed or not
- ▶ Views the abstraction text on the specified APAR on installation media or installed system
- ▶ Lists the included fileset updates in APARs

To examine whether specified APARs are already installed on the system, you can use the following command (multiple APAR keywords can be specified if you enclose them in double quotes).

```
# instfix -ik "IX71385 IX71988"
    There was no data for IX71385 in the fix database.
    All filesets for IX71988 were found.
```

In this case, the APAR IX71988 is already applied in this system, though the APAR IX71385 is not applied in this system.

The **instfix** command can also be used to extract information about fixes (called APAR description) already installed and fixes on a media (if you specified the install source media or directory using the -d flag).

For example:

```
# instfix -aik IY01781
IY01781 Abstract: Behavior of ptf_discard has changed

IY01781 Symptom Text:
  ptf_discard() behavior has changed from 43D to 43K and
  causes problem with unmount file systems.
```

```
-----
    All filesets for IY01781 were found.
```

To confirm the contents of the specified APARs, you can use the **instfix** command is as follows.

```
# instfix -ivk IY01781
IY01781 Abstract: Behavior of ptf_discard has changed

    Fileset bos.mp:4.3.2.10 is applied on the system.
    Fileset bos.up:4.3.2.10 is applied on the system.
    All filesets for IY01781 were found.
```

In this case, the APAR IY01781 includes the fileset updates for bos.mp with both 4.3.2.10 levels.

To install all fixes on the media in the tape drive, enter:

```
# instfix -T -d /dev/rmt0.1 | instfix -d /dev/rmt0.1 -f-
```

The first part of this command lists the fixes on the media, and the second part of this command uses the list as an input file. The `-f` flag specifies the input file containing key words or fixes.

Note: The e-fix (emergency fix) is not installable using the `instfix` command, since it is not provided in installable bff format yet. To install the e-fixes contact IBM software service personnel for appropriate installation instructions. The instruction states how to examine the integrity of the e-fix package, how to stop your application instance, how to replace modules, and how to confirm the installation results.

3.1.3 Installation reports

Previously, the use of `smitty installp` menus, the `installp` command, and the `instfix` command were discussed. Depending on what you choose, the installation reports have to be handled in different ways.

SMIT logging

Most SMIT tasks create or add to the `smit.log` and `smit.script` files in your home directory. (These are the same files appended when you run a SMIT session.) The time, name of the task, and the command (flags and parameters included) are added to the end of the `smit.script` file in a format that can easily be used to create executable shell scripts.

The `installp` or the `instfix` command used by SMIT will be logged in `$HOME/smit.log`.

The `installp` command logging

If you choose to use the `installp` command from the command line, you have the possibility to use redirection, as in the example with `instfix`, or you can use the `-e` flag to specify a log file. For example:

```
# installp -ae /tmp/thomasc/installp.log -d /dev/cd0 "bos.dosutil 4.3.3.0"
```

This command would install `bos.dosutil` in an applied state and send all the familiar installation messages to the log file specified as well as report the command syntax used and the time the installation started and finished. Another


```
+ [ false = true ]
+ exit 0
Finished processing all filesets. (Total time: 4 secs).
.....
.....
```

When the target `all` is used in the `installp` command, it means all available filesets on the media.

The `instfix` command log

If you choose to use the `instfix` command from the command line, remember that `instfix` uses `stdout` and `stderr` by default, so you have to manipulate these if you want an installation report. For example:

```
# instfix -k IY07536 -d /tmp/thomasc > /tmp/thomasc/inst.log 2>&1
```

3.2 How to install using different scenarios

As pointed out in the previous discussion, software is divided into smaller and more manageable parts. The following discusses the different methods used.

3.2.1 Bundles

There are several system-defined bundles that are used to make the installation process easier. The system-defined bundles includes `Server`, `Client`, `App-Dev`, `Media-Defined`, `Graphics-Startup`, `Pers-Prod`. Each of the `App-Dev`, `Client`, and `Server` bundles has two files in the `/usr/sys/inst.data/sys_bundles` directory. The first one defines the default definition of the bundle, which has the name extension of `.def`. The second file is the actual bundle file that is recognized by `smitty easy_install` menus. This file has the file extension of `.bnd`. If the `cfgmgr -f` command finds a graphical adapter at IPL, the `Graphics-Startup` bundle will be used for installation of filesets required for use of the graphical console.

By editing the `.bnd` file for the respective bundle, you can include or exclude filesets in the bundle installation. If, after an installation, you want to remove only a single product that was installed with the bundle, you can do this by using `smitty remove` or:

```
# installp -u bos.dosutil
```

Since the bundle is defined as the collection of the installable filesets, there is no easy way to uninstall a bundle in a single step.

Media-defined bundle

Another bundle is the Media-Defined bundle. This bundle is used when a bundle definition is available on the installation media. This feature can be used by application developers to prepare for an `easy_install`.

How to manipulate bundles

You have the option to create your own bundles. It is as simple as creating a file with an editor, for example `vi`, and listing the filesets to be installed. The file must have `.bnd` as the extension. When using the `smitty easy_install` menu (Figure 3-3), you will find your bundle in the list. In this example, it is named `my_own_bundle`.

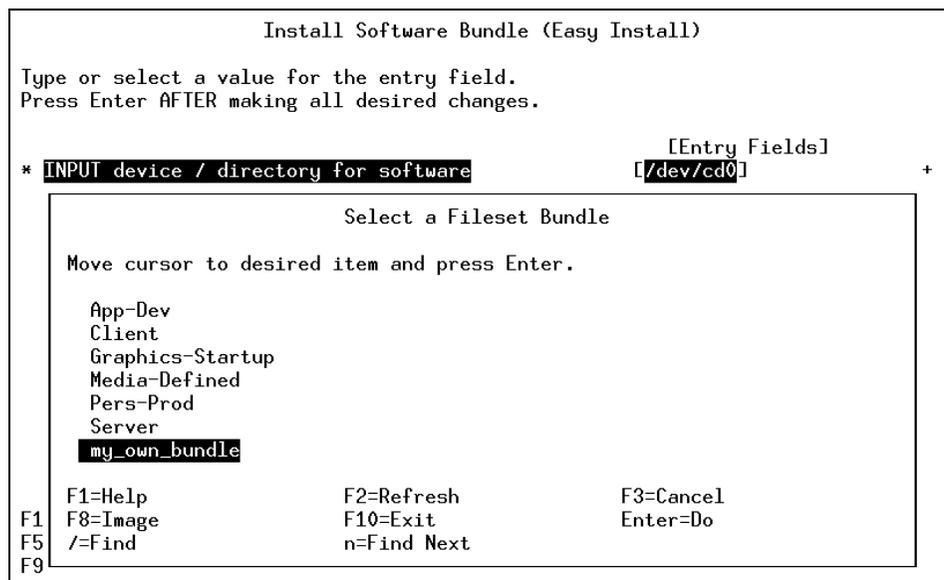


Figure 3-3 Installing user-defined bundle

3.2.2 Filesets

To install the fileset `bos.dosutil`, the following command can be used:

```
# installp -acd /dev/cd0 -X bos.dosutil
```

In this case, the installation will be in *committed* mode, defined by the `-ac` flags. The installation media is `/dev/cd0`, as defined by the `-d` flag, and finally, the `-X` flag attempts to expand any file systems where there is insufficient space to do the installation. If you use SMIT installation menus for your installation with default settings in the menu, then the syntax used will appear as follows:

```
# installp -acNqwd /dev/cd0 bos.dosutil
```

The options and flags shown in Table 3-2 are some of the more helpful one.

Table 3-2 Flags for `installp` used by SMIT

Flag	Description
-ac	Commits
-B	Considers updates only (does not install base fileset)
-g	Includes requisites
-N	Overrides saving of existing files
-q	Quiet mode
-w	Does not place a wildcard at end of fileset name
-X	Attempts to expand file system size if needed
-d	Input device

The commit state should only be used at a base level installation. The reason to use the apply state when updating software will be covered in “How to update a fileset” on page 40.

SMIT and the `installp` command only handle filesets in the bff format. All filesets on the AIX installation media are, by default, in the bff format. If you want to copy installation image files from a installation media to disk for future use, you need to use the `bfcreate` command to make the files usable for the `installp` command. The `bfcreate` command calls the `inutoc` command, which creates a table of contents file (.toc) in the specified directory. For example:

```
# bfcreate -d /dev/cd0 -t /tmp/thomasc -w tmp/thomasc bos.dosutil
```

In this example, the `bfcreate` command takes the install image file, `bos.dosutil`, from `/dev/cd0` and places it in the `/tmp/thomasc` directory, defined by the `-t` flag. While doing this, `bfcreate` creates temporary files, and the `-w` flag indicates that the `/tmp/thomasc` directory can be used for this. Finally, the `inutoc` command will be automatically called to generate a table of contents file. The contents of the directory will appear as follows:

```
# ls -la
total 136
drwxr-xr-x  2 root  system    512 Jun 14 09:56 .
drwxrwxrwt 11 bin   bin      1024 Jun 14 10:17 ..
-rw-r--r--  1 root  system    202 Jun 14 09:56 .toc
-rw-r--r--  1 root  system  57344 Jun 14 09:56 bos.dosutil.4.3.3.0.I
```

There are several other useful flags for the `bfcreate` command, which is beyond the scope of this document.

The preview function

One of the most valuable features of the `installp` command, and the SMIT menus using the `installp` command, is the possibility to perform a preview of an action by running all pre-installation checks for the specified action. This `-p` flag is only valid with `apply`, `commit`, `reject`, and `remove` (`-a`, `-c`, `-r`, and `-u`) flags. The pre-installation check includes verifying prerequisites and checking if the amount of needed space in the target file system is sufficient. To preview an apply of `bos.dosutil`, the command would look as follows:

```
# installp -ape /tmp/thomasc/installp.log -d /dev/cd0 bos.dosutil
```

In SMIT, the choice is done by changing `PREVIEW` only to `yes` (Figure 3-4).

```

                                Install and Update from ALL Available Software
Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* INPUT device / directory for software      /dev/cd0
* SOFTWARE to install                        []
PREVIEW only? (install operation will NOT occur)  yes
COMMIT software updates?                    yes
SAVE replaced files?                         no

PREVIEW only? (install operation will NOT occur)

Press Enter or Cancel to return to the application.
Field cannot be edited.
Use tab key to sequence through valid selections or return
to application and use PF4 or Esc+4 to generate a list.

yes

F F1=Help          F2=Refresh      F3=Cancel
F F8=Image         F10=Exit       Enter=Do
F
```

Figure 3-4 Preview option in SMIT

How to update a fileset

During and after installation, there are four major actions that can be taken with optional software products and service updates. Optional software and service updates can be *applied*, *committed*, *rejected*, and *removed*. Whether a particular action can be taken depends on whether actions have previously been taken.

You can use the `ls1pp` command to check the actual status value of a fileset, for example, the `-h` flag, which will provide us with the installation history of the specified fileset.

```
# ls1pp -h bos.sysmgmt.trace
Fileset      Level      Action      Status      Date      Time
```

```

-----
Path: /usr/lib/objrepos
bos.sysmgmt.trace
      4.3.3.0  COMMIT      COMPLETE    06/15/00    09:57:28
      4.3.3.11 APPLY       COMPLETE    06/15/00    09:58:54

Path: /etc/objrepos
bos.sysmgmt.trace
      4.3.3.0  COMMIT      COMPLETE    06/15/00    09:57:33
      4.3.3.11 APPLY       COMPLETE    06/15/00    09:58:57

```

The use of Commit or Apply has been mentioned so far. As mentioned before, you cannot install a base level of a fileset in any mode other than Commit. The installation will not take any notice of the use of `-a` flag with the `installp` command. It will automatically use the `-ac` flags. See “The `installp` command” on page 33 for further examples of usage of `installp -a` and `installp -ac`. Table 3-3 provides assistance for the Commit or Apply decision.

Table 3-3 Commit or apply?

Fileset level	Install options	Removal	Status after removal
Base level fileset	Commit	Remove/reinstall	Removed
Update of fileset	Commit	Remove/reinstall	Removed
	Apply	Reject	Rejected

By using the Apply option when upgrading a fileset, you keep the former version of that software product in the `/usr/lpp/PackageName` directory so that if you want to return to the former version, you can do so without having to reinstall it. When doing this through SMIT, you need to remember the mutual dependencies in the SMIT menu. When applying a fileset, you should change the SAVE replaced files to yes as shown in Figure 3-5 on page 42.

Enabling the SAVE replaced fileset to yes is a safer option in production use, but it also consumes additional disk space to store the older level fileset modules. Keep it mind, if you use these options frequently, after some weeks or months of normal operations, these saved older level fileset modules should be removed by committing the applied fileset, before applying newer filesets updates.

```

Install and Update from ALL Available Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* INPUT device / directory for software      /dev/cd0
* SOFTWARE to install                       [bos.dosutil]      +
PREVIEW only? (install operation will NOT occur)  no      +
COMMIT software updates?                     no      +
SAVE replaced files?                          yes      +

                                SAVE replaced files?

Press Enter or Cancel to return to the application.
Field cannot be edited.
Use tab key to sequence through valid selections or return
to application and use PF4 or Esc+4 to generate a list.

yes

F F1=Help           F2=Refresh       F3=Cancel
F F8=Image          F10=Exit        Enter=Do
F

```

Figure 3-5 Saving previous version when updating fileset

How to reject applied software

If you find out that a fileset was not compatible with some of your applications, or if the fileset introduced some other problems, you want to be able to remove the fileset. If you have applied the fileset, it is a simple procedure of executing the following command:

```
# installp -r bos.sysmgmt.trace 4.3.3.11
```

When rejecting an applied fileset, you need to give the level of the fileset as a parameter to the **installp** command. When using SMIT, you can do this under the software maintenance menus, as shown in Figure 3-6 on page 43.

```
Software Installation and Maintenance

Move cursor to desired item and press Enter.

Install and Update Software
List Software and Related Information
Software Maintenance and Utilities
Network Installation Management
System Backup Manager

F1=Help      F2=Refresh   F3=Cancel   F8=Image
F9=Shell     F10=Exit    Enter=Do
```

Figure 3-6 Software maintenance and utilities

Under this submenu you can, among other things, choose either to commit an applied fileset or reject an applied fileset, as shown in Figure 3-7.

```
Software Maintenance and Utilities

Move cursor to desired item and press Enter.

Commit Applied Software Updates (Remove Saved Files)
Reject Applied Software Updates (Use Previous Version)
Remove Installed Software

Copy Software to Hard Disk for Future Installation

Check Software File Sizes After Installation
Verify Software Installation and Requisites

Clean Up After Failed or Interrupted Installation

F1=Help      F2=Refresh   F3=Cancel   F8=Image
F9=Shell     F10=Exit    Enter=Do
```

Figure 3-7 SMIT menu for rejecting an applied fileset

Broken or inconsistent software

Filesets in the broken state have encountered a problem while installing an update to the fileset. To correct the problem, the update should be reinstalled, making sure that these options in the menu are set as indicated in Table 3-4.

Table 3-4 Choice of options when reinstalling software

Option	Value
Commit software updates	Yes
Save replaced files	No
Automatically install requisite software	Yes

Filesets or fixes in the inconsistent state have / (root) and /usr parts that are out of synchronization, usually because installation was interrupted or failed. To recover, try to reinstall the software. Alternately, remove the fileset or reject the update.

Install same or older version

If you end up with a fileset in the committed state that you do not want to use, the only way to remove it is to do a reinstallation of the base level of the fileset. This can be done through SMIT under all of the installation menus, as shown in Figure 3-8.

```

                                Install Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* INPUT device / directory for software      /dev/cd0
* SOFTWARE to install                        [_all_latest]      +
PREVIEW only? (install operation will NOT occur)  no      +
COMMIT software updates?                      yes      +
SAVE replaced files?                          no      +
AUTOMATICALLY install requisite software?       yes      +
EXTEND file systems if space needed?            yes      +
OVERWRITE same or newer versions?              yes      +
VERIFY install and check file sizes?          no      +
Include corresponding LANGUAGE filesets?       yes      +
DETAILED output?                              no      +
Process multiple volumes?                     yes      +
ACCEPT new license agreements?                no      +
Preview new LICENSE agreements?                no      +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do
```

Figure 3-8 Install newer version of software

This can also be done by use of the `-F` flag in combination with the normal `-a` for apply or `-ac` for commit. For example:

```
# installp -acFd /dev/cd0 bos.dosutil
```

The `-F` flag can be used to force the installation of a software product even if there is an existing a previously installed version of the software product that is the same as or newer than the version currently being installed. The `-F` flag is not valid with update packages or the `-g` flag. When you use the `-F` flag the `-l` flag is implicit.

How to clean up after a failed or interrupted installation

If the installation, for some reason, has been interrupted, you can have status reports on the fileset indicating a problem. For example, the following output was generated after sending a SIGKILL signal to an installation process:

```
# ls1pp -h bos.dosutil
Fileset      Level      Action      Status      Date        Time
-----
Path: /usr/lib/objrepos
bos.dosutil
           4.3.3.0  APPLY       PENDING     06/15/00    14:16:49

# ls1pp -l bos.dosutil
Fileset      Level      State      Description
-----
Path: /usr/lib/objrepos
bos.dosutil
           4.3.3.0  APPLYING   DOS Utilities
```

When you see this kind of output you should use the `lppchk` command to check what exactly is the consistency problem with the fileset. For example:

```
# lppchk -v
lppchk: The following filesets need to be installed or corrected to bring
the system to a consistent state:

bos.dosutil 4.3.3.0 (APPLYING)
```

The `lppchk` command verifies that files for an installable software product (fileset) match the software vital product data (SWVPD) database information for file sizes, checksum values, or symbolic links. As shown, the `-v` flag is used to verify that the `/` (root), `/usr`, and `/usr/share` parts of the system are valid with each other. In other words, this flag verifies that all software products installed on the `/` (root) file system are also installed on the `/usr` file system.

When you run into this kind of problem, you need to clean up after the installation. This is done in the Clean Up After Failed or Interrupted Installation submenu, as shown in Figure 3-7 on page 43, or by issuing the `installp`

command with the `-C` flag. This will automatically search for filesets in states other than applied or committed and remove them. For example:

```
# installp -C
```

3.3 Uninstall software

To remove a software product, use the `installp` command as in the following example:

```
# installp -u bos.dosutil
```

Or you can use the `smitty remove`, a preferred method, as shown in Figure 3-9.

```
Remove Installed Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

* SOFTWARE name          [Entry Fields]
PREVIEW only? (remove operation will NOT occur)  [bos.dosutil] +
REMOVE dependent software?                        no      +
EXTEND file systems if space needed?              no      +
DETAILED output?                                  no      +

F1=Help          F2=Refresh          F3=Cancel          F4=List
Esc+5=Reset      Esc+6=Command       Esc+7=Edit         Esc+8=Image
Esc+9=Shell      Esc+0=Exit          Enter=Do
```

Figure 3-9 Removing software through SMIT

To remove a software product and its dependent software, use the `installp` command as in the following example:

```
# installp -u -f bos.dosutil -g
```

Or you can use the `smitty remove` command as in Figure 3-10.

```
Remove Installed Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* SOFTWARE name                  [bos.help.msg.en_US.com] +
PREVIEW only? (remove operation will NOT occur)  yes      +
REMOVE dependent software?      yes      +
EXTEND file systems if space needed?           no      +
DETAILED output?                  no      +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell    F10=Exit       Enter=Do
```

Figure 3-10 Removing installed software and dependent software through SMIT

3.4 Summary

As a general summary of this chapter, there is but one single piece of advice:

Understand your task.

Find out what you are to install, choose the right installation option, and finally, if you have to repair a failed installation, choose the right flag to reject, uninstall, or clean up.

3.5 Command summary

The following section provides a list of the key commands discussed in this chapter.

3.5.1 The installp command

As seen below, the **installp** command is a very useful and powerful tool. More on the **installp** command can be found in the *AIX Version 5.0 Installation Guide*, SC23-4112.

To install with apply only or with apply and commit:

```
installp [ -a | -ac [ -N ] ] [ -eLogFile ] [ -V Number ] [ -dDevice ] [ -b ]  
[ -S ] [ -B ] [ -D ] [ -I ] [ -p ] [ -Q ] [ -q ] [ -v ] [ -X ]  
[ -F | -g ] [ -O { [ r ] [ s ] [ u ] } ] [ -tSaveDirectory ] [ -w ]  
[ -zBlockSize ] { FilesetName [ Level ]... | -f ListFile | all }
```

To commit applied updates:

```
installp -c [ -eLogFile ] [ -VNumber ] [ -b ] [ -g ] [ -p ] [ -v ] [ -X ]  
[ -O { [ r ] [ s ] [ u ] } ] [ -w ] { FilesetName [ Level ]... | -f  
ListFile | all }
```

To reject applied updates:

```
installp -r [ -eLogFile ] [ -VNumber ] [ -b ] [ -g ] [ -p ] [ -v ] [ -X ]  
[ -O { [ r ] [ s ] [ u ] } ] [ -w ] { FilesetName [ Level ]... |  
-f ListFile }
```

To uninstall (remove) installed software:

```
installp -u [ -eLogFile ] [ -VNumber ] [ -b ] [ -g ] [ -p ] [ -v ] [ -X ]  
[ -O { [ r ] [ s ] [ u ] } ] [ -w ] { FilesetName [ Level ]... |  
-f ListFile }
```

To clean up a failed installation:

```
installp -C [ -b ] [ -eLogFile ]
```

To list all installable software on media:

```
installp { -l | -L } [ -eLogFile ] [ -d Device ] [ -B ] [ -I ] [ -q ]  
[ -zBlockSize ] [ -O { [ s ] [ u ] } ]
```

To list all customer-reported problems fixed with software or display all supplemental information:

```
installp { -A|-i } [ -eLogFile ] [ -dDevice ] [ -B ] [ -I ] [ -q ]  
[ -z BlockSize ] [ -O { [ s ] [ u ] } ] { FilesetName [ Level ]... | -f  
ListFile | all }
```

To list installed updates that are applied but not committed:

```
installp -s [ -eLogFile ] [ -O { [ r ] [ s ] [ u ] } ] [ -w ] { FilesetName  
[ Level ]... | -fListFile | all }
```

Table 3-5 provides a general summary of some useful **installp** flags.

Table 3-5 Commonly used flags of the installp command

Flag	Description
-ac	Commits
-g	Includes requisites
-N	Overrides saving of existing files
-q	Quiet mode
-w	Does not place a wildcard at end of fileset name
-X	Attempts to expand file system size if needed
-d	Inputs device
-l	Lists of installable filesets
-c	Commits an applied fileset
-C	Cleans up after an failed installation
-u	Uninstalls
-r	Rejects an applied fileset
-p	Previews installation
-e	Defines an installation log
-F	Forces overwrite of same or newer version

3.5.2 The instfix command

The **instfix** command is used to install fixes and to check for installed fixes. More on the **instfix** command can be found in the *AIX Version 5.0 Installation Guide*, SC23-4112.

The syntax of **instfix** is as follows:

```
instfix [ -T ] [ -s String ] [ -S ] [ -k Keyword | -f File ] [ -p ]
[ -d Device ] [ -i [ -c ] [ -q ] [ -t Type ] [ -v ] [ -F ] ] [ -a ]
```

Table 3-6 gives a general summary of some useful **instfix** flags.

Table 3-6 Commonly used flags of the instfix command

Flag	Description
-T	Displays the entire list of fixes present on the media

Flag	Description
-i	Shows if the fix is installed; should be used with -k or -f
-k	Keyword; the APAR number is a keyword
-f	Specifies the input file containing keywords or fixes
-p	Displays filesets associated with keywords

3.5.3 The ls1pp command

The **ls1pp** command displays information about installed filesets or fileset updates.

The syntax of **ls1pp** is as follows:

```
ls1pp { -d | -E | -f | -h | -i | -l | -L | -p } [ -a ] [ -c ] [ -J ] [ -q ]
[ -I ] [ -O { [ r ] [ s ] [ u ] } ] [ [ FilesetName ... | FixID ... | all ]
```

With the **-w** flag, the following syntax is used:

```
ls1pp -w [ -c ] [ -q ] [ -O { [ r ] [ s ] [ u ] } ] [ FileName ... | all ]
```

Table 3-7 gives a general summary of some useful **ls1pp** flags.

Table 3-7 Commonly used flags of the ls1pp command

Flags	Description
-l	Displays the name, most recent level, state, and description of the specified fileset.
-p	Displays requisite information for the specified fileset.
-i	Displays the product information for the specified fileset.
-a	Displays all the information about filesets specified when combined with other flags. This flag shows all updates when combined with the -l flag and all history when combined with the -h flag. This flag cannot be specified with the -f flag.
-d	Displays filesets that are dependents of the specified software. A dependent fileset is one that has the specified software as a prerequisite, corequisite, ifrequisite, or installed requisite.

3.6 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

1. Before installation occurs, which of the following should be used to view the contents of the Media-Defined bundle shipped as a part of AIX?
 - A. The **inutoc** command
 - B. The **installp** command
 - C. The SMIT's List Software on the Installation Media menu
 - D. View contents of files located in `/usr/sys/inst.data/sys_bundles`
2. A system administrator needs to install the Personal Productivity bundle, without the DOS utilities fileset (`bos.dosutil`). Which of the following procedures is the best way for a system administrator to successfully perform the installation?
 - A. Install the bundle and then remove the DOS utilities fileset.
 - B. Use the **chgcfs** command with the `-d` option for `bos.dosutil`.
 - C. Use the ODM edit (**odme**) utility to remove the product definition from the system's ODM database.
 - D. Remove the `bos.dosutil` name from `/usr/sys/inst.data/sys_bundles/Pers-Prod.bnd` for the Personal Productivity bundle.
3. Which of the following procedures is most appropriate to remove a committed optional program product from an AIX system?
 - A. Use the Remove Software Products menu in SMIT.
 - B. Use **installp -C** to remove optional program products.
 - C. Use the reject option for **installp** to remove each unwanted program product installed on a system.
 - D. There is no systematic way to remove optional program products.

3.6.1 Answers

The following are the preferred answers to the questions provided in this section.

1. C
2. D
3. A

3.7 Exercises

In this exercise, the system has been installed using AIX Version 4.3.3 maintenance level 2. Start with choosing a fileset well suited for this exercise (not system critical and with some fix levels available, for example, bos.sysmgt.trace). Verify that the previous version of your chosen fileset is in committed state before doing this exercise. Do not test this in a production system.

1. Download an updated fileset (for example, bos.sysmgt.trace 4.3.3.11) from the Web. This can be done from:

<http://techsupport.services.ibm.com/server/nav?fetch=a4ojc>

Install the fileset in applied state. Verify the state of the fileset. Reject the fileset.

2. Install the fileset in applied state once more, and during the installation send SIGKILL to the **insta11p** PID. Check for installation problems. Clean up after the failed installation.



System boot

The following topics are discussed in this chapter:

- ▶ A general boot process
- ▶ Differences between MCA and PCI systems
- ▶ Boot phase 1
- ▶ Boot phase 2
- ▶ Boot phase 3
- ▶ Common boot problem scenarios and how to fix them

Boot problems are among the most common problems. When looking at the subject of installation and system recovery, it is important that you have good knowledge of the boot process. This chapter begins with a general overview of the boot process, then expands on the details and discusses the process with their LED codes in further detail. A summary of the LED codes can be found in 4.7.2, “LED codes” on page 74.

4.1 General overview of the boot process

Both hardware and software problems can cause the system to halt in the boot process. The boot process is also dependent on what hardware platform is used. In the initial startup phase, there are some important differences between MCA and PCI systems, and these differences will determine the way to handle a hardware-related boot problem. These differences will be covered in 4.2, “BIST - POST” on page 55.

A general representation of the boot process is shown in Figure 4-1.

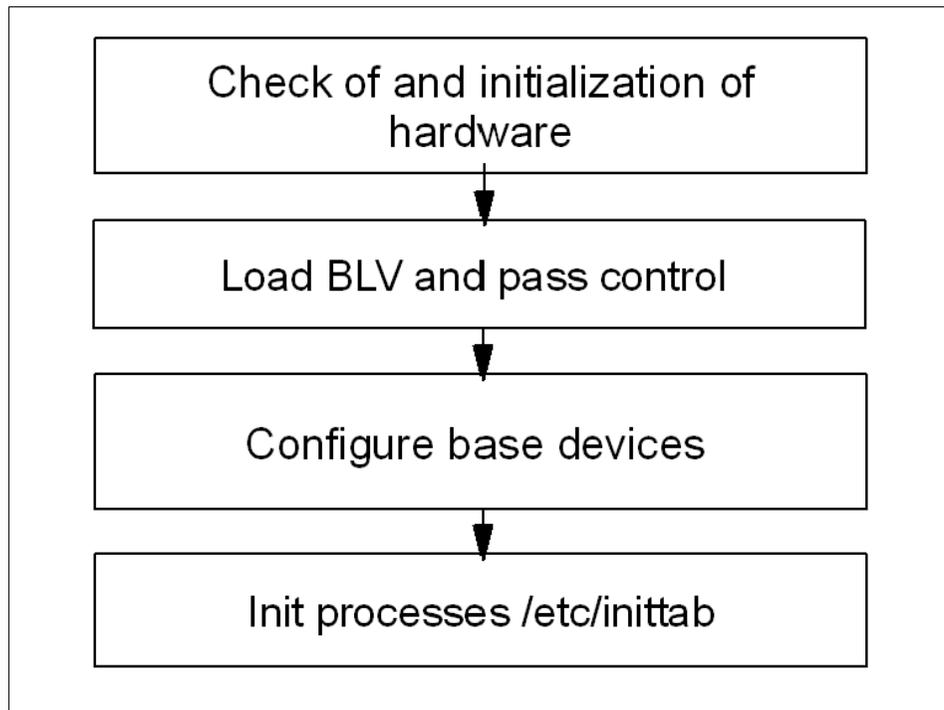


Figure 4-1 General boot order

The initial hardware check is to verify that the primary hardware is OK. This phase is divided into two separate phases on an MCA system, first a built-in self test (BIST), and the second, a power-on self test (POST). On PCI systems, it is handled by a single POST. After this, the system attempts to load the boot logical volume (BLV) into a RAM file system (RAMFS) and passes control to the BLV. Since there is no Logical Volume Manager (LVM) device driver loaded into memory, the system has to find the correct BLV location without using LVM. The correct address information (offset and length) is written in the first physical disk sector (512 byte block) on the boot disk. This sector is called boot record.

Note: Content of BLV:

- ▶ Soft ROS (PCI architecture models only)
Soft ROS performs certain system initialization on the PCI architecture models that is not provided by their ROS, but required by AIX.
- ▶ AIX kernel
The kernel is always loaded from the BLV. There is a copy of the kernel in /unix (soft link to /usr/lib/boot/unix_mp or unix_up). This version is used to build the hd4 file system where the kernel image is read during system boot.
- ▶ rc.boot
This is the configuration script that will be called three times by the init process during boot.
- ▶ Reduced ODM
Device support is provided only to devices marked as base devices.
- ▶ Boot commands
For example, **cfgmgr** and **bootinfo**.

Because the rootvg is not available at this point, all the information needed for boot is included in the BLV used for creation of the RAMFS in the memory. After this, the init process is loaded and starts to configure the base devices. This is called boot phase 1 (init executes the rc.boot script with an argument of 1).

The next step, called phase 2, aims at activating rootvg, and this is probably the phase where the most common boot problems occur, for example, a file system or the jfslog is corrupt. Next, the control is passed to the rootvg **init** command and the RAMFS is released.

Finally, the init process, now loaded from disk (not the BLV init), executes the rc.boot script with an argument of 3 to configure the remaining devices. This final stage is done from /etc/inittab.

4.2 BIST - POST

As mentioned before, there are differences between the classic RS/6000 system with MCA architecture and the PCI systems that are delivered today. The MCA system will be discussed first.

4.2.1 MCA systems

At a system startup of an MCA system, the first thing that happens is a Built-In Self-Test (BIST). These tests reside in EPROM chips, and the hardware tested by BIST are mainly components on the motherboard. LED codes shown during this phase of the startup will be in the range of 100–195, depending on the status of the hardware. After this, the POST will be initialized.

The task of the POST is to find a successful hardware path to a BLV. All hardware that is required to load a boot image is tested. The LED codes at this stage are in the range of 200–2E7, and both hardware and software problems can cause a halt in the startup process at this stage.

On an MCA system, the load of the BLV starts with checking the bootlist. The bootlist is defined by the key position. When the key is in normal position applications will be started as well as network services. This is done when the init process reads /etc/inittab and executes the configuration scripts referenced in the file. A normal boot is represented by the runlevel 2. The /etc/inittab file is discussed in further detail in 4.1, “General overview of the boot process” on page 54, and 4.5.1, “The /etc/inittab file” on page 70. To manipulate the boot list for normal mode, use the following command:

```
# bootlist -m normal hdisk0 hdisk1 rmt0 cd0
```

This will cause the system to first search hdisk0 for a usable BLV. If there is no BLV at hdisk0, then hdisk1 will be searched, and so on.

The service boot list is used when booting the system for maintenance tasks. No applications or network services will be started. To check what the service bootlist looks like, use the -o option, which was introduced with AIX Version 4.2.1, as follows:

```
# bootlist -m service -o
fd0
cd0
rmt0
hdisk2
ent0
```

Another feature introduced with AIX Version 4.2 is the use of generic device names. Instead of pointing out the specified disk, with hdisk0 or hdisk1, you can use the generic definition of SCSI disks. For example:

```
# bootlist -m service cd rmt scdisk
```

This will cause the system to probe any CD, then probe any tape drive, and finally, probe any SCSI disk for a BLV. The actual probing of the disk is a check of sector 0 for a boot record, which, in turn, will point out the boot image.

Changes to the boot list can also be done through the diag menus. At the Function Selection menu, select **Task Selections**, as shown in Figure 4-2.

```
FUNCTION SELECTION801002

Move cursor to selection, then press Enter.

Diagnostic Routines
  This selection will test the machine hardware. Wrap plugs and
  other advanced functions will not be used.
Advanced Diagnostics Routines
  This selection will test the machine hardware. Wrap plugs and
  other advanced functions will be used.
Task Selection(Diagnostics, Advanced Diagnostics, Service Aids, etc.)
  This selection will list the tasks supported by these procedures.
  Once a task is selected, a resource menu may be presented showing
  all resources supported by the task.
Resource Selection
  This selection will list the resources in the system that are supported
  by these procedures. Once a resource is selected, a task menu will
  be presented showing all tasks that can be run on the resource(s).

F1=Help           F10=Exit          F3=Previous Menu
```

Figure 4-2 Function selection menu in diag

In the list of tasks, choose **Display or Change Bootlist**, as shown in Figure 4-3 on page 58.

```
TASKS SELECTION LIST 801004

From the list below, select a task by moving the cursor to
the task and pressing 'Enter'.
To list the resources for the task highlighted, press 'List'.

[MORE...18]
  Display Firmware Device Node Information
  Display Hardware Error Report
  Display Hardware Vital Product Data
  Display Microcode Level
  Display Previous Diagnostic Results
  Display Resource Attributes
  Display Service Hints
  Display Software Product Data
  Display System Environmental Sensors
  Display Test Patterns
  Display or Change Bootlist
  Download Microcode
[MORE...12]

F1=Help           F4=List           F10=Exit          Enter
F3=Previous Menu
```

Figure 4-3 Task selection menu in diag

Finally, you have to choose whether to change the Normal mode bootlist or the Service mode bootlist, as shown in Figure 4-4.

```
DISPLAY/ALTER BOOTLIST 802590

Select an option, then press Enter.

  Normal mode bootlist
  This selection allows displaying, altering, or erasing
  the normal mode bootlist.
  Service mode bootlist
  This selection allows displaying, altering, or erasing
  the service mode bootlist.

F3=Cancel           F10=Exit
```

Figure 4-4 Display/alter bootlist menu in diag

At this point, a lot of things can cause a boot problem. The boot list could point to a device that does not have a BLV, or the devices pointed to are not accessible because of hardware errors.

The following sections cover some problems that can cause a halt. All problems at this stage of the startup process have an error code defined, which is shown in the LED display on the front panel.

LED 200

The LED code 200 is connected to the secure key position. When the key is in the secure position, the boot will stop until the key is turned either to the normal position or the service position. The boot will then continue.

LED 299

A LED code of 299 shows that the BLV will be loaded. If this LED code is passed, then the load has been successful. If, after passing 299, you get a stable 201, then you have to recreate the BLV as discussed in “How to recreate the BLV” on page 59.

MCA LED codes

Table 4-1 provides a list of the most common LED codes on MCA systems. More of these can be found in the *Message Guide and References*, which is part of the AIX Version 4 base documentation.

Table 4-1 Common MCA LED codes

LED	Description
100–195	Hardware problem during BIST.
200	Key mode switch in secure position.
201	1. If LED 299 passed, recreate BLV. 2. If LED 299 has not passed, POST encountered a hardware error.
221 721 221–229 223–229 225–229 233–235	The bootlist in NVRAM is incorrect (boot from media and change the bootlist), or the bootlist device has no bootimage (boot from media and recreate the BLV), or the bootlist device is unavailable (check for hardware errors).

How to recreate the BLV

When the LED code indicates that the BLV cannot be loaded, you should start by checking for hardware problems, for example, cable connections. The next step is to start the system in maintenance mode from an external media, for example,

an AIX installation CD. Use the **Access this Volume Group** action and start a shell menu for recreation of the BLV (this menu is also used if the boot problem was due to an incorrect bootlist). Execute the following command if you want to recreate the BLV on hdisk0:

```
# bosboot -ad /dev/hdisk0
```

Another scenario where you might want to create a BLV with the **bosboot** command is with a mirrored rootvg. Just mirroring this volume group does not make the disks containing the secondary copy bootable, since the boot record is not a logical volume (AIX mirroring works only on logical volumes). You still have to define the disks in the bootlist and execute the **bosboot** command on the secondary copy. The **bosboot** command creates both the BLV and the boot record.

The following is a short summary on how to access the maintenance menus.

1. Boot the system from media.
2. At the installation menu, choose **Start Maintenance for System Recovery**.
3. On the next menu, choose **Access a Root Volume Group**.
4. A list of accessible disks is shown. Choose the rootvg disk.
5. Finally, choose **Access this Volume Group and start a shell** when you want to recreate the BLV. Change the bootlist or forgotten root password.

Choose **Access this Volume Group and start a shell before mounting file systems** if file systems or the jfslog in the rootvg are corrupt.

4.2.2 PCI systems

When booting PCI systems, there are important differences from the MCA systems. It has already been mentioned that there is an absence of BIST. Another difference is the absence of the key switch. Modern PCI systems use a logical keymode switch, which is handled by the use of function key. Also, the diag function is missing on some older PCI systems. The following section discusses how to change the bootlist and discusses the support of the normal and service boot options on PCI systems.

Changing the bootlist on PCI systems

All PCI systems have System Management Services (SMS) menus. On most systems, these menus can be accessed by pressing F1 or 1 when the console is initiated (the use of F1 or 1 depends on the use of graphical display or ASCII terminal). At this time, a double beep is heard. Depending on the PCI model, there are three or four choices on the SMS main menu. One of these is named boot. Under this menu, you can define the bootlist. The SMS main menu from a

43P-140 is shown in Figure 4-5. Newer PCI systems also have an additional selection called multiboot.



Figure 4-5 SMS main menu

Changing the boot order can also be done with the **bootlist** command. For example, to boot from either CD-ROM or hdisk0, enter:

```
# bootlist -m normal cd0 hdisk0
```

Normal boot and service boot on PCI systems

Some PCI systems do not support service mode, for example, the 7248-43P. The only way to boot in another mode, such as maintenance mode, is to change the normal bootlist. This can be done with the **bootlist -m normal** command if the system is accessible. If the system is not accessible, this can be done by booting from media and changing the bootlist through the SMS menu.

All PCI systems have a default bootlist. On modern PCI systems, this default bootlist can be accessed by using the F5 function key. On older PCI systems, this cannot be done. Instead, a single bootlist is provided that can be reset to the default values by removing the battery for about 30 seconds. This is because the bootlist is stored in NVRAM, and the NVRAM is only non-volatile as long as the battery is maintaining the memory.

Newer PCI architecture machines (for example, the 43P-150) support a service bootlist. The simplest way to find out if a particular system supports the service boot option is to execute:

```
# bootlist -m service -o
0514-220 bootlist: Invalid mode (service) for this model
```

The error message 0514-220 indicates that the system does not support the service boot option.

All new PCI systems support the following key allocation standard:

- ▶ F1 or 1 on ASCII terminal: Starts System Management Services
- ▶ F5 or 5 on ASCII terminal: Boots diag; uses default boot list of fd, cd, scdisk, or network adapter
- ▶ F6 or 6 on ASCII terminal: Boots diag; uses of custom service boot list

POST LED codes on PCI systems

On old PCI systems, such as the 7020-40P or the 7248-43P, the LED display is missing, so there will be no LED codes to help solve boot problems. Fortunately, this has been changed on modern PCI systems, but the error codes generated during this phase of the system startup differs from model to model. The only way to figure out the exact meaning of an error code is to refer to the Service Guide delivered with the system. IBM provides a Web page where Service Guides for most PCI systems are available in HTML and PDF format. The URL is:

http://www.ibm.com/servers/eserver/pseries/library/hardware_docs/index.html

4.3 Boot phase 1

So far, the system has tested the hardware, found a BLV, created the RAMFS, and started the init process from BLV. The rootvg has not yet been activated. From now on, the boot sequence is the same on both MCA systems and PCI systems.

The workflow for boot phase 1 is shown in Figure 4-6 on page 63.

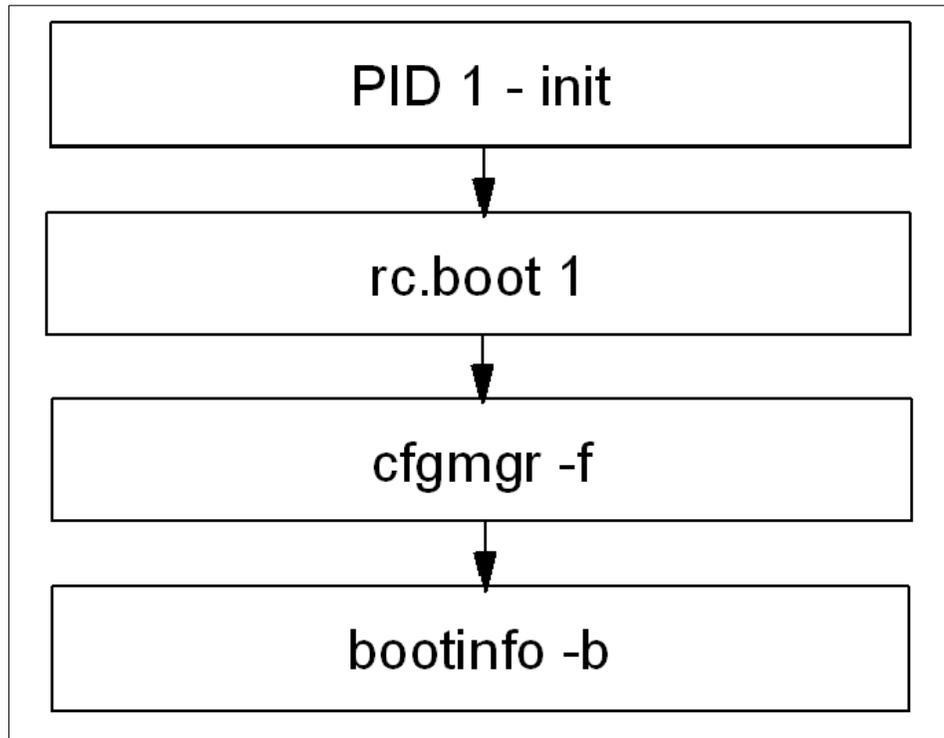


Figure 4-6 Boot phase 1

During this phase, the following steps are taken.

1. The `init` process started from RAMFS executes the boot script `rc.boot 1`. At this stage, the `restbase` command is called to copy the reduced ODM from the BLV into the RAMFS. If this operation fails, you will see a LED code of 548.
2. After this, `cfgmgr -f` reads the `Config_Rules` class from the reduced ODM. In this class, devices with the attribute `phase=1` will be considered base devices. Base devices are all devices that are necessary to access rootvg. The process invoked with `rc.boot 1` attempts to configure devices so that rootvg can be activated in the next `rc.boot` phase.
3. At the end of boot phase 1, the `bootinfo -b` command is called to determine the last boot device. At this stage, the LED shows 511.

4.4 Boot phase 2

In boot phase 2 the rc.boot script is passed the parameter 2. The first part of this phase is shown in Figure 4-7.

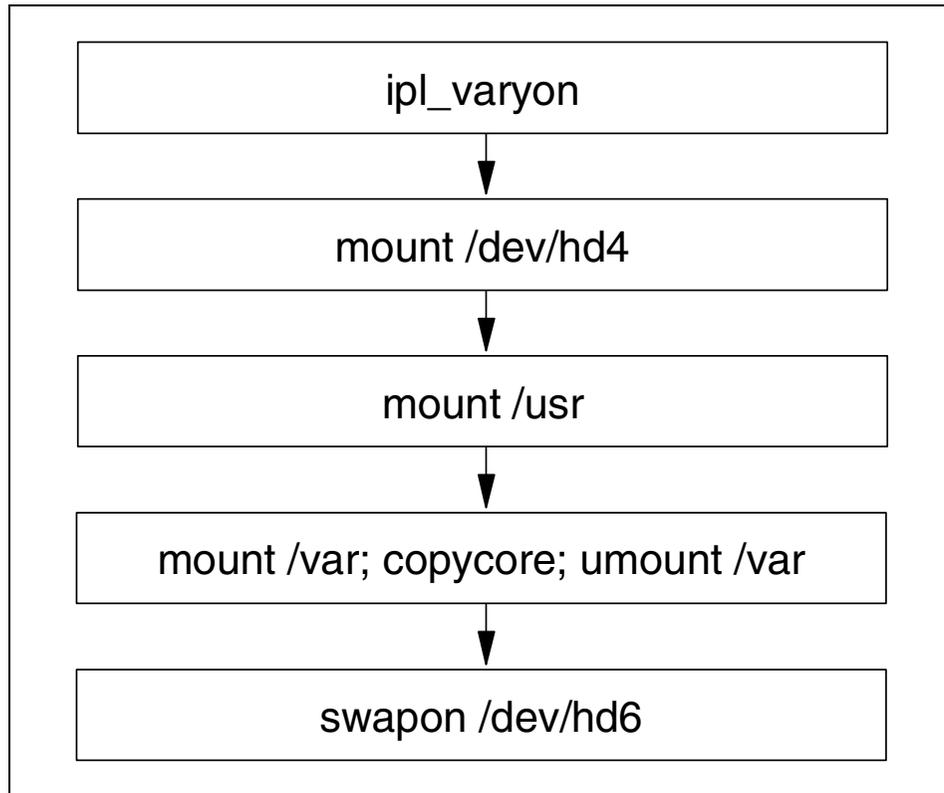


Figure 4-7 Boot phase 2, part one

During this phase, the following steps are taken.

1. Rootvg will be varied on with the special `ipl_varyon` utility. If this command is not successful, one of the following LED codes will appear: 552, 554, or 556.
2. After the successful execution of `ipl_varyon`, the root file system (`/dev/hd4`) will be mounted on a temporary mount point (`/mnt`) in RAMFS. If this fails, 555 or 557 will appear in the LED display.
3. Next, the `/usr` and `/var` file system is mounted. If this fails, the LED 518 appears. The mounting of `/var` at this point enables the system to copy an eventual dump from default dump devices, `/dev/hd6`, to the default copy directory, `/var/adm/ras`. This is done with the **copycore** command.
4. After this, rootvg's primary paging space, `/dev/hd6`, will be activated.

The second part of this phase is shown in Figure 4-8.

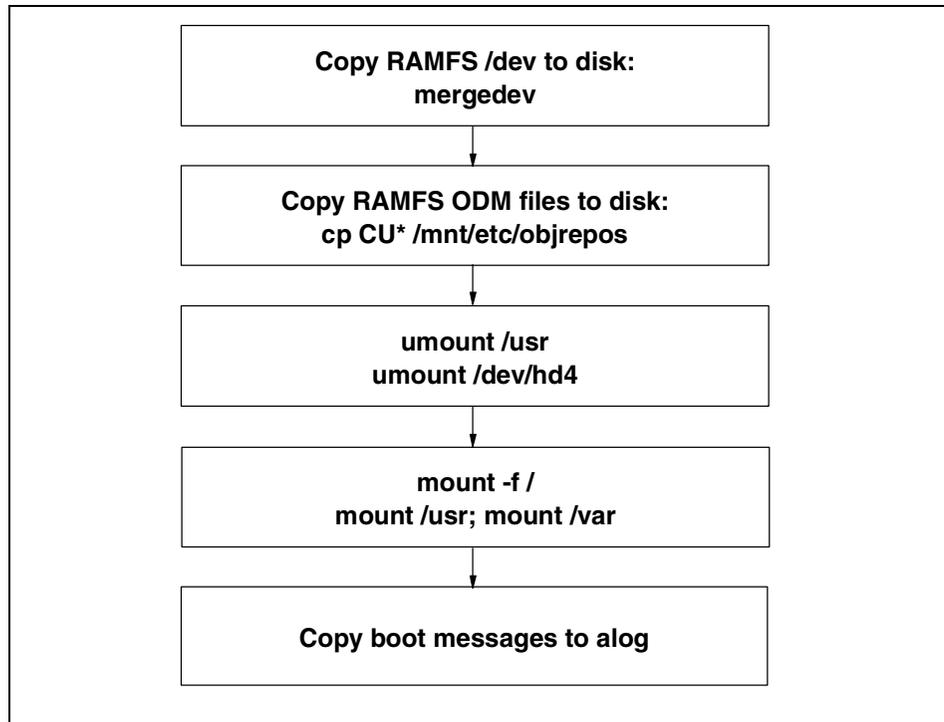


Figure 4-8 Boot phase 2, part two

5. Next, the copy of rootvg's RAMFS ODM and /dev directories will occur (**mergedev**). This is possible because the temporary mount point, /mnt, is used for the mounted root file system.
6. Next, the /usr and /var from the RAMFS are unmounted.
7. Finally, the root file system from rootvg (disk) is mounted over the root file system from the RAMFS. The mount points for the rootvg file systems become available. Now the /var and /usr file systems from the rootvg can be mounted again on their ordinary mount points.
8. There is no console available at this stage, so all boot messages will be copied to the error log.

As mentioned, there are a lot of different problem possibilities in this phase of the boot. The following section discusses how to correct some of them.

4.4.1 LED 551, 555, or 557

There can be several reasons for a system to halt with LED codes 551, 555, or 557. For example:

- ▶ A damaged file system
- ▶ A damaged Journaled File System (JFS) log device
- ▶ A failing disk in the machine that is a member of the rootvg

To diagnose and fix these problems, you will need to boot from a bootable media, access the maintenance menus, choose **Access a Volume Group and start a shell before mounting file systems**, and then do one or all of the following actions.

To ensure file system integrity, run the **fsck** command to fix any file systems that may be corrupted:

```
# fsck -y /dev/hd1
# fsck -y /dev/hd2
# fsck -y /dev/hd3
# fsck -y /dev/hd4
# fsck -y /dev/hd9var
```

To correct the function of the log device, run **logform** on /dev/hd8 to recreate the log device:

```
# /usr/sbin/logform /dev/hd8
```

Or, if the BLV is corrupted, recreate the BLV and update the bootlist:

```
# bosboot -a -d /dev/hdisk0
# bootlist -m normal hdisk0
```

4.4.2 LED 552, 554, and 556

A LED code of 552, 554, or 556 during a standard disk-based boot indicates that a failure occurred during the varyon of the rootvg volume group. This can be caused by:

- ▶ A damaged file system
- ▶ A damaged Journaled File System log device
- ▶ A damaged IPL-device record or a damaged IPL-device magic number (the magic number indicates the device type)
- ▶ A damaged copy of the Object Data Manager (ODM) database on the boot logical volume
- ▶ A hard disk in the inactive state in the root volume group

► A damaged superblock

To diagnose and fix the problem, you will need to boot from the installation media and navigate the menus to access the volume group and start a shell before mounting the file systems.

If **fsck** indicates that block 8 could not be read when used, as shown in 4.4.1, “LED 551, 555, or 557” on page 66, the file system is probably unrecoverable. The easiest way to fix an unrecoverable file system is to recreate it. This involves deleting it from the system and restoring it from a backup. Note that `/dev/hd4` (root file system) cannot be recreated. If `/dev/hd4` is unrecoverable, you must reinstall AIX.

Another possible reason for these error codes is a corrupted superblock. If you boot in maintenance mode and get error messages, such as not an AIX file system, or not a recognized file system type, it is probably due to a corrupted superblock in the file system.

Each file system has two superblocks, one in logical block 1 and a copy in logical block 31. To copy the superblock from block 31 to block 1 for the root file system, issue the following command:

```
# dd count=1 bs=4k skip=31 seek=1 if=/dev/hd4 of=/dev/hd4
```

To correct the function of the log device, run **logform** on `/dev/hd8` to recreate the log device:

```
# /usr/sbin/logform /dev/hd8
```

A corrupted ODM in the BLV is also a possible cause for these LED codes. To create a correct one, run the following commands that remove the system's configuration and save it to a backup directory:

```
# /usr/sbin/mount /dev/hd4 /mnt
# /usr/sbin/mount /dev/hd2 /usr
# /usr/bin/mkdir /mnt/etc/objrepos/bak
# /usr/bin/cp /mnt/etc/objrepos/Cu* /mnt/etc/objrepos/bak
# /usr/bin/cp /etc/objrepos/Cu* /mnt/etc/objrepos
# /usr/sbin/umount all
# exit
```

After this, you must copy this new version of the ODM in the RAMFS to the BLV. This is done with the **savebase** command. Before that, make sure you place it on the disk used for normal boot by executing:

```
# lslv -m hd5
```

Save the clean ODM database to the boot logical volume. For example:

```
# savebase -d /dev/hdisk0
```

Finally, recreate the BLV and reboot the system. For example:

```
# bosboot -ad /dev/hdisk0
# shutdown -Fr
```

4.4.3 LED 518

The 518 LED code has an unclear definition in the *Message Guide and References* (which is part of the AIX Version 4 base documentation), which says:

Display Value 518

Remote mount of the / (root) and /usr file systems during network boot did not complete successfully.

This is not the entire problem. If the system runs into problems while mounting the /usr from disk (locally, not network mount) you will get the same error. Fix this problem the same way as any other rootvg file system corruption is fixed.

4.4.4 The alog command

Up until this stage, the system has not yet configured the console, so there is no stdout defined for the boot processes. At this stage, the **alog** command comes to good use.

The **alog** command can maintain and manage logs. All boot information is sent through the **alog**. To look at the boot messages, use the following command:

```
# alog -ot boot
***** no stderr *****
-----
Time: 12          LEDs: 0x538
invoking top level program -- "/usr/lib/methods/definetc > /dev/null
2>&1;opt=`/u
sr/sbin/lssattr -E -l inet0 -a bootup_option -F value`
      if [ $opt = "no" ];then nf=/etc/rc.net
      else nf=/etc/rc.bsdnet
      fi;$nf -2;x=$?;test $x -ne 0&&echo $nf failed. Check for invalid
command
s >&2;exit $x"
Time: 21          LEDs: 0x539
return code = 0
***** no stdout *****
```

At this point, the `bootup_option` is checked to determine if a BSD-style configuration of TCP/IP services is to be used, or if the default of ODM supported configuration should be used. During this stage, the LED codes 538 and 539 are shown as shown in the preceding example.

4.5 Boot phase 3

So far, the following boot tasks have been accomplished:

- ▶ Hardware configuration performed during BIST and/or POST
- ▶ The load of BLV
- ▶ Phase 1 where base devices are configured to prepare the system for activating the rootvg
- ▶ Phase 2 where rootvg is activated

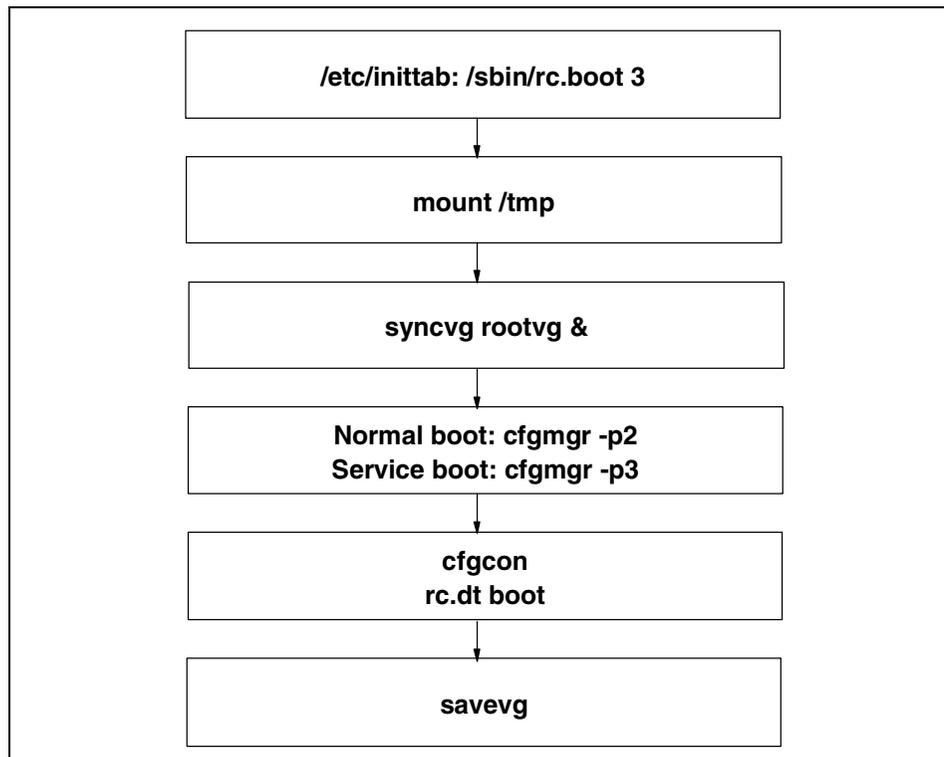


Figure 4-9 Boot phase 3

Finally, phase 3, initiated by the init process loaded from rootvg, is shown in Figure 4-9.

1. Phase 3 is started in `/etc/inittab`.
2. First it will mount `/tmp`.

3. After this, the rootvg will be synchronized. This can take some time. That is why the **syncvg rootvg** is executed as a background process. At this stage, the LED code 553 is shown.
4. At this stage, the **cfgmgr -p2** for normal boot and the **cfgmgr -p3** for service mode is also executed. **cfgmgr** reads the Config_rules file from ODM and checks for devices with phase=2 or phase=3.
5. Next, the console will be configured. LED codes shown when configuring the console are shown in the Note box below.
6. Finally, the synchronization of the ODM in the BLV with the ODM from the / (root) file system is done by the **savebase** command.

When **cfgcons** is called, different LED codes are shown depending on which device is configured.

Note: The cfgcon LED codes:

- ▶ c31: Console not yet configured. Provides instructions to select console.
- ▶ c32: Console is a lft terminal.
- ▶ c33: Console is a tty.
- ▶ c34: Console is a file on the disk.

4.5.1 The /etc/inittab file

The /etc/inittab file supplies configuration scripts to the init process. In Figure 4-10 on page 71, the highlighted line is the rc.boot with parameter 3 executed. The **init** command's role is a general process dispatcher. Typical processes dispatched by the **init** command are daemons and the shell.

```

(C) COPYRIGHT International Business Machines Corp. 1989, 1993
: All Rights Reserved
: Licensed Materials - Property of IBM
:
: US Government Users Restricted Rights - Use, duplication or
: disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
:
: Note - initdefault and sysinit should be the first and second entry.
:
init:2:initdefault:
brc::sysinit:/sbin/rc.boot 3 >/dev/console 2>&1 # Phase 3 of system boot
powerfail::powerfail:/etc/rc.powerfail 2>&1 | alog -tboot > /dev/console # Power
Failure Detection
rc:2:wait:/etc/rc 2>&1 | alog -tboot > /dev/console # Multi-User checks
fbcheck:2:wait:/usr/sbin/fbcheck 2>&1 | alog -tboot > /dev/console # run /etc/fi
rstboot
srcmstr:2:respawn:/usr/sbin/srcmstr # System Resource Controller
rctcpip:2:wait:/etc/rc.tcpip > /dev/console 2>&1 # Start TCP/IP daemons
rcnfs:2:wait:/etc/rc.nfs > /dev/console 2>&1 # Start NFS Daemons
cron:2:respawn:/usr/sbin/cron
pio-be:2:wait:/usr/lib/lpd/pio/etc/piobinit >/dev/null 2>&1 # pb cleanup
qdaemon:2:wait:/usr/bin/startsrc -sqdaemon
writesrv:2:wait:/usr/bin/startsrc -swritesrv

```

Figure 4-10 Example of rc.boot 3 in /etc/inittab

The /etc/inittab file is composed of entries that are position dependent and have the following format:

```
Identifier:RunLevel:Action:Command
```

The first line in /etc/inittab (initdefault) defines what runlevel is to be considered as a default run level. In this example the runlevel is 2, which means a normal multi-user boot. If this line is missing, you are prompted at boot to define the run level. You can also use the **telinit** command to direct a particular run level.

The **telinit** command, which is linked to the **init** command, directs the actions of the **init** command. The **telinit** command takes a one-character argument and signals the **init** command by way of the **kill** subroutine to perform the appropriate action. The **telinit** command sets the system at a specific run level.

When the **init** command finds a record in the /etc/inittab file with a value of a, b, or c in the run level field, it starts the process. However, it does not kill any processes at the current run level; processes with a value of a, b, or c in the run level field are started in addition to the processes already running at the current system level.

Table 4-2 The `telinit` command examples

Command	Description
<code>telinit q</code>	Requests the <code>init</code> command to reexamine the <code>/etc/inittab</code>
<code>telinit s</code>	Requests the <code>init</code> command to enter maintenance mode
<code>telinit 2</code>	Requests process to be run in a multiuser environment

The `rc.boot` line is to be executed on all run levels (this equals runlevel 0123456789). The action defined, `sysinit`, has to finish before continuing with the next line in `/etc/inittab`. From `rc.boot 3`, among a lot of other things, the `rootvg` is synchronized, the `mirrord` is started, and the `/tmp` directory mounted.

4.5.2 LED 553

A LED code of 553 is caused when the `/etc/inittab` cannot be read. To recover from a LED 553, check `/dev/hd3` and `/dev/hd4` for space problems and erase files if necessary. Check the `/etc/inittab` file for corruption and fix it if necessary. Typical examples of syntax errors found in `/etc/inittab` that are seen at the support centers are incorrectly defined entries in the file. When editing `/etc/inittab`, the `inittab` commands should be issued. For example:

- ▶ `mkinitab`
- ▶ `chinitab`

It is helpful to remember that `/etc/inittab` is very sensitive to even the smallest syntax error. A misplaced dot can halt the system boot.

4.5.3 LED c31

LED c31 is not really an error code, but the system is waiting for input from you on the keyboard. This is usually encountered when booting from CD-ROM or `mksysb` tape. This is normally the dialog to select the system console.

If a console has not been configured, the system pauses with this value in the three-digit LED and then displays instructions for choosing a console on the native display. System initialization and configuration continue after you choose a console. If you do not see these instructions on your primary attached console, then one of the following causes could be the reason:

- ▶ An incorrect console configuration
- ▶ A missing device driver
- ▶ A hardware problem with the console device or machine

4.6 pSeries servers

Although new pSeries servers have the same basic boot steps, the LEDs display a little differently. Please consult your specific pSeries Service Guide for more information. Following is an example of IBM @server pSeries 690 boot phases:

The IPL process starts when AC power is connected to the system. The IPL process has the phases discussed in the following sections.

4.6.1 Phase 1: Service processor initialization

Phase 1 starts when AC power is connected to the system and ends when 0K is displayed in the media subsystem operator panel. 8xxx checkpoints are displayed during this phase. Several 9xxx codes may also be displayed. Service processor menus are available at the end of this phase by striking any key on the console keyboard.

4.6.2 Phase 2: Hardware initialization by the service processor

Phase 2 starts when system power-on is initiated by pressing the power-on button on the media subsystem operator panel. 9xxx checkpoints are displayed during this time. 91FF, the last checkpoint in this phase, indicates that the transition to phase 3 is taking place.

4.6.3 Phase 3: System firmware initialization

On a full system partition, at phase 3, a system processor takes over control and continues initializing partition resources. During this phase, checkpoints in the form Exxx are displayed. E105, the last checkpoint in this phase, indicates that control is being passed to the AIX boot program. On a partitioned system, there is a global system-wide initialization phase 3, during which a system processor continues the initialization process. Checkpoints in this phase are of the form Exxx. This global phase 3 ends with a LPAR. . . on the operator panel. As a logical partition begins a partition-initialization phase 3, one of the system processors assigned to that partition continues initialization of resources assigned to that partition. Checkpoints in this phase are also of the form Exxx. This partition phase 3 ends with an E105 displayed on the partition's virtual operator panel on the HMC, indicating control has been passed to that logical partition's AIX boot program. For both the global and partition phase 3, location codes may also be displayed on the physical operator panel and the partition's virtual terminal, respectively.

4.6.4 Phase 4: AIX boot

When AIX starts to boot, checkpoints in the form 0xxx and 2xxx are displayed. This phase ends when the AIX login prompt displays on the AIX console.

4.7 Summary

The next sections provide short summaries of the boot phases and some common LED codes.

4.7.1 Boot phases

BIST and POST are used to test hardware and to find a successful hardware path to a BLV.

Boot phase 1 (init rc.boot 1) is used to configure base devices.

Boot phase 2 (init rc.boot 2) is used to activate the rootvg.

Boot phase 3 (init /sbin/rc.boot 3) is used to configure the rest of the devices.

4.7.2 LED codes

LED codes shown during POST on a MCA system are listed in Table 4-3.

Table 4-3 MCA POST LED

LED	Reason/action
100–195	Hardware problem during BIST.
200	Key mode switch in secure position.
201	1. If LED 299 passed, recreate the BLV. 2. If LED 299 has not passed, POST encountered a hardware error.
221 721 221–229 223–229 225–229 233–235	The bootlist in NVRAM is incorrect (boot from media and change the bootlist), or the bootlist device has no boot image (boot from media and recreate the BLV), or the bootlist device is unavailable (check for hardware errors).

LED codes shown during boot phase 2 are shown in Table 4-4 on page 75.

Table 4-4 Boot phase 2 LED codes

LED	Reason/action
551 555 557	1. Corrupted file system (<code>fsck -y device</code>). 2. Corrupted jfslog (<code>/usr/sbin/logform /dev/hd8</code>). 3. Corrupted BLV (<code>bosboot -ad device</code>).
552 554 556	The <code>ipl_varyon</code> failed. Except for the reason mentioned above (551, 555, or 557): 1. Corrupted ODM (backup ODM, recreate with <code>savebase</code>). 2. Superblock dirty (Copy in superblock from block 31).
518	<code>/usr</code> cannot be mounted. 1. If <code>/usr</code> should be mounted over the network (check for network problem). 2. If <code>/usr</code> is to be mounted locally (fix the file system).

LED codes shown during boot phase 3 are shown in Table 4-5.

Table 4-5 Boot phase 3 LED codes

LED	Reason/action
553	Syntax error in <code>/etc/inittab</code> .
c31	Define the console.

4.8 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

- In order to log important events during a system boot process and view boot information, which of the following AIX commands must be performed?
 - `alog`
 - `errpt`
 - `logger`
 - `bootlist`

2. The rootvg volume group has just been mirrored. In order to make hdisk1 bootable, which of the following procedures should be performed?
 - A. Update bootlist and reboot.
 - B. Add hdisk1 to bootlist and reboot.
 - C. Add hdisk1 to bosboot and reboot.
 - D. Run **bosboot** and update the bootlist.

3. After a server fails to boot in normal mode with a LED 551 displayed, debug the problem. LED 551 indicates the root volume group is being varied on. Boot the system in service mode using bootable media and select the maintenance shell. Run the **fsck** command to clean up the root volume group:

```
fsck -y /dev/hd1
fsck -y /dev/hd2
fsck -y /dev/hd3
fsck -y /dev/hd4
fsck -y /dev/hd9var
```

Type exit to continue the process of accessing the rootvg and mount the file systems. Determine the boot disk by issuing the **lslv -m hd5** command. Which of the following procedures should be performed prior to rebooting the system in normal mode?

- A. Edit /etc/inittab and correct any erroneous entries.
 - B. Delete unnecessary files in the /dev/hd3 (/tmp) file system.
 - C. Clear the log device with the **/usr/sbin/logform /dev/hd8** command.
 - D. Create a new boot image with the **bosboot -a -d/dev/hdisk#** command.
4. After a RS/6000 fails to boot in normal mode with an LED 552 displayed, the system administrator begins to debug the problem. LED 552 indicates the *varyon* of the root volume group failed. The system administrator boots the system in service mode using bootable media and selects the maintenance shell. The system administrator then runs the **fsck** command to clean up the root volume group:

```
fsck -y /dev/hd1
fsck -y /dev/hd2
fsck -y /dev/hd3
fsck -y /dev/hd4
fsck -y /dev/hd9var
```

The system administrator receives a message that 'block 8' cannot be read. **fsck** reports that the file system is not an AIX Version 4 file system or a recognized type. The system administrator receives a message that the file system has an unknown log record type.

The system administrator then attempts to recover the superblock by typing:

```
dd if=/dev/hdn of=/dev/hdn count=1 bs=4k skip=31 seek=1
```

Which of the following procedures should be performed next?

- A. Run **fsck -y dev/hd8**.
 - B. Rebuild the bootlist with the command **bootlist -m normal hdiskn**.
 - C. Clear the log device with the command **/usr/sbin/logform /dev/hd8**.
 - D. Create a new boot image with the command **bosboot -a -d /dev/hdiskn**.
5. Given the information provided in the following example, which of the following procedures will successfully resolve the problem before rebooting in normal mode?

After an AIX server fails to boot in normal mode with an LED 553 displayed, begin to debug the problem. LED 553 indicates that IPL phase 1 is complete. Boot the system in service mode using bootable media and select the maintenance shell, then gather the following information:

```
# df /dev/hd4
Filesystem 512-blocks   Free %Used    Iused %Iused  Mounted on
/dev/hd4    16384      6976   58%     1023   25%      /
# df /dev/hd3
Filesystem 512-blocks   Free %Used    Iused %Iused  Mounted on
/dev/hd3    32768    14096   57%      369   10%     /tmp
# df /dev/hd2
Filesystem 512-blocks   Free %Used    Iused %Iused  Mounted on
/dev/hd2    671744   176552   74%    13775   17%     /usr
-----
: @(#)49 1.28.2.7 src/bos/etc/inittab/inittab, cmdoper, bos411,
:9430C411a 7/26 /94 16:27:45
:
: COMPONENT_NAME: CMDOPER
:
: ORIGINS: 3, 27
:
: (C) COPYRIGHT International Business Machines Corp. 1989, 1993
: All Rights Reserved
: Licensed Materials - Property of IBM
:
: US Government Users Restricted Rights - Use, duplication or
: disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
:
: Note - initdefault and sysinit should be the first and second entry.
:
logsymp:2:once:/usr/lib/ras/logsymptom # for system dumps
init:2:initdefault:
```

```

brc::sysinit:/sbin/rc.boot 3 >/dev/console 2>&1 # Phase 3 of system boot
powerfail::powerfail:/etc/rc.powerfail 2>&1 | alog -tboot > /dev/console
#Power Failure Detection
rc:2:wait:/etc/rc 2>&1 | alog -tboot > /dev/console # Multi-User checks
fbcheck:2:wait:/usr/sbin/fbcheck 2>&1 | alog -tboot > /dev/console
#run /etc/firstboot
srcmstr:2:respawn:/usr/sbin/srcmstr # System Resource Controller
rctcpip:2:wait:/etc/rc.tcpip > /dev/console 2>&1 # Start TCP/IP daemons
cron:2:respawn:/usr/sbin/cron
piobe:2:wait:/usr/lib/lpd/pio/etc/pioint >/dev/null 2>&1 # pb cleanup
qdaemon:2:wait:/usr/bin/startsrc -sqdaemon
writesrv:2:wait:/usr/bin/startsrc -swritesrv
uprintfd:2:respawn:/usr/sbin/uprintfd
cons:0123456789:respawn:/usr/sbin/getty /dev/console
infod:2:once:startsrc -s infod
-----
# ls -l /etc/environment
-rw-rw-r-- 1 root system 1477 Jun 15 14:39 /etc/environment
# ls -l /etc/profile
-r-xr-xr-x 1 bin bin 1397 Jan 4 1995 /etc/profile
# ls -l /.profile
-rwxr---- 1 root system 265 Jan 4 1995 /.profile
# ls -l /bin/bsh
-r-xr-xr-x 2 bin bin 341108 Aug 24 10:51 /bin/bsh
# ls -l /bin/sh
-r-xr-xr-x 4 bin bin 230116 Aug 14 22:57 /bin/sh

```

- A. Edit /etc/inittab and correct any erroneous entries.
 - B. Delete unnecessary files in the / (root) file system.
 - C. Change the permissions on /etc/environment so the root user can run the file.
 - D. Run **fsck** on the root VG and recreate the log device with the **/usr/sbin/logform /dev/hd8** command.
6. A workstation has a bootlist that is set for network boot. An attempt is made to access SMS menus to change the bootlist to the local disk. A discovery is made that someone has set an SMS supervisory password that is not recorded. Which of the following procedures will enable a system administrator to gain access to the SMS menu?
- A. Reinstall the system firmware.
 - B. Boot from AIX installation media, then reinstall SMS.
 - C. Remove the battery from the system for at least one minute and then reboot.
 - D. Call IBM HW and ask for the over-write password based on system serial number.

4.8.1 Answers

The following are the preferred answers to the questions provided in this section.

1. A
2. D
3. D
4. C
5. A
6. C

4.9 Exercises

The following exercises provide sample topics for self study. They will help ensure comprehension of this chapter.

Do not perform these exercises on an existing file system or on a production system:

1. Create a file system for this exercise and copy in some files to the file system. Then destroy the first superblock. This can be done by copying 4 KB from `/dev/zero` to block one on your logical volume, for example:

```
# dd count=1 bs=4k seek=1 if=/dev/zero of=/dev/thomasclv
```

Try to mount the file system and run **fsck** on the file system to determine the problem. Finally, fix the problem as described in this chapter.

2. While still on your test system, with verified **mksysb** at hand, make a backup of `/etc/inittab`. Remove the first line (first uncommented line, that is, the `sysinit` stanza) and try to reboot. You are, at reboot, prompted for what?

After the boot has finished, edit the `/etc/inittab` and change a dot to a comma or a colon to semicolon on a line with `action=wait`. What happens? What is the LED code displayed? What do you have to do to fix this?



Disks and file systems

In AIX, the logical volume manager (LVM) provides a hierarchy of structures to manage fixed disk storage. Each individual fixed-disk drive is called a physical volume (PV) and has a logical name, such as `hdisk0`, and a corresponding device special file name, such as `/dev/hdisk0`. A RAID array is named `/dev/hdiskX`, but is a special case of a collection of disks. Every physical volume in use belongs to a volume group (VG). All of the physical volumes in a volume group are divided into physical partitions (PPs) of the same size.

Within each volume group, one or more logical volumes (LVs) is defined. Logical volumes are groups of information located on physical volumes. Data on logical volumes appears to be contiguous to the user but can be discontinuous on the physical volume. This allows file systems, paging space, and other logical volumes to be resized or relocated, span multiple physical volumes, and have their contents mirrored for greater flexibility and availability in the storage of data.

Each logical volume consists of one or more logical partitions (LPs). Each logical partition corresponds to at least one physical partition. If mirroring is specified for the logical volume, additional physical partitions are allocated to store the additional copies of each logical partition. Although the logical partitions are numbered consecutively, the underlying physical partitions are not necessarily consecutive or contiguous.

Logical volumes can serve a number of system purposes, such as paging, but each logical volume that holds an ordinary system, user data, or programs contains a single journaled file system (JFS). Each JFS consists of a pool of

page-size (4 KB) blocks. When data is to be written to a file, one or more additional blocks is allocated to that file. These blocks may or may not be contiguous with one another or with other blocks previously allocated to the file.

After installation, a system has one volume group (the rootvg volume group) consisting of a base set of logical volumes required to start the system and any others you specify to the installation script. Any other physical volumes you have connected to the system can be added to a volume group (using the **extendvg** command). You can add the physical volume either to the rootvg volume group or to another volume group (defined by using the **mkvg** command). Logical volumes can be tailored using commands or the SMIT interface.

5.1 Volume groups

If the physical hardware is the foundation for a system, then the volume groups are the cornerstones. Volume group management is one of the most important considerations in managing your system. It is from here that the administrator or support professional creates the future path of the system. Not planning how to manage or create volume groups in the beginning could lead to complications in the future.

5.1.1 Creating a volume group

Creating a volume group is the first step in setting up a system. The command used to create a volume group is `mkvg`. However, it is easier to create a volume group using SMIT.

Some considerations when creating a new volume group are as follows:

- ▶ The physical volume is checked to verify that it is not already in another volume group. If the system believes the physical volume belongs to a volume group that is varied on, it exits. But, if the system detects a description area from a volume group that is not varied on, it prompts you for confirmation in continuing with the command. The previous contents of the physical volume are lost, so you must be cautious when using the override function.
- ▶ Before you attempt to add a new volume group, check the root file system (with the `df` command) to ensure that 1 to 2 MB of free space exists. If you run the `mkvg` command on a full rootvg file system, the `mkvg` command fails when it tries to varyon the newly created volume group. If this occurs, add or free 2 MB of space in the `/etc` directory for file updates.
- ▶ To use the `mkvg` command, you must either have root user authority or be a member of the system group.
- ▶ For disks greater than 4 GB, make sure that the physical partition size (`-s`) is set to a large enough value so that the 1016 physical partitions per PV limit is not violated. With the default value (4 MB), along with the default maximum number of physical partitions (1016), only a disk up to 4 GB can be accommodated. Or, with AIX Version 4.3.2, use a large enough factor value (`-t`) such that $\text{factor} \times 1016$ is greater than the number of partitions that would need to be created with given partition size and disk size. For example, a partition size of at least 16 MB would be needed to create a volume group with a 10 GB disk. Or, with a factor size of 2, a smaller partition size of 8 MB can be used. However, this limits the total number of disks that can be added to the volume group. If a factor value is used, a maximum of $32/\text{factor}$ disks can be included in the volume group.

- ▶ Whenever a volume group is created, the operating system automatically does a varyon. However, if you create a volume group with the `-c` flag (for use in HACMP clusters), the system will not auto varyon the volume group at the end of the Concurrent Capable volume group creation. Instead, the `mkvg` command notifies you to manually `varyonvg` the volume group in either non-concurrent or concurrent mode. Note that, to varyon the volume group with concurrent mode, you have to configure HACMP.

The following examples show how these steps are performed.

Check for existing volume groups using the `lsvg` command:

```
# lsvg
rootvg
volg01
volg02
```

Check for available physical volumes using the `lspv` command:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg02
hdisk3      000bc6fdbff972f4    None
```

The PV `hdisk3` is available for a new user VG. In the following example, the volume group, `volg03`, will be created on the physical volume `hdisk3`.

The `chdev -l hdisk3 -a pv=yes` command can be used to assign a PVID to a disk that does not have one. It has no effect if the disk is already a physical volume, as shown in the example.

Check the attributes of the disk as follows:

```
# lsattr -E -l hdisk3
pvid          000bc6fdbff972f40000000000000000 Physical volume identifier False
queue_depth 3                               Queue DEPTH                False
size_in_mb   9100                          Size in Megabytes          False
```

Notice the size of `hdisk3` in megabytes. It has a capacity of 9100 MB.

SMIT can also be used (Figure 5-1 on page 85) as follows:

```
# smitty vg
```

```

                                Volume Groups

Move cursor to desired item and press Enter.

List All Volume Groups
Add a Volume Group
Set Characteristics of a Volume Group
List Contents of a Volume Group
Remove a Volume Group
Activate a Volume Group
Deactivate a Volume Group
Import a Volume Group
Export a Volume Group
Mirror a Volume Group
Unmirror a Volume Group
Synchronize LVM Mirrors
Back Up a Volume Group
Remake a Volume Group
List Files in a Volume Group Backup
Restore Files in a Volume Group Backup

F1=Help          F2=Refresh      F3=Cancel      F8=Image
F9=Shell        F10=Exit       Enter=Do

```

Figure 5-1 The Volume Groups screen

Select the **Add a Volume Group** option as shown in Figure 5-2.

```

                                Add a Volume Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
VOLUME GROUP name                [volg03]
Physical partition SIZE in megabytes 16 +
* PHYSICAL VOLUME names           [hdisk3] +
Activate volume group AUTOMATICALLY  yes +
  at system restart?
Volume Group MAJOR NUMBER          [] ++#
Create VG Concurrent Capable?      no +
Auto-varyon in Concurrent Mode?     no +

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset        F6=Command     F7=Edit        F8=Image
F9=Shell        F10=Exit       Enter=Do

```

Figure 5-2 The Add a Volume Group selection screen

Enter the volume group name (volg03) and the hdisk number (hdisk3). The next step is to select the physical partition size. The physical partition size is selected based on the default maximum size of the disk being used. Figure 5-2 shows a selected physical partition size of 16 MB.

```

                                Add a Volume Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
VOLUME GROUP name                [volg03]
Physical partition SIZE in megabytes 16      +
* PHYSICAL VOLUME names          [hdisk3]    +
Activate volume group AUTOMATICALLY  yes     +
  at system restart?
Volume Group MAJOR NUMBER          []      +##
                                     +
                                     +
                                ARE YOU SURE?
Continuing may delete information you may want
to keep. This is your last chance to stop
before continuing.
  Press Enter to continue.
  Press Cancel to return to the application.
  █
F1 F1=Help           F2=Refresh       F3=Cancel
F5 F8=Image         F10=Exit          Enter=Do
F9

```

Figure 5-3 The Add a Volume Group warning screen

The system prompts for confirmation on the creation of the volume disk. Press **Enter** to continue, and a screen similar to Figure 5-4 on page 87 will appear.

```
COMMAND STATUS

Command: OK          stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

volg03
█

F1=Help          F2=Refresh          F3=Cancel          Esc+6=Command
Esc+8=Image      Esc+9=Shell         Esc+0=Exit         /=Find
n=Find Next
```

Figure 5-4 The Add a Volume Group Command Status screen

The Command Status screen indicates successful completion of the volume group creation. Press F10 to exit to the system prompt to check the volume group with the commands shown in the following examples.

Note: The volume group could also have been created using the following command:

```
# mkgv -f -y volg03 -s16 hdisk3
```

Check the physical volume and notice that hdisk3 has been assigned to the volume group volg03.

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg02
hdisk3      000bc6fdbff972f4    volg03
```

5.1.2 Migrating data to another disk in the same volume group

When migrating data from one physical volume to another in the same volume group, use the **migratepv** command. There are a few considerations before doing this:

- ▶ The **migratepv** command cannot migrate a striped logical volume. To do this, use the **cp1v** and **rm1v** commands.
- ▶ The new physical volume must be part of the same volume group.
- ▶ Ensure that all file systems on the logical volume are unmounted.
- ▶ Only the root user has authority to use the **migratepv** command.

The following commands show examples of migrating data to other disks in the same volume group.

Check which physical volumes are in the system and to which volume groups they are assigned by using the **lspv** command:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg01
hdisk3      000bc6fdbff972f4    volg02
```

Check the **hdisk1** and **hdisk2** that belong to the **volg01** volume group:

```
# lspv hdisk1
PHYSICAL VOLUME:    hdisk1                VOLUME GROUP:    volg01
PV IDENTIFIER:      000bc6fdbff75ee2    VG IDENTIFIER     000bc6fdc07cc316
PV STATE:           active
STALE PARTITIONS:   0                    ALLOCATABLE:     yes
PP SIZE:            16 megabyte(s)    LOGICAL VOLUMES: 2
TOTAL PPs:          542 (8672 megabytes)  VG DESCRIPTORS:  2
FREE PPs:           482 (7712 megabytes)
USED PPs:           60 (960 megabytes)
FREE DISTRIBUTION:  109..48..108..108..109
USED DISTRIBUTION:  00..60..00..00..00
```

```
# lspv hdisk2
PHYSICAL VOLUME:    hdisk2                VOLUME GROUP:    volg01
PV IDENTIFIER:      000bc6fdbff92812    VG IDENTIFIER     000bc6fdc07cc316
PV STATE:           active
STALE PARTITIONS:   0                    ALLOCATABLE:     yes
PP SIZE:            16 megabyte(s)    LOGICAL VOLUMES: 1
TOTAL PPs:          542 (8672 megabytes)  VG DESCRIPTORS:  1
FREE PPs:           541 (8656 megabytes)
USED PPs:           1 (16 megabytes)
FREE DISTRIBUTION:  109..107..108..108..109
```

```
USED DISTRIBUTION: 00..01..00..00..00
```

The information on `hdisk1` in `volg01` will be moved to `hdisk2` in the same `volg01` volume group. Use the `migratepv` command as follows:

```
# migratepv hdisk1 hdisk2
```

After migrating the physical volume, check the physical volumes:

```
# lspv hdisk1
PHYSICAL VOLUME:   hdisk1                VOLUME GROUP:   volg01
PV IDENTIFIER:    000bc6fdbff75ee2          VG IDENTIFIER    000bc6fdc07cc316
PV STATE:         active
STALE PARTITIONS: 0                    ALLOCATABLE:    yes
PP SIZE:          16 megabyte(s)         LOGICAL VOLUMES: 0
TOTAL PPs:        542 (8672 megabytes)    VG DESCRIPTORS: 2
FREE PPs:         542 (8672 megabytes)
USED PPs:         0 (0 megabytes)
FREE DISTRIBUTION: 109..108..108..108..109
USED DISTRIBUTION: 00..00..00..00..00
```

```
# lspv hdisk2
PHYSICAL VOLUME:   hdisk2                VOLUME GROUP:   volg01
PV IDENTIFIER:    000bc6fdbff92812          VG IDENTIFIER    000bc6fdc07cc316
PV STATE:         active
STALE PARTITIONS: 0                    ALLOCATABLE:    yes
PP SIZE:          16 megabyte(s)         LOGICAL VOLUMES: 3
TOTAL PPs:        542 (8672 megabytes)    VG DESCRIPTORS: 1
FREE PPs:         481 (7696 megabytes)
USED PPs:         61 (976 megabytes)
FREE DISTRIBUTION: 109..47..108..108..109
USED DISTRIBUTION: 00..61..00..00..00
```

The migration has successfully completed. The `hdisk1` can now be removed from the volume group.

5.1.3 Importing a volume group from another system

When a disk or set of disks containing data that must be moved to another system, it is important to take the right steps. Generally once all the users of the data have been identified, any applications using the disks have been stopped, and the file systems unmounted, the following steps should be taken:

- ▶ Run the `varyoffvg` command to take the volume group offline.
- ▶ Run the `exportvg` command to remove the volume group from the system.
- ▶ Run the `rmdev -l hdiskn -d` command to delete the physical volume definition (this command does not remove any data from the physical fixed disk storage).

If a new disk is added into an existing environment, and this disk has data on it that needs to be accessed, the new disk will need to be imported. However, this is not the only requirement. The new disk needs to be configured to the system first using the **cfgmgr** command. The **cfgmgr** command configures devices and optionally installs device software into the system.

Note: Before importing the volume group, you should verify that your file systems are named uniquely. The AIX LVM will change the logical volume names to uniquely assigned name to import this volume group successfully. However, you may have difficulty mounting non-uniquely named file systems without editing the `/etc/filesystems` file to correct conflicts.

Once the disk has been set up in the new environment, the data needs to be imported using the **importvg** command. The **importvg** command imports a new volume group definition from a set of physical volumes.

Once the volume group has been imported, the **varyonvg** command may need to be run depending on the AIX version (AIX Version 4.3 or later automatically does the **varyonvg** option). The **varyonvg** command activates the volume group specified by the `VolumeGroup` parameter and all associated logical volumes.

Before mounting the file system, it will need to be checked using the **fsck** command. This may or may not give errors, but it should be run to ensure file system integrity. The **fsck** command checks file system consistency and interactively repairs the file system.

The following is an example of this scenario.

A disk has been added to the environment, and the user wants to have the data on that disk available for use. These are the steps to follow:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg01
hdisk3      000bc6fdbff972f4    volg02
```

Run the configuration manager and detect the new disk:

```
# cfgmgr
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg01
hdisk3      000bc6fdbff972f4    volg02
hdisk4      000bc6ddc0c0973a    None
```

Notice that `hdisk4` is not assigned to a volume group:

```
# lspv hdisk4
0516-320 : Physical volume 000bc6ddc0c0973a is not assigned to
          a volume group.
```

Import the `hdisk` into the volume group `volg03`:

```
# importvg -y volg03 hdisk4
0516-530 sync1vodm: Logical volume name loglv00 changed to loglv02.
volg03
```

The volume group on `hdisk4` has been imported into the volume group `volg03`, and it has also renamed the `jfslog`. Vary the volume group on:

```
# varyonvg volg03
```

List the volume group to check what the logical volume is:

```
# lsvg -l volg03
volg03:
LV NAME          TYPE      LPs  PPp  PVs  LV STATE      MOUNT POINT
loglv02          jfslog    1    1    1    closed/syncd  N/A
lv02             jfs       1    1    1    closed/syncd  /test
```

Run the `fsck` to check the file system on the logical volume `lv02`:

```
# fsck /test
** Checking /dev/lv02 (/test)
** Phase 0 - Check Log
log redo processing for /dev/lv02
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
559 files 8944 blocks 23824 free
```

Mount the file system `/test` (providing there is no `/test` file system already defined):

```
# mount /test
```

Check that the file system is mounted:

```
# mount
node          mounted          mounted over    vfs      date            options
-----
...
              /dev/hd3         /tmp            jfs      Jun 11 16:41    rw,log=/dev/hd8
              /dev/hd1         /home          jfs      Jun 11 16:42    rw,log=/dev/hd8
```

```

        /dev/lvo101      /filsys01      jfs      Jun 15 13:44
rw,log=/dev/loglv00
        /dev/lvo102      /filsys02      jfs      Jun 15 13:44
rw,log=/dev/loglv00
        /dev/lvo103      /filsys03      jfs      Jun 15 14:18
rw,log=/dev/loglv01
        /dev/lv02        /test          jfs      Jun 15 18:03
rw,log=/dev/loglv02

```

The system is now available for the user. If the volume group volg03 is to be integrated into the volume group volg02 use the **cp1v** command, as described in “Using the cplv and rmlv commands” on page 106.

5.1.4 Replacing a disk with another

If a disk in a volume group must be replaced with another disk, the **replacepv** command can be used. The **replacepv** command replaces allocated physical partitions and the data they contain from SourcePhysicalVolume to DestinationPhysicalVolume. The specified source physical volume cannot be the same as DestinationPhysicalVolume.

The allocation of the new physical partitions follows the policies defined for the logical volumes that contain the physical partitions being replaced.

Note: The DestinationPhysicalVolume size must be at least the size of the SourcePhysicalVolume.

The **replacepv** command cannot replace a SourcePhysicalVolume with stale logical volume unless this logical volume has a non-stale mirror.

An example of this is shown as follows.

Check the system to see if any disks are available as follows:

```

# lspv
hdisk0      000bc6fdc3dc07a7      rootvg
hdisk1      000bc6fdbff75ee2      uservg
hdisk2      000bc6fdbff92812      None
hdisk3      000bc6fdbff972f4      None
hdisk4      000bc6fd672864b9      None

```

Replace hdisk1 with hdisk2 as follows:

```

# replacepv hdisk1 hdisk2
0516-1232 replacepv:

```

If this program is terminated before the completion due to a system crash or Ctrl+C, and you want to continue afterwards, look for a /tmp/replacepvXXXXX file in /tmp for the actual PID value. Execute the command **replacepv -R /tmp/replacepv26606** (where 26606 is an example number).

Check the system as follows:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    None
hdisk2      000bc6fdbff92812    uservg
hdisk3      000bc6fdbff972f4    None
hdisk4      000bc6fd672864b9    None
```

During the **replacepv** operation, hdisk1 was removed from the volume group and replaced with hdisk2.

5.1.5 Reducing a volume group

If a disk needs to be removed from a system due to volume group reduction or due to incorrect removal, the **reducevg** command needs to be used. The **reducevg** command removes physical volumes from a volume group.

Note: If there are allocated physical partitions of logical volumes in these physical volumes, you cannot remove the physical volume from the volume group. First, you have to move the allocated physical partitions of logical volumes on these physical volumes to other physical volumes.

The following commands provide an example of this task.

In this example, hdisk2 has been removed without following proper procedure.

Display a list of physical volumes in a system using the **lspv** command. Below is the listing *before* hdisk2 was removed:

```
# lspv
hdisk0      00015f8f9d9d874a    rootvg
hdisk1      00015f8fe7288c7b    None
hdisk2      00015f8fa0bec62a    redbookvg
hdisk3      00015f8fa0bed102    redbookvg
```

When the disk was removed, the server was rebooted. The new configuration shows that hdisk2 is missing:

```
# lspv
hdisk0      00015f8f9d9d874a    rootvg
hdisk1      00015f8fe7288c7b    None
hdisk3      00015f8fa0bed102    redbookvg
```

Running the **lsdev** command shows that the disk has a status of Defined:

```
# lsdev -Cc disk
hdisk0 Available 11-08-00-2,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 11-08-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk2 Defined 27-09-00-8,0 16 Bit LVD SCSI Disk Drive
hdisk3 Available 27-09-00-9,0 16 Bit LVD SCSI Disk Drive
```

Remove the physical volume **hdisk2** from the volume group **redbookvg** using the **reducevg** command:

```
# reducevg redbookvg hdisk2
```

Remove **hdisk2** logically from the system.

```
# rmdev -d1 hdisk2
hdisk2 deleted
```

Check to make sure that **redbookvg** now only resides on one disk only, **hdisk3**:

```
# lsvg -p redbookvg
redbookvg:
PV_NAME          PV STATE          TOTAL PPs   FREE PPs   FREE DISTRIBUTION
hdisk3           active            542         542        109..108..108..108.
.109
```

5.1.6 Reorganizing a volume group

To reorganize a volume group, the **reorgvg** command is used. The **reorgvg** command reorganizes the placement of allocated physical partitions within the volume group according to the allocation characteristics of each logical volume. Use the logical volume parameter to reorganize specific logical volumes; highest priority is given to the first logical volume name in the logical volume parameter list and lowest priority is given to the last logical volume in the parameter list. The volume group must be varied on and must have free partitions before you can use the **reorgvg** command. The following are a few considerations for using the **reorgvg** command:

- ▶ The **reorgvg** command does not reorganize the placement of allocated physical partitions for any striped logical volumes.
- ▶ At least one free physical partition must exist on the specified volume group for the **reorgvg** command to run successfully.
- ▶ To use this command, you must either have root user authority or be a member of the system group.

- ▶ If you enter the **reorgvg** command with the volume group name and no other arguments, it will only reorganize the first logical volume in the volume group. The first logical volume is the one listed by the **lsvg -l VolumeGroupName** command. In AIX 4.2 or higher, whether you enter the **reorgvg** command with the volume group name and no other arguments, the entire volume group is reorganized.

The following commands provide an example of this task.

Select the volume group to be reorganized, in this case, rootvg, and then list the logical volumes in the rootvg volume group:

```
# lsvg -l rootvg
rootvg:
LV NAME          TYPE      LPs  PPs  PVs  LV STATE    MOUNT POINT
hd5              boot      2    2    1    closed/syncd N/A
hd6              paging    72   72   1    open/syncd   N/A
hd8              jfslog    1    1    1    open/syncd   N/A
hd4              jfs       1    1    1    open/syncd   /
hd2              jfs       75   75   1    open/syncd   /usr
hd9var           jfs       1    1    1    open/syncd   /var
hd3              jfs       3    3    1    open/syncd   /tmp
hd1              jfs       1    1    1    open/syncd   /home
```

Run the **reorgvg** command on the rootvg volume group:

```
# reorgvg rootvg
0516-962 reorgvg: Logical volume hd6 migrated.
0516-962 reorgvg: Logical volume hd8 migrated.
0516-962 reorgvg: Logical volume hd4 migrated.
0516-962 reorgvg: Logical volume hd2 migrated.
0516-962 reorgvg: Logical volume hd9var migrated.
0516-962 reorgvg: Logical volume hd3 migrated.
0516-962 reorgvg: Logical volume hd1 migrated.
```

If only the logical volume hd4 was to be reorganized, the command is as follows:

```
# reorgvg rootvg hd4
0516-962 reorgvg: Logical volume hd4 migrated.
```

5.2 Logical volumes

The logical volume is created within a volume group. Although there may be more than one logical volume per volume group, there can be only one file system per logical volume. The file system is super-imposed on the logical volume. The size of a logical volume can, for example, be increased without the file system size having to be increased. However, if you increase the file system

size, the logical volume size may need to be increased. This is done automatically until the maximum number of logical partitions setting is exceeded.

5.2.1 Creating a logical volume and setting the attributes

As just described, a logical volume is created in a volume group. The volume group is limited to up to 256 logical volumes per volume group by default. When you use the -B flag of the `mkvg` command during creation of the volume group, you create a BigVG capable volume group. The maximum will be 512 per volume group. In this section, a logical volume, `logvol01`, will be created on the volume group `volg01`. Although you can use the `mk1v` command to create a logical volume, using a SMIT dialog is much easier.

The following commands provide an example of this task.

Check the system to see what logical volumes are assigned to the system:

```
# lsvg -o | lsvg -i -l
volg01:
LV NAME          TYPE          LPs   PPs   PVs  LV STATE    MOUNT POINT
rootvg:
LV NAME          TYPE          LPs   PPs   PVs  LV STATE    MOUNT POINT
hd5              boot          1     1     1    closed/syncd N/A
hd6              paging        64    64    1    open/syncd   N/A
hd8              jfslog        1     1     1    open/syncd   N/A
hd4              jfs           3     3     1    open/syncd   /
hd2              jfs           32    32    1    open/syncd   /usr
hd9var           jfs           1     1     1    open/syncd   /var
hd3              jfs           2     2     1    open/syncd   /tmp
hd1              jfs           3     3     1    open/syncd   /home
lv00             jfs           4     4     1    open/syncd   /usr/welcome_arcade
lv01             jfs           2     2     1    open/syncd   /usr/welcome
```

Check the volume group, `volg01`, for availability and size:

```
# lsvg volg01
VOLUME GROUP:   volg01                VG IDENTIFIER:  000bc6fdca261bbf
VG STATE:       active                PP SIZE:        16 megabyte(s)
VG PERMISSION:  read/write           TOTAL PPs:      1084 (17344 megabytes)
MAX LVs:        256                FREE PPs:       1084 (17344 megabytes)
LVs:           0                USED PPs:       0 (0 megabytes)
OPEN LVs:       0                QUORUM:         2
TOTAL PVs:     2                VG DESCRIPTORS: 3
STALE PVs:     0                STALE PPs:      0
ACTIVE PVs:    2                AUTO ON:        yes
MAX PPs per PV: 1016           MAX PVs:        32
```

Run **smitty lv** to enter the SMIT Logical Volumes menu, as shown in Figure 5-5.

```
# smitty lv
```

```
Logical Volumes

Move cursor to desired item and press Enter.

List All Logical Volumes by Volume Group
Add a Logical Volume
Set Characteristic of a Logical Volume
Show Characteristics of a Logical Volume
Remove a Logical Volume
Copy a Logical Volume

F1=Help      F2=Refresh   F3=Cancel    Esc+8=Image
Esc+9=Shell  Esc+0=Exit   Enter=Do
```

Figure 5-5 The SMIT Logical Volumes screen

Select the **Add a Logical Volume** option and press Enter, as shown in Figure 5-6 on page 98.

```

                                Add a Logical Volume

Type or select a value for the entry field.
Press Enter AFTER making all desired changes.

* VOLUME GROUP name                [Entry Fields]
                                     [volg01]
                                     +

F1=Help          F2=Refresh        F3=Cancel        F4=List
Esc+5=Reset      Esc+6=Command    Esc+7=Edit      Esc+8=Image
Esc+9=Shell      Esc+0=Exit      Enter=Do

```

Figure 5-6 The Add a Logical Volume screen

Type in the volume group name or press F4 to select the volume group name. In Figure 5-7, the Add a Logical Volume menu can be seen.

```

                                Add a Logical Volume

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
Logical volume NAME                  [logvol01]
* VOLUME GROUP name                  volg01
* Number of LOGICAL PARTITIONS      [2]
PHYSICAL VOLUME names                [hdisk1 hdisk2]
Logical volume TYPE                  [ ]
POSITION on physical volume          center
RANGE of physical volumes            minimum
MAXIMUM NUMBER of PHYSICAL VOLUMES  [2]
to use for allocation                #
Number of COPIES of each logical    1
partition                            +
Mirror Write Consistency?           yes
Allocate each logical partition copy yes
[MORE...11]

F1=Help          F2=Refresh        F3=Cancel        F4=List
Esc+5=Reset      Esc+6=Command    Esc+7=Edit      Esc+8=Image
Esc+9=Shell      Esc+0=Exit      Enter=Do

```

Figure 5-7 The Add a Logical Volume menu 1 of 2

The name of the logical volume has been entered as well as the physical volumes it will utilize for the logical volume. You do not have to specify the physical volume names, except to make the LVM choose the physical partition allocation policy other than the default.

As shown in this case, the logical volume has been set up to use the center position for the allocation of the logical volume on the physical volume. The default for this setting is center.

The intra-disk allocation policy choices are based on the five regions of a disk where physical partitions can be located. The five regions are: Outer edge, inner edge, outer middle, inner middle, and center. The edge partitions have the slowest average seek times, which generally result in longer response times for any application that uses them. The center partitions have the fastest average seek times, which generally result in the best response time for any application that uses them. There are, however, fewer partitions on a physical volume at the center than at the other regions (unless when mirrored with mirror write consistency check).

The inter-disk allocation policy determines the number (or range) of physical volumes that the physical partitions will be allocated across. If the minimum inter-physical volume setting is selected, the physical partitions assigned to the logical volume are located on a single physical volume to enhance availability. If the maximum inter-physical volume is selected (Range=maximum), the physical partitions are located on multiple physical volumes to enhance performance.

For non-mirrored logical volumes, use the minimum settings to provide the greatest availability (access to data in case of hardware failure). The minimum setting indicates that one physical volume should contain all the original physical partitions of this logical volume, if possible. If the allocation program must use two or more physical volumes, it uses the minimum number while remaining consistent with other parameters.

The strict allocation policy determines whether copies of a logical partition share the same physical volume. The strict option does not allow copies of a logical partition to share the same physical volume. The non-strict option allows copies of a logical partition to share the same physical volume. The super strict option does not allow physical partitions from one copy of a logical partition to share the same physical volume as physical partitions from another copy of other logical partitions of the same logical volume.

In the case of multiple logical partitions, the inter-physical volume settings and the intra-physical volume settings must be considered for availability and speed requirements.

Figure 5-8 illustrates the regions that can be used for allocating physical partitions in a physical volume.

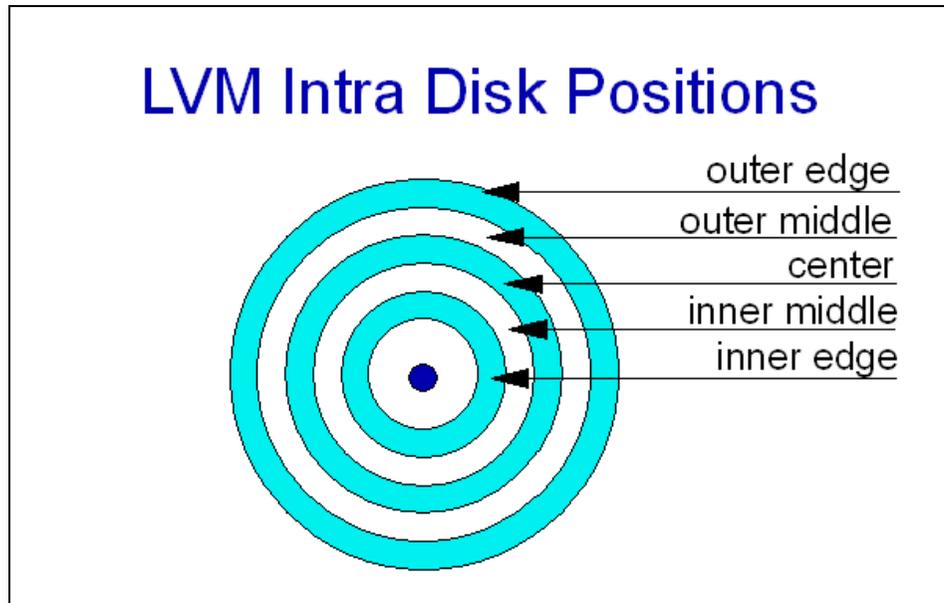


Figure 5-8 Five regions of a disk

Note: To ensure a JFS is allocated on the center region, you have to create a logical volume on the center region first, then create the JFS on it, or, if the space is available, change the attributes and reorg.

Figure 5-9 on page 101 is the second half of the screen as described in Figure 5-7 on page 98, and the significant portion here is stripe size, which is not being set up for this example since this attribute is only used in the striped logical volumes.

```

                                Add a Logical Volume

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[MORE...11]                                [Entry Fields]
Mirror Write Consistency?                  yes          +
Allocate each logical partition copy      yes          +
  on a SEPARATE physical volume?
RELOCATE the logical volume during        yes          +
  reorganization?
Logical volume LABEL                       []
MAXIMUM NUMBER of LOGICAL PARTITIONS     [512]        #
Enable BAD BLOCK relocation?              yes          +
SCHEDULING POLICY for reading/writing     parallel     +
  logical partition copies
Enable WRITE VERIFY?                       no          +
File containing ALLOCATION MAP              []
Stripe Size?                              [Not Striped] +
[BOTTOM]

F1=Help          F2=Refresh          F3=Cancel          F4=List
Esc+5=Reset      Esc+6=Command       Esc+7=Edit        Esc+8=Image
Esc+9=Shell      Esc+0=Exit          Enter=Do

```

Figure 5-9 The Add a Logical Volume menu 2 of 2

After the command is run a screen similar to Figure 5-10 is shown.

```

                                COMMAND STATUS

Command: OK          stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

logvol01
█

F1=Help          F2=Refresh          F3=Cancel          Esc+6=Command
Esc+8=Image      Esc+9=Shell         Esc+0=Exit         /=Find
n=Find Next

```

Figure 5-10 The Command Status screen for the logical volume

The logical volume has been created. Press F10 to exit to the command prompt and check the logical volume. The following example shows the results of this action:

```
# lslv logvo101
LOGICAL VOLUME:      logvo101          VOLUME GROUP:      volg01
LV IDENTIFIER:      000bc6fdca261bbf.1    PERMISSION:        read/write
VG STATE:           active/complete  LV STATE:          closed/syncd
TYPE:               jfs              WRITE VERIFY:      off
MAX LPs:            512              PP SIZE:           16 megabyte(s)
COPIES:             1                SCHED POLICY:     parallel
LPs:                2                PPs:              2
STALE PPs:         0                BB POLICY:         relocatable
INTER-POLICY:       minimum          RELOCATABLE:      yes
INTRA-POLICY:       center           UPPER BOUND:      2
MOUNT POINT:        N/A             LABEL:            None
MIRROR WRITE CONSISTENCY: on
EACH LP COPY ON A SEPARATE PV ?: yes
```

The logical volume is now created and ready for the creation of file systems.

The INTER-POLICY setting controls the behavior of allocation between physical volumes in a VG. You can use this attribute to command the LVM to precisely adhere to the physical partition allocation policy for the logical volume.

The **iostat** command is used for monitoring system input/output device loading by observing the time the physical disks are active in relation to their average transfer rates. The **iostat** command generates reports that can be used to change system configuration to better balance the input/output load between physical disks and adapters. For example, to display a continuous disk report at two second intervals for the disk with the logical name hdisk0, enter:

```
# iostat -d hdisk0 2
```

Disks:	% tm_act	Kbps	tps	Kb_read	Kb_wrtn
hdisk0	0.0	0.0	0.0	49	561
hdisk0	0.0	0.0	0.0	0	0
hdisk0	0.0	0.0	0.0	0	0
hdisk0	0.0	0.0	0.0	0	0
hdisk0	0.0	0.0	0.0	0	0
hdisk0	0.0	0.0	0.0	0	0

5.2.2 Different LVs associated with a JFS

In every volume group that has logical volumes created, there are various logical volumes that are associated with the system as shown, in the following example:

```
# lsvg -o | lsvg -i -l
volg01:
```

LV NAME	TYPE	LPs	PPs	PVs	LV STATE	MOUNT POINT
logvol01	jfs	2	2	1	closed/syncd	N/A
rootvg:						
LV NAME	TYPE	LPs	PPs	PVs	LV STATE	MOUNT POINT
hd5	boot	2	2	1	closed/syncd	N/A
hd6	paging	40	40	1	open/syncd	N/A
hd8	jfslog	1	1	1	open/syncd	N/A
hd4	jfs	2	2	1	open/syncd	/
hd2	jfs	118	118	1	open/syncd	/usr
hd9var	jfs	1	1	1	open/syncd	/var
hd3	jfs	3	3	1	open/syncd	/tmp
hd1	jfs	11	11	1	open/syncd	/home

The following logical volumes are all journaled file systems belonging to the rootvg volume groups hd4, hd2, hd9var, hd3, and hd1.

The logical volume hd8 is the JFS log. The function of the JFS log is to record modifications in the JFS file system.

The logical volume hd5 is the boot logical volume. Recreate it and rewrite boot image if it is accidentally corrupted or missing.

The hd6 is the paging space for the rootvg volume group by default. Paging space, sometimes called swap space, is a regular logical volume of the type paging. It is fixed disk storage that allows a system to operate with virtual memory larger than the installed physical memory.

The logvol01 logical volume is the user-defined logical volume belonging to the user defined volume group volg01.

5.2.3 Moving a logical volume from one physical volume to another

There may be instances where the logical volumes of a volume group need to be migrated from one disk to another. This must not be confused with migrating a physical volume in the same volume group as described in 5.1.2, “Migrating data to another disk in the same volume group” on page 88. This method is used for moving the specified logical volumes on the physical volume, not every physical partition on the physical volume.

Using the `migratepv` command

When migrating a logical volume from one physical volume to another using the `migratepv` command, there are a few considerations before doing this:

- ▶ The `migratepv` command cannot migrate a striped logical volume. To do this, use the `cp1v` and `rm1v` commands.
- ▶ The new physical volume must be part of the same volume group.

- ▶ Ensure that the file systems on all logical volumes are unmounted.
- ▶ Only the root user has authority to use the `migratepv` command.

The following commands provide an example of this task.

In the following example, the logical volume `lv01` on the volume group `volg1` situated on `hdisk1` will be migrated to `hdisk2`:

```
# lslv lv01
LOGICAL VOLUME:    lv01                      VOLUME GROUP:    volg1
LV IDENTIFIER:    000bc6fdbff8e05a.1          PERMISSION:      read/write
VG STATE:         active/complete            LV STATE:        opened/syncd
TYPE:             jfs                        WRITE VERIFY:    off
MAX LPs:          512                        PP SIZE:         16 megabyte(s)
COPIES:           1                         SCHED POLICY:   parallel
LPs:              10                        PPs:             10
STALE PPs:        0                         BB POLICY:       relocatable
INTER-POLICY:     minimum                   RELOCATABLE:    yes
INTRA-POLICY:     middle                    UPPER BOUND:    32
MOUNT POINT:      /fs1                      LABEL:           /fs1
MIRROR WRITE CONSISTENCY: on
EACH LP COPY ON A SEPARATE PV ?: yes

# lspv hdisk1
PHYSICAL VOLUME:    hdisk1                    VOLUME GROUP:    volg1
PV IDENTIFIER:      000bc6fdbff75ee2          VG IDENTIFIER     000bc6fdbff8e05a
PV STATE:           active
STALE PARTITIONS:   0                       ALLOCATABLE:     yes
PP SIZE:            16 megabyte(s)           LOGICAL VOLUMES: 2
TOTAL PPs:          542 (8672 megabytes)     VG DESCRIPTORS:  2
FREE PPs:           531 (8496 megabytes)
USED PPs:           11 (176 megabytes)
FREE DISTRIBUTION:  109..97..108..108..109
USED DISTRIBUTION:  00..11..00..00..00

# lspv hdisk2
PHYSICAL VOLUME:    hdisk2                    VOLUME GROUP:    volg1
PV IDENTIFIER:      000bc6fdbff92812          VG IDENTIFIER     000bc6fdbff8e05a
PV STATE:           active
STALE PARTITIONS:   0                       ALLOCATABLE:     yes
PP SIZE:            16 megabyte(s)           LOGICAL VOLUMES: 0
TOTAL PPs:          542 (8672 megabytes)     VG DESCRIPTORS:  1
FREE PPs:           542 (8672 megabytes)
USED PPs:           0 (0 megabytes)
FREE DISTRIBUTION:  109..108..108..108..109
USED DISTRIBUTION:  00..00..00..00..00
```

Migrate the logical volume from one disk to another:

```
# migratepv -l lvol1 hdisk1 hdisk2
```

Check that the logical volume has migrated:

```
# lslv lvol1
LOGICAL VOLUME:      lvol1                VOLUME GROUP:      volg1
LV IDENTIFIER:      000bc6fdbff8e05a.1  PERMISSION:        read/write
VG STATE:           active/complete  LV STATE:          opened/syncd
TYPE:               jfs                WRITE VERIFY:      off
MAX LPs:            512                PP SIZE:           16 megabyte(s)
COPIES:             1                  SCHED POLICY:     parallel
LPs:                10                PPs:              10
STALE PPs:          0                  BB POLICY:         relocatable
INTER-POLICY:       minimum            RELOCATABLE:      yes
INTRA-POLICY:       middle             UPPER BOUND:      32
MOUNT POINT:        /fs1              LABEL:            /fs1
MIRROR WRITE CONSISTENCY: on
EACH LP COPY ON A SEPARATE PV ?: yes

# lspv hdisk1
PHYSICAL VOLUME:    hdisk1                VOLUME GROUP:      volg1
PV IDENTIFIER:      000bc6fdbff75ee2  VG IDENTIFIER      000bc6fdbff8e05a
PV STATE:           active
STALE PARTITIONS:   0                  ALLOCATABLE:       yes
PP SIZE:            16 megabyte(s)        LOGICAL VOLUMES:   1
TOTAL PPs:          542 (8672 megabytes)  VG DESCRIPTORS:    2
FREE PPs:           541 (8656 megabytes)
USED PPs:           1 (16 megabytes)
FREE DISTRIBUTION:  109..107..108..108..109
USED DISTRIBUTION:  00..01..00..00..00

# lspv hdisk2
PHYSICAL VOLUME:    hdisk2                VOLUME GROUP:      volg1
PV IDENTIFIER:      000bc6fdbff92812  VG IDENTIFIER      000bc6fdbff8e05a
PV STATE:           active
STALE PARTITIONS:   0                  ALLOCATABLE:       yes
PP SIZE:            16 megabyte(s)        LOGICAL VOLUMES:   1
TOTAL PPs:          542 (8672 megabytes)  VG DESCRIPTORS:    1
FREE PPs:           532 (8512 megabytes)
USED PPs:           10 (160 megabytes)
FREE DISTRIBUTION:  109..98..108..108..109
USED DISTRIBUTION:  00..10..00..00..00
```

The logical volume has now been migrated from hdisk1 to hdisk2.

Note: The following step is used only if you are migrating the BLV in a rootvg volume group:

```
migratepv -l hd5 SourceDiskNumber DestinationDiskNumber
```

Next, you will get a message warning you to perform the **bosboot** command on the destination disk. You must also perform a **mkboot -c** command to clear the boot record on the source. Do the following:

```
bosboot -a -d /dev/DestinationDiskNumber
```

Then:

```
bootlist -m normal DestinationDiskNumber
```

Then:

```
mkboot -c -d /dev/SourceDiskNumber
```

Using the **cp1v** and **rm1v** commands

The **cp1v** command is used to copy a logical volume to either a new or existing logical volume. You can use the **cp1v** command to copy a non-striped or striped logical volume. Considerations when using the **cp1v** command are:

- ▶ The **cp1v** command will fail if a new logical volume is created and the volume group is variedon in concurrent mode.
- ▶ If you are copying a striped logical volume, and the destination logical volume does not exist, an identical copy, including the striped block size and striping width of the source logical volume, is created, and then the data is copied.
- ▶ If you are copying a striped logical volume and you have created the destination logical volume with the **mk1v** command using a different stripe block size and striping width, or the destination is not a striped logical volume, the new characteristics are maintained, and the data is copied from the source logical volume.
- ▶ To use the **cp1v** command, you must either have root user authority or be a member of the system group.

The **rm1v** command is used to remove a logical volume from a volume group. Considerations when using the **rm1v** command are:

- ▶ This command removes information about the specified logical volumes.
- ▶ To use the **rm1v** command, you must either have root user authority or be a member of the system group.

5.2.4 Mirroring logical volumes

Mirroring the logical volumes is important because:

- ▶ Maintaining an active mirrored copy on another disk ensures continuous operation in the event that a disk experiences failure.
- ▶ For rootvg, such maintenance provides the ability to boot from more than one disk in the event that another boot disk has failed. In some cases, the ability to boot from alternate disks may require some user intervention.

The following commands demonstrate the task. In this case rootvg is used to illustrate the procedure. The following steps assume that you have rootvg contained on hdisk0 and is attempting to mirror the rootvg to a new disk, hdisk1.

1. Extend rootvg to hdisk1 by executing the following:

```
# extendvg rootvg hdisk1
```

You may encounter the following error message:

```
0516-050 Not enough descriptor space left in this volume group. Either try
adding a smaller PV or use another volume group.
```

In this case, you may not add hdisk1 to rootvg for mirroring. You may attempt to mirror rootvg's logical volumes to another disk that already exists in rootvg, or you may attempt to add a smaller disk to the rootvg. If neither option is possible, then mirroring rootvg cannot be performed on this system.

2. Disable QUORUM by executing the following:

```
# chvg -Qn rootvg
```

3. Mirror the logical volumes that make up the AIX operating system by executing the following:

```
# mklvcopy hd1 2 hdisk1
# mklvcopy hd2 2 hdisk1
# mklvcopy hd3 2 hdisk1
# mklvcopy hd4 2 hdisk1
# mklvcopy hd5 2 hdisk1
# mklvcopy hd6 2 hdisk1
# mklvcopy hd8 2 hdisk1
# mklvcopy hd9var 2 hdisk1
# mklvcopy hd10opt 2 hdisk1
```

This command should also be run against any user-created logical volumes in rootvg that you would like to have mirrored.

If you have other paging devices, rootvg and non-rootvg, it is recommended that you also mirror those logical volumes in addition to hd6. If hd5 consists of more than one logical partition, then, after mirroring hd5, you must verify that the mirrored copy of hd5 resides on contiguous physical partitions.

This can be verified with the following command:

```
# ls1v -m hd5
```

If the mirrored hd5 partitions are not contiguous, you must delete the mirror copy of hd5 (on hdisk1) and rerun the **mk1vcopy** command for hd5 using the -m option. You should consult documentation on the usage of the -m option for the **mk1vcopy** command.

4. Synchronize the newly created mirrors with the following command:

```
# syncvg -v rootvg
```

5. Initialize all boot records and devices by executing the following command:

```
bosboot -a -d /dev/hdisk0
```

A message appears confirming that the **bosboot** command run successfully.

```
bosboot: Boot image is 13260 512 byte blocks.
```

6. Initialize the boot list by executing the following:

```
# bootlist -m normal hdisk0 hdisk1
```

Even though this command identifies the list of possible boot disks, it does not guarantee that the system will boot from the alternate disk in all cases involving failures of the first disk. In such situations, it may be necessary to boot from the installation or maintenance media: Select **maintenance**, reissue the **bootlist** command leaving out the failing disk, and then reboot. On some models, firmware provides a utility for selecting the boot device at boot time. This may also be used to force the system to boot from the alternate disk.

7. Shut down and reboot the system by executing the following command:

```
# shutdown -Fr
```

This is so that the Quorum OFF functionality takes effect.

The **smitty mirrorvg** fast path or **mirrorvg** command can be used to mirror the rootvg volume group or other volume groups. For example, to create another copy of the rootvg volume group, enter:

```
# mirrorvg rootvg hdisk0 hdisk1
0516-1124 mirrorvg: Quorum requirement turned off, reboot system for this
to take effect for rootvg.
0516-1126 mirrorvg: rootvg successfully mirrored, user should perform
bosboot of system to initialize boot records. Then, user must modify
bootlist to include: hdisk0 hdisk1.
```

Run the **bosboot** and **bootlist** commands.

Removing logical volume mirror

Following is an example of removing mirror from hd6. First check if your paging space (hd6) is mirrored.

```
# lslv -m hd6
hd6:N/A
LP   PP1  PV1                PP2  PV2                PP3  PV3
0001 0110 hdisk0        0110 hdisk1
```

PV1 and PV2 listing hdisk0 and hdisk1, respectively, indicate that this logical volume is mirrored.

To remove a copy from hdisk1, run:

```
# rmlvcopy hd6 1 hdisk1
# lslv -m hd6
hd6:N/A
LP   PP1  PV1                PP2  PV2                PP3  PV3
0001 0110 hdisk0
```

This shows only one copy of hd6.

To remove mirror from the entire rootvg volume group use **smitty unmirrorvg** fast path or enter the following commands to keep one copy of rootvg:

```
# unmirrorvg rootvg
0516-1246 rmlvcopy: If hd5 is the boot logical volume, please run 'chpv -c
<disk name>'
    as root user to clear the boot record and avoid a potential boot
    off an old boot image that may reside on the disk from which this
    logical volume is moved/removed.
0516-1132 unmirrorvg: Quorum requirement turned on, reboot system for this
to take effect for rootvg.
0516-1144 unmirrorvg: rootvg successfully unmirrored, user should perform
bosboot of system to reinitialize boot records. Then, user must modify
bootlist to just include: hdisk0.
# chpv -c hdisk1
```

Run the **bosboot** and **bootlist** commands.

Replacing a disk

Remove mirroring as described in “Removing logical volume mirror” on page 109 so that all the physical partitions on that disk are deallocated. To remove the disk from the volume group and remove the logical definition, enter:

```
# reducevg -f rootvg hdisk1
# rmdev -dl hdisk1
hdisk1 deleted
```

Or you may have to remove the disk from ODM by running:

```
# odmdelete -q value=00015f8fe7288c7b0000000000000000 -o CuAt  
0518-307 odmdelete: 2 objects deleted.
```

At that point, the disk can be replaced and mirroring has to be recreated following the procedure above.

5.3 Journalled file systems

A file system is a general representation of a storage space on UNIX. Although there are several file system implementations on AIX (JFS, NFS, AFS, and so on), only the journalled file system (JFS) is a native local file system implementation on AIX (the newest release of AIX, AIX 5L, also supports an enhanced native local file system called JFS2). A JFS is created on top of a logical volume. There are some considerations when you are creating a JFS:

- ▶ If performance is a key consideration, then the logical volume must be created in the center of a disk (unless you are using mirrored disks with MWCC, then the outer edge is best). A file system cannot reside in the center of the physical volume if the physical partitions are placed on the middle or edges of the physical volume.
- ▶ When increasing the size of the file system it will automatically increase the size of the logical volume if required. However, increasing the size of a logical volume will not automatically increase the size of the file system.

There is a lot of information about file systems. The following discussion starts at an intermediate technical level.

5.3.1 Metadata

Data and metadata must not be confused. Data is the actual user data within a file system. Metadata is the structural data associated with the file system, such as the organization of directories, i-node tables, and links. Metadata is not data supplied by a user; it is information about the structure of user data.

5.3.2 Adding a file system

The `crfs` command creates a file system on a logical volume within a previously created volume group. A new logical volume is created for the file system unless the name of an existing logical volume is specified using the `-d` flag. An entry for the file system is put into the `/etc/filesystems` file.

The System Management Interface Tool (SMIT) `smit crfs` fast path can also be used to run this command, as in the following example.

```

                                Add a Standard Journaled File System

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* LOGICAL VOLUME name          █          +
* MOUNT POINT                  []          +
Mount AUTOMATICALLY at system restart?  no          +
PERMISSIONS                    read/write  +
Mount OPTIONS                  []          +
Start Disk Accounting?         no          +
Fragment Size (bytes)          4096        +
Number of bytes per inode       4096        +
Allocation Group Size (MBytes)   8          +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell    F10=Exit         Enter=Do

```

Figure 5-11 The SMIT `crfs` screen

Tip: Enter `smit lv` and select the **Add a Logical Volume** option to create the previously defined logical volume before the file system can be created. Alternatively, use the `mk1v` command at the command line.

5.3.3 Mounting and unmounting a file system

To mount a newly created JFS, the `mount` command is used. To unmount a file system, the `umount` or `unmount` commands are used. Either of these commands can be used, as they are identical in function.

Unmounting a file system

To unmount the file system, use the `umount` or `unmount` commands. If, however, a file system cannot be unmounted, use the `fuser` command to determine who has a file open on the device. You cannot use the `umount` command on a device in use. A device is in use if any file is open for any reason or if a user's current directory is on that device.

In the following example, the file system test01 needs to be unmounted:

```
# umount /test01
umount: 0506-349 Cannot unmount /dev/lv02: The requested resource is busy.
```

The file system cannot be unmounted, as it is busy. Check which resource is locking the device:

```
# fuser -u /dev/lv02
/dev/lv02: 13984c(user01)
```

The resource has been identified as user01. The user can either log out of the system, or, if needed, the running resource can be terminated using the `kill -1` or `fuser -k` commands.

Mounting a file system over an existing file system

If a file system is mounted over a directory (called a *mount point*) held in another file system, beneath the mount point hierarchy in the mounted-over file system, the original file system data is hidden by the newly mounted file system. The originally mounted file system's data will be unavailable for use.

An example of where you would over mount is in a development environment where you are working with different levels of the same code and do not wish to change your make files. In the next example, there are two logical volumes, lv01 and lv02. The file system on lv01 is mounted on /filesys01, and the file system on lv02 is not yet mounted:

```
# mount
node          mounted          mounted over    vfs    date          options
-----
/dev/hd4      /
/dev/hd2      /usr
/dev/hd9var   /var
/dev/hd3      /tmp
/dev/hd1      /home
/dev/lv00     /usr/welcome_arcade jfs    Jun 11 16:42
rw,log=/dev/hd8
/dev/lv01     /usr/welcome       jfs    Jun 11 16:42 rw,log=/dev/hd8
/dev/lv0101   /filsys01          jfs    Jun 14 13:31
rw,log=/dev/loglv00
```

If a listing of the mounted file system filsys01 is done, the directory containing data fs1dir is displayed:

```
# ls -l filsys01
total 16
drwxr-sr-x 20 root    sys      2560 Jun 14 13:32 fs1dir
drwxrwx--- 2 root    system   512 Jun 14 13:28 lost+found
```

If the file system on logical volume lvol02 is now mounted over the same mount point as that of lvol01:

```
# mount /dev/lvol02 /filsys01
```

When checked it can be noted that both file systems are still shown as being mounted:

```
# mount
node          mounted      mounted over  vfs      date      options
-----
/dev/hd4      /
/dev/hd2      /usr
/dev/hd9var   /var
/dev/hd3      /tmp
/dev/hd1      /home
/dev/lv00     /usr/welcome_arcade jfs      Jun 11 16:42
rw,log=/dev/hd8
/dev/lv01     /usr/welcome      jfs      Jun 11 16:42 rw,log=/dev/hd8
/dev/lvol01   /filsys01         jfs      Jun 14 13:31
rw,log=/dev/loglv00
/dev/lvol02   /filsys01         jfs      Jun 14 13:44
rw,log=/dev/loglv01
```

If a listing of the mounted file system filsys01 is done, the directory containing data fs2dir is displayed:

```
# ls -l /filsys01
total 32
drwxr-sr-x  5 root    sys      9728 Jun 14 13:37 fs2dir
drwxrwx---  2 root    system   512 Jun 14 13:28 lost+found
```

The way to fix this is to unmount the logical volume lvol02 and then mount it on a new mount point:

```
# umount /dev/lvol02
```

Once unmounted, the data on lvol01 will be available again, as it was not overwritten:

```
# mount
node          mounted      mounted over  vfs      date      options
-----
/dev/hd4      /
/dev/hd2      /usr
/dev/hd9var   /var
/dev/hd3      /tmp
/dev/hd1      /home
/dev/lv00     /usr/welcome_arcade jfs      Jun 11 16:42
rw,log=/dev/hd8
/dev/lv01     /usr/welcome      jfs      Jun 11 16:42 rw,log=/dev/hd8
```

```

                /dev/lvo101      /filsys01      jfs      Jun 14 13:31
rw,log=/dev/loglv00

# ls -l /filsys01
total 16
drwxr-sr-x  20 root      sys      2560 Jun 14 13:32 fs1dir
drwxrwx---   2 root      system   512 Jun 14 13:28 lost+found

```

5.3.4 Reducing the size of a file system

There may be occasion to shrink a file system, although this may not be common. Before shrinking a file system, the following needs to be checked:

- ▶ For a file system in rootvg, image.data must exist and be current. If this file does not exist or you feel it is out of date, use the **mkszfile** command to create this file.
- ▶ For a file system that may exist on a user volume group the file `/tmp/vgdata/vgname/vgname.data` must exist. If this file does not exist, then the **mkvgdata VGName** command must be run, which will create this file.

Following is an example of the information for the /usr file system from the image.data file showing volume group, logical volume, and file system information:

```

##Command used for vg_data; /usr/sbin/lsvg

vg_data:
    VGNAME= rootvg
    PPSIZE= 16
    ...
##Command used for lv_data; /usr/sbin/lslv
...
lv_data:
    VOLUME_GROUP= rootvg
    LV_SOURCE_DISK_LIST= hdisk0
    ...
    LOGICAL_VOLUME= hd2
    ...
    MAX_LPS= 32512
    ...
    LPS= 430
    ...
    MOUNT_POINT= /usr
    ...
    PP_SIZE= 16
    ...
    PP= 430
    ...

```

```

        LV_MIN_LPS= 54
...
##Commands used for fs_data; /usr/bin/df and /usr/sbin/lvsjfs
...
fs_data:
    FS_NAME= /usr
    FS_SIZE= 14090240
    FS_MIN_SIZE= 1748712
    FS_LV= /dev/hd2
    FS_FS= 4096
    FS_NBPI= 4096
    FS_COMPRESS= no
    FS_BF= false
...

```

The following is an example of /tmp/vgdata/vgname/vgname.data:

```

##Command used for vg_data; /usr/sbin/lsvg

vg_data:
    VGNAME= volg01
    PPSIZE= 16
    VARYON= yes
    VG_SOURCE_DISK_LIST= hdisk1
...
lv_data:
    VOLUME_GROUP= volg01
    LV_SOURCE_DISK_LIST= hdisk1
    LV_IDENTIFIER= 000fa17d2d4bda99.3
    LOGICAL_VOLUME= lv00
    VG_STAT= active/complete
    TYPE= jfs
    MAX_LPS= 512
    COPIES= 1
    LPs= 40
    STALE_PPs= 0
    INTER_POLICY= minimum
    INTRA_POLICY= middle
    MOUNT_POINT= /u
    MIRROR_WRITE_CONSISTENCY= on
...
    PP_SIZE= 16
    SCHED_POLICY= parallel
    PP= 40
...
    LV_MIN_LPS= 37
...
##Commands used for fs_data; /usr/bin/df and /usr/sbin/lvsjfs

```

```

fs_data:
    FS_NAME= /u
    FS_SIZE= 1310720
    FS_MIN_SIZE= 1188384
    FS_LV= /dev/lv00
    FS_FS= 4096
    FS_NBPI= 4096
...

```

On the above system the following is the current disk usage as per the **df** command:

```

# df -k
Filesystem      1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4         16384           8656  48%     1281   16% /
/dev/hd2        7045120        6170764  13%    31736   2% /usr
/dev/hd9var     16384          14416  13%     399   10% /var
/dev/hd3        16384          15792   4%      28    1% /tmp
/dev/hd1        16384          15764   4%      41    2% /home
/dev/lv00       655360         61168  91%     4528   3% /u

```

On investigation into the reason for the high utilization of the /u file system, the following is found:

```

# ls -l /u
total 804520
drwxrwxrwx  2 root    system      512 Jul 05 17:59 lost+found
-rwxrwxrwx  1 user01  sys        411904000 Jul 05 18:28 tempfile1
drwxrwsrwx  7 root    sys         512 Jul 05 18:17 userdir1

```

The file /u/tempfile1 is no longer needed and can be removed.

Remove the file using the **rm** command and then check the new disk usage as per the **df** command:

```

# df -k
Filesystem      1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4         16384           6288  62%     1283   16% /
/dev/hd2        7045120        6170764  13%    31736   2% /usr
/dev/hd9var     16384          14412  13%     400   10% /var
/dev/hd3        16384          15800   4%      28    1% /tmp
/dev/hd1        16384          15764   4%      41    2% /home
/dev/lv00       655360         463820  30%     4527   3% /u

```

Run the **mkvgdata** command to create a new /tmp/vgdata/vgname/vgname.data file. In this case, the file name is /tmp/vgdata/volg01/volg01.data. For the root volume group, the **mkszfile** command needs to be run:

```

# mkvgdata volg01

```

If the `/tmp/vgdata/volg01/volg01.data` is checked, the following can be seen:

```
lv_data:
    VOLUME_GROUP= volg01
    LV_SOURCE_DISK_LIST= hdisk1
    LV_IDENTIFIER= 000fa17d2d4bda99.3
    ...
    LABEL= /u
    MAPFILE=
    LV_MIN_LPS= 12
...
fs_data:
    FS_NAME= /u
    FS_SIZE= 1310720
    FS_MIN_SIZE= 383080
...
```

Notice that `LV_MIN_LPS=` has reduced from 37 to 12 and that `FS_MIN_SIZE=` has reduced from 1188384 to 383080. These numbers must be large enough to hold the current data, otherwise the process will fail.

Calculate any requirement for a new `PP_SIZE`. Typically a smaller disk will be more efficiently used with a smaller `PP_SIZE`. Old disks under 300 MB use a 2 MB `PP_SIZE` by default; all others use a 4 MB size by default. You have to conform to the general rule for making a volume group on a physical volume 1016 x `PP_SIZE` greater than the real capacity of the physical volume.

Save the volume group using the `savevg` command. During a `savevg`, the `/tmp/vgdata/vgname/vgname.data` file can be created using the `-i` flag, which has the same effect as running the `mkvgdata` command:

```
# savevg -f/dev/rmt0 -v volg01
```

```
...
```

```
The total size is 161611492 bytes.
```

```
Backup finished on Thu Jul 6 10:00:25 CDT 2000; there are 218300 blocks on 1 volumes.
```

```
0512-038 savevg: Backup Completed Successfully
```

Note: The `savevg` command uses the `backup` command to save the contents of the journaled file systems held in the target volume group. It never saves the contents of the raw device type logical volume (logical volumes not used for journaled file system).

Check the volume groups and the disks associated with the volume groups:

```
# lsvg
rootvg
volg01
```

```
# lspv
hdisk0      000fa17debb9c93a  rootvg
hdisk1      000fa17dc9bf376f  volg01
```

Unmount the file system:

```
# umount /u
```

To restore the contents of the backup tape of the volume group volg01 taken by the **savevg** command, you have to remove the physical volume and volume group, since the **restvg** command does not allow restoring to the existing volume group and the physical volumes. The **-f** option flag of the **reducevg** command deletes the physical volume forcibly if there are any logical volumes that exist on the physical volume:

```
# reducevg -d -f volg01 hdisk1
rmlv: Logical volume logvol01 is removed.
rmlv: Logical volume loglv00 is removed.
rmlv: Logical volume lv00 is removed.
ldeletepv: Volume Group deleted since it contains no physical volumes.
```

Now the system needs to be restored using the **restvg** command with the **-s** flag. The **-s** flag (meaning shrink) specifies that the logical volumes be created at the minimum size possible to accommodate the file systems. This size is specified by the value of the **LV_MIN_LPS** field of the **lv_data** stanza of the **vgname.data** file where the **vgname** is **/tmp/vgdata/volg01/volg01.data**:

```
# restvg -s -q -f/dev/rmt0
```

```
Will create the Volume Group:   volg01
Target Disks:   hdisk1
Allocation Policy:
    Shrink Filesystems:   yes
    Preserve Physical Partitions for each Logical Volume:   no
```

```
volg01
...
x          10970 ./u/...
The total size is 161611492 bytes.
The number of restored files is 4516.
```

After the restore, the file system may need to be mounted by using the **mount** command.

The file system can now be checked for disk usage:

```
# df -k
Filesystem      1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4         16384           6080   63%     1283   16% /
/dev/hd2        7045120        6170764  13%     31736   2% /usr
```

/dev/hd9var	16384	14408	13%	400	10%	/var
/dev/hd3	16384	15796	4%	29	1%	/tmp
/dev/hd1	16384	15764	4%	41	2%	/home
/dev/lv00	196608	23452	89%	4527	10%	/u

Note the file system has been shrunk from a size of 655360 to 196608. The file system can now grow as required.

The system can also be restored using the **smitty restvg** command, as shown in Figure 5-12. Note the SHRINK the file systems? option is set to yes.

```

Remake a Volume Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Restore DEVICE or FILE          [/dev/rmt0]          +/
  SHRINK the filesystems?         yes                  +
  PHYSICAL VOLUME names           []                  +
    (Leave blank to use the PHYSICAL VOLUMES listed
     in the vgname.data file in the backup image)
  Use existing MAP files?         yes                  +
  Physical partition SIZE in megabytes []              +#
    (Leave blank to have the SIZE determined
     based on disk size)
  Number of BLOCKS to read in a single input []         #
    (Leave blank to use a system default)

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit      F8=Image
F9=Shell     F10=Exit      Enter=Do

```

Figure 5-12 The SMIT restvg screen

When shrinking a root volume group file system the process is similar with a few exceptions:

- ▶ Use the **mksysb** backup to back up the rootvg volume group.
- ▶ User volume groups will have to be exported using the **exportvg** command.
- ▶ The system would have to be restored. A panel will be available to allow you to shrink the file systems.
- ▶ Once the system has been restored, check to see if your user volume groups need to be imported by using the **importvg** command.

5.3.5 Removing a file system

To remove a file system from the system, the **rmfs** command is used as follows.

To remove the file system **filsys02**, which resides on the logical volume **lvol02**, the following steps are followed:

```
# mount
node          mounted          mounted over    vfs      date          options
-----
/dev/hd4      /                  /               jfs      Jun 11 16:41  rw,log=/dev/hd8
/dev/hd2      /usr              /usr            jfs      Jun 11 16:41  rw,log=/dev/hd8
/dev/hd9var   /var              /var            jfs      Jun 11 16:41  rw,log=/dev/hd8
/dev/hd3      /tmp              /tmp            jfs      Jun 11 16:41  rw,log=/dev/hd8
/dev/hd1      /home             /home           jfs      Jun 11 16:42  rw,log=/dev/hd8
/dev/lv00     /usr/welcome_arcade jfs             Jun 11 16:42
rw,log=/dev/hd8
/dev/lv01     /usr/welcome      /usr/welcome    jfs      Jun 11 16:42  rw,log=/dev/hd8
/dev/lvol01   /filsys01         /filsys01       jfs      Jun 14 13:31
rw,log=/dev/loglv00
/dev/lvol02   /filsys02         /filsys02       jfs      Jun 14 17:17
rw,log=/dev/loglv01
```

Unmount the file system that will be removed:

```
# umount /filsys02
```

Remove the file system (the **-r** flag of the **rmfs** command specifies that the mount point is also removed):

```
# rmfs -r /filsys02
rmlv: Logical volume lvol02 is removed.
# lslv -L lvol02
0516-306 lslv: Unable to find lvol02 in the Device
Configuration Database.
```

As can be seen when the file system is removed, the underlying logical volume will also be removed, and it also removes the corresponding entry in the **/etc/filesystems** file.

5.3.6 Creating a new JFS log

If a journal file system is used, then a **jfslog** is necessary. The **jfslog** is a logical volume used for maintaining the file system integrity of file systems associated with this **jfslog**. To create, mount, and use journal file systems, you have to have at least one **jfslog** per volume group. More than one journal file system residing in a volume group can share the same **jfslog**. To format the **jfslog** on the existing logical volume explicitly, you can use the **logform** command. This step is also done automatically when creating a new journal file system.

To configure the jfslog on the system, you can use the following method. The jfslogs shown in the examples are logical volume type *jfslog*.

The logical volume loglv00 is the JFS log for the volume group volg01:

```
# lsvg -o | lsvg -i -l
testvg1:
LV NAME          TYPE      LPs  PPs  PVs  LV STATE    MOUNT POINT
loglv01          jfslog    1    1    1    open/syncd  N/A
lv02             jfs       10   10   1    open/syncd  /test01
volg01:
LV NAME          TYPE      LPs  PPs  PVs  LV STATE    MOUNT POINT
logvo101        jfs       10   10   1    open/syncd  /userfs01
loglv00          jfslog    1    1    1    open/syncd  N/A
rootvg:
...
lv01             jfs       2    2    1    open/syncd  /usr/welcome
```

If while mounting a file system the following or a similar error should occur:

```
# mount /userfs01
mount: 0506-323 Cannot get information about log device /dev/loglv00.
```

Remove the corrupt logical volume loglv00 from the system:

```
# rmlv loglv00
Warning, all data contained on logical volume loglv00 will be destroyed.
rmlv: Do you wish to continue? y(es) n(o)? y
rmlv: Logical volume loglv00 is removed.
```

Use the **mk1v** command to create a jfslog file system loglv00 for the volume group volg01:

```
# mk1v -t jfslog -y loglv00 volg01 1
loglv00
```

Run the **logform** command to initialize the logical volume loglv00 for use as a JFS log:

```
# logform /dev/loglv00
logform: destroy /dev/loglv00 (y)?y
```

Check the integrity of the file system using the **fsck** command. Fix the superblock error if needed:

```
# fsck /userfs01

** Checking /dev/rlogvo101 (/userf)
** Phase 0 - Check Log
log redo processing for /dev/rlogvo101
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
```

```

** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
Superblock is marked dirty; FIX? y
1809 files 159704 blocks 167976 free
***** Filesystem was modified *****

```

Mount the file system and check that it is mounted:

```

# mount /userfs01
# mount

```

node	mounted	mounted over	vfs	date	options
/dev/hd4	/		jfs	Jun 19 14:50	rw,log=/dev/hd8
/dev/hd2	/usr		jfs	Jun 19 14:50	rw,log=/dev/hd8
/dev/hd9var	/var		jfs	Jun 19 14:50	rw,log=/dev/hd8
/dev/hd3	/tmp		jfs	Jun 19 14:50	rw,log=/dev/hd8
/dev/hd1	/home		jfs	Jun 19 14:51	rw,log=/dev/hd8
/dev/lv00	/usr/welcome_arcade		jfs	Jun 19 14:51	
rw,log=/dev/hd8					
/dev/lv01	/usr/welcome		jfs	Jun 19 14:51	rw,log=/dev/hd8
/dev/lv02	/test01		jfs	Jun 19 14:51	
rw,log=/dev/loglv01					
/dev/logvo101	/userfs01		jfs	Jun 19 16:00	
rw,log=/dev/loglv00					

5.4 Managing a paging space

A paging space is fixed disk storage working as a backing store to hold not currently accessed data in virtual memory. If some processes or kernels require this data, then it is read from the corresponding paging space disk blocks into memory (called a paged in). On AIX, a paging space is a logical volume with LV TYPE set to paging. The default paging space hd6 is created upon initial installation of AIX in the rootvg volume group, and activated automatically every time the system is booted.

This section describes the following topics.

- ▶ How to create a new paging space
- ▶ How to expand an existing paging space
- ▶ How to remove a paging space
- ▶ How to shrink an existing paging space size

5.4.1 How to create a new paging space

On AIX, creating a new paging space is a simple task, since a newly created paging space is activated at the moment of the creation. Before creating the paging space, you should examine the following points.

- ▶ How much paging space is needed for your system?

It depends on the virtual memory usage of your application. Some applications require three or four times the physical memory to work smoothly.

- ▶ Is there enough free disk space to create a new paging space on your system?

If your rootvg volume group has only 100 MB or less of free space, it is not recommended to add a new paging space in the rootvg. This amount of space is usually required for system maintenance tasks (for example, applying PTFs).

- ▶ Is there any possibility of affecting user application performance?

Creating a new paging space on a physical volume consumes certain disk I/O bandwidth on your system if the paging space is used heavily. If your application is I/O intensive and sensitive, we do not recommend creating a new paging space on the same physical volume that holds your application data. In this case, you might consider an additional disk I/O path to create a new paging space by adding a new adapter and disks.

To add a new paging space, you can use the following method. This task is also achieved by using the **mkps** command.

Use the **smit 1vm** command, then select **Paging Space** followed by **Add Another Paging Space**. When you are prompted, select the volume group you would like to add a paging space to, from the list. The results of these actions are shown in Figure 5-13 on page 124.

```

                                Add Another Paging Space

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Volume group name                rootvg
SIZE of paging space (in logical partitions) [16] #
PHYSICAL VOLUME name             +
Start using this paging space NOW? yes      +
Use this paging space each time the system is
  RESTARTED?                      yes      +

F1=Help      F2=Refresh      F3=Cancel      F4=List
Esc+5=Reset  F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Figure 5-13 The SMIT Add Another Paging Space dialog

Then press Enter to generate the paging space.

There is no field to specify the logical volume name of this newly created paging space. By default, AIX will sequentially assign the name paging00, paging01, to pagingnn for paging space. The following points are for your consideration.

- ▶ The size of the paging space is specified as a multiplication of the PP_SIZE of the volume group. In this example, it is 16 MB.
- ▶ If you want to use this paging space now, alter the Start using this paging space NOW? field to yes, as shown.
- ▶ If you want to activate this paging space across boot, alter the Use the paging space each time the system is RESTARTED? to yes, as shown in Figure 5-13.

To confirm the result of this task, please type the `lsps -a` command like this:

```

# lsps -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6        hdisk0          rootvg      1024MB 0    yes  yes  lv
paging00   hdisk0          rootvg      256MB  0    yes  yes  lv

```

5.4.2 How to expand an existing paging space

To expand an existing paging space, you can use the following method. This task is also achieved by using the `chps` command. It remains activated after expansion.

Use the `smit lvm` command and select **Paging Space** then **Change/Show Characteristics of a Paging Space**. You will be prompted to select the paging space from the list. The results of these actions are shown in Figure 5-14.

```
Change / Show Characteristics of a Paging Space

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Paging space name                 hd6
Volume group name                 rootvg
Physical volume name              hdisk0
NUMBER of additional logical partitions [16] #
Or NUMBER of logical partitions to remove [] #
Use this paging space each time the system is restarted? yes +

F1=Help      F2=Refresh      F3=Cancel      F4=List
Esc+5=Reset  F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Figure 5-14 The SMIT Change/Show Characteristics of a Paging Space dialog

Press Enter to execute the dialog.

Note that the size is specified as a difference, so if you have an existing paging space with 16 PPs (assume it corresponds to 256 MB), and you wish to have a paging space equal to 512 MB, then you should specify 16 PPs (256 MB); the result is 512 MB (256 MB + 256 MB).

To confirm the result of this task, type the `lspv -a` command as shown in the following example:

```
# lspv -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6         hdisk0          rootvg      1024MB 0    yes  yes  lv
paging00    hdisk0          rootvg      512MB  0    yes  yes  lv
```

5.4.3 How to remove an existing paging space

In AIX 5L, you can use the **swapoff** command to move all pages out of a paging space which that will be removed. After pages are moved out, use the **rmpps** command to remove the paging space. In the following example, paging00 is to be removed:

```
# swapoff /dev/paging00
# rmpps paging00
rmlv: Logical volume paging00 is removed.
```

In AIX 4.3.3 or earlier, before removing the existing paging space, all the references to this paging space should be removed. This requires the paging space to be deactivated first, and a subsequent reboot of the system. In AIX 5L and higher versions, a system reboot is no longer required to remove a paging space.

To deactivate an existing paging space, you can use the following method. This task is also achieved by using the **chpps** command.

Using the same SMIT dialog shown in Figure 5-14 on page 125, modify the “Use this paging space each time the system is RESTARTED?” field to no. Press Enter to run the dialog and reboot your system.

```
# shutdown -Fr
```

To confirm the result of this task, type the **lpps -a** command as shown in the following example.

```
# lpps -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6        hdisk0          rootvg      1024MB 0    yes  yes  lv
paging00   hdisk0          rootvg      512MB 0    no   no   lv
```

Remove this paging space using SMIT or by using the **rmpps** command. This step is not shown in this example, but was completed successfully.

To confirm the result of this task, enter the **lpps -a** command as shown in the following example.

```
# lpps -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6        hdisk0          rootvg      1024MB 0    yes  yes  lv
```

5.4.4 How to shrink an existing paging space size

With the introduction of AIX 5L, managing paging space became much simpler. You can reduce the size of paging space dynamically. For example, remove one logical partition from paging00 by running:

```
# chps -d 1 paging00
shrinks: Temporary paging space paging01 created.
shrinks: Paging space paging00 removed.
shrinks: Paging space paging00 recreated with new size.
```

The **chps** command creates the temporary paging spaces as needed in AIX 5L.

In versions prior to AIX 5L, there are no provisions to shrink an existing paging space size directly. You have to remove it, then recreate it with a smaller size, which requires a system reboot. AIX 5L has enhancements that allow for shrinking a paging space and avoids the requirements for a system reboot.

If you want to shrink the default paging space (/dev/hd6) size, then the following procedures should be followed.

1. Create a new paging space (typically, /dev/paging00) that has enough size to continue the system operation.
2. Change the default paging space /dev/hd6 to deactivate after the system reboot.
3. Modify the /sbin/rc.boot file since activating (swapon) the default paging space is hard-coded in this startup procedure script. It must activate the newly created paging space from the previous step. You should create a backup of this script first.
4. Change the default paging space using the **sysdumpdev** command.
5. Recreate the BLV to reflect the above modification.
6. Reboot the system.
7. Remove the default paging space /dev/hd6.
8. Recreate the default paging space /dev/hd6 with your specified size. You cannot use the **mkps** command in here, since it always assigns the name *pagingnn*.
9. Change the additional paging space you created in the previous step to deactivate after the system reboot (assume it is /dev/paging00).
10. Restore the original /sbin/rc.boot.
11. Change the default paging space using the **sysdumpdev** command.
12. Recreate the BLV to reflect the above modification.
13. Reboot the system.

14. Remove the additional paging space (assume it is /dev/paging00).

These previous steps are shown in the following example.

```
# mkps -a -n -s 64 rootvg
paging00
# lspv -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6         hdisk0          rootvg      2048MB 0    yes  yes  lv
paging00    hdisk0          rootvg      1024MB 1    yes  yes  lv
# chps -a n hd6
# cp /sbin/rc.boot /sbin/rc.boot.orig
# vi /sbin/rc.boot
```

Change the paging space entry in the /sbin/rc.boot file from swapon /dev/hd6 to swapon /dev/paging00.

```
# sysdumpdev -P -p /dev/paging00
primary /dev/paging00
secondary /dev/sysdumpnull
copy directory /var/adm/ras
forced copy flag TRUE
always allow dump FALSE
dump compression OFF
# bosboot -d /dev/hdisk0 -a
bosboot: Boot image is 6748 512 byte blocks.
# shutdown -Fr
# lspv -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6         hdisk0          rootvg      2048MB 0    no   no   lv
paging00    hdisk0          rootvg      1024MB 1    yes  yes  lv
# rmps hd6
rmlv: Logical volume hd6 is removed.
# mklv -b n -t paging -y hd6 rootvg 64
# swapon /dev/hd6
# chps -a n paging00
# cp -p /sbin/rc.boot.orig /sbin/rc.boot
# sysdumpdev -P -p /dev/hd6
primary /dev/hd6
secondary /dev/sysdumpnull
copy directory /var/adm/ras
forced copy flag TRUE
always allow dump FALSE
dump compression OFF
# bosboot -d /dev/hdisk0 -a
# shutdown -Fr
# rmps paging00
# lspv -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6         hdisk0          rootvg      1024MB 0    yes  yes  lv
```

5.5 Working with physical volumes

In AIX, fixed disk storage device is represented as a logical instance called a physical volume. Usually it has a name `hdiskN`, where `N` is a sequence of numbers starting from zero. The physical volume is not used directly by any applications, even if your application does raw I/O. To use fixed disk storage in AIX, you must configure a volume group, which is composed of one or more physical volumes.

5.5.1 Removing a physical disk drive from a system

When removing the disk from the system, it is done at a volume group and physical volume level.

In the following system there are two volume groups, `rootvg` and `volg01`. The `volg01` volume group has three disks assigned to it, and the disk `hdisk3` needs to be removed from the volume group. There is only one logical volume (`logvol01`) on the `volg01` volume group, and the file system is `/userfs01`. If there are logical volumes residing on the disk to be removed, use the `migratepv` command as described in 5.2.3, “Moving a logical volume from one physical volume to another” on page 103:

```
# lsvg
rootvg
volg01

# lspv
hdisk0      000bc6fdc3dc07a7   rootvg
hdisk1      000bc6fdbff75ee2   volg01
hdisk2      000bc6fdbff92812   volg01
hdisk3      000bc6fdbff972f4   volg01

# lsvg -l volg01
volg01:
LV NAME      TYPE      LPs  PPs  PVs  LV STATE      MOUNT POINT
logvol01     jfs       10   10   1    open/syncd    /userfs01
loglv00      jfslog    1    1    1    open/syncd    N/A

# mount
node         mounted      mounted over  vfs      date      options
-----
/dev/hd4     /
/dev/hd2     /usr
/dev/hd9var  /var
/dev/hd3     /tmp
/dev/hd1     /home
/dev/lv01    /usr/welcome
jfs         Jun 19 18:08 rw,log=/dev/hd8
jfs         Jun 19 18:09 rw,log=/dev/hd8
jfs         Jun 19 18:09 rw,log=/dev/hd8
```

```
                /dev/logvo101  /userfs01      jfs    Jun 19 18:09
rw,log=/dev/loglv00
```

The first step in the process is to unmount the file system:

```
# umount /userfs01
```

Remove the physical volume from the volume group:

```
# reducevg -df volg01 hdisk3
```

Ensure the physical volume is no longer associated with a volume group:

```
# lspv
hdisk0          000bc6fdc3dc07a7    rootvg
hdisk1          000bc6fdbff75ee2    volg01
hdisk2          000bc6fdbff92812    volg01
hdisk3          000bc6fdbff972f4    None
```

Remove the disk from the system:

```
# rmdev -d -l hdisk3
hdisk3  deleted
```

Ensure that the disk has been removed:

```
# lsdev -Ccdisk
hdisk0 Available 30-58-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 30-58-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
```

Mount the file system to make the system available for use:

```
# mount /userfs01
# mount
```

node	mounted	mounted over	vfs	date	options
/dev/hd4	/		jfs	Jun 19 18:08	rw,log=/dev/hd8
/dev/hd2	/usr		jfs	Jun 19 18:08	rw,log=/dev/hd8
/dev/hd9var	/var		jfs	Jun 19 18:08	rw,log=/dev/hd8
/dev/hd3	/tmp		jfs	Jun 19 18:08	rw,log=/dev/hd8
/dev/hd1	/home		jfs	Jun 19 18:09	rw,log=/dev/hd8
/dev/logvo101	/userfs01		jfs	Jun 20 09:33	

```
rw,log=/dev/loglv00
```

5.5.2 Recovering an incorrectly removed disk

If a disk needs to be recovered that was incorrectly removed from the system, and the system was rebooted, the **sync1vadm** command will need to be run.

This is an advanced topic condensed to demonstrate the requirement to use several commands to recover from incorrectly removed disks. It is possible that customer situations will require several more steps than what is shown in this sample.

In the following example, a disk has been incorrectly removed from the system, and the logical volume control block needs to be rebuilt.

The disks in the system before the physical volume was removed were:

```
# lsdev -Cc disk
hdisk0 Available 30-58-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 30-58-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk3 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
```

The allocation of the physical volumes before the disk was removed was:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk2      000bc6fdbff92812    volg01
hdisk3      000bc6fdbff972f4    volg01
```

The logical volumes on the volume group are:

```
# lsvg -l volg01
volg01:
LV NAME      TYPE      LPs  PPp  PVs  LV STATE      MOUNT POINT
logvol01     jfs       1000 1000  2    open/syncd    /userfs01
loglv00      jfslog    1     1    1    open/syncd    N/A
```

The logical volume distribution on the physical volumes is:

```
# lslv -l logvol01
logvol01:/userfs01
PV          COPIES      IN BAND      DISTRIBUTION
hdisk1      542:000:000 19%          109:108:108:108:109
hdisk3      458:000:000 23%          109:108:108:108:025
```

After a reboot, the system disk configuration appears as follows:

```
# lspv
hdisk0      000bc6fdc3dc07a7    rootvg
hdisk1      000bc6fdbff75ee2    volg01
hdisk3      000bc6fdbff972f4    volg01
```

When trying to mount the file system on the logical volume, an error may appear as follows:

```
# mount /userfs01
```

mount: 0506-324 Cannot mount /dev/logvol01 on /userfs01: There is an input or output error.

To get the logical volume synchronized correctly, the following needs to be done:

```
# syncldm -v volg01
syncldm: Physical volume data updated.
syncldm: Logical volume logvol01 updated.
syncldm: Warning, lv control block of loglv00 has been over written.
0516-622 syncldm: Warning, cannot write lv control block data.
syncldm: Logical volume loglv00 updated.
```

The system can now be repaired. If the file system data was spread across all the disks, including the failed disk, it may need to be restored from the last backup.

5.5.3 General tips on boot problems

If, for some inexplicable reason, the system will not boot and all the errors have been checked, ensure that the following have been checked:

- ▶ Check that any changes made in the `/etc/inittab` file are not causing the boot problem.
- ▶ Check that any changes made in the `/etc/environment` file are not causing the boot problem.
- ▶ Check that the root file system or one of the other file systems in the rootvg volume group have sufficient space.
- ▶ Check that a log file is not full or corrupt, thus not allowing the system to boot.
- ▶ Check the `/etc/filesystems` file to ensure that the changes made are not causing the system boot to fail.
- ▶ Make sure any mirrored disks have a boot image and are contained in the bootlist. Use the `bosboot` and `bootlist` commands.

5.6 Command summary

The following section provides a list of the key commands discussed in this chapter. For a complete reference of the following commands, consult the AIX product documentation.

5.6.1 The `cfgmgr` command

The `cfgmgr` command configures devices and optionally installs device software into the system.

The **cfgmgr** command syntax is as follows:

```
cfgmgr [ -f | -s | -p Phase ] [ -i Device ] [ -l Name ] [ -v ]
```

The commonly used flags are provided in Table 5-1.

Table 5-1 Commonly used flags of the **cfgmgr** command

Flag	Description
-f	Specifies that the cfgmgr command executes the phase 1 configuration rules. This flag is not valid at run time (after system start).
-i Device	Specifies the location of the installation medium.
-l Name	Specifies the named device to configure along with its children.
-p Phase	Specifies that the cfgmgr command executes the specified phase.
-s	Specifies that the cfgmgr command executes the phase 2 configuration rules.
-v	Specifies verbose output. The cfgmgr command writes information about what it is doing to standard output.

5.6.2 The **bosboot** command

The **bosboot** command creates the boot image that interfaces with the machine boot Read-Only Storage (ROS) Erasable Programmable Read-Only Memory (EPROM).

The **bosboot** command syntax is as follows.

For general use:

```
bosboot -Action [ -d Device ] [ -Options ... ]
```

To create a device boot image:

```
bosboot -a [ -d Device ] [ -p Proto ] [ -k Kernel ] [ -U ] [ -I | -D ]  
[ -l LVdev ] [-L] [ -M { Norm | Serv | Both } ] [ -O Number ] [ -T Type ]  
[ -b FileName ] [ -q ]
```

To copy a device boot image:

```
bosboot -w FileName [ -d Device ] [ -q ]
```

To create a ROS emulation boot image:

```
bosboot -r FileName [ -d Device ] [ -l LVdev ] [ -M { Norm | Serv | Both } ]  
[ -q ]
```

The commonly used flags are provided in Table 5-2.

Table 5-2 Commonly used flags of the bosboot command

Flag	Description
-a	Creates complete boot image and device.
-d Device	Specifies the boot device. This flag is optional for hard disk.

5.6.3 The bootlist command

The **bootlist** command allows the user to display and alter the list of possible boot devices from which the system may be booted. When the system is booted, it will scan the devices in the list and attempt to boot from the first device it finds containing a boot image.

The **bootlist** command syntax is as follows:

```
bootlist [ { -m Mode } [ -r ] [ -o ] [ [ -i ] | [ [ -f File ]
[ Device [ Attr=Value ... ] ... ] ] ] ]
```

The commonly used flags are provided in Table 5-3.

Table 5-3 Commonly used flags of the bootlist command

Flag	Description
<i>Device</i>	Provides the names of the specific or generic devices to include in the boot list.
-f File	Indicates that the device information is to be read from the specified file name.
-i	Indicates that the device list specified by the -m flag should be invalidated.
-m Mode	Specifies which boot list to display or alter. Possible values for the mode variable are normal, service, both, or prevboot.
-o	Indicates that the specified boot list is to be displayed after any specified alteration is performed. The output is a list of device names. This flag applies only to AIX Version 4.2.1 or higher.
-r	Indicates that the specified boot list is to be displayed after any specified alteration is performed. The output is hardware-platform dependent. It may be a hexadecimal dump of the boot list or a list of device names. This is normally used for problem determination.

5.6.4 The **chps** command

The **chps** command changes attributes of a specific paging space. The `PagingSpace` parameter specifies the name of the paging space to be changed.

The **chps** command syntax is as follows:

```
chps [ -s LogicalPartitions ] [ -a { y | n } ] PagingSpace
```

The commonly used flags are provided in Table 5-4.

*Table 5-4 Commonly used flags of the **chps** command*

Flag	Description
-a	Specifies to use a paging space at the next system restart.
y	Specifies that the paging space is active at subsequent system restarts.
n	Specifies that the paging space is inactive at subsequent system restarts.
-s LogicalPartitions	Specifies the number of logical partitions to add.

5.6.5 The **cp1v** command

The **cp1v** command copies the contents of a logical volume to a new logical volume.

The **cp1v** command syntax is as follows.

To copy to a new logical volume:

```
cp1v [ -v VolumeGroup ] [ -y NewLogicalVolume | -Y Prefix ] SourceLogicalVolume
```

To copy to an existing logical volume:

```
cp1v -e DestinationLogicalVolume [ -f ] SourceLogicalVolume
```

The commonly used flags are provided in Table 5-5 on page 136.

Table 5-5 Commonly used flags of the `cplv` command

Flag	Description
-e	Specifies that the DestinationLogicalVolume exists and that a new logical volume should not be created. If the DestinationLogicalVolume is smaller than the SourceLogicalVolume, the extra logical partitions are not copied. When you use this flag, any data already in the DestinationLogicalVolume is destroyed. For this reason, user confirmation is required unless the -f flag is added. The type characteristic of the DestinationLogicalVolume must be copy to prevent inadvertently overwriting data. To change the type characteristic, use the <code>chlv</code> command.
-f	Copies to an existing logical volume without requesting user confirmation.
-v VolumeGroup	Specifies the volume group where the new logical volume resides. If this is not specified, the new logical volume resides in the same volume group as the SourceLogicalVolume.
-y NewLogicalVolume	Specifies the name to use, in place of a system-generated name, for the new logical volume. Logical volume names must be unique system-wide names and can range from one to 15 characters.
-Y Prefix	Specifies a prefix to use in building a system-generated name for the new logical volume. The prefix must be less than or equal to 13 characters. A name cannot begin with a prefix already defined in the PdDv class in the Device Configuration Database for other devices or a name already used by another device.

5.6.6 The `fsck` command

The `fsck` command checks and interactively repairs inconsistent file systems. You should run this command before mounting any file system. You must be able to read the device file on which the file system resides. Normally the file system is consistent, and the `fsck` command merely reports on the number of files, used blocks, and free blocks in the file system. If the file system is inconsistent, the `fsck` command displays information about the inconsistencies found and prompts you for permission to repair them when the -y flag is not used.

Note: If you run the **fsck** command on a mounted file system, it might report the following messages:

```
Bad Inode Map (NOT SALVAGED)
Bad Block Map (NOT SALVAGED)
```

This is a normal situation (you cannot determine if this file system is corrupted or not from this evidence only), especially on a frequently written file system such as `/var`. You can see the **fsck** command warning about this in the following:

```
Checking a mounted filesystem does not produce dependable results.
```

To determine whether this file system is corrupted, you have to unmount the file system, then run the **fsck** command on it.

The **fsck** command syntax is as follows:

```
fsck [ -n ] [ -p ] [ -y ] [ -dBlockNumber ] [ -f ] [ -ii-NodeNumber ]
[ -o Options ] [ -tFile ] [ -V VfsName ] [ FileSystem1 - FileSystem2 ... ]
```

The commonly used flags are provided in Table 5-6.

Table 5-6 Commonly used flags of the fsck command

Flag	Description
-dBlockNumber	Searches for references to a specified disk block. Whenever the fsck command encounters a file that contains a specified block, it displays the i-node number and all path names that refer to it.
-f	Performs a fast check. Under normal circumstances, the only file systems likely to be affected by halting the system without shutting down properly are those that are mounted when the system stops. The -f flag prompts the fsck command to not check file systems that were unmounted successfully. The fsck command determines this by inspecting the <code>s_fmmod</code> flag in the file system superblock. This flag is set whenever a file system is mounted and cleared when it is unmounted successfully. If a file system is unmounted successfully, it is unlikely to have any problems. Because most file systems are unmounted successfully, not checking those file systems can reduce the checking time.
-ii-NodeNumber	Searches for references to a specified i-node. Whenever the fsck command encounters a directory reference to a specified i-node, it displays the full path name of the reference.
-n	Assumes a no response to all questions asked by the fsck command. It does not open the specified file system for writing.

Flag	Description
-o Options	<p>Passes comma-separated options to the fsck command. These options are assumed to be file system implementation-specific, except that the following are currently supported for all file systems:</p> <ul style="list-style-type: none"> ▶ mountable Causes the fsck command to exit with success, returning a value of 0, if the file system in question is mountable (clean). If the file system is not mountable, the fsck command exits returning with a value of 8. ▶ mytype Causes the fsck command to exit with success (0) if the file system in question is of the same type as either specified in the <code>/etc/filesystems</code> file or by the <code>-V</code> flag on the command line. Otherwise, 8 is returned. For example, fsck -o mytype -V jfs / exits with a value of 0 if / (the root file system) is a journaled file system.
-p	<p>Does not display messages about minor problems but fixes them automatically. This flag does not grant the wholesale license that the <code>-y</code> flag does and is useful for performing automatic checks when the system is started normally. You should use this flag as part of the system startup procedures whenever the system is being run automatically. Also allows parallel checks by group. If the primary superblock is corrupt, the secondary superblock is verified and copied to the primary superblock.</p>
-tFile	<p>Specifies a file parameter as a scratch file on a file system other than the one being checked if the fsck command cannot obtain enough memory to keep its tables. If you do not specify the <code>-t</code> flag and the fsck command needs a scratch file, it prompts you for the name of the scratch file. However, if you have specified the <code>-p</code> flag, the fsck command is unsuccessful. If the scratch file is not a special file, it is removed when the fsck command ends.</p>
-V VfsName	<p>Uses the description of the virtual file system specified by the <code>VFSName</code> variable for the file system instead of using the <code>/etc/filesystems</code> file to determine the description. If the <code>-V VfsName</code> flag is not specified on the command line, the <code>/etc/filesystems</code> file is checked and the <code>vfs=Attribute</code> of the matching stanza is assumed to be the correct file system type.</p>
-y	<p>Assumes a yes response to all questions asked by the fsck command. This flag lets the fsck command take any action it considers necessary. Use this flag only on severely damaged file systems.</p>

Flag	Description
-C	Specifies that all future dumps will be compressed before they are written to the dump device. The -C flag applies to only AIX Version 4.3.2 and higher versions.
-c	Specifies that dumps will not be compressed. The -c flag applies to only AIX Version 4.3.2 and higher versions.

5.6.7 The fuser command

The **fuser** command lists the process numbers of local processes that use the local or remote files specified by the file parameter. For block special devices, the command lists the processes that use any file on that device.

The **fuser** command syntax is as follows:

```
fuser [ -c ] [ -f ] [ -k ] [ -u ] File ...
```

Each process number is followed by a letter indicating how the process uses the file:

- c** Uses the file as the current directory
- p** Uses the file as the parent of the current directory (only when in use by the system)
- r** Uses the file as the root directory

The commonly used flags are provided in Table 5-7.

Table 5-7 Commonly used flags of the fuser command

Flag	Description
-k	Sends the SIGKILL signal to each local process. Only the root user can kill a process of another user.
-u	Provides the login name for local processes in parentheses after the process number.
-c	Reports on any open files in the file system containing File.
-f	Reports on open instances of file only.

Note: The -c and -f flags apply to AIX Version 4.3 and higher releases.

5.6.8 The migratepv command

The **migratepv** command moves allocated physical partitions from one physical volume to one or more other physical volumes.

The **migratepv** command syntax is as follows:

```
migratepv [ -i ] [ -l LogicalVolume ] SourcePhysicalVolume  
DestinationPhysicalVolume ...
```

The commonly used flags are provided in Table 5-8.

Table 5-8 Commonly used flags of the migratepv command

Flag	Description
-i	Reads the DestinationPhysicalVolume parameter from standard input.
-l LogicalVolume	Moves only the physical partitions allocated to the specified logical volume and located on the specified source physical volume.

5.6.9 The mkboot command

The **mkboot** command creates the boot image, the boot record, and the service record. This command is *not* a user-level command and is not supported in AIX Version 4.2 or higher.

The **mkboot** command syntax is as follows:

```
mkboot -d Device [ -b ] [ -D ] [ -c ] [ -h ] [ -i ] [ -I ] [ -l LVDev ]  
{ -k Kernel | -e Expander } [-L] [-s] [-r] [-p Offset] [-w ]  
-f FileSystem
```

The commonly used flags are provided in Table 5-9.

Table 5-9 Commonly used flags of the mkboot command

Flag	Description
-c	Zeros out the boot record on the device. This flag is optional.
-d Device	Specifies the device required for the IPL record. This flag is required.

5.6.10 The mklv command

The **mklv** command creates a new logical volume within a volume group.

The **mk1v** command syntax is as follows:

```
mk1v [ -a Position ] [ -b BadBlocks ] [ -c Copies ] [ -d Schedule ]
[ -e Range ] [ -i ] [ -L Label ] [ -m MapFile ] [ -r Relocate ]
[ -s Strict ] [ -t Type ] [ -u UpperBound ] [ -v Verify ]
[ -w MirrorWriteConsistency ] [ -x Maximum ] [ -y NewLV | -Y Prefix ]
[ -S StripeSize ] VG Number [ PV ]
```

The commonly used flags are provided in Table 5-10.

Table 5-10 Commonly used flags of the **mk1v** command

Flag	Description
-a Position	<p>Sets the intra-physical volume allocation policy (the position of the logical partitions on the physical volume). The position variable can be one of the following:</p> <ul style="list-style-type: none"> ▶ <i>m</i>: Allocates logical partitions in the outer middle section of each physical volume. This is the default position. ▶ <i>c</i>: Allocates logical partitions in the center section of each physical volume. ▶ <i>e</i>: Allocates logical partitions in the outer edge section of each physical volume. ▶ <i>ie</i>: Allocates logical partitions in the inner edge section of each physical volume. ▶ <i>im</i>: Allocates logical partitions in the inner middle section of each physical volume.
-b BadBlocks	<p>Sets the bad-block relocation policy. The relocation variable can be one of the following:</p> <ul style="list-style-type: none"> ▶ <i>y</i>: Causes bad-block relocation to occur. This is the default. ▶ <i>n</i>: Prevents bad-block relocation from occurring.
-c Copies	<p>Sets the number of physical partitions allocated for each logical partition. The copies variable can be set to a value from 1 to 3; the default is 1.</p>
-d Schedule	<p>Sets the scheduling policy when more than one logical partition is written. The schedule variable can be one of the following:</p> <ul style="list-style-type: none"> ▶ <i>p</i>: Establishes a parallel scheduling policy. This is the default for scheduling policy. ▶ <i>s</i>: Establishes a sequential scheduling policy.

Flag	Description
-e Range	<p>Sets the inter-physical volume allocation policy (the number of physical volumes to extend across using the volumes that provide the best allocation). The range value is limited by the UpperBound variable (set with the -u flag), and can be one of the following:</p> <ul style="list-style-type: none"> ▶ <i>x</i>: Allocates across the maximum number of physical volumes. ▶ <i>m</i>: Allocates logical partitions across the minimum number of physical volumes. This is the default range.
-i	<p>Reads the PhysicalVolume parameter from standard input. Use the -i flag only when PhysicalVolume is entered through standard input.</p>
-L	<p>Sets the logical volume label. The default label is None. The maximum size of the label file is 127 characters.</p> <p>Note: If the logical volume is going to be used as a journaled file system (JFS), then the JFS will use this field to store the mount point of the file system on that logical volume for future reference.</p>
-m MapFile	<p>Specifies the exact physical partitions to allocate. Partitions are used in the order given in the MapFile parameter. Used partitions in the MapFile parameter are not legal since the new logical volume cannot occupy the same physical space as a previously allocated logical volume. All physical partitions belonging to a copy are allocated before allocating for the next copy of the logical volume. The MapFile parameter format is:</p> <ul style="list-style-type: none"> ▶ <i>PVname:PPnum1[-PPnum2]</i>: In this example, <i>PVname</i> is a physical volume name (for example, hdisk0) as specified by the system. It is one record per physical partition or a range of consecutive physical partitions. <i>PPnum</i> is the physical partition number. ▶ <i>PVname</i>: Name of the physical volume as specified by the system. ▶ <i>PPnum</i>: Physical partition number.
-r Relocate	<p>Sets the reorganization relocation flag. For striped logical volumes, the relocate parameter must be set to n (the default for striped logical volumes). The relocate parameter can be one of the following:</p> <ul style="list-style-type: none"> ▶ <i>y</i>: Allows the logical volume to be relocated during reorganization. This is the default for relocation. ▶ <i>n</i>: Prevents the logical volume from being relocated during reorganization.

Flag	Description
-s Strict	<p>Determines the strict allocation policy. Copies of a logical partition can be allocated to share or not to share the same physical volume. The strict parameter is represented by one of the following:</p> <ul style="list-style-type: none"> ▶ <i>y</i>: Sets a strict allocation policy, so copies for a logical partition cannot share the same physical volume. This is the default for allocation policy. ▶ <i>n</i>: Does not set a strict allocation policy, so copies for a logical partition can share the same physical volume.
-S StripeSize	<p>Specifies the number of bytes per striped. Must be a power of two, between 4 KB and 128 KB (for example 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, or 128 KB).</p>
-t Type	<p>Sets the logical volume type. The standard types are <i>jfs</i> (file systems), <i>jfslog</i> (journal file system logs), and <i>paging</i> (paging spaces), but a user can define other logical volume types with this flag. You cannot create a striped logical volume of type <i>boot</i>. The default is <i>jfs</i>. If a log is manually created for a file system, the user must run the logform command to clean out the new <i>jfslog</i> before the log can be used. For example, to format the logical volume <i>logdev</i>, enter:</p> <pre>logform /dev/logdev</pre> <p>Where <i>/dev/logdev</i> is the absolute path to the logical volume.</p>
-u UpperBound	<p>Sets the maximum number of physical volumes for new allocation. The value must be between one and the total number of physical volumes. The default is the total number of physical volumes in the volume group. When used to create a striped logical volume using the <i>-S</i> flag, the <i>-u</i> flag sets the number of physical volumes being striped across.</p>
-v Verify	<p>Sets the write-verify state for the logical volume. Causes (<i>y</i>) all writes to the logical volume to either be verified with a follow-up read or prevents (<i>n</i>) the verification of all writes to the logical volume. The verify parameter is represented by one of the following:</p> <ul style="list-style-type: none"> ▶ <i>n</i>: Prevents the verification of all write operations to the logical volume. This is the default for the <i>-v</i> flag. ▶ <i>y</i>: Causes the verification of all write operations to the logical volume.
-w MirrorWriteCo nsistency	<ul style="list-style-type: none"> ▶ <i>y</i>: Turns on mirror write consistency, which insures data consistency among mirrored copies of a logical volume during normal I/O processing. ▶ <i>n</i>: No mirror write consistency (see the <i>-f</i> flag of the syncvg command).

Flag	Description
-x Maximum	Sets the maximum number of logical partitions that can be allocated to the logical volume. The default value is 512. The number represented by the number parameter must be equal to or less than the number represented by the maximum variable. The maximum number of logical partitions per logical volume is 32,512.
-y NewLogicalVolume	Specifies the logical volume name to use instead of using a system-generated name. Logical volume names must be unique system-wide names and can range from one to 15 characters. If the volume group is varied on in concurrent mode, the new name should be unique across all the concurrent nodes the volume group is varied on. The name cannot begin with a prefix already defined in the PdDv class in the Device Configuration Database for other devices.
-Y Prefix	Specifies the prefix to use instead of the prefix in a system-generated name for the new logical volume. The prefix must be less than or equal to 13 characters. The name cannot begin with a prefix already defined in the PdDv class in the Device Configuration Database for other devices, nor be a name already used by another device.

Note: The -c, -d, -e, -m, -s, and -w flags are not valid when creating a striped logical volume using the -S flag.

5.6.11 The mkps command

The **mkps** command adds additional paging space to the system.

The **mkps** command syntax is as follows.

To add a logical volume for additional paging space:

```
mkps [ -a ] [ -n ] [ -t lv ] -s LogicalPartitions VolumeGroup
[ PhysicalVolume ]
```

To add additional paging space on an NFS server:

```
mkps [ -a ] [ -n ] -t nfs ServerHostName ServerFileName
```

The commonly used flags are provided in Table 5-11.

Table 5-11 Commonly used flags of the mkps command

Flag	Description
-a	Specifies that the paging space is configured at subsequent restarts.

Flag	Description
-n	Activates the paging space immediately.
-s LogicalPartitions	Specifies the size of the paging space and the logical volume to be made in logical partitions.
-t	Specifies the type of paging space to be created. One of the following variables is required: <ul style="list-style-type: none"> ▶ <i>lv</i>: Specifies that a paging space of type logical volume should be created on the system. ▶ <i>nfs</i>: Specifies that a paging space of type NFS should be created on the system.

5.6.12 The `replacepv` command

The `replacepv` command replaces a physical volume in a volume group with another physical volume.

The `replacepv` command syntax is as follows:

```
replacepv [ -f ] {SourcePhysicalVolume | SourcePhysicalVolumeID }
DestinationPhysicalVolume
replacepv [ -R ] dir_name [ DestinationPhysicalVolume ]
```

The commonly used flags are provided in Table 5-12.

Table 5-12 Commonly used flags of the `replacepv` command

Flag	Description
-f	Forces to replace a SourcePhysicalVolume with the specified DestinationPhysicalVolume unless the DestinationPhysicalVolume is part of another volume group in the Device Configuration Database or a volume group that is active.
-R dir_name	Recovers <code>replacepv</code> if it is interrupted by Ctrl+c, a system crash, or a loss of quorum. When using the -R flag, you must specify the directory name given during the initial run of <code>replacepv</code> . This flag also allows you to change the DestinationPhysicalVolume.

5.6.13 The `restvg` command

The `restvg` command restores the user volume group and all its containers and files as specified in the `/tmp/vgdata/vgname/vgname.data` file (where `vgname` is the name of the volume group) contained within the backup image created by the `savevg` command.

The **restvg** command restores a user volume group. The bosinstall routine reinstalls the root volume group (rootvg). If the **restvg** command encounters a rootvg volume group in the backup image, the **restvg** command exits with an error.

If a yes value has been specified in the EXACT_FIT field of the logical_volume_policy stanza of the /tmp/vgdata/vgname/vgname.data file, the **restvg** command uses the map files to preserve the placement of the physical partitions for each logical volume. The target disks must be of the same size or larger than the source disks specified in the source_disk_data stanzas of the vgname.data file.

The **restvg** command syntax is as follows:

```
restvg [ -b Blocks ] [ -f Device ] [ -q ] [ -s ] [-n ] [-p PPsize ]
[ DiskName ]
```

The commonly used flags are provided in Table 5-13.

Table 5-13 Commonly used flags of the restvg command

Flag	Description
-b Blocks	Specifies the number of 512-byte blocks to read in a single input operation. If this parameter is not specified the default of 100 is used by the restore command. Larger values result in larger physical transfers to tape devices.
-f Device	Specifies the device name of the backup media. The default is /dev/rmt0.
-q	Specifies that the usual prompt not be displayed before the restoration of the volume group image. If this flag is not specified, the prompt displays the volume group name and the target disk-device names.
-s	Specifies that the logical volumes be created at the minimum size possible to accommodate the file systems. This size is specified by the value of the LV_MIN_LPS field of the lv_data stanza of the vgname.data file (where vgname is the name of the volume group). The -s flag overrides the values of the SHRINK and EXACT_FIT fields in the logical_volume_policy stanza of the vgname.data file. The -s flag causes the same effect as values of SHRINK=yes and EXACT_FIT=no would cause.
-n	Specifies that the existing MAP files are ignored. The -n flag overrides the value of the EXACT_FIT field in the logical_volume_policy stanza of the vgname.data file.

Flag	Description
-p PPSize	Specifies the number of megabytes in each physical partition. If not specified, <code>restvg</code> uses the best value for the PPSize, dependent upon the largest disk being restored to. If this is not the same as the size specified in the <code>vname.data</code> file, the number of partitions in each logical volume will be appropriately altered with respect to the new PPSize. If a PPSize is specified that is smaller than appropriate for the disk sizes, the larger PPSize will be used. If a PPSize is specified that is larger than appropriate for the disk sizes, the specified larger PPSize will be used.
<i>DiskName</i>	Specifies the names of disk devices to be used instead of the disk devices listed in the <code>vname.data</code> file. Target disk devices must be defined as empty physical volumes; that is, they must contain a physical volume identifier and must not belong to a volume group. If the target disk devices are new, they must be added to the system using the <code>mkdev</code> command. If the target disk devices belong to a volume group, they must be removed from the volume group using the <code>reducevg</code> command.

5.6.14 The `sysdumpdev` command

The `sysdumpdev` command changes the primary or secondary dump device designation in a system that is running.

The `sysdumpdev` command syntax is as follows:

```
sysdumpdev -P { -p Device | -s Device } [ -q ]
sysdumpdev [ -p Device | -s Device ] [ -q ]
sysdumpdev [ -d Directory | -D Directory | -e | [ -k | -K ] | -l | -L | -p
Device | -q | -r Host:Path | -s Device | -z ]
```

The commonly used flags are provided in Table 5-14.

Table 5-14 Commonly used flags of the `sysdumpdev` command

Flag	Description
-d Directory	Specifies the directory the dump is copied to at system boot. If the copy fails at boot time, the <code>-d</code> flag ignores the system dump.

Flag	Description
-D Directory	<p>Specifies the Directory the dump is copied to at system boot. If the copy fails at boot time, using the -D flag allows you to copy the dump to an external media.</p> <p>Note: When using the -d Directory or -D Directory flags, the following error conditions are detected:</p> <ul style="list-style-type: none"> ▶ Directory does not exist. ▶ Directory is not in the local journaled file system. ▶ Directory is not in the rootvg volume group.
-e	Estimates the size of the dump (in bytes) for the current running system.
-k	Requires the key mode switch to be in the service position before a dump can be forced with the reset button or the dump key sequences. This is the default setting.
-K	<p>The reset button or the dump key sequences will force a dump with the key in the normal position or on a machine without a key mode switch.</p> <p>Note: On a machine without a key mode switch, a dump cannot be forced with the reset button nor the key switch without this value set.</p>
-l	Lists the current value of the primary and secondary dump devices, copy directory, and forcecopy attribute.
-L	Displays statistical information about the most recent system dump. This includes date and time of last dump, number of bytes written, and completion status.
-P	Makes permanent the dump device specified by -p or -s flags. The -P flag can only be used with the -p or -s flags.
-p Device	Temporarily changes the primary dump device to the specified device. The device can be a logical volume or a tape device. For a network dump, the device can be a host name and a path name.
-q	Suppresses all messages to standard output. If this flag is used with the -l, -r, -z, or -L flag, the -q command will be ignored.
-r Host:Path	Frees space used by the remote dump file on server host. The location of the dump file is specified by the path.
-s Device	Temporarily changes the secondary dump device to the specified device. The device can be a logical volume or a tape device. For a network dump, the device can be a host name and a path name.

Flag	Description
-z	Determines if a new system dump is present. If one is present, a string containing the size of the dump in bytes and the name of the dump device will be written to standard output. If a new system dump does not exist, nothing is returned. After the <code>sysdumpdev -z</code> command is run on an existing system dump, the dump will no longer be considered recent.

5.7 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

- There are two disks, `hdisk0` and `hdisk1`, in the same volume group. In order to move one of the logical volumes, `hd9var`, from `hdisk0` to `hdisk1`, which of the following commands should be used?
 - `cp hd9var hdisk1`
 - `chlv hd9var hdisk0 hdisk1`
 - `chpv hd9var hdisk0 hdisk1`
 - `migratepv -l hd9var hdisk0 hdisk1`
- Which of the following statements best describes AIX journaled file systems (JFS) prior to AIX Version 4.3.0?
 - Data access time is improved by JFS logging.
 - A JFS log records file system metadata.
 - JFS logging is optional for file systems.
 - JFS logging provides database transaction logging.
- Two disks, `hdisk0` and `hdisk1`, are both included in the root volume group. The following information shows the current contents of `hdisk1`:

```
LV NAME  LPs PPs  DISTRIBUTION  MOUNT POINT
hd1      1   1    01..00..00..00..00/home
hd3      8   8    00..00..00..08..00/tmp
```

Which of the following commands should be used to move all data from `hdisk1` to `hdisk0`?

- `migratepv hdisk1 hdisk0`
- `cplv for each logical volume`
- `migratepv -l hd3 hdisk1 hdisk0`
- `mk1vcopy for each logical volume contained on hdisk1`

4. A user has accidentally removed the boot logical volume from the system. In order to recreate the boots logical volume, which of the following procedures should be performed?
 - A. Create hd5 on a physical partition in the rootvg.
 - B. Create a logical volume called hd5 within the rootvg, and reboot the system.
 - C. Create a file system mounted on /blv and copy the /unix file into the /blv directory.
 - D. Reconfigure a logical volume called hd5 on hdisk0, with the same size as the root file system.
5. A system administrator has a system where the file systems are much larger than they need to be. In order to allow the file systems to be created with the smallest size when the backup is restored, which of the following procedures should occur before or after performing the **mksysb** system backup?
 - A. Before performing the **mksysb** backup, edit the image.data file and set EXACT_FIT=yes.
 - B. Before performing the **mksysb** backup, edit the bosinst.data file and set INSTALL_METHOD=reduce.
 - C. Before performing the **mksysb** backup, edit the bosinst.data file and set EXISTING_SYSTEM_OVERWRITE=yes.
 - D. After booting from the **mksysb**, choose the menu option to shrink file systems.
6. A system administrator must perform disaster recovery on a smaller system. The original rootvg drive was 9 GB and is being replaced with a 4.5 GB drive. Which of the following procedures should be performed before the backup to make sure that the **mksysb** restores on the 4.5 GB disk?
 - A. Edit the image.data file and set FS_COMPRESS=yes for each file system.
 - B. Edit the image.data file and set UPPER_BOUND=4 for each logical volume.
 - C. Edit the image.data file and change the PP_SIZE for each logical volume to 8 MB.
 - D. Edit the image.data file and divide the logical partition number for each logical volume by 2.

7. A volume group, `appvg`, on `hdisk4` contains all of the application data. The disk that contains this volume group must be moved to another system. Which of the following groups of commands, in the order specified, will remove the disk and volume group definitions from the current system?
- A. `exportvg appvg / varyoffvg appvg / rmdev -l hdisk4`
 - B. `reducevg appvg / varyoffvg appvg / exportvg appvg`
 - C. `reducevg appvg / exportvg appvg / varyoffvg appvg`
 - D. `varyoffvg appvg / exportvg appvg / rmdev -l hdisk4 -d`
8. The disk, `hdisk4`, contains the volume group, `appvg`. The disk is physically attached to the system. Now the disk and volume group must be configured so they are available to the system. Which of the following group of commands, in the order given, should be run to successfully complete the configuration?
- A. `mkdev / mkvg / varyonvg`
 - B. `cfgmgr / importvg / varyonvg`
 - C. `cfgmgr / varyonvg / importvg`
 - D. `extendvg / importvg / varyonvg`
9. A system administrator is trying to develop a logical volume strategy based on availability, performance, and cost that is appropriate to the site. Which of the following statements is *not* true?
- A. Performance is the average speed at which data is accessed. Policies such as write-verify and mirroring add to system processing load, and thus degrade performance.
 - B. Specifying mirrored logical volumes with a parallel-scheduling policy may improve I/O read-operation performance because multiple copies allow the system to direct the read operation to the copy that can be most quickly accessed.
 - C. The parallel-scheduling policy starts the write operation for all the physical partitions in a logical partition at the same time. When the write operation to a partition that takes the longest to finish is complete, the write operation is considered complete.
 - D. The sequential-scheduling policy performs writes to multiple copies or mirrors in order. In sequential scheduling, the physical partitions are written to in sequence. The system does not wait for the write operation to complete before starting the write operation for the next one.

5.7.1 Answers

The following are the preferred answers to the questions provided in this section.

1. D
2. B
3. A
4. A
5. D
6. C
7. D
8. B
9. D

5.8 Exercises

The following exercises provide sample topics for self study. They will help ensure comprehension of this chapter.

1. On a test system, migrate a logical volume from one disk to another.
2. On a test system, create a boot image.
3. On a test system, reduce a volume group.
4. Create a new JFS log on a new hdisk.
5. Import an old disk from another system into the current system.



System backup and recovery

The following topics are discussed in this chapter:

- ▶ Backup procedures
- ▶ Restore procedures
- ▶ Troubleshooting backup and restore procedures

At most sites, information stored on computers is worth more than the computers themselves. Protecting this information is one of the system administrator's most important tasks.

6.1 Backup

This section will cover backup and restore procedures of your data. When do you need backups? When your worst fear happens, and you have lost all of your data, backups will help you to survive. Keep it in your mind and make frequent backups.

6.1.1 Backup of the rootvg volume group

In AIX, the files and directories are divided into two categories, for system and user. For the system, most AIX components (BOS, LPPs, miscellaneous configuration files, and so on) are held in the rootvg volume group. This category is not changed frequently except for the /var file system contents (usually hold the system and application logs). For the user data, it could be held in either the rootvg volume group or the user-defined volume groups. This category is changed frequently, since the user applications frequently change the data each day or each hour.

It is a good idea to place user category data in a volume group other than the rootvg volume group. This will provide a simple backup operation policy for the rootvg volume group. Frequent backups of the separate volume groups should be created.

To back up the contents of the rootvg volume group, you can use the **mksysb** command, commonly used for creating a bootable tape.

Features of **mksysb** backup are:

- ▶ Backup of the rootvg group only.
- ▶ Backs up only mounted file systems.
- ▶ The image backup is performed using the **backup** command.
- ▶ Provides facilities for non-interactive install.
- ▶ Saves system paging space definitions.
- ▶ Saves logical volumes polices.

AIX Version 3 uses the **tar** command during a **mksysb** for rootvg.

6.1.2 The data layout of a mksysb tape

The **mksysb** tape has a special layout designed for easy BOS installation and recovery. The layout of **mksysb** tape is shown in Figure 6-1 on page 155.

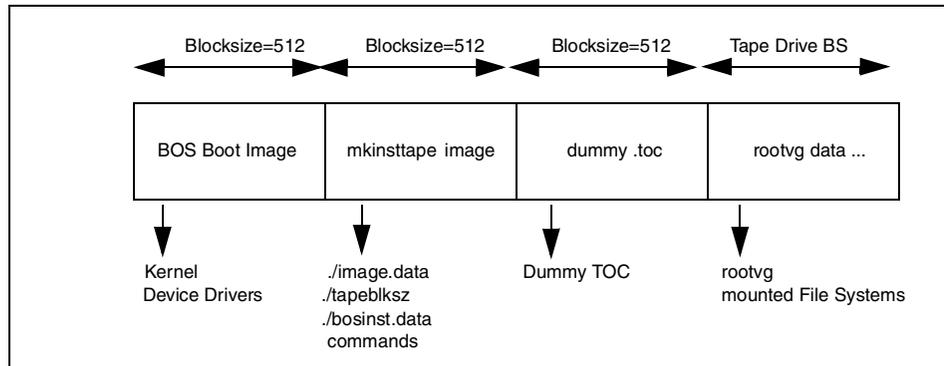


Figure 6-1 Contents of mksysb image

The BOS boot image contains a copy of the system's kernel and device drivers needed to boot from **mksysb** tape. It is created by the **bosboot** command.

There are two important files in mkinsttape (the second image of the **mksysb** tape): `/image.data` and `/bosinst.data`. The `image.data` file contains information describing the image installed during the BOS installation process. This information includes the sizes, names, maps, and mount points of logical volumes and file systems in the rootvg. The **mkszfile** command generates the `image.data` file. The `image.data` file is arranged in stanza format. Each stanza contains one or more fields. The most important fields for you are:

- | | |
|---------------------|--|
| SHRINK | When set to YES, causes the system to create the smallest file systems required to contain all the data in the file system. |
| BOSINST_FILE | Provides the full path name of a file or command to execute after BOS install completes. |
| EXACT_FIT | When set to YES, causes the system to place logical volumes on the disk according to the physical partition maps that were generated with the <code>-m</code> flag of the mksysb command. |

You can customize this file before using the **mksysb** command.

The `/bosinst.data` file allows you to specify the requirements at the target system and how the user interacts with the target system. This file contains the customized BOS install procedures and dictates how the BOS install program will behave. You can customize this file before using the **mksysb** command or use a procedure to customize this after the image backup is done.

Take note that the **mksysb** command uses the **backup** command to save the contents of JFS data held in the rootvg volume group. It does not save raw data.

6.1.3 Back up non-rootvg volume group

If you need to make a backup of the non-rootvg volume group, use the **savevg** command. It works the same as the **mksysb** command except that it does not create a bootable tape. The first three images on the tape are dummy replacements for the images normally found on a bootable tape (refer to Figure 6-1 on page 155). The **savevg** command uses the data file `/tmp/vgdata/uservg/uservg.data` created by the **mkvgdata** command. This file contains information about a user volume group, and is similar to an `image.data` file. After making changes in `uservg.data`, you can remake the user volume group using the **restvg** command.

To back up the uservg volume group to the tape drive, enter the following:

```
# savevg -if /dev/rmt0 uservg
```

To back up the uservg volume group to the tape drive excluding the files listed in the `/etc/exclude.uservg` file, enter the following:

```
# savevg -ief /dev/rmt0 uservg
```

6.1.4 Back up file systems

The **backup** command is the preferred command for making backups of an AIX system. There are two backup techniques available:

1. Back up by file name:
 - a. To back up all the files and subdirectories in the `/home` directory using full path names, use:

```
# find /home -print | backup -if /dev/rmt0
```
 - b. To back up all the files and subdirectories in the current directory using relative path names, use:

```
# cd /home  
# find . -print | backup -if /dev/rmt0
```
2. A back up of the file system by i-node has the advantage of making incremental backups.

To back up the file system `/home` at the 0 level, enter the following:

```
# backup -0 -uf /dev/rmt0 /home
```

To generate a system backup on the tape and to create `/image.data`, use the command:

```
# mksysb -i /dev/rmt0
```

To generate a system backup and exclude the `/home` directory, use:

```
# echo /home > /etc/exclude.rootvg  
# mksysb -e /dev/rmt0
```

Note: Before running the `mksysb` command, ensure that enough space is available in the `/tmp` file to store a boot image. It typically needs 8 to 12 MB of free space in `/tmp`. Use `-X` with `mksysb` to dynamically increase the `/tmp` file system.

6.2 Recovery

This section discusses the recovery procedures of volume groups and file systems.

6.2.1 Recovery of rootvg

You can use the `mksysb` tape in two situations, as described in the following:

1. Reinstalling or cloning a system:
 - a. Check if your system will boot from tape:

```
# bootinfo -e
```

If a 1 is returned, a boot from tape is supported. If a 0 is returned, a boot from tape is not supported.
 - b. Put the `mksysb` tape in the tape drive.
 - c. Check booting order for service mode:

```
# bootlist -m service -o
```

If it is necessary, change it to `rmt0`:

```
# bootlist -m service rmt0
```
 - d. Reboot the system.
2. For extracting data from `mksysb` tape, see 6.2.4, “Extracting data from `mksysb` tape” on page 158.

6.2.2 Restore a non-rootvg volume group

To restore a backup of the user volume group use the **restvg** command from a **savevg** image. This utility can be used to shrink the size of file systems within the **uservg** volume group. All you have to do is modify the `/tmp/vgdata/uservg/uservg.data` file.

To restore the volume group **uservg** image from the tape onto the disks specified in the `/tmp/vgdata/uservg/uservg.data` file, enter the following:

```
# restvg -f /dev/rmt0
```

6.2.3 Restore a file system

The **restore** command is used to restore data backed up with the **backup** command.

restore -Tvf /dev/rmt0 will show the table of contents for the tape and list the files on it.

restore -xvf /dev/rmt0 /home/ostach restores individual files from backup made by the **backup -i** command.

restore -rvf /dev/rmt0 restores an entire file system archive.

6.2.4 Extracting data from mksysb tape

If you need to verify the **mksysb** tape or extract some data from it, follow the procedures below:

1. Issue the **restore** command.

a. Rewind the tape using the command:

```
# tctl -f /dev/rmt0 rewind
```

b. Move the tape forward to the third tape marker (beginning of fourth image):

```
# tctl -f /dev/rmt0.1 fsf 3
```

c. Restore data:

```
# restore -xqvf /dev/rmt0.1 /tmp/my_file
```

2. SMIT fast path:

```
# smitty lsmksysb
```

6.2.5 A customized mksysb restore

The `bosinst.data` file is located in the `/var/adm/ras` directory.

A modified copy of this file can be used during restore to enable customized actions. It is best to create a copy of the original, such as `bosinst.data.orig`, for safe keeping.

You can customize the `bosinst.data` file using a text editor. Commonly modified fields are described in the following.

CONSOLE	Specifies the full path name of the device you want to use as the console. For example, <code>/dev/lft0</code> .
INSTALL_METHOD	Specifies a method of installation: Migrate, preserve, or overwrite. The default value is initially blank. The installation program assigns a value, depending on which version of AIX was previously installed.
PROMPT	Specifies whether the installation program uses menus from which you make choices.
CUSTOMIZATION_FILE	Specifies the path name of a customization file you create. The customization file is a script that starts immediately after the installation program concludes.
HDISKNAME	Specifies the path name of the target disk.

You can tailor these actions through the following procedure.

1. Create a file named signature consisting of one word: *data*.
2. Back up the edited `bosinst.data` file and the new signature file to diskette:

```
# ls ./bosinst.data signature | backup -qiv /dev/fd0
```
3. Put the diskette in the diskette drive of the target machine you are installing.
4. Boot the target machine from the install tape and install AIX. The BOS installation program will use the diskette file rather than the default `bosinst.data` file shipped with the installation media.

If you would like to exclude some files or directories from the backup, put these file names into the `/etc/exclude.rootvg` file and use the `-e` flag with the `mksysb` command.

Another way to selectively restore a `mksysb` is to modify the `image.data` file. For example, if a `mksysb` restore is required without `/home`, the following procedure can be used.

Note: Access to another AIX system and an AIX-formatted diskette is required to perform this procedure.

The procedure is:

1. On another AIX system, place the **mksysb** tape in the tape drive. Change the block size of the tape drive to 512 by running:

```
# chdev -l rmt0 -a block_size=512
rmt0 changed
```

2. Create a temporary directory in /tmp called newdata and **cd** to /tmp/newdata. Rewind the **mksysb** tape by running:

```
# tctl -f /dev/rmt0 rewind
```

3. Restore the image.data file from the second image of the **mksysb** tape. Enter:

```
# restore -s2 -xqv /dev/rmt0.1 ./image.data
New volume on /dev/rmt0.1:
Cluster size is 51200 bytes (100 blocks).
The volume number is 1.
The backup date is: Fri Aug 23 12:05:24 CDT 2002
Files are backed up by name.
The user is root.
x          6931 ./image.data
The total size is 6931 bytes.
The number of restored files is 1.
```

4. In order to skip restoring /home from this **mksysb**, remove the **lv_data** section and **fs_data** for /home:

```
lv_data:
    VOLUME_GROUP= rootvg
    LV_SOURCE_DISK_LIST= hdisk0
    LV_IDENTIFIER= 00015f8f00004c00000000efd755423d.8
    LOGICAL_VOLUME= hd1
    VG_STAT= active/complete
    TYPE= jfs
    MAX_LPS= 512
    COPIES= 1
    LPs= 1
    STALE_PPs= 0
    INTER_POLICY= minimum
    INTRA_POLICY= center
    MOUNT_POINT= /home
    MIRROR_WRITE_CONSISTENCY= on/ACTIVE
```

```
.
.
.
```

```
fs_data:
```

```
FS_NAME= /home
FS_SIZE= 32768
FS_MIN_SIZE= 1128
FS_LV= /dev/hd1
FS_FS= 512
FS_NBPI= 4096
FS_COMPRESS= no
FS_BF= false
FS_AGSIZE= 8
```

5. After making the changes to the `image.data` file, create a signature file. The signature file is required by the restore procedure.

```
# echo data > signature
```

6. After formatting a diskette, back up the two files (`image.data` and `signature`) to the diskette drive.

```
# find . -print |backup -iqvf /dev/rfd0
Backing up to /dev/rfd0.
Cluster 9216 bytes (18 blocks).
Volume 1 on /dev/rfd0
a      0 .
a      6931 ./image.data
a      5 ./signature
The total size is 6936 bytes.
Backup finished on Fri Aug 23 14:27:08 CDT 2002; there are 18 blocks on 1
volumes.
```

7. Reboot the *target system* off of the `mksysb` tape with the diskette in the floppy drive and follow the normal `mksysb` restore procedures. The edited version of the `image.data` file will be used to restore the system.

6.3 Other UNIX backup commands

If you are familiar with other UNIX backup commands, you can use them in AIX as well.

- ▶ The `tar` command is a traditional UNIX archive command. It can generate a tar file as well as tapes or diskettes, `tar` backups files, directories, and subdirectories. It does not support AIX ACLs. Although AIX 5L allows creation of files larger than 2 GB, the `tar` command is not enabled for files greater than 2 GB in size due to limitations imposed by XPG/4 and POSIX.2 standards. It does not preserve the sparse nature of any file that is sparsely allocated. Any file that was originally sparse before the restoration will have all space allocated within the file system for the size of the file (see 6.6.2, “Sparse files” on page 166, for details).

To create a **tar** backup of the home directory, enter:

```
# tar -cvf /dev/rmt0 /home/ostach
```

To create a **tar** backup of a relative path name, enter:

```
# tar -cvf /dev/rmt0 file1 file2
```

To list contents of a **tar** archive file, enter:

```
# tar -tvf /tmp/mybackup.tar
```

To restore from a backup **tar**:

```
# tar -xvf /dev/rmt0
```

- ▶ The **cpio** command reads or writes from stdin and stdout. It is a common UNIX command, so it is very useful for moving information between systems. To back up data in the current directory to the /tmp/file.cpio file using **cpio**, enter:

```
# find . -print | cpio -ov >/tmp/file.cpio
```

To restore data from the **cpio** archive file, /tmp/file.cpio, use:

```
# cpio -idv < /tmp/file.cpio
```

6.4 Sysback

IBM Operational Support Services for RS/6000 system backup and recovery (Sysback) provides easy backup and recovery of either individual files or your entire system. Sysback is available on a limited basis from IBM and as a services offering. It provides the following functions.

- ▶ Replicates your AIX operating system
- ▶ Improves performance and capacity of your backups
- ▶ Utilizes the AIX interface, System Management Interface Tool (SMIT)
- ▶ Provides backup and recovery from local and remote devices

With the IBM Operational Support Services for RS/6000 system backup and recovery software, Sysback, you get a flexible backup method that handles the recovery and replication of your Advanced Interactive Executive (AIX) systems. Sysback is a single product for backing up and restoring user data, applications, and your operating system.

Sysback performs backup, listing, verification, and restoration from various backup types including:

- ▶ A full system backup (including multiple volume groups)
- ▶ Select volume groups

- ▶ Select file systems
- ▶ Select files or directories
- ▶ Raw logical volumes

6.4.1 Sysback installation

You can install Sysback using your current hardware configurations or you can change the configuration. Some of the installation features include the abilities to:

- ▶ Install locally from a tape drive or remotely from a tape drive or files.
- ▶ Boot over the network, which removes your local media from the process.
- ▶ Preserve previous volume groups and logical volume placements.
- ▶ Create logical volume types.
- ▶ Change logical volume and file system sizes and placement.
- ▶ Create, import, or exclude individual volume groups.

6.4.2 Virtual devices

Sysback increases your backup performance and volume by allowing each user to group multiple tape drives into a single virtual device. Each user can select from the following backup types:

- Sequential** Performs an automatic backup on the next device when a prior tape reaches the end of its volume
- Parallel** Stripes data across multiple devices in parallel, thus completing backups in as little as one-third the time
- Multi-copy** Creates multiple, identical backups in the same time it takes to create a single backup without the AIX Sysback software

6.4.3 Other features

Using the AIX user-friendly interface, System Management Interface Tool (SMIT), Sysback also provides:

- ▶ Sequential auto-loader support
- ▶ Software data compression
- ▶ User and host security features
- ▶ Selective file restoration

- ▶ Recreation of volume groups, logical volumes, and file systems from backup media
- ▶ Incremental backups
- ▶ Multiple tape volume support
- ▶ An on-screen progress indicator
- ▶ Exclusion of files or directories
- ▶ The ability to append multiple backups to the same tape

6.4.4 Optional features

The Sysback Offline Mirror Backup feature allows you to perform backups on offline mirrors of data. The backup can include a single file system or logical volume to a complete system. When complete, the offline copies are returned to their original file system or logical volume. The files are also compared, and only the changed files are updated from the active copies.

6.5 Tivoli Storage Manager

Another backup and recovery product is the Tivoli Storage Manager. The Tivoli Storage Manager product set is an enterprise-wide solution integrating automated network backup, archive and restore, storage management, and disaster recovery. The Tivoli Storage Manager product set is ideal for heterogeneous, data-intensive environments, supporting over 35 platforms and over 250 storage devices across LANs, WANs, and SANs, plus providing protection for leading databases and e-mail applications.

The first version of Tivoli Storage Manager Version 3.7.1, the successor to ADSTAR Distributed Storage Manager (ADSM) Version 3.1, was announced in September 1999. Besides the rebranding of the entire product with Version 3.7.1, several enhancements for both server and client, as well as complementary products, introduced a new age for backup and restore solutions.

Feature improvements for server and client exploit Storage Area Network (SAN) technology and respond to the necessity of speeding up backup and restore operations. Tivoli Storage Manager utilizes SAN library sharing, giving multiple servers the ability to share an automated library in a high-performance Storage Area Network (SAN) configuration. The LAN-free client data transfer feature reduces network traffic and improves bandwidth by backing up and restoring data directly to and from SAN-attached disk or tape storage.

Utilizing Tivoli Data Protection products, Tivoli Storage Manager supports most enterprise management, database, and groupware applications. Tivoli Data

Protection interfaces directly with these applications using their backup-certified utilities and interfaces, ensuring a comprehensive storage management solution.

Tivoli Storage Manager Version 3.7.3 was released in April 2000 and Version 4.1 in July 2000. The major new functions introduced in these versions were:

- ▶ Windows 2000 exploitation and support of client and server
- ▶ New AIX and Windows 2000 platform support for tape library sharing in a SAN environment for SCSI-connected libraries
- ▶ Client tape support for backup set restore on the local client system without interaction with the Tivoli Storage Manager server
- ▶ New Version 3.7 backup-archive clients, including the newly supported Linux platform
- ▶ Backup-archive client and server enhancements for the special backup requirements of mobile systems
- ▶ LAN-free client data transfer for the API client in a SAN environment
- ▶ Tivoli Storage Manager hardware integration to implement server-free backup solutions for enhanced disk subsystems like IBM ESS and EMC Symmetrix

Consult your local IBM branch for information on the Tivoli Storage Manager product.

6.6 Troubleshooting

This section will cover problems that can occur during backup procedures.

6.6.1 Remote backup

It is possible to create **tar** backup images on remote tape drives. To make server3 files backed up on a server1 tape drive, perform the following steps.

```
# hostname
server3
# cd /test
# tar cBf - . | rsh server1 "dd ibs=512 obs=512 of=/dev/rmt0"
5080+0 records in.
5080+0 records out.
```

The **-B** flag for **tar** tells it to create 512-byte blocks, which is necessary on pipes. The **-f -** combination makes **tar** send output to stdout. The **rsh** takes stdin and writes it to **/dev/rmt0** using 512-byte blocks.

For a similar example of using the **backup** command facility to back up the content of /tmp as standard input to a tape drive of a remote system named server4, use:

```
# find /tmp -print | backup -ivqf - | rsh server4 "dd of=/dev/rmt0 bs=512"
```

Note: Set up the /etc/hosts.equiv and \$HOME/.rhosts files to create relationships between systems to work with **rsh**.

6.6.2 Sparse files

A file is a sequence of indexed blocks of arbitrary size. The indexing is accomplished through the use of direct mapping or indirect index blocks from the file's i-node. Each index within a file's address range is not required to map to an actual data block.

A file that has one or more indexes that are not mapped to a data block is referred to as being *sparsely allocated* or a *sparse* file. A sparse file will have a size associated with it, but it will not have all of the data blocks allocated to fulfill the size requirements. To identify if a file is sparsely allocated, use the **fileplace** command. It will indicate all blocks in the file that are not currently allocated. The syntax is:

```
fileplace [ { -l | -p } [ -i ] [ -v ] ] File
```

6.6.3 Block size determination

The block size problem is usually encountered while moving a tape from one system to another.

The following are a two methods to find out the block size:

► Method 1

Use the **dd** command to read a single block from the device and find out what block size is used for the archive:

```
dd if=/dev/rmt0 bs=128k count=1 | wc -c
```

This will return to you the size in bytes of the block being read. Assuming that your backup was made with the same physical block size, you can change your device to use this block size.

► Method 2

Use the **tcopy** command as follows to find out the block size:

```
# tcopy /dev/rmt0
tcopy : Tape File: 1; Records: 1 to 7179 ; size:512
tcopy : Tape File: 1; End of file after :7179 records; 3675648 bytes
```

```
tcopy : Tape File: 2; Records: 1 to 2900 ; size:512
tcopy : Tape File: 2; End of file after 2900 records, 76890 bytes
...
```

This will give you a list of all the files found on the media with their byte count and the block size (shown in bold) used.

6.6.4 Tape block size determination scenario

The following procedure describes problems with the tape block size. First, the `/etc/hosts` file is backed up to the tape using the `tar` command with a fixed tape block size of 512 bytes. After this, the block size is changed to 2048, and the tape is attempted to be read (note the error shown in bold).

```
# chdev -l rmt0 -a block_size=512
rmt0 changed
# lsattr -El rmt0 -a block_size
block_size 512 BLOCK size (0=variable length) True
# tar cvf /dev/rmt0 /etc/hosts
a /etc/hosts 4 blocks.
# chdev -l rmt0 -a block_size=2048
rmt0 changed
# tar tvf /dev/rmt0
tar: 0511-193 An error occurred while reading from the media.
There is an input or output error.
```

To determine the tape block size, use the procedure described in 6.6.3, “Block size determination” on page 166, or use `block_size=0`.

```
# chdev -l rmt0 -a block_size=512
rmt0 changed
# tar tvf /dev/rmt0
-rw-rw-r--  0 0    1817 Jun 09 16:24:17 2000 /etc/hosts
# chdev -l rmt0 -a block_size=0
rmt0 changed
# tar tvf /dev/rmt0
tar: The block size is 1 bytes.
-rw-rw-r--  0 0    1817 Jun 09 16:24:17 2000 /etc/hosts
```

6.7 Common backup and restore errors

Table 6-1 on page 168 provides a list of several common errors relating to backup and restore.

Table 6-1 Common backup and restore errors

Error code	Description	Probable causes and recovery procedures
0512-016	Attempt to create a bootable tape failed	Ensure media not write protected. Check and extend file system space if necessary. Ensure enough space in /tmp and /usr. Check status of tape drive. Restore the original directories. Check and recreate missing /unix link.
0503-350	Error occurred while running the restore command	installp fails due to a damaged tape or drive. installp fails due to insufficient space in /var, /tmp. installp fails due to insufficient paging space.
0511-051	Read failed	Block size of restore device incorrect; incorrect value specified with bs, ibs or obs flag; incorrect value specified with fskip flag; tape is damaged.
0511-080	Path name is too long	During the backup process, the find command could not locate data because the path name specified was too long.
0511-133	Data read error using restore -sNumber -f Device	Use the tcopy command to check the number of records on the tape, and specify a correct value with the skip flag.

6.8 Command summary

The following section provides a list of the key commands discussed in this chapter. For a complete reference of the following commands, consult the AIX product documentation.

6.8.1 The **mksysb** command

The **mksysb** command creates an installable image of the root volume group either in a file or onto a bootable tape. The command has the following syntax:

```
mksysb [ -e ] [ -p ] [ -X ] [ -i | -m ] Device | File
```

The commonly used flags are provided in Table 6-2 on page 169.

Table 6-2 Commonly used flags of the `mksysb` command

Flag	Description
-e	Excludes files listed in the <code>/etc/exclude.rootvg</code> file from being backed up
-i	Calls the <code>mkszfile</code> command, which generates the <code>/image.data</code> file
-m	Generates map files
-p	Disables software packing of the files as they are backed up
-X	Specifies to automatically expand the <code>/tmp</code> file system if necessary

6.8.2 The `savevg` command

The `savevg` command finds and backs up all files belonging to a specified volume group. The command has the following syntax:

```
savevg [ -e ] [ -f Device ] [ -i | -m ] [ -p ] [ -v ] [ -X ] VGName
```

The commonly used flags are provided in Table 6-3.

Table 6-3 Commonly used flags of the `savevg` command

Flag	Description
-e	Excludes files specified in the <code>/etc/exclude.vgname</code> file from being backed up by this command.
-f <i>Device</i>	Specifies the device or file name on which the image is to be stored. The default is the <code>/dev/rmt0</code> device.
-i	Creates the data file by calling the <code>mkvgdata</code> command.
-m	Creates the data file with map files.
-p	Disables software packing of the files as they are backed up.
-X	Specifies to automatically expand the <code>/tmp</code> file system if necessary.

6.8.3 The `restvg` command

The `restvg` command restores the user volume group and all its containers and files. The command has the following syntax:

```
restvg [ -f Device ] [ -s ] [-n ] DiskName ... ]
```

The commonly used flags are provided in Table 6-4 on page 170.

Table 6-4 Commonly used flags of the `restvg` command

Flag	Description
-n	Specifies that the existing MAP files are ignored.
-f <i>Device</i>	Specifies the device or file name on which the image is to be stored. The default is the <code>/dev/rmt0</code> device.
-s	Specifies that the logical volumes be created at the minimum size possible to accommodate the file systems.

6.8.4 The backup command

The **backup** command backs up files and file systems. The command has the following syntax:

```
backup -i [ -p [ -e RegularExpression ] ] [ -f Device ] [ -l Number ] [ -o ] [ -q ] [ -v ]
```

The commonly used flags are provided in Table 6-5.

Table 6-5 Commonly used flags of the `backup` command

Flag	Description
-f <i>Device</i>	Specifies the output device
-i	Specifies that files be read from standard input and archived by file name
-m	Creates the data file with map files
-q	Indicates that the removable medium is ready to use
-v	Causes the backup command to display additional information about the backup

6.8.5 The restore command

The **restore** command extracts files from archives created with the **backup** command. The command has the following syntax:

```
restore -x -T [ v q ] [ -f Device ] [ File ... ]
```

The commonly used flags are provided in Table 6-6 on page 171.

Table 6-6 Commonly used flags of the restore command

Flag	Description
-f <i>Device</i>	Specifies the device or file name on which the image is to be stored. The default is the /dev/rmt0 device.
-T	Displays information about the backup archive.
-q	Indicates that the removable medium is ready to use.
-v	Causes the backup command to display additional information about the backup.
-x	Restores individually named files specified by the file parameter.

6.8.6 The tar command

The **tar** command manipulates archives by writing files to, or retrieving files from, an archive storage medium. The command has the following syntax:

```
tar { -c | -t | -x } [ -B ] [ -v ] [ -f Archive] [ File | Directory ...
```

The commonly used flags are provided in Table 6-7.

Table 6-7 Commonly used flags of the tar command

Flag	Description
-c	Creates a new archive and writes the files specified by one or more file parameters to the beginning of the archive
-t	Lists the files in the order in which they appear in the archive
-B	Forces input and output blocking to 20 blocks per record
-f <i>Archive</i>	Uses the Archive variable as the archive to be read or written
-v	Lists the name of each file as it is processed
-x	Restores individually named files specified by the File parameter

6.9 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

1. Which of the following is an example of a backup by relative path name?
 - A. `tar -cvf /dev/rfd0 /file1`
 - B. `tar -cvf /dev/rfd0 file1 file2`
 - C. `tar -cvf /dev/rfd0 /home/file1`
 - D. `tar -cvf /dev/rfd0 /home/file1/file2`
2. Which of the following commands should be used to view the contents of a tape in `/dev/rmt0` that was created using the `backup` command?
 - A. `restore -T`
 - B. `restore -Tf /dev/rmt0`
 - C. `restore -xf /dev/rmt0`
 - D. `backup -tvf /dev/rmt0`
3. Before upgrading a system to the latest maintenance release, a system administrator attempts to create a `mksysb` backup of the current system. During the backup process, the following error message is received:

```
0512-016: Attempt to create a bootable tape failed
bosboot -d /dev/device -a failed with return code 52
```

Which of the following commands should be run to diagnose this problem?
 - A. `errpt -a | grep TAPE_ERR6`
 - B. `ls -la /usr/lib/boot/`
 - C. `fsck -n filesystems 2>&1 /tmp/mksysb.log`
 - D. `mkboot -d /dev/hdisk0 -c`
4. Which of the following is backed up with the `mksysb` command when a system consists of the volume groups `rootvg`, `vol2`, and `vol3`?
 - A. All files in `rootvg`, `vol2`, and `vol3`
 - B. All locally mounted file systems in the `rootvg`
 - C. All files in the root volume group
 - D. All files in the root volume group, excluding the `odm` files

5. A system administrator is running a remote archive to the machine whose host name is Best. The following command was used:

```
find /tmp -print | backup -ivqf | rsh Best "dd of=/dev/rmt0 \    bs=512"
```

After the files are listed, the following error message occurs:

```
backup: permission denied
```

Which of the following options is the most probable cause?

- A. The machine Best does not exist.
- B. The \$HOME/.rhosts file on the machine Best is not set up properly.
- C. The system administrator is trying to back up files on a local machine without having access.
- D. The bs field on the **dd** command is different from what the block size of the tape drive indicates.

6.9.1 Answers

The following are the preferred answers to the questions provided in this section.

1. B
2. B
3. B
4. B
5. B

6.10 Exercises

The following exercises provide sample topics for self study. They will help ensure comprehension of this chapter.

1. Create a bootable tape by using the **mksysb** command.
2. Restore files from the second image of the **mksysb** tape by using the **tct1** and **restore** commands.
3. On a test system that does not affect any user, save a user volume group.
4. Try to restore a user volume group by shrinking the file systems.
5. Make a file system backup by using the **backup** command.



Network installation management

Network Installation Management (NIM) is used to manage remote sites on a network. The types of machines you can manage are *stand-alone*, *diskless*, and *dataless*. A stand-alone machine is one that can boot (start up) by itself. Diskless and dataless systems cannot boot by themselves; they must use remote resources to boot. Diskless systems have no disk drive. Dataless systems have a local disk drive but they cannot boot from it.

Using NIM, you can install a group of machines with a common configuration or customize an installation for the specific needs of a given machine. The number of machines you can install simultaneously depends on the throughput of your network, the disk access throughput of the installation servers, and the platform type of your servers. Validate that the filesets required for NIM are installed:

```
# ls|pp -l |grep bos.sysmgmt.nim
bos.sysmgmt.nim.client      5.1.0.25 COMMITTED Network Install Manager -
bos.sysmgmt.nim.master     5.1.0.25 COMMITTED Network Install Manager -
bos.sysmgmt.nim.spot       5.1.0.25 COMMITTED Network Install Manager - SPOT
```

The NIM environment comprises client and server machines. A server provides resources (for example, files and programs required for installation) to another machine. A machine that is dependent on a server to provide resources is known as a client. In this guide and reference, any machine that receives NIM resources

is a client, although the same machine can also be a server in the overall network environment.

The preferred method when installing logical partitions is the NIM installation method as this method allows multiple, similar installations, in the shortest time period.

Table 7-1 Types of machines managed by NIM

Standalone	Diskless	Dataless
Local disks	No local disks	Limited local disks
Local boot image	External boot image	External boot image
Local resources	External paging space	Local paging space, /tmp and /home
Mounts file system locally	Mounts file systems from remote server	Mounts file systems from remote server
Independent of network servers for operation	Dependent on network servers for operation	Dependent on network servers for operation

For more information on NIM, refer to *NIM from A to Z in AIX 4.3*, SG24-5524, and *AIX 5L Version 5.2 Installation in a Partitioned Environment*, SC23-4382.

7.1 Debugging NIM boot images

If a client machine is unable to network boot from its boot server, there may be a problem in one or more of the following stages of the network boot:

- ▶ Establishing network communication between the client and server
- ▶ Obtaining the boot image from the server
- ▶ Running the boot image on the client

7.1.1 Establishing network communication

The following steps may be used to establish network communication:

1. Before initiating the network boot on the client, perform a ping test from the client bootp menus.
2. If the ping test fails, verify that the client, server, and gateway addresses are specified correctly.
3. If the addresses are correct, try to ping the server from a different machine in the client's subnet.
4. If the server can be pinged from another machine, the network adapter on the boot client may be faulty.
5. If the server cannot be pinged from another machine in the client's subnet, there may be routing problems between the client and the server, or else network communications on the server may be faulty. Perform network debugging procedures to determine the source of the problem.

7.1.2 Obtaining the boot image from the server

The following steps may be used to obtain a boot image from the server.

1. If the ping test is successful, perform a network boot of the client. When a network boot is initiated on a client, a bootp request packet is sent from the client to the server.

The server then replies with a packet to the client. The client machine displays the number of packets sent and received for the bootp request. If a packet is sent from the client, but none is received, another packet will be sent.

2. If bootp packets continue to be sent but not received, the boot server may not be responding to the request.

3. From the bootp server, view the `/etc/bootptab` file on the server. It should contain an entry for the client machine with the following information:

```
hostname of client  
bf=boot file  
ip=client ip address  
ht=network type  
sa=boot server address  
sm=client subnet mask  
ha=network adapter hardware address (required only if bootp requests are  
sent by broadcasting)
```

If an entry does not exist, either the `nim` command used to set up the current operation failed, or the machine was reset before the boot operation could occur.

Rerun the NIM `bos_inst`, `diag`, or `maint_boot` operation to prepare the server for the client boot request.

If the entry exists in `/etc/bootptab`, verify that the specified data is correct. If a field contains incorrect data, the information that was used to define the machine or network in the NIM database was probably wrong. Correct this problem by resetting the client machine, correcting the incorrect data in the client or network definition, retrying the NIM operation, and rebooting the client.

4. If the `/etc/bootptab` file is correct, verify that the `inetd` daemon is running. If it is not running, start it and retry the network boot from the client. If the `inetd` daemon is running, it should automatically start the `bootpd` daemon when the bootp request is received at the server.
5. If the `bootpd` daemon is not started, verify that the `bootpd` entry in the `/etc/inetd.conf` file is not commented out. If it is commented out, uncomment it and restart `inetd` with the `refresh -s inetd` command. Retry the network boot from the client.
6. If a bootp reply is still not received at the client, manually start the `bootpd` daemon in debug mode:
 - a. Comment out the `bootpd` entry from the `/etc/bootptab` file on the server.
 - b. Stop all running `bootpd` processes.
 - c. Restart `inetd` using the `refresh -s inetd` command.
 - d. Start `bootpd` from the command line using the `/usr/sbin/bootpd -s -d -d -d` command.

7. Retry to network boot from the client. If no output is displayed from the running **bootpd** command, the client bootp request is not reaching the server. Verify that the addresses specified in the bootp menus are correct. If they are correct, perform network debugging procedures to determine why the packet is not reaching the server.
8. If the server receives the client bootp request, then running the **bootpd** command displays output matching the client data in the `/etc/bootptab` file. Verify that the specified addresses are correct. This information is sent back to the client in the bootp reply.
9. If the client is still not receiving the bootp reply, perform network debugging procedures to determine why the reply packet is not reaching the client.
10. After the client receives the bootp reply, it will use TFTP to transfer the boot image from the server.
11. The number of TFTP packets transferred to the client will be displayed at the client machine.
12. The boot image has been successfully retrieved at the client machine when the LED shows 299 on rs6k-platform machines or when the bottom third of the screen turns gray on other platforms.
13. If the tftp transfer of the boot image does not complete successfully, the client may be trying to get the wrong boot image. Verify that the client definition in the NIM database shows the correct platform and kernel type. If the data is incorrect, correct it, reset the client machine, rerun the NIM operation, and reboot the client over the network.
14. Verify that the `/tftpboot` directory on the boot server contains a link with the client name to the correct boot image. If the link does not exist, reset the client machine, rerun the NIM operation, and reboot the client over the network.
15. If the link with the client name is pointing to the correct boot image and the tftp transfer of the boot image does not complete successfully, the boot image may be corrupted.

Recreate the boot image by performing a NIM check operation with the force flag on the SPOT. If the client is not an rs6k-platform machine, also make sure the client has the latest version of the firmware installed.

7.1.3 Running the boot image on the client

After the client machine has successfully received the boot image from the server, the most common errors encountered are hangs with the LED showing 608, 611, or 613. Some machines may not have LED displays. Debugging such problems on these machines will require using debug-enabled boot images.

Error 608

Attempting to retrieve the client.info file with tftp. Note that a flashing 608 indicates that multiple attempts to retrieve the client_info file are occurring.

Error 611

The /etc/exports file either does not have the correct permissions for the client to mount /usr or does not have any entry for /usr.

Error 613

The system is unable to set up the routing tables.

7.2 Error 611: Network error

If you get an error 611 during boot, the problem is that the /etc/exports file either does not have the correct permissions for the client to mount /usr or does not have any entry for /usr.

To rectify this problem type the following command:

```
exportfs
```

Using the **exportfs** command, verify that each entry is filled in correctly.

If you receive the following error message when adding the network boot client, then the client has been configured, but you must correct the problem with the export before booting the client.

```
Configuring network boot for client thumper.aix.dfw.ibm.com
/usr/ro,root=thumper.aix.dfw.ibm.com
Starting NFS and BOOTP services..
exports: /usr: sub-directory (/usr/lpp) already exported
/usr not found in etc/exports
```

Execute the **exportfs** command.

If you continue to have problems trying to export a /usr parent directory, check the /etc/exports file, and look for an existing /usr subdirectory that is not a file system. If you have a /usr subdirectory that is not a file system you will not be allowed to export /usr until you remove it from the /etc/exports/file. An example would be:

```
/usr/sys/inst.images -rw
/usr/lpp              -root=mars,access=mars
```

In the above example the `/usr/sys/inst.images` is its own file system, but `/usr/lpp` is still part of the `/usr` file system. Execute the following command to remove `/usr/lpp` from the `/etc/exports` file.

```
/usr/sbin/rmfsxp -d /usr/lpp -B
```

Re-add the network boot client by running the `smitty sysback` command. Verify by executing the `exportfs` command.

7.3 The nim command

The `nim` command performs an operation on a NIM object. The type of operation performed is dependent on the type of object specified by the `ObjectName` parameter. Possible operations include initializing environments and managing resources. You can use the `lsnim` command to display the list of supported operations.

The `nim` command syntax is as follows.

```
nim {-o Operation} [-F] [-t Type] [-a Attribute=Value . . .] {ObjectName}
```

Table 7-2 provides a list of the flags and their descriptions. The most important functions are the flag subfunctions.

Table 7-2 Commonly used flags of the nim command

Flag	Flag description
<code>-a Attribute = Value . . .</code>	Assigns the specified value to the specified attribute. Use the <code>lsnim -q Operation -t Type</code> command to get a list of valid attributes for a specific operation.
<code>-F</code>	Overrides some safety checks, performs a force.

Flag	Flag description
-o <i>Operation</i>	<p>Specifies an operation to perform on a NIM object. The possible operations are:</p> <p>allocate Allocates a resource for use</p> <p>bos_inst Performs a BOS installation</p> <p>change Changes an object's attributes</p> <p>check Checks the status of a NIM object</p> <p>cust Performs software customization</p> <p>deallocate Deallocates a resource</p> <p>define Defines an object</p> <p>diag Allows a machine to boot a diag image</p> <p>dkls_init Initializes a diskless machine's env</p> <p>dtls_init Initializes a dataless machine's env</p> <p>fix_query Lists the fix information for a given APAR or keyword</p> <p>lppchk Verifies installed filesets on NIM machines and SPOTs</p> <p>lspp Lists licensed program information about an object</p> <p>maint Performs software maintenance</p> <p>remove Removes an object</p> <p>reset Resets an object's NIM state</p> <p>sync_roots Synchronizes root directories for diskless and dataless clients for a specific SPOT</p> <p>unconfig Unconfigures the NIM master fileset</p> <p>maint_boot Lets machine boot in maintenance mode</p> <p>showlog Displays a client's install, boot, or custom log, or a SPOT's install log from the NIM master</p> <p>showres Displays the contents of a resource</p> <p>alt_disk_install Performs an Alternate Disk Installation</p>

Flag	Flag description
-t <i>Type</i>	<p>Specifies the type of the NIM object for define operations. The possible types are listed below.</p> <p>Resource types:</p> <p>bosinst_data Config file used during base system installation</p> <p>dump Parent directory for client dump files</p> <p>fix_bundle Fixes (keyword) input file for the cust or fix_query operation</p> <p>home Parent directory for client /home directories</p> <p>image_data Config file used during base system installation</p> <p>installp_bundle Installp bundle file</p> <p>lpp_source Source device for optional product images</p> <p>mksysb AIX mksysb image</p> <p>paging Parent directory for client paging files</p> <p>root Parent directory for client / directories</p> <p>script Executable file that is run on a client</p> <p>shared_home /home directory shared by clients</p> <p>spot Shared Product Object Tree - equivalent to /usr file system</p> <p>tmp Parent directory for client /tmpdirectory</p> <p>exclude_files Files excluded from a mksysb image</p> <p>resolv_conf Name-server configuration file</p> <p>Machine types:</p> <p>diskless All file systems and resources remote</p> <p>dataless Local paging, dump; remote /,/usr; others remote or local</p> <p>standalone Local file systems and resources</p> <p>master Machine that controls the NIM environment</p> <p>Network types:</p> <p>tok Token-ring network</p> <p>ent Ethernet network</p> <p>fddi FDDI network</p> <p>generic Other TCP/IP networks</p> <p>Group types:</p> <p>mac_group Group of machines</p> <p>res_group Group of resources</p>

The following are examples of using the **nim** command.

To define a resource that is a directory containing installable images, is located on server altoid, and has a path name of /usr/sys/inst.images, and name the resource images, enter:

```
# nim -o define -t lpp_source -a server=altoid \  
-a location=/usr/sys/inst.images images
```

To remove a resource named dump_files, enter:

```
# nim -o remove dump_files
```

Machines on the BLDG905 network use the gateway905 gateway to reach the OZ network. Machines on the OZ network use the gatewayOZ gateway to reach the BLDG905 network. To add a route between two networks named BLDG905 and OZ, enter:

```
# nim -o change -a routing1="OZ gateway905 gatewayOZ" BLDG905
```

To have NIM check on the usability of a SPOT named myspot, enter:

```
# nim -o check myspot
```

To have NIM check on the usability of a SPOT named myspot and force results from busy clients, add the -F flag. This is not normally needed with check; however, use it to get the best results.

```
# nim -F -o check myspot
```

To allocate resources to a diskless workstation with the name of syzygy and SPOT attribute value of spot1, enter:

```
# nim -o allocate -a spot=spot1 syzygy
```

To deallocate a lpp_source named images from the stand-alone machine client1, enter:

```
# nim -o deallocate -a lpp_source=images client1
```

To install the machine bluefish using the resources spot1, images1, bosinst_data1, and rconf1, first allocate the resources by entering:

```
# nim -o allocate -a spot=spot1 -a lpp_source=images1 \  
-a bosinst_data=bosinst_data1 -a resolv_conf=rconf1 bluefish
```

Then, perform the BOS installation by entering:

```
# nim -o bos_inst bluefish
```

To install a software product into a spot, spot1, such that the image for the installable option, adt, resides in the lpp_source, images, enter:

```
# nim -o cust -a lpp_source=images -a filesets=adt spot1
```

To uninstall the software products bos.INed and adt from a spot, spot1, enter:

```
# nim -o maint -a installp_flags="-u" -a filesets="bos.INed adt" spot1
```

To initialize the environment for a diskless workstation with the name of syzygy, using the resources spot1, root1, dump1, and paging1, first allocate the resources by entering:

```
# nim -o allocate -a spot=spot1 -a root=root1 -a dump=dump1 \  
-a paging=paging1 syzygy
```

Then initialize the resources for the client machine by entering:

```
# nim -o dkls_init syzygy
```

To initialize the environment for a dataless workstation with the name of syzygy, using the resources spot1, root1, and dump1, first allocate the resources by entering:

```
# nim -o allocate -a spot=spot1 -a root=root1 -a dump=dump1 syzygy
```

Then initialize the resources for the client machine by entering:

```
# nim -o dtls_init syzygy
```

To list information about fixes installed on client Standalone1 for 20 APAR numbers, create the file /tmp/apar.list with one APAR number per line, as shown:

```
IX123435  
IX54321  
IX99999  
...
```

Then enter:

```
# nim -o define -t fix_bundle -allocation=/tmp/apar.list \  
-aserver=master fix_bun  
# nim -o allocate -a fix_bundle=fix_bun Standalone1  
# nim -o fix_query Standalone1
```

To allocate all resources applicable to stand-alone machines from the NIM resource group res_grp1 to the machine mac1, enter:

```
# nim -o allocate -a group=res_grp1 mac1
```

To install the machine bluefish while allocating the resources spot1, images1, bosinst_data1, and rconf1 automatically when the bos_inst operation starts, enter:

```
# nim -o bos_inst -a spot=spot1 -a lpp_source=images1 \  
-a bosinst_data=bosinst_data1 -a resolv_conf=rconf1 bluefish
```

To define default routes for the networks net1 and net2 that use default gateways gw1 and gw2, respectively, enter the following two commands:

```
# nim -o change -a routing1="default gw1" net1  
# nim -o change -a routing1="default gw2" net2
```

To install the machine catfish with the contents of the installp_bundle bundle1 while allocating this resource and the lpp_source images1 when the customize operation runs, enter:

```
# nim -o cust -a installp_bundle=bundle1 -a lpp_source=images1 catfish
```

To define a machine group named DisklessMacs1 with members that are NIM diskless machines named disks1, disks2, and disks3, enter:

```
# nim -o define -t mac_group -a add_member=disks1 \  
-a add_member=disks2 -a add_member=disks3 DisklessMacs1
```

To initialize the group of diskless machines defined by the machine group dtgrp1 while allocating the required and optional resources defined by the resource group dk_resgrp1, when the dkls_init operation runs, enter:

```
# nim -o dkls_init -a group=dtgrp1 dk_resgrp1
```

To exclude the member named dataless1 from operations on the machine group DatalessMacs1, and then initialize the remaining members while allocating the dataless resources defined by the resource group named DatalessRes1, enter the following two commands:

```
# nim -o select -a exclude=dataless2 DatalessMacs1  
# nim -o dtls_init -a group=DatalessMacs1 DatalessRes1
```

To verify the file checksums for all packages beginning with the name bos on NIM targets in the group of stand-alone machines macgrp1, and displaying detailed error information and updating the software database to match the actual file checksum when inconsistencies are found, enter:

```
# nim -o lppchk -a lppchk_flags='-c -m3 -u' -a filesets='bos*' macgrp1
```

Since the lppchk operation runs in the background on group members by default, to view the output from the lppchk operation enter:

```
# nim -o showlog -a log_type=lppchk macgrp1
```

From the master, to uninstall the software products `bos.INed` and `adt` from a stand-alone machine, `stand1`, such that `installp_bundle bundle2` contains the names of the installable options, enter:

```
# nim -o maint -a installp_flags="-u" -a installp_bundle=bundle2 stand1
```

To enable the NIM stand-alone client `stand1` to boot in maintenance mode, enter:

```
# nim -o maint_boot stand1
```

To view the boot logs of the machines defined by the group `Disk1sMac1`, enter:

```
# nim -o showlog -a log_type=boot Disk1sMac1
```

To list all problems fixed by software on the `lpp_source lpp_source1`, use:

```
# nim -o showres -a instfix_flags="T" lpp_source1
```

To clone a `rootvg` on client `austin` to `hdisk2`, but only run `phase1` and `phase2` (leaving the `/alt_inst` file systems mounted), enter:

```
# nim -o alt_disk_install -a source=rootvg -a disk='hdisk2'\
-a phase=12 austin
```

7.4 The `showmount` command

The **`showmount`** command displays a list of all clients that have remotely mounted a file system from a specified machine in the *host* system. This information is maintained by the `mountd` daemon on the host system. This information is saved in the `/etc/rmtab` file in case the server abnormally terminates. The default value for the host system is the value returned by the **`hostname`** command.

The **`showmount`** command syntax:

```
showmount [ -a ] [ -d ] [ -e ] [ Host ]
```

The flags for the **`showmount`** command are described in Table 7-3.

Table 7-3 Commonly used flags of the `showmount` command

Flag	Description
-a	Prints all remote mounts in the format <i>HostName:Directory</i> , in which <i>HostName</i> is the name of the client and <i>Directory</i> is a directory path name that has been remotely mounted
-d	Lists only directories that have been remotely mounted by clients
-e	Prints the list of exported directories

To check which file systems are allowed to be exported, type:

```
# lsfs
Name           Nodename  Mount Pt          VFS  Size  Options  Auto
Accounting
/dev/hd4       --        /                 jfs  24576 --       yes
no
/dev/hd1       --        /home             jfs  360448 --       yes
no
/dev/hd2       --        /usr              jfs  663552 --       yes
no
/dev/hd9var    --        /var              jfs  401408 --       yes
no
/dev/hd3       --        /tmp              jfs  65536  --       yes
no
/dev/lv00      --        /export/lpp_source jfs  884736 rw       yes
no
/dev/lv01      --        /export/spot      jfs  442368 rw       yes
no
```

Run the **showmount** command to display a list of all clients that have remotely mounted a file system.

```
# showmount
raven.itsc.austin.ibm.com
```

Run **showmount -a** to see which directories have been remotely mounted with the host details.

```
# showmount -a
raven.itsc.austin.ibm.com:/export/spot/aix432spot/usr
```

The **showmount -d** command is similar to the **showmount -a** command except it does not show the host details.

```
# showmount -d
/export/spot/aix432spot/usr
```

The **showmount -e** command displays a list of exported directories.

```
# showmount -e
export list for aix4xdev:
/richfs          raven.itsc.austin.ibm.com
/export/redbooks (everyone)
```

The **showmount -e raven** command displays a list of exported directories on the remote host raven.

```
# showmount -e raven
export list for raven:
/elaine (everyone)
```

7.5 Producing debug output for NIM BOS installations

Due to problems in the network or in the NIM configuration, clients may fail to boot properly or install. When this happens, it may be necessary to produce debug information in order to determine the cause of the problem. If a client machine fails to configure properly from the network boot image, debug output from the boot image can be obtained by building the debug-enabled image and attaching a TTY to the client system. This will display the commands and output that is run as the client is configured before further processing is done by AIX.

If the system has been booted from the network boot image, but failures are still occurring during a BOS installation, it may be necessary to collect debug information from the BOS install program. The commands and output from the BOS install program will automatically be displayed on the TTY if the boot image was built debug-enabled. If the boot image was not built for debugging, output can be obtained by either setting a value in a `bosinst.data` file or by entering special codes at the installation menus.

When problems arise during a NIM BOS installation, you will most likely see system *hangs*. Viewing the debug output can be extremely useful, because you will be able to see the commands that failed. The problem may be a misconfiguration of the network adapter or an inability to perform an operation from the client to the server. By examining the debug output, you can see what failed and make corrections to avoid the error in the future.

You will see the **showled** command running in the debug output. This command displays status values on the LEDs on the front of the machine. Frequently, known problems and solutions are referenced by the LED value that is displayed when a problem occurs. Some machines do not have LEDs for displaying such information. Therefore, when debugging problems on such machines, special attention should be given to observing the values that the **showled** commands are displaying.

Obtaining debug information from a network installation can save you time in determining the root cause of a problem. Usually, the problem will be an incorrect definition in the NIM environment that can be found without the debug information. However, with the debug information, you can significantly reduce the scope of the investigation. Without obtaining debug information, an error message itself may not produce enough information. For example, to display detailed information about a NIM object named `melanie_miho_spot`, enter:

```
# lsnim -l melanie_miho_spot
```

This command may produce the following error:

```
0042-011 lsnim: unable to access the /etc/niminfo file
```

This could indicate multiple issues; the NIM master may not have been configured properly, the NIM client may not have access to the NIM master, the `/etc/niminfo` file may have been corrupted or does not exist, etc.

7.5.1 Producing debug output from a network boot image

To produce the debug output from a network boot image:

1. Create debug versions of the network boot images using the check operation from the Web-based System Manager or SMIT interfaces or by using the following command:

```
nim -Fo check -a debug=yes SPOTName
```

Where *SPOTName* is the name of your spot.

2. Get the address for entering the debugger.

From Web-based System Manager:

- a. From the NIM resources container, double-click the spot resource to open the properties notebook 2.
- b. Click the **Boot Image Information** tab in the properties notebook to obtain the address.

Alternatively, you can use the following command to get the address:

```
lsnim -a enter_dbg SPOTName
```

This command displays information about the Network Installation Management environment, where the `-a` flag specifies the specific attribute `enter_dbg` and where *SPOTName* is the name of your SPOT. The displayed output will be similar to the following:

```
spot1:  
enter_dbg = "chrp.mp 0x001840d4"  
enter_dbg = "chrp.up 0x00160b7c"  
enter_dbg = "rs6k.mp 0x001840d4"  
enter_dbg = "rs6k.up 0x00160b7c"  
enter_dbg = "rspc.mp 0x001840d4"  
enter_dbg = "rspc.up 0x00160b7c"
```

Write down the `enter_dbg` address for the client you are going to boot. For example, if your client is an rs6k-uniprocessor machine, you would write down the address 160b7c.

3. Attach a TTY device to your client system (port 1).
4. Set up and perform the NIM operation that will require the client to boot over the network. Boot the client over the network.

5. After the client gets the boot image from the SPOT server, the debug screen will appear on the TTY. At the > prompt, enter:

```
st Enter_dbg_Value 2
```

Where *Enter_dbg_Value* is the number you wrote down in step 2 as your machine type's *enter_dbg* value. Specifying a 2 at the address of the *enter_dbg* value prints the output to your tty.

6. Type *g* for go and press Enter to start the boot process.
7. Use Ctrl+s to temporarily stop the process as you watch the output on the tty. Use Ctrl+q to resume the process.
8. To rebuild your boot images in non-debug mode, use the following command:

```
nim - Fo check SPOTName
```

Where *SPOTName* is the name of your SPOT.

If the boot image is left in debug mode, every time a client is booted from these boot images, the machine will stop and wait for a command at the debugger > prompt. If you attempt to use these debug-enabled boot images and there is not a tty attached to the client, the machine will appear to be hanging for no reason.

7.5.2 Producing debug output from the BOS install program

There are two ways to obtain debug output from the BOS install program. Method A involves entering a special value at one of the installation menus. Method B uses a *bosinst_data* resource to tell the installation program to display debug output.

Method A

To produce debug output when not using a *bosinst_data* resource:

1. To enable debugging for the BOS install program, start by performing all the processing you would normally do to install a client.
2. Since you are not using a *bosinst_data* resource, you will be prompted to supply information about the installation to the BOS install program.
3. Select your console.
4. Select your language.
5. The Welcome to Base Operating System Installation and Maintenance menu is displayed. Instead of selecting one of the options, enter 911 at the prompt and press Enter.

- Continue the normal procedure for selecting options and specifying data until the installation begins. Debug output will be sent to the client's display as the installation proceeds.

Method B

To produce debug output when using a `bosinst_data` resource:

- To enable debugging for the BOS install program, set the value `BOSINST_DEBUG = yes` in the `control_flow` stanza of the `bosinst.data` file that you are using for your `bosinst_data` resource.

A minimum `bosinst.data` file for debugging purposes would contain the following lines:

```
control_flow: BOSINST_DEBUG = yes
```

- In addition to the processing you would normally do to install a client, include the modified `bosinst_data` resource as a resource for the operation. After the client boots over the network, it will use the `bosinst_data` resource to obtain settings for the installation. If the only data specified in your `bosinst.data` file is `BOSINST_DEBUG = yes`, you will be prompted for the remaining required information before the installation will continue. Debug output will be sent to the client's display as the installation continues.

7.6 Common NIM errors

Table 7-4 provides a list of the most common NIM errors.

Table 7-4 Common NIM codes

Error code	Description	Solution
LED C54	<code>mksysb</code> NIM installation hangs	Produce debug output using the 911 method (refer to <i>AIX Version 5.1 Network Installation Management Guide and Reference</i> , SC23-4385, for details).
0042-012	Command may only be executed on master	Execute the command on the master server only.
0042-014	Internal NIM error has occurred	Perform the remove operation on the NIM object followed by the appropriate define operation.
0042-019	Option argument missing	Specify the missing argument for the <code>nim</code> command.

7.7 Quiz

The following assessment questions will help verify your understanding of the topics discussed in this chapter.

1. Company ABC recently ordered 30 new model F50 systems, which they are custom installing with AIX Version 4.3.1 as branch office servers. The system administrator decided to use NIM to install the new servers with the base operating system and all of the applications the systems will need. The install type is spot, using a non /usr spot named 431spot. The installation of the first system is hanging at 611 LED, but all the file systems appear to be properly exported from the NIM server. Which of the following commands will build debug boot images?
 - A. `nim -o check -a debug=yes 431spot`
 - B. `nimclient -o check -a debug=yes 431spot`
 - C. `nim -o debug 431spot`
 - D. `lsnim -a debug=yes 431spot`
2. Company ABC is migrating all of the plant servers from AIX Version 4.2.1 to AIX Version 4.3.1 using NIM. The 240 server in the customer services department is hanging during the install, and the LED is at 611, indicating a failure mounting a network file system from the NIM server. What can the system administrator do to determine which file system the NIM client is unable to mount.
 - A. Unmount the NFS file systems exported from the server to the client.
 - B. Put the tftp daemon on the server in debug mode.
 - C. Restart the install using NIM debug boot images.
 - D. Restart the NFS services on the NIM server.
3. Which of the following is *not* a valid NIM object type?
 - A. Diskless
 - B. Resource
 - C. Standalone
 - D. bosinst_data

7.7.1 Answers

The following are the preferred answers to the questions provided in this section.

1. A
2. C
3. B

7.8 Exercises

The following exercises provide sample topics for self study. They will help ensure comprehension of this chapter.

1. How do you debug a NIM install? Which command would you use?
2. Debug NIM error code 611.
3. Determine if the NIM spot and lpp_source are exported. What command would you use?

Abbreviations and acronyms

ABI	Application Binary Interface	BIND	Berkeley Internet Name Daemon
AC	Alternating Current	BIST	Built-In Self-Test
ACL	Access Control List	BLAS	Basic Linear Algebra Subprograms
ADSM	ADSTAR Distributed Storage Manager	BLOB	Binary Large Object
ADSTAR	Advanced Storage and Retrieval	BLV	Boot Logical Volume
AFPA	Adaptive Fast Path Architecture	BOOTP	Boot Protocol
AFS	Andrew File System	BOS	Base Operating System
AH	Authentication Header	BSC	Binary Synchronous Communications
AIX	Advanced Interactive Executive	CAD	Computer-Aided Design
ANSI	American National Standards Institute	CAE	Computer-Aided Engineering
APAR	Authorized Program Analysis Report	CAM	Computer-Aided Manufacturing
API	Application Programming Interface	CATE	Certified Advanced Technical Expert
ARP	Address Resolution Protocol	CATIA	Computer-Graphics Aided Three-Dimensional Interactive Application
ASCI	Accelerated Strategic Computing Initiative	CCM	Common Character Mode
ASCII	American National Standards Code for Information Interchange	CD	Compact Disk
ASR	Address Space Register	CDE	Common Desktop Environment
ATM	Asynchronous Transfer Mode	CDLI	Common Data Link Interface
AuditRM	Audit Log Resource Manager	CD-R	CD Recordable
AUI	Attached Unit Interface	CD-ROM	Compact Disk-Read Only Memory
AWT	Abstract Window Toolkit	CE	Customer Engineer
BCT	Branch on Count	CEC	Central Electronics Complex
BFF	Backup File Format	CFD	Computational Fluid Dynamics
BI	Business Intelligence	CGE	Common Graphics Environment

CHRP	Common Hardware Reference Platform	DFS	Distributed File System
CISPR	International Special Committee on Radio Interference	DHCP	Dynamic Host Configuration Protocol
CLIO/S	Client Input/Output Sockets	DIMM	Dual In-Line Memory Module
CLVM	Concurrent LVM	DIP	Direct Insertion Probe
CMOS	Complimentary Metal-Oxide Semiconductor	DIT	Directory Information Tree
CMP	Certificate Management Protocol	DIVA	Digital Inquiry Voice Answer
COFF	Common Object File Format	DLT	Digital Linear Tape
COLD	Computer Output to Laser Disk	DMA	Direct Memory Access
CPU	Central Processing Unit	DMT	Directory Management Tool
CRC	Cyclic Redundancy Check	DN	Distinguished Name
CSID	Character Set ID	DNS	Domain Naming System
CSR	Customer Service Representative	DOE	Department of Energy
CSS	Communication Subsystems Support	DOI	Domain of Interpretation
CSU	Customer Set-Up	DOS	Disk Operating System
CWS	Control Workstation	DPCL	Dynamic Probe Class Library
DAD	Duplicate Address Detection	DRAM	Dynamic Random Access Memory
DAS	Dual Attach Station	DS	Differentiated Service
DASD	Direct Access Storage Device	DSA	Dynamic Segment Allocation
DAT	Digital Audio Tape	DSE	Diagnostic System Exerciser
DBCS	Double Byte Character Set	DSMIT	Distributed SMIT
DBE	Double Buffer Extension	DSU	Data Service Unit
DC	Direct Current	DTE	Data Terminating Equipment
DCE	Distributed Computing Environment	DW	Data Warehouse
DDC	Display Data Channel	EA	Effective Address
DDS	Digital Data Storage	EC	Engineering Change
DE	Dual-Ended	ECC	Error Checking and Correcting
DES	Data Encryption Standard	EEPROM	Electrically Erasable Programmable Read Only Memory
DFL	Divide Float	EFI	Extensible Firmware Interface
DFP	Dynamic Feedback Protocol	EHD	Extended Hardware Drivers
		EIA	Electronic Industries Association

EISA	Extended Industry Standard Architecture	FRCA	Fast Response Cache Architecture
ELA	Error Log Analysis	FRU	Field Replaceable Unit
ELF	Executable and Linking Format	FSRM	File System Resource Manager
EMU	European Monetary Union	FTP	File Transfer Protocol
EOF	End of File	FTP	File Transfer Protocol
EPOW	Environmental and Power Warning	GAI	Graphic Adapter Interface
ERRM	Event Response resource manager	GAMESS	General Atomic and Molecular Electronic Structure System
ESID	Effective Segment ID	GPFS	General Parallel File System
ESP	Encapsulating Security Payload	GPR	General-Purpose Register
ESSL	Engineering and Scientific Subroutine Library	GUI	Graphical User Interface
ETML	Extract, Transformation, Movement, and Loading	GUID	Globally Unique Identifier
F/C	Feature Code	HACMP	High Availability Cluster Multi Processing
F/W	Fast and Wide	HACWS	High Availability Control Workstation
FC	Fibre Channel	HCON	IBM AIX Host Connection Program/6000
FCAL	Fibre Channel Arbitrated Loop	HDX	Half Duplex
FCC	Federal Communication Commission	HFT	High Function Terminal
FCP	Fibre Channel Protocol	HIPPI	High Performance Parallel Interface
FDDI	Fiber Distributed Data Interface	HIPS	High Performance Switch
FDPR	Feedback Directed Program Restructuring	HiPS LC-8	Low-Cost Eight-Port High Performance Switch
FDX	Full Duplex	HMC	Hardware Management Console
FIFO	First In/First Out	HostRM	Host Resource Manager
FLASH EPROM	Flash Erasable Programmable Read-Only Memory	HP	Hewlett-Packard
FLIH	First Level Interrupt Handler	HPF	High Performance FORTRAN
FMA	Floating point Multiply Add operation	HPSSDL	High Performance Supercomputer Systems Development Laboratory
FPR	Floating Point Register	HP-UX	Hewlett-Packard UNIX
FPU	Floating Point Unit	HTML	Hyper-text Markup Language
		HTTP	Hypertext Transfer Protocol

Hz	Hertz	IRQ	Interrupt Request
I/O	Input/Output	IS	Integrated Service
I²C	Inter Integrated-Circuit Communications	ISA	Industry Standard Architecture, Instruction Set Architecture
IAR	Instruction Address Register	ISAKMP	Internet Security Association Management Protocol
IBM	International Business Machines	ISB	Intermediate Switch Board
ICCCM	Inter-Client Communications Conventions Manual	ISDN	Integrated-Services Digital Network
ICE	Inter-Client Exchange	ISMP	InstallShield Multi-Platform
ICElib	Inter-Client Exchange library	ISNO	Interface Specific Network Options
ICMP	Internet Control Message Protocol	ISO	International Organization for Standardization
ID	Identification	ISV	Independent Software Vendor
IDE	Integrated Device Electronics	ITSO	International Technical Support Organization
IDS	Intelligent Decision Server	JBOD	Just a Bunch of Disks
IEEE	Institute of Electrical and Electronics Engineers	JDBC	Java Database Connectivity
IETF	Internet Engineering Task Force	JFC	Java Foundation Classes
IHV	Independent Hardware Vendor	JFS	Journaled File System
IIOB	Internet Inter-ORB Protocol	JTAG	Joint Test Action Group
IJG	Independent JPEG Group	KDC	Key Distribution Center
IKE	Internet Key Exchange	L1	Level 1
ILS	International Language Support	L2	Level 2
IM	Input Method	L2	Level 2
INRIA	Institut National de Recherche en Informatique et en Automatique	LAN	Local Area Network
IP	Internetwork Protocol (OSI)	LANE	Local Area Network Emulation
IPL	Initial Program Load	LAPI	Low-Level Application Programming Interface
IPSec	IP Security	LDAP	Lightweight Directory Access Protocol
IrDA	Infrared Data Association (which sets standards for infrared support including protocols for data interchange)	LDIF	LDAP Directory Interchange Format
		LED	Light Emitting Diode
		LFD	Load Float Double

LFT	Low Function Terminal	MODS	Memory Overlay Detection Subsystem
LID	Load ID	MP	Multiprocessor
LLNL	Lawrence Livermore National Laboratory	MPC-3	Multimedia PC-3
LP	Logical Partition	MPI	Message Passing Interface
LP64	Long-Pointer 64	MPOA	Multiprotocol over ATM
LPI	Lines Per Inch	MPP	Massively Parallel Processing
LPP	Licensed Program Product	MPS	Mathematical Programming System
LPR/LPD	Line Printer/Line Printer Daemon	MST	Machine State
LRU	Least Recently Used	MTU	Maximum Transmission Unit
LTG	Logical Track Group	MWCC	Mirror Write Consistency Check
LV	Logical Volume	MX	Mezzanine Bus
LVCB	Logical Volume Control Block	NBC	Network Buffer Cache
LVD	Low Voltage Differential	NCP	Network Control Point
LVM	Logical Volume Manager	ND	Neighbor Discovery
MAP	Maintenance Analysis Procedure	NDP	Neighbor Discovery Protocol
MASS	Mathematical Acceleration Subsystem	NFB	No Frame Buffer
MAU	Multiple Access Unit	NFS	Network File System
MBCS	Multi-Byte Character Support	NHRP	Next Hop Resolution Protocol
Mbps	Megabits Per Second	NIM	Network Installation Management
MBps	Megabytes Per Second	NIS	Network Information System
MCA	Micro Channel Architecture	NL	National Language
MCAD	Mechanical Computer-Aided Design	NLS	National Language Support
MDI	Media Dependent Interface	NT-1	Network Terminator-1
MES	Miscellaneous Equipment Specification	NTF	No Trouble Found
MFLOPS	Million of Floating point Operations Per Second	NTP	Network Time Protocol
MII	Media Independent Interface	NUMA	Non-Uniform Memory Access
MIP	Mixed-Integer Programming	NUS	Numerical Aerodynamic Simulation
MLR1	Multi-Channel Linear Recording 1	NVRAM	Non-Volatile Random Access Memory
MMF	Multi-Mode Fibre	NWP	Numerical Weather Prediction
		OACK	Option Acknowledgment
		OCS	Online Customer Support

ODBC	Open DataBase Connectivity	PKR	Protection Key Registers
ODM	Object Data Manager	PMTU	Path MTU
OEM	Original Equipment Manufacturer	POE	Parallel Operating Environment
OLAP	Online Analytical Processing	POP	Power-On Password
OLTP	Online Transaction Processing	POSIX	Portable Operating Interface for Computing Environments
ONC+	Open Network Computing	POST	Power-On Self-test
OOUI	Object-Oriented User Interface	POWER	Performance Optimization with Enhanced Risc (Architecture)
OSF	Open Software Foundation, Inc.	PPC	PowerPC
OSL	Optimization Subroutine Library	PPM	Piecewise Parabolic Method
OSLp	Parallel Optimization Subroutine Library	PPP	Point-to-Point Protocol
P2SC	POWER2 Single/Super Chip	PREP	PowerPC Reference Platform
PAM	Pluggable Authentication Mechanism	PSE	Portable Streams Environment
PAP	Privileged Access Password	PSSP	Parallel System Support Program
PBLAS	Parallel Basic Linear Algebra Subprograms	PTF	Program Temporary Fix
PCI	Peripheral Component Interconnect	PTPE	Performance Toolbox Parallel Extensions
PDT	Paging Device Table	PTX	Performance Toolbox
PDU	Power Distribution Unit	PV	Physical Volume
PE	Parallel Environment	PVC	Permanent Virtual Circuit
PEDB	Parallel Environment Debugging	PVID	Physical Volume Identifier
PEX	PHIGS Extension to X	QMF	Query Management Facility
PFS	Perfect Forward Security	QoS	Quality of Service
PGID	Process Group ID	QP	Quadratic Programming
PHB	Processor Host Bridges	RAID	Redundant Array of Independent Disks
PHY	Physical Layer	RAM	Random Access Memory
PID	Process ID	RAN	Remote Asynchronous Node
PID	Process ID	RAS	Reliability, Availability, and Serviceability
PIOFS	Parallel Input Output File System	RDB	Relational DataBase
		RDBMS	Relational Database Management System

RDISC	ICMP Router Discovery	SCSI	Small Computer System Interface
RDN	Relative Distinguished Name	SCSI-SE	SCSI-Single Ended
RDP	Router Discovery Protocol	SDLC	Synchronous Data Link Control
RFC	Request for Comments	SDR	System Data Repository
RIO	Remote I/O	SDRAM	Synchronous Dynamic Random Access Memory
RIP	Routing Information Protocol	SE	Single Ended
RIPL	Remote Initial Program Load	SEPBU	Scalable Electrical Power Base Unit
RISC	Reduced Instruction-Set Computer	SGI	Silicon Graphics Incorporated
RMC	Resource Monitoring and Control	SGID	Set Group ID
ROLTP	Relative Online Transaction Processing	SHLAP	Shared Library Assistant Process
RPA	RS/6000 Platform Architecture	SID	Segment ID
RPC	Remote Procedure Call	SIT	Simple Internet Transition
RPL	Remote Program Loader	SKIP	Simple Key Management for IP
RPM	Redhat Package Manager	SLB	Segment Lookaside Buffer
RSC	RISC Single Chip	SLIH	Second Level Interrupt Handler
RSCT	Reliable Scalable Cluster Technology	SLIP	Serial Line Internet Protocol
RSE	Register Stack Engine	SLR1	Single-Channel Linear Recording 1
RSVP	Resource Reservation Protocol	SM	Session Management
RTC	Real-Time Clock	SMB	Server Message Block
RVSD	Recoverable Virtual Shared Disk	SMIT	System Management Interface Tool
SA	Secure Association	SMP	Symmetric Multiprocessor
SACK	Selective Acknowledgments	SMS	System Management Services
SAN	Storage Area Network	SNG	Secured Network Gateway
SAR	Solutions Assurance Review	SOI	Silicon-on-Insulator
SAS	Single Attach Station	SP	IBM RS/6000 Scalable POWER parallel Systems
SBCS	Single-Byte Character Support	SP	Service Processor
ScaLAPACK	Scalable Linear Algebra Package	SPCN	System Power Control Network
SCB	Segment Control Block		

SPEC	System Performance Evaluation Cooperative	UCS	Universal Coded Character Set
SPI	Security Parameter Index	UDB EEE	Universal Database and Enterprise Extended Edition
SPM	System Performance Measurement	UDI	Uniform Device Interface
SPOT	Shared Product Object Tree	UIL	User Interface Language
SPS	SP Switch	ULS	Universal Language Support
SPS-8	Eight-Port SP Switch	UP	Uniprocessor
SRC	System Resource Controller	USB	Universal Serial Bus
SRN	Service Request Number	USLA	User-Space Loader Assistant
SSA	Serial Storage Architecture	UTF	UCS Transformation Format
SSC	System Support Controller	UTM	Uniform Transfer Model
SSL	Secure Socket Layer	UTP	Unshielded Twisted Pair
STFDU	Store Float Double with Update	UUCP	UNIX-to-UNIX Communication Protocol
STP	Shielded Twisted Pair	VESA	Video Electronics Standards Association
SUID	Set User ID	VFB	Virtual Frame Buffer
SUP	Software Update Protocol	VG	Volume Group
SVC	Switch Virtual Circuit	VGDA	Volume Group Descriptor Area
SVC	Supervisor or System Call	VGSA	Volume Group Status Area
SWVPD	Software Vital Product Data	VHDCI	Very High Density Cable Interconnect
SYNC	Synchronization	VLAN	Virtual Local Area Network
TCE	Translate Control Entry	VMM	Virtual Memory Manager
Tcl	Tool Command Language	VP	Virtual Processor
TCP/IP	Transmission Control Protocol/Internet Protocol	VPD	Vital Product Data
TCQ	Tagged Command Queuing	VPN	Virtual Private Network
TGT	Ticket Granting Ticket	VSD	Virtual Shared Disk
TLB	Translation Lookaside Buffer	VSM	Visual System Manager
TOS	Type Of Service	VSS	Versatile Storage Server
TPC	Transaction Processing Council	VT	Visualization Tool
TPP	Toward Peak Performance	WAN	Wide Area Network
TSE	Text Search Engine	WLM	Workload Manager
TSE	Text Search Engine	WTE	Web Traffic Express
TTL	Time To Live		

XCOFF	Extended Common Object File Format
XIE	X Image Extension
XIM	X Input Method
XKB	X Keyboard Extension
XLF	XL Fortran
XOM	X Output Method
XPM	X Pixmap
XSSO	Open Single Sign-on Service
XTF	Extended Distance Feature
XVFB	X Virtual Frame Buffer

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 207.

- ▶ *AIX 5L Differences Guide*, SG24-5765
- ▶ *AIX Logical Volume Manager, from A to Z: Introduction and Concepts*, SG24-5432
- ▶ *AIX 5L Performance Tools Handbook*, SG24-6039
- ▶ *IBM @server Certification Study Guide - AIX Communications*, SG24-6186
- ▶ *IBM @server Certification Study Guide - AIX Performance and System Tuning*, SG24-6184
- ▶ *IBM @server Certification Study Guide - AIX Problem Determination Tools and Techniques*, SG24-6185
- ▶ *IBM @server Certification Study Guide - pSeries AIX System Administration*, SG24-6191
- ▶ *IBM @server Certification Study Guide - pSeries AIX System Support*, SG24-6199
- ▶ *IBM @server Certification Study Guide - pSeries HACMP for AIX*, SG24-6187
- ▶ *IBM @server Certification Study Guide - RS/6000 SP*, SG24-5348
- ▶ *Managing AIX Server Farms*, SG24-6606
- ▶ *NIM from A to Z in AIX 4.3*, SG24-5524
- ▶ *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496
- ▶ *TCP/IP Tutorial and Technical Overview*, GG24-3376

Other resources

These publications are also relevant as further information sources:

- ▶ *AIX 5L Version 5.2 Installation in a Partitioned Environment*, SC23-4382
- ▶ *AIX Version 4.3 System Management Concepts: Operating System and Devices*, SC23-4311
- ▶ *AIX Version 5.0 Installation Guide*, SC23-4112
- ▶ *AIX Version 5.1 Network Installation Management Guide and Reference*, SC23-4385
- ▶ *Message Guide and References*, which is part of the AIX Version 4 base documentation
- ▶ *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538
- ▶ *SSA Adapters: User's Guide and Maintenance Information*, SA33-3272
- ▶ The following types of documentation are located through the Internet at the following URL:

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

- User guides
- System management guides
- Application programmer guides
- All command reference volumes
- Files reference
- Technical reference volumes used by application programmers

Referenced Web sites

These Web sites are also relevant as further information sources:

- ▶ IBM certification test information
<http://www.ibm.com/certify/tests/info.shtml>
- ▶ IBM eServer pSeries and IBM RS/6000 server hardware documentation
http://www.ibm.com/servers/eserver/pseries/library/hardware_docs/index.html
- ▶ IBM eServer pSeries support
<http://techsupport.services.ibm.com/server/support?view=pSeries>
- ▶ IBM eServer pSeries support - AIX 4.3 OS, Java, compilers
<http://techsupport.services.ibm.com/server/nav?fetch=a40jc>

- ▶ IBM Redbooks home page
<http://www.redbooks.ibm.com>
- ▶ IBM Storage
<http://www.storage.ibm.com>
- ▶ Open Group Web site
<http://www.opengroup.org/onlinepubs/9629799/toc.htm>
- ▶ Professional certification programs from IBM
<http://www.ibm.com/certify>
- ▶ UNIX servers (pSeries)
http://www-132.ibm.com/content/home/store_IBMPublicUSA/en_US/eServer/pSeries/pSeries.html

How to get IBM Redbooks

You can order hardcopy Redbooks, as well as view, download, or search for Redbooks at the following Web site:

ibm.com/redbooks

You can also download additional materials (code samples or diskette/CD-ROM images) from that site.

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.

Index

Symbols

\$HOME/.rhosts 166
.dtprofile
 files
 .dtpfile 22
/bosinst.data 155
/etc/exclude.rootvg 159
/etc/filesystems 110
/etc/hosts 22
/etc/hosts.equiv 166
/etc/resolv.conf 22
/image.data 155

Numerics

0503-350 168
0511-051 168
0511-080 168
0511-133 168
0512-016 168
911 191

A

alog 65, 68
alt_disk_install command 17, 25
alternate disk installation 16
 cloning rootvg 18
APAR 31
Authorized Program Analysis Report 31

B

backup 154
 commands
 backup 156
 cpio 162
 tar 161
 file system 156
 volume group
 non-rootvg 156
 rootvg 154
BIST 54–55
 LED codes 56
BLV 54

 content 55
 how to recreate 59
blvset command 12, 24
boot 70
 /etc/inittab 55, 69–70
 /etc/inittab figure 71
 /mnt 64
 alog 65, 68
 BIST 54–55
 BLV 54
 bootinfo 63
 bootlist 56, 60, 66
 bosboot 60, 66, 68
 cfgmgr 63, 70
 Config_Rules 63
 content of the BLV 55
 copycore 64
 diag menu figure 1 57
 from CD-ROM 8
 general workflow 54
 general workflow figure 54
 identifier 71
 init 55
 ipl_varyon 64
 LED codes
 100 - 195 56
 200 - 2E7 56
 BIST 56
 POST 56
 mergedev 65
 normal 56
 PCI normal boot 61
 PCI service boot 61
 phase1 55, 62, 74
 phase1 workflow figure 62
 phase2 55, 64, 74
 phase2 figure 1 64
 phase2 figure 2 65
 phase3 69, 74
 phase3 figure 69
 POST 54–55
 RAMFS 54
 rc.boot 74
 restbase 63

- runlevel 56, 71
- savebase 70
- service 56
- SMS 60
- SMS main menu figure 61
- syncvg 70
- boot logical volume
 - BLV 54
 - content 55
 - how to recreate 59
- boot problem tips 132
- bootinfo 157
- bootinfo -p 20
- bootlist 157
- bootlist command
 - flags 134
 - syntax 134
- bootpd daemon 178
- bosboot 155
- bosboot command 122
 - flags 134
 - syntax 133
- bundle 31, 37
 - app-dev 31
 - bundle install figure 38
 - client 31
 - create your own 38
 - graphics-startup 31
 - media-defined 31, 38
 - pers-prod 31
 - server 31

C

- CDE common installation problems 22
- cfgcon 70
- cfgmgr command 21, 24, 90
 - flags 133
 - syntax 133
- chdev command 84
- chps command 122
 - flags 135
 - syntax 135
- Cloning 19
- cloning rootvg 18
- command syntax
 - showmount 187
- commands
 - alog 65, 68
 - alt_disk_install 17, 25
 - backup 156
 - bffcreate 39
 - blvset 12, 24
 - bootinfo 63, 157
 - bootlist 56, 66, 157
 - bosboot 60, 66, 68, 122, 155
 - cfgcon 70
 - cfgmgr 21, 24, 63, 70, 90
 - chdev 84
 - chps 122
 - copycore 64
 - cplv 106
 - dd 166
 - df 116, 118
 - exportfs 180
 - fsck 66, 91, 121
 - fuser 112
 - importvg 91
 - install_assist 14
 - installp 30, 33, 35, 38, 45–46, 48
 - instfix 30, 34, 37, 49
 - inutoc 39
 - ipl_varyon 64
 - logform 66–67, 121
 - lppchk 45
 - lsattr 84
 - lsdev 130–131
 - lsfs 188
 - lslpp 33, 40, 45, 50
 - lslv 102, 104–105, 120, 131
 - lsps 122
 - lspv 84, 87–88, 90, 104
 - lsvg 84, 91, 96, 131
 - mergedev 65
 - migratepv 89, 105
 - mklv 121–122
 - mkps 122
 - mksysb 119, 154, 157
 - mkszfile 116, 155
 - mkvgdata 116, 156
 - mount 112–113
 - oslevel 15
 - reducevg 118, 130
 - refresh 178
 - replacepv 92
 - restbase 63
 - restore 31, 158
 - restvg 118, 158

- rmdev 130
- rmfs 120
- rmlv 106, 121
- rmps 122, 126
- savebase 67, 70
- savevg 117, 156
- showmount 187
- swapoff 126
- swapon 122
- synclvdm 132
- syncvg 70
- sysdumpdev 122
- tcopy 166
- tctl 158
- telinit 71
- umount 91, 112–113, 130
- varyonvg 91
- Config_Rules 63
- configuration assistant taskguide 14
- copycore 64
- cplv command 106
 - flags 136
 - syntax 135
- creating paging space 122

D

- dd 166
- df command 116, 118
- DTSOURCEPROFILE in .dtprofile 22

E

- emergency fix 31
- environment variable
 - DTSOURCEPROFILE 22
- exportfs 180

F

- figures
 - /etc/inittab 71
 - diag alter bootlist menu 58
 - diag function selection menu 57
 - diag task selection menu 58
 - general boot order 54
 - installing user defined bundle 38
 - preview option in smitty 40
 - saving previous version of fileset 42
 - smitty installp 32

- smitty menu for rejecting fileset 43
- smitty software maintenance menu 43
- smitty software remove menu 46
- SMS main menu 61
- workflow phase 2 figure 2 65
- workflow phase1 62
- workflow phase2 figure 1 64
- workflow phase3 69

file

- image.data 159
- rc.boot 72

file system

- backup 156
- removing 120
- restore 158

file systems 110

- creating JFS log 120
- mounting 111–112
- over mounted 112
- reducing size 114
- unmounting 111

files 30

- \$HOME/.rhosts 166
- /bosinst.data 155
 - customize 159
- /etc/exclude.rootvg 159
- /etc/filesystems 120
- /etc/hosts 22
- /etc/hosts.equiv 166
- /etc/inittab 55, 69–70
- /etc/resolv.conf 22
- /image.data 155
- /tmp/vgdata/vgname/vgname.data 116
- /usr/sys/inst.images/sys_bundles 22
- image.data 114
- rc.boot 55, 63–64
- resolv.conf 22
- smit.log 35

fileset 31, 38

- applied 40
- broken 44
- cleanup 45
- commit or apply table 41
- committed 40
- install same or older version 44
- reinstallation 44
- rejected 40
- removed 40

- fileset inconsistent 44

- fixed disk structures 81
- fsck command 91, 121
 - flags 137
 - syntax 137
- fuser command 112
 - file usage 139
 - flags 139
 - syntax 139

G

- generic device names 56

I

- image.data file 114
- importvg command 91
- init
 - phase1 55
 - phase2 55
 - phase3 55
- INST_DEBUG 36
- install_assist see configuration assistant taskguide 14
- installation
 - additional hardware 20
 - alternate disk installation 16
 - boot up from CD-ROM 8
 - bundle 21
 - configuration assistant taskguide 14
 - determining maintenance level 15
 - e-fix 31
 - failed 45
 - fileset 31
 - interrupted 45
 - licensed program package 31
 - methods 8, 32
 - migration 8, 10, 26
 - new and complete overwrite 8, 10
 - preservation 8, 13–14, 26
 - missing device driver 21
 - online documentation 23
 - options 30
 - post installation problems 22
 - problems with
 - CDE 22
 - TCP/IP 22
 - PTF 31
 - reports 35
 - settings 9

- startup 8
- trusted computing base (TCB) 9
- uninstall 46

J

- journaled file system (JFS) 81

L

- LED codes
 - 100 - 195 56
 - 200 59
 - 200 - 2E7 56
 - 299 59
 - 511 63
 - 518 64, 68
 - 548 63
 - 551 66
 - 552 64, 66
 - 553 72
 - 554 64, 66
 - 555 64, 66
 - 556 64, 66
 - 557 64, 66
 - BIST 56
 - c31 70, 72
 - c32 70
 - c33 70
 - c34 70
 - Common MCA LED code table 59, 74
 - common phase2 LED codes 75
 - common phase3 LED codes 75
 - POST 56
- logform command 121
- logical partitions (LPs) 81
- logical volume manager (LVM) 81
- logical volumes 95
 - copying 106
 - creating 96
 - migrating 103
 - paging space 122
 - removing 106
- logical volumes (LVs) 81
- LPP 31
- lppchk 45
- lppchk command 45
- lsattr command 84
- lscfg -vp | grep Arch 20
- lsdev command 130–131

- lsfs 188
- lslpp 33, 40
- lslpp command 33, 50
- lsiv command 102, 104–105, 120, 131
- lsps command 122
- lspv command 84, 87–88, 90, 104
- lsvg command 84, 91, 96, 131
- LVM 81

M

- magic number 66
- maintenance level 15
- MCA
 - BIST 55
 - LED 56
 - 200 59
 - 299 59
 - LED code table 59, 74
 - POST 55
- Metadata description 110
- migratepv command 89, 105
 - flags 140
 - syntax 140
- migration installation 10
 - installation steps 11
 - missing option 12
- missing device driver 21
- mkboot command
 - flags 140
 - syntax 140
- mklv command 121–122
 - flags 141
 - syntax 141
- mkps command 122
 - flags 144
 - syntax 144
- mksysb 19, 154, 157
- mksysb command 119
- mksysb tape
 - extracting data 158
 - layout 154
- mkszfile 155
- mkszfile command 116
- mkvgdata 156
- mkvgdata command 116
- mount command 91, 112–113

N

- network boot
 - problems 177
- Network Interface Management
 - see NIM 175
- new and complete overwrite installation 10
- NIM 175
- NIM debug, 911 191
- normal boot 56
 - PCI 61

O

- online documentation installation 23
- oslevel command 15

P

- PCI systems 60
 - bootlist 60
 - normal boot 61
 - service boot 61
 - service guide URL 62
 - SMS 60
 - SMS main menu figure 61
- phase1 55, 62, 74
 - bootinfo 63
 - cfgmgr 63
 - Config_Rules 63
 - init 55
 - restbase 63
 - workflow figure 62
- phase2 55, 64, 74
 - alog 65
 - copycore 64
 - figure 1 64
 - figure 2 65
 - init 55
 - ipl_varyon 64
 - mergedev 65
 - rootvg 55
- phase3 69, 74
 - /etc/inittab 55, 69–70
 - /etc/inittab figure 71
 - cfgcon 70
 - cfgmgr 70
 - init 55
 - rc.boot 72
 - savebase 70
 - syncvg 70

- workflow figure 69
- physical partitions
 - intra-disk policy 99
- physical partitions (PPs) 81
- physical volume (PV) 81
- physical volumes
 - incorrectly removed 130
 - removing 129
- POST 54–55
 - LED codes 56
- preservation installation 13
- program temporary fix 31
- ptf 31

R

- RAMFS 54
- rc.boot 74
 - /etc/inittab 55
 - init 55
- recovery 157
- Redbooks Web site 207
 - Contact us xviii
- reducevg command 118, 130
- refresh 178
- reject fileset 43
- replacepv command 92
- resintallation 44
- resolv.conf 22
- restore 158
 - file system 158
 - volume group
 - non-rootvg 158
 - rootvg 157
- restvg 158
- restvg command 118
 - flags 146
 - syntax 146
- rmdev command 130
- rmfs command 120
- rmlv command 106, 121
- rmpps command 122, 126
- rootvg
 - phase2 55
- runlevel 56, 71

S

- savevg 156
- savevg command 117

- service boot 56
 - PCI 61
- showmount 187
- SMIT fast path
 - smit install_bundle 22
- smit lv 97
- smit vg 84
- smitty
 - ismksysb 158
- smitty menus
 - easy_install 37
 - installp 32
 - preview only 40
 - reject fileset 43
 - remove 37
 - saving previous version of fileset 42
 - software maintenance 43
 - software removal 46
 - update_by_fix 31
- superblock 67
- swapoff command 126
- swapon command 122
- synclvodm command 132
- sysdumpdev command 122
 - flags 147
 - syntax 147

T

- tables
 - boot phase2 LED codes 75
 - boot phase3 LED codes 75
 - commit or apply 41
 - common MCA LED codes 59, 74
 - installp flags 33, 39
 - reinstallation options 44
- tape
 - block size 166
- tcopy 166
- TCP/IP installation problems 22
- tctl 158
- telinit command 71
- trusted computing base 9

U

- umount command 112–113, 130

V

- varyonvg command 91
- volume group
 - non-rootvg
 - backup 156
 - restore 158
 - rootvg
 - backup 154
 - restore 157
- volume group (VG) 81
- volume groups 83
 - associated logical volumes 102
 - creating 83
 - data migration 88
 - moving disks 89
 - reducing 93
 - reorganize 94

W

- wake-up alternate volume group 19
- workflow 54





IBM server Certification Study Guide - AIX 5L Installation and System Recovery



Developed specifically for the purpose of preparing for AIX certification

Makes an excellent companion to classroom education

For experienced AIX professionals

This IBM Redbook is designed as a study guide for professionals wishing to prepare for the AIX 5L Installation and System Recovery certification exam as a selected course of study in order to achieve the IBM server Certified Advanced Technical Expert - pSeries and AIX 5L certification.

This redbook is designed to provide a combination of theory and practical experience needed for a general understanding of the subject matter. It also provides sample questions that will help in the evaluation of personal progress and provide familiarity with the types of questions that will be encountered in the exam.

This publication does not replace practical experience, nor is it designed to be a stand-alone guide for any subject. Instead, it is an effective tool that, when combined with education activities and experience, can be a very useful preparation guide for the exam. Whether you are planning to take the AIX 5L Installation and System Recovery certification exam, or you just want to validate your AIX skills, this redbook is for you.

This publication was updated to include the new content included in Test 233, which was based on AIX 5L Version 5.1.

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**