Bull

HACMP 4.4
Installation Guide

AIX

ORDER REFERENCE
86 A2 56KX 02

Bull

HACMP 4.4
Installation Guide

AIX

Software

August 2000

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

ORDER REFERENCE
86 A2 56KX 02

Printed in France

Suggestions and criticisms concerning the form, content, and presentation of this book are invited. A form is provided at the end of this book for this purpose.

To order additional copies of this book or other Bull Technical Publications, you are invited to use the Ordering Form also provided at the end of this book.

Trademarks and Acknowledgements

Year 2000

The product documented in this manual is Year 2000 Ready.

The information in this document is subject to change without notice. Groupe Bull will not be liable for errors contained herein, or for incidental or consequential damages in connection with the use of this material.

Contents

Chapter 3 Using the Quick Configuration Utility 3-1

Part 3 Installing HACMP Cluster Configurations: Standard Method

Chapter 4 Configuring Cluster Networks and Performance Tuning 4-1

Chapter 5 Installing Shared Disk Devices 5-1

Chapter 6 Defining Shared LVM Components 6-1

Chapter 7 Additional AIX Administrative Tasks 7-1

Chapter 8 Installing HACMP for AIX Software 8-1

Chapter 9 Upgrading an HACMP Cluster 9-1

About This Guide

This guide provides information necessary to install and configure High Availability Cluster Multi-Processing for AIX, Version 4.4 (HACMP for AIX) software in your cluster environment.

Who Should Use This Guide

This guide is intended for system and network administrators, and for customer engineers responsible for:

- Planning hardware and software resources for an HACMP for AIX cluster environment
- Configuring networks
- Defining physical and logical storage
- Installing and configuring an HACMP cluster

As a prerequisite to installing HACMP for AIX software, you should be familiar with:

- System components (including disk devices, cabling, and network adapters)
- The AIX operating system, including the Logical Volume Manager subsystem
- The System Management Interface Tool (SMIT)
- Communications, including the TCP/IP subsystem

Before You Begin

Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide* contains blank copies of the worksheets referred to in this guide. These worksheets help you plan, install, configure, and maintain an HACMP cluster. Complete the required worksheets before installing the HACMP for AIX software.

The examples that rely on SMIT assume you are using AIX from an ASCII display. SMIT is also available within the AIXwindows environment.

Alternatively, you can use the online planning worksheets. See Appendix B of the *HACMP for AIX Planning Guide* for information.

How To Use This Guide

This guide is divided into five parts.

Part 1: Overview

Part 1 lists the tasks involved in installing and configuring HACMP for AIX software using either of two configuration methods: Quick or Standard. Steps required for each method are listed in the following chapter:

- Chapter 1, Installing an HACMP Cluster: List of Steps

Part 2: Installing HACMP Cluster Configurations: Quick Method

Part 2 describes the tasks involved in installing and configuring an HACMP cluster using the HACMP for AIX Quick Configuration method. This part includes the following chapters:

- Chapter 2, Setting Up a Quick Configuration Cluster
- Chapter 3, Using the Quick Configuration Utility

Part 3: Installing HACMP Cluster Configurations: Standard Method

Part 3 describes the tasks involved in installing and configuring an HACMP cluster using the Standard method.

Chapters 4 through 7 describe the required AIX system configuration tasks that must be done before installing the HACMP for AIX software. Chapters 8 through 10 describe how to install and upgrade HACMP for AIX software (overwrite a previous software version and convert the cluster configuration) and how to use the verification procedure to check for installation errors or problems. Chapters 11 through 16 describe procedures for defining and configuring an HACMP cluster.

- Chapter 4, Configuring Cluster Networks and Performance Tuning
- Chapter 5, Installing Shared Disk Devices
- Chapter 6, Defining Shared LVM Components
- Chapter 7, Additional AIX Administrative Tasks
- Chapter 8, Installing HACMP for AIX Software
- Chapter 9, Upgrading an HACMP Cluster
- Chapter 10, Verifying Cluster Software
- Chapter 11, Defining the Cluster Topology
- Chapter 12, Configuring Cluster Resources
- Chapter 13, Verifying the Cluster Topology
- Chapter 14, Customizing Cluster Events and Log Files
- Chapter 15, Setting Up Clinfo on Server Nodes
- Chapter 16, Supporting AIX Error Notification (optional)

Part 4: Installing and Configuring Cluster Clients

Part 4 describes the steps for installing and configuring HACMP for AIX software on clients. This part includes the following chapter:

- Chapter 17, Installing and Configuring Clients

Part 5: Appendixes

Part 5 contains reference information. The appendixes include:

- Appendix A, Supporting IP Address Takeover
- Appendix B, Installing and Configuring Cluster Monitoring with Tivoli
- Appendix C, Image Cataloger Demo
- Appendix D, CLMarket Demo
- Appendix E, HACMP for AIX and SNMP Utilities
- Appendix F, Installing and Configuring HACMP for AIX on RS/6000 SPs

An index follows Part 5.

Highlighting

The following highlighting conventions are used in this guide:

Italic Identifies variables in command syntax, new terms and concepts, or indicates emphasis.

Bold Identifies routines, commands, keywords, files, directories, menu items, and other items whose actual names are predefined by the system.

`Monospace` Identifies examples of specific data values, examples of text similar to what you might see displayed, examples of program code similar to what you might write as a programmer, messages from the system, or information that you should actually type.

ISO 9000

ISO 9000 registered quality systems were used in the development and manufacturing of this product.

Related Publications

The following books provide additional information about HACMP for AIX:

- *Release Notes* in **/usr/lpp/cluster/doc/release_notes** describe hardware and software requirements
- *HACMP for AIX, Version 4.4: Concepts and Facilities*, order number 86 A2 54KX 02
- *HACMP for AIX, Version 4.4: Planning Guide*, order number 86 A2 55KX 02
- *HACMP for AIX, Version 4.4: Administration Guide*, order number 86 A2 57KX 02
- *HACMP for AIX, Version 4.4: Troubleshooting Guide*, order number 86 A2 58KX 02
- *HACMP for AIX, Version 4.4: Programming Locking Applications*, order number 86 A2 59KX 02
- *HACMP for AIX, Version 4.4: Programming Client Applications*, order number 86 A2 60KC 02
- *HACMP for AIX, Version 4.4: Master Index and Glossary*, order number 86 A2 65KX 02
- *HACMP for AIX, Version 4.4: Enhanced Scalability Installation and Administration Guide Volumes I and II*, order numbers 86 A2 62KX 02 and 86 A2 89KX 01

Ordering Publications

To order additional copies of this guide, use order number 86 A2 56KX 02.

Part 1 Overview

This part outlines the steps involved in the installation process.

Chapter 1 Installing an HACMP Cluster: List of Steps

This chapter provides an overview of how to install and configure the HACMP for AIX, Version 4.4 software in a cluster environment. It also provides steps for installing and configuring HACMP for AIX clients in cluster environments.

Prerequisites

Read the *HACMP for AIX Planning Guide* before installing the HACMP for AIX software. It contains the necessary worksheets and diagrams to consult as you proceed through the installation and configuration steps listed in this chapter. If you have not completed these worksheets and diagrams, return to the appropriate chapters in the *HACMP for AIX Planning Guide* and do so before continuing.

Upgrading an HACMP for AIX Cluster to Version 4.4

If you are upgrading an existing HACMP cluster to HACMP for AIX, Version 4.4, read Chapter 9, Upgrading an HACMP Cluster, for more information.

Steps for Installing and Configuring an HACMP Cluster

This section identifies the steps required to set up, install, and configure an HACMP cluster using either of two configuration methods: Quick or Standard. Steps for each method are divided into the following major areas:

- Preparing AIX for an HACMP cluster—setting up hardware and software in AIX
- Installing and configuring an HACMP cluster—configuring the cluster to handle resources according to your specifications.

Whichever configuration method you choose, proper planning is essential before installing and configuring the HACMP for AIX software. See the *HACMP for AIX Planning Guide* for more information on planning your cluster configuration.

Note: You can use the HACMP for AIX VSM utility (**xhacmpm**) to perform many configuration tasks after installing the HACMP for AIX software. The utility is an X Windows tool for creating cluster configurations using icons that represent cluster resources. See the *HACMP for AIX Administration* guide, Appendix D, VSM Graphical Configuration Application, for more information on using **xhacmpm**.

Preparing AIX for an HACMP Cluster: Quick Method

Step 1: Read the Quick Configuration Cluster Diagrams Section

In this step you read the Quick Configuration Cluster Diagrams section in Chapter 2, Setting Up a Quick Configuration Cluster for a description of the required hardware and type of resource allocation for each predefined cluster configuration.

Step 2: Install the Hardware for the Chosen Configuration

In this step you install the hardware for the chosen predefined configuration.

Step 3: Read the Quick Configuration Prerequisites and Preconfiguration Guidelines

In this step you read the Quick Configuration Prerequisites and Guidelines on page 2-9 to ensure that your HACMP for AIX software installation and cluster configuration will be implemented correctly.

Step 4: Configure the Prerequisite Network Adapters

In this step you configure the prerequisite network adapters for your chosen predefined configuration.

Step 5: Preconfigure Network and Disk Subsystems Devices (optional)

In this step you preconfigure network and disk subsystems devices for your chosen predefined configuration.

Step 6: Read Chapter 7, Additional AIX Administrative Tasks

In this step you read Chapter 7, Additional AIX Administrative Tasks, to ensure that all requirements for proper performance of the HACMP cluster are met for each cluster node.

Step 7: Install the HACMP for AIX, Version 4.4 Software

In this step you install the HACMP for AIX software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software.

Step 8: Verify the Software Installation

In this step you verify that the software used in the cluster is current by following the instructions in Chapter 10, Verifying Cluster Software.

Step 9: Complete the *HACMP for AIX Planning Guide* worksheets (optional)

In this step you complete the worksheets in Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide* for your chosen configuration. The worksheets provide a useful record of your cluster configuration in case you later add to or change it.

When you have finished these tasks, you are ready to run the Quick Configuration Utility. See Chapter 3, Using the Quick Configuration Utility, for more information.

Installing and Configuring an HACMP Cluster: Quick Method

Step 1: Invoke the Quick Configuration Utility

In this step you invoke the Quick Configuration utility using the **xclconfig** command. Help information is provided throughout the utility.

Step 2: Select One of Five Cluster Configurations

In this step you select one of five predefined configurations provided by the utility.

Step 3: Define the IP Address for Each Node

In this step you define the IP address of each cluster node to the Quick Configuration utility.

Step 4: Customize the Cluster Configuration (optional)

In this step you customize your chosen cluster configuration, if you do not want to use the current system settings as determined by the Quick Configuration utility.

Step 5: Verify the Cluster Configuration

In this step you run the HACMP for AIX **/usr/sbin/cluster/diag/clverify** utility to ensure that the required hardware and cluster topology are available and properly configured.

Step 6: Apply the Cluster Configuration

In this step you cause the system to synchronize (copy) the HACMP ODM classes, as defined in the temporary ODM, to the defined cluster nodes.

When you have finished these tasks, see the chapter on starting and stopping cluster services in the *HACMP for AIX Administration Guide* to start your cluster.

Preparing AIX for an HACMP Cluster: Standard Method

Step 1: Configure Networks

In this step you configure network adapters as described in Chapter 4, Configuring Cluster Networks and Performance Tuning.

Step 2: Install Shared Disk Devices

In this step you install the shared external disks for your HACMP cluster as described in Chapter 5, Installing Shared Disk Devices.

Step 3: Define Shared LVM Components

In this step you create the shared volume groups, logical volumes, and filesystems for your cluster as described in Chapter 6, Defining Shared LVM Components.

Step 4: Perform Additional AIX Administrative Tasks

In this step you review or edit various AIX files to ensure a proper configuration for I/O pacing, network options, NFS, and for various host files as described in Chapter 7, Additional AIX Administrative Tasks.

Installing and Configuring an HACMP Cluster: Standard Method

Step 1: Install HACMP for AIX Software

In this step you install the HACMP for AIX software on each cluster node. Chapter 8, Installing HACMP for AIX Software, describes this step for a new install. Chapter 9, Upgrading an HACMP Cluster, describes this step for an upgrade.

After installing HACMP for AIX software, read the *HACMP for AIX, Version 4.4 Release Notes* in the **/usr/lpp/cluster/doc** directory for additional information about the software's functionality.

Step 2: Verify Cluster Software

In this step you use the **/usr/sbin/cluster/diag/clverify** utility to verify that other software installed on cluster nodes is compatible with the HACMP for AIX software. Chapter 10, Verifying Cluster Software, describes this step.

Step 3: Define the Cluster Topology

In this step you define the components of your HACMP cluster as described in Chapter 11, Defining the Cluster Topology.

Step 4: Configure Cluster Resources

In this step you identify the scripts that start and stop server applications running on cluster nodes, configure the resource groups, and set the HACMP for AIX node variables. Chapter 12, Configuring Cluster Resources, describes these steps.

Step 5: Verify the Cluster Configuration

In this step you use the **/usr/sbin/cluster/diag/clverify** utility to verify that the cluster is configured properly. Chapter 13, Verifying the Cluster Topology, describes this step.

Step 6: Customize Cluster Events

In this step you customize your environment to process cluster events as described in Chapter 14, Customizing Cluster Events and Log Files.

Step 7: Set up the Cluster Information Program

In this step you edit the **/usr/sbin/cluster/etc/clhosts** file and the **/usr/sbin/cluster/etc/clinfo.rc** script. Chapter 15, Setting Up Clinfo on Server Nodes, describes these steps.

Step 8: Enable AIX Error Notification Facility Support (optional)

In this step you use the AIX Error Notification facility to identify and respond to failures within an HACMP cluster. Chapter 16, Supporting AIX Error Notification, describes this step.

The installation is complete when you have performed these tasks.

Steps for Installing and Configuring an HACMP Client

After installing the HACMP for AIX software on each cluster node, the following steps allow you to install the software on clients and to enable it to provide users with cluster status information. These steps are described in Chapter 17, Installing and Configuring Clients.

Step 1: Install the Base System on Clients

In this step you install the base high availability software on a client.

Step 2: Edit the /usr/sbin/cluster/etc/clhosts File

In this step you edit the **/usr/sbin/cluster/etc/clhosts** file to provide the HACMP for AIX server addresses needed for clients to communicate with cluster nodes.

Step 3: Edit the /usr/sbin/cluster/etc/clinfo.rc Script

In this step you review the importance of editing the **/usr/sbin/cluster/etc/clinfo.rc** script to ensure that the ARP cache is updated as a result of a cluster event.

Step 4: Update Non-Clinfo Clients

In this step you update the ARP cache on non-Clinfo clients.

Step 5: Reboot the Clients

In this step you reboot each cluster client. See the chapter on starting and stopping cluster services in the *HACMP for AIX Administration Guide* for more information.

Specified Operating Environment

This section describes the required and supported hardware for HACMP, as of version 4.3.1.

Hardware Requirements

HACMP 4.3.1 works with RS/6000 uniprocessors, SMP servers, and SP systems in a "no-single-point-of-failure" server configuration. HACMP 4.3.1 supports the RS/6000 models designed for server applications and meet the minimum requirements for internal memory, internal disk, and I/O slots. The following RS/6000 models and their corresponding upgrades are supported in HACMP 4.3.1:

- PCI Desktop Systems, Models 140, 150, 240, and 260
- PCI Deskside Systems, Models E20, E30, F30, F40, and F50
- PCI Rack Systems, Models H10, S70, S7A, and S80
- Entry Systems, Models 25S, 250, and 25T
- Compact Server Systems, Models C10 and C20
- Desktop Systems, Models 370, 380, 390, 397, and 39H
- Deskside Systems, Models 570, 57F, 580, 58F, 58H, 590, 59H, 591, and595
- Rack Systems, Models 98B, 98E, 98F, 990, 99E, 99F, 99J, 99K, R10,R20, R21, R24, R50, R5U, S70, S7A, H50, and H70

- Symmetric Multiprocessor Server Systems, Models G30, J30, R30, R3U,G40, J40, R40, R4U, J50, R4U, S70, and S7A
- SP Systems, Models 204, 205, 206, 207, 208, 209, 20A, 2A4, 2A5, 2A7,2A8, 2A9, 2AA, 304, 305, 306, 307, 308, 309, 30A, 3A4, 3A5, 3A7, 3A8, 3A9,3AA, 3B4, 3B5, 3B7,3B8, 3B9, 3BA, 404, 405, 406, 407, 408, 409, 40A, 500, 50H, 550, and55H, including the 604 High Nodes, 604E High Nodes, and the Power2 SuperChip (P2SC) nodes

Any supported RS/6000 can be joined with any other supported RS/6000 in an HACMP 4.3.1 configuration. The Models 250 and 25T can be used in the HACMP4.3.1 server configuration, but due to slot limitations, a "single-point-of-failure" is unavoidable in shared-disk or shared-network resources.

HACMP 4.3.1 executing in a concurrent access configuration requires one of the following devices:

- IBM 7131 SSA Multi-Storage Tower Model 405 (supports up to eight nodes; no CD-ROMs or tapes can be installed)
- IBM 7133 SSA Disk Subsystem Models 020, 600, D40 and T40 (supports up to eight nodes)
- IBM 7135 RAIDiant Array Models 110 and 210 (supports up to four nodes; dual controllers recommended)
- IBM 7137 Disk Array Subsystem Models 413, 414, 415, 513, 514, or 515 (supports up to four nodes)
- IBM 2105 Versatile Storage Server (VSS) Models B09 and 100 (supports up to four nodes)

Certain non-IBM RAID systems can operate in concurrent I/O access environments. IBM will not accept Authorized Program Analysis Reports (APARs) if the non-IBM RAID offerings do not work properly with HACMP 4.3.1.

The minimum configuration and sizing of each machine is highly dependent on the user’s database package and other applications.

Actual configuration requirements are highly localized according to the required function and performance needs of individual sites. In configuring a cluster, particular attention must be paid to:

- Fixed-disk capacity and mirroring (Logical Volume Manager (LVM) and database)
- Slot limitations and their effect on creating a single-point-of-failure
- Client access to the cluster
- Other LAN devices (routers, bridges) and their effect on the cluster
- Replication of I/O adapters/subsystems
- Replication of power supplies
- Other network software

Whenever a process takes over resources after a failure, consideration must be given to work partitioning. For example, if processor "A" is expected to take over for failed processor "B" and continue to perform its original duties, "A" must be configured with enough resources to perform the work of both.

HACMP 4.3.1 Device Support

At this time, the following adapters are supported in the HACMP 4.3.1 environment.Refer to individual hardware announcements for the levels of AIX that are supported.

Communications Adapters

- PCI/ISA
- 2920 IBM PCI Token-Ring Adapter
- 2931 ISA 8-Port Asynchronous Adapter
- 2932 ISA 8-Port Asynchronous Adapter
- 2933 ISA 128-Port Asynchronous Controller
- 2741 PCI FDDI-Fiber Single-Ring Upgrade
- 2742 PCI FDDI-Fiber Dual-Ring Upgrade
- 2743 PCI FDDI-Fiber Single-Ring Upgrade
- 2944 128-Port Asynchronous Controller, PCI bus
- 2943 8-Port Asynchronous EIA-232/RS-422, PCI bus Adapter
- 2963 Turboways 155 PCI UPT ATM Adapter
- 2968 PCI Ethernet 10/100 Adapter
- 2969 PCI Gigabit Ethernet Adapter
- 2979 PCI AutoLANStreamer Token-Ring Adapter
- 2985 PCI Ethernet BNC/RJ-45 Adapter
- 2986 PCI Ethernet 10/100 Adapter
- 2987 PCI Ethernet AUI/RJ-45 Adapter
- 2988 Turboways 155 PCI MMF ATM Adapter
- 4959 Token Ring PCI Adapter
- 8396 RS/6000 SP System Attachment Adapter

ATM Hardware Address Takeover is limited to adapters connected to the same switch.

MCA

- 1904 Fibre Channel Adapter
- 2402 Network Terminal Accelerator Adapter
- 2403 Network Terminal Accelerator Adapter
- 2723 FDDI-Fiber Dual-Ring Upgrade
- 2724 FDDI-Fiber Single-Ring Adapter
- 2725 FDDI-STP Single-Ring Adapter
- 2726 FDDI-STP Dual-Ring Upgrade
- 2930 8-Port Async Adapter - EIA-232
- 2964 10/100 Mbps Ethernet Adapter - UNI
- 2972 AutoLANStreamer Token-Ring Adapter
- 2980 Ethernet High-Performance LAN Adapter

- 2989 Turboways 155 ATM Adapter
- 2992 Ethernet/FDX 10 Mbps TP/AUI MC Adapter
- 2993 Ethernet BNC MC Adapter
- 2994 10/100 Mbps Ethernet Adapter - SMP
- 4018 High-Performance Switch (HPS) Adapter-2
- 4020 Scalable POWERParallel Switch Adapter

Disk Adapters

PCI

- 6205 PCI Dual Channel Ultra2 SCSI Adapter
- 6206 PCI SCSI-2 Single-Ended Ultra-SCSI Adapter
- 6207 PCI SCSI-2 Differential Ultra-SCSI Adapter
- 6208 PCI SCSI-2 Single-Ended Fast/Wide Adapter
- 6209 PCI SCSI-2 Differential Fast/Wide Adapter
- 6215 PCI SSA Adapter
- 6225 Advanced SerialRAID Adapter

MCA

2412 Enhanced SCSI-2 Differential Fast/Wide Adapter/A 2415 SCSI-2 Fast/Wide Adapter/A 2416 SCSI-2 Differential Fast/Wide Adapter/A 2420 SCSI-2 Differential High-Performance External I/O Controller 6212 High Performance Subsystem Adapter/A (40/80 Mbps)

- 6214 SSA 4-Port Adapter
- 6216 Enhanced SSA 4-Port Adapter
- 6219 MCA SSA Adapter

For compatibility with subsystems not listed below, refer to the individual hardware announcements.

External Storage Subsystems

- IBM 2105 Versatile Storage Server (VSS) Models B09 and 100 (supports up to four nodes)
- IBM 7131 SCSI Multi-Storage Tower Model 105 (supports up to four nodes; no CD-ROMs or tapes can be installed)
- IBM 7131 SSA Multi-Storage Tower Model 405 (supports up to eight nodes; no CD-ROMs or tapes can be installed)
- IBM 7133 SSA Disk Subsystem Models 020, 600, D40 and T40 (supports up to eight nodes)
- IBM 7135 RAIDiant Array Models 110 and 210 (supports up to four nodes; dual controllers recommended)
- IBM 7137 Disk Array Subsystem Models 413, 414, 415, 513, 514, and 515 (supports up to four nodes)
- IBM 7204 External Disk Drive Models 317, 325, 339, 402, 404, and 418 (supports up to four nodes)

- IBM 2105 Versatile Storage Server (VSS) Models B09 and 100 and Enterprise Storage Server (ESS) Models E10 and E20 (supports up to four nodes)

Router Support

The IBM RS/6000 SP Switch Router 9077-04S can be used in cluster configurations where the router is used to provide communications to client systems.

The router is not supported in the communications path between nodes in an HACMP cluster.

Rack-Mounted Storage Subsystems

IBM 7027 High Capacity Storage Drawer Model HSC (supports up to two nodes; no CD-ROMS or tapes installed)

IBM 7027 High Capacity Storage Drawer Model HSD (supports up to four nodes; no CD-ROMS or tapes installed)

Part 2 Installing HACMP Cluster Configurations: Quick Method

Use the Quick Configuration Utility for two-node clusters.

Chapter 2 Setting Up a Quick Configuration Cluster

The Quick Configuration utility, **xclconfig**, is an X Window System application that simplifies the task of configuring a two-node HACMP cluster. It lets you automate the configuration of one of five predefined two-node cluster configurations.

This chapter describes the predefined cluster configurations and provides instructions on tasks you must complete before installing HACMP for AIX software and before using the Quick Configuration utility. The chapter also provides guidelines for optional preconfigurations.

As a prerequisite to using the Quick Configuration utility to install and configure your cluster, you must have the hardware required for your chosen predefined cluster configuration.

Steps for Setting Up a Quick Configuration

This chapter provides instructions for the first five steps listed below. Refer to other chapters as instructed to complete the set up for your quick configuration.

1. Read the section Quick Configuration Cluster Diagrams on page 2-2. This section describes the required hardware and resource allocation type for each predefined cluster configuration.
2. Install the hardware for the chosen configuration.
3. Read the section Quick Configuration Prerequisites and Guidelines on page 2-9.
4. Configure the prerequisite network adapters.
5. Preconfigure network and disk subsystems devices.
6. Read Chapter 7, Additional AIX Administrative Tasks, to ensure that all requirements for the proper performance of the HACMP cluster are met for each node.
7. Install the HACMP for AIX, Version 4.4 software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software. See Chapter 9, Upgrading an HACMP Cluster, for information about upgrading an existing cluster configuration.
8. Verify that the software installed on a cluster node is compatible with the installed version of HACMP for AIX software by following the instructions in Chapter 10, Verifying Cluster Software.
9. (optional) Complete the worksheets in the *HACMP for AIX Planning Guide* for your configuration. This information provides a useful record of your cluster configuration in case you later add to or change it.

When you have finished these tasks, you are ready to run **xclconfig**, which resides in the **/usr/sbin/cluster** directory. See Chapter 3, Using the Quick Configuration Utility, for more information.

Note: The C-SPOC utility lets you use cluster-wide commands to manage your predefined configuration. See Appendix A of the *HACMP for AIX Administration Guide* for a list of C-SPOC commands you can use. See Chapter 6 of the *HACMP for AIX Concepts and Facilities Guide* for a general description of the C-SPOC utility.

Quick Configuration Cluster Diagrams

The cluster diagrams in this section show a cluster's function and structure. Each cluster diagram identifies one of four predefined default configurations and includes the following:

- The default cluster name
- The default node names
- The types of resources used by highly available applications
- The IP addresses shared by the nodes
- The method of shared disk access
- The network topology.

Default Cluster Name

The default cluster name is *Cluster1* for all predefined configurations. When using the Quick Configuration utility, you can change this name to customize the cluster.

Default Node Names

The default node names for the predefined configurations are *Node1* and *Node2*. After entering the IP addresses of the actual nodes you want to include in the configuration, the utility derives and uses the node names corresponding to those addresses.

Resource Groups

The default resource group names for the predefined configurations are *Group1* and *Group2* (where applicable). You can change these names to customize the cluster when using the utility. Descriptions of resource group type associations to a configuration follow:

- Configuration 1 uses one rotating resource group, called *Group1*.
- Configurations 2, 4, and 5 use two cascading resource groups, called *Group1* and *Group2*.
- Configuration 3 uses one concurrent access resource group, called *Group1*.

Resources

The predefined configurations use the following names for default resources included in *Group1* and *Group2*, respectively:

- Application servers: *AppServer1*, *AppServer2*
- IP service labels: *node1_svc*, *node2_svc*
- Volume groups: *vg1*, *vg2*
- Filesystems: *fs1*, *fs2*.

Shared IP Address

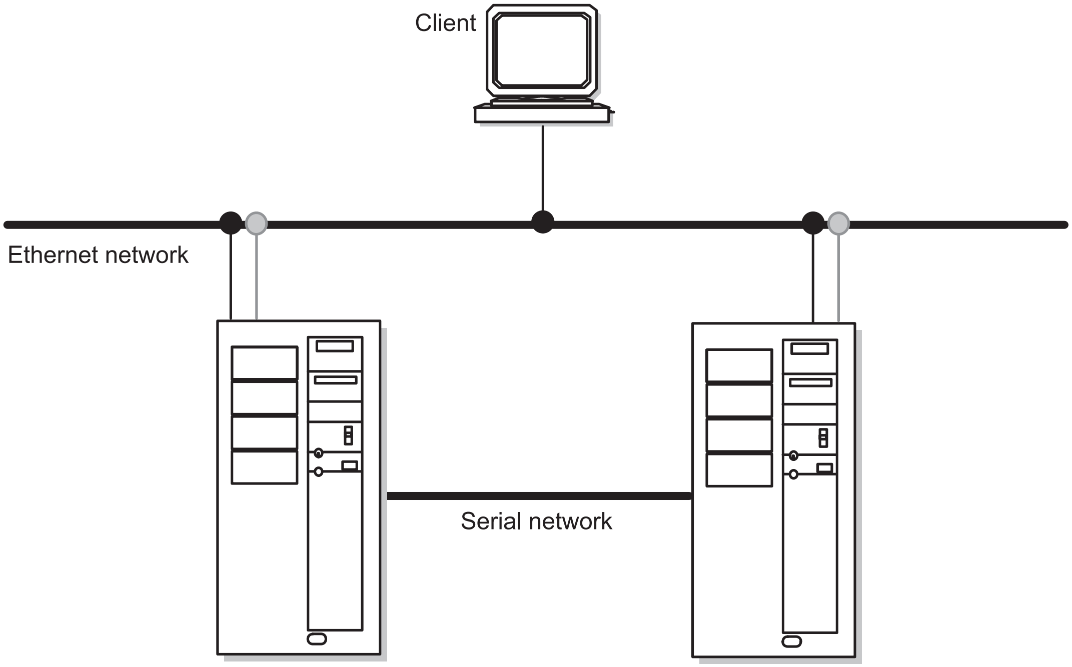
Configuration 1 uses a shared IP address. The default name is *shared_svc*. You may be required to enter the name associated with this address.

Methods of Shared Disk Access

Configuration 3 uses concurrent access mode; the other configurations use non-concurrent access mode.

Quick Configurations: Cluster Network Topology

The five predefined HACMP for AIX quick configurations use one public Ethernet network (default name is *Ethernet1*) and one private serial network (default name is *RS232*) for communication between the two cluster nodes. Each node has both a service and a standby adapter attached to the Ethernet network.



Quick Configurations: Cluster Network Topology

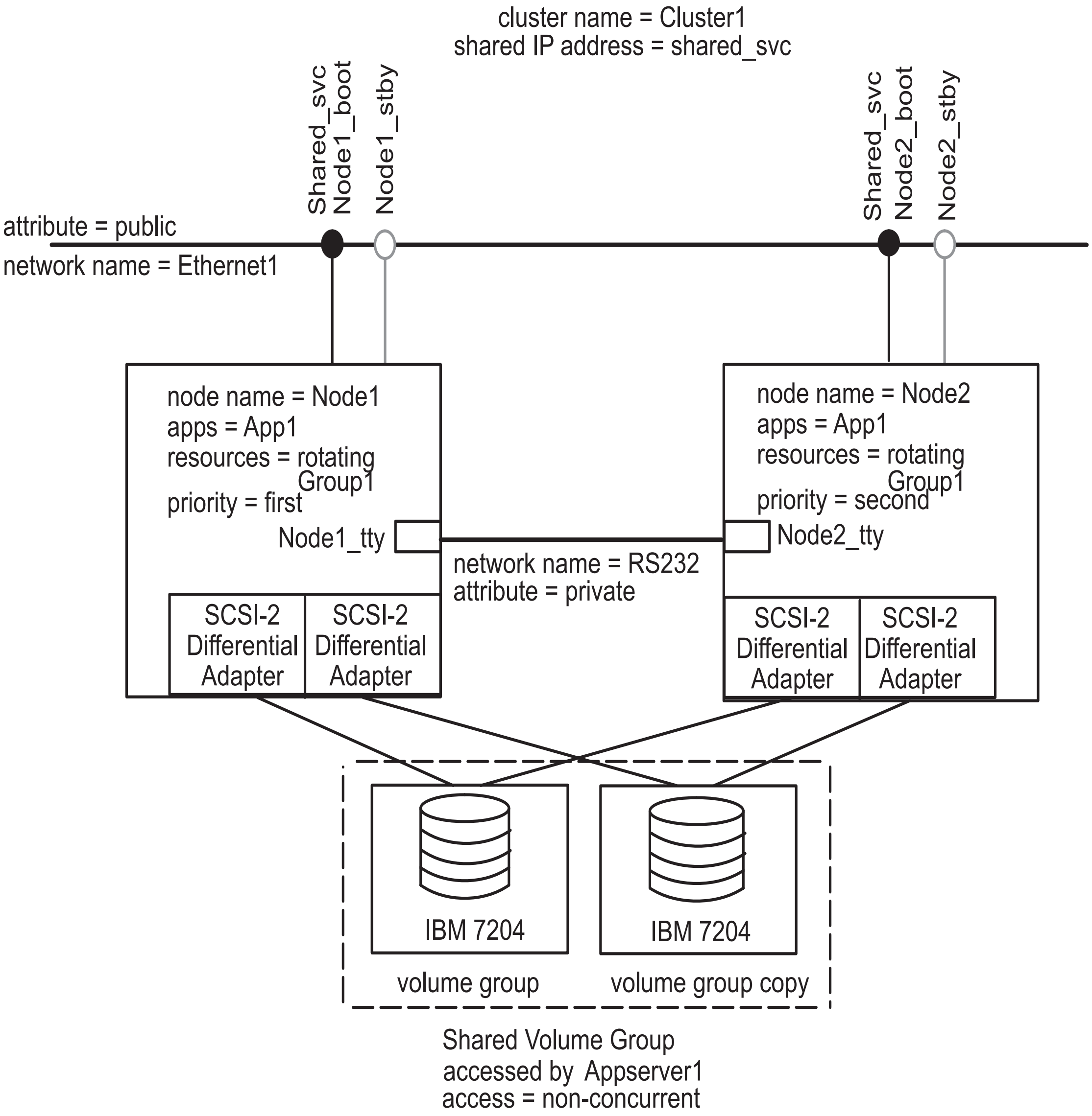
The information in the following diagrams describes the hardware and software for each predefined cluster configuration in general terms. Consult the *HACMP for AIX Planning Guide* for details on supported devices.

Configuration 1: Two IBM 7204 Disk Subsystems, One Rotating Resource Group

This configuration requires:

- Two nodes, two Ethernet adapters per node.
- One Ethernet network connecting the two nodes and their clients.

- One RS232 link (private network) between the two nodes.
- Two IBM 7204 disk subsystems shared between the two nodes, each connected to two SCSI adapters. Each IBM 7204 disk subsystem is on a separate SCSI bus.
- Two SCSI adapters per node (for external disks).
- One volume group, mirrored from one IBM 7204 disk subsystem to the other.
- One rotating resource group, Group1, consisting of an application, the volume group, and the IP address for the server function.



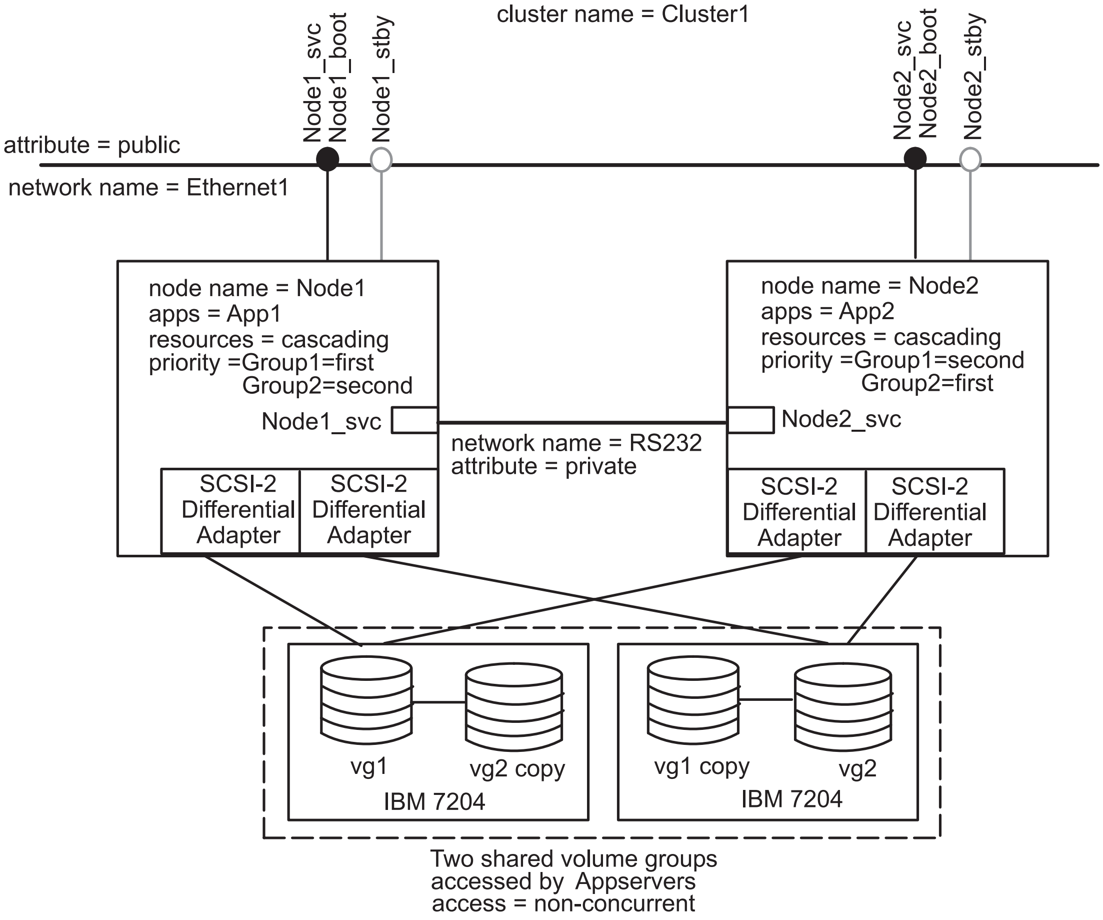
Quick Configuration 1: Cluster Diagram with Network Topology:

Shared IBM 7204 Disk Subsystem and Bus Configuration

Configuration 2: Two IBM 7204 Disk Subsystems, Two Cascading Resource Groups

This configuration requires:

- Two nodes, two Ethernet adapters per node.
- One Ethernet network connecting the two nodes and their clients.
- One RS232 link (private network) between the two nodes.
- Two IBM 7204 disk subsystems shared between the two nodes, each connected to two SCSI adapters (two are on each SCSI bus).
- Two SCSI adapters per node (for external disks).
- Two volume groups, mirrored from one IBM 7204 disk subsystem to the other.
- Two cascading resource groups, each consisting of an application, one of the shared volume groups, and the IP address for the server application. Each node is the initial owner of one of the resource groups.



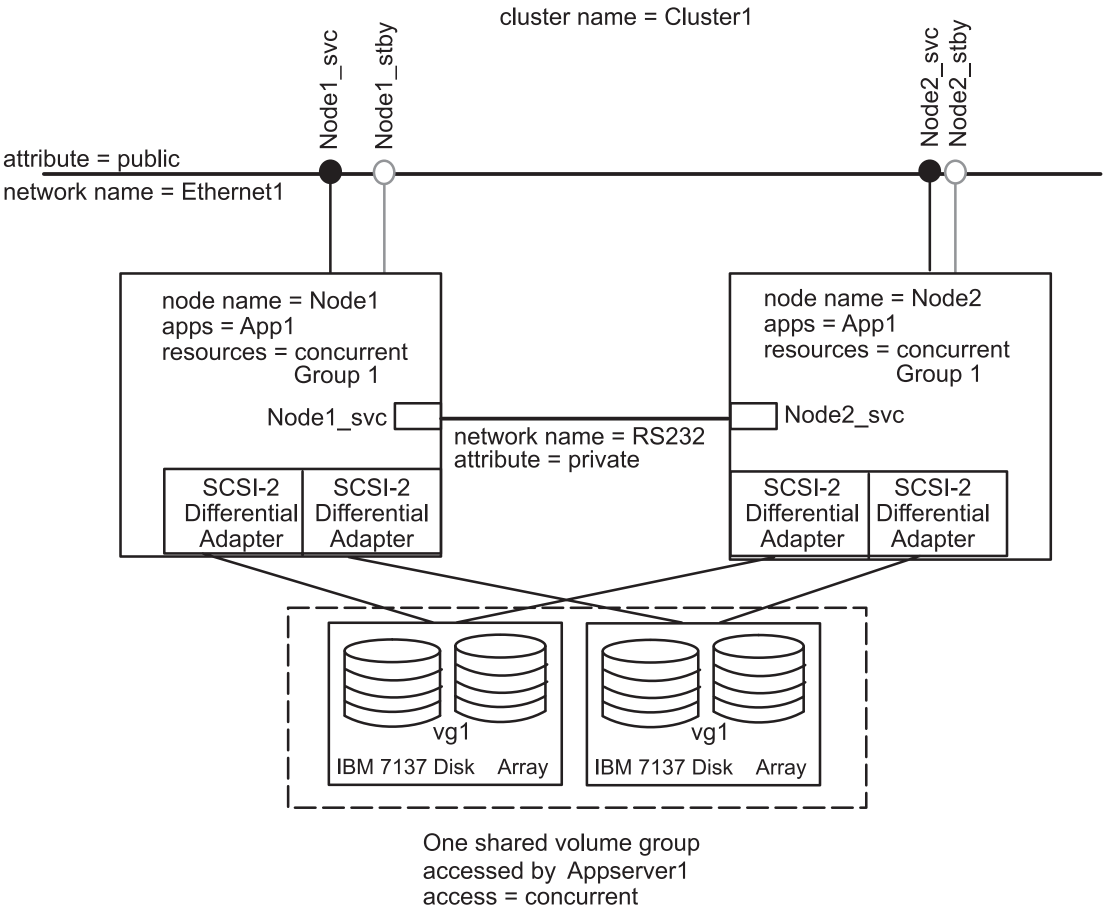
Quick Configuration 2: Cluster Diagram with Network Topology:

Shared IBM 7204 Disk Subsystems and Bus Configuration

Configuration 3: Two IBM 7137 Disk Arrays, One Concurrent Access Resource Group

This configuration requires:

- Two nodes, two Ethernet adapters per node.
- One Ethernet network connecting the two nodes and their clients.
- One RS232 link (private network) between the two nodes.
- Two IBM 7137 Disk Arrays shared between the two nodes, connected to two SCSI adapters in each node.
- Two SCSI adapters per node (for external disks).
- One shared volume group.
- One concurrent resource group, Group 1, consisting of an application and the volume group.



Quick Configuration 3: Cluster Diagram with Network Topology:

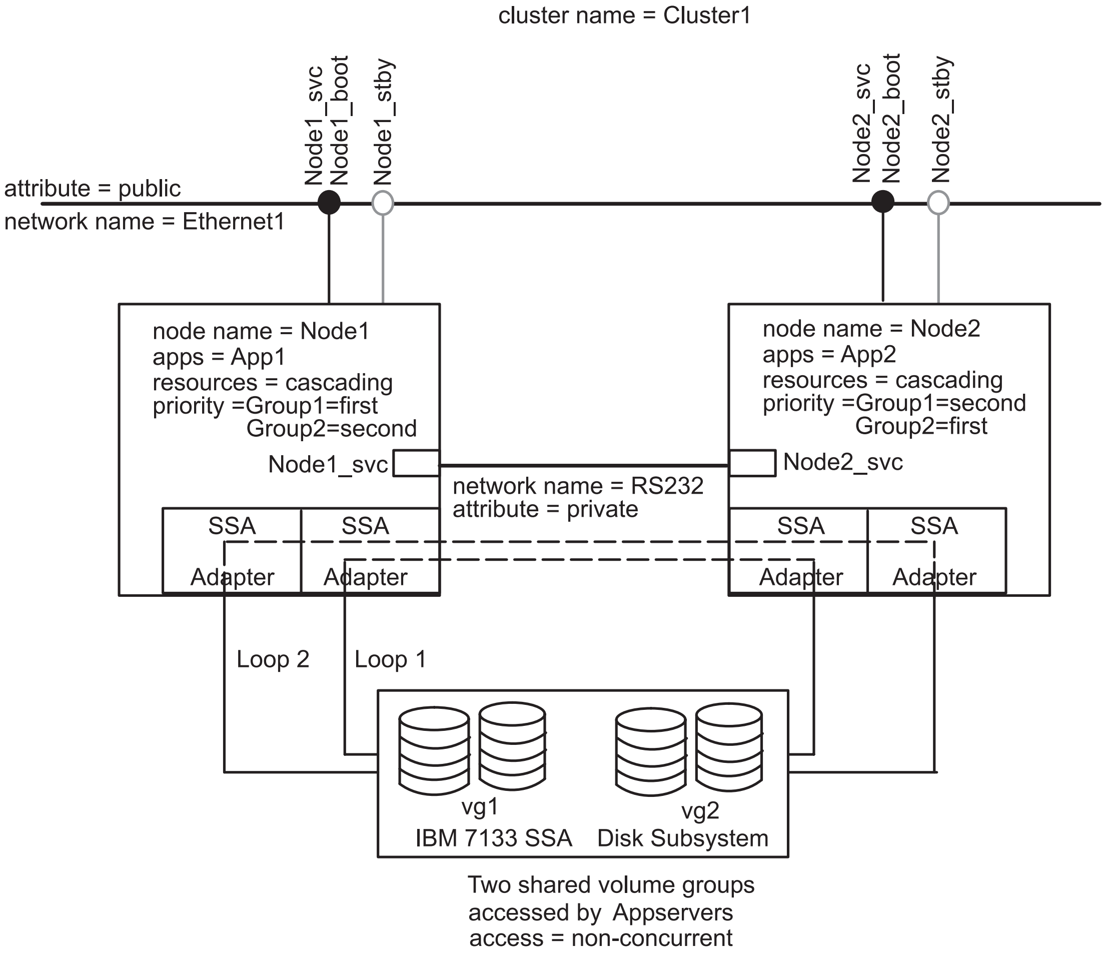
Shared IBM 7137 Disk Array Configuration

Note: The IBM 7131-405 Disk Subsystem also can be used in concurrent access mode.

Configuration 4: One IBM 7133 Serial Storage Architecture (SSA) Disk Subsystem, Two Cascading Resource Groups

This configuration requires:

- Two nodes, two Ethernet adapters per node.
- One Ethernet network connecting the two nodes and their clients.
- One RS232 link (private network) between the two nodes.
- One IBM 7133 SSA Disk Subsystem shared between the two nodes. One SSA adapter in each node is connected to the disk subsystem to make a closed loop configuration. The IBM 7133 SSA Disk Subsystem contains four physical disks.
- Two SSA adapters per node.
- Two volume groups, with single mirrors, on the IBM 7133 SSA Disk Subsystem.
- Two cascading resource groups, each consisting of an application, one of the shared volume groups, and the IP address for the server application. Each node is the initial owner of one of the resource groups.



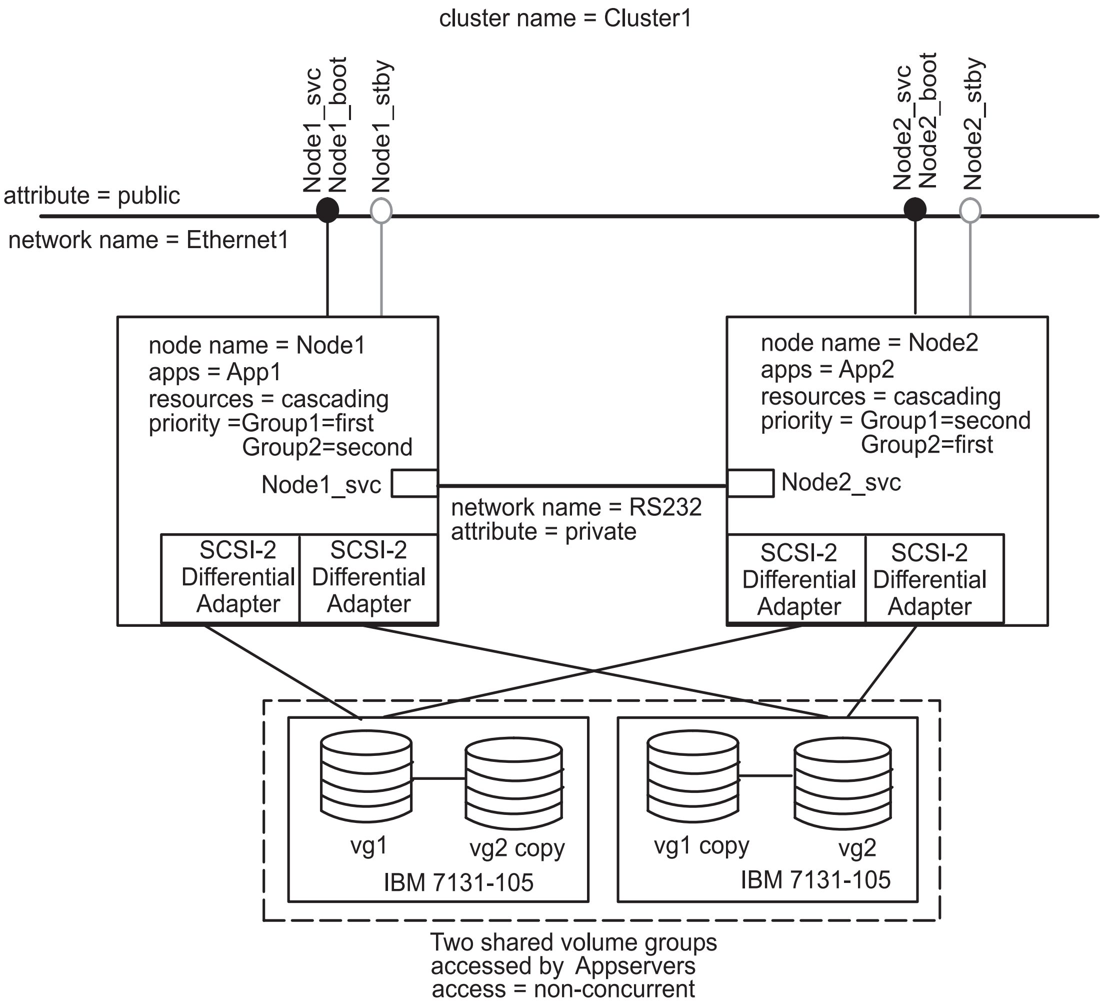
Quick Configuration 4: Cluster Diagram with Network Topology:

Shared IBM 7133 SSA Disk Subsystem Configuration

Configuration 5: Two IBM 7131-105 Disk Subsystems, Two Cascading Resource Groups

This configuration requires:

- Two nodes, two Ethernet adapters per node.
- One Ethernet network connecting the two nodes and their clients.
- One RS232 link (private network) between the two nodes.
- Two IBM 7131-105 disk subsystems shared between the two nodes, each connected to two SCSI adapters. Each IBM 2131-105 disk subsystem is on a separate SCSI bus.
- Two SCSI adapters per node (for external disks).
- Two volume groups, mirrored from one IBM 7131-105 disk subsystem to the other.
- Two cascading resource groups, each consisting of an application, one of the shared volume groups, and the IP address for the server application. Each node is the initial owner of one of the resource groups



Quick Configuration 5: Cluster Diagram with Network Topology

Shared IBM 7131-105 SCSI Disk Subsystem Configuration

Installing the Hardware for a Quick Configuration

Install the hardware for the chosen configuration following the guidelines in the cluster diagram. Consult your AIX documentation for both the hardware and software setup of disk devices.

Installing Shared Disk Devices

To install disk devices, follow the instructions for the device specified in Chapter 5, Installing Shared Disk Devices.

Installing an RS232 Serial Line

An RS232 serial network connecting the two nodes in a Quick Configuration is required. The serial network allows Cluster Managers on cluster nodes to continuously exchange keepalive packets should the TCP/IP-based subsystem, networks, or network adapters fail. Thus, the serial network prevents nodes from becoming isolated and from attempting to take over shared resources. See Chapter 4, Configuring Cluster Networks and Performance Tuning, for more information on configuring an RS232 serial line.

Quick Configuration Prerequisites and Guidelines

In addition to correctly installing the hardware required to match one of the five predefined configurations, make sure you meet the prerequisites for communication, as described in this section.

You can choose to have the Quick Configuration utility configure devices on nodes as necessary for a predefined configuration, or you can choose to preconfigure certain critical parameters for the utility to use.

If you preconfigure the devices, the Quick Configuration utility verifies that they are compatible with the chosen configuration. The Quick Configuration utility does not, however, automatically fix any errors it finds; instead, it simply informs you of those errors. You must then manually correct the error and rerun the utility.

Prerequisites for Quick Configuration Communication

To use the Quick Configuration utility, you must fulfill the following communication requirements:

- Each cluster node to be configured must have at least one of its Ethernet adapters properly configured so that it can communicate with the system to be used as the configurator (the node you will use to run the Quick Configuration utility to do the cluster configuration). You can use one cluster node as the configurator, or specify another system.

  Use the **smit mktcpip** fastpath to define the IP label (IP address) and network mask for the adapter.

  The hostname should match the adapter label of the Ethernet network’s service adapter because some applications may depend on the hostname (though this is not required for HACMP for AIX).

- Use the **smit tty** fastpath to define the device on each node that will be connected to the RS232 line.
- Each node must have an entry for the configurator in both the **/etc/hosts** and the **/.rhosts** file to allow the configurator to run commands on each node.

  Note: If your cluster uses a **nameserver** configuration, use SMIT to create the **/etc/resolv.conf** file, which also must have an entry for the configurator.

- Each disk must have a Physical Volume Identifier (PVID) so that the Quick Configuration utility can automatically configure shared volume groups, logical volumes, and filesystems. The Quick Configuration utility checks PVIDs to determine which disks are shared among the cluster nodes.

  When a disk is newly installed or defined, a PVID may not yet be associated with the disk. Associate a PVID with a disk before attempting to apply one of the Quick Configuration utility configurations.

  Use the **smit chgdsk** fastpath to associate a PVID with a disk.

  Note: This procedure changes the AIX LVM entries but not the HACMP ODM.

- Use the Quick Configuration utility only when all nodes are in their initial state; that is, no fallover has occurred and the nodes are on their configured service/boot addresses. Using the Quick Configuration utility on a cluster after a fallover can produce unpredictable results.

Once these prerequisites are met, the Quick Configuration utility configures any further necessary network devices. These include the Ethernet service, boot, and standby adapters. The utility aims to choose defaults that fit the HACMP for AIX software requirements and that do not interfere with site-specific parameters. (The latter aim, however, cannot be guaranteed).

Guidelines for Preconfiguring Network Parameters

The Quick Configuration utility follows a series of steps, each of which has a corresponding screen. After choosing the configuration, you next assign IP addresses to the template nodes, *Node1* and *Node2*. You are prompted to enter an IP address for each node. When you do this, the physical nodes are mapped to the configuration template. The utility proceeds to configure the network and disk subsystems using either default or preconfigured parameters.

The following section describes how the Quick Configuration utility chooses the default network parameters, then shows how you can preconfigure the devices so the utility uses your preconfigured parameters instead of the defaults.

Ethernet Service Addresses

For Configurations 2 through 5, the addresses you enter into the Quick Configuration utility Select IP Address screen are service addresses. The utility derives the IP label associated with the address and uses it as the service IP label, and also as the node name.

For Configuration 1, in which the service IP label is shared between the two nodes, the IP addresses you enter into the Quick Configuration utility Select IP Address screen are used as boot addresses. The IP label associated with a boot address is used as the node name. The utility prompts you for the IP label of the shared service address.

Ethernet Boot Addresses

For Configurations 2 through 5, the Quick Configuration utility defines the IP label for the boot adapter by adding an extension of *_boot* to the service IP label. For example, if the service IP label is *clam*, then the boot IP label is *clam_boot*.

If a matching entry exists in the **/etc/hosts** or **/etc/resolv.conf** files, the corresponding IP address will be used. If not, the Quick Configuration utility determines an unused IP address (using **ping**) by adding or subtracting a number to or from the service IP address.

Note: The boot and service addresses must be on the same logical subnet.

For Configuration 1, the boot IP address is determined as defined in the previous section.

Ethernet Standby Addresses

For all configurations, the Quick Configuration utility determines the standby IP label by adding an extension of *_stby* to the node name.

If a matching entry exists in the **/etc/hosts** or **/etc/resolv.conf** files, the corresponding IP address is used. If not, the Quick Configuration utility determines an unused IP address (using **ping**) by adding or subtracting a number to or from the network portion of the service IP address.

Note: The standby addresses must be on a logical subnet different from the service and boot address subnet.

RS232 Service Devices

The Quick Configuration utility places the defined **tty** devices in a pick list on each node’s Customization Window RS232 Interfaces Service IP Label box. You must select the correct **tty** device from the pick list.

If you are unsure of the **tty** device attached to a node, refer to the section Configuring an RS232 Serial Line on page 4-10 to determine if the **tty** device you intend to use is defined or available. If the **tty** device is not defined, follow the instructions in that section to define it.

Using Defaults or Preconfiguring Network Devices

If the Quick Configuration utility configures a network device using an IP address and an IP label that it generates, it adds the necessary entries to each node’s **/etc/resolv.conf** and **/.rhosts** files. If the Quick Configuration utility uses predefined IP addresses and IP labels, it verifies the existence of such entries on each node and, if entries are not found, it issues warnings to the user to add the entries to the files as necessary.

If desired, you can preconfigure the Ethernet network interfaces by defining a service, boot, and standby IP label in each node’s **/etc/hosts** file. For example, you might define the IP labels *clam*, *clam_boot*, and *clam_standby* and, given a network mask of 255.255.255.0, their associated IP addresses (for example, *100.100.100.1*, *100.100.100.2*, *100.100.101.1*, respectively). The Quick Configuration utility then automatically uses these values and, if necessary, configures the devices.

If you choose to preconfigure network devices, see Chapter 4, Configuring Cluster Networks and Performance Tuning, for additional information and instructions.

Prerequisites for Quick Configuration Disk Subsystems

The only prerequisite required for disk subsystems is that the hardware be properly installed. In fact, for any of the five predefined configurations, it is much easier to let the Quick Configuration utility configure the volume groups, logical volumes, and optional filesystems than to do any preconfiguration.

When you select a configuration, the Quick Configuration utility checks for the proper hardware configuration and creates the necessary volume groups. The utility prompts you to enter the volume group names, then it creates the logical volumes and properly mirrors them. Next it optionally creates the filesystems (except for Configurations 3 and 5, which use a concurrent access volume group on which filesystems may not exist). In the process of creating filesystems, the utility prompts you to choose the filesystem mount point. The size of the filesystem will be the maximum allowed for the size of the volume groups created, with space for mirrored copies.

If you find that the default parameters are not appropriate for your particular installation, the volume groups, logical volumes, and filesystems can be predefined. If the Quick Configuration utility finds an existing volume group and associated filesystem on a disk or set of disks shared between the nodes, no configuration is performed, except for the possible import of the volume group to one of the nodes if necessary. The Quick Configuration utility uses the preconfigured entities as part of the configuration.

The Quick Configuration utility also verifies that the disk subsystem parameters are properly configured for an HACMP for AIX cluster environment, and it informs you of problems detected. You must manually correct any errors found and rerun the verification as necessary.

If you choose to preconfigure disk subsystem parameters, see Chapter 6, Defining Shared LVM Components, for additional information and instructions.

Guidelines for Preconfiguring Application Servers

If you plan to use applications that you want to configure to run in your cluster environment, the following start and stop script templates in the **/usr/sbin/cluster/local** directory can be modified to include site specific information:

```
start_AppServer1, start_AppServer2
stop_AppServer1, stop_AppServer2
```

These templates initially are empty and are customizable. See Chapter 12, Configuring Cluster Resources, for more information about configuring application servers, and refer to your Application and Application Server worksheets in Appendix A of the *HACMP for AIX Planning Guide* when configuring application resources.

Where You Go From Here

To complete the remaining required steps *before* running the Quick Configuration utility:

- Read Chapter 7, Additional AIX Administrative Tasks, to ensure that all requirements for proper performance of the HACMP cluster are met for each node.
- Install the HACMP for AIX, Version 4.4 software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software. Chapter 9, Upgrading an HACMP Cluster, describes how to upgrade your existing HACMP for AIX software to Version 4.4.
- Verify the software used in the cluster following the instructions in Chapter 10, Verifying Cluster Software.
- (optional, but strongly recommended) Complete the worksheets in Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide* for your configuration. This information provides a useful record of your configuration in case you later add to or change it.

When you have finished these tasks, you are ready to run the Quick Configuration utility. See Chapter 3, Using the Quick Configuration Utility, for more information.

Chapter 3 Using the Quick Configuration Utility

This chapter describes how to use the HACMP for AIX Quick Configuration utility, **xclconfig**, to install and configure a two-node HACMP cluster using one of the five predefined configurations discussed in Chapter 2, Setting Up a Quick Configuration Cluster.

For information about managing a two-node cluster, see the *HACMP for AIX Concepts and Facilities Guide* for a description of the C-SPOC utility. See the *HACMP for AIX Administration Guide* for information on using the utility.

Prerequisites

- Install and configure the hardware and software as described in Chapter 2, Setting Up a Quick Configuration Cluster.
- Install the HACMP for AIX, Version 4.4 software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software. Chapter 9, Upgrading an HACMP Cluster, describes how to upgrade your existing HACMP for AIX software to Version 4.4.
- Verify the cluster software using the instructions in Chapter 10, Verifying Cluster Software.
- (optional) Complete the planning worksheets in Appendix A of the *HACMP for AIX Planning Guide* for your configuration.

Overview

The Quick Configuration utility is an X Window System application designed to simplify the task of configuring an HACMP cluster. You choose to configure one of four predefined, two-node cluster configurations. The utility automates the process for you.

You can customize the chosen configuration by changing the names for the cluster, nodes, networks, and resource groups. Extensive on-line help information is provided for each step. This chapter complements that information.

This chapter also provides a description of the utility, explaining how to use the buttons, menus, and icons on the displays; and it walks you through the steps involved in using the utility, showing the display for each of the following steps:

1. Invoking the utility with the **xclconfig** command
2. Selecting one of the five cluster configurations
3. Defining the IP address for each node
4. Customizing the cluster configuration (optional)
5. Verifying the cluster configuration
6. Applying the cluster configuration.

Step 1: Invoking the Quick Configuration Utility

Before invoking the Quick Configuration utility, be sure that you have completed the prerequisite tasks. You must have chosen the desired predefined configuration and installed and configured the necessary hardware and software on each node. Keep your worksheets handy for reference.

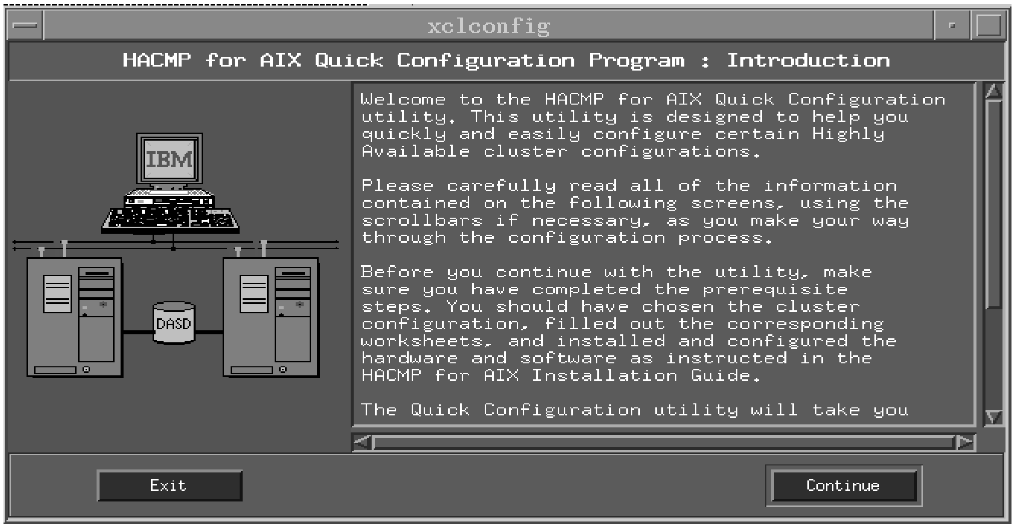
Make sure that both cluster nodes allow remote root access for the configurator. The configurator can be one of the nodes to be configured, or it can be another node that has the HACMP for AIX software installed, and thus has network connectivity to the cluster nodes. The HACMP for AIX software must be installed on the cluster nodes, but not yet running.

To invoke the Quick Configuration utility, enter:

```
xclconfig
```

Note: Do not run xclconfig in the background, as the utility may hang when it reaches the IP address selection stage.

The **xclconfig** window appears. The Introduction screen shows an HACMP cluster. The Help Information provides introductory information on the application’s function.



HACMP for AIX Quick Configuration Program : Introduction

Welcome to the HACMP for AIX Quick Configuration utility. This utility is designed to help you quickly and easily configure certain Highly Available cluster configurations.

Please carefully read all of the information contained on the following screens, using the scrollbars if necessary, as you make your way through the configuration process.

Before you continue with the utility, make sure you have completed the prerequisite steps. You should have chosen the cluster configuration, filled out the corresponding worksheets, and installed and configured the hardware and software as instructed in the HACMP for AIX Installation Guide.

The Quick Configuration utility will take you

Mouse Buttons

Unless otherwise noted, use the first (left) mouse button to select items and to click on buttons.

xclconfig Window

When the **xclconfig** window first appears, it displays the Introductory screen. The window itself includes the following features:

- The title bar
- The form area
- The Help Information area
- The control area.

Title Bar

The title bar is at the top of the window. The title bar tells you where you are in the sequence of steps, from the Introduction through the Apply screen. The initial title caption displays “HACMP for AIX Quick Configuration Program: Introduction.” As you proceed through the application, the title bar caption changes to show the name of the current screen.

Form Area

The form area occupies the left side of the window. This is the area where different forms (usually a box or a prompt) appear as you step through the configuration process. On the Introductory screen, this area shows a picture of a two-node HACMP cluster. On the succeeding screens, you will be prompted to enter information (customize the configuration) or perform actions.

Help Information Area

The Help Information area occupies the right side of the window. It shows help text related to the current step. This text area is resizable, and it also contains vertical and horizontal scroll bars. Whenever a new form occupies the form area, the Help Information area displays the corresponding text to guide you through the functions on the screen.

When commands are run, such as those used for retrieving and applying data, the command output is echoed to the **/tmp/xclconfig.log** file; this file is replaced with new data each time you run the **xclconfig** utility. Errors are printed to the screen from which you are running the utility. Messages informing you of the basic actions being performed appear in the Help Information Area. For example, “The configuration template is being read...” appears after you choose a configuration.

Additional Help Windows

Message and prompt boxes may appear under certain conditions. Message boxes inform you of errors or conditions which require you to take an action before continuing. Prompt boxes appear if you need to enter or choose information needed by the program.

Control Area

The control area at the bottom of the window contains navigational buttons: **Exit**, **Previous Menu**, and **Continue**.

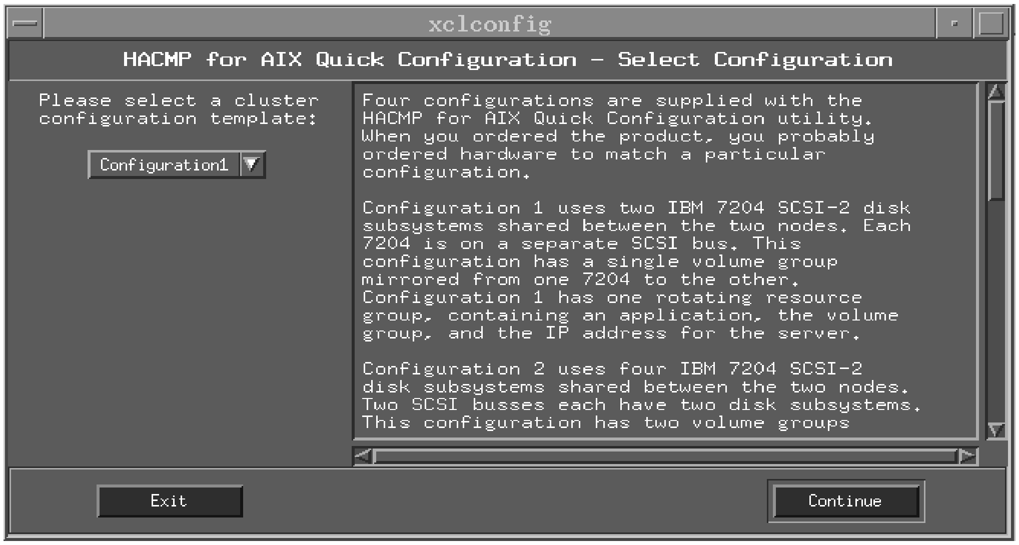
- Use the **Exit** button at any time to exit the application. If you have made changes when you choose to exit the application, the utility prompts you to save the data as a new configuration. You may define the new configuration name, and you can access it later. Note however that the **xclconfig** utility can only read in the predefined configuration templates, so it will not be able to read in the saved file. You must use the **clsnapshot** or **xhacmpm** utility to read in this file. See the chapter on saving cluster configurations in the *HACMP for AIX Administration Guide* for more information.
- Use the **Previous Menu** button to return to a previous screen. This action does not undo any changes made, unless you go all the way back to the Choose Configuration screen (screen 2). Then you must start over.

- Use the **Continue** button to move to the next screen. Clicking on this button generally commits changes you have made. If you change your mind or realize you made a mistake, use the **Previous Menu** button and make the changes, then click on the **Continue** button to apply the changes. The configuration is not written to the ODM until you click on the **Apply** button in the final step of the process.

ACTION:
Click on the **Continue** button to move to the next screen.

Step 2: Selecting a Cluster Configuration

The **xclconfig** window now displays the second screen. The title bar reads “HACMP for AIX Quick Configuration – Select Configuration.” Configuration 1 is shown in a box, with an arrow next to it



HACMP for AIX Quick Configuration - Select Configuration

Please select a cluster configuration template:

Configuration1

Four configurations are supplied with the HACMP for AIX Quick Configuration utility. When you ordered the product, you probably ordered hardware to match a particular configuration.

Configuration 1 uses two IBM 7204 SCSI-2 disk subsystems shared between the two nodes. Each 7204 is on a separate SCSI bus. This configuration has a single volume group mirrored from one 7204 to the other. Configuration 1 has one rotating resource group, containing an application, the volume group, and the IP address for the server.

Configuration 2 uses four IBM 7204 SCSI-2 disk subsystems shared between the two nodes. Two SCSI busses each have two disk subsystems. This configuration has two volume groups

The Help Information provides basic instructions for selecting a configuration, reminding you to choose the configuration that matches your hardware and your worksheets. If you scroll down, you can read additional text describing each configuration.

The Select Configuration Form prompts you to select one of the four predefined configurations. Clicking on the arrow or on the box itself provides a list of the available configurations. The configurations are described below. You should have already installed and configured the hardware corresponding to one of the following configurations:

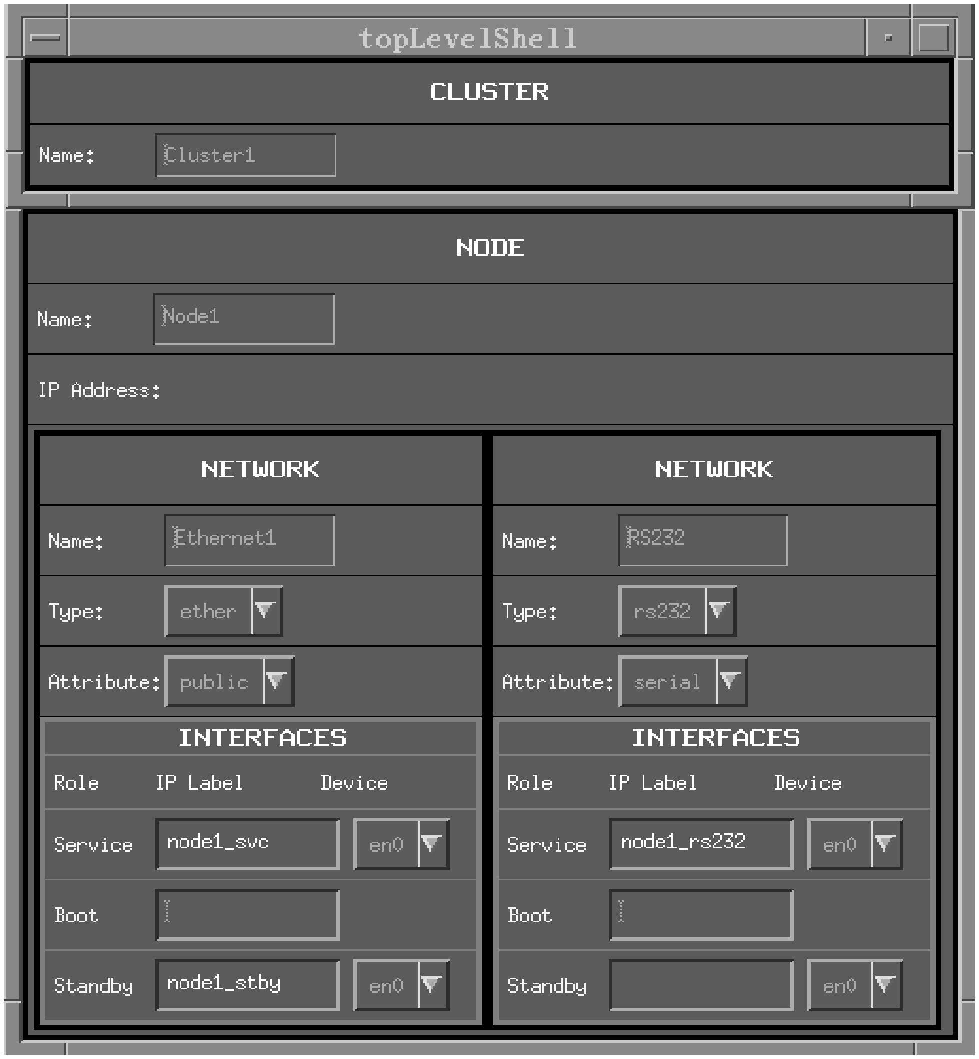
- Configuration 1: Two IBM 7204 SCSI-2 disk subsystems, one rotating resource group
- Configuration 2: Two IBM 7204 SCSI-2 disk subsystems, two cascading resource groups
- Configuration 3: Two IBM 7137 SCSI-2 Disk Arrays, one concurrent access resource group
- Configuration 4: One IBM 7133 SSA disk subsystem, two cascading resource groups
- Configuration 5: Two IBM7131-105 Disk Subsystems, two cascading resource groups

ACTION:

Select your configuration and then click on the *Continue* button. The chosen configuration is stored in a temporary file, and the Customization windows are displayed with the default values for the chosen configuration.

Note: The Customization windows display all the configurable cluster parameters at once. Four or five windows may appear on your display. These are the cluster, node (two), and resource group (one or two) windows. Move them to a convenient space on your display. Be aware that closing a Customization window will cause **xclconfig** to halt when attempting to perform the remaining steps listed in this chapter.

The following figures show samples of the cluster, node, and resource group windows.



CLUSTER

Name: Cluster1

NODE

Name: Node1

IP Address:

NETWORK

Name: Ethernet1

Type: ether

Attribute: public

INTERFACES

Role	IP Label	Device
Service	node1_svc	en0
Boot		
Standby	node1_stby	en0

NETWORK

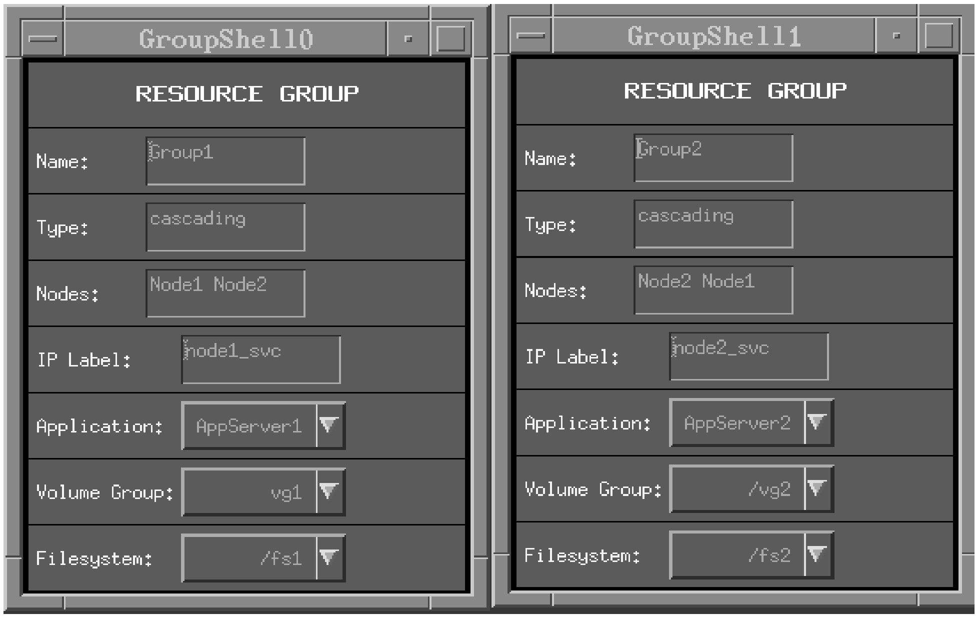
Name: RS232

Type: rs232

Attribute: serial

INTERFACES

Role	IP Label	Device
Service	node1_rs232	en0
Boot		
Standby		en0



GroupShell0

RESOURCE GROUP

Name:	Group1
Type:	cascading
Nodes:	Node1 Node2
IP Label:	node1_svc
Application:	AppServer1
Volume Group:	vg1
Filesystem:	/fs1

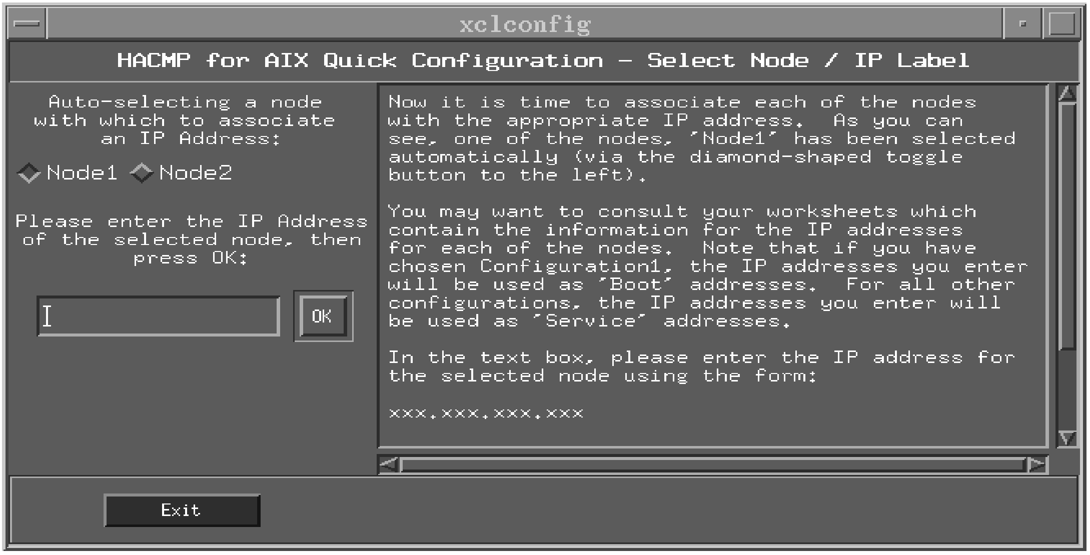
GroupShell1

RESOURCE GROUP

Name:	Group2
Type:	cascading
Nodes:	Node2 Node1
IP Label:	node2_svc
Application:	AppServer2
Volume Group:	/vg2
Filesystem:	/fs2

Step 3: Defining IP Addresses

The **xclconfig** window now displays the third screen. The title bar reads “HACMP for AIX Quick Configuration – Select Node / IP Label.” Each node is represented by a toggle button, one of which is selected automatically.



xclconfig

HACMP for AIX Quick Configuration – Select Node / IP Label

Auto-selecting a node
with which to associate
an IP Address:

◇ Node1 ◇ Node2

Please enter the IP Address
of the selected node, then
press OK:

```
Now it is time to associate each of the nodes
with the appropriate IP address.  As you can
see, one of the nodes, 'Node1' has been selected
automatically (via the diamond-shaped toggle
button to the left).

You may want to consult your worksheets which
contain the information for the IP addresses
for each of the nodes.  Note that if you have
chosen Configuration1, the IP addresses you enter
will be used as 'Boot' addresses.  For all other
configurations, the IP addresses you enter will
be used as 'Service' addresses.

In the text box, please enter the IP address for
the selected node using the form:

xxx.xxx.xxx.xxx
```

The Help Information provides basic instructions for associating your physical nodes with the abstract nodes on the screen. Refer to your worksheets and fill in the correct IP addresses for the nodes in your system.

The Select IP Address Form prompts you to define the selected node's IP address. Type in the IP address at the prompt; then click on the **OK** button. The other node is now selected. Define its IP address and click on the **OK** button.

The IP address entered is used as the service address for a node in a non-rotating configuration (2, 3, 4), and as a boot address for a node in a rotating configuration (1).

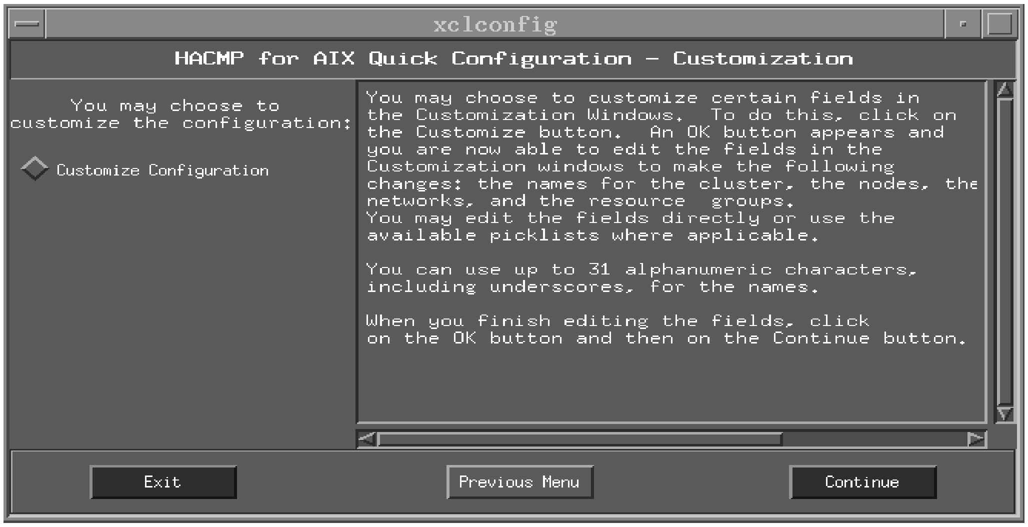
The IP address must be preconfigured as an HACMP IP address.

ACTION:
Select each node in turn and assign the proper IP address according to your configuration; then click on the **Continue** button. This button appears after you assign IP addresses. The program retrieves information from each node to fill in the selection values in the Customization windows.

Note: The Node Customization windows now display the chosen IP addresses for each node (and the actual node names); the Resource Group Customization windows display preconfigured or default volume group and filesystem configuration information.

Step 4: Customizing the Cluster Configuration

The **xclconfig** window now displays the fourth screen. The title bar reads "HACMP for AIX Quick Configuration – Customization." The form area contains one toggle button. The Customization Form lets you select the toggle button to customize the configuration.


xclconfig
HACMP for AIX Quick Configuration – Customization
You may choose to customize the configuration:
Customize Configuration
Exit
Previous Menu
Continue

```
You may choose to customize certain fields in
the Customization Windows.  To do this, click on
the Customize button.  An OK button appears and
you are now able to edit the fields in the
Customization windows to make the following
changes: the names for the cluster, the nodes, the
networks, and the resource  groups.
You may edit the fields directly or use the
available picklists where applicable.

You can use up to 31 alphanumeric characters,
including underscores, for the names.

When you finish editing the fields, click
on the OK button and then on the Continue button.
```

The Help Information provides instructions for customizing your system.

Using the Defaults

If you want to use the current system settings as determined by the Quick Configuration utility and as reflected in the Customization windows, click on the **Continue** button. These parameters are saved to the temporary file. The Apply process then writes this configuration to the HACMP for AIX ODM when you complete Step 6 by clicking on the **Apply** button.

Customizing the Current System Settings

If you want to customize your system by changing one or more of the current fields, click on the **Customize Configuration** toggle button. The fields you can select for editing in the Customization windows become active, and an *OK* button appears under the **Customize Configuration** toggle button.

The information gathered from the nodes defined in the configuration is offered as pick values in the editable fields on the Customization windows. Default parameters are displayed for all items for which preconfigured data was not found.

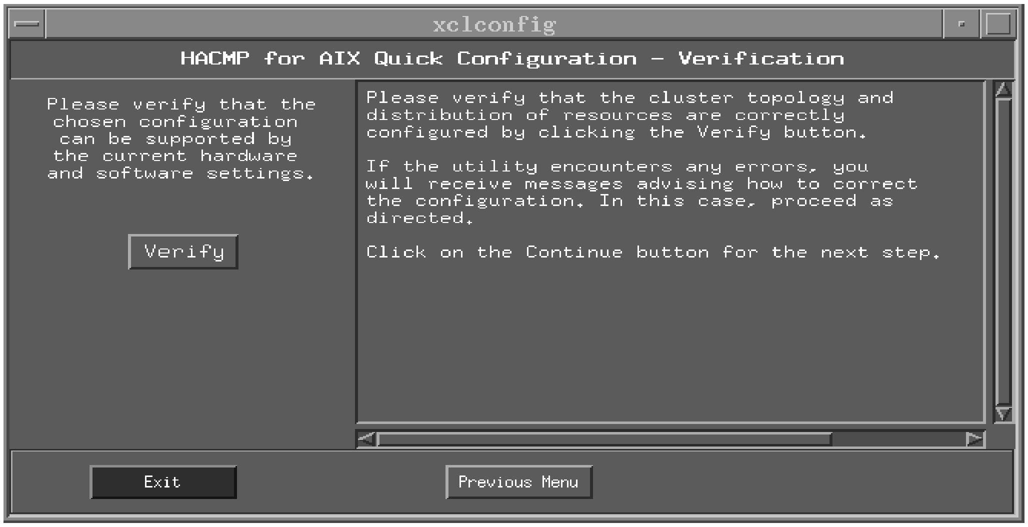
You can edit these fields or select values available in the picklists. Make needed changes according to the information on your worksheets. When you finish, click on the **OK** button; then click on the **Continue** button to save the values to the temporary file.

ACTION:
Click on the **Customize Configuration** toggle and then click on the **OK** button. Next click on the **Continue** button. The configuration is read from file and stored in a temporary ODM directory.

Step 5: Verifying the Cluster Configuration

The **xclconfig** window now displays the fifth screen. The title bar reads “HACMP for AIX Quick Configuration – Verification.” The form area contains a single **Verify** button.

The Help Information provides instructions for verifying your system.



Please verify that the chosen configuration can be supported by the current hardware and software settings.

Please verify that the cluster topology and distribution of resources are correctly configured by clicking the Verify button.

If the utility encounters any errors, you will receive messages advising how to correct the configuration. In this case, proceed as directed.

Click on the Continue button for the next step.

When you click on the **Verify** button, the system runs the HACMP for AIX **/usr/sbin/cluster/diag/clverify** utility. Output is displayed in the Help Information area, and is also logged to the **/tmp/qconfig/clverify.log** file. If errors are noted, correct them as instructed. Verification succeeds if the required hardware and cluster topology are available and properly configured.

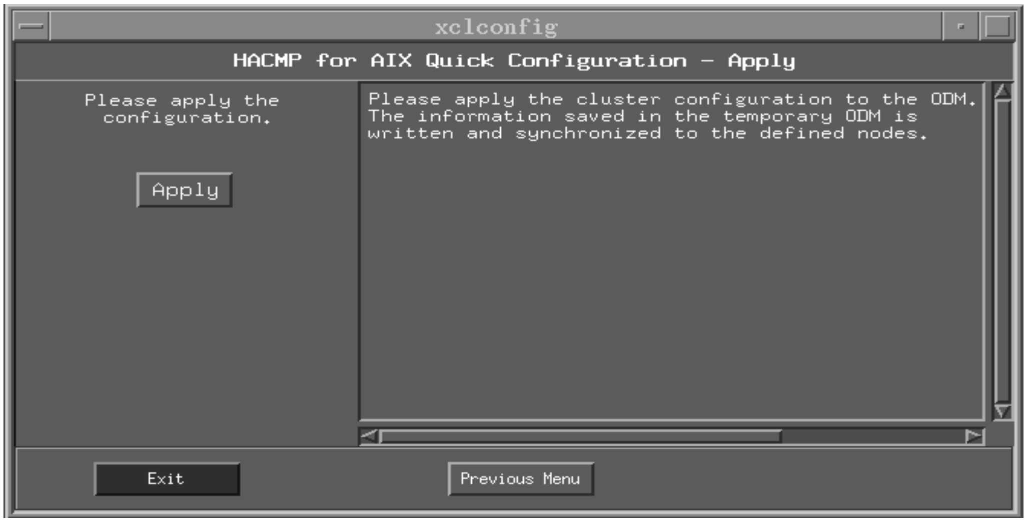
ACTION:

Click on the **Verify** button. If no errors exist, click on the **Continue** button. If errors exist, correct them as instructed and click on the **Verify** button again. Then click on the **Continue** button.

Step 6: Applying the Cluster Configuration

The **xclconfig** window now displays the sixth screen. The title bar reads “HACMP for AIX Quick Configuration – Apply.” The form area contains a single **Apply** button.

The Help Information provides instructions for applying the configuration to your system.



Please apply the configuration.

Please apply the cluster configuration to the ODM. The information saved in the temporary ODM is written and synchronized to the defined nodes.

The Apply Form prompts you to click on the **Apply** button.

When you click on the **Apply** button, the system synchronizes the HACMP for AIX ODM classes (as defined in the temporary ODM) to the defined nodes.

ACTION:
Click on the **Apply** button.

Where You Go From Here

When you have successfully completed the configuration process, do the following to complete the cluster configuration:

- Set up **clinfo** on cluster nodes. See Chapter 15, Setting Up Clinfo on Server Nodes.
- (optional) Set up **clinfo** on clients. See Chapter 17, Installing and Configuring Clients.
- (optional) Use the C-SPOC utility described in Chapter 2 of the *HACMP for AIX Administration Guide* to start or stop cluster services on multiple nodes.

Part 3 Installing HACMP Cluster Configurations: Standard Method

This part contains instructions for installing and configuring an HACMP for AIX cluster.

Chapter 4 Configuring Cluster Networks and Performance Tuning

This chapter lists the required steps to configure TCP/IP network adapters (including SOCC optical links and SLIP lines), RS232 serial lines, and supported network types. Consult your *AIX System Management Guide* for general assistance.

Prerequisites

Complete the networking worksheets discussed in Chapter 3, Planning TCP/IP Networks, and Chapter 4, Planning Serial Networks, of the *HACMP for AIX Planning Guide*.

Overview

The steps in setting up the network support for an HACMP cluster include:

- Configuring network adapters
- Configuring a SOCC optical link
- Configuring a SLIP line, if required
- Configuring serial networks
- Configuring for Asynchronous Transfer Mode (ATM)
- Configuring RS/6000 SP Switches

Related Tasks

This chapter lists the steps you must complete to configure a network adapter in the AIX environment. After configuring an adapter, you must also define it to the HACMP cluster topology. See Chapter 11, Defining the Cluster Topology, for more information. Other related tasks include setting network options and editing certain configuration files. See Chapter 7, Additional AIX Administrative Tasks, for more information.

Configuring Ethernet, Token-Ring, and FDDI Adapters

Complete the following steps for each service and standby adapter listed on your completed copies of the *TCP/IP Networks Adapter Worksheet*. Repeat the procedure for each node in your cluster.

1. Use the **smit mktcpip** fastpath to define the IP label, IP address, and network mask for each adapter.

   It is important that you specify through SMIT the IP label, IP address, and network mask for each adapter in order to store these values in the HACMP ODM. Some functions use these ODM values and will not work properly if they are not present.

When using the SMIT interface to define an adapter, the HOSTNAME field changes the default hostname. For instance, if you configure the first adapter as *clam_svc,* and then configure the second adapter as *clam_stby,* the default hostname at system boot is *clam_stby*. To avoid this problem, configure the adapter with the desired default hostname last.

Also, the hostname should match the adapter label of the primary network's service adapter because some applications may depend on the hostname (though this is not required by the HACMP for AIX software).

2. Use the **smit chinet** fastpath to configure each adapter, for which IP address takeover might occur, to boot from the boot adapter address and not from its service adapter address. This step applies if you are using IP address takeover or rotating resource groups. Refer to your completed copies of the *TCP/IP Network Adapter Worksheet* for the required information.

3. After configuring the network interfaces on a node, record the interface names on that node's *TCP/IP Network Adapter Worksheet.* For a listing of “Available” and “Defined” adapters for the node, enter:

   ```
   lsdev -Cc if
   ```

   At this point, all interfaces used by the HACMP for AIX topology should be “Available.” List the adapters marked “Available” in the **Interface Name** field on the *TCP/IP Network Adapter Worksheet.*

Defining Ethernet, Token-Ring, and FDDI Adapters

After you have installed and tested Ethernet, Token-Ring, or FDDI adapters, you must define them to the HACMP cluster topology. Chapter 11, Defining the Cluster Topology, describes how to define an adapter in an HACMP cluster.

Note: The netmask for all adapters in an HACMP network must be the same to avoid communication problems between standby adapters after an adapter swap. The communication problem occurs when the standby adapter assumes its original address but retains the netmask of the takeover address.

Configuring a SOCC Link

Configuring a Serial Optical Channel Converter (SOCC link) requires the following steps:

1. Use the **smit ops** fastpath to configure the SOCC adapter.
2. Use the **smit mktcpip** fastpath to configure TCP/IP on the SOCC adapter. An “so” indicates a SOCC link.
3. Set the **START Now** field to **no**.

After configuring all the network interfaces for a node, record the network interface names on that node's *TCP/IP Network Adapter Worksheet.* For a listing of “Available” and “Defined” adapters for the node, enter:

```
lsdev -Cc if
```

At this point, all interfaces used by the HACMP for AIX system should be “Available.” List the adapters marked “Available” in the **Interface Name** field on the *TCP/IP Network Adapter Worksheet.*

Note: If configured nodes with SOCC links are rebooted, this command may show that the cluster came up without the links configured. If this happens, see the *HACMP for AIX Troubleshooting Guide* for more information.

Defining SOCC Adapters

After you have installed and tested the SOCC adapters, you must define them to the HACMP cluster topology. Chapter 11, Defining the Cluster Topology, describes how to define an adapter in an HACMP cluster.

Configuring a SLIP Line

Configuring a Serial Line Internet Protocol (SLIP) line requires the following steps:

1. Use the **smit tty** fastpath to create a **tty** device.
2. Use the **smit inet** fastpath to configure the SLIP line.
3. Test communication over the SLIP line.
    - Run the **netstat -i** command to make sure the SLIP line is recognized. You should see the device listed as **sl1**.
    - On the first node, enter:

      ```
      ping IP_address_of_other_node
      ```

      where *IP_address_of_other_node* is the address in dotted decimal that you configured as the destination address for the other node.
    - Do the same on the second node, entering the destination address of the first node:

      ```
      ping IP_address_of_other_node
      ```

You must perform the first two steps on each node connected to the SLIP line, and then test the SLIP line. After configuring the SLIP network interfaces for a node, record the SLIP interface names on that node's *TCP/IP Network Adapter Worksheet.*

Defining the SLIP Line

After you have installed and tested a SLIP line, you must define it to the HACMP cluster topology. Chapter 11, Defining the Cluster Topology, describes how to define a SLIP line as an IP address in an HACMP cluster.

When you define the device as an adapter to the HACMP cluster topology, consider using the **sl** number assigned by AIX to the serial device. If this device is assigned *sl1*, then the adapter name should end with the characters "sl1" (for example, *clam_sl1*).

Configuring Serial Networks

This section summarizes the steps required to configure a SCSI-2 bus, an SSA connection using Multi-Initiator RAID adapters, or a raw RS232 serial line as a serial network in an HACMP cluster. Each step is explained as it occurs in the planning and installation process.

A serial network allows Cluster Managers to continuously exchange keepalive packets should the TCP/IP-based subsystem, networks, or network adapters fail. Thus, the serial network prevents nodes from becoming isolated and from attempting to take over shared resources. The serial connection can be a SCSI-2 Differential bus using target mode SCSI, a TMSSA loop, or a raw RS232 serial line. See Chapter 5, Installing Shared Disk Devices, for directions on configuring target mode SCSI and SSA connections. See page 19-6, for information on enabling and testing these connections.

Note: On the SP thin or wide nodes there are no serial ports available. Therefore, any HACMP configurations that require a tty network need to make use of a serial adapter card (8-port async EIA-232 adapter, FC/2930), available on the SP as an RPQ.

Also see the chapter on planning serial networks in the *HACMP for AIX Planning Guide* for a more general discussion of the purpose and types of serial networks.

Supported Serial Networks

The HACMP for AIX software supports three types of serial networks: the SCSI-2 Differential bus (using target mode SCSI), target mode SSA connection, and the raw RS232 serial line.

Target Mode SCSI

You can configure a SCSI-2 bus as an HACMP for AIX serial network only if you are using SCSI-2 Differential devices that support target mode SCSI. SCSI-1 Single-Ended and SCSI-2 Single-Ended devices do not support serial networks in an HACMP cluster; neither do PCI buses. The advantage of using the SCSI-2 Differential bus is that it eliminates the need for a dedicated serial port at each end of the connection, and for associated RS232 cables. See the next section, Configuring Target Mode SCSI Connections on page 4-5, for more information.

Target Mode SSA

You can configure a target mode SSA connection between nodes sharing disks connected to SSA on Multi-Initiator RAID adapters (FC 6215 and FC 6219). The adapters must be at Microcode Level 1801 or later.

You can define a serial network to HACMP that connects nodes on an SSA loop.

RS232 Serial Line

If you are using shared disk devices other than SCSI-2 Differential devices, you must use a raw RS232 serial line as the serial network. Note that each point-to-point RS232 serial network requires a dedicated serial port at each end. See the section later in this chapter, Configuring an RS232 Serial Line on page 4-10, for more information.

Configuring Target Mode SCSI Connections

This section describes how to configure a target mode SCSI-2 Differential bus as a serial network. The SCSI disks must be installed, cabled to the processors, and powered on.

Note: Neither PCI SCSI-2 differential busses nor SE busses support target mode SCSI.

Checking the Status of SCSI Adapters and Disks

To define a target mode SCSI connection, each SCSI adapter (controller) on nodes that will share disks on the SCSI bus must have a unique ID and must be “Defined,” known to the system but not yet available. Additionally, all disks assigned to an adapter must also be “Defined” but not yet available.

Note: The uniqueness of adapter SCSI IDs ensures that TMSCSI devices created on a given node do not reflect the SCSI IDs of adapters on other nodes connected to the same bus.

To check the status of SCSI adapters you intend to use, enter:

```
lsdev -C | grep scsi
```

If an adapter is “Defined,” see Defining Target Mode SCSI Devices in AIX on page 4-6 to configure the target mode connection.

To check the status of SCSI disks on the SCSI bus, enter:

```
lsdev -Cc disk
```

If either an adapter or disk is “Available,” follow the steps in the procedure below to return both the adapter (and its disks) to a defined state so that the adapters can be configured for target mode SCSI and made available.

Returning Adapters and Disks to a Defined State

Use the following command to make “Defined” each available disk associated with an adapter:

```
rmdev -l hdiskx
```

where *hdiskx* is the hdisk to be made “Defined.”

For example:

```
rmdev -l hdisk3
```

Next, run the following command to return the SCSI adapter to a “Defined” state:

```
rmdev -l scsix
```

where *scsix* is the adapter to be made “Defined.”

If using an array controller, run the **rmdev** command to make “Defined” a disk and a controller, respectively, as follows:

```
rmdev -l darx
rmdev -l dacx
```

When all controllers and disks are “Defined,” see the next section, Defining Target Mode SCSI Devices in AIX on page 4-6, to enable the Target Mode connection.

Note: Target mode SCSI is automatically configured if you are using the SCSI-2 Differential Fast/Wide Adapter. Skip ahead to the section on Defining the Target Mode Connection to HACMP on page 4-8.

Defining Target Mode SCSI Devices in AIX

The steps in defining a target mode SCSI device are:

1. Enable the target mode interface for the SCSI adapter.
2. Configure (make available) the devices.

Complete both steps on one node, then on the remaining nodes.

Enabling Target Mode Interface

To enable the target mode interface:

1. Enter:

   ```
   smit devices
   ```

   SMIT displays a list of devices.
2. Select **SCSI Adapter** and press Enter.
3. Select **Change/Show Characteristics of a SCSI Adapter** and press Enter.

   SMIT prompts you to identify the SCSI adapter.
4. Select the appropriate adapter and press Enter to display the next screen.
5. Set the **Enable TARGET MODE interface** field to **yes** to enable the target mode interface on the device (the default value is **no**).
6. Press Enter to commit the value.
7. Press F10 to exit SMIT.

Configuring the Target Mode SCSI Device

After enabling the target mode interface, you must run **cfgmgr** to create the initiator and target devices and make them available:

1. Enter:

   ```
   smit devices
   ```

   SMIT displays a list of devices.
2. Select **Install/Configure Devices Added After IPL** and press Enter to display the next screen.
3. Press Enter.
4. Press F10 to exit SMIT after the **cfgmgr** command completes.
5. Run the following command to ensure that the devices are paired correctly:

   ```
   lsdev -Cc tmsci
   ```

   At this point, a target mode SCSI device is generated that points to the other cluster nodes that share the SCSI bus. For example, given that Node A and Node B have SCSI IDs of 6 and 7, respectively, running the **lsdev -Cc tmscsi** command on either node will return the following output:

```
On Node A: lsdev -Cc tmscsi
tmscsi0 Available 00-04-00-70 SCSI I/O Controller Initiator Device
```

On Node B: lsdev -Cc tmscsi

```
tmscsi0 Available 00-04-00-60 SCSI I/O Controller Initiator Device
```

Note: The adapter SCSI ID on the node from which you enabled the interface will not be listed.

Repeat the above procedure for the other node to be connected to the SCSI-2 bus.

Target Mode Files

After you have configured the target mode connection on each adapter, two special files per target mode interface exist in the **/dev** directory: the **/dev/tmscsi*xx*.im** and **/dev/tmscsi*xx*.tm** files, where ***xx*** is a device number the system sequentially assigns per tmscsi connection. For example:

```
/dev/tmscsi0.im, /dev/tmscsi0.tm
/dev/tmscsi1.im, /dev/tmscsi1.tm
```

The file with the**.im** extension is the initiator, which transmits data. The file with the **.tm** extension is the target, which receives data. If you have four nodes, each node will have three pairs of files.

Testing the Target Mode Connection

For the target mode connection to work, initiator and target devices must be paired correctly. To ensure that devices are paired and that the connection is working after enabling the target mode connection on both nodes:

1. Enter the following command on one node connected to the bus:

   ```
   cat < /dev/tmscsinn.tm
   ```

2. Enter the following command on the other node connected to the bus:

   ```
   cat filename > /dev/tmscsinn.im
   ```

where *nn* must be the device number of the sending node and `filename` is a file. The contents of the specified file are displayed on the node on which you entered the first command above; however, the tmscsi device numbers will not necessarily be the same. You must rely on the **lsdev -Cc tmscsi** output to determine the pair connection.

Note: After a system reboot, the first execution of the **cat** command will not work. Instead, the target device receives a unit attention that notifies the initiator that the device has been reset. This notification puts the devices in sync. Afterwards, the **cat** command will work.

Note: If the SCSI bus is disconnected while running as a target mode SCSI network, the network will not properly reintegrate when the bus is reconnected. HACMP for AIX should be shut down on a node that has had the SCSI bus detached from it before reattaching the SCSI bus to that node. See the *HACMP for AIX Troubleshooting Guide* for more information if you experience the following problem: “Network Will Not Properly Reintegrate When Reconnecting the Bus.”

Defining the Target Mode Connection to HACMP

After configuring and testing the target mode connection, you must define it as a serial network to the HACMP for AIX cluster environment. The following steps describe how to use the **Configure Adapters** option on the Cluster Topology screen to define a target mode connection in the HACMP cluster.

To define the target mode connection (on each adapter) as a serial network to the HACMP for AIX software:

1. Select **Configure Adapters** from the Cluster Topology menu.
2. Select **Add an Adapter** and press Enter to display the next screen.
3. Enter field values as follows:

   Adapter Label Enter the name of the target mode adapter. The label for a target mode adapter must be unique. For example, *clam_tmscsi2*.

   Network Type Pick **tmscsi** on the pop up pick list. Use the lsdev -Cc tmscsi command to identify the proper target mode SCSI device number, which is not necessarily the same number at both ends. You must define both ends of the tmscsi network. For example, define tmscsi2 to Node A, with the tmscsi device pointing to Node B, and vice versa.

   Network Name Enter the name of the network that refers to the target mode connection. This name is arbitrary but must be used consistently. For example, *tmscsi2*.

   Network Attribute Set this field to **serial.**

   Adapter Function Set this field to **service**.

   Adapter Identifier Enter the device name: **/dev/tmscsi*x***. The device number *x* should match the number entered for the adapter label.

   Adapter Hardware Address Leave this field blank.

   Node Name Enter the name of the node connected to the adapter.

4. Press Enter. The system adds these values to the HACMP for AIX ODM and returns you to the Configure Adapters menu.

Note that when you run the **clverify** utility, it checks to see that defined **/dev/tmscsix** devices exist in the ODM for all devices defined to the HACMP for AIX environment. See the *HACMP for AIX Troubleshooting Guide* for more information on using this utility.

5. Repeat the above procedure to define the other adapter connected by the SCSI-2 bus.

Configuring Target Mode SSA Connections

This section describes how to configure a target mode SSA connection between nodes sharing disks connected to SSA on Multi-Initiator RAID adapters (FC 6215 and FC 6219). The adapters must be at Microcode Level 1801 or later.

You can define a serial network to HACMP that connects all nodes on an SSA loop.

Changing Node Numbers on Systems in SSA Loop

By default, node numbers on all systems are zero. In order to configure the target mode devices, you must first assign a unique non-zero node number to all systems on the SSA loop.

1. To change the node number use the following command.

```
chdev -l ssar -a node_number=#
```

2. To show the system's node number use the following command.

```
lsattr -El ssar
```

Configuring Target Mode SSA Devices

After enabling the target mode interface, you must run **cfgmgr** to create the initiator and target devices and make them available. To configure the devices and make them available:

1. Enter:

```
smit devices
```

SMIT displays a list of devices.

2. Select **Install/Configure Devices Added After IPL** and press Enter.
3. Press F10 to exit SMIT after the **cfgmgr** command completes.
4. Run the following command to ensure that the devices are paired correctly:

```
lsdev -Cc tmssa
```

Repeat the above procedure (enabling and configuring the target mode SSA device) for other nodes connected to the SSA adapters.

Target Mode Files

Configuring the target mode connection creates two special files in the **/dev** directory of each node, the **/dev/tmssa#.im** and **/dev/tmssa#.tm** files. The file with the **.im** extension is the initiator, which transmits data. The file with the **.tm** extension is the target, which receives data

Testing the Target Mode Connection

For the target mode connection to work, initiator and target devices must be paired correctly. To ensure that devices are paired and that the connection is working after enabling the target mode connection on both nodes:

1. Enter the following command on a node connected to the SSA disks.

```
cat < /dev/tmssa#.tm
```

where # must be the number of the target node. (This command hangs and waits for the next command.)

2. On the target node, enter the following command:

   ```
   cat filename > /dev/tmssa#.im
   ```

   where # must be the number of the sending node and *filename* is a file.

The contents of the specified file are displayed on the node on which you entered the first command.

3. You can also check that the tmssa devices are available on each system using the following command:

   ```
   lsdev -C | grep tmssa
   ```

Defining the Target Mode SSA Serial Network to HACMP

Take the following steps to configure the Target Mode SSA serial network in the HACMP cluster.

1. Select **Configure Adapters** from the Cluster Topology menu.
2. Select **Add an Adapter** and press Enter.

   Enter the fields as follows.

Adapter Label	Unique label for adapter. For example, *adp_tmssa_1*
Network Type	Pick *tmssa* from the pop up pick list.
Network Name	Arbitrary name used consistently for all adapters on this network. For example, *tmssa_1*
Network Attribute	Set this field to **Serial**
Adapter Function	Set this field to **service**.
Adapter Identifier	Enter the device name */dev/tmssa#*
Adapter Hardware Address	Leave this field blank.
Node Name	Enter the name of the node the adapter is connected to.

3. Press Enter. The system adds these values to the HACMP for AIX ODM and returns you to the **Configure Adapters** menu.
4. Repeat the above procedure to define the other adapters connected on the SSA loop.

Configuring an RS232 Serial Line

This section describes how to configure an RS232 serial line as a serial network in an HACMP cluster. Using a serial network to connect two nodes is strongly recommended for an HACMP for AIX environment. Before configuring the RS232 serial line, however, you must have physically installed the line between the two nodes. The HACMP for AIX serial line (a 25-pin null-modem, serial to serial cable) can be used to connect the nodes. The cable is available in the following lengths:

- 3.7 meter serial to serial port cable (FC3124)
- 8 meter serial to serial port cable (FC3125).

Note: The 7013-S70, 7015-S70, and 7017-S70 do not support the use of native serial ports in an HACMP RS232 serial network. Configuration of an RS232 serial network in an S70 system requires a PCI multi-port Async card.

Steps to Configuring an RS232 Serial Line

Remember the following when configuring an RS232 serial line:

1. Ensure that you have physically installed the RS232 serial line between the two nodes before configuring it. The HACMP for AIX serial line, a null-modem line, is built like a standard SLIP line.
2. Use the following command to check the status of each serial port you intend to use after installing the RS232 serial line:

   ```
   lsdev -Cc tty
   ```

   If the **tty** device is neither defined nor available, it will not be listed by the **lsdev** command. Use the **smit tty** fastpath to define the device.

   If the **tty** device is defined but not available, or if you have questions about its settings, use the **rmdev** command to delete the **tty** device:

   ```
   rmdev -l ttyx -d
   ```

   where *ttyx* is the targeted **tty** device (for example, *tty1*).
3. Use the **smit tty** fastpath to define the device on each node that will be connected to the RS232 line. Removing and then defining the **tty** device makes it available with the default settings (which are appropriate for the communication test described below).
4. Set the **ENABLE login** field to **off** to prevent **getty** processes from spawning on this device. Refer to the section below, Defining the tty Device on page 4-11.
5. Test communication over the serial line after creating the **tty** device. See the section Testing the Serial Connection on page 4-12 for more information about testing serial networks.

Defining the tty Device

To create a **tty** device on each node to be connected to the RS232 line:

1. Enter **smit tty**. SMIT displays the TTY screen.
2. Select **Add a TTY** and press Enter. The SMIT Add a TTY screen appears, prompting you for a **tty** type.
3. Select **tty rs232 Asynchronous Terminal** and press Enter.

   SMIT prompts you to identify the parent adapter.
4. Select the parent adapter and press Enter.

   The parent adapter you select is the adapter to which the RS232 cable is connected.

5. Enter field values as follows:

PORT number	Press F4 to list the available port numbers. Select the appropriate port number and press Enter. The port that you select is the port to which the RS232 cable is connected.
ENABLE login	Make sure this field is set to **off** to prevent **getty** processes from spawning on this device. Enter the following command to set this option from the command line: `chdev -l tty0 -a ttyprog_action=off`.

6. Press Enter to commit the values.

7. Press F10 to exit SMIT.

Repeat this procedure for the other node that will be connected to the RS232 line.

Testing the Serial Connection

To test communication over the serial line after creating the **tty** device on both nodes:

1. On the first node, enter:

   ```
   stty < /dev/ttyx
   ```
   where */dev/ttyx* is the newly added **tty** device. The command line on the first node should hang until the second node receives a return code.

2. On the second node, enter:

   ```
   stty < /dev/ttyx
   ```
   where */dev/ttyx* is the newly added **tty** device.

   If the nodes are able to communicate over the serial line, both nodes display their **tty** settings and return to the prompt.

   Note: This is a valid communication test of a newly added serial connection before the HACMP for AIX **/usr/sbin/cluster/clstrmgr** daemon has been started. This test yields different results after the **/usr/sbin/cluster/clstrmgr** daemon has been started, since this daemon changes the initial settings of the **tty** devices and applies its own settings. The original settings are restored when the HACMP for AIX software exits.

Defining the RS232 Serial Line to HACMP for AIX

After you have installed and tested the RS232 serial line, define it as a serial network to the HACMP cluster. The following steps describe how to use the Add an Adapter screen to define an RS232 serial line to the HACMP for AIX cluster environment.

To associate a network adapter with a cluster node:

1. Select **Configure Adapters** from the Cluster Topology menu and press Enter.

2. Select **Add an Adapter** and press Enter to display the following screen.

3. Enter field values as follows:

   Adapter Label Enter the name of the serial adapter. The label for a serial adapter must be unique (for example, *caviar_tty1*).Also, make sure that the serial adapters connected to the same RS232 line have different names.

   Network Type Pick the type **RS232** from the pop up pick list.

   Network Name Enter the name of the network connected to this adapter. Refer to the *Serial Network Adapter Worksheet.*

   Network Attribute Set this field to **serial**.

   Adapter Function Set this field to **service**.

   Adapter Identifier Enter the full path name of the tty device (for example, */dev/tty1*).

   Adapter Hardware Address Leave this field blank.

   Node Name Enter the name of the node connected to this adapter.

4. Press Enter. The system adds these values to the HACMP for AIX ODM and displays the Configure Adapters menu.
5. Repeat this procedure to define the other adapter connected by the RS232 line.

See the chapter on planning serial networks in the *HACMP for AIX Planning Guide* for more information on serial networks.

Configuring for Asynchronous Transfer Mode (ATM)

Asynchronous Transfer Mode (ATM) is a connection-oriented, private network. A point-to-point connection between two systems makes an ATM network similar to Frame Relay, X.25, and SLIP connections rather than to non-connection methods such as Ethernet, Token-Ring, and FDDI.

HACMP for AIX supports both the Classical IP and LAN Emulation forms of the ATM protocol. Classical IP implements the TCP/IP protocols directly on top of the ATM protocol. ATM Classical IP networks are significantly different in functionality and methodology than more traditional networks, like Token-Ring and Ethernet. TCP/IP does not operate the same on TCP/IP networks using Classical IP. One key difference is that ATM does not support a network broadcast.

ATM LAN Emulation provides an extra layer of software that hides ATM characteristics from the upper level protocols by using a pseudo device driver that acts like Token Ring or Ethernet. You can use ATM LAN Emulation to bridge existing Ethernet or Token-Ring networks—particularly switched, high-speed Ethernet—across an ATM backbone network.

Configuring For ATM Classical IP

Because ATM is a connection-oriented technology and IP is a datagram-oriented technology, mapping IP addresses over ATM Classical IP is complex. For PVC-type ATM connections, each IP station must be manually configured. SVC-type ATM networks are dynamic and are divided into logical IP subnets (LIS). An LIS is similar to a traditional LAN segment.

The ATM Classical IP standard requires one ARP server per LIS, where each ARP server resolves IP addresses into the corresponding ATM hardware address without using broadcasting. Each IP station must be configured with the ATM addresses of the ARP servers that serve the IP subnet configured on that IP station. This section provides details about configuring ATM Classical IP ARP servers and clients.

The current ATM Classical IP support in HACMP for AIX reflects the following restrictions:

HACMP Configuration Restrictions

ATM Classical IP networks must be defined to HACMP as private networks because ATM Classical IP does not support broadcasting.

ATM Classical IP Configuration Restrictions

There are two basic components to configuring ATM Classical IP networks:

- Configuring the ATM Classical IP ARP servers
- Configuring the ATM Classical IP ARP clients.

ATM Classical IP networks require the use of an ARP server to handle resolutions from ATM hardware addresses to IP addresses. One ARP server exists for each defined subnetwork. Thus, two ARP servers are required for each ATM Classical IP network configured in an HACMP cluster—one for the service subnet and one for the standby subnet. The ATM ARP client must be configured to include the following information about the ATM ARP server:

- The MAC address of the ATM ARP server
- The ATM ARP server and ATM ARP client must be on the same subnet.

An ATM ARP server can be an HACMP for AIX client, but it cannot be an HACMP server because ATM ARP clients hard code the ATM link address of the ARP server (including the MAC address). Adapter swaps and IPAT, therefore, do not work on HACMP nodes that are also ATM ARP servers.

ATM can support multiple interfaces per ATM adapter. Adapters used in HACMP servers, however, should use only one interface per adapter. ATM switches can be used as ATM ARP servers only if the switch automatically updates its ARP cache after an HACMP adapter event.

Thus, ATM networks must be configured as Switched Virtual Circuits through an ATM switch, with the ATM ARP server configured on a system that is not a member of the cluster nor on the ATM switch itself (except as noted in the previous paragraph).

Configuring ATM ARP Servers for Use by HACMP Nodes

Before configuring an ATM ARP server, install the ATM adapters and the switch as described in your ATM product documentation. When installation is complete, do the following:

1. Configure an ATM ARP server for the HACMP service subnet
2. Configure an ATM ARP server for the HACMP standby subnet

3. Determine the ATM server address for each ATM server

Configuring ATM ARP Servers for HACMP Service Subnetworks

To configure an ARP server for the HACMP “service” subnetwork:

1. Enter:

   ```
   smitty chinet
   ```

   SMIT displays a list of Available Network Interfaces.

2. Select **at0** as the ATM network interface. This interface will serve as the ARP server for the subnetwork 192.168.110 as shown in the following example of the **Change/Show an ATM Interface** screen.

Network Interface Name	at0
INTERNET ADDRESS (dotted decimal0	192.168.110.28
Network MASK (hex or dotted decimal)	255.255.255.0
Connection Type	svc_s
ATM Server Address	
Alternate Device	
Idle Timer	60
Current STATE	up

Note: The **Connection Type** field is set to **svc_s** to indicate that the interface is used as an ARP server.

3. Press F10 to exit SMIT.

Configuring ATM ARP Servers for HACMP Standby Subnetworks

To configure an ARP server for an HACMP “standby” subnetwork:

1. Repeat Steps 1 through 3 of the preceding procedure to configure an ATM ARP server for a “standby” subnetwork, selecting **at1** as the network interface and other options as shown in the following example:

Network Interface Name	at1
INTERNET ADDRESS (dotted decimal0	192.168.111.28
Network MASK (hex or dotted decimal)	255.255.255.0
Connection Type	svc_s
ATM Server Address	
Alternate Device	

Idle Timer	60
Current STATE	up

Note: The interface name is (at1) for the standby adapter; the **Connection Type** designates the interface as an ARP server, **svc_s**. The “standby” subnet is 192.168.111.

Obtaining an ARP Server’s Hardware Addresses:

To show an ARP server’s hardware addresses, use the **arp** command as follows on the ATM ARP server:

```
arp -t atm -a svc
```

A display similar to the following appears:

```
SVC IP Addr       ATM address

arpserver1        (192.168.110.28)
                  47.0.5.80.ff.el.0.0.0.f2.1a.21.e.8.0.5a.99.82.95.0

aprserver2        (192.168.111.28)
                  47.0.5.80.ff.el.0.0.0.f2.1a.21.e.8.0.5a.99.82.95.1
```

Note: The ATM arp server address is the 20-byte hardware address of the ATM arp server used for the subnet of an internet address.

ARP server addresses need to be defined to ATM ARP clients, as explained in the following section.

Configuring ATM ARP Clients on HACMP Cluster Nodes

To configure ATM ARP clients on cluster nodes:

1. On each cluster node, configure the service and standby ATM adapters in AIX to use the “service” and “standby” ATM ARP servers previously configured.
2. Test the configuration.
3. Define the ATM network to HACMP.

Configuring the HACMP Cluster Nodes as ATM ARP Clients

Use the **smitty chinet** command to configure two ATM interfaces, one on each adapter (at0 on atm0 for “service”, and at1 on atm1 for “standby”).

Configuring the “service” subnet

Indicate the following values for these interfaces:

Network Interface Name	at0
INTERNET ADDRESS (dotted decimal)	192.168.110.30

Network MASK (hex or dotted decimal)	255.255.255.0
Connection Type	svc_c
ATM Server Address	47.0.5.80.ff.el.0.0.0.f2.1a.39.65.0.20.48.1a.39.65.0
Alternate Device	
Idle Timer	60
Current STATE	up

The **Connection Type** field is set to svc_c to indicate that the interface is used as an ATM ARP client. Because this ATM ARP client configuration is being used for the HACMP “service” subnet, the **INTERNET ADDRESS** must be a host on the 192.168.110 subnet. The ATM server address is the 20-byte address which identifies the ATM ARP server being used for the 192.168.110 subnet.

Note: If IPAT is enabled for the HACMP-managed ATM network, the INTERNET ADDRESS represents the “boot” address. If IPAT is not enabled, the INTERNET ADDRESS represents the “service” address.

Configuring the “standby” subnet

Indicate the following values for these interfaces:

Network Interface Name	at1
INTERNET ADDRESS (dotted decimal)	192.168.111.30
Network MASK (hex or dotted decimal)	255.255.255.0
Connection Type	svc_c
ATM Server Address	47.0.5.80.ff.el.0.0.0.f2.1a.39.65.0.20.48.1a.39.65.1
Alternate Device	
Idle Timer	60
Current STATE	up

The **Connection Type** field is set to svc_c to indicate that the interface is used as an ATM ARP client. Because this ATM ARP client configuration is being used for the HACMP “standby” subnet, the **INTERNET ADDRESS** must be a host on the 192.168.111 subnet. The ATM server address is the 20-byte address which identifies the ATM ARP server being used for the 192.168.111 subnet.

Testing Communication Over the Network

To test communication over the network after configuring ARP servers and clients:

1. Run the **netstat -i** command to make sure the ATM network is recognized. You should see the device listed as **at1**.
2. Enter the following command on the first node:

   ```
   ping IP_address_of_other_node
   ```

   where *IP_address_of_other_node* is the address in dotted decimal that you configured as the destination address for the other node.
3. Repeat Steps 1 and 2 on the second node, entering the destination address of the first node as follows:

   ```
   ping IP_address_of_other_node
   ```

Defining the ATM Network to HACMP

After you have installed and tested an ATM (Classic IP?) network, you must define it to the HACMP cluster topology as a “private” network. Chapter 11, Defining the Cluster Topology, describes how to define an ATM network in an HACMP cluster.

ATM LAN Emulation

ATM LAN emulation provides an emulation layer between protocols such as Token-Ring or Ethernet and ATM. It allows these protocol stacks to run over ATM as if it were a LAN.

ATM LAN emulation does not require an ARP server or underlying ATM interfaces. A LAN emulation server must be configured to use ATM LAN emulation. LAN emulation servers reside in the ATM switch. Configuring the switch varies with the hardware being used. Once you have configured your ATM switch and a working ATM network, you can configure adapters for ATM LAN emulation.

Note: You must load **bos.atm** from AIX on each machine if you have not already done so.

To configure ATM LAN emulation through SMIT, take the following steps:

1. Enter the SMIT fastpath `atmle_panel.`

   SMIT displays the ATM LAN Emulation menu.
2. Select **Add an ATM LE Client**.
3. Choose one of the adapter types (Ethernet or Token-Ring). A popup appears with the adapter selected (Ethernet in this example). Press Enter.
4. SMIT displays the **Add an Ethernet ATM LE Client** screen. Make entries as follows:

Local LE Client’s LAN MAC Address (dotted hex) Assign a hardware address like the burned in address on actual network cards. Address must be unique on the network to which it is connected.

Automatic Configuration via LECS **No** is the default. Toggle if you want **yes**.

If no, enter the LES ATM address (dotted hex) Enter the 20-byte ATM address of the LAN Emulation server.

If yes, enter the LECS ATM address (dotted hex) If the switch is configured for LAN Emulation Configuration Server either on the well-known address, or on the address configured on the switch, enter that address here.

Local ATM Device Name Press F4 for a list of available adapters.

Emulated LAN Type Ethernet/IEEE 802.3 (for this example)

Maximum Frame Size (bytes)

Emulated LAN name (optional) Enter a name for this virtual network.

5. Once you make these entries, press Enter. Repeat these steps for other ATM LE clients.
6. The ATM LE Clients should be visible to AIX as network cards when you execute the `lsdev -Cc adapter` command.
7. Each virtual adapter has a corresponding interface that must be configured, just like a real adapter of the same type, and it should behave as such.

Defining the ATM LAN Emulation Network to HACMP

After you have installed and tested an ATM LAN Emulation network, you must define it to the HACMP cluster topology as a public network. Chapter 11, Defining the Cluster Topology, describes how to define networks and adapters in an HACMP cluster.

You will define these virtual adapters to HACMP just as if they were real adapters. They have all the functions of Ethernet or Token-Ring adapters, such as hardware address swapping.

Configuring RS/6000 SP Switches

The RS/6000 SP Switch, used by an SP node, serves as a network device for configuring multiple clusters and connecting clients. The switch is not required for an HACMP installation. If the HACMP for AIX software is installed, however, the SP Switch Network Module default settings are sufficient to allow it to operate effectively in an HACMP cluster environment. For more information about changing these settings, see the chapter on changing the cluster topology in the *HACMP for AIX Administration Guide.*

Basic points to remember about using an SP Switch in an HACMP for AIX configuration:

- ARP must be enabled for the SP Switch network so that IP address takeover can work.
- All SP Switch addresses must be defined on a private network.
- In an HACMP for AIX configuration on an RS/6000 SP, SP Switch boot and service addresses will be alias addresses on the css0 IP interface. The css0 base IP address can be configured as a service address if the network will only be used by HACMP for heartbeat traffic. The base IP address cannot be configured as a service address if the Switch will be

used with IP address takeover. Standby adapters are not used for SP Switch IP address takeover. The alias boot and service addresses will appear as **ifconfig alias** addresses to the css0 IP interface and to the base IP address.

- The netmask associated with the css0 base IP address will be used as the netmask for all HACMP for AIX SP Switch network adapters.
- It is recommended that AIX Error notification be used to recover/fallover in the event of switch adapter errors.
- When configuring the SP with multiple networks to use Enhanced Security, you must, to avoid single points of failure, configure each network for Kerberos authentication. You can do this either at initial setup and installation of Kerberos on the SP, or later when you are customizing the nodes.

For more information about the RS/6000 SP and SP Switches, see Appendix F, Installing and Configuring HACMP for AIX on RS/6000 SPs.

Configuring Cluster Performance Tuning

Cluster nodes sometimes experience extreme performance problems, such as large I/O transfers, excessive error logging, or lack of memory. When this happens, the Cluster Manager can be starved for CPU time. It might not reset the “deadman switch” within the time allotted. Misbehaved applications running at a priority higher than the cluster manager can also cause this problem.

The deadman switch is the AIX kernel extension that halts a node when it enters a hung state that extends beyond a certain time limit. This enables another node in the cluster to acquire the hung node’s resources in an orderly fashion, avoiding possible contention problems. If the deadman switch is not reset in time, it can cause a system panic and dump under certain cluster conditions.

Setting the following tuning parameters correctly may avoid some of the performance problems noted above. It is highly recommended to set the two AIX parameters; this may preclude having to change the HACMP Network Modules Failure Detection Rate.

- AIX high and low watermarks for I/O pacing
- AIX **syncd** frequency rate
- HACMP Network Module Failure Detection Rate (Custom)
  - HACMP cycles to failure
  - HACMP heartbeat rate.

Setting Advanced Performance Tuning Parameters

In HACMP 4.4, you can configure these related parameters directly from HACMP SMIT.

Note: You must set the two AIX parameters on each cluster node. Network module settings are propagated to all nodes when you set them on one node and then synchronize the cluster topology.

Setting I/O Pacing

Although the most efficient high- and low-water marks vary from system to system, an initial high-water mark of **33** and a low-water mark of **24** provides a good starting point. These settings only slightly reduce write times and consistently generate correct fallover behavior from the HACMP for AIX software.

See the *AIX Performance Monitoring & Tuning Guide* for more information on I/O pacing.

To change the I/O pacing settings:

1. Enter `smitty hacmp` > **Cluster Configuration > Advanced Performance Tuning Parameters > Change/Show I/O Pacing**
2. Configure the entry fields with the recommended HIGH and LOW watermarks:

HIGH water mark for pending write I/Os per file	**33** is recommended for most clusters. Possible values are 0 to 32767.
LOW water mark for pending write I/Os per file	**24** is recommended for most clusters. Possible values are 0 to 32766.

Setting Syncd Frequency

The **syncd** setting determines the frequency with which the I/O disk-write buffers are flushed. Frequent flushing of these buffers reduces the chance of deadman switch time-outs.

The AIX default value for **syncd** as set in **/sbin/rc.boot** is 60. It is recommended to change this value to 10. Note that the I/O pacing parameters setting should be changed first. The utility updates **/sbin/rc.boot**, kills the old **syncd** process, then starts the new one with the new value.

To change the **syncd** frequency setting:

1. Enter `smitty hacmp` > **Cluster Configuration > Advanced Performance Tuning Parameters > Change/Show syncd frequency**
2. Configure the entry fields with the recommended **syncd** frequency:

syncd frequency in seconds	**10** is recommended for most clusters. Possible values are 0 to 32767.

Changing the Failure Detection Rate of a Network Module

Warning: I/O pacing must be enabled before changing the failure detection rate of a Network Module; it regulates the number of I/O data transfers. Also keep in mind that the setting for the Failure Detection Rate is network specific, and may vary. You should also try adjusting the **syncd** rate and test the system thoroughly before changing the attributes of a network module.

Be sure to consult the chapter on Planning Networks in the *HACMP Planning Guide*, for information on heartbeat settings for each type of network module and how these settings interact with the deadman switch before changing the defaults.

To change the attributes of a network module:

1. Stop cluster services on all cluster nodes.
2. Enter `smitty hacmp`
3. Select **Cluster Configuration** > **Advanced Performance Tuning Parameters** > **Change/Show Network Modules** and press Enter.

   SMIT displays a list of defined network modules.
4. Select the network module you want to change and press Enter.

   SMIT displays the attributes of the network module, with the current values.

Network Module Name Name of network type, for example, ether.

New Network Module Name []

Description For example, Ethernet Protocol

Address Type **Address** or **Device**. Toggle to select the correct type.

Path Actual pathname of the network module, for example **/usr/sbin/cluster/nims/nim_ether**

Parameters

Failure Detection Rate The default is **Normal**. Choices are **Fast**, **Slow**, and **Custom**. The failure cycle and the heartbeat interval determine how soon a failure can be detected. The time needed to detect a failure can be calculated using this formula: (heartbeat interval) * (failure cycle).

Failure Cycle The current setting is the default for the network module selected. (Default for ether is 2.) This is the number of successive heartbeats that can be missed before the interface is considered to have failed. If you select the **Custom** option in the Failure Detection Rate field, you can enter a number from 1 to 21474.

Heartbeat Rate The current setting is the default for the network module selected. This parameter tunes the interval (in tenths of a second) between heartbeats for the selected network module. If you select the **Custom** option in the Failure Detection Rate field, you can enter a number from 1 to 21474.

5. To change the heartbeat rate and failure cycle fields, you must first select **Custom** for the Failure Detection Rate field. Make the desired changes and press Enter. SMIT executes the command to modify the values of these attributes in the ODM.
6. On the local node, synchronize the cluster topology. Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option.

   The configuration data stored in the DCD on each cluster node is updated and the changed configuration becomes the active configuration when cluster services are started.

Chapter 5 Installing Shared Disk Devices

This chapter lists steps for installing the following shared IBM disks and arrays in an HACMP cluster:

- SCSI-2 Differential and SCSI-2 Differential Fast/Wide disks
- SCSI-2 SE disks (Portmadog only)
- IBM 7135-110 and 7135-210 RAIDiant and IBM 7137 Disk Arrays
- IBM 9333 serial disk subsystems
- IBM 7133 or 7131-405 Serial Storage Architecture (SSA) Disk Subsystems.

Adapter Feature Code	6214	6215	6216	6219
Minimum Microcode Level	2401	1801	2402	1801

The chapter also explains how to configure:

- A target mode SCSI connection between nodes sharing a disk connected to a SCSI-2 Differential bus. (Note that neither PCI SCSI-2 differential busses nor SE busses support target mode SCSI.)
- A target mode SSA connected to SSA on Multi-Initiator RAID adapters (FC 6215 and FC 6219). You can define a serial network to HACMP that connects all nodes on an SSA loop.

Prerequisites

- Read the chapter on planning shared disk devices in the *HACMP for AIX Planning Guide* before installing shared disk devices. This chapter describes how to lay out the shared disk configuration for your system.
- Consult your AIX documentation for both the hardware and software setup of disk devices.
- Install the appropriate disk adapters. The installation procedures outlined in this chapter assume you have already installed these adapters. To install an adapter, if you have not done so, follow the procedure outlined in the documentation you received with the unit.

Related Tasks

As you install the disks, record the shared disk configuration on the disk worksheets in Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide*. You refer to the completed worksheets when you define the cluster resources following the instructions provided in Chapter 12, Configuring Cluster Resources, later in this guide.

Note: Connect shared SCSI disks to the same SCSI bus as the nodes sharing the disks. You can use RAID arrays in both concurrent and non-concurrent modes, but SCSI disks only in non-concurrent mode.

Installing Shared IBM SCSI-2 Differential and Differential Fast/Wide Disks

Complete the following steps to install shared IBM SCSI-2 Differential or IBM SCSI-2 Differential Fast/Wide disks. Differences in procedures are noted as necessary. Refer to the *Shared SCSI-2 Differential or Differential Fast/Wide Disks* worksheet.

If installing a disk array, see the section Installing an IBM SCSI Disk Array in an HACMP Cluster on page 5-6.

When installing shared SCSI-2 Differential or Differential Fast/Wide disks:

1. Review the shared disk configuration diagram you drew while planning disk storage needs. Note the type of SCSI bus installation, SCSI-2 Differential or SCSI-2 Differential Fast/Wide, in the **Type of SCSI Bus** section of the worksheet. The type of SCSI bus determines the types of cables and terminators needed, and the device name.
2. Fill in the node name of each node that will be connected to the shared SCSI bus in the **Node Name** field. You previously recorded this name on the *TCP/IP Networks* worksheet.
3. Enter a unique name for each SCSI adapter used in the disk configuration in the **SCSI Adapter Label** field. For example, use *ha1, ha2,* and so on.
4. Record the I/O slot that each SCSI adapter uses in the **Slot Number** field.
5. To determine the slot number, enter:

   ```
   lscfg | grep scsi
   ```

6. Enter the logical name of each adapter in the **Logical Name** field.

   Identify the SCSI ID of each SCSI adapter for each cluster node. At boot time, when AIX configures the adapter, it assigns the adapter SCSI ID 7, by default. Note that the integer value in the logical device name (for example, the *1* in *scsi1*) is not an ID, but simply part of the name given to the configured device.

   To determine the SCSI IDs of the disk adapters, use the **lsattr** command, as in the following example to find the ID of the adapter scsi1:

   For SCSI-2 Differential adapters:

   ```
   lsattr -E -l scsi1 | grep id
   ```

   For SCSI-2 Differential Fast/Wide adapters:

   ```
   lsattr -E -l ascsi1 | grep external_id
   ```

   Do not use wildcard characters or full pathnames on the command line for the device name designation.

7. Verify that each SCSI device connected to the shared SCSI bus has a unique ID.

   A common configuration is to set the SCSI ID of the adapters on the nodes to be higher than the SCSI IDs of the shared devices. (Devices with higher IDs take precedence in SCSI bus contention arbitration.) For example, in a two-node cluster, the adapter on one node can be SCSI ID 6, and the adapter on the other node can be SCSI ID 7. The external disk SCSI IDs should later be set to an integer from 0 through 5.

Note: You may want to set the SCSI IDs of the adapters to 5 and 6 to avoid a possible conflict when booting one of the systems in service mode from a **mksysb** tape of other boot devices, since this will always use an ID of 7 as the default. Also, the IBM SCSI-2 Differential Fast/Wide Adapter cannot be assigned SCSI IDs 0,1, or 2. It can be assigned IDs from 3 to 15.

If the SCSI IDs are not unique, shut down AIX and power down all the nodes. Power the nodes back up, one at a time, and use the following command to specify a unique SCSI ID for each device:

For SCSI-2 Differential adapters:

```
chdev -l 'basename' -a 'id=x' -P
```

For SCSI-2 Differential Fast/Wide adapters:

```
chdev -l 'basename' -a 'external_id=x' -P
```

where *x* is the new value of the attribute.

For example, to change the SCSI ID of adapter scsi1 to 6, enter:

```
chdev -l 'scsi1' -a 'id=6' -P
```

Record these values in the **SCSI Device ID: Adapter** field.

8. Shut down both nodes (you may have already shut down one node if you changed the adapter ID) so that you can set the SCSI IDs for the external disks and connect the cables. To shutdown the nodes, enter:

   ```
   shutdown -F
   ```

9. Set the SCSI ID of each disk device to a unique digit.

   Recall that each device connected to the SCSI bus (each SCSI adapter and each physical disk) must have a unique SCSI ID. Each ID within the chain must be an integer value from 0 through 7 (standard SCSI-2 Differential) or from 0 through 15 (SCSI-2 Differential Fast/Wide). Refer to your worksheet for the values previously assigned to the adapters. For example, if the standard adapters have IDs of 6 and 7, assign values from 0 through 5.

   Enter the SCSI ID of each drive in the **SCSI Device ID: Shared Drive** fields on the worksheet.

10. Connect the disks together with device-to-device SCSI cables appropriate for the type of SCSI bus you are installing. The order in which you connect the disks does not matter.

11. Remove any external SCSI terminators from the device-to-device cables.

12. Attach the male end of a Y-cable to the adapter. Terminate the short leg of the Y-cable using a terminator or connect it to a third or fourth node to the shared SCSI bus.

    Note: The bus must be terminated on the short leg of a Y-cable connected to two (and only two) of the nodes on the shared bus. Connect the longer leg of the Y-cable to a Y-cable to Device cable. Repeat at the cluster node that is connected to the last shared device.

13. Power on all external disks. Reboot the AIX operating system on both nodes.

Verifying the Installation

At boot time (or whenever the **/etc/cfgmgr** command is used), AIX configures all the devices connected to the system. It first configures all the devices connected to the I/O bus, such as adapter cards. When it configures a SCSI host adapter, AIX assigns a logical name to the adapter of the form *scsix,* where *x* is an integer that uniquely identifies the adapter. For example, an adapter could be named *scsi0* or *scsi1* (Fast/Wide Adapters are named *ascsix*).

After configuring the SCSI adapter card, AIX probes the SCSI bus connected to the adapter and configures each target device connected to the bus. When AIX configures a SCSI disk, it assigns a logical name to the disk of the form *hdiskx,* where *x* is an integer that uniquely identifies the disk. For example, a disk could be named *hdisk1* or *hdisk2.* You use these logical names to refer to the disks when creating shared-disk volume groups and logical volumes.

Note: Different nodes connected to the same SCSI bus can assign different logical names to the same physical disk. For example, node A can assign the logical name *hdisk3* to a SCSI disk and node B can assign the name *hdisk2* to the same disk.

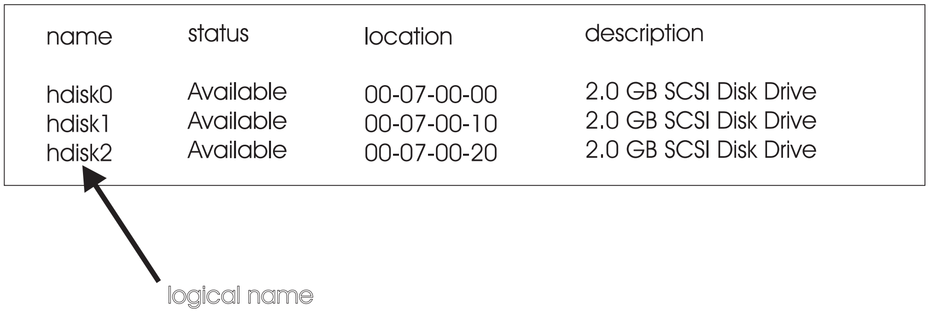
When AIX configures the SCSI disks at boot time, the LED on the external drive should light. If not, a hardware problem may exist, such as an improper termination, a loose cable, or a bent cable pin.

To verify that AIX has configured the shared SCSI bus as you expected, use the following procedure.

1. To determine the logical names of the physical disks, enter:

   ```
   lsdev -Cc disk -H
   ```

   The first column of the resulting display lists the logical names of the SCSI disks.



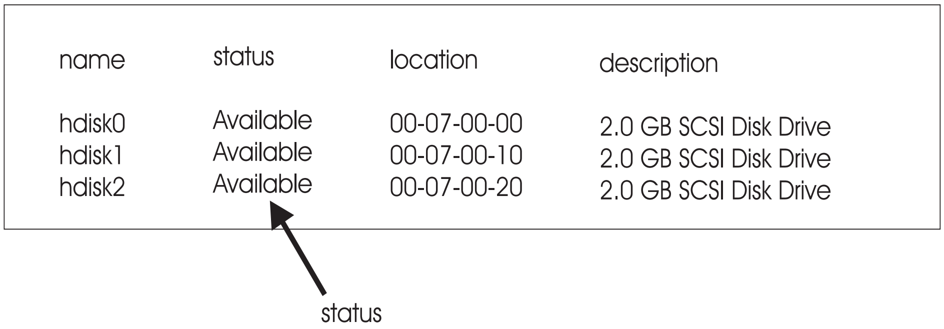
```
name      status       location       description

hdisk0    Available    00-07-00-00    2.0 GB SCSI Disk Drive
hdisk1    Available    00-07-00-10    2.0 GB SCSI Disk Drive
hdisk2    Available    00-07-00-20    2.0 GB SCSI Disk Drive
```

   Record the name of each external SCSI disk in the **Shared Drives: Logical Device Name** field, and the size in the **Shared Drives: Size** field. Be aware that the nodes can assign different names to the same physical disk. Note these situations on the worksheet.

   You can also use the **lspv** command to compare PVIDs listed for each node. If a PVID is not listed, you probably have a software or hardware configuration problem.

2. Verify that all disks have a status of “Available,” as shown in the second column of the resulting display.



```
name      status      location      description

hdisk0    Available   00-07-00-00   2.0 GB SCSI Disk Drive
hdisk1    Available   00-07-00-10   2.0 GB SCSI Disk Drive
hdisk2    Available   00-07-00-20   2.0 GB SCSI Disk Drive
```

3. If any disks have the status “Defined”, use the following command to change their status to “Available”:

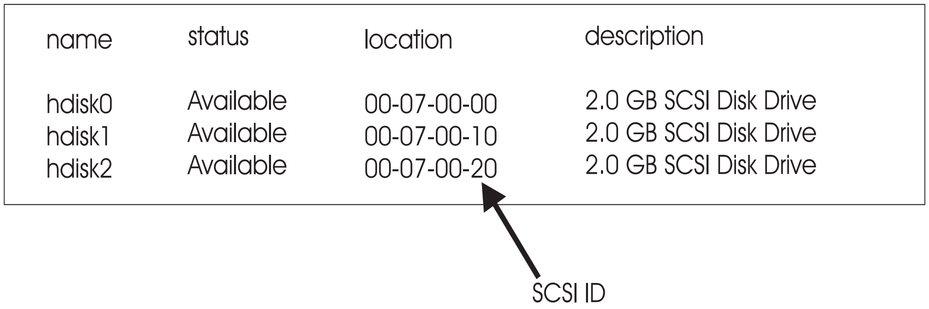
   ```
   mkdev -l hdiskx
   ```

   where *hdiskx* is the logical name of the unlisted disk.

4. Verify the SCSI IDs of the logical disks; enter:

   ```
   lsdev -Cc disk -H
   ```

   The third column of the display generated by the **lsdev -Cc disk -H** command is the location code in the format AA-BB-CC-DD. The first digit (the first D) of the DD field is the SCSI ID. These numbers should match the numbers you set on each disk unit.



```
name      status      location      description

hdisk0    Available   00-07-00-00   2.0 GB SCSI Disk Drive
hdisk1    Available   00-07-00-10   2.0 GB SCSI Disk Drive
hdisk2    Available   00-07-00-20   2.0 GB SCSI Disk Drive
```

   For SCSI-2 Differential Fast/Wide disks, a comma follows the SCSI ID field of the location code because the IDs can require two digits, such as 00-07-00-12,0.

At this point your IBM SCSI-2 Differential or Differential Fast/Wide disk installation is complete.

Installing an IBM SCSI Disk Array in an HACMP Cluster

Complete the following steps to install a shared IBM 7135 RAIDiant Disk Array or an IBM 7137 Disk Array.

As you install an IBM disk array, record the key information about the shared disk configuration on the *Shared IBM SCSI Disk Arrays* worksheet. See Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide* for a copy of this worksheet. Then complete a separate worksheet for each shared disk array SCSI bus. You refer to these worksheets when you define the node environment following the instructions provided in Chapter 12, Configuring Cluster Resources, later in this guide.

Note: The Disk Array software must be installed before physically connecting the hardware. The following procedure assumes that you have already installed the host adapter following the device's documentation for software and physical connections.

IBM 7135 RAIDiant Disk Array

Each node will be configured with two adapters dedicated to the RAIDiant Disk Array, so you can create two separate SCSI-2 Differential busses, connecting an adapter on each node to one of the array controllers. The following steps to install an IBM disk array are valid for both SCSI-2 Differential and SCSI-2 Differential Fast/Wide disks. Differences in procedures are noted as necessary.

Note: You can connect the IBM 7135-210 RAIDiant Disk Array to *only* High Performance SCSI-2 Differential Fast/Wide adapters, while the 7135-110 RAIDiant Array *cannot* use those High Performance Fast/Wide adapters.

IBM 7137 Disk Array

The IBM 7137 Disk Array is a single controller subsystem and thus does not support a dual adapter configuration.

Installing an IBM Disk Array

To install an IBM disk array in an HACMP cluster:

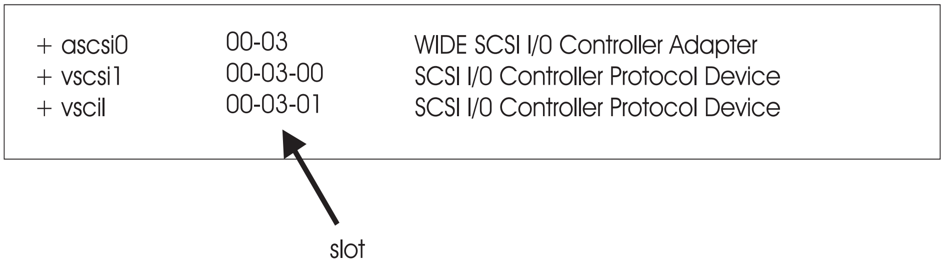
1. Review the shared disk configuration diagram you drew while planning disk storage needs.
2. Fill in the name of each node connected to this shared SCSI-2 Differential bus in the **Node Name** field.
3. Name each SCSI-2 Differential adapter used in this shared SCSI bus and record the name in the **SCSI Adapter Label** field of the configuration worksheet. For example, you could name the adapters *rha1, rha2,* and so on.
4. Record the I/O slot of each SCSI-2 Differential adapter used in this shared SCSI bus in the **Slot Number** field of the configuration worksheet.

   To determine the slot number, use the **lscfg** command, as in the following example:

   ```
   lscfg | grep scsi
   ```

In the display generated by the command, the second column lists the location code of the adapter in the format AA-BB. The last digit of that value (the last B) is the I/O slot number.

(The following examples display the information for a configuration using the SCSI-2 Differential Fast/Wide adapters.)



```
+ ascsi0      00-03        WIDE SCSI I/0 Controller Adapter
+ vscsi1      00-03-00     SCSI I/0 Controller Protocol Device
+ vscil       00-03-01     SCSI I/0 Controller Protocol Device
```

5. Record the logical device name of each adapter in the **Logical Name** field. The first column of the display generated by the **lscfg** command lists the logical name of the SCSI adapters.
6. Determine that each device connected to this shared SCSI bus has a unique SCSI ID. The first time AIX configures an adapter, it assigns the adapter card the SCSI ID 7, by default. Because each adapter on a shared SCSI bus must have a unique SCSI ID, you must change the SCSI ID of one or more of the adapters used in the shared SCSI bus. A common configuration is to let one of the nodes keep the default SCSI ID 7 and assign the adapters on the other cluster nodes the next lower SCSI IDs in sequence, such as 6 and 5. The array controller SCSI IDs should later be set to an integer starting at 0 and going up. Make sure no array controller has the same SCSI ID as any adapter. See step 8 for more information.

Note: You may want to set the SCSI IDs of the adapters to 5 and 6 to avoid a possible conflict when booting one of the systems in service mode from a **mksysb** tape of other boot devices, since this will always use an ID of 7 as the default. Also, the IBM SCSI-2 Differential High Performance Fast/Wide Adapter cannot be assigned SCSI IDs 0, 1, or 2. The IBM SCSI-2 Differential Fast/Wide Adapter cannot be assigned SCSI IDs 0 or 1.

Note that the integer value in the logical device name (for example, the *1* in *vscsi1*) is not a SCSI ID, but simply part of the name given to the configured device.

To determine the SCSI IDs of the disk adapters, use the **lsattr** command, specifying the logical name of the adapter as an argument. In the following example, the SCSI ID of the Fast/Wide adapter named *ascsi0* is obtained:

```
lsattr -E -l ascsi0 | grep external_id
```

Do not use wildcard characters or full pathnames on the command line for the device name designation.

In the resulting display, the first column lists the attribute names. The integer to the right of the **id** (**external_id**) attribute is the adapter SCSI ID.

To change the ID of a SCSI adapter, power down all but one of the nodes along with all shared devices. On the powered-on node, use the **chdev** command to change the SCSI ID of the adapter from 7 to 6, as in the following example:

```
chdev -l ’ascsi0’ -a ’external_id=6’ -P
```

Verify that each SCSI device adapter in the daisy chain has a unique ID and record these values in the **SCSI Device ID: Adapter** field of the configuration worksheet.

7. Shut down both nodes so that you can set the SCSI IDs for the array controllers and connect the cables. Use the following command to shutdown the nodes:

   ```
   shutdown -F
   ```

8. Refer to the documentation for your disk array to find out how to set SCSI IDs. Assign each controller on the Disk Array a SCSI ID that is unique on this shared SCSI bus. Refer to your worksheet for the values previously assigned to the adapters. For example, if the adapters have IDs of 6 and 7, you can assign the array controllers any SCSI ID from 0 through 5.
9. Record the SCSI ID of each array controller in the **SCSI Device ID: Array Controller** fields on the worksheet.
10. Connect the cables.
11. Repeat the above procedure for the second shared SCSI bus.
12. Power on the Disk Array and all nodes; then reboot AIX on each node.

    At this point the physical connection of the Disk Array shared bus is complete. If you are using a 7135 RAIDiant Disk Array, the next section provides steps to verify the installation. As you verify the installation, record additional information about the installation on the *Shared IBM SCSI Disk Arrays* worksheet. This information is needed when creating shared disk volume groups and logical volumes that include the 7135 RAIDiant Disk Array.
13. Run **cfgmgr** on one node at a time to complete the installation.

Note: Typically, the IBM 7137 Disk Array, with four physical drives, configures to one hdisk (LUN)

Verifying the Installation of an IBM 7135 RAIDiant Disk Array

At boot-time, AIX configures all the devices that are connected to the I/O bus, including the SCSI adapters. AIX assigns each adapter a logical name of the form *scsix,* where *x* is an integer. For example, an adapter could be named *scsi0* or *scsi1* (Fast/Wide Adapters are named *ascsix*). After AIX configures the SCSI adapter, it probes the SCSI bus and configures all target devices connected to the bus. For an IBM RAIDiant Disk Array, AIX does the following:

1. Configures the disk array controllers (for example, **dac0** and **dac1**)
2. Configures the RAIDiant pseudo-device (**dar0**)
3. Configures the hard disks defined on the array (**hdiskx**).

Configuring the Disk Array Controllers

After configuring the SCSI-2 host adapters, AIX configures the disk array controllers (DACs) on the 7135 RAIDiant Disk Array. The LED code *844* indicates that AIX is configuring the DAC. AIX assigns each DAC on the 7135 RAIDiant Disk Array a logical name of the form *dacx,* where *x* is an integer that uniquely identifies the array controller.

Configuring the Disk Array Router

After configuring the array controllers, AIX configures the RAIDiant pseudo-device, called a disk array router (DAR). Each adapter/array controller pair defines a path between the host and the 7135 RAIDiant Disk Array. The disk array router represents these paths as a single pseudo-device to AIX. The RAIDIANT software manages the switching between these paths. AIX assigns each DAR a logical name of the form *darx*, where *x* is an integer that uniquely identifies the pseudo-device. AIX creates one instance of the DAR for each 7135 RAIDiant Disk Array.

Configuring the Hard Disks

After configuring the DAR, AIX configures the logical units (LUNs) defined on the array. Each LUN on the array is configured as a hard disk. The LED code *845* indicates that AIX is configuring the hard disks on the array. AIX assigns each LUN a logical name of the form *hdiskx*, where *x* is an integer that uniquely identifies the LUN to the system. You use these logical names to refer to the disks when creating shared-disk volume groups and logical volumes.

Different nodes in an HACMP cluster can refer to the same LUN by different logical names. For example, *hdisk1* on node A can be regarded as *hdisk3* on node B.

Defining LUNs on the RAIDiant Disk Array

The 7135 RAIDiant Disk Array comes preconfigured with LUNs, depending on the number of drives it contains and the size of the drives. In addition, the LUNs are usually assigned a default RAID level of 5. Refer to the documentation you received with your 7135 RAIDiant Disk Array for the specific LUN composition of your unit.

You can modify the default setup, creating LUNs of different sizes that support different RAID levels. You also can create multiple LUNs on the same set of disk drives. To modify the configuration of LUNs on a RAIDiant Disk Array, consult the documentation for your system.

To verify the installation of a 7135 RAIDiant Disk Array:

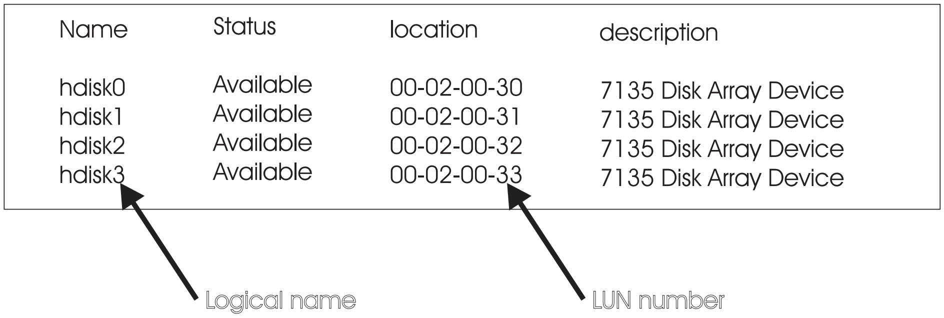
1. Verify that AIX created the device definitions (hdisks) that you expected.

   To determine the logical names of the LUNs on the RAIDiant Disk Array, use the **lsdev** command, as in the following example:

   ```
   lsdev -Cc disk -H
   ```

   The example illustrates how this command lists the hard disks created for a four-LUN RAIDiant Disk Array. The display includes the location code of the hard disk in the form AA-BB-CC-DD. The last digit of the location code included in the display represents the LUN number. (With other SCSI-2 disks, this number is always 0 because these disks do not support multiple LUNs.)

   The first column lists the logical names of the LUNs on the 7135 RAIDiant Disk Array.



```
Name      Status      location       description

hdisk0    Available   00-02-00-30    7135 Disk Array Device
hdisk1    Available   00-02-00-31    7135 Disk Array Device
hdisk2    Available   00-02-00-32    7135 Disk Array Device
hdisk3    Available   00-02-00-33    7135 Disk Array Device
```

Record the logical name of each LUN in the **Shared LUNs: Logical Device Name** field, and the size in the **Shared Drives: Size** field. Be aware that the nodes can assign different names to the same physical disk. Note these situations on the worksheet.

2. Verify that all disks have a status of "Available."

   If a disk has a status of Defined, instead of Available, check the cable connections and then use the **mkdev** command to make the disk available. Enter:

   mkdev -l *hdiskx*

   where `hdiskx` is the logical name of the defined disk.

3. Verify the SCSI IDs of the array controllers. To determine that AIX has the correct SCSI IDs for the array controllers, obtain a listing of the array controllers using the **lscfg** command:

   lscfg | grep dac

   Note: Since these controllers are on separate SCSI busses, they can have the same SCSI ID.

   The SCSI ID in the display should match the numbers you set on each array controller.

   Record the logical name of each array controller in the **Array Controller Logical Name** field of the configuration worksheet.

4. Verify that the Disk Array Router pseudo-device has been created to represent multiple paths between the host and the RAIDiant Disk Array. Use the **lsdev** command as follows:

   ```
   lsdev -Ct dar
   ```

At this point, your 7135-110 or 7135-210 RAIDiant Disk Array installation is complete.

Installing Shared IBM 9333 Serial Disk Subsystems

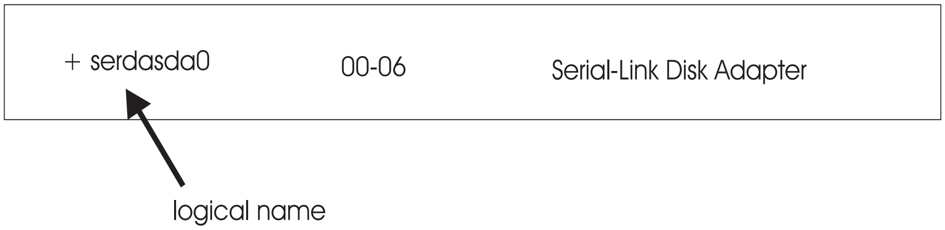
Complete the following steps to install a shared IBM 9333 serial disk subsystem. As you install the disks, record the shared disk configuration on the *Shared IBM 9333 Serial Disk* worksheet. Use a separate worksheet for each set of shared IBM 9333 serial disks. You will refer to the completed worksheets when you define the node environment.

When installing the disk subsystem:

1. Review any drawings or notes that you made while planning your disk storage needs. Notice the number of adapter-to-controller cables your cluster requires and make sure you have this number available.
2. Fill in the node name of each node connected to the shared IBM 9333 serial disk subsystem in the **Node Name** field.
3. Assign a unique name to each adapter used in the disk configuration in the **IBM 9333 Adapter Label** field. For example, use *ha1*, *ha2*, and so on.
4. Enter the logical device name of each adapter in the **Logical Name** field.

   To get the logical device name, at each node enter:

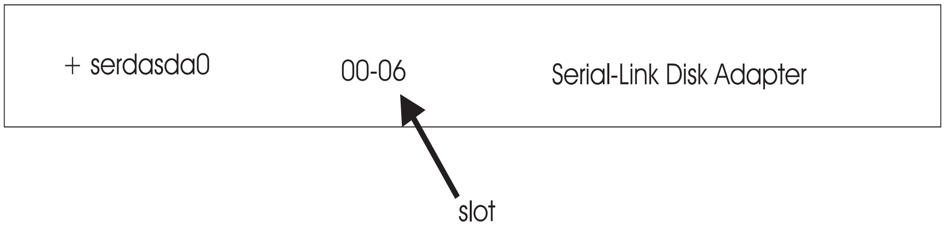
   ```
   lscfg | grep serdasda
   ```
   The first column of the resulting display lists the logical device names of the 9333 adapters.



```
+ serdasda0          00-06          Serial-Link Disk Adapter
```

5. Record the I/O slot that each adapter uses in the **Slot Number** field. Record this slot for each node. The slot number is an integer value that can range from 1 to 16.

   The second column of the existing display lists a value of the form AA-BB. The last digit of that value (the last B) is the I/O slot number.



```
+ serdasda0          00-06          Serial-Link Disk Adapter
```

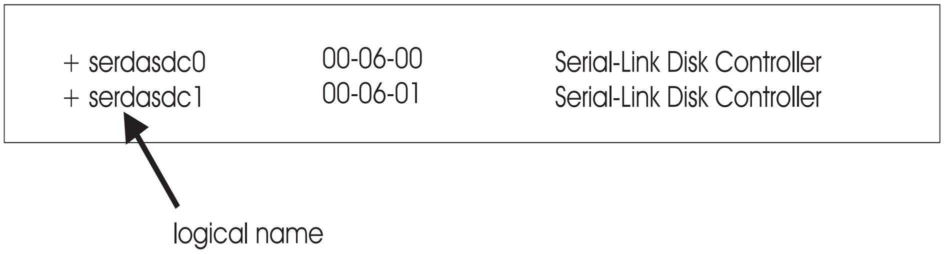
6. Label each shared-disk drawer with a unique identifier, and record the identifier in the **IBM 9333 Drawer/Desk Label** field. For example, use *drawer1*, *desk2*, and so on.
7. Connect the cables from the adapters to the controllers.

8. Record the integer value signifying the adapter I/O connector to which each drawer is connected in the **Adapter I/O Connector** field. Record this value for each node. The integer value is either *0* or *1*.

9. Record the logical name for the IBM 9333 serial-link controller in the **Controller** field for each cluster node. To get the name, enter:

   ```
   lscfg  | grep serdasdc
   ```

   The first column of the resulting display lists the logical names of the 9333 serial-link controllers.



```
+ serdasdc0        00-06-00        Serial-Link Disk Controller
+ serdasdc1        00-06-01        Serial-Link Disk Controller
```

10. Halt AIX on each node; enter:

    ```
    shutdown -F
    ```

    Do not attempt to restart the system or turn off the power before the final shutdown message displays; otherwise, you might damage a filesystem.

11. Power up and boot AIX on the nodes.

Verifying the Installation

AIX configures the disks by checking the serial line for the devices (in this case, disks) connected to it and by labeling each physical disk volume with a logical name (for example, *hdisk1*, *hdisk2*, *hdisk3*). You use these labels to refer to the disks when creating shared-disk volume groups and logical volumes.

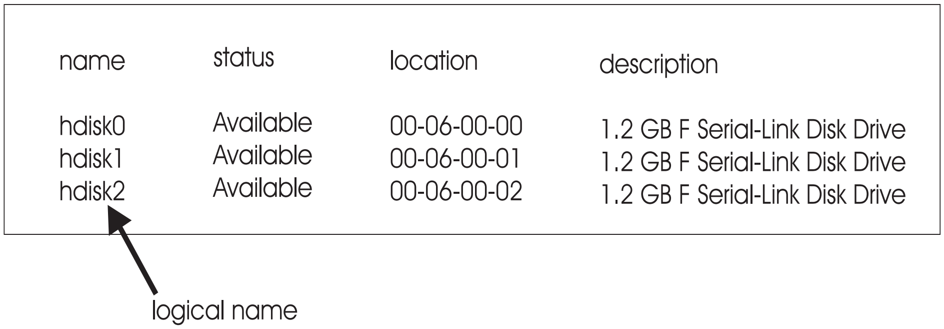
If the LED displays 870 continuously, a connection problem may exist.

To verify that AIX has configured the shared bus as you expected:

1. Determine the logical device name and size of each physical volume and record the values on the worksheet. On each node, enter:

   ```
   lsdev -Cc disk -H
   ```

   The first column of the resulting display lists the logical names of the disks.



```
name      status      location       description

hdisk0    Available   00-06-00-00    1.2 GB F Serial-Link Disk Drive
hdisk1    Available   00-06-00-01    1.2 GB F Serial-Link Disk Drive
hdisk2    Available   00-06-00-02    1.2 GB F Serial-Link Disk Drive
```

logical name

Enter the name in the **Logical Device Name** field.

Record the size of each external disk in the **Size** field.

2. Verify that all disks have a status of “Available.”
3. If a disk has the status “Defined,” use the following command to make it available:

   ```
   mkdev -l hdiskx
   ```
   where *hdiskx* is the logical name of the unavailable disk.

At this point the IBM 9333 serial disks are configured for your cluster.

Installing Shared IBM SSA Disk Subsystems

Complete the following steps to install a shared IBM 7133 or 7131-405 SSA disk subsystem.

As you install the disks, record the shared disk configuration on the *Shared IBM SSA Disk Subsystems* worksheet. Use a separate worksheet for each set of shared IBM SSA disks. You will refer to the completed worksheets when you define the node environment.

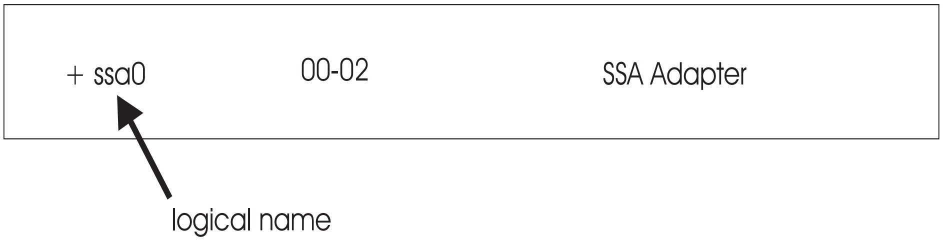
Note: The IBM 7133 SSA adapter card supports four external connectors that allow one of the adapter's two dual-ports to be connected externally to a system unit.

When installing the disk subsystem:

1. Review any drawings or notes that you made while planning your disk storage needs.
2. Fill in the node name of each node connected to the shared IBM 7133 SSA disk subsystem in the **Node Name** field.
3. Assign a unique name to each adapter used in the disk configuration in the **SSA Adapter Label** field. For example, use *ha1, ha2*, and so on.
4. Enter the logical device name of each adapter in the **Logical Name** field.

   To get the logical device name, at each node enter:

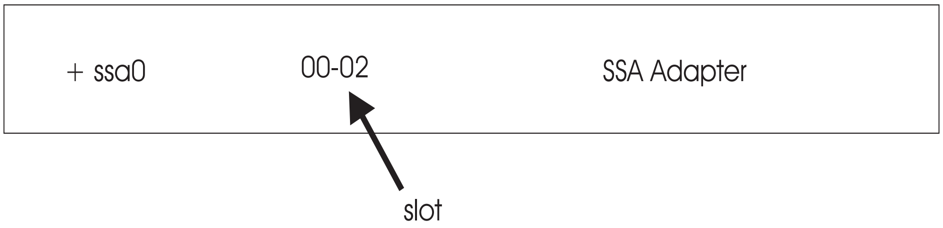
   ```
   lscfg | grep ssa
   ```
   The first column of the resulting display lists the logical device names of adapters.



```
+ ssa0          00-02          SSA Adapter
```

5. Record the slot that each adapter uses in the **Slot Number** field. Record this slot for each node. The slot number is an integer value that can range from 1 to 16.

   The second column of the existing display lists a value of the form AA-BB. The last digit of that value (the last B) is the slot number.



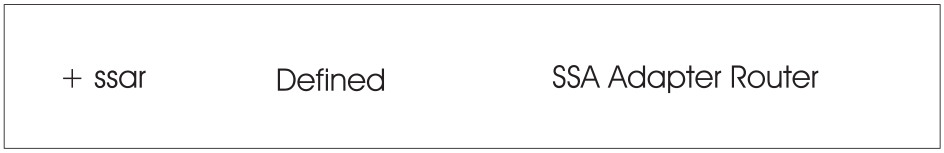
```
+ ssa0          00-02          SSA Adapter
```

6. Identify the dual-port on the SSA adapter to connect to the system unit and record this number on the worksheet.
7. Connect the cables.

Configuring the SSA Adapter Router

At configuration time, AIX configures an SSA adapter router (ssar). The ssar is only a conceptual configuration aid and is always in a “Defined” state. It cannot be made “Available.” You can list the ssar with the following command:

```
lsdev -C | grep ssar
```



```
+ ssar        Defined        SSA Adapter Router
```

8. Halt AIX on each node; enter:

   ```
   shutdown -F
   ```
   Do not attempt to restart the system or turn off the power before the final shutdown message displays; otherwise, you might damage a filesystem.

9. Power up and boot AIX on the nodes.

Verifying the Installation

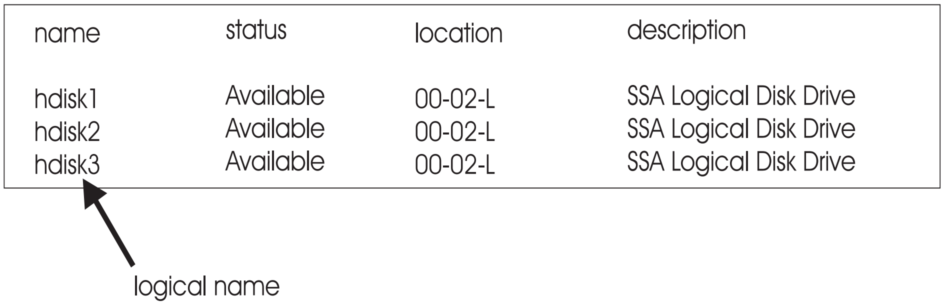
AIX configures the disks by checking the serial line for the devices (in this case, disks) connected to it and by labeling each physical disk volume with a logical name (for example, *hdisk1*, *hdisk2*, *hdisk3*). You use these labels to refer to the disks when creating shared-disk volume groups and logical volumes.

If the LED displays 870 continuously, a connection problem may exist.

To verify that AIX has configured the disks as you expected:

1. Determine the logical device name of each physical volume and record the values on the worksheet. On each node, enter:

   ```
   lsdev -Cc disk | grep -i ssa
   ```
   The first column of the resulting display lists the logical names of the disks.

   ```
   name        status       location     description

   hdisk1      Available    00-02-L      SSA Logical Disk Drive
   hdisk2      Available    00-02-L      SSA Logical Disk Drive
   hdisk3      Available    00-02-L      SSA Logical Disk Drive
   ```

   Enter the name in the **Logical Device Name** field.

2. Verify that all disks have a status of “Available.”

3. Use the following command for unavailable disks:

   ```
   mkdev -l hdiskx
   ```
   where `hdiskx` is the logical name of the unavailable disk.

At this point, the IBM SSA disks are configured for your cluster.

Configuring Target Mode SCSI Connections

This section describes how to configure a target mode SCSI connection between nodes sharing disks connected to a SCSI-2 Differential bus. Before you can configure a target mode SCSI connection, all nodes that share the disks must be connected to the SCSI bus, and all nodes and disks must be powered on.

Note: Neither PCI SCSI-2 differential busses nor SE busses support target mode SCSI.

Checking the Status of SCSI Adapters and Disks

To define a target mode SCSI connection, each SCSI adapter on nodes that share disks on the SCSI bus must have a unique ID and must be "Defined," known to the system but not yet available. Additionally, all disks assigned to an adapter must also be "Defined" but not yet available.

Note: The uniqueness of adapter SCSI IDs ensures that tmscsi devices created on a given node reflect the SCSI IDs of adapters on other nodes connected to the same bus.

To check the status of SCSI adapters you intend to use, enter:

```
lsdev -C | grep scsi
```

If an adapter is "Defined," see "Enabling Target Mode SCSI Devices in AIX" below to configure the target mode connection.

To check the status of SCSI disks on the SCSI bus, enter:

```
lsdev -Cc disk
```

If either an adapter or disk is "Available," follow the steps in the procedure below to return both the adapter (and its disks) to a defined state so that they can be configured for target mode SCSI and made available.

Returning Adapters and Disks to a Defined State

For a SCSI adapter**,** use the following command to make "Defined" each available disk associated with an adapter:

```
rmdev -l hdiskx
```

where *hdiskx* is the hdisk to be made "Defined."

For example:

```
rmdev -l hdisk3
```

Next, run the following command to return the SCSI adapter to a "Defined" state:

```
rmdev -l scsix
```

where *scsix* is the adapter to be made "Defined."

If using an array controller, you use the same command to return a router and a controller to a “Defined” state. However, make sure to perform these steps after changing the disk and before changing the adapter. The following lists these steps in this order:

```
rmdev -l hdiskx
rmdev -l darx
rmdev -l dacx
rmdev -l scsix
```

When all controllers and disks are “Defined,” see “Enabling Target Mode SCSI Devices in AIX” to enable the Target Mode connection.

Note: Target mode SCSI is automatically configured if you are using the SCSI-2 Differential Fast/Wide Adapter. Skip ahead to the section on “Follow-up Task.”Enabling Target Mode SCSI Devices in AIX

To define a target mode SCSI device:

1. Enable the target mode interface for the SCSI adapter.
2. Configure (make available) the devices.

Complete both steps on one node, then on the second node.

Enabling Target Mode Interface

To enable the target mode interface:

1. Enter:

   ```
   smit devices
   ```
   SMIT displays a list of devices.
2. Select **SCSI Adapter** and press Enter.
3. Select **Change/Show Characteristics of a SCSI Adapter** and press Enter.

   SMIT prompts you to identify the SCSI adapter.
4. Set the **Enable TARGET MODE interface** field to **yes** to enable the target mode interface on the device (the default value is no).

   At this point, a target mode SCSI device is generated that points to the other cluster nodes that share the SCSI bus. Note, however, that the SCSI ID of the adapter on the node from which you enabled the interface will not be listed.
5. Press Enter to commit the value.
6. Press F10 to exit SMIT.

Configuring the Target Mode SCSI Device

After enabling the target mode interface, you must run **cfgmgr** to create the initiator and target devices and make them available. To configure the devices and make them available:

1. Enter

   ```
   smit devices
   ```
   SMIT displays a list of devices.
2. Select **Install/Configure Devices Added After IPL** and press Enter.
3. Press F10 to exit SMIT after the **cfgmgr** command completes.

4. Run the following command to ensure that the devices are paired correctly:

   ```
   lsdev -Cc tmsci
   ```

Repeat the above procedure (enabling and configuring the target mode SCSI device) for other nodes connected to the SCSI-2 bus.

Target Mode Files

Configuring the target mode connection creates two special files in the **/dev** directory of each node, the **/dev/tmscsi*nn*.im** and **/dev/tmscsi*nn*.tm** files. The file with the**.im** extension is the initiator, which transmits data. The file with the **.tm** extension is the target, which receives data.

Testing the Target Mode Connection

For the target mode connection to work, initiator and target devices must be paired correctly. To ensure that devices are paired and that the connection is working after enabling the target mode connection on both nodes:

Enter the following command on a node connected to the bus.

```
cat < /dev/tmscsinn.tm
```

where *nn* must the logical name representing the target node. (This command hangs and waits for the next command.) On the target node, enter the following command:

```
cat filename > /dev/tmscsinn.im
```

where *nn* must be the logical name of the sending node and *filename* is a file.

The contents of the specified file are displayed on the node on which you entered the first command.

Note: Target mode SCSI devices are not always properly configured during the AIX boot process. Ensure that all tmscsi initiator devices are available on all cluster nodes before bringing up the cluster. Use the "`lsdev -Cc tmscsi`" command to ensure that all devices are available. See the *HACMP for AIX Troubleshooting Guide* for more information regarding problems with target mode SCSI devices.

Note: If the SCSI bus is disconnected while running as a target mode SCSI network, you must shut down HACMP for AIX before reattaching the SCSI bus to that node. *Never attach to a running system.*

Follow-Up Task

After you have installed and tested the target mode SCSI bus, you must define the target mode connection as a serial network to the HACMP for AIX cluster environment. See Testing the Target Mode Connection on page 4-7 for more information.

Configuring Target Mode SSA Connections

This section describes how to configure a target mode SSA connection between nodes sharing disks connected to SSA on Multi-Initiator RAID adapters (FC 6215 and FC 6219). The adapters must be at Microcode Level 1801 or later.

You can define a serial network to HACMP that connects all nodes on an SSA loop.

Changing Node Numbers on Systems in SSA Loop

By default, node numbers on all systems are zero. In order to configure the target mode devices, you must first assign a unique non-zero node number to all systems on the SSA loop.

1. To change the node number use the following command.
   ```
   chdev -l ssar -a node_number=#
   ```
2. To show the system's node number use the following command.
   ```
   lsattr -El ssar
   ```

Configuring Target Mode SSA Devices

After enabling the target mode interface, you must run **cfgmgr** to create the initiator and target devices and make them available. To configure the devices and make them available:

1. Enter
   ```
   smit devices
   ```
   SMIT displays a list of devices.
2. Select **Install/Configure Devices Added After IPL** and press Enter.
3. Press F10 to exit SMIT after the **cfgmgr** command completes.
4. Run the following command to ensure that the devices are paired correctly:
   ```
   lsdev -Cc tmssa
   ```

Repeat the above procedure (enabling and configuring the target mode SSA device) for other nodes connected to the SSA adapters.

Target Mode Files

Configuring the target mode connection creates two special files in the **/dev** directory of each node, the **/dev/tmssa#.im** and **/dev/tmssa#.tm** files. The file with the **.im** extension is the initiator, which transmits data. The file with the **.tm** extension is the target, which receives data.

Testing the Target Mode Connection

For the target mode connection to work, initiator and target devices must be paired correctly. To ensure that devices are paired and that the connection is working after enabling the target mode connection on both nodes:

1. Enter the following command on the initiator node connected to the SSA disks.
   ```
   cat < /dev/tmssa#.tm
   ```
   where # must be the number of the target node. (This command hangs and waits for the next command.)

2. On the target node, enter the following command:

   ```
   cat filename > /dev/tmssa#.im
   ```

   where # must be the number of the sending node and *filename* is a file.

The contents of the specified file are displayed on the node on which you entered the first command.

3. You can also check that the tmssa devices are available on each system using the following command:

   ```
   lsdev -C | grep tmssa
   ```

Follow-Up Task

After you have installed and tested the target mode SSA, you must define the target mode connection as a serial network to the HACMP for AIX cluster environment. See Defining the Target Mode Connection to HACMP on page 4-8 for more information.

Chapter 6 Defining Shared LVM Components

This chapter describes how to define the LVM components shared by cluster nodes in an HACMP for AIX cluster environment.

Prerequisites

- Complete the shared volume group worksheets following the instructions in Chapter 6, Planning Shared LVM Components, of the *HACMP for AIX Planning Guide*.
- Set up your shared disk configuration following the instructions in Chapter 5, Installing Shared Disk Devices.

Overview

Creating the volume groups, logical volumes, and filesystems shared by the nodes in an HACMP cluster requires that you perform steps on all nodes in the cluster. In general, you define the components on one node (referred to in the text as the source node) and then import the volume group on the other nodes in the cluster (referred to as destination nodes). This ensures that the ODM definitions of the shared components are the same on all nodes in the cluster.

Non-concurrent access environments typically use journaled filesystems to manage data, while concurrent access environments use raw logical volumes. This chapter provides different instructions for defining shared LVM components in non-concurrent access and concurrent access environments.

TaskGuide for Creating Shared Volume Groups

The TaskGuide is a graphical interface that simplifies the task of creating a shared volume group within an HACMP cluster configuration. The TaskGuide presents a series of panels that guide the user through the steps of specifying initial and sharing nodes, disks, concurrent or non-concurrent access, volume group name, physical partition size, and cluster settings. The TaskGuide can reduce errors, as it does not allow a user to proceed with steps that conflict with the cluster's configuration. Online help panels give additional information to aid in each step.

The TaskGuide for creating a shared volume group was introduced in HACMP 4.3.0. In version 4.4, the TaskGuide has two enhancements: it automatically creates a JFS log, as you would need to do manually when creating a shared volume group without the TaskGuide. In addition, it now displays the physical location of available disks.

Note that you may still want to rename and mirror the default JFS log after creating the shared volume group, as discussed on page 6 -5.

TaskGuide Requirements

Before starting the TaskGuide, make sure:

- You have a configured HACMP cluster in place.
- You are on a graphics capable terminal.
- You have set the display to your machine using your IP address or an alias, for example:

```
export DISPLAY=<your IP address>:0.0
```

Starting the TaskGuide

If you have the TaskGuide filesets installed and your display set properly, you can start the TaskGuide from the command line by typing

```
/usr/sbin/cluster/tguides/bin/cl_ccvg
```

or you can use the SMIT interface as follows:

1. Type `smit hacmp`
2. From the SMIT main menu, choose **Cluster System Management > Cluster Logical Volume Manager >Taskguide for Creating a Shared Volume Group**

   After a pause, the TaskGuide “Welcome” panel appears.
3. Proceed through the panels to create or share a volume group.

   In the last panel, you have the option to cancel or to back up and change what you have entered. If you are satisfied with your entries, click **Apply** to create the shared volume group.

Defining Shared LVM Components for Non-Concurrent Access

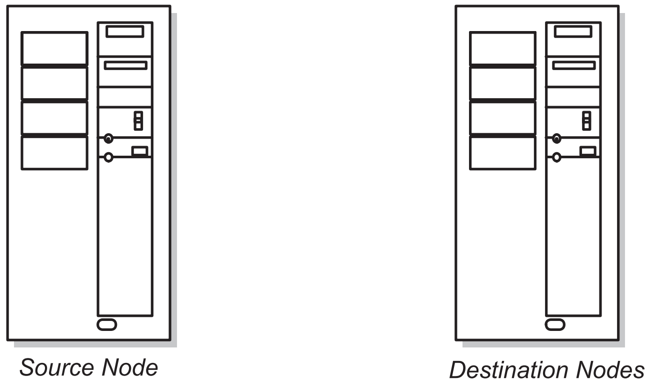
Non-concurrent access environments typically use journaled filesystems to manage data. (In some cases, a database application running in non-concurrent access environments may bypass the journaled filesystem and access the raw logical volume directly.)

The key consideration, however, is whether a non-concurrent access environment uses mirrors. Shared logical volumes residing on non-RAID disk devices should be mirrored in AIX to eliminate the disk as a single point of failure. Shared volume groups residing on a IBM 7135, IBM 9333, SSA with RAID enabled, or IBM 2105 Versatile Storage Server RAID devices should not be AIX mirrored; the disk array provides its own data redundancy.

Note: The discussion of the RAID disk device assumes you are using RAID level 1, 3, or 5. RAID level 0 does not provide data redundancy and therefore *is not* recommended for use in an HACMP for AIX configuration.

The following figures lists the tasks you complete to define the shared LVM components for non-concurrent access environments using either non-RAID disk subsystems or RAID disk devices. Each task is described throughout the pages following the figures. Refer to your completed copies of the shared volume group worksheets as you define the shared LVM components.

Non-Concurrent Access Using Non-RAID Subsystems

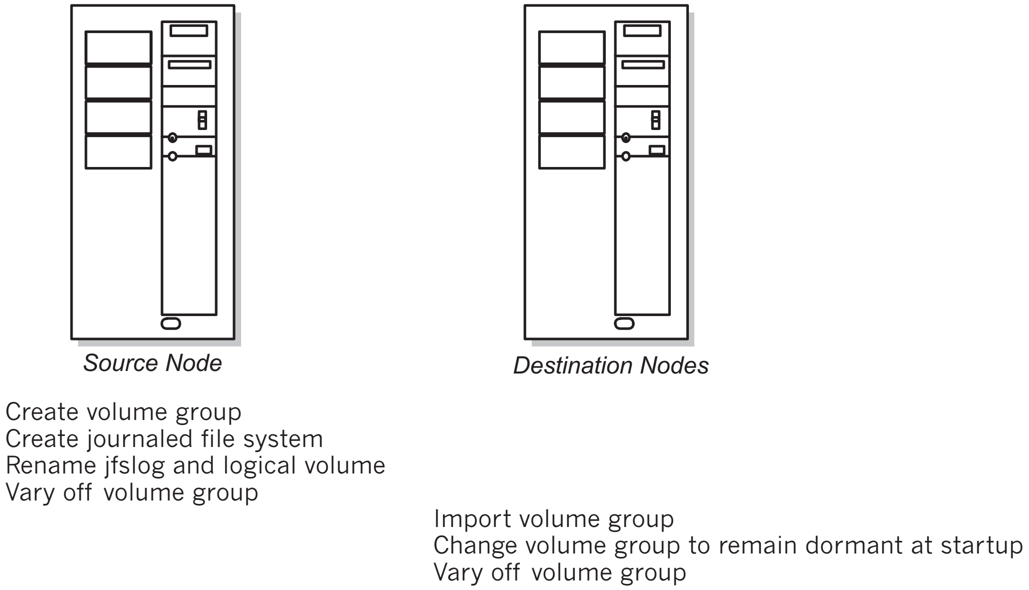


Create volume group
Create journaled file system
Rename jfslog and logical volume
Mirror jfslog and logical volume
Vary off volume group

Import volume group
Change volume group to remain dormant at startup
Vary off volume group

Defining Shared LVM Components in non-RAID, Non-Concurrent Access Environments

Non-Concurrent Access Using RAID Disk Devices


Source Node
Destination Nodes
Create volume group
Create journaled file system
Rename jfslog and logical volume
Vary off volume group
Import volume group
Change volume group to remain dormant at startup
Vary off volume group

Defining Shared LVM Components in RAID Device, Non-Concurrent Access Environments

Again, the difference between the procedures is that non-RAID disks require AIX mirrors, whereas RAID devices provide their own mirroring or similar redundancy.

Creating a Shared Volume Group on the Source Node

This section covers how to create a shared volume group on the source node using the HACMP SMIT interface. Use the **smit mkvg** fastpath to create a shared volume group. Use the default field values unless your site has other requirements, or unless you are specifically instructed otherwise here.

VOLUME GROUP name
The name of the shared volume group should be unique within the cluster.

Activate volume group AUTOMATICALLY at system restart?
Set to **no** so that the volume group can be activated as appropriate by the cluster event scripts.

ACTIVATE volume group after it is created?
Set to **yes**.

Volume Group MAJOR NUMBER
If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must make sure to use the same major number on all nodes. Use the **lvlstmajor** command on each node to determine a free major number common to all nodes.

Creating a Shared Filesystem on the Source Node

Use the **smit crjfs** fast path to create the shared filesystem on the source node. When you create a journaled filesystem, AIX creates the corresponding logical volume. Therefore, you do not need to define a logical volume. You do, however, need to later rename both the logical volume and the log logical volume for the filesystem and volume group.

Mount AUTOMATICALLY at system restart?	Make sure this field is set to **no**.
Start Disk Accounting	Make sure this field is set to **no**.

Renaming jfslogs and Logical Volumes on the Source Node

AIX assigns a logical volume name to each logical volume it creates. Examples of logical volume names are */dev/lv00* and */dev/lv01*. Within an HACMP cluster, the name of any shared logical volume must be unique. Also, the journaled filesystem log (**jfslog**) is a logical volume that requires a unique name in the cluster.

To make sure that logical volumes have unique names, rename the logical volume associated with the filesystem and the corresponding **jfslog** logical volume. Use a naming scheme that indicates the logical volume is associated with a certain filesystem. For example, *lvsharefs* could name a logical volume for the */sharefs* filesystem.

1. Use the **lsvg -l** ***volume_group_name*** command to determine the name of the logical volume and the log logical volume (**jfslog**) associated with the shared volume groups. In the resulting display, look for the logical volume name that has type **jfs**. This is the logical volume. Then look for the logical volume name that has type **jfslog**. This is the log logical volume.
2. Use the **smit chlv** fastpath to rename the logical volume and the log logical volume.

   After renaming the **jfslog** or a logical volume, check the **/etc/filesystems** file to make sure the **dev** and **log** attributes reflect the change. Check the **log** attribute for *each* filesystem in the volume group and make sure that it has the new **jfslog** name. Check the **dev** attribute for the logical volume you renamed and make sure that it has the new logical volume name.

Adding Copies to Logical Volume on the Source Node

To add logical volume copies on a source node:

1. Use the **smit mklvcopy** fastpath to add copies to a logical volume. Add copies to both the **jfslog** log logical volume and the logical volumes in the shared filesystems. To avoid space problems, first mirror the **jfslog** log logical volume and then the shared logical volumes.

   The copies should reside on separate disks that are controlled by different disk adapters and are located in separate drawers or units, if possible.

   Note: These steps do not apply to RAID devices, which provide their own mirroring of logical volumes. Continue with “Test Filesystem.”

2. Verify the number of logical volume copies. Enter:

   ```
   lsvg -l volume_group_name
   ```

In the resulting display, locate the line for the logical volume for which you just added copies. Notice that the number in the physical partitions column is *x* times the number in the logical partitions column, where *x* is the number of copies.

3. To verify the placement of logical volume copies, enter:

   ```
   lspv -l hdiskx
   ```

   where *hdiskx* is the name of each disk to which you assigned copies. That is, you enter this command for each disk. In the resulting display, locate the line for the logical volume for which you just added copies. For copies placed on separate disks, the numbers in the logical partitions column and the physical partitions column should be equal. Otherwise, the copies were placed on the same disk and the mirrored copies will not protect against disk failure.

Testing a Filesystem

To run a consistency check on each filesystem's information:

1. Enter:

   ```
   fsck /filesystem_name
   ```

2. Verify that you can mount the filesystem by entering:

   ```
   mount /filesystem_name
   ```

3. Verify that you can unmount the filesystem by entering:

   ```
   umount /filesystem_name
   ```

Varying Off a Volume Group on the Source Node

After completing the previous tasks, use the **varyoffvg** command to deactivate the shared volume group. You vary off the volume group so that it can be properly imported onto a destination node and activated as appropriate by the cluster event scripts. Enter the following command:

```
varyoffvg volume_group_name
```

Importing a Volume Group onto the Destination Node

This section covers how to import a volume group onto destination nodes using the SMIT interface. You can also use the TaskGuide utility for this task.The TaskGuide uses a graphical interface to guide you through the steps of adding nodes to an existing volume group. For more information on the TaskGuide, see the TaskGuide description and instructions on page 6 -1.

Importing the volume group onto the destination nodes synchronizes the ODM definition of the volume group on each node on which it is imported.

You can use the **smit importvg** fastpath to import the volume group.

VOLUME GROUP name	Enter the name of the volume group that you are importing. Make sure the volume group name is the same name that you used on the source node.

PHYSICAL VOLUME name Enter the name of a physical volume that resides in the volume group. Note that a disk *may have* a different physical name on different nodes. Make sure that you use the disk name as it is defined on the destination node.

ACTIVATE volume group after it is imported? Set the field to **yes**.

Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must make sure to use the same major number on all nodes. Use the **lvlstmajor** command on each node to determine a free major number common to all nodes.

Changing a Volume Group’s Startup Status

By default, a volume group that has just been imported is configured to automatically become active at system restart. In an HACMP for AIX environment, a volume group should be varied on as appropriate by the cluster event scripts. Therefore, after importing a volume group, use the SMIT Change a Volume Group screen to reconfigure the volume group so that it is not activated automatically at system restart.

Use the **smit chvg** fastpath to change the characteristics of a volume group.

Activate volume group automatically at system restart? Set this field to **no**.

A QUORUM of disks required to keep the volume group online? This field is site-dependent. See Chapter 6, Planning Shared LVM Components, of the *HACMP for AIX Planning Guide* for a discussion of quorum in an HACMP cluster.

Varying Off the Volume Group on the Destination Nodes

Use the **varyoffvg** command to deactivate the shared volume group so that it can be imported onto another destination node or activated as appropriate by the cluster event scripts. Enter:

```
varyoffvg volume_group_name
```

Defining Shared LVM Components for Concurrent Access

Concurrent access does not support filesystems. Instead, you must use logical volumes.

This section describes the procedure for defining shared LVM components for a concurrent access environment. Concurrent access is supported on the following devices:

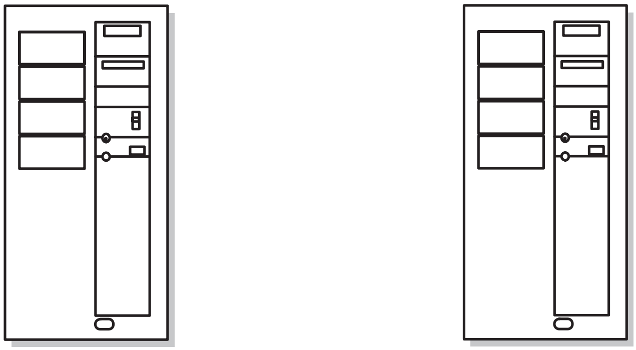
- IBM 7135-110 and 210 RAIDiant Disk Array
- IBM 7137 Disk Arrays
- IBM 9333 serial disk subsystems

- IBM 7133 or 7131-405 SSA Disk Subsystems
- IBM 2105-B09 and 100 Versatile Storage Servers

Refer to your completed copies of the shared volume group worksheets as you define the shared LVM components.

Creating a Concurrent Access Volume Group on a Source Node

The figure below summarizes the steps you must complete on the source and destination nodes in an HACMP cluster to create a concurrent capable volume group that HACMP can vary on in concurrent access mode.



Source Node	*Destination Nodes*
1. Complete prerequisite tasks	1. Complete prerequisite tasks
2. Create concurrent access volume group	
	3. Import volume group
4. Vary on the volume group in non-concurrent mode	
5. Create logical volumes	
	6. Vary off volume group
	7. Change volume group to remain dormant at startup
8. Complete follow-up tasks	8. Complete follow-up tasks

Creating a Concurrent Access Volume Group

The above steps are described in detail throughout this section.

Step 1. Complete Prerequisite Tasks

The physical volumes (hdisks) should be installed, configured, and available. You can verify the disks' status using the **lsdev -Cc disk** command.

Step 2. Create a Concurrent Access Volume Group on Source Node

The procedure used to create a concurrent access volume group varies depending on which type of device you are using: serial disk subsystem or RAID disk subsystem.

Warning: If you are creating (or plan to create) concurrent volume groups on SSA devices, be sure to assign unique non-zero node numbers through the ssar on each cluster node before using the failed drive replacement procedure described in Chapter 5 of the *HACMP for AIX Administration Guide.* If you plan to specify SSA disk fencing in your concurrent resource group, the node numbers are assigned when you synchronize resources. If you do not specify SSA disk fencing, assign node numbers using the following command: `chdev -l ssar -a node_number=x`, where x is the number to assign to that node. You must reboot the system to effect the change.

Creating a Concurrent Access Volume Group on Serial Disk Subsystems

To use a concurrent access volume group, defined on a serial disk subsystem such as an IBM 9333 or IBM 7133 disk subsystem, you must create it as a *concurrent capable* volume group. A concurrent capable volume group can be activated (varied on) in either non-concurrent mode or concurrent access mode. To define logical volumes on a concurrent capable volume group, it must be varied on in non-concurrent mode.

Use the **mkvg** command, specifying the **-c** flag, to create the concurrent capable volume group, as in the following example:

```
mkvg -n -s 4 -c -y myvg hdisk1 hdisk2
```

You can also use SMIT to build the **mkvg** command by using the following procedure:

1. To create a concurrent capable volume group, enter:

   ```
   smit mkvg
   ```

   SMIT displays the Add a Volume Group screen. Enter the specific field values as follows:

   VOLUME GROUP name Specify name of volume group.

   Physical partition SIZE in megabytes Accept the default.

   PHYSICAL VOLUME NAMES Specify the names of the physical volumes you want included in the volume group.

   Activate volume group AUTOMATICALLY at system restart? Set this field to **no** so that the volume group can be activated as appropriate by the cluster event scripts.

   ACTIVATE volume group after it is created? Set this field to **no.**

   Volume Group MAJOR NUMBER Accept the default.

   Create VG concurrent capable? Set this field to **yes** so that the volume group can be activated in concurrent access mode by the HACMP for AIX event scripts.

Auto-varyon concurrent mode? Set this field to **no** so that the volume group can be activated as appropriate by the cluster event scripts.

2. Press Enter.

   SMIT responds:

   ARE YOU SURE?
3. Press Enter.
4. Press F10 to exit SMIT after the command completes.

Creating a Concurrent Access Volume Group on RAID Disk Subsystems

To create a concurrent access volume group on a RAID disk subsystem, such as an IBM 7135 disk subsystem or IBM 2105 Versatile Storage Server, you follow the same procedure as you would to create a non-concurrent access volume group. A concurrent access volume group can be activated (varied on) in either non-concurrent mode or concurrent access mode. To define logical volumes on a concurrent access volume group, it must be varied on in non-concurrent mode.

Use the **smit mkvg** fastpath to create a shared volume group. Use the default field values unless your site has other requirements, or unless you are specifically instructed otherwise.

VOLUME GROUP name The name of the shared volume group should be unique within the cluster.

Activate volume group AUTOMATICALLY at system restart? Set to no **so** that the volume group can be activated as appropriate by the cluster event scripts.

ACTIVATE volume group after it is created? Set to **yes**.

Volume Group MAJOR NUMBER Use the default (which is the next available number in the valid range).

Create VG concurrent capable? Set this field to **no**.

Step 3. Import Volume Group Information on Destination Nodes

For this step, you use the **importvg** command or the SMIT interface. You can also use the TaskGuide graphical interface to guide you through the process. For more information on using the TaskGuide, refer back to the section TaskGuide for Creating Shared Volume Groups on page 6-1

To import volume group information to destination nodes:

On each destination node, import the volume group, using the **importvg** command, as in the following example:

```
importvg -y vg_name physical_volume_name
```

Specify the name of any disk in the volume group as an argument to the **importvg** command. By default, AIX automatically varies on non-concurrent capable volume groups when they are imported. AIX does *not* automatically vary on concurrent capable volume groups when they are imported.

You can also build the **importvg** command through SMIT using the procedure below.

To import a concurrent capable volume group using SMIT:

1. Enter:

   ```
   smit importvg
   ```

   SMIT displays the Import a Volume Group SMIT screen.

2. Enter specific field values as follows. For other fields, use the defaults or the appropriate entries for your operation:

   VOLUME GROUP name Enter the name of the volume group that you are importing. Make sure the volume group name is the same name that you used on the source node.

   PHYSICAL VOLUME name Enter the name of one of the physical volumes that resides in the volume group. *Note that a disk may have a different hdisk number on different nodes. Make sure that you use the disk name as it is defined on the destination node.* Use the **lspv** command to map the hdisk number to the PVID. The PVID uniquely identifies a disk.

   ACTIVATE volume group after it is imported? Set the field to **no**.

   Volume Group MAJOR NUMBER Accept the default.

   Make this VG concurrent capable Accept the default.

   Make default varyon of VG concurrent Accept the default.

3. Press Enter to commit the information. Press F10 to exit SMIT and return to the command line.

If your cluster uses SCSI external disks (including RAID devices) and the import of the volume group fails, check that no reserve exists on any disk in the volume group by executing the following command, only after installing the HACMP for AIX software as described in Chapter 8, Installing HACMP for AIX Software:

```
/usr/sbin/cluster/events/utils/cl_scdiskreset /dev/hdiskn ...
```

For example, if the volume group consists of *hdisk1* and *hdisk2*, enter:

```
/usr/sbin/cluster/events/utils/cl_scdiskreset /dev/hdisk1 /dev/hdisk2
```

Step 4. Vary On the Concurrent Capable Volume Group in Non-concurrent Mode

Use the **varyonvg** command to activate a volume group in non-concurrent mode. To create logical volumes, the volume group must be varied on in non-concurrent access mode. For example, to vary on the concurrent capable volume group **myvg** in non-concurrent access mode, enter the following command:

```
varyonvg myvg
```

You can also use SMIT to build the **varyonvg** command by using the following procedure.

1. To vary on a concurrent capable volume group in non-concurrent mode, enter:

   ```
   smit varyonvg
   ```

   SMIT displays the Add a Volume Group SMIT screen. Enter the specific field values as follows.

VOLUME GROUP name	Specify name of volume group.
RESYNCHRONIZE stale physical partitions?	Set this field to **no**.
Activate volume group in SYSTEM MANAGEMENT mode?	Accept the default.
FORCE activation of the volume group?	Accept the default.
Varyon volume group in concurrent mode	Accept the default. To create logical volumes on the volume group, it must be varied on in non-concurrent mode.

2. Press Enter

   SMIT responds:

   ARE YOU SURE?

3. Press Enter.
4. Press F10 to exit SMIT after the command completes.

Step 5. Create Logical Volumes on Concurrent Capable Volume Group on Source Node

Create logical volumes on the volume group, specifying logical volume mirrors to provide data redundancy. If the volume group is varied on in concurrent access mode, you will not be able to create logical volumes. A concurrent capable volume group must be varied on in non-concurrent access mode to create logical volumes on it.

For more information about creating logical volumes, see the *AIX System Management Guide: Operating System and Devices*.

1. To create a logical volume, enter:

   ```
   smity mklv
   ```

   SMIT displays the Add a Logical Volume menu.

2. You must specify the size of the logical volume as the number of logical partitions. Accept default values for all other fields except the following:

   Logical volume name Specify name of logical volume. Name must be the same on all cluster nodes.

   VOLUME GROUP name Specify name of shared volume group.

   PHYSICAL VOLUME names? Specify the physical volumes you want the logical volume to include.

   Number of COPIES of each logical partition Specify 1, 2, or 3 mirror copies.

   If You defined the volume group on an IBM 7135-110 or IBM 7135-210 disk array, do not create mirror copies. Instead, use the data redundancy provided by RAID levels 1, 3, or 5.

   Mirror Write Consistency Specify the value to **no.**

   Enable BAD BLOCK relocation? Specify the value **no**.

Step 6. Vary Off Volume Group on Source Node

After creating the logical volume, vary off the volume group using the **varyoffvg** command so that it can be varied on by the HACMP for AIX scripts. Enter:

```
varyoffvg volume_group_name
```

Step 7. Change Volume Group to Remain Dormant at Startup on Destination Nodes

By default, AIX configures an imported volume group to automatically become active at system restart. In the HACMP for AIX system, a volume group should be varied on as appropriate by the HACMP for AIX scripts. Therefore, after importing a volume group, you must reconfigure the volume group so that it remains dormant at startup.

To change the startup state of a volume group, enter:

```
chvg -a n volume_group_name
```

You can also build the **chvg** command through SMIT using the following procedure:

1. Use the **smit chvg** fastpath to change the characteristics of a volume group.

   Enter the specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

   Set the **Activate volume group automatically at system restart?** field to **no**.

2. Press Enter to commit this change.

3. Press F10 to exit SMIT and return to the command line.

Step 8. Complete Follow-up Tasks

Verify that the HACMP for AIX scripts can activate the concurrent capable volume group as a concurrent cluster resource.

Chapter 7 Additional AIX Administrative Tasks

This chapter discusses several general tasks necessary to ensure that your HACMP for AIX cluster environment works as planned.

AIX Considerations

Consider or check the following issues to ensure that AIX works as expected in an HACMP cluster.

- I/O pacing
- User and group IDs
- Network option settings
- **/etc/hosts** file and nameserver edits
- **/.rhosts** file edits
- NFS configurations
- Managing applications using the SPX/IPX protocol.

I/O Pacing

AIX users have occasionally seen poor interactive performance from some applications when another application on the system is doing heavy input/output. Under certain conditions I/O can take several seconds to complete. While the heavy I/O is occurring, an interactive process can be severely affected if its I/O is blocked or if it needs resources held by a blocked process.

Under these conditions, the HACMP for AIX software may be unable to send keepalive packets from the affected node. The Cluster Managers on other cluster nodes interpret the lack of keepalive as node failure, and the I/O-bound node is “failed” by the other nodes. When the I/O finishes, the node resumes sending keepalives. Its packets, however, are now out of sync with the other nodes, which then kill the I/O-bound node with a RESET packet.

You can use I/O pacing to tune the system so that system resources are distributed more equitably during high disk I/O. You do this by setting high- and low-water marks. If a process tries to write to a file at the high-water mark, it must wait until enough I/O operations have finished to make the low-water mark.

While enabling I/O pacing may have a slight performance effect on very I/O intensive processes, it is required for an HACMP cluster to behave correctly during large disk writes. If you anticipate heavy I/O on your HACMP cluster, you should enable I/O pacing.

See Configuring Cluster Performance Tuning on page 4-20 for information on setting these AIX parameters.

Checking User and Group IDs

If a node in a cluster fails, users should be able to log on to the surviving nodes without experiencing problems caused by mismatches in the user or group IDs. To avoid mismatches, make sure that user and group information is propagated to nodes as necessary. Also ensure that the **/etc/passwd** and **/etc/security** files are the same on all nodes because user and group IDs should be the same on all nodes. For information about checking and managing C-SPOC user accounts and groups, see the chapter on managing user accounts and groups in a cluster in the *HACMP for AIX Administration Guide.*

Checking Network Option Settings

The default settings for the following network options should be changed as described in this section:

- thewall
- routerevalidate

Changing thewall Network Option

To ensure that HACMP for AIX requests for memory are handled correctly, you can set (on every cluster node) “thewall” network option to be higher than its default value. The suggested value for his option is shown below:

```
thewall = 5120
```

1. To change this default value, add the following line to the end of the **/etc/rc.net** file:
   ```
   no -o thewall=5120
   ```
2. After making this change, monitor mbuf usage using the **netstat -m** command and increase or decrease “thewall” option as needed.
3. To list the values of other network options that are currently set on a node, enter:
   ```
   no -a
   ```

Changing routerevalidate Network Option

Changing hardware and IP addresses within HACMP changes and deletes routes. Due to the fact that AIX caches routes, it is required that you set the “routerevalidate” network option as follows:

```
routerevalidate=1
```

This setting ensures the maintenance of communication between cluster nodes.

- To change the default value, add the following line to the end of the **/etc/rc.net** file:
  ```
  no -o routerevalidate=1
  ```

Editing the /etc/hosts File and Nameserver Configuration

Make sure all nodes can resolve all cluster addresses. See the chapter on planning TCP/IP networks (the section Using HACMP with NIS and DNS) in the *HACMP for AIX Planning Guide*, for more information on nameserving and HACMP.

Edit the **/etc/hosts** file (and the **/etc/resolv.conf** file, if using the **nameserver** configuration) on each node in the cluster to make sure the IP addresses of all clustered interfaces are listed.

For each boot address, make an entry similar to the following:

```
100.100.50.200 crab_boot
```

Also, make sure that the **/etc/hosts** file on each node has the following entry:

```
127.0.0.1   loopback localhost
```

cron and NIS Considerations

If your HACMP cluster nodes use NIS services which include the mapping of the **/etc/passwd** file and IPAT is enabled, users that are known only in the NIS-managed version of the **/etc/passwd** file will not be able to create crontabs. This is because **cron** is started via the /**etc/inittab** file with run level 2 (for example, when the system is booted), but **ypbind** is started in the course of starting HACMP via the **rcnfs** entry in **/etc/inittab**. When IPAT is enabled in HACMP, the run level of the **rcnfs** entry is changed to **-a** and run via the **telinit -a** command by HACMP.

In order to let those NIS-managed users create crontabs, you can do one of the following:

1. Change the runlevel of the **cron** entry in **/etc/inittab** to **-a** and make sure it is positioned after the **rcnfs** entry in **/etc/inittab**. This solution is recommended if it is acceptable to start **cron** after HACMP has started.
2. Add an entry to the **/etc/inittab** file like the following script with runlevel **-a.** Make sure it is positioned after the **rcnfs** entry in **/etc/inittab**. The important thing is to kill the **cron** process, which will respawn and know about all of the NIS-managed users. Whether or not you log the fact that **cron** has been refreshed is optional.

```
#! /bin/sh
 # This script checks for a ypbind and a cron process. If both
 # exist and cron was started before ypbind, cron is killed so
 # it will respawn and know about any new users that are found
 # in the passwd file managed as an NIS map.
 echo "Entering $0 at ‘date‘" >> /tmp/refr_cron.out
 cronPid=‘ps -ef |grep "/etc/cron" |grep -v grep |awk \
 ’{ print $2 }’‘
 ypbindPid=‘ps -ef | grep "/usr/etc/ypbind" | grep -v grep | \
 if [ ! -z "${ypbindPid}" ]
 then
     if [ ! -z "${cronPid}" ]
     then
         echo "ypbind pid is ${ypbindPid}" >> /tmp/refr_cron.out
         echo "cron pid is ${cronPid}" >> /tmp/refr_cron.out
         echo "Killing cron(pid ${cronPid}) to refresh user \
               list" >> /tmp/refr_cron.out
         kill -9 ${cronPid}
         if [ $? -ne 0 ]
         then
             echo "$PROGNAME: Unable to refresh cron." \
             >>/tmp/refr_cron.out
             exit 1
        fi
     fi
 fi
 echo "Exiting $0 at ‘date‘" >> /tmp/refr_cron.out
 exit 0
```

Editing the /.rhosts File

Make sure that each node's service adapters and boot addresses are listed in the **/.rhosts** file on each cluster node. Doing so allows the **/usr/sbin/cluster/utilities/clruncmd** command and the **/usr/sbin/cluster/godm** daemon to run. The **/usr/sbin/cluster/godm** daemon is used when nodes are configured from a central location.

For security reasons, IP label entries you add to the **/.rhosts** file to identify cluster nodes should be deleted when you no longer need to log on to a remote node from these nodes. The cluster synchronization and verification functions use **rcmd** and **rsh** and thus require these **/.rhosts** entries. These entries are also required to use C-SPOC commands in a cluster environment. The **/usr/sbin/cluster/clstrmgr** daemon, however, does not depend on **/.rhosts** file entries.

Note: The **/.rhosts** file is not required on SP systems running the HACMP Enhanced Security. This feature removes the requirement of TCP/IP access control lists (for example, the **/.rhosts** file) on remote nodes during HACMP configuration.

DCE Authentication

As of AIX 4.3.1, Kerberos Version 5 (DCE authentication) can be used on non-SP RS/6000 systems. However, Kerberos 5 is the *only* method of authentication, you will not be able to alter, synchronize, or verify the HACMP configuration. Note that you will not be able to explicitly move a resource group, since that is a type of reconfiguration. If DCE (i.e. only Kerberos V5) is enabled as an authentication method for the AIX remote commands, you can still use HACMP, but must perform the following steps:

1. Prior to configuring the HACMP cluster, enable Kerberos V4 or Standard AIX as an authentication method for the AIX remote commands on the cluster nodes, and create the remote command authorization files (either **/.klogin** or **/.rhosts** files) for the root user on those nodes. This provides the ability for root to **rsh** among those nodes.
2. Configure the HACMP cluster.
3. Remove the remote command authorization files created in step 1 on the HACMP cluster nodes.
4. Disable the Kerberos V4 or Standard AIX authentication method enabled in step 1 on the HACMP cluster nodes.

Reliable NFS Server Capability

An HACMP 4.4 two-node cluster can now take advantage of AIX extensions to the standard NFS functionality that enable it to handle duplicate requests correctly and restore lock state during NFS server fallover and reintegration. This support was previously only available in the HANFS feature. More detail can be found in the HACMP for AIX Planning Guide, Using NFS with HACMP on page 12-11.

Checking NFS Configurations

To ensure that NFS works as expected in an HACMP cluster, check your configuration in reference to the following topics:

- Major numbers on shared volume groups

- Exporting NFS filesystems
- NFS mounting
- SNMP configurations.

Major Numbers on Shared Volume Group

When a node detaches from the cluster, NFS clients attached to the cluster operate as they do when a standard NFS server fails and reboots.

To prevent problems with NFS filesystems in an HACMP cluster, make sure that each shared volume group has the same major number on all nodes. If the major numbers are different, client sessions will not be able to recover when a takeover node re-exports an NFS filesystem after an owner node has left the cluster. Client sessions are not able to recover because the filesystem exported by the takeover node appears to be different from the one exported by the owner node.

The **lvlstmajor** command lists the free major numbers on a node. Use this command on each node to find a major number that is free on all cluster nodes. Then record that number in the **Major Number** field on the *Shared Volume Group/Filesystem (Non-Concurrent Access)* worksheet in Appendix A, Planning Worksheets, of the *HACMP for AIX Planning Guide* for a non-concurrent access configuration.

NFS Exporting Filesystems and Directories

The process of NFS-exporting filesystems and directories in HACMP for AIX is different from that in AIX. Remember the following when NFS-exporting in HACMP:

Specifying Filesystems and Directories to NFS Export

While in AIX, you list filesystems and directories to NFS-export in the **/etc/exports** file, in HACMP for AIX, you must put these in a resource group.

Specifying Export Options for NFS-Exported Filesystems and Directories

If you want to specify special options in for NFS-exporting in HACMP, you can create a **/usr/sbin/cluster/etc/exports** file. This file has the same format as the regular **/etc/exports** file used in AIX.

Note: Use of this alternate exports file is optional. HACMP checks the **/usr/sbin/cluster/etc/exports** file when NFS-exporting a filesystem or directory. If there is an entry for the filesystem or directory in this file, HACMP will use the options listed. If the filesystem or directory for NFS-export is not listed in the file, or, if the user has not created the **/usr/sbin/cluster/etc/exports** file, the filesystem or directory will be NFS-exported with the default option of root access for all cluster nodes.

Configuring the Optional /usr/sbin/cluster/etc/exports File

In this step you add the directories of the shared filesystems to the exports file. Complete the following steps for each filesystem you want to add to the exports file. Refer to your NFS-Exported Filesystem Worksheet.

1. Enter the **smit mknfsexp** fastpath to display the **Add a Directory to Exports List** screen.
2. In the **EXPORT directory now**, **system restart or both** field, enter **restart**.

3. In the **PATHNAME of alternate Exports file** field, enter **/usr/sbin/cluster/etc/exports**. This step will create the alternate exports file which will list the special NFS export options.
4. Add values for the other fields as appropriate for your site, and press Enter. Use this information to update the **/usr/sbin/cluster/etc/exports** file.
5. Press F3 to return to the **Add a Directory to Exports List** screen, or F10 to exit SMIT.
6. Repeat steps 1 through 4 for each filesystem or directory listed in the **FileSystems/Directories to Export** field on your planning worksheets.

   Note: Remember that this alternate exports file does not specify *what* will be exported, only *how* it will be exported. To specify what to export, you must put it in a resource group.

NFS Mounting

Review the section Using NFS with HACMP on page 12-11 in the *HACMP for AIX Planning Guide* before setting up your NFS mount points.

SNMP Configurations

See Appendix E, HACMP for AIX and SNMP Utilities, for a discussion of the Simple Network Management Protocol (SNMP) and a description of the relationship between the HACMP for AIX SNMP-based utilities and other SNMP-based utilities that run on RS/6000 and SMP platforms.

Managing Applications That Use the SPX/IPX Protocol

If running applications or hardware that use the SPX/IPX protocol stack to establish a STREAMS connection through an adapter to another node, be aware that the SPX/IPX stack that is loaded to support applications like netware (or hardware, like the 7318 model P10 terminal server) is not unloaded by the **ifconfig detach** command, as is the IP stack, during a takeover or when attempting a hardware address swap. Thus, the **rmdev** command used in cl_swap_HW_address fails with a “device busy” error prior to changing the adapter hardware address.

To avoid this error, use the following command to unload the SPX/IPX stack:

```
strload -uf /etc/dlpi.conf
```

This command unloads the dlpi device driver. Likewise, when using 7318 P10-style ports, you may want to unload other drivers using commands similar to the following:

```
strload -uf /etc/xtiso.conf
strload -uf /etc/netware.conf
```

These commands must be issued after the SPX/IPX stack is loaded and before HACMP runs cl_swap_HW_address. For example, run these commands before the cl_swap_HW_address script is called by release_service_address or acquire_takeover_address.

You can run these commands in pre or post-event scripts that run prior to cl_swap_HW_address, or you can add them to the stop_server script that stops the application using the protocol. The specific scripts you use depend on when SPX/IPX-related drivers are loaded at your site.

Chapter 8 Installing HACMP for AIX Software

This chapter describes how to install the HACMP for AIX, Version 4.4 Licensed Program Product (LPP) on cluster nodes (servers). The chapter contains instructions for a new installation.

If you are installing HACMP for AIX on an RS/6000 SP node, see Appendix B or consult the *SP Installation and Migration Guide*, order number GA22-7347.

If you are installing HACMP for AIX on a client only, see Chapter 17, Installing and Configuring Clients.

Read the *HACMP for AIX Planning Guide* before installing the Version 4.4 software. It contains the worksheets and diagrams necessary to plan an HACMP for AIX installation and configuration.

Prerequisites

- Each cluster node must have AIX Version 4.3.2 or higher installed.
- The AIX optional component, *bos.adt*, is mandatory for the HACMP for AIX software to work. If you do not install this AIX component, the HACMP for AIX software will not function correctly.
- Each server requires its own HACMP for AIX software license.
- A root user must perform the installation.
- The **/usr** directory must have 50 megabytes (MB) of free disk space for a full install.
    - If you are not planning to install optional software you can plan for less space: HAView =14MB, TaskGuide= 400KB, VSM= 5.2MB. You should also choose to install the message catalogs for the language you will be using, rather than all message catalogs (Japanese message catalogs use 1.6MB).
- HAView requires that you install the following NetView filesets:

  on a server:

    - nv6000.base.obj 4.1.2.0 (or higher) and nv6000.database.obj 4.1.2.0 (or higher)

  or

    - nv6000.base.obj 5.0.0.0 or higher

  on a client:

    - nv6000.base.obj 4.1.2.0 (or higher)

  or

    - nv6000.client.obj 4.1.2.0 (or higher)

Contact your IBM Sales representative for information on obtaining NetView software.

- If you plan to monitor your cluster through Tivoli Framework, you must install the following Tivoli software:
    - Tivoli Framework version 3.6 or higher (on TMR and cluster nodes)

- Tivoli Distributed Monitoring version 3.5.1 or higher (on TMR and cluster nodes)
- Tivoli Application Extension Facility (AEF) version 3.6 or greater (on TMR only)

Contact your IBM sales representative for information on obtaining Tivoli software.

Overview of HACMP Software Installation

The HACMP for AIX software is distributed on media which contain the following

- The base High Availability Subsystem images, some of which you must install on all servers and clients. This feature provides the base services for cluster membership, system management, configuration integrity and control, fallover, and recovery. It also includes cluster status and monitoring facilities for programmers and system administrators.

  Note: If you are using a separate client to provide cluster status and monitoring facilities, see Chapter 17, Installing and Configuring Clients.

- The Concurrent Resource Manager (CRM) images. This feature optionally adds concurrent shared-access management for supported RAID and SSA disk subsystems. Concurrent access is provided at the raw logical volume level. Applications that use the CRM must be able to control access to the shared data. The CRM includes the High Availability Subsystem which provides distributed locking facilities to support access to shared data.

  Note: In HACMP for AIX Version 4.3.1 environments, concurrent access is available using only an IBM 7135-110 or 210 Disk Array, an IBM 7137 Disk Array, IBM 2105-B09 and 100 Versatile Storage Servers, an IBM 7133 SSA disk subsystem, or an IBM 9333 disk subsystem. RAID devices from other manufacturers may not support concurrent access.

If you are using an earlier version of the HACMP for AIX software and want to save your cluster configuration, see the chapter on saving and restoring cluster configurations in the *HACMP for AIX Administration Guide*. If you do not want to save your configuration, run the **smit install_remove** utility before installing the Version 4.4 software. If your previous version is committed, however, you can perform an upgrade (as described in Chapter 9, Upgrading an HACMP Cluster) to replace it.

The final step in installing the HACMP for AIX software is to reboot each server in your HACMP for AIX environment.

Checking the State of a Previous Software Version

To see if an earlier version of the HACMP/6000 or HACMP for AIX software exists and if your cluster configuration is committed, enter the following command:

```
lslpp -h “cluster*”
```

If the word “COMMIT” appears beside each software image, see Chapter 9, Upgrading an HACMP Cluster, to perform an upgrade.

Removing a Previous Software Version

If you do not want to save your existing HACMP/6000 or HACMP for AIX software or cluster configuration, run the **smit install_remove** utility on cluster nodes and clients before installing the Version 4.4 software.

To remove a previous version of the software and your cluster configuration:

1. Enter the following command:

```
smit install_remove
```

2. Enter field values as follows:

SOFTWARE name	Enter **cluster*** to remove all server and client software (or concurrent access software, if installed), or press F4 for a popup window listing all installed software. Use the arrow keys to locate all software you want to remove; then press F7 to select it. Press Enter after making all selections. Your selections appear in this field.
PREVIEW only?	Enter **no**.
REMOVE dependent software?	Enter **yes**.
DETAILED output?	Enter **no.**

3. Continue with the new software installation.

Use the **smit install_selectable_all** fastpath to install the base High Availability Subsystem on all nodes and clients. Separate procedures for servers and clients are described in the following sections.

Installation Choices

You must install the HACMP for AIX software on each server machine. You can install the software from an installation server, from the installation media, or from a hard disk to which the software has been copied.

At this point, you either want to install the HACMP for AIX software for the first time, or you want to upgrade an earlier version that you do not want to save. Follow the steps in this section to perform a first-time install.

If you are using an earlier version of the HACMP for AIX software and do not want to save your existing cluster configuration, see Removing a Previous Software Version on page 8-3 to remove it before installing Version 4.4. If the earlier version has been committed, do an upgrade as described in Chapter 9, Upgrading an HACMP Cluster.

Installation Server

To install the HACMP for AIX software in a cluster environment, you can create an HACMP for AIX installation server (containing all HACMP for AIX software installable images) on one node and then load the images onto the remaining cluster nodes. Creating an installation server lets you load the HACMP for AIX software onto other nodes faster from the server than from other media types. For instructions on creating an installation server, see the *AIX Installation Guide* or the *AIX Network Installation Management Guide and Reference*.

The organization of cluster images on the base High Availability Subsystem media allows you to make individual or multiple image selections through SMIT when installing the HACMP for AIX software. The installable HACMP for AIX, Version 4.4 images and their associated modules on the base High Availability Subsystem media include the following:

- /usr/sys/inst.images/cluster.base

```
cluster.base.client.lib       HACMP Base Client Libraries
cluster.base.client.rte       HACMP Base Client Runtime
cluster.base.client.utils     HACMP Base Client Utilities
cluster.base.server.diag      HACMP Base Server Diags
cluster.base.server.events    HACMP Base Server Events
cluster.base.server.rte       HACMP Base Server Runtime
cluster.base.server.utils     HACMP Base Server Utilities
```

- /usr/sys/inst.images/cluster.cspoc

```
cluster.cspoc.rte             HACMP CSPOC Runtime commands
cluster.cspoc.cmds            HACMP CSPOC commands
cluster.cspoc.dsh             HACMP CSPOC dsh and perl
```

- /usr/sys/inst.images/cluster.adt

```
cluster.adt.client.demos              HACMP Client Demos
cluster.adt.client.samples.demos      HACMP Client Demos Samples
cluster.adt.client.samples.clinfo     HACMP Client clinfo Samples
cluster.adt.client.samples.clstat     HACMP Client clstat Samples
cluster.adt.client.include            HACMP Client includes
cluster.adt.client.samples.libcl      HACMP Client libcl Samples
cluster.adt.server.samples.images     HACMP Sample Images
cluster.adt.server.demos              HACMP Server Demos
cluster.adt.server.samples.demos      HACMP Server Sample Demos
```

- /usr/sys/inst.images/cluster.man.en_US.data

```
cluster.man.en_US.cspoc.data  HACMP CSPOC Man pages
cluster.man.en_US.client.data HACMP Client Man pages
cluster.man.en_US.server.data HACMP Server Man pages
cluster.man.en_US.haview.data HACMP HAView Man pages
```

- /usr/sys/inst.images/cluster.msg.en_US

```
cluster.msg.en_US.cspoc       HACMP CSPOC Messages
cluster.msg.en_US.client      HACMP Client Messages
cluster.man.en_US.haview.data HACMP HAView Messages
```

- /usr/sys/inst.images/cluster.vsm

```
cluster.vsm                   HACMP X11 Dependent
```

- /usr/sys/inst.images/cluster.haview

```
cluster.haview                HACMP HAView
cluster.haview.client         HACMP HAView Client
cluster.haview.server         HACMP HAView Server
```

- /usr/sys/inst.images/cluster.man.en_US.haview.data

```
cluster.man.en_US.haview.data HACMP HAView Manpages
```

- /usr/sys/inst.images/cluster.msg.en_US.haview

```
cluster.msg.en_US.haview      HACMP HAView Messages
```

- /usr/sys/inst.images/cluster.taskguides

```
cluster.taskguides.shrvolgrp     HAES Shr Vol Grp Task Guides
```

The installable images on the CRM installation media are listed below.

- `/usr/sys/inst.images/cluster.clvm`

`cluster.clvm`	`HACMP for AIX Concurrent Access`

- `/usr/sys/inst.images/cluster.hc`

`cluster.hc.rte`	`Application Heart Beat Daemon`

The installation media requires the following software:

`bos.rte.lvm.usr.4.3.2.0`	`AIX Run-time Executable`

In addition, if you plan to monitor the cluster with Tivoli, install these hativoli filesets:

- `cluster.hativoli.client`
- `cluster.hativoli.server`
- `cluster.msg.en_US.hativoli`

HAView Installation Notes

HAView requires TME 10 NetView for AIX. Install NetView before installing HAView.

The HAView fileset includes a server image and a client image. If NetView is installed using a client/server configuration, the HAView server image should be installed on the NetView server, and the client image on the NetView client. Otherwise, you can install both the HAView client and server images on the NetView server.

Note: It is recommended that you install the HAView components on a node outside the cluster. Installing HAView outside the cluster minimizes the probability of losing monitoring capabilities during a cluster node failure.

For more information on using HAView to monitor a cluster, see the *HACMP for AIX Administration Guide*, Chapter 3: Monitoring an HACMP Cluster.

Tivoli Installation Notes

If you want to have your HACMP cluster monitored by Tivoli Framework, you must install Tivoli Framework software as listed in the section Prerequisites on page 8-1, and also install the **cluster.hativoli** lpp filesets.

In addition, there are a number of installation steps and prerequisites you must be aware of in order to effectively monitor your cluster through Tivoli. See Appendix B, Installing and Configuring Cluster Monitoring with Tivoli, for full instructions.

Copying HACMP Software to Hard Disk

Complete the following steps to copy the HACMP for AIX software to your hard disk.

1. Place the HACMP tape into the tape drive and enter the **smit bffcreate** fastpath to display the Copy Software to Hard Disk for Future Installation screen.
2. Enter the name of the tape drive in the **INPUT device / directory for software** field and press Enter.

   If you are unsure of the input device name, press F4 to list available devices. Select the proper drive and press Enter. That value is entered into the **INPUT device/directory** field as the valid input device.

3. Press Enter to display the next screen.
4. Enter field values as follows:

SOFTWARE name	Enter **cluster*** or **all** to copy all server and client images, or press F4 for a software listing.
DIRECTORY for storing software	Change the value to the storage directory accessed by all nodes using HACMP.

5. Enter values for the other fields as appropriate for your site.
6. When you are satisfied with the entries, press Enter. SMIT responds: `ARE YOU SURE?`
7. Press Enter again.

Once the HACMP software has been copied to your system, install the software by following the instructions given in the following section. After you copy the software, read the HACMP 4.4 **release_notes** file in the **/usr/lpp/cluster/doc** directory.

Installing Base HACMP on Server Node

To install the base High Availability Subsystem on a server node:

1. Insert the installation medium and enter:

   ```
   smit install_selectable_all
   ```

2. Enter the device name of the installation medium or Install Directory in the **INPUT device / directory for software** field and press Enter.

   If you are unsure about the input device name or about Install Directory, press F4 to list available devices. Then select the proper drive or directory and press Enter. The correct value is entered into the **INPUT device/directory** field as the valid input device.

3. Press Enter to display the Install/Update From All Available Software screen.
4. Enter field values as follows:

SOFTWARE to install	Enter **cluster*** or **all** to install all server and client images, or press F4 for a software listing. If you press F4, a popup window appears, listing all installable software. Use the arrow keys to locate all software modules associated with the following Version 4.4 cluster images: **cluster.base, cluster.cspoc, cluster.adt, cluster.man.en_US, cluster.vsm, cluster.haview**, and c**luster.man.en_US.haview.data**, **cluster.taskguides**. The **cluster.base** image (which contains the HACMP for AIX run-time executables) and the **cluster.cspoc** image are required and must be installed on all servers. If you select either **cluster.base** or **cluster.cspoc**, the other will be installed automatically. Next press F7 to select either an image or a module. Then press Enter after making all selections. Your selections appear in this field. Note that selecting **cluster.base** installs the base High Availability Subsystem and all associated messages.

PREVIEW ONLY?	Change the value to **no**.
OVERWRITE same or newer versions?	Leave this field set to **no**. Set it to yes if you are reinstalling or reverting to Version 4.4 from a newer version of the HACMP for AIX software.
AUTOMATICALLY Install requisite software	Set this field to **no** if the prerequisite software for Version 4.4 is installed or if the **OVERWRITE same or newer versions?** field is set to **yes**; otherwise, set this field to **yes** to install required software.

5. Enter values for the other fields as appropriate for your site.
6. When you are satisfied with the entries, press Enter. SMIT responds:

   ```
   ARE YOU SURE?
   ```
7. Press Enter again.

You are then instructed to read the HACMP 4.4 **release_notes** file in the **/usr/lpp/cluster/doc** directory for further instructions.

Installing the Concurrent Resource Manager

To install the concurrent access feature on cluster nodes, complete the procedure in this section.

Note: In HACMP for AIX, Version 4.4 environments, concurrent access is available using only an IBM 7135-110 or 210 Disk Array, an IBM 7137 Disk Array, IBM 2105-B09 and 100 Versatile Storage Servers, an IBM 7133 SSA disk subsystem, or an IBM 9333 disk subsystem. RAID devices from other manufacturers may not support concurrent access.

Use the **smit install_selectable_all** fastpath to load the concurrent access install image on a node. See the section Installation Choices on page 8-3 for a list of software images to install. Depending on the AIX level installed on your system, not all images are required.

To install the concurrent access software on a server:

1. Insert the installation media and enter:

   ```
   smit install_selectable_all
   ```
2. Enter the device name of the installation media or Install Directory in the **INPUT device / directory for software** field and press Enter.

   If you are unsure about the input device name or about Install Directory, press F4 to list available devices. Then select the proper media or directory and press Enter. The correct value is entered into the **INPUT device/directory** field as the valid input device.
3. Press Enter. SMIT refreshes the screen.

4. Enter field values as follows:

   SOFTWARE to install
   Change the value in this field to **cluster.clvm.** Note that the run-time executables for the HACMP for AIX software and associated images are automatically installed when you select this image.

   Note: If using Oracle Parallel Server, you must also install **cluster.hc**

   PREVIEW ONLY?
   Change the value to **no.**

   OVERWRITE same or newer versions?
   Leave this field set to **no**. Set it to yes if you are reinstalling or reverting to Version 4.4 from a newer version of the HACMP for AIX software.

   AUTOMATICALLY Install requisite software
   Set this field to **no** if the prerequisite software for Version 4.4 is installed or if the **OVERWRITE same or newer versions?** field is set to **yes**; otherwise, set this field to **yes** to install required software.

5. Enter values for other fields appropriate for your site.
6. Press Enter when you are satisfied with the entries. SMIT responds:

   ```
   ARE YOU SURE?
   ```

7. Press Enter again.

You are then instructed to read the HACMP 4.4 **release_notes** file in the **/usr/lpp/cluster/doc** directory for further instructions.

Problems During the Installation

If you experience problems during an installation, the installation program automatically performs a cleanup process. If for some reason the cleanup is not performed after an unsuccessful installation, do the following:

1. Enter:

   ```
   smit install
   ```

2. Select **Install and Update Software** from the SMIT Software Installation and Maintenance menu that appears.
3. Select **Clean Up After an Interrupted Installation**.
4. Review the SMIT output (or examine the **/smit.log** file) for the interruption's cause.
5. Fix any problems and repeat the installation process.

Processor ID Licensing Issues

The Concurrent Resource Manager is licensed to the processor ID of a cluster node. Many of the **clvm** or concurrent access commands validate the processor ID against the license file. A mismatch will cause the command to fail, with an error message indicating the lack of a license.

Restoring a system image from a **mksysb** tape created on a different node or replacing the planar board on a node will cause this problem. In such cases, you must recreate the license file by removing and reinstalling the **cluster.clvm** component of the current release from the original installation images.

Rebooting Servers

The final step in installing the HACMP for AIX software is to reboot each server in your HACMP for AIX environment.

Warning: If you are creating (or plan to create) concurrent volume groups on SSA devices, be sure to assign unique non-zero node numbers through the ssar on each cluster node before using the failed drive replacement procedure described in Chapter 5 of the *HACMP for AIX Administration Guide.* If you plan to specify SSA disk fencing in your concurrent resource group, the node numbers are assigned when you synchronize resources. If you do not specify SSA disk fencing, assign node numbers using the following command: `chdev -l ssar -a node_number=x`, where x is the number to assign to that node. You must reboot the system to effect the change.

Chapter 9 Upgrading an HACMP Cluster

This chapter provides instructions for upgrading HACMP cluster software and configuration. It also includes instructions for migrating from HANFS 4.3.1 to HACMP 4.4.

Preparing for an Upgrade

This section identifies the following prerequisite tasks you must perform to prepare for an upgrade to HACMP for AIX, Version 4.4:

- Read the *HACMP for AIX Planning Guide* before starting an upgrade to the Version 4.4 software. It contains the worksheets and diagrams necessary to plan an HACMP for AIX installation and configuration.

  If you have not completed these worksheets and diagrams, return to the appropriate chapters in the *HACMP for AIX Planning Guide* and do so before continuing. You should transfer all information about your existing installation, plus any changes you plan to make after the upgrade, to these new worksheets.
- Ensure that each cluster node has its own HACMP for AIX license.
- It is recommended to perform a full system backup to your preferred media.
- If using SCSI disks, ensure that adapter SCSI IDs for the shared disk are not the same on each node. During an operating system upgrade, disk adapters are reset to the default value of seven. This setting can cause a SCSI ID conflict on the bus that prevents proper access to the shared disk.
- Ensure that the **/usr** filesystem has a minimum of 50 megabytes of free disk space for nodes in a non-concurrent environment and 51 megabytes for nodes in a concurrent access environment. If you are installing only the HACMP for AIX software for clients, a minimum of three megabytes is recommended.
  - If you are not planning to install optional software you can plan for less space: HAView =3MB, TaskGuide= 400KB, VSM= 5.2MB. You should also choose to install the message catalogs for the language you will be using, rather than all message catalogs (Japanese message catalogs use 1.6MB).
- Perform the installation process as the root user.

Before upgrading to HACMP for AIX, Version 4.4:

1. Archive any localized script and configuration files to prevent losing them during an upgrade.
2. Commit your current HACMP for AIX Version 4.* software (if it is applied but not committed) so that the HACMP for AIX 4.4 software can be installed over the existing version. To see if your configuration is already committed, enter:

   ```
   lslpp -h “cluster.*”
   ```

   If the word “COMMIT ” is displayed under the Action header for all the cluster filesets, continue to the next step. If not, run the **smit install_commit** utility before installing the Version 4.4 software.

SMIT displays the next screen. Enter field values as follows:

SOFTWARE name	Set this value to **cluster.***.
COMMIT old version if above version used it?	Leave this field set to **yes**.
EXTEND filesystem if space needed?	Leave this field set to **yes**.

3. Make a **mksysb** backup on each node. This saves a backup of the AIX root volume group.

   Note: If using SCSI disks, and for some reason you do restore a **mksysb** back onto your system, you will need to reset the SCSI IDs on the system.

4. Save the current configuration, using the cluster snapshot utility, and save any customized event scripts in a directory of your own.

   (To review cluster snapshot instructions, see the chapter on saving and restoring cluster configurations in the *HACMP for AIX Administration Guide.*)

Steps for Upgrading an HACMP Cluster

This section summarizes the steps required to upgrade an HACMP cluster from an earlier release of the HACMP for AIX or HANFS software to HACMP Version 4.4. Refer to the corresponding section for details on each step.

Note: When upgrading an HACMP cluster, you should not leave the cluster at mixed versions of HACMP for long periods of time. New functionality supplied with Version 4.4 are available only when all nodes have been upgraded and the cluster has been synchronized. You cannot synchronize a mixed-version cluster.

Complete the following steps to upgrade your cluster:

Step 1: Upgrade the AIX Operating System

In this step, you upgrade the AIX operating system to a minimum level of AIX 4.3.3.

Note: If you are running HACMP for AIX Version 4.2.2 or 4.3 on AIX 4.3.3, or HANFS for AIX, see special upgrade instructions at the end of the section Upgrading Your Existing Software Version.

Step 2: Upgrade Your Existing HACMP Software Version

In this step, you upgrade your existing HACMP or HANFS software version 4.2.2, 4.3, or 4.3.1) to HACMP for AIX, Version 4.4. Earlier software versions are no longer supported and should be removed prior to upgrading to HACMP for AIX 4.4.

If you plan to upgrade from HACMP Versions 4.2.2 through 4.3.1 to Version 4.4, you must perform a migration installation of AIX to upgrade to AIX Version 4.3.3 on all cluster nodes.

The HACMP upgrade process is described in the section Upgrading Existing HACMP Software to Version 4.4 on page 9-4.

The HANFS migration to HACMP process is described in the section Migrating from HANFS to HACMP 4.4 on page 9-9.

Step 3: Verifying Cluster Configuration

In this step, you use the **/usr/sbin/cluster/diag/clverify** utility to verify that the cluster is configured properly. See the *HACMP for AIX Administration Guide* for more information on the **clverify** utility.

Step 4: (optional) Adding to or Changing the Cluster

In this step, you make further changes or additions to the system as defined on the planning worksheets. For example, you may want to include newly added nodes in the resource chain for one or more resource groups. Consult the appropriate chapters in the *HACMP for AIX Administration Guide* for the changes you can make.

You can also modify Version 4.2.2, 4.3, or 4.3.1 cluster snapshot files so they can be applied to a Version 4.4 cluster. See the information on using the cluster snapshot utility in the *HACMP for AIX Administration Guide* for instructions.

Step 5: Rebooting Cluster Nodes and Clients

The upgrade is complete when you have performed the steps identified above. You must then reboot the system to make the cluster topology active.

Step 6: Upgrading Applications Using Clinfo API

Check the *HACMP for AIX Programming Client Applications* guide for updated information on the Clinfo C and C++ API routines. Make any necessary changes, then recompile and link your applications.

Step 7: Upgrading Recompiling and Linking Applications Using Locking API

Check the *HACMP for AIX Programming Locking Applications* guide for updated information on the Cluster Lock Manager (CLM) API routines. Make any necessary changes, then recompile and link your applications.

Upgrading the AIX Operating System

Before attempting to install and upgrade the HACMP for AIX, Version 4.4 software in your cluster environment, you must first upgrade the AIX operating system to a minimum level of AIX 4.3.3. See your *AIX Installation Guide* for instructions on upgrading the AIX operating system in your cluster, if necessary.

Upgrading Existing HACMP Software to Version 4.4

After upgrading the AIX operating system, install the HACMP for AIX, Version 4.4 software on all cluster nodes and clients. Keep in mind, however, that to replace your current version's ODM object classes with the Version 4.4 ODM object classes, you must perform an upgrade when installing either the Version 4.4 base High Availability Subsystem or the Concurrent Resource Manager (CRM) software. See Chapter 8, Installing HACMP for AIX Software for information about first-time installation of the HACMP for AIX software.

If your site is currently running an earlier version of the HACMP for AIX software in its cluster environment, *except for Version 4.2.2, 4.3 or 4.3.1 already running on AIX 4.3.3,* the procedures in this section describe how to upgrade your existing HACMP software to HACMP for AIX, Version 4.4. If you are already running AIX 4.3.3, see the special section at the end of this section.

If you are migrating from HANFS, see the section Migrating from HANFS to HACMP 4.4 on page 9-9.

Note: Although your objective in performing a migration installation is to keep the cluster operational and to preserve essential configuration information, do not run your cluster with mixed versions of the HACMP for AIX software for an extended period of time.

Supported Upgrades of HACMP

If you wish to convert to HACMP version 4.4 from versions earlier than those listed below, you must first do an installation upgrade to one of the following supported versions. Since a conversion from the versions below to version 4.4 are supported upgrades, you will then be able to convert to HACMP 4.4. For example, to convert from HACMP 4.2.1 to HACMP 4.4 you must first do an installation upgrade to HACMP 4.2.2. (Refer to the *HACMP for AIX Guide*, Version 4.2.2 for information on this specific upgrade.) You will then be able to migrate to HACMP 4.4

HACMP conversion utilities provide easy conversion between the HACMP versions and products listed below.

- HACMP 4.2.2 to HACMP 4.4
- HACMP 4.3.1 to HACMP 4.4
- HACMP ES 4.2.2 to HACMP ES 4.4
- HACMP ES 4.3.1 to HACMP ES 4.4
- HACMP 4.4 to HACMP ES 4.4
- HANFS 4.3.1 to HACMP 4.4

cl_convert and clconvert_snapshot

The HACMP conversion utilities are **cl_convert** and **clconvert_snapshot**.

Upgrading HACMP software to the newest version involves converting the ODM from a previous release to that of the current release. When you install HACMP, **cl_convert** is run automatically. However, if installation fails, you must run **cl_convert** from the command line. The **clconvert_snapshot** is not run automatically during installation, and must always be run from the command line.

Note: Root user privilege is required to run a conversion utility. You must know the HACMP version from which you are converting in order to run these utilities.

The **cl_convert** utility logs conversion progress to the **/tmp/clconvert.log** file so that you can gauge conversion success. This log file is regenerated each time **cl_convert** or **clconvert_snapshot** is executed.

Run **clconvert_snapshot** to upgrade cluster snapshots.

For more information on **cl_convert** and **clconvert_snapshot**, refer to the respective man pages, or to the *HACMP for AIX Administration Guide* Appendix K, HACMP for AIX Commands.

Upgrading from Version 4.2.2 through 4.3.1 to Version 4.4

The following procedure applies to upgrading a two-node or multi-node cluster running version 4.2.2 through 4.3.1 to Version 4.4 when the installed AIX version is less than 4.3.3.

To perform a rolling AIX migration installation and HACMP upgrade from Version 4.2.2 through Version 4.3.1 to Version 4.4, complete the following steps:

Upgrade AIX on One Node

1. If you wish to save your cluster configuration, see the chapter Saving and Restoring Cluster Configurations in the *HACMP for AIX Administration Guide.*
2. Shut down the first node (gracefully with takeover) using the smit **clstop** fastpath. For this example, shut down Node A. Node B will takeover Node A’s resources and make them available to clients.

   See the chapter Starting and Stopping Cluster Services in the *HACMP for AIX Administration Guide* for more information about stopping cluster services.
3. Turn the maintenance key on Node A to the Service position and boot the node from the AIX 4.3.3 installation media. In this position, the node will boot from tape or from CD-ROM if the media contain the proper boot image.
4. Select the **Migration Installation** option from the Change Method of Installation screen after the system boots to upgrade the AIX version on the node to 4.3.3. Complete the AIX installation as described in your *AIX Installation Guide*.

   The Migration Installation option preserves the current version of the HACMP for AIX software and upgrades the existing base operating system to AIX 4.3.3. Product (application) files and configuration data also are saved.
5. Return the maintenance key to the Normal position. AIX reboots the node automatically.

Perform AIX Installation Verification Checks

6. Check that all external disks are available on Node A and that the **lspv** command shows PVIDs for each disk, and lists the volume groups.
   - If PVIDs and VGs are *not* displayed for the disks, you may need to remove the disks and reconfigure them. *If this reconfiguration is necessary, the cluster will be down for a short period while you do this task. This is usually not necessary.*
     - Import all desired volume groups from other nodes onto Node A using the **smit importvg** fastpath.

   Note: Be sure to set the volume groups to *not* autovaryon at start using the **smit chvg** command, and set ownership of all logical volumes using the **chown** command. Also, see the section Major Numbers on Shared Volume Group on page 7-5 to determine the free major numbers available on a node.

   See the chapter Maintaining Shared LVM Components of the *HACMP for AIX Administration Guide* for more information on importing and exporting volume groups and changing major numbers.

7. If using SCSI disks, check that the SCSI ID of the shared disk adapter is unique and is not equal to 7. A SCSI ID conflict can occur if SCSI ID 7 is in use by the shared adapter when the HACMP cluster is restarted.

Install HACMP 4.4

8. After verifying that the disks are correctly configured after upgrading AIX, Install the HACMP for AIX 4.4 software on Node A. Insert the installation medium and enter the fastpath:

   ```
   smit install_selectable_all
   ```

   (Or through SMIT, go to the **Install and Update from All Available Software** screen.)

9. Enter the device name of the installation medium or Install Directory in the **INPUT device / directory for software** field and press Enter.

   If you are unsure about the input device name or about Install Directory, press F4 to list available devices. Then select the proper drive or directory and press Enter. The correct value is entered into the **INPUT device/directory** field as the valid input device.

10. Press Enter to display the **Install/Update From All Available Software** screen. It is recommended to use F4 to list the software. This way you can install either the English or the Japanese message catalogs, and you can omit optional software if so desired.

11. Enter field values as follows:

    SOFTWARE to install Enter **cluster*** or **all** to install all server and client images, or press F4 for a software listing. If you press F4, a popup window appears, listing all installable software. Use the arrow keys to locate all software modules associated with the following Version 4.4 cluster images: **cluster.adt, cluster.base, (cluster.msg<*LANG*> cluster.man.<*LANG*>), cluster.cspoc (cluster.msg<*LANG*>.cspoc** and **cluster.cpsoc.man.<*LANG*>), cluster.haview**, **cluster.man.<*LANG*>.haview.data**, **cluster.taskguides, cluster.vsm**.

    The **cluster.base** (which contains the HACMP for AIX run-time executables) and the **cluster.cspoc** images are required and must be installed on all servers. If you select either **cluster.base** or **cluster.cspoc**, the other will be installed automatically (along with their message catalogs and man pages, if *LANG* is set to something other than C).

    Next press F7 to select either an image or a module. Then press Enter after making all selections. Your selections appear in this field. Note that selecting **cluster.base** installs the base High Availability Subsystem and all associated messages.

    PREVIEW ONLY? It is recommended to change the value to **yes** to ensure that all required prerequisite software is installed. Once you have made sure the install will pass the prereq check, come back to this screen and enter **no** to start installation.

    OVERWRITE same or newer versions? Leave this field set to **no**. Set it to yes if you are reinstalling or reverting to Version 4.4 from a newer version of the HACMP for AIX software.

    AUTOMATICALLY Install requisite software Set this field to **no** if the prerequisite software for Version 4.4 is installed or if the **OVERWRITE same or newer versions?** field is set to **yes**; otherwise, set this field to **yes** to install required software.

12. Enter values for the other fields as appropriate for your site.

13. When you are satisfied with the entries, press Enter. SMIT responds:

    ```
    ARE YOU SURE?
    ```
    Press Enter again.

14. The installation process automatically runs the **cl_convert** program. It removes the current HACMP objects from **/etc/objrepos** and saves them to HACMP<*ODM.NAME*>.OLD It creates new HACMP ODM object classes for Version 4.4 in **/etc/objrepos**.

15. Reboot the node.

16. Start the Version 4.4 software on Node A using the **smit clstart** fastpath. Check to ensure that the node successfully joins the cluster.

> **Warning:** If the node running Version 4.4 fails while the cluster is in this state, the surviving nodes running the previous version may not successfully mount the filesystems that were not properly unmounted due to Node A's failure.

17. Repeat Steps 2 through 16 on remaining cluster nodes, one at a time.

    > **Warning:** In a multi-node cluster, do not synchronize the node configuration or the cluster topology until the last node has been upgraded.

    When the last node has been upgraded to both AIX 4.3.3 and HACMP 4.4, the cluster install/upgrade process is complete.

Check Upgraded Configuration

18. If using tty devices, check that the **tty** device is configured as a serial network using the **smit chgtty** fastpath.
19. In order to verify and synchronize the configuration (if desired), you must have **/.rhosts** files on cluster nodes. If it does not exist, create the **/.rhosts** file on Node A using the following command:

    ```
    /usr/sbin/cluster/utilities/cllsif -x >> /.rhosts
    ```

    This command will append information to the **/.rhosts** file instead of overwriting it.
20. Verify the cluster topology on all nodes using the **clverify** utility.
21. Check that custom event scripts are properly installed.
22. (optional) Synchronize the node configuration and the cluster topology from Node A to all nodes. To save time, you can skip cluster verification during topology synchronization.
23. Go to the section Making Additional Changes to the Cluster on page 9-11.
24. It is recommended that you test the upgraded cluster to ensure proper behavior.

Client-only Migration

If you are migrating from an HACMP for AIX, Version 4.2.2 through 4.3.1 server node to a client-only node running Version 4.4, first remove the existing server portion of HACMP. If, after upgrading AIX, you install the cluster.base.client.* filesets on a node running an earlier version of HACMP for AIX without de-installing the server, the results are unpredictable.

To determine if there is a mismatch between the HACMP client and server software installed on a node,issue the following command to list the installed software:

```
lslpp -L “cluster*”
```

Examine the list and make sure that all cluster filesets are at 4.4.

If you determine that there is a mismatch between the client and server, de-install the server and then repeat the installation of the client software.

Upgrading from Version 4.2.2, 4.3, or 4.3.1 on AIX 4.3.3 to HACMP Version 4.4

If your cluster is currently running Version 4.2.2, 4.3, or 4.3.1 of the HACMP for AIX software on AIX Version 4.3.3, use the following procedure to upgrade to HACMP for AIX Version 4.4.

1. If you wish to save your cluster configuration, see the chapter Saving and Restoring Cluster Configurations in the *HACMP for AIX Administration Guide.*
2. Commit your current HACMP for AIX software on all nodes.
3. Shut down one node (gracefully with takeover) using the smit **clstop** fastpath. For this example, shut down Node A. Node B will takeover Node A's resources and make them available to clients.

   See the chapter Starting and Stopping Cluster Services in the *HACMP for AIX Administration Guide* for more information on stopping cluster services.
4. Install HACMP for AIX version 4.4. See Chapter 8, Installing HACMP for AIX Software, starting with the section Installation Choices, for instructions.

   The **cl_convert** utility automatically updates the HACMP ODM object classes to the 4.4 version.

   Note: If IP address swapping is being used on this node, that is, a boot address is defined for this node, check to ensure that the HACMP changes to **/etc/inittab** and **/etc/rc.net** exist as specified in Appendix A before rebooting the node.
5. Reboot Node A.
6. Start the HACMP for AIX software on Node A using the **smit clstart** fastpath and verify that Node A successfully joins the cluster.
7. Repeat Steps 2 through 5 on remaining cluster nodes, one at a time.
8. After all nodes have been upgraded to HACMP for AIX 4.4, synchronize the node configuration and the cluster topology from Node A to all nodes.
9. Verify the cluster topology on all nodes using the **clverify** utility.
10. Go to the section Making Additional Changes to the Cluster below.
11. Complete a test phase on the cluster before putting it into production.

Migrating from HANFS to HACMP 4.4

HACMP 4.4 provides HANFS users a node-by-node migration path from HANFS 4.3.1 to HACMP 4.4. HACMP 4.4 now supports the NFS export behavior of the HANFS cluster.

You can perform a migration from a running HANFS 4.3.1 cluster to a running HACMP 4.4 cluster *without bringing the cluster offline*, thereby keeping all cluster resources running during the migration process.

Prerequisites

In order to perform HANFS 4.3.1 to HACMP 4.4 node-by-node migration:

- Both nodes in the cluster must have HANFS version 4.3.1 installed and committed. (If you are running an older versions of HANFS, you must upgrade to 4.3.1 before migrating to HACMP 4.4).
- Both nodes in the cluster must be up and running the HANFS software.
- The cluster must be in a stable state.

Note: As in any migration, do not attempt to make any changes to the cluster topology or configuration once you have started the migration process. You can make any necessary changes after both nodes have finished the migration process.

Procedure for HANFS to HACMP 4.4 Node-by-Node Migration

Prerequisite steps:

1. If your version of HANFS is less than 4.3.1, upgrade both nodes to HANFS 4.3.1.
2. Upgrade AIX to Version 4.3.3 on both nodes.

Take the following steps to perform a node-by-node migration from HANFS 4.3.1 to HACMP 4.4:

1. Stop cluster services on one of the nodes running HANFS 4.3.1 using the “graceful with takeover” method.
2. Install HACMP 4.4 on the node. See Chapter 8, Installing HACMP for AIX Software, for instructions.

   Note: The HACMP installation utility first checks the current version of HANFS. If your HANFS software is an earlier version than 4.3.1, you will see an error message and the installation will be aborted. If you are running version 4.3.1, you will see a message indicating that a migration has been requested and is about to proceed.

3. After you install the HACMP software, reboot the node.
4. Using the SMIT **Start Cluster Services** screen, start the HACMP software.

   Two events take place when you start HACMP:

   - The HACMP Cluster Manager communicates with the HANFS Cluster Manager on the other node.
   - The HACMP software reacquires the resources assigned to the node.

5. Repeat steps 1 through 4 for the other cluster node.

Backout Procedure

If the migration process fails (a node crash, for example):

- If the first node fails, you must deinstall all HACMP and/or HANFS software, reinstall the HANFS 4.3.1 software, and resynchronize the cluster from the other HANFS cluster node.
- If the second node fails, you must deinstall all HACMP and/or HANFS software, reinstall the HACMP 4.4 software, and resynchronize the cluster from the other (already migrated) HACMP 4.4 cluster node.

Making Additional Changes to the Cluster

Make any further changes or additions to the system as planned for in the worksheets. For example, you may want to include newly added nodes in the resource chain for one or more resource groups. Consult the appropriate chapters in the *HACMP for AIX Administration Guide* for those changes.

Verifying the Cluster Configuration

When you have finished configuring your system, run the **/usr/sbin/cluster/diag/clverify** utility to ensure the cluster is configured properly. The chapter Verifying the Cluster Topology in the *HACMP for AIX Administration Guide* describes this step.

Rebooting Servers

The final step in upgrading the HACMP for AIX software is to reboot each server in your cluster environment.

Upgrading Clinfo Applications to HACMP 4.4 for AIX

Check the *HACMP for AIX Programming Client Applications* guide for updated information on the Clinfo C and C++ API routines. Make any changes necessary to your applications; then recompile and link the applications using the Clinfo library.

Note: If you do not want to change your applications, simply link them.

Upgrading Locking Applications to HACMP 4.4 for AIX

Check the *HACMP for AIX Programming Client Applications Guide* for updated information on the Cluster Lock Manager API routines. Make any changes necessary to your applications; then recompile and link the applications using the Cluster Lock Manager library.

Note: If you do not want to change your applications, simply link them.

Chapter 10 Verifying Cluster Software

This chapter describes how to verify that all HACMP-specific modifications to AIX software are correct.

Prerequisites

- Complete all the hardware and software installation and configuration tasks described earlier in this guide.
- Install the HACMP for AIX software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software.

Using the /usr/sbin/cluster/diag/clverify Utility

Use the **/usr/sbin/cluster/diag/clverify** utility on each cluster node to verify that the correct HACMP-specific modifications to AIX system files exist.

You should run this utility before starting up HACMP for AIX.

You can use the **clverify** utility in either of two modes: interactive or directly from the command line. Using the interactive mode, you step through the list of valid options until you get to the specific program you want to run. The interactive mode also includes a help facility. If you know the complete syntax the utility needs for verifying a given feature, you can enter the command and its required options at the system command-line prompt to run the program directly.

The following sections show how to use the utility in interactive mode. See the chapter on verifying a cluster configuration in the *HACMP for AIX Administration Guide* for a complete description of this utility.

Verifying Cluster Software

To ensure that an HACMP cluster works properly, you must verify that all the HACMP for AIX-specific modifications to AIX system files are correct. The cluster software verification procedure automates this task for you.

To run the software verification procedure interactively:

1. Type:

   ```
   clverify
   ```

   The command returns a list of command options and the `clverify` prompt.

```
----------------------------------------------------------
To get help on a specific option, type: help <option>
To return to previous menu, type: back
To quit the program, type: quit
----------------------------------------------------------
Valid Options are:
software
```

```
cluster


clverify>
```

2. Type:

   ```
   software
   ```

   The following option and an updated prompt appear:

   ```
   Valid Options are:
   lpp


   clverify.software>
   ```

3. To verify that the HACMP-specific modifications to AIX system files are correct, and to log the results in a file called **verify_hacmp**, use the -R flag, as follows.Type:

   ```
   lpp -R verify_hacmp
   ```

   When the program completes, read the **verify_hacmp** file. If no problems exist, no messages are logged.

   Note: If you receive messages about configuration errors but have not yet configured the cluster, ignore them.

4. Type CTRL-C or **quit** to return to the system prompt.

For a complete description of the clverify utility, see the chapter on verifying a cluster configuration in the *HACMP for AIX Administration Guide*.

Chapter 11 Defining the Cluster Topology

This chapter describes how to define the HACMP for AIX cluster topology.

Prerequisites

- Complete the worksheets discussed in the *HACMP for AIX Planning Guide.*
- Install and configure the hardware and software as described earlier in Part 3 of this guide.
- Install the HACMP for AIX software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software.
- Verify the cluster software following the instructions in Chapter 10, Verifying Cluster Software.

Overview

The cluster topology comprises the following components:

- The cluster definition
- The cluster nodes
- The network adapters
- The network modules.

You define the cluster topology by entering information about each component in HACMP-specific ODM classes. You enter the HACMP for AIX ODM data by using the HACMP for AIX SMIT interface or the VSM utility, **xhacmpm**. The **xhacmpm** utility is an X Windows tool for creating cluster configurations using icons to represent cluster components. For more information about the **xhacmpm** utility, see the *HACMP for AIX Administration* guide, Appendix D.

Note: The SP Switch network module can support multiple clusters; therefore, its settings should remain at their default values to avoid affecting HACMP for AIX event scripts. If you must change these settings, see the chapter on changing the cluster topology in the *HACMP for AIX Administration Guide* for more information.

Making the Cluster Topology Active

The cluster topology becomes active the next time the Cluster Manager starts. At startup, the Cluster Manager initializes the topology with the values defined in the HACMP for AIX ODM. See the Starting and Stopping Cluster Services chapter in the *HACMP for AIX Administration Guide* for more information on starting and stopping cluster services.

Changing the Cluster Topology

This chapter describes how to define the initial cluster configuration. If you later want to change the cluster definition—for example, define a new network—follow the steps in the chapter Changing the Cluster Topology in the *HACMP for AIX Administration Guide.*

Defining the Cluster

The cluster ID and name identifies a cluster in an HACMP for AIX environment. The cluster ID and name must be unique for each cluster defined. Complete the following steps to define the cluster ID and name. Refer to your completed network planning worksheets for the values.

To define a cluster ID and name:

1. Enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration** > **Cluster Topology > Configure Cluster > Add a Cluster Definition** and press Enter to display the Add a Cluster Definition screen.
3. Enter field values as follows:

   Cluster ID Enter a positive integer in the range 1 to 99999. The specified value must be unique to your site.

   Cluster Name Enter an ASCII text string that uniquely identifies the cluster. The cluster name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. It can be different from the hostname.

4. Press Enter.

   The HACMP for AIX software uses this information to create the cluster entries for the ODM.

5. Press F3 until you return to the Cluster Topology screen, or F10 to exit SMIT.

Defining Nodes

After defining the cluster name and ID, define the cluster nodes.

To define the cluster nodes:

1. Select **Configure Nodes** from the Cluster Topology menu and press Enter.
2. Select **Add Cluster Nodes** and press Enter to display the following screen.
3. Enter the name for each cluster node in the **Node Names** field.

   Enter an ASCII text string that identifies the node. The node name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. Leave a space between names. If you specify a duplicate name, the operation fails. Press Enter to define the nodes. Node names can be different from the hostnames.

4. Press F3 until you return to the Cluster Topology screen, or F10 to exit SMIT.

Adding or Changing a Node Name after the Initial Configuration

If you want to add or change a node name after the initial configuration, use the Change/Show Cluster Node Name screen. See the chapter on changing the cluster topology of the *HACMP for AIX Administration Guide* for more information.

Defining Adapters

To define the adapters after defining the node names, first consult your planning worksheets for both TCP/IP and serial networks listed. Now complete the following steps:

1. Select **Configure Adapters** from the Cluster Topology menu and press Enter.
2. Select **Add an Adapter** and press Enter to display the following screen.
3. Enter field values as follows:

Adapter IP Label Enter the IP label (the name) of the adapter you have chosen as the service address for this adapter. Adapter labels can be any ASCII text string consisting of alphabetical and numeric characters, underscores, and hyphens. Adapter labels typically are less than eight characters long (31 characters is the limit).

If the cluster uses IP address takeover or rotating resources, each adapter that can have its IP address taken over must have a boot adapter (address) label defined for it. Use a consistent naming convention for boot adapter labels. (You will choose the Add an Adapter option again to define the boot adapter when you finish defining the service adapter.) You can use hyphens in adapter labels; however, the **/usr/sbin/cluster/diag/clverify** utility flags adapter labels that contain hyphens each time it runs.

Network Type Indicate the type of network to which this adapter is connected. Pre-installed network modules are listed on the pop-up pick list. See Configuring Network Modules on page 11-5 for more information on network modules. For an SP switch, the network type must be *hps*.

Network Name Enter an ASCII text string that identifies the network. The network name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. The network name is arbitrary, but must be used consistently.

If several adapters share the same physical network, make sure you use the same network name for each of these adapters.

Network Attribute Indicate whether the network is **public**, **private**, or **serial**. Press TAB to toggle the values. Ethernet, Token-Ring, FDDI, and SLIP are public networks. SOCC, ATM, and an SP Switch are private networks. RS232 lines, target mode SSA loops, and target mode SCSI-2 busses are serial networks.

Adapter Function

Indicate whether the adapter’s function is **service**, **standby**, or **boot**. Press TAB to toggle the values. A node has a single service adapter for each public or private network. A serial network has a single service adapter.

A node can have none, one, or more standby adapters for each public network. Serial and private networks do not have standby adapters, with the exception of ATM networks. ATM networks must be defined as private, and therefore standby adapters are supported.

In an HACMP for AIX environment on the RS/6000 SP, the ethernet adapters can be configured as service adapters but *should not* be configured for IP address takeover. Regarding the SP Switch, network, boot, and service addresses used for IP address takeover are ifconfig alias addresses used on the css0 network. See Configuring the SP Switch Network on page E-10 for more information on adapter functions in an SP Switch environment.

Keep in mind that the netmask for all adapters in an HACMP network must be the same to avoid communication problems between standby adapters after an adapter swap and after the adapter is reconfigured with its original standby address.

Adapter Identifier

Enter the IP address in dotted decimal format or a device file name. IP address information is required for non-serial network adapters only if the node’s address cannot be obtained from the domain name server or the local **/etc/hosts** file (using the adapter IP label given).

You must enter device file names for serial network adapters. RS232 serial adapters must have the device file name **/dev/tty*n***. Target mode SCSI serial adapters must have the device file name **/dev/tmscsi*n***.Target mode SSA adapters must have the device file name **/dev/tmssa*n.im or* /dev/tmssa*n.tm***

Adapter Hardware Address

(optional) Enter a hardware address for the adapter. The hardware address must be unique within the physical network. Enter a value in this field *only* if: You are currently defining a service adapter, *and* the adapter has a boot address, *and* you want to use hardware address swapping. See the chapter on planning TCP/IP networks of the *HACMP for AIX Planning Guide* for more information on hardware address swapping. This facility is supported for Ethernet, Token-Ring, FDDI, and ATM adapters. It does not work with the SP Switch. The hardware address is 12 digits for Ethernet, Token-Ring and FDDI; and 14 digits for ATM.

Node Name Define a node name for all adapters except for those whose addresses may be shared by nodes participating in the resource chain for a rotating resource configuration. These adapters are rotating resources. The event scripts use the user-defined configuration to associate these service addresses with the proper node. In all other cases, addresses are associated with a particular node (service, boot, and standby).

4. Press Enter. The system adds these values to the HACMP for AIX ODM and displays the Configure Adapters menu.
5. Define all the adapters, then press F3 until you return to the Cluster Topology screen, or F10 to exit SMIT.

Note: Although it is possible to have only one physical network adapter (no standby adapters), this constitutes a potential single point of failure condition and is not recommended for an HACMP for AIX configuration. The instructions listed here assume you have at least one standby adapter for each public network.

Adding or Changing Adapters after the Initial Configuration

If you want to change the information about an adapter after the initial configuration, use the Change/Show an Adapter screen. See the chapter on changing the cluster topology of the *HACMP for AIX Administration Guide* for more information.

Configuring Network Modules

Each supported cluster network in a configured HACMP cluster has a corresponding cluster network module. Each network module monitors all I/O to its cluster network.

Note: The Network Modules are pre-loaded when you install the HACMP for AIX software. You do not need to enter information in the Network Module SMIT screens unless you want to change some field associated with a network module, such as the failure detection rate.

Each network module maintains a connection to other network modules in the cluster. The Cluster Managers on cluster nodes send messages to each other through these connections. Each network module is responsible for maintaining a working set of service adapters and for verifying connectivity to cluster peers. The network module also is responsible for reporting when a given link actually fails. It does this by sending and receiving periodic heartbeat messages to or from other network modules in the cluster.

Currently, network modules support communication over the following types of networks:

- Serial (RS232)
- Target-mode SCSI
- Target-mode SSA
- IP

- Ethernet
- Token-Ring
- FDDI
- SOCC
- SLIP
- SP Switch
- ATM.

It is highly unlikely that you will add or remove a network module. For information about changing a characteristic of a Network Module, such as the failure detection rate, see the chapter on changing the cluster topology of the *HACMP for AIX Administration Guide*.

Synchronizing the Cluster Definition Across Nodes

Synchronization of the cluster topology ensures that the ODM data on all cluster nodes is the same. The HACMP for AIX ODM entries must be the same on each node in the cluster. If the definitions are not synchronized across nodes, the HACMP for AIX software generates a run-time error at cluster startup.

Note: Even if you have a cluster defined with only one node, you must still synchronize the cluster.

The processing performed in synchronization varies depending on whether the cluster manager is active on the local node. If the cluster manager is not active on the local node when you select this option, the ODM data in the system default configuration directory (DCD) on the local node is copied to the ODMs stored in the DCDs on all cluster nodes. The cluster manager is typically not running when you synchronize the initial cluster configuration.

If the cluster manager is active on the local node, the ODM data stored in the DCDs on all cluster nodes are synchronized. In addition, the configuration data stored in the ACD on each cluster node is overwritten by the new configuration data, which becomes the active configuration. If the cluster manager is active on some cluster nodes but not on the local node, the synchronization operation is aborted.

Note: Before attempting to synchronize a cluster configuration, ensure that all nodes are powered on, that the HACMP for AIX software is installed, and that the **/etc/hosts** and **/.rhosts** files on all nodes include all HACMP for AIX boot and service IP labels.

Note: The **/.rhosts** file is not required on SP systems running HACMP Enhanced Security. This feature removes the requirement of TCP/IP access control lists (for example, the **/.rhosts** file) on remote nodes during HACMP configuration.

Complete the following steps to synchronize a cluster definition across nodes:

1. Enter:

   ```
   smit hacmp
   ```

2. From the **Cluster Configuration** menu, select **Cluster Topology** > **Synchronize Cluster Topology** and press Enter.

   SMIT displays the Synchronize Cluster Topology screen.

3. Enter field data as follows:

   Skip Cluster Verification — By default, this field is set to **no** and the cluster topology verification program is run. To save time in the cluster synchronization process, you can toggle this entry field to **yes**. By doing so cluster verification will be skipped.

   Ignore Cluster Verification Errors — By choosing **yes**, the result of the cluster verification is ignored and the configuration is synchronized even if verification fails.

   By choosing **no**, the synchronization process terminates; view the error messages in the system error log to determine the configuration problem.

   Emulate or Actual — If you set this field to **Emulate**, the synchronization is an emulation and does not affect the Cluster Manager. If you set this field to **Actual**, the synchronization actually occurs, and any subsequent changes affect the Cluster Manager. **Actual** is the default value.

4. After you specify values and press Enter, SMIT displays a screen asking if you want to continue with the synchronization. If you want to proceed, press Enter. The cluster topology definition (including all node, adapter, and network module information) is copied to the other nodes in the cluster.

Chapter 12 Configuring Cluster Resources

This chapter describes how to register server applications with the HACMP for AIX software to make them highly available, how AIX Fast Connect, AIX Connections, and CS/AIX work under HACMP, how to configure cluster resources in a resource group, and how to define run-time parameters for cluster nodes.

Prerequisites

- Complete the worksheets discussed in the *HACMP for AIX Planning Guide.*
- Complete the hardware and software installation and configuration tasks as described earlier in Part 3 of this guide.
- Install the HACMP for AIX software on each cluster node following the instructions in Chapter 8, Installing HACMP for AIX Software.
- Define the cluster topology following the instructions in Chapter 11, Defining the Cluster Topology.
- Write the application server start and stop scripts.

  Note that this chapter does not discuss writing application-specific start and stop scripts. See your vendor documentation for information on starting and stopping a particular application. Also see the Applications and HACMP appendix in the *HACMP for AIX: Planning Guide*.
- Register the application servers following the steps in the section Configuring Application Servers on page 11-1.

Configuring Application Servers

An *application server* is a cluster resource made highly available by the HACMP software, such as the HACMP for AIX Image Cataloger demo that runs on a cluster node. Client applications query the application server, which in turn accesses a database on a shared external disk, and then responds to the client's request.

When you configure an application server in an HACMP cluster, you:

- Provide (associate) a meaningful name for the server application. For example, you could give the HACMP for AIX Image Cataloger demo a name such as *imagedemo.* You then use this name to refer to the application server when you define the cluster nodes (which is the next step in the installation process).
- Add the location of the application server's start and stop scripts to the HACMP for AIX ODM.

The steps in this section describe how to configure an application server. Configuring an application server registers it with the HACMP for AIX software. After completing these steps on a single node, the HACMP for AIX software copies the information to all cluster nodes when you synchronize the nodes, described in the section on page 1 2 -15.

To configure application servers:

1. Enter the following to start HACMP for AIX system management:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration > Cluster Resources > Define Application Servers > Add an Application Server**.
3. Enter field values as follows. Refer to the Application Server Worksheet you filled out earlier.

   Server Name Enter an ASCII text string that identifies the server (for example, *imagedemo*). You use this name to refer to the application server when you define it as a resource during node configuration. The server name can include alphabetic and numeric characters and underscores. Use no more than 31 characters.

   Start Script Enter the full pathname of the script that starts the server (for example, **/usr/sbin/cluster/events/utils/start_imagedemo**). This script is called by the cluster event scripts. This script must be in the same location on each cluster node that might start the server.

   Stop Script Enter the full pathname of the script that stops the server (for example, **/usr/sbin/cluster/events/utils/stop_imagedemo**). This script is called by the cluster event scripts. This script must be in the same location on each cluster node that may stop the server.

4. Press Enter to add this information to the HACMP for AIX ODM.
5. Press F10 after the command completes to leave SMIT and return to the command line.

Changing Application Servers

To change an application server or its start and stop scripts after the initial configuration, see the chapter on Changing Resources and Resource Groups of the *HACMP for AIX Administration Guide* for more information.

Configuring AIX Fast Connect Resources

AIX Fast Connect allows client PCs running Windows, DOS, and OS/2 operating systems to request files and print services from an AIX server. Fast Connect supports the transport protocol NetBIOS over TCP/IP. You can configure AIX Fast Connect resources using the SMIT interface.

The AIX Fast Connect application is integrated with HACMP already so you can configure it, via the SMIT interface, as highly available resources in resource groups. This is one of the applications that does not need to be associated with an application server or special scripts. This section contains information you may need before you are ready to add AIX Fast Connect resources in the SMIT screen.

Be sure to read the sections in the *Planning Guide* that cover planning for Fast Connect and other aspects of resource and resource group configuration.

Prerequisites

Before you can configure Fast Connect resources in HACMP, make sure:

- (*If you are converting from AIX Connections*) You have unconfigured any AIX Connections services and deinstalled the AIX Connections software. For more information, see *Converting from AIX Connections to AIX Fast Connect* on page 12-10.
- You have installed the Fast Connect server on all nodes in the cluster.
- For cascading and rotating resource groups, you have assigned the *same* netBIOS names to each node when configuring the Fast Connect server. This action will minimize the steps needed for the client to connect to the server after fallover.
- For concurrently configured resource groups, you have assigned *different* netBIOS names across nodes.
- AIX print queue names match for all nodes in the cluster if Fast Connect printshares are to be highly available.

Configuration Notes for Fast Connect

When configuring Fast Connect as a cluster resource in HACMP, remember that:

- You cannot configure both Fast Connect and AIX Connections in the same resource group or on the same node. You should de-install AIX Connections software before configuring Fast Connect as a resource.
- When starting cluster services, the Fast Connect server must be stopped on all nodes, so that HACMP can take over the starting and stopping of Fast Connect resources properly.
- You must define any filesystems associated with Fast Connect fileshares when you configure the resource group in SMIT.
- In concurrent configurations, you should define a second, non-concurrent, resource group to control any filesystem that must be available for the Fast Connect nodes. Having a second resource group configured in a concurrent cluster keeps the AIX filesystems used by Fast Connect cross-mountable and highly available in the event of a node failure.
- Fast Connect cannot be configured in a mutual takeover configuration. Make sure there are no nodes participating in more than one Fast Connect resource groups at the same time.

For instructions on using SMIT to configure Fast Connect services as resources, see the section Configuring Resource Groups on page 12-9.

Verification of Fast Connect

After completing your resource configuration, you synchronize cluster resources. During this process, if Fast Connect resources are configured in HACMP, the **clverify** utility verifies:

- That the Fast Connect server application exists on all participating nodes in a resource group.
- That the Fast Connect fileshares are in filesystems that have been defined as resources on all nodes in the resource group.
- That Fast Connect resources are not configured in a mutual takeover form; that is, there are no nodes participating in more than one Fast Connect resource group.
- That AIX Connections and Fast Connect resources do not exist in the same resource group or on the same node.

Configuring AIX Connections Services

AIX Connections software lets you share files, printers, applications, and other resources between AIX workstations and PC and Mac clients. You can still take advantage of AIX's multi-user and multi-tasking facilities, scalability, file and record locking features, and other security features. The AIX Connections application is integrated with HACMP so that you can configure it as a resource in your HACMP cluster, making the protocols handled by AIX Connections—IPX/SPX, Net BEUI, and AppleTalk—highly available in the event of node or adapter failure.

This section contains information you may need before adding AIX Connections services in the SMIT screen.

Configuration Notes for AIX Connections

Keep these considerations in mind as you configure AIX Connections:

- Make sure you have copied the AIX Connections installation information to all nodes on which the program might be active.
- On start-up, HACMP starts all AIX Connections services and the network protocols specified in resource groups active on the local node. HACMP does not stop those active on remote nodes, however. You need to see that any AIX Connections services that shouldn't be active on node start-up are not.
- Before you boot up, comment out any **tnstart** commands for HACMP in the initialization scripts (**/etc/rc.lsserver**, **/etc/rc.nwserver**, or **/etc/rc.macserver**), or modify the **/etc/inittab** file to not call these scripts. This prevents the possibility of two nodes broadcasting the same name in shared disk volumes.
- You must configure a realm's AIX Connections to use the service adapter, not the standby

For instructions on configuring AIX Connections resources in SMIT, see the section Configuring Resource Groups on page 12-9.

Adapter Failure Considerations

Keep the following considerations in mind regarding AIX Connections and adapter failures:

- You must define one interface for each network allowing adapter swap. Each IPX/SPX and AppleTalk interface references a physical network adapter card. Each NetBIOS interface, on the other hand, references a Local Area Network Adapter (LANA), and each LANA references a physical network adapter card.
- When an IPX/SPX or AppleTalk protocol is running, it broadcasts all services over all interfaces. You need to remove an interface from the configuration to deactivate it. Again, LANAs work differently. You can define a LANA and still deactivate it.
- Even though you need to set up the interfaces anyway for AIX Connections to work, because of the limitations of the IPX/SPX and AppleTalk protocols, HACMP always moves the interfaces to the service adapter on the network.
- The only way to change a protocol's network adapter is to change the configuration file and restart the protocol.

- A configuration file may not report accurately on a protocol's network adapters for either of two reasons: the configuration file either might have been changed without restarting the protocol, or it might not report the way the user originally entered the information.

For instructions on adding AIX Connections services to a resource group, see the section Configuring Resources for Resource Groups on page 12-9.

Verification of AIX Connections Configuration

After you have configured your resources, you synchronize cluster resources across all nodes. During this process, the **clverify** command checks the following AIX Connections configuration information:

- Realm/service pairs are configured correctly on all nodes in the resource group
- Volume references of realm/service pairs are configured on all nodes in the resource group
- Printer references of realm/service pairs are configured on all nodes in the resource group
- Attach points of realm/service pairs are configured on all nodes in the resource group
- AIX Connections initialization files (**rc.lsserver**, **rc. macserver**, and **rc.nwserver**) contain no **tnstart** commands that aren't commented out. (See the Configuration Notes section above for more information.)

If problems are detected, the realm/service configuration check produces an error, the rest warnings.

Shell Commands for AIX Connections

For information about AIX Connections-specific commands, see the appendix on HACMP for AIX Commands, in the *HACMP for AIX Administration Guide*.

Configuring CS/AIX Communications Links

CS/AIX communication links are integrated with HACMP already so you can configure them, via the SMIT interface, as highly available resources in resource groups. They do not need to be associated with application servers or special scripts.

An HACMP for AIX CS/AIX communication link contains CS/AIX configuration information which is specific to a given node and network adapter. This configuration information enables an RS/6000 computer to participate in an SNA network that includes mainframes, PCs and other workstations. You can configure CS/AIX Communications Links resources using the SMIT interface.

See the *HACMP for AIX Planning Guide* for installation considerations, supported networks, CS/AIX protocols, and CS/AIX product versions

Creating a CS/AIX Communications Link

These steps describe how to configure a highly available CS/AIX communications link. After completing these steps on a single node, the HACMP for AIX software copies the information to all cluster nodes when you synchronize the nodes, described in the section Synchronizing Cluster Resources on page 12-15.

To configure a CS/AIX Communications Link:

1. Enter the following to start HACMP for AIX system management:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration > Cluster Resources > Define Highly Available Communications Links > Define Communication Links**. This brings you to the main AIX SMIT menu for CS/AIX system configuration. Press F1 for help on entries required for configuring these links. A valid CS/AIX configuration must exist before a CS/AIX DLC profile can be made highly available.
3. Press F3 to return to the **Define Highly Available Communications Links** screen. Select **Make Communications Links Highly Available > Add a Highly Available Communications Link**.
4. Enter field values as follows:

   DLC Name — Identify the CS/AIX DLC profile to be made highly available. Pick F4 to see a list of the DLC names.

   Port — Enter ASCII text strings for the names of any CS/AIX ports to be started automatically.

   Link Station — Enter ASCII text strings for the names of the CS/AIX link stations.

   Note: CS/AIX 4.2 does not offer this field entry. CS/AIX 5.0 does offer this field entry.

   Service — Enter the full pathname of an application start script. This start script starts any application layer processes that use the communication link. This field is optional.

5. Press Enter to add this information to the HACMP for AIX ODM.
6. Press F10 after the command completes to leave SMIT and return to the command line.
7. Once you have defined a highly available CS/AIX communications link, you then must configure it as a resource in a resource group. See Configuring Resources for Resource Groups on page 12-9 for information on how this is done.

Changing a CS/AIX Communications Link

To change or remove a highly available CS/AIX communications link, see the chapter on Changing Resources and Resource Groups of the *HACMP for AIX Administration Guide* for more information.

CS/AIX Communications Links as Highly Available Resources

CS/AIX connections are protected during adapter and node failures. This section describes the HACMP for AIX actions that take place for each of these failures.

Note: HACMP for AIX handles the stopping of DLC profiles during an adapter or node failure differently depending on whether CS/AIX Version 5.0 or 4.2 is used.

CS/AIX Version 5.0 allows individual DLC profiles to be stopped. This allows HACMP for AIX to only stop the DLC profiles which are using the adapter being taken out of service. Communication on other DLC profiles is unaffected.

CS/AIX Version 4.2 does not allow stopping of individual DLC profiles. HACMP for AIX must stop all of the active DLC profiles, on all active adapters.

Adapter Fallover and Recovery

When a service adapter over which CS/AIX is running fails, HACMP will take the following actions: Stop the LU2 or LU6.2 sessions; stop the link stations; if CS/AIX version 4.2, stop the CS/AIX server; modify the DLC profiles to use an available standby adapter; verify CS/AIX; restart CS/AIX if stopped; start the link stations; start the sessions; and start the applications.

Depending on the number of DLC profiles, this process make take several minutes. CS/AIX may be unavailable during the time it takes to recover from an adapter failure. Clients or applications connected before the failure may have to reconnect.

Node Fallover and Recovery

CS/AIX profiles are defined as resource in a resource group to HACMP. When a node fails, the resource group is taken over in the normal fashion, and the CS/AIX DLC profiles are restarted on the takeover node. Any identified resources of that DLC profile, such as link stations and service applications are started on the takeover node.

Network Fallover and Recovery

Network failures are handled as they would in a non CS/AIX environment. When a network failure occurs, HACMP for AIX detects an IP network down and logs an error message in the */tmp/hacmp.out* file. Even though the CS/AIX network is independent of the IP network, it is assumed that an IP network down event indicates that the C/AIX network is also down. Recovery is achieved through customer customization, as is consistent with HACMP for AIX network fallover strategy.

Verification of CS/AIX Communications Links

The **clverify** command will check the consistency of DLC profiles between nodes which have a fallover relationship. An error messages is generated if a DLC profile is not available on a node which participates in a resource group containing that profile. There is no checking for invalid CS/AIX configuration information; it is assumed that the system administrator has properly configured CS/AIX.

Configuring Resources—Overview

The HACMP for AIX software provides a highly available environment by identifying a set of cluster-wide resources essential to uninterrupted processing, and then by defining relationships among nodes that ensure these resources are available to client processes. Resources include the following hardware and software:

- Disks
- Volume groups
- Filesystems
- Network addresses
- Application servers

In the HACMP for AIX software, you define each resource as part of a *resource group*. This allows you to combine related resources into a single logical entity for easier configuration and management. You then configure each resource group to have a particular kind of relationship with a set of nodes. Depending on this relationship, resource groups can be defined as one of three types: cascading, concurrent access, or rotating.

Furthermore, a cascading resource group attribute, Cascading without Fallback (CWOF), allows you to modify the behavior of a cascading resource group configuration. For more information on CWOF, refer to Cluster Resources and Resource Groups on page 1-9 of the *HACMP for AIX Concepts and Facilities Guide*.

Refer to your resource group worksheets to complete the steps for configuring resources. See Chapter 7, Planning Application servers and Resource Groups, of the *HACMP for AIX Planning Guide* if you have not completed these worksheets.

After configuring the cluster topology, you must configure resources and set up the cluster node. This involves:

- Configuring resource groups and node relationships to behave as desired
- Adding individual resources to each resource group
- Setting up run-time parameters per node
- Synchronizing cluster nodes.

This remainder of this chapter provides instructions for these steps.

Changing the Resource and Node Configuration

After you have configured your resources and topology, when you want to make changes to the configuration, follow the steps in Chapter 7, Changing Resources and Resource Groups, of the *HACMP for AIX Administration Guide.*

Configuring Resource Groups

Do the following steps for each resource group:

1. Name the resource group.
2. Define the node/resource relationship for the resource group.
3. Define the resource chain for the resource group, and assign priorities to the participating nodes.
4. Configure the resource group (assign individual resources to the resource group).

Creating Resource Groups

To create resource groups:

1. Enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration** > **Cluster Resources** > **Define Resource Groups > Add a Resource Group** and press Enter.
3. Enter the field values as follows:

   Resource Group Name Enter an ASCII text string that identifies the resource group. The resource group name can include alphabetic or numeric characters and underscores. Use no more than 31 characters. Duplicate entries are not allowed.

   Node Relationship Toggle the entry field between **Cascading**, **Concurrent**, and **Rotating**.

   Participating Node Names Enter the names of the nodes that you want to be members of the resource chain for this resource group. Enter the node names in order from highest to lowest priority (left to right). Leave a space between node names.

   Priority is ignored for concurrent resource groups.

4. Press Enter to add the resource group information to the HACMP for AIX ODM.
5. Press F3 after the command completes until you return to the Cluster Resources screen, or F10 to exit SMIT.

Configuring Resources for Resource Groups

Once you have defined resource groups, you configure each by assigning cluster resources to one resource group or another. You can configure resource groups even if a node is powered down; however, SMIT cannot list possible shared resources for the node (making configuration errors likely).

Converting from AIX Connections to AIX Fast Connect

If you previously configured the AIX Connections application as a highly available resource, and you now wish to switch to AIX Fast Connect, you should take care to examine your AIX Connections planning and configuration information before removing it from the resource group. Remember that you cannot have both of these applications configured at the same time in the same resource group, so you must unconfigure all AIX Connections realm/service pairs before configuring Fast Connect fileshares and print queues.

The following instructions are repeated in the *HACMP for AIX Administration Guide*, in the chapter on changing resources and resource groups.

Keep in mind that AIX Fast Connect does not handle the AppleTalk and NetWare protocols that AIX Connections is able to handle. Fast Connect is primarily for connecting with clients running Windows operating systems. Fast Connect uses NetBIOS over TCP/IP.

Follow these steps when converting from AIX Connections to Fast Connect:

1. Refer to your original planning worksheet for AIX Connections, where you listed the participating nodes and the realm/service pairs you planned to configure. Compare this information to your Fast Connect planning worksheet so you can be sure you are not leaving anything out.

   If you do not have your planning sheet, note the information you see in the AIX Connections Services field when you go into SMIT to remove the AIX Connections realm/service pairs from the resource group.

2. Start the Fast Connect server on each node and verify that you can connect to the shared directories and files on each node in turn.
3. In SMIT, go to the Change/Show Resource Groups screen, as described in the section below.
4. Select the AIX Connections Resources field and remove all specified realm/service pairs.
5. Select Fast Connect Services and specify the resources you wish to configure in the resource group. If you are specifying Fast Connect fileshares, make sure you have defined their filesystems in the Filesystems field earlier in the SMIT screen.
6. Synchronize the cluster as usual after you have made all changes. Instructions for synchronizing begin on page 1 2 -15.

General Considerations for Configuring Resources

Keep the following points in mind when configuring resources:

- You cannot configure a resource group until you have completed the information on the **Add a Resource Group** screen.
- If you plan to configure AIX Fast Connect services, you must remove any configured AIX Connections services from the resource group first. You cannot have both in the same resource group.
- If you configure a cascading resource group with an NFS mount point, you must also configure the resource to use IP Address Takeover. If you do not do this, takeover results are unpredictable.You should also set the field value **Filesystems Mounted Before IP Configured** to **true** so that the takeover process proceeds correctly.

- When setting up a cascading resource with an IP Address takeover configuration, each cluster node should be configured in no more than (*N* + 1) resource groups on a particular network. Here, *N* is the number of standby adapters on a particular node and network.
- Failure to use kerberos or **/.rhosts** with a Cascading without Fallback resource group will result in resource group failure.
- HACMP limits the number of nodes participating in a Cascading without Fallback resource group to two.

Configuring Resources for a Resource Group Using SMIT

To define the resources that will be part of a resource group:

1. From the **Cluster Resources** SMIT screen, select the **Change/Show Resources/Attributes for a Resource Group** option and press Enter.

   SMIT displays a picklist of defined resource groups.

2. Pick the desired resource group.
3. Press Enter and SMIT displays the **Configure a Resource Group** screen.
4. Enter values that define all the resources you want to add to this resource group. If the participating nodes are powered on, you can press F4 to list the shared resources. If a resource group/node relationship has not been defined, or if a node is not powered on, F4 displays the appropriate warnings.

Resource Group Name Reflects the choice you made on the previous screen; the resource group to configure.

Node Relationship Reflects the fallover strategy entered when you created the resource group.

Participating Node Names Reflects the names of the nodes that you entered as members of the resource chain for this resource group. Node names are listed in order from highest to lowest priority (left to right), as you designated them.

Service IP Label If IP address takeover is being used, list the IP labels to be taken over when this resource group is taken over. Press F4 to see a list of valid IP labels. These include addresses which rotate or may be taken over.

Filesystems Identify the filesystems to include in this resource group. Press F4 to see a list of the filesystems. When you enter a filesystem in this field, the HACMP for AIX software determines the correct values for the Volume Groups and Raw Disk PVIDs fields. If you will be configuring Fast Connect fileshares, be sure their filesystems are configured here.

Filesystems Consistency Check Identify the method of checking consistency of filesystems, **fsck** (default) or **logredo** (for fast recovery).

Filesystems Recovery Method

Identify the recovery method for the filesystems, **parallel** (for fast recovery) or **sequential** (default).

Do *not* set this field to **parallel** if you have shared, nested filesystems. These must be recovered sequentially. (Note that the cluster verification utility, **clverify**, does not report filesystem and fast recovery inconsistencies.)

Filesystems/Directories to Export

Identify the filesystems or directories to be exported. The filesystems should be a subset of the filesystems listed above. The directories for export should be contained in one of the filesystems listed above. Press F4 for a list.

Filesystems/Directories to NFS Mount

Identify the filesystems or directories to NFS mount. All nodes in the resource chain will attempt to NFS mount these filesystems or directories while the owner node is active in the cluster.

Network for NFS Mount

(This field is optional.)

Choose a previously defined IP network where you want to NFS mount the filesystems. The F4 key lists valid networks.

This field is relevant only if you have filled in the previous field. The **Service IP Label** field should contain a service label which is on the network you choose.

Note: You can specify more than one service label in the **Service IP Label** field. It is highly recommended that at least one entry be an IP label on the network chosen here.

If the network you have specified is unavailable when the node is attempting to NFS mount, it will seek other defined, available IP networks in the cluster on which to establish the NFS mount.

Volume Groups

Identify the shared volume groups that should be varied on when this resource group is acquired or taken over. Press F4 to see a list of shared volume groups.

If you have previously entered values in the **Filesystems** field, the appropriate volume groups are already known to the HACMP for AIX software.

If you are using raw logical volumes in non-concurrent mode, you only need to specify the volume group in which the raw logical volume resides to include the raw logical volumes in the resource group.

Concurrent Volume Groups

Identify the shared volume groups that can be accessed simultaneously by multiple nodes. Press F4 to see a list of shared volume groups.

Raw Disk PVIDs

Press F4 for a listing of the PVIDs and associated hdisk device names.

If you have previously entered values in the **Filesystems** or **Volume groups** fields, the appropriate disks are already known to the HACMP for AIX software.

If you are using an application that directly accesses raw disks, list the raw disks here.

AIX Connections Services

Press F4 to choose from a list of all realm/service pairs that are common to all nodes in the resource group. You can also type in realm/service pairs. Use **%** as a divider between service name and service type; do not use a colon. *Note that you cannot configure both AIX Connections and AIX Fast Connect in the same resource group.*

AIX Fast Connect Resources

Press F4 to choose from a list of Fast Connect resources common to all nodes in the resource group. If you configure Fast Connect fileshares, make sure you have defined their filesystems in the resource group in the Filesystems field. *Note that you cannot configure both AIX Connections and AIX Fast Connect in the same resource group. See the section Converting from AIX Connections to AIX Fast Connect on page 12-10 for further notes on this.*

Application Servers

Indicate the application servers to include in the resource group. Press F4 to see a list of application servers. See the section Configuring Application Servers on page 12-1 for information on defining application servers.

Highly Available Communications Links

Indicate the communications links to include in the resource group. Press F4 to see a list of communications links. See the section Configuring CS/AIX Communications Links on page 12-5 for information on defining communications links.

Miscellaneous Data

A string you want to place into the topology, along with the resource group information. It is accessible by the scripts, for example, *Database1*.

Inactive Takeover Activated

Set this variable to control the *initial acquisition* of a resource group by a node when the node/resource relationship is cascading. This variable does not apply to rotating or concurrent resource groups.

If Inactive Takeover is **true**, then the first node in the resource group to join the cluster acquires the resource group, regardless of the node’s designated priority.

If Inactive Takeover is **false**, the first node to join the cluster acquires only those resource groups for which it has been designated the highest priority node.

The default is **false**.

Cascading without Fallback Enabled

Set this variable to determine the fallback behavior of a cascading resource group.

When the CWOF variable is set to **false**, a cascading resource group will fallback as a node of higher priority joins or reintegrates into the cluster.

When CWOF is **true**, a cascading resource group will not fallback as a node of higher priority joins or reintegrates into the cluster. It migrates from its owner node only if the owner node fails. It will not fallback to the owner node when it reintegrates into the cluster.

Note: You may find it useful to review the *HACMP for AIX Concepts Guide* section, Fallover vs. Fallback on page 1-10.

The default for CWOF is **false**.

9333 Disk Fencing Activated

By default, 9333 disk fencing is disabled in a concurrent access environment. To enable 9333 disk fencing, set the field to **true.** Once set, the values in a fence register can typically only be changed by power-cycling the 9333 unit. The fence register is immune to all other “reset” conditions.

Certain occurrences (for example, powering the disk up or down or killing the Cluster Manager) could leave the 9333 disks fenced out from a node (becoming) responsible for managing them. Therefore, the HACMP for AIX software provides a command to clear fence register contents in the same way that power-cycling the disks would. If a node needs access to a disk that is fenced out, you can clear the fence registers for that disk to allow the node access to disk resources. Use the command provided on the Cluster Recovery Aids SMIT screen to do this in *extraordinary* circumstances only.

SSA Disk Fencing Activated

By default, SSA disk fencing is disabled in a concurrent access environment. SSA disk fencing is only available for concurrent access configurations. To enable SSA disk fencing, set this field to **true.** Once set, the values in a fence register cannot be changed by power-cycling the SSA unit. Use the Cluster Recovery Aids SMIT screen to clear the fence registers. For more information on SSA disk fencing, see Chapter 5, Planning Shared Disk Devices, of the *HACMP for AIX Planning Guide*.

Filesystems Mounted Before IP Configured This field specifies whether, on fallover, HACMP takes over volume groups and mounts filesystems before or after taking over the failed node's IP address or addresses.

The default is **false**, meaning the IP address is taken over first. Similarly, upon reintegration of a node, the IP address is acquired before the filesystems.

Set this field to **true** if the resource group contains filesystems to export. This is so that the filesystems will be available once NFS requests are received on the service IP address.

5. After entering field values, synchronize cluster resources.
6. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

   If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 12-15.

Synchronizing Cluster Resources

The act of synchronizing cluster nodes sends the information contained on the current node to all defined cluster nodes.

Note: All configured nodes must be on their boot addresses when a cluster has been configured and the nodes are synchronized for the first time. Any node not on its boot address will not have its **/etc/rc.net** file updated with the HACMP for AIX entry; this causes problems for the reintegration of this node into the cluster.

If a node attempts to join a cluster when another node is out-of-sync with other active cluster nodes, it will be denied. You must ensure that other nodes are synchronized to the joining member.

To synchronize cluster nodes:

1. Select **Synchronize Cluster Resources** from the Cluster Resources menu. Set the appropriate values in the fields that appear.

   SMIT prompts you to confirm that you want to synchronize cluster resources.

   Note: To save time, you can skip cluster verification during synchronization by setting the **Skip Cluster Verification** field in SMIT to **yes**.

2. Press Enter to synchronize the resource group configuration and all cluster nodes.

When synchronization is complete, press F3 until you return to the HACMP for AIX main menu, or F10 to exit SMIT.

Configuring Run-Time Parameters

To define the run-time parameters for a node:

1. From the Cluster Resources SMIT screen, select the **Change/Show Run Time Parameters** option and press Enter. SMIT displays a picklist of cluster nodes. You define run-time parameters individually for each node.
2. Select a node name and press Enter. SMIT displays the Change/Show Run Time Parameters screen with the node name displayed.
3. Enter field values as follows:

   Debug Level Cluster event scripts have two levels of logging. The **low** level logs errors encountered while the script executes. The **high** level logs all actions performed by the script. The default is **high**.

   Host uses NIS or Name Server If the cluster uses Network Information Services (NIS) or name serving, set this field to **true**. The HACMP for AIX software disables these services before entering reconfiguration, and enables them after completing reconfiguration. The default is **false.**

4. Press Enter to add the values to the HACMP for AIX ODM.
5. Press F3 until you return to the Cluster Resources menu, or F10 to exit SMIT.

Chapter 13 Verifying the Cluster Topology

This chapter describes the process of verifying the cluster topology, including the cluster and node configurations. This process ensures that all nodes agree on the cluster topology and on the assignment of resources.

Prerequisites

- Define the cluster topology following the instructions in Chapter 11, Defining the Cluster Topology
- Register the application servers, configure resources, and define the node environment following the instructions in Chapter 12, Configuring Cluster Resources.

Overview

After defining the cluster topology and configuration, run the Cluster Verification utility, **/usr/sbin/cluster/diag/clverify**, on one node to check that all cluster nodes agree on the cluster configuration and the assignment of HACMP for AIX resources. Run this utility before starting the HACMP for AIX software.

The **clverify** utility cluster option contains the following programs and options:

- The **topology** program verifies that all nodes agree on the cluster topology. It contains the following options:
  - The **check** option checks for agreement on cluster, node, network, and adapter information. This option is not available through SMIT.
    - For example, it checks for invalid characters in cluster names, node names, network names, adapter names and resource group names
  - The **sync** option allows you to synchronize the cluster topology if necessary, forcing agreement with the local node’s definition.
- The **config** program verifies that networks are configured correctly and that all nodes agree on the ownership of resources. It contains the following options:
  - The **networks** option checks for the valid configuration of adapters and serial lines, and for netmask consistency on all cluster nodes.
    - also checks each cluster node to determine whether multiple RS232 serial networks exist on the same tty device
  - The **resources** option
    - checks for agreement among all nodes on the ownership of defined resources (filesystems, volume groups, disks, application servers, and events) and on the distribution of resources in case of a takeover
    - checks, more specifically, for the existence and defined ownership of filesystems to be taken over
    - checks the volume group and disks where the filesystems reside to verify that the takeover information matches the owned resources information

    - checks the major device numbers for NFS-exported directories
    - checks to ensure that application servers are configured correctly
    - prints out diagnostic information about custom snapshot methods, custom verification methods, custom pre/post events, and redirection of cluster log files.
  - The **All** option checks the network topology and resources, and runs all custom-defined verification methods.

If you have configured Kerberos on your system, the **clverify** utility also verifies that:

- All IP labels listed in the configuration have the appropriate service principals in the **.klogin** file on each node in the cluster.
- All nodes have the proper service principals.
- Kerberos is installed on all nodes in the cluster.
- All nodes have the same security mode setting.

You can run the Cluster Verification procedure through SMIT (except for the **topology check** option), or you can use run **clverify** utility interactively or directly from the command line. Using the SMIT interface to verify the cluster is described in this chapter.

For information about using **clverify** interactively or about adding, changing, or removing custom-defined verification methods that perform specific checks, see the chapter on verifying a cluster configuration in the *HACMP for AIX Administration Guide*.

Also see the **clverify** man page for details about using this utility.

Verifying the Cluster Topology Using SMIT

After defining the cluster topology, run the Cluster Verification procedure on one node to check that all nodes agree on the assignment of cluster resources.

To verify a cluster and node configuration:

1. Enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration** and press Enter.
3. Select **Cluster Verification** and press Enter.
4. Select **Verify Cluster** and press Enter.

   Fill in the fields as follows:

   Base HACMP Verification Methods
   By default, both the cluster topology and resources verification programs are run. You can toggle this entry field to run either program, or you can select **none** to specify a custom-defined verification method in the Define Custom Verification Method field.

Define Custom Verification Method Enter the name of a custom-defined verification method. You can also press F4 for a list of previously defined verification methods. By default, if no methods are selected, the clverify utility will check the topology, resources, and all custom verification methods in alphabetical order.

The order in which verification methods are listed determines the sequence in which selected methods are run. This sequence remains the same for subsequent verifications until different methods are selected.

Log File to store output Enter the name of an output file in which to store verification output. By default, verification output is stored in the **smit.log** file.

5. Press Enter. A screen similar to the following appears.

```
                         COMMAND STATUS


Command: OK           stdout: yes            stderr: no

      Before command completion
      below.

      (result from clverify command)

       (list of any configuration errors from clverify command)
```

If you receive error messages, make the necessary corrections and run the verification procedure again.

Checking Cluster Topology

Run the following command to verify that all nodes agree on the cluster topology:

```
clverify cluster topology check
```

When the program completes, check the output. If a cluster topology problem exists, a message similar to the following appears:

```
ERROR: Could not read local configuration
ERROR: Local Cluster ID XXX different from Remote Cluster ID XXX.
ERROR: Nodes have different numbers of networks
```

Synchronizing Cluster Topology

If all nodes do not agree on the cluster topology and you are sure you want to define the cluster as it is defined on the local node, you can force agreement of cluster topology onto all nodes by synchronizing the cluster configuration.

To synchronize a cluster configuration across nodes:

1. Enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration>Cluster Topology>Synchronize Cluster Topology**
3. Press Enter.

   The following fields appear:

   Ignore Cluster Verification Errors By choosing **yes**, the result of the cluster verification is ignored and the configuration is synchronized even if verification fails.

   By choosing **no**, the synchronization process terminates; view the error messages in the system error log to determine the configuration problem.

   Emulate or Actual If you set this field to **Emulate**, the synchronization is an emulation and does not affect the Cluster Manager. If you set this field to **Actual**, the synchronization actually occurs, and any subsequent changes affect the Cluster Manager. **Actual** is the default value.

   Skip Cluster Verification By default, this field is set to **no** and the cluster topology verification program is run. To save time in the cluster synchronization process, you can toggle this entry field to **yes**.

4. Specify **yes** or **no** in the **Ignore Cluster Verification Errors.** Set the **Emulate or Actual** field to **Actual**. Leave the default setting of **no** in the **Skip Cluster Verification** field.

   Press enter

   The cluster definition (including all node, adapter, and network method information) is copied to the other cluster nodes.

5. Press F10 to exit SMIT.

Skipping Cluster Verification During Synchronization

Note that to save time during cluster synchronization, you can skip cluster verification. Cluster verification is optional only when a cluster is *inactive*. Even if one node is active, cluster verification will be run.

You can skip cluster verification using SMIT or at the command line. To skip verification of cluster topology using the SMIT interface, choose **Yes** at the **Skip Cluster Verification** field**.**

Chapter 14 Customizing Cluster Events and Log Files

This chapter explains how to customize event processing for an HACMP cluster by including pre- and post-processing and event notification. It also describes the process of customizing cluster log files.

Prerequisites

- Read the chapter on tailoring cluster event processing in the *HACMP for AIX Planning Guide* before reading this chapter.
- Complete the planning worksheets discussed in the *HACMP for AIX Planning Guide.*
- Define the cluster topology following the instructions in Chapter 11, Defining the Cluster Topology.
- Define the cluster resources following the instructions in Chapter 12, Configuring Cluster Resources.

Customizing Event Processing

An HACMP for AIX cluster environment is event-driven. An event is a change of status within a cluster that the Cluster Manager recognizes and processes. When the Cluster Manager detects a change in cluster status, it executes a script designated to handle the event and its subevents. You can specify additional processing to customize event handling for your site if needed.

The HACMP for AIX software provides a script for each event and subevent. By default, the Cluster Manager calls the corresponding event script supplied with the HACMP for AIX software for a specific event. If the actions taken by the default scripts are sufficient for your site, you do not need to do anything further to configure events.

To tailor event processing to your environment, however, use the HACMP for AIX event customization facility.

The HACMP for AIX software provides an event customization facility that allows you to tailor event processing to your site. Use this facility to include the following types of customization:

- Adding, changing, removing custom cluster events
- Pre- and post-event processing. Note that you must register any user-defined scripts with the HACMP for AIX software as Custom Cluster Events.
- Event notification
- Event recovery and retry.

Pre- and Post-Event Processing

To tailor event processing to your environment, specify commands or user-defined scripts that execute before and after a specific event is generated by the Cluster Manager. For pre-processing, for example, you may want to send a message to specific users, informing them to stand by while a certain event occurs. For post-processing, you may want to disable login for a specific group of users if a particular network fails.

Event Notification

You can specify a command or user-defined script that provides notification (for example, mail) that an event is about to happen and that an event has just occurred, along with the success or failure of the event.

Event Recovery and Retry

You can specify a command that attempts to recover from an event command failure. If the retry count is greater than zero and the recovery command succeeds, the event script command is rerun. You can also specify the number of times to attempt to execute the recovery command.

Notes on Customizing Event Processing

You must declare a shell (for example #**! /bin/sh**) at the beginning of each script executed by the notify, recovery, and pre/post-event processing commands.

Notify, recovery, and pre- and post-event processing do not occur when the force option of the **node_down** event is specified.

Warning: Be careful not to kill any HACMP processes as part of your script!
As a precaution, you can add this line to your scripts:

```
grep -v hacmprd_run_rcovcmd
```

Event Emulator

To test the effect of running an event on your cluster, HACMP for AIX provides a utility to run an emulation of an event. This emulation lets you predict a cluster's reaction to an event as though the event actually occurred. The emulation runs on all active nodes in your cluster, and the output is stored in an output file. You can select the path and name of this output file using the **EMU_OUTPUT** environment variable or use the default **/emuhacmp.out** file on the node that invoked the Event Emulator.

For more information on event emulation, see these chapters: Administrative Facilities in the *HACMP for AIX Concepts and Facilities* guide and Monitoring an HACMP Cluster in the *HACMP for AIX Administration Guide*.

Node Events

The Cluster Manager recognizes node_up and node_down events.

node_up Events

A node_up event can be initiated by a node joining the cluster at cluster startup, or rejoining the cluster after having previously left the cluster.

Joining the Cluster at Cluster Startup

This section describes the steps taken by the Cluster Manager on each node when the cluster starts and the initial membership of the cluster is established. It shows how the Cluster Managers establish communication among the member nodes, and how the cluster resources are distributed as the cluster membership grows.

First Node Joins the Cluster

1. The HACMP for AIX software is started on Node A. Node A then broadcasts a message indicating that it is ready to join the cluster on all configured networks to which it is attached.
2. Node A interprets the lack of a response to mean that it is the first node in the cluster.
3. Node A initiates a node_up_local event, which processes the node environment resource configuration information. When the event processing has completed, Node A becomes a member of the cluster.

   All resource groups defined for Node A are available for clients at this point.

   If Node A is defined as part of a resource chain for several resource groups and the Inactive Takeover node environment variable is set to TRUE, it will take control of all these resource groups. If this variable is set to FALSE, it will take control of only those resource groups for which it has the highest priority.

   If Node A is defined as part of a rotating resource configuration, it takes control of the first resource group per network listed in the node environment.

   If Node A is defined as part of a concurrent access resource configuration, it makes those concurrent resources available.

Second Node Joins the Cluster

4. The HACMP for AIX software is started on Node B. Node B broadcasts a message indicating that it is ready to join the cluster on all configured networks to which it is attached.
5. Node A receives the message and sends an acknowledgment.
6. Node A adds Node B to a list of “active nodes,” starts keepalive communications with Node B, and puts a node_up_remote event in its event queue.
7. Node B receives the acknowledgment from Node A. The message includes information identifying Node A as the only other member of the cluster. (If there were other members, Node B would receive the list of members.)

8. Node A processes the node_up_remote event and sends a message to let other nodes know when it is finished.

   Processing the node_up_remote event may include releasing resources currently held by Node A, if both nodes are in the resource chain for one or more resource groups and Node B has a higher priority for one or more of those resources.

9. Meanwhile, Node B has been monitoring and sending keepalives, and waiting to receive messages about changes in the cluster membership. When Node B receives the message that Node A has finished its node_up processing, it processes its own node_up_local event and notifies Node A when it is done.

   During its node_up processing, Node B claims all resource groups configured for it (see step 3 above).

10. Both nodes process a node_up_complete event simultaneously.

    At this point, Node B includes Node A in its "member nodes" and "keepalive" lists.

11. Node B sends a "new member" message to all possible nodes in the cluster.

12. When Node A gets the message, it moves Node B from its "active nodes" list to its "member nodes" list.

    At this point, all resource groups configured for Node A and Node B are available to cluster clients.

Remaining Nodes Join the Cluster

13. As the HACMP for AIX software is started on each remaining cluster node, steps 4 through 9 are repeated, with each member node sending and receiving control messages, and processing events in the order outlined. Note especially that all nodes must verify the node_up_complete event before completing the processing of the event and moving the new node to the "cluster member" list.

    As new nodes join, already active nodes determine (as part of the processing of the node_up event) who their immediate (logical) neighbors are. Each node maintains keepalive communication with its immediate neighbors. For example, Node D exchanges keepalive information with Nodes C and E. Node A exchanges keepalive messages with Node B and with whatever node has the name beginning with the highest alphabetical letter.

Rejoining the Cluster

When a node rejoins the cluster, the Cluster Managers running on the existing nodes initiate a node_up event to acknowledge that the returning node is up. When these nodes have completed their event processing, the new node then processes a node_up event so that it can resume providing cluster services.

This processing is necessary to ensure the proper balance of cluster resources. As long as the existing Cluster Managers first acknowledge a node rejoining the cluster, they can release any resource groups belonging to that node if necessary. Whether or not the resource groups are actually released in this situation depends on how the resource groups are configured for takeover. The new node can then start its operations.

Sequence of node_up Events

The following listing describes the sequence of node_up events.

node_up

This event occurs when a node joins the cluster. Depending on whether the node is local or remote, this event initiates either a node_up_local or node_up_remote event.

node_up_local

This script acquires the service address (or shared address), gets all its owned (or shared) resources, and takes the resources. This includes making disks available, varying on volume groups, mounting filesystems, exporting filesystems, NFS-mounting filesystems, and varying on concurrent access volumes groups.

acquire_service_addr

(If configured for IP address takeover.) Configures boot addresses to the corresponding service address, and starts TCP/IP servers and network daemons by running the **telinit -a** command.

acquire_takeover_addr

The script checks to see if a configured standby address exists then swaps the standby address with the takeover address.

get_disk_vg_fs

Acquires disk, volume group, and filesystem resources.

node_up_remote

Causes the local node to release all resources taken from the remote node and to place the concurrent volume group in concurrent mode. Some of the scripts called by node_up_local include:

release_takeover_addr

(If configured for IP address takeover.) Identifies a takeover address to be released because a standby adapter on the local node is masquerading as the service address of the remote node. Reconfigures the local standby adapter to its original address (and hardware address, if necessary).

stop_server

Stops application servers belonging to the reintegrating node.

release_vg_fs

Releases volume groups and filesystems belonging to a resource group that the remote node will be taking over.

cl_deactivate_nfs

Unmount NFS filesystems.

node_up_complete

This event occurs only after a node_up event has successfully completed. Depending on whether the node is local or remote, this event initiates either a node_up_local_complete or node_up_remote_complete event.

node_up_local_complete

Calls the start_server script to starts application servers. This event occurs only after a node_up_local event has successfully completed.

node_up_remote_complete

Allows the local node to do an NFS mount only after the remote node is completely up. This event occurs only after a node_up_remote event has successfully completed.

node_down Events

Cluster nodes exchange keepalives with peer nodes so that the Cluster Manager can track the status of the nodes in the cluster. A node that fails or is stopped purposefully no longer sends keepalives. The Cluster Managers then post a node_down event. Depending on the cluster configuration, the peer nodes then take the necessary actions to get critical applications up and running and to ensure data remains available.

A node_down event can be initiated by a node:

- Being stopped “gracefully”
- Being stopped “gracefully with takeover”
- Being stopped “forcefully”
- Failing

Graceful Stop

In a *graceful stop,* the HACMP for AIX software stops on the local node after the node_down_complete event releases some or all of the stopped node’s resources. The other nodes run the node_down_complete event with the “graceful” parameter and do not take over the resources of the stopped node.

Graceful with Takeover Stop

In a *graceful with takeover stop*, the HACMP for AIX software stops after the node_down_complete event on the local node releases some or all of its resource groups. The surviving nodes in the resource chain take over these resource groups.

Forced Stop

In a *forced stop,* the HACMP for AIX software stops immediately on the local node. The node_down event is not run on this node, but it sends a message to the other nodes to view it as a graceful stop. The Cluster Managers on remote nodes process node_down events, but do not take over any resource groups. The stopped node retains control of its resource groups.

Node Failure

When a node fails, the Cluster Manager on that node does not have time to generate a node_down event. In this case, the Cluster Managers on the surviving nodes recognize a node_down event has occurred (when they realize the failed node is no longer communicating) and trigger node_down events.

The node_down_remote event initiates a series of subevents that reconfigure the cluster to deal with that failed node. Based upon the cluster configuration, surviving nodes in the resource chain will take over the resource groups.

Sequence of node_down Events

The following listing describes the sequence of node_down events.

node_down

This event occurs when a node intentionally leaves the cluster or fails. Depending on whether the exiting node is local or remote, this event initiates either the node_down_local or node_down_remote event, which in turn initiates a series of subevents.

node_down_local

Processes the following events:

stop_server

Stops application servers.

release_takeover_addr

(If configured for IP address takeover.) Identifies a takeover address to be released because a standby adapter on the local node is masquerading as the service address of the remote node. Reconfigures the local standby with its original IP address (and hardware address, if necessary).

release_vg_fs

Releases volume groups and filesystems that are part of a resource group the local node took from the remote node.

release_service_addr

(If configured for IP address takeover.) Detaches the service address and reconfigures the service adapter to its boot address.

node_down_remote

Unmounts highly available NFS filesystems and places any concurrent volume group in non-concurrent mode if the local node is the only surviving node in the cluster accessing that volume group. If the failed node did not go down gracefully, this event calls the necessary subevent scripts.

acquire_service_addr

(If configured for IP address takeover.) Checks the configured boot adapter, configures boot addresses to the corresponding service or shared address, and starts TCP/IP servers and network daemons by running the **telinit -a** command.

acquire_takeover_addr

(If configured for IP address takeover.) Checks for a configured standby address currently seen as up by the Cluster Manager, and then does a standby_address to takeover_address swap (and hardware address, if necessary.

get_disk_vg_fs

Acquires disk, volume group, and filesystem resources as part of a takeover.

node_down_complete

This event occurs only after a node_down event has successfully completed. Depending on whether the node is local or remote, this event initiates either a node_down_local_complete or node_down_remote_complete event.

node_down_local_complete

Instructs the Cluster Manager to exit when the local node has left the cluster. This event occurs only after a node_down_local event has successfully completed.

node_down_remote_complete

Starts takeover application servers if the remote node did not stop gracefully. This event occurs runs only after a node_down_remote event has successfully completed.

start_server

Starts application servers.

Network Events

The Cluster Manager recognizes the network_down and network_up events.

Sequence of Network Events

The following listing shows the network events. By default, these events take no action, since each site must determine its own needs.

network_down

This event occurs when the Cluster Manager determines a network has failed. A network_down event can take one of two forms:

Local network_down, where only a particular node has lost contact with a network.

Global network_down, where all of the nodes connected to a network have lost contact with a network. It is assumed in this case that a network-related failure has occurred rather than a node-related failure.

The network_down event mails a notification to the system administrator, but takes no further action since appropriate actions depend on the local network configuration.

network_down_complete

This event occurs only after a network_down event has successfully completed. The default network_down_complete event processing takes no actions since appropriate actions depend on the local network configuration.

network_up

This event occurs when the Cluster Manager determines a network has become available for use. The default network_up event processing takes no actions since appropriate actions depend on the local network configuration.

network_up_complete This event occurs only after a network_up event has successfully completed. The default network_up_complete event processing takes no actions since appropriate actions depend on the local network configuration.

Network Adapter Events

The Cluster Manager reacts to the failure, unavailability, or joining of network adapters by initiating one of the following events. (For exceptions, see note below on single adapter situations.):

swap_adapter This event occurs when the service adapter on a node fails. The swap_adapter event exchanges or swaps the IP addresses of the service and a standby adapter on the same HACMP network and then reconstructs the routing table.

swap_adapter_complete This event occurs only after a swap_adapter event has successfully completed. The swap_adapter_complete event ensures that the local ARP cache is updated by deleting entries and pinging cluster IP addresses.

swap_address This event occurs when a user requests to move a service/boot address to an available standby adapter on the same node and network.

swap_address_complete This event occurs only after a swap_address event has successfully completed. The swap_address_complete event ensures that the local ARP cache is updated by deleting entries and pinging cluster IP addresses.

fail_standby This event occurs if a standby adapter fails or becomes unavailable as the result of an IP address takeover. The fail_standby event displays a console message indicating that a standby adapter has failed or is no longer available.

join_standby This event occurs if a standby adapter becomes available. The join_standby event displays a console message indicating that a standby adapter has become available.

Failure of a Single Adapter Does Not Generate Events

Be aware that if you have only one adapter active on a network, the Cluster Manager does not generate a failure event for that adapter. “Single adapter” situations include:

- One-node clusters
- Multi-node clusters with only one node active
- Failure of all but one adapter on a network, one at a time.

For example, starting a cluster with all service or standby adapters disconnected produces results as follows:

1. *First node up*: No failure events are generated.
2. *Second node up*: One failure event is generated.
3. *Third node up*: One failure event is generated.
4. And so on.

Whole Cluster Status Events

By default, the Cluster Manager recognizes a six-minute time limit for reconfiguring a cluster and processing topology changes. If the time limit is reached, the Cluster Manager initiates one of the following events:

config_too_long This event occurs when a node has been in reconfiguration for more than six minutes. The event periodically displays a console message.

unstable_too_long This event occurs when a node has been unstable (processing topology changes) for more than six minutes. The event periodically displays a console message.

reconfig_topology_start This event marks the beginning of a dynamic reconfiguration of the cluster topology.

reconfig_topology_complete This event indicates that a cluster topology dynamic reconfiguration has completed.

reconfig_resource_acquire This event indicates that cluster resources that are affected by dynamic reconfiguration are being acquired by appropriate nodes.

reconfig_resource_release This event indicates that cluster resources affected by dynamic reconfiguration are being released by appropriate nodes.

reconfig_resource_complete This event indicates that a cluster resource dynamic reconfiguration has completed.

Configuring Custom Cluster Events

To add customized cluster events, take the following steps.

1. To start system management for HACMP for AIX, enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration** > **Cluster Custom Modification** > **Define Custom Cluster Events**. SMIT displays the menu choices for adding, changing, or removing a custom event.

Adding Customized Cluster Events

To add a customized event:

1. Select **Add a Custom Cluster Event** from the menu.
2. Enter the field values as follows:

   Cluster Event Name The name can have a maximum of 32 characters.

   Cluster Event Description Enter a short description of the event.

   Cluster Event Method Enter the full pathname of the script to execute.

3. Define the cluster event method to the system (see the following section).

Customizing Pre- or Post-Event Processing

Complete the following steps to change the processing for an event. The changes you can make include pointing the Cluster Manager to a different script to process the event, or using the event customization facility to specify pre- or post- processing event scripts. You only need to complete these steps on a single node and then synchronize the ODM data on the other cluster nodes. The HACMP for AIX system propagates the information to the other nodes.

1. To start system management for HACMP for AIX, enter:

   ```
   smit hacmp
   ```

2. From the main menu, select **Cluster Configuration** > **Cluster Resources** >**Cluster Events** >**Change/Show Cluster Events**. When you press Enter, SMIT displays the list of events.
3. Select a specific event or subevent that you want to configure and press Enter. SMIT displays the node name, event name, description, and default event command.
4. Enter field values as follows:

   Event Command Enter the name of the command that processes the event. HACMP for AIX provides a default script. If additional functionality is required, make any changes by adding pre- or post-event processing scripts or commands of your own design, rather than modifying the default scripts or writing new ones.

Notify Command *This field is optional.* Enter the full pathname of a user-supplied script to run both before and after a cluster event. This script can notify the system administrator that an event has occurred. The arguments passed to the command are: The event name, one keyword (either start or complete), the exit status of the event (if the keyword was complete), and the same arguments passed to the event command.

Pre-Event Command *This field is optional.* The field has a picklist of pre-defined custom cluster event names; you can enter more than one custom event name. Use the F7 key to get an alphabetized list of custom event names, or enter the custom event names in the desired order, separated by commas. The names must correspond to a custom event name already defined in the HACMPcustom ODM.

This command is run before the cluster event command executes. This command provides pre-processing before a cluster event occurs. The arguments passed to this command are the event name and the arguments passed to the event command.

Post-Event Command *This field is optional.* The field has a picklist of pre-defined custom cluster event names; you can enter more than one custom event name. Use the F7 key to get an alphabetized list of custom event names, or enter the custom event names in the desired order, separated by commas. The names must correspond to a custom event name already defined in the HACMPcustom ODM.

This command is run after the cluster event command executes. This command provides post-processing after a cluster event. The arguments passed to this command are the event name, event exit status, and the arguments passed to the event command.

Recovery Command *This field is optional.* Enter the full pathname of a user-supplied script or AIX command to execute to attempt to recover from a cluster event command failure. If the recovery command succeeds and the retry count is greater than zero, the cluster event command is rerun. The arguments passed to this command are the event name and the arguments passed to the event command.

Recovery Counter *This field is optional.* Enter the number of times to run the recovery command. Set this field to zero if no recovery command is specified, and to at least one if a recovery command is specified.

5. Press Enter to add this information to the ODM on the local node.
6. Synchronize your changes across all cluster nodes by selecting the **Synchronize Cluster Resources** option off the **Cluster Resources** SMIT screen. Press F10 to exit SMIT.

Note: Synchronizing does not propagate the actual new or changed scripts; you must add these to each node manually.

Customizing Log Files

You can redirect a cluster log from its default directory to a directory of your choice. Should you redirect a log file to a directory of your choice, keep in mind that the requisite (upper limit) disk space for most cluster logs is 2MB. 14MB is recommended for **hacmp.out**.

Note: Logs should not be redirected to shared filesystems or NFS filesystems. Having logs on those filesystems may cause problems if the filesystem needs to unmount during a fallover event.

To redirect a cluster log from its default directory to another destination, take the following steps:

1. Enter

   ```
   smitty hacmp
   ```

2. Select **Cluster System Management** > **Cluster Log Management** > **Change/Show Cluster Log Directory**

   SMIT displays a picklist of cluster log files with a short description of each.

Log	Description
cluster.mmdd	Cluster history files generated daily
cm.log	Generated by clstrmgr activity
cspoc.log	Generated by C-SPOC commands
dms_loads.out	Generated by deadman switch activity
emuhacmp.out	Generated by event emulator scripts
hacmp.out	Generated by event scripts and utilities

3. Select a log that you want to redirect.

   SMIT displays a screen with the selected log's name, description, default pathname, and current directory pathname. The current directory pathname will be the default pathname if you do not change it.

   Edit the final field to change the default pathname. The example below shows the **cluster.mmdd** log file screen.

Custom Log Name	**cluster.mmdd**
Cluster Log Description	Cluster history files generated daily
Default Log Destination Directory	**/usr/sbin/cluster/history**
Log Destination Directory	The default directory name appears here. To change the default, enter the desired directory pathname.

4. Press F3 to return to the screen to select another log to redirect, or return to the Cluster System Management screen to proceed to the screen for synchronizing cluster resources.
5. After you change a log destination directory, a prompt appears reminding you to synchronize cluster resources from this node (cluster log ODMs must be identical across the cluster). The cluster log destination directories as stored on this node will be synchronized to all nodes in the cluster.

Log destination directory changes will take effect when you synchronize cluster resources, or if the cluster is not up, the next time cluster services are restarted.

Sample Custom Scripts

Two situations where it is useful to run custom scripts are illustrated here:

- Making **cron** jobs highly available
- Making print queues highly available.

Making cron jobs Highly Available

To help maintain the HACMP environment, you need to have certain **cron** jobs execute only on the cluster node that currently holds the resources. If a **cron** job executes in conjunction with a resource or application, it is useful to have that **cron** entry fallover along with the resource. It may also be necessary to remove that **cron** entry from the **cron** table if the node no longer possesses the related resource or application.

The following example shows one way to use a customized script to do this:

The example cluster is a two node hot standby cluster where node1 is the primary node and node2 is the backup.

Node1 normally owns the shared resource group and application. The application requires that a **cron** job be executed once per day but only on the node that currently owns the resources.

To ensure that the job will be run even if the shared resource group and application have fallen over to node2, create two files as follows:

1. Assuming that the root user is executing the **cron** job, create a file *root.resource* and another called *root.noresource* in a directory on a non-shared filesystem on node1. Make these files resemble the **cron** tables that reside in the directory **/var/spool/crontabs**.

   The *root.resource* table should contain all normally executed system entries and entries pertaining to the shared resource or application.

   The *root.noresource* table should contain all normally executed system entries but no entries pertaining to the shared resource or application.
2. Copy the files to the other node so that both nodes have a copy of the two files.
3. On both systems, the following command should be executed at system startup:

   ```
   crontab root.noresource
   ```

   This will ensure that the **cron** table for root has only the “no resource” entries at system startup.

4. You can use either of two methods to activate the *root.resource* **cron** table. The first method is the simpler of the two.
   - Execute *crontab root.resource* as the last line of the application start script. In the application stop script, the first line should then be *crontab root.noresource*. By executing these commands in the application start and stop scripts, you are ensured that they will activate and deactivate on the proper node at the proper time.
   - Execute the **crontab** commands as a post_event to node_up_complete and node_down_complete.
     - Upon node_up_complete on the primary node, execute *crontab root.resources* .
     - On node_down_complete execute *crontab root.noresources*.

     The takeover node must also use the event handlers to execute the correct **cron** table. Logic must be written into the node_down_complete event to determine if a takeover has occurred and to execute the *crontab root.resources* command. On a reintegration, a pre-event to node_up must determine if the primary node is coming back into the cluster and then execute *a crontab root.noresource* command.

Making Print Queues Highly Available

In the event of a fallover, the print jobs currently queued can be saved and moved over to the surviving node.

The print spooling system consists of two directories: /**var/spool/qdaemon** and **/var/spool/lpd/qdir**. One directory contains files containing the data (content) of each job. The other contains the files consisting of information pertaining to the print job itself. When jobs are queued, there are files in each of the two directories. In the event of a fallover, these directories do not normally fallover and thus the print jobs are lost.

The solution for this problem is to define two filesystems on a shared volume group. You might call these filesystems **/prtjobs** and **/prtdata**. When HACMP for AIX starts, these filesystems are mounted over **/var/spool/lpd/qdir** and **/var/spool/qdaemon**.

Write a script to perform this operation as a post event to node_up. The script should do the following:

- Stop the print queues
- Stop the print queue daemon
- Mount **/prtjobs** over **/var/spool/lpd/qdir**
- Mount **/prtdata** over **/var/spool/qdaemon**
- Restart the print queue daemon
- Restart the print queues.

In the event of a fallover, the surviving node will need to do the following:

- Stop the print queues
- Stop the print queue daemon
- Move the contents of **/prtjobs** into **/var/spool/lpd/qdir**
- Move the contents of **/prtdata** into /**var/spool/qdaemon**
- Restart the print queue daemon
- Restart the print queues.

To do this, write a script called as a post-event to node_down_complete on the takeover. The script needs to determine if the node_down is from the primary node.

Chapter 15 Setting Up Clinfo on Server Nodes

This chapter describes how to edit files and scripts for the proper use of the Cluster Information Program (Clinfo) on server nodes. Clinfo must be run on server nodes in order to receive cluster status information from the clstat utility, described in the chapter Monitoring an HACMP Cluster in the *HACMP for AIX Administration Guide*.

Prerequisites

- Read the chapters Tailoring Cluster Event Processing and Planning HACMP for AIX Clients of the *HACMP for AIX Planning Guide* before reading this chapter.
- Install the HACMP for AIX, Version 4.4 Licensed Program Product (LPP) on the cluster nodes. See Chapter 8, Installing HACMP for AIX Software, for more information.

Overview of the Cluster Information Program

Clinfo is an SNMP-based monitor. SNMP is an industry-standard set of standards for monitoring and managing TCP/IP-based networks. SNMP includes a protocol, a database specification, and a set of data objects. A set of data objects forms a Management Information Base (MIB). SNMP provides a standard MIB that includes information such as IP addresses and the number of active TCP connections. The actual MIB definitions are encoded into the agents running on a system. The standard SNMP agent is the SNMP daemon, **snmpd**.

The HACMP for AIX software provides the HACMP for AIX MIB, associated with and maintained by the HACMP for AIX management agent, the Cluster SMUX peer daemon (**clsmuxpd**). Clinfo retrieves information from the HACMP for AIX MIB through **clsmuxpd**.

Installing the HACMP for AIX software on a client machine enables the client to receive messages from Clinfo about events and actions taken by the high availability software running on the cluster. The client can take predefined automatic steps in response to some situations handled by the high availability software. It can also print messages making the users logged in to the client aware of actions they may need to take to maintain connectivity, or it may simply inform them of the cluster state.

For information on starting and stopping Clinfo and on using Clinfo in asynchronous mode, see the chapter on starting and stopping cluster services of the *HACMP for AIX Administration Guide*.

Editing the /usr/sbin/cluster/etc/clhosts File

For the Clinfo daemon (**clinfo**) to get the information it needs, you must edit the **/usr/sbin/cluster/etc/clhosts** file. This file should contain hostnames (addresses) of any HACMP for AIX nodes (servers) with which **clinfo** can communicate, including servers from clusters accessible through logical connections.

Note: If a client is located in a network that has both HACMP and HACMP/ES clusters, the following rules apply to the **clhosts** file:

- If the client has HACMP software installed, the **clhosts** file must be configured with only HACMP hostnames. If the files contains any HACMP/ES hostnames, **clinfo** will abort.
- If the client has HACMP/ES software installed, the **clhosts** file may contain both HACMP and HACMP/ES hostnames.

As installed, the **/usr/sbin/cluster/etc/clhosts** file on an HACMP for AIX server node contains a loopback address. The **clinfo** daemon first attempts to communicate with a **clsmuxpd** process locally. If it succeeds, **clinfo** then acquires an entire cluster map, including a list of all HACMP for AIX server interface addresses. From then on, **clinfo** uses this list rather than the provided loopback address to recover from a **clsmuxpd** communication failure.

If **clinfo** does not succeed in communicating with a **clsmuxpd** process locally, however, it only can continue trying to communicate with the local address. For this reason, you should replace the loopback address with all HACMP for AIX service addresses accessible through logical connections to this node. The loopback address is provided only as a convenience.

Important: Do not include standby addresses in the **clhosts** file, and do not leave this file empty. If either of these conditions exist, neither **clinfo** nor the **/usr/sbin/cluster/clstat** utility will work properly.

An example **/usr/sbin/cluster/etc/clhosts** file follows:

```
cowrie_en0_cl83# cowrie service
140.186.91.189 #  limpet service
floyd_en0_cl83 #  floyd service
squid_en0_cl83 #  squid service
```

Editing the /usr/sbin/cluster/etc/clinfo.rc Script

The **/usr/sbin/cluster/etc/clinfo.rc** script, executed whenever a cluster event occurs, updates the system's ARP cache. If you are not using the hardware address swapping facility, a copy of the **clinfo.rc** script must exist on each node and client in the cluster in order for all ARP caches to be updated and synchronized. Flushing the ARP cache typically is not necessary if the HACMP for AIX hardware address swapping facility is enabled because hardware address swapping maintains the relationship between a network address and a hardware address.

Note: In a switched Ethernet network you may need to flush the ARP cache to ensure that the new MAC address is communicated to the switch, or use the procedure "MAC Address Is Not Communicated to the Ethernet Switch" described in the *HACMP for AIX Troubleshooting Guide* to ensure that the MAC address is communicated correctly.

The HACMP for AIX software is distributed with a template version of the **clinfo.rc** script. You can use the script as distributed, add new functionality to the script, or replace it with a custom script.

If you are not using hardware address swapping, the ARP functionality must remain. Edit the **/usr/sbin/cluster/etc/clinfo.rc** file on each server node by adding the IP label or IP address of each system that will be accessing shared filesystems managed by HACMP to the PING_CLIENT_LIST list. Then run the **clinfo** daemon.

Note: You can also set the PING_CLIENT_LIST in the file **/etc/cluster/ping_client_list**. This method ensures that the list of clients to ping will not be overlaid by future changes to **clinfo.rc.**

The format of the **clinfo** call to **clinfo.rc**:

```
clinfo.rc {join,fail,swap} interface_name
```

When **clinfo** gets a cluster_stable event, or when it connects to a new **clsmuxpd**, it receives a new map. It next checks for changed states of interfaces.

- If a new state is UP, **clinfo** calls `clinfo.rc join interface_name.`
- If a new state is DOWN, **clinfo** calls `clinfo.rc fail interface_name.`
- If **clinfo** receives a node_down_complete event, it calls **clinfo.rc** with the fail parameter for each interface currently UP.
- If **clinfo** receives a fail_network_complete event, it calls **clinfo.rc** with the fail parameter for all associated interfaces.
- If **clinfo** receives a swap_complete event, it calls `clinfo.rc swap interface_name`.

See the chapter on tailoring cluster event processing of the *HACMP for AIX Planning Guide* for complete information on cluster events and tailoring scripts.

See the sample Clinfo client program in the *Programming Client Applications* guide for a sample client application that uses the Clinfo C API within the context of a customized **clinfo.rc** script.

Note: If you decide not to use **clinfo**, please refer to Chapter 17, Installing and Configuring Clients, to install the HACMP for AIX Version 4.4 LPP on clients and to configure clients for an HACMP cluster.

Rebooting the Clients

The final step in installing the HACMP for AIX software on a client is to reboot each client in your cluster topology.

Chapter 16 Supporting AIX Error Notification

This chapter describes how to use the AIX Error Notification facility to identify and respond to failures in an HACMP cluster.

Error Notification

The AIX Error Notification facility detects errors matching predefined selection criteria and responds in a programmed way. The facility provides a wide range of criteria you can use to define an error condition. These errors are called *notification objects.*

Each time an error is logged in the system error log, the error notification daemon determines if the error log entry matches the selection criteria. If it does, an executable is run. This executable, called a *notify method*, can range from a simple command to a complex program. For example, the notify method might be a mail message to the system administrator or a command to shut down the cluster.

The following sections discuss the Error Notification facility as it applies to the HACMP for AIX software, and they include specific examples of how you might use this facility with the software. Refer to the AIX InfoExplorer facility for a broader discussion of the Error Notification facility.

Using Error Notification in an HACMP for AIX Environment

Use the Error Notification facility to add an additional layer of high availability to the HACMP for AIX software. Although the combination of the HACMP for AIX software and the inherent high availability features built into the AIX operating system keeps single points of failure to a minimum, failures still exist that, although detected, are not handled in a useful way.

Take the example of a cluster where an owner node and a takeover node share a SCSI disk. The owner node is using the disk. If the SCSI adapter on the owner nodes fails, an error may be logged, but neither the HACMP for AIX software nor the AIX Logical Volume Manager responds to the error. If the error has been defined to the Error Notification facility, however, an executable that shuts down the node with the failed adapter could be run, allowing the surviving node to take over the disk.

Note: If you are using SP nodes, see Appendix B for information about specific errors that should be provided to AIX error notification for SP Switch failures.

Defining an Error Notification Object and Notify Method

To define an error notification object and its corresponding notify method:

1. Enter the following to start system management for the HACMP for AIX software:

   ```
   smit hacmp
   ```

2. Select **RAS Support** > **Error Notification > Add a Notify Method** and press Enter.

   Here you define the notification object and its associated notify method.

3. Enter values for the following fields:

Notification Object Name Enter a user-defined name that identifies the error.

Persist across system restart? Set this field to **yes** if you want this notification object to survive a system reboot. If not, set the field to **no**.

Process ID for use by Notify Method Specify a process ID for the notify method to use. Objects that have a process ID specified should have the **Persist across system restart** field set to **no**.

Select Error Class Identify the class of error log entries to match. Valid values are: **None** No error class **All** All error classes **Hardware** Hardware error class **Software** Software error class **Errlogger** Messages for the **errlogger** command

Select Error Type Identify the severity of error log entries to match. Valid values are: **None** No entry types to match **All** Match all error types **PEND** Impending loss of availability **PERM** Permanent **PERF** Unacceptable performance degradation **TEMP** Temporary **UNKN** Unknown

Match Alertable Errors? Indicate whether the error is alertable. This descriptor is provided for use by alert agents associated with network management applications. Valid alert descriptor values are: **None** No errors to match **All** Match all alertable errors **TRUE** Matches alertable errors **FALSE** Matches non-alertable errors

Select Error Label Enter the label associated with a particular error identifier as defined in the **/usr/include/sys/errids.h** file. If you are unsure about an error label, press F4 for a listing. Specify **All** to match all error labels.

Resource Name Indicate the name of the failing resource. For the hardware error class, a resource name is a device name. For the software error class, the resource name is the name of the failing executable. Specify **All** to match all resource names.

Resource Class Indicate the class of the failing resource. For the hardware error class, the resource class is the device class. The resource error class does not apply to software errors. Specify **All** to match all resource classes.

Resource Type Enter the type of failing resource. For the hardware error class, a resource is the device type by which a resource is known in the devices object. Specify **All** to match all resource types.

Notify Method Enter the name of the executable that should run whenever an error matching the defined selection criteria occurs. The following keywords are automatically expanded by the error notification daemon as arguments to the notify method:

$1 Sequence number from the error log entry
$2 Error ID from the error log entry
$3 Error class from the error log entry
$4 Error type from the error log entry
$5 Alert flag values from the error log entry
$6 Resource name from error log entry
$7 Resource type from the error log entry
$8 Resource class from the error log entry

4. Press Enter to create the notification object and method.
5. Press F10 to exit SMIT.

You now have defined a notification object and a corresponding notify method. The next time the error occurs, the system detects the error and responds as directed.

Examples

The following examples suggest how the Error Notification facility can be used in an HACMP for AIX cluster environment. A completed **Add a Notify Method** screen is shown for each example.

Example 1: Permanent Software Errors

In this example, the notification object is any permanent software error. The notify method is a mail message to the system administrator indicating that an error has occurred. The message includes the error ID and resource name.

Notification Object Name	SOFT_ERR.
Persist across system restart?	**yes**
Process ID for use by Notify Method	
Select Error Class	**Software**
Select Error Type	**PERM**
Match Alertable Errors?	
Select Error Label	.
Resource Name	
Resource Class	
Resource Type	

Notify Method	echo Permanent error occurred. Error ID + $2 Resource Name = $6 | mail sysadmin

Example 2: Permanent Hardware Errors (SCSI Adapter Device Driver)

In this example, the notification object is a permanent hardware error on the SCSI adapter device driver. The notify method performs a **halt -q** to shut down the node.

Notification Object Name	SCSI_adapter
Persist across system restart?	**yes**
Process ID for use by Notify Method	
Select Error Class	**Hardware**
Select Error Type	**PERM**
Match Alertable Errors?	**None**
Select Error Label	**SCSI_ERR1**
Resource Name	**scsi0**
Resource Class	**adapter**
Resource Type	**adapter**
Notify Method	**`halt -q`**

Automatic Error Notification

Using SMIT screen options in HACMP version 4.3.1 and above, you can configure error notification automatically for the cluster resources listed below, list currently defined automatic error notify entries for the same cluster resources, or remove previously configured automatic error notify methods. Before you configure Automatic Error Notification, you must have a valid HACMP configuration.

Warning: Automatic error notification should be configured only when the cluster is not running.

Choosing to add error notify methods automatically runs the **cl_errnotify** utility which turns on error notification on all nodes in the cluster for the following devices:

- All disks in the rootvg volume group
- All disks in HACMP volume groups, concurrent volume groups, and filesystems (To avoid single points of failure, the JFS log must be included in an HACMP volume group.)
- All disks defined as HACMP resources
- The SP switch adapter (for clusters with SP nodes).

Automatic error notification applies to selected hard, non-recoverable error types: disk, disk adapter, and SP switch adapter errors. No media errors, recovered errors, or temporary errors are supported by this utility.

Executing automatic error notification assigns one of two error notification methods for all the error types noted:

- **cl_failover** is assigned if a disk or an adapter (including SP switch adapter) is determined to be a single point of failure and its failure should cause the cluster resources to fall over. In case of a failure of any of these devices, this method logs the error to **hacmp.out** and shuts down the cluster software on the node. It first tries to do a graceful shutdown with takeover; if this fails, it calls **cl_exit** to shut down the node.
- **cl_logerror** is assigned for all other error types. In case of a failure of any of these devices, this method logs the error to **hacmp.out**.

You can also use the utility to list currently defined auto error notification entries in your HACMP cluster configuration and to delete all automatic error notify methods.

Configuring Automatic Error Notification

To configure automatic error notification, take the following steps:

1. Be sure the cluster is not running.
2. Open the SMIT main HACMP menu by typing `smit hacmp.`
3. From the main menu, choose **RAS Support > Error Notification > Configure Automatic Error Notification.**
4. Select the **Add Error Notify Methods for Cluster Resources** option from the following list:

   List Error Notify Methods for Cluster Resources — Lists all currently defined auto error notify entries for certain cluster resources: HACMP defined volume groups, concurrent volume groups, filesystems, and disks; rootvg; SP switch adapter (if present). The list is output to the screen.

   Add Error Notify Methods for Cluster Resources — Error notification methods are automatically configured on all relevant cluster nodes.

   Delete Error Notify Methods for Cluster Resources — Error notification methods previously configured with the **Add Error Notify Methods for Cluster Resources** option are deleted on all relevant cluster nodes.

5. (optional) Since error notification is automatically configured for all the listed devices on all nodes, you must make any modifications to individual devices or nodes manually, after running this utility. To do so, choose the **Error Notification** option in the **RAS Support** SMIT screen. See the earlier section in this chapter.

   Note: If you make any changes to cluster topology or resource configuration, you may need to reconfigure automatic error notification. When you run **clverify** after making any change to the cluster configuration, you will be reminded to reconfigure error notification if necessary.

Listing Error Notify Methods

To see the automatic error notify methods that currently exist for your cluster configuration, take the following steps:

1. From the main HACMP menu, choose **RAS Support > Error Notification > Configure Automatic Error Notification.**
2. Select the **List Error Notify Methods for Cluster Resources** option. The utility lists all currently defined automatic error notification entries with these HACMP components: HACMP defined volume groups, concurrent volume groups, filesystems, and disks; rootvg; SP switch adapter (if present). The list is output to a screen similar to that shown below, in which the cluster nodes are named *sioux* and *quahog*:

```
                      COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

sioux:
sioux: HACMP Resource           Error Notify Method
sioux:
sioux: hdisk0           /usr/sbin/cluster/diag/cl_failover
sioux: hdisk1           /usr/sbin/cluster/diag/cl_failover
sioux: scsi0            /usr/sbin/cluster/diag/cl_failover
quahog:
quahog: HACMP Resource          Error Notify Method
quahog:
quahog: hdisk0          /usr/sbin/cluster/diag/cl_failover
quahog: scsi0           /usr/sbin/cluster/diag/cl_failover
```

Deleting Error Notify Methods

To delete automatic error notification entries previously assigned using this utility, take the following steps:

1. From the main menu, choose **RAS Support > Error Notification > Configure Automatic Error Notification.**
2. Select the **Delete Error Notify Methods for Cluster Resources** option. Error notification methods previously configured with the **Add Error Notify Methods for Cluster Resources** option are deleted on all relevant cluster nodes.

Error Log Emulation

After you have added one or more notify methods, you can test your methods by emulating an error. This allows you to learn whether your pre-defined notify method carries out the intended action.

To emulate an error log entry:

1. From the main HACMP SMIT menu, choose **RAS Support > Error Notification > Emulate Error Log Entry**.

   The **Select Error Label** box appears, showing a picklist of the notification objects for which notify methods have been defined.

2. Select a notification object and press return to begin the emulation.

   As soon as you press the return key, the emulation process begins: the emulator inserts the specified error into the AIX error log, and the AIX error daemon runs the notification method for the specified object.

3. When the emulation is complete, you can view the error log by typing the **errpt** command to be sure the emulation took place. The error log entry has either the resource name EMULATOR, or a name as specified by the user in the **Resource Name** field during the process of creating an error notify object.

   You can now determine whether the specified notify method was carried out.

   Note: Remember that the actual notify method will be run. Whatever message, action, or executable you defined will occur. Depending on what it is, you may need to take some action, for instance, to restore your cluster to its original state.

Only the root user is allowed to run an error log emulation.

Part 4 Installing and Configuring Cluster Clients

This part contains instructions for installing and configuring HACMP for AIX on client nodes.

Chapter 17 Installing and Configuring Clients

This chapter describes how to install the HACMP for AIX, Version 4.4 Licensed Program Product (LPP) on clients, and how to configure clients for an HACMP cluster.

Prerequisites

- Read Planning HACMP for AIX Clients in the *HACMP for AIX Planning Guide* before reading this chapter.
- Install the HACMP for AIX, Version 4.4 LPP on the cluster nodes. See Chapter 8, Installing HACMP for AIX Software, for a new install, or Chapter 9, Upgrading an HACMP Cluster, to upgrade the cluster configuration.
- Read the *HACMP for AIX LPP 4.4 Release Notes* in **/usr/lpp/cluster/doc/release_notes** for additional information on installing the HACMP for AIX software.

Overview

Installing the HACMP for AIX software on each RS/6000, SP, or SMP client that will run the Clinfo daemon enables the clients to receive messages about events and actions taken by the high availability software running in the cluster. The client can take predefined, automatic steps in response to some situations handled by the high availability software, and it can print messages to inform users logged in to a client of the cluster state and thus make them aware of actions required to maintain connectivity.

Note: Do not install HACMP for AIX on a diskless, dataless client. This configuration is not supported

Installing and configuring the HACMP for AIX software on each client consists of the following steps:

1. Install the base high availability system software on all clients. The SPO Feature Code medium contains the base high availability system.
2. Edit the **/usr/sbin/cluster/etc/clhosts** file on each client to specify IP addresses on cluster nodes that run **clsmuxpd**.
3. Edit the **/usr/sbin/cluster/etc/clinfo.rc** script on each client to specify IP addresses on cluster nodes that run **clsmuxpd**.

   Note: You can also set the PING_CLIENT_LIST in the file **/etc/cluster/ping_client_list**. This method ensures that the list of clients to ping will not be overlaid by future changes to **clinfo.rc.**

4. Reboot each client in your cluster topology.

If you do not plan to use hardware address swapping and if your clients are not running **clinfo** or access cluster nodes through a router, see the section Updating Non-Clinfo Clients and Routers Accessed by Any Client on page 17-4 to ensure that clients' and routers' ARP caches reflects the new hardware address of nodes after a node_up, node_down, or swap_adapter event.

Note: The maximum number of clusters Clinfo can monitor is eight.

Installing the Base System Client Images

For a new installation, the **/usr** directory should have a minimum of three megabytes (MB) of space available. If you are upgrading the software, see Chapter 9, Upgrading an HACMP Cluster for instructions on this step.

To install the base high availability software on a client:

1. Place the HACMP for AIX tape into the tape drive and enter:

   ```
   smit install_selectable_all
   ```

   If you are not sure of the name of the input device, press F4 to list the available devices. Select the proper drive and press Enter. That value is entered into the **INPUT device/directory** field as the valid input device.

2. Press Enter. SMIT refreshes the screen.

3. Enter field values as follows:

SOFTWARE to install	Press F4 for a software listing. A popup window appears, listing all installed software. Use the arrow keys to choose the following required filesets: **cluster.base.client.lib**, **cluster.base.client.rte**, and **cluster.base.client.utils**. You can also use the arrow keys to select any other client filesets, which are optional. Press Enter after making all selections. Your selections appear in this field. If you select at least one required base client module, all other required client modules are installed automatically.

4. Enter values for other fields as appropriate for your site.

5. Press Enter when you are satisfied with the entries. SMIT responds:

   ```
   ARE YOU SURE?
   ```

6. Press Enter again.

You are then instructed to read the HACMP 4.4 **release_notes** file in the **/usr/lpp/cluster** directory for further instructions.

Editing the /usr/sbin/cluster/etc/clhosts File on Clients

As installed, the **clhosts** file on an HACMP for AIX client node contains no hostnames or addresses. You must provide the HACMP for AIX server addresses at installation time. This file should contain all boot and service names (or addresses) of HACMP for AIX servers accessible through logical connections to this client node. Upon startup, **clinfo** uses these names to attempt communication with a **clsmuxpd** process executing on an HACMP for AIX server.

Warning: Do not include standby addresses in the **clhosts** file, and do not leave this file empty. If either of these conditions exist, neither **clinfo** nor **clstat** works properly.

An example **/usr/sbin/cluster/etc/clhosts** file follows:

```
cowrie_en0_cl83       #  cowrie service
140.186.91.189        #  limpet service
floyd_en0_cl83        #  floyd service
squid_en0_cl83        #  squid service
```

Editing the /usr/sbin/cluster/etc/clinfo.rc Script

The **/usr/sbin/cluster/etc/clinfo.rc** script, executed whenever a cluster event occurs, updates the system's ARP cache. If you are not using the hardware address swapping facility, a copy of the **clinfo.rc** script must exist on each node and client in the cluster in order for all ARP caches to be updated and synchronized. Flushing the ARP cache typically is not necessary if the HACMP for AIX hardware address swapping facility is enabled because hardware address swapping maintains the relationship between a network address and a hardware address.

Note: In a switched Ethernet network, you may need to flush the ARP cache to ensure that the new MAC address is communicated to the switch, or use the procedure, "MAC Address Is Not Communicated to the Ethernet Switch," described in the *HACMP for AIX Troubleshooting Guide* to ensure that the MAC address is communicated correctly.

The HACMP for AIX software is distributed with a template version of the **clinfo.rc** script. You can use the script as distributed, add new functionality to the script, or replace it with a custom script.

Note: If you are not using hardware address swapping, the ARP functionality must remain.

The format of the **clinfo** call to **clinfo.rc**:

```
clinfo.rc {join,fail,swap} interface_name
```

When **clinfo** gets a cluster_stable event, or when it connects to a new **clsmuxpd**, it receives a new map. It next checks for changed states of interfaces.

- If a new state is UP, **clinfo** calls `clinfo.rc join interface_name.`
- If a new state is DOWN, **clinfo** calls `clinfo.rc fail interface_name.`
- If **clinfo** receives a node_down_complete event, it calls **clinfo.rc** with the fail parameter for each interface currently UP.

- If **clinfo** receives a fail_network_complete event, it calls **clinfo.rc** with the fail parameter for all associated interfaces.
- If **clinfo** receives a swap_complete event, it calls `clinfo.rc swap interface_name`.

See the chapter on tailoring event processing of the *HACMP for AIX Planning Guide* for complete information on cluster events and tailoring scripts.

See the sample Clinfo client program of the *HACMP for AIX Programming Client Applications* guide for a sample client application that uses the Clinfo C API within the context of a customized **clinfo.rc** script.

Finally, if you have written applications that use the Clinfo API and plan to use Sumps, you may need to make changes to your application. See either the *HACMP for AIX Programming Client Applications* guide or the *HACMP for AIX Programming Locking Applications* guide for updated information on the library routines. Then recompile and link your application.

Updating Non-Clinfo Clients and Routers Accessed by Any Client

This section explains how to update the ARP cache of locally-attached, non-Clinfo clients and of routers that sit between cluster nodes and clients, whether or not these clients are running **clinfo**. Keep in mind, however, that the ARP cache processing described in this section is necessary only if IPAT is configured and if hardware address swapping is not enabled.

Note: Adding the script outlined below as a post-event script to the acquire_takeover_addr_event is necessary only if **clinfo** is not running on the takeover node, or if **clinfo** is running but updates to **clinfo.rc** suggested in the previous section have not been implemented.

Updating the ARP Cache After IP Address Takeover

When IP address takeover occurs, the ARP caches of routers and gateways that sit between cluster nodes and clients, and the ARP caches of locally-attached non-clinfo clients must be updated to reflect the adapter hardware address of the takeover node.

To perform this update, write a shell script to run on the surviving node that contains the following commands. Use the post-processing facility to call this shell script after the acquire_takeover_addr_event.

To update the ARP cache on non-Clinfo clients:

1. Add a route to the router or system you wish to update, using the relocated interface:

```
route add destination_IP_address -interface failed_node_service_IP_address
```

This creates a route that specifically uses the interface.

2. Delete the possible ARP entry for the router or system being accessed:

```
arp -d router_IP_address
```

3. Ping the router or system you wish to update again:

```
ping -c1 router_IP_address
```

This updates the ARP cache.

4. Delete the route you previously added:

```
route delete router_IP_address failed_node_service_IP_address
```

This procedure forces the **ping** to go out over the interface with the relocated IP address. The address resolution triggered by the **ping** will provide the router or system you are pinging with the new hardware address now associated with this IP address.

Rebooting the Clients

The final step in installing the HACMP for AIX software on a client is to reboot each client in your cluster topology.

Part 5 Appendixes

This part contains appendixes.

Appendix A Supporting IP Address Takeover

This appendix summarizes the steps required to configure an HACMP cluster to support IP address takeover. The checklist points to where each step is explained in the main text as it occurs during the planning and installation process. Use this appendix as a reference to ensure you have completed all required steps.

Note: In an HACMP for AIX SP Switch network on the SP, integrated Ethernet adapters cannot be used, and no standby adapters are configured. If a takeover occurs, the service address is aliased onto another node’s service address. See Appendix J, Installing and Configuring HACMP for AIX on RS/6000 SPs, for complete information on adapter functions in an SP Switch environment.

Defining Boot Addresses: List of Steps

To define boot addresses and enable IP address takeover:

1. Select a boot address and an associated name for each service adapter on each cluster node for which IP address takeover might occur. See Chapter 3, Planning TCP/IP Networks (the section Defining Boot Addresses) of the *HACMP for AIX Planning Guide* for a complete description of this step.
2. Add the boot addresses to the **/etc/hosts** file and the **nameserver** configuration (if applicable) on each cluster node. See the section Editing the /etc/hosts File and Nameserver Configuration on page 7-2 of this book for a complete description of this step.
3. Use the **smit chinet** fastpath to reconfigure each node (for which IP address takeover might occur) to boot on its boot address rather than on its service address. You must know the IP address of the service adapter. See Chapter 4, Configuring Cluster Networks and Performance Tuning, for information on configuring network adapters.
4. Reboot each system reconfigured in the previous step.
5. Select **Configure Adapters** from the Cluster Topology menu to define the boot addresses to the HACMP cluster environment. See Chapter 11, Defining the Cluster Topology, for a complete description of this step.
6. Select **Change/Show Resources/Attributes for a Resource Group** from the Cluster Resources menu to define the service IP labels to be taken over when the given resource group is taken over. See the section Configuring Resources for Resource Groups on page 12-9 of this book for a complete description of this step.

HACMP for AIX Edits the /etc/inittab and /etc/rc.net Files

When you complete step 6, the HACMP for AIX software performs two additional steps. First, it runs the **/usr/sbin/cluster/utilities/clchipat** utility to edit the **/etc/inittab** file on each node for which IP address takeover might occur. The **/usr/sbin/cluster/utilities/clchipat** utility changes the **rc.tcpip-** and **inet-**dependent entries, and then adds entries to start network daemons. Next, the HACMP for AIX software edits the **/etc/rc.net** file so that the Configuration Manager (**cfgmg**r) does not use the ODM entries to configure network adapters.

Note: If you disable IP address takeover on a node, the HACMP for AIX software restores both the **/etc/inittab** and **/etc/rc.net** files to their original state.

The /etc/inittab File

The **/etc/inittab** file is read by the **init** command. The **/usr/sbin/cluster/events/node_up** script issues the **telinit -a** command when a node joins the cluster to bind various network daemons to the correct service address.

When IP address takeover is defined, the system edits **/etc/inittab** to change the **rc.tcpip-** and **inet-**dependent entries from runlevel “2” (the default multi-user level) to runlevel “a”. Entries that have run level “a” are processed only when the **telinit** command requests them to be run.

Entries Edited

The following entries in the **/etc/inittab** file change from runlevel “2” to runlevel “a”:

- rctcpip
- rcnfs
- qdaemon
- writesrv
- rcncs
- aicd.

You may need to change additional entries in the **/etc/inittab** file. Change any daemon that binds to a network address to run level “a”.

Entries Added

In addition, the HACMP for AIX software adds entries to the **/etc/inittab** file. The first entry is shown below:

```
harc:2:wait:/usr/sbin/cluster/etc/harc.net #HACMP for AIX network startup
```

This entry calls the **/usr/sbin/cluster/etc/harc.net** script, which starts network daemons called by HACMP for AIX utilities.

The second entry added to the **/etc/inittab** file is shown below:

```
clinit:a:wait:touch /usr/sbin/cluster/.telinit #HACMP for AIX Last entry
```

This entry creates the **/usr/sbin/cluster/.telinit** file.

The /etc/rc.net File

The second step performed by the HACMP for AIX software is to edit the **/etc/rc.net** file on each node for which IP address takeover might occur. The **/etc/rc.net** file is edited so that it can be called only by the Cluster Manager while the HACMP for AIX software is running, and not by the AIX **cfgmgr**.

As it normally executes, the **cfgmgr** command runs the **/etc/rc.net** file. The **/etc/rc.net** command gets information about network parameters from the HACMP for AIX ODM and uses this information to configure network adapters. If a boot address is defined in the ODM, any calls to **cfgmgr** while the Cluster Manager is running on its service address reconfigure the service adapter to use the boot address.

Note: If a network adapter fails and a standby adapter is configured as a service adapter, the **cfgmgr** attempts to reconfigure the failed adapter as the service adapter.

To prevent this reconfiguration from happening, the HACMP for AIX software places an identifier in the call to the **/etc/rc.net** file which causes it to exit immediately when called by **cfgmgr**, but to run to completion when called by the Cluster Manager.

The system places the following code after the initial line of pound symbols `(#)` in the **/etc/rc.net** script:

```
######################################################################
# HACMP for AIX
# HACMP for AIX These lines added by HACMP for AIX software
[ "$1" = "-boot" ] && shift || { ifconfig 1o0 127.0.0.1 up; exit 0; } #HACMP for AIX
# HACMP for AIX
```

In an HACMP for AIX cluster, the **/usr/sbin/cluster/etc/rc.cluster** file calls the **/etc/rc.net** file to configure the network. If this file is called with the **-boot** parameter and if a boot address is defined for a node, then **/usr/sbin/cluster/etc/rc.cluster** calls **/etc/rc.net** with the **-boot** parameter, and the network configuration completes. (This is how the HACMP for AIX entry in the **/etc/inittab** file is defined to work.)

To run the **/etc/rc.net** script manually, call it with the **-boot** parameter as follows:

```
rc.net -boot
```

If you modify the **/etc/rc.net** file, and then it is called without the **-boot** parameter, it exits immediately.

Appendix B Installing and Configuring Cluster Monitoring with Tivoli

This appendix contains instructions for making an HACMP cluster known to Tivoli in order to monitor the cluster through the Tivoli management console.

Overview

You can monitor the state of an HACMP cluster and its components through your Tivoli Framework enterprise management system. Using various windows of the Tivoli interface, you can monitor the following aspects of your cluster:

- Cluster state and substate
- Configured networks and network state
- Participating nodes and node state

In order to set up this monitoring, you must do a number of installation and configuration steps in order to make Tivoli aware of the HACMP cluster and to ensure proper functioning of IP address takeover.

For more information on using Tivoli to monitor your cluster once installation is complete, see the chapter on monitoring your cluster in the *HACMP for AIX Administration Guide.*

Installing and Configuring Cluster Monitoring with Tivoli

The rest of this appendix covers prerequisites and procedures for setting up your cluster to be monitored by Tivoli.

Prerequisites and Considerations

When planning and configuring Cluster Monitoring with Tivoli, keep the following points in mind:

- The Tivoli Management Region (TMR) should be located on an AIX node *outside* the cluster.
- The HACMP cluster nodes must be configured as managed nodes in Tivoli.
- The Tivoli Framework, Distributed Monitoring, and AEF components must be installed on the Tivoli Management Region node and on each cluster node. (See page F-4 for details.)
- To ensure accurate monitoring of IP address takeover, the ideal configuration is to have a separate network dedicated to communication between the TMR and the cluster nodes. If you do *not* have a separate network dedicated to Tivoli, you must take additional steps to ensure that IPAT functions properly. These steps include defining an extra subnet and an alias IP address. This additional subnet is used for the TMR node IP address and for the cluster node's alias IP address. (See note below on subnet considerations.)

Memory and Disk Requirements for Cluster Monitoring with Tivoli

The memory required for individual Distributed Monitors for cluster components varies depending on the size of the cluster and the number of components being monitored. Consult your Tivoli documentation for more information.

Installation of the **hativoli** filesets requires 400 KB of disk space. Check your Tivoli documentation for additional disk space requirements.

Subnet Considerations for Cluster Monitoring with Tivoli

In order to ensure the proper monitoring of IP address takeover in a Tivoli-monitored cluster, you must create an alias to the standby adapter of each cluster node. You must include this alias in the **/etc/hosts** file and also in the Tivoli **/etc/wlocalhost** file.

The subnet of this alias must be *different* than the node’s service and standby adapters, and the *same* as the subnet of the non-cluster Tivoli Management Region node. Note that if you already have Tivoli set up, you may need to change its address to match the alias.

Here is an example of what you might insert into the **/etc/hosts** file for a Tivoli-monitored cluster node named HAnode and a Tivoli server node named TMRnode. HAnode has service, standby, and alias IP addresses; TMRnode has a service IP address.

The netmask for this example network is 255.255.255.0

Adapter Label	Address
HAnode_svc	10.50.**20**.88
HAnode_stby	10.50.**25**.88
HAnode_alias	10.50.**21**.89
TMRnode	10.50.**21**.10

In this example, the alias address and the TMR address are on the same subnet, and this subnet is *in addition to* the two already used for the cluster node’s service and standby adapters.

Steps for Installing and Configuring Cluster Monitoring with Tivoli

Preparing to monitor a cluster with Tivoli involves several stages and prerequisite tasks.

The table below provides an overview of all of the steps you will take. Use this table to familiarize yourself with the “big picture” of the installation and configuration steps. Then refer to the sections that follow for details on each step.

This sequence of steps assumes an environment in which:

- Tivoli has already been installed and set up.
- The Tivoli configuration is being modified to monitor an HACMP cluster for the first time.
- You do not have a separate network dedicated to monitoring the HACMP cluster.

	Step	Details on page...
1	Ensure that Tivoli software is installed and running on the TMR node and on cluster nodes. **Note:** If you are doing a fresh installation of Tivoli, see the steps below related to creating an alias to each node's standby adapter. You may want to perform these steps as you install Tivoli to avoid unnecessary work later.	B-4
2	On the TMR, create a Policy Region and Profile Manager for HACMP monitoring.	B-4
3	Define the cluster nodes as Tivoli clients (managed nodes)	B-5
4	Define other necessary managed resources	B-4
5	Subscribe the cluster nodes to the Tivoli Profile Manager.	B-5
6	If you haven't done so already, install the HACMP software, and install the three **cluster.hativoli** filesets on the TMR and the cluster nodes.	B-5
7	If you haven't done so already, configure the HACMP cluster and synchronize.	B-5
8	Define the IP address alias and enter it in the **/etc/hosts** file on each node. **Note:** At this point, you may need to change the IP address of the TMR so that the TMR can communicate with the alias IP address on the cluster nodes. Refer to your Tivoli documentation or customer support for additional help.	B-6
9	Verify that the Tivoli **/etc/wlocalhost** file exists and add the IP address alias to it.	B-6
10	Create the **ipaliases.conf** file and include the network name connecting Tivoli with the cluster, and the name of each cluster node with its alias label.	B-6
11	Run the **ifconfig** command to define the alias to network interface.	B-6
12	Start the Tivoli **oserv** process on each node.	B-6
13	If you have prior customizations of node properties in Tivoli, make sure they are saved.	B-6
14	Run the **/usr/sbin/hativoli/bin/install** script.	B-7
15	Re-synchronize cluster resources from the same node you used in the previous step.	B-8

	Step	Details on page...
16	Run the **/usr/sbin/hativoli/AEF/install** script on the TMR.	B-7
17	Run the **/usr/sbin/hativoli/bin/install_aef_client** script on all cluster nodes.	B-7
18	Start cluster services on the cluster nodes.	B-7
19	Start Tivoli on the TMR node (if not already running).	B-7

The following sections provide further details about each of the installation steps.

Ensuring the Required Tivoli Software

The following Tivoli software must be installed before installing the Tivoli-related HACMP filesets:

- Tivoli Framework 3.6 (on TMR and cluster nodes)
- Tivoli Application Extension Facility (AEF) 3.6 (on TMR only)
- Tivoli TME 10 Distributed Monitoring 3.5 (on TMR and cluster nodes)
- Tivoli TME 10 Distributed Monitoring 3.5.1 (on TMR and cluster nodes)

Creating a Cluster Policy Region and Profile Manager

The first step is to create a Policy Region and Profile Manager to handle the HACMP cluster information.

Consult your Tivoli documentation or online help if you need instructions for performing these Tivoli tasks.

Defining HACMP Cluster Nodes as Tivoli Managed Nodes

You also must configure each HACMP cluster node as a subscriber (client) node to an HACMP Profile on the Tivoli Management Region (TMR). Each configured node is then considered a “managed node” that appears in the Tivoli Policy Region window. Each managed node maintains detailed node information in its local Tivoli database, which the TMR accesses for updated node information.

Note that since the TMR does not recognize HACMP automatically, you must enter the name of an adapter known to the cluster node you are defining as a client. Do this in the Add Clients window.

Note: If you already have Tivoli configured, and want to configure IP address takeover (without a separate dedicated network), remember that you must configure an alias to the standby adapter of each cluster node. In addition, you must change the IP address of the TMR node to match the alias IP address you assigned for the standby adapter.

Follow the procedure you would follow to install any nodes for Tivoli to manage. Refer to Tivoli documentation and online help for instructions.

Defining Administrators

Define the cluster nodes as Login Names in the Administrators screen. Consult your Tivoli documentation or online help if you need instructions for performing Tivoli tasks.

Defining Other Managed Resources

At this stage, you define some other resources to be managed in addition to the cluster nodes, such as the profile manager and indicator collection, as follows:

1. In the TME Desktop initial window, click on the newly-created policy region.

   The Policy Region window appears.

2. From the Policy Region window, select **Properties** > **Managed Resources.**

   The Set Managed Resources window appears.

3. From the Available Resources list, double-click on the following items to move them to the Current Resources list:

   - **ManagedNode**
   - **IndicatorCollection**
   - **ProfileManager**
   - **SentryProfile**
   - **TaskLibrary**

4. Click **Set & Close** to continue.

Adding Nodes as Subscribers to the Profile Manager

1. Double-click on the new Profile Manager icon.

   The Profile Manager window appears.

2. Select **ProfileManager > Subscribers...**

3. In the Subscribers window, move your cluster node names from the Available to become Subscribers list to the Current Subscribers list.

4. Click **Set Subscriptions & Close**.

5. Return to the main TME Desktop window.

Installing the HACMP Cluster Monitoring (hativoli) Filesets

Your HACMP software includes three Tivoli-related optional filesets named **cluster.hativoli.client**, **cluster.hativoli.server**, and **cluster.msg.en_US.hativoli** that you can select in the SMIT Install screen. Verify that these are installed on both the Tivoli server node and the HACMP cluster nodes.

Configure the HACMP Cluster

You must have your HACMP software installed and the cluster configured at this point.

Setting up IP Aliasing and Checking the /etc/hosts and /etc/wlocalhost Files

If you do not have a separate dedicated network connecting Tivoli to the cluster nodes, you must define an alias to the IP address of each cluster node's standby adapter with a netmask that matches that of the TMR adapter. This ensures that IPAT activity can be accurately monitored by Tivoli. Make sure this alias is included in both the **/etc/hosts** file and the Tivoli **/etc/wlocalhost** files. Note that if the **/etc/wlocalhost** file was not created earlier in Tivoli, you must create it now.

Note that you may need to change the IP address of the TMR so that it has the same subnet as the cluster nodes' IP address aliases.

Refer back to the section Subnet Considerations for Cluster Monitoring with Tivoli on page B-2 for more details and an example showing what the IP addresses might look like.

To create the alias, use the ifconfig command as follows:

```
ifconfig <standby interface> alias <alias name> netmask <subnet mask>
```

where the *standby interface* is the interface (e.g. en2) of the adapter over which you want to communicate with Tivoli, and *alias* is a chosen IP address or host name (e.g. NodeA_alias).

Creating the ipaliases.conf File

If you are using IPAT without a dedicated network, you must create a file called **/usr/sbin/hativoli/ipaliases.conf** and copy it to each cluster node. This file must contain the network name you will be using for the IP aliasing, the name of each cluster node with its alias label. For example:

```
network=token21
node1 node1_alias
node2 node2_alias
node3 node3_alias
```

Starting the oserv Process

Start the Tivoli **oserv** process on all nodes. Note that the **oserv** process will not start if the alias IP address is not configured.

Note: The Tivoli **oserv** process must be running at all times in order to update the cluster information accurately. It is recommended that you set up a way to monitor the state of the **oserv** process.

To start **oserv**, run the following command on each node:

```
/etc/Tivoli/oserv.rc start
```

Saving Prior Node Properties Customizations

If you previously customized the node properties displayed in the Tivoli Cluster Managed Node window, they will be lost when the **hativoli** scripts are installed.

HACMP automatically saves a copy of your parent dialog. If you need to restore earlier customizations, find the saved file in **/usr/sbin/hativoli/ParentDialog.dsl.save**.

Running the hativoli Install Scripts

You now run three additional install scripts as follows. Note the node(s) on which you run each script, and note that you must synchronize cluster resources after step one.

1. Run **/usr/sbin/hativoli/bin/install** *on any ONE cluster node*

   You are prompted to select the Region, the Profile Manager, and the Indicator Collection, which you set up earlier on the TMR.

   There is a delay of up to 10 minutes while the system creates and distributes profiles and indicators, and adds custom post-events to your HACMP configuration on this node.

2. *Important:* From the same node, re-synchronize cluster resources so that the custom post-event scripts are added to all cluster nodes.

3. Run /**usr/sbin/hativoli/AEF/install** *on the TMR node*

4. Run **/usr/sbin/hativoli/AEF/install_aef_client** *on ALL cluster nodes*

Added Post-event Scripts

Running the first **hativoli** install script and then synchronizing cluster resources automatically adds on each node a set of post-event scripts. These scripts are necessary to enable IP address takeover to work as needed. Post-events are added to the following HACMP events:

- swap_adapter_complete
- node_down_complete
- node_up_complete
- reconfig_resource_complete
- reconfig_topology_complete
- fail_standby

The post-event scripts will be appended to any other post-event scripts you may have configured previously for these events.

Starting Cluster Services

Start cluster services on each cluster node.

Starting Tivoli

If Tivoli is not already running, start Tivoli by performing these steps on the TMR node

1. Make sure access control has been granted to remote nodes by running the **xhost** command with the plus sign (+) or with specified nodes. This will allow you to open a SMIT window from Tivoli.

   If you want to grant access to all computers in the network, type:

   ```
   xhost +
   ```

   or, if you want to grant access to specific nodes only:

   ```
   xhost <computers to be given access>
   ```

2. Also to ensure later viewing of SMIT windows, set DISPLAY=<*TMR node*>.

3. Run the command **. /etc/Tivoli/setup_env.sh** if it was not run earlier.
4. Type **tivoli** to start the application.

The Tivoli graphical user interface appears, showing the initial TME Desktop window.

Note that there may be a delay as Tivoli adds the indicators for the cluster.

Deinstalling Cluster Monitoring with Tivoli

To discontinue cluster monitoring with Tivoli, you must perform the following steps to delete the HACMP-specific information from Tivoli.

Perform the following steps:

1. Run a deinstall through the SMIT interface, deinstalling the three **hativoli** filesets on all cluster nodes and the TMR.
2. If it is not already running, invoke Tivoli on the TMR:
    1. type **. /etc/Tivoli/setup_env.sh**
    2. Type **tivoli**
3. In the Policy Region for the cluster, go to HATivoli Properties.
4. Select the Modify Properties task.

    A window appears containing task icons.
5. Choose **Edit > Select All** to select all tasks, and then **Edit > Delete** to delete.

    The Operations Status window at the left shows the progress of the deletions.
6. Return to the Properties window and delete the Modify Properties task icon.
7. Open the Profile Manager.
8. Choose **Edit > Profiles > Select All** to select all HACMP Indicators.
9. Choose **Edit > Profiles > Delete** to delete the Indicators.
10. Unsubscribe the cluster nodes from the Profile Manager:
    1. In the Profile Manager window, choose Subscribers.
    2. Highlight each HACMP node on the left, and click to move it to the right side.
    3. Click **Set & Close** to unsubscribe the nodes.

Where You Go From Here

If the installation procedure has been completed successfully, Tivoli can now begin monitoring your cluster.

See Chapter 35, Monitoring an HACMP Cluster, in the *HACMP for AIX Administration Guide* for information on monitoring your HACMP cluster through the Tivoli management console.

Appendix C Image Cataloger Demo

This appendix describes the HACMP for AIX, Version 4.4 Image Cataloger demo. This demo should help you to better understand the HACMP for AIX software by allowing you to observe and demonstrate HACMP for AIX operations.

Note: The HACMP for AIX Image Cataloger demo is not supported on RS/6000 SP systems.

What is the Image Cataloger?

The Image Cataloger is an X Window System application that demonstrates client/server computing, high availability, and the functionality of the HACMP for AIX Cluster Lock Manager. It provides clients in a highly-available, clustered environment access to digital images through a graphical user interface, and it guarantees resource data integrity through the services of the Cluster Lock Manager.

Image Cataloger Operation

The Image Cataloger contains two programs that define its operation and demonstrate high availability and the functionality of the HACMP for AIX Cluster Lock Manager: Image Cataloger (**imcat**) and Image Server (**imserv**).

The **imcat** program runs on HACMP for AIX clients. It requests digital images in a cluster’s image library from the **imserv** program and displays them on X Window System clients. The **imcat** program predetermines each image’s display time and requests another image only after this time elapses.

The **imserv** program runs on HACMP for AIX server nodes. It responds to image requests from client **imcat** programs and in turn requests locks on images from the Cluster Lock Manager daemon (**cllockd**) so that no two clients can access images simultaneously. The **imserv** program returns the requested image to the **imcat** program once an image lock is granted.

System Requirements

To run the Image Cataloger in an HACMP cluster, you need:

- Two HACMP for AIX servers
- An HACMP for AIX client (X Window capable)
- HACMP for AIX system software
- Image Cataloger software (installed or on diskette).

Before installing the Image Cataloger on a server, be sure that you have at least 1.3 MB of available disk space. Also, be sure that cluster hardware and the HACMP for AIX software are operational and configured correctly before running the demo.

Starting the Image Cataloger

This section describes how to start the Image Cataloger on both a server and a client. Before starting the demo, ensure that the HACMP for AIX software is installed correctly and running with the Cluster Lock Manager enabled on all servers and **clinfo** running on clients.

Starting the Demo from a Server

To start the Image Cataloger demo server process:

1. Register the Image Cataloger demo.
2. Configure the demo in a resource group.
3. Verify Object Data Manager (ODM) definition changes.
4. Enable the Cluster Lock Manager option.

Registering the Demo

Before the Image Cataloger can display images, you must register the demo as an application server.

To register the demo:

1. Enter:

   ```
   smit hacmp
   ```
2. Select **Cluster Configuration** and press Enter.
3. Select **Cluster Resources** and press Enter to display the Cluster Resources screen.
4. Select **Define Application Servers** from the Cluster Resources menu and press Enter.
5. Select **Add an Application Server** and press Enter.
6. Specify a server name and paths for both the demo start and stop scripts on the Add an Application Server and press Enter. The screen shows the default start and stop scripts. Note the possible options to the start and stop scripts listed below.

   The scripts search the IMSERV_IMAGE_LOCATION environment variable for the directory containing the images. Images are stored by default in the **/usr/lpp/cluster/samples/demos/image/images** directory.

   If you designate another directory for storing images, you must specify the **-d** option with the name of the directory as an option to the start and stop scripts.

   Specify the service IP address (IP label) with the **-a** option for the server running the demo. IP address takeover will not work correctly if this option is not specified.

   An example of the start script using both options is shown below:

   ```
   /usr/sbin/cluster/events/utils/start_imagedemo -d mydir -a jim_svc
   ```

   The stop script allows one option, the **-a** *service_address* option. If the start script uses this option, the same address should be specified with the stop script if you only want this instance stopped. Otherwise, all instances of the image_demo will be stopped.
7. Press F3 until you return to the HACMP for AIX menu.

Configuring the Demo

After you register the demo as an application server, configure the demo in a resource group and define the node on which it will run to have the highest priority.

To configure the demo in a resource group:

1. Select **Cluster Configuration** from the HACMP for AIX menu and press Enter.
2. Select **Cluster Resources** and press Enter.
3. Select **Define Resource Groups** and press Enter.
4. Select **Add a Resource Group** and press Enter.
5. Enter the field values as follows:

   Resource Group Name Enter the name image_demo.

   Node Relationship Toggle the entry field to **cascading**.

   Participating Node Names Enter the names of the nodes that you want to be members of the resource chain for this resource group. Enter the node names in order from highest to lowest priority (left to right). Leave a space between node names.

6. Press Enter to add the Resource Group information to the ODM.
7. Press F3 until you return to the Cluster Resources screen after the command completes.
8. Select **Change/Show Resources/Attributes for a Resource Group** and press Enter to display a list of defined resource groups.
9. Select the image_demo resource group and press Enter.

   SMIT returns the following screen with the **Resource Group Name**, **Node Relationship**, and **Participating Node Names** fields filled in.

   If the participating nodes are powered on, you can press F4 to get a listing of shared resources. If a resource group/node relationship has not been defined, or if a node is not powered on, F4 displays the appropriate warnings.

10. Enter the field values as follows:

    Resource Group Name Reflects the choice you made on the previous screen. The resource group to configure.

    Node Relationship Reflects the fallover strategy entered when you created the resource group.

    Participating Node Names Reflects the names of the nodes that you entered as members of the resource chain for this resource group. Node names are listed in order from highest to lowest priority (left to right), as you designated them.

    Service IP Label If IP address takeover is being used, list the IP label to be taken over when this resource group is taken over. Press F4 to see a list of valid IP labels. These include addresses which rotate or may be taken over.

Filesystems

The filesystems included in this resource group. Press F4 to see a list of the filesystems.

Filesystems Consistency Check

Identify the method of checking consistency of filesystems, **fsck** (default) or **logredo** (for fast recovery).

Filesystems Recovery Method

Identify the recovery method for the filesystems, **parallel** (for fast recovery) or **sequential** (default).

Filesystems/Directories to Export

Identify the filesystems or directories to be exported. The filesystems should be a subset of the filesystems listed above. The directories for export should be contained in one of the filesystems listed above. Press F4 for a list.

Filesystems/Directories to NFS Mount

Identify the filesystems or directories to NFS mount. All nodes in the resource chain will attempt to NFS mount these filesystems or directories while the owner node is active in the cluster.

Network for NFS Mount

(This field is optional.)

Choose a previously defined IP network where you want to NFS mount the filesystems. The F4 key lists valid networks.

This field is relevant only if you have filled in the previous field. The **Service IP Label** field should contain a service label which is on the network you choose.

Note: You can specify more than one service label in the **Service IP Label** field. It is highly recommended that at least one entry be an IP label on the network chosen here.

If the network you have specified is unavailable when the node is attempting to NFS mount, it will seek other defined, available IP networks in the cluster on which to establish the NFS mount.

Volume Groups

The shared volume groups that should be varied on when this resource group is acquired or taken over. Press F4 to see a list of shared volume groups. If you are using raw logical volumes in non-concurrent mode, you only need to specify the volume group in which the raw logical volume resides in order to include the raw logical volumes in the resource group.

Concurrent Volume Groups

Volume groups that can be concurrently accessed if you configure the demo for concurrent mode.

Raw Disk PVIDs

Press F4 for a listing of the PVIDs and associated hdisk device names.

If you are using an application that directly accesses raw disks, list the raw disks here.

AIX Connections Services

For systems that have AIX connection services, PC connectivity. Not applicable to setting up the Image Cataloger.

AIX Fast Connect Services
Provides file and print services to PCs. Not applicable to setting up the Image Cataloger.

Application Servers
Enter the image_demo as the application server to include in the resource group.

Highly Available Communication Links
For systems that have SNA. Not applicable to setting up the Image Cataloger.

Miscellaneous Data
A string you want to place into the environment along with the resource group information. It is accessible by the scripts. For example, *Imagedatabase*.

Inactive Takeover
Leave this variable set to FALSE.

Cascading without Fallback Enabled
Leave this variable set to FALSE.

9333 Disk Fencing Activated
Leave this variable set to FALSE unless using the disk subsystem in concurrent mode.

SSA Disk Fencing Activated
Leave this variable set to FALSE unless using the disk subsystem in concurrent mode.

Filesystems Mounted Before IP Configured
This field specifies whether, on fallover, HACMP takes over volume groups and mounts filesystems before or after taking over the failed node's IP address or addresses.

The default is **false**, meaning the IP address is taken over first. Similarly, upon reintegration of a node, the IP address is acquired before the filesystems.

Set to **true** if the resource group contains filesystems to export. This is so that the filesystems will be available once NFS requests are received on the service IP address.

11. Press Enter to add the values to the HACMP for AIX ODM.
12. Press F3 until you return to the HACMP for AIX menu.

Verifying ODM Definition Changes

Whenever the node environment and resource configuration changes, information in the HACMP for AIX ODM reflects the changes.

To verify that you have configured the Image Cataloger demo:

1. Select **Cluster Configuration** from the HACMP for AIX menu and press Enter.
2. Select **Cluster Resources** and press Enter.
3. Select **Show Cluster Resources** and press Enter.
4. Select **Show Resource Information by Node** and press Enter. SMIT displays a list.
5. Select the image_demo resource group and press Enter. The COMMAND STATUS screen appears.

6. Press F3 until you return to the HACMP for AIX menu, or press F10 to exit SMIT.

Note: The first group of fields is repeated for each resource group before the Run Time Parameters and remaining fields appear.

Enabling the Cluster Lock Manager Option

Before the Image Cataloger can display images on servers or on clients, the HACMP for AIX software must be started and each node's cluster lock services must be enabled. Once enabled, the Cluster Lock Manager can grant **imserv** requests for locks on images.

If you specify an incorrect image directory in your node environment or you do not enable the Cluster Lock Manager daemon (**cllockd**), **imserv** fails to start. You cannot run the Image Cataloger.

To enable a node's cluster lock services when starting the HACMP for AIX software:

1. Select **Cluster Services** from the HACMP for AIX menu and press Enter.
2. Select **Start Cluster Services** and press Enter.
3. Select **Startup Cluster Lock Services?** and press Tab to toggle this option to **true**.
4. Select **Startup Cluster Information Daemon?** and press Tab to toggle this option to **true**. Starting the clinfo daemon allows it to monitor cluster status information.
5. Press Enter after setting the options. The COMMAND STATUS screen appears. It displays OK when options have been enabled; otherwise it displays Failed.
6. Press F3 until you return to the HACMP for AIX menu.

Starting the Demo from a Client

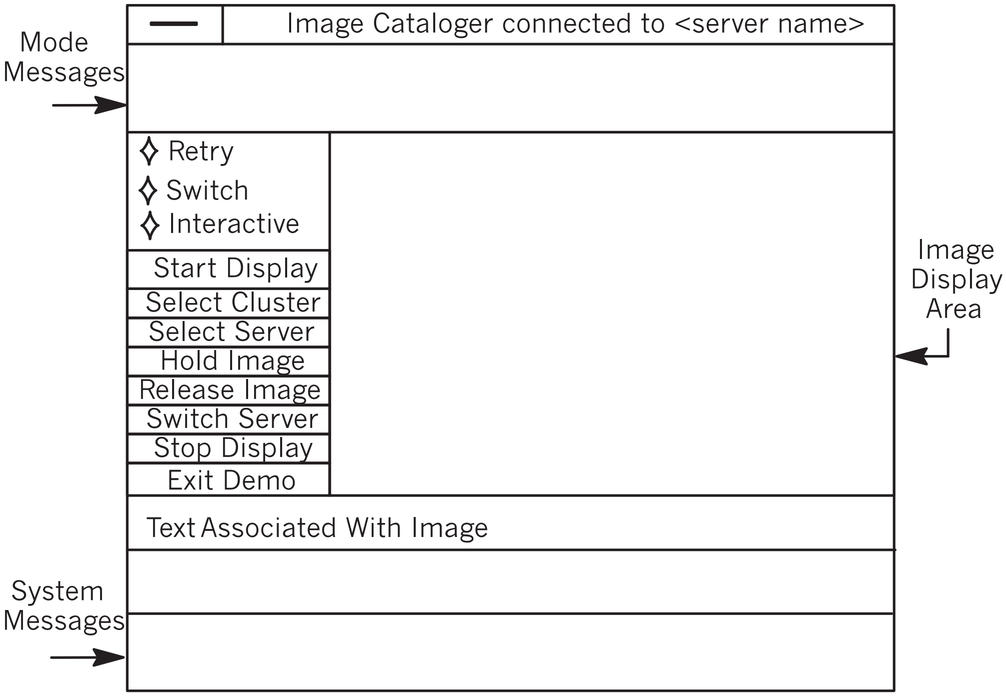
Once the HACMP cluster is up and stable and the application server containing the **imserv** program is running, you can start the Image Cataloger on a client.

To start the Image Cataloger on a client:

1. Start AIXwindows.
2. Use the **cd** command to change to the following directory:

   ```
   cd /usr/sbin/cluster/demos/image
   ```
3. Enter the following command:

   ```
   ./imcat
   ```
   The Image Cataloger window appears.



You can now use the Image Cataloger to select configuration options and control buttons that determine how the Image Cataloger displays images.

Using the Image Cataloger

This section describes the Image Cataloger recovery options and control buttons, gives steps for displaying images, and shows you how to create or convert images using the **/usr/sbin/cluster/demos/image/imcreate** and **/usr/sbin/cluster/demos/image/imconvert** utilities.

Selecting Recovery Options

The following options let you configure the Image Cataloger for fallover situations.

Retry

The Retry option causes a client **imcat** program to attempt to reconnect to a failed node's IP address after a successful fallover. If the Internet Protocol address takeover (IPAT) option is enabled on HACMP for AIX cluster nodes, **imcat** will reconnect to the designated IP address. The **imcat** program tries to reconnect several times before returning to its initial display state. Once the fallover node restarts the failed node's **imserv** program, **imcat** programs on clients can again display images.

Switch

Should a node in the cluster fail, the Switch option automatically connects a client's **imcat** program to the **imserv** program on the next available node. If **imserv** is not running on other nodes, the **imcat** program returns to its initial blank display state.

Interactive

The **Interactive** option lets you select either the Retry or Switch options if you did not choose a recovery option before a node failure occurred.

Selecting Control Buttons

The following buttons let you control the Image Cataloger. See the next section to begin displaying images.

Start Display

Click this button after you select a cluster and server. This button causes the Image Cataloger to start displaying images from the server's image directory.

Select Cluster

Click this button to cause the *clusterselect_popup* menu to appear. In the popup menu, you can select a cluster from a list of available clusters displayed. Click OK to confirm your selection and close the popup menu.

Select Server

Click this button to cause the *connecthostselect_popup* menu to appear. In the popup menu, you can select a server to connect to from the list of available servers displayed. Click OK to confirm your selection and close the popup menu.

Hold Image

Click this button to hold a displayed image in an X Window. Holding an image denies other clients access to the image. You can edit a text associated with the graphic if necessary.

Release Image

Click this button to release the currently held image so that other **imcat** programs can access it. After releasing the image, text changes are written to the .text file associated with the image.

Switch Server

Click this button to select the next available server in the cluster. The Switch Server button lets you change servers at any time.

Stop Display

Click this button to cause the Image Cataloger to stop displaying images.

Exit Demo

Click this button to exit the Image Cataloger demo.

Displaying Images

To display images using the Image Cataloger:

1. Select a recovery option (Retry or Switch) for use with HACMP for AIX fallover configurations.
2. Click Select Cluster to display a list of available clusters. A popup window appears.
3. Select a cluster, then click OK.
4. Click Select Server to display a list of available nodes in the cluster. A popup window appears.
5. Select a server, then click OK.
6. Click Start Display to begin displaying images.
7. Click Stop Display to stop displaying images.

Creating and Converting Images

This section describes two Image Cataloger utilities:

- **/usr/sbin/cluster/demos/image/imcreate**
- **/usr/sbin/cluster/demos/image/imconvert**.

The **imcreate** utility lets you create demo images in a Journaled Filesystem (JFS or FILE) access format, and the **imconvert** utility lets you create images in a raw logical volume (RAW) access format and convert between FILE and RAW formats.

Using the imcreate utility

The **imcreate** utility lets you change an X11 (**.xbm**) bitmap image into a FILE formatted Ximage that the **imserv** program can access and return to the **imcat** program for display.

To create an Ximage:

1. Locate or obtain an image in X11 bitmap format. This step may require assistance because various methods exist for creating an **.xbm** file.

   Because the **.xbm** file must be 392 pixels wide to display properly, you should know the bitmap format information and pixel dimensions for each image before using **imcreate**. Read the file’s first line to verify the pixel width. If the file is more than 392 pixels long, **imcreate** truncates additional pixels.
2. Start AIXwindows.
3. Enter the following command:

   ```
   imcreate filename.xbm
   ```
   where *filename.xbm* is the name of the X11 bitmap file with its **.xbm** extension. A window containing the image appears.
4. Review the image for clarity, then click anywhere inside the window to create an Ximage. The **imcreate** utility creates an Ximage consisting of the following files before returning you to the shell prompt: *filename.data*, the actual image data; and *filename.args*, a binary image description. If you add text associated with an image in the Image Text Window during display, a *filename.text* file is added to the group of files.

Using the imconvert utility

This **imconvert** utility lets you convert an Ximage to or from a RAW or FILE access format. You must use this utility to create and store images on a raw logical volume.

RAW - A RAW format merges Ximage files into one file on a raw logical volume. The **imconvert** utility converts Ximages to this format to store them. The **imserv** program references an **images.dir** index file of all images when the **imcat** program requests an image for display. The index file is on the raw logical volume which is identified by the IMSERV_IMAGE_LOCATION environment variable in concurrent mode operations.

FILE - A FILE format groups Ximage files (.args, .data, and .text) as separate files in a directory. The **imconvert** utility converts Ximages to this format to store them in a filesystem on disks. An index of all images is stored in an **images.dir** file. This file is identified by the IMSERV_IMAGE_LOCATION environment variable when the **imserv** program requests an image for display.

Note: When designating a raw logical volume, be sure to indicate the “r” prefix in the IMSERV_IMAGE_LOCATION. For example, if you create a raw logical volume named *logv*, the IMSERV_IMAGE_LOCATION environment variable will be **/dev/rlogv**.

Command Options and Syntax

The following table describes options you can use with the **imconvert** command.

Option	Description
-d *level*	Defines the level of debug output you can see. A higher number means more printed output.
-i *leve*l	Defines the level of informational output to be shown. A higher number means more information.
-l	Lists all images in the image directory.
-f *image_location*	Defines the source (from) location of images to be converted. The source location can be RAW or FILE.
-t *image_location*	Defines the target (to) location for storing converted images. The target location can be RAW or FILE.

Option	Description
-a *image_name*	Finds and adds an image to the image directory. If the image location is in FILE format, **imconvert** searches for the ARGS, DATA, and TEXT files before adding an entry to the **images.dir** file. If the image location is in RAW format, imconvert searches the directory for a deleted image that matches *image_name* and clears its deleted state.
-r *image_name*	Deletes (removes) an image from the image directory. If the image location is RAW, **imconvert** marks the image as deleted, creating a hole in the directory that can be filled using the Pack option.
-p	Packs the image directory, removing holes in a RAW image directory. This action removes deleted images from the file. Deleted images cannot be recovered after a pack operation.

Use the following syntax with the **imconvert** command.

To list images in the **images.dir** file, enter:

```
imconvert -l image_location
```

To convert an Ximage in the default image directory to a RAW format and store them on the raw logical volume, for example */dev/rv1,* enter:

```
imconvert -f image_location -t /dev/rv1
```

To convert images from one format to another (RAW and FILE), enter:

```
imconvert [-d LEVEL -i LEVEL] -f image_location -t image_location
```

To operate on a set of images (RAW and FILE), enter:

```
imconvert [-d LEVEL -i LEVEL] [-l] [-a image] [-r image] image_location
```

To check on a set of images after conversion to either the RAW or FILE format, enter:

```
imconvert -l [image_location]
```

Note: To display converted images stored in your images directory, be sure to add the name of each image to the **images.dir** file using a text editor. To prevent a particular image from displaying, delete it from this file.

Observing HACMP for AIX Operations Using the Image Cataloger

This section describes HACMP cluster operations you can observe using the Image Cataloger.

Node Failure

To demonstrate node failure and IP address takeover, run two Image Catalogers on a client in “Retry” mode, accessing image servers on separate nodes in a cluster configured for IPAT. When one node fails, both **imcat** programs momentarily halt. After the surviving node processes the fallover, **imcat** programs become active again, accessing the image server running on the surviving node. They both can display images again.

Note: Because IPAT causes a surviving node to masquerade as the failed node when a fallover occurs, the label in the X-Window continues to indicate a connection to the failed node.

Fallover and Disk Takeover

To demonstrate fallover and disk takeover, be sure that:

1. The HACMP for AIX software has been started.
2. Each client is running **clinfo**, **clstat**, and the **imcat** programs.

After the **imcat** programs start, enter a **halt -q** or similar command on one of the servers. The Image Cataloger stops and waits for an image lock. In the **clstat** window, Clinfo indicates that a server node has failed. Notice the output of **/tmp/hacmp.out** on the surviving node as disk takeover occurs. If the Image Cataloger is set for the **SWITCH** option, the Image Cataloger continues to display images after a delay. On the surviving node, enter **df** to ensure that disk takeover occurred. Use the **netstat -in** command to see that the standby adapter is masquerading as the failed node (if you enabled IP address takeover).

Reboot the failed server and start the HACMP for AIX software. Once the Cluster Manager and associated scripts have run, click the **SWITCH SERVER** button and re-attach the Image Cataloger to the previously failed node.

Reintegration

When a failed node rejoins a cluster, the **imcat** program connected to the surviving node hesitates while the cluster stabilizes. After a failed node reintegrates into a cluster, the **imserv** program restarts on that node and the Image Catalogers on both servers continue to display images. If you are in **RETRY** mode with IPAT, the program switches back to the previously failed node. If you are in **SWITCH** mode, the program continues to run on the takeover node.

Cluster Locking

To demonstrate HACMP for AIX cluster locking features, click the **Hold Image** button on one of the catalogers while an image is present. Note that although the Image Cataloger initially displays images randomly, the sequence followed after you click this button becomes the same for both servers. When an **imcat** program (other than the one displaying the image) attempts to access the held image, access is denied until you click the **Release Image** button.

Concurrent Access

The **imconvert** utility, together with **imserv's** ability to access raw logical volumes, lets you access images on nodes in a concurrent access environment. This HACMP for AIX functionality is referred to as a concurrent access operation.

Note: To demonstrate concurrent access, you must use a disk technology that supports concurrent access. See the chapter on planning shared disk devices in the *HACMP for AIX Planning Guide* for a list of supported disks.

To see if the concurrent access functionality is installed on the system, enter:

```
lslpp -al cluster.clvm
```

If the software is installed, the system will list

```
cluster.clvm
```

Note: You must configure the **image_demo** as a concurrent resource group before demonstrating concurrent access. Be sure to have the image data on a raw logical volume for concurrent access.

To demonstrate concurrent access, run two **imcat** programs on a client. Each program should be connected to separate servers. After the programs start, click the **Hold Image** button in one of the Image Cataloger windows. The **imcat** program on the other server continues to run. When the active program attempts to display the image being held by the other **imcat** program, it waits until the lock on the image is released before displaying it.

To show that the same data area is being accessed, try adding text to the held image. When the image is released, the other **imcat** program is allowed to display the image with the added text.

Appendix D CLMarket Demo

This appendix describes the HACMP for AIX CLMarket demo. This demo should help you to better understand the HACMP for AIX software by allowing you to observe and demonstrate HACMP for AIX operations.

What is CLMarket?

The CLMarket demo is an X Window System-based application that demonstrates HACMP for AIX functionality within a concurrent access environment. The CLMarket demo simulates a supermarket's cash register, allowing you to perform purchase transactions on a shared product database. Using CLMarket, you can show how networks can be configured for recovery should a transaction fail to complete or should a networked server fail altogether.

Note: To demonstrate concurrent access, you must use a disk technology that supports concurrent access. See the chapter on planning shared disks in the *HACMP for AIX Planning Guide* for a list of supported disks.

CLMarket Operation

The CLMarket demo consists of two applications: a **checkout client** ("front end") and a **market server** ("back end").

The **checkout client** runs on networked client systems and communicates with the cluster-based **market server** over a local area network. When a **checkout client**, which simulates a cash register, requests information on a specific product, the **market server**, which manages a product database containing records for each product in a store's inventory, satisfies the client's request to look up a product in this inventory and buy it. Each database record contains the following product information:

- Universal Price Code (UPC)
- Description
- Price
- Quantity.

The **checkout client** lets you read this information and initiate purchase requests, each of which can be handled by a **market server** application running on a cluster node. The **market server** provides the client application with a library of Remote Procedure Calls (RPCs) to access the product database and to decrement the quantity of a particular item. The RPCs communicate with the HACMP for AIX Cluster Lock Manager (CLM) on the primary node, the first node in the cluster that becomes the resource owner, to serialize read and write access to the shared disk through the **cllockd** daemon.

The **cllockd** daemon coordinates requests for shared data in a concurrent access environment and ensures data integrity among server nodes accessing the data. It grants access (lock) permissions to secondary servers attempting to retrieve product data from the disk. If **cllockd** grants a secondary server lock permission to a resource, the product information is retrieved from the database so that the **market server** can pass it to the **checkout client** for display.

Together the **checkout client** and the **market server** applications let you demonstrate the concurrent access of raw logical volumes on a shared disk.

System Requirements

To run the CLMarket demo in an HACMP cluster, you need:

- An HACMP cluster
- An HACMP for AIX client (X Window System and server capable)
- HACMP for AIX software (both tapes).
- CLMarket software.

Before installing the CLMarket demo, be sure that you have at least 1.3 megabytes (MB) of available disk space. Also, be sure that cluster hardware and the HACMP for AIX software are operational and configured correctly before running the demo.

Starting the CLMarket Demo

This section describes the steps necessary for starting the CLMarket demo on both a server and a client.

Starting the Demo from a Server

Starting the CLMarket demo from a server requires that you complete the following steps:

1. Create two logical volumes, **lvmarket** and **lvlogmarket**, on a concurrent volume group.
2. Initialize the logical volumes.
3. Load the CLMarket product database.
4. Register the demo.
5. Configure the demo.
6. Verify Object Data Manager (ODM) definition changes.
7. Enable the Cluster Lock Manager option.

Each step is described on the following pages.

Creating Logical Volumes

See Defining Shared LVM Components for Concurrent Access on page 6-7 in this guide for instructions on creating shared logical volumes in a concurrent access environment.

As you create logical volumes for use with the CLMarket demo, keep the following information in mind:

- Logical partitions of four MB can contain up to 65,536 product records, each 64 bytes in size.
- The CLMarket demo provides the following default names for logical volumes:
  - `DBASE_FILE = "/dev/rlvmarket"`
  - `LOG.FILE = "/dev/rlvlogmarket"`

The DBASE_FILE name represents the logical volume used as the product database, and the LOG.FILE name represents the logical volume where the integrity of product database information is maintained.

Using the default names for initializing the logical volumes, for loading the CLMarket product database, and for registering and configuring the CLMarket demo makes running the demo easier. If you choose to define other logical volume names, be sure to modify the start and stop scripts before registering the demo. Also, be sure to use logical volume names consistently when completing the remaining steps in this appendix.

Note: If you define names other than the default names for logical volumes, remember to specify the raw logical volume names: rlv*name* and r*loglv*, when modifying the demo start and stop scripts.

Initializing the Logical Volumes

Before you can store data on raw logical volumes, you must initialize them. To initialize the product database and the log file:

1. Use the **cd** command to change to the following directory:

   ```
   cd /usr/sbin/cluster/demos/clmarket
   ```

2. Enter the following command and press Enter:

   ```
   ./marinit /dev/rlvmarket /dev/rlvlogmarket
   ```

A message similar to the following appears after initializing the product database:

```
STATUS: PRODUCT DATABASE INITIALIZED
.
.
.
STATUS: PRODUCT DATABASE CLOSED
.
.
.
STATUS: LOGFILE CLOSED
```

Loading the Product Database

The CLMarket demo comes with a flat ASCII file (**mar.database**) that can be used to represent the product database. This file is stored in the following directory: **/usr/lpp/cluster/samples/demos/clmarket/utils**. You load this file onto the market database (/**dev/rlvmarket**) logical volume.

To load the product database, enter:

```
./mara2db /dev/rlvmarket < /usr/lpp/cluster/samples/demos/clmarket/utils/mar.database
```

where **mara2db** is the command used to convert an ASCII file to a database format.

You can create and load your own ASCII file to represent the product database if you choose not to use the **mar.database** file provided with the CLMarket demo. If you create your own database file, be sure that it conforms to the database format for storing records.

To check the **rlvmarket** database format, use the **mardb2a** command to convert the database to an ASCII display on your terminal. To convert the **rlvmarket** database to an ASCII display, enter:

```
./mardb2a  /dev/rlvmarket | more
```

Registering the Demo

You must register the CLMarket demo as an application server before you can use it to demonstrate the HACMP for AIX concurrent access feature.

To register the demo:

1. Enter:

   ```
   smit hacmp
   ```

2. Select **Cluster Configuration** and press Enter.
3. Select **Cluster Resources** and press Enter to display the Cluster Resources screen.
4. Select **Define Application Servers** from the Cluster Resources menu and press Enter.
5. Select **Add an Application Server** and press Enter.
6. Specify on the Add an Application Server screen a server name and paths for both the demo start and stop scripts and press Enter.

   The scripts search the logical volumes for the default directories containing the product database and the log file. You can optionally specify a different database directory, using the **-d** flag, or a different log file directory, by using the **-l** flag. If you specify either of these optional directories, remember to use raw logical volume names, such as, rlv*mylv* and r*mylogfile*.

7. Press F3 until you return to the HACMP for AIX menu.

Configuring the Demo

After you register the demo as an application server, configure the demo as an owned resource belonging to the node on which it will run.

To configure the demo:

1. From the HACMP for AIX menu, select **Cluster Configuration** > **Cluster Resources > Define Resource Groups > Add a Resource Group** and press Enter.
2. Enter the field values as follows:

   Resource Group Name Enter the name clmarketdemo.

   Node Relationship Toggle the entry field to Concurrent.

   Participating Node Names Enter the names of the nodes that you want to be members of the resource chain for this resource group. Leave a space between node names. Priority is ignored for concurrent resource groups.

3. Press Enter to add the Resource Group information to the ODM.
4. Press F3 after the command completes until you return to the Cluster Resources screen.
5. Select **Change/Show Resources for a Resource Group** and press Enter to display a list of defined resource groups. SMIT displays a list of defined resource groups.
6. Select the clmarketdemo resource group and press Enter.

   SMIT returns the following screen with the **Resource Group Name**, **Node Relationship**, and **Participating Node Names** fields filled in. If the participating nodes are powered on, press F4 to get a listing of shared resources.
7. Enter the field values as follows:

   Resource Group Name Reflects the choice you made on the previous screen. The resource group to configure.

   Node Relationship Reflects the type of information entered when you created the resource group.

   Participating Node Names Reflects the names of the nodes that you entered as members of the resource chain for this resource group. Node names are listed in order from highest to lowest priority (left to right), as you designated them. (Priority does not apply to concurrent node relationships.)

   Service IP Label Ignore for clmarket demo.

   HTY Service IP Label Ignore for clmarket demo.

   Filesystems Ignore for clmarket demo.

   Filesystems Consistency Check Ignore for clmarket demo.

   Filesystems Recovery Method Ignore for clmarket demo.

   Filesystems to Export Ignore for clmarket demo.

   Filesystems to NFS Mount Ignore for clmarket demo.

   Volume Groups Ignore for clmarket demo.

   Concurrent Volume Groups Identify the shared volume groups that can be accessed simultaneously by multiple nodes (*market_vg* in this example).

   Raw Disk PVIDs Ignore for clmarket demo.

   Application Servers Enter “clmarketdemo” as the application server to include in the resource group.

Miscellaneous Data A string you want to place into the environment along with the resource group information. It is accessible by the scripts. For example, *Database1*.

Inactive Takeover Leave this set to FALSE.

9333 Disk Fencing Activated Set this field to **true** when using the disk subsystem in concurrent mode.

SSA Disk Fencing Activated Set this field to **true** when using the disk subsystem in concurrent mode.

Filesystems Mounted Before IP Configured Leave this set to FALSE.

8. Press Enter to add the values to the HACMP for AIX ODM.
9. Press F3 until you return to the HACMP for AIX menu.

Verifying ODM Definition Changes

Whenever the node environment and resource configuration changes, information in the HACMP for AIX ODM reflects the changes.

To verify that you have configured the CLMarket demo:

1. Select **Cluster Configuration** > **Cluster Resources** > **Show Cluster Resources** > **Show Resource Information by Node**.
2. Select the resource group desired from the list.
3. Select **clmarketdemo** and press Enter. The COMMAND STATUS screen appears.
4. Press F3 until you return to the HACMP for AIX menu, or F10 to exit SMIT.

Enabling the Cluster Lock Manager Option

To start the CLMarket demo, you must enable each node’s cluster lock services and start the HACMP for AIX software. Once the services are enabled and the HACMP for AIX software is started, the Cluster Lock Manager grants **checkout client** requests for locks on database records. You cannot run the demo if you do not enable the Cluster Lock Manager daemon (**cllockd**) the **market server** will fail to start.

To enable a node’s cluster lock services when starting the HACMP for AIX software:

1. Select **Cluster Services** > **Start Cluster Services**.
2. Select **Startup Cluster Lock Services?** and press Tab to toggle this option to **true**.
3. Select **Startup Cluster Information Daemon?** and press Tab to toggle this option to **true**. Starting the **clinfo** daemon lets it monitor cluster status information.
4. Press Enter after setting the options. The COMMAND STATUS screen appears. It displays OK when options have been enabled; otherwise it displays Failed.
5. Press F3 until you return to the HACMP for AIX menu, or F10 to exit SMIT.

Starting the Demo from a Client

To start the CLMarket client process:

1. Use the **cd** command to change to the following directory:

   ```
   cd /usr/sbin/cluster/demos/clmarket
   ```

2. Enter the following command:

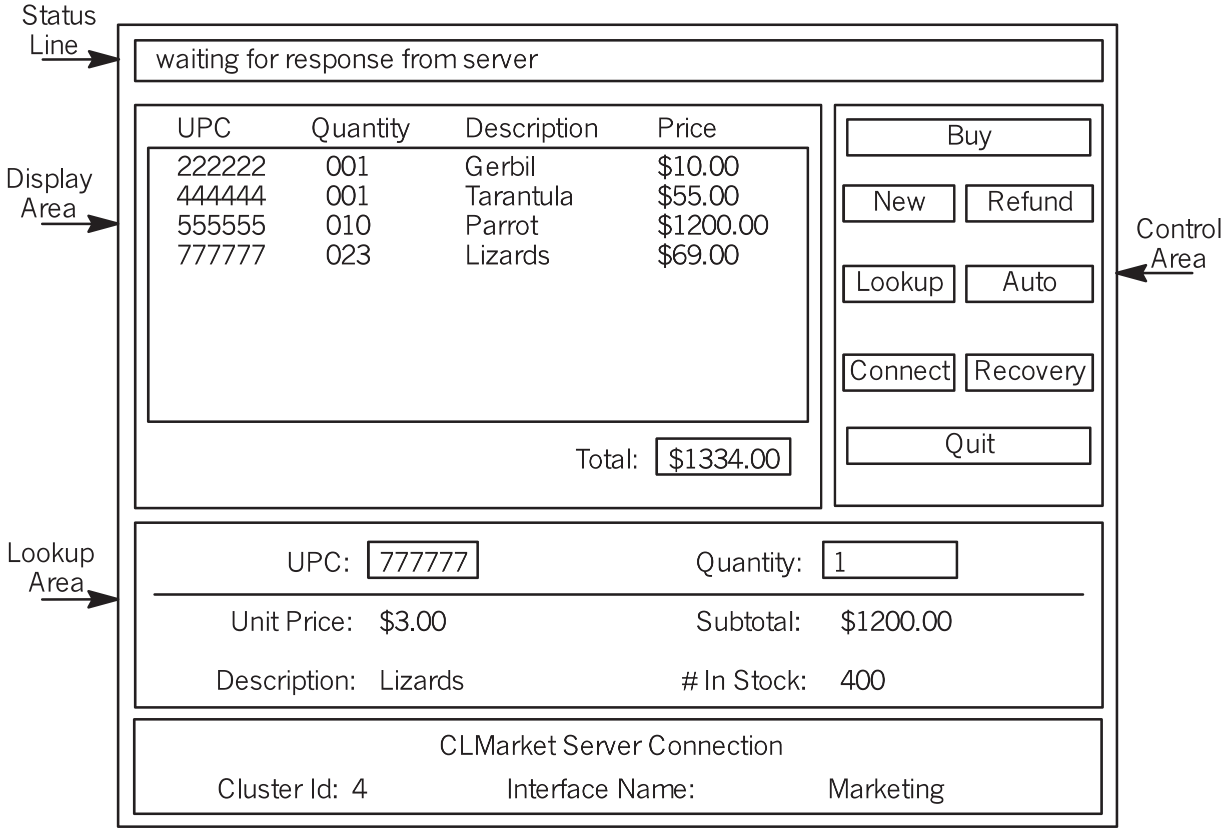
   ```
   ./marxclient
   ```

   The CLMarket demo starts and displays the window shown in the next section.

Using the Checkout Client

This section describes the CLMarket **checkout client** window and its control buttons. As stated earlier, the CLMarket **checkout client** resembles a cash register that you can use to demonstrate the functionality of the HACMP for AIX concurrent access environment.

When you start the CLMarket demo, the following **checkout client** window appears. Use this window to select a recovery option and perform transactions on a shared product database. The **checkout client** window is divided as shown in the following graphic.



waiting for response from server

UPC	Quantity	Description	Price
222222	001	Gerbil	$10.00
444444	001	Tarantula	$55.00
555555	010	Parrot	$1200.00
777777	023	Lizards	$69.00

Total: $1334.00

UPC: 777777 Quantity: 1

Unit Price: $3.00 Subtotal: $1200.00

Description: Lizards # In Stock: 400

CLMarket Server Connection

Cluster Id: 4 Interface Name: Marketing

Status Line

The status line displays messages specific to a client's activities and information about the current state of the server connection. Recovery attempts in progress during a server failure are also displayed.

Display Area

The display area provides an active list of items you purchase from the product database and their costs. Each entry in the list contains a UPC code, a description of the item, the number of items purchased, and a total price for the items purchased.

Lookup Area

The lookup area displays information about a product in the database. This information includes a UPC code, a unit price for the item purchased, a description of the item, the quantity of the item in stock, and a subtotal of all items in stock. The quantity field is used to specify the number of a particular item to buy; for example, how many lizards or gerbils. The field is reset to one when item information is displayed during a lookup.

Control Area

The control area contains buttons you can use to operate the demo. Each button is described in the following table.

Button	Description
Buy	Click this button after entering a UPC and a quantity in the lookup area. This button causes the client to check for a specific item in the product database. If the item exists and the request can be satisfied, the quantity in the database is decremented accordingly, and a new entry describing the purchase is added to the display area. The total field is also updated to reflect any new purchases. Information about remaining items of this kind are displayed in the lookup area.
New	Click this button to clear the display area and total field. Clicking this button does not affect information in the product database.
Refund	Click this button after selecting an item from the display list. This button causes the selected entry to be duplicated in the display area with a minus sign next to the price. The minus sign signifies that the amount for the specific entry has been credited against the total. The quantity in the product database is updated to reflect the credit. The originally selected entry is updated with an **R** beside the price. The **R** signifies that the cost of the item has been refunded.
Lookup	Click this button after entering a valid UPC. This button causes the checkout client to query the database for information about the specified product. The information is displayed in the lookup area at the bottom of the checkout client window.

Button	Description
Auto	Click this button to toggle the CLMarket demo between manual and automatic purchase mode. In automatic purchase mode, the checkout client performs automatic lookups of random items in the product database and initiates purchases of each item. The display area scrolls automatically as purchases are made. The manual purchase mode lets you decide which items the checkout client will query before you purchase the item.
Connect	Click this button to cause a connection popup menu to appear. In the popup menu, you can select a server (containing a product database) to connect to from the list of available servers displayed. Click **OK** to confirm your selection and close the popup menu.
Recovery	Click this button to cause a recovery popup menu to appear. In the popup menu, you can select a recovery option (**Fallover** or **Retry**) to be executed if the market server or the node fails. If you select Fallover (default), the market server switches to another server. If you select Retry, you can specify the time (in seconds) he market server tries to reconnect the failed node or to another node in the cluster. Click OK to confirm your selection and close the popup menu.
Quit	Click this button to exit the CLMarket demo.

Appendix E HACMP for AIX and SNMP Utilities

This appendix discusses the Simple Network Management Protocol (SNMP) and describes the relationship between the HACMP for AIX SNMP-based utilities and other SNMP-based utilities that run on the RS/6000 and SMP platforms.

This guide does not discuss the SNMP standard in depth. See the appropriate AIX documentation for more detailed information about SNMP.

Overview

SNMP is an industry-standard set of standards for monitoring and managing TCP/IP-based networks. SNMP includes a protocol, a database specification, and a set of data objects. A set of data objects forms a Management Information Base (MIB). SNMP provides a standard MIB that includes information such as IP addresses and the number of active TCP connections. The actual MIB definitions are encoded into the agents running on a system. The standard SNMP agent is the SNMP daemon, **snmpd**.

MIB-2 is the Internet standard MIB that defines over 100 TCP/IP specific objects, including configuration and statistical information such as:

- Information about interfaces
- Address translation
- IP, ICMP (Internet-control message protocol), TCP, UDP.

SNMP can be extended through the use of the SNMP Multiplexing protocol (the SMUX protocol) to include *enterprise-specific* MIBs that contain information relating to a discrete environment or application. A management agent (a SMUX peer daemon) retrieves and maintains information about the objects defined in its MIB, and passes this information on to a specialized network monitor or network management station.

The HACMP for AIX software, NetView for AIX, and Systems Monitor for AIX all include the following SMUX daemons: **clsmuxpd**, **trapgend**, and **sysinfod**, respectively. You must be aware of possible conflicts between these daemons.

HACMP for AIX SNMP Components

The HACMP for AIX software provides the HACMP for AIX MIB, associated with and maintained by the HACMP for AIX management agent, the Cluster SMUX peer daemon (**clsmuxpd**). The HACMP for AIX software also provides two cluster monitor programs, the Cluster Information Program (Clinfo) and **clstat**.

Cluster SMUX Peer Daemon (clsmuxpd)

The **clsmuxpd** manages the HACMP for AIX enterprise-specific (generic type 6) MIB. The MIB is stored in the **hacmp.defs** file.

The source file is **hacmp.my**. It is compiled (with other standard MIBs) by the **mosy** command to generate the **hacmp.defs** file.

Cluster Information Program (Clinfo)

Clinfo is a cluster monitor program. It requests information about the current cluster state from **clsmuxpd**, and it updates a shared memory segment that is accessible to Clinfo clients (applications that use Clinfo API functions).

By default, Clinfo receives information from **clsmuxpd** by polling. The time between polling is set by a command line argument to Clinfo, which defaults to 15. Clinfo also can receive information asynchronously via traps. Clinfo's response to traps is to send a request for more information to **clsmuxpd**; it does not parse the trap message data itself. Instead, it employs a trap-directed polling policy.

To enable Clinfo to receive traps, you must invoke it with the **-a** option. Since Clinfo is started through the System Resource Controller (SRC), the best way to do this is by entering:

```
chssys -s clinfo -a "-a"
```

Then use the **lssrc** command to verify the change. Enter:

```
lssrc -Ss clinfo | awk -F: ’{print $3}’
```

Traps provide more timely information to Clinfo clients. The trap function is totally transparent to these clients–they simply register to receive various events and are notified by Clinfo via signals when those events occur. Note, however, that Clinfo's polling interval is doubled when traps are enabled.

SNMP Community Names and Clinfo

The default SNMP community name for Clinfo is “public.” You can override this by using the following command to force the SRC to start Clinfo with the **-c** command line switch by entering:

```
chssys -s clinfo -a "-c abdcef"
```

where `abcdef` is an SNMP community name defined as such in the **snmpd.conf** file.

Then use the **lssrc** command to verify the change. Enter:

```
lssrc -Ss clinfo | awk -F: ’{print $3}’
```

The /usr/sbin/cluster/clstat Utility

The **/usr/sbin/cluster/clstat** utility runs on both ASCII and X terminals. The display automatically corresponds to the capability of the system. If you want to run an ASCII display on an X-capable machine, however, you can do so by specifying the **-a** option.

clstat is a Clinfo client. It uses the Clinfo C API to get cluster information from the shared memory segment maintained by Clinfo. It does not register to receive events, but uses the Clinfo polling method.

The LPP contains both executables and source code for the clstat utility. If you want to recompile clstat, run the **make** command in the directory **/usr/lpp/cluster/samples/clstat.**

NetView for AIX

NetView for AIX is a network manager which includes both a GUI and daemons that support the SNMP protocol. It can be used in IBM RS/6000 environments to provide an effective tool for monitoring and managing networks. It supports the loading and browsing of enterprise-specific MIBs, and it can be enabled to receive SNMP trap information.

The **trapgend** daemon is the SMUX peer agent provided with the NetView for AIX program that converts alertable errors to SNMP traps. On RS/6000 processors running an AIX version of 4.1 or greater, system errors are logged by the AIX error logging facilities in the **/dev/error** special file. An object installed by the NetView for AIX program in each system's Object Data Manager (ODM) directs the AIX error logging daemon **(errdemon**) to notify the trap-notify process when alertable errors are logged. These alertable errors are forwarded by the trap-notify process to the **trapgend** daemon, which converts them to SNMP traps. Using the SMUX protocol, **trapgend** forwards the traps to the AIX SNMP agent process, **snmpd**. The **snmpd** daemon then forwards the traps to the NetView for AIX program's **trapd** daemon.

For more information about using this product, see the NetView for AIX documentation.

Systems Monitor for AIX

Systems Monitor for AIX runs the **sysinfod** SMUX peer daemon which monitors the following characteristics:

- Machine name, type, and processor ID
- Devices installed on the machine
- Operating system configuration
- Status of subsystems, paging devices, and filesystems
- Network traffic
- Ethernet, Token-Ring, and X.25 adapter information
- Active processes
- Users
- CPU and device utilization.

If trap filtering is enabled on this agent system, the **sysinfod** daemon receives SNMP traps on port 162. By default, the **snmpd** daemon sends all received SNMP traps to the **sysinfod** daemon for filtering. The traps are evaluated by the **sysinfod** daemon, and those traps meeting the filter criteria are forwarded to the manager system.

See the Systems Monitor for AIX documentation for more information about using this product.

Systems Monitor Startup Options for HACMP Compatibility

If you are using the Systems Monitor for AIX along with HACMP for AIX on your system, you should start the **sysinfod** with the **-H** option. This option allows the HACMP for AIX cl_swap_HW_address utility to function correctly. If the **sysinfod** is not started with the **-H** option, it keeps the adapter busy all the time it is active, and this prevents the cl_swap_HW_address utility from removing the device when it tries swapping the HW address.

Trap Conflicts Between SMUX Peer Daemons

A single SNMP agent (**snmpd** daemon) can send the same trap to multiple SNMP managers; this agent is configured in the **/etc/snmpd.conf** file. Only one SNMP manager, however, (for example, NetView for AIX) can run on a given network station, because only one TCP/IP program at a time can listen on a particular port. There is no way to work around this limitation.

In the case of NetView for AIX, the **trapd** daemon listens on port 162 and forwards traps to NetView for AIX. In turn, NetView for AIX can forward traps to multiple NetView for AIX applications that have registered with NetView for AIX. **trapgend** can generate traps for AIX system error-log-related events. The variables in the private portion of **trapgend** are described in the file **/usr/etc/nm/mibs/ibm-nv6ksubagent.mib.**

When the **sysinfod** daemon is installed on an NetView for AIX manager, trap reception is disabled for filtering. This feature is set in the configuration file **/usr/adm/sm6000/config/install.config**. However, when the **sysinfod** daemon is installed on a node without the manager installed, trap reception is enabled using the same file. You can install NetView for AIX on a node where the **sysinfod** daemon is already installed and trap reception is enabled. This causes the NetView for AIX **trapd** daemon to fail to start since the **sysinfod** daemon is using the port.

Both the NetView for AIX manager and the **sysinfod** daemon cannot share this port. You must disable filtering on this node by way of the **/usr/adm/sm6000/config/install.config** configuration file. In this way, when you start the **sysinfod** daemon, it has trap reception and filtering disabled.

Similarly, Clinfo cannot be enabled to receive traps from **clsmuxpd** (activated by the **-a** flag) if **trapgend** is also running. If the NetView for AIX trap daemons are started first, Clinfo will immediately exit with a smux_connect error. If Clinfo is started first with the **-a** option, most of the NetView for AIX daemons will not start.

Appendix F Installing and Configuring HACMP for AIX on RS/6000 SPs

This appendix describes installation and configuration considerations for using HACMP for AIX, Version 4.4 software on RS/6000 SP Systems.

Overview

The HACMP for AIX software provides high-availability functions on RS/6000-based products, including the SP platform. Before attempting to install and configure HACMP for AIX on the SP, you should be familiar with manuals in the HACMP for AIX documentation set, especially the *HACMP for AIX Planning Guide* and the *HACMP for AIX Administration Guide*.

Related Publications

You can find information about IBM RS/6000 SP books in an on-line library of documentation covering AIX, RS/6000, and related products on the World Wide Web. Enter the following URL:

```
http://www.rs6000.ibm.com/aix/library
```

Installing HACMP for AIX on an SP System

Although there are no unique HACMP for AIX filesets (install images) for the SP, refer to the instructions in Chapter 8, Installing HACMP for AIX Software to determine which HACMP for AIX filesets you need.

Additionally, the following software must be installed on the SP control workstation and nodes:

- RS/6000 SP version 3, release 1 of the AIX Parallel System Support Programs (PSSP) or greater.
- Latest service level of the PSSP software you plan to install on your system.

Warning: **DCE security and HACMP:** PSSP 3.2 offers new options for enhanced security. However, if you use these options with an HACMP cluster, you must perform additional steps to ensure the proper functioning of HACMP. Refer to the section Configuring Cluster Security on page F-3 for further details.

You must install the HACMP for AIX software on all nodes of the SP system that will participate in an HACMP cluster. It is recommended that the HACMP for AIX client image be installed on the control workstation such that the workstation can monitor cluster status.

Once the HACMP for AIX install images are available to each node and to the control workstation (either via NFS mount or a local copy), log onto each node and install the HACMP for AIX software by following the instructions in Chapter 8, Installing HACMP for AIX Software. After installing the software, read the HACMP for AIX release notes in the **/usr/lpp/cluster/doc** directory.

Assuming all necessary HACMP for AIX filesets are in the **/usr/sys/inst.images/hacmp** directory on the SP control workstation (and that you have a **.toc** file created in this directory), perform the following procedure on the control workstation:

Note: If you do not have a **.toc** file in the **/usr/sys/inst.images/hacmp** directory, enter either the **inutoc** command or the following command to create it:

```
installp -ld/spdata/sys1/install/lppsource
```

1. Create a file called **/HACMPHOSTS** that contains hostnames of nodes in the SP frame that will have the HACMP for AIX software installed.
2. Export the Working Collective (WCOLL) environment variable using the following command**:**

   ```
   export WCOLL=/HACMPHOSTS
   ```

3. Ensure that all hosts listed in the **/HACMPHOSTS** file are up (that is, each host responds) by entering the following command:

   ```
   /usr/lpp/spp/bin/SDRGetObjects host_responds
   ```

   where **SDRGetObjects** is an SP command that retrieves information from the SP System Data Repository to the SP database. A host response of 1 indicates that the node on which the HACMP for AIX software is installed and responding properly to the **HATS** daemon.
4. Enter the following command to mount the filesystem (from the control workstation) onto all nodes:

   ```
   dsh -ia /etc/mount CWNAME:/usr/sys/inst.images/hacmp /mnt
   ```

   where *CWNAME* is the hostname of the control workstation.
5. Enter the following command to install the HACMP for AIX software on the nodes:

   ```
   dsh -ia "/etc/installp -Xagd /mnt LPP_NAME"
   ```

   where *LPP_NAME* is the name of the product/fileset you need to install. This must be done for each fileset you need to install. Note that you can install “all” filesets.
6. Enter the following command to verify that HACMP was successfully installed on each node:

   ```
   dsh -ia "/etc/installp -s | grep cluster"
   ```

7. Reboot the nodes on which you installed the HACMP for AIX software.

Configuring the RS/6000 SP for HACMP for AIX

The following sections describe SP system changes that should be done for HACMP for AIX. Consult the *SP Installation Guide* and *SP Administration Guide* for more information concerning SP management.

Network Options

Consult the *IBM RS/6000 SP Administration Guide* and add the proper network options in the **tuning.cust** file on all SP HACMP for AIX nodes for the SP Switch network. Here's an example of the options for the SP nodes with HACMP for AIX:

```
no -o ipforwarding=0
no -o ipsendredirects=0
no -o thewall=16834
```

Consult the *IBM RS/6000 SP Administration Guide* for more information about changing other network options for maximizing performance based on your expected SP worktype and workload.

Configuring Cluster Security

Kerberos is a network authentication protocol used on the SP. Based on a secret-key encryption scheme, Kerberos offers a secure authentication mechanism for client/server applications.

In addition, PSSP 3.2 provides the option of running an RS/6000 SP system with an enhanced level of security, and as of AIX 4.3.1, you can use DCE authentication rather than Kerberos 4 authentication. However, these options may affect your HACMP functionality. Please read the following sections before planning your cluster security.

Kerberos

Kerberos eliminates the need for the traditional TCP/IP access control lists (**.rhosts** files) that were used in earlier HACMP security implementations by centralizing command authority via one authentication server, normally configured to be the SP control workstation. Rather than storing hostnames in a file (the **/.rhosts** approach), Kerberos issues dually encrypted *authentication tickets*. Each ticket contains two encryption keys: One key is known to both the client user and to the ticket-granting service, and one key is known to both the ticket-granting service and to the target service that the client user wants to access. For a more detailed explanation of Kerberos and the security features of the SP system, refer to the *IBM Parallel System Support Programs for AIX Administration Guide*.

By setting up all network IP labels in your HACMP configuration to use Kerberos authentication, you reduce the possibility of a single point of failure. You can configure Kerberos for a cluster automatically by running a setup utility called **cl_setup_kerberos**. Alternatively, you can perform the process manually. Because the utility-based approach is faster and less prone to error, it is usually preferable to the manual method.

To configure Kerberos on the SPs within an HACMP cluster, you must perform these general steps (detailed procedures appear in the following sections):

Step	What you do...
1	Make sure that HACMP has been properly installed on all nodes in the cluster. For more information, see Chapter 8, Installing HACMP for AIX Software.
2	Configure the HACMP cluster topology information on one node in the cluster. Note that because the **cl_setup_kerberos** utility needs an initial Kerberized **rcmd** path to each node in the cluster and to the control workstation, you must include the SP Ethernet as part of the configuration. Note that on the SP **setup_authent** is usually used to configure Kerberos on the entire SP system. **setup_authent** creates **rcmd** (used for **rsh** and **rcp**) service principals for all network IP labels listed in the System Data Repository (SDR). The SDR does not allow multiple IP labels to be defined on the same interface. However, HACMP requires that multiple IP labels be defined for the same interface during IPAT configurations. HACMP also requires that **godm** (Global ODM) service principals be configured on all IP labels for remote ODM operations. For these reasons, each time the nodes are customized after the SP **setup_authent** script is run (via **setup_server** or alone), you must rerun the **cl_setup_kerberos** script or manually reconfigure the systems to use Kerberos.
3	Create new Kerberos service principals and configure all IP labels for Kerberos authentication. You can choose to perform these tasks automatically (see Configuring Kerberos Automatically on page F-4) or manually (see Configuring Kerberos Manually on page F-5).
4	Set the cluster security mode to **Enhanced**, then synchronize the cluster topology. See Setting a Cluster’s Security Mode on page F-9.
5	Delete (or at least edit) the **cl_krb_service** file, which contains the Kerberos service principals password you entered during the configuration process. At the very least, you should edit this file to prevent unauthorized users from obtaining the password and possibly changing the service principals.
6	Consider removing unnecessary **.rhosts** files. With Kerberos configured, HACMP does not require the traditional TCP/IP access control lists provided by these files (but other applications might). You should consult your cluster administrator before removing any version of this file.

Configuring Kerberos Automatically

The **cl_setup_kerberos** utility automatically creates new Kerberos service principals in the Kerberos Authentication Database by copying the IP labels from the **cl_krb_service** file. It extracts the service principals and places them in a new Kerberos services file, **cl_krb-srvtab**; creates a **cl_klogin** file that contains additional entries required by the **.klogin** file; updates the **.klogin** file on the control workstation and on all nodes in the cluster; concatenates the **cl_krb-srvtab** file to each node’s **/etc/krb-srvtab** file.

To run the **cl_setup_kerberos** utility:

Note: Make sure that you have already installed HACMP on at least one node, and that you have configured the topology information before you perform this procedure.

1. Verify that there is a valid **/.k** file on the control workstation. This file stores the Kerberos Authentication Password so that batched commands can be run. If the **/.k** file is not present, issue the following command locally on the control workstation:

   ```
   /usr/lpp/ssp/kerberos/etc/kstash
   ```

2. Run **cl_setup_kerberos** from the configured node. (The utility is found in the **/usr/sbin/cluster/sbin** directory.)

   Note: You must be *within* the directory to run this command successfully. It is not sufficient to define the PATH correctly; the only way to run the cl_setup_kerberos command correctly is from within the **/usr/sbin/cluster/sbin** directory.

   cl_setup_kerberos extracts the HACMP IP labels from the configured node and creates a file, **cl_krb_service**, that contains all of the IP labels and additional format information required by the **add_principal** Kerberos utility. It also creates the **cl_adapters** file that contains a list of the IP labels required to extract the service principals from the authentication database.

3. When prompted, enter a Kerberos password for the new principals:

   ```
   Password:
   ```

   Note: This password is added to the **cl_krb_service** file. This can be the same as the Kerberos Administration Password, but doesn’t have to be. Follow your site’s password security procedures.

Configuring Kerberos Manually

To properly configure Kerberos on all HACMP-configured networks, you must perform the following general steps:

Step	What you do
1	Add an entry for each new Kerberos service principal to the Kerberos Authentication Database. See Adding New Service Principals to the Authentication Database on page F-6.
2	Update the **krb-srvtab** file by extracting each newly added instance from the Kerberos Authentication Database. See Updating the krb-srvtab File on page F-6.
3	Add the new service principals to each node’s **/.klogin** file. See Adding Kerberos Principals to Each Node’s .klogin File on page F-7.
4	Add the new service principals to each node’s **/etc/krb.realms** file. See Adding Kerberos Principals to Each Node’s /etc/krb.realms File on page F-8.

Adding New Service Principals to the Authentication Database

To add new service principals to the Kerberos Authentication Database for each network interface:

1. On the control workstation, start the **kadmin** utility

   ```
   kadmin
   ```

   A welcome message appears.

2. At the `admin:` prompt type the **add_new_key** command with the name and instance of the new principal:

   ```
   admin: ank service_name.instance
   ```

   where

   `service_name` is the service (**godm** or **rcmd**) and *`instance`* is the address label to be associated with the service. Thus, using the service **godm** and address label **i1_sw** the command is:

   ```
   admin: ank godm.i1_sw
   ```

3. When prompted, enter the Kerberos Administration Password.

   ```
   Admin password: password
   ```

4. When prompted, enter a Kerberos password for the new principal.

   ```
   Password for service_name.instance: password
   ```

   Note: The password can be the same as the Kerberos Administration Password, but doesn’t have to be. Follow your site’s password security procedures.

5. Verify that you have indeed added the new principals to the Kerberos database.

   ```
   kdb_util dump /tmp/testdb
   cat /tmp/testdb
   ```

   Remove this copy of the database when you have finished examining it.

   ```
   rm /tmp/testdb
   ```

Updating the krb-srvtab File

To update the **krb-srvtab** file and propagate new service principals to the HACMP cluster nodes:

1. Extract each new service principal for each instance you added to the Kerberos Authentication Database for those nodes you want to update. (This operation creates a new file in the current directory for each instance extracted.)

   ```
   usr/lpp/ssp/kerberos/etc/ext_srvtab -n i1_sw i1_en i1_tr
   ```

2. Combine these new files generated by the **ext_srvtab** utility into one file called *`node_name-new-srvtab`*:

   ```
   cat i1_sw-new-srvtab i1_en-new-srvtab i1_tr-new-srvtab
   > node_name-new-srvtab
   ```

   The new file appears in the directory where you typed the command.

   Note: Shared labels (used for rotating resource groups) need to be included in every **krb-srvtab** file (for nodes in that rotating resource group), so you must concatenate each shared-label srvtab file into each *`node_name-new-srvtab`* file.

3. Copy each *node_name-new-srvtab* file to its respective node.
4. Make a copy of the current **/etc/krb-srvtab** file so that it can be reused later if necessary:

   ```
   cp /etc/krb-srvtab /etc/krb-srvtab-date
   ```

   (where *date* is the date you made the copy).
5. Replace the current **krb-srvtab** file with the new *node_name-new-srvtab* file:

   ```
   cp node_name-new-srvtab /etc/krb-srvtab
   ```

6. Verify that the target node recognizes the new principals by issuing the following command on it:

   ```
   ksrvutil list
   ```

   You should see all the new principals for each network interface on that node; if not, repeat this procedure.

Adding Kerberos Principals to Each Node’s .klogin File

To add the new Kerberos principals to the **/.klogin** file on each HACMP cluster node:

1. Edit the **/.klogin** file on the control workstation to add the principals that were created for each network instance:

   ```
   vi /.klogin
   ```

   Here is an example of the **/.klogin** file for two nodes, i and j. ELVIS_IMP is the name of the realm that will be used to authenticate service requests. Each node has the SP Ethernet, a Token Ring service, and an Ethernet service adapter.

   ```
   root.admin@ELVIS_IMP
   rcmd.i1@ELVIS_IMP
   rcmd.i1_ensvc@ELVIS_IMP
   rcmd.i1_trsvc@ELVIS_IMP
   rcmd.j1@ELVIS_IMP
   rcmd.j1_ensvc@ELVIS_IMP
   rcmd.j1_trsvc@ELVIS_IMP
   godm.i1@ELVIS_IMP
   godm.i1_ensvc@ELVIS_IMP
   godm.i1_trsvc@ELVIS_IMP
   godm.j1@ELVIS_IMP
   godm.j1_ensvc@ELVIS_IMP
   godm.j1_trsvc@ELVIS_IMP
   ```

2. Copy the **/.klogin** file from the control workstation to each node in the cluster.

To verify that you set this up correctly, issue a Kerberized **rsh** command on all nodes using one of the newly defined interfaces. For example:

```
/usr/lpp/ssp/rcmd/bin/rsh i1_ensvc date
```

To eliminate single points of failure, you should add Kerberos **rcmd** and **godm** principals for every interface configured in HACMP.

Adding Kerberos Principals to Each Node’s /etc/krb.realms File

To add the new Kerberos principals to the **/etc/krb.realms** file on each HACMP cluster node:

1. Edit the **/etc/krb.realms** file on the control workstation and add the principals that were created for each network instance.

   ```
   vi /etc/krb.realms
   ```

   Here is an example of the **krb.realms** file for two nodes, i and j. ELVIS_IMP is the name of the realm that will be used to authenticate service requests. Each node has the SP Ethernet, a Token-Ring service, and an Ethernet service adapter.

   ```
   root.admin ELVIS_IMP
   i1 ELVIS_IMP
   i1_ensvc ELVIS_IMP
   i1_trsvc ELVIS_IMP
   j1 ELVIS_IMP
   j1_ensvc ELVIS_IMP
   j1_trsvc ELVIS_IMP
   i1 ELVIS_IMP
   i1_ensvc ELVIS_IMP
   i1_trsvc ELVIS_IMP
   j1 ELVIS_IMP
   j1_ensvc ELVIS_IMP
   j1_trsvc ELVIS_IMP
   ```

2. Copy the **/etc/krb.realms** file from the control workstation to each node in the cluster.

PSSP 3.2 Enhanced Security Options

PSSP 3.2 provides the option of running your RS/6000 SP system with an enhanced level of security. This function removes the dependency PSSP has to internally issue **rsh** and **rcp** commands as a root user from a node. When this function is enabled, PSSP does not automatically grant authorization for a root user to issue **rsh** and **rcp** commands from a node. Be aware that if you enable this option, some procedures may not work as documented. To run HACMP, an administrator must grant the authorizations for a root user to issue **rsh** and **rcp** commands that PSSP would otherwise grant automatically. See the redbook *Exploiting RS/6000 SP Security: Keeping it Safe*, SG24-5521-00, for a description of this function and a complete list of limitations.

If the enhanced security feature is enabled, the administrator could authorize a root user to issue rsh and **rcp** commands using the following steps:

1. Prior to the altering, synchronizing, or verifying an HACMP cluster configuration, the administrator of each node in the HACMP cluster must create (or update) the root user’s **rsh** authorization file (either **/.klogin** or **/.rhosts**) to allow the root users on the other nodes in the cluster to issue **rsh** commands to that node. Also, the appropriate AIX remote command authentication method would have to be enabled on the nodes of the HACMP cluster (if the method was not already enabled).
2. Perform any desired alteration, verification, and synchronization of the HACMP cluster configuration.
3. (Optional step if desired by the node administrators): Remove the authorization file entries added in step 1. Disable the authentication method (if enabled in step 1).

DCE Authentication

As of PSSP 3.2 and AIX 4.3.1, you are allowed the option of using DCE authentication rather than Kerberos 4 authentication. If you do this, you will not be able to alter, synchronize, or verify the HACMP configuration. Note that you will not be able to explicitly move a resource group, since that is a type of reconfiguration. If DCE (i.e. only Kerberos V5) is enabled as an authentication method for the AIX remote commands, you can still use HACMP, but must perform the following steps:

1. Prior to configuring the HACMP cluster, enable Kerberos V4 or Standard AIX as an authentication method for the AIX remote commands on the cluster nodes, and create the remote command authorization files (either **/.klogin** or **/.rhosts** files) for the root user on those nodes. This provides the ability for root to **rsh** among those nodes.
2. Configure the HACMP cluster.
3. Remove the remote command authorization files created in step 1 on the HACMP cluster nodes.
4. Disable the Kerberos V4 or Standard AIX authentication method enabled in step 1 on the HACMP cluster nodes.

Setting a Cluster’s Security Mode

You can set or change the security mode of all nodes in the cluster from the Change/Show Cluster Security SMIT screen. Because the cluster security mode is part of the HACMPcluster ODM, any changes you make are viewed as topology changes. This means that you must synchronize topology to propagate your security mode changes to all other nodes in the cluster. (You also verify the security setting by synchronizing/verifying the cluster topology.) For the same reason, you cannot dynamically reconfigure a cluster’s topology and resource configuration simultaneously.

To set the security mode of all nodes in a cluster to **Enhanced**:

1. From the SMIT Cluster Configuration screen, select **Cluster Security** > **Change/Show Cluster Security**.

   The Change/Show Cluster Security SMIT screen appears.
2. Set the security mode to **Enhanced**.
3. Synchronize the cluster topology. For more information, see Chapter 13, Verifying the Cluster Topology.

Automount Daemon

For SP installations that require the automount daemon (AMD) on HACMP nodes, a modification is needed to insure that AMD starts properly (with NFS available and running) on node bootup. This is due to the way HACMP for AIX manages the **inittab** file and run levels upon startup.

To enable AMD on nodes that have HACMP for AIX installed, add the following line as the last line of the file **/usr/sbin/cluster/etc/harc.net**:

```
startsrc -s nfsd
```

Cluster Verification and Synchronization

When running HACMP for AIX verification and synchronization, make sure the SP Switch network is up; otherwise, you receive network down error messages.

Configuring the SP Switch Network

The process for configuring the HACMP for AIX Version 4.4 software on SP nodes is similar to configuring it on an RS/6000 system but with additional information needed for configuring the SP Switch. Follow the instructions given in the *HACMP for AIX Planning Guide* and the *HACMP for AIX Installation Guide* for all installations, and follow these instructions for configuring the SP Switch.

Note: You must define an additional network besides the SP Switch, in order to avoid synchronization errors. Since the SP Switch uses IP aliasing, defining the network does not update the CuAt ODM database with an HACMP-defined adapter. During synchronization, HACMP looks for entries for adapters in the CuAt ODM database. If it finds none, synchronization fails.

Since only one SP Switch adapter per SP node is present, it is a single point of failure for that node. It is recommended that you use AIX error notification to promote adapter failures to node failures. Errors that should be provided to AIX error notification are “HPS_FAULT9_ER” and “HPS_FAULT3_RE”. To test the notify methods produced by these errors, use the error log emulation utility. SP Switch errors are included in the set of errors for which you can use the Automatic Error Notification feature.

For more information, see Chapter 16, Supporting AIX Error Notification.

HACMP for AIX Eprimary Management for the SP Switch

HACMP for AIX 4.2.1 and 4.2.2 allowed either the HPS (older version) or the SP (newer version) of the switch. HACMP for AIX 4.3.0 and up support only the SP switch.

The Eprimary node is the designated node for the switch initialization and recovery.

The SP switch cannot be managed by HACMP. The SP switch can configure a secondary Eprimary and reassign the Eprimary automatically on failure. This is handled by the SP software, outside of HACMP.

Upgrading From HPS Switch to SP Switch

You must upgrade to the SP Switch before installing HACMP 4.4 on an RS/6000 SP. If you are currently running HACMP Eprimary management with an HPS switch, you should run the HACMP for AIX script to unmanage the Eprimary BEFORE upgrading the switch.

To check whether the Eprimary is set to be managed:

```
odmget -q’name=EPRIMARY’ HACMPsp2
```

If the switch is set to MANAGE, before changing to the new switch, run the script:

```
/usr/sbin/cluster/events/utils/cl_HPS_Eprimary unmanage
```

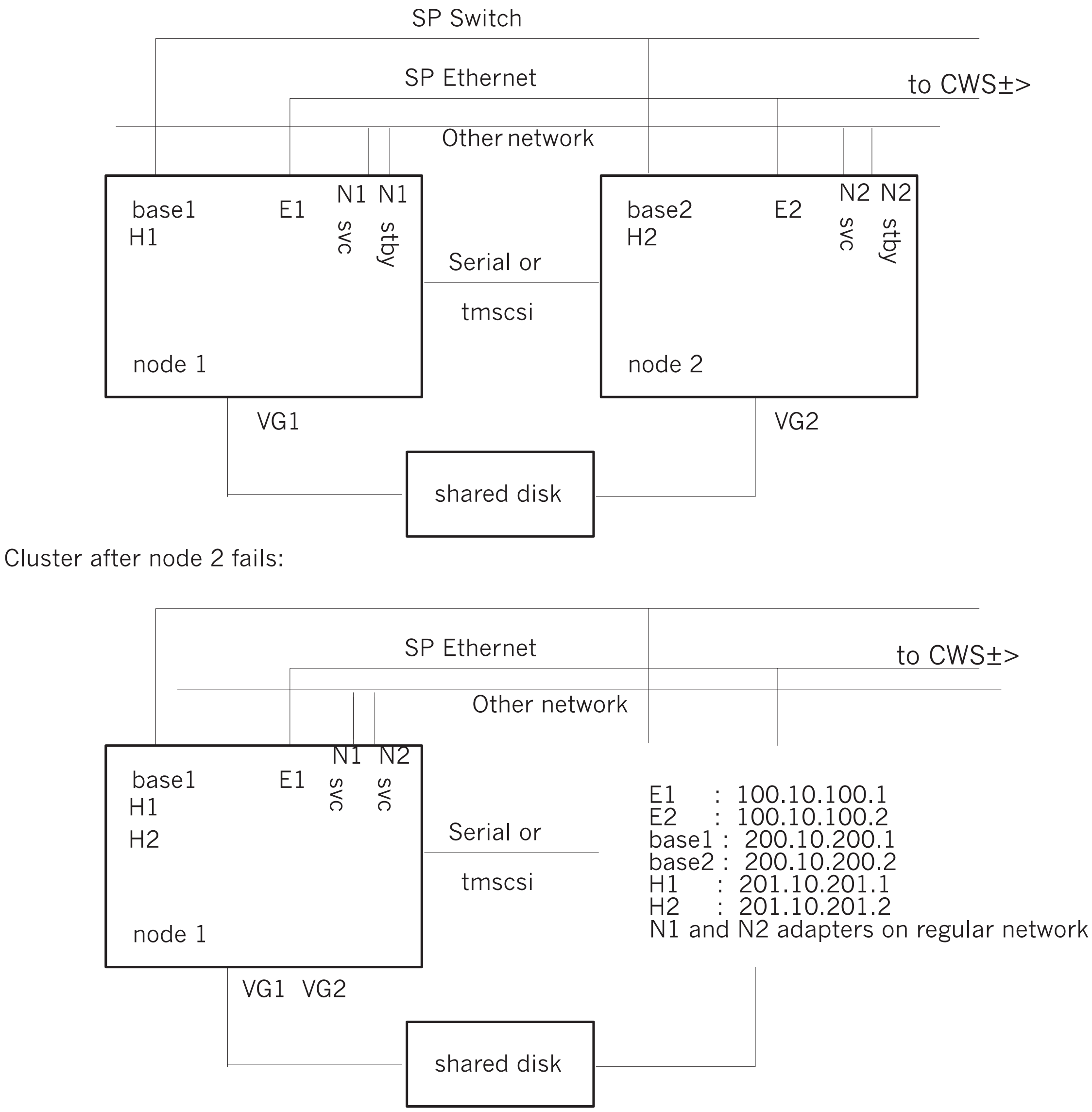
Configuring the SP Switch Base IP Address as Service Adapter

The SP Switch adapter css0 base IP address can be used as a service adapter as long as IP Address Takeover is *not* configured for the switch. It must not be used in an IP Address Takeover configuration.

The base address cannot be modified.

Configuring HACMP for AIX for IP Address Takeover on an SP Switch Network

The SP Switch is the first network to make use of IP address aliasing with HACMP for AIX to permit IP address takeover (IPAT). See the following figure for a general illustration of IP address takeover on the SP Switch. In the figure, Node 1 and Node 2 are SP nodes. E1 and E2 are SP Ethernet (Reliable Ethernet) IP addresses. The base1 and base2 labels reflect SP Switch base IP addresses. H1 and H2 are SP Switch alias HACMP for AIX service addresses. VG1 and VG2 are volume groups. N1 and N2 are adapters on other (Ethernet, FDDI) networks.



Sample HACMP for AIX Two-Node Cluster Configuration on the SP Machine.

Note: The HACMP boot addresses are not included in this figure. Boot addresses are also aliases; they are different from the SP switch base IP address.

Considerations for IPAT with the SP Switch

Keep the following points in mind when configuring the HACMP for AIX 4.4 software for the SP Switch using aliasing for an IPAT configuration:

- HACMP for AIX SP Switch boot and service addresses must be alias addresses on the SP Switch css0 IP interface. The css0 interface can have more than one alias IP address; therefore, it can support IP takeover addresses. At present, only one boot address can be defined per SP Switch css0 interface.

You can configure HACMP to have the switch adapter take over up to seven additional node IP addresses using aliasing. These HACMP for AIX “alias HPS service addresses” appear as “ifconfig alias” addresses on the css0 interface when viewing the node interfaces.

- SP Switch boot and service addresses must be different from the css0 base IP address in order to configure IP Address takeover.
- Address Resolution Protocol (ARP) must be enabled for the SP Switch network in order for IPAT to work on the SP Switch. ARP can be configured by an SP customize operation, or during initial SP setup. A method to update the SP Switch to use ARP is presented in SP Switch Address Resolution Protocol (ARP) on page F-13.
- Standby adapter addresses are not used for SP Switch IP address takeover.
- The SP Switch alias addresses for IPAT can be configured as a part of a cascading or rotating resource group.

Note: In the case of a major SP Switch failure, the aliases HACMP/ES needs for switch IP address takeover may be deleted when the **Eclock** command runs **rc.switch.** For this reason, if you are configuring IPAT with the SP Switch, you should create an event script for either the network_down or the network_down_complete event to add back the aliases for css0.

SP Switch Address Resolution Protocol (ARP)

If your SP nodes are already installed and the switch network is up on all nodes, you can verify whether ARP is enabled. On the control workstation, enter the following command:

```
dsh -av "/usr/lpp/ssp/css/ifconfig css0"
```

If NOARP appears as output from any of the nodes, you must enable ARP to use IP takeover on the SP Switch. ARP must be enabled on all SP nodes connected to the SP Switch.

Warning: Before you perform the following steps, be sure to back up **CuAt**. If user error causes **CuAt** to become corrupt, the SP nodes may be corrupted and will have to be re-installed. You will need to copy your backup of **CuAt** to **/etc/objrepos/CuAt** prior to rebooting the system. Be careful! If you feel this is too risky, customize the nodes to turn ARP on (see the *SP Administration Guide* for help with this procedure).

To enable ARP on all the nodes, follow these steps carefully. Enter all commands from the control workstation. Ensure all nodes are up. The quotation marks shown in the commands must be typed.

1. Create a copy of the **CuAt** file on all nodes:

   ```
   dsh -av "cp /etc/objrepos/CuAt /etc/objrepos/CuAt.save"
   dsh -av "odmget -q ’name=css and attribute=arp_enabled’ CuAt |
         sed s/no/yes/ > /tmp/arpon.data"
   dsh -av "odmchange -o CuAt -q’name=css and attribute=arp_enabled’
         /tmp/arpon.data"
   ```

2. Verify that the previous commands worked:

   ```
   dsh -av "odmget -a ’name=css and name=arp_enabled’ CuAt | grep value"
   ```

   You should see an entry reporting “value=yes” from every node.

3. Remove the temporary file from all nodes:

   ```
   dsh -av rm /tmp/arpon.data
   ```

4. Shut down and reboot the nodes:

   ```
   dsh -av "shutdown -Fr"
   ```

Handling Global Network Failure

The SP Switch is a highly available network. All nodes have four paths to each other through the switch network. Fault isolation and recovery is automatic. However, extremely rare failures will result in SP Switch outage on all nodes, or global network failure. The following section is intended to help in dealing with this situation.

Warning: Do not promote a global network_down failure to node failure. A node failure causes all the nodes to be powered off.

Global Network Failure Detection and Action

Several options exist for detecting failure and invoking user defined scripts to verify the failure and recover.

The switch power off will be seen as a HPS_FAULT9_ER recorded on each node, followed by HPS_FAULT6_ER (fault service daemon terminated). By modifying the AIX error notification strategies, it is possible to call a user script to detect the global switch failure and perform some recovery action. The user script would have to do the following:

- Detect global network failure (switch power failure or fault service daemon terminated on all nodes).
- Take recovery action, such as moving workload to another network, or reconfiguring a backup network.

Note: In order to recover from a major switch failure (power off, for example), you must issue **Eclock** and **Estart** commands to bring the switch back on-line. The **Eclock** command runs **rc.switch,** which deletes the aliases HACMP/ES needs for SP Switch IP address takeover. It is recommended to create an event script for either the network_down or the network_down_complete event to add back the aliases for css0.

SP Switch errors are included in the errors you can configure with the HACMP Automatic Error Notification feature. For more information on defining notify methods in the AIX Error Notification facility, see Chapter 16, Supporting AIX Error Notification.

Other Network Issues

You should give thought to the following issues when using an RS/6000 SP machine:

- No IPAT support on the SP administrative Ethernet
- Serial or non-IP network considerations.

IP Address Takeover Not Supported on the SP Administrative Ethernet

Since some of the SP software requires that an IP address on the SP administrative Ethernet (en0 adapter) is associated with a specific node, IP address takeover, as defined in cascading or rotating resource groups, cannot be configured on this network. You should configure this adapter so the network will be monitored by HACMP for AIX (this is done by configuring the SP Ethernet adapter as part of a cascading resource group where the adapter labels are not part of the resource group), but it must not be configured for IP address takeover (do not configure a boot address). In addition, no owned or takeover resources can be associated with this adapter.

Serial or Non-IP Network Considerations

It is strongly recommended (but not required) that a non-IP network be present between nodes that share resources, in order to eliminate TCP/IP (**inetd)** on one node as a single point of failure. At present, target mode SCSI (tmscsi) target mode SSA (tmssa) or serial (tty) networks are supported by HACMP for AIX.

- On the SP there are no serial ports available on thin or wide nodes. Therefore, any HACMP for AIX configurations that require a tty network need to make use of a serial adapter card (8-port async EIA-232 adapter, FC/2930), available on the SP as an RPQ.
- For 7135 and SCSI configurations, tmscsi or tmssa can be used with I/O pacing to provide the serial network, as described in this book.

Index

Vos remarques sur ce document / Technical publication remark form

Titre / Title : Bull HACMP 4.4 Installation Guide

Nº Reférénce / Reference Nº : 86 A2 56KX 02

Daté / Dated : August 2000

ERREURS DETECTEES / ERRORS IN PUBLICATION

AMELIORATIONS SUGGEREES / SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION

Vos remarques et suggestions seront examinées attentivement.
Si vous désirez une réponse écrite, veuillez indiquer ci-après votre adresse postale complète.

Your comments will be promptly investigated by qualified technical personnel and action will be taken as required.
If you require a written reply, please furnish your complete mailing address below.

NOM / NAME : ______________________________ Date : __________

SOCIETE / COMPANY : ______________________________

ADRESSE / ADDRESS : ______________________________

____

Remettez cet imprimé à un responsable BULL ou envoyez-le directement à :

Please give this technical publication remark form to your BULL representative or mail to:

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

Technical Publications Ordering Form

Bon de Commande de Documents Techniques

To order additional publications, please fill up a copy of this form and send it via mail to:
Pour commander des documents techniques, remplissez une copie de ce formulaire et envoyez-la à :

BULL CEDOC
ATTN / MME DUMOULIN
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

Managers / Gestionnaires :

Mrs. / Mme :	**C. DUMOULIN**	+33 (0) 2 41 73 76 65
Mr. / M :	**L. CHERUBIN**	+33 (0) 2 41 73 63 96
FAX :		+33 (0) 2 41 73 60 19
E–Mail / Courrier Electronique :		srv.Cedoc@franp.bull.fr

Or visit our web site at: / Ou visitez notre site web à:
`http://www-frec.bull.com` (PUBLICATIONS, Technical Literature, Ordering Form)

CEDOC Reference # N^o Référence CEDOC	**Qty** Qté	**CEDOC Reference #** N^o Référence CEDOC	**Qty** Qté	**CEDOC Reference #** N^o Référence CEDOC	**Qty** Qté
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	

[_ _] : **no revision number means latest revision** / pas de numéro de révision signifie révision la plus récente

NOM / NAME : ______________________________ Date : __________

SOCIETE / COMPANY : ______________________________

ADRESSE / ADDRESS : ______________________________

____

PHONE / TELEPHONE : ____________________ FAX : ____________________

E–MAIL : ______________________________

For Bull Subsidiaries / Pour les Filiales Bull :

Identification: ______________________________

For Bull Affiliated Customers / Pour les Clients Affiliés Bull :

Customer Code / Code Client : ______________________________

For Bull Internal Customers / Pour les Clients Internes Bull :

Budgetary Section / Section Budgétaire : ______________________________

For Others / Pour les Autres :

Please ask your Bull representative. / Merci de demander à votre contact Bull.

PLACE BAR CODE IN LOWER
LEFT CORNER

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

ORDER REFERENCE
86 A2 56KX 02

Bull

Utiliser les marques de découpe pour obtenir les étiquettes.
Use the cut marks to get the labels.

AIX

HACMP 4.4
Installation Guide

86 A2 56KX 02

AIX

HACMP 4.4
Installation Guide

86 A2 56KX 02

AIX

HACMP 4.4
Installation Guide

86 A2 56KX 02