IBM

Cluster Systems Management Cookbook for pSeries

- AIX 5L running on the management server
- AIX and Linux nodes in the cluster
- Migration scenarios included

Dino Quintero
Thomas Braunbeck
Ong Swee Thye
Andrei Vlad
Peter Zutenis

Redbooks

ibm.com/redbooks

International Technical Support Organization

Cluster Systems Management Cookbook for pSeries

November 2004

SG24-6859-01

Note: Before using this information and the product it supports, read the information in "Notices" on page ix.

Second Edition (November 2004)

This edition applies to AIX 5L Version 5, Release 3, Fix 10, and Cluster Systems Management (CSM) for AIX Version 1, Release 4, Modification 0, Fix 10.

Contents

Notices ix
Trademarks x

Preface xi
The team that wrote this redbook. xi
Become a published author xii
Comments welcome. xiii

Chapter 1. Introduction 1
1.1 What is Cluster Systems Management (CSM)? 2
1.2 CSM 1.4.0 highlights 2
1.3 CSM differentiators 2
1.4 Why consider CSM? 3
1.5 What is covered in this redbook 4

Chapter 2. New features in CSM 1.4.0 5
2.1 New hardware support 6
2.2 New software support 6
2.3 CSM Highly Available Management Server 7
2.3.1 Hardware support 7
2.3.2 Software support. 7
2.4 Monitoring Service Focus Point using CSM 7
2.5 Changes in Linux distribution support 8
2.6 1024 way xSeries nodes support 8
2.7 Hardware control command 8

Chapter 3. CSM installation scenario 9
3.1 Assumptions and planning 10
3.2 Our environment 10
3.3 Installation of the CSM management server 11
3.3.1 Installation of AIX 12
3.3.2 Apply maintenance to AIX. 12
3.3.3 Installation and customization of NIM 13
3.3.4 Install and apply maintenance to the CSM server. 16
3.3.5 Customize the CSM server 20
3.4 Installation of nodes 23
3.4.1 Define nodes to CSM 23
3.4.2 Node definition 23
3.4.3 Create nodegroups 25

3.4.4 Validate hardware control 25
3.4.5 Get adapter information 26
3.5 Set up cluster file management. 27
3.6 NIM steps 27
3.6.1 Define NIM clients 27
3.6.2 NIM machine groups 28
3.6.3 NIM CSM preparation 28
3.6.4 Add OpenSSL and OpenSSH bundles 28
3.6.5 NIM resource allocation 30
3.7 Network boot node 31
3.7.1 Verify resources 32
3.7.2 Initiate netboot. 33
3.7.3 Monitoring the install 33
3.8 Integration of Linux pSeries nodes 35
3.8.1 Check the node software 36
3.8.2 Copy Linux packages to the correct directory 36
3.8.3 Verify node definitions. 37
3.8.4 Create CSM node groups 37
3.8.5 Set up the cluster configuration. 38
3.8.6 Add Linux nodes to the cluster 38
3.8.7 Verify nodes 38

Chapter 4. CSM advanced features implementation 39
4.1 Building a Linux repository for pSeries 40
4.1.1 Setting up the network installation server 40
4.1.2 Red Hat repository 45
4.1.3 SUSE LINUX SLES8 repository 45
4.1.4 SUSE LINUX SLES9 repository 46
4.1.5 Creating the autoinstallation files 47
4.1.6 Install the CSM client on Linux 58
4.2 Install Linux using the AIX NIM server 60
4.2.1 Switch the netboot method from BOOTP to DHCP. 60
4.2.2 Prepare AIX 5L V5.x NIM machine definitions 61
4.2.3 Activating NIM clients to create entries in /etc/bootptab 62
4.2.4 Converting BOOTP information to DHCP syntax 62
4.2.5 Deallocate NIM clients to remove entries from /etc/bootptab 64
4.2.6 Deactivate BOOTP service from inetd 64
4.2.7 Configure and activate DHCP. 64
4.2.8 Booting Linux on pSeries 66
4.3 CSM performance, scalability, and challenges 66
4.3.1 Firewall 66
4.3.2 WAN network 68

Chapter 5. CSM migration scenario 71
5.1 General planning for migrations 72
5.1.1 Where we want to go 72
5.1.2 Where we come from 73
5.1.3 Gather prerequisites information 74
5.1.4 Checking for latest updates available 75
5.1.5 Planning for recovery 76
5.2 Migrate CSM on the management server 78
5.2.1 Get the necessary software 79
5.2.2 Check for prerequisites on client nodes and HMCs 80
5.2.3 Back up the management server 80
5.2.4 Commit previous installed updates 80
5.2.5 Update and migrate the software 81
5.2.6 Update the open source software packages 82
5.2.7 Reboot the management server 82
5.2.8 Accept the terms and conditions of the CSM license agreement . . . 82
5.2.9 Copy updated CSM files into /csminstall 82
5.2.10 Verify the installation 82
5.2.11 Perform a backup of the migrated system 83
5.3 Migrate CSM on AIX client nodes 83
5.3.1 Check for application software prerequisites 84
5.3.2 Back up the client nodes 84
5.3.3 Commit previously installed updates 84
5.3.4 Stop the applications on the client nodes 84
5.3.5 Update and migrate the nodes 85
5.3.6 Reboot the client nodes 88
5.3.7 Updatenode (when necessary) 88
5.3.8 Verify CSM functions 89
5.3.9 Start up and verify the applications on the client nodes 89
5.3.10 Back up client nodes 89
5.4 Migrate AIX on the client nodes 90
5.4.1 Alternate disk migration installation 90
5.4.2 Traditional AIX migration using NIM 93
5.5 Migrate AIX on the management server 96
5.5.1 Get the necessary software 97
5.5.2 Check for prerequisites on client nodes and HMCs 98
5.5.3 Back up the management server 98
5.5.4 Perform the migration 98
5.5.5 Apply updates to the migrated system 98
5.5.6 Update the open source software packages 99
5.5.7 Post migration actions 99

Chapter 6. CSM cluster administration 101

6.1 Use all available information sources 102
6.2 Maintaining NIM lpp_sources for one AIX release 104
6.3 Verify the AIX level of SPOT and backup 106
6.4 Node status in csmstat command output 107
6.5 Hardware control problems 109
6.5.1 Invalid hardware control point message 110
6.5.2 Error connecting to hardware control point 112
6.5.3 Power status undefined. 112
6.5.4 Power on for a client node is not working 114
6.5.5 IBM.HWCTRLRM trace and Java controls 115

Chapter 7. CSM High Availability Management Server 117
7.1 When to deploy CSM HA MS 118
7.2 HA MS installation requirement. 118
7.2.1 Software support 118
7.2.2 Storage support. 118
7.2.3 Network consideration. 119
7.2.4 Remote shell consideration. 119
7.2.5 Network Installation Management (NIM) consideration 120
7.2.6 Symbolic link consideration. 120
7.3 HA MS implementation scenarios 120
7.3.1 Environment setup 120
7.3.2 CSM MS installation on the backup MS 122
7.4 HA MS configuration 124
7.4.1 HA MS installation. 124
7.4.2 Set up NTP on both management servers 125
7.4.3 Shared disk configuration 126
7.4.4 Create the HA MS definition file 130
7.4.5 Customized scripts 131
7.4.6 Activate the failover notification. 131
7.5 HA MS operations 133
7.5.1 Start HA MS 134
7.5.2 Stopping the HA MS 134
7.5.3 Verify HA MS is active. 135
7.5.4 Manual HA MS failover 136
7.5.5 Active MS shutdown 138
7.5.6 Active MS crash 138
7.5.7 Reintegration. 138
7.5.8 Pause the HA MS 139
7.5.9 Things to do after an active MS failover 139
7.6 HA MS administration 140
7.6.1 Update HA MS definition file 141
7.6.2 Back up the HA MS definition file 141

7.6.3 HS MS PTF installation 141
7.7 HA MS log and debug information 142
7.7.1 HA MS logs 142
7.7.2 Audit log 142
7.7.3 Debugging tips 143
7.8 High Availability Management Server Q&A 148

Appendix A. Simple user management and authentication via LDAP . . 151
Installation and configuration of AIX Directory Server 152
Install and configure LDAP on a CSM management server 152
Setup of AIX LDAP client 155
Test the AIX LDAP server and AIX client 156
Setup of Red Hat Linux client 158
Setup of SUSE LINUX client 160
Configuration of LDAP server replicas 162
Additional information 165

Appendix B. Managing IBM WebSphere Application Server with CSM . 167
Combined CSM cluster and WebSphere cell topology 168
Using CSM node groups 169
Setting up a WebSphere Application Server cell 170
Installing WebSphere Application Server and third-party software on application server nodes 171
File synchronization across the cell 171
Managing the entire WebSphere infrastructure 172
Summary 174

Abbreviations and acronyms 175

Related publications 177
IBM Redbooks 177
Other publications 177
Online resources 177
How to get IBM Redbooks 178
Help from IBM 179

Index 181

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:

AIX®
AIX 5L™
BladeCenter™
DB2®
DRDA®
@server®
HACMP™
IBM®
iSeries™
LoadLeveler®
OS/2®
POWER™
POWER4™
POWER5™
PowerPC®
pSeries®
Redbooks™
Redbooks (logo) ™
Tivoli®
WebSphere®
xSeries®

The following terms are trademarks of other companies:

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

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

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

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

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, and service names may be trademarks or service marks of others.

Preface

This IBM® Redbook is a practical cookbook that provides up-to-date information about Cluster Systems Management (CSM) for AIX® 5L™ for a pSeries® environment. The book provides information about the latest CSM for AIX 5L enhancements including implementation techniques, installation changes, installation tools, system management tools, monitoring tools, hardware control, file distribution, problem determination, and management server high availability.

The book summarizes the latest news in CSM 1.4.0. It contains a Q&A chapter with questions and answers our team discussed during the residency, a CSM installation scenario, a CSM advanced chapter, CSM migration scenarios, and a CSM cluster administration chapter. We include information about how to manage Linux® nodes on pSeries hardware, including operating system installation and node management in a mixed cluster environment.

This Redbook is targeted to technical professionals (consultants, IT architects, and IT specialists) who are responsible for providing pSeries clustering solutions.

The team that wrote this redbook

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

Dino Quintero is a Consulting IT Specialist at ITSO in Poughkeepsie, New York. Before joining ITSO, he worked as a Performance Analyst for the Enterprise Systems Group and as a Disaster Recovery Architect for IBM Global Services. His areas of expertise include disaster recovery and pSeries clustering solutions. He is certified on pSeries system administration and pSeries clustering technologies. He is also an IBM Certified Professional on pSeries technologies. Currently, he leads technical teams delivering Redbook solutions on pSeries clustering technologies and delivering technical workshops worldwide.

Thomas Braunbeck is a Support Professional in Germany. He has 15 years of experience in AIX and 10 years of experience in PSSP and related software. He holds a degree in Computer Science.

Ong Swee Thye is an Advisory IT specialist working in ITS in IBM Singapore. He has six years of support and implementation experience in AIX, RS6000/SP, HACMP™, and CSM. His areas of expertise include AIX, HACMP, PSSP, CSM, Linux, TSM, and Veritas Netbackup. He holds a Degree of Bachelor of Electrical and Electronic Engineering from Nanyang Technological University.

Andrei Vlad is an IT specialist for IBM Global Services Romania. He has six years of experience in Linux and is an AIX Certified Engineer. He has been involved in the design and implementation of several large UNIX®-based projects, including support and special customization for a variety of applications. His areas of expertise include Linux, AIX, TCP/IP, and OS/2®. Currently, he is Technical Coordinator for Linux Center of Competence in Romania.

Peter Zutenis is a Advisory I/T Specialist with IBM Australia who joined IBM six years ago and is an AIX Certified Systems Administrator. He has nearly 25 years of experience in Information Technology, mainly in Systems Administration and Management. He works with the pSeries Systems Sales team in Australia and is mainly invovled with post-sales implementations. His areas of expertise include AIX, PSSP, Linux, CSM, and Tivoli® Storage Manager (TSM).

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

Octavian Lascu
International Technical Support Organization, Austin Center

Paul Swiatocha Jr., Norm Nott, Bruce Potter, Vimala Govindan, Jennifer Cranfill, Mark Gurevich, John Simpson, Brian Croswell, Shujun Zhou, Janet Ellsworth, Les Vigotty, Ling Gao, Pat Meehan, Ning-Wu Wang, Joseph Hughey, Greg Behrend
IBM Poughkeepsie

Paul Finley
IBM Austin

Jean-Michel Berail
IBM France

Martin Schwenke
IBM Australia

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 because 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 e-mail to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. JN9B Building 905 Internal Zip 2834
  11501 Burnet Road
  Austin, Texas 78758-3493

1

Introduction

This introduction summarizes Cluster Systems Management (CSM). It includes information to show readers' considerations and reasons why CSM can be considered an important component of their cluster. It also provides CSM highlights that differentiate it from any other competitor cluster software management in the market.

This introduction includes:

- 1.1, "What is Cluster Systems Management (CSM)?" on page 2
- 1.2, "CSM 1.4.0 highlights" on page 2
- 1.3, "CSM differentiators" on page 2
- 1.4, "Why consider CSM?" on page 3
- 1.5, "What is covered in this redbook" on page 4

1.1 What is Cluster Systems Management (CSM)?

CSM for AIX, xSeries® Linux, and pSeries Linux provides a robust, powerful, and centralized way to manage a large number of pSeries and xSeries machines from a single point of control.

CSM can lower the overall cost of IT ownership by simplifying the tasks of installing, operating, and maintaining clusters of servers.

CSM provides one consistent interface for managing both AIX and Linux and has the flexibility to manage across multiple hardware platforms, various network topologies, and different geographic sites.

CSM is designed to scale to a large number of servers and to protect performance by providing very efficient monitoring and reduced network traffic. Automatic error detection is another key feature of CSM that can help with problem avoidance, rapid resolution, and recovery.

1.2 CSM 1.4.0 highlights

For information about new features of CSM 1.4.0, see Chapter 2, “New features in CSM 1.4.0” on page 5.

1.3 CSM differentiators

What differentiates CSM from other cluster management solutions:

- CSM supports the installation and updating of software for xSeries Linux, pSeries Linux, and AIX. Using CSM to control the installs of Linux and AIX can get the cluster operating very quickly.
- CSM provides power status, hardware status, and information and diagnostic probes.
- CSM provides hardware monitoring and application monitoring events that can be responded to automatically.
- CSM manages AIX and Linux xSeries or pSeries machines from a single point of control with a consistent interface for the administrator to see the entire state of the cluster.
- CSM provides automatic setup of security for the cluster infrastructure using OpenSSH or remote shell.
- The CSM monitoring infrastructure can be customized easily to suit customer needs.

- CSM has a modular design and use of open source tools.
- CSM can run commands across the cluster or a subset of the cluster.
- CSM can synchronize files across the cluster or a subset of the cluster.

1.4 Why consider CSM?

CSM should be considered if a single point of control of your pSeries and xSeries systems is required. CSM provides many features and benefits to assist in the management over a range of many servers.

Such benefits are:

- Node groups

  Node groups can be created within a cluster and managed as distinct entities. Node groups can be defined as static lists of nodes or as a dynamic list based on a characteristic of a node, such as operating system or type. CSM ships with some predefined groups that suit many different configurations.

- Distributed command execution

  The **dsh** command enables systems administrators to run commands in parallel across all nodes or a subset of nodes in the cluster.

- Configuration file management

  A configuration file manager (CFM) is provided to synchronize and maintain consistency in files across nodes in the cluster. This helps to maintain the cluster in a consistent state. Files can be changed on the CSM management server and then distributed to nodes in the cluster. CFM can distribute to all nodes or to a subset of nodes based on node groups or host name. Pre and post scripts can also be run during CFM execution.

- Distributed monitoring

  An administrator can set up monitoring for various conditions across nodes or node groups in the cluster and have actions run in response to events that occur in the cluster. Predefined conditions are supplied with CSM that enable the administrator to start monitoring right away. The monitoring infrastructure can be extended by the customer for customized monitoring.

- Hardware control

  The remote hardware control capability enables the administrator to power on, power off, reboot, query, and use a remote console across a variety of hardware platforms.

- Diagnostic probes

  The administrator can run diagnostic probes provided by CSM to automatically perform “health checks” of particular software functions.

- Security

  CSM offers two approaches to security. Distributed command and CFM use `ssh` or `rsh` under the covers. The administrator can choose the method at install time.

- Interfaces

  CSM provides a complete and consistent command line interface. On AIX, the Web-based System Manager GUI provides the ability to manage clusters and the individual OS from a single console.

1.5 What is covered in this redbook

This IBM Redbook introduces and covers the following topics:

- The installation of a CSM cluster on pSeries in Chapter 3, “CSM installation scenario” on page 9.
- The implementation of some CSM advanced features in Chapter 4, “CSM advanced features implementation” on page 39.
- The migration of the cluster to newer releases of software in Chapter 5, “CSM migration scenario” on page 71.
- Some CSM cluster administration tips in Chapter 6, “CSM cluster administration” on page 101.
- The installation of High Availability Management Server (HA MS) in Chapter 7, “CSM High Availability Management Server” on page 117.

We also explain how to integrate Linux servers with AIX servers running on the same platform, pSeries, and using the latest CSM on the latest AIX to manage all servers from one point. The servers may be chosen from a variety of hardware, such as IBM eServer POWER4™, IBM eServer POWER5™, or IBM xSeries JS20 Blades, from low-end servers to the high-end server with multiple LPARs on the latest POWER5 systems.

In addition, we discuss installation and configuration of the HA MS. In a complex environment with AIX servers and Linux servers, you may need an HA MS for continuous availability of the CSM management server.

This book gives you the possibility to explore new ideas and the ability and the knowledge to put them into practice. For us, the writers, it was a great opportunity to share our experiences and to exchange new ideas.

2

New features in CSM 1.4.0

This chapter describes the changes and new features in Cluster Systems Management V1.4.0 for AIX 5L.

The following topics are discussed:

- 2.1, “New hardware support” on page 6.
- 2.2, “New software support” on page 6.
- 2.3, “CSM Highly Available Management Server” on page 7.
- 2.4, “Monitoring Service Focus Point using CSM” on page 7.
- 2.5, “Changes in Linux distribution support” on page 8.
- 2.6, “1024 way xSeries nodes support” on page 8.
- 2.7, “Hardware control command” on page 8.

2.1 New hardware support

CSM 1.4 supports the following new hardware on IBM eServer pSeries:

- POWER5 520
- POWER5 550
- POWER5 570
- POWER5 590
- POWER5 595
- POWER5 720
- BladeCenter™ JS20

For the complete list of hardware support for CSM 1.4, see “Planning for CSM” in *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919.

2.2 New software support

CSM 1.4 on an AIX management server can run both AIX 5L V5.2 and V5.3. The installation code for CSM 1.4 is packaged with AIX 5L 5.3 base media and in AIX 5L 5.2 maintenance level 04 (ML4).

Customers can upgrade from existing CSM Version 1.3 to 1.4 using this media. The license key for the management server is upward compatible. For example, if a customer has a CSM 1.3 license key, after upgrading to CSM 1.4, the customer could enroll the license by using the CSM 1.3 license key.

> **Note:** The CSM 1.4 installation packages require Reliable Scalable Cluster Technology (RSCT) 2.3.4 or higher for AIX 5L 5.2 and RSCT 2.4 or higher for AIX 5L 5.3.

The management server running CSM 1.4 can manage nodes running earlier levels of CSM; however, all functions provided by CSM 1.4 might not be available to the clients. For example, the Kerberos Version 5 support provided by CSM requires the managed nodes to be installed with CSM 1.3.3 and higher.

The CSM cluster is capable of scaling on each POWER4 Hardware Management Console (HMC) to up to 32 physical systems and 64 operating systems. For each POWER5 Hardware Management Console, the scaling is up to 16 physical systems and 64 operating systems.

2.3 CSM Highly Available Management Server

The IBM CSM Highly Available Management Server (HA MS) is an optional feature on Linux, AIX 5.2, and AIX 5.3. It is designed to allow automatic failover of the management server to a backup management server. It prevents the management server from being a single point of failure in the CSM cluster.

The CSM HA MS uses the shared disk to store the management server data and Tivoli System Automation Manager to control the failure.

The CSM HA MS does not require IBM High Availability Cluster Multi-Processing (HACMP) for its implementation.

> **Note:** There is no “try and buy” option for HA MS; you must purchase a separate HA MS CD.

2.3.1 Hardware support

CSM HA MS requires two shared disks attached to the management server. One is used to store the CSM MS data and the other is used during RSCT Peer Domain (RPD) partition to determine the tie breaker.

The CSM HA MS is supported on any pSeries hardware that supports Management Server. The following are the supported pSeries shared disks:

- DS4300
- DS4400
- DS4500
- 2104 Expandable Storage Plus
- SSA

2.3.2 Software support

CSM HA MS can be installed on the following platforms:

- Red Hat EL 3 on xSeries
- Red Hat AS3 update 3 on a POWER™ server
- AIX 5.2 with maintenance level 04
- AIX 5.3 with CSM 1.4.0.10

2.4 Monitoring Service Focus Point using CSM

CSM 1.4 can be used to monitor the Service Focus Point (SFP) that runs on the Hardware Management Console (HMC) by providing a set of pre-defined conditions, responses, and sensors. The objective is to provide a centralized location to view and monitor the Service Focus Point events that are created at various Hardware Management Consoles.

For the monitoring, the Hardware Management Console has to be defined as a non-node device in the CSM cluster. The remote HMC command is executed via the communication between CSM and HMC through dsh using ssh as the remote shell. The result is collected and displayed on the CSM console.

2.5 Changes in Linux distribution support

CSM 1.4 support on Linux distribution has changed. When managing Linux nodes from the AIX management server, the following CSM client packages are supported:

- Red Hat Enterprise Linux (EL) Version 3
- Red Hat Linux 2.1
- SUSE LINUX Enterprise Server 8

For hardware and software requirements for Linux on POWER server, see *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919.

2.6 1024 way xSeries nodes support

CSM 1.4 supports up to 1024 way xSeries nodes within a cluster. Changes have been made to the internal CSM command to improve performance and scaling support for xSeries hardware.

2.7 Hardware control command

The following hardware control commands have been enhanced in CSM 1.4 to return UUID for each item in a new field:

- `lshwinfo`
- `lshwstat`

The **`rpower`** command has been updated to be able to perform a query output using the return host name. A new “sort” option switch has been added to the command and is useful on a system with a large number of nodes where we can find a specific node in a arbitrary output returned by the **`rpower`** query command.

3

CSM installation scenario

In this chapter, we describe how we set up our pSeries Cluster Systems Management (CSM) cluster in our laboratory environment. This cluster is to be used for the demonstration of systems management, migration of AIX, migration of CSM, and some advanced features such as High Availability Management Server (HA MS) that will be discussed in later chapters. The intent of this chapter is to show how we set up our base pSeries CSM cluster for these exercises.

The topics that are discussed include:

- 3.1, “Assumptions and planning” on page 10.
- 3.3, “Installation of the CSM management server” on page 11.
- 3.4, “Installation of nodes” on page 23.
- 3.8, “Integration of Linux pSeries nodes” on page 35.

> **Note:** We used the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.3.3*, SA22-7919-06 during the installation of our cluster. This is a wise practice as procedures may change with new releases of software and hardware.

3.1 Assumptions and planning

We have made some assumptions for the planning and installation of our pSeries CSM cluster. As a starting point for our cluster exercises, we want to have the AIX CSM management server installed with AIX 5.2 at ML03, and CSM 1.3.3 updated to the latest PTF level of 1.3.3.4. We assumed that:

- The Hardware Management Console (HMC) is installed and operational, and any managed systems are already defined to it.
- The cluster network is in place and functional.

Note: Further detail on planning for CSM clusters can be found in the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.3.3*, SA22-7919-06, the *Introduction to pSeries Provisioning*, SG24-6389, and the *Introduction to CSM 1.3 for AIX 5L*, SG24-6859.

3.2 Our environment

Table 3-1 describes the initial software levels we desired for our cluster.

Table 3-1 Initial software levels

Node	OS type	OS level	CSM level
ms_01	AIX	5.2.0-03	1.3.3.4
node_01	AIX	5.1.0-03	1.1.2
node_02	AIX	5.1.0-03	1.1.2
node_03	AIX	5.2.0-03	1.3.3.4
node_04	AIX	5.2.0-03	1.3.3.4
node_05	Linux	RH EL-AS 3	1.3.3.3
node_06	AIX	5.2.0.-03	1.3.3.4
node_07	Linux	SLES 9.0	N/A
node_08	Linux	SLES 8.1	1.3.3.3

Note: node_07 is running SUSE LINUX SLES 9. This is not supported by CSM 1.3.3. However, after we migrate the cluster to CSM 1.4.0.3, it is supported.

Figure 3-1 shows our lab environment.

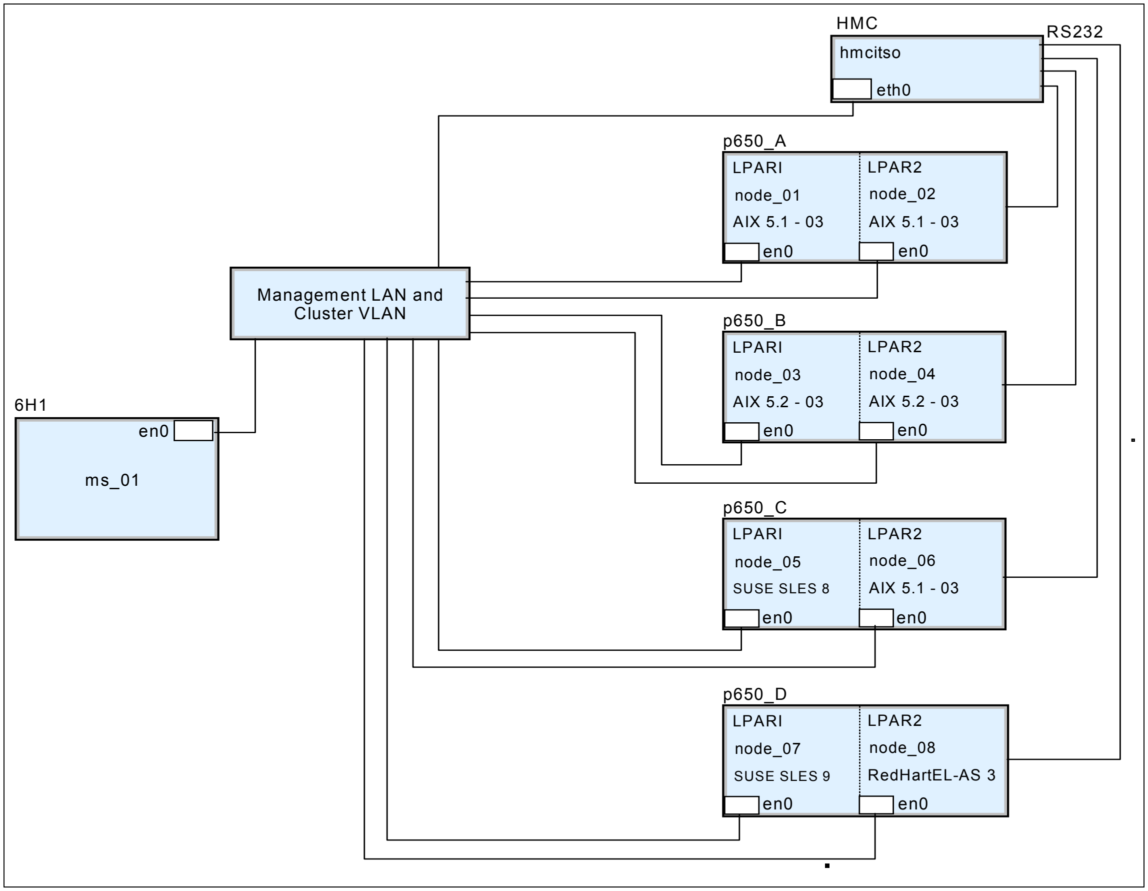


Figure 3-1 Diagram of our CSM cluster

3.3 Installation of the CSM management server

The following sections describe the steps we undertook to install our CSM management server. These steps include:

- 3.3.1, "Installation of AIX" on page 12
- 3.3.2, "Apply maintenance to AIX" on page 12
- 3.3.3, "Installation and customization of NIM" on page 13
- 3.3.4, "Install and apply maintenance to the CSM server" on page 16
- 3.3.5, "Customize the CSM server" on page 20

3.3.1 Installation of AIX

We performed a standard install of AIX using the AIX 5.2 CDs. The CDs we already have incorporate Maintenance Level 01. However, for this exercise, we want to have AIX 5.2 installed at Maintenance Level 03, therefore maintenance will have to be applied after the initial install of AIX.

> **Note that:**
>
> - AIX 5.2 Maintenance Level 03 updates CSM to Version 1.3.3. CSM is at level 1.3.1 at AIX 5.2 Maintenance Level 01.
> - AIX 5.2 Maintenance Level 04 updates CSM to Version 1.4.0. For this reason we avoided ML04 as we want to remain with CSM 1.3.3 for now.

Our management server is an IBM pSeries 7026-6H1 and is 64-bit capable. We selected the options to install the 64-bit kernel and the JFS2 file system.

After the install, we then performed some basic customization to complete the AIX install of the management server. These included:

1. Set up the primary and secondary dump areas.
2. Mirror the root volume group.
3. Increase the maxuproc to 1024.
4. Set up Transmission Control Protocol / Internet Protocol (TCP/IP) suitable for our lab environment. This included host name and TCP/IP addresses. The network in our lab was very simple. There was just one network for both the cluster VLAN and the management VLAN. A production cluster is likely to have separate LANs for the cluster VLAN and management VLAN.
5. Set the root password.
6. Install perfagent.tools, openssl and openssh.

> **Note:** We did not install the openssl64.* file sets. These did not work correctly for us. We used the openssl.* file sets and had no issues.

3.3.2 Apply maintenance to AIX

After the successful installation of AIX 5.2 at ML01, the application of the maintenance to bring the system up to ML03 was performed. This also brought the level of the CSM client file sets to V1.3.3.

Maintenance for AIX can be downloaded from:

`http://www.ibm.com/support/`

For more information about installing and maintaining AIX, refer to *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389.

3.3.3 Installation and customization of NIM

This section is not meant to be a detailed discussion of how to install and set up Network Installation Manager (NIM), but it gives an overview of the processes we followed in setting up our NIM server. See *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, for more in-depth detail about NIM master and client installation.

There are two ways to install NIM onto the management server. We can set up the NIM environment manually or we can choose the eznim method. We have to install the NIM master and Shared Product Object Tree (SPOT) file sets, as the CSM management server will become a NIM master for our cluster. We installed the NIM master file sets from the base media and applied maintenance to them. Application of maintenance is very important to bring the NIM master to the same maintenance level as the operating system. Failure to do this may cause failure of the NIM master. The file sets to install and update are:

- bos.sysmgt.nim.master
- bos.sysmgt.nim.spot

Note: NIM is optional in a CSM cluster. We highly recommend that you install NIM to ease installation and maintenance of AIX 5L within the cluster.

For our cluster, we decided to use the eznim method for setting up NIM because it is a very simple but effective way to install a NIM master. It automates the various tasks that are required in setting up a NIM server and is very good for those who may have little experience in setting up a NIM master. Of course, we could have chosen any method we liked. Those experienced with NIM can choose to set it up any way that suits them.

We initially wanted to install NIM into the same file system that CSM uses (/csminstall), but eznim is not very flexible in where it can place the master. We decided to install the NIM master to the eznim default file system: /export/eznim. The high-level directory name can be chosen when setting up via eznim.

Note: If you plan to install CSM High Availability Management Server (HA MS), and you require NIM failover, then HA MS has a restriction that NIM is to be installed to /csminstall/AIX.

Tip: Before starting the setup of NIM via eznim, make sure that the root user ID has the file size set to `unlimited` with the command:

```
ulimit -f unlimted
```

If you are creating the file systems for NIM manually, ensure that the file systems are JFS2, or if using JFS, ensure that they are large-file enabled.

To start the setup of NIM via the eznim process, we issued the **`smitty eznim`** command as shown in Example 3-1.

Example 3-1 Display of smitty eznim

```
Select Configure as a NIM Master
Select Setup the NIM Master environment.
[ Select or specify software source                  [cd0]              +
  to initialize environment

  Select Volume Group for resources                  [rootvg]           +

  Select Filesytem for resources                     [/export/eznim]
```

For our NIM master, we used the default parameters displayed on the SMIT panels. Press Enter, and the eznim process performs this for the NIM master:

1. Setup of the system as a NIM master and start the nimesis daemon
2. A mksysb image
3. An AIX 520 spot
4. A bosinst.data resource
5. A boot resource

In addition to the file system structure eznim built for us, we created two extra directories. These were /export/eznim/extras and /export/eznim/lpp_source. The /export/eznim/extras directory is used to hold additional software, and the /export/eznim/lpp_source directory holds the various file sets for each AIX release.

> **Note:** Building the SPOT from a CD with eznim will not create a lpp_source in the newly created NIM master. To create an lpp_source for AIX, use **`smitty nim_bffcreate`** to copy the file sets to the /export/eznim/lpp_source/ directory. Make sure that Create an LPP_SOURCE and LPP_SOURCE Name are specified. We recommend that you create directories here identifying each release of AIX. We chose to use the standard as follows: AIX52_01, AIX52_03, AIX53, and so on. The directory we used for our lpp_source was /export/eznim/lpp_source/AIX52.
>
> If a directory containing all of the AIX bff file sets is used for input to the eznim master creation, the eznim process will create an lpp_source for you.

After the setup of NIM via eznim has completed, it is still necessary to update the SPOT to the same maintenance level as the operating system. After the eznim process has set up NIM, the SPOT will be at the same maintenance level as the base CD from which the base NIM file sets were installed. We updated the SPOT using the lpp_source resource containing the AIX 5.2 ML03 file sets. This step of applying maintenance to the SPOT is often forgotten and is a very common cause of issues when trying to install systems that are at a higher maintenance level than the SPOT.

In the end, our NIM master is shown in Example 3-2.

As our lab progresses, we add more resources to NIM such as:

- New maintenance levels for AIX and CSM
- New releases of AIX and CSM
- Optional software packages
- Node definitions

Example 3-2 Final definition of our NIM master

```
master:
   class               = machines
   type                = master
   max_nimesis_threads = 20
   if_defined          = chrp.mp.ent
   comments            = machine which controls the NIM environment
   platform            = chrp
   netboot_kernel      = mp
   if1                 = master_net ms_01 000629DC2595
   cable_type1         = N/A
   Cstate              = ready for a NIM operation
   prev_state          = ready for a NIM operation
   Mstate              = currently running
   serves              = 5200-03master_sysb
   serves               = 520spot_res
   serves               = aix520
   serves               = aix52_ml03
   serves               = bid_ow
   serves               = boot
   serves               = nim_script
   registration_port    = 1059
   reserved             = yes
```

3.3.4 Install and apply maintenance to the CSM server

The CSM server software is supplied on the AIX Base CDs in the “try and buy” format. The software can be installed off the base CDs and converted to a full license without having to reconfigure CSM. CSM can be used for 60 days with full functionality. To convert, the full license can be applied and the “try and buy” cluster functionality is retained without the need for reconfiguration.

The following sections outline the steps we used to install CSM on our cluster. We install CSM 1.3.3 and migrate this version to 1.4.0.10 later in our exercise.

Open source programs

Two open source programs are required for CSM to operate that are not supplied on the base CDs and must be obtained from the Internet. These two programs are openCIMOM and autoupdate.

openCIMOM

This open source program is used to support CSM hardware control functions. If we are not using CSM for hardware control, then this program is not required.

Tip:

- When installing CSM 1.3.3, you must use Version 0.7. Version 0.8 will not work and will cause problems with power control.
- If you are installing Version 1.4 of CSM, then you must use Version 0.8.

To install openCIMOM, execute the following steps:

1. Download the openCIMOM Red Hat Package Manager (RPM) package from:

   http://www.ibm.com/servers/aix/products/aixos/linux/download.html

2. Copy the file to a temporary directory. In our case we copied it to /export/eznim/extras/open_source.
3. Install the RPM package by issuing the **rpm** command:

   ```
   rpm -ivh /export/eznim/extras/open_source/openCIMOM-0.7-4.aix5.1.noarch.rpm
   ```

Note: Do not be fooled by the reference to AIX 5.1 in the RPM file name. It will work fine under AIX 5.2.

autoupdate

This open source program is used to manage RPMs on Linux nodes from the CSM management server. If we plan not to have Linux nodes, then this RPM is not required. This file set is not installed onto the management server, but onto any Linux node as part of the CSM Linux node install. This RPM will be copied to the correct location later on in the CSM setup for Linux process, so we just left it in the temporary location for now.

Refer to 3.8.1, “Check the node software” on page 36 for more about autoupdate.

Updated PATH and MANPATH

We then update the PATH and MANPATH for root as per the install guide.

Create /csminstall filesystem

We create and mount a 1 GB file system called /csminstall. The size of this file system will vary depending on how many nodes and type, but 1 GB is a good starting point.

Update /etc/hosts

We then add our nodes to the /etc/hosts file. Example 3-3 shows an example.

Example 3-3 Excerpt of /etc/hosts file

```
192.168.100.69  hmcitso # HMC console
192.168.100.182 ms_01   # CSM Primary Server
192.168.100.181 ms_02   # CSM Backup Server
192.168.100.183 node_01 # AIX 5.1 ML03
192.168.100.184 node_02 # AIX 5.1 ML03
192.168.100.185 node_03 # AIX 5.2 ML01
```

Install the CSM server software

Before installing the CSM management server, we first check to make sure that we have the necessary prerequisites installed, checking the install guide for the correct software levels. Be aware that the management server must be AIX 5.2. The levels of Reliable Scalable Cluster Technology (RSCT) and the CSM client software that are installed with the Maintenance Level of AIX must be checked carefully.

Install the CSM server file sets

We install the CSM server onto our management server by following the instructions in the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.3.3*, SA22-7919-06. We issue this command:

```
geninstall -IaX -d /dev/cd0 csm.server csm.diagnostics csm.gui.dcem \
csm.gui.websm R:expect R:tcl R:tk R:conserver
```

> **Attention:** The *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.3.3*, SA22-7919-06, instructs you to use CD 1 to install the CSM Server. This is a typographical error. Use CD 2 to install the CSM server.

Example 3-4 shows some of the output from our installation of the CSM server.

Example 3-4 Output from the CSM server install

```
conserver                   ##################################################
expect                      ##################################################
tcl                         ##################################################
tk                          ##################################################

+-----------------------------------------------------------------------------+
                    Pre-installation Verification...
+-----------------------------------------------------------------------------+
Verifying selections...done
Verifying requisites...done
Results...

checking for open source prerequsites....
/opt/csm/csmbin/addsvcent -s ibm_cspd -n 8435 -p tcp
/opt/csm/csmbin/addsvcent -s ibm_hcdh -n 9095 -p tcp
/opt/csm/csmbin/addsvcent -s ibm_hcdx -n 9096 -p tcp
Removing existing cfmupdatenode cron job.
Installing cfmupdatenode cron job.
Filesets processed:  1 of 2  (Total time:  8 secs).

installp:  APPLYING software for:
        csm.gui.websm 1.3.1.0
Please wait...

        /usr/sbin/rsct/install/bin/ctposti
0513-071 The ctrmc Subsystem has been added.
0513-059 The ctrmc Subsystem has been started. Subsystem PID is 311522.

        PERL_BADLANG=0 /opt/csm/install/csmserverposti
Waiting for ERRM to start...
ERRM is started.
```

```
Executing the "/opt/csm/bin/predefined-condresp -m" command.
Executing the "/opt/csm/bin/predefined-nodegroups -m" command.
Executing the "/usr/bin/chrsrc-api -s "IBM.AuditLog::Name='ERRM'::MaxSize::20""
command.
done

Installation Summary
--------------------
Name                        Level           Part        Event       Result
-------------------------------------------------------------------------------
csm.server                  1.3.1.0         ROOT        APPLY       SUCCESS
csm.gui.websm               1.3.1.0         USR         APPLY       SUCCESS
```

Apply maintenance to the CSM server

The version of the CSM server that is installed off ML01, the base CDs, is 1.3.1. We must apply Maintenance Level 03 to the CSM Server before we can proceed.

To apply the maintenance, change directory to the NIM directory that contains the AIX ML03 file sets.

```
cd /export/eznim/lpp_source/AIX52_03
smitty update_all
```

It is good practice to do a preview update to confirm that all prerequisites for the update are available. After the update, the CSM server should be at level 1.3.3.0. We check this by using the **`lslpp`** command.

Check the IBM CSM cluster Web site for the latest updates for CSM 1.3.3:

http://techsupport.services.ibm.com/server/cluster/fixes/csmfixhome.html

We download the latest fix level for CSM 1.3.3 (which, during our lab, was 1.3.3.4). Uncompress and untar the file into a new directory called /export/eznim/extras/csm1334. Again, we change to the directory containing the software, and run a **`smitty update_all`** to update CSM to 1.3.3.4.

When we preview the update_all operation, the preview operation informs us that we need an update to rsct.core.errm. We download the necessary RSCT updates the IBM AIX fix Web site into the same directory as the CSM updates.

We then run inutoc . to update the .toc file and re-run the preview update. This time the preview runs clean and we apply the fixes successfully.

Example 3-5 on page 20 shows the final levels of the CSM and RSCT software that we used for our initial cluster install.

An update to conserver supplied with the CSM 1.3.3.4 fixes is located in the /export/eznim/extras/csm1334/RPMS/ppc directory. We run an **rpm -Fvh** against the conserver RPM.

Example 3-5 Software levels after the CSM server install

```
Installation Summary
--------------------
Name                        Level           Part        Event       Result
-------------------------------------------------------------------------------
rsct.core.utils             2.3.3.4         USR         APPLY       SUCCESS
rsct.core.utils             2.3.3.4         ROOT        APPLY       SUCCESS
rsct.core.sr                2.3.3.2         USR         APPLY       SUCCESS
rsct.core.sr                2.3.3.2         ROOT        APPLY       SUCCESS
rsct.core.sec               2.3.3.1         USR         APPLY       SUCCESS
rsct.core.sec               2.3.3.1         ROOT        APPLY       SUCCESS
rsct.core.rmc               2.3.3.3         USR         APPLY       SUCCESS
rsct.core.rmc               2.3.3.3         ROOT        APPLY       SUCCESS
rsct.core.fsrm              2.3.3.1         USR         APPLY       SUCCESS
rsct.core.fsrm              2.3.3.1         ROOT        APPLY       SUCCESS
rsct.core.errm              2.3.3.2         USR         APPLY       SUCCESS
rsct.core.errm              2.3.3.2         ROOT        APPLY       SUCCESS
csm.server                  1.3.3.4         USR         APPLY       SUCCESS
csm.server                  1.3.3.4         ROOT        APPLY       SUCCESS
csm.core                    1.3.3.3         USR         APPLY       SUCCESS
csm.core                    1.3.3.3         ROOT        APPLY       SUCCESS
csm.dsh                     1.3.3.3         USR         APPLY       SUCCESS
csm.client                  1.3.3.2         USR         APPLY       SUCCESS
csm.client                  1.3.3.2         ROOT        APPLY       SUCCESS
csm.gui.dcem                1.3.3.1         USR         APPLY       SUCCESS
```

We can now consider the CSM software installed.

3.3.5 Customize the CSM server

We now have to configure our CSM server by following these steps:

1. Adjust the configuration attributes.
2. Accept the license.
3. Copy the CSM files to the correct directories.
4. Store the hardware control point user ID.
5. Verify the CSM server installation.

Adjust the configuration attributes

Display the current configuration of the CSM Server by using the **csmconfig** command without any flags. To change its values, use the **csmconfig** command as in Example 3-6 on page 21. Specify the attribute to change and the new value of that attribute. Table 3-2 shows the values we changed for our CSM cluster. We decided to use the OpenSSH commands for our cluster, so we have to modify the attributes accordingly.

Example 3-6 csmconfig attribute change

```
# csmconfig RemoteCopyCmd=/usr/bin/scp
# csmconfig RemoteShell=/usr/bin/ssh
```

Table 3-2 Changed csmconfig attributes

Flag	Default value	New value
RemoteCopyCmd	/usr/bin/rcp	/usr/bin/scp
RemoteShell	/usr./bin/rsh	/usr/bin/ssh

Accept the license

CSM is a licensed software. We need to apply our license or use the 60-day “try-and-buy” license, which is included on the AIX CDs. The full license is supplied on the CSM CD. We apply our license using the **`csmconfig`** command with the -L flag:

```
csmconfig -L /tmp/full_license_file
```

The license agreement is displayed and we follow the questions to accept the agreement.

Note: When a license file with the -L flag is *not* specified, the try-and-buy license is installed.

Copy the CSM files to the correct directories

We copy the CSM files to the proper /csminstall subdirectories using **`cmsconfig`**:

```
# cmsconfig -c
#
```

Store the hardware control point user ID

We use the Hardware Management Console (HMC) that our p650 nodes are attached to for hardware control. To enable correct functionality of the hardware control, we store the user ID and password that CSM will use to communicate to the HMC. We issue the **`systemid`** command as shown in Example 3-7.

Example 3-7 The systemid command

```
# systemid hmcitso hscroot
```

```
Password:
Verifying, please re-enter password:
systemid: Entry updated.
#
```

It prompts for the password (`hscroot`) on the HMC with the host name `hmcitso`.

Tip: To enable the management server to communicate with the HMC, you must make sure that remote execution has been enabled on the HMC.

Note: The user ID on the HMC must exist. The password and user ID that are entered with the **`systemid`** command must match the user ID and password on the HMC. The default user ID for an HMC is `hscroot` with password `abc123`. This is likely to have been changed on an existing HMC.

Verify the CSM server installation

To verify that our CSM management server has been defined correctly, we run a CSM probe:

```
# probemgr -p ibm.csm.ms -l 0
#
```

A lot of output is produced by this command. We receive a warning that IP forwarding was not turned on, but in our situation this was acceptable. The last line of the output in Example 3-8 shows that probe ibm.csm.ms ran successfully.

Example 3-8 Last few line of the probemgr output

```
Running probe /opt/diagnostics/probes/ibm.csm.ms.
Probe ibm.csm.ms returned the following information.
ibm.csm.ms:trace:Checking that directory /opt/csm exists.
ibm.csm.ms:trace:Checking that directory /opt/csm/bin exists.
ibm.csm.ms:trace:Checking that directory /opt/csm/csmbin exists.
ibm.csm.ms:trace:Checking that directory /usr/sbin/rsct/bin exists.
ibm.csm.ms:trace:Checking that directory /var/log/csm exists.
ibm.csm.ms:trace:Checking for packages : csm.client*, csm.core*,
csm.diagnostics
*, csm.dsh*, csm.gui.dcem*, csm.gui.websm*, csm.msg.*_*, csm.server*.
ibm.csm.ms:trace:Check if the CFM cronjob is enabled.
Probe ibm.csm.ms was run successfully.
```

Our CSM server is now ready for use.

3.4 Installation of nodes

Installation of our nodes was a straightforward process. There are many steps required for the initial install of the nodes, and these are outlined in the sections that follow.

3.4.1 Define nodes to CSM

The first step in installing our nodes was to define the attributes of our nodes. This step, at a minimum, requires the collection of some attributes prior to defining the node to the CSM server. These attributes are discussed in 3.4.2, “Node definition” on page 23.

3.4.2 Node definition

The easiest way to add nodes to CSM is to create a node stanza file. Example 3-9 shows a preview of our stanza file.

Example 3-9 A stanza from our node stanza file

```
node_03:
ConsoleMethod=hmc
ConsoleServerName=hmcitso
HWControlNodeId=p650_B_LPAR1
HWControlPoint=hmcitso
LParID=001
PowerMethod=hmc
```

We create a temporary file called /tmp/nodedef with one stanza for each node. Table 3-3 explains each line in the stanza.

Table 3-3 Node stanza definitions

Stanza entry	Explanation
node_03	Host name of node.
ConsoleMethod=hmc	The HMC is to be used for opening serial consoles on this node.
ConsoleServerName=hmcitso	The host name of Console server, which in this case is the HMC.
HWControlNodeId=p650_B_LPAR1	The name of the LPAR (as shown on the HMC) that node_03 will run in.

Stanza entry	Explanation
HWControlPoint=hmcitso	As this node is controlled from the HMC, then the host name of the HMC is entered here.
LParID=001	The LParID number of the LPAR that node_03 will run in. This can be obtained from the HMC by displaying the properties of the LPAR.
PowerMethod=hmc	For pSeries, the Power Method must be set to hmc or csp. As the node is managed from an HMC, then the power method is hmc.

After our stanza file is built, we run the **definenode** command against this file.

When the command completes successfully, we then run the **lsnode** command as shown in Example 3-10 to confirm that are nodes were defined. For a more detailed output we run the **lsnode -l** command.

Example 3-10 Running the definenode and lsnode command

```
# definenode -f /tmp/nodedef
Defining CSM Nodes:
Defining Node "node_01"("192.168.100.183")
Defining Node "node_02"("192.168.100.184")
Defining Node "node_03"("192.168.100.185")
Defining Node "node_04"("192.168.100.186")
Defining Node "node_05"("192.168.100.187")
Defining Node "node_06"("192.168.100.188")
Defining Node "node_07"("192.168.100.189")
Defining Node "node_08"("192.168.100.190")
#
# lsnode
node_01
node_02
node_03
node_04
node_05
node_06
node_07
node_08
#
```

Change node attributes

After the nodes are defined to the CSM database, we have to alter some nodes by hand to suit the mixed operating environment we want to install. By default, many CSM node attributes in the CSM database inherit the value of the CSM management server, but this does not suit some of our nodes. We use the **chnode** command as shown in Example 3-11 to alter the CSM attributes. In our case, node_01 and node_02 are AIX 5.1 nodes, and node_05,node_06 and node_08 are Linux nodes. Refer to Table 3-1 on page 10 for our cluster software setup.

Example 3-11 Use of the chnode command to change CSM node attributes

```
chnode -n node_01,node_02 InstallDistributionVersion=5.1.0
chnode -n node_05,node_07,node_08 InstallOSName=Linux
```

3.4.3 Create nodegroups

This step is optional, but we found it useful to create two groups to suit our cluster. The two groups we created were dynamic groups where one group contained all AIX 5.1 nodes and the other group contained AIX 5.2 nodes. We used the **nodegrp** command to create the aix51 and aix52 groups. Example 3-12 shows this command in action.

Note: Many default node groups are installed as part of the CSM server and these may suit your needs. Use the **nodegrp** command without any flags to show all node groups. The **nodegrp** command with the -p flag displays which nodes are in each group.

Example 3-12 The nodegrp command used to create dynamic groups

```
# nodegrp -w "InstallDistributionVersion=='5.1.0'" aix51
# nodegrp -w "InstallDistributionVersion=='5.2.0'" aix52
```

3.4.4 Validate hardware control

Before we can continue installing our cluster, we must ensure that hardware control is functioning correctly on the cluster. The installation steps that follow rely on hardware control. If hardware control is not operating correctly, then it must be fixed before continuing with the cluster install; otherwise, the cluster install will fail. We use the **rpower** command to control the powering on and off of nodes and to display the power status of nodes. Example 3-13 shows the use of the **rpower** command, which is working just fine in this example.

Example 3-13 Use of rpower to control power to nodes

```
# rpower -a query
```

```
node_01 off
node_02 off
node_03 off
node_04 off
node_05 off
node_06 off
node_07 off
node_08 off
#
# rpower -n node_01 on
node_01 on complete rc=0
# rpower -n node_01 query
node_01 on
# rpower -n node_01 off
node_01 off complete rc=0
rpower -n node_01 query
node_01 off
```

3.4.5 Get adapter information

The **getadapters** command can be used to obtain the Media Access Control (MAC) addresses of the node's Ethernet adapters. We are interested in only the first adapter that responds to the management server. We decide to run **getadapters** in two passes. The first execution of **getadapters**, shown in Example 3-14, obtains the MAC addresses of the first Ethernet adapter that is capable of 100 full duplex operation of all nodes in the AIXNodes group. The output is written into the /tmp/adapters file. At this stage, the CSM database has not been updated.

After the /tmp/adapters file is built, we add the MAC addresses to the CSM database with the **getadapters** command using the -w flag.

> **Note:** You can run the **getadapters** command to obtain MAC addresses and update the CSM database in one pass. We decided to be cautious and wanted to check the contents of the /tmp/adapters file prior to adding the MAC addresses to the CSM database.

Example 3-14 Use of getadapters to obtain MAC addresses and add to CSM

```
# getadapters -N AIXNodes-D -t ent -s 100 -d full -z /tmp/adapters
# getadapters -w -f /tmp/adapters
```

3.5 Set up cluster file management

CSM uses a file management system known as Configuration File Management (CFM). A directory called /cfmroot is the base for the distribution of files to nodes within the cluster. CFM can be set up to perform some sophisticated file management, but for the purposes of our cluster, we set up a very simple file distribution.

We want the /etc/hosts files installed onto each node as part of the CSM customizing process. To achieve this, we create a symbolic link on the management server from /etc/hosts to /cfmroot/etc/hosts. When the CSM client code is installed to the node, CFM will follow the symbolic link and copy the /etc/hosts that is on the management server to the node.

Note: Files can be distributed to the nodes at any time using the **`cfmupdatenode`** command. As part of the CSM server install, a cronjob is setup to run **`cfmupdatenode`** at midnight every day.

3.6 NIM steps

The following steps outline the necessary procedures to define the CSM nodes as NIM clients. We could have defined these nodes using NIM commands, but CSM provides some commands to ease this process.

3.6.1 Define NIM clients

We use the **`lsnim -l`** command (Example 3-15) to show the result of the **`csm2nimnodes`** command, which we used to define the CSM nodes to NIM.

Note: The **`csm2nimnodes`** command can be run only if the CSM server and the NIM master are on the same system.

Example 3-15 The csm2nimnodes command execution

```
# csm2nimnodes -N AIXNodes
#
#lsnim -l node_03
node_03:
   class          = machines
   type           = standalone
   platform       = chrp
   netboot_kernel = mp
   if1            = master_net node_03 0002556F2037 ent
   cable_type1    = N/A
   Cstate         = ready for a NIM operation
   prev_state     = ready for a NIM operation
   Mstate         = not running
```

3.6.2 NIM machine groups

We decided to add the CSM node group AIXNODES to NIM as a NIM client group. We used the **csm2nimgrps** command to achieve this (Example 3-16).

> **Note:** The **csm2nimgrps** command can be run only if the CSM server and NIM master are on the same system.

Example 3-16 csm2nimgrps execution

```
# csm2nimgrps -N AIXNodes
# lsnim -l AIXNodes
AIXNodes:
   class   = groups
   type    = mac_group
   member1 = node_01
   member2 = node_02
   member3 = node_03
   member4 = node_04
   member5 = node_06
```

3.6.3 NIM CSM preparation

The command **csmsetupnim** is used to create node configuration files and defines a script to NIM that is then allocated to the node. This script is used to install CSM onto the node after NIM has finished installing AIX onto the node.

> **Note:** Every time you install or reinstall a node, you must run this command.

This is how we used the **csmsetupnim** command:

```
# csmsetupnim -N AIXNodes
#
```

3.6.4 Add OpenSSL and OpenSSH bundles

As we have specified in our CSM configuration to use OpenSSH for the cluster shell, we must ensure that it is installed as part of the NIM install of AIX. OpenSSL must be installed before attempting to install OpenSSH. The following steps outline how we set up for the install of OpenSSH and OpenSSL using NIM bundle files. The file sets and RPMs for these two products must be copied to the same lpp_source as the operating system level that is to be installed.

1. Created the bundle files using **vi**. (The bundle files can be located in any directory; we used /export/eznim/extras/openssh.) When the bundle files are to be defined in NIM we have to enter the correct location of the bundle files.

   Example 3-17 shows the bundle files.

Example 3-17 OpenSSL and OpenSSH bundle files

```
#cat /export/eznim/extras/openssh/openssl.bnd
R:openssl*
# cat /export/eznim/extras/openssh/openssh.bnd
I:openssh.base
I:openssh.license
I:openssh.man.en_US
I:openssh.msg.en_US
I:openssh.msg.EN_US
#
```

2. Copy the RPMs for OpenSSL to /export/eznim/lpp_source/AIX52/RPMS/ppc.

> **Note:** You can get the OpenSSL software off the AIX Toolbox for Linux Applications for Power Systems CD that is supplied with the AIX CDs. You can use the **smitty bffcreate** command to extract the RPMs from the CD and copy them to the correct directory.

3. The file sets for OpenSSH must be copied into the same lpp_source directory of the version of AIX that is being installed. In our case, the file sets were copied into /export/eznim/lpp_source/AIX52/installp/ppc.

> **Note:** The file sets for OpenSSH can be obtained from the IBM AIX Bonus Pack CD; however, this version is very old. Download the latest version from:
>
> http://www.ibm.com/developerworks/opensource/index.html

4. Define the bundle files in NIM using **smitty nim_res**. We chose simple names, as shown in Example 3-18.

   After the bundles are defined to NIM, they can be included as installp_bundle resources for the bos_inst operation.

Example 3-18 OpenSSH and OpenSSL bundles

```
# lsnim -l openssl
openssl:
   class       = resources
   type        = installp_bundle
   comments    = Openssl Bundle
   Rstate      = ready for use
   prev_state  = unavailable for use
   location    = /export/eznim/extras/openssh/openssl.bnd
   alloc_count = 0
   server      = master
#
# lsnim -l openssh
openssh:
   class       = resources
   type        = installp_bundle
   comments    = Openssh Bundle
   Rstate      = ready for use
   prev_state  = unavailable for use
   location    = /export/eznim/extras/openssh/openssh.bnd
   alloc_count = 0
   server      = master
```

Tip: When defining the openssl and openssh bundles, ensure that you define the openssl bundle first. This ensures that openssl is installed before openssh.

3.6.5 NIM resource allocation

We created an initial mksyb for our install by doing an rte install of AIX to a node. This enabled us to install AIX, apply maintenance to it, and update the CSM client code to the latest levels. We also installed some optional file sets such as perfagent.tools. We had to run the **`smitty nim_update_all`** several times against each lpp_source to ensure that the latest maintenance was installed.

After we were satisfied with our initial base AIX image we backed, it up using **`smitty nim_mkres`**. This enabled us to create a mksysb of the node and create a resource within NIM. We will use this image to install our nodes.

We use the **`smitty nim_mac_res`** command to allocate the resource to our nodes for the install. Example 3-19 shows the resources that we allocated to our nodes.

Example 3-19 NIM resources that were allocated

```
bid_ow                  bosinst_data                                  ¦
520spot_res             spot                                          ¦
aix520                  lpp_source            AIX 52 Base cd with ml01 ¦
openssl                 installp_bundle       Openssl Bundle           ¦
openssh                 installp_bundle       Openssh Bundle           ¦
aix520_03_base          mksysb                                        ¦
csmprereboot_script     script                                        ¦
```

Tip: Before running the install, we checked the bid_ow (/export/eznim/bid_ow) bosinst.data resource to make sure the file was accurate. We altered two settings:

```
ENABLE_64BIT_KERNEL = yes
CREATE_JFS2_FS = yes
```

This was to ensure that the 64-bit kernel was installed and that JFS2 file systems were built.

After the NIM resources were allocated successfully to the nodes, we set up the nodes for a NIM bos_inst operation using the **smitty nim_mac_op** command. We set the options for a mksysb install as in Example 3-20.

Example 3-20 Setting up bos_inst for a node

```
                      Perform a Network Install

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

                                                       [Entry Fields]
  Target Name                                        node_03
  Source for BOS Runtime Files                       mksysb              +
  installp Flags                                    [-agX]
  Fileset Names                                     []
  Remain NIM client after install?                   yes                 +
  Initiate Boot Operation on Client?                 no                  +
  Set Boot List if Boot not Initiated on Client?     no                  +
  Force Unattended Installation Enablement?          no                  +
    ACCEPT new license agreements?                  [yes]                +
```

We modified these entry fields:

- Source for BOS Runtime Files?
- Initiate Boot Operation on Client?
- ACCEPT new license agreements?

3.7 Network boot node

After the NIM bos_inst operation completes successfully, the node can be installed.

3.7.1 Verify resources

Before netbooting the node for an install, it is prudent to confirm that all resources are ready for a netboot. We checked the following basic items:

- Issue the l**s -l** command to check the contents of the /tftpboot file system. There should be a symbolic link to the network boot kernel and an .info file for the node as shown in Example 3-21.

Example 3-21 Output of ls -l of /tftpboot

```
ls -l /tftpboot
-rw-r--r--   1 root     system      8403007 Oct 21 16:30
52Hspot_res.chrp.mp.ent
lrwxrwxrwx   1 root     system           33 Oct 22 17:52 node_03 ->
/tftpboot/52Hspot_res.chrp.mp.ent
-rw-r--r--   1 root     system         1096 Oct 22 17:52 node_03.info
```

- Run the **lsnim -l node_03** command to check its status with NIM (Example 3-22).

Example 3-22 Output of the lsnim -l command

```
node_03:
   class          = machines
   type           = standalone
   platform       = chrp
   netboot_kernel = mp
   if1            = master_net node_03 0002556F2037 ent
   cable_type1    = N/A
   Cstate         = BOS installation has been enabled
   prev_state     = ready for a NIM operation
   Mstate         = currently running
   boot           = boot
   lpp_source     = AIX520
   nim_script     = nim_script
   spot           = AIX52_spot
   cpuid          = 000197BA4C00
```

- Check /etc/bootptab. There is a entry for node_03 as shown in Example 3-23.

Example 3-23 cat of /etc/bootptab

```
cat /etc/bootptab
node_03:bf=/tftpboot/node_03:ip=192.168.100.185:ht=ethernet:ha=0002556F2037:sa=
192.168.100.181:sm=255.255.255.0:
```

- Run **showmount -e** to ensure that the NFS setup had run as expected. Example 3-24 on page 33 shows the output.

Example 3-24 The showmount -e command

```
# showmount -e
export list for ms_01:
/export/eznim/lpp_source/AIX52 node_03
/export/eznim/spot/AIX520_spot/usr node_03
/export/nim/scripts/node_03.script node_03
```

3.7.2 Initiate netboot

To initiate the install of the node, we issued the **netboot** command as shown:

```
# netboot -n node_03
```

Note: The **netboot** command takes quite a while to run. The command only returns to the command prompt when the node has successfully netbooted AIX for the install. The command waits for the open firmware processes to complete on the node.

3.7.3 Monitoring the install

While the install is underway, it is often useful to monitor if for any problems that may occur. Here are two ways:

- **rconsole** command

  The **rconsole** command opens a serial console link to the node. We issued the following command in from a shell on the management server:

  ```
  rconsole -t -n node_03
  ```

  Note: The -t flag of the **rconsole** command opens a serial link in your current shell. Without the -t flag, **rconsole** attempts to open a new window to an x-server and will fail in a pure tty environment. Of course, if an x-windows is operating you may wish to allow a new window to be opened.

  The keystroke sequence Ctrl+E+c terminates the **rconsole** session.

Note: By default, the **rconsole** command opens a write-enabled session. As the install process requires a write-enabled session (there can be only one), the system automatically converts your write session to a read-only session.

- **csmstat** command

  The **-l** (lower-case L) of the **csmstat** command shows the LED values of the node as it executes the install. Example 3-25 shows sample output.

Example 3-25   csmstat -l output

```
-------------------------------------------------------
Hostname           LCD1             LCD2
-------------------------------------------------------
node_01
node_02
node_03            c54
node_04
node_05            Linux ppc64   2.~
node_06
node_07            SuSE Linux ppc64~
node_08            Linux ppc64   2.~
```

After the node has completed installing successfully, it should be integrated into the cluster. Proceed to verify the installation. Several commands are available that enable us to verify the successful install of our node:

- **csmsta**t

  The **csmstat** command without any flags shows the status of the node at a glance. Example 3-26 shows that the node is powered on and the network interface is online. This command was entered after we had installed several of our nodes.

Example 3-26   csmstat showing node status

```
# csmstat
---------------------------------------------------------------------------------
Hostname           HWControlPoint     Status    PowerStatus    Network-Interfaces
---------------------------------------------------------------------------------
node_01            hmcitso            on        on             en0-Online
node_02            hmcitso            on        on             en0-Online
node_03            hmcitso            on        on             en0-Online
node_04            hmcitso            on        on             en0-Online
node_05            hmcitso            off       off            unknown
node_06            hmcitso            on        on             en0-Online
node_07            hmcitso            off       off            unknown
node_08            hmcitso            on        on             eth0-Online
#
```

- **lsnode**

  We also used the **lsnode** command to verify the state of the node. The command **lsnode -n node_03 -a Mode** returned a value of `Managed`. This tells us that the node is now managed by the management server.

  Issuing the command **lsnode -p -n node_03** returned a value of `alive`. The node is online. See Example 3-27.

Example 3-27 Use of the lsnode command to verify status

```
# lsnode -n node_03 -a Mode
node_03:  Managed
# lsnode -p -n node_03
node_03:  1 (alive)
#
```

- **dsh**

  The **dsh** command can be used to verify that the secure shell has been set up correctly. We issued the **dsh** command to the node to execute the date command as in Example 3-28.

Example 3-28 Use of dsh to verify ssh

```
# dsh -w node_03 date
node_03: Wed Oct 13 17:52:51 CDT 2004
#
```

As these tests were successful, we can consider the node installed.

3.8 Integration of Linux pSeries nodes

The integration of Linux for pSeries nodes is a different process compared to the nodes that are running AIX. One of the major differences is that NIM Version 5.2 cannot install Linux onto our nodes. Linux must already be installed using the procedures for the distribution of Linux that is to be on the node. We installed various distributions and versions of Linux into three LPARs on our p650 nodes.

Refer to the IBM redbook *Deploying Linux on IBM eServer pSeries Clusters*, SG24-7014, for more about installing Linux on pSeries.

After Linux is operational on the node, CSM is installed onto the node using the following steps:

1. Check the node software.
2. Copy Linux packages to the correct directory.
3. Verify node definitions.

4. Create CSM node groups.
5. Set up the cluster configuration.
6. Add Linux nodes to the cluster.
7. Verify nodes.

3.8.1 Check the node software

We initially decided to install the CSM client at level 1.3.3 on our Linux nodes. This restricted us to Red Hat EL-AS 3 for PPC and SUSE LINUX SLES 8.1 for PPC.

As node_07 is running a distribution of Linux (SUSE LINUX SLES 9) that is not supported by CSM Version 1.3.3, we have to wait until we migrate our cluster to CSM version 1.4.0.3 before we can integrate it.

We obtained the CSM client code for pSeries Linux from the IBM cluster fixes Web page:

http://techsupport.services.ibm.com/server/cluster/fixes

We downloaded csm-linux-1.3.3.3.ppc64.tar.gz, gunzipped it, and untarred the file to the /export/eznim/extras/csm_plinux_1333 directory.

We also obtained an RPM for autoupdate. The autoupdate software is required to do software maintenance on the node from the CSM management server. Autoupdate is optional; however, warning messages will be issued during the integration of the Linux nodes into the cluster if it is not installed.

The RPM for autoupdate can be obtained from:

http://freshmeat.net/projects/autoupdate

We downloaded autoupdate-5.3.10-1.noarch.rpm and placed it into the /export/eznim/extras/csm_plinux_1333 directory with the CSM RPMs.

3.8.2 Copy Linux packages to the correct directory

The CSM RPMs and open source RPMs must be copied to the correct directory structure within the /csminstall file system.

We use the **copycsmpkgs** command to accomplish this. This command must be run for each version of CSM and Linux that we have installed.

This command builds the correct directory structure and copies the CSM RPMs, open source RPMs and some RPMs of the Linux distribution CD. We actually copied the respective Red Hat and SUSE LINUX CDs to a file system, so there was no need for us to mount the CDs. The **copycsmpkgs** command prompts us to mount the distribution CDs if the command cannot find the necessary RPM files that it needs. Example 3-29 shows output for **copycsmpkgs** commands we ran. We executed the command with our current working directory set to /export/eznim/extras/csm_plinux_1333. The files for each distribution were mounted to /mnt/RedHat or /mnt/SLES depending on which distribution we were setting up.

Example 3-29 The copycsmpkgs command

```
#copycsmpkgs -p .:/mnt/RedHat InstallDistributionName=RedHatEL-AS
InstallOSName=Linux InstallCSMVersion=1.3.3 InstallDistributionVersion=3
InstallPkgArchitecture=ppc64
#
# copycsmpkgs -p .:/mnt/SLES InstallDistributionName=SLES InstallOSName=Linux
InstallCSMVersion=1.3.3 InstallDistributionVersion=8.1
InstallPkgArchitecture=ppc64
#
```

The attributes passed to the **copycsmpkgs** command marry up with the node attributes of the nodes we wish to integrate into our cluster.

3.8.3 Verify node definitions

Now that the /csminstall file system directory and files infrastructure is in place, we have to alter some node attributes so that the CSM management server can correctly integrate the Linux nodes into the cluster. We altered the node attributes in the CSM database using the **chnode** command, as shown in Example 3-30.

Example 3-30 Using chnode to alter node attributes

```
# chnode -n node_05 InstallDistributionVersion=8.1 InstallPkgArchitecture=ppc64
InstallCSMVersion=1.3.3 InstallDistributionName=SLES InstallOSName=Linux
#
# chnode -n node_08 InstallDistributionName=RedHatEL-AS
InstallDistributionVersion=3 InstallPkgArchitecture=ppc64 InstallOSName=Linux
InstallCSMVersion=1.3.3
```

3.8.4 Create CSM node groups

We decided not to create any extra optional nodegroups for our Linux nodes because the predefined node groups suited our needs. Refer to 3.4.3, “Create nodegroups” on page 25 for more about creating node groups.

3.8.5 Set up the cluster configuration

We had previously set up CFM to copy the /etc/hosts file from the CSM management server for our AIX nodes. The same applies to the Linux nodes. No specific setup for Linux was done with CFM. If we had Linux-specific files to be distributed, we could still place them in /cfmroot and use the group feature of CFM to restrict the file to Linux nodes. Table 3-4 shows a simple example.

Table 3-4 CFM setup

Filename on the management server	Filename on node	Comments
/cfmroot/etc/hosts._LinuxNodes	/etc/hosts	Only distributes to nodes within the LinuxNodes node group.
/cfmroot/etc/hosts	/etc/hosts	Distributes to all nodes in the cluster.

Fore more about using CFM, see the *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.3.3,* SA22-7934.

3.8.6 Add Linux nodes to the cluster

Before integrating the node into the cluster, be sure that name resolution is working correctly between them. When we installed our Linux nodes, we made sure that the host name matched the name of the node as defined in the CSM database, that IP was set up correctly on the Linux node, and that the /etc/hosts file contained an entry for the CSM management server.

After ensuring that everything is ready for the node to be integrated into the CSM cluster, issue the **updatenode** command as shown:

```
# updatenode -v -n node_08
#
```

We specify the -v flag to see what activity is occurring on the node throughout the updatenode process. Any errors that are encountered will be displayed.

We also checked the install log on the node after the updatenode completed. The file is located in the /var/log/csm/install.log on the node.

3.8.7 Verify nodes

After the update node has completed successfully, we can verify the node using the same commands as for AIX nodes. See 3.7.3, “Monitoring the install” on page 33 for more about verifying the newly integrated Linux node.

4

CSM advanced features implementation

This chapter describes how AIX servers and Linux servers can coexist in a CSM managed environment. Both types of servers can be managed using an AIX management server (MS). In this chapter, we assume that an AIX environment already exists, that you are adding some Linux PPC servers to the same network, and that you want them to be managed by an existing AIX MS server.

This chapter covers:

- 4.1, “Building a Linux repository for pSeries” on page 40
- 4.2, “Install Linux using the AIX NIM server” on page 60
- 4.3, “CSM performance, scalability, and challenges” on page 66

4.1 Building a Linux repository for pSeries

Installing Linux on a pSeries server is similar to installing Linux on an xSeries server. The planning is almost the same, but you have to decide what distribution of Linux to use, the disk configuration, the network configuration, and the software configuration.

However, there are some differences, and most of them are due to the hardware. A pSeries server offers multiple choices regarding the hardware configuration. A pSeries can be configured in full server mode, logically partitioned mode (LPAR), and virtual LPAR on the new IBM eServer p5 systems. The IBM Blade JS20 can also be used in the same environment.

Note: In this book, we assume that you are familiar with Linux and pSeries servers. The redbook *Deploying Linux on IBM eServer pSeries Clusters*, SG24-7014, has more information about installing Linux on PPC, and this chapter is based on that redbook.

In the following sections, we describe how to set up the repository for Red Hat Version 3.0 update 3, SUSE LINUX Enterprise Server Version 8 Service Pack 3 (SLES8 SP3), and SLES9. As we write this book, the Red Hat Advanced Server (AS) V4 is still in beta version.

Note: In this book, we work with Linux for PowerPC® only (PPC).

A high-level view of the Linux installation includes the following steps:

1. Set up the network installation (for example: BOOTP, NFS, console access).
2. Build the Linux repository. All of the files that must be installed from the network are located in this repository.
3. Migrate the installation server to AIX 5.3 management server using Network Installation Manager (NIM).

The migration step is the last step in our scenario, but we have to prepare the configuration on Linux Installation Manager (LIM) server in such way as to ease the migration process.

4.1.1 Setting up the network installation server

We use a Linux server for network installation (in our case, a Red Hat AS 3.0 U3). We call the Linux server for network installation Linux Install Manager (LIM). Using a Linux server to install another Linux server on an IBM pSeries system is the *supported* method. After we set up the Linux installation server, we migrate the LIM to an AIX server and use NIM for network installation for all nodes, both Linux and AIX. This way, we can have a single CSM management server and a single network installation.

> **Note:** Using NIM on AIX 5.3 to install Linux servers is not yet supported.

To install Linux on a PPC successfully, you have to follow the steps, not necessary in this order:

- Configure the network boot method; you can use BOOTP or DHCP.
- Configure the TFTP server.
- Configure the network transfer method; you can use NFS, HTTP, FTP, or Samba.
- Create the repository for each Linux distribution and version.
- Create the autoinstallation file, kickstart, or AutoYaST.
- Create, if necessary, the pre-installation and post-installation scripts.
- Install the Linux OS.
- Apply patches, if necessary.
- Install the CSM client.
- Reboot the server, if necessary.

Network boot

On the LIM server, we create a central path for the network boot server. We decided to use a separated partition and mount it as /lim, but it can be any path you want.

DHCP server

The pSeries servers support network booting using BOOTP. The DHCP server supports BOOTP protocol and has more features than BOOTP. We are using DHCP server Version 3.0pl1-23 on the LIM server. In this book, we assume that you have experience with DHCP and know how to configure the DHCP server. If not, refer to:

http://www.isc.org/index.pl?/sw/dhcp/

Example 4-1 shows the DHCP configuration file that we used.

Example 4-1 DHCP configuration file

```
#cat /etc/dhcpd.conf
default-lease-time 600;
max-lease-time 7200;
ddns-update-style none;
always-reply-rfc1048 true;

subnet 192.168.100.0 netmask 255.255.255.0 {
```

```
         range 192.168.100.192 192.168.100.196;
        option routers 192.168.100.60;

host node_08 {
              hardware ethernet 00:02:55:6f:1f:ef;
              fixed-address 192.168.100.190;
              filename "netboot.img";
            }

host node_07 {
              hardware ethernet 00:02:55:3a:07:db;
              fixed-address 192.168.100.189;
              filename "install.sles9";
            }

host node_05 {
              hardware ethernet 00:02:55:6f:1c:35;
              fixed-address 192.168.100.187;
              filename "install.sp3.sles8";
            }

host node_09 {
              hardware ethernet 00:02:55:6a:4e:36;
              fixed-address 192.168.100.191;
              filename "install.sles9.test";
            }

        }
```

Note:

- In order for the BOOTP to work, you have to define a fixed network address for every node (server). You can find the MAC address for the Ethernet adapter in the SMS menu, under Adapter information page.
- yaboot does not work over the network, so you cannot send parameters with the kernel.
- You can use the DHCP server for other purposes, but after we migrate to the NIM server we use the DHCP only for BOOTP.
- The DHCP server has to listen on port UDP 67.
- Because the Linux kernel is bigger than 2 MB, you must insert a permanent ARP entry for the MAC address; otherwise the kernel does not load. For example:

  ```
  #arp -s 92.168.100.191 00:02:55:6a:4e:36
  ```

TFTP server

In order for the TFTP server to work, you must configure the path to the tftpboot directory. In our case, we chose the path /lim/tftpboot. To verify that the tftpboot is running, you can check whether it listens on the port UDP 69. In the /lim/tftpboot path, we install the kernel files.

NFS server

We use the NFS server for node installation through the network because the AIX NIM server works only with NFS so it is easy to migrate to AIX. To set up the NFS server on LIM, edit the /etc/exportfs file. Example 4-2 shows our file.

Example 4-2 The NFS configuration file

```
#cat /etc/exports
/lim/SLES8/ *(rw,no_root_squash,insecure)
/lim/SLES9/ *(rw,no_root_squash,insecure)
/lim/rh-ppc/ *(rw,no_root_squash,insecure)
```

For every distribution, we use a separate directory in order to move or migrate each distribution easily to the NIM server. For the NFS server to start properly on the LIM server, the portmap service has to run. To check whether the NFS is running, mount the shared directories on the same server. If it is not working, check the configuration. For more information about the NFS server, refer to:

http://nfs.sourceforge.net/nfs-howto/

HTTP server

We recommend that when building the Linux repository you use HTTP install because it is easy to debug if something goes wrong. With HTTP you have the access logs, and you can see exactly which files are transferred over the network. For example, it is easy to create the repository SLES8 with Service Pack 3 install at the same time. In the access logs from the HTTP server, you can see whether the files are downloaded in the right order or whether something might be missing, as shown in Example 4-3.

Example 4-3 HTTP access log

```
//sles8/media.1/media HTTP/1.0" 200 31 "-" "Wget/1.8.2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET //sles8/yast/order
HTTP/1.0" 200 117 "-" "Wget/1.8.2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET //sles8/yast/instorder
HTTP/1.0" 200 117 "-" "Wget/1.8.2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/media.1/info.txt HTTP/1.0" 404 324 "-" "Wget/1.8.
2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/media.1/media HTTP/1.0" 200 31 "-" "Wget/1.8.2"
```

```
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/media.1/products HTTP/1.0" 200 24 "-" "Wget/1.8.2
"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/content HTTP/1.0" 200 546 "-" "Wget/1.8.2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/suse/setup/descr/selections HTTP/1.0" 404 335 "-"
 "Wget/1.8.2"
192.168.100.191 - - [06/Oct/2004:19:53:59 -0400] "GET
//sles8/service-pack-3/CD1/suse/setup/descr/packages HTTP/1.0" 404
```

Note: In the HTTP access log, the first tab after the file transfer tab is the error code of the command. In Example 4-3, in the last line after the file transfer tab is the number 404, which means that the HTPP server count not find the file. The number 200 means the file is found. For more information about HTTP logs, refer to:

http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html

In order to successfully install from an HTTP server, you have to create an alias from the DocumentRoot to the directory where the repository is. If you are using Red Hat as the installation server, create a link as shown in Example 4-4. The URL for the HTTP installation is http://<IP Address>/rh-ppc/ in our case.

Example 4-4 Creating a link for the HTTP repository

```
[root@lnx html]# cd /var/www/html/
[root@lnx html]# ll
total 0
lrwxrwxrwx    1 root     root           12 Oct  6 11:18 rh-ppc -> /lim/rh-ppc/
lrwxrwxrwx    1 root     root           10 Oct  6 10:07 sles8 -> /lim/SLES8
lrwxrwxrwx    1 root     root           11 Oct  5 18:17 sles9 -> /lim/SLES9/
```

If you are using SUSE LINUX as the HTTP installation, create an alias in the Web server configuration file /etc/apache2/default-server.conf, as shown in Example 4-5. In the configuration file, /SLES9-ppc is the alias and /srv/SLES9-ppc is the full path of the location on the disk. The URL for the HTTP installation is http://<IP Address>/SLES9-ppc/.

Example 4-5 SUSE LINUX HTTP install configuration

```
Alias /SLES9-ppc /srv/SLES9-ppc

<Directory "/srv/SLES9-ppc">
 Options Indexes MultiViews
 AllowOverride None
 Order allow,deny
 Allow from all
</Directory>
```

4.1.2 Red Hat repository

To create the Red Hat repository, follow these steps:

1. Define the repository directory, in our case /lim/rh-ppc.
2. Copy all CDs into the /lim/rh-ppc directory.

Example 4-6 shows the repository for Red Hat.

Example 4-6 Red Hat repository

```
#ll
total 264
-rwxr-xr-x    1 root     root          248 Sep  2 22:54 autorun
-rw-r--r--    1 root     root        32768 Sep  2 22:28 boot_image
drwxr-xr-x    2 root     root         4096 Sep  2 23:07 etc
-rw-r--r--    1 root     root         7207 Sep 26  2003 EULA
-rw-r--r--    1 root     root        18416 Sep 26  2003 GPL
drwxr-xr-x    2 root     root         4096 Sep  2 23:07 images
-rw-r--r--    1 root     root         2072 Oct 11 14:43 ks.cfg
drwxr-xr-x    4 root     root         4096 Sep  2 23:07 ppc
-rw-r--r--    1 root     root        11586 Sep  2 22:54 README-Accessibility
-rw-r--r--    1 root     root         5117 Sep  2 22:54 README-en
drwxr-xr-x    4 root     root         4096 Sep  2 23:04 RedHat
-rw-r--r--    1 root     root        38159 Sep  2 22:54 RELEASE-NOTES-en
-rw-r--r--    1 root     root        14800 Sep  2 22:54 RELEASE-NOTES-U1-en
-rw-r--r--    1 root     root        18015 Sep  2 22:54 RELEASE-NOTES-U2-en
-rw-r--r--    1 root     root        21107 Sep  2 22:54 RELEASE-NOTES-U3-en
-rw-r--r--    1 root     root         1910 Jul 15  2003 RPM-GPG-KEY
-rw-r--r--    1 root     root         1706 Sep  2 22:54 RPM-GPG-KEY-beta
drwxr-xr-x    2 root     root         4096 Sep  2 23:13 SRPMS
-r--r--r--    1 root     root        15716 Sep  2 23:23 TRANS.TBL
```

4.1.3 SUSE LINUX SLES8 repository

SUSE LINUX includes the update packages as a separate CD called ServicePack. As this book is being written, the last ServicePack for SLES8 is Version 3.

We want to install SLES8 and SP3 at the same time so we do not have to reboot the server and it does not need human intervention. To do this, build the SLES8 repository in the following way:

1. Define the repository directory; in our case it is /lim/SLES8.

2. Copy the SLES8 CD1 into /lim/SLES8/CD1, CD2 into /lim/SLES8/CD2, and CD3 into /lim/SLES8/CD3.
3. Copy SP3 CD1 into /lim/SLES8/SP3/CD1 and CD2 into /lim/SLES8/SP3/CD2.
4. Copy the /lim/SLES8/CD1/boot directory to /lim/SLES8/boot.
5. Copy the /lim/SLES8/CD1/content file to /lim/SLES8/content.
6. Copy the /lim/SLES8/SP3/CD1/driverupdate file to /lim/SLES8/driverupdate.
7. Copy the /lim/SLES8/CD1/media.1 directory to /lim/SLES8/media.1.
8. Create these files: /lim/SLES8/yast/order and /lim/SLES8/yast/instorder.

Example 4-7 shows the order file, and Example 4-8 shows the instorder file.

Example 4-7 The SLES8 order file

```
/SP3/CD1        /SP3/CD1
/CD1    /CD1
```

Example 4-8 The SLES8 instorder file

```
/CD1    /CD1
/SP3/CD1        /SP3/CD1
```

Note:

- In the files /lim/SLES8/yast/order and /lim/SLES8/yast/instorder, the space between the first column and the second has to be one tab.
- If you are installing using HTTP server, you can create symbolic links for the files in root of the repository, but it you want to use NFS you have to copy the files, not create symbolic links.
- In the /lim/SLES8/SP3/CD1 directory, create a symbolic link setup that points to ppc/update/SuSe-SLES/8/rpm/setup. Otherwise the SP3 installation does not work and it installs only the base level.

4.1.4 SUSE LINUX SLES9 repository

As this book is being written, SUSE LINUX SLES9 does not have a ServicePack, so it is easier to build the repository. Follow these steps:

1. Copy SLES9 CD1 to /lim/SLES9/SUSE-SLES-Version-9/CD1.
2. Copy SLES CD2 through CD6 to /lim/SUSE-CORE-Version-9/CD2 through /lim/SUSE-CORE-Version-9/CD6.

3. Copy the file /lim/SLES9/SUSE-SLES-Version-9/CD1/content to /lim/SLES9/content.
4. Copy the file /lim/SLES9/SUSE-SLES-Version-9/CD1/control.xml to /lim/SLES9/control.xml.
5. Copy the directory /lim/SLES9/SUSE-SLES-Version-9/CD1/boot to /lim/SLES9/boot.
6. Copy the directory /lim/SLES9/SUSE-SLES-Version-9/CD1/media.1 to /lim/SLES9/media.1.
7. Create files /lim/SLES9/yast/order and /lim/SLES9/yast/instorder.

Example 4-9 shows the order file, and Example 4-10 shows the instorder file.

Example 4-9 The SLES9 order file

```
/SUSE-SLES-Version-9/CD1	/SUSE-SLES-Version-9/CD1
/SUSE-CORE-Version-9/CD1	/SUSE-CORE-Version-9/CD1
```

Example 4-10 The SLES9 instorder file

```
/SUSE-SLES-Version-9/CD1
/SUSE-CORE-Version-9/CD1
```

Note:

- In the files /lim/SLES9/yast/order and /lim/SLES9/yast/instorder, the space between the first column and the second has to be one tab.
- If you are installing using HTTP server, you can create symbolic links for the files in root of the repository, but it you want to use NFS you have to copy the files, not create symbolic links.

4.1.5 Creating the autoinstallation files

Creating autoinstallation files (ks.cfg for Red Hat or autoyast.xml for SUSE LINUX) can take a long time, especially if you want create complicated configurations. In this book, we need an autoinstallation file just for installing Linux remotely via the network. It is beyond the scope of this book to create complicated autoinstallation files for different flavors of Linux, although you may need such a configuration in your real environment, so visit the following Web sites for additional information:

- For SUSE LINUX

  http://www.suse.de/~nashif/autoinstall/

- For Red Hat

  http://www.redhat.com/docs/manuals/linux/RHL-9-Manual/custom-guide/part-install-info.html

> **Attention:** Most of the documentation about autoinstallation files on the Internet is for general use and it applies best to Intel® servers not on PPC. So you may need to adapt the documentation to fit your requirements.

PPC PRep boot

The PowerPC platform boots differently from the i386 platform, so you have to create a partition called *PPC PRep Boot*. The partition ID is 41 and has to be less than 10 MB. Usually 4 MB is enough. In this partition, the kernel is installed by the Linux installation. Because of this partition, there are some limitations:

- The partition PPC PRep boot cannot be mirrored.
- You cannot create RAID 1 for booting partition.
- As this book is being written, kickstart does not know to create the PPC PRep Boot partition. You have to use the autopart option for the root disk.

> **Note:** If you have more disks in the server, the autopart option in the kickstart file will use more than one disk. It will create a swap partition on a separate disk and create the /(root) partition as large as the disk. Our work-around solution is to use one disk at installation time.

> **Attention:** If for some reason the installation fails and the disk does not have a boot label, it disappears from the boot list in the SMS, which appears to have no disks in the system. To fix this issue, you have to relabel the boot or reinstall Linux.

Kickstart file

Example 4-11 shows the kickstart file we are using, ks.cfg.

Example 4-11 The ks.cfg file

```
# cat ks.cfg
#Generated by Kickstart Configurator
#System  language
lang en_US
#Language modules to install
langsupport --default=en_US
#System keyboard
keyboard us
#System mouse
```

```
mouse generic3ps/2
#Sytem timezone
timezone America/New_York
#Root password
rootpw --iscrypted $1$VoJeNV7b$4N5sr.fFEinBLTsqhOUec/
#Reboot after installation
reboot
#Use text mode install
text
#Install Red Hat Linux instead of upgrade
install
#Use Web installation
nfs --server=192.168.100.180 --dir=/lim/rh-ppc/
#System bootloader configuration
bootloader --location=mbr
#Clear the Master Boot Record
zerombr yes
#Partition clearing information
clearpart --all --initlabel
#Disk partitioning information
autopart
#part  None --fstype "PPC PRep Boot" --size=8 --asprimary
#part /boot --fstype ext3 --size 100 --ondisk sda
#part / --fstype ext3 --size 10000 --ondisk sda
#part swap --size 1024 --asprimary --ondisk sda
#System authorization infomation
auth  --useshadow  --enablemd5
#Network information
network --bootproto=dhcp --device=eth0 --hostname node_08
#network --bootproto=dhcp --device=eth1
#Firewall configuration
firewall --disabled
#Do not configure XWindows
skipx
#Package install information
%packages --resolvedeps
@ X Window System
@ KDE Desktop Environment
@ Graphical Internet
@ Text-based Internet
@ Server Configuration Tools
@ Web Server
@ Mail Server
@ Windows File Server
@ DNS Name Server
@ FTP Server
@ Network Servers
@ Development Tools
@ Kernel Development
@ Administration Tools
@ System Tools
@ Printing Support
%post --nochroot
echo "192.168.100.69  hmcitso # HMC
192.168.100.182 ms_01   # CSM MS Primary Server
192.168.100.181 ms_02   # CSM MS Backup Server
192.168.100.183 node_01 # AIX 5.1 ML03
192.168.100.184 node_02 # AIX 5.1 ML03
192.168.100.185 node_03 # AIX 5.2 ML01
192.168.100.186 node_04 # AIX 5.2 ML03
192.168.100.187 node_05 # SUSE SLES 8
192.168.100.188 node_06 # AIX 5.2 ML03
192.168.100.189 node_07 # RedHat
192.168.100.190 node_08 # SUSE SLES 8
192.168.100.180 lnx     # Linux Install
192.168.100.191 node_09 # p690_lpar3" >> /mnt/sysimage/etc/hosts
```

During the installation, we noticed the following:

- The VNC or SSH installation does not work; only the console installation works.
- It is possible to create volume groups in the kickstart file. However, you have to create a custom partitioning and you cannot specify the PPC PRep boot partition, unless you create a pre-install script that does the disk partitioning.
- If you want to modify the root partition with a post-install script with the --nochroot option, you have to add /mnt/sysimage in the path as we did in our kickstart file.
- In order for Red Hat to boot from the network using the kickstart configuration, you have to use the prompt in the SMS firmware and pass the kickstart parameters as shown. If you have more Ethernet adapters, the ksdevice parameter has to be entered.

  ```
  boot net ks=nfs:192.168.100.180:/lim/rh-ppc ksdevice=eth0
  ```

  > **Note:** The root password for this kickstart server is `itsoadmin`.

SLES8 AutoYaST

Example 4-12 shows the AutoYaST file (in our case sles8.lvm.all).

Example 4-12 The SLES8 AutoYaST file

```
# cat sles8.lvm.all
<?xml version="1.0"?>
<!DOCTYPE profile SYSTEM "/usr/share/YaST2/include/autoinstall/profile.dtd">
<profile xmlns="http://www.suse.com/1.0/yast2ns"
```

```
xmlns:config="http://www.suse.com/1.0/configns">
<configure>
<networking>
<dns>
<dhcp_hostname config:type="boolean">false</dhcp_hostname>
<dhcp_resolv config:type="boolean">false</dhcp_resolv>
<domain>local</domain>
<hostname>node_05</hostname>
</dns>
<interfaces config:type="list">
<interface>
<bootproto>dhcp</bootproto>
<device>eth0</device>
<module>e100</module>
<startmode>onboot</startmode>
<wireless>no</wireless>
</interface>
</interfaces>
<routing>
<ip_forward config:type="boolean">false</ip_forward>
<routes config:type="list">
<route>
<destination>default</destination>
<device>-</device>
<gateway>192.168.100.60</gateway>
<netmask>-</netmask>
</route>
</routes>
</routing>
</networking>
<scripts>
<post-scripts config:type="list">
<script>
<filename>custom</filename>
<interpreter>shell</interpreter>
<source>
<![CDATA[# Very basic example of a custom script
#echo "export PATH=$PATH:/opt/ibmcmp/xlf/8.1/bin" > /etc/profile.d/xlf.sh
echo "node_05" > /etc/HOSTNAME
echo "192.168.100.69  hmcitso # HMC console
192.168.100.182 ms_01   # CSM Primary Server
192.168.100.181 ms_02   # CSM Backup Server
192.168.100.183 node_01 # AIX 5.1 ML03
192.168.100.184 node_02 # AIX 5.1 ML03
192.168.100.185 node_03 # AIX 5.2 ML01
192.168.100.186 node_04 # AIX 5.2 ML03
192.168.100.187 node_05 # SUSE SLES 8
192.168.100.188 node_06 # AIX 5.2 ML03
192.168.100.189 node_07 # RedHat
192.168.100.190 node_08 # SUSE SLES 8
192.168.100.180 lnx     # SUSE SLES 8
192.168.100.191 node_09 # P690 lpar3 Linux" >> /etc/hosts
mkdir /media/nfs
mount ms_01:/csminstall/Linux /media/nfs
nohup /media/nfs/csm.install &
umount /media/nfs
]]>
</source>
</script>
</post-scripts>
</scripts>
<users config:type="list">
<user>
<encrypted config:type="boolean">true</encrypted>
<user_password>hRoWvYp76SFiU</user_password>
<username>root</username>
</user>
#<user>
#<encrypted config:type="boolean">true</encrypted>
#<user_password>3VBnRzl.nZu06</user_password>
#<username>admin</username>
#</user>
</users>
</configure>
<install>
<general>
<clock>
<timezone>US/Eastern</timezone>
</clock>
<keyboard>
<keymap>english-us</keymap>
</keyboard>
<language>en_US</language>
<mode>
<confirm config:type="boolean">false</confirm>
<forceboot config:type="boolean">false</forceboot>
<interactive_boot config:type="boolean">false</interactive_boot>
<reboot config:type="boolean">false</reboot>
</mode>
<mouse>
<id>00_ps2</id>
</mouse>
</general>
<partitioning config:type="list">
<drive>
<device>/dev/sda</device>
<initialize config:type="boolean">true</initialize>
<partitions config:type="list">
```

```
<partition>
<crypt_fs config:type="boolean">false</crypt_fs>
<format config:type="boolean">false</format>
<partition_id config:type="integer">65</partition_id>
<partition_type>primary</partition_type>
<size>4MB</size>
</partition>
<partition>
<crypt_fs config:type="boolean">false</crypt_fs>
<crypt_key></crypt_key>
<filesystem config:type="symbol">swap</filesystem>
<format config:type="boolean">true</format>
<mount>swap</mount>
<partition_id config:type="integer">130</partition_id>
<partition_type>primary</partition_type>
<size>1024MB</size>
</partition>
<partition>
<crypt_fs config:type="boolean">false</crypt_fs>
<crypt_key></crypt_key>
<filesystem config:type="symbol">reiser</filesystem>
<format config:type="boolean">true</format>
<mount>/</mount>
<partition_id config:type="integer">131</partition_id>
<partition_type>primary</partition_type>
<size>10GB</size>
</partition>
</partitions>
<use>all</use>
</drive>
</partitioning>
<software>
<addons config:type="list">
        <addon>64bit</addon>
        <addon>Basis-Devel</addon>
        <addon>Kde-Desktop</addon>
        <addon>X11</addon>
       <addon>YaST2_modules</addon>
        <addon>analyze</addon>
        <addon>auth</addon>
        <addon>cross-ppc64</addon>
        <addon>dhcp_dns</addon>
        <addon>file_print</addon>
        <addon>mail_news</addon>
        <addon>sles_admin</addon>
      </addons>
      <base>default</base>
</software>
```

```
</install>
</profile>
```

Note: The partition ID number in the AutoYaST file is decimal, and in Linux is hexadecimal, so you only have to transform the numbers.

During the installation we noticed the following:

- It is possible to create logical volumes right from the installation.
- It is possible to create the PPC PRep boot partition in AutoYaST.
- SLES 8 SP3 comes with a very interesting tool that can “burn” command line arguments directly into a SUSE LINUX kernel, using a built-in static 512 characters. This is described in the /ppc/netboot/ directory of the SP3 CD. To create a customized kernel for the unattended installation, copy the install kernel into the /lim/tftpboot directory and issue the command as shown in Example 4-13. This is very handy because we do not have to go to Open Firmware any more. Just boot through TFTP and the kernel will start linuxrc with the right arguments.

Example 4-13 Burn the kernel

```
/lim/tftpboot:# ...../sp3/ppc/netboot/mkzimage_cmdline -a 1 -c \
-s “insmod=e100 install=nfs://192.168.100.180/lim/SLES8 \
autoyast=nfs://192.168.100.180:/lim/SLES8/sles8.lvm.all install.sp3.sles8
```

To recover a “neutral” kernel, simply run as shown in Example 4-14. This is a far-reaching utility, as it can be used to pass any kind of argument to. At the time of writing, RHAS 3 does not provide an equivalent utility.

Example 4-14 Clean the kernel

```
/tftpboot:# ....../mkzimage_cmdline -a 0 -c install
```

Note: The root password for this AutoYaST is *itsoadmin.*

SLES9 AutoYaST

The AutoYaST configuration is very similar to the SUSE LINUX SLES 8 AutoYaST configuration. Example 4-15 shows the AutoYaST configuration.

Example 4-15 SLES9 AutoYaST configuration file

```
# cat  ../SLES9/autoyast
<?xml version="1.0"?>
<!DOCTYPE profile SYSTEM "/usr/share/autoinstall/dtd/profile.dtd">
```

```
<profile xmlns="http://www.suse.com/1.0/yast2ns"
xmlns:config="http://www.suse.com/1.0/configns">
  <configure>
    <networking>
      <dns>
        <dhcp_hostname config:type="boolean">false</dhcp_hostname>
        <dhcp_resolv config:type="boolean">false</dhcp_resolv>
        <domain>localdomain</domain>
        <hostname>node_07</hostname>
      </dns>
      <interfaces config:type="list">
        <interface>
          <bootproto>dhcp</bootproto>
          <broadcast>192.168.100.255</broadcast>
          <device>eth0</device>
          <ipaddr>192.168.100.189</ipaddr>
          <netmask>255.255.255.0</netmask>
          <network>192.168.100.0</network>
          <startmode>onboot</startmode>
        </interface>
      </interfaces>
      <modules config:type="list">
        <module_entry>
          <device>static-0</device>
          <module>e100</module>
          <options></options>
        </module_entry>
      </modules>
      <routing>
        <ip_forward config:type="boolean">false</ip_forward>
      </routing>
    </networking>
<scripts>
<post-scripts config:type="list">
<script>
<chrooted config:type="boolean">true</chrooted>
<filename>add_hostfile</filename>
<interpreter>shell</interpreter>
<source>
<![CDATA[# Very basic example of a custom script
#echo "export PATH=$PATH:/opt/ibmcmp/xlf/8.1/bin" > /etc/profile.d/xlf.sh
echo "node_07" > /etc/HOSTNAME
echo "192.168.100.69  hmcitso # HMC console
192.168.100.182 ms_01   # CSM Primary Server
192.168.100.181 ms_02   # CSM Backup Server
192.168.100.183 node_01 # AIX 5.1 ML03
192.168.100.184 node_02 # AIX 5.1 ML03
192.168.100.185 node_03 # AIX 5.2 ML01
192.168.100.186 node_04 # AIX 5.2 ML03
192.168.100.187 node_05 # SUSE SLES 8
192.168.100.188 node_06 # AIX 5.2 ML03
192.168.100.189 node_07 # SLES9
192.168.100.190 node_08 # RedHat
192.168.100.180 lnx     # RedHat
192.168.100.191 node_09 # P690 lpar3 Linux" >> /etc/hosts
/etc/init.d/portmap restart
mkdir /media/nfs
mount ms_01:/csminstall/Linux /media/nfs
nohup /media/nfs/csm.install &
umount /nfs/cdrom
]]>
</source>
</script>
</post-scripts>
</scripts>
<users config:type="list">
<user>
<encrypted config:type="boolean">true</encrypted>
<user_password>hRoWvYp76SFiU</user_password>
<username>root</username>
</user>
</users>
  </configure>
  <install>
    <general>
      <clock>
        <hwclock>UTC</hwclock>
        <timezone>US/Eastern</timezone>
      </clock>
      <keyboard>
        <keymap>english-us</keymap>
      </keyboard>
      <language>en_US</language>
      <mode>
        <confirm config:type="boolean">false</confirm>
        <forceboot config:type="boolean">true</forceboot>
      </mode>
      <mouse>
        <id>none</id>
      </mouse>
    </general>
    <partitioning config:type="list">
      <drive>
        <device>/dev/sda</device>
        <initialize config:type="boolean">true</initialize>
        <partitions config:type="list">
          <partition>
            <crypt>twofish256</crypt>
            <format config:type="boolean">false</format>
            <loop_fs config:type="boolean">false</loop_fs>
            <mount></mount>
            <partition_id config:type="integer">65</partition_id>
            <partition_type>primary</partition_type>
            <size>4M</size>
          </partition>
          <partition>
            <crypt>twofish256</crypt>
            <filesystem config:type="symbol">reiser</filesystem>
            <format config:type="boolean">true</format>
            <loop_fs config:type="boolean">false</loop_fs>
            <mount>/</mount>
            <partition_id config:type="integer">131</partition_id>
            <partition_type>primary</partition_type>
            <size>10G</size>
          </partition>
<partition>
<crypt_fs config:type="boolean">false</crypt_fs>
<crypt_key></crypt_key>
<filesystem config:type="symbol">swap</filesystem>
<format config:type="boolean">true</format>
<mount>swap</mount>
<partition_id config:type="integer">130</partition_id>
<partition_type>primary</partition_type>
<size>1024MB</size>
</partition>
        </partitions>
        <use>all</use>
      </drive>
    </partitioning>
    <software>
      <addons config:type="list">
        <addon>Base-System</addon>
        <addon>Basis-Devel</addon>
        <addon>Basis-Sound</addon>
        <addon>File-Server</addon>
        <addon>Gnome</addon>
        <addon>HA</addon>
        <addon>Kde-Desktop</addon>
        <addon>LAMP</addon>
        <addon>LSB</addon>
        <addon>Linux-Tools</addon>
        <addon>Mail_News-Server</addon>
        <addon>Print-Server</addon>
        <addon>SuSE-Documentation</addon>
        <addon>Various-Tools</addon>
        <addon>WBEM</addon>
        <addon>X11</addon>
        <addon>YaST2</addon>
        <addon>analyze</addon>
        <addon>auth</addon>
        <addon>dhcp_DNS-Server</addon>
      </addons>
      <base>default</base>
    </software>
  </install>
</profile>
#
```

Note: The root password for this AutoYaST is `itsoadmin`.

4.1.6 Install the CSM client on Linux

In this paragraph, we describe how to install the CSM client using the post-installation scripts that Red Hat and SUSE LINUX provide. This way, we can create a fully unattended Linux installation on pSeries.

In this chapter, we assume that the CSM server installation and configuration are complete. (Refer to 3.8, “Integration of Linux pSeries nodes” on page 35.)

Note: The CSM client install for Linux is supported as described in 3.8, “Integration of Linux pSeries nodes” on page 35. This paragraph describes a different but unsupported way of installing CSM on Linux nodes.

We set up the CSM unattended installation only for SUSE LINUX servers because SUSE LINUX SLES 8 and SLES 9 do not reboot after install. For Red Hat, additional scripts must be created.

Note: Red Hat presents a challenge because the NFS daemon rpc.statd uses bindresvport() to obtain an arbitrary UDP port number in the range 600-1023. Sometimes rpc.statd grabs port 657, which RMCD needs. The bindresvport function call completely overrides /etc/services where rmcd port numbers are defined. The workaround is:

```
#/etc/init.d/nfslock stop
#startsrc -s ctrmc
#/etc/init.d/nfslock start
```

On the CSM server, NFS exports the /csminstall/csm and /csminstall/Linux directories.

In the AutoYaST file, we add a script (shown in Example 4-16) that practically runs a script from the CSM server in the background. It is important to run the script in the background so that the unattended Linux installation does not stop and wait for the script to finish.

Example 4-16 Post script in AutoYaST

```
mount ms_01:/csminstall/Linux /media/nfs
nohup /media/nfs/csm.install &
umount /media/nfs
```

Before starting the unattended installation, the node definition in CSM must be set up properly, as shown in Example 4-17.

Example 4-17 Modified CSM definition

```
#chnode -n node_07 AllowManageRequest=1
#chnode -n node_07 Mode=installing
#lsnode  -a AllowManageRequest,Mode -l node_07
 AllowManageRequest = 1 (allow node management request)
 Mode = installing
```

Example 4-18 shows the CSM install script. We could not install RPMs in the post script execution phase, so we have to wait in the background for the Linux installation to finish and then run the CSM installation. That is why we use the command **`sleep 180`** in the csm.install script, which also exchanges the ssh keys with the MS server in order to log in using ssh without asking for a password.

Example 4-18 csm.install script

```
#!/bin/bash
sleep 180
mkdir -p /var/opt/csm/mnt/csm
mkdir -p /var/opt/csm/mnt/Linux
mkdir -p /opt/csm/install
mkdir -p /var/log/csm
mkdir -p /root/.ssh
mount ms_01:/csminstall/csm /var/opt/csm/mnt/csm
mount ms_01:/csminstall/Linux /var/opt/csm/mnt/Linux
cp -a /var/opt/csm/mnt/csm/config/.ssh/auth* /root/.ssh/
chmod 600 /root/.ssh/auth*
cp -a /var/opt/csm/mnt/csm/defs /opt/csm/
cp -a /var/opt/csm/mnt/csm/pkgdefs /opt/csm/
cp -a /var/opt/csm/mnt/csm/config/`hostname`.config_info
/opt/csm/install/configinfo
cd /var/opt/csm/mnt/csm
./makenode -m ms_01 -n `hostname` -k -v
cd /
```

```
umount /var/opt/csm/mnt/Linux
umount /var/opt/csm/mnt/csm/
umount /media/nfs
```

4.2 Install Linux using the AIX NIM server

This chapter is intended to be technical documentation to show how a network installation service for Linux can be set up for multiple hardware platforms on an AIX 5L™ Version 5.3 NIM server. Regarding the nature of the process, it will work similar to other hardware platforms, namely: IA-64, AMD Opteron™, UltraSPARC, Intel® Xeon™, and Intel® Pentium® 4 supporting Intel® EM64T.

With the previous version of AIX 5L (AIX 5L Version 5.2), it was not possible to accomplish the whole process by OS-owned means. The major problem was that RFC 2349 (negotiation of timeout and tsize parameters between server and client) was not implemented for the tiny-FTP subsystem (TFTP and TFTPD) which plays an important role in the network installation process. The RFC 2349 is now supported with the TFTP subsystem of AIX 5L Version 5.3. With the new version of AIX 5L, everything to build a multiple platform installation server comes with the operating system except the Linux installation sources.

The setup is independent from the hardware platform and should work just the same for every hardware platform that supports either BOOTP or PXE boot netboot protocol.

Note: The information in this chapter should be considered as a guideline on how to install Linux using an AIX NIM server. The installation method described in this chapter is not supported.

4.2.1 Switch the netboot method from BOOTP to DHCP

The standard netboot protocol for pSeries systems is BOOTP, but netboot requests for Intel-based Linux clients can only be satisfied by using DHCP. This means that both services would have to be available at the same time to serve both platforms as a netboot server. But AIX 5L V5.3 allows only one of both netboot methods to be active. The advantage of the AIX 5L V5.3 DHCP service is that it can handle both kinds of netboot requests but only if the native BOOTP service via inetd is disabled and put under the control of DHCP.

Since DHCP does not understand the BOOTP configuration file syntax, it is necessary to convert it to DHCP style. Section 4.2.2, “Prepare AIX 5L V5.x NIM machine definitions” on page 61 describes how to use existing NIM machine definitions and convert them from BOOTP to DHCP syntax, how the BOOTP configuration is transformed, and how DHCP must be configured to serve as a netboot server for AIX 5L V5.x NIM clients.

> **Note:** The customization steps below assume that there is an existing NIM server environment in place. This document cannot be used for an initial setup of a NIM environment. Refer to the appropriate documentation to set up the correct environment before continuing with this customization.

4.2.2 Prepare AIX 5L V5.x NIM machine definitions

To be able to easily transform from BOOTP to DHCP, you must have the MAC address of the NIM interface configured in the NIM machine definition, as shown in Example 4-19**.** If you do not see a MAC address in the interface line if1, then you must configure it by following the steps described in this section. Otherwise, you may skip to 4.2.4, “Converting BOOTP information to DHCP syntax” on page 62.

Example 4-19 Getting the MAC address

```
# lsnim –l lpar15150
lpar15150:
class = machines
type = standalone
platform = chrp
netboot_kernel = mp
if1 = nim_net lpar15150 0002554F8EA0 ent
cable_type1 = N/A
Cstate = ready for a NIM operation
prev_state = ready for a NIM operation
Mstate = not running
Cstate_result = reset
```

If you do not have the MAC address already configured in the NIM machine definition, it can be added easily. First, gather the MAC address information from the client as shown in Example 4-20.

Example 4-20 Listing the MAC address

```
root@lpar15150 #> netstat -in
Name Mtu Network Address Ipkts Ierrs Opkts Oerrs Coll
en0 1500 link#2 0.2.55.4f.8e.9f 2168 0 432 0 0
en0 1500 9.155.29.12 9.155.29.154 2168 0 432 0 0
en1 1500 link#3 0.2.55.4f.8e.a0 0 0 7 0 0
en1 1500 10.42.7 10.42.7.11 0 0 7 0 0
lo0 16896 link#1 495 0 791 0 0
lo0 16896 127 127.0.0.1 495 0 791 0 0
```

```
lo0 16896 ::1 495 0 791 0 0
root@lpar15150 #>
```

After the MAC address information is collected from each NIM client, it can be added it to the NIM machine definition as shown (where `nim_net` is the NIM network name and `lpar15150` is the NIM machine name):

```
#nim -o change -a "if1=nim_net lpar15150 0002554F8EA0" lpar15150
```

If you are not able to change the NIM machine definition, it may be due to its current NIM state (maybe due to some pending NIM operations). In this case, make sure that all NIM operations concerning this client have been finished and perform a reset operation on the NIM client.

4.2.3 Activating NIM clients to create entries in /etc/bootptab

There is a tool that does the conversion from NIM BOOTP definitions to DHCP syntax, but before that all NIM clients have to be activated in the /etc/bootptab file. Unless you do not have your NIM clients at all times activated for BOOTP in the /etc/bootptab file, you need to perform a NIM operation that does exactly this for you. The easiest way is to set up the client temporarily in diag mode. You will have to do this for each NIM client, so you might prefer the command line as shown (where `aix52002_spot` is the NIM spot name is and `lpar15150` is the NIM machine name):

```
# nim -o diag -a spot=aix52002_spot lpar15150
```

After this has been done for all NIM clients, your /etc/bootptab file should show entries similar to those shown in Example 4-21.

Example 4-21 The /etc/bootptab file

```
# cat /etc/bootptab
...
lpar15150:bf=/tftpboot/lpar15150:ip=10.42.7.11:ht=ethernet:ha=0002554F8EA0:sa=1
0.42.7.9:sm=255.255.255.0:
lpar25150:bf=/tftpboot/lpar25150:ip=10.42.7.12:ht=ethernet:ha=0002554F8E9F:sa=1
0.42.7.9:sm=255.255.255.0:
```

4.2.4 Converting BOOTP information to DHCP syntax

The conversion of the BOOTP syntax to DHCP syntax is easily accomplished through a conversion tool that is part of the bos.net.tcp.server file set. The command is called **`bootptodhcp`** and it is located in the /usr/sbin directory. Example 4-22 on page 63 shows the command options, and a detailed description can be found in the man pages.

Example 4-22 bootptodhcp command

```
bootptodhcp -help
bootptodhcp: Not a recognized flag: h
Usage: bootptodhcp [-d dhcpsd file] [[-b bootptab file] | -r hostname ]
        -b bootptab file, defaults to /etc/bootbtab
        -d dhcpsd configuration file, defaults to /etc/dhcpsd.cnf
        -r remove the specified hostname from the /etc/dhcpsd.cnf
        -r and -b are exclusive.
        The -r flag tells the command to remove the specified hostname
        from the dhcpsd file.  If the -r flag is not included,
        the command converts the bootptab file and appends it to
        the dhcpsd file.
```

For our purposes, we use the default bootptab file, but we redirect the output to a different file than the dhcpsd configuration file because the DHCP is not set up yet. So the command to be executed is:

```
# bootptodhcp –d /tmp/bootp2dhcp.out
```

This generates a file that contains the BOOTP information converted to DHCP syntax so that the DHCP service can serve the future BOOTP requests as shown in Example 4-23.

Example 4-23 The DHCP file

```
#cat /tmp/bootp2dhcp.out
# BOOTP CLIENT: lpar25150
client 1 0002554F8E9F 10.42.7.12
{
   option 1 255.255.255.0
   option sa 10.42.7.9
   option bf "/tftpboot/lpar25150"
}
# BOOTP CLIENT: lpar15150
client 1 0002554F8EA0 10.42.7.11
{
   option 1 255.255.255.0
   option sa 10.42.7.9
   ption bf "/tftpboot/lpar15150"
}
```

You should see all of your NIM clients listed and all entries should look similar, with the MAC address and the client's NIM interface IP address in the first line, and the NIM server IP address and the location of the boot file in brackets.

There is no problem that the bootfile option bf is represented by a symbolic link. DHCP does not check whether this is a valid file. Netboot fails if there is no valid file behind this link or if this link does not exist anyway.

4.2.5 Deallocate NIM clients to remove entries from /etc/bootptab

With the conversion completed, we can deallocate the NIM clients, which removes the entries from /etc/bootptab, and perform a NIM reset operation for all NIM clients to clean up the /etc/bootptab file as shown:

```
# nim -o reset -a force=yes lpar15150
# nim -Fo deallocate -a subclass=all lpar15150
```

4.2.6 Deactivate BOOTP service from inetd

Edit the /etc/inetd.conf file to comment out the bootps service by putting a comment at the beginning of the line, and save the file.

When the line of the BOOTP service bootps is deactivated, you should refresh inetd to make it aware of the changes by running the **`refresh –s inetd`** command. The native BOOTP service is now permanently disabled.

4.2.7 Configure and activate DHCP

DHCP is the method of choice to serve netboot requests for both platforms. It is capable of answering DHCP netboot requests from PXE boot clients as well as BOOTP requests from pSeries systems. The following setup of the DHCP service should not be used as a general DHCP server. Nor is it recommended to put the AIX and Linux net install clients under the control of a general DHCP server that might be already running somewhere in the network.

> **Note:** This DHCP service is exclusively configured to serve as netboot server. All clients have dedicated IP addresses configured to their MAC addresses, so there should be no spare addresses to serve other DHCP clients. This DHCP service is not highly available and thus is not recommended for use for other purposes than its designated function.

The setup of the DHCP server is easy. All keywords and values are explained within the configuration file in /etc/dhcpsd.cnf. By default, all settings are commented out with ‘#’ and must be activated step by step. Example 4-24 on page 65 relates to the logging behavior of dhcpsd, and shows settings that work. For debugging purposes all log items may be activated, but for normal service mode `ACTION` and `INFO` are sufficient.

Example 4-24 Log configuration

```
numLogFiles 4
logFileSize 100
logFileName /tmp/dhcp/dhcpsd.log
#logItem SYSERR
#logItem OBJERR
#logItem PROTERR
#logItem WARNING
#logItem EVENT
logItem ACTION
logItem INFO
#logItem ACNTING
#logItem TRACE
```

Create the directory as shown:

```
# mkdir /tmp/dhcp; chmod 755 /tmp/dhcp
```

As shown Example 4-25, the default lease time is set to infinite because the IP address is dedicated to one MAC address and therefore one client, so it makes no sense to limit the `leaseTimeDefault`. For the same reason, it is not necessary to check for lease expiration, so `leaseExpireInterval` is not activated. The third keyword is the most interesting and it activates the BOOTP service in DHCP. With this option, DHCP can understand and answer BOOTP requests as well.

Example 4-25 DHCP configuration

```
leaseTimeDefault -1
#leaseExpireInterval 3 minutes
supportBOOTP yes
```

DHCP needs at least one network and one subnet defined to work properly. At the end of the configuration file, add a small network as shown in Example 4-26.

Example 4-26 Dummy network

```
#
# Leave some space between the stuff above and your configuration
#
# DHCP needs at least one network defined to start!
# This is a dummy network to satisfy this requisite
network 10.0.0.0 255.255.255.253
# DHCP needs at least one subnet defined to start!
# This is a dummy subnet to satisfy this requisite
subnet 10.0.0.0 10.0.0.1-10.0.0.2
```

To finalize the configuration of the DHCP server, the output from the BOOTP to DHCP conversion must be added. In our example, the conversion was written to the /tmp/bootp2dhcp.out file. The configuration is added as shown:

```
# cat /tmp/bootp2dhcp.out >> /etc/dhcpsd.cnf
```

4.2.8 Booting Linux on pSeries

An advantage of the pSeries systems is that the boot order may be altered either remotely through the service processor or through an appropriate command from the operating system level (for example, the **bootlist** command).

To boot Linux from an AIX server, modify the Linux server /etc/dhcp.cnf configuration file with entries similar to those the AIX clients have, and pass the file as a kernel image. For the next steps, read 4.1.1, “Setting up the network installation server” on page 40.

> **Note:** All of the information in this chapter may not work in all environments, and it is not recommended for production environments.

4.3 CSM performance, scalability, and challenges

In this chapter, we try to address the performance, scalability, and challenges of CSM. These topics are complex in a real and large environment when the managed nodes or management server are either behind a firewall or on a WAN network or both. Due to the residency time limitation, this chapter is just a guideline.

How well a software performs is subject to interpretation. In some environments, the performance is acceptable, and in others the performance is poor or even unacceptable.

4.3.1 Firewall

If the CSM environment is large and in different geographic areas, it is common to have a firewall between the sites.

CSM was ultimately designed to be flexible. CSM can work when behind or split on different sides of a firewall as long as the management server (MS) can reach each node using TCP/IP. This includes using Network Address Translation (NAT) on a firewall. Figure 4-1 on page 67 shows a possible scenario of the firewall that has several LANs. The LAN 192.168.100.0/24 is where the management server and some nodes are. Other nodes are in the LANs

192.168.200.0/24, 192.168.210.0/24 and 192.168.220.0/24. In this scenario, the two systems can use direct routing and NAT.

For direct routing, we need only add routes between the MS and all of the nodes and clear the ports in the firewall. If we want to use NAT in all ways, we make sure that the IP mapping is working and that the name resolution is configured properly. For example, in network 192.168.100.0/24, node_04 must be resolved with an IP from the same class (as shown in Figure 4-1), and in the network 192.168.210.0/24, the management server must be resolved with an IP from the same class, as shown in the examples on page 67.

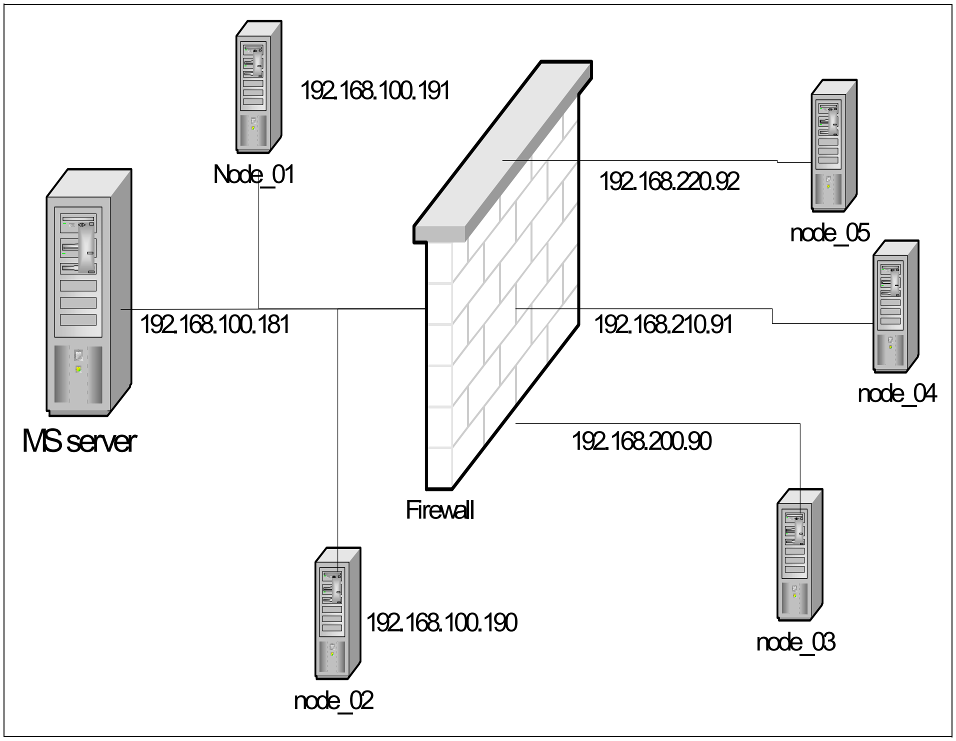


Figure 4-1 Firewall scenario

If we want to use NAT in only one network, we have to adapt the /etc/hosts file on only that network. For the rest of the networks we can use the real names and the real IP address because we can use masquerade or NAT hidden functions, depending on the firewall software.

Example 4-27 Two-way NAT: MS /etc/hosts file

```
#cat /etc/hosts
192.168.100.181    ms_01
```

```
192.168.100.191    node_01
192.168.100.190    node_02
192.168.100.90     node_03
192.168.100.91     node_04
192.168.100.92     node_05
```

Example 4-28 Two-way NAT: node_03 /etc/hosts file

```
#cat /etc/hosts
192.168.200.181    ms_01
192.168.200.191    node_01
192.168.200.190    node_02
192.168.200.90     node_03
192.168.200.91     node_04
192.168.200.92     node_05
```

Example 4-29 Two-way NAT: node_04 /etc/hosts file

```
#cat /etc/hosts
192.168.210.181    ms_01
192.168.210.191    node_01
192.168.210.190    node_02
192.168.210.90     node_03
192.168.210.91     node_04
192.168.210.92     node_05
```

4.3.2 WAN network

When the servers are connected using WAN, sometimes it is necessary to be able to manage all the servers from one point. In this case, CSM can help, but it requires planing before deploying. When CSM is to be installed in an environment similar to Figure 4-2 on page 69, take into consideration:

- All RSCT and CSM communication is based on UDP protocol. In the WAN environment, the IP packets sometimes get lost and the UDP does not wait for acknowledgment. If the UDP packet is lost, it is up to the application to request the re-sending of the UDP packet. The RSCT does not re-send the packets, and the heartbeat packets are UDP packets. If more UDP packets are lost (depending on the settings), the RSCT and CSM may conclude that the node is not responding and issue an alert. So in a real environment, the UDP packets should get a higher priority in the router queue.
- RSCT communication and identification is based on secure keys. These keys must be generated and exchanged. Generating a secure key consumes CPU, and exchanging the keys between the MS and managed node consumes bandwidth. In a large environment when the managed nodes are across WANs, the key exchange could consume considerable bandwidth and

become a bottleneck. This situation usually applies when an MS has to be changed or restored from the backup or when HA MS is used. See question 10 of 7.8, “High Availability Management Server Q&A” on page 148.

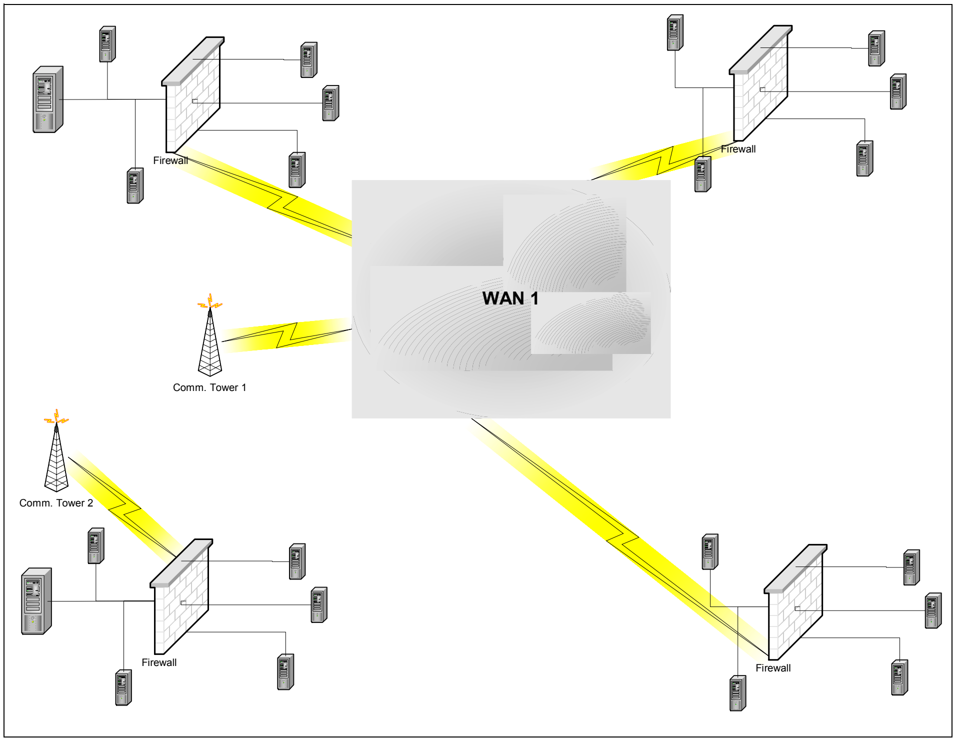


Figure 4-2   Large CSM environment

5

CSM migration scenario

This chapter describes the migration of Cluster Systems Management (CSM) and AIX on the management server and the client nodes for the cluster setup in Chapter 3, “CSM installation scenario” on page 9. We cover the migration of CSM and AIX on the CSM management server and the CSM client nodes, and show each step necessary for a successful migration.

> **Note:** This chapter is no replacement for the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919. Always use the CSM product documentation to plan and perform a migration.

In this chapter, the following topics are discussed:

- 5.1, “General planning for migrations” on page 72.
- 5.2, “Migrate CSM on the management server” on page 78.
- 5.3, “Migrate CSM on AIX client nodes” on page 83.
- 5.4, “Migrate AIX on the client nodes” on page 90.
- 5.5, “Migrate AIX on the management server” on page 96.

5.1 General planning for migrations

Planning is the most important step of any action performed on a computer system. For a cluster system, with different operating systems, different versions of an operating system, consisting of different hardware platforms, planning is even more important. There are applications running on the cluster nodes, and a migration of the operating system to a newer release would require a migration of these applications as well. However, in many cases planning is not done well. In too many cases an “on-the-fly” planning is performed, which usually causes problems and additional downtime. We will show the general planning activities in this section, which are:

- Where we want to go
  In this step, the level of software to migrate to is defined.
- Where we come from
  Determine the current level of software running on the systems.
- Gather prerequisites information
  Determine which prerequisites must be solved to perform the migration.
- Checking for latest updates available
  We recommend installing the latest updates to the migrated software.
- Planning for recovery
  Determine which options are available for recovery of a failed migration.

5.1.1 Where we want to go

This first planning step should be easy:

- For application testing, a node at a new AIX release is needed.
- For high availability, CSM 1.4 on the management server is required.
- New hardware to be added to the cluster requires the latest AIX release.
- Software support for the currently used release ends.

For our example we have two requests:

- A High Availability Management Server (HA MS) should be installed.
- One AIX client node must be migrated to AIX 5.3. This node must be ready in four weeks.

Based on the two requests, we can now define our destination:

- CSM 1.4 is the first release of CSM that supports HA MS. This must be installed on the management server.
- To migrate the AIX client node to AIX 5.3, the Network Installation Manager (NIM) master, which is our management server, must be at AIX 5.3.

- One AIX client node must be migrated to AIX 5.3.

AIX 5.3 is not needed now, so we can do the necessary migration on the management server in two steps. We will first migrate CSM to 1.4 and later migrate AIX to 5.3. The AIX migration on the client node will be done after the management server gets migrated to AIX 5.3.

5.1.2 Where we come from

Unlike a single system, in a cluster we need to be aware which levels of software are running on all cluster systems, because migrating the controlling systems may require updates on the client nodes. This includes the Hardware Management Console (HMC).

In a well-managed cluster, this information is known and documented. However, we never should trust documentations created in the past, so we must verify the software levels. The **dsh** and **dshbak** commands in conjunction with the CSM node groups can be used to gather all necessary information as shown in Example 5-1.

Example 5-1 Use of dsh and dshbak to gather the oslevel from AIX client nodes

```
root@ms_01:/ # dsh -v -N AIXNodes oslevel -r|dshbak -c
HOSTS --------------------------------------------------------------------------
node_01, node_02
--------------------------------------------------------------------------------
5100-03

HOSTS --------------------------------------------------------------------------
node_03, node_04, node_06
--------------------------------------------------------------------------------
5200-03
```

For clients running AIX, information should be gathered by using the following commands:

- `oslevel -r`
- `lslpp -Lc bos.mp bos.mp64 bos.up`
- `bootinfo -K`
- `lslpp -Lc csm\*`
- `lslpp -Lc rsct\*`

Determine the applications that are running on the nodes and the software level of these applications. The commands to use depend on the application.

For clients running Linux, information should be gathered by using the following commands:

- `uname -a`
- `cat /proc/version`
- `rpm -qa |grep csm`
- `rpm -qa |grep rsct`
- `rpm -qa |grep src`

Determine the applications that are running on the nodes and the software level of these applications. The commands to use depend on the application.

On the management server, gather the same data as on the AIX nodes, plus:

- `rpm -qa`
- `lslpp -Lc sam.core`

For the Hardware Management Console (HMC), use the Web-based System Manager to get the current code version. From an AIX system running AIXwindows, use the **wsm** command. For PCs running Microsoft® Windows NT®/2000/XP or Linux, the WebSM remote client software is available on the HMC. To download it, use a browser to go to:

```
http://your_HMC/remote_client.html
```

Execute this image to install the WebSM remote client, then start it. You will be prompted for the host name, user name, and password. Enter this information for your HMC to log on. Open **Management Environment** → **HMC host name** → **Software Maintenance** → **HMC**. The level of HMC software is shown in the status section.

5.1.3 Gather prerequisites information

Always use the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919, to determine which prerequisites are needed. In general, the following information should be checked for a migration of CSM on the management server:

- Required AIX level
- Required Reliable Scalable Cluster Technology (RSCT) level
- Required open source software
- Required software level for the Hardware Management Console (HMC)
- Requirements for client nodes

We suggest that you create a list of prerequisites for the migration scenario. For our example scenario used in this chapter, we start with the migration to CSM 1.4, which has following prerequisites:

- AIX 5.2 ML04 with RSCT 2.3.4 or later or AIX 5.3 with RSCT 2.4.0 or later.
- The following open source software is required:
  - conserver-7.2.4-1
  - tcl-8.3.3-1
  - tk-8.3.3-1
  - expect-5.32-1
  - openCIMOM-0.8-1 (required for hardware control only)
  - sg3_utils-1.06-2.ppc (required for managing Red Hat EL 3 AS QU1 on POWER-based architecture nodes only). Download this package from:

    `http://people.redhat.com/pknirsch`
  - autoupdate-4.8.4-1.noarch.rpm or higher (required for managing Linux nodes only). Download this package from:

    `http://freshmeat.net/projects/autoupdate`
- HMC for pSeries Release 3 Version 1.0 or later is required. To download updates if needed, go to the HMC support for pSeries and iSeries™ site at:

  `https://techsupport.services.ibm.com/server/hmc`
- The requirements for client nodes running AIX are:
  - AIX 5.1 ML03 with CSM 1.1.0 or later
  - AIX 5.2 with CSM 1.3.1 or later
  - AIX 5.3 with CSM 1.4 (only possible if the management server is at AIX 5.3)
- Client nodes running Linux should be at CSM 1.3.1 or later.
- CSM 1.4 client nodes running Linux must use one of these distributions:
  - Red Hat Linux AS 2.1
  - Red Hat EL 3 (AS/ES/WS)
  - SUSE LINUX Enterprise Server 8 (8.1)

5.1.4 Checking for latest updates available

Migrating a program product to a new release uses the installation images that are found on the product media. The update level for these images in many cases is 0, which means no updates are included. For example: Migration CSM to 1.4 using the CSM install images shipped with AIX migrates the CSM file sets to:

- csm.core 1.4.0.0
- csm.client 1.4.0.0
- csm.diagnostics 1.4.0.0
- csm.dsh 1.4.0.0
- csm.msg 1.4.0.0
- csm.server 1.4.0.0
- csm.gui.dcem 1.4.0.0
- csm.gui.websm 1.4.0.0

We recommend checking for the latest available updates and installing them. The same applies to AIX. All updates are downloadable from the Internet. Visit IBM support Fix Central, which is the central entry point to get all available software updates, at:

`http://www-912.ibm.com/eserver/support/fixes/fcgui.jsp`

However, the direct link can be used to get to the corrective service for CSM:

`http://techsupport.services.ibm.com/server/cluster/fixes/csmfixhome.html`

All necessary updates should be downloaded to the management server before the migration is performed.

5.1.5 Planning for recovery

Reliable backups of the operating systems, system configuration data, and applications including the application data must be available for recovery in case of disk problems or other disasters. These backups can also be used to recover a system after a failed migration. Depending on the software to be migrated and the way this is done, there may be other and faster ways to recover.

Updating software on an AIX system

Commit previous applied updates using the **`smit commit`** command. For the update operation, use the **`smit update_all`** command and set these options:

▸ COMMIT software updates?	`no`
▸ SAVE replaced files?	`yes`

This way an update can be rejected easily using the AIX function Reject Applied Software Updates. This reject is faster to restore a system backup and can help to reduce downtime, but it does not replace the need for a reliable backup.

A second option is the alternate disk rootvg cloning procedure. This creates a copy of the rootvg on a second disk or set of disks. The update can be performed on one copy. In case of the need to fall back to the previous software level, the second copy of the rootvg can be booted. Refer to the *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, or the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887, for more details.

Updating CSM on the management server running AIX

In addition to the actions outlined in “Updating software on an AIX system” on page 76, the CSM configuration data should be saved using the **csmbackup** command. Refer to the *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.4*, SA22-7934 for more information on the **csmbackup** command. We recommend to save a copy of this backup on external media.

Migrating software on an AIX system

A migration is an overwrite of installed software by a new version or release. Configuration data is preserved or converted to a format necessary for the new version or release by the migration. There is no reject option provided. Restoring the system backup is one way to fall back to the previous AIX version or release. However, there are options which can be used to prevent a system backup restore. Please refer to the chapters titled Alternate Disk rootvg Cloning and Alternate Disk Migration Installation in the *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389 or the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887.

Migrating CSM on the management server running AIX

Save the CSM configuration data using the **csmbackup** command. Refer to the *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.4*, SA22-7934 for more information. To fall back from a failed migration reinstall the previous used CSM level, then restore the backup of the CSM configuration data using the **csmrestore** command.

As a second option, the alternate disk rootvg cloning procedure can be used to create a second copy on the rootvg. On one copy the CSM migration is performed. In case of the need to fall back to the previous CSM level, the second copy of the rootvg can be booted. Refer to the *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389 or the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887 for more details.

Migrating CSM on AIX client nodes

To fall back to a previous CSM release on client nodes running AIX, perform following steps:

1. Deinstall CSM on the client nodes using the **installp -u csm*** command.

2. Install the previous CSM release on the client nodes.
   This can be done using the NIM cust operation or by NFS mounting the lpp_source containing the desired CSM release and install it from there.
3. Use the **chnode** command on the CSM management server to set the InstallCSMVersion to the installed release.
4. On the CSM management server, run the **updatenode** command for these nodes to ensure that they are configured properly.

Migration CSM on Linux client nodes

To fall back to a previous CSM release on client nodes running Linux, perform the following steps:

1. Deinstall CSM on the client nodes using the **rpm -e package_name** command. CSM for Linux ships the necessary RSCT packages and it may be necessary to uninstall these RSCT packages as well.

> **Note:** Other software may also use RSCT, and a deinstallation of the RSCT packages may destroy the RSCT configuration for these software products. Recovery procedures for the other software products must be in place prior the deinstall of RSCT.

- Use the **chnode** command on the CSM management server to set the InstallCSMVersion to the release to fall back to.
- On the CSM management server, run the **updatenode** command for these nodes to install the CSM release set in InstallCSMVersion.

5.2 Migrate CSM on the management server

In this section, we show the migration to CSM 1.4 on the CSM management server. The migration consists of following steps:

1. Get the necessary software.
2. Check for prerequisites on client nodes and HMCs.
3. Back up the management server.
4. Commit previous installed updates.
5. Update and migrate the software.
6. Update the open source software packages.
7. Reboot the management server.
8. Accept the terms and conditions of the CSM license agreement.
9. Copy updated CSM files into /csminstall.
10. Verify the installation.
11. Perform a backup of the migrated system.

Note: This chapter is no replacement for the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919. Always use the CSM product documentation to plan and perform a migration.

Our migration steps are not in the same order as in the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919. We copy the downloaded software to the management server first and back up right before the migration is performed. In case of a fallback using the backup, we want to have the downloaded software back. In addition, we perform a post-migration full backup of the management server. Imagine a hardware problem a few hours after migration with no such backup available. Restoring the pre-migration backup means that the whole migration would have to be performed again.

5.2.1 Get the necessary software

We use the data gathered in 5.1.2, “Where we come from” on page 73 and 5.1.3, “Gather prerequisites information” on page 74 to determine what software must be migrated or updated for this migration to be successful. Table 5-1 shows the result.

Table 5-1 Software levels for CSM migration on the management server

File set	Installed level	Required level	Available level
CSM	1.3.3.x	1.4.0.0	1.4.0.3
bos.mp64	5.2.0.30	5.2.0.40	5.2.0.43
rsct.core.auditrm	2.3.3.0	2.3.4.0	2.3.4.0
rstc.core.errm	2.3.3.2	2.3.4.0	2.3.4.1
rsct.core.fsrm	2.3.3.1	2.3.4.0	2.3.4.0
rsct.core.gui	2.3.3.0	2.3.4.0	2.3.4.1
rsct.core.hostrm	2.3.3.0	2.3.4.0	2.3.4.1
rsct.core.rmc	2.3.3.3	2.3.4.0	2.3.4.2
rsct.core.sec	2.3.3.1	2.3.4.0	2.3.4.1
rsct.core.sensorrm	2.3.3.0	2.3.4.0	2.3.4.1
rsct.core.sr	2.3.3.2	2.3.4.0	2.3.4.1

File set	Installed level	Required level	Available level
rsct.core.utils	2.3.3.4	2.3.4.0	2.3.4.2
openCIMOM	0.7-4	0.8-1	0.8-1
expect	5.32-1	5.32-1	5.34
tcl	8.3.3-1	8.3.3-1	8.3.3
tk	8.3.3-1	8.3.3-1	8.3.3
conserver	7.2.4-1	7.2.4-1	7.2.4

We decide to get the latest updates for CSM, bos.mp64, and all RSCT file sets. We create a new lpp_source using the AIX 5.2 ML04 updates and place the latest available updates into it. We stay at the required level of expect.

> **Note:** Be sure to rebuild the .toc file using the `inutoc` command after you put the updates into the lppsource directory.

5.2.2 Check for prerequisites on client nodes and HMCs

Next, we check whether updates are required on the client nodes. Nodes running AIX 5.1 must be at CSM 1.1.0 or later. Nodes running AIX 5.2 or Linux must be at CSM 1.3.1 or later. All of our client nodes meet these requirements, so updating them is not necessary.

The HMC must be at HMC for pSeries Release 3 Version 1.0 or later. Our HMC is at Release 3 Version 2.6. No update of HMC code is necessary to migrate CSM to 1.4 on the management server.

5.2.3 Back up the management server

Back up the CSM configuration data using the **`csmbackup`** command. Then create a full-system backup using the **`smit mksysb`** command followed by a backup of the data stored in none rootvg volume groups using the **`smit savevg`** command.

5.2.4 Commit previous installed updates

We recommend committing all previously installed updates prior to installing the AIX updates required to migrate to CSM 1.4. This way a reject of the updates to fall back to the current update level can be performed by rejecting all applied and not committed updates. Use **`installp -c all`** to commit all applied updates.

5.2.5 Update and migrate the software

In this step, we update AIX and RSCT and migrate CSM to 1.4. We suggest using SMIT to perform this step as it saves a log in the smit.log and smit.script files. In case of errors, this log file can be reviewed to find the cause of the problem. Removing the old smit.log and smit.script files before the update is started makes it easier to find the right entries. Use the command sequence **`cd;rm smit.log smit.script`** to remove the smit log files. To start the update and migrate CSM in one step, use the **`smit update_all`** command, then enter the name of the directory where the updates and CSM 1.4 install images are located. In the menu, ensure that the following options are set as shown in Example 5-2:

- COMMIT software updates?
- SAVE replaced files?
- ACCEPT new license agreements?

Example 5-2 smit update_all options to use for CSM migration

```
Update Installed Software to Latest Level (Update All)

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

                                                        [Entry Fields]
* INPUT device / directory for software               /export/eznim/lpp_sou>
* SOFTWARE to update                                  _update_all
  PREVIEW only? (update operation will NOT occur)     no                     +
  COMMIT software updates?                            no                     +
  SAVE replaced files?                                yes                    +
  AUTOMATICALLY install requisite software?           yes                    +
  EXTEND file systems if space needed?                yes                    +
  VERIFY install and check file sizes?                no                     +
  DETAILED output?                                    no                     +
  Process multiple volumes?                           yes                    +
  ACCEPT new license agreements?                      yes                    +
  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
```

Inspect the smit.log file for error messages after the update operation completes.

5.2.6 Update the open source software packages

The only open source software we need to update is openCIMOM. This can be done by executing the command:

```
rpm -Uvh openCIMOM-0.8-1.aix5.2.noarch.rpm.
```

5.2.7 Reboot the management server

Because of the update to AIX, the management server has to be rebooted to activate the new AIX kernel. To reboot the management server, use the command:

```
shutdown -Fr.
```

5.2.8 Accept the terms and conditions of the CSM license agreement

The terms and conditions of the CSM license agreement must be accepted after the migration to CSM 1.4. This is done by running the **`csmconfig -L`** command, which displays the license agreement first. The license agreement then can be accepted by typing the digit `1`.

> **Note:** A new license key file does not have to be provided with the -L flag when you migrate to CSM 1.4.0.

5.2.9 Copy updated CSM files into /csminstall

Execute the command **`csmconfig -c`** to copy updated CSM files into the /csminstall directory. These files are used when CSM management scripts are run on the client nodes.

5.2.10 Verify the installation

A verification of the AIX update and CSM migration should be performed after the reboot. In addition, the CSM functionality should be tested. We suggest running at least the following commands:

- **`lppchk -v`**
  This command confirms that the software installation is consistent and checks for missing prerequisites. A successful execution returns no message. Any message indicates an error or warning and needs further review.
- **`oslevel -r`**
  This command should return `5200-04`, indicating that we updated AIX to AIX 5.2 ML04.

- **`probemgr -p ibm.csm.ms -l 0`**
  This command checks the configuration of the management server. It should return this message: `Probe ibm.csm.ms was run successfully`.
- Execute the commands **`csmstat`**, **`rpower -a query`**, and **`rpower -la query`** and verify that they return the correct output.

5.2.11 Perform a backup of the migrated system

We suggest taking a full backup of the system at this point. Refer to 5.2.3, “Back up the management server” on page 80. This new backup is no replacement for the backup taken before the migration. Both backups should be kept for some time, so you can fall back to the pre-migration state as well as the post-migration state if necessary by restoring a backup.

5.3 Migrate CSM on AIX client nodes

This chapter describes the migration to CSM 1.4 on AIX nodes. This migration requires RSCT to be updates to 2.3.4.x and AIX to be updated to 5.2 ML04. We use the lppsource that was set up already in 5.2.1, “Get the necessary software” on page 79 to perform the necessary updates as well as the migration of CSM. The following steps must be performed:

1. Check for application software prerequisites.
2. Back up the client nodes.
3. Commit previously installed updates.
4. Stop the applications on the client nodes.
5. Update and migrate the nodes.
6. Reboot the client nodes.
7. Updatenode (when necessary).
8. Verify CSM functions.
9. Start up and verify the applications on the client nodes.
10. Back up client nodes.

> **Note:** All software maintenance on CSM client nodes can be performed from the CSM management server if NIM is configured and the client nodes are defined as NIM clients.

> **Important:** Prior to migrating all client nodes belonging to a node group, the migration should be tested using a test node or group of test nodes. We suggest also testing the fallback. If no test nodes are available, we suggest migrating one production node first to ensure that all migration steps are running without problems before the whole node group is migrated.

5.3.1 Check for application software prerequisites

We need to know what applications are running on the CSM client nodes. The update to AIX 5.2 ML04 that we want to perform also may require updates to the application software. In our example, the nodes to be migrated are running IBM LoadLeveler® for AIX 5L Version 3 Release 2. The LoadLeveler documentation lists the following prerequisites:

- AIX 5L 5.2 or later compatible releases
- RSCT 2.3.1.0 or a later compatible release for configuring LoadLeveler to support RSCT security services

Updating AIX to 5.2 ML04 and RSCT 2.3.4 should not cause problems with LoadLeveler. The applications run by the users using LoadLeveler are already tested on AIX 5.2 ML04 and AIX 5.3.

5.3.2 Back up the client nodes

Ensure that reliable backups for the client nodes are available. However, alternate disk rootvg cloning as discussed in 5.1.5, “Planning for recovery” on page 76 may be an additional option for fast recovery. This is only an AIX and RSCT update. In case of problems with the new update level, we can reject these updates. Only CSM, then migrated to 1.4, cannot be rejected and must be reinstalled as described in “Migrating CSM on AIX client nodes” on page 77.

5.3.3 Commit previously installed updates

We recommend committing all previously installed updates prior to installing the AIX updates required to migrate to CSM 1.4. This way a reject of the updates to fall back to the current update level can be performed by rejecting all applied and not committed updates. Use the command **`dsh -n nodename installp -c all`** to commit all applied updates on one node. Use the command **`dsh -vN nodegrp installp -c all`** to commit applied software on all nodes belonging to the node group nodegrp.

5.3.4 Stop the applications on the client nodes

Prior to actually running the migration, the applications running on the nodes should be stopped. You need to determine the right way to stop these applications. After the commands to stop the applications run, ensure that the application stopped to avoid application problems.

5.3.5 Update and migrate the nodes

For our example, we use NIM to perform the update and migration to show how easy such a task can be in a CSM cluster. We already have a static CSM node group defined containing all our nodes running LoadLeveler. The name of this node group is LLNodes. This CSM node group got defined to NIM using the **`csm2nimgrps -N LLNodes`** command. To set up the update and migration using NIM, execute the **`smitty nim_task_inst`** command on the CSM management server. This displays the menu shown in Example 5-3.

Example 5-3 NIM Install and Update Software menu

```
Install and Update Software

Move cursor to desired item and press Enter.

  Install the Base Operating System on Standalone Clients
  Install Software
  Update Installed Software to Latest Level (Update All)
  Install Software Bundle
  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
```

Select the function **Update Installed Software to Latest Level (Update All)**, and a target selection window (Example 5-4) is displayed.

Example 5-4 NIM target select window

```
+--------------------------------------------------------------------------+
  ¦                    Select a TARGET for the operation                     ¦
  ¦                                                                          ¦
  ¦ Move cursor to desired item and press Enter.                             ¦
  ¦                                                                          ¦
  ¦   AIXNodes        groups          mac_group                              ¦
  ¦   LLNodes         groups          mac_group                              ¦
  ¦   520spot_res     resources       spot                                   ¦
  ¦   510spot_res     resources       spot                                   ¦
  ¦   node_06         machines        standalone                             ¦
  ¦   node_03         machines        standalone                             ¦
  ¦   node_01         machines        standalone                             ¦
  ¦   node_02         machines        standalone                             ¦
  ¦   node_04         machines        standalone                             ¦
  ¦   ms_02           machines        standalone                             ¦
  ¦   master          machines        master                                 ¦
  ¦                                                                          ¦
  ¦ F1=Help                 F2=Refresh              F3=Cancel                ¦
  ¦ F8=Image                F10=Exit                Enter=Do                 ¦
F1¦ /=Find                  n=Find Next                                      ¦
F9+--------------------------------------------------------------------------+
```

Move the cursor to the LLNodes node group and press Enter to select this node group as the target for the update_all operation. This displays a selection window to select the lpp_source used for the update_all operation (Example 5-5).

Example 5-5 NIM lppsource selection window

```
+--------------------------------------------------------------------------+
  ¦           Select the LPP_SOURCE containing the install images            ¦
  ¦                                                                          ¦
  ¦ Move cursor to desired item and press Enter.                             ¦
  ¦                                                                          ¦
  ¦   aix52_ml03     resources       lpp_source                              ¦
  ¦   aix520         resources       lpp_source                              ¦
  ¦   aix510_03      resources       lpp_source                              ¦
  ¦   csm1334        resources       lpp_source                              ¦
  ¦   AIX52_04       resources       lpp_source                              ¦
  ¦                                                                          ¦
  ¦ F1=Help                 F2=Refresh              F3=Cancel                ¦
  ¦ F8=Image                F10=Exit                Enter=Do                 ¦
F1¦ /=Find                  n=Find Next                                      ¦
F9+--------------------------------------------------------------------------+
```

The lpp_source created in 5.2, “Migrate CSM on the management server” on page 78 is named AIX52_04. We select this lpp_source. Now the update options menu (Example 5-6) is displayed. It looks very much like the SMIT menu shown for the update_all operation in 5.2.5, “Update and migrate the software” on page 81 that is used to update AIX and migrate CSM on the management server. Ensure that the `installp Flags` section is set as shown in Example 5-6.

Example 5-6 NIM Update All options menu

```
Update Installed Software to Latest Level (Update All)

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

                                                        [Entry Fields]
* Installation Target                                 LLNodes
* LPP_SOURCE                                          AIX52_04
  Software to Install                                 update_all

  Customization SCRIPT to run after installation     []                      +
    (not applicable to SPOTs)
  Force                                                  no                    +

  installp Flags
    PREVIEW only?                                       [no]                   +
    COMMIT software updates?                            [no]                   +
    SAVE replaced files?                                [yes]                  +
    AUTOMATICALLY install requisite software?           [yes]                  +
    EXTEND filesystems if space needed?                 [yes]                  +
    OVERWRITE same or newer versions?                   [no]                   +
    VERIFY install and check file sizes?                [no]                   +
    ACCEPT new license agreements?                      [yes]                  +
    Preview new LICENSE agreements?                     [no]                   +

  Group controls (only valid for group targets):
    Number of concurrent operations                     []                      #
    Time limit (hours)                                  []                      #

  Schedule a Job                                        [no]                   +
  YEAR                                                  []                      #
  MONTH                                                 []                     +#
  DAY (1-31)                                            []                     +#
  HOUR (0-23)                                           []                     +#
  MINUTES (0-59)                                        []                     +#


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

Note the Group controls section, which can be used to limit the number of concurrent operations and to set a time limit. This is useful if a big number of client nodes belong to the selected node group to reduce the peak load on the NIM server node (the CSM management server in our case) and the network. For more about NIM, see the *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, or the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887. Use the **`lsnim`** command to monitor the progress of the operation.

When the update and migration completes, the Cstate of the nodes changes to `ready for a NIM operation`. Use the command **`nim -o showlog -alog_type=niminst nodename | more`** to review the log for one node.

If the NIM operation was performed on many nodes, it would be useful to create a shell script as shown in Example 5-7 on page 88 to perform the logs review.

Example 5-7 Run nim showlog for all nodes belonging to a CSM node group

```
#!/usr/bin/ksh
Nodes=`nodegrp -p LLNodes`
for Node in $Nodes
do
  echo "niminst log for node $Node"
  echo "======================================"
  nim -o showlog -alog_type=niminst $Node
  echo;echo;echo;echo
done
```

5.3.6 Reboot the client nodes

We need to reboot the updated nodes because a new kernel and new device driver code was installed by the update. The NIM log for each updated node in this case is shown in Example 5-8.

Example 5-8 Message in NIM log indicating that a reboot is required

```
installp:  * * *  A T T E N T I O N ! ! !
        Software changes processed during this session require this system
        and any of its diskless/dataless clients to be rebooted in order
        for the changes to be made effective.
```

Use the command **`dsh -N NodeGrp "nohup shutdown -Fr&"`** on the CSM management server to shut down and reboot all client nodes belonging to the node group NodeGrp.

> **Note:** The use of **`nohup`** prevents the shutdown process from being killed during the shutdown process. This would leave the system in a state where it can only be accessed using the system console or **`rconsole`** command.

5.3.7 Updatenode (when necessary)

The **`updatenode`** command distributes configuration files, runs user-provided customization scripts, exchanges RSCT public keys, and performs other update functions. We suggest running the **`updatenode`** command for the migrated and updated client nodes using the **`updatenode -N NodeGrp`** command.

5.3.8 Verify CSM functions

We suggest testing the CSM functions after rebooting the updated and migrated nodes. This includes **dsh**, **rpower**, **rconsole**, and the CSM monitoring functions. The following commands should be executed on the CSM management server:

- **dsh -N NodeGrp date**
  A basic test that dsh to all nodes in the node group is working.
- **dsh -N NodeGrp uptime**
  Check the uptime value for each node to ensure all nodes rebooted.
- **rpower -N NodeGrp query**
  Query the power status for each node in the node group.
- **rpower -lN NodeGrp query**
  Query the power status including the LCD value for each node in the node group.
- **csmstat**
  Ensure that Status, PowerStatus, and Network-Interfaces are shown correctly for all nodes.

Start WebSM and connect to the CSM management server. Open the **CSM Cluster view** → **Distributed Monitoring** → **Overview and Tasks**. Select **Monitor conditions** to see which conditions are currently enabled for visual monitoring. Back in the CSM Cluster view, select **Groups**. Right-click on the **NodeGrp** node group, and in the pop-up select **Monitored Resources**. Trigger a condition on at least one node beloging to the node group to see that the monitoring function is working. In our example, the free space in the /tmp file system is monitored. The command **dsh -n one_node dd if=/dev/zero of=/tmp/t** will fill the /tmp file system. This should trigger the monitoring. It may take some time until the event occurs, depending on the monitor interval used for this condition. Remove the test file /tmp/t after successful testing using the **dsh -n one_node rm /tmp/t** command and the event should be cleared soon after.

5.3.9 Start up and verify the applications on the client nodes

Now the applications on the updated and migrated nodes can be started. At least some basic tests should be performed on the applications to ensure that there is no major problem with these applications.

5.3.10 Back up client nodes

We suggest creating new backups for the updated and migrated client nodes. These backups should not overwrite the backups taken in 5.3.2, “Back up the client nodes” on page 84 as it may be necessary to back out the updates and migrated code by restoring these backups.

5.4 Migrate AIX on the client nodes

In this section, we show examples of AIX migrations on client nodes. We migrate a client node running AIX 5.1 ML03 to AIX 5.2 ML04. The CSM client software will be migrated automatically in this step, because CSM is shipped with AIX and the necessary CSM install images are in the used lpp_source. The same applies to RSCT. The lpp_source we use is the one we set up in 5.2.1, “Get the necessary software” on page 79 for the AIX update and CSM migration of the management server.

> **Attention:** Care must be taken for the applications that are running on the nodes to ensure that these applications are working on the new AIX release.

It may be necessary to run the **`updatenode -kn Node_Name`** command after a migration. We recommend taking a new backup after a successful migration. AIX and application functions should be tested prior to putting the node back into production. After migrating client nodes, use the **`lsnode -l`** command to verify the InstallCSMVersion and InstallDistributionVersion node attributes. Use the **`chnode -n`** command to change either of them to the correct CSM version and AIX version installed on the node.

5.4.1 Alternate disk migration installation

In this section, we show an example of an alternate disk migration. This migration type enables a migration of AIX on a running system to a second, previous unused, set of disks. The migration itself requires no downtime. The result of this migration type is a second rootvg, which holds the migrated AIX. A change of the bootlist and a reboot of the system activates the new AIX release. Fallback is equally simple. Change the bootlist back to the old rootvg disks and reboot.

> **Note:** The first step of the migration is the creation of a copy of the active rootvg to other disks. The migration is done on this copy. Changes to files located in rootvg after this copy will be lost in the migrated system.

Check the alternate disk migration requirements

There are some requirements and limitations for an alternate disk migration. See *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, for all requirements and limitations. The requirements are:

- The NIM master must run at AIX 5.1 ML03 or later.
  The CSM management server is the NIM master and must be at AIX 5.2 or later.

- The file set bos.alt_disk_install.rte must be installed on the NIM master.
  This can be checked using the **lslpp -Lc bos.alt_disk_install.rte** command.
- The file set bos.alt_disk_install.rte must be installed on the SPOT that is used for the migration.
  The **nim -o lslpp -a lslpp_flags="-Lc bos.alt_disk_install.rte" SPOT_name** command can be used to verify this requirement.
- The AIX level installed on the NIM master, the lpp_source, and SPOT must be the same.
  There is no **oslevel**-like command for an lpp_source or SPOT. However, we can use the **lslpp** command to get a list of all installed software on the NIM master and the **nim -o lslpp** command to get a list of all installed software in the SPOT. The level of the lpp_source can be checked using the **ls** command. In our case, the NIM master got updated to AIX 5.2 ML04 using an lpp_source we created. The SPOT created from this lpp_source is at the same level as the lpp_source used to create it. So we meet this requirement.
- The client node that is to be migrated must be at AIX 4.3.3 or later.
  CSM client nodes must be at AIX 5.1 ML03 or later.
- The client node must have free disks large enough to clone the rootvg and an additional 500 MB to perform the migration.
  To list all disks on a client node, use the command **dsh -n node_name lspv**. This command shows the name of the volume group that a disk is assigned to or no name for a free disk. The **dsh -n node_name lsvg rootvg** command shows how much disk space is currently used for rootvg. Use the **dsh -n node_name lsattr -El disk_name** command to get the size of the free disks.
- The client must be a registered NIM client to the master.
  All of our CSM client nodes running AIX are registered NIM clients.
- The NIM master must be able to execute remote commands using **rsh** on the client.
  This can be tested using the **rsh node_name date** command. If this is not working, then check that a valid .rhosts file exists on the client node and rsh is enabled in /etc/inetd.conf.
- The client node must have at least 128 MB of memory installed.
  The command **dsh -n node_name bootinfo -r** displays the available memory in KB.
- A network between the NIM master and client, which can facilitate large amounts of NFS traffic, must be available.
  We will use the management LAN used by CSM.
- The client hardware must support the AIX level we want to migrate to.

The limitations are:

- The client must not have Trusted Computing Base (TCB) enabled.
  To verify the TCB setting, run the **tcbck** command on the node.
- All NIM resources used for the alternate disk migration must be local to the NIM master.
  Use the command **lsnim -l <NIMresource>** to check the server name for a NIMresource.
- The client node may experience a performance decrease during the migration because of additional disk I/O, NFS activity, and CPU load.
- NFS tuning may be required to optimize migration performance.

Back up the client nodes

Repeat and repeat again: "Ensure that reliable backups for the client nodes are available. This includes a backup of the application data and none rootvg volume groups."

Performing the alternate disk migration

Use the **nimadm** command to start the migration. For our example, we migrate the client node node_01. The rootvg is on hdisk0. A free disk, big enough to hold a copy of rootvg and the additional required space for the migration, is available. The name of this disk is hdisk1. The command **nimadm -c node_01 -s spot_res -l lpp_source -d hdisk1 -Y -B** starts the AIX migration on node_01. Licenses are accepted (-Y flag) and the bootlist is not set to the migrated AIX (-B flag). A full log of the migration is stored in a file in the directory /var/adm/ras/alt_mig on the NIM master. This log file should be reviewed for possible errors after the migration process completes. To boot the migrated AIX on the client node, the bootlist must be set, followed by a reboot as shown in Example 5-9.

Example 5-9 Activate migrated rootvg after alternate disk migration

```
root@node_01:/ # lspv
hdisk0         000197ca5abe26ae                    rootvg
hdisk1         000197ba699d5dbb                    altinst_rootvg
root@node_01:/ # bootlist -m normal hdisk1 hdisk0
root@node_01:/ # bootlist -m normal -o
hdisk1
hdisk0
root@node_01:/ # shutdown -Fr
```

> **Note:** We suggest setting the bootlist to contain both disks. In case the boot from hdisk1 fails or there are other major problems with the AIX on it, simply remove the disk from the system and it will boot the old release of AIX from hdisk0.

Run **updatenode -kn node_01** after booting the new AIX release. On the client node, the naming of the volume groups changed as shown in Example 5-10.

Example 5-10 Volume groups after booting the migrated AIX

```
root@node_01:/ # lspv
hdisk0          000197ca5abe26ae                    old_rootvg
hdisk1          000197ba699d5dbb                    rootvg          active
```

Example 5-11 shows the fall back to the old AIX release. This is done by changing the bootlist and rebooting.

Example 5-11 Fall back to old AIX release

```
root@node_01:/ # bootlist -m normal hdisk0 hdisk1
root@node_01:/ # bootlist -m normal -o
hdisk0
hdisk1
root@node_01:/ # shutdown -Fr
rebooting ...
root@node_01:/ # lspv
hdisk0          000197ca5abe26ae                    rootvg
hdisk1          000197ba699d5dbb                    altinst_rootvg
```

This way either one of the two AIX releases can be activated by setting the bootlist accordingly and rebooting. The time to activate the new AIX release or to fall back to the old release is reduced to the time a reboot takes.

5.4.2 Traditional AIX migration using NIM

The traditional AIX migration of AIX requires more downtime than the alternate rootvg migration installation described in 5.4.1, “Alternate disk migration installation” on page 90. However, it has fewer restrictions and limitations than the alternate rootvg migration installation. As long as the client node hardware is supported by the new AIX release, the traditional AIX migration using NIM can be used for migration. Refer to *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, or *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887, for more information about the AIX installation and migration procedures.

The traditional AIX migration is like a new installation. The system to be migrated boots off a boot image at the release level the system will migrate to. The rootvg is activated and the current configuration will be saved. This includes a list of all installed software packages. The files belonging to the old AIX release are removed and the new AIX release is installed, including all of the software packages installed on the previous release. New software packages may be installed, and software packages that are no longer supported or needed are uninstalled. Finally the saved configuration is restored into the new AIX release. This is a very high-level view of migration that shows that the old release gets removed, the new release gets installed, and the system configuration is preserved. The only way to revert such a migration is the restoration of a backup taken from the system prior the migration. However, creating an alternate rootvg by cloning the existing rootvg to another disk can be used for recovery. Refer to the chapter about installing to an alternate disk on a NIM client (cloning or mksysb) in the *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389, or the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887. In case of problems with the migration or the new AIX release, the alternate rootvg, which is still at the old release, can be used to boot the system.

> **Attention:** Care must be taken with the applications running on the nodes to ensure that these applications are working on the new AIX release.

Back up the client nodes

Ensure that reliable backups for the client nodes are available. This includes a backup of the application data and none rootvg volume groups.

Set up NIM for the AIX migration on client nodes

The required NIM resources for migration of AIX on client nodes are an lpp_source and a SPOT at the AIX release the nodes are migrated to. We already created the lpp_source and SPOT resources for AIX 5.2 ML04 in 5.2, “Migrate CSM on the management server” on page 78. However, to be able to perform an unintended migration, a bosinst_data file is necessary.

The major difference between a bosinst_data resource used to install a node and perform a migration operation on a node is the value of the INSTALL_METHOD field. For an installation it is set to `overwrite`; for a migration it is set to `migrate`. We suggest exactly setting the target_disk_data stanza to point to the right disks holding the AIX that we want to migrate. The bosinst_data resource that is used for the initial installation of an AIX system is stored in the bosinst.data file in the /var/adm/ras directory of that system. Copy this file from the node to be migrated to the NIM master and use it to create a new NIM bosinst_data resource to be used for the migration. For an unintended migration, the value for PROMPT must be set to `no`. Change the INSTALL_METHOD to `migrate` and verify the stanza for target_disk_data. Example 5-12 shows a target_disk_data stanza.

Example 5-12 The target_disk_data stanza

```
target_disk_data:
    PVID = 000197ca5abe26ae
    SAN_DISKID = none
    CONNECTION = scsi0//14,0
    LOCATION = 14-08-00-14,0
    SIZE_MB = 34715
    HDISKNAME = hdisk0
```

To validate the target_disk_data shown in Example 5-12 on page 94, log on to the node and execute the following commands:

- **lspv|grep hdisk0**
  The volume group and the physical volume identifier (PVID) are shown. They should match the value in the target_disk_data stanza.
- **lsdev -Cl hdisk0**
  The location code should match the value in the target_disk_data stanza.
- **lsdev -Cl scsi0**
  This command returns the location address of the scsi0 adapter. The first part should match the location of the hdisk.
- **lsparent -Cl hdisk0**
  The output of this command shows the parent device of hdisk0.

This bosinst.data file is used to create a NIM bostinst_data resource. The **smit nim_mkres** command can be used to perform the resource creation.

Perform the AIX migration

Allocate the necessary NIM resources to the node using **smit nim_mac_res**. Other resources, such as script or installp_bundle, can also be allocated. For our example, we allocated the lpp_source, SPOT, and bosinst_data resources. Verify the allocated resources using **lsnim -l Node_name** as shown in Example 5-13.

Example 5-13 Allocated NIM resources for AIX migration

```
root@ms_01:/ # lsnim -l node_01
node_01:
   class          = machines
   type           = standalone
   platform       = chrp
   netboot_kernel = mp
   if1            = master_net node_01 0002556F1FE3 ent
   cable_type1    = N/A
   Cstate         = ready for a NIM operation
   prev_state     = currently running
   Mstate         = currently running
   bosinst_data   = node_01_migrate
   lpp_source     = lpp_AIX52H
   spot           = 52Hspot_res
   cpuid          = 000197CA4C00
   control        = master
   Cstate_result  = reset
```

Now the node can be set to migrate using **smit nim_mac**. From the menu, select **Perform Operations on Machines**, then select the node to be migrated and the operation bos_inst, and set the options in the option menu as in Example 5-14.

Example 5-14 NIM install options for an AIX migration

```
Perform a Network Install

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

                                                        [Entry Fields]
  Target Name                                          node_01
  Source for BOS Runtime Files                         rte                    +
  installp Flags                                      [-agX]
  Fileset Names                                       []
  Remain NIM client after install?                     yes                    +
  Initiate Boot Operation on Client?                   no                     +
  Set Boot List if Boot not Initiated on Client?       no                     +
  Force Unattended Installation Enablement?            no                     +
    ACCEPT new license agreements?                    [yes]                   +
```

Stop all applications on the node, then shut down the node. On the CSM management server, run the command **netboot -n Node_Name** to start the migration. The migration can be monitored using **lsnim -Fl Node_Name** or by watching the console output. To open a read-only console to node Node_Name, execute **rconsole -r -t -n Node_Name** on the CSM management server. After a successful migration, the system will boot the new AIX release.

5.5 Migrate AIX on the management server

In this chapter, we show the migration of AIX to AIX 5.3 on the CSM management server. The migration of AIX on the CSM management server can be compared to 5.2, “Migrate CSM on the management server” on page 78. If the management server is running at AIX 5.2 ML03 or earlier, then the AIX migration includes the migration of CSM. The necessary CSM 1.4 install images are on the AIX installation CDs. The migration of CSM is automatically performed during the AIX migration. The migration of AIX consists of the following steps:

1. Get the necessary software.
2. Check for prerequisites on client nodes and HMCs.
3. Back up the management server.
4. Perform the migration.
5. Apply updates to the migrated system.
6. Update the open source software packages.
7. Post migration actions.

Note: This chapter is no replacement for the *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919. Always use the CSM product documentation to plan and perform a migration.

5.5.1 Get the necessary software

The AIX 5.3 installation CDs are required to migrate the CSM management server. The current AIX 5.3 installation CDs contain AIX 5.3.0.0, RSCT 2.4.0.0, and CSM 1.4.0.0. There will be installation CDs for future AIX 5.3 maintenance levels, for example 5.3 ML01. These CDs will contain a higher level of AIX, RSCT, and CSM. For 5.3 ML01, it should be AIX 5.3.0.10, RSCT 2.4.1.0, and CSM 1.4.0.10. The AIX 5.3 ML01 is not available as this book is being written. We will migrate to AIX 5.3 without a maintenance level and will update AIX, RSCT, and CSM after the migration to the latest available levels. With the information gathered in 5.1.2, “Where we come from” on page 73 and 5.1.3, “Gather prerequisites information” on page 74, the list of necessary software (Table 5-2) can be created.

Table 5-2 Software levels for CSM migration on the management server

File set	Installed level	Required level	Available level
CSM	1.4.0.3	1.4.0.0	1.4.0.3
bos.mp64	5.2.0.43	5.3.0.0	5.3.0.2
rsct.core.auditrm	2.3.4.0	2.4.0.0	2.4.0.0
rstc.core.errm	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.fsrm	2.3.4.0	2.4.0.0	2.4.0.0
rsct.core.gui	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.hostrm	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.rmc	2.3.4.2	2.4.0.0	2.4.0.2
rsct.core.sec	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.sensorrm	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.sr	2.3.4.1	2.4.0.0	2.4.0.1
rsct.core.utils	2.3.4.2	2.4.0.0	2.4.0.2
openCIMOM	0.8-1	0.8-1	0.8-1

File set	Installed level	Required level	Available level
expect	5.32-1	5.32-1	5.34
tcl	8.3.3-1	8.3.3-1	8.3.3
tk	8.3.3-1	8.3.3-1	8.3.3
conserver	7.2.4-1	7.2.4-1	7.2.4

After the migration of AIX, we update to the latest available software, except for expect, where we decide to stay at the current level. Get the latest available fixes and store them on the CSM management server.

5.5.2 Check for prerequisites on client nodes and HMCs

Next we need to check whether updates are required on the client nodes. Nodes running AIX 5.1 must be at CSM 1.1.0 or later. Nodes running AIX 5.2 or Linux must be at CSM 1.3.1 or later. All of our client nodes meet these requirements. Updating the client nodes is not necessary.

The HMC must be at HMC for pSeries Release 3 Version 1.0 or later. Our HMC is at Release 3 Version 2.6. No update of HMC code is necessary to migrate CSM to 1.4 on the management server.

5.5.3 Back up the management server

Back up the CSM configuration data using the **`csmbackup`** command. Then create a full system backup using the command **`smit mksysb`** followed by a backup of the data stored in none rootvg volume groups using the **`smit savevg`** command.

5.5.4 Perform the migration

To migrate AIX on the CSM management server, booting from the AIX 5.3 installation CDs is necessary. For instruction on the AIX migration, see the *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887. The management server is not operational during the AIX migration. However, the client nodes can stay up and running in production.

5.5.5 Apply updates to the migrated system

The CSM management server will reboot after the AIX migration completes. We now install the updates we got in 5.5.1, “Get the necessary software” on page 97 at the management server. Use the command **`smit update_all`**, select the directory where the updates are stored, and set the update options as shown in

Example 5-2 on page 81. This update will require another reboot of the management server. However, we will first update the open source software and reboot after this update.

5.5.6 Update the open source software packages

If CSM was migrated to 1.4 during the migration of AIX (AIX was 5.2 ML03 or an earlier version), then the open source software package openCIMOM must be updated. This can be done by executing the command **`rpm -Uvh openCIMOM-0.8-1.aix5.2.noarch.rpm`**.

5.5.7 Post migration actions

Continue with 5.2.7, “Reboot the management server” on page 82 to 5.2.11, “Perform a backup of the migrated system” on page 83 to complete the migration of AIX on the CSM management server.

6

CSM cluster administration

This chapter discusses administration topics for a Cluster Systems Management (CSM) cluster that are not covered in the CSM or AIX product documentation. Our intention is to provide useful information for problem determination and show practical ways for more effective administration of a CSM cluster.

In this chapter, the following topics are discussed:

- 6.1, “Use all available information sources” on page 102
- 6.2, “Maintaining NIM lpp_sources for one AIX release” on page 104
- 6.3, “Verify the AIX level of SPOT and backup” on page 106
- 6.4, “Node status in csmstat command output” on page 107
- 6.5, “Hardware control problems” on page 109

6.1 Use all available information sources

The administration of a computer system is not easy these days. New hardware and new software releases, both providing new functions and features, are released in short cycles. The requirement to add the new hardware and software to an existing CSM cluster will result in constant changes to the cluster software. New releases of AIX or Linux and new releases of CSM must be installed to support the new hardware. These constant changes increase the possibility of discovering problems. Errors in the new software or cluster configuration may cause down time.

There are ways to reduce these kinds of problems, starting with good planning and documentation. Chapter 5, “CSM migration scenario” on page 71 shows ways to successfully update and migrate the software in a cluster, including possible fallback procedures.

Another way to reduce problems is to stay informed. IBM provides a wide range of information that should be used. This includes a CSM mailing list, where questions can be asked or problems and suggestions can be reported, and the subscription service for pSeries servers, where you can subscribe to different topics and have new information sent to you by e-mail.

The following is a list of useful Web sites and a brief description of the content found on them:

- `http://techsupport.services.ibm.com/server/cluster/home.html`

  Technical support for clustering provides links to:

  - Corrective service for:
    - Cluster Systems Management (CSM)
    - CSM High Availability Management Server (HA MS)
    - General Parallel Files Systems (GPFS)
    - LoadLeveler
    - ESSL and Parallel ESSL
    - PSSP to CSM transition tools
  - Hints and tips for clusters
    This page links to frequently asked questions (FAQ), the CSM mailing list, the ibm.software.csm newsgroup, and CSM-related redbooks. We suggest subscribing to the CSM mailing list.
  - Cluster System Management (CSM) library
    This page has all CSM documentation in HTML and PDF formats.

  - Problem reporting
    Following this link gets you to the pSeries support home page.

- `http://www-912.ibm.com/eserver/support/fixes/fcgui.jsp`

  Fix Central provides the latest fixes and updates for your system's software, hardware, and operating system.

- `https://techsupport.services.ibm.com/server/hmc`

  Hardware Management Console (HMC) support for pSeries and iSeries provides documentation and updates for HMC software.

- `http://publib.boulder.ibm.com/clresctr/windows/public/rsctbooks.html`

  This page offers the Reliable Scalable Cluster Technology (RSCT) documentation.

- `http://www16.boulder.ibm.com/pseries/en_US/infocenter/base/aix52.htm`

  This page contains the AIX 5.2 product documentation.

- `http://publib.boulder.ibm.com/infocenter/pseries/index.jsp?topic=/com.ibm.aix.doc/infocenter/base/aix53.htm`

  From this page the AIX 5L Version 5.3 documentation can be accessed.

- `http://www.ibm.com/servers/aix/products/aixos/linux/download.html`

  The AIX Toolbox for Linux Applications download page, alphabetical listing.

- `https://techsupport.services.ibm.com/server/pseries.subscriptionSvcs`

  The subscription service for pSeries servers is located on this page. A subscription to receive e-mail for following topics is available:

  - Security advisories
  - Maintenance release information
  - Critical fixes
  - Latest software fixes
  - Installation tips
  - PTFs in error
  - High impact
  - Microcode updates
  - Hardware Management Console (HMC)

- `http://www.redbooks.ibm.com/`

  This is the IBM Redbooks home page.

6.2 Maintaining NIM lpp_sources for one AIX release

One requirement of the Network Installation Manager (NIM) is that the AIX level of an lpp_source and the Shared Product Object Tree (SPOT) have to be the same as the AIX level in system backup to be restored or the node to be booted into maintenance. Trying to install a mksysb taken from a system running AIX 5.2 ML04 using lpp_source and SPOT at AIX 5.2 ML02 likely will fail. In an environment with client nodes running at different levels of an AIX release, an lpp_source and a SPOT for each of these levels must be available to be able to restore client nodes backups or perform maintenance boots of the client nodes.

Creating an lpp_source for each AIX level (for example, AIX 5.2 ML01, 5.2 ML02, 5.2 ML03, and so on) from the installation media requires some disk space. The use of hard links can reduce the amount of disk space needed for different levels of one AIX release. This requires that all lpp_sources are stored in one file system located on one logical volume because hard links work only within a file system. Links crossing file system boundaries are always symbolic links. Symbolic links cannot be used because the lpp_source is mounted to the client node during the installation. A symbolic link on the client then points to a location that does not exist on the client, and the installation will fail.

For our example with the different AIX 5.2 level, there is one lpp_source with the AIX 5.2 ML01 installation images at /nim/lppsource/AIX52ML01. To create a new lpp_source for AIX 5.2 ML02 at /nim/lppsource/AIX52ML02 using the AIX 5.2 ML01 installation images and the AIX 5.2 ML02 updates, we may simply copy the AIX 5.2 ML01 lpp_source to the AIX 5.2 ML02 location and apply the ML02 updates to it. A copy requires disk space. We use hard links to have the files that are stored under /nim/lppsource/AIX52ML01 available at the location for the AIX 5.2 ML02 lpp_source. The shell script lnlppsource shown in Example 6-1 can be used to create the directory structure and link all files from an existing lpp_source directory.

Example 6-1 The lnlppsource script

```
#!/usr/bin/ksh
# takes one parameter: the top directory of the source lppsource
# The directory structure from the source is setup and all files
# in there are hard linked to the current directory

source=$1
if [ ! -d "$source" ]
then
   echo "directory $source not found"
   exit 1
fi

# first create all subdirectories in cwd
```

```
(cd $source;find . -type d -print)|while read dir
do
  mkdir $dir
done

# then link all files from source
(cd $source;find . -type f  -print)|cut -c 3-|while read file
do
  echo "ln $source/$file $file"
  ln $source/$file $file
done

# remove any .toc file we may have linked from the source
find . -name .toc -exec rm {} \;
```

To create a new lpp_source at /nim/lppsource/AIX52ML02 from an existing lpp_source at /nim/lppsource/AIX52ML01, and the updates at /nim/updates/AIX52ML02, execute the following commands:

```
mkdir /nim/lppsource/AIX52ML02
cd /nim/lppsource/AIX52ML02
lnlppsource /nim/lppsource/AIX52ML01
cd installp/pcc
lnlppsource /nim/updates/AIX52ML02
nim -o define -t lpp_source -a server=master -a
location=/nim/lppsource/AIX52ML02 -a arch=power AIX52ML02
lsnim -Fl AIX52ML02
```

Ensure that the simages = yes flag is set for the lpp_source.

Note: The AIX updates used in this example are downloaded from IBM Support Fix Central and placed into /nim/updates/AIX52ML02. These updates do not have the same directory structure as the lpp_source build from the installation media. The updates have to be linked into ./installp/ppc of the new lpp_source.

The lpp_source AIX52ML02 can now be used to build a SPOT.

Attention: Support for new hardware may be available with new maintenance levels. The necessary device drivers are not part of the updates. These device drivers are found only on the updated installation media, which can be ordered from IBM.

6.3 Verify the AIX level of SPOT and backup

As stated in 6.2, “Maintaining NIM lpp_sources for one AIX release” on page 104, the SPOT and the system backup must be at the same levels for a successful restore. During creation, NIM stores the *oslevel* information of a SPOT and a backup into the NIM object. The **lsnim -l SPOT_name** and l**snim -l backup_name** commands can be used to verify the stored oslevel. However, NIM does not update the oslevel stored for a mksysb if the backup is overwritten by a new backup.

Consider a client node, initially installed using AIX 5.2 ML01. After the installation and setup of the client node, a system backup was created and added to NIM as a mksysb resource. The mksysb resource location is /nim/backup/nodename. Weekly system backups of the node are done by mounting the /nim/backup directory from the NIM master to the node, renaming the old backup to `nodename.old` and creating a new one using the original used file name. Then the node is updated to AIX 5.2 ML04. The backups that are created after this update require a SPOT at AIX 5.2 ML04 for a successful installation. However, the NIM mksysb resource still has the oslevel information taken from the initial system backup that was used to create the mksysb resource. A hard disk failure on the client node requires a restore of the last system backup. If you are not sure which AIX level was used on the client, use the **lsnim -l backup_name** command to find out. Our **lsnim -l backup_name** shows `5200-01`, so it uses the AIX 5.2 ML01 SPOT to restore the backup. The restore of this backup will most likely fail.

The real AIX level of the client nodes system backup is saved in the ./image.data of the backup. This file can be restore using the **restore -xqvf backup_filename** ./image.data command. For our example, with /nim/backup/nodename as the location of the mksysb NIM resource, use the following commands to restore the ./image.data file:

```
mkdir /tmp/hugo
cd /tmp/hugo
restore -xqvf /nim/backup/nodename ./image.data
```

Open the image.data file using the **view** command and search for the OSLEVEL_R and OSLEVEL fields.

To correct the AIX level information stored in NIM for the mksysb resource, remove the NIM mksysb resource using the **nim -o remove mksysb_name** command and recreate it using **smit nim** → **Perform NIM Administration Tasks** → **Manage Resources** → **Define a Resource**.

> **Attention:** Do not specify the -arm_image=yes option on the remove operation. This deletes the backup image.

6.4 Node status in csmstat command output

The **csmstat** command can be used to get a quick overview on the client nodes. Example 6-2 shows its output.

Example 6-2 Output of csmstat command with all nodes online

```
root@ms_01:/ # csmstat
-------------------------------------------------------------------------------
Hostname           HWControlPoint    Status    PowerStatus   Network-Interfaces
-------------------------------------------------------------------------------
node_01            hmcitso           on        on            en0-Online
node_02            hmcitso           on        on            en0-Online
node_03            hmcitso           on        on            en0-Online
node_04            hmcitso           on        on            en0-Online
node_05            hmcitso           on        on            eth0-Online
node_06            hmcitso           on        on            en0-Online
node_07            hmcitso           on        on            eth0-Online
node_08            hmcitso           on        on            eth0-Online
```

However, the **csmstat** command output may look like Example 6-3, indicating a problem.

Example 6-3 Output of csmstat command with one node in Status = off

```
root@ms_01:/ # csmstat
-------------------------------------------------------------------------------
Hostname           HWControlPoint     Status   PowerStatus   Network-Interfaces
-------------------------------------------------------------------------------
node_01            hmcitso            off       on           unknown
node_02            hmcitso            on        on           en0-Online
node_03            hmcitso            on        on           en0-Online
node_04            hmcitso            on        on           en0-Online
node_05            hmcitso            on        on           eth0-Online
node_06            hmcitso            on        on           en0-Online
node_07            hmcitso            on        on           eth0-Online
node_08            hmcitso            on        on           eth0-Online
```

node_01 can be reached over its en0 network interface using the **ping**, **telnet**, and **ssh** commands. All necessary subsystems belonging to the rsct_rm group are up on the CSM management server and on node_01. However, for some reason the CSM management server cannot get the required information from node_01.

We are using the **lsrsrc** command and the CT_CONTACT environment variable to query the resources defined for node_01 from the management server. Set and export CT_CONTACT to the client node the following **lsrsrc** commands should address. Then use the **lsrsrc** command on the management server to list resources from the problem client node node_01. Example 6-4 shows a working example for node_03 followed by the results returned from node_01.

Example 6-4 Using lsrsrc to read resources from client nodes

```
root@ms_01:/ # export CT_CONTACT=node_03
root@ms_01:/ # lsrsrc
class_name
"IBM.Association"
"IBM.ATMDevice"
"IBM.AuditLog"
"IBM.AuditLogTemplate"
"IBM.Condition"
"IBM.EthernetDevice"
"IBM.EventResponse"
"IBM.FDDIDevice"
"IBM.Host"
"IBM.FileSystem"
"IBM.PagingDevice"
"IBM.PhysicalVolume"
"IBM.Processor"
"IBM.Program"
"IBM.TokenRingDevice"
"IBM.Sensor"
"IBM.Sfp"
"IBM.ServiceEvent"
"IBM.ManagementServer"
"IBM.NetworkInterface"
"IBM.HostPublic"
"IBM.DRM"
"IBM.WLM"
"IBM.LPAR"
root@ms_01:/ # lsrsrc "IBM.NetworkInterface"
Resource Persistent Attributes for: IBM.NetworkInterface
resource 1:
        Name             = "en0"
        DeviceName       = "ent0"
        IPAddress        = "192.168.100.185"
        SubnetMask       = "255.255.255.0"
        Subnet           = "192.168.100.0"
        CommGroup        = ""
        HeartbeatActive  = 0
        Aliases          = {}
        ActivePeerDomain = ""
        NodeNameList     = {"node_03"}
root@ms_01:/ # export CT_CONTACT=node_01
root@ms_01:/ # lsrsrc "IBM.NetworkInterface"
```

```
2610-418 Permission is denied to access the resources or resource class
specified in this command.
```

The error returned trying to read the IBM.NetworkInterface from node_01 indicates a permission problem. The command **/usr/sbin/rsct/bin/ctsthl -l** on both the CSM management server and the client node can be used to compare the keys added for the node. The key should have the same value on both the management server and the node. The command **updatenode node_01** will get the keys into sync.

However, the open question is: Why are the keys out of sync? We found that in case the Dynamic Logical Partition (DLPAR) functionality is not working for a Logical Partition (LPAR), the following command sequence was suggested to get DPLAR working again. On the HMC and in the LPAR run:

```
/usr/sbin/rsct/install/bin/recfgct
/usr/sbin/rsct/bin/rmcctrl -z
/usr/sbin/rsct/bin/rmcctrl -A
/usr/sbin/rsct/bin/rmcctrl -p
```

The commands used as shown above reconfigure RSCT, including the security setup. Using the **updatenode** command after the reconfiguration of RSCT on the node updates the management server with the node's new security key.

6.5 Hardware control problems

The hardware control commands **rconsole** and **rpower** are sensitive to configuration changes made on the hardware control point, which is the HMC in our case and following examples. These configuration changes include changes to the LPAR definitions. Deleting LPARs, creating new LPARs, and redefining just deleted LPARs may be enough to see some mysterious hardware control problems. Commands such as **netboot** and **getadapters** depend on correct working hardware control and will fail if the **rconsole** and **rpower** commands are failing. Example 6-5 shows some hardware control problems a CSM administrator may face.

Example 6-5 Hardware control problems

```
root@ms_01:/ # rconsole -t -n node_01
[Enter `^Ec?' for help]
.hmc_s1:  2651-636 [node_01] Invalid hardware control point address specified
"itsohmc"

root@ms_01:/ # rconsole -t -n node_02
[Enter `^Ec?' for help]
```

```
.hmc_s1:  2651-675 [node_02] Error connecting to Hardware Control Point
"hmcitso"
rconsole:  2651-864 Remote console xinit command DISPLAY=; CSM_CHECK=RCONSOLE;
CSM_CHECK=node_02; /opt/csm/bin/hmc_s1 -w -s node_02 failed for node node_02

root@ms_01:/ # rconsole -t -n node_02
[Enter `^Ec?' for help]
csp_console:  2651-872 The ConsolePortNum attribute is not defined in the CSM
database for node node_02
rconsole: 2651-993 Issuing the command "/opt/csm/bin/csp_console hmcitso
node_02 -1 -1 " gave a return code of 255.  The routine will continue.

root@ms_01:/ # rpower -a query
2651-636 [node_01] Invalid hardware control point address specified "itsohmc"
node_02 on
node_03 on
node_04 on
node_05 on
node_06 on
node_07 on
node_08 on

root@ms_01:/ # rpower -a query
node_01 on
node_03 on
node_04 on
node_05 on
node_06 on
node_07 on
node_08 on
node_02 undefined
root@ms_01:/ # rpower -n node_02 query
node_02 on
```

All errors shown in this example are the result of configuration or usage errors. Many errors display messages, including error numbers. *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.4*, SA22-7934 contains a list of all CSM error numbers with an explanation and a suggested user response to solve the problem. General information about hardware control can be found in the hardware control chapter in *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918.

6.5.1 Invalid hardware control point message

The commands **rpower** and **rconsole** return with errors, as shown in Example 6-6 on page 111.

Example 6-6 Error 2651-636

```
root@ms_01:../csm/bin # rpower -n node_01 query
2651-636 [node_01] Invalid hardware control point address specified "itsohmc"

root@ms_01:../csm/bin # rconsole -t -n node_01
[Enter `^Ec?' for help]
.hmc_s1:  2651-636 [node_01] Invalid hardware control point address specified
"itsohmc"
```

The message for error 2651-636 states that the hardware control point address is invalid. The hardware control point for a node is stored in the HWControlPoint field of the node definition. The **`lsnode -n node_name -a HWControlPoint`** command can be used to check the definition, as shown in Example 6-7.

Example 6-7 lsnode -n node_01 -a HWControlPoint command output

```
root@ms_01:../csm/bin # lsnode -n node_01 -a HWControlPoint
node_01:  itsohmc
```

The HWControlPoint is the name of the HMC controlling node_01. It is currently set to `itsohmc`. However, this is the wrong name for the HMC. Its correct name is hmcitso. The command **`chnode -n node_01 HWControlPoint=hmcitso`** can be used to change the HWControlPoint for node_01.

However, a name resolution problem may cause the 2651-636 error, as shown in Example 6-8.

Example 6-8 Error 2651-636 because of name resolution problems

```
root@ms_01:../csm/bin # rpower -a query
2651-636 [node_02] Invalid hardware control point address specified "hmcitso"
2651-636 [node_03] Invalid hardware control point address specified "hmcitso"
2651-636 [node_04] Invalid hardware control point address specified "hmcitso"
2651-636 [node_05] Invalid hardware control point address specified "hmcitso"
2651-636 [node_06] Invalid hardware control point address specified "hmcitso"
2651-636 [node_07] Invalid hardware control point address specified "hmcitso"
2651-636 [node_08] Invalid hardware control point address specified "hmcitso"
node_01 on
root@ms_01:../csm/bin # lsnode -a HWControlPoint
node_01:  itsohmc
node_02:  hmcitso
node_03:  hmcitso
node_04:  hmcitso
node_05:  hmcitso
node_06:  hmcitso
node_07:  hmcitso
node_08:  hmcitso
```

```
root@ms_01:../csm/bin # host itsohmc
itsohmc is 192.168.100.69
root@ms_01:../csm/bin # host hmcitso
host: 0827-801 Host name hmcitso does not exist.
```

The name resolution for hmcitso must be fixed. In our case, the entry in /etc/hosts for itsohmc is wrong (as well as the HWControlPoint for node_01). hmcitso has the address 192.168.100.69. It may be necessary to stop and restart the IBM.HWCTRLRM subsystem after correcting a name resolution problem. This is done by using **`stopsrc -c -s IBM.HWCTRLRM`** followed by **`startsrc -s IBM.HWCTRLRM`**.

6.5.2 Error connecting to hardware control point

The error `2651-675 Error connecting to Hardware Control Point` in most cases is caused by a wrong user ID or wrong password defined for the hardware control point. A change of the password used for hscroot on the HMC is in many cases the cause for the error. The CSM management server also needs to know the new password. The command **`systemid`** can be used to store the correct user ID and password to access the hardware control point. See *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.4*, SA22-7934, for more information about the **`systemid`** command. Changing the user ID or password for a hardware control point may require a restart of the IBM.HWCTRLRM subsystem. This restart can be done using the commands **`stopsrc -c -s IBM.HWCTRLRM`** and **`startsrc -s IBM.HWCTRLRM`**.

6.5.3 Power status undefined

In a cluster environment with changing LPAR definitions on client systems, the node definitions in CSM could get out of sync with the LPAR definitions made on the HMC.

Consider the following situation: A new system is set up, and the plan is to have four production LPARs and one small LPAR for testing defined on it. Two of the production LPARs are in production. So they are defined, added to CSM as client nodes, and AIX is installed on them. Application installation and setup on these new systems is done by the application team. The CSM administrator now defines the LPAR for testing and adds all free system resources to this LPAR. This test system then is added to CSM as a client node. Example 6-9 shows the LPAR definitions on the system.

Example 6-9 LPAR definitions for new system

```
[hscroot@hmcitso bin]$ lssyscfg -r lpar -m itso_p690 --all
Name                 id   DLPAR  State          Profile               OpPanel
```

```
FullSystemPartition  000  NO     Not Available  PowerOnNormalProfile
p690_LPAR1           001  YES    Running        LPAR
p690_LPAR2           002  YES    Running        LPAR
p690_Test            003  YES    Running        LPAR
```

Later, the two other production LPARs have to be defined. The test LPAR got all free system resources assigned and it seems to be easier to delete this test LPAR, create the two new LPARs, and re-create the test LPAR. The CSM definitions for the test system defined on the test LPAR are not deleted and will be reused. Example 6-10 shows the LPAR definitions after this change.

Example 6-10 Changed LPAR definitions

```
[hscroot@hmcitso bin]$ lssyscfg -r lpar -m itso_p690 --all
Name                 id   DLPAR  State          Profile               OpPanel
FullSystemPartition  000  NO     Not Available  PowerOnNormalProfile
p690_LPAR1           001  YES    Running        LPAR
p690_LPAR2           002  YES    Running        LPAR
p690_LPAR3           003  YES    Running        LPAR
p690_LPAR4           004  YES    Running        LPAR
p690_Test            005  YES    Running        LPAR
```

Note the ID change for p690_Test. The two new production systems are added to CSM. In CSM, we now end up with a wrong LParID for the test system. Example 6-11 shows the **lsnode** output.

Example 6-11 CSM definitions for the new client nodes

```
root@ms_01:/ # lsnode -a HWType,HWModel,HWSerialNum,HWControlNodeId,LParID
Prod_01: 7040, 61R, 022BE1A, p690_LPAR1, 001
Prod_02: 7040, 61R, 022BE1A, p690_LPAR2, 002
Test_01: 7040, 61R, 022BE1A, p690_Test, 003
Prod_03: 7040, 61R, 022BE1A, p690_LPAR3, 003
Prod_04: 7040, 61R, 022BE1A, p690_LPAR4, 004
```

The result of an **rpower -a query** with the LParID set as in Example 6-11 is shown in Example 6-12.

Example 6-12 Output of rpower -a query with duplicate LParID

```
root@ms_01:/ # rpower -a query
Prod_01 on
Prod_02 on
Test_01 on
Prod_04 on
Prod_03 undefined
```

```
root@ms_01:/ # rpower -n Prod_03 query
Prod_03 on
```

The LParID for the CSM client node Test_01 must be corrected using the command **chnode -n Test_01 -a LParID=005**. We suggest verifying the CSM node definitions after changes to LPAR definitions are made to ensure that the CSM configuration is still correct.

6.5.4 Power on for a client node is not working

In an environment with LPARs and DLPARs, a power on (power query) of a client node gives a return code of 0, but the node stays powered off as shown in Example 6-13.

Example 6-13 Power on a node using the rpower command

```
root@ms_01:/ # rpower -la query
node_01 off   LCDs are blank
node_02 on   LCDs are blank
node_03 on   LCDs are blank
node_04 on   LCDs are blank
node_05 on   LCD1 = Linux ppc64   2.4.21-111-pseri   LCD2 is blank
node_06 on   LCDs are blank
node_07 on   LCD1 = SuSE Linux ppc64   2.6.5-7.97-pseri   LCD2 is blank
node_08 on   LCD1 = Linux ppc64   2.4.21-20.EL   LCD2 is blank
root@ms_01:/ # rpower -n node_01 on
node_01 on complete rc=0
root@ms_01:/ # rpower -la query
node_01 off   LCDs are blank
node_02 on   LCDs are blank
node_03 on   LCDs are blank
node_04 on   LCDs are blank
node_05 on   LCD1 = Linux ppc64   2.4.21-111-pseri   LCD2 is blank
node_06 on   LCDs are blank
node_07 on   LCD1 = SuSE Linux ppc64   2.6.5-7.97-pseri   LCD2 is blank
node_08 on   LCD1 = Linux ppc64   2.4.21-20.EL   LCD2 is blank
```

The client node node_01 stays powered off. The **rpower** command issues a power on of the node to the HMC controlling the node. However, the **rpower** command does not wait for the power on to complete. With CSM 1.4 a new flag -w was added to the **rpower** command, causing **rpower** to wait for the result of the **on, off**, **reboot**, **cec_on**, and **cec_off** subcommands. In Example 6-14 on page 115, we show how the **rpower** command with the -w flags works.

Example 6-14 Power on nodes using the rpower command with -w flag

```
root@ms_01:/ # rpower -w -n node_01,node_02 on
node_02 on complete rc=0
2651-664 [node_01] CIMException
```

However, the error message from the **rpower** command is not very useful. We suggest using Web-based System Manager, connecting to the HMC controlling the problem node, and trying to perform the failing operation there. In our example, this is the activation of an LPAR. Here we see the error: HSCL03EA There is an insufficient number of processors: Obtained - 0, Required - 1. There is no free CPU resource for the LPAR left. The LPAR configuration should be checked and corrected.

6.5.5 IBM.HWCTRLRM trace and Java controls

Each resource manager writes a trace file. The location for the trace file is /var/ct/IW/log/mc/Resource_Manager. The trace file for IBM.HWCTRLRM is in the directory /var/ct/IW/log/mc/IBM.HWCTRLRM. The traces are in binary format and must be formatted using the **/usr/sbin/rsct/bin/rpttr** command. This trace may provide useful information to debug hardware control problems. Usage examples for **rpttr** include:

- `rpttr /var/ct/IW/log/mc/IBM.HWCTRLRM/trace`
  Writes the formatted trace to stdout.
- `rpttr /var/ct/IW/log/mc/IBM.HWCTRLRM/trace >/tmp/hwc.trace`
  Writes the formatted trace to the /tmp/hwc.trace file
- `rpttr -f /var/ct/IW/log/mc/IBM.HWCTRLRM/trace`
  Shows records as they are added to the trace

In addition to the trace file, a Java trace can be enabled for IBM.HWCTRLRM. This is done by setting the environment variable HC_JAVA_VERBOSE to the file where the trace will be stored. IBM.HWCTRLRM is started under control of the system resource controller (SRC) subsystem. To enable the Java trace, execute the following commands:

```
stopsrc -c -s IBM.HWCTRLRM
ps -ef|grep IBM.HWCTRLRM
```

(Ensure that IBM.HWCTRLRM did terminate.)

```
startsrc -s IBM.HWCTRLRM -e HC_JAVA_VERBOSE=/tmp/jni.txt
```

The files /tmp/jni[hmc].txt and /var/log/csm/hmc[ip.of.HMC].java_trace are created. The file /tmp/jni[hmc].txt holds Java-relevant trace data. The file /var/log/csm/hmc[ip.of.HMC].java_trace holds the functional trace for

IBM.HWCTRLRM. To turn the Java trace off, stop and restart IBM.HWCTRLRM using these commands:

```
stopsrc -c -s IBM.HWCTRLRM
ps -ef|grep IBM.HWCTRLRM
```

(Ensure that IBM.HWCTRLRM did terminate.)

```
startsrc -s IBM.HWCTRLRM
```

On AIX, different versions of Java can be installed. The command **java -version** returns the current default Java version used. The HC_JAVA_PATH environment variable can be used to force IBM.HWCTRLRM to use a different Java version installed on the system. Example 6-15 shows how IBM.HWCTRLRM can be forced to used Java 1.3.1 on a system with Java 1.4 as the default Java version.

Example 6-15 Using a different Java version for IBM.HWCTRLRM

```
8:root@cws5et:/ # java -version
java version "1.4.1"
Java(TM) 2 Runtime Environment, Standard Edition (build 1.4.1)
Classic VM (build 1.4.1, J2RE 1.4.1 IBM AIX build ca1411-20040110 (JIT enabled:
jitc))
8:root@cws5et:/ # ls -d /usr/java*
/usr/java131  /usr/java14
8:root@cws5et:/ # stopsrc -c -s IBM.HWCTRLRM
0513-044 The IBM.HWCTRLRM Subsystem was requested to stop.
8:root@cws5et:/usr # ps -ef|grep HWC
    root 30062  7000   0   Sep 20      - 14:51 /usr/sbin/rsct/bin/IBM.HWCTRLRMd
8:root@cws5et:/ # ps -ef|grep HWC
    root 19188 16228   1 00:28:06  pts/8  0:00 grep HWC
8:root@cws5et:/ # startsrc -s IBM.HWCTRLRM -e HC_JAVA_PATH=/usr/java131
0513-059 The IBM.HWCTRLRM Subsystem has been started. Subsystem PID is 31642.
8:root@cws5et:/ # lssrc -s IBM.HWCTRLRM
Subsystem         Group            PID          Status
 IBM.HWCTRLRM     rsct_rm          31644        active
8:root@cws5et:/ # ps ewwwu 31644
USER       PID %CPU %MEM   SZ  RSS    TTY STAT    STIME  TIME COMMAND
root     31644  0.0  0.0 2952 2900      - A    00:28:57  0:00
/usr/sbin/rsct/bin/IBM.HWCTRLRMd TERM=dumb AUTHSTATE=compat SHELL=/usr/bin/ksh
HOME=/ USER=root
PATH=/usr/bin:/etc:/usr/sbin:/usr/ucb:/usr/bin/X11:/sbin:/usr/java14/jre/bin:/u
sr/java14/bin TZ=MEZ-1MES-2,M3.5.0,M10.5.0/03:00:00 LANG=en_US
LOCPATH=/usr/lib/nls/loc LC__FASTMSG=true ODMDIR=/etc/objrepos LOGNAME=root
LOGIN=root HC_JAVA_PATH=/usr/java131
NLSPATH=/usr/lib/nls/msg/%L/%N:/usr/lib/nls/msg/%L/%N.cat
HC_POWER_STATUS_MODE=0
```

We suggest using the Java version shipped with AIX to run IBM.HWCTRLRM.

7

CSM High Availability Management Server

High Availability Management Server (HA MS) is designed to provide high availability protection for the CSM management server in business-critical clusters. It prevents the CSM management server from being the single point of failure through an automatic failover capability to a backup management server. Thus, it improves reliability, availability, and serviceability of the CSM management server.

HA MS is an optional feature of CSM that must be purchased separately. It can be implemented at any time during the life of the CSM cluster.

In this chapter, the following topics are discussed:

- 7.1, “When to deploy CSM HA MS” on page 118.
- 7.2, “HA MS installation requirement” on page 118.
- 7.3, “HA MS implementation scenarios” on page 120.
- 7.4, “HA MS configuration” on page 124.
- 7.5, “HA MS operations” on page 133.
- 7.6, “HA MS administration” on page 140.
- 7.7, “HA MS log and debug information” on page 142.

7.1 When to deploy CSM HA MS

CSM HA MS typically is used in the following scenarios:

1. There is a business need for the management server to be high available at all times.
2. The management server is employed for business-critical monitoring and these monitors must be highly available.

7.2 HA MS installation requirement

HA MS requires the backup management server to have the same hardware and software configuration as the primary management server. Both management servers must be connected to the same networks.

See the chapter about CSM High Availability Management Server in *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918, for details about CSM HA MS installation requirements.

7.2.1 Software support

The HA MS is supported on AIX 5L V5.2 with maintenance level 04. HA MS 1.4.0.10 is supported on AIX 5L V5.2 and AIX 5L V5.3 if the prerequisites shown in Table 7-1 are met.

Table 7-1 HA MS requisite matrix

	AIX 5.2	**AIX 5.3**
csm.server	1.4.0.10	1.4.0.10
csm.hams	1.4.0.10	1.4.0.10
sam.core	1.2.0.2	1.2.0.2
rsct.basic	2.3.5.0	2.4.0.0
rsct.core	2.3.5.0	2.4.0.0

7.2.2 Storage support

HA MS requires two shared disks to be connected and accessible by both management servers.

The following disk storage systems are supported on pSeries server:

- DS4300
- DS4400
- DS4500
- SSA

7.2.3 Network consideration

Both management servers (MS) must connect to the same cluster VLAN and management VLAN. This enables the backup management server to continue to communicate to the nodes after active management server failover.

The HA MS does not take over the cluster VLAN IP and management VLAN IP that are connected the active MS. This is because of RSCT restrictions that do not allow IP address takeover for a CSM management server's connection to the nodes or hardware contol points.

HA MS can provide IP address takeover for the public IP address. This is the IP address that users on the public network use to access the CSM management server. The IP address takeover is implemented using the IP alias on the public adapter. However, this is an optional configuration for HA MS.

7.2.4 Remote shell consideration

HA MS requires that the two MS servers can talk to each other by way of the remote shell command (**dsh**) that is defined within the **csmconfig** command. The setup of the remote shell authentication for **dsh** is done by HA MS. However, the remote shell that is used must be one of the remote shells that is supported by the management server.

HA MS can set up the remote shell automatically if the following **csmconfig** attributes have been set to the same value on both management servers:

- RemoteCopyCmd
- RemoteShell
- SetupRemoteShell
- SetupKRB5

The following values are used in our setup:

- RemoteCopyCmd=/usr/bin/scp
- RemoteShell=/usr/bin/ssh
- SetupRemoteShell=1
- SetupKRB5=0

7.2.5 Network Installation Management (NIM) consideration

HA MS provides the ability for the NIM server to fail over to a backup MS. However, we did not test this function in our environment.

For more about the configuration and setup for NIM master failover, see the CSM High Availability Management Server chapter in *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918.

> **Note:** NIM lpp_source and SPOT must be defined in the /csminstall/AIX directory. When HA MS starts, it might take a longer time to move the resources from /csminstall directory to shared disk if you have multiple copies of lpp_sources defined.

7.2.6 Symbolic link consideration

HA MS does not allow any symbolic links made in the /csminstall, /cfmroot, /etc/opt/csm, or /var/opt/csm file systems stored on the HA MS shared disk to store file systems locally. This is to prevent the symbolic link from pointing to the wrong file or file system when the active MS fails over to the backup MS.

For example, we should copy /etc/hosts to /cfmroot/etc/hosts instead of creating a symbolic link from /etc/hosts to /cfmroot/etc/hosts. When the HA MS starts, a warning message is given if such symbolic links exist.

7.3 HA MS implementation scenarios

There are two scenarios for HA MS implementation:

- CSM is already installed on the primary management server.
- CSM has not been installed on either management server.

7.3.1 Environment setup

HA MS requires two identical IBM servers to be configured as the primary and backup management servers for the CSM cluster.

The following configurations are used in our setup:

- Cluster VLAN and Management VLAN are on the same VLAN.
- Public VLAN is not configured, so IP Address Takeover is not implemented.
- DS4500 disk storage is used.
- Secure Shell (ssh) is configured as the remote shell.

- NIM failover is not implemented.
- AIX 5.2 with maintenance level 04 is installed.
- CSM server V1.4.0.3 and HA MS V1.4.0.2 are installed.
- CSM is already installed on the primary management server.
- Backup management server was installed from scratch.
- Network Time Protocol (NTP) is implemented only between management servers. In production environment, we recommend implementing NTP on management servers and nodes.
- We have stored the /csminstall file system in a separate logical volume on the AIX management server.

Note: Make sure that the /csminstall file system is not automatically mounted during the system reboot because this would interfere with the HA MS.

Note: We have successfully tested HA MS 1.4.0.2 on AIX 5.2 with ML04. The HA MS was upgraded to HA MS 1.4.0.10 on AIX 5.3. HA MS was upgraded and tested successfully on AIX 5.3.

The configuration shown in Figure 7-1 on page 122 was configured and tested. The management LAN and Cluster LAN are on the same network. There is no public LAN in our configuration.

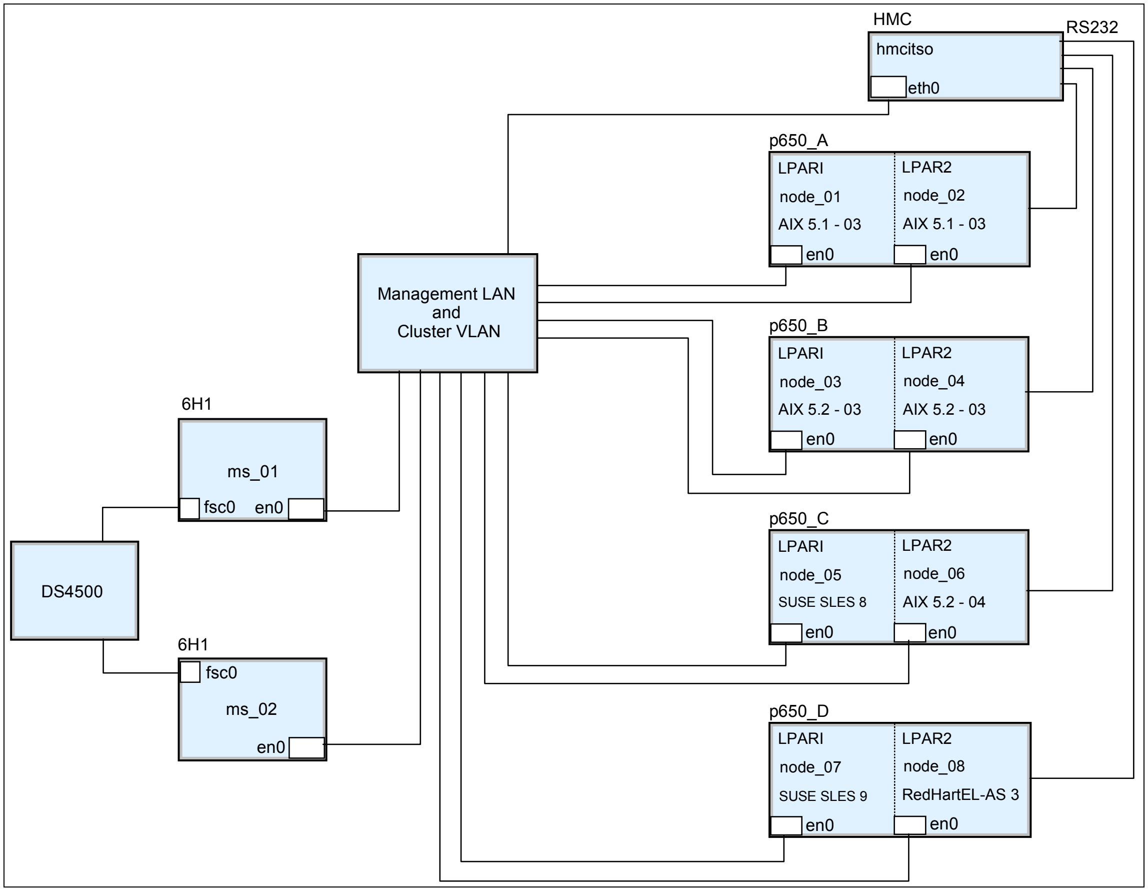


Figure 7-1 CSM HA MS configuration

7.3.2 CSM MS installation on the backup MS

There are two methods of installing the backup management server:

- Install from scratch.
- Clone existing MS system image and install it into the backup MS.

Install from scratch

Follow these steps to install from scratch:

1. Install AIX 5.2 from AIX 5.2 operating system media.
2. Install AIX 5.2 maintenance level 04.
3. Set up an IP connection for the VLAN on the backup MS.

4. Install CSM server. See 3.3, “Installation of the CSM management server” on page 11.

> **Note:** You only have to install CSM and its related file sets. The CSM configuration will be done by the HA MS when it starts.

Clone active MS system image to a backup MS

To perform the clone of the active MS system image to the backup management server, the stages involved are:

- Perform an mksysb backup of the active management server.
- Perform an mksysb install of the backup MS from the NIM master.
- Set up an IP connection for the VLAN on the backup MS.

Verify that the remote shell is configured on the backup MS

HA MS requires that the two management servers be able to **dsh** to each other using the remote shell defined with the **csmconfig** command.

To check that **ssh** is set up as the remote shell on the backup MS (ms_02), issue:

```
csmconfig
```

Example 7-1 shows the sample output from the command.

Example 7-1 csmconfig on ms_02

```
AddUnrecognizedNodes = 0 (no)
ClusterSNum =
ClusterTM = 9078-160
DeviceStatusFrequency = 12
DeviceStatusSensitivity = 8
ExpDate =
LicenseProductVersion = 1.3
PowerStatusMode = 0 (Mixed)
RegSyncDelay = 1
RemoteCopyCmd = /usr/bin/scp
RemoteShell = /usr/bin/ssh
SetupKRB5 = 0
SetupRemoteShell = 1 (yes)
```

7.4 HA MS configuration

The basic steps for HA MS implementation are:

1. Install HA MS.
2. Configure the shared disk.
3. Configure HA MS.

7.4.1 HA MS installation

HA MS is an optional feature of CSM that is found on a separate HA MS CD that is available for purchase.

Install the following software on both management servers using the HA MS CD:

- csm.hams
- sam.core

HA MS prerequisites

HA MS uses RSCT peer domain (RPD) for communication between both management servers. The RSCT package is contained in the rsct.basic package.

Install the rsct.basic package from the AIX 5.2 CD 3 media. Issue the command:

```
geninstall -IaX -d /dev/cd0 rsct.basic
```

You can also use the **smit installp** command to install the rsct.basic file sets.

Install HA MS packages

To install csm.hams and sam.core, insert the HA MS CD and issue the following command:

```
geninstall -IaXY -d /dev/cd0 sam.core csm.hams
```

Alternatively, use the **smit installp** command to install the HA MS file sets.

> **Note:** The HA MS and SAM licenses are accepted automatically during the csm.hams and sam.core installation. You must accept the license during installation.

Install HA MS program temporary fix (PTF)

The following PTF for HA MS was installed:

- csm.hams.1.4.0.2
- sam.core.rte.1.2.0.1

7.4.2 Set up NTP on both management servers

The RSCT registry is stored locally on each management server. The registry cannot be stored on the shared disk because the disk is accessible only by one of the management servers at any time. HA MS keeps track of the registry changes by storing the text equivalent of the registry on the shared disk. Therefore, the two management servers must be time-synchronized so that the HA MS can verify the latest changes of the registry.

Set up the NTP master on the primary MS (ms_01)

Edit /etc/ntp.conf and put in the entry as shown in Example 7-2.

Example 7-2 /etc/ntp.conf on ms_01

```
#broadcastclient
server 127.127.1.0
driftfile /etc/ntp.drift
tracefile /etc/ntp.trace
```

To start the xntpd daemon, issue the following command:

```
startsrc -s xntpd
```

Set up the NTP client on the secondary MS (ms_02)

Edit /etc/ntp.conf and put in the entry as shown in Example 7-3.

Example 7-3 /etc/ntp.conf on ms_02

```
#broadcastclient
server 192.168.100.182
driftfile /etc/ntp.drift
tracefile /etc/ntp.trace
```

To start the xntpd daemon, issue the following command:

```
startsrc -s xntpd
```

Verify the time synchronization on both management servers

There are two ways to verify timed synchronization on both management servers:

1. Use the **date** command on both servers; each should return the same date and time.
2. Use the **ntpq** command to verify.

   On ms_01 MS, issue the **ntpq** command as shown in Example 7-4 on page 126.

Example 7-4 ntpq command on ms_01

```
root@ms_01:/ # ntpq
ntpq> as
ind assID status  conf reach auth condition  last_event cnt
===========================================================
  1 49044  9614   yes   yes  none  sys.peer   reachable  1
```

On the ms_02 management server, issue **ntpq** as shown in Example 7-5.

Example 7-5 ntpq command on ms_02

```
root@ms_02:/ # ntpq
ntpq> asp
ind assID status  conf reach auth condition  last_event cnt
===========================================================
  1 52991  9624   yes   yes  none  sys.peer   reachable  2
```

The time on both servers is synchronized when the condition field status shows `sys.peer`.

Note: If the time difference between the management servers is greater than five minutes, the xntp daemon will not be able to perform the time synchronization. Issue the **ntpdate** command to perform a manual time synchronization between servers.

7.4.3 Shared disk configuration

Two shared disks are connected to both management servers. One of the shared disks is used to store CSM data. The size must be at least 1.5 times larger than the data store on the /csminstall directory.

The second disk is used to identify the active management server with the TieBreaker during an RSCT RPD partition. The TieBreaker disk can be very small as there is no data stored. It is used only for SCSI reserves.

Note: These conditions must be followed when configuring the shared disks:

- Both shared disks should be accessible by both management servers.
- The shared volume group should not be automatically varied on when either management server is rebooted.
- The shared file systems on the shared disk should not be automatically mounted when either management server is rebooted.
- The shared logical volume names and shared file system names should be unique on both management servers.

We assume that the following disks on the disk storage have been assigned to the management servers:

- hdisk3 on ms_01 and hdisk2 on ms_02 for shared data volume group
- hdisk8 on ms_01 and hdisk7 on ms_02 for TieBreaker disk

Create shared data volume group (VG)

The creation of the shared data volume group is performed on the ms_01 management server.

Issue the following command to create the volume group csmvg:

```
mkvg -y'csmvg' -s'32' '-f' '-n' -V'80' hdisk3
```

The following volume group attributes are used:

- Volume group name: csmvg
- Physical volume name: hdisk3
- Volume group major number: 80
- Activate the volume group automatically at system restart: no

Alternatively, you can use the **smit mkvg** command to create the volume group.

Create the TieBreaker shared disk

When the disk is newly assigned, there is no physical volume ID (PVID) on the physical disk.

To assign a PVID to the disk on the ms_01 management server without creating a volume group on the disk, issue the command:

```
chdev -l hdisk8 -a pv=yes
```

Issue **lspv** to verify. Example 7-6 on page 128 shows the sample output.

Example 7-6 lspv command on ms_01

```
hdisk0          000c4d1fcb6ff7d9                    rootvg          active
hdisk1          000c4d1f568c162a                    rootvg          active
hdisk2          000c4d1f9cd9d144                    csmvg
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
hdisk8          000c4d1f9cd9da2d                    None
```

Create a shared logical volume on the shared VG

The creation of the shared logical volume is performed on the ms_01 management server. We need to create an unique jfs2log device and a jfs2 logical volume.

To create the unique jfs2log device issue:

```
mklv -y'csmlog' -t'jfs2log' csmvg 1
```

The log device must be formatted. Issue the command:

```
logform -y /dev/csmlog
```

To create the logical volume for the shared file system, issue the command:

```
mklv -y'hams_lv' -t'jfs2' csmvg 35
```

The command creates the logical volume with the following attributes:

- Logical volume name: hams_lv
- Number of logical partitions: 35
- Logical volumn type: jfs2

Create a shared file systems on the shared VG

The creation of the shared file system is performed on the ms_01 management server.

To create the enhanced journaled file system using the predefined logical volume, issue the command:

```
crfs -v jfs2 -d'hams_lv' -m'/var/opt/hams/mnt' -A'
```

The command creates the file system with the following attributes:

- Predefined logical volume: /dev/hamslv
- File system mount point: /var/opt/hams/mnt
- Mount automatically at system restart: no

Import the shared data VG information to the backup MS

The shared disk information must be imported to the backup MS. The steps are:

1. Vary off the shared volume group on ms_01.

   Issue the **varyoffvg** command on ms_01:

   ```
   varyoffvg csmvg
   ```

2. Detect the physical disks on ms_02.

   Issue the **cfgmgr** command on ms_02:

   ```
   cfgmgr -v
   ```

> **Note:** If the disks already exist on ms_02, remove the disks and run the **cfgmgr** command to re-detect the physical disk information.

3. Verify that the physical disks are detected on ms_02.

   Issue the **lspv** command on ms_02.

Issue **lspv** command to verify the configuration. Example 7-7 shows the sample output.

Example 7-7 lspv command on ms_02

```
hdisk0          000c4d2fcb1dc34a                    rootvg          active
hdisk1          000c4d1f9cd9d144                    None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          000c4d1f9cd9da2d                    None
```

4. Import the shared volume group, csmvg, on ms_02.

   Issue the following command to import the shared volume group, csmvg, with major number, 80, on ms_02 management server:

   ```
   importvg -y'csmvg' -V'80' hdisk1
   ```

   Alternatively, you can use **smit importvg** to import the volume group.

5. Update the shared volume group, csmvg, on ms_02 MS.

   Update the shared volume group, csmvg, to not automatically vary on when the system reboots.

   Issue the **chvg** command to change the volume group attribute:

   ```
   chvg -an csmvg
   ```

Verify that shared file system is accessible on both systems

Verify that the shared file system created on the shared volume group is accessible on both management servers. The **varyonvg** and **mount** commands are used for the verification.

7.4.4 Create the HA MS definition file

HA MS is supported on both AIX and Linux systems, but in our environment, HA MS is implemented and tested only on an AIX management server. AIX-related options for the HA MS definition file are discussed in this section.

For the more about the HA MS definition file, see the CSM High Availability Management Server chapter in *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918.

Options required for AIX HA MS

The default HA MS definition file (hamsdef) is stored in the /opt/csm/hams directory. The hamsdef file is used when the HA MS is started.

The following options must be configured for HA MS in AIX:

- IsPrimary= 1 for MS primary; 0 for MS backup.
- CSMInstallPartition= The logical volume where /csminstall mounted.
- FileSystemType= The type of file system on the shared disk.
- NIMNetworkInterface= The network interface to be used by the NIM master.
- PartitionDeviceName= The logical volume used for shared CSM data.
- PhysicalVolume= The device name of the physical volume used for TieBreaker.
- VolumeGroup= The shared volume group name.
- ClusterVLANAdapter= Optional. Ethernet adapter that is connected to the cluster VLAN.
- ManagementVLANAdapter= Optional. Ethernet adapter that is connected to the management VLAN.
- PublicIPAdapter= Optional. Ethernet Adapter that is connected to the Public IP VLAN.
- PublicIPAddress= Optional. The service IP alias that is used for IPAT.
- PublicIPNetmask= Optional. The netmask for the service IP alias.

Environment HA MS definition file

Example 7-8 on page 131 shows the hamsdef file used in our environment.

Example 7-8 hamsdef file

```
ms_01:
    IsPrimary=1
    FileSystemType=jfs2
    VolumeGroup=csmvg
    PhysicalVolume=/dev/hdisk8
    PartitionDeviceName=/dev/hams_lv
    CSMInstallPartition=/dev/csm_lv
    NIMNetworkInterface=en0
    ClusterVLANAdapter=en0
    ManagementVLANAdapter=en0

ms_02:
    IsPrimary=0
    FileSystemType=jfs2
    VolumeGroup=csmvg
    PhysicalVolume=/dev/hdisk7
    PartitionDeviceName=/dev/hams_lv
    NIMNetworkInterface=en0
    ClusterVLANAdapter=en0
    ManagementVLANAdapter=en0
```

Replicate the HA MS definition file to the backup MS

The hamsdef file does not have to be replicated to the backup MS. However, to avoid confusion and ease of management, we recommend that you copy the hamsdef file to the /opt/csm/hams/ directory on the backup management server.

On ms_01:

```
scp /opt/csm/hams/hamsdef ms_02:/opt/csm/hams/hamsdef
```

7.4.5 Customized scripts

HA MS enables a user to put in user-defined, pre-failover and post-failover scripts to perform user-customized tasks. These scripts must be in the /opt/csm/hams/scripts directory and executable on both management servers.

For details, see *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918.

7.4.6 Activate the failover notification

A simple way to notify when the active management server failover occurs is to configure the ActiveManagementServer ERRM condition with BroadcastEventsAnyTime response.

The ActiveManagementServer condition watches the HAMode attribute stored in csmconfig. It broadcasts a message to all login sessions if the active management server fails over to the backup management server (Example 7-9).

Example 7-9 Broadcast message when active MS fails over

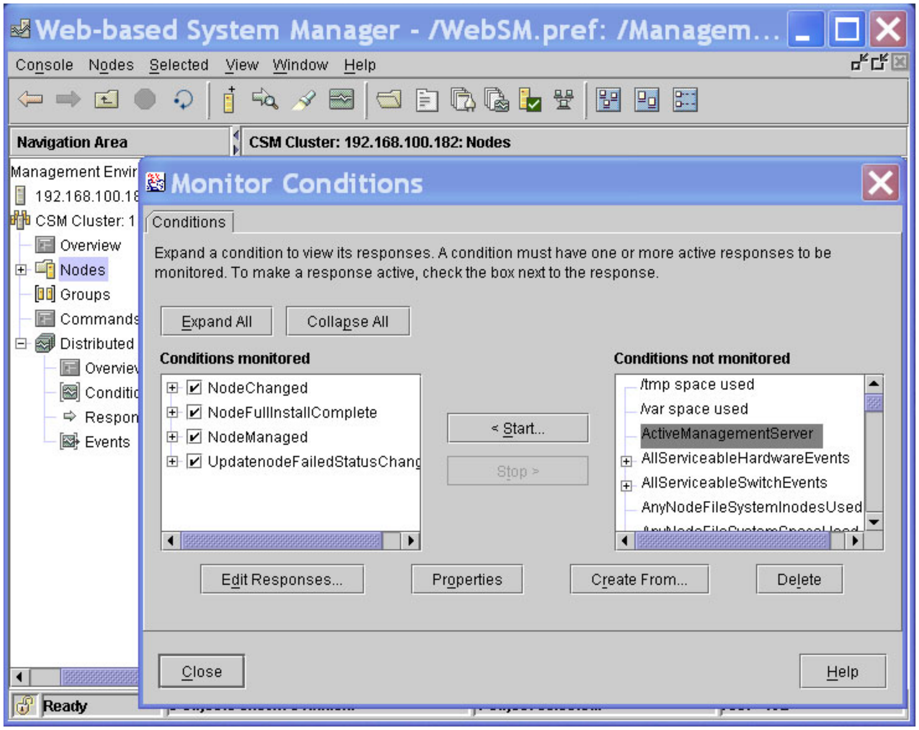
```
Broadcast message from root@ms_02 (tty) at 16:50:07 ...

Informational Event occurred for Condition ActiveManagementServer on the
resource Unknown_Resource_Name of the resource class IBM.DmsCtrl at Friday
10/15/04 16:50:07.  The resource was monitored on ms_02 and resided on {ms_02}.
```

There are two ways to configure the condition:

- Using the command line:

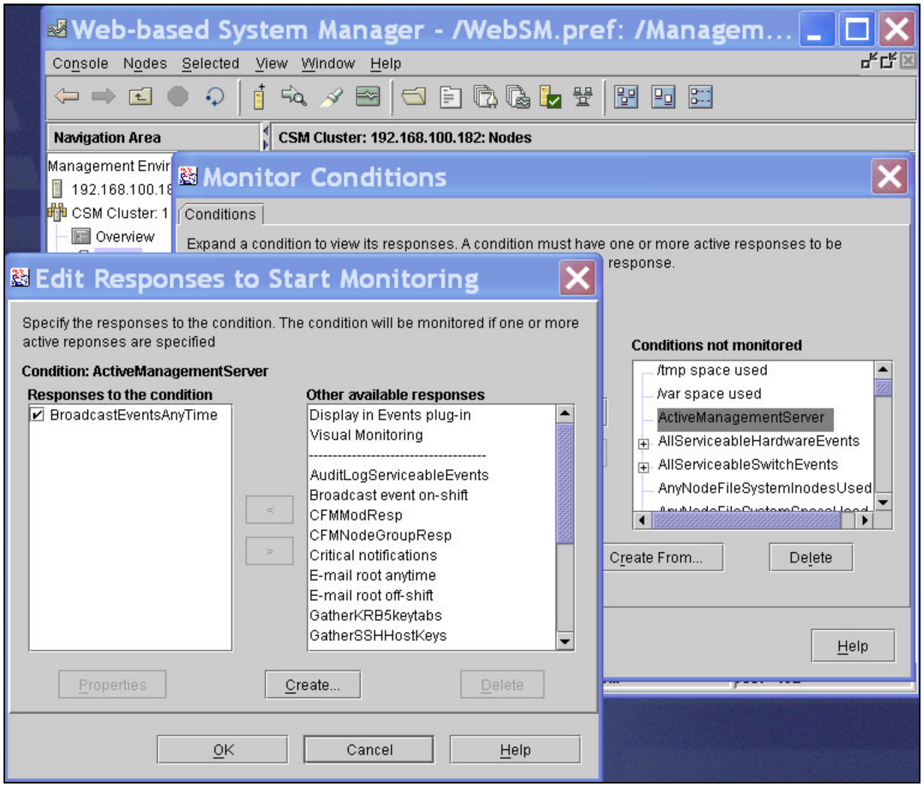
  ```
  startcondresp ActiveManagementServer BroadcastEventsAnyTime
  ```

- Using Web-based Systems Management (WebSM): Activate WebSM on the remote client, add a monitor condition as shown in Figure 7-2, and add a response as shown in Figure 7-3 on page 133.



Expand a condition to view its responses. A condition must have one or more active responses to be monitored. To make a response active, check the box next to the response.

Figure 7-2 Add a monitor condition



Expand a condition to view its responses. A condition must have one or more active responses to be response.

Specify the responses to the condition. The condition will be monitored if one or more active reponses are specified

Condition: ActiveManagementServer

Responses to the condition

☑ BroadcastEventsAnyTime

Other available responses

Display in Events plug-in
Visual Monitoring

AuditLogServiceableEvents
Broadcast event on-shift
CFMModResp
CFMNodeGroupResp
Critical notifications
E-mail root anytime
E-mail root off-shift
GatherKRB5keytabs
GatherSSHHostKeys

Conditions not monitored

/tmp space used
/var space used
ActiveManagementServer
AllServiceableHardwareEvents
AllServiceableSwitchEvents
AnyNodeFileSystemInodesUsed

Figure 7-3 Add a response to a monitor condition

7.5 HA MS operations

HA MS is an independent component of CSM. After HA MS starts, it makes the CSM data and functionality highly available. It automatically fails over to the inactive management server when the active management server is unavailable.

The following conditions have to be checked before HA MS starts:

1. Both management servers must be time synchronized.
2. The shared volume group on both management servers must be offline.
3. CSM data is on a local disk of the active management server.
4. Issue the `lsrpdomain` command on both management servers to verify that the HA MS domain is not active.

7.5.1 Start HA MS

To start the HA MS, issue the command:

```
hams -s /opt/csm/hams/hamsdef
```

Alternatively, you can start without specifying the hamsdef file. The command will start with the default hamsdef file located in /opt/csm/hams directory.

To start the HA MS with the default hamsdef file issue:

```
hams -s
```

When HA MS successfully starts, the following directories will be moved to the shared file systems (/var/opt/hams/mnt):

- /csminstall
- /var/opt/csm
- /etc/opt/csm

Link files will be created for these directories to the shared file system as follows:

- /csminstall → /var/opt/hams/mnt/csminstall
- /var/opt/csm → /var/opt/hams/mnt/var
- /etc/opt/csm → /var/opt/hams/mnt/etc

Tivoli's System Automation Manager (SAM) takes over control of the shared disks. It monitors the shared disks and controls the HA MS failover.

> **Note:** HA MS can only be started from the primary management server. If you try to start HA MS from the backup management server, you will receive an error message at the console. Example 7-10 shows the message.

Example 7-10 Error message on inactive management server

```
root@ms_02:../csm/hams # hams -s hamsdef
hams: 2658-016 The primary management server ms_01 is not the local machine
ms_02. HA MS must be started on the primary management server, since it has
valid copies of the CSM data. Please run the start command on the primary
management server.
```

7.5.2 Stopping the HA MS

To stop the HA MS, issue the command:

```
hams -S
```

The command stops all HA MS functionality and copies the following files from the shared file systems to the local machine:

- /var/opt/hams/mnt/csminstall
- /var/opt/hams/mnt/var
- /var/opt/hams/mnt/etc

The shared volume group will be varied off.

> **Note:** Execute the command on the active management server only. During the HA MS stop, it copies the CSM data files back to it original location on the local disk. It may copy to the backup MS if it executed on the backup MS.

7.5.3 Verify HA MS is active

To verify the status of the HA MS, issue:

```
hams -l
```

This displays the status of the HA MS. Example 7-11 shows a sample output of the command on a running HA MS cluster.

Example 7-11 hams -l on active HA MS

```
Active MS = ms_01

Displaying Member Resource information:
Class:Resource:Node[ManagedResource]     Mandatory MemberOf     OpState
IBM.Application:Active_CSM               True      Active_MS    Online
IBM.Application:Shared_Disk_Dep          True      Active_MS    Online
IBM.Application:Active_CSM_Req           True      Active_MS    Online
IBM.Application:Shared_Disk              True      Active_MS    Online
IBM.Application:Save_CSM_Data            True      Active_MS    Online
IBM.Application:Active_CSM_Dep           True      Active_MS    Online
IBM.Application:Shared_Disk_Req          True      Active_MS    Online
IBM.Application:Active_VLAN_Connection   True      Active_MS    Online
IBM.Application:Inactive_CSM             True      Inactive_MS  Online
IBM.Application:Gather_CSM_Nodelist      True      Inactive_MS  Online
IBM.Application:Inactive_VLAN_Connection True      Inactive_MS  Online
```

To monitor the resource group status, issue:

```
lsrg -Ab
```

The resource groups for active and inactive management servers will display as shown in Example 7-12 on page 136.

Example 7-12 Display HA MS resource group

```
Displaying Resource Group information:
All Attributes

Resource Group 1:
        Name                              = Active_MS
        MemberLocation                    = Collocated
        Priority                          = 5
        AllowedNode                       = ALL
        NominalState                      = Online
        ExcludedList                      = {}
        ActivePeerDomain                  = CSM_HAMS
        OpState                           = online
        TopGroup                          = Active_MS
        MoveStatus                        = [None]
        ConfigValidity                    =
        AutomationDetails[CompoundState] = Satisfactory

Resource Group 2:
        Name                              = Inactive_MS
        MemberLocation                    = Collocated
        Priority                          = 0
        AllowedNode                       = ALL
        NominalState                      = Online
        ExcludedList                      = {}
        ActivePeerDomain                  = CSM_HAMS
        OpState                           = Online
        TopGroup                          = Inactive_MS
        MoveStatus                        = [None]
        ConfigValidity                    =
        AutomationDetails[CompoundState] = Satisfactory
```

On both management servers, **lsrpdomain** shows that the HA MS domain is active (Example 7-13).

Example 7-13 lsrpdomain on both management servers

```
Name     OpState RSCTActiveVersion MixedVersions TSPort GSPort
CSM_HAMS Online  2.3.4.2           No            12347  1234
```

7.5.4 Manual HA MS failover

To perform a controlled forced HA MS failover, issue this command on the active management server:

```
hams -m
```

This command can be used to verify the functionality of the HA MS after the implementation. It verifies that the HA MS is able to fail over to the backup MS.

To monitor the HA MS failover, use these two commands:

```
hams -l
lsaudrec |grep _HAM
```

When the failover is complete, the output for the **hams -l** command shows the status of the Active MS field changed to backup MS.

Example 7-14 shows the sample output of **hams -l** when the HA MS fails over from active MS (ms_01) to backup MS (ms_02).

Example 7-14 Controlled HA MS failover to backup MS

```
Active MS = ms_02

Displaying Member Resource information:
Class:Resource:Node[ManagedResource]     Mandatory MemberOf     OpState
IBM.Application:Active_CSM               True      Active_MS    Online
IBM.Application:Shared_Disk_Dep          True      Active_MS    Online
IBM.Application:Active_CSM_Req           True      Active_MS    Online
IBM.Application:Shared_Disk              True      Active_MS    Online
IBM.Application:Save_CSM_Data            True      Active_MS    Online
IBM.Application:Active_CSM_Dep           True      Active_MS    Online
IBM.Application:Shared_Disk_Req          True      Active_MS    Online
IBM.Application:Active_VLAN_Connection   True      Active_MS    Online
IBM.Application:Inactive_CSM             True      Inactive_MS  Online
IBM.Application:Gather_CSM_Nodelist      True      Inactive_MS  Online
IBM.Application:Inactive_VLAN_Connection True      Inactive_MS  Online
```

Example 7-15 shows sample output of the **lsaudrec** command when the active MS fails over to the backup management server.

Example 7-15 Lsaudrec output when active MS fails over to the backup MS

```
root@ms_02:/ # lsaudrec|grep _HAM
10/26/04 14:06:32     CLOG Info     [_HAM] Resource "Inactive_VLAN_Connection"
has been started.
10/26/04 14:06:36      CLOG Info     [_HAM] Resource "Inactive_CSM" has been
started.
10/26/04 14:06:42      CLOG Info     [_HAM] Resource "Gather_CSM_Nodelist" has
been started.
10/26/04 14:09:23      CLOG Info     [_HAM] Resource "Active_CSM_Dep" has been
started.
10/26/04 14:09:23     CLOG Info     [_HAM] Resource "Shared_Disk_Dep" has been
started.
```

```
10/26/04 14:09:23      CLOG Info     [_HAM] Resource "Active_VLAN_Connection"
has been started.
10/26/04 14:09:34      CLOG Info     [_HAM] Resource "Shared_Disk" has been
started.
10/26/04 14:09:38      CLOG Info     [_HAM] Resource "Shared_Disk_Req" has been
started.
10/26/04 14:10:45      CLOG Info     [_HAM] Resource "Active_CSM" has been
started.
10/26/04 14:10:48      CLOG Info     [_HAM] Resource "Active_CSM_Req" has been
started.
10/26/04 14:11:10      CLOG Info     [_HAM] Resource "Gather_CSM_Nodelist" has
been started.
10/26/04 14:11:42      CLOG Info     [_HAM] Resource "Save_CSM_Data" has been
started.
10/26/04 14:11:49      CLOG Info     [_HAM] Resource "Gather_CSM_Nodelist" has
been stopped.
10/26/04 14:11:53      CLOG Info     [_HAM] Resource "Inactive_VLAN_Connection"
has been stopped.
```

7.5.5 Active MS shutdown

When the HA MS is active and in the ready state (use **hams -l** to verify), performing a graceful shutdown on the active management server will cause a failover for the CSM resources to the backup management server automatically. The backup management server will take control and become the active MS.

It will remain as the active management server until there is another HA MS failover.

> **Note:** Graceful shutdown could be in any of the following scenarios: **shutdown -Fr**, **shutdown -F**, or **reboot**.

7.5.6 Active MS crash

When the HA MS is active and the active management server crashes or powers off abnormally, the CSM resources will fail over to the backup management server automatically. The backup management server will become the active management server until there is another HA MS failover.

7.5.7 Reintegration

When the failed server reboots, it automatically joins the HA MS domain without user intervention.

However, in the event that the server does not join the domain, verify that the IBM.RecoveryRM resource is active on both management servers:

```
lssrc -a IBM.RecoveryRM
```

To start the resource manually:

```
startsrc -s IBM.RecoveryRM
```

Note: IBM.RecoveryRM must be active for the inactive MS to reintegrate to HA MS domain.

To verify that the management server has reintegrated to the HA MS domain, use the **`hams -l`** and **`lsrpdomain`** commands.

7.5.8 Pause the HA MS

To temporarily stop the HA MS and prevent it from failing over, issue the command:

```
hams -p
```

The function is useful when we need to upgrade the CSM or RSCT on the management server without performing the HA MS stop. If you wish to upgrade RSCT or SAM, you must stop the RPD after pausing HA MS. To do this, run **`stoprpdomain CSM_HAMS`** on either management server. Before restarting HA MS, restart the RPD with **`startrpdomain CSM_HAMS`**.

When HA MS is paused, CSM functionalities can only be used on the active management server as it holds the valid copy of the CSM data.

Note: When HA MS is paused, all HA MS resources are offline and the processes are stopped.

To resume the HA MS, issue the following command on the active management server:

```
hams -s
```

Note: The HA MS can be paused on either primary or backup management servers.

7.5.9 Things to do after an active MS failover

In our test environment, we have a mixture of CSM clients. The CSM clients on different operating systems are tested as shown in Table 7-2 on page 140.

Table 7-2 CSM client version versus operating system

CSM client version	Operating systems
1.1.2.0	AIX 5.1 ML04
1.4.0.2	AIX 5.2 ML04
1.4.0.3-79	SUSE SLES 8
1.4.0.3-79	SUSE SLES 9
1.3.3.3-71	Red Hat EL AS 3

When the HA MS first starts and performs the first failover to the backup MS, the **ssh** authentication is not present on the backup MS. You must run the following command to update and set the remote shell authentication to all nodes:

```
updatenode -avk
```

Note: Root password for each node might be needed if this is the initial access to the nodes. Enter the password when prompted.

We advise checking the status of the nodes in the cluster using the **csmstat** command after the failover. If the status does not return on a particular node, run the following command to fix it:

```
updatenode -vk -n client_node_name
```

In our test, the HA MS software with the CSM client V1.2.0 did not return the status of the node. We need to run the **updatenode** command to fix the **csmstat** status after the failover to the backup MS occurs. It works for all other listed CSM clients.

7.6 HA MS administration

For HA MS administration, see *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918.

The following topics are discussed:

- 7.6.1, “Update HA MS definition file” on page 141
- 7.6.2, “Back up the HA MS definition file” on page 141
- 7.6.3, “HS MS PTF installation” on page 141

7.6.1 Update HA MS definition file

The HA MS definition file (hamsdef) is a text file. It can be updated any time. However, the updated option on the file will not take effect immediately. To load the changes to the HA MS, use:

```
hams -s /opt/csm/hams/hamsdef
```

Note: The HA MS will stop, then reload.

7.6.2 Back up the HA MS definition file

There are two ways to back up the HA MS definition:

1. Back up the /opt/csm/hams/hamsdef file on the active management server.
2. Issue the following command on the active management server and save the output:

   ```
   hams -F
   ```

7.6.3 HS MS PTF installation

The HA MS should not be upgraded when it is running.

There are two ways to install the HA MS PTF on the management server:

- Stop the HA MS.
- Pause the HA MS.

Stop the HA MS and install the HA MS PTF

Issue the HA MS stop command to stop the HA MS before installing the HA MS PTF. When the command is executed, all the CSM data will be copied to the local disk on the active management server.

The HA MS PTF could be installed on both management servers. Restart the HA MS after the installation.

Pause the HA MS for the HA MS PTF installation

Issue the HA MS pause command to pause the HA MS. When the HA MS is paused, stop the RPD with the command **stoprpdomain CSM_HAMS**; then you can install the HA MS PTF on both management servers. Issue the **startrpdomain CSM_HAMS** command, then **hams -s** to resume the HA MS.

We recommend the following prior to any PTF installation or upgrade:

- Perform a csmbackup for CSM database.
- Back up the /opt/csm/hams/hamsdef file.

Note: The HA MS PTF upgrade saves a backup copy of the hamsdef file as /opt/csm/hams/hamsdef.installp_save.

7.7 HA MS log and debug information

The HA MS log files are stored in the /var/log/csm/hams directory on both management servers.

7.7.1 HA MS logs

Whenever **hams** is executed, it creates a hams.log in the /var/log/csm/hams directory. It contains detailed message about the tasks that were performed.

The hams.log file is unique to each management server. It keeps four backups per management server.

Other logs stored in /var/log/csm/hams directory are:

- Active_CSM.log
- Inactive_CSM.log
- Active_VLAN_Connection.log
- Inactive_VLAN_Connection.log
- Shared_Disk.log
- Shared_Disk_monitor.log

7.7.2 Audit log

Another source of activity logging for HA MS troubleshooting is the audit log on each management server. To check the audit log, issue the command:

```
lsaudrec |grep _HAM
```

Example 7-16 shows a sample output.

Example 7-16 A sample output for audit log

```
10/15/04 15:15:08      CLOG Info     [_HAM] Resource "Active_CSM_Dep" has been
started.
10/15/04 15:15:08      CLOG Info     [_HAM] Resource "Shared_Disk_Dep" has been
started.
10/15/04 15:15:08      CLOG Info     [_HAM] Resource "Active_VLAN_Connection"
has been started.
10/15/04 15:15:20      CLOG Info     [_HAM] Resource "Shared_Disk" has been
started.
10/15/04 15:15:24      CLOG Info     [_HAM] Resource "Shared_Disk_Req" has been
```

```
started.
10/15/04 15:15:47      CLOG Info     [_HAM] Resource "Active_CSM" has been
started.
10/15/04 15:15:51      CLOG Info     [_HAM] Resource "Active_CSM_Req" has been
started.
10/15/04 15:16:47      CLOG Info     [_HAM] Resource "Save_CSM_Data" has been
started.
```

7.7.3 Debugging tips

HA MS logs are stored on both management servers. When troubleshooting a problem, we have to check the logs on both servers. The logs files are located in the /var/log/csm/hams directory.

Steps to determine the problem

Here are some simple procedures for troubleshooting problems:

1. If the HA MS start fails, check the /var/log/csm/hams/hams.log on the management server.
2. Restart the hams using the **`hams -sv`** command. The command displays verbose messages to standard output.

   > **Note:** You can run the **`hams -s`** command many times, but for each execution of the command, it will run the HA MS stop command followed by the HA MS start command.

3. Issue the **`hams -l`** command to verify that the HA MS resources are online. If any of the resources are offline, run the **`hams -lv`** command to determine the state of the resources.

   Example 7-17 shows a sample of the output.

Example 7-17   hams -lv sample output

```
root@ms_01:../csm/hams # hams -lv
Running cmd: CT_MANAGEMENT_SCOPE=1 lsrsrc-api -i -s
IBM.DmsCtrl::::RemoteShell::SetupRemoteShell::RemoteCopyCmd::HAMode::SetupKRB5
2>&1
Running cmd: /bin/lslpp -Lcq bos.rte | /usr/bin/cut -d: -f 3 2>&1
Loading /opt/csm/install/defs/AIX.pm.
Running cmd: /usr/sbin/ifconfig -a
Running cmd: CT_MANAGEMENT_SCOPE=2 /usr/bin/lsrsrc-api -D ':|:' -i -s
IBM.Application::"Name IN ('Active_CSM')"::Name::NodeNameList::OpState 2>&1
Active MS =
Name     OpState RSCTActiveVersion MixedVersions TSPort GSPort
CSM_HAMS Online  2.3.4.2           No            12347  12348
```

```
Name  OpState RSCTVersion
ms_02 Online  2.3.4.2
ms_01 Online  2.3.4.2

Displaying Member Resource information:
Class:Resource:Node[ManagedResource]     Mandatory MemberOf     OpState
IBM.Application:Active_CSM               True      Active_MS    Offline
IBM.Application:Active_CSM_Dep           True      Active_MS    Offline
IBM.Application:Shared_Disk_Req          True      Active_MS    Offline
IBM.Application:Active_VLAN_Connection   True      Active_MS    Offline
IBM.Application:Active_CSM_Req           True      Active_MS    Offline
IBM.Application:Shared_Disk              True      Active_MS    Offline
IBM.Application:Save_CSM_Data            True      Active_MS    Offline
IBM.Application:Shared_Disk_Dep          True      Active_MS    Offline
IBM.Application:Inactive_CSM             True      Inactive_MS  Online
IBM.Application:Gather_CSM_Nodelist      True      Inactive_MS  Online
IBM.Application:Inactive_VLAN_Connection True      Inactive_MS  Online
Resource Persistent and Dynamic Attributes for: IBM.Application
Name::NodeNameList::OpState::
"Active_CSM"::{"ms_01","ms_02"}::2::
"Active_CSM"::{"ms_01"}::2::
"Active_CSM"::{"ms_02"}::2::
"Inactive_CSM"::{"ms_01","ms_02"}::1::
"Inactive_CSM"::{"ms_01"}::2::
"Inactive_CSM"::{"ms_02"}::1::
"Active_CSM_Dep"::{"ms_01","ms_02"}::2::
"Active_CSM_Dep"::{"ms_01"}::2::
"Active_CSM_Dep"::{"ms_02"}::2::
"Active_CSM_Req"::{"ms_01","ms_02"}::2::
"Active_CSM_Req"::{"ms_01"}::2::
"Active_CSM_Req"::{"ms_02"}::2::
"Shared_Disk_Dep"::{"ms_01","ms_02"}::2::
"Shared_Disk_Dep"::{"ms_01"}::2::
"Shared_Disk_Dep"::{"ms_02"}::2::
"Shared_Disk"::{"ms_01","ms_02"}::2::
"Shared_Disk"::{"ms_01"}::2::
"Shared_Disk"::{"ms_02"}::2::
"Shared_Disk_Req"::{"ms_01","ms_02"}::2::
"Shared_Disk_Req"::{"ms_01"}::2::
"Shared_Disk_Req"::{"ms_02"}::2::
"Save_CSM_Data"::{"ms_01","ms_02"}::2::
"Save_CSM_Data"::{"ms_01"}::2::
"Save_CSM_Data"::{"ms_02"}::2::
"Public_IP_Dep"::{"ms_01","ms_02"}::2::
"Public_IP_Dep"::{"ms_01"}::2::
"Public_IP_Dep"::{"ms_02"}::2::
"Public_IP_Req"::{"ms_01","ms_02"}::2::
```

The example shows that the resources are offline and includes the details of each resource group. Each resource group should have a state of 1 and 2 for active MS and backup MS, respectively.

Any resource group that exhibits problems will show a state of 2 on both management servers or a state of 3 on either management server. The status is indicated on the last column of each resource.

4. To monitor each resource, you can use the command:

   ```
   /opt/csm/csmbin/hams/resource_name -mv
   ```

   The resource output is logged to the /var/log/csm/hams/resource_name log.

5. To restart any resource, you can use the following command with CT_MANAGEMENT_SCOPE set to the HA MS environment:

   ```
   CT_MANAGEMENT_SCOPE=2 resetrsrc -s “Name IN ‘resource_name’” -V\
   IBM.Application
   ```

   If the environment is not set, the **resetrsrc** command will not start the HA MS resource.

How to recover the HA MS if stopped on the backup MS

In our scenarios, we stop the HA MS on the backup management server. The HA MS stopped successfully, but the CSM data was not copied to the local disk. The hamsdef file does not contain the /csminstall information for the backup MS. HA MS stops and the following directories are still linked to the shared storage disk:

- /csminstall → /var/opt/hams/mnt/csminstall
- /var/opt/csm → /var/opt/hams/mnt/var
- /etc/opt/csm → /var/opt/hams/mnt/etc

When **hams -S** is executed successfully as shown in Example 7-18, the /csminstall directory still links to the shared file system.

Example 7-18 csminstall directory

```
lrwxrwxrwx   1 root     system          29 Oct 27 08:18 csminstall ->
/var/opt/hams/mnt/csminstall/
```

Two ways to recover from this scenario:

- Starting HA MS on the primary management server
- Manual recovery

Note: Manual recovery is not recommended. Only use it if the HA MS is not able to do a self-recovery.

Start HA MS on the primary management server

We can start HA MS on the management server, and let HA MS perform the automatic recovery.

When the **hams -s** command is issued on the active management server (ms_02), the command fails. This is shown in Example 7-19.

Example 7-19 Start hams on ms_02 management server

```
root@ms_02:../hams/resources # hams -s /opt/csm/hams/hamsdef
hams: 2658-016 The primary management server ms_01 is not the local machine
ms_02. HA MS must be started on the primary management server, since it has
valid copies of the CSM data. Please run the start command on the primary
management server.
```

It failed, as we have defined the primary MS to be ms_01 and backup MS to be ms_02 in our HA MS definition file.

Run the **hams -s** command on ms_01 management server. HA MS will start and bring back the resource to the primary MS (ms_01).

We issue the **hams -l** command to verify the HA MS status on ms_01 management server as shown in Example 7-20.

Example 7-20 HA MS status on ms_01 management server

```
root@ms_01:/ # hams -l
Active MS = ms_01

Displaying Member Resource information:
Class:Resource:Node[ManagedResource]     Mandatory MemberOf     OpState
IBM.Application:Shared_Disk_Dep          True      Active_MS    Online
IBM.Application:Shared_Disk_Req          True      Active_MS    Online
IBM.Application:Active_VLAN_Connection   True      Active_MS    Online
IBM.Application:Active_CSM_Dep           True      Active_MS    Online
IBM.Application:Shared_Disk              True      Active_MS    Online
IBM.Application:Save_CSM_Data            True      Active_MS    Online
IBM.Application:Active_CSM_Req           True      Active_MS    Online
IBM.Application:Active_CSM               True      Active_MS    Online
IBM.Application:Inactive_VLAN_Connection True      Inactive_MS  Online
IBM.Application:Gather_CSM_Nodelist      True      Inactive_MS  Online
IBM.Application:Inactive_CSM             True      Inactive_MS  Online
```

We can stop the HA MS on ms_01. HA MS will stop and copy CSM data to its local directories.

Manual recovery

Perform the following steps manually:

1. On ms_02, unmount the shared file system:

   ```
   umount /var/opt/hams/mnt/
   ```

2. On ms_02, deactivate the shared volume group:

   ```
   varyoffvg csmvg
   ```

3. On ms_01, remove the following linked files:

   – /csminstall

   – /var/opt/csm

   – /etc/opt/csm

4. On ms_01, activate the shared volume group:

   ```
   varyonvg csmvg
   ```

5. On ms_01, mount the shared file system:

   ```
   mount /var/opt/hams/mnt/
   ```

6. On ms_01, mount the /csminstall local directory

   ```
   mount /csminstall
   ```

7. On ms_01, copy the following data from the shared file system:

   ```
   cp -rp /var/opt/hams/mnt/csminstall/* /csminstall/
   cp -rp /var/opt/hams/mnt/var/* /var/opt/csm/
   cp -rp /var/opt/hams/mnt/etc/* /etc/opt/csm/
   ```

8. On ms_01, restart the IBM.HWCTRLRM daemon:

   ```
   stopsrc -s IBM.HWCTRLRM
   startsrc -s IBM.HWCTRLRM
   ```

9. On ms_01, perform an **updatenode** command to gather client information:

   ```
   updatenode -ak
   ```

Note: Step 9 may not be needed. Execute **csmstat -a** to verify the cluster status; if the status is not shown, then execute the **updatenode** command.

7.8 High Availability Management Server Q&A

In this section, we answer some questions regarding the High Availability Management Server. Not all of the questions have an answer, but we provide answers to as many questions as possible.

1. *Question:* How much of the CSM install of the backup management server do I have to do?

Answer: When installing the backup management server for HA MS, you only need to install the CSM server software and set the csmconfig attributes to be the same as the primary management server's csmconfig attributes. HA MS will failover all of the CSM data and functionality when needed.

2. *Question*: I store the /csminstall file system in a separate logical volume on my AIX management server. Are there any special considerations for this environment?

Answer: Yes, you should set the CSMInstallPartition attribute in the hamsdef file to the device name of the logical volume (for example, CSMInstallPartition=/dev/csmlv). Note that the *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918, indicates that CSMInstallPartition is a Linux-only attribute; however, it has been expanded in CSM HA MS 1.4.0.1 to be used as an AIX attribute as well. Before you start the HA MS, make sure there are no running processes in the /csminstall file system with the **fuser** command because HA MS will have to unmount it to move the data to the shared disk. Also make sure that the /csminstall file system is not automatically mounted during the OS reboot as this will interfere with the HA MS.

3. *Question:* What is IPAT and what values do I put in the Public* attributes in the hamsdef file?

Answer: IPAT stands for IP address takeover. HA MS can perform IPAT of a public IP address of your management server as long as it is done on adapters that are not used to connect to the CSM cluster or management VLANs. HA MS IPAT functionality is performed with IP aliasing. The PublicIPAddress attribute in the hamsdef should be filled in with the IP address that you wish to alias onto the adapter of the active management server (not the base address of the adapter, but the one that you want to transfer between the machines). The PublicIPNetmask attribute is the netmask of the PublicIPAddress, and the PublicIPAdapter is the adapter the IP address should be aliased onto. When HA MS fails over the CSM functionality, the IP in the PublicIPAddress attribute will be aliased onto the new active management server; this way, users of the CSM management server do not need to know which is the active management server and they can just connect to the aliased IP address.

4. *Question:* What mount point should I create for the CSM shared data file system?

Answer: The file system should have a mount point of /var/opt/hams/mnt but it should not be automatically mounted during an OS reboot on either management server.

5. *Question:* I just created the volume group, logical volume, and file system for the CSM data shared disk on the primary management server. How do I get this defined on the backup management server?

Answer: You can import the volume group on the backup management server with the **`importvg -y <VolumeGroup> <PhysicalVolume>`** command. You should ensure that you can mount the file system on both management servers (not at the same time) before starting HA MS.

6. *Question:* Which management server should I start HA MS on? Does it matter?

Answer: It does matter which management server you start HA MS on. You should start HA MS on the machine with the valid CSM data. This is normally the primary management server, or the management server that you last stopped HA MS on.

7. *Question:* HA MS did not start correctly, what do I do now?

Answer: First try to determine why the HA MS had trouble starting. Save off the HA MS log files that are stored in the /var/log/csm/hams/ and the output of **`lsaudrec|grep _HAM`** on each management server. Then stop HA MS on the same machine that you tried to start it on (if the HA MS says that it is not configured, that is fine). Make sure HA MS is stopped completely before trying to restart it. One of the most common problems with HA MS startup involves running processes in a file system that HA MS is trying to unmount. If this problem occurs, kill the processes with opened files, stop HA MS, and then restart it on the same machine.

8. *Question:* HA MS is in an unknown state, and I cannot run CSM commands on either management server, what should I do?

Answer: First, save the HA MS log files that are stored in /var/log/csm/hams/ and the output of **`lsaudrec | grep _HAM`** on each management server. Then you should try to stop HA MS on the machine that was last the active management server. If you have trouble stopping HA MS and cannot run CSM commands, you can override the block on CSM by exporting the following environment variable CSM_HAMS_CONTROL=1, and running **`csmconfig HAMode=0`**. If you need hardware control, you will have to restart IBM.HWCTRLRM with the following commands:

- `stopsrc -s IBM.HWCTRLRM`
- Kill any leftover Java processes that show up in the following command:

  ```
  /bin/ps -eo pid,args | grep hcdaemon
  ```
- `startsrc -s IBM.HWCTRLRM`
- After you diagnose and fix the problem, stop HA MS on the machine that you want to be the active management server (you may have to set **csmconfig HAMode=3** first if **hams -Sv** says that HA MS is not configured), and then restart HA MS on the same machine.

9. *Question:* The remote shell is not set up between the inactive management server and the nodes when I run **updatenode -k**. What is wrong?

Answer: HA MS will only set up the remote shell from the inactive management server to the nodes when it sets up the remote shell for the active management server. This means that running **updatenode -k** will not always set up the remote shell for this environment because it does not attempt remote shell setup for the current management server unless remote shell authentication is not working. If this is the case, you can break the remote shell for the active management server, and then rerun **updatenode -k** to set it up for both management servers, or just run **updatenode -k** after failover. If you are having other problems with remote shell in the HA MS environment, ensure you have the same remote shell settings in csmconfig on both management servers.

10. *Question:* What happens with the nodes when the primary management server fails?

Answer: When the primary MS fails, the backup MS issues the **mgmtsvr** command to all of the nodes. The backup MS takes over the management of the nodes in the cluster. The down side is that if a node is not up and running at the moment the secondary MS issues the **mgmtsrv** command, the node is not updated, and when the node comes up it “thinks” that its MS has failed. If the number of nodes is high or the connection to the nodes is slow, the migration could take several minutes. If the node is not updated during the MS failover, manual intervention is needed. The node has to be updated manually by issuing from the MS the **updatenode** command.

A

Simple user management and authentication via LDAP

In this appendix, we describe how to setup User Management and Authentication using the IBM version of the Lightweight Directory Access Protocol (LDAP) product called IBM Directory Server Version 4.1.

We describe how to set up the LDAP server on AIX 5.2, the DB2® backend database that the LDAP servers use, and how to set up clients on AIX and Linux. All the software to set up the AIX LDAP environment can be found on the AIX 5.2 CDs. Linux will use the openLDAP client that is installed with most distributions.

It is often essential within a clustered environment to keep the UNIX user IDs and groups (uids/gids) and passwords in sync between the nodes. This appendix details the setup of a simple LDAP server that enables this to be done. LDAP can be used for storing many things, but we restrict it to a simple user management and authentication within the cluster. We use the simple authentication method for communication between the client and server because of the closed nature of private cluster networks and to keep the setup example simple.

> **Note:** The procedures that are outlined in this appendix apply only to IBM Directory Server 4.1 and AIX 5.2. Newer versions of this software have different install procedures.

Installation and configuration of AIX Directory Server

This section describes how to install the LDAP software and the DB2 backend software. We decided to install the LDAP server onto our CSM management server that has the host name of ms_01.

Later, we plan to set up peer-to-peer replication to our second CSM management server with the host name of ms_02. This will allow for continued availability of the LDAP environment if one of the CSM management servers fails.

Our AIX and Linux nodes will be clients of these LDAP servers.

Install and configure LDAP on a CSM management server

Before installing the LDAP file sets, we ensured that the AIX install is “clean”: If any remains of a previous LDAP/DB2 install are hanging around, then the **mksecldap** command might fail. In fact, it is worse because the installation will appear to work fine, but the resulting LDAP server will not function.

Install file sets

Install the ldap.client and ldap.server file sets from the AIX 5.2 CDs. This also installs the DB2 prerequisite file sets automatically.

Note: We installed the ldap.client as well as the ldap.server onto the CSM management server because we wished the CSM management server to be an LDAP client of its own LDAP server.

Use the **smitty install** command to install the LDAP file sets. Select just the ldap.client and ldap.server file sets. All other file sets that are needed by these two file sets will be installed automatically. Example A-1 on page 153 shows which file sets will be installed.

Example: A-1 LDAP file sets installed on the server

```
  ldap.client.adt            4.1.0.0  COMMITTED  IBM Directory Client SDK
  ldap.client.dmt            4.1.0.0  COMMITTED  IBM Directory Client DMT
  ldap.client.java           4.1.0.0  COMMITTED  IBM Directory Client Java
  ldap.client.rte            4.1.0.0  COMMITTED  IBM Directory Client Runtime
  ldap.html.en_US.config     4.1.0.0  COMMITTED  IBM Directory Install/Config
  ldap.html.en_US.man        4.1.0.0  COMMITTED  IBM Directory Man Pages - U.S.
  ldap.msg.en_US             4.1.0.0  COMMITTED  IBM Directory Messages - U.S.
  ldap.server.admin          4.1.0.0  COMMITTED  IBM Directory Server
  ldap.server.cfg            4.1.0.0  COMMITTED  IBM Directory Server Config
  ldap.server.com            4.1.0.0  COMMITTED  IBM Directory Server Framework
  ldap.server.rte            4.1.0.0  COMMITTED  IBM Directory Server Runtime
  ldap.client.rte            4.1.0.0  COMMITTED  IBM Directory Client Runtime
  ldap.server.admin          4.1.0.0  COMMITTED  IBM Directory Server
  ldap.server.cfg            4.1.0.0  COMMITTED  IBM Directory Server Config
  ldap.server.com            4.1.0.0  COMMITTED  IBM Directory Server Framework

  db2_07_01.client          7.1.0.40  COMMITTED  Client Application Enabler
  db2_07_01.cnvucs          7.1.0.40  COMMITTED  Code Page Conversion Tables -
  db2_07_01.conn            7.1.0.40  COMMITTED  Connect
  db2_07_01.conv.jp         7.1.0.40  COMMITTED  Code Page Conversion Tables -
  db2_07_01.conv.kr         7.1.0.40  COMMITTED  Code Page Conversion Tables -
  db2_07_01.conv.sch        7.1.0.40  COMMITTED  Code Page Conversion Tables -
  db2_07_01.conv.tch        7.1.0.40  COMMITTED  Code Page Conversion Tables -
  db2_07_01.cs.drda         7.1.0.40  COMMITTED  Communication Support - DRDA
  db2_07_01.cs.ipx          7.1.0.40  COMMITTED  Communication Support - IPX
  db2_07_01.cs.rte          7.1.0.40  COMMITTED  Communication Support - TCP/IP
  db2_07_01.cs.sna          7.1.0.40  COMMITTED  Communication Support - SNA
  db2_07_01.das             7.1.0.40  COMMITTED  Administration Server
  db2_07_01.db2.engn        7.1.0.40  COMMITTED  Engine
  db2_07_01.db2.rte         7.1.0.40  COMMITTED  Run-time Environment
  db2_07_01.db2.samples     7.1.0.40  COMMITTED  Sample Database Source
  db2_07_01.elic            7.1.0.40  COMMITTED  Product Signature for UDB
  db2_07_01.tspf            7.1.0.40  COMMITTED  Transformer Stored Procedure
```

Configure the LDAP server

After installing the LDAP and DB2 file sets, we configured the LDAP server. The LDAP server and AIX clients are configured using the **mksecldap** command.

When the **mksecldap** command is used to set up the LDAP server, it creates a DB2 instance called ldapdb2. It also creates a DB2 database called ldapdb2. It then populates the database with the LDAP tree and, optionally, migrates any users to the LDAP server. The **mksecldap** command sets up an AIX user ID called ldapdb2 with a home directory of /home/ldapdb2. This is where the LDAP DB2 database is stored.

We created the LDAP server using the command:

```
mklsecldap -s -a cn=root -p password -S rfc2307aix
```

Table A-1 explains some of the flags used with the **mksecldap** command.

Table A-1 LDAP server creation flags with mksecldap

Flag	Description
-s	This flag tells mksecldap to set up an LDAP server.
-a	This is the admin user ID domain name you wish to use to administer the LDAP server. We chose cn=root, but you can use any other IDs, such as cn=admin or cn=foo. This user ID does not have to exist in AIX.
-p	The password to use for the admin user created with the -a flag.
-S	The schema type to load to the LDAP database. There is the choice of three schemas: 1. AIX: This supports only AIX clients. 2. rfc2307: This supports only POSIX clients. 3. rfc2307aix: This supports POSIX and AIX clients. We chose schema rfc2307aix to support both AIX and Linux clients.

Example A-2 shows the output from the execution of the **mksecldap** command.

Example: A-2 Output from the mksecldap command

```
# mksecldap -s -a cn=root -p password -S rfc2307aix
 Password for administrator DN cn=root has been set.

IBM Directory Server Configuration complete.
 Creating the directory DB2 default database.
 This operation may take a few minutes.

Configuring the database.
Creating database instance: ldapdb2.
Created database instance: ldapdb2.
Starting database manager for instance: ldapdb2.
Started database manager for instance: ldapdb2.
Creating database: ldapdb2.
Created database: ldapdb2.
Updating configuration for database: ldapdb2.
Updated configuration for database: ldapdb2.
Completed configuration of the database.

IBM Directory Server Configuration complete.
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type PREOPERATION is successfully loaded from libDSP.a.
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type AUDIT is successfully loaded from /lib/libldapaudit.a.

Plugin of type AUDIT is successfully loaded from
/usr/ccs/lib/libsecldapaudit.a(shr.o).
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type DATABASE is successfully loaded from /lib/libback-rdbm.a.
Non-SSL port initialized to 389.
Local UNIX socket name initialized to /tmp/s.slapd.
#
```

Setup of the Web interface

We are running a small Apache Web server on our CSM management server. To assist with the management of the LDAP server, we configured the Apache Web server with the **ldapcfg** command:

```
ldapcfg -s apache -f /usr/local/apache/conf/httpd.conf
```

After the **ldapcfg** command was run and Apache restarted, we could access the directory server administration Web page at:

```
http://ms_01/ldap/index.html
```

The setup of the first LDAP server is complete.

Setup of AIX LDAP client

We now make our AIX nodes clients of the LDAP servers on ms_01 and ms_02.

To set up as an LDAP client, we first installed the ldap.client file set onto the node. We then issued the **mksecldap** command to configure the client. We issued this command on each of our AIX nodes:

```
# mksecldap -c -h ms_01 -a cn=root -p password -t 0
#
```

Table A-2 shows details of the **mksecldap** flags that we used in the creation of the AIX LDAP client.

Table A-2 LDAP client creation using the mksecldap command

Flag	Description
-c	Create a client.

Flag	Description
-h	The host names of the LDAP servers.
-a	The administrative user domain name to bind to the LDAP server.
-p	The administrative user's password.
-t	Cachetimeout is disabled. We disabled the cachetimeout parameter by setting it to 0 (zero). This enables the LDAP client to see changes made to the LDAP server more quickly. The default value is 5 minutes.

After we set up the LDAP client, we configured AIX to use the LDAP server for user authentication. We edited the /etc/security/user file and changed the default stanza SYSTEM entry as shown in Example A-3. We left the other lines in the default: stanza untouched.

Example: A-3 Modify /etc/security/user

```
# vi /etc/security/user
default:

        SYSTEM = "compat or LDAP"
```

Tip: If OpenSSH is being used on the client, you must restart the sshd daemon to allow OpenSSH to authenticate against the LDAP server. Restart sshd by executing:

```
# stopsrc -s sshd
# startsrc -s sshd
```

The AIX LDAP client setup should now be complete.

Test the AIX LDAP server and AIX client

We tested our newly created AIX LDAP directory server and client by running a few simple tests.

Setup user

We created a new user by issuing the **`mkuser -R LDAP testuser`** command.

Set passwd

We then set a password for the new user: **`passwd -R LDAP testuser`**.

Note: The user will be prompted to change their password at the first logon attempt on AIX. This is because the passwordFlag field within the LDAP server has been set to ADMCHG. Linux does not appear to honor this flag, so a Linux user will not be prompted to change the password on the first logon.

In traditional file-based user definitions, the ADMCHG flag can be removed by using the **`pwdadm -c`** command. As of AIX 5.3, the **`pwdadm`** command supports the -R flag, so this flag can be reset on AIX 5.3 systems with the **`pwdadm -R LDAP -c`** command, if so desired.

List user

We then listed the user with the **`lsuser testuser`** command. Of interest is the registry field: This should show `LDAP`.

ssh login

We then logged on with the new user via **`ssh`** by issuing the command: **`ssh testuser@ms_01`**.

rmuser

We then deleted the user from the LDAP server: **`rmuser -R LDAP testuser`**.

Setup of Red Hat Linux client

To enable LDAP authentication on our pSeries Red Hat Enterprise Linux Advance Server nodes, we utilize the openLDAP client that is installed with this distribution. Check whether the client is installed by issuing the **rpm** command as shown in Example A-4.

Example: A-4 Determine whether openldap client is installed

```
$ rpm -qa|grep openldap
openldap-2.0.27-17
openldap-clients-2.0.27-17
openldap-devel-2.0.27-17
$
```

We set up the Linux node to use the LDAP server by using the **authconfig** command. This is the command we issued to configure the client:

```
# authconfig --enableldap --enableldapauth --ldapserver ms_01,ms_02 \
--ldapbasedn cn=aixdata --kickstart
```

Note: We used the --kickstart option of the **authconfig** command because this suppresses the authconfig menu being launched.

The authconfig setup updates the /etc/ldap.conf, /etc/pam.d/system-auth, and /etc/nssswitch.conf files to enable use of the LDAP server. Although **authconfig** does a good job in setting up the LDAP client side of things, it does not quite do everything so a few changes still have to be made.

Changes to /etc/ldap.conf

The statements shown in Example A-5 were added to the /etc/ldap.conf file.

Example: A-5 Statements added to /etc/ldap.conf

```
# The distinguished name to bind to the server with.
# Optional: default is to bind anonymously.
binddn cn=root

# The credentials to bind with.
# Optional: default is no credential.
bindpw password

# The distinguished name to bind to the server with
# if the effective user ID is root. Password is
# stored in /etc/ldap.secret (mode 600)
rootbinddn cn=root
pam_password crypt
```

As can be seen in Example A-5 on page 158, there is a comment regarding the /etc/ldap.secret file. This file has to be created and its file access characteristics changed using **chmod** 600. The contents of this file is the administrative domain name password. One other interesting note is that the `pam_password` line must be set to `crypt`. The default is md5, which is not supported by AIX for password encryption. If this option is not set to crypt and a user changes his password in Linux, the password will be invalid on an AIX system because it would have been stored in the LDAP server database with md5 encryption.

PAM/NSS issues

There is an issue with LDAP authentication with the interaction of PAM and NSS. The high-level description of the problem is that if, for whatever reason, the LDAP servers are unavailable to the Linux system, no user will be able to log on to the system, including users who are defined in the local system files (/etc/passwd).

For more information about this issue, refer to:

http://meltin.net/people/martin/publications/polythenepam.html

The workaround to this problem is to modify PAM by adding a plug-in to the /etc/pam.d/system-auth file. This plug-in is called pam_localuser.so. Example A-6 shows the modified /etc/pam.d/system-auth file we used on our pSeries Red Hat Linux systems.

Example: A-6 /etc/pam.d/system-auth file

```
#%PAM-1.0
# This file is auto-generated.
# User changes will be destroyed the next time authconfig is run.
auth        required      /lib/security/$ISA/pam_env.so
auth        sufficient    /lib/security/$ISA/pam_unix.so likeauth nullok
auth        sufficient    /lib/security/$ISA/pam_ldap.so use_first_pass
auth        required      /lib/security/$ISA/pam_deny.so

account     required      /lib/security/$ISA/pam_unix.so
account     sufficient    /lib/security/$ISA/pam_localuser.so
account     required      /lib/security/$ISA/pam_ldap.so

password    required      /lib/security/$ISA/pam_cracklib.so retry=3 type=
password    sufficient   /lib/security/$ISA/pam_unix.so nullok use_authtok md5
shadow
password    sufficient    /lib/security/$ISA/pam_ldap.so use_authtok
password    required      /lib/security/$ISA/pam_deny.so

session     required      /lib/security/$ISA/pam_limits.so
session     required      /lib/security/$ISA/pam_unix.so
session     optional      /lib/security/$ISA/pam_ldap.so
```

Tip: Make sure there is a working copy of the /etc/pam.d/system-auth file before making any changes. Any error in this file may disable all user access to the Linux system.

Note: Running the **`authconfig`** command again to completion will wipe out any manual changes that have been made to /etc/pam.d/system-auth.

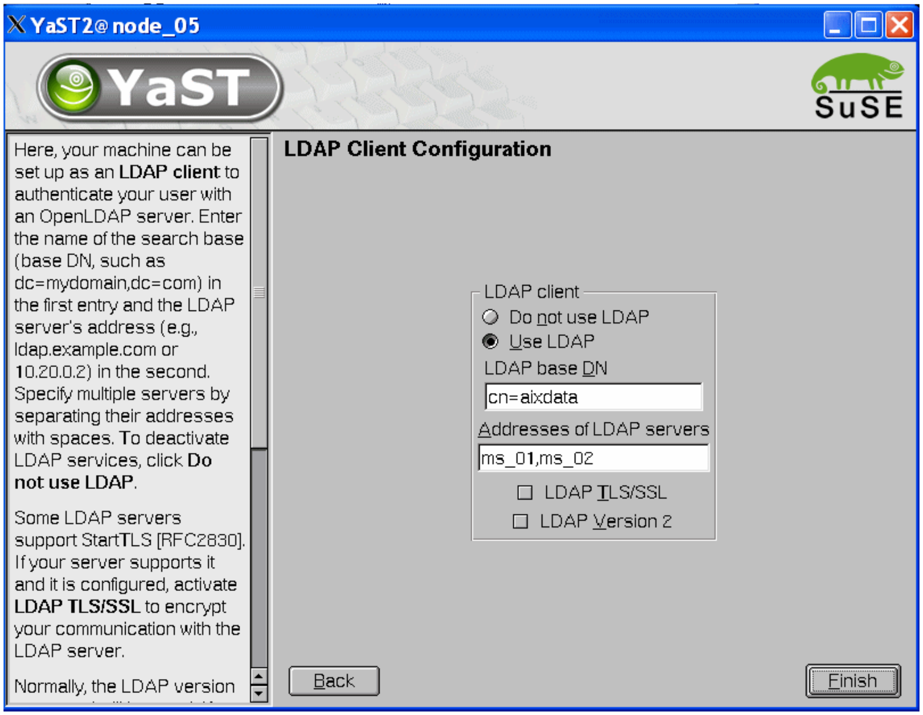
The setup of the LDAP client on Red Hat Linux is now complete. We tested the client by listing all users that the system can resolve by issuing the **`getent passwd`** command. All local and LDAP users will be listed if the setup is correct.

Note: There are weaknesses in the security of some of the files configured (/etc/ldap.conf on linux). You may want to set up SASL authentication mechanisms if stronger security is desired.

Setup of SUSE LINUX client

The setup of the LDAP client on SUSE differs from the Red Hat distribution. The **`yast2`** tool is used to initially configure the LDAP client, and some changes are required to the /etc/openldap/ldap.cfg file.

We started the **`yast2`** tool and selected **Network/Advanced** → **LDAP client**. Figure A-1 on page 161 shows the LDAP client configuration screen.


YaST2@ node_05

Here, your machine can be set up as an **LDAP client** to authenticate your user with an OpenLDAP server. Enter the name of the search base (base DN, such as dc=mydomain,dc=com) in the first entry and the LDAP server's address (e.g., ldap.example.com or 10.20.0.2) in the second. Specify multiple servers by separating their addresses with spaces. To deactivate LDAP services, click **Do not use LDAP**.

Some LDAP servers support StartTLS [RFC2830]. If your server supports it and it is configured, activate **LDAP TLS/SSL** to encrypt your communication with the LDAP server.

Normally, the LDAP version

LDAP Client Configuration

LDAP client

○ Do not use LDAP

● Use LDAP

LDAP base DN: cn=aixdata

Addresses of LDAP servers: ms_01,ms_02

□ LDAP TLS/SSL

□ LDAP Version 2

Figure A-1 Configuration of LDAP client with yast2

After the LDAP client setup is configured, some modifications will have to be made to the /etc/openldap/ldap.conf file. Example A-7 details these changes.

Note: On SUSE LINUX SLES 9, the /etc/ldap.conf file requires modifying as per Example A-5 on page 158.

Example: A-7 Additions to /etc/openldap/ldap.conf

```
binddn  cn=root
bindpw  itsoadmin
rootbinddn      cn=root
pam_password    crypt
```

The /etc/ldap.secret file also must be created containing the administrative user ID domain name password. This file access is changed with the command **chmod** 600.

> **Note:** There are weaknesses in the security of some of the configured files (/etc/ldap.conf on Linux). You may want to set up SASL authentication mechanisms if stronger security is desired.

Configuration of LDAP server replicas

We want to set up a replica of our LDAP server database because it will protect the LDAP environment from failure if one machine fails. There are two ways to set up replication between two LDAP servers: master/slave or peer-to-peer. We decided to use peer-to-peer because the master/slave setup can have issues sharing updates if the master server is down.

Install LDAP server on the second machine (ms_02)

We installed our second LDAP server onto our second CSM management server, ms_02. First, we ensured that there is not an existing LDAP client configured on ms_02. If an LDAP client is active, it will interfere with the running of **mksecldap**.

The second thing we checked was to ensure that the /home directory was not shared with ms_01. The LDAP server database resides in /home/ldapdb2 and if this directory already exists (for example, if the /home directory is NFS-shared), the **mksecldap** command will fail.

We installed the second LDAP server following the process detailed in “Configure the LDAP server” on page 153. We made sure we used the same values for the -a and -p flags and included the -u NONE flag for the second LDAP server.

> **Important:** When creating the second LDAP server, be sure to use the -u NONE option of the **mksecldap** command. Users will be migrated from the first LDAP server in later steps.

Stop the LDAP server on the second machine

We stopped the LDAP server (the slapd daemon) on ms_02, but left the DB2 tasks running by issuing this command:

```
# ps -efa|grep slapd
ldap 434392      1   2 08:20:10      - 16:10 /bin/slapd -f/etc/slapd32.conf
# kill 434292
```

Update the slapd configuration on the first machine (ms_01)

We then updated the /etc/slapd32.conf file on ms_01 as per Example A-8 on page 163. This stanza can be put at the end of the file.

Example: A-8 Update /etc/slapd32.conf on machine one

```
dn: cn=Master Server, cn=Configuration
cn: Master Server
ibm-slapdPeerDn: cn=peer
ibm-slapdPeerPW: password_ms_01
objectclass: ibm-slapdReplication
objectclass:top
```

Note: The password specified in ibm-slapdPeerPW marries up with the stanza statement replicaCredentials that is added to the second LDAP server in Example A-11 on page 164.

Update LDAP server on first machine

We then had to configure the LDAP server on ms_01. We placed the statements as shown in Example A-9 into a temporary file, then loaded the file to the LDAP server on ms_01 using the **ldapadd** command. We first added the statements to a file called /tmp/ms_01.ldif.

Note: The replicaCredentials in Example A-9 shows the password on the ms_02 system. This marries up with the ibm-slapdPeerPW password in the /etc/slapd32.conf on ms_02, which is updated in Example A-10 on page 164.

Example: A-9 LDIF that is added to machine one (ms_01)

```
dn: cn=ms_02, cn=localhost
cn: ms_02
objectclass: replicaObject
replicabindDN: cn=peer
replicaCredentials: password_ms_02
replicaPort: 389
replicaHost: ms_02

#ldapadd -f /tmp/ms_01.ldif -D cn=root -w password
#
```

Copy users from ms_01 to ms_02 LDAP server

To populate the second LDAP server on ms_02, we must export the users from the first LDAP server and load to them to the second LDAP server. We use **db2ldif** and **ldfi2db** to achieve this.

We executed the following steps:

- On ms_01 we ran the command:

```
db2ldif -o /tmp/ldap.ldif
```

- We then transferred the file /tmp/ldap.ldif to ms_02.
- On ms_02 we ran the command:

```
ldif2db -i /tmp/ldap.ldif
```

Update the slapd configuration on the second machine

We then updated the /etc/slapd32.conf file on ms_02 as in Example A-10. This stanza can be put at the end of the file.

> **Note:** The password that is specified in ibm-slapdPeerPW will marry up with the stanza that is added to the first LDAP server in Example A-9 on page 163

Example: A-10 Update /etc/slapd32.conf on machine two

```
dn: cn=Master Server, cn=Configuration
cn: Master Server
ibm-slapdPeerDn: cn=peer
ibm-slapdPeerPW: password_ms_02
objectclass: ibm-slapdReplication
objectclass: top
```

Update the LDAP server on the second machine

We then had to configure the LDAP server on ms_02 with which machine to replicate to. We placed the statements in Example A-11 into a temporary file and loaded the file to the LDAP server on machine one using the **ldapadd** command. Before running **ldapadd**, we started the LDAP server on ms_02 by using the **slapd** command. We added the statements to a file called /tmp/ms_02.ldif. Example A-11 also shows the LDIF statements and Example A-12 on page 165 shows how we started the LDAP server and added the LDIF statements.

> **Note:** The replicaCredentials in Example A-11 shows the password on the ms_01 system. This marries up with the ibm-slapdPeerPW password in the /etc/slapd32.conf on ms_01, which is updated in Example A-8 on page 163.

Example: A-11 LDIF that is added to machine two (ms_02)

```
dn: cn=ms_01, cn=localhost
cn: ms_01
objectclass: replicaObject
replicabindDN: cn=peer
replicaCredentials: password_ms_01
replicaPort: 389
replicaHost: ms_01
```

Example: A-12   Starting the LDAP server and adding the users

```
# /bin/slapd -f/etc/slapd32.conf
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type PREOPERATION is successfully loaded from libDSP.a.
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type AUDIT is successfully loaded from /lib/libldapaudit.a.
Plugin of type AUDIT is successfully loaded from
/usr/ccs/lib/libsecldapaudit.a(shr.o).
Plugin of type EXTENDEDOP is successfully loaded from libevent.a.
Plugin of type EXTENDEDOP is successfully loaded from libtranext.a.
Plugin of type DATABASE is successfully loaded from /lib/libback-rdbm.a.
Non-SSL port initialized to 389.
Local UNIX socket name initialized to /tmp/s.slapd.
#
#ldapadd -f /tmp/ms_02.ldif -D cn=root -w password
#
```

The setup of replication should now be complete. We will test out replication with the setup of the AIX clients.

Additional information

This appendix explained how to set up a very simple LDAP server for a CSM Cluster. For more information about this subject, visit this site:

http://www.ibm.com/servers/aix/whitepapers/ldap_client.html

B

Managing IBM WebSphere Application Server with CSM

CSM provides a complete, integrated set of administration tools to help manage a cluster of systems. The value of CSM, however, is not limited to managing systems; CSM can also help manage clustered middleware systems such as database, messaging, and application server clusters. This appendix highlights some ways that CSM can be used to help manage an IBM WebSphere® Application Server multi-machine environment (called a cell).

WebSphere is the industry-leading J2EE and Web Services technology-based application platform offering from IBM. Using WebSphere Application Server Network Deployment Edition (ND) Version 5.1, multiple WebSphere Application Servers can be administered from a single node, the deployment manager node, with application-level clustering and failover services available for each application server in the cell. The node topology for a WebSphere Application Server ND cell is similar to that of a CSM cluster, as shown in Figure B-1 on page 168.

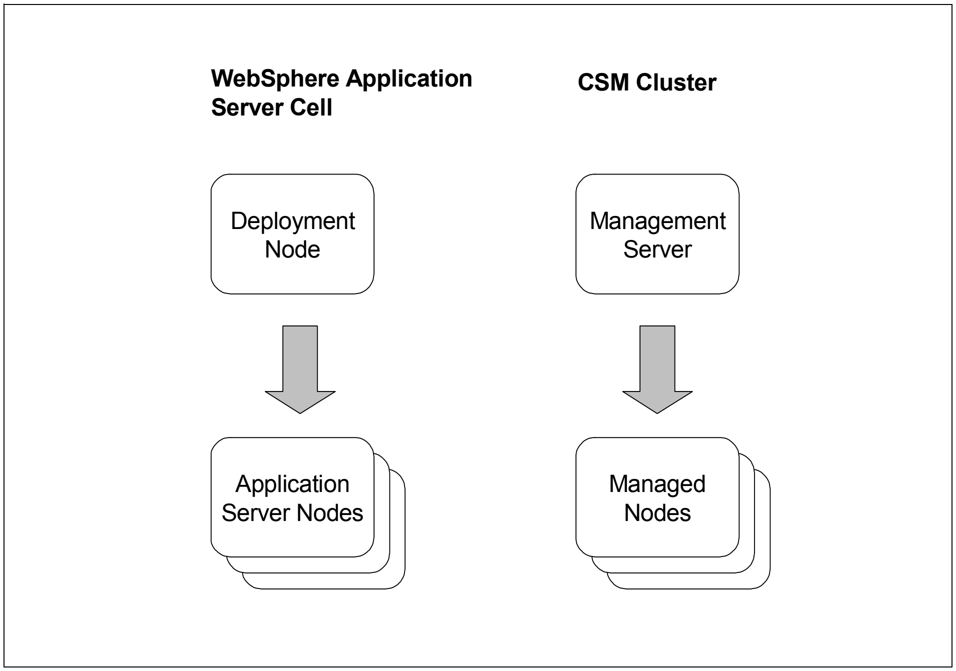


Figure B-1 WebSphere cell and CSM cluster topologies

This similarity enables CSM to complement the administration interface of WebSphere, but not replace it. Both topologies have the notion of a single management node and one or more managed nodes. The difference is that CSM addresses cluster administration at a system level, where WebSphere addresses cell administration at the application server level. However, many WebSphere cell administration tasks require manual repetitive steps across all cluster nodes and control only application server nodes, and not the Web server nodes, database nodes, or other enterprise server nodes that are part of the total WebSphere application infrastructure. Using CSM can help bridge this administration gap.

Combined CSM cluster and WebSphere cell topology

As shown in option (a) of Figure B-2 on page 169, the deployment manager is installed on the CSM management server, and each application server node is a managed node in the CSM cluster. This is a straightforward mapping of the deployment manager node to the management server, and the application servers to managed nodes. As shown in option (b) of Figure B-2 on page 169, the deployment manager is installed on separate managed nodes in the cluster. In this case, a CSM management server manages both the deployment manager nodes and the application server nodes as managed nodes in a CSM cluster.

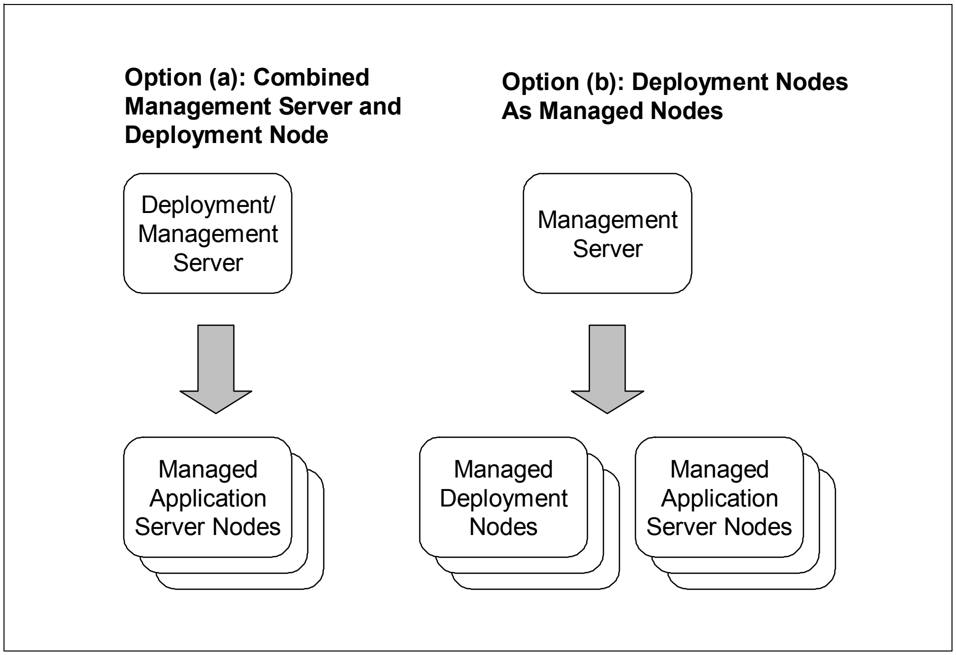

Option (a): Combined Management Server and Deployment Node
Option (b): Deployment Nodes As Managed Nodes
Deployment/ Management Server
Management Server
Managed Application Server Nodes
Managed Deployment Nodes
Managed Application Server Nodes

Figure B-2 Topologies for a CSM-managed WebSphere cell

Both topologies permit CSM and WebSphere administration of the cell. In option (a), the combined CSM management server and WebSphere deployment manager creates a single point of administration for the WebSphere cell. In option (b), the WebSphere deployment manager is separate from the CSM management server. This improves flexibility, as the deployment manager software will not be subject to any software or hardware requirements of the CSM management server and allows for multiple WebSphere deployment managers to be manged in a single CSM cluster.

For the examples that follow, the topology in option (a) of Figure B-2 is assumed. It is also assumed that the cluster platform is AIX or Linux, and the appropriate version of WebSphere Application Server (AIX or Linux) is used.

Using CSM node groups

A CSM node group contains nodes that share a common trait. They can be user-defined (static) or automatically defined by CSM (dynamic). For the examples that follow, each application server node is defined as a CSM managed node and grouped into the user-defined CellNodes node group.

Additional node groups can also be defined as an administrator sees fit. In a WebSphere Application Server cell, this may mean a group for nodes where a specific J2EE application is deployed, or a group for nodes that have database client software installed. Many CSM commands accept a node group parameter. By strategically defining node groups, a cell administrator can manage many different aspects of a WebSphere cell in a powerful way.

Setting up a WebSphere Application Server cell

CSM can help to set up a WebSphere Application Server cell. Consider this high-level procedure from the IBM WebSphere Application Server ND 5.1 InfoCenter that describes a subset of steps to set up a WebSphere Application Server cell:

1. Install the base WebSphere Application Server product (and third-party software) on each application server node.
2. Install the WebSphere Application Server Network Deployment product on the deployment manager node.
3. Execute the **startManager** command on the deployment manager node to start the deployment manager.
4. Execute the **addNode** command script on every application server that will be federated into the cell.
5. Secure files on each application server node and the network deployment node.

Many of these steps can benefit from automation; however, they are still subject to repetition as they must be executed on each node in the cell (the more nodes, the more repetitions).

CSM commands can help. The **dsh** command can be used in step 4 to execute the **addNode** command on each application server node, without having to open a terminal session to each node. Because only one node can be federated into the cell at a time, the fanout for **dsh** can be set to 1 to meet this requirement. Securing files in step 5 can also be accomplished with **dsh**. The WebSphere Application Server InfoCenter lists the recommended permissions that important files should have. These permissions can be set with one execution of the **dsh** command instead of opening a terminal session to each application server node and repeatedly setting permissions for the critical files.

CSM provides a comprehensive installation and update feature for software on nodes in the cluster. In this capacity, CSM can help install WebSphere Application Server and third-party software on the application server nodes (usually an application server must interact with other enterprise systems using third party software). Step 1, “Installing WebSphere Application Server and third-party software on application server nodes” on page 171, is considered in more detail below.

Installing WebSphere Application Server and third-party software on application server nodes

In the existing approach, WebSphere Application Server installation is accomplished using the supplied InstallShield for Multiplatforms wizard. An automated, silent installation using the wizard can also be configured with a response file. Although the automation eliminates interaction with the wizard, the install script still has to be executed individually on each application server node. Each third-party software component (such as database client or messaging client) will also have to be installed using its supplied install program.

CSM can assist as follows:

- WebSphere Application Server file sets (or RPMs) can be included as part of a NIM or Kickstart/AutoYAST configuration, and installed using the **installnode** command on specific application server nodes or on the CellNodes group, for a complete OS, CSM, and WebSphere installation.
- Third-party software file sets (or RPMs), if available, can also be included as part of the NIM or Kickstart/AutoYAST configurations.
- Alternatively, the **dsh** command can be used to NFSmount the installation directory and run a silent install on nodes where the operating system and CSM have already been installed. Non-interactive third-party software installation scripts can also be executed, if available.

With the WebSphere Application Server and third-party software file sets (or RPMs) part a NIM or Kickstart/AutoYAST configuration for the application server nodes, the **installnode** command can be used to complete and monitor the installation from the management server/deployment node, without using multiple sessions to each application server. By using **dsh**, the silent install and third-party installation programs can be started on each application server from the management server without individual sessions to each node.

File synchronization across the cell

After the WebSphere Application Server cell has been defined, additional configuration tasks may be required on the newly federated application server nodes. Alternatively, an existing cell may need updates to its configuration. Either way, this may involve updating configuration files on the application server nodes. The WebSphere Application Server deployment manager provides a facility to update and distribute its configuration files to each node in the cluster.

To ensure integrity, this facility should continue to be used to effect changes in the WebSphere cell configuration.

Outside of the WebSphere Application Server configuration management facility, the Configuration File Manager (CFM) can be used to synchronize the configuration files on each application server node. This may include system, third-party, or user-defined business application configuration files. By maintaining the files under CFM, changes can be made to the files and propagated to all application server nodes. Using pre-update and post-update scripts, the WebSphere-managed application that uses the configuration file can be stopped before the update and started after the update, if required. CFM can also be configured to automatically update a file after it has been changed, combining the edit and update process in one step. For example, a configuration file on each application server node may define the level of logging for a J2EE application running in WebSphere. If this file is under CFM control, the logging detail can be changed by editing the file and using the **`cfmupdate`** command to synchronize it on each of the CellNodes.

If the WebSphere Application Server and third-party software is installed using installnode, CFM can be configured to automatically update the configuration files on an application server node as part of the installnode process. For example, if node01 is an application server node configured in the CellNodes group, and it has just been installed using installnode, any file in CFM that is configured to be synchronized across the CellNodes group will get updated on node01 as part of the installnode process. The benefit here is that software installation and configuration can be completed in a combined step.

Managing the entire WebSphere infrastructure

The scope of the WebSphere cell administration tools is limited to the application server nodes that are federated in the cell. Depending on the type of applications that are deployed to the WebSphere cell, other enterprise server software may be part of the WebSphere infrastructure, and necessary prerequisites for the deployed applications. Database servers, Web servers, messaging servers, and other servers have to be managed as part of the WebSphere cell. CSM provides several tools to help control WebSphere, third-party and user-defined application processes, and to monitor entities in the cell. Here are some examples.

Starting and stopping processes

In order to use the WebSphere Application Server administration tools to manage the cell, the node agent process must be started on each application server node. WebSphere cell administration tools do not provide a way to start the node agents.

The **dsh** command can be used to start and stop the node agents from the management server/deployment node, instead of opening a terminal session to each application server node and starting the node agent. Even if the node agent is part of the node's inittab, using dsh is helpful to control the node agent program, on all or some of the nodes in the cell. The **dsh** command can also be used to start or stop any user-defined applications or clients that support a deployed J2EE application.

Managing Web servers

In most J2EE application infrastructures, a Web server forwards an HTTP request to an application server for processing. A WebSphere cell topology may include one or more Web server nodes. In this case, the WebSphere Application Server plug-in must be installed and configured on each Web server node to enable Web server requests to be forwarded to the application server cell. WebSphere cell administration tools provide the ability to generate the plug-in configuration file, but not to distribute it. Control of the Web server is also outside of the scope of WebSphere cell administration, even though the Web server is part of the WebSphere infrastructure. CSM can help in the following way:

- A node group can be defined (WebServerNodes) to manage all Web server nodes as a single entity.
- Web server file sets (or RPMs) can be included as part of a NIM or Kickstart/AutoYAST configuration for Web server nodes.
- The **dsh** command can be used to start, stop, or restart the Web server processes.
- CFM can be used to synchronize the plug-in file configuration for the WebServerNodes node group. If the file is updated, pre-update and post-update scripts can be used to stop the Web server and start it, respectively, reloading the plug-in configuration.

A Web server node is one example of an enterprise server node outside of the WebSphere cell that can be managed with CSM. The ideas here can be extended to any critical server in the infrastructure.

Monitoring cell processes and system resources

CSM Monitoring can be used to define conditions to monitor in the cell and responses to execute when the condition is true. Some things to consider monitoring in a cell:

- Log files
  A monitor can be created to archive any system, WebSphere, or user-defined application log files from each application server node to the management server when they reach a certain size. This effectively creates a ring log system.

- File system usage
If an application makes heavy use of a file system, or to protect against core dumps consuming disk space, a monitor can be created to clean up the file system when they reach a certain utilization threshold.
- Process availability
Monitors can be created to monitor the status of critical processes in the cell. Conditions such as a process terminating, or using extraneous memory, can be remedied with a preconfigured response.
- Centralized logging with the audit log. Many business applications require logging for audit purposes. The CSM audit log can be used by a user-defined business application as a centralized logging facility. Because it is designed to be efficient and is available on all CSM managed application server nodes in the cell, it is a powerful alternative to custom application log development.
- Monitoring as an application event framework. The CSM monitoring infrastructure can be used as an event framework for deployed applications in the cell. Conditions can be defined that when true will execute a response in the business application (for example, account expiry event fires, generating an e-mail message to customer).

Summary

Although CSM is traditionally considered as a product to manage clustered systems, its ability to manage the software that is running in the cluster should not be overlooked. CSM can never replace the administrative function that is part of IBM WebSphere Application Server ND to manage a cell. It can, however, complement it. This appendix has highlighted some of the ways that CSM can be used to help administer a WebSphere Application Server cell and the applications that run in the infrastructure.

Abbreviations and acronyms

AIX	Advance Interactive Executive
CFM	Configuration File Manager
CSM	Cluster Systems Management
DLAR	Dynamic Logical Partition
HA MS	High Availability Management Server
HACMP	High Availability Cluster Multi-Processing
HMC	Hardware Management Console
IBM	International Business Machines Corporation
ITSO	International Technical Support Organization
LIM	Linux Installation Manager
LPAR	Logical Partition
MAC	Media Access Control
ML	Maintenance Level
MS	Management Server
NIM	Network Installation Management
NTP	Network Time Protocol
PPC	PowerPC
PTF	Program Temporary Fix
RPD	RSCT Peer Domain
RSCT	Reliable Scalable Cluster Technology
SAM	Tivoli System Automation Manager
SFP	Service Focus Point
SPOT	Shared Product Object Tree
TCB	Trusted Computing Base
TCP/IP	Transmission Control Protocol/Internet Protocol
URL	Universal Resource Locator
VG	Volume Group
WebSM	Web-based Systems Manager

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 178. Note that some of the documents referenced here may be available in softcopy only.

- *Deploying Linux on IBM eServer pSeries Clusters*, SG24-7014
- *Introduction to CSM 1.3 for AIX 5L*, SG24-6859
- *Introduction to pSeries Provisioning*, SG24-6389

Other publications

These publications are also relevant as further information sources:

- *AIX 5L Version 5.2 Installation Guide and Reference*, SC23-4389
- *AIX 5L Version 5.3 Installation Guide and Reference*, SC23-4887
- *IBM Cluster Systems Management for AIX 5L Administration Guide Version 1.4*, SA22-7918
- *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.3.3*, SA22-7919-06
- *IBM Cluster Systems Management for AIX 5L Planning and Installation Guide Version 1.4*, SA22-7919
- *IBM Cluster Systems Management for AIX 5L Command and Technical Reference Version 1.3.3,* SA22-7934

Online resources

These Web sites and URLs are also relevant as further information sources:

- NFS information

  http://nfs.sourceforge.net/nfs-howto/

- HTTP logs

  http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html

- SUSE LINUX Web site

  http://www.suse.com

- YaST Autoinstaller

  http://www.suse.de/~nashif/autoinstall/

- Red Hat Web site

  http://www.redhat.com

- Red Hat Linux Customization Guide

  http://www.redhat.com/docs/manuals/linux/RHL-9-Manual/custom-guide/part-install-info.html

- AIX 5L Web site

  http://www.ibm.com/servers/aix/

- Download the openCIMOM Red Hat Package Manager (RPM) package from

  http://www.ibm.com/servers/aix/products/aixos/linux/download.html

- IBM CSM Web site for the latest updates for CSM

  http://techsupport.services.ibm.com/server/cluster/fixes/csmfixhome.html

- File sets for OpenSSH can be obtained from:

  http://www-124.ibm.com/developerworks/projects/opensshi

- CSM client code for pSeries Linux from the IBM Cluster fixes Web page

  http://techsupport.services.ibm.com/server/cluster/fixes

- RPMs updates for AutoUpdate can be obtained from

  http://freshmeat.net/projects/autoupdate

- Configuring the DHCP server

  http://www.isc.org/index.pl?/sw/dhcp/

- IBM Fix Central Web site

  http://www-912.ibm.com/eserver/support/fixes/fcgui.jsp

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications, and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

C

cell processes, monitoring
  as an application event framework 174
  centralized logging with the audit log 174
  file system usage 174
  log files 173
  process availability 174
Cluster Systems Management (CSM) 101
command
  arp 42
  authconfig 158, 160
  bootinfo 73
  bootlist 66
  bootptodhcpc 62
  cat 74
  cfgmgr 129
  cfmupdatenode 27
  chdev 127
  chmod 159
  chnode 25, 37, 78, 90, 111, 114
  chvg 129
  copycsmpkgs 36
  cp -rp /var/opt/hams/mnt/csminstall/* /csminstall/ 147
  cp -rp /var/opt/hams/mnt/etc/* /etc/opt/csm/ 147
  cp -rp /var/opt/hams/mnt/var/* /var/opt/csm/ 147
  csm2nimgrps 28, 85
  csm2nimnodes 27
  csmbackup 77, 80, 98
  csmconfig 20–21, 123
  csmconfig -c 82
  csmrestore 77
  csmsetupnim 28
  csmstat 34, 83, 89, 107, 140
  csmstat -a 147
  ctsthl -l 109
  date 125
  db2ldif 163
  definenode 24
  dsh 3, 8, 35, 73, 84, 89, 123, 170, 173
  dshbak 73
  fuser 148
  geninstall 18, 124
  getadapters 26, 109
  getent passwd 160
  hams 142
  hams -F 141
  hams -l 135, 137
  hams -m 136
  hams -p 139
  hams -S 134
  hams -s 134, 139, 141
  hams -s /opt/csm/hams/hamsdef 141
  hams -sv 143
  importvg 129, 149
  installp 77, 80
  inutoc 80
  java -version 116
  ldapadd 163–164
  ldapcfg 155
  ldfi2db 163
  logform 128
  lppchk 82
  ls 32
  lsaudrec 137
  lsaudrec lgrep _HAM 137, 142
  lsdev 95
  lshwinfo 8
  lshwstat 8
  lslpp 19, 73, 91
  lsnim 32, 87, 106
  lsnim -l 27
  lsnode 24, 35, 90, 113
  lsparent 95
  lspv 95, 127
  lsrg -Ab 135
  lsrpdomain 133, 136
  lsrsrc 107
  lssrc -a IBM.RecoveryRM 139
  lsuser testuser 157
  mgmtsvr 150
  mkdir 105
  mklsecldap 154
  mklv 128
  mksecldap 152–153, 155, 162
  mkuser -R LDAP testuser 156

mkvg 127
mount 130
mount /csminstall 147
mount /var/opt/hams/mnt/ 147
netboot 33, 96, 109
nim 87, 91
nim -o 105
nimadm 92
nodegrp 25
nohup 88
ntpq 125
oslevel 73, 82
passwd -R LDAP testuser 156
ping 107
probemgr 83
ps 115
pwdadm -c 157
rconsole 33, 88–89, 96, 109
reboot 138
recfgct 109
refresh -s inetd 64
resetrsrc 145
resource_name -mv 145
restore 106
rmcctrl -A 109
rmcctrl -p 109
rmcctrl -z 109
rpm 16, 74, 78, 158
rpm -Fvh 20
rpower 8, 25, 83, 89, 109
rpower -a query 113
rpttr 115
rsh 4
showmount 32
shutdown 82
shutdown -F 138
shutdown -Fr 138
sleep 59
smit 76
smit commit 76
smit importvg 129
smit installp 124
smit mksysb 80
smit mkvg 127
smit nim_mac 96
smit nim_mkres 95
smit savevg 80
smitty bffcreate 29
smitty eznim 14
smitty install 152
smitty nim_mac_res 30
smitty nim_mkres 30
smitty nim_update_all 30
smitty update_all 19
ssh 4, 8, 107, 140, 157
startcondresp 132
startsrc 112, 115, 125
startsrc -s IBM.HWCTRLRM 147
startsrc -s IBM.RecoveryRM 139
stopsrc 112, 115, 150
stopsrc -s IBM.HWCTRLRM 147
systemid 21–22, 112
tcbck 92
telnet 107
ulimit 14
umount 147
uname 74
updatenode 38, 78, 88, 90, 93, 109, 147
updatenode -ak 147
updatenode -avk 140
updatenode -k 150
varyoffvg 129
varyoffvg csmvg 147
varyonvg 130
varyonvg csmvg 147
vi 29
view 106
wsm 74
Configuration File Management (CFM) 27
Configuration File Manager (CFM) 3, 172
crfs 128
CSM Highly Available Management Server (CSM HA MS) 7
csmconfig 119

D
daemon
nimesis 14
rpc.statd 58
slapd 162
sshd 156
xntpd 125
directory
/cfmroot 27
/csminstall 21
/csminstall/AIX 120
/export/eznim/extras 14

/export/eznim/lpp_source 14
/home/ldapdb2 153
/lim 41
/lim/tftpboot 54
/opt/csm/hams 130
/ppc/netboot/ 54
dsh 119
Dynamic Logical Partition (DLPAR) 109

F
file
.info 32
.rhosts 91
.toc 80
/etc/apache2/default-server.conf 44
/etc/bootptab 62
/etc/dhcpsd.cnf 64
/etc/exportfs 43
/etc/hosts 17, 38, 112
/etc/inetd.conf 64
/etc/ldap.conf 158
/etc/ldap.secret 159
/etc/nssswitch.conf 158
/etc/openldap/ldap.cfg 160
/etc/security/user 156
/etc/slapd32.conf 162
/var/log/csm/hams 142
autoinstallation 47
bootptab 63
bosinst.data 94
etc/ntp.conf 125
etc/pam.d/system-auth 158
hamsdef 130
ks.cfg 48
opt/csm/hams/hamsdef 141

G
General Parallel Files Systems (GPFS) 102

H
hams -s /opt/csm/hams/hamsdef 134
Hardware Management Console (HMC) 6, 10, 73, 103
High Availability Cluster Multi-Processing (HACMP) 7
High Availability Management Server (HA MS) 72, 117

L
Lightweight Directory Access Protocol (LDAP) 151
Linux Installation Manager (LIM) 40
log
/var/log/csm/install.log 38
Active_CSM.log 142
Active_VLAN_Connection.log 142
hams.log 142
Inactive_CSM.log 142
Inactive_VLAN_Connection.log 142
Shared_Disk.log 142
Shared_Disk_monitor.log 142
smit.log 81
Logical Partition (LPAR) 109

M
management server (MS) 39
Media Access Control (MAC) 26
mklv 128

N
Network Address Translation (NAT) 66
Network Installation Manager (NIM) 13, 72, 104
Network Time Protocol (NTP) 121

P
probe
ibm.csm.ms 22

R
Redbooks Web site 178
Contact us xiii
Reliable Scalable Cluster Technology (RSCT) 6, 17, 103
RPD 124
RSCT peer domain 124
RSCT Peer Domain (RPD) 7

S
script
/opt/csm/hams/scripts 131
csm.install 59
Service Focus Point (SFP) 7
Shared Product Object Tree (SPOT) 13, 104

T
Tivoli System Automation Manager (SAM) 134
Transmission Control Protocol / Internet Protocol (TCPIP) 12
Trusted Computing Base (TCB) 92

IBM Redbooks

Cluster Systems Management Cookbook for pSeries

(0.2”spine)
0.17”<->0.473”
90<->249 pages

Cluster Systems Management Cookbook for pSeries

Redbooks

AIX 5L running on the management server

AIX and Linux nodes in the cluster

Migration scenarios included

This IBM Redbook is a practical cookbook that provides up-to-date information about Cluster Systems Management (CSM) for AIX 5L for a pSeries environment. The book provides information about the latest CSM for AIX 5L enhancements, including implementation techniques, installation changes, installation tools, system management tools, monitoring tools, hardware control, file distribution, problem determination, and management server high availability.

This book summarizes the latest news in CSM 1.4.0. It contains a Q&A chapter, a CSM installation scenario, a CSM advanced chapter, CSM migration scenarios, and a CSM cluster administration chapter. We include information about how to manage Linux nodes on pSeries hardware including operating system installation and node management in a mixed cluster environment.

This Redbook is targeted to technical professionals (consultants, IT architects, and IT specialists) who are responsible for providing pSeries clustering solutions.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-6859-01 ISBN 0738490997