High Availability Cluster
Multi-Processing for AIX

Installation Guide

Version 5.4.1

SC23-5209-01

Second Edition (October 2007)

Before using the information in this book, read the general information in Notices for HACMP Installation Guide.

This edition applies to HACMP for AIX v.5.4.1 and to all subsequent releases of this product until otherwise indicated in new editions.

Contents

About This Guide 9

Chapter 1: Overview of the Installation Process 13

Overview of the Installation Tasks 13
Step 1: Installing HACMP on Server Nodes 13
Step 2: Installing HACMP on Client Nodes 13
Overview of the Cluster Creation Tasks 13
Step 1: Creating an HACMP Cluster 13
Step 2: Configuring an HACMP Cluster 14
Overview of the Migration Tasks 14
Where to Find Documentation 14

Chapter 2: Creating Shared LVM Components 15

Prerequisites 15
Logical Volumes 15
Configuring HACMP to Use NFS Version 4 17
Step 1: Configuring NFS and Changing to Version 4 17
Step 2: Configuring the NFS Local Domain 18
Step 3: Exporting, Mounting, and File Systems 18
Step 4: Editing the /usr/es/sbin/cluster/etc/exports file 18
Step 5: Removing Entries from /etc/exports 19
Where You Go from Here 19

Chapter 3: Upgrading an HACMP Cluster 21

Overview 21
Prerequisites 21
Terminology Overview 22
Identifying Your Upgrade Path 23
Supported Upgrade Paths 23
Upgrading from Pre-5.1 Versions of HACMP 23
Choosing Your Upgrade Scenario 23
Planning for an Upgrade 24
Updating the Current Cluster Configuration 24
Checking Cluster Condition, Settings, and Worksheets 25
Reviewing the Cluster and Node Version in the HACMP Configuration Database 25

Checking Types of Networks 25
Migrating Resource Groups Configured without Sites 26
Upgrading Applications Configured Using a Smart Assist 28
Ensuring Applications Use Proprietary Locking Mechanism 28
Addressing Concurrent Access (AIX) Migration Issues 28
Priority Override Location and Persistence 28
Notes on Upgrading from HACMP 5.1 29
HACMP 5.4.1 Inter-Site Selective Fallover for Resource Group Recovery 30
Upgrade Do's and Don'ts 31
Upgrade Do's 31
Upgrade Don'ts 31
Checking That the Software Is Committed 32
Performing a Rolling Migration to HACMP 5.4.1 32
Step 1: Stop Cluster Services on a Node Hosting the Application . 32
Step 2: Install the HACMP 5.4.1 Software 33
Step 3: Reboot the Node 33
Step 4: Start Cluster Services on the Upgraded Node 33
Step 5: Repeat Steps for Other Cluster Nodes 34
Step 6: Verify the Upgraded Cluster Definition 34
Upgrading to HACMP 5.4.1 Using a Snapshot 34
Step 1: Creating a Snapshot 34
Step 2: Stopping All Nodes 34
Step 3: Removing the Existing Version of the HACMP Software . 35
Step 4: Installing the HACMP 5.4.1 Software 35
Step 5: Converting and Applying a Saved Snapshot 35
Step 6: Verifying the Upgraded Cluster Definition 36
Upgrading to HACMP 5.4.1 on an Offline Cluster 37
Applying a PTF 37
Step 1: Stop Cluster Services on a Node Hosting the Application . 37
Step 2: Install the PTF Update 38
Step 3: Start Cluster Services on the Upgraded Node 39
Step 4: Repeat Steps for Other Cluster Nodes 39
Step 5: Verify the Upgraded Cluster Definition 40
Additional Migration Tasks 40
Recompiling Clinfo Clients after Migrating to HACMP 5.4.1 40
Using a Web-Based Interface for Configuration and Management 40
Resetting HACMP Tunable Values 41
Verifying the Success of Your Migration 41
Verifying Software Levels Installed Using AIX Commands 42
Automatically Saved Files 42
Verifying the Upgraded Cluster Definition 42
Verifying All Cluster Filesets Have Migrated 43
Testing HACMP on the Upgraded Cluster 43
Running AIX Commands on the Migrated Cluster 43
Troubleshooting Your Migration 45
Error: "config_too_long" 45
Error: "cldare cannot be run" 45
Recovering Your Previous Version of HACMP 46

Recovering from a Conversion Failure 46
Recovering Configuration Information 46

Chapter 4: Installing HACMP on Server Nodes 47

Prerequisites 47
Supported Hardware 47
Disk Space Requirements 48
Required Versions of AIX and RSCT 48
Security Fileset Requirements 50
HAView Requirements 50
Tivoli Requirements 51
English and Japanese Message Catalogs 51
Contents of the Installation Medium 52
HACMP Installable Images 52
HACMP Installation Choices 55
Installing from an Installation Server 55
Installing from a Hard Disk 56
Copying HACMP Software to Hard Disk 56
Installing HACMP from the Hard Disk 56
Installing HACMP on an SP from the Hard Disk 57
Installing from the Installation Medium 58
Completing the Installation 60
Verifying the Software Installation 60
Addressing Problems during the Installation 61
Cleaning up after an Unsuccessful Installation 61
Resolving System ID Licensing Issues 61
Removing the HACMP Software 61
Installing and Configuring WebSMIT 62
Planning for WebSMIT 62
Prerequisites Tasks for Installing WebSMIT 63
Steps for Installing and Configuring WebSMIT 63
Upgrading from a Previous Configuration 65
Where You Go from Here 66

Chapter 5: Installing HACMP on Client Nodes 69

Prerequisites 69
Overview 69
Installing and Configuring HACMP on Client Systems 69
Step 1: Installing the Base System Client Images 70
Step 2: Copying the clhosts.client File onto Client Nodes 70
Step 3: Editing the clinfo.rc Script 71
Step 4: Updating the ARP Cache for Clients Not Using Clinfo . . . 71
Step 5: Rebooting the Clients 71

Chapter 6: Configuring Installed Hardware 73

Configuring Network Interface Cards 73
Ethernet, Token-Ring, and FDDI Interface Cards 73
Completing the TCP/IP Network Interface Worksheets 74
Configuring Point-to-Point Networks 74
Configuring RS232 Serial Connections 75
Configuring an RS232 Serial Connection in AIX 75
Testing the Serial Connection 77
Defining the Serial Connection to HACMP 78
Configuring an SP Switch 78
Configuring for Asynchronous Transfer Mode (ATM) 79
Support of Classic IP and LANE on the Same Interface Card 79
Configuring ATM Classic IP 79
Configuring Classic IP for HACMP Cluster Networks 79
Testing the Configuration 80
Configuring ATM ARP Servers for Use by HACMP Nodes 81
Configuring ATM ARP Servers for Service Subnetworks 81
Configuring ATM ARP Servers for Non-Service Subnetworks . . . 81
Configuring ATM ARP Clients on Cluster Nodes 82
Testing Communication over the Network 84
Defining the ATM Network to HACMP 84
Configuring ATM LAN Emulation 84
Defining the ATM LAN Emulation Network to HACMP 85
Configuring Shared External Disk Devices 88
Configuring Shared Fibre Channel Disks 88
Configuring Shared SCSI Disks 89
Configuring IBM SCSI Disk Arrays 91
Configuring Target Mode SCSI Connections 92
Configuring the Status of SCSI Adapters and Disks 92
Configuring Target Mode SCSI Devices in AIX 93
Testing the Target Mode Connection 94
Defining the Target Mode SCSI Connection to HACMP 95
Configuring Target Mode SSA Connections 95
Changing Node Numbers on Systems in SSA Loop 95
Configuring Target Mode SSA Devices in AIX 96
Testing the Target Mode Connection 96
Configuring Shared IBM SSA Disk Subsystems 97
Defining the Target Mode SSA Connection to HACMP 97
Installing and Configuring Shared Tape Drives 98
Installing Shared SCSI Tape Drives 98
Installing and Configuring Shared Fibre Tape Drives 100
Configuring the Installation of a Shared Tape Drive 101

Chapter 7: Defining Shared LVM Components 103

Overview 103
Defining Shared LVM Components for Non-Concurrent Access 104
Defining Shared LVM Components 104

Creating a Shared Volume Group on Source Node 104
Creating a Shared File System on Source Node 105
Renaming jfslogs and Logical Volumes on Source Node 105
Adding Copies to Logical Volume on Source Node 106
Testing a File System 107
Varying Off a Volume Group on the Source Node 107
Collecting Information on Current Volume Group Configuration 107
Importing a Volume Group onto Destination Nodes 107
Changing a Volume Group's Startup Status 110
Varying Off Volume Group on Destination Nodes 110
Defining LVM Components for Concurrent Access 110
Creating a Concurrent Access Volume Group on a Source Node 111

Chapter 8: Configuring AIX for HACMP 117

I/O Considerations 117
I/O Pacing 117
Syncd Frequency 118
Networking Considerations 118
Checking User and Group IDs 118
Configuring Network Options 118
Changing routerevalidate Network Option 119
Updating the /etc/hosts File and nameserver Configuration 119
Setting up NIS-Managed Users to Create Crontabs 119
Enabling the AIX Automounter Daemon 120
Planning HACMP File Collections 120
Default HACMP File Collections 121
Types of Error Notification 124
AIX Error Notification Facility 124
HACMP Automatic Error Notification 125
Error Notification Method Used for Volume Group Loss 127
Emulation of Error Log Entries 129
SP-Specific Considerations 130
SP Switch Address Resolution Protocol (ARP) 130
SP Switch Global Network Failure Detection and Action 131

Chapter 9: Creating a Basic HACMP Cluster 133

Overview 133
HACMP Cluster Definition 134
Resource Group Policies 134
Prerequisites 135
TCP/IP Connectivity 135
Copy of the Application 135
Start Scripts and Stop Scripts 135
Volume Groups 135
Service IP Label/Address 136
HACMP Software 136
Planning a Two-Node Cluster 136

Using the Two-Node Cluster Configuration Assistant 138
User Privileges 138
Existing Clusters 138
Using the Standalone Assistant 139
Using the SMIT Assistant 139
Logging for the Two-Node Cluster Configuration Assistant 140
Preventing Single Points of Failure 140
Where You Go from Here 141

Appendix A: Cluster Monitoring with Tivoli 143

Appendix B: OEM Disk, Volume Group, and File Systems Accommodation151

Appendix C: GPFS Cluster Configuration 177

Appendix D: HACMP and SNMP Utilities 187

Index 195

About This Guide

This guide provides information necessary to plan and install the High Availability Cluster Multi-Processing for AIX software, version 5.4.1.

The following table provides version and manual part numbers for the *Planning and Installation Guide*.

HACMP Version	Book Name	Book Number
5.4.1	*Installation Guide*	SC23-5209-01
5.4	*Installation Guide*	SC23-5209-00
5.3 update 8/2005	*Planning and Installation Guide*	SC23-4861-07
5.3	*Planning and Installation Guide*	SC23-4861-06
5.2 update 10/2005	*Planning and Installation Guide*	SC23-4861-05
5.2 update 10/2004	*Planning and Installation Guide*	SC23-4861-04
5.2	*Planning and Installation Guide*	SC23-4861-03
5.1 update 6/2004	*Planning and Installation Guide*	SC23-4861-02

Who Should Use This Guide

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

- Planning hardware and software resources for an HACMP™ cluster
- Installing and configuring an HACMP cluster
- Maintaining and troubleshooting an HACMP cluster.

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

- IBM System p™ 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.

Highlighting

This guide uses the following highlighting conventions:

Italic	Identifies new terms or 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.

HACMP Publications

The HACMP software comes with the following publications:

- *HACMP for AIX Release Notes* in **/usr/es/sbin/cluster/release_notes** describe issues relevant to HACMP on the AIX platform: latest hardware and software requirements, last-minute information on installation, product usage, and known issues.
- *HACMP on Linux Release Notes* in **/usr/es/sbin/cluster/release_notes.linux/** describe issues relevant to HACMP on the Linux platform: latest hardware and software requirements, last-minute information on installation, product usage, and known issues.
- *HACMP for AIX: Administration Guide,* SC23-4862
- *HACMP for AIX: Concepts and Facilities Guide,* SC23-4864
- *HACMP for AIX: Installation Guide,* SC23-5209
- *HACMP for AIX: Master Glossary,* SC23-4867
- *HACMP for AIX: Planning Guide,* SC23-4861
- *HACMP for AIX: Programming Client Applications*, SC23-4865
- *HACMP for AIX: Troubleshooting Guide,* SC23-5177
- *HACMP on Linux: Installation and Administration Guide,* SC23-5211
- *HACMP for AIX: Smart Assist Developer's Guide,* SC23-5210
- *IBM International Program License Agreement.*

HACMP/XD Publications

The HACMP Extended Distance (HACMP/XD) software solutions for disaster recovery, added to the base HACMP software, enable a cluster to operate over extended distances at two sites. HACMP/XD publications include the following:

- *HACMP/XD for Geographic LVM (GLVM): Planning and Administration Guide,* SA23-1338
- *HACMP/XD for HAGEO Technology: Concepts and Facilities Guide, SC23-1922*
- *HACMP/XD for HAGEO Technology: Planning and Administration Guide,* SC23-1886
- *HACMP/XD for Metro Mirror: Planning and Administration Guide,* SC23-4863.

HACMP Smart Assist Publications

The HACMP Smart Assist software helps you quickly add an instance of certain applications to your HACMP configuration so that HACMP can manage their availability. The HACMP Smart Assist publications include the following:

- *HACMP Smart Assist for DB2 User's Guide, SC23-5179*
- *HACMP Smart Assist for Oracle User's Guide, SC23-5178*
- *HACMP Smart Assist for WebSphere User's Guide, SC23-4877*
- *HACMP for AIX 5L: Smart Assist Developer's Guide, SC23-5210*
- *HACMP Smart Assist Release Notes.*

IBM AIX Publications

The following publications offer more information about IBM technology related to or used by HACMP:

- *RS/6000 SP High Availability Infrastructure*, SG24-4838
- *IBM AIX v.5.3 Security Guide,* SC23-4907
- *IBM Reliable Scalable Cluster Technology for AIX and Linux: Group Services Programming Guide and Reference,* SA22-7888
- *IBM Reliable Scalable Cluster Technology for AIX and Linux: Administration Guide,* SA22-7889
- *IBM Reliable Scalable Cluster Technology for AIX: Technical Reference*, SA22-7890
- *IBM Reliable Scalable Cluster Technology for AIX: Messages*, GA22-7891.

Accessing Publications

Use the following Internet URLs to access online libraries of documentation:

AIX, IBM eServer pSeries, and related products:

http://www.ibm.com/servers/aix/library

AIX v.5.3 publications:

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

WebSphere Application Server publications:

Search the IBM website to access the WebSphere Application Server Library

DB2 Universal Database Enterprise Server Edition publications:

http://www.ibm.com/cgi-bin/db2www/data/db2/udb/winos2unix/support/v8pubs.d2w/en_main#V8PDF

Tivoli Directory Server publications:

http://publib.boulder.ibm.com/tividd/td/IBMDirectoryServer5.1.html

IBM Welcomes Your Comments

You can send any comments via e-mail to hafeedbk@us.ibm.com. Make sure to include the following in your comment or note:

- Title and order number of this book
- Page number or topic related to your comment.

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

Trademarks

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

- AFS
- AIX
- DFS
- @server
- eServer Cluster 1600
- Enterprise Storage Server
- HACMP
- IBM
- NetView
- RS/6000
- Scalable POWERParallel Systems
- Series p
- Series x
- Shark
- SP
- WebSphere
- Red Hat Enterprise Linux (RHEL)
- SUSE Linux Enterprise Server
- RPM Package Manager for Linux and other Linux trademarks.

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

Linux is a registered trademark in the United States and other countries and is licensed exclusively through the GNU General Public License.

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

Chapter 1: Overview of the Installation Process

This chapter provides an overview of the installation process, and presupposes that you have completed the planning process as described in the *Planning Guide*. This chapter contains the following sections:

- Overview of the Installation Tasks
- Overview of the Cluster Creation Tasks
- Overview of the Migration Tasks
- Where to Find Documentation.

Overview of the Installation Tasks

After completing the planning steps, you are ready to install the HACMP software. This section provides an overview of the installation process.

Step 1: Installing HACMP on Server Nodes

In this step, you install HACMP on all server nodes. Chapter 4: Installing HACMP on Server Nodes describes this step of the installation process.

Step 2: Installing HACMP on Client Nodes

In this step, you install HACMP on all client nodes. Chapter 5: Installing HACMP on Client Nodes describes this step of the installation process.

After installing HACMP, you are ready to create an HACMP cluster.

Overview of the Cluster Creation Tasks

After installing the HACMP software, you are ready to create an HACMP cluster. This section provides an overview of the cluster creation process.

Step 1: Creating an HACMP Cluster

HACMP supplies the following utilities that enable you to quickly create a cluster:

- Online Planning Worksheets
- Two-Node Cluster Configuration Assistant
- WebSphere, DB2 UDB or Oracle Smart Assists.

Using the Online Planning Worksheets Application

During the cluster-planning process, the Online Planning Worksheets application enables you to enter configuration data and save it to a cluster definition file. At the end of the planning process, you can use the cluster definition file to immediately configure your cluster. For information about Online Planning Worksheets, see the chapter on Using Online Planning Worksheets in the *Planning Guide*.

Using the Two-Node Cluster Configuration Assistant

You can easily create a basic two-node cluster using the Two-Node Cluster Configuration Assistant. For information about the Two-Node Cluster Configuration Assistant, see Chapter 9: Creating a Basic HACMP Cluster.

Using Smart Assist for WebSphere, DB2 or Oracle

You can also configure a cluster with a WebSphere, DB2 UDB, or Oracle application. For information, see the corresponding HACMP Smart Assist guide.

Step 2: Configuring an HACMP Cluster

In this step, you perform tasks as described in the following chapters:

- Chapter 6: Configuring Installed Hardware
- Chapter 7: Defining Shared LVM Components
- Chapter 8: Configuring AIX for HACMP

After you create a basic cluster and configured supporting components, you are ready to configure and monitor your cluster as described in the *Administration Guide*.

Overview of the Migration Tasks

If you are migrating an existing installation of HACMP, follow the instructions from Chapter 3: Upgrading an HACMP Cluster, and refer to Chapter 4: Installing HACMP on Server Nodes as needed.

Where to Find Documentation

All documentation is install in the directory:

`/usr/share/man/info/en_US/cluster/HAES`

Chapter 2: Creating Shared LVM Components

Setting up shared LVM components for an HACMP cluster depends on both of the following:

- The type of shared disk device
- The method of shared disk access.

This chapter describes the following:

- Prerequisites
- Logical Volumes
- Configuring HACMP to Use NFS Version 4
- Where You Go from Here.

Note: If you are planning an IBM General Parallel File System (GPFS) cluster, see the network requirements in Appendix C: GPFS Cluster Configuration.

If you are planning to use OEM disks, volume groups, or file systems in your cluster (including Veritas volumes) see Appendix B: OEM Disk, Volume Group, and File Systems Accommodation.

Prerequisites

At this point, you should have completed the planning steps described in the *Planning Guide.*

You should also be familiar with how to use the Logical Volume Manager (LVM). For information about AIX LVM, see the *AIX System Management Guide*.

Logical Volumes

A *logical volume* is a set of logical partitions that AIX makes available as a single storage unit—that is, the logical view of a disk. A *logical partition* is the logical view of a physical partition. Logical partitions may be mapped to one, two, or three physical partitions to implement mirroring.

In the HACMP environment, logical volumes can be used to support a journaled file system or a raw device.

Specify the **super strict** disk allocation policy for the logical volumes in volume groups for which forced varyon is specified. This configuration:

- Guarantees that copies of a logical volume always reside on separate disks
- Increases the chances that forced varyon will be successful after a failure of one or more disks.

If you plan to use forced varyon for the logical volume, apply the **superstrict** disk allocation policy for disk enclosures in the cluster.

To specify the **superstrict** disk allocation policy in AIX:

1. In SMIT, go to **Add a Shared Logical Volume**, or **Change a Shared Logical Volume.**
2. Select **Allocate each logical partition copy on a separate physical volume?**
3. When using the **superstrict** disk allocation policy, specify the correct number of physical volumes for this logical volume. *Do not use the default* setting of 32 physical volumes.

Default NFS Mount Options for HACMP

When performing NFS mounts, HACMP uses the default options `hard`, `intr`.

To set `soft` mounts or any other options on the NFS mounts:

1. Enter `smit mknfsmnt`
2. In the **MOUNT now, add entry to /etc/filesystems or both?** field, select the **file systems** option.
3. In the **/etc/filesystems entry will mount the directory on system RESTART** field, accept the default value of **no**.

This procedure adds the options you have chosen to the **/etc/filesystems** entry created. The HACMP scripts then read this entry to pick up any options you may have selected.

Creating and Configuring NFS Mount Points on Clients

An NFS mount point is required to mount a file system via NFS. In a non-concurrent resource group, all the nodes in the resource group NFS mount the file system. The NFS mount point must be outside the directory tree of the local mount point.

Once you create the NFS mount point on all nodes in the resource group, configure the **NFS Filesystem to NFS Mount** attribute for the resource group.

To create NFS mount points and to configure the resource group for the NFS mount:

1. On each node in the resource group, create an NFS mount point by executing the following command:

   ```
   mkdir /mountpoint
   ```

   where *mountpoint* is the name of the local NFS mount point over which the remote file system is mounted.
2. In the **Change/Show Resources and Attributes for a Resource Group** SMIT panel, the **Filesystem to NFS Mount** field must specify both mount points.

   Specify the NFS mount point, then the local mount point, separating the two with a semicolon. For example:

   ```
   /nfspoint;/localpoint
   ```

   If there are more entries, separate them with a space:

   ```
   /nfspoint1;/local1 /nfspoint2;/local2
   ```
3. (*Optional*) If there are nested mount points, nest the NFS mount points in the same manner as the local mount points so that they match up properly.
4. (*Optional*) When cross-mounting NFS file systems, set the **Filesystems Mounted before IP Configured** field in SMIT for the resource group to **true**.

Configuring HACMP to Use NFS Version 4

HACMP supports NFS protocol Version 4 (NFS V4). To ensure that HACMP properly identifies NFS file systems mounted for NFS V4, you must:

1. Correctly set up NFS V4 configuration.
2. Make this configuration consistent on all nodes.

The fields needed to configure HACMP to use NFS V4 are described in this section. For more information about configuring NFS V4, see the AIX documentation as listed in About This Guide.

To correctly configure HACMP for NFS V4, follow the steps in these sections:

Step 1: Configuring NFS and Changing to Version 4

Step 2: Configuring the NFS Local Domain

Step 3: Exporting, Mounting, and File Systems

Step 4: Editing the /usr/es/sbin/cluster/etc/exports file

Step 5: Removing Entries from /etc/exports.

Step 1: Configuring NFS and Changing to Version 4

For HACMP to recognize NFS V4, you change the NFS version on one node in the cluster in AIX first, and then on the rest of the nodes.

To change the NFS version on one node in the cluster:

1. Enter the fastpath `smitty nfs`
2. In SMIT, select **Network File System (NFS) > Configure NFS on This System > Change Version 4 Server Root Node** and press Enter.
3. Enter field values on the **Change Version 4 Server Root Node panel** as follows:

Root Node Directory	Enter the root node directory, for example [/].
Change number now, system restart, or both	Select **both** from the picklist so that changes requested take place immediately and for every subsequent system restart.

You must also change the NFS version on each node in the cluster in AIX.

To change the NFS version on each node in the cluster:

1. Enter the fastpath `smitty nfs`
2. In SMIT, select **Network File System (NFS) > Configure NFS on This System > Change Version 4 Server Public Node** and press Enter.
3. Enter field values on the **Change Version 4 Server Public Node** panel as follows:

Public Node Directory	Enter the public node directory, for example [/].
Change number now, system restart, or both	Select **both** from the picklist so that changes requested take place immediately and for every subsequent system restart.

Step 2: Configuring the NFS Local Domain

To set the Local Domain on each node using AIX SMIT:

1. Enter the fastpath `smitty nfs`
2. In SMIT, select **Network File System (NFS) > Configure NFS on This System > Configure NFS Local Domain > Change NFS Local Domain** and press Enter.
3. Enter the following field value on the **Display Current NFS Local Domain** panel as follows:

NFS Local Domain	Specify the new NFS local domain of the system.

Step 3: Exporting, Mounting, and File Systems

NFS exports are configured into HACMP through Resource Groups. After NFS file systems become part of resource groups that belong to an active HACMP cluster, HACMP exports, unexportis, cross-mounts and unmounts the file systems, during cluster events, such as fallover of a resource group containing the file system to another node in the cluster.

There are two ways how the exports can be added and mounts can be specified.

1. Using NFS Configuration Assistant. This is designed to help you configure, view, change, or delete Resource Group(s) with NFS exports. The Configuration Assistant creates a new Resource Group with the specified NFS exports and mounts. For more information, see NFS Configuration Assistant in Chapter 14: Managing Cluster Resources in the *Administration Guide*.

2. Using the **Resource Group Extended Attributes Configuration** screen. This is used to add, modify, or delete NFS exports and mounts to an already existing Resource Group. For more information about configuring Resource Groups, see Chapter 5: Configuring HACMP Resource Groups (Extended) in the *Administration Guide*.

Step 4: Editing the /usr/es/sbin/cluster/etc/exports file

Modify the HACMP **/usr/es/sbin/cluster/etc/exports** file on each node in the cluster to add the IP addresses for the network. You may edit the file on one node and copy it to other cluster nodes. You can also use the HACMP File Collection facility to keep this file in sync on all of the nodes of the cluster. For more information, see Managing HACMP File Collections in Chapter 7: Verifying and Synchronizing an HACMP Cluster in the *Administration Guide*.

To modify the **/usr/es/sbin/cluster/etc/exports** file on each HACMP cluster node:

1. Edit the **/usr/es/sbin/cluster/etc/exports** file on the control workstation:

   ```
   vi /usr/es/sbin/cluster/etc/exports
   ```

2. For each file system, there should be a line that looks like this:

   ```
   /fs/fs3big
   -vers=4,sec=sys:krb5p:krb5i:krb5:dh:none,rw,root=192.168.20.1:19
   2.168.20.1:192.168.20.2:192.168.20.3:192.168.21.1:192.168.21.2:192.16
   8.21.
   3:192.168.30.1:192.168.30.2:192.168.30.3
   ```

   where root is a colon-separated list of all the IP addresses for the specified network.

Note: Using this alternate exports file is optional. HACMP checks the **/usr/es/sbin/cluster/etc/exports** file when NFS exports a file system or directory. If there is an entry for the file system or directory in this file, HACMP uses the options listed, except that HACMP 5.4.1 and later might ignore the version option as described in Steps for Adding Resources and Attributes to Resource Groups (Extended Path) in Chapter 5: Configuring HACMP Resource Groups (Extended) in the *Administration Guide*. If the file system or directory for NFS-export is not listed in the file, or if the alternate file does not exist, the file system or directory is NFS-exported with the default option of root access for all cluster nodes.

Step 5: Removing Entries from /etc/exports

Modify the AIX **/etc/exports** file (*not* the HACMP **/usr/es/sbin/cluster/etc/exports** file) on each HACMP cluster node by removing all the entries.

To remove all the entries in the **/etc/exports** file on each HACMP cluster node run the command:

```
cat /dev/null > /etc/exports
```

Where You Go from Here

After you create your shared LVM components, complete the following steps to configure an HACMP server:

1. If you are upgrading an HACMP cluster configuration, see Chapter 3: Upgrading an HACMP Cluster.
2. If you need to install HACMP on your server nodes, see Chapter 4: Installing HACMP on Server Nodes.
3. Set up the Cluster Information Program.

   Copy the **clhosts.client** file to each client node as **/usr/es/sbin/cluster/etc/clhosts** and edit the **/usr/es/sbin/cluster/etc/clinfo.rc** script as described in Chapter 5: Installing HACMP on Client Nodes.
4. Ensure that the network interfaces and shared external disk devices are ready to support an HACMP cluster.

   See Chapter 6: Configuring Installed Hardware.
5. Define shared LVM components, including creating shared volume groups, logical volumes, and file systems for your cluster. Also define enhanced concurrent volume groups.

   See Chapter 7: Defining Shared LVM Components.
6. Customize your AIX environment for HACMP.

   See Chapter 8: Configuring AIX for HACMP.
7. If you want to set up a basic two-node configuration, use the Two-Node Cluster Configuration Assistant. For information about using the Two-Node Cluster Configuration Assistant, see Chapter 9: Creating a Basic HACMP Cluster.

8. (*Optional*) Configure the HACMP cluster using the Online Planning Worksheets application.

   See the chapter on Using Online Planning Worksheets in the *Planning Guide*.

   If you decided *not* to use the application, use the paper planning worksheets you filled out using the *Planning Guide*, and follow the steps in the *Administration Guide* to configure your cluster using SMIT or WebSMIT.

9. Test your configuration using the Cluster Test Tool.

   See Chapter 8: Testing and HACMP Cluster in the *Administration Guide*.

Chapter 3: Upgrading an HACMP Cluster

This chapter provides instructions for upgrading an existing HACMP cluster configuration to HACMP 5.4.1.

Overview

Before performing an upgrade, read these sections:

- Prerequisites
- Terminology Overview
- Identifying Your Upgrade Path
- Planning for an Upgrade
- Upgrade Do's and Don'ts.

To perform an upgrade, refer to these sections:

- Performing a Rolling Migration to HACMP 5.4.1
- Upgrading to HACMP 5.4.1 Using a Snapshot
- Upgrading to HACMP 5.4.1 on an Offline Cluster
- Applying a PTF.

After performing a migration, read these sections:

- Verifying the Success of Your Migration
- Troubleshooting Your Migration
- Recovering Your Previous Version of HACMP.

Prerequisites

To understand the concepts in this section, you should have a basic knowledge of the following:

- HACMP (from high-level concepts to low-level tasks, such as planning, maintenance, and troubleshooting), since the procedures in this chapter build on that knowledge.
- The version of HACMP from which you are upgrading. In particular, you should know how to use the cluster snapshot utility.
- The version of HACMP you are installing. It is helpful to have copies of any paper worksheets you filled in as part of the planning process with information about the cluster nodes, networks, subnets, IP labels, applications, resource groups, and pre- and post-event scripts that are currently used.

  For information on the planning worksheets, see Appendix A in the *Planning Guide*.

Terminology Overview

This chapter uses the following terms and acronyms:

Term	Definition
cllockd	Cluster Lock Manager daemon. This daemon was last supported in HACMP 5.1.
HACMP (HAS)	HACMP, also known as HACMP classic, is an option of HACMP without Enhanced Scalability. This option exists for all versions of HACMP before HACMP 5.1. Starting with HACMP 5.1, the HACMP classic option was merged with the HACMP/ES option and both renamed to HACMP 5.1.
HACMP/ES	HACMP/ES is an option of HACMP with Enhanced Scalability. The Enhanced Scalability options exists for all versions, starting with HACMP 4.1.
Hybrid cluster or mixed cluster	Nodes in a cluster running two different versions of HACMP.
Migration	The process of upgrading an existing HACMP cluster to the HACMP current level. (Used interchangeably with the term “upgrade.”)
ODM	Object Data Model, also known as the HACMP configuration database.
Offline migration	A type of upgrade where HACMP is brought offline on all nodes prior to performing the upgrade. During this time, resources are *not* available.
Rolling migration	A type of upgrade from one HACMP version to another during which cluster services are *not* stopped on all nodes in the cluster. Cluster services are stopped on one node at a time, that node is upgraded and reintegrated into the cluster before the next node is upgraded. Starting with HACMP 5.4, when using rolling migration to upgrade HACMP software on an individual node, you may choose to keep the applications and the resources running continuously on that node, though they will *not* be highly available during the upgrade.
Snapshot conversion	A type of upgrade from one HACMP version to another during which you take a snapshot of the current cluster configuration, stop cluster services on all nodes, install the next version of HACMP, and then convert the snapshot by running the **clconvert_snapshot** utility. See also Upgrading Using a Snapshot.
SWVPD	Software Vital Product Data, a set of installable software product filesets.

Identifying Your Upgrade Path

To identify your upgrade path, determine the version of HACMP from which you are migrating and whether it is appropriate for cluster service to be running.

Supported Upgrade Paths

HACMP 5.4.1 supports the upgrade scenarios from HACMP versions 5.1, 5.2, and 5.3 using either rolling migration or snapshot conversion.

Upgrading from Pre-5.1 Versions of HACMP

To upgrade to HACMP 5.4.1 from versions *earlier* than 5.1, you must first upgrade to one of the supported versions and then upgrade to HACMP 5.4.1.

For example, to upgrade from HACMP/ES 4.5, the following upgrade paths must be executed:

1. Upgrade from HACMP/ES 4.5 to HACMP 5.1
2. Upgrade from HACMP 5.1 to HACMP 5.4.1.

For up-to-date information about any available APARs, see the IBM Web site. Apply all applicable APARs after you upgrade to the new version of HACMP.

Choosing Your Upgrade Scenario

Next, identify the way in which you want to perform the upgrade with cluster services running or with them stopped. Both methods are described in this section.

Note: If you are upgrading from a previous release to HACMP 5.4.1, you must reboot the node after the installation. If you are upgrading HACMP 5.4.1 with a PTF, you are *not* required to reboot the node.

Upgrading While Keeping the HACMP Cluster Services Running

You can perform a nondisruptive upgrade of HACMP software, keeping cluster services running and the application continuously available throughout the upgrade if you are applying a PTF to HACMP 5.4.1.

To upgrade the cluster while keeping cluster services running:

1. Stop cluster services on one cluster node and choose the Move Resource Group option.
2. Upgrade the HACMP software on the node.
3. Reintegrate the node into the cluster by restarting cluster services on the node.
4. Repeat steps 1-3 for all nodes in the cluster.

Upgrading While Stopping the HACMP Cluster Services

If you have a maintenance window during which you can temporarily stop cluster services on all nodes, you can upgrade all nodes to HACMP 5.4.1. Although you can upgrade without using a snapshot, taking a snapshot beforehand is always recommended.

Upgrading Using a Snapshot

An upgrade using a snapshot is also referred to as a *snapshot conversion*.

To upgrade the cluster using a snapshot:

1. In an active cluster that predates HACMP 5.4.1, create and save the cluster snapshot.
2. Stop cluster services on all cluster nodes.
3. Remove the current version of the HACMP software on all nodes.
4. Install the HACMP 5.4.1 cluster software on all nodes.
5. Convert the previously saved snapshot by using the **clconvert_snapshot** utility. For information on this utility, see Appendix C: HACMP for AIX Commands in the *Administration Guide*.
6. Apply the snapshot to the cluster with the newly installed version of HACMP.
7. Verify and start cluster services one node at a time.

Upgrading without Using a Snapshot

Upgrade the cluster *without using a snapshot*: To upgrade the cluster without using a snapshot, follow the steps as described in the section Upgrading to HACMP 5.4.1 on an Offline Cluster.

To upgrade the cluster without using a snapshot:

1. Stop the HACMP cluster services on all cluster nodes.
2. Upgrade the HACMP software on each node.
3. Start cluster services on one node at a time.

Planning for an Upgrade

To properly plan for an upgrade to HACMP 5.4.1, follow the steps listed in the following sections.

Updating the Current Cluster Configuration

To update the current cluster configuration, do the following:

1. Check that all nodes in the cluster are up and running the same and most recent version of the HACMP software. Check the IBM web site for the latest HACMP APARs and PTFs available for the current version.
2. Review the installation prerequisites for HACMP 5.4.1 and ensure that the system being upgraded meets these requirements. See the IBM web site for the latest software levels.
3. If needed, upgrade the AIX operating system and RSCT before upgrading HACMP.

Checking Cluster Condition, Settings, and Worksheets

Do the following:

1. Use **clstat** to review the cluster state and to make certain that the cluster is in a stable state. For more information on the utility **clstat**, see the section on Monitoring Clusters with **clstat** in the *Administration Guide*.
2. Review the **/etc/hosts** file on each node to make certain it is correct.
3. Take a snapshot of each node configuration.
4. Save the planning worksheets (paper or online) as a reference. Transfer to new worksheets all information about your existing installation and any changes you plan to make after the upgrade.
5. Ensure that each cluster node has its own HACMP license. Otherwise, contact an IBM representative about licensing HACMP.
6. Ensure you have privileges to perform the installation as the root user, or ask your system administrator to make the appropriate changes.

Reviewing the Cluster and Node Version in the HACMP Configuration Database

To review the version of your cluster:

1. Run `odmget HACMP <cluster name>` or `odmget HACMP <node name>`.

   The versions in **HACMPcluster** and **HACMPnode** classes in the HACMP configuration database (ODM) for previous versions of HACMP are shown in the following table. After migration to HACMP 5.4.1, the cluster versions noted in the table will be equal to 9.

   The following table shows the versions of HACMP in the HACMP configuration database:

HACMP Version	In HACMPcluster	In HACMPnode
HACMP 5.3	cluster_version = 8	version = 8
HACMP 5.2	cluster_version = 7	version = 7
HACMP 5.1	cluster_version = 6	version = 6
HACMP/ES 4.5	cluster_version = 5	version = 5
HACMP 4.4.1	cluster_version = 4	version = 4

Checking Types of Networks

Make sure that HACMP 5.4.1 supports the types of networks that you plan to use. Remove or change unsupported types before you upgrade the HACMP software.

If your previous configuration includes unsupported network types and you attempt to upgrade a node to HACMP 5.4.1, the installation will fail and an error message will notify you to change the unsupported network type.

Migrating Resource Groups Configured without Sites

Refer to the following table for information on how existing rotating, cascading, and concurrent resource groups are migrated to HACMP 5.4.1:

Resource Group Definition before HACMP 5.2	Resource Group Definition in HACMP 5.4.1
Cascading resource group Inactive Takeover = False Cascading without Fallback (CWOF)= False	• **Startup:** Online on Home Node Only • **Fallover:** Fallover to Next Priority Node in the List • **Fallback:** Fallback to Higher Priority Node in the List
Cascading resource group Inactive Takeover = True CWOF = False	• **Startup:** Online on First Available Node • **Fallover:** Fallover to Next Priority Node in the List • **Fallback:** Fallback to Higher Priority Node in the List
Cascading resource group Inactive Takeover = False CWOF = True	• **Startup:** Online on Home Node Only • **Fallover:** Fallover to Next Priority Node in the List • **Fallback:** Never Fallback
Cascading resource group Inactive Takeover = True CWOF = True	• **Startup:** Online on First Available Node • **Fallover:** Fallover to Next Priority Node in the List • **Fallback:** Never Fallback
Rotating resource group	• **Startup:** Online Using Node Distribution Policy • **Fallover:** Next Priority Node in the List • **Fallback:** Never Fallback
Concurrent resource group	• **Startup:** Online on All Available Nodes • **Fallover:** Bring Offline (On Error Node Only) • **Fallback:** Never Fallback

New Resource Groups

If you upgraded from pre-5.3 versions of HACMP and plan to add new resource groups to an HACMP 5.4.1 cluster, refer to the mapping table in this chapter for information about the combinations of startup, fallover and fallback policies for resource groups.

Resource Group Distribution Policy during Migration

In HACMP 5.3 and above, the resource group distribution policy is a cluster-wide attribute that has only one option, the node-based distribution policy. This ensures that only one resource group that uses this distribution policy is brought online on a node during a node startup.

The following statements apply to newly-created resource groups that are based on rotating resource groups in releases prior to HACMP 5.1, or to upgraded rotating resource groups:

- Rotating resource groups migrated from HACMP 5.1 have the node-based distribution policy.
- If you are planning to use a single-adapter network that will be configured with IPAT via Replacement, set the startup policy for your resource group to Online using Distribution Policy.

Note: The network-based distribution policy previously available was deprecated in HACMP 5.3. When you migrate it, it is converted to the node-based policy.

HACMP Configuration Database Security Changes May Affect Migration

The HACMP Configuration Database (ODM) has the following security enhancements:

- Ownership. All HACMP ODM files are owned by the root user and the **hacmp** group. In addition, all HACMP binaries that are intended for use by non-root users are owned by root user and the **hacmp** group.
- Permissions. The **hacmpdisksubsystem** file is set with `600` permissions. Most of the other HACMP ODM files are set with `640` permissions (the root user can read and write, while the **hacmp** group can only read). All HACMP binaries that are intended for use by non-root users are installed with `2555` permissions (readable and executable by all users, with the **setgid** bit turned on so that the program runs as **hacmp** group).

During the installation, HACMP creates the **hacmp** group on all nodes. By default, the **hacmp** group has permission to read the HACMP ODMs, but does *not* have any other special authority. For security reasons, do *not* to expand the authority of the **hacmp** group.

If you use programs that access the HACMP ODMs directly, you may need to rewrite them if they are intended to be run by non-root users:

- All access to the ODM data by non-root users should be handled via the provided HACMP utilities.
- In addition, if you are using the PSSP File Collections facility to maintain the consistency of **/etc/group**, the new **hacmp** group that is created at installation time on the individual cluster nodes may be lost when the next file synchronization occurs.

  To prevent overwriting your **hacmp** group, before installing HACMP 5.4.1, either:

  - Turn off the PSSP File Collections synchronization of the **/etc/group** file

  OR

  - Include the **hacmp** group in the master **/etc/group** file and propagate this change to all cluster nodes.

Upgrading Applications Configured Using a Smart Assist

The framework for applications configured with Smart Assists has been enhanced in HACMP 5.4.1. DB2 and WebSphere applications configured in HACMP 5.3 with the DB2 and WebSphere Smart Assists are converted to use the new infrastructure during an upgrade. After the upgrade, you can manage these applications (change/show/remove) using the new HACMP 5.4.1 SMIT Standard and Initialization panels.

Applications previously configured with either the Two-Node Configuration Assistant or the Oracle Smart Assist are *not* converted to the new infrastructure. Continue to use the HACMP Extended SMIT path for managing those applications.

Ensuring Applications Use Proprietary Locking Mechanism

Before upgrading to HACMP 5.4.1, *make sure that applications use their proprietary locking mechanism*. The Cluster Lock Manager (**cllockd**, or **cllockdES**) *is not supported* in HACMP 5.4.1. Check with the application vendor about the concurrent access support.

Installing HACMP 5.4.1 removes the Lock Manager files and definitions from a node. After a node is upgraded, the Lock Manager state information in SNMP and **clinfo** shows that the Lock Manager is inactive on a node, whether or *not* the Lock Manager is running on a back-level node.

Addressing Concurrent Access (AIX) Migration Issues

Before migrating concurrent volume groups to HACMP 5.4.1, decide whether or *not* to convert them to enhanced concurrent mode, the default for volume groups created in AIX v.5.2 and up.

Note that:

- Enhanced concurrent mode is supported only on AIX v.5.2 and up.
- SSA concurrent mode is *not* supported on 64-bit kernels.
- All nodes in a cluster must use the same form of concurrent mode for a specified volume group.

This means that if your cluster included SSA disks in concurrent mode on AIX 4.3.3, and then in preparation for upgrading to HACMP 5.4.1, you upgrade a node to AIX v.5.2 and chose the 64-bit kernel, you will *not* be able to bring the concurrent volume group online on all cluster nodes. If you have SSA disks in concurrent mode, you cannot run 64-bit kernels until you have converted all volume groups to enhanced concurrent mode.

Priority Override Location and Persistence

If you are upgrading from a previous release, HACMP 5.4.1 handles Priority Override Location (POL) and persistence differently than earlier releases did:

- The Priority Override Location (POL) setting is *not* used.

  The POL setting is *not* used for resource groups that you move from one node to another. In general, in HACMP 5.4.1, if you move a resource group from one node to another, it remains on its new node until you need to move it again.

- If you reboot the cluster (which you seldom need to do in HACMP 5.4.1), the group returns to the node that is originally configured as its highest priority node in the nodelist (if the group has a fallback policy that tells it to fall back).
- If you do *not* reboot the cluster, the group remains on the node to which you moved it, and, if it has a fallback policy to fall back, then it falls back to its "acting" highest priority node.

- The persistence after a reboot is *not* retained.

  If a resource group in your cluster has a fallback policy with the option Persist Across Cluster Reboot and resides on the node to which you moved it before an upgrade, when you upgrade to HACMP 5.4.1, then the resource group remains on its destination node after the upgrade. In this case, you did *not* reboot.

  However, if you reboot the cluster, the group returns to the node that is its originally configured highest priority node.

Note: If you want the group to be permanently hosted on the node its originally configured highest priority node, change the highest priority node in the nodelist for the resource group.

Notes on Upgrading from HACMP 5.1

If you are upgrading from HACMP 5.1 to HACMP 5.4.1, the procedures are the same as described in the section Performing a Rolling Migration to HACMP 5.4.1 and Upgrading to HACMP 5.4.1 Using a Snapshot. However, some issues and terminology that are specific to upgrading from HACMP 5.1 are described below.

Environment Variables Are Not Carried Over after Migration from Earlier Versions

This problem may affect you if you used environment variables in your pre- and post-event scripts in releases prior to HACMP 5.1. In general, the shell environment variables declared in **/etc/environment** are much more strictly managed under HACMP in versions 5.1 and higher.

If, for instance, you previously used HACMP 4.5 (also known as HACMP classic), and had pre- and post-event scripts that used environment variables, such as LOGNAME, be aware that once you start your application with HACMP/ES, that is HACMP v.5.1 or higher, those environment variables are overwritten (empty) and become unavailable to you. In other words, upon an upgrade to HACMP v.5.1 or higher, the environment variables will no longer be set in your pre- and post-event scripts.

Note that in general, when writing your HACMP pre- or post-events or using your pre- and post-evens written for previous versions of HACMP, none of the shell environment variables defined in **/etc/environment** are available to your program. If you need to use any of these variables, explicitly source them by including this line in your script:
“`. /etc/environment`”.

See the chapter on Cluster Events in the *Planning Guide* for more information on pre- and post-event scripts.

Security Mode Name Change

The Enhanced security mode in HACMP has been renamed to Kerberos security. The function remains the same.

Migrating HACMP 5.1 Dynamic Node Priority Fallback Policy

If your configuration includes an HACMP 5.1 resource group with a dynamic node priority policy for fallover, migration or conversion to HACMP 5.4.1 changes this setting. The **cl_convert** utility changes the fallover policy for that resource group to Fallover to Next Priority Node when migration or conversion is complete. During migration, the default node fallover policy is used.

Migrating HACMP 5.1 Existing Resource Groups to HACMP 5.4.1

In HACMP 5.4.1, all types of groups are referred to as resource groups. During an upgrade, the cascading, rotating and concurrent resource groups that you configured in HACMP 5.1 are converted to resource groups with a specific set of startup, fallover and fallback policies. All functional capabilities of resource groups are retained in HACMP 5.4.1. For resource groups mapping information, see Migrating Resource Groups Configured without Sites.

Migrating HACMP 5.1 Resource Groups Configured with Sites

Starting with HACMP 5.2, inter-site management policy names changed. This table summarizes the mapping between the inter-site policies in HACMP 5.1 and HACMP 5.4.1 for resource groups configured with sites:

Inter-Site Policy in HACMP 5.1	Inter-Site Policy in HACMP 5.4.1
Cascading	Prefer Primary Site
Rotating	Online on Either Site
Concurrent	Online on Both Sites

HACMP 5.4.1 Inter-Site Selective Fallover for Resource Group Recovery

Selective fallover of resource groups between sites is disabled by default when you upgrade to HACMP 5.4.1 from a previous release. This is the pre-5.4.1 release behavior for non-IGNORE site management policy. A particular instance of a resource group can fall over within one site, but cannot move between sites. If no nodes are available on the site where the affected instance resides, that instance goes into ERROR or ERROR_SECONDARY state. It does *not* stay on the node where it failed. This behavior applies to both primary and secondary instances.

Note that in HACMP 5.3 and above, though the Cluster Manager does *not* initiate a selective fallover across sites by default, it still moves the resource group if a node_down or node_up event occurs, and you can manually move a resource group between sites.

For a new install of HACMP 5.4.1, inter-site resource group recovery is enabled by default.

You can change the default behavior after the migration and installation of HACMP 5.4.1 is complete, using the HACMP SMIT path:

Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Customize Resource Group and Resource Recovery > Customize Inter-Site Resource Group Recovery.

Upgrade Do's and Don'ts

This section lists the major tasks to do during or before an upgrade, and it also lists what *not* to do.

Upgrade Do's

Make sure that you do the following:

1. Take a cluster snapshot and save it to the **/tmp** directory as well as to another machine and CD.
2. Save a copy of any event script files to the **/tmp** directory as well as to another machine and CD.
3. Ensure that the same level of cluster software (including PTFs) are on all nodes before beginning a migration.
4. Ensure that the cluster software is committed (and *not* just applied). See Checking That the Software Is Committed.

Upgrade Don'ts

During any type of upgrade, do *not* do the following:

- Do *not* save your cluster configuration or customized event scripts under these directory paths: **/usr/sbin/cluster,** /**usr/es/sbin/cluster** or **/usr/lpp/cluster**. Data in these directories may be lost during an upgrade. Instead, copy files to a separate machine or to a CD.
- Do *not* synchronize the cluster.
- When migrating from HACMP 5.3 to 5.4.1, do *not* stop a node and place resource groups in an UNMANAGED state.
- Do *not* attempt a DARE or a C-SPOC command. For example, do *not* change node priority in a resource group, add or remove a network, update an LVM component, or add users to a cluster.
- Do *not* leave the cluster in a hybrid state for an extended period of time.

When migrating an *active* cluster, one node at a time (a rolling migration), the use of commands and functions are restricted as follows when the cluster has mixed versions (is in a *hybrid* state):

- Do *not* change the cluster topology or configuration.
- Do *not* verify and synchronize the cluster configuration.
- Do *not* use any **System Management (C-SPOC)** functions except for the **Manage HACMP Services** functions.
- Do *not* use the **Problem Determination Tools** > **View Current State** function.
- Do *not* use the **Extended Configuration > Snapshot Configuration** > **Add a Cluster Snapshot** option or run the **clsnapshot** command.
- Do *not* use the **Problem Determination Tools > Recover From HACMP Script Failure** option or run the **clruncmd** command, except when running the command or SMIT option from the target node specified by the command.

Checking That the Software Is Committed

Before upgrading the cluster to HACMP 5.4.1, ensure that the current software installation is committed (*not* just applied).

To ensure that the software is already committed:

1. Run the `lslpp -h cluster.*` command.
2. If the word **APPLY** displays under the action header, enter `smit install_commit` before installing the HACMP software.

   SMIT displays the **Commit Applied Software Updates (Remove Saved Files)** panel.
3. Enter field values as follows:

SOFTWARE name	From the picklist, select all cluster filesets.
COMMIT old version if above version used it?	Set this field to **yes**.
EXTEND filesystem if space needed?	Set this field to **yes**.

Performing a Rolling Migration to HACMP 5.4.1

This section describes how to upgrade an HACMP cluster to HACMP 5.4.1 while keeping cluster services running on the nodes. During the upgrade window, the cluster will be running in a hybrid state where HACMP 5.3 cluster components run in unison with HACMP 5.4.1 cluster components in processing any specific cluster events that may occur.

Before continuing, review the section Security Mode Name Change.

Step 1: Stop Cluster Services on a Node Hosting the Application

Once all HACMP cluster nodes that have a previous version installed are up and the cluster is stable, stop cluster services on node A using the **graceful with takeover** option. (Starting with HACMP 5.4, this is known as stopping cluster services and moving the resources groups to other nodes.) If needed, install a new version of RSCT.

1. Enter `smit hacmp`
2. Use the **System Management (C-SPOC) > Manage HACMP Services > Stop Cluster Services** SMIT menu to stop cluster services.
3. Select **takeover** for **Shutdown mode**.
4. Select local node only and press Enter.

For example, if application X is running on node A, stopping node A gracefully with takeover causes the resource group containing the application to fall over to node B. After fallover is complete, the upgrade of node A to HACMP 5.4.1 can continue.

Step 2: Install the HACMP 5.4.1 Software

On node A, install HACMP 5.4.1, which converts the previous HACMP configuration database (ODMs) to the HACMP 5.4.1 format. This installation process uses the **cl_convert** utility and creates the **/tmp/clconvert.log** file. A previously created version of the file is overwritten.

To install the HACMP software:

1. Enter `smit install`
2. In SMIT, select **Install and Update Software > Update Installed Software to Latest Level (Update All)** and press Enter.
3. Enter the values for **Preview only?** and **Accept new license agreements?** For all other field values, select the defaults.

Preview only?	Select **no**
Accept new license agreements?	Select **yes**

4. Press Enter.

Step 3: Reboot the Node

Reboot the node with the `shutdown -Fr` command.

Note: If you are applying a PTF to HACMP 5.4.1, this command is *not* needed.

Step 4: Start Cluster Services on the Upgraded Node

Start cluster services on node A. Node A is running HACMP 5.4.1 while nodes B, C, and others are running HACMP with an earlier version.

To start cluster services on a single upgraded node:

1. Enter `smit clstart`
2. Enter field values as follows and press Enter

Start now, on system restart or both	Select **now**
Start Cluster Services on these nodes	Select **local node** (Default)
Manage Resource Groups Automatically/Manually	Select **Automatically** HACMP brings resource group(s) online according to the resource groups' configuration settings and the current cluster state, and starts monitoring the resource group(s) and applications for availability.
BROADCAST message at startup?	Select **false**
Startup Cluster Information Daemon?	Select **true**
Ignore verification errors?	Select **false**

Automatically correct errors found during cluster start?	Whatever value is in these fields will not make sense since it is a mixed cluster.

Note: Verification is *not* supported on a mixed version cluster. Run verification only when all nodes have been upgraded.

Step 5: Repeat Steps for Other Cluster Nodes

Repeat steps 1–3 for the remaining cluster nodes, one node at a time.

Step 6: Verify the Upgraded Cluster Definition

Verification provides errors or warnings to ensure that the cluster definition is the same on all nodes. You can verify and synchronize a cluster only when all nodes in the cluster are running the same version of the software.

To verify the cluster:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Verification and Synchronization** and press Enter.
3. Change **Verify Changes Only** to **yes**.

Upgrading to HACMP 5.4.1 Using a Snapshot

This section describes how to upgrade a cluster with an earlier version of HACMP to HACMP 5.4.1 using a cluster snapshot that was created on the configuration with an earlier version of HACMP. This upgrade path requires cluster services to be down simultaneously on all nodes; as a result, your applications *will not be* highly available during the upgrade window.

Step 1: Creating a Snapshot

While all HACMP 5.3 cluster nodes are up and the cluster is stable, create a **snapshot.odm** file on node A. In addition, copy the **snapshot.odm** file to the **/tmp** directory.

For instructions on creating a snapshot, see the chapter on Saving and Restoring Cluster Configurations in the *Administration Guide*.

Step 2: Stopping All Nodes

Stop all nodes, one at a time, gracefully (Starting with HACMP 5.4, stopping gracefully will be referred to as stopping cluster services and taking resource groups offline.)

1. Enter `smit hacmp`
2. Use the **System Management (C-SPOC) > Manage HACMP Services > Stop Cluster Services** SMIT menu to stop cluster services.

Step 3: Removing the Existing Version of the HACMP Software

Before you remove the HACMP software from a system, make sure that HACMP cluster services are stopped. You cannot remove the HACMP software when the cluster is running: The system displays an error message and prevents removal of core filesets.

To remove the existing version:

1. Ensure that cluster services have been stopped on all nodes.
2. Enter `smit install_remove`
3. Enter field values in the **Remove Installed Software** SMIT panel as follows and press Enter:

SOFTWARE name	Enter `cluster` to remove all server and client software, or use the picklist to select individual options. Your selections appear in this field.
REMOVE dependent software?	Select **no**.
EXTEND filesystems if space needed?	Select **yes**.
DETAILED output?	Select **no**.

Note: If HAGEO is installed, you must also remove it from your system. Use the same steps as listed above, using `hageo` as the software name.

Step 4: Installing the HACMP 5.4.1 Software

To install the HACMP 5.4.1 software on each cluster node:

1. Enter `smit install_all`
2. In SMIT, select **cluster.es.server.rte** and any additional component you would like to install. Press Enter.
3. Enter the values for **Preview only?** and **Accept new license agreements?** For all other field values, select the defaults.

Preview only?	Select **no**
Accept new license agreements?	Select **yes**

4. Press Enter.

Step 5: Converting and Applying a Saved Snapshot

After you install HACMP 5.4.1 on all cluster nodes, convert and apply the snapshot on the same node where the cluster snapshot was added.

Use the **clconvert_snapshot** utility to convert the cluster snapshot. The **clconvert_snapshot** utility converts your existing **snapshot.odm** file to the HACMP 5.4.1 format and saves a backup **snapshot.odm.old** file in the HACMP format that was previously installed.

In the following example, *version* is the HACMP version number and *snapshotfile* is the name of your snapshot file. The snapshot file is stored in the directory specified by the $SNAPSHOTPATH variable that by default is **/usr/es/sbin/cluster/snapshots**:

```
clconvert_snapshot  -v version -s snapshotfile
```

For example, if converting a snapshot called *my530snapshot* use:

```
clconvert_snapshot -v 5.3 -s my530snapshot
```

For information on the **clconvert_snapshot** utility, see Saving and Restoring Cluster Configurations in the *Administration Guide*.

Step 6: Verifying the Upgraded Cluster Definition

After the HACMP software is installed on all nodes, and all nodes have been rebooted, verify and synchronize the cluster configuration. Verification provides errors or warnings to ensure that the cluster definition is the same on all nodes. You can verify and synchronize a cluster only when all nodes in the cluster are running the same version of the software.

To verify the cluster:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Verification and Synchronization** and press Enter.
3. Change **Verify Changes Only** to **yes**.
4. Start nodes one at a time.

Post-Upgrade Check: Troubleshooting a Stalled Snapshot Application

If you upgrade to HACMP 5.4.1 using the previously created snapshot, upgrading to the new version may stall if the cluster configuration in the snapshot is *not* 100% accurate according to verification. (Also, dynamic cluster reconfiguration does *not* work if **verification** finds any errors).

If you apply a snapshot and see an error, review the log files to check if it can be automatically corrected in HACMP 5.4.1 by **the verification utility**. If the error is included in the list, to continue the upgrade process, force apply the snapshot and run the cluster synchronization and verification process, with the option **Automatically Correct Errors during the Cluster Verification** set to **Interactively**.

Note: Be careful when applying the snapshot forcefully; only use this option if you are sure that the error you encountered can be automatically corrected.

You may see the following warnings and errors:

```
WARNING: "The NFS mount/Filesystem specified for resource group rg1 is
using incorrect syntax for specifying an NFS cross mount: /mnt/fs1".

ERROR: "Disk Heartbeat Networks have been defined, but no Disk Heartbeat
devices. You must configure one device for each node in order for a Disk
Heartbeat network to function".
```

In these cases, apply the snapshot forcefully to continue an upgrade process to HACMP 5.4.1. Although the upgrade process via a snapshot fails (there is no automatic corrective action for them), the cluster remains intact, therefore, force applying the snapshot is safe.

Upgrading to HACMP 5.4.1 on an Offline Cluster

It is possible to bring cluster services down on all of the nodes and then migrate the cluster definitions individually on each node. This process is supported for the following versions:

- HACMP 5.1
- HACMP 5.2
- HACMP 5.3.

To bring a cluster offline and upgrade the HACMP software on the cluster nodes, complete the following procedure:

1. Stop cluster services on all nodes.
2. Ensure that cluster services have been stopped on all nodes.
3. Install the new HACMP software on each node.
4. Review the **/tmp/clconvert.log** file to ensure that a conversion of the HACMP ODMs has occurred.
5. Start cluster services, one node at a time, and ensure that each node successfully joins the cluster.

See also sections on Performing a Rolling Migration to HACMP 5.4.1, Testing An Upgraded Cluster, and Upgrading to HACMP 5.4.1 Using a Snapshot.

Applying a PTF

This section describes how to apply a Program Temporary Fix (PTF)—a correction for a software problem reported in an APAR. Starting with HACMP 5.4, you can apply an HACMP 5.4.1 PTF update on an individual node using *rolling migration.* During the upgrade your critical applications and resources continue to run on that node, though they will *not* be highly available.

Before continuing, review the section Security Mode Name Change.

Step 1: Stop Cluster Services on a Node Hosting the Application

Once all HACMP 5.4.1 cluster nodes are up and the cluster is stable, bring node A offline without bringing its associated resource groups offline by stopping cluster services using the **Unmanage Resource Groups** option. (Prior to HACMP 5.4, this was known as forced down.)

1. Enter `smit hacmp`
2. In SMIT, select the **System Management (C-SPOC) > Manage HACMP Services > Stop Cluster Services** SMIT menu to stop cluster services. Use the **Unmanage Resource Groups** option.

3. Enter field values. The example below displays only the settings that are different from the default settings.

Stop now, on system restart or both	Select to stop cluster services **now**
BROADCAST cluster shutdown?	Indicate whether you want to send a broadcast message to users before the cluster services stop. If you specify **true**, a message is broadcast on all cluster nodes.
Unmanage Resource Groups	Cluster services are stopped immediately. Resources that are online on the node are *not* released. Applications continue to run, though they are *not* highly available. If you stop the cluster services with this option, failures are *not* detected. This applies to hardware resources such as disks and adapters as well as any monitored applications. If any resources defined to HACMP are disrupted, they should be restored to their prior state before restarting cluster services. For example, if application X is running on node A, when stopping cluster services on node A using the **Unmanage Resource Groups** option, the resource group and the application continue to run on that node, though they are *not* monitored or managed by HACMP.

For more information on stopping cluster services, see Understanding Stopping Cluster Services in the *Administration Guide*.

Note: HACMP cluster services are *not* monitoring the applications at some periods during the rolling migration—applications continue to run, but HACMP cluster services are suspended on the nodes—your applications may potentially fail at this time. If your application *must* be highly available, we recommend that you do *not* keep the application in the `UNMANAGED` state for a long period of time.

Step 2: Install the PTF Update

On node A, apply the PTF update, which updates the HACMP 5.4.1 configuration database (ODMs). This installation process uses the **cl_convert** utility and creates the **/tmp/clconvert.log** file. A previously created version of the file is overwritten.

To install the HACMP software:

1. Enter `smit install`
2. In SMIT, select **Install and Update Software > Update Installed Software to Latest Level (Update All)** and press Enter.

3. Enter the values for **Preview only?** and **Accept new license agreements?** For all other field values, select the defaults.

Preview only?	Select **no**
Accept new license agreements?	Select **Yes**

4. Press Enter.

Step 3: Start Cluster Services on the Upgraded Node

Start cluster services on node A. Node A is running the updated HACMP version while nodes B, C, and others are running the previous version of HACMP 5.4.1. HACMP starts monitoring the running applications and resources on Node A.

To start cluster services on a single upgraded node:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Manage HACMP Services > Start Cluster Services** and press Enter.
3. Enter field values as follows and press Enter

Start now, on system restart or both	Select **now**
Start Cluster Services on these nodes	Select **local node**
Manage Resource Groups Automatically/Manually	Select **Automatically** HACMP brings resource group(s) online according to the resource groups' configuration settings and the current cluster state, and starts monitoring the resource group(s) and applications for availability.
BROADCAST message at startup?	Select **false**
Startup Cluster Information Daemon?	Select **true**
Ignore verification errors?	Select **false**
Automatically correct errors found during cluster start?	Whatever value is in these fields will not make sense since it is a mixed cluster.

Note: Verification is *not* supported on a mixed version cluster. Run verification only when all nodes have been upgraded.

Note: After applying a PTF update, a warning may display stating the Cluster Manager daemon did *not* start. Use the command **lssrc -ls clstrmgrES** to verify the Cluster Manager started successfully.

Step 4: Repeat Steps for Other Cluster Nodes

Repeat steps 1–3 for the remaining cluster nodes, one node at a time.

Step 5: Verify the Upgraded Cluster Definition

Once all nodes are up and the cluster is stable, run cluster verification on the upgraded HACMP cluster.

Verification provides errors or warnings to ensure that the cluster definition is the same on all nodes. You can verify and synchronize a cluster only when all nodes in the cluster are running the same version of the software.

To verify the cluster:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Verification and Synchronization** and press Enter.
3. Change **Verify Changes Only** to **yes**.

Additional Migration Tasks

After you complete the upgrade to HACMP 5.4.1, you may need complete additional tasks such as:

- Recompiling Clinfo Clients after Migrating to HACMP 5.4.1
- Using a Web-Based Interface for Configuration and Management
- Resetting HACMP Tunable Values

Recompiling Clinfo Clients after Migrating to HACMP 5.4.1

Recompiling existing Clinfo applications is *not* required. However, in HACMP 5.4.1:

- **CL_MAXNAMELEN** is now 256 characters
- There are changes related to resource group information.

If you wish to incorporate these changes in your applications then make the desired modifications, recompile and link the applications using the Clinfo library. For updated information about the Clinfo C and C++ API routines, see *Programming Client Applications*.

Using a Web-Based Interface for Configuration and Management

HACMP 5.4.1 includes a web-based user interface, called WebSMIT, that provides consolidated access to the SMIT functions of configuration and management as well as to interactive cluster status and documentation.

To use the WebSMIT interface, you must configure and run a web server process on the cluster node(s) to be administered. See the **/usr/es/sbin/cluster/wsm/README** file for full information on basic web server configuration, the default security mechanisms in place when HACMP 5.4.1 is installed, and the configuration files available for customization.

Also see Installing and Configuring WebSMIT in Chapter 4: Installing HACMP on Server Nodes.

Resetting HACMP Tunable Values

In HACMP 5.4.1, you can change the settings for a list of tunable values that were changed during the cluster maintenance and reset them to their default settings, or installation-time cluster settings.

The installation-time cluster settings are equivalent to the values that appear in the cluster after manually installing HACMP.

Note: Resetting the tunable values does *not* change any other aspects of the configuration, while installing HACMP removes all user-configured configuration information including nodes, networks and resources.

List of Tunable Values

The following values can be reset:

- User-supplied information:
  - Network module tuning parameters, such as failure detection rate, grace period and heartbeat rate. HACMP resets these parameters to their installation-time default values.
  - Cluster event customizations, such as all changes to cluster events. Note that resetting changes to cluster events does *not* remove any files or scripts that the customization used, only the HACMP's knowledge of pre- and post-event scripts.
  - Cluster event rules. Any changes made to the event rules database are reset to their installation-time default values.
  - HACMP command customizations. Any changes to the default set of HACMP commands are reset to their installation-time defaults.
- Automatically generated and discovered information:

  Typically, you cannot see this information. HACMP rediscovers or regenerates this information when the cluster services are restarted or during the next cluster synchronization.

  HACMP resets the following:
  - Local node names stored in the cluster definition database
  - Netmasks for all cluster networks
  - Netmasks, interface names and aliases for heartbeat (if configured) for all cluster interfaces
  - SP switch information generated during the latest **node_up** event (this information is regenerated at the next **node_up** event)
  - Instance numbers and default log sizes for the RSCT subsystem.
  - For information about how to reset the tunable values using SMIT, see the *Administration Guide.*

Verifying the Success of Your Migration

Now that all nodes have been migrated, you need to complete the following tasks described in the following sections to ensure that everything is working correctly:

- Verifying Software Levels Installed Using AIX Commands
- Automatically Saved Files
- Verifying the Upgraded Cluster Definition
- Verifying All Cluster Filesets Have Migrated
- Testing HACMP on the Upgraded Cluster
- Running AIX Commands on the Migrated Cluster.

Verifying Software Levels Installed Using AIX Commands

Verify the software installation by using the **lppchk** AIX command, and check the installed directories to see that expected files are present.

The **lppchk** command verifies that files for an installable software product (fileset) match the Software Vital Product Data (SWVPD) database information for file sizes, **checksum** values, or symbolic links.

Run the commands **lppchk -v** and **lppchk -c "cluster.*"**

Both commands return nothing if the installation is OK.

Automatically Saved Files

The following files in the **/usr/lpp/save.config** directory are automatically saved during the upgrade process:

```
/usr/lpp/save.config/usr/es/sbin/cluster/events/node_up.rp
/usr/lpp/save.config/usr/es/sbin/cluster/events/node_down.rp
```

In addition, the following files are saved during the upgrade process and removed from the system at the end of migration:

```
 /lpp/cluster/objrepos/HACMPnpp
 /lpp/cluster/objrepos/HACMPude
```

If you are upgrading from HACMP/ES 4.5, HACMP also saves the following file in the **/usr/lpp/save.config** directory during the upgrade process:

```
/usr/lpp/save.config/usr/es/sbin/cluster/events/rg_move.rp
```

WARNING: Until the cluster is migrated to HACMP 5.4.1, *do not delete any of the files listed above*.

During the migration process, the SNMP MIBs are inconsistent between the nodes running HACMP 5.4.1 and the nodes running HACMP/ES 4.5. When migration is complete, all nodes use the same version of the MIBs.

Verifying the Upgraded Cluster Definition

After the HACMP 5.4.1 software is installed on all of the nodes in the cluster and cluster services restored, verify and synchronize the cluster configuration. Verification ensures that the cluster definition is the same on all nodes. You can verify and synchronize a cluster only when all nodes in the cluster are running the same version of the software.

To verify the cluster:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Verification and Synchronization > Verify Changes only** and press Enter.

Verifying All Cluster Filesets Have Migrated

Previous documentation APARs may *not* be successfully converted resulting in the inability to synchronize the cluster. Execute the following to verify all cluster filesets are at the expected level:

```
lslpp –l | "cluster.*"
```

Testing HACMP on the Upgraded Cluster

You can also use the Cluster TestTool available in HACMP 5.4.1 to test your cluster. For information about this tool see Chapter 8: Testing and HACMP Cluster in the *Administration Guide*.

Running AIX Commands on the Migrated Cluster

To determine which daemons are active on a cluster node, you can use the options on the SMIT **System Management (C-SPOC) > Manage HACMP Services** > **Show Cluster Services** panel or **lssrc** command as follows:

1. Run **`lssrc –ls topsvcs`**. The result show the internal state of the RSCT Daemon as follows:

```
Subsystem         Group            PID     Status
topsvcs           topsvcs          24726   active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
Ethernet_1_0   [ 0]    3    3  S 192.168.245.42
192.168.245.72
Ethernet_1_0   [ 0]             (192.168.245.40 )
Ethernet_1_0   [ 0] en1          0x417fc659      0x417fc93f
HB Interval = 1 secs. Sensitivity = 10 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 465057 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 681363 ICMP 0 Dropped: 0
NIM's PID: 26012
Token_Ring_1_0 [ 1]    3    3  S 192.168.248.40
192.168.248.70
Token_Ring_1_0 [ 1] tr0          0x417fc61a      0x417fc940
HB Interval = 1 secs. Sensitivity = 10 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 443287 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 661396 ICMP 0 Dropped: 0
NIM's PID: 22400
rs232_0        [ 2]    2    0  D 255.255.0.1
rs232_0        [ 2] tty0
HB Interval = 2 secs. Sensitivity = 5 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 88465 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 0 ICMP 0 Dropped: 0
NIM's PID: 27412
tmssa_1       [ 4]    2    2  S 255.255.2.4    255.255.2.4
tmssa_1        [ 4] ssa1         0x817fc94a      0x817fc951
HB Interval = 2 secs. Sensitivity = 5 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 319657 ICMP 0 Errors: 460 No mbuf: 0
Packets received: 319661 ICMP 0 Dropped: 0
NIM's PID: 27818
tmssa_2       [ 5]    2    2  S 255.255.2.5    255.255.2.5
tmssa_2        [ 5] ssa3         0x817fc620      0x817fc626
HB Interval = 2 secs. Sensitivity = 5 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 319855 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 319864 ICMP 0 Dropped: 0
NIM's PID: 14906
  2 locally connected Clients with PIDs:
haemd( 29630) hagsd( 30326)
  Dead Man Switch Enabled:
     reset interval = 1 seconds
     trip  interval = 20 seconds
  Configuration Instance = 20
  Default: HB Interval = 1 secs. Sensitivity = 4 missed
beats
  Daemon employs no security
  Segments pinned: Text Data.
  Text segment size: 730 KB. Static data segment size: 639
KB.
  Dynamic data segment size: 3849. Number of outstanding
malloc: 313
  User time 430 sec. System time 526 sec.
  Number of page faults: 0. Process swapped out 0 times.
  Number of nodes up: 3. Number of nodes down: 0.
```

2. Execute the **/usr/es/sbin/cluster/utilities/clshowsrv -v** utility, which produces results similar to the following:

```
Status of the RSCT subsystems used by HACMP:
Subsystem         Group            PID     Status
topsvcs           topsvcs          32182   active
grpsvcs           grpsvcs          33070   active
grpglsm           grpsvcs          30886   active
emsvcs            emsvcs           32628   active
emaixos           emsvcs           31942   active
ctrmc             rsct             14458   active

Status of the HACMP subsystems:
Subsystem Group            PID     Status
clcomdES         clcomdES         15778   active
clstrmgrES       cluster          32792   active

Status of the optional HACMP subsystems:
Subsystem       Group            PID     Status
clinfoES        cluster          31210   active
```

Troubleshooting Your Migration

For help with specific errors during your migration, see the following sections:

- Error: “config_too_long”
- Error: “cldare cannot be run”.

Error: “config_too_long”

When the migration process has completed and the HACMP filesets are being uninstalled, you may see a **config_too_long** message.

This message appears when the cluster manager detects that an event has been processing for more than the specified time. The **config_too_long** messages continue to be appended to the **hacmp.out** log until the event completes. If you observe these messages, periodically check that the event is indeed still running and has *not* failed.

Error: “cldare cannot be run”

Making configuration changes is not supported during the migration process. If you try to change the cluster topology or resources when migration is incomplete, the synchronization process fails. You receive a message similar to the following:

```
cldare: Migration from HACMP 5.1 to HACMP 5.4.1 Detected. cldare cannot
be run until migration has completed.
```

When migration is complete, you can apply changes or remove them.

To remove changes, restore the active HACMP configuration database:

1. Enter `smit hacmp`
2. In SMIT, select **Problem Determination Tools > Restore HACMP Configuration Database from Active Configuration**.

Recovering Your Previous Version of HACMP

If you want to completely undo your migration, see the following sections:

- Recovering from a Conversion Failure
- Recovering Configuration Information.

Recovering from a Conversion Failure

When you install HACMP, the **cl_convert** command runs automatically to convert the HACMP configuration database from a previous version of HACMP to that of the current version. If the installation fails, run **cl_convert** to convert the database.

In a failed conversion, run **cl_convert** using the **-F** flag. For example, to convert from HACMP/ES 5.1 to HACMP 5.3, use the **-F** and **-v** (version) flags as follows:

```
cl_convert -F -v 5.1.0
```

To run a conversion utility the following is required:

- Root user privileges
- The HACMP version from which you are converting.

The **cl_convert** utility records the conversion progress to the **/tmp/clconvert.log** file so you can gauge conversion success.

Recovering Configuration Information

The post-installation output informs you to merge site-specific configuration information into the newly installed files:

```
Some configuration files could not be automatically merged into the
system during the installation. The previous versions of these files
have been saved in a configuration directory as listed below. Compare
the saved files and the newly installed files to determine if you need
to recover configuration data. Consult your product documentation to
determine how to merge the data.

Configuration files that were saved in /usr/lpp/save.config:

/usr/es/sbin/cluster/etc/rc.cluster
/usr/es/sbin/cluster/samples/clinfo.rc
/usr/es/sbin/cluster/samples/pager/sample.txt
/usr/es/sbin/cluster/etc/clinfo.rc
/usr/es/sbin/cluster/utilities/clexit.rc
/usr/es/sbin/cluster/etc/clhosts
/usr/es/sbin/cluster/etc/rc.shutdown
/usr/es/sbin/cluster/diag/clconraid.dat
/usr/es/sbin/cluster/etc/hacmp.term
/etc/cluster/lunreset.lst
/etc/cluster/disktype.lst
```

Chapter 4: Installing HACMP on Server Nodes

This chapter lists the prerequisites for the HACMP software and describes how to install it. This chapter contains the following sections:

- Prerequisites
- Contents of the Installation Medium
- HACMP Installation Choices
- Installing from an Installation Server
- Installing from a Hard Disk
- Installing from the Installation Medium
- Completing the Installation
- Addressing Problems during the Installation
- Removing the HACMP Software
- Installing and Configuring WebSMIT
- Where You Go from Here

If you are upgrading an HACMP installation, see Chapter 3: Upgrading an HACMP Cluster.

If you are installing HACMP on a client, see Chapter 5: Installing HACMP on Client Nodes.

Prerequisites

Before you install the HACMP software:

- Read chapters 1–8 of this guide and complete the planning worksheets and diagrams.
- Ensure that your system meets the requirements listed in this chapter.

Supported Hardware

To ensure that your system meets the guidelines established for HACMP, contact your sales representative, or see the IBM sales guide using the following URL:

http://www.ibm.com/common/ssi

For information about the disks supported for concurrent resource manager, see the chapter Planning Shared Disk and Tape Devices in the *Planning Guide.*

Notes that RAID devices from other manufacturers may *not* support concurrent access. See Appendix B: OEM Disk, Volume Group, and File Systems Accommodation for more information.

Disk Space Requirements

An HACMP installation has the following space requirements:

- The **/usr** directory must have 82 MB of free space for a full installation.

  If you are *not* planning to install optional software, you can plan for less space (for example, HAView =14 MB, hativoli =140 KB). Also, install only the message catalogs for the language you will be using, rather than all message catalogs. (Japanese message catalogs use 1.6 MB.)
- The / (root) directory must have 710 KB of free space (beyond any need to extend the **/usr** directory).

Required Versions of AIX and RSCT

HACMP 5.4.1 requires the versions of AIX and RSCT listed in the following table. For more information about these requirements, see also:

- Requirements When Upgrading AIX
- Required AIX BOS Components
- Requirements for RSCT

AIX Version	RSCT Version	RSCT Filesets
AIX v.5.3 plus ML 1	2.4.2	• rsct.compat.basic.hacmp 2.4.2.0 • rsct.compat.clients.hacmp 2.4.2.0 • rsct.core.sec 2.4.2.1 • rsct.core.rmc 2.4.2.1
AIX v.5.2 plus ML 5	2.3.6	• rsct.compat.basic.hacmp 2.3.6.0 • rsct.compat.clients.hacmp 2.3.6.0 • rsct.core.sec 2.3.6.1 • rsct.core.rmc 2.3.6.1

Requirements When Upgrading AIX

If AIX is an earlier version, upgrade it before installing HACMP. If you install AIX on an SP system, upgrade PSSP to version 3, release 5.

1. If you are upgrading from HACMP 4.5 running on AIX v.5.2, install APAR IY50233 first.
2. If you upgrade the AIX operating system, ensure that adapter SCSI IDs or SSA node numbers 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 7. This setting can cause a SCSI ID conflict on the bus that prevents proper access to the shared disk.
3. Disable import of volume groups. If you upgrade AIX to version 5.2 or greater, all volume groups are imported and varied on, on the node being upgraded. This causes loss of disk access by other nodes in the concurrent resource group in the cluster.

   To prevent this disruption when migrating to AIX v.5.2 or greater, use the option to disable import of volume groups.

4. Disable "restore a **mksysb**". If you upgrade AIX v.5.2 or greater, and are using or planning to use enhanced concurrent volume groups, the "restore a **mksysb**" function will open the disk(s) with reserve. Therefore, to avoid this, disable the "restore a **mksysb**" option.
5. Make sure that you upgrade the AIX operating system before you install HACMP.

   When you migrate the operating system to AIX v.5.2 ensure that you have HACMP 5.1 plus the APAR IY55542 for SNMP v 3 compatibility running on AIX v.5.1.

   Note: The version of **/etc/snmpd.conf** depends on which version of AIX you are using. For AIX v.5.2 and up, the default version used in HACMP is **snmpdv3.conf**.

   It is recommended to migrate from and to the same version of SNMP to ensure that SNMP-based applications function correctly. Once the migration has completed, you can switch to a different version of SNMP if you prefer.

   For example, if you are migrating from an environment using SNMP v.1, and you are upgrading to AIX v.5.3, then before migrating HACMP, execute the following series of commands:

   ```
   stopsrc -s snmpd
   /usr/sbin/snmpv3_ssw -1
   startsrc -s snmpd
   ```

   See the AIX documentation for full information on the **snmpd.conf** file. Version 3 has some differences from Version 1.
6. There are no known problems when upgrading to AIX v.5.3.

Required AIX BOS Components

The following AIX base operating system (BOS) components are required for HACMP:

AIX BOS Component	AIX v.5.3	AIX v.5.2
bos.adt.lib	5.3.0.10	5.2.0.10
bos.adt.libm	5.3.0.10	5.2.0.13
bos.adt.syscalls	5.3.0.10	5.2.0.11
bos.net.tcp.client	5.3.0.10	5.2.0.14
bos.net.tcp.server	5.3.0.10	5.2.0.14
bos.rte.SRC	5.3.0.10	5.2.0.10
bos.rte.libc	5.3.0.10	5.2.0.14
bos.rte.libcfg	5.3.0.10	
bos.rte.libcur	5.3.0.10	5.2.0.10

AIX BOS Component	AIX v.5.3	AIX v.5.2
bos.rte.libpthreads	5.3.0.10	5.2.0.12
bos.rte.odm	5.3.0.10	5.2.0.11
bos.rte.lvm.rte (required only using Concurrent Logical Volume Manager for concurrent access)		
bos.clvm.enh (required only using Concurrent Logical Volume Manager for concurrent access)		

Requirements for NFSv4

The **cluster.es.nfs** fileset that comes with the HACMP 5.4.1 installation medium installs the NFSv4 support for HACMP, along with a new NFS Configuration Assistant.

To install this fileset, the following BOS NFS components must also be installed on the system:

bos.net.nfs.server 5.3.7.0

bos.net.nfs.client 5.3.7.0

Requirements for RSCT

Install the RSCT images before installing HACMP. Ensure that each node has the same version of RSCT.

To determine if the appropriate filesets are installed and their level, issue the following commands:

```
/usr/bin/lslpp -l rsct.compat.basic.hacmp
/usr/bin/lslpp -l rsct.compat.clients.hacmp
/usr/bin/lslpp -l rsct.basic.rte
```

If these filesets are *not* present, install the appropriate version of RSCT as described in Required Versions of AIX and RSCT, and apply any updates.

Security Fileset Requirements

If you plan to use message authentication or encryption for HACMP communication between cluster nodes, the following filesets must be installed on each node:

- **rsct.crypt.des—**for data encryption with DES message authentication
- **rsct.crypt.3des—**for data encryption standard Triple DES message authentication:
- **rsct.crypt.aes256—**for data encryption with Advanced Encryption Standard (AES) message authentication.

You can install these filesets from the AIX Expansion Pack CD-ROM.

HAView Requirements

HAView requires Tivoli NetView. Install NetView before installing HAView.

Contact your IBM sales representative for information on obtaining Tivoli NetView software.

The HAView fileset includes a server image and a client image. If NetView is installed using a client/server configuration, the HAView server and client images should be installed on the NetView server, and the client image on the NetView client.

Note: Installing HAView outside the cluster lessens the likelihood of losing monitoring capabilities during a cluster node failure.

For more information on using HAView to monitor a cluster, see Chapter 10: Monitoring an HACMP Cluster in the *Administration Guide*.

For information about installing HACMP on a client system, see Chapter 5: Installing HACMP on Client Nodes in this guide.

Tivoli Requirements

If you plan to monitor your cluster through Tivoli Distributed Monitoring, install the following Tivoli software:

- Tivoli Managed Framework (on TMR and cluster nodes)
- Tivoli Application Extension Facility (AEF) (on TMR only)
- Tivoli Distributed Monitoring (on TMR and cluster nodes)
- Unix Monitors
- Universal Monitors.

Contact your IBM sales representative for information on obtaining Tivoli Framework and Distributed Monitoring software.

Note that the Tivoli Management Region (TMR) node should be outside the HACMP cluster.

In addition, install the **cluster.hativoli** filesets.

Review the installation steps and prerequisites to effectively monitor your cluster through Tivoli. For more information, see Appendix A: Cluster Monitoring with Tivoli.

English and Japanese Message Catalogs

US English and Japanese message catalogs are available. Set the LANG variable to an appropriate locale value so that SMIT help is available in the desired language.

Note that in SMIT, pressing F1 displays help information in the correct language only *if the LANG variable is set correctly*. HACMP supports the following locales:

- en_US
- En_US
- ja_JP
- Ja_JP.

Also ensure that the correct base system locale is installed. To list the installed locales, type:

```
locale -a
```

To list the active locale, type:

```
locale
```

Since the active locale is determined by the LANG environment variable setting, the locale will be en_US if LANG=en_US.

Ensure that the proper HACMP message catalogs for the chosen language have been installed. To list the message catalogs, type:

```
lslpp -l "cluster.msg*"
```

Contents of the Installation Medium

The HACMP software installation medium contains the HACMP Enhanced Scalability Subsystem images, some of which you install on all cluster nodes and clients. This facility provides the 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.

HACMP includes the Concurrent Logical Volume Manager as an installation option. This adds concurrent shared-access management for supported disk subsystems. Concurrent access is provided at the raw logical volume level. Applications that use the Concurrent Logical Volume Manager must be able to control access to the shared data.

The HACMP Application Plug-in software (on the HACMP installation medium) contains the network service plug-in images. This plug-in fileset provides example scripts to start and stop the network service scripts for Domain Name Server (DNS), Dynamic Host Configuration Protocol (DHCP), and printer services. Each network service has start and stop scripts bundled in a fileset. These scripts are provided as examples that may be customized for your environment.

Several prerequisites must be completed before setup begins. A setup wizard is included in each fileset to assist with the setup after installation.

The plug-in filesets are as follows:

- **cluster.es.plugins.dns**. This plug-in fileset provides scripts to start and stop the DNS server process, a script to confirm that configuration files are present and stored in a shared file system, and scripts called by the monitoring functions of HACMP that check on DNS server process life.
- **cluster.es.plugins.dhcp**. The DHCP Network Service Plug-in provides scripts to start and stop the DHCP server process, a script to ensure that configuration files are present and stored in a shared file system, and scripts called by the monitoring functions of HACMP that check on DHCP server process life.
- **cluster.es.plugins.printserver.** This plug-in fileset provides scripts to start and stop the print server process, a script to ensure that configuration files are present and stored in a shared file system, and scripts called by the monitoring functions of HACMP that check on print server process life.

HACMP Installable Images

The organization of cluster images on the HACMP medium allows you to make individual or multiple image selections through SMIT when installing the HACMP software.

Note: 1. The RSCT images are prerequisites for HACMP and are packaged with all versions of AIX.

2. The **cluster.es** and **cluster.es.cspoc** images contain HACMP runtime executable files and are required.

3. The plug-in software is optional demo software.

The following list shows the installable HACMP images with the filesets for each image listed under the image:

cluster.es

`cluster.es.server.rte`	Base Server Runtime
`cluster.es.client.lib`	Client Libraries
`cluster.es.client.rte`	Client Runtime
`cluster.es.client.utils`	Client Utilities
`cluster.es.server.testtool`	Cluster Test Tool
`cluster.es.server.diag`	Server Diags
`cluster.es.server.events`	Server Events
`cluster.es.server.utils`	Server Utilities
`cluster.es.server.cfgast`	Two-Node Configuration Assistant
`cluster.es.client.wsm`	Web-based SMIT

cluster.es.cspoc

`cluster.es.cspoc.cmds`	CSPOC Commands
`cluster.es.cspoc.rte`	CSPOC Runtime Commands
`cluster.es.cspoc.dsh`	CSPOC dsh

cluster.es.cfs

`cluster.es.cfs.rte`	Cluster File System Support

cluster.es.worksheets

`cluster.es.worksheets`	Online Planning Worksheets

cluster.adt.es

`cluster.adt.es.client.samples.clinfo`	Client CLINFO Samples
`cluster.adt.es.client.samples.clstat`	Client Clstat Samples
`cluster.adt.es.client.include`	Client Include Files
`cluster.adt.es.client.samples.libcl`	Client LIBCL Samples
`cluster.adt.es.java.demo.monitor`	Web-based Monitor Demo

cluster.man.en_US.es

`cluster.man.en_US.es.data`	Man Pages—U. S. English

cluster.msg.en_US.es

`cluster.msg.en_US.es.server`	Recovery Driver Messages—U. S. English
`cluster.msg.en_US.es.client`	Client Messages—U. S. English
`cluster.msg.en_US.cspoc`	CSPOC Messages—U. S. English

cluster.haview

`cluster.haview.client`	HAView Client
`cluster.haview.server`	HAView Server

cluster.man.en_US.haview.data

`cluster.man.en_US.haview.es.data`	HAView Man Pages

cluster.msg.en_US.haview

`cluster.msg.en_US.es.haview`	HAView Messages

User documentation image and filesets:

cluster.doc.en_US.es

`cluster.doc.en_US.es.html`	HTML Documentation—U. S. English
`cluster.doc.en_US.es.pdf`	PDF Documentation—U. S. English

HACMP Tivoli image and filesets:

cluster.hativoli

`cluster.hativoli.client`	HACMP Tivoli Client
`cluster.hativoli.server`	HACMP Tivoli Server

cluster.msg.en_US.hativoli

`cluster.msg.en_US.hativoli`	HATivoli Messages—U. S. English

Application plug-in software image and filesets:

cluster.es.plugins

`cluster.es.plugins.dns`	DNS Network Service Plug-in
`cluster.es.plugins.dhcp`	DHCP Network Service Plug-in
`cluster.es.plugins.printserver`	Printer Server Network Service Plug-in

Note: There are filesets corresponding to En_US, ja_JP, and Ja_JP for `cluster.doc...`, `cluster.msg...`, and `cluster.man...`.

HACMP Installation Choices

Install the HACMP software on each cluster node (server) and on any client machines that run the **clinfo** daemon. You can install the software in the following ways:

- From a Network Installation Management (NIM) server
- From a hard disk to which the software has been copied
- Directly from the installation medium.

Note that you must accept the license agreement as you install. Each node requires an HACMP software license.

A user with root privileges must perform the installation.

If you are installing HACMP on a client system, see Chapter 5: Installing HACMP on Client Nodes.

Installing from an Installation Server

To install the HACMP software in a cluster environment, you can create an HACMP installation server (containing all HACMP 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 software onto other nodes faster from the server than from other media. HACMP

supports the Network Installation Management (NIM) program including the Alternate Disk Migration option. For instructions on creating an installation server, see the *AIX Installation Guide* or the *AIX Network Installation Management Guide and Reference*.

After installing the HACMP software, verify the software installation. For information about verifying the software installation, see the section Verifying the Software Installation.

Installing from a Hard Disk

To install the HACMP software from your hard disk, first copy the software from the installation medium to the hard disk.

After installing the HACMP software, verify the software installation. For information about verifying the software installation, see the section Verifying the Software Installation.

Copying HACMP Software to Hard Disk

To copy the HACMP software to your hard disk:

1. Place the HACMP CD into the CD-ROM drive.
2. Enter `smit bffcreate`

   The Copy Software to Hard Disk for Future Installation panel appears.
3. Enter the name of the CD-ROM drive in the **INPUT device / directory for software** field and press Enter.

   Select the proper drive from the list and press Enter. That value is entered into the **INPUT device/directory** field as the valid input device.
4. Press Enter to display the **Copy Software to Hard Disk for Future Installation** panel.
5. Enter field values as follows:

SOFTWARE name	Make a selection from the picklist, or select **all** to copy all server and client images.
DIRECTORY for storing software	Change the value to the storage directory accessed by all nodes using HACMP.

6. Enter values for the other fields as appropriate for your site.
7. When you are satisfied with the entries, press Enter.

   SMIT prompts you to confirm your selections.
8. Press Enter again to copy the software.

Installing HACMP from the Hard Disk

After the HACMP software has been copied to your system (that is *not* an SP system), install the software by following the instructions in the section Directly Installing HACMP.

If you are installing the software on the SP, see the section Installing HACMP on an SP from the Hard Disk.

Installing HACMP on an SP from the Hard Disk

Install the HACMP software on all nodes of the SP system that will be used to form an HACMP cluster. If a node or control workstations is used to monitor the cluster, HACMP client code must be installed on those systems.

Use the SP **dsh** command to speed installation of the LPP filesets (images).

To install HACMP on an SP system:

1. Create a file called **/HACMPHOSTS** that contains host names of nodes in the SP frame that will have the HACMP 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 the hosts respond) using the following command:

   ```
   /usr/lpp/ssp/bin/SDRGetObjects host_responds
   ```

   where **SDRGetObjects** refers to the SP database. Host responds should indicate a 1 for all nodes that respond.
4. Mount the file system (from the control workstation) onto all nodes, enter the command:

   ```
   dsh /etc/mount CWNAME:STORAGE_DIRECTORY /mnt
   ```

   where *CWNAME* is the hostname of the control workstation and *STORAGE_DIRECTORY* is the directory where the software is stored.
5. Install the HACMP software on the nodes:
   - *To install filesets one at a time,* issue the following command:

     ```
     dsh "/etc/installp -Xagd /mnt LPP_NAME"
     ```

     where *LPP_NAME* is the name of the package/fileset to install. Do this for each fileset you need to install.

     or
   - *To install all packages and filesets* included in the HACMP software, issue one of the following commands:

     ```
     dsh "/etc/installp -Xagd /mnt cluster*"
     ```

     or

     ```
     dsh  /usr/lib/instl/sm_inst installp_cmd -a -Q -d '/mnt'
     -f '_all_latest' -cNgXG
     ```

     WARNING: Make sure that the **_all_latest** directory contains only HACMP filesets. If the **_all_latest** directory contains filesets other than those for the HACMP software, the preceding command installs those files as well.
6. To ensure that HACMP was successfully installed on each node, enter:

   ```
   dsh "usr/sbin/lslpp -l cluster*"
   ```

For SP nodes installed through the SP install process, you can use that process or the customize process to install HACMP and PTFs. See the *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*.

HACMP files are installed in the **/usr/es/sbin/cluster** directory.

Installing from the Installation Medium

If you install the HACMP software from the CD-ROM, you install the software directly onto each cluster node.

To install the Concurrent Logical Volume Manager, include the **cluster.es.clvm** image when you select which software to install.

Installing RSCT Files for AIX

To install the required RSCT files:

1. Insert the CD into the CD-ROM drive and enter:

   ```
   smit install_all
   ```

   SMIT displays the first **Install and Update from ALL Available Software** panel.
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 the install directory, select one from the picklist.
3. Either enter **all** for SOFTWARE to install or select items from a picklist.

   For information about other fields, see the section Directly Installing HACMP.
4. Press Enter to begin the installation.
5. After the installation of finishes, install the HACMP software as described in the following section Directly Installing HACMP.

Directly Installing HACMP

To install the HACMP software on a server node from the installation medium:

1. Insert the CD into the CD-ROM drive and enter:

   ```
   smit install_all
   ```

   SMIT displays the first **Install and Update from ALL Available Software** panel.
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 the install directory, select one from the picklist.
3. Enter field values as follows. Press F1 for help on any field.

Note: Use F4 to list the software before proceeding with the installation. This way you can install either the English or the Japanese message catalogs, and you can omit optional software if you wish.

INPUT device / directory for software

This field shows the device or directory you specified earlier.

SOFTWARE to install

Select an option from the picklist, or enter **all** to install all server and client images. For a list of filesets, also see the section HACMP Installable Images.

Install the following required images (which contain the HACMP runtime executables) on all servers:

- **cluster.es**
- **cluster.cspoc**

Make sure to install the level of RSCT required for your release of AIX. See Required Versions of AIX and RSCT.

If you are installing the Two-Node Cluster Configuration Assistant, install **cluster.es.server.cfgast**.

If you are installing the Cluster Test Tool, install **cluster.es.server.testtool**. (If **cluster.es i**s selected, the Cluster Test Tool is installed automatically.)

Your selections appear in this field. Note that selecting **cluster.es** and **cluster.cspoc** installs HACMP and all associated messages.

PREVIEW only?

If set to **yes,** the preview option will check and ensure that installation prerequisites are met (for example, that required software is installed and sufficient disk space is available). Press F1 for details.

When you are ready to perform the actual installation, set this field to **no**.

COMMIT software updates?

This field applies only when installing software updates (PTFs). See F1 help for details.

SAVE replaced files?

Use the default.

This field applies only when installing software updates (PTFs). If you select **no** to **commit software updates?** you must select **yes** for this field. See F1 help for details.

AUTOMATICALLY install requisite software

Use the default.

Set this field to **no** if the prerequisite software for HACMP is already installed or if the **OVERWRITE same or newer versions?** field is set to **yes**; otherwise, set this field to **yes** to install required software. See F1 help for details.

EXTEND filesystems if space needed?

Select **yes** if you have adequate hard disk space, **no** if you have limited space. See F1 help for details.

OVERWRITE same or newer versions?	Use the default. For normal new installations, leave this field set to **no**. Set it to **yes** if you are reinstalling. If you set this field to **yes**, you must set the **Automatically install requisite software** field to **no**. See F1 help for details.
VERIFY install and check file sizes?	Use the default. Select **yes** if you want the system to perform some checks on the software you installed. See F1 help for details.
DETAILED output?	Select **yes** if you want a detailed log of all installation messages.
Process multiple volumes?	Select this option if you want to enable the processing of multiple-volume CDs. See F1 for information.
ACCEPT new license agreements?	Select **yes** for this item to proceed with installation. If you select **no**, installation may stop with a warning that one or more filesets require software license agreements. You accept the license agreement only once for each node.
Preview new license agreements?	Select **yes** to view the text of the license agreements. The text appears in the current window in the language defined on your system.

4. When you are satisfied with the entries, press Enter.

   SMIT prompts you to confirm your selections.

5. Press Enter again to install.
6. After the installation completes, verify the installation as described in the section Completing the Installation in this chapter.

Be sure to read the HACMP **release_notes** file in the **/usr/es/sbin/cluster/** directory for further instructions and the latest information on requirements or known issues.

Completing the Installation

After you install the HACMP software, complete the tasks described in this section.

Verifying the Software Installation

To complete the installation after the HACMP software is installed:

1. Verify the software installation by using the AIX command **lppchk**, and check the installed directories for the expected files.

   The **lppchk** command verifies that files for an installable software product (fileset) match the Software Vital Product Data (SWVPD) database information for file sizes, checksum values, or symbolic links.

2. Run the commands **lppchk -v** and **lppchk -c "cluster.*"**

If the installation is OK, both commands should return nothing.

3. Reboot each HACMP cluster node and client.

Addressing Problems during the Installation

Review the following sections if you encounter problems during an HACMP installation.

Cleaning up after an Unsuccessful Installation

If you experience problems during the installation, the installation program automatically performs a cleanup process. If the cleanup does *not* perform automatically, you can perform a cleanup manually.

To perform a manual cleanup after an unsuccessful installation:

1. To display the **Installation and Maintenance** menu, enter `smit install`
2. Select **Install and Update Software**.
3. Select **Clean Up After a Interrupted Installation**.
4. Review the SMIT output (or examine the **/smit.log** file or **/smit.log.<*locale*>** file) for the interruption's cause.
5. Fix any problems and repeat the installation process.

Resolving System ID Licensing Issues

The Concurrent Logical Volume Manager is licensed to the system ID (hardware system identifier) of a cluster node. Many of the **clvm** or concurrent access commands validate the system ID against the license file. A mismatch causes the command to fail with an error message indicating the lack of a license.

Note: The licensing of the Concurrent Resource Manager to the system ID is *not related* to the licensing of the HACMP software.

Removing the HACMP Software

Before you remove the HACMP software from a system, stop cluster services. You cannot remove the software while the cluster is running.

To stop cluster services, enter the fastpath `smitty clstop` and press Enter.

For more information about stopping cluster services see Chapter 9: Starting and Stopping Cluster Services in the *Administration Guide*.

To remove the HACMP software and your cluster configuration on cluster nodes and clients:

1. Enter `smit install_remove`

   The **Install/Remove** panel appears.

2. Enter the following values for these fields:

SOFTWARE name	Enter **cluster*** to remove all server and client software, or you can select options from the picklist. Your selections appear in this field.
REMOVE dependent software?	Select **no**.
EXTEND filesystems if space needed?	Select **yes**.
DETAILED output?	Select **no**.

3. When you are satisfied with the entries, press Enter.

   SMIT prompts you to confirm your selections.

4. Press Enter again to remove the software.

Installing and Configuring WebSMIT

WebSMIT is a Web-based user interface that provides consolidated access to the SMIT functions of configuration, management, status, and the HACMP documentation. Starting with HACMP 5.4, you can use WebSMIT to navigate the running cluster and view graphical displays of cluster topology and resource group dependencies. With cluster services running on one node in the cluster, you can view cluster and node status.

Planning for WebSMIT

You can run WebSMIT on a single node; however, if that node goes down, WebSMIT will be unavailable. To provide better availability, you can set up WebSMIT to run on multiple nodes. Since WebSMIT is retrieving and updating information from the HACMP cluster, that information should be available from all nodes in the cluster.

Typically, you set up WebSMIT to be accessible from the cluster's internal network that is *not* reachable from the Internet. Network access is required between the browser and cluster node that serves as a Web server. To run WebSMIT on a node at a remote site, it is highly recommended that you ensure HTTP(S) connectivity to that node; it is *not* handled automatically by WebSMIT or HACMP.

Most HACMP cluster utilities require root authority. Allowing execution of such programs through a Web browser requires careful analysis and planning to reduce security risks.

HACMP does *not* supply the Web server software, but it does supply a default configuration. See the **/usr/es/sbin/cluster/wsm/README** file for information on WebSMIT, the default security mechanisms in place when HACMP 5.4.1 is installed and how to create a more secure Web server installation.

We recommend that you plan to create a new user for the sole purpose of using WebSMIT. This allows you to distinguish between root logins and WebSMIT session logins, and permits a user other than root to configure HACMP using WebSMIT without having knowledge of the root password.

You have the option of specifying a group of users that have read-only access to WebSMIT. Users with read-only access may view the configuration and cluster status, and may navigate through the WebSMIT screens but cannot execute commands or make any changes to the configuration.

Prerequisites Tasks for Installing WebSMIT

Perform the following steps before installing WebSMIT:

1. You must first install a Web server on one or more cluster nodes. You can use any Apache-compliant server, such as the IBM HTTP Server, or the SSL-enabled Apache server. See the **/usr/es/sbin/cluster/wsm/README** file for information.
2. Install HACMP 5.4.1 on all nodes in the cluster. For installation information, see Chapter 9: Installing HACMP on Server Nodes.

Steps for Installing and Configuring WebSMIT

Here are the steps to install and configure WebSMIT:

What You Do	Description
Step 1: Set up a Web server	After the prerequisite tasks are completed, proceed directly to the overview of WebSMIT, discussion of security, and how WebSMIT can be integrated with your existing configuration. If you are *not* currently running a Web server, first plan the basic implementation of running a Web server within your enterprise using the information supplied in this section.
Step 2: Install the WebSMIT fileset	Install the WebSMIT fileset on a node running HACMP 5.4.1. To increase availability, you can setup WebSMIT to run on multiple nodes. WebSMIT is installed in the directory **/usr/es/sbin/cluster/wsm**.

What You Do	Description
Step 3: Set up users	Parameters in the configuration file **/usr/es/sbin/cluster/wsm/wsm_smit.conf** default to enhanced security. Documentation on how to modify parameters is within the configuration file itself. Optionally, you may edit the sample post-install script variables and those in the **/usr/es/sbin/cluster/wsm/wsm_smit.conf** file to: • Create new users for the sole purpose of using WebSMIT. This allows you to distinguish between root logins and WebSMIT session logins, and permits a user other than root to configure HACMP using WebSMIT without having knowledge of the root password. You can also view the **/etc/security/lastlog**, to distinguish between root logins and WebSMIT session logins. The tty for a WebSMIT user is /WebSMIT. • Specify a set of users who have read-only access. Users with read-only access may view the configuration and cluster status, and may navigate through the WebSMIT screens but cannot execute commands or make any changes to the configuration. To specify users that have read-only access, edit the **/usr/es/sbin/cluster/wsm/wsm_smit.conf** file as follows: 1. Enable authentication by setting the REQUIRE_AUTHENTICATION variable to 1. 2. Add the usernames of all users to the ACCEPTED_USERS list. 3. Add the usernames, of the users that will have read-only access, to the READ_ONLY_USERS list. The READ_ONLY_USERS list does *not* provide additional usernames for authentication purposes, but will merely specify which of those users listed in ACCEPTED_USERS have limited access.
Step 4: Change the status refresh interval (optional)	The polling interval (STATUS_UPDATE_INTERVAL) is the period between the last status update completion and the next status update request, as opposed to the time between requests (value must be in seconds). HACMP updates the cluster status displayed in WebSMIT at a regular polling interval set in the **wsm_smit.conf** file.

What You Do	Description
Step 5: Set up the WebSMIT interface	Installing HACMP does *not* enable the WebSMIT graphical interface. You can set up WebSMIT from the command line using the **websmit_config** command. The **websmit_config** command can be used to quickly and easily configure the locally installed HTTP server to host WebSMIT. Of course, this configuration can also be performed manually, if desired, as might be the case when the HTTP server is already being used for another task. The **websmit_config** command automatically detects which HTTP server is in use, and creates a private configuration for the discovered server. If the IBM HTTP Server is in use, you have the option of creating an SSL key file by supplying a pass phrase using the **-p** flag. This is because it does not ship with a self-signed certificate, as Apache does.
Step 6: Start the Web server	Use the **websmitctl** command to start (or stop) the Web server HTTP daemon or get the status of an HTTP process group.
Step 7: Set up the browser	Once the Web server process is running and providing access to the WebSMIT facilities, you can monitor and manage your cluster using a Web browser. The platform running the browser is independent of the cluster nodes. WebSMIT supports most popular browsers (including Mozilla, FireFox, and Internet Explorer). See the **/usr/es/sbin/cluster/wsm/README** file for information about tested/supported versions. To set up the Web browser, Javascript must be enabled.

WebSMIT should now be available. Launch any supported browser (does not have to be on a cluster node) and load the following URL (where *hostname* is the hostname of the cluster node where you installed and configured WebSMIT):

```
* https://<hostname>:42267
```

Note that WebSMIT uses the special port 42267 to keep WebSMIT from interfering with any existing Web server configuration on the system. See the **/usr/es/sbin/cluster/wsm/README** file for more information, including some troubleshooting tips.

Upgrading from a Previous Configuration

Note: If you are configuring WebSMIT for the first time, this section does not apply. See Steps for Installing and Configuring WebSMIT.

If you are upgrading from a previous configuration that was integrated with the system-wide server configuration, we strongly recommend removing that configuration and creating a new configuration using the provided template files and sample configuration script. This will make the WebSMIT configuration self-contained and eliminate the need to manually reconfigure WebSMIT in the future.

A self-contained WebSMIT configuration means that the Web server configuration only serves WebSMIT, it does not need to be integrated into any existing configuration, and it will not interfere with non-WebSMIT Web server configurations.

If you have upgraded to HACMP 5.4.1, and before the upgrade, had a working configuration for WebSMIT in HACMP 5.3 that was integrated with a system-wide Web server configuration, do the following:

1. Remove WebSMIT from your existing Web server configuration and reconfigure it using one of the provided templates and the sample configuration script. This is described in the "Configuration" section of the **README** file. This option is strongly recommended. In most cases, this will eliminate the need to manually reconfigure the Web server for WebSMIT, greatly reducing the risk of human error.
2. Integrate the configuration in the appropriate template file:

   ```
   httpd.wsm.conf.ihs.template
   httpd.wsm.conf.apache.template
   httpd.wsm.conf.apache2.template
   httpd.wsm.conf.httpd.template
   ```

   Ensure the following files still have permissions 4511:

   ```
   /usr/es/sbin/cluster/wsm/cgi-bin/wsm_cmd_exec
   /usr/es/sbin/cluster/wsm/utils/sperl
   ```

   Those file permission changes are just one step of several taken by the sample setup script (**websmit_config**). Refer to that script for other actions that may need to be taken.

Where You Go from Here

After you install the HACMP software, complete the following steps to configure an HACMP server:

1. Set up the Cluster Information Program.

   Copy the **clhosts.client** file to each client node as **/usr/es/sbin/cluster/etc/clhosts** and edit the **/usr/es/sbin/cluster/etc/clinfo.rc** script as described in Chapter 5: Installing HACMP on Client Nodes.
2. Ensure that the network interfaces and shared external disk devices are ready to support an HACMP cluster.

   See Chapter 6: Configuring Installed Hardware.
3. Define shared LVM components, including creating shared volume groups, logical volumes, and file systems for your cluster. Also define enhanced concurrent volume groups.

   See Chapter 7: Defining Shared LVM Components.
4. Customize your AIX environment for HACMP.

See Chapter 8: Configuring AIX for HACMP.

5. If you want to set up a basic two-node configuration, use the Two-Node Cluster Configuration Assistant. For information about using the Two-Node Cluster Configuration Assistant, see Chapter 9: Creating a Basic HACMP Cluster.
6. (*Optional*) Configure the HACMP cluster using the Online Planning Worksheets application.

   See the chapter on Using Online Planning Worksheets in the *Planning Guide*.

   If you decided *not* to use the application, use the paper planning worksheets you filled out during the planning process, and follow the steps in the *Administration Guide* to configure your cluster using SMIT or WebSMIT.
7. Test your configuration using the Cluster Test Tool.

   See the chapter on Testing an HACMP Cluster in the *Administration Guide*.

Chapter 5: Installing HACMP on Client Nodes

This chapter describes how to install and configure the HACMP software on client systems. It also describes how to edit files and scripts related to the Cluster Information Program (Clinfo). This chapter contains the following sections:

- Prerequisites
- Overview
- Installing and Configuring HACMP on Client Systems

Prerequisites

Before you install the HACMP software on client systems:

- Make sure that the system has at least 2.6 MB available for the installation.
- Read the chapter on Planning for HACMP Clients in the *Planning Guide*.
- Read the Release Notes for HACMP in **/usr/es/sbin/cluster/release_notes** for additional information on installing the HACMP software.
- If you are installing HACMP on Linux, read the Linux Release Notes for HACMP in **/usr/es/sbin/cluster/release_notes.linux/** for issues relevant to HACMP on the Linux platform.
- If you are using HAView, see section HAView Requirements in Chapter 4: Installing HACMP on Server Nodes.
- Install HACMP on the server nodes. For more information, see Chapter 4: Installing HACMP on Server Nodes.)

Overview

Installing the HACMP software on each client that runs 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.

Installing and Configuring HACMP on Client Systems

To install and configure the HACMP software on each client, complete these steps:

Step 1: Installing the Base System Client Images

Step 2: Copying the clhosts.client File onto Client Nodes

Step 3: Editing the clinfo.rc Script

Step 4: Updating the ARP Cache for Clients Not Using Clinfo

Step 5: Rebooting the Clients.

Step 1: Installing the Base System Client Images

For a new installation, the **/usr** directory requires a minimum of 2.6 MB of available space.

To install the base high availability software on a client:

1. Place the HACMP CD into the CD-ROM drive and enter:

   ```
   smit install_selectable_all
   ```

   SMIT displays the **Install Selectable All** panel.

   - If you are *not* sure of the name of the input device, select one from the picklist.
   - 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 panel.
3. In the **SOFTWARE to install** field, use the picklist to select the client software modules associated with the following cluster images: **cluster.es**, **cluster.msg**, and **cluster.adt.es**.

   Note: Note that if you select at least one client module associated with an installable image, 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 prompts you to confirm your selections.

6. Press Enter again.
7. Read the HACMP **release_notes** file in the **/usr/es/sbin/cluster/** directory for further instructions.
8. Copy the **/usr/es/sbin/cluster/etc/clhosts.client** file from the server to each client node, renaming it as **/usr/es/sbin/cluster/etc/clhosts**.
9. (*Optional*) Edit the **clinfo.rc** script as described in the section Step 3: Editing the clinfo.rc Script.
10. Reboot the client.

Repeat this procedure for each client system.

Step 2: Copying the clhosts.client File onto Client Nodes

During the verification phase of running the Two-Node Cluster Configuration Assistant, HACMP creates and populates a **clhosts.client** file on all HACMP servers.

Copy this **/usr/es/sbin/cluster/etc/clhosts.client** file from a HACMP server node to all client nodes, renaming it **clhosts** (remove the **.client** extension).

Step 3: Editing the clinfo.rc Script

The **/usr/es/sbin/cluster/etc/clinfo.rc** script updates the ARP cache on the local node whenever an event occurs. If you are *not* using IPAT via IP Replacement with Alternate Hardware Addresses, a copy of the **/usr/es/sbin/cluster/etc/clinfo.rc** script must exist on each server node and client in the cluster in order for all ARP caches to be updated.

The HACMP software is distributed with a template version of the **/usr/es/sbin/cluster/etc/clinfo.rc** script. You can use the script as distributed, modify it, or replace it with a custom script.

Note: If you are *not* using IPAT via IP Replacement with Alternate Hardware Addresses, the ARP function must remain in the **clinfo.rc** script.

The format of the **clinfo** call to **clinfo.rc**:

```
clinfo.rc {join,fail,swap} interface_name
```

The **clinfo** utility obtains information about the interfaces and their current state, and 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***.

For complete information about cluster events and customizing scripts, see the chapter on Planning for Cluster Events in the *Planning Guide*.

For a sample client application that uses the Clinfo API within the context of a customized **clinfo.rc** script, see *Programming Client Applications*.

If you have custom applications that use the Clinfo API and plan to use symmetric multi-processors, you may need to modify your application. Afterwards, recompile and link your application. For updated information about the library routines, see the *Programming Client Applications.*

Step 4: Updating the ARP Cache for Clients Not Using Clinfo

On clients that are *not* running **clinfo**, you will need HACMP to update the client's ARP cache by pinging the client from a cluster node when an IP address has moved. Add the client IP label you want to notify to the PING_CLIENT_LIST variable in the **clinfo.rc** script on each node. Now, whenever a cluster event occurs, **clinfo.rc** runs the following command for each host specified in PING_CLIENT_LIST:

```
ping -c1 $host
```

Step 5: Rebooting the Clients

The final step in installing the HACMP software on a client is to reboot the client.

Chapter 6: Configuring Installed Hardware

This chapter describes how to ensure that network interface cards (NICs), shared external disk devices, and shared tape drives are ready to support an HACMP cluster. This chapter contains the following sections:

- Configuring Network Interface Cards
- Configuring Point-to-Point Networks
- Configuring RS232 Serial Connections
- Configuring an SP Switch
- Configuring for Asynchronous Transfer Mode (ATM)
- Configuring ATM Classic IP
- Configuring ATM ARP Servers for Use by HACMP Nodes
- Configuring ATM LAN Emulation
- Configuring Shared External Disk Devices
- Configuring Target Mode SCSI Connections
- Configuring Target Mode SSA Connections
- Installing and Configuring Shared Tape Drives.

Before you read this chapter, you should have already installed the devices following the instructions in the relevant AIX documentation. For the latest information about the hardware supported for use with the version of AIX that you are using, see the IBM website.

Note: For information on installing and configuring OEM disks, see Appendix B: OEM Disk, Volume Group, and File Systems Accommodation.

Configuring Network Interface Cards

This section describes how to ensure that network interface cards (NICs) are configured properly to support the HACMP software. For each node, ensure that the settings of each NIC match the values on the completed copies of the TCP/IP Network Interface Worksheet, described in the *Planning Guide*. Special considerations for a particular NIC type are discussed in the following sections.

Ethernet, Token-Ring, and FDDI Interface Cards

Consider the following guidelines when configuring NICs:

- When using the **smit mktcpip** fastpath to define a NIC, the **hostname** field changes the default hostname. For instance, if you configure the first NIC as *n0-svc,* and then configure the second NIC as *n0-nsvc,* the default hostname at system boot is *n0-nsvc.*

To avoid this problem, do *not* change the value in the **hostname** field, and enter the rest of the information specific to that NIC. The hostname should match the NIC label of the primary network's service NIC because some applications may depend on the hostname (although this is *not* required for HACMP).

- If you are using HACMP IP Address Takeover via IP Replacement, ensure that each NIC that will host a service IP label is set to use a non-service IP label at boot. Use the **smit chinet** or **smit chghfcs** fastpath to reconfigure these NICs, if necessary, to a non-service IP label. Refer to your completed copies of the TCP/IP Network Interface Worksheet. For more information, see the section Planning for IP Address Takeover via IP Aliases in the chapter on Planning Cluster Network Connectivity in the *Planning Guide*.

Completing the TCP/IP Network Interface Worksheets

After reviewing all network interfaces for a node, record the network interface names on that node's TCP/IP NETWORK Interface Worksheet. To display a list of available and defined NICs for the node, enter the following command:

```
lsdev -Cc if
```

At this point, all interfaces used by the HACMP software should be available. List the NICs marked as **Available** in the **Interface Name** field on the TCP/IP Network Interface Worksheet.

When you initially add a new NIC to the cluster, HACMP discovers the NIC interface name from the AIX configuration. However, if you later change any part of the existing NIC definition in HACMP, ensure that the NIC interface name known to HACMP is the same as the NIC definition in AIX. If they do *not* match, change the NIC interface name in HACMP to match the definition in AIX.

Configuring Point-to-Point Networks

A point-to-point network is an ideal way to connect the nodes in an HACMP cluster. The point-to-point network allows Cluster Managers to continue to exchange keepalive packets should the TCP/IP-based subsystem, networks, or network NICs fail. Thus, the point-to-point network prevents the nodes from becoming isolated and from attempting to take over shared resources. The HACMP software supports four types of point-to-point networks:

- RS232
- Disk heartbeating (over enhanced concurrent mode disks)
- Target mode SCSI
- Target mode SSA.

For more information on point-to-point networks, see the chapter Planning Cluster Network Connectivity in the *Planning Guide.*

For information about configuring point-to-point networks in HACMP, see Chapter 4: Configuring HACMP Cluster Topology and Resources (Extended) in the *Administration Guide*.

Configuring RS232 Serial Connections

This section describes how to configure an RS232 serial cable as a serial network in an HACMP cluster. Use a serial network when connecting two nodes in an HACMP environment. For a cluster of more than two nodes, configure the serial network in a ring configuration (NodeA<—>NodeB NodeB<—>NodeC NodeC<—>NodeA).

Before configuring the RS232 serial connection, physically install the cables between the nodes. To connect the nodes, use a fully pinned out, 25-pin null-modem (turnaround), serial-to-serial cable. You can order an HACMP serial cable from IBM. 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: Many systems have special serial port requirements. Refer to the documentation for your system. For more information on serial ports for HACMP, see the chapter on Planning Cluster Network Connectivity in the *Planning Guide*.

Configuring an RS232 Serial Connection in AIX

To configure an RS232 serial connection:

1. Ensure that you have physically installed the RS232 serial cable between the two nodes before configuring it.
2. Use the following command to review the status of each serial port you intend to use after installing the RS232 serial cable:

   ```
   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 cable. Removing and then defining the **tty** device makes it available with the default settings, which are appropriate for the communication test described here.
4. Set the **ENABLE login** field to **DISABLE** to prevent **getty** processes from spawning on this device. Refer to the following section, Defining the tty Device.
5. Test communication over the serial cable after creating the **tty** device. For more information about testing serial networks, see the section Testing the Serial Connection.

Defining the tty Device

To create a **tty** device on each node that is connected to the RS232 cable:

1. Enter `smit tty`

   The TTY panel appears.

2. Select **Add a TTY** and press Enter. The SMIT Add a TTY panel 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	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 **disable** to prevent **getty** processes from spawning on this device. (You can also set this option using the following command: `chdev -l tty0 -a login='disable')`
Flow Control to use	Leave the default of **xon**, as Topology Services will disable the **xon** setting when it begins using the device. If **xon** is *not* available, then use **none**. Topology Services cannot disable **rts**, and that setting has (in rare instances) caused problems with the use of the adapter by Topology Services.

6. Press Enter to commit the values.

   Repeat this procedure for the other node that is connected to the RS232 cable.

Note: Regardless of the baud rate setting of the tty when it is created, all RS232 networks used by HACMP are brought up by RSCT with a default baud rate of 38400. Some RS232 networks that are extended to longer distances will require that the baud rate be lowered from the default of 38400.

For more information, see the section Changing an RS232 Network Module Baud Rate in the chapter on Managing the Cluster Topology in the *Administration Guide*.

Testing the Serial Connection

To ensure that the RS232 cable is properly configured and transmits data, run the following test after creating the **tty** device on both nodes.

Run this test while the **tty** device is *not* in use. If the cluster is active, remove the serial network dynamically from the configuration before running the test. Also, verify that the **tty** device is *not* in use by any other process.

To determine if the device is in use, run the **fuser** command:

```
fuser /dev/tty0
```

The output lists the PID of any process which uses the device.

If the device is in use by RSCT, the output shows that a process hats_rs232_nim is accessing the device. After the network has been dynamically removed from the cluster configuration, no such process should exist.

In rare cases, the hats_rs232_nim process may *not* terminate during a dynamic removal of the network or a stop of the cluster services. In these cases, you should call IBM support. However, it is safe to terminate any leftover hats_nim_rs232 process if the cluster is inactive on the local node.

Use the **fuser** command to terminate a process which accesses the **tty** device:

```
fuser -k /dev/tty0
```

Running the sttyTest

The stty test determines whether the serial connection allows the transmission of communications.

Running the stty Test on TTYs with RTS Flow Control Set

To perform the stty test:

1. On the receiving side, run:

   ```
   (stty raw -echo; cat > outputfilename) < /dev/tty2
   ```

2. On the sending side, run:

   ```
   (stty raw -echo < /dev/tty1; cat filetobesent ; sleep 5) > /dev/tty1
   ```

Running the stty Test on TTY's with XON or No Flow Control Set:

To perform the stty test:

1. On the receiving side (node 2), run:

   ```
   (stty raw -echo ixon ixoff; cat > outputfilename) < /dev/tty2
   ```

2. On the sending side, run

```
(stty raw -echo ixon ixoff < /dev/tty1; cat filetobesent; sleep 5) > /dev/tty1
```

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

If the data is transmitted successfully from one node to another, then the text from the **/etc/hosts** file from the second node appears on the console of the first node. Note that you can use any text file for this test, and do *not* need to specifically use the **/etc/hosts** file.

Defining the Serial Connection to HACMP

After you install and test the serial connection, define the connection as a point-to-point network to HACMP. For information about how to configure a serial network, see the *Administration Guide*.

Configuring an SP Switch

The SP Switch used by an SP node serves as a network device for configuring multiple clusters, and it can also connect clients. This switch is *not* required for an HACMP installation. When installed, the SP Switch default settings are sufficient to allow it to operate effectively in an HACMP cluster environment.

The HPS switch (older version of the switch) differs from the SP switch (newer version of the switch): *The SP switch does not allow HACMP to control the Eprimary*. You must upgrade to the SP switch before installing HACMP. If you are currently running HACMP Eprimary management with an HPS switch, run the HACMP script to unmanage the Eprimary *before* upgrading the switch.

To ensure that Eprimary is set to be managed, enter the following:

```
odmget -q'name=EPRIMARY' HACMPsp2
```

If the switch is set to **manage**, before changing to the new switch, run the script:

```
/usr/es/sbin/cluster/events/utils/cl_HPS_Eprimary unmanage
```

Keep in mind the following points about the SP Switch in an HACMP configuration:

- ARP must be enabled for the SP Switch network so that IP Address Takeover can work.
- HACMP SP Switch Base and service IP labels are alias addresses on the SP Switch css0 IP interface. The css0 base IP address is unused and should *not* be configured for IP Address Takeover via IP Replacement. However, for IPAT via IP Aliases, the css0 base IP address should be configured as an HACMP base address. Non-service IP labels are *not* allowed for SP Switch IP address takeover. The alias service IP labels appear as **ifconfig alias** addresses on the css0 interface.
- Service IP labels must be defined on a different subnet than the HACMP base IP label.
- The netmask associated with the css0 base IP address is used as the netmask for all HACMP SP Switch network interfaces.

For more information on SP switch networks, see the section on Planning for the SP Switch Network in the chapter on Planning Cluster Network Connectivity in the *Planning Guide*.

If you migrated a cluster that contains an SP Switch, see Chapter 3: Upgrading an HACMP Cluster.

Configuring for Asynchronous Transfer Mode (ATM)

Asynchronous Transfer Mode (ATM) is a networking technology and protocol suite based on packet switching. It is a connection-oriented protocol that uses virtual paths and channels to transfer data between devices.

HACMP supports two ATM protocols, Classic IP and LAN Emulation (LANE), for the configuration of cluster networks. Cluster networks defined on Classic IP are of cluster network type **atm**. Cluster networks defined on LANE are of the corresponding LAN type, that is, **ether** for LANE Ethernet, and **token** for LANE Token Ring.

ATM switches typically have inherent capabilities for fault tolerance. See the documentation for those products to determine how those recovery actions may integrate with HACMP.

Support of Classic IP and LANE on the Same Interface Card

ATM allows the configuration of multiple network interfaces and protocols over the same ATM device (`atm#`). HACMP allows multiple ATM clients to be configured on the same ATM device. Clients can belong to the Classic IP or LANE protocol types.

Note that interfaces that are configured over the same ATM device do *not* increase redundancy. To protect against single points of failure, each ATM network requires separate physical adapters for the service and non-service IP labels.

Configuring ATM Classic IP

An ATM Classic IP interface is either a Classic IP client or a Classic IP server. The server performs ATM address resolution for all clients and the connection setup between clients. Each logical IP subnet requires its own server. Clients maintain their own ARP cache. For packets sent to an IP address that is *not* contained in the ARP cache, the client sends a request to the server of its subnet, which sends a broadcast to determine the ATM address.

The current ATM Classic IP support in HACMP has the following restrictions:

- A node belonging to the HACMP cluster cannot perform as a Classic IP *server* for a Classic IP network. Only Classic IP *clients* can be defined as HACMP interfaces.
- The use of Alternate Hardware Addresses is *not* supported on ATM networks.

Configuring Classic IP for HACMP Cluster Networks

A cluster network consisting of service and non-service IP labels requires that you have two Classic IP servers configured, one for each IP subnet.

Before you can configure ATM Classic IP for cluster networks, the following must already be configured:

- All device IP labels of a cluster network to belong to the same IP subnet
- All non-service IP labels to belong to a different IP subnet
- A Classic IP server for each IP subnet.

The ATM addresses of the servers must be known. Now you can configure ATM Classic IP clients on cluster nodes.

To configure ATM Classic IP clients on cluster nodes:

1. On each cluster node, configure the Service and non-service ATM interfaces in AIX to use the “Service” and “non-service” Classic IP servers previously configured.
2. Test the configuration.
3. Define the ATM network to HACMP.

Testing the Configuration

If the ATM interface cards on which the interfaces have been configured are connected to the ATM network, and the ATM network is functional, the clients will register with the ATM switch and connect with the Classic IP server of the subnet to which they belong.

To test Classic IP client communication over the network, confirm the following:

1. The IP addresses on all nodes are reachable. Use the **ping** command to confirm this.
2. The ATM device is up and running. Use the **ifconfig** command to review the status of the ATM device.

   ```
   >ifconfig at1
   at1: flags=e000861<UP,NOTRAILERS,RUNNING,SIMPLEX,GROUPRT>
           inet 192.168.111.30 netmask 0xffffff00
   ```

   If the RUNNING flag is set, the interface has connected with its Classic IP server, and is operational.
3. The Classic IP client is registered with the Classic IP server. Use the **arp** command to confirm the registration of the client with its server. See the example of the output of the **arp** command after step 4. The client has registered with its server (at1), server_192_168_111.
4. The ATM TCP/IP layer is functional. Use the **arp** command to confirm this.

   The following example shows output of the **arp** command. The ATM layer is functional, since the first 13 bytes of the hardware address of the client at1 correspond to the address of the ATM switch.

   ```
   > arp -t atm -a


   SVC - at0 on device atm0 -
   ==========================
   at1(192.168.110.30)
   47.0.5.80.ff.e1.0.0.0.f2.1a.39.65.42.20.48.1a.39.65.0
     IP Addr        VPI:VCI Handle ATM Address
     server_192_168_110 (198.168.110.99) 0:761   12
   47.0.5.80.ff.e1.0.0.0.f2.1a.39.65.88.88.88.88.10.00.0

   SVC - at1 on device atm1 -
   ==========================
   at1(192.168.111.30)
   47.0.5.80.ff.e1.0.0.0.f2.1a.39.65.42.35.42.1f.36.221
     IP Addr        VPI:VCI Handle ATM Address
     server_192_168_111 (198.168.111.99) 0:772 23
   47.0.5.80.ff.e1.0.0.0.f2.1a.39.65.88.88.88.88.11.00.1
   ```

Configuring ATM ARP Servers for Use by HACMP Nodes

Before configuring an ATM ARP server, install the ATM interfaces 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 Non-service subnet.
3. Determine the ATM server address for each ATM server.

Configuring ATM ARP Servers for Service Subnetworks

To configure an ARP server for the HACMP service subnetwork:

1. Enter `smit 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 fields on the **Change/Show an ATM Interface** panel.

Network Interface Name	at0
INTERNET ADDRESS (dotted decimal)	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.

Configuring ATM ARP Servers for Non-Service Subnetworks

To configure an ARP server for an HACMP non-service subnetwork:

1. Enter `smit chinet`

   SMIT displays a list of available Network Interfaces.
2. Select **at1** as the network interface.

The following example shows the values of the fields on the **Change/Show an ATM Interface** panel:

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	atm0
Idle Timer	60
Current STATE	up

Note: The interface name is (at1) for the non-service interface; the **Connection Type** designates the interface as an ARP server, **svc-s**. The **Alternate Device** field is set to atm0. This setting puts at1 on atm0 with at0. The non-service subnet is 192.168.111.

Configuring ATM ARP Clients on Cluster Nodes

To configure ATM ARP clients on cluster nodes:

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

Configuring the Cluster Nodes as ATM ARP Clients

To configure HACMP cluster nodes as ATM ARP clients:

1. Use the **smit chinet** command to configure two ATM interfaces, one on each adapter:
   - at0 on atm0 for service
   - at1 on atm1 for non-service
2. To configure the *service* subnet, specify values for the following fields for these interfaces:

Network Interface Name	at0
INTERNET ADDRESS (dotted decimal0	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.21.e.8.0.5a.99.82.95.0

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 in the 192.168.110 subnet. The ATM server address is the 20-byte address that 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 non-service address. If IPAT is *not* enabled, the **INTERNET ADDRESS** represents the service address.

3. To configure the *non-service* subnet, specify values for the following fields for these interfaces:

Network Interface Name	at1
INTERNET ADDRESS (dotted decimal0	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.21.e.8.0.5a.99.82.95.1
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 non-service subnet, the **INTERNET ADDRESS** must be in the 192.168.111 subnet. The ATM server address is the 20-byte address that 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. It is 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 network, define it to the HACMP cluster topology as a network. For information about how to define an ATM network in an HACMP cluster, see the *Administration Guide*.

Configuring 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. You can use ATM LAN emulation to bridge existing Ethernet or Token-Ring networks—particularly switched, high-speed Ethernet—across an ATM backbone network.

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 the lpp **bos.atm** from AIX on each machine if you have *not* already done so.

To configure ATM LAN emulation:

1. Enter `atmle_panel`

   SMIT displays the ATM LAN Emulation menu.
2. Select **Add an ATM LE Client**.
3. Select 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** panel. 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	Select **yes** if you want automatic configuration. The default is **No**.

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	Review the 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 LE network, define it to the HACMP cluster topology as a public network. For information about how to define networks and interfaces in an HACMP cluster, see Chapter 1: Administering an HACMP Cluster in the *Administration Guide*.

You will define these virtual adapters to HACMP just as if they were real interfaces, except you cannot use Hardware Address Takeover (HWAT).

Special Considerations for Configuring ATM LANE Networks

Cluster verification does *not* check whether the IP Address is configured on the interface stored in the HACMPadapter object in the HACMP configuration database. When configuring a cluster adapter, the interface stanza in HACMPadapter is *not* specified. During topology synchronization or when applying a snapshot, the interface stanza gets assigned a value corresponding to the AIX network configuration at that moment. If the IP Address is assigned to a different interface later, HACMPadapter no longer contains the correct information about the corresponding cluster adapter. Depending on the configuration, such an error may go unnoticed until a certain cluster event occurs, and then cause the cluster manager to exit fatally.

Therefore, after any changes to the cluster or AIX network configuration, the cluster topology should be synchronized.

This mistake is likely to occur when configuring a cluster network over ATM. For example, you may try to correct the network configuration after a verification error or warning by deleting a network interface and moving an IP address to a different interface. A topology synchronization must be done to update the interface stanza in the HACMPadapter object in the HACMP configuration database.

Avoiding Default Gateways

When configuring TCP/IP for clients over ATM networks, no default gateways should be configured that would cause packets to be sent over networks other than ATM networks. If you have an incorrect ATM configuration or an ATM hardware failure, clients' attempts to connect with their corresponding servers by sending packets out to the gateway would be unsuccessful. The generated traffic could severely impact the performance of the network that connects to the gateway and affect how your cluster performs.

Configuring ATM LANE Networks for HACMP

For HACMP networks that are configured over emulated LANs (ELAN), configure all interfaces of a given cluster network over clients that belong to the same ELAN.

Conceptually, all clients that belong to the same ELAN correspond to LAN adapters connected to the same wire. They can communicate with each other if they are on the same subnet. Multiple IP subnets can exist on the same ELAN.

Clients that belong to different ELANs generally cannot communicate with each other if bridging is *not* configured.

Bridged ELANs are *not* supported for an HACMP network configuration. If you intend to use a bridged ELAN configuration, ask an ATM network representative whether it conforms to the requirements of HACMP. If the interfaces of a cluster network do *not* belong to the same ELAN, the cluster may *not* generate a network-related event if there is a loss of connectivity. For example, if the interfaces on different nodes are assigned to LANE clients that belong to different ELANs, it is possible that no network-related cluster events would be generated indicating this configuration error, even though HACMP clients would *not* be able to connect to them.

The following figure shows a cluster consisting of two nodes (A, B) and three networks (net-1, net-2, net-3). One network is configured on clients belonging to ELAN 1. The adapters are on two subnets. Networks net-2 and net-3 are configured on clients belonging to ELAN 2.

Note: This configuration would *not* be supported by all ATM switches.

ATM LANE Cluster Configuration

Switch Dependent Limitations for Emulated LANs in HACMP

Two clients that belong to the same ELAN and are configured over the same ATM device cannot register with the LAN Emulation Server (LES) at the same time. Otherwise, it cannot be determined when a client registered with the switch.

This is true for the A-MSS router, used standalone or in the IBM Nways 8260 or IBM Nways 8265 switching hub. The number of clients per ELAN and adapters that are allowed to register with the LAN Emulation server concurrently may be a user-configurable setting.

If this limitation is applicable to your configuration, avoid configuring multiple clients that belong to the same ELAN over the same ATM device. In particular, when configuring a cluster network over LANE, keep in mind the following:

- No two cluster interfaces can be configured on clients that belong to the same ELAN and that are configured over the same ATM device.
- For any cluster interface that is configured over LANE, no other LANE client can belong to the same ELAN configured over the same ATM device.

If this limitation is violated, a network related cluster event indicating a loss of connectivity may be generated, most likely after a cluster event that involves changes to clients on the same ATM network interface.

Configuring ELANs

When configuring ELANs, ensure the following:

- All LANE clients in the entire ATM network—*not only* those used for HACMP—should be configured correctly and registered with the switch.
- The switch log should *not* indicate any errors.
- It is *not* sufficient that each cluster interface configured over LANE is reachable from all other nodes. All interfaces of one cluster network should be configured on clients belonging to the correct ELAN.

Configuring Shared External Disk Devices

This section describes how to ensure that shared external disk devices are configured properly for the HACMP software. Separate procedures are provided for Fibre Channel, SCSI disk devices, IBM SCSI disk arrays, and IBM Serial Storage Architecture (SSA) disk subsystems.

Note: For information on installing and configuring OEM disks, see Appendix B: OEM Disk, Volume Group, and File Systems Accommodation.

Configuring Shared Fibre Channel Disks

Use the following procedure to ensure access to Fibre Channel disks and to record the shared disk configuration on the Fibre Channel Disks Worksheet, as described in the *Planning Guide*. Use a separate worksheet for each node. You refer to the completed worksheets when you configure the cluster.

Note: The fileset **ibm2105.rte** must be installed on a node to properly display information for Fibre Channel devices in an IBM 2105 Enterprise Storage System.

To ensure that a node has access to shared Fibre Channel disks and to complete the Fibre Channel Disks Worksheet:

1. Record the name of the node in the **Node Name** field.
2. Ensure that the Fibre Channel adapters on the cluster nodes are available. Run the **lsdev -Cc adapter** command, for example:

```
lsdev -Cc adapater | grep fcs

fcs0    Available 31-08    FC Adapter
.
.
.
```

3. Record the names of the Fibre Channel adapters in the **Fibre Channel Adapter** field. Allow sufficient space to list the disks associated with each adapter.

4. Ensure that the disks associated with that adapter are available by running the **lsdev -C** and **grep** for the location shown in the output from the **lsdev -Cc adapter** command, for example:

```
lsdev -C | grep 31-08
fcs0       Available 31-08         FC Adapter
hdisk1     Available 31-08-01      Other FC SCSI Disk Drive
fscsi0     Available 31-08-01      FC SCSI I/O Controller Protocol
Device
hdisk2     Available 31-08-01      Other FC SCSI Disk Drive
hdisk3     Available 31-08-01      Other FC SCSI Disk Drive
hdisk4     Available 31-08-01      Other FC SCSI Disk Drive
hdisk5     Available 31-08-01      Other FC SCSI Disk Drive
hdisk6     Available 31-08-01      Other FC SCSI Disk Drive
hdisk7     Available 31-08-01      Other FC SCSI Disk Drive
hdisk8     Available 31-08-01      Other FC SCSI Disk Drive
hdisk9     Available 31-08-01      Other FC SCSI Disk Drive
hdisk10    Available 31-08-01      Other FC SCSI Disk Drive
hdisk11    Available 31-08-01      Other FC SCSI Disk Drive
hdisk12    Available 31-08-01      Other FC SCSI Disk Drive
```

   Output may vary depending on the disk device in use.

5. Record the names of the Fibre Channel disks associated with each Fibre Channel adapter in the **Disks Associated with Adapter** list.

If the list of disks is different than expected, or if disks are *not* available:

1. Make sure that the volume groups are configured correctly on the storage system.

   See the product documentation for the disk device for information about configuring volume groups.

2. Make sure that the Fibre Channel switch is set up correctly and that zoning is configured appropriately.

   See the product documentation for the switch for information about configuring zoning.

3. Check the World Wide Name by adapter:

```
lscfg -vpl fcs0
```

   Look for an entry similar to the following:

```
Network Address.............10000000C92EB183
```

4. Ensure that the World Wide Name is configured correctly on both the disk device and the Fibre Channel switch.

Configuring Shared SCSI Disks

As you review the installation, record the shared disk configuration on the Shared SCSI Disk Worksheet, as described in the *Planning Guide*. Use a separate worksheet for each set of shared SCSI disks. You refer to the completed worksheets when you configure the cluster.

To ensure that a SCSI disk is installed correctly and to complete the Shared SCSI Disk Worksheet:

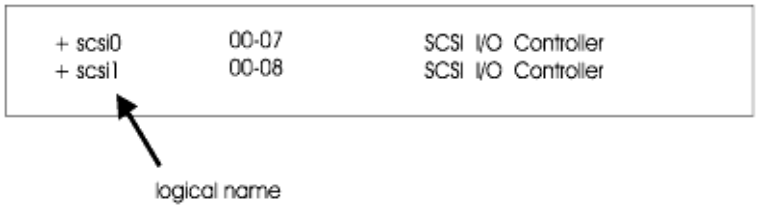
1. Record the node name of each node that connected to the shared SCSI bus in the **Node Name** field.

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

To determine the logical name, use the command:

```
lscfg | grep scsi
```

In the command output, the first column lists the logical name of the SCSI adapters, for example scsi0, as shown in the following figure:



3. Record the I/O slot (physical slot) that each SCSI adapter uses in the **Slot Number** field. Use the **lscfg -vpl** command, for example:

   ```
   lscfg -vpl scsi0
   ```

4. Record the SCSI ID of each **SCSI adapter** on each node in the **Adapter** field. 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*:

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

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

   In the resulting output, the first column lists the attribute names. The integer to the right of the **id** attribute is the adapter SCSI ID:

5. Record the SCSI IDs of the physical disks in the **Shared Drive** fields. Use the command **lsdev -Cc disk -H**

   The third column of the command output is a numeric location with each row in the format AA-BB-CC-DD. The first digit (the first D) of the DD field is the SCSI ID.

   Note: For SCSI disks, a comma follows the SCSI ID field of the location code since the IDs can require two digits, such as 00-07-00-12,0 for a disk with an ID of 12.

```
name            status          location        description
hdisk00         Available       00-07-00-00     73 GB SCSI Disk Drive
hdisk01         Available       00-07-00-10     73 GB SCSI Disk Drive
hdisk02         Available       00-07-00-20     73 GB SCSI Disk Drive
```

6. At this point, ensure 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.) For example, the adapter on one node can have SCSI ID 6 and the adapter on the other node can be SCSI ID 5, and the external disk SCSI IDs should be an integer from 0 through 4. Do *not* use SCSI ID 7, as this is the default ID given when a new adapter is installed.
7. Determine the logical names of the physical disks. The first column of the data generated by the **lsdev -Cc disk -H** command lists the logical names of the SCSI disks.

   For sample command output, see step 5.

   Record the name of each external SCSI disk in the **Logical Device Name** field, and the size in the **Size** field - the size is usually part of the description. *Be aware that the nodes can assign different names to the same physical disk. Note these situations on the worksheet.*
8. Ensure that all disks have a status of **Available**. The second column of the output generated by the **lsdev -Cc disk -H** command shows the status.

   If a disk has a status of **Defined** (instead of **Available**) ensure that the cable connections are secure, and then use the **mkdev** command to make the disk available.

At this point, you have verified that the SCSI disk is configured properly for the HACMP environment.

Configuring IBM SCSI Disk Arrays

Ensure that the disk device is installed and configured correctly. Record the shared disk configuration on the Shared IBM SCSI Disk Arrays Worksheet, as described in the *Planning Guide*. Use a separate worksheet for each disk array. You refer to the completed worksheets when you define the cluster topology.

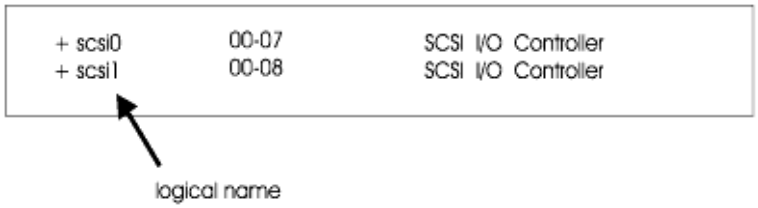
To confirm the configuration and complete the Shared IBM SCSI Disk Arrays Worksheet:

1. Record the node name of each node that connected to the shared SCSI bus in the **Node Name** field.
2. Enter the logical name of each adapter in the **Logical Name** field.

   To determine the logical name, use the command:

   ```
   lscfg | grep scsi
   ```

   In the command output, the first column lists the logical name of the SCSI adapters, for example scsi0, as shown in the following figure:



3. Record the I/O slot (physical slot) that each **SCSI adapter** uses in the **Slot Number** field. Use the **lscfg -vpl** command, for example:

   ```
   lscfg -vpl scsi0
   ```

4. Record the SCSI ID of each **SCSI adapter** on each node in the **Adapter** field. 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*:

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

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

   In the resulting output, the first column lists the attribute names. The integer to the right of the **id** attribute is the adapter SCSI ID:

At this point, you have ensured that the disk array is configured properly for the HACMP software.

Configuring Target Mode SCSI Connections

This section describes how to configure a target mode SCSI connection between nodes sharing disks connected to a SCSI 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: Only parallel SCSI adapters support target mode SCSI.

Configuring the Status of SCSI Adapters and Disks

To define a target mode SCSI connection (tmscsi), 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 review the status of SCSI adapters you intend to use, enter the following:

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

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

To review the status of SCSI disks on the SCSI bus, enter the following:

```
lsdev -Cc disk
```

If either an adapter or disk is **Available**, follow the steps in the section Returning Adapters and Disks to a Defined State to return both the adapter (and its disks) to a **Defined** state so that they can be configured for target mode SCSI and set to an **Available** state.

Returning Adapters and Disks to a Defined State

For a SCSI adapter, use the following command to make each **Available** disk associated with an adapter defined:

```
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 you are 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 scsix
```

When all controllers and disks are **Defined**, see the section Configuring Target Mode SCSI Devices in AIX to enable the Target Mode connection.

Note: Target mode SCSI may be automatically configured depending on the SCSI adapter in use. In this case, skip ahead to the section on Defining the Target Mode SCSI Connection to HACMP.

Configuring 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.

Note: Only parallel SCSI adapters support target mode SCSI.

Enabling the 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.

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. Exit SMIT after the **cfgmgr** command completes.
4. Run the following command to ensure that the devices are paired correctly:

   ```
   lsdev -Cc tmscsi
   ```

Repeat the procedures in the sections Enabling the Target Mode Interface and Configuring the Target Mode SCSI Device for other nodes connected to the SCSI bus.

Configuring the target mode connection creates two target mode files in the **/dev** directory of each node:

- **/dev/tmscsinn.im.** The initiator file that transmits data
- **/dev/tmscsinn.tm.** The target file that 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 bus.

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

   where *nn* must be the logical name representing the target node. (This command hangs and waits for the next command.)
2. 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 *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, shut down HACMP before reattaching the SCSI bus to that node. *Never attach a SCSI bus to a running system.*

Defining the Target Mode SCSI Connection to HACMP

After you install and test the target mode SCSI bus, define the target mode connection as a point-to-point network to HACMP. For information about how to configure a target mode SCSI network, see the *Administration Guide*.

Configuring Target Mode SSA Connections

This section describes how to configure a target mode SSA (tmssa) connection between HACMP 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 point-to-point network to HACMP that connects all nodes on an SSA loop.

Changing Node Numbers on Systems in SSA Loop

By default, SSA node numbers on all systems are zero.

To configure the target mode devices:

1. Assign a unique non-zero SSA node number to all systems on the SSA loop.

   Note: The ID on a given SSA node should match the HACMP node ID that is contained in the **node_id** field of the HACMPnode entry.

   The following command retrieves the HACMP node ID:

   ```
   odmget -q "name = node_name" HACMPnode
   ```

2. To change the SSA node number use the following command:
   chdev -l ssar -a node_number=*number*
3. To show the system's SSA node number use the following command.

   ```
   lsattr -El ssar
   ```

Configuring Target Mode SSA Devices in AIX

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

To create the initiator and target devices:

1. Enter `smit devices`

   SMIT displays a list of devices.

2. Select **Install/Configure Devices Added After IPL** and press Enter.

3. Exit SMIT after the **cfgmgr** command completes.

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

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

Repeat the procedures for enabling and configuring the target mode SSA devices for other nodes connected to the SSA adapters.

Configuring the target mode connection creates two target mode files in the **/dev** directory of each node:

- **/dev/tmssan.im** where n represents a number. The initiator file that transmits data
- **/dev/tmssan.tm** where n represents a number. The target file that 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 any short ascii file.

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

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

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

Configuring Shared IBM SSA Disk Subsystems

When planning a shared IBM SSA disk subsystem, record the shared disk configuration on the Shared IBM Serial Storage Architecture Disk Subsystems Worksheet, as described in the *Planning Guide*. Use a separate worksheet for each set of shared IBM SSA disk subsystems.

To complete a Shared IBM Serial Storage Architecture Disk Subsystems Worksheet:

1. Record the node name of each node connected to the shared IBM SSA disk subsystem in the **Node Name** field.
2. Record the logical device name of each adapter in the **Adapter Logical Name** field.

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

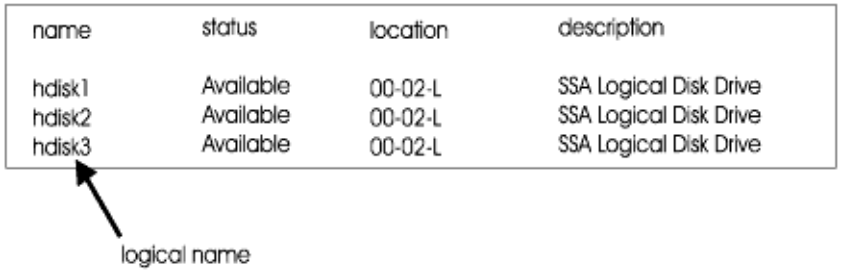
   ```
   lscfg | grep ssa
   ```

   The first column of command output lists the logical device names of the SSA adapters.
3. For each node, record the slot that each adapter uses in the **Slot Number** field. The slot number value is an integer value from 1 through 16. Use the **lscfg -vpl** command, for example:

   ```
   lscfg -vpl sssa0
   ```
4. Determine the logical device name and size of each physical volume and record the values. On each node run the command:

   ```
   lsdev -Cc disk | grep -i ssa
   ```

   The first column command output lists the logical names of the disks.



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

   Record the size of each external disk in the **Size** field.
5. Ensure that all disks have a status of **Available**. The second column of the existing output indicates the disk status.

   If a disk has a status of **Defined** (instead of **Available**) ensure that the cable connections are secure, and then use the **mkdev** command to make the disk available. At this point, you have verified that the IBM SSA disk is configured properly for the HACMP software.

Defining the Target Mode SSA Connection to HACMP

After you install and test the SSA target mode connection, define the connection as a point-to-point network to HACMP. For information about how to configure a target mode network, see the *Administration Guide*.

Installing and Configuring Shared Tape Drives

HACMP supports both SCSI and Fibre Channel tape drives. Instructions for each are included here. For more general information, see the chapter on Planning Shared Disk and Tape Devices in the *Planning Guide*.

As you install an IBM tape drive, record the key information about the shared tape drive configuration on the Shared IBM SCSI or Fibre Channel Tape Drives Worksheet. Complete a separate worksheet for each shared tape drive.

Installing Shared SCSI Tape Drives

Complete the procedures in this section to install a shared IBM SCSI tape drive.

Prerequisites

Be sure to install the appropriate SCSI 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.

Installing an IBM SCSI Tape Drive

To install an IBM SCSI tape drive and complete a Shared IBM SCSI Tape Drive Worksheet, as described in the *Planning Guide*:

1. Review the shared tape configuration diagram you drew while planning tape drive storage needs.
2. Record the name of each node connected to this shared SCSI bus in the **Node Name** field.
3. Name each SCSI adapter used in this shared SCSI bus and record the name in the **SCSI Adapter Label** field of the configuration worksheet. For example, AIX may name the adapters *scsi0, scsi1,* and so on.
4. Record the I/O slot of each SCSI 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 command output, 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.
5. Record the logical device name of each adapter in the **Logical Name** field. The first column of the **lscfg** command output 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 higher SCSI IDs in sequence, such as 8 and 9. The tape drive SCSI IDs should later be set to an integer starting at 0 and going up. Make sure no SCSI device on the same bus 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 host adapters to 8 and 9 to avoid a possible conflict when booting one of the systems in service mode from a **mksysb** tape or other boot device, since this will always use an ID of 7 as the default. For limitations specific to a type of SCSI adapter, see the documentation for the device.

Note that the integer value in the logical device name (for example, the *1* in *scsi1*) is *not* a SCSI ID, but simply part of the name given to the configured device.

To review and change SCSI IDs:

- To determine the SCSI IDs of the tape drive 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 *scsi0* is obtained:

  ```
  lsattr -E -l scsi0 | grep external_id
  ```

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

  In the command output, 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 'scsi0' -a 'external_id=6' -P
  ```
- Ensure that each SCSI 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 all nodes connected to the SCSI bus so that you can set the SCSI IDs for the tape drives and connect the cables. Use the **shutdown** command to shutdown the nodes.
8. Assign each controller on the tape drive a SCSI ID that is unique on this shared SCSI bus.

   See the documentation for your tape drive to find out how to set SCSI IDs.

   Refer to your worksheet for the values previously assigned to the adapters. For example, if the host adapters have IDs of 8 and 9, you can assign the tape drives any SCSI ID from 0 through 6.
9. Record the SCSI ID of each tape drive in the **SCSI Device ID: Tape Drive** fields on the worksheet.
10. Connect the cables.
11. Power on the tape drive and all nodes; then reboot each node.
12. Run **cfgmgr** on one node at a time to complete the installation.

Logical Device Names

All nodes that are connected to the SCSI tape drive must have the same Logical Device Name (for example,**/dev/rmt0**). If the names differ (**/dev/rmt0** and **/dev/rmt1**, for example) perform the steps in the following procedure.

To configure logical device names:

1. On the nodes with the lower numbers, put the tape device in the **Defined** state with the **rmdev -l** command. For example, run **rmdev -l rmt0**.
2. Enter `smit tape`
3. From the **Tape Drive** panel, select **Change/Show Characteristics of a Tape Drive**.
4. Select the drive you put in the **Defined** state and note its characteristics in the various fields.
5. From the **Tape Drive** panel, select the **Add a Tape Drive**.
6. Use the information gathered to create a new Logical Device Name.
7. Review all nodes to assure that the Logical Device Name is in the **Available** state and that the external tape drive has this same Logical Name.

Installing and Configuring Shared Fibre Tape Drives

Complete the procedures in this section to install a shared IBM Fibre tape drive.

Prerequisites

Be sure to install the appropriate Fibre Channel 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.

The IBM Fibre Channel tape drive also requires the installation of IBM AIX Enhanced Tape and Medium Changer Device Drives (Atape). Follow the installation directions that come with this device driver.

Installing an IBM Fibre Channel Tape Drive

To install an IBM Fibre Channel tape drive and complete a Shared IBM Fibre Tape Drive Worksheet, as described in the *Planning Guide*:

1. Review the shared tape configuration diagram you drew while planning tape drive storage.
2. Record the name of each node connected to this shared Fibre Channel bus in the **Node Name** field.
3. Name each Fibre Channel adapter used in this shared Fibre bus and record the name in the **Fibre Adapter Label** field. For example, AIX may name the adapters *fcs0, fcs1,* and so on.
4. Record the I/O slot of each Fibre Channel adapter used in this shared Fibre bus in the **Slot Number** field.

   To determine the slot number, use the **lscfg** command, as in the following example. Note that FC must be capital letters:

   ```
   lscfg | grep FC
   ```

In the command output, 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.

```
fcs0    04-02       FC Adapter

fscsi0  04-02-02    FC SCSI I/O Controller Protocol Drive
```

5. Record the logical device name of each adapter in the **Logical Name** field. The first column of the output generated by the **lscfg** command lists the logical name of the Fibre channel adapters.
6. Connect the cables.
7. Power on the tape drive and all nodes; then reboot each node.
8. Run **cfgmgr** on one node at a time to complete the installation.

Logical Device Names

All nodes that are connected to the Fibre tape drive must have the same Logical Device Name (**/dev/rmt0**, for example). If they differ (**/dev/rmt0** and **/dev/rmt1**, for example), perform the steps in the following procedure:

To configure logical device names:

1. On the node with the lower numbers, put the tape device in the **Defined** state with the **rmdev -l** command. For example, run **rmdev -l rmt0**.
2. Enter `smit tape`
3. From **Tape Drive** panel, select **Change/Show Characteristics of a Tape Drive**.
4. Select the drive you have put in the **Defined** state and note its characteristics as **Defined** in the various fields.
5. From **Tape Drive** panel, select **Add a Tape Drive**.
6. Use the information gathered to create a new Logical Device Name.
7. Review both nodes to assure that the Logical Device Name is in the **Available** state and that the external tape drive has this same Logical Name.

Configuring the Installation of a Shared Tape Drive

During the boot process, 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 *scsi* or *fscsix,* where *x* is an integer. For example, an adapter could be named *scsi0* or *fscsi0* (Fast/Wide Adapters are named *scsix*). After AIX configures the SCSI adapter, it probes the SCSI bus and configures all target devices connected to the bus.

To confirm the installation of a tape drive:

1. Ensure that AIX created the device definitions that you expected.

   To determine the logical name of the tape drive, use the **lsdev** command as follows:

   ```
   lsdev -Cc tape
   ```

   For example, the **lsdev** command may return output that resembles the following:

   ```
   Name   Status          Location        Description
   ```

```
rmt0   Available       00-02-00-00     Other SCSI Tape Device
```

2. Ensure that the logical device name for your tape device has a status of **Available**.

   If the tape drive has a status of **Defined** (instead of **Available**) ensure that the cable connections are secure, and then use the **mkdev** command to make the tape available. Enter:

   ```
   mkdev -l rmtx
   ```

   where **rmtx** is the logical name of the defined tape drive.

At this point, your tape drive installation is complete.

For information on how to configure tape drives as resources in resource groups, see the *Administration Guide*.

Chapter 7: Defining Shared LVM Components

This chapter describes how to define the LVM components shared by the nodes in an HACMP cluster. This chapter contains the following sections:

- Overview
- Defining Shared LVM Components for Non-Concurrent Access
- Defining LVM Components for Concurrent Access.

Overview

Creating the volume groups, logical volumes, and file systems shared by multiple nodes in an HACMP cluster requires that you perform steps on all those nodes. In general, you define the components on one node (referred to in the text as the source node) and then import the volume group onto each of the other nodes that will be part of the resource group (referred to as destination nodes). This ensures that the definitions in the HACMP configuration database for the shared components are the same on all the nodes that may access them.

Using SMIT, you can collect information about all volume groups available either on a local node, or on all nodes in the cluster defined to the resource group. Later, you can automatically import these volume groups onto all the destination nodes, if needed. As you configure shared resources in SMIT, when you add a volume group to a resource group, you can select the option to automatically import the volume group onto all the destination nodes in the participating resource group.

HACMP non-concurrent access environments typically use journaled file systems to manage data, while concurrent access environments use raw logical volumes. This chapter provides instructions for both types of environments.

While configuring file systems for a resource group, you can select one or more individual file systems to be mounted in the particular resource group. In addition, you can specify that all file systems in the specified shared volume groups to be mounted at runtime. Using this option allows you to add or delete file systems in the particular shared volume group without resynchronizing the cluster resources after each update.

Defining Shared LVM Components for Non-Concurrent Access

HACMP non-concurrent access environments typically use journaled file systems to manage data, although some database applications may bypass the journaled file system and access the logical volume directly.

The key consideration is whether the environment uses mirrors. Shared logical volumes residing on SCSI devices, and IBM 7133 SSA devices should be mirrored in AIX to eliminate the disk as a single point of failure. Shared volume groups residing on an IBM 2105 E10 and E20 Enterprise Storage Server RAID devices and IBM TotalStorage DS6000 and DS8000 should *not* be AIX mirrored; these systems provide their own data redundancy.

The following figures list the tasks you complete to define the shared LVM components with and without mirrors. 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.

Defining Shared LVM Components

You can define shared LVM components with or without mirroring.

To define shared LVM components on the source node:

1. Create the volume group.
2. Create the journaled file system.
3. Rename the **jfslog** and the logical volume, so that they are unique across the cluster.
4. Mirror the **jfslog** and logical volume (only with AIX mirroring on *one node* in the cluster).
5. Vary off the volume group.

To define shared LVM components (with or without AIX mirroring) *on the other nodes* that will access this volume group:

1. Import the volume group.
2. Change the volume group options so that it does *not* vary on at system startup.
3. Vary off the volume group.

Creating a Shared Volume Group on Source Node

To create a shared volume group on a source node:

1. Enter `smit mkvg`
2. Use the default field values unless your configuration has other requirements, or unless you are specifically instructed otherwise here. (Only fields with specific HACMP instructions are shown.)

Note: If you are creating a mirrored volume group, select two or three times as many disks for the volume group as you will need for your data, depending on the number of mirrors you select.

VOLUME GROUP name	The name of the shared volume group must be unique within the cluster and distinct from the service IP label/address and resource group names; it should relate to the application it serves, as well as to any corresponding device, such as `websphere_service_address`.
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.
Volume Group MAJOR NUMBER	If you will be using NFS, 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 File System on Source Node

To create a shared file system on a source node:

1. Enter `smit crjfs`

   When you create a journaled file system, AIX creates the corresponding logical volume. Therefore, you do *not* need to define a logical volume.

2. Rename both the logical volume and the log logical volume for the file system and volume group.
3. Review the settings for the following fields:

Mount AUTOMATICALLY at system restart?	Make sure this field is set to **no**.
Start Disk Accounting	Set this field to **no** unless you want disk accounting.

Renaming jfslogs and Logical Volumes on 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 file system 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 file system and the corresponding **jfslog** logical volume. Use a naming scheme that indicates the logical volume is associated with a certain file system. For example, *lvsharefs* could name a logical volume for the */sharefs* file system

To rename **jfslogs** and logical volumes:

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 command output:
   - 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.
3. 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 file system 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 Source Node

Note: This step is *not* needed for disk subsystems that provide their own mirroring, such as the IBM 2105 Enterprise Storage Server and IBM Total Storage DS6000 and DS8000.

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 file systems. 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.
2. Review the number of logical volume copies. Enter:

   `lsvg -l volume_group_name`

   In the command output, 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 review 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 command output, 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 File System

To run a consistency check on each file system's information:

1. Enter:
   `fsck /filesystem_name`
2. Ensure that you can mount the file system by entering:
   `mount /filesystem_name`
3. Ensure that you can unmount the file system 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 the destination node and activated as appropriate by the HACMP event scripts. Enter the following command:

`varyoffvg volume_group_name`

Collecting Information on Current Volume Group Configuration

HACMP can collect information about all volume groups available across all nodes in the cluster as part of the discovery process. HACMP collects information about all shared volume groups that are currently available on the nodes in the cluster. It then determines which volume groups can be imported to the other nodes in the resource group. HACMP filters out those volume groups that are already included in any other resource groups. You can use this information later to import discovered volume groups onto other nodes in the resource group that do *not* have these volume groups.

To collect the information about volume group configuration:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration** > **Discover HACMP-related Information from Configured Nodes** and press Enter. The information on current volume group configuration is collected, as part of the general configuration information.

Importing a Volume Group onto Destination Nodes

Importing the volume group onto the destination nodes synchronizes the definition in the HACMP configuration database for the volume group on each node on which it is imported.

When adding a volume group to the resource group, you can manually import a volume group onto the destination nodes or automatically import it onto all the destination nodes in the resource group.

Importing a Volume Group Automatically

You can set up automatic import of a volume group.

To automatically import a volume group:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Group Configuration > Change/Show all Resources for a Resource Group** and press enter.

This option enables HACMP to automatically import shareable volume groups onto all the destination nodes in the resource group. Automatic import allows you to create a volume group and then add it to the resource group immediately, without manually importing it onto each of the destination nodes in the resource group.

Before automatically importing volume groups, make sure that you have collected the information on appropriate volume groups, using the **Discover HACMP-related Information from Configured Nodes** option.

Note: Each volume group is assigned a major number when it is created. When HACMP automatically imports a volume group, the major number already assigned to the volume group will be used if it is available on all the destination nodes. Otherwise, any free major number will be used. NFS fallover requires that the major number be the same on all nodes. For more information, see Using NFS with HACMP in the *Planning Guide.*

Prerequisites and Notes

For HACMP to import available volume groups, ensure that the following conditions are met:

- Volume group names must be the same across cluster nodes, and unique to the cluster.
- Logical volumes and file systems must have unique names.
- All physical disks must be known to AIX and have PVIDs assigned.
- A volume group is available for auto import onto other nodes only if the physical disks on which it resides are available to all of the nodes in the resource group.

Importing a Volume Group Automatically

To add volume groups to a resource group via automatic import:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Group Configuration > Change/Show all Resources for a Resource Group** and press enter.
3. Select the resource group for which you want to define a volume group and press Enter. A panel appears with fields for all types of resources.
4. The **Automatically Import Volume Groups** flag is set to **False** by default. Set this flag to **True**.
5. In the **Volume Groups** field, select the volume groups from the picklist or enter volume groups names.

If you ran the discovery process before starting this procedure, review the list of all existing volume groups cluster-wide (excluding **rootvg**). The list includes all shared volume groups in the resource group *and* the volume groups that are currently available for import onto the resource group nodes. HACMP filters out from the cluster-wide list those volume groups that are already included in any of the resource groups.

6. Press Enter. HACMP determines whether the volume groups that you entered or selected in the **Volume Groups** field need to be imported to any of the nodes that you defined for the resource group, and proceeds to import them, as needed.

Final State of Automatically Imported Volume Groups

When HACMP automatically imports volume groups, the final state (varied on or varied off) depends on the initial state of the volume group and whether the resource group is online or offline when the import occurs.

In all cases, the volume group will end up varied *on* after the resource group is started or after the cluster resources are synchronized, even if it is varied off at some point during the import process.

The following table shows the initial condition of the volume group after its creation, the state of the resource group when the import occurs, and the resulting state of the imported volume group:

Initial Volume Group State	Resource Group State	Auto Imported Volume Group State
Varied ON	OFFLINE	Remains varied ON
Varied ON	ONLINE	Varied ON
Varied OFF	OFFLINE	Varied OFF until the resource group is started
Varied OFF	ONLINE	Varied ON

Importing a Volume Group Manually

If you do *not* want HACMP to import your volume group automatically upon adding it to the resource group, in SMIT make sure that the HACMP **Automatically Import Volume Groups** flag is set to **False** (this is the default).

To set values for the volume group:

1. Enter `smit importvg`
2. Enter field values as follows:

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.

Volume Group MAJOR NUMBER	If you are using NFS, 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 environment, a volume group should be varied on as appropriate by the cluster event scripts. Therefore, after importing a volume group, use the **Change a Volume Group** panel to reconfigure the volume group so that it is *not* activated automatically at system restart.

To change the volume group's startup status:

1. Enter `smit chvg`
2. Select the volume group from the list.
3. Enter these field values as follows:

Activate volume group Automatically at system restart?	Set this field to **no**.
A QUORUM of disks required to keep the volume group on-line?	This field is site-dependent. See the chapter on Planning Shared Disk and Tape Devices in the *Planning Guide* for a discussion of quorum.

Varying Off Volume Group on 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 LVM Components for Concurrent Access

Using concurrent access with HACMP requires installing an additional fileset. For more information, see the section Prerequisites in Chapter 4: Installing HACMP on Server Nodes.

Concurrent access does *not* support file systems. Instead, use logical volumes.

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

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

With HACMP 5.1 and up, you can create enhanced concurrent volume groups. Enhanced concurrent volume groups can be used for both concurrent and non-concurrent access.

Creating a Concurrent Access Volume Group on a Source Node

Take the following steps on the source and destination nodes in an HACMP cluster to create a volume group that HACMP can vary on in concurrent access mode.

Step 1. Complete Prerequisite Tasks

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

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

AIX v.5.2 and up supports enhanced concurrent mode and creates concurrent volume groups in enhanced concurrent mode by default.

Creating a Concurrent Access Volume Group

To use a concurrent access volume group, 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.

To create a concurrent capable volume group from the command line, use the **mkvg** command. The following example creates an enhanced concurrent volume group on hdisk1 and hdisk2:

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

The **mkvg** flags do the following:

-n	Ensures that the volume group does *not* vary on at boot time.
-s	Specifies the partition size.
-C	Creates an enhanced concurrent volume group.
-y	Specifies the volume group name.

You can also use SMIT to run the **mkvg** command.

To create a concurrent access volume group:

1. Enter `smit cl_convg`
2. Select **Create a Concurrent Volume Group.**

   SMIT provides picklists for the node name and physical volume ids (PVIDs) of the disks to be used.

   The **Add a Volume Group** panel displays. Some fields will already be filled and cannot be changed.
3. Enter field values as follows:

VOLUME GROUP name	Specify name of volume group.
Physical partition SIZE in megabytes	Accept the default.
Volume Group Major Number	Accept the default.

4. Press Enter.

   SMIT prompts you to confirm your selections.

5. Press Enter.

Step 3. Import Concurrent Capable Volume Group Information on Destination Nodes

Importing the volume group into the destination nodes synchronizes the definition in the HACMP configuration database for the volume group on each node on which it is imported.

When you add a volume group to the resource group, you may choose to import a volume group manually onto the destination nodes or you may automatically import it onto all the destination nodes in the resource group.

Run the discovery process to make picklists available. See the section Collecting Information on Current Volume Group Configuration for information.

The list of volume groups will include only the concurrent capable volume groups. This list will *not* include **rootvg**, and any volume groups already defined to another resource group.

Final State of Automatically Imported Volume Groups

When HACMP automatically imports volume groups, the final state (varied on or varied off) depends on the initial state of the volume group and whether the resource group is online or offline when the import occurs.

In all cases, the volume group will end up varied ON after the resource group is started or after the cluster resources are synchronized, even if it is varied off at some point during the import process.

This table shows the initial condition of the volume group after its creation, the state of the resource group when the import occurs, and the resulting state of the imported volume group:

Initial Volume Group State	Resource Group State	Auto Imported Volume Group State
Varied ON	OFFLINE	Remains varied ON
Varied ON	ONLINE	Varied ON
Varied OFF	OFFLINE	Varied OFF until the resource group is started
Varied OFF	ONLINE	Varied ON

Importing a Volume Group Manually

In SMIT, if you do *not* want HACMP to import your volume group automatically upon adding it to the resource group, make sure that the **Automatically Import Volume Groups** flag is set to **False** (this is the default), and use the **importvg** fastpath.

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

```
importvg -C -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 that follows.

To import a concurrent capable volume group using SMIT:

1. Enter `smit importvg`

   The **Import a Volume Group** SMIT panel appears.

2. Enter 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.

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 software as described in Chapter 4: Installing HACMP on Server Nodes:

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

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

```
/usr/es/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 run the **varyonvg** command by using the following procedure.

1. To vary on a concurrent capable volume group in non-concurrent mode, enter `smit varyonvg`

   The **Activate a Volume Group** SMIT panel appears.

2. Enter field values as follows:

VOLUME GROUP name	Specify the 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. (AIX v.5.2 field shown here; default is **no**.)

3. Press Enter.

   SMIT prompts you to confirm your selections.

4. Press Enter.

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*.

To create logical volumes on concurrent-capable volume group on source node:

1. Enter `smit cl_conlv`
2. 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. The name must be the same on all cluster nodes.
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. Specifying 1 means no mirroring.
Mirror Write Consistency	Set this value to **no.**

ENABLE BAD BLOCK relocation	Set this field to **no** to disable bad block relocation for concurrent environments (applies to RAID devices).

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 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 become active automatically at system restart. In the HACMP system, a volume group should be varied on as appropriate by the HACMP scripts. Therefore, if you did *not* use the **smit cl_** fastpaths to do your configuration, you must reconfigure the volume group so that it remains dormant at startup.

To change the startup state of a volume group from the command line, enter:

```
chvg -a n volume_group_name
```

To use SMIT to change the startup state of a volume group:

1. Enter `smit chvg`
2. Set the **Activate volume group Automatically at system restart?** field to **no**. For other fields use the defaults or the appropriate entries for your operation.
3. Press Enter to commit this change.
4. Exit SMIT.

Step 8. Complete Follow-up Tasks

Complete any follow-up tasks as needed. For example, ensure that the HACMP scripts can activate the concurrent capable volume group as a concurrent cluster resource.

Chapter 8: Configuring AIX for HACMP

This chapter discusses several general tasks necessary for ensuring that your HACMP environment works as planned. This chapter contains the following sections:

- I/O Considerations
- Networking Considerations
- Planning HACMP File Collections
- Types of Error Notification
- SP-Specific Considerations.

Consult the *AIX System Management Guide and the Performance Management Guide* for more information on these topics.

Note: You can change the settings for a list of tunable values that were changed during the cluster maintenance and reset them to their default settings, that is, to the values that appear in the cluster after manually installing HACMP.

Resetting the tunable values does *not* change any other aspects of the configuration, while installing HACMP removes all user-configured configuration information including nodes, networks and resources.

For more information on using tunable values, see the *Administration Guide*.

I/O Considerations

This section discusses configuration considerations for I/O pacing and **syncd** frequency.

I/O Pacing

In certain circumstances, when one application on the system is doing heavy input/output, I/O can take several seconds to complete for another application. During heavy I/O activity, 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 software may be unable to send keepalive packets from the affected node. The RSCT software on other cluster nodes interprets the lack of keepalive packets 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. However, its packets 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 to redistribute system resources more equitably during periods of high disk I/O activity. 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.

By default, AIX is installed with high- and low-water marks set to **zero**, which disables I/O pacing.

WARNING: I/O pacing and other tuning parameters should only be set to values other than defaults after a system performance analysis indicates that doing so will lead to both the desired and acceptable side effects. *Before changing these settings*, read the section Setting I/O Pacing in Chapter 1: Troubleshooting HACMP Clusters in the *Troubleshooting Guide*.

Although enabling I/O pacing may have only 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, enable I/O pacing.

Although the most efficient settings for 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 slightly reduce write times and consistently generate correct fallover behavior from the HACMP software.

For more information on I/O pacing, see the *AIX Performance Monitoring & Tuning Guide*.

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. Change this value to 10. Note that the I/O pacing parameter setting should be changed first. You do *not* need to adjust this parameter again unless time-outs frequently occur.

Networking Considerations

This sections discusses networking-related considerations when setting up HACMP.

Checking User and Group IDs

If a node 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. User and group IDs should be the same on all nodes.

Configuring Network Options

HACMP requires that the **nonlocsrcroute**, **ipsrcroutesend**, **ipsrcrouterecv**, and **ipsrcrouteforward** network options be set to 1; these are set by RSCT's **topsvcs** startup script. If these options are set to anything besides 1, HACMP issues a warning that RSCT will change them.

The **verification utility** ensures that the value of each network option is consistent across the nodes of a cluster for the following options:

- **tcp_pmtu_discover**
- **udp_pmtu_discover**
- **ipignoreredirects**
- **routerevalidate.**

Changing routerevalidate Network Option

Changing hardware and IP addresses within HACMP changes and deletes routes. Because AIX caches routes, setting the **routerevalidate** network option is required 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
```

Updating the /etc/hosts File and nameserver Configuration

Make sure all nodes can resolve all cluster addresses. If you are using NIS or DNS, review the section Using HACMP with NIS and DNS in the *Planning Guide*.

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 HACMP service and non-service IP label, make an entry similar to the following:

```
100.100.50.200 payroll-service
```

Also, make sure that the **/etc/hosts** file on each node has the following entry:

```
127.0.0.1 loopback localhost
```

Setting up NIS-Managed Users to Create Crontabs

If your HACMP cluster nodes use Network Information Service (NIS) that includes 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.

To let those NIS-managed users create crontabs, you can do one of the following:

- If it is acceptable to start **cron** after HACMP has started, change the runlevel of the **cron** entry in **/etc/inittab** to **-a**, and make sure it follows the **rcnfs** entry.

  or

- Add an entry to the **/etc/inittab** file that resembles the following sample script with runlevel **-a.** Ensure that it follows the **rcnfs** entry, to terminate the **cron** process properly, 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.

Sample Script

```
#! /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
```

Enabling the AIX Automounter Daemon

For installations that require the AIX automounter daemon on HACMP nodes, a modification is needed to ensure that automounter starts properly (with NFS available and running) on node boot. This is due to the way HACMP manages the **inittab** file and run levels upon startup.

To enable the automounter on nodes that have HACMP installed, add the following line as the last line of the file **/usr/es/sbin/cluster/etc/harc.net**:

```
chnfs -g on -x 1
startsrc -s nfsd
```

Note: The **chnfs** command change is needed only if NFSv4 is being used.

Planning HACMP File Collections

Certain AIX and HACMP configuration files located on each cluster node must be kept in sync (be identical) for HACMP to behave correctly. Such files include event scripts, application scripts, and some system and node configuration files.

Using the HACMP File Collection feature, you can request that a list of files be kept in sync across the cluster automatically. You no longer have to copy an updated file manually to every cluster node, confirm that the file is properly copied, and confirm that each node has the same version of it. With HACMP file collections enabled, HACMP can detect and warn you if one or more files in a collection is deleted or has a zero value on one or more cluster nodes.

Default HACMP File Collections

When you install HACMP, it sets up the following file collections:

- **Configuration_Files**
- **HACMP_Files**.

Contents of the HACMP Configuration_Files Collection

Configuration_Files is a container for the following essential system files:

- **/etc/hosts**
- **/etc/services**
- **/etc/snmpd.conf**
- **/etc/snmpdv3.conf**
- **/etc/rc.net**
- **/etc/inetd.conf**
- **/usr/es/sbin/cluster/netmon.cf**
- **/usr/es/sbin/cluster/etc/clhosts**
- **/usr/es/sbin/cluster/etc/rhosts**
- **/usr/es/sbin/cluster/etc/clinfo.rc** .

Contents of the HACMP_Files Collection

HACMP_Files is a container for user-configurable files in the HACMP configuration. HACMP uses this file collection to reference all of the user-configurable files in the HACMP Configuration Database classes.

The **HACMP_Files** collection references the following Configuration Database fields:

Configuration Database Class	Configuration Database Field	Description
HACMPevent:	notify	Event notify script
HACMPevent:	pre	Pre-event script
HACMPevent:	post	Post-event script
HACMPevent:	recv	Recovery script
HACMPserver:	start	Application server start script
HACMPserver:	stop	Application server stop script
HACMPmonitor:	value, when name=NOTIFY_METHOD	Application monitor notify script
HACMPmonitor:	value, when name=CLEANUP_METHOD	Application monitor cleanup script

HACMPmonitor:	value, when name=RESTART_METHOD	Application monitor restart script
HACMPpager:	filename	remote notification text message file
HACMPsna:	app_svc_file	SNA link start and stop scripts
HACMPx25:	app_svc_file	X.25 link start and stop scripts
HACMPtape:	start_script_name	Tape start script
HACMPtape:	stop_script_name	Tape stop script
HACMPude:	recovery_prog_path	User-defined event recovery program
HACMPcustom:	value	Custom snapshot method script

Note: This collection excludes the **HACMPEvent:cmd** event script. Do *not* modify or rename the HACMP event script files. Also, do *not* include HACMP event scripts in any HACMP file collection.

Notes on Using the Default File Collections

- Neither of these file collections is enabled by default. If you prefer to include some user-configurable files in another collection instead of propagating all of them, leave the **HACMP_Files** collection disabled.
- When copying a file to a remote node, the local node's owner, group, modification time stamp, and permission settings are maintained on the remote node. That is, the remote node inherits these settings from the local node.

  Permissions for all files in the **HACMP_Files** collection is set to *execute*, which helps to prevent problems if you have *not* yet set execute permission for scripts on all nodes. (This is often the cause of an event failure.)
- You cannot rename the **HACMP_Files** collection or delete it. You cannot add files to the collection or remove them. You can add a file that is already included in the **HACMP_Files** collection (for example, an application start script) to another file collection. However, in any other case, a file can only be included in *one* file collection.
- You can add a file that is already included in the **HACMP_Files** collection (for example, an application start script) to another file collection, without receiving the following error message, where XXX _Files is the name of the previously defined collection:

  `This file is already included in the <XXX_Files> collection).`
- You can add and remove files or delete the **Configuration_Files** collection.

Options for Propagating an HACMP File Collection

Propagating a file collection copies the files in a file collection from the current node to the other cluster nodes. Use one of the following methods to propagate an HACMP file collection:

- Propagate the file collection at any time manually. You can propagate files in a file collection from the HACMP File Collection SMIT menu on the local node (the node that has the files you want to propagate).
- Set the option to propagate the file collection whenever cluster verification and synchronization is executed. The node from which verification is run is the propagation node. (This is set to **No** by default.)
- Set the option to propagate the file collection automatically after a change to one of the files in the collection. HACMP checks the file collection status on each node (every 10 minutes by default) and propagates any changes. (This is set to **No** by default.)

  One timer is set for all file collections. You can change the timer. The maximum is 1440 minutes (24 hours) and the minimum is 10 minutes.

You can set up and change file collections on a running cluster. However, note that if you add a node dynamically, the file collection on that node may have files that are *not* in sync with the files on the other cluster nodes. If the file collection on the node being added is set for automatic propagation upon cluster verification and synchronization, the files on the node just added are updated properly. If this flag is *not* set, you must manually run the file collection propagation from one of the other nodes.

For information about configuring file collections, see the *Administration Guide*.

Completing the HACMP File Collection Worksheet

To plan your file collections, complete an HACMP File Collection Worksheet:

1. Enter the **File Collection name** in the appropriate field. The name can include alphabetic and numeric characters and underscores. Use no more than 32 characters. Do *not* use reserved names. For a list of reserved names, see the section List of Reserved Words in the *Administration Guide*.
2. Enter a description for the file collection. Use no more than 100 characters.
3. Specify **Yes** if HACMP should propagate files listed in the current collection before every cluster verification or synchronization. (**No** is the default.)
4. Specify **Yes** if HACMP should propagate files listed in the current collection across the cluster automatically when a change is detected. (**No** is the default.)
5. Enter the list of files to include in this collection. These are the files to be propagated on all cluster nodes.
6. If you plan to use the automatic check option, add the time limit here in minutes (the default is 10 minutes). This time limit applies to all file collections for which you select the automatic check.

Types of Error Notification

This section discusses error notification by AIX and by HACMP.

AIX Error Notification Facility

Although the HACMP software does *not* monitor the status of disk resources, it does provide a SMIT interface to the AIX Error Notification facility. The AIX Error Notification facility allows you to detect an event *not* specifically monitored by the HACMP software—for example a disk adapter failure—and to program a response to the event.

Permanent hardware errors on disk drives, controllers, or adapters may impact the fault resiliency of data. By monitoring these errors through error notification methods, you can assess the impact of a failure on the cluster's ability to provide high availability. A simple implementation of error notification would be to send a mail message to the system administrator to investigate the problem further. A more complex implementation could include logic to analyze the failure and decide whether to continue processing, stop processing, or escalate the failure to a node failure and have the takeover node make the volume group resources available to clients.

Implement an error notification method for *all* errors that affect the disk subsystem. Doing so ensures that degraded fault resiliency does *not* remain undetected.

Note that, if you want HACMP to react to a volume group failure on a node, you have an option to configure a customized AIX Error Notification method for this specific error, which would cause a **node_down** event or move the affected resource group(s) to another node.

You can customize resource recovery for a volume group that fails due to an LVM_SA_QUORCLOSE error. This error occurs if you use mirrored volume groups with quorum enabled. For this case, you can choose to do either of the following:

- Let the HACMP selective fallover function move the affected resource group
- Send a notification using the AIX Error Notification utility
- Continue using your pre- and post-event scripts for this type of recovery.

If you previously had a pre- or post-event configured to handle these cases, assess how they are working with the selective fallover function. For more information on how HACMP handles this particular error, see the section Error Notification Method Used for Volume Group Loss.

However, HACMP does *not* react to any other type of volume group errors automatically. In all other cases, you still need to configure customized error notification methods, or use AIX Automatic Error Notification methods to react to volume group failures.

For information about using this utility to assign error notification methods in one step to a number of selected disk devices, see the section HACMP Automatic Error Notification. For more information about recovery from volume group failures, see the *Troubleshooting Guide*.

HACMP Automatic Error Notification

Before you configure Automatic Error Notification, you must have a valid HACMP configuration.

Using a SMIT option, you can:

- Configure automatic error notification for cluster resources.
- List currently defined automatic error notification entries for the same cluster resources.
- Remove previously configured automatic error notification methods.

You can also use the Automatic Error Notification utility to view currently defined auto error notification entries in your HACMP cluster configuration and to delete all Automatic Error Notification methods.

WARNING: The cluster must be down when configuring automatic error notification. If the cluster is up, a warning is issued and SMIT fails.

If you add error notification methods, the AIX **cl_errnotify** utility runs automatically. This utility 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 file systems
- All disks defined as HACMP resources
- The SP switch network interface card.

To avoid single points of failure, the JFS log must be included in an HACMP volume group.

Automatic error notification applies to selected hard, non-recoverable error types: disk, disk adapter, and SP switch adapter errors. This utility does *not* support media errors, recovered errors, or temporary errors.

Note: You do *not* need to set up automatic error notification for the 2105 IBM Enterprise Storage System (ESS). These systems use the Subsystem Device Driver, which enables the hardware to handle failures itself and automatically switch to another path if it loses one.

For more information, search the IBM website for TotalStorage support, Storage software, Support for Subsystem Device Driver or see the following URL:

http://www.ibm.com/servers/storage/support/software/sdd/

If you do set up automatic error notification it will simply log errors, and *not* initiate fallover action, since the Subsystem Device Driver handles this. However, if all PVIDs are *not* on VPATHS, the error notification fails. Messages are logged to the **cpoc.log** and **smit.log** files.

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 a network interface card (including SP switch NIC) 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 will stop cluster services with the Move Resource Groups option 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**.

The **cl_logerror** script is specified in the notification method instead of the **cl_failover** script for the following system resources:

- Disks that contain unmirrored logical volumes and, therefore, are considered single points of failure
- Disks that are part of volume groups or file systems defined as resources in non-concurrent resource groups.

This prevents unnecessary **node_down** events.

Configuring Automatic Error Notification

To configure automatic error notification:

1. Ensure that the cluster is *not* running.
2. Enter `smit hacmp`
3. In SMIT, select **Problem Determination Tools > HACMP Error Notification > Configure Automatic Error Notification**. The SMIT menu contains the following items:

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, file systems, and disks; **rootvg**. 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.

4. Select the **Add Error Notify Methods for Cluster Resources** option from the list.
5. (*Optional*) Since error notification is automatically configured for all the listed devices on all nodes, make any modifications to individual devices or nodes manually, after running this utility.

If you make any changes to cluster topology or resource configuration, you may need to reconfigure automatic error notification. When you run **verification** *after making any change to the cluster configuration*, you will be reminded to reconfigure error notification if necessary.

Listing Error Notification Methods

To see the automatic error notification methods that exist for your cluster configuration:

1. Enter `smit hacmp`
2. In SMIT, select **Problem Determination Tools > HACMP Error Notification > Configure Automatic Error Notification** and press Enter.
3. 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, file systems, and disks; **rootvg**; SP switch adapter (if present). The following example shows output for cluster nodes 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. Enter `smit hacmp`
2. In SMIT, select **Problem Determination Tools > HACMP Error Notification > Configure Automatic Error Notification** and press Enter.
3. 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 Notification Method Used for Volume Group Loss

If quorum is lost for a volume group that belongs to a resource group on a cluster node, HACMP selectively moves the affected resource group to another cluster node (unless you have customized resource recovery to select notify instead).

For this action, HACMP uses an automatic error notification method to inform the Cluster Manager about the failure of a volume group. The system checks whether the LVM_SA_QUORCLOSE error appeared in the AIX error log file on a cluster node and informs the Cluster Manager to selectively move the affected resource group. HACMP uses this error notification method only for mirrored volume groups with quorum enabled.

You do *not* need to set up automatic error notification for the 2105 IBM Enterprise Storage System (ESS). These systems use the Subsystem Device Driver, which enables the hardware to handle failures itself and automatically switch to another path if it loses one.

If you do set up automatic error notification it will simply log errors and *not* initiate fallover action since the Subsystem Device Driver handles this. However, if all PVIDs are *not* on VPATHS, the error notification fails. Messages are logged to the **cspoc.log** and to the **smit.log**.

Note: *Do not modify* the error notification method used by HACMP to react to a volume group loss. HACMP issues a warning and takes no action if you attempt to customize this notification method or use it to protect against the failure of other types of resources.

The automatically configured AIX Error Notification method is launched if it finds:

- The error LVM_SA_QUORCLOSE in the AIX error log on a cluster node.
- The appropriate entries in the **errnotify** class in the HACMP configuration database on that node. **errnotify** entries are created during synchronization of the cluster resources.

The AIX Error Notification method updates the Cluster Manager. The Cluster Manager then tries to move the resource group that has been affected by a volume group loss to another node in the cluster.

If fallover does *not* occur, check that the LVM_SA_QUORCLOSE error appeared in the AIX error log. When the AIX error log buffer is full, new entries are discarded until space becomes available in the buffer, and, therefore, AIX Error Notification does *not* update the Cluster Manager to selectively move the affected resource group. For information about increasing the size of the error log buffer, see the AIX documentation listed in About This Guide.

Note: You can change the default selective fallover action to be a notify action instead. For more information, see the *Administration Guide*.

Note: If the AIX **errdaemon** is *not* running on a cluster node, HACMP has no means to detect the “loss of quorum” error in the AIX log file, and therefore, cannot selectively move a resource group if it contains a failed volume group. In this case, HACMP issues a warning.

The automatically configured Error Notification method works correctly if the following requirements are met:

- Do *not* modify this error notification method.
- Synchronize the cluster after making changes to the cluster configuration. A notification script used for a volume group failure should correspond to the current configuration of cluster resources. Otherwise, HACMP issues a warning during verification and takes no action to selectively move the affected resource group.

- Besides the **errnotify** entries created by HACMP for selective fallover, the **errnotify** class in the HACMP configuration database may also contain other entries related to the same AIX error labels and resources. However, the selective fallover utility provides the most effective recovery mechanism to protect a resource group from the failure of a single resource.

The notification method that is run in the case of a volume group failure provides the following information in the **hacmp.out** log file:

- AIX error label and ID
- Name of the affected system resource (resource group)
- Node's name on which the error occurred.

You can test the error notification methods generated by the selective fallover facility by emulating an error for each volume group in SMIT.

To test error notification:

1. Enter `smit hacmp`
2. In SMIT, select **Problem Determination Tools > HACMP Error Notification > Emulate Error Log Entry**.
3. Select from the picklist the error notification object that was generated by the selective fallover facility for each volume group.

For more information about how HACMP handles volume group failures, see the section Selective Fallover for Handling Resource Groups in the Appendix B: Resource Group Behavior During Cluster Events in the *Administration Guide*.

Emulation of Error Log Entries

After you have added one or more error notification methods to the AIX Error Notification facility, test your methods by emulating an error. By inserting an error log entry into the AIX error device file **(/dev/error)**, you cause the AIX error daemon to run the appropriate specified notify method. This enables you to determine whether your predefined action is carried through.

To emulate an error log entry:

1. Enter `smit hacmp`
2. In SMIT, select **Problem Determination Tools > HACMP 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. The list includes error notification objects generated by both the Automatic Error Notification facility and by the selective fallover facility for volume group loss. (See the previous section Error Notification Method Used for Volume Group Loss for the description of these methods).
3. Select a notification object and press Enter to begin the emulation.

   As soon as you press Enter, 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.

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 the user specified in the **Resource Name** field during the process of creating an error notify object.

You will now be able to 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.

SP-Specific Considerations

The SP switch has some specific requirements for ARP and network failure notification.

SP Switch Address Resolution Protocol (ARP)

In HACMP 4.5 and up, ARP is enabled for SP Switch networks by default. You can ensure that the network is configured to support gratuitous ARP in HACMP.

To view the setting for gratuitous ARP:

1. Enter `smit cm_config_networks`
2. In SMIT, select **Change a Network Module Using Custom Values.**
3. Select HPS from the picklist, then make sure that the **Supports Gratuitous ARP** setting is set to **true**.

If you are using SP Switch networks in an earlier version of HACMP, manually enable ARP on all SP nodes connected to the SP Switch. If your SP nodes are already installed and the switch network is up on all nodes, you can check whether ARP is enabled. On the control workstation, enter the following command:

```
dsh -av "/usr/lpp/ssp/css/ifconfig css0"
```

If NOARP appears on output from any of the nodes, enable ARP to use IP address takeover on the SP Switch using the following method.

To enable AP to use IP address takeover on an SP switch:

1. Enter `smitty node_data`
2. In SMIT, select **Additional Adapter Information**.
3. On this panel, set the **Enable ARP Settings for the css Adapter** to **yes**, and press Enter.
4. Proceed with customizing the nodes.

SP Switch Global Network Failure Detection and Action

Several options exist for detecting failure and calling user defined scripts to confirm 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 (such as, switch power failure or fault service daemon terminated on all nodes).

  Note: If the local network failure is detected, the Cluster Manager takes selective recovery actions for resource groups containing a service IP label connected to that network. The Cluster Manager tries to move only the resource groups affected by the local network failure event, rather than all resource groups on a particular node.

- Take recovery action, such as moving workload to another network or reconfiguring a backup network.
- To recover from a major switch failure (power off, for example), issue **Eclock** and **Estart** commands to bring the switch back online. The **Eclock** command runs **rc.switch,** which deletes the aliases HACMP needs for SP Switch IP address takeover. Create an event script for either the **network_down** or the **network_down_complete** event to add back the aliases for css0.

Sample SP Switch Notify Method

In the following example of the **Add a Notify Method** panel, you specify an error notification method for an SP Switch.

Notification Object Name	HPS_ER9
Persist across system restart?	yes
Process ID for use by Notify Method	
Select Error Class	All
Select Error Type	PERM
Match Alertable Errors?	None
Select Error Label	HPS_FAULT9_ER
Resource Name	all
Resource Class	all
Resource Type	all
Notify Method	**usr/es/sbin/cluster/utilities/clstop -grsy**

Chapter 9: Creating a Basic HACMP Cluster

This chapter describes how to create a basic two-node cluster by using the Two-Node Cluster Configuration Assistant. This chapter contains the following sections:

- Overview
- Prerequisites
- Planning a Two-Node Cluster
- Using the Two-Node Cluster Configuration Assistant
- Preventing Single Points of Failure
- Where You Go from Here.

Overview

The Two-Node Cluster Configuration Assistant enables you to configure a basic two-node HACMP cluster quickly. Using the Assistant requires very little experience with HACMP. The Assistant guides you through the configuration process of entering data into several entry fields, and then HACMP discovers the remaining configuration information for you.

Note: Do *not* use the Assistant to reconfigure a cluster during HACMP migration.

Note: You can also configure a cluster with a WebSphere, DB2 UDB, or Oracle application. For information, see the corresponding HACMP Smart Assist guide.

The Two-Node Cluster Configuration Assistant is available in two forms:

- A SMIT-based Assistant. For information about this application, see the section Using the SMIT Assistant.
- A standalone Java application. For information about this application, see Using the Standalone Assistant.

HACMP Cluster Definition

The Two-Node Cluster Configuration Assistant creates an HACMP cluster definition with the following characteristics:

- At least one network that connects the two cluster nodes. This network uses IP Address Takeover (IPAT) via Aliases configured as a result of HACMP topology discovery.

  This allows a single NIC to support more than one service IP label. HACMP places a service IP label from the failed node onto a network interface of the takeover node *as an IP alias*, while *keeping* the interface's original IP label and hardware address. IPAT uses the IP aliasing network capabilities of AIX.

  Note: The Assistant configures multiple networks if the subnets to support them exist.

- A non-IP disk heartbeating network over an enhanced concurrent volume group (if available).
- A single resource group that includes the following resources:
  - Two nodes, identified as **localnode** and **remotenode**

    The **localnode**, the node from which you run the Assistant, is assigned the higher priority.
  - A service IP label that you specify in the Assistant
  - An application start script that you specify in the Assistant
  - An application stop script that you specify in the Assistant
  - An application server that you specify in the Assistant

    This application server is associated with the start script and the stop script that you configure for the application.
  - All shareable volume groups discovered by HACMP.

    These volume groups are configured on at least one of the cluster nodes, and both cluster nodes share the disks.

  For more information about resource groups, see the chapter Planning Resource Groups in the *Planning Guide*.

Note: HACMP uses the application server name specified in the Assistant to create the cluster name and the resource group name. For example, if you name the application server db1, the cluster name is db1_cluster, and the resource group name is db1_group.

Resource Group Policies

The Assistant creates an HACMP cluster that has a single resource group. This resource group has the following policies:

- Startup policy: The resource group goes online on the higher priority node.
- Fallover policy: The resource group falls over to the next priority node.
- Fallback policy: The resource group does *not* fall back.

For more information about resource groups, see the chapter Planning Resource Groups in the *Planning Guide*.

Prerequisites

Before using the Two-Node Cluster Configuration Assistant to configure an HACMP cluster definition, make sure that:

- The node running the Assistant has start and stop scripts for the application(s) to be made highly available.
- Both nodes have TCP/IP connectivity to each other.
- Both nodes are physically connected to all disks configured within the volume groups.
- Both nodes have the HACMP software and the same version of the RSCT software.
- Both nodes have a copy of the application that is to be highly available.
- The **etc/hosts** file on both nodes is configured with a service IP label/address to be specified in the Assistant.

TCP/IP Connectivity

To set up a cluster, the nodes require a network connection over TCP/IP.

Copy of the Application

The application should be configured and ready for use on each node.

Ensure that the application licensing requirements are met. Some vendors require a unique license for each processor that runs an application, which means that you must license-protect the application by incorporating processor-specific information into the application when it is installed. As a result, even though the HACMP software processes a node failure correctly, it may be unable to restart the application on the takeover (remote) node because of a restriction on the number of licenses for that application available within the cluster. To avoid this problem, be sure that you have a license for each system unit in the cluster that may run the application.

For additional information about potential licensing issues, see the section on Application Planning in the *Planning Guide*.

Start Scripts and Stop Scripts

When you run the Assistant, an application start script and an application stop script should be present on the node running the Assistant. If the scripts exist on only one node, the Assistant copies the start and stop scripts to the second node.

If you want the scripts to be different on each node, create scripts on each node.

Volume Groups

Ensure that both nodes are physically connected to all disks that are configured within the volume groups on the node running the Assistant. These disks should appear in the AIX configuration on each node.

At least one of the shared volume groups should be an enhanced concurrent volume group. This way, HACMP can set up a non-IP network over the disk for heartbeating traffic.

Note: The Assistant imports volume group definitions of any shareable volume group to both nodes.

If the cluster does *not* have an enhanced concurrent mode volume group, establish a separate disk heartbeating network outside of the Assistant. After you finish running the Assistant, it displays a message if a disk heartbeating network could *not* be created. For information about heartbeating networks, see the chapter on Planning Cluster Network Connectivity in the *Planning Guide*.

Service IP Label/Address

Ensure that an **/etc/hosts** file on both of the nodes has a service IP label that is required by the application server configured. The service IP label is the IP label over which services are provided. HACMP keeps the service IP label available. Clients use the service IP label to access application programs, and HACMP uses it for disk heartbeating traffic. The service IP label should be on a different subnet than the IP label used at boot time on the interface.

The **/etc/hosts** file on one of the nodes must contain all IP labels and associated IP addresses. The Assistant populates the **/etc/hosts** file on the second node if the file does *not* list the IP labels and associated IP addresses.

Note: The cluster configured by the Assistant uses AIX support for multiple IP aliases on a NIC.

HACMP Software

Ensure that the HACMP software is installed on the two nodes to be included in the cluster.

For information about installing HACMP, see Chapter 4: Installing HACMP on Server Nodes.

Note: Running the Two-Node Cluster Configuration Assistant standalone application requires Java version 1.3 or higher. AIX v.5.2 and up includes Java version 1.3 or higher.

Planning a Two-Node Cluster

Before you start the Two-Node Cluster Configuration Assistant, make sure you have the following information available. You can use the Two-Node Cluster Configuration Worksheet as described in the *Planning Guide* to record this information.

Local Node	The node on which the application that is to be highly available typically runs. You run the Two-Node Cluster Configuration Assistant from this node to set up the cluster definition.

Takeover (Remote) Node	The node on which the application will run should the local node be unable to run the application.
Communication Path to Takeover Node	A resolvable IP label (this may be the hostname), IP address, or fully qualified domain name on the takeover node. This path is used to initiate communication between the local node and the takeover node. Examples of communication paths are NodeA, 10.11.12.13, and NodeC.ibm.com. In the SMIT Assistant, use the picklist display of the hostnames and addresses in the **/etc/hosts** file that are *not* already configured for HACMP. The Assistant uses this path for IP network discovery and automatic configuration of the HACMP topology.
Application Server	A label for the application that is to be made highly available. This label is associated with the names of the start script and the stop script for the application. The server name can include alphabetic and numeric characters and underscores. The application server name may contain no more than 24 characters. (Note that the number of characters is fewer than the number allowed in other SMIT fields to define an application. This is because the application server name in the Assistant is used as part of the name of the cluster and the resource group.)
Application Start Script	The name of the script that is called by the cluster event scripts to start the application server. The script name includes the name of the script, its full pathname, and is followed by any arguments. This script must be in the same location on each cluster node that may start the server. However, the contents of the script can differ. The application start script name may contain no more than 256 characters.
Application Stop Script	The full pathname of the script (followed by arguments) that is called by the cluster event scripts to stop the server. This script must be in the same location on each cluster node that may stop the server. However, the contents of the script can differ. The application stop script name may contain no more than 256 characters.
Service IP Label	The IP label/IP address to be kept highly available. In the SMIT Assistant, use the picklist display IP labels/addresses.

Using the Two-Node Cluster Configuration Assistant

Run the Assistant from the node that is to initially run the application server to be made highly available. Both the Assistant standalone application and the SMIT-based Assistant provide detailed information to help you enter the correct information for the Assistant. The Assistant standalone application also provides tool tips for command buttons, panel controls, and data entry fields. As you complete panels in the Assistant, onscreen messages describe the progress of your configuration.

Note: If the Assistant encounters errors during the configuration process, it does *not* save the configuration.

If the Assistant encounters errors during synchronization and verification, it saves the cluster configuration even though it may be invalid. Typically, you can correct most verification errors within AIX.

Before you use the Assistant, to make sure that you have the applicable information available, consider completing a Two-Node Cluster Configuration Worksheet, as described in the *Planning Guide*.

User Privileges

Only users with root privilege can use the Assistant. Although a user who does *not* have root privilege can start the Assistant, that user cannot use the application to make configuration changes.

Existing Clusters

If there is an existing HACMP cluster on the node on which you are running the Assistant, the Assistant removes the cluster definition and saves a snapshot of the cluster definition that was removed. The snapshot file created by the Assistant uses the following naming convention:

```
clustername_YYYYMMDDhhmm
```

For example, a snapshot named **db2_cluster_200512011254** indicates that the snapshot of the configuration for the cluster db2_cluster was saved at 12:54 P.M. on December 1, 2005.

The Assistant creates a snapshot only for valid HACMP configurations. Information about this activity is saved to the log file for the Configuration Assistant. If you are using the SMIT Assistant, a status message appears after you complete the entry fields to show that a snapshot was created and the full pathname of that file.

The system running the Assistant requires sufficient disk space to save a snapshot. If you are using the SMIT Assistant, it prompts you for an alternate location for the file if there is insufficient disk space.

For information about the log file for the Configuration Assistant, see the section Logging for the Two-Node Cluster Configuration Assistant.

For information about using a cluster snapshot, see the *Administration Guide*.

Using the Standalone Assistant

If you completed a Two-Node Cluster Configuration Worksheet as part of the planning process, refer to that information when entering information in the Assistant.

To use the Two-Node Cluster Configuration Assistant:

1. Enter the **cl_configassist** command to start the Two-Node Cluster Configuration Assistant.

   The **cl_configassist** command resides in the **/usr/es/sbin/cluster/utilities** directory. For more information about this command, see its man page.
2. Read the overview information on the HACMP Two-Node Cluster Configuration Assistant panel. Make sure that you have the information listed.
3. Click **Start**.
4. Enter values in the following panels and click **Next** on each panel when finished:
   - Step 1 of 4: Topology Configuration panel
   - Step 2 of 4: Application Server Configuration panel
   - Step 3 of 4: Resource Configuration panel

   For information about the values to enter, see the section Planning a Two-Node Cluster.
5. On the Step 4 of 4: Verification and Synchronization panel, review the verification and synchronization messages.

   The verification and synchronization for the cluster in this step is the same as the standard verification and synchronization for an HACMP cluster.
6. Click **Finished** if the verification and synchronization completes successfully.

   or

   Click **Exit**, if there are verification or synchronization errors. You can resolve any errors, then select the **Back** button to correct entries or run the Assistant again.

   Note: The **Finished** button is available if verification and synchronization complete successfully. The **Exit** button is available if there are verification and synchronization errors.

Using the SMIT Assistant

If you completed a Two-Node Cluster Configuration Worksheet, as described in the *Planning Guide*, refer to that information when entering information in the Assistant.

To use the SMIT-based Assistant:

1. Enter `smit hacmp` and select **Initialization and Standard Configuration > Configuration Assistants > Two-Node Configuration Assistant**.

   or

   Enter the fastpath `smitty cl_configassist`

   The **Two-Node Configuration Assistant** panel appears.
2. Enter values for the following fields and Press **Enter**:
   - **Communication Path to Takeover Node**
   - **Application Server**

- **Application Start Script**
- **Application Stop Script**
- **Service IP Label**

For information about the values to enter, see the section Planning a Two-Node Cluster.

SMIT provides configuration status as each command is run. At command completion, SMIT automatically runs verification and synchronization. SMIT displays information about the verification and synchronization process.

If verification and synchronization does *not* finish successfully, review the messages to determine the problem. Remedy any issues and then run the Assistant again.

Logging for the Two-Node Cluster Configuration Assistant

The Two-Node Cluster Configuration Assistant logs debugging information to the **cl_configassist.log** file. This log file is stored in the **/var/hacmp/utilities/** directory by default and can be redirected to another location in the same way that you redirect other HACMP log files. For information about redirecting log files, see the *Troubleshooting Guide*.

The Assistant stores up to 10 copies of the log file to assist with troubleshooting activities. The log files are numbered to differentiate one from the other, for example **cl_configassist.log**, **cl_configassist.log.1**, and **cl_configassist.log.2**.

The **clverify.log** and **smit.log** files provide additional information for troubleshooting issues.

Preventing Single Points of Failure

A robust HACMP cluster requires configuring cluster components to prevent one from becoming a single point of failure. Review the following components to ensure that one does *not* become a single point of failure in your cluster:

- Shared disks

  For information about preventing a disk from being a single point of failure, see the chapter Planning Shared Disk and Tape Devices in the *Planning Guide*.

- Shared Logical Volume Manager components

  For information about preventing a volume group from being a single point of failure, see the chapter Planning Shared LVM Components in the *Planning Guide*.

- Networks

  If possible, the Assistant creates a non-IP disk heartbeating network to ensure connectivity between cluster nodes. If the Assistant cannot create a disk heartbeating network, it displays a message to this effect. In this case, set up a non-IP network to transmit disk heartbeating messages.

  For information about preventing the network from being a single point of failure, see the chapter Planning Cluster Network Connectivity in the *Planning Guide*.

Where You Go from Here

After you have a basic cluster configuration in place, set up cluster security. For information about configuring cluster security, see the *Administration Guide*.

You can also customize your cluster by:

- Configuring cluster events

  See Chapter 8: Configuring AIX for HACMP.
- Setting up cluster monitoring

  See Chapter 10: Monitoring an HACMP Cluster in the *Administration Guide*.

Appendix A: Cluster Monitoring with Tivoli

This appendix contains instructions for making an HACMP cluster known to Tivoli in order to monitor and administer the cluster through the Tivoli management console. This process is described in the following sections:

- Overview
- Before You Start
- Steps for Installing and Configuring Cluster Monitoring with Tivoli
- Removing Cluster Monitoring with Tivoli
- Where You Go from Here.

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
- Resource group location and state
- Individual resource location (*not* state).

In addition, you can perform the following *cluster administration* tasks from within Tivoli:

- Start cluster services on specified nodes.
- Stop cluster services on specified nodes.
- Bring a resource group online.
- Bring a resource group offline.
- Move a resource group to another node.

Setting up this monitoring requires a number of installation and configuration steps to make Tivoli aware of the HACMP cluster and to ensure proper monitoring of IP address takeover.

After you complete the installation, you can use Tivoli to monitor and administer your cluster as described in the section Monitoring Clusters with Tivoli Distributed Monitoring in Chapter 10: Monitoring an HACMP Cluster in the *Administration Guide.*

Before You Start

Before installing and configuring cluster monitoring with Tivoli, make sure you have fulfilled the prerequisite conditions.

Prerequisites and Considerations

When planning and configuring Cluster Monitoring with Tivoli, keep the following points in mind:

- The Tivoli administrator must be set to **root**@*persistent_label* for each node in the cluster before you start installing HATivoli.
- 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.

  For more information, see the section Step 1: Installing Required Tivoli Software.
- For proper monitoring of IP address takeover activity, 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 is monitored properly. You must use a persistent node IP label on each node and make sure it is on the same subnet as the TMR.

  For more information about these requirements, see the sections Using Persistent Node IP Labels with HATivoli and Subnet Considerations for Cluster Monitoring with Tivoli.
- The verification utility does *not* check for the following conditions (you must do so manually):
  - Whether every node is installed with the Tivoli cluster monitoring software
  - Whether the **oserv** process is running on all of your nodes.

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 for Tivoli.

Using Persistent Node IP Labels with HATivoli

In many cases, it is *not* possible to install separate physical adapters dedicated to communication with the non-cluster Tivoli Management Region node. Without a dedicated Tivoli network, proper monitoring of IPAT through Tivoli requires that each node in the cluster have a node-bound IP address defined that does *not* belong to any resource group and does *not* move between nodes.

You need to assign persistent node IP labels to a service or non-service interface on each node in HACMP, and configure Tivoli to use these persistent node IP labels. When the Tivoli **oserv** process starts, Tivoli uses the persistent IP node labels already configured in HACMP.

For information about how to configure persistent node IP labels in HACMP, see the section Configuring HACMP Persistent Node IP Labels/Addresses in Chapter 4: Configuring HACMP Cluster Topology and Resources (Extended) in the *Administration Guide*.

Note: If you already have Tivoli set up, you may need to change the TMR's IP address to match the subnet of the monitored nodes' persistent node IP labels.

Subnet Considerations for Cluster Monitoring with Tivoli

For proper monitoring of IPAT, make sure the service address of the TMR node is on the same subnet as the persistent IP alias assigned to the monitored nodes. Depending on whether you are using IPAT via IP Replacement or IPAT via Aliases, other subnet considerations for the persistent node IP label vary as described in the following sections.

Subnet Requirements for Non-Aliased Networks

For non-aliased networks, (networks in the cluster that use standard IPAT), the following requirements for subnets apply:

- The subnet of a node's persistent node IP label must be *different* from the subnet of the node's *non-service* labels.
- The subnet of a node's persistent node IP label may be the *same* as the subnet of the node's *service* labels.
- For cluster monitoring through Tivoli, the subnet of the monitored node's persistent node IP label must be the *same* as the subnet of the non-cluster TMR node.

Subnet Requirements for Aliased Networks

For aliased networks (networks in the cluster that use IPAT via IP Aliases), the following requirements for subnets apply:

- The subnet of the persistent node IP label must be *different* from the subnet of the node's non-service IP labels.
- For cluster monitoring through Tivoli, the subnet of the monitored node's persistent node IP label must be the *same* as the subnet of the non-cluster TMR node.

For details on IPAT via IP Aliases, see the section IP Address Takeover via IP Aliases in the chapter on Planning Cluster Network Connectivity in the *Planning Guide*.

Steps for Installing and Configuring Cluster Monitoring with Tivoli

Preparing to monitor a cluster with Tivoli involves several stages and prerequisite tasks.

The following table 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 for monitoring the HACMP cluster and, therefore, need to take steps to ensure proper IPAT monitoring.

The following sections provide further details about each of the installation steps.

Step 1: Installing Required Tivoli Software

The following Tivoli software must be installed and running on the TMR node and on cluster nodes before installing the Tivoli-related HACMP filesets. See the IBM website for information on the latest supported versions:

- Tivoli Managed Framework (on TMR and cluster nodes)
- Tivoli Application Extension Facility (AEF) (on TMR only)
- Tivoli Distributed Monitoring (on TMR and cluster nodes)
- Unix Monitors
- Universal Monitors.

Note: If you are doing a fresh installation of Tivoli, see the steps related to the node IP aliases. You may want to ensure now that the IP address of the Tivoli TMR is on the same subnet as the node IP aliases.

Step 2: Creating a Cluster Policy Region and Profile Manager

On the TMR, create a Policy Region and Profile Manager for HACMP monitoring to handle the HACMP cluster information.

Consult your Tivoli documentation or online help if you need instructions for performing these Tivoli tasks.

Step 3: Defining HACMP Cluster Nodes as Tivoli Managed Nodes

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, 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 before adding HACMP nodes, and want to monitor IP address takeover, change the IP address of the TMR node to match the subnet of the persistent node IP address aliases you assigned to the cluster nodes. For details, see the section Subnet Considerations for Cluster Monitoring with Tivoli.

Follow the procedure you would follow to install any nodes for Tivoli to manage. Refer to Tivoli documentation and online help for instructions.

Step 4: Defining Administrators

Define the cluster nodes as Login Names in the Administrators panel. Consult your Tivoli documentation or online help if you need instructions for performing Tivoli tasks.

Step 5: 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.

Step 6: Adding Nodes as Subscribers to the Profile Manager

Subscribe the cluster nodes to the Tivoli 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.

Step 7: Installing the HACMP Cluster Monitoring (hativoli) Filesets

If you have *not* done so already, install the HACMP software, and install the three **cluster.hativoli** filesets on the both the Tivoli server node and the HACMP cluster nodes. The **cluster.hativoli** filesets that you can select during the SMIT installation are:**cluster.hativoli.client**, **cluster.hativoli.server**, and **cluster.msg.en_US.hativoli**.

Step 8: Configuring the HACMP Cluster

If you have *not* done so already, configure the HACMP cluster and synchronize. Make sure to configure a persistent node IP alias on each node in the cluster.

Step 9: Placing the Node IP Alias in the /etc/hosts and /etc/wlocalhosts Files

Make sure each node has a persistent IP alias with a subnet that matches that of the TMR adapter. Make sure this alias is included in the **/etc/hosts** file, the **ipaliases.conf** file (see next section), and the Tivoli **/etc/wlocalhost** files. Note that if the **/etc/wlocalhost** file was *not* created earlier in Tivoli, create it now.

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.

Step 10: Creating the ipaliases.conf File

To monitor IPAT without a dedicated network, create a file called **/usr/es/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 aliases, and the name of each cluster node with its alias label. For example:

```
network=token21
node1 node1-alias
node2 node2-alias
node3 node3-alias
```

Step 11: 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. Set up a way to monitor the state of the **oserv** process. For information about defining an HACMP application monitor, see Chapter 4: Configuring HACMP Cluster Topology and Resources (Extended) in the *Administration Guide*.

To start **oserv**, run the following command on each node:

```
/etc/Tivoli/oserv.rc start
```

Upon reintegration of a node, HACMP automatically sets the alias to an adapter (if applicable) and then starts the **oserv** process. For HACMP to do this, the failed node must be part of a non-concurrent resource group.

Step 12: Saving Previous 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 so you should make sure they are saved.

HACMP automatically saves a copy of your parent dialog. If you need to restore earlier customizations, find the saved file in **/usr/es/sbin/hativoli/ParentDialog.dsl.save**.

Step 13: Running Additional 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 synchronize cluster resources after step 1.

1. Run **/usr/es/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 may be a delay of up to several minutes while the system creates and distributes profiles and indicators.

2. Run /**usr/es/sbin/hativoli/AEF/install** *on the TMR node.*

3. Run **/usr/es/sbin/hativoli/AEF/install_aef_client** *on ALL cluster nodes.*

Step 14: Starting Cluster Services

Start cluster services on each cluster node.

Step 15: 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.

Removing Cluster Monitoring with Tivoli

To discontinue cluster monitoring with Tivoli, perform the following steps to delete the HACMP-specific information from Tivoli:

1. Run a uninstall through the SMIT interface to remove the three **hativoli** filesets on all cluster nodes and the TMR.

2. If it is *not* already running, call Tivoli on the TMR:
   - Enter **. /etc/Tivoli/setup_env.sh**
   - Enter **tivoli**

3. From the Policy Region for the cluster, open Modify HATivoli Properties task library.

4. A window appears containing task icons.

5. Select **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 HATivoli Properties task icon.
7. Repeat steps 3 through 6 for the Cluster Services task library.
8. Open the Profile Manager.
9. Select **Edit > Profiles > Select All** to select all HACMP Indicators.
10. Select **Edit > Profiles > Delete** to delete the Indicators.
11. Unsubscribe the cluster nodes from the Profile Manager as follows:
    - In the Profile Manager window, select Subscribers.
    - Highlight each HACMP node on the left, and click to move it to the right side.
    - 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.

For information about monitoring and administering an HACMP cluster through the Tivoli management console, see the chapter on Monitoring an HACMP Cluster, in the *Administration Guide*.

Appendix B: OEM Disk, Volume Group, and File Systems Accommodation

This appendix describes how you can customize HACMP software to integrate Original Equipment Manufacturer (OEM) disks, volume groups, and file systems in an HACMP cluster. The appendix contains these sections:

- Integrating OEM Disks in an HACMP Cluster
- Integrating OEM Volume Groups in an HACMP Cluster
- Integrating OEM File Systems in an HACMP Cluster.

Integrating OEM Disks in an HACMP Cluster

HACMP lets you use either storage disks manufactured by IBM or OEM physical storage disks as part of a highly available infrastructure, as long as the disks are defined in AIX and are part of AIX LVM volume groups.

Depending on the type of OEM disk, custom methods allow you (or an OEM disk vendor) either to tell HACMP that an unknown disk should be treated the same way as a known and supported disk type, or to specify the custom methods that provide the low-level disk processing functions supported by HACMP for that particular disk type.

This section contains the following topics:

- Overview
- Handling Disks in HACMP:Serial Access to Volume Groups and Concurrent Access to Volume Groups
- Disks Supported by HACMP
- Supporting Disks with Known Characteristics
- Customizing HACMP Disk Processing
- Configuring OEM Disk Methods in SMIT.

Overview

Custom methods for OEM disks provide a means for configuring and supporting OEM disks in an HACMP cluster. These methods depend on the types of disks that are supported by the HACMP software, and on the different ways they are handled. A level of similarity between certain types of disks allows you to configure an OEM disk that is unknown to HACMP as you would a similar type of disk known to HACMP. Alternatively, you can select one or more of the disk processing methods used by HACMP to safely deal with an OEM disk.

The disk processing methods include:

- Identifying *ghost* disks

  Ghost disks are duplicate profiles of the disks created during disk configuration processing. Ghost disks must be removed to allow AIX to vary on the volume group residing on the original disks.
- Determining whether another node in the cluster is holding a disk reserve
- Breaking a disk reserve
- Making a disk available for use by another node.

Handling Disks in HACMP

HACMP handles disks according to AIX LVM disk requirements. The AIX Logical Volume Manager (LVM) by default is *not* designed to handle multiple nodes accessing the same set of disks, which are referred to as *shared*. Therefore, HACMP performs special disk processing for shared disks in a cluster. When nodes in the HACMP cluster join and leave the cluster, HACMP brings the shared disks into a state where the LVM commands work correctly.

In the LVM, one or more disks, also known as physical volumes, can be grouped together to form a volume group. Logical volumes can be built on top of a volume group and file systems can be built on top of logical volumes. See AIX LVM documentation for more information.

HACMP supports two modes of access to volume groups in a cluster: serial access mode and concurrent access mode.

Serial Access to Volume Groups

In serial access mode, the volume group is active on one node at a time. From the LVM's point of view, this is an ordinary volume group. Since the LVM does *not* handle multiple node access to the disks on which the volume group resides, when such a volume group is active on one of the nodes, the LVM instructs the device driver layer on that node to establish a reserve on the disks. A disk reserve is a hardware-level command that tells the disk to accept commands only from the SCSI adapter that sent the reserve request. This restriction prevents other cluster nodes from inadvertently modifying the volume group.

The reserve function is absolutely critical for disks that are used to hold Journaled File System(s) (JFS). JFS function by aggressively caching file data and file system-specific information in real memory. JFS have no mechanism by which one node can inform another node that its cache is no longer valid. Therefore, if multiple nodes access a disk holding a JFS, and the file system is mounted and being accessed on those nodes, the result would be invalid data being returned on read requests, a node failure, and corruption of the JFS.

A disk reserve remains active on the node even if the node that established the disk reserve fails. For serial access volume groups, HACMP must ensure that, in the event of a node failure, a takeover node can access the disks. To do this, the takeover node must break the disk reserve established by the failed node. To break the disk reserve, the takeover node uses special low-level commands, typically specific to the attachment mechanism and the disk subsystem.

Concurrent Access to Volume Groups

In concurrent access, the volume groups residing on shared disks are activated and can be simultaneously accessed in read/write mode by all active nodes in the cluster.

Note: AIX introduced *enhanced concurrent mode*. When concurrent volume groups are created on AIX, they are automatically created as enhanced concurrent mode, except when SSA disks are used. Convert your SSA and RAID concurrent volume groups to enhanced concurrent mode volume groups. For details, see the chapter Planning Shared LVM Components in the *Planning Guide*.

Accessing disks in concurrent mode requires that software packages running above the LVM coordinate their operations across the cluster to ensure data integrity. The AIX Journaled File System is *not* supported on volume groups accessed in concurrent mode. Consult the provider of any middleware to determine whether that middleware can reliably function in concurrent mode.

For concurrent access volume groups, HACMP must ensure that no disk reserve is placed on the disk, so that access by other nodes is possible.

The question arises as to whether any disk can be used in concurrent mode. The answer is no, because of a practical requirement to mirror data in an HACMP cluster, so that data is *not* be lost if any of the disks fails. Even though most disks could be accessed without establishing a disk reserve on them, many disks still cannot be used in concurrent mode, because the LVM cannot simultaneously update all the copies of a single logical partition. Even when the writes on a disk are done in parallel, they may *not* actually be completed at the same time, due to other activity on the disk, or the physical characteristics of the disks. If one node writes data and the other node reads it, there is no guarantee that the reader gets the latest copy. Additionally, if the LVM cannot update its structures on a disk, it ceases to write to that disk and marks it as failed. In a multi-node cluster, this situation creates a single point of failure.

The following section describes how these practical problems with concurrent access volume groups are handled in the HACMP software.

Disks Supported by HACMP

HACMP supports four kinds of disks:

- Enhanced Concurrent Mode Disks
- True Concurrent Mode Disks
- RAID Concurrent Mode Disks
- Serial Access Mode Disks.

Enhanced Concurrent Mode Disks

AIX v.5.2 and up support *enhanced concurrent mode*. In enhanced concurrent mode, instances of the Concurrent Logical Volume Manager (CLVM) coordinate changes between nodes through the Group Services component of the Reliable Scalable Cluster Technology (RSCT) facility in AIX. Group Services protocols flow over the communications links between the cluster nodes.

With enhanced concurrent mode:

- Any disk supported by HACMP for attachment to multiple nodes can be included in an enhanced concurrent mode volume group.
- The same capabilities for online changes of volume group and logical volume structure that have always been supported by AIX for SSA concurrent mode volume groups are available for enhanced concurrent mode volume groups.
- Enhanced concurrent mode volume groups require the Concurrent Resource Manager. This provides the cluster initialization that is needed for CLVM to provide enhanced concurrent mode support.

True Concurrent Mode Disks

True concurrent mode disks are managed across the cluster by the Concurrent Logical Volume Manager (CLVM), an extension to the AIX LVM that is enabled when HACMP is installed.

HACMP software supports one type of true concurrent mode IBM disks: SSA disks. This type of disks has the covert channel facility, which CLVM uses to allow multiple instances of the CLVM on the nodes across the cluster. This makes it possible to keep the local LVMs and the disk contents synchronized. These disks also provide a "fencing" function, which is the multi-node equivalent of disk reserves: It allows the specified subset of cluster nodes to access the shared disks while blocking out other nodes. For this type of disks, HACMP automatically fences out failed nodes to prevent inadvertent modification of the disks.

For true concurrent mode disks, the CLVM can keep the individual LVMs on the cluster nodes synchronized. The CLVM also allows you to make changes to the volume group configuration (such as adding or extending a logical volume, or adding or deleting a disk on a node), and propagate the changes to all other cluster nodes. The CLVM uses the covert channel facility to coordinate updates across the cluster, and to inform other nodes that they must reread the LVM structures on the disk to pick up the latest information. This ensures that the nodes in the cluster will *not* pick up stale data by reading an obsolete disk mirror.

RAID Concurrent Mode Disks

RAID concurrent mode disks, like true concurrent mode disks, can be accessed in read/write mode by multiple cluster nodes, and the shared volume groups residing on these disks can be varied on multiple nodes in a cluster.

The RAID concurrent mode disks do *not* have a covert channel. The CLVM provides no direct management or coordination of LVM updates across nodes in a cluster.

The following two functions allow support for RAID concurrent mode disks in an HACMP cluster:

- Static configuration. A RAID disk is an abstraction of a number of physical disks, with data and parity information scattered across the physical disks in a manner defined by RAID implementation. Since the RAID controller manages the physical disks, the LVM perceives a RAID disk device as a certain number of hdisks that is smaller than the number of present physical spindles. The RAID controller handles most data redundancy concerns; the LVM's capabilities for data mirroring and online modification of a volume group are *not* used.

- Lack of mirroring. As mentioned previously, the RAID controller handles data redundancy; LVM does no mirroring. As a result, it is *not* possible for a node to read the wrong mirror copy.

Note: It is possible to mirror a RAID disk, although this is generally considered unnecessary and is *not* supported when the disk is used in a concurrent access volume group.

Note: RAID concurrent mode disks provide no communication mechanisms to keep the individual LVMs on the cluster nodes synchronized. The lack of a covert channel in RAID concurrent mode disks prevents dynamic modifications to the volume group structure (for example, expanding a logical volume).

For information about RAID disks, see the following URL:

http://www.acnc.com/raid.html

Serial Access Mode Disks

Only one node at a time can access serially accessed disks in read/write mode, and the shared volume groups residing on these disks can be varied on only one node in a cluster. The LVM perceives these disks as normal disks. The actual `hdisks` as perceived by the LVM can be either physical spindles (a configuration sometimes referred to as "JBOD"—Just a Bunch of Disks—to contrast it with RAID), or RAID disks. For JBOD configurations, configure the LVM to provide mirroring for data reliability.

The following shell script contains most of the processing for HACMP disk takeover: **/usr/sbin/cluster/events/utils/cl_disk_available**.

Supporting Disks with Known Characteristics

All the disks supported by HACMP adhere to some form of the SCSI standard. However, this standard has undergone considerable evolution over time and allows for many implementation choices. Even so, many disks behave in similar ways. HACMP provides mechanisms that will allow you, while configuring a cluster, to direct HACMP to treat an unknown disk exactly the same way as another disk it supports. You do this by making simple additions to ASCII files using any available text editor. In particular, changes to the following files will be necessary:

/etc/cluster/conraid.dat

Use this file to tell the CRM HACMP that a particular disk is a RAID disk that can be used in concurrent mode. The file contains a list of disk types, one disk type per line. The value of the Disk Type field for a particular `hdisk` is returned by the following command:

```
lsdev -Cc disk -l <hdisk name> -F type
```

This command returns a response similar to the following:

```
scsd
```

This file is referenced by **/usr/sbin/cluster/events/utils/cl_raid_vg**.

/etc/cluster/lunreset.lst

Use this file to tell HACMP that a particular disk supports LUN resets, even though its response to the SCSI inquiry command indicates that the disk does *not* support the SCSI command set. The file contains the **Vendor Identification** field. A SCSI inquiry command returns the value of this field. Consult the disk manufacturer for the value of this field.

/etc/cluster/disktype.lst

Use this file to tell HACMP that it can process a particular disk the same way it processes disks that it supports. The file contains a series of lines of the following form:

```
<PdDvLn field of the hdisk><tab><supported disk type>
```

To determine the value of the **PdDvLn** field for a particular hdisk, enter the following command:

```
lsdev -Cc disk -l <hdisk name> -F PdDvLn
```

This command returns a response similar to the following:

```
disk/scsi/scsd
```

The supported disk types are:

Disk Name in HACMP	Disk Type
SCSIDISK	SCSI -2 Disk
SSA	IBM Serial Storage Architecture
FCPARRAY	Fibre Attached Disk Array
ARRAY	SCSI Disk Array
FSCSI	Fibre Attached SCSI Disk

This file is referenced by **/usr/sbin/cluster/events/utils/cl_disk_available.**

HACMP does *not* modify the previously described files after they have been configured and are *not* removed if the product is uninstalled. This ensures that customized modifications are unaffected by the changes in HACMP. By default, the files initially contain comments explaining their format and usage.

Keep in mind that the entries in these files are classified by disk type, *not* by the number of disks of the same type. If there are several disks of the same type attached to a cluster, there should be only one file entry for that disk type. In addition, unlike other configuration information, HACMP does *not* automatically propagate these files across nodes in a cluster. It is your responsibility to ensure that these files contain the appropriate content on all cluster nodes. Use the HACMP File Collections facility to propagate this information to all cluster nodes. For more information on the HACMP File Collections facility, see *Administration Guide.*

Customizing HACMP Disk Processing

Some disks may behave sufficiently differently from those supported by HACMP so that it is *not* possible to achieve proper results by telling HACMP to process these disks exactly the same way as supported disk types. For these cases, HACMP provides finer control. While doing cluster configuration, you can either select or specify one of the specific methods to be used for the steps in disk processing.

HACMP supports the following disk processing methods:

- Identifying ghost disks
- Determining whether a disk reserve is being held by another node in the cluster
- Breaking a disk reserve
- Making a disk available for use by another node.

HACMP allows you to specify any of its own methods for each step in disk processing, or to use a customized method, which you define. This allows you to use a custom defined method only for configuring a disk, and to use the existing HACMP methods for other disk operations.

The following methods are supported by HACMP for each of the steps in disk processing:

Identifying Ghost Disks

Ghost disks are duplicate profiles of the disks created during disk configuration processing. You must remove ghost disks in order to allow AIX to vary on the volume group residing on the original disks.

A ghost disk is created when the disk configuration method is run while a disk is being reserved by another node. In this case, the configuration method cannot read the **PVID** of the reserved disk to identify it as an existing disk. This can easily happen when:

- A node fails
- HACMP takes over the failed disks. During the takeover, the other node breaks any outstanding disk reserve on the failed node and establishes its own reserve on the disks.
- The failed node reboots. The disk configuration method run at boot time cannot read the disks due to the disk reserve held by another node.

Ghost disks are a problem for two reasons:

- The presence of superfluous disk profiles is confusing to the system
- If the ghost disk is in an available state and the **PVID** is returned by **lsdev**, then any attempt to make the real disk available will fail, since the two disks refer to the same device. This prevents use of the real disk for operations such as varyon.

Identifying a ghost disk is supported in HACMP in the following way:

1. The method is passed as an input hdisk name, such as *hdisk42*.
2. The method writes to standard output a list of disk names that are ghost disks for the given disk.
3. If there are no ghost disks for the given disk, nothing is returned.

A return code of 0 indicates success; any other value will cause processing to be terminated for the given disk.

HACMP supports values of *SCSI2* and *SCSI3* for the name of this method. If either of those values is specified, HACMP uses the existing disk processing that is being used to identify ghost disks for IBM SCSI-2 or SCSI-3 disks:

- *SCSI2* ghost disks are those with different names from the given disk, but identical parent and location attributes in **CuDv**.
- *SCSI3* ghost disks are those with different names from the given disk, but identical parent and location attributes in **CuDv**; and identical **lun_id**, **scsi_id** and **ww_name** attributes in **CuAt**.

Determining Whether a Disk Reserve Is Held by Another Node

A disk reserve restricts access to the disk by a specific node. If the node that established the reserve has failed, HACMP takes special steps to remove that disk reserve.

This method is passed as an input hdisk name, such as *hdisk42*. It passes back a return code with one of the following values:

- 0—The disk is *not* reserved, and it is readable and writable.
- 1—A disk reserve is currently held by another node.
- 2—An error has occurred. HACMP should terminate processing for this disk.

Any other value will be interpreted as equivalent to 2.

HACMP supports a value of *SCSI_TUR* for the name of this method. If this value is specified, existing processing in HACMP is used to determine whether a disk reserve is being held. The processing opens the disk device and sends a SCSI Test Unit Ready command via `ioctl()`. A response of `reservation conflict` means that there is a disk reserve held by another node.

Breaking a Disk Reserve

Disk reserves can be broken in parallel. Any serialization requirements must be handled by user-supplied methods.

This method is passed as input two formal parameters:

- The name of the owning adapter for the disk, such as *scsi1*. This is determined by HACMP by running the command:

  ```
  lsdev -Cc disk -l <hdisk name> -F parent
  ```
- The hdisk name, such as *hdisk42*.

A return code of 0 (zero) indicates success; any other value causes HACMP to terminate processing for this disk.

HACMP supports values of *TARGET, PSCSI,* and *FSCSI* for the name of this method. If one of these values is specified, HACMP uses existing processing for this method. The processing breaks the disk reserves for these types of disks:

- *TARGET*. A SCSI target ID reset will be sent by executing

  ```
  openx(<hdisk name>, ,SC_FORCED_OPEN)
  ```

  Specifying the *TARGET* value is appropriate for SCSI devices with only one LUN per SCSI ID.

- *PSCSI*. The disk is treated as a parallel SCSI device. In particular, the SCSI and LUN ID information is retrieved from the position attribute in the **CuDV** entry for the *hdisk*. The disk is opened, and either a LUN reset or a target reset is sent, depending on whether a SCSI inquiry command reports this device as being *SCSI-2* or *SCSI-3*. Note, that an entry in **/etc/cluster/lunreset.lst** causes a LUN reset to be sent to a device that identifies itself as *SCSI-2*.
- *FSCSI*. The disk is treated as a fiber-attached SCSI device. In particular, the LUN ID is retrieved from the **lun_id** field in the **CuAt** entry for this device.

Making a Disk Available for Use by Another Node

This method allows for making a disk available in the AIX state, as returned by **lsdev** command. A device must be in an available state for the LVM to be able to access it and to vary on a volume group.

This method is passed as an input hdisk name, such as *hdisk42*.

A return code of zero indicates success; any other value causes HACMP to terminate processing for this disk.

HACMP supports a value of *MKDEV* for the name of this method. If this value is specified, HACMP uses existing processing for this method and attempts to run the following command up to four times:

```
mkdev -l <hdisk name>
```

Configuring OEM Disk Methods in SMIT

You can add, change, and remove custom disk processing methods for a specific OEM disk using SMIT. You can select existing HACMP-supported custom disk methods or use your own custom disk methods.

Note: There are no SMIT panels for modifying the following files:

/etc/cluster/conraid.dat

/etc/cluster/lunreset.lst

/etc/cluster/disktype.lst

These are text files that can be modified with an editor. For more information on how to change these files, see Supporting Disks with Known Characteristics.

Adding Custom Disk Methods

To add a custom disk method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration** > **Extended Resource Configuration** > **HACMP Extended Resources Configuration > Configure Custom Disk Methods > Add Custom Disk Methods** and press Enter.

3. Enter field values that define the disk processing methods you want to specify for the particular OEM disk:

Disk Type (PdDvLn field from CuDv)	Enter the identifier for the particular disk type. It is the value of the **PdDvLn** field of the **CuDv** entry for the disk. This value can be retrieved from the **CuDv** class in the HACMP configuration database with the following command: `odmget -q "name = <hdisk name>" CuDv`
Method to identify ghost disks	You can select a method from the picklist, or you can enter the full path name of a custom method that HACMP should use to identify ghost disks. Select the appropriate method and press Enter. Ghost disks are duplicate profiles of the disks created during disk configuration processing. Ghost disks must be removed to allow AIX to vary on the volume group residing on the original disks.
Method to determine if a reserve is held	You can select a method from the picklist, or you can enter the full path name of a custom method that HACMP should use to determine whether another node holds a reserve on the specified disk. Select the appropriate method and press Enter. A disk reserve restricts access to the disk by a specific node. If the node that established the reserve has failed, HACMP takes special steps to remove that disk reserve.
Method to break a reserve	You can select a method from the picklist, or you can enter the full path name of a custom method that HACMP should use to break a reserve on the specified disk. Select the appropriate method and press Enter. A disk reserve restricts access to the disk by a specific node. If the node that established the reserve has failed, HACMP takes special steps to remove that disk reserve.
Break reserves in parallel	Select **true** or **false**. The default is **false**. Some disk processing methods can be safely run in parallel. This may provide a performance advantage in cluster configurations with a large number of disks.

Methods to make the disk available	You can select a method from the picklist, or you can enter the full path name of a custom method that HACMP should use to make this disk available. Select the appropriate method and press Enter. Once a disk becomes accessible and any reserve has been removed from that disk, it must be made available in order for AIX to access that disk.

Note: The custom disk processing method that you specify for a particular OEM disk is added only to the local node. This information is *not* propagated to other nodes; you must copy this custom disk processing method to each node manually or use the HACMP File Collections facility.

Once you have made the selections, the information is applied to the disks defined on the local node.

4. Configure the same custom disk processing method on each other node in the cluster and synchronize the cluster resources. The cluster verification process ensures that the method that you configured exists and is executable on all nodes. The synchronization process ensures that the entries in the HACMP configuration database are the same on all nodes, but will *not* synchronize the methods named in the database entries.

Changing Custom Disk Methods

To change a custom disk method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Disk Methods > Change/Show Custom Disk Methods** and press Enter.

   SMIT displays a picklist with the names of the specified disk processing methods.
3. Select a name of a particular disk method and press Enter. SMIT displays the current information.
4. Enter new information in the fields you want to change and press Enter.

Note: The characteristics of the method are changed on the local node but they are *not* updated on remote nodes. Custom disk methods should be updated manually on other nodes in the cluster, or you can use the HACMP File Collections facility for this purpose.

Removing Custom Disk Methods

To remove a custom disk method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Disk Methods > Remove Custom Disk Methods** and press Enter.

SMIT displays a picklist with the names of the specified disk processing methods.

3. Select a disk method you want to remove and press Enter. SMIT prompts you to confirm your selection. If you choose to continue, the corresponding entry in **HACMPdisktype** will be deleted.

   Note: HACMP deletes the characteristics of the disk methods on the local node but does *not* update them on remote nodes. Manually delete custom disk methods on other nodes in the cluster.

Integrating OEM Volume Groups in an HACMP Cluster

You can configure OEM volume groups in AIX and use HACMP as an IBM High Availability solution to manage such volume groups, their corresponding file systems, and application servers. (Applications servers are defined as start and stop scripts for the applications supported by the volume groups).

For information on file systems, see Integrating OEM File Systems in an HACMP Cluster.

Note: Different OEMs may use different terminology to refer to similar constructs. From now on, this appendix uses the term *volume groups* to refer to OEM and Veritas volume groups.

In particular, HACMP 5.4.1 *automatically* detects and provides the methods for volume groups created with the Veritas Volume Manager (VxVM) using Veritas Foundation Suite (VFS) v. 4.0. Note, Veritas Foundation Suite is also referred to as Veritas Storage Foundation (VSF). This documentation uses VFS.

Use the following table to identify the terminology used for the storage components by IBM and Veritas. Other manufacturers may use similar terminology:

AIX Logical Volume Manager (LVM)	Veritas Volume Manager (VxVM)
Physical Volumes	Disk Media
Logical Volumes	Volumes
Volume Groups	Disk Groups

For general information on OEM volume groups and file systems, see the corresponding vendor storage documentation. In particular, for Veritas volume groups see the latest *Veritas Volume Manager Migration Guide for AIX.*

This section contains the following topics:

- Overview
- Supporting Veritas Volume Groups
- Configuring OEM Volume Groups Methods in SMIT.

Overview

Depending on the type of OEM volume, custom methods in HACMP allow you (or an OEM vendor) either to tell HACMP that a volume unknown to AIX LVM should be treated the same way as a known and supported volume or to specify the custom methods that provide the volume groups processing functions supported by HACMP.

You can either use custom methods for volume groups offered by HACMP, or create your own custom methods to use for the non-IBM volume groups in the cluster. These functions are performed within the normal HACMP event processing.

OEM Volume Group Functions Performed by HACMP

When HACMP identifies OEM volume groups of a particular type, it provides the following built-in processing functions for them or lets you specify your own custom methods for any one of these functions:

- List volume groups of a specified type defined on the nodes.
- List physical and logical disks comprising a specified volume group.
- Bring a volume group online and offline.
- Determine a volume group status.
- Verify volume groups configuration.
- Provide a location of log files and other debugging information. You can then view this information by using the AIX **snap -e** command.

HACMP enables you to manage OEM volume groups in an HACMP cluster by bringing them online and offline as needed for cluster event processing.

Before bringing a volume group online or offline, HACMP checks the volume group type to determine whether it must use the corresponding built-in custom method to manage the volume group and file system and bring it online or offline as needed.

OEM Volume Groups and File Systems as Resources in an HACMP Resource Group

You can include OEM disks, volume groups, and file systems in an HACMP resource group. HACMP recognizes OEM volume groups and file systems and handles them as the AIX LVM volume groups and file systems by listing, activating, checking and taking them offline when it is necessary for the cluster event processing tasks.

You can also mix the OEM disks, volume groups and file systems with AIX volume groups and file systems in a single HACMP resource group.

When you include OEM or Veritas volume groups or file systems into HACMP resource groups, **Automatically Import Volume Groups** field should be set to **false**. Also OEM volume groups should be set to *not* automatically varyon when the node is restarted; OEM file systems should be set to *not* automatically mount when the node is restarted.

To view OEM disks, volume groups and file systems used in the HACMP cluster, you can use the HACMP SMIT interface.

Prerequisites

In order for HACMP to handle OEM volume groups and file systems in your configuration, the volume groups and file systems must adhere to these conditions:

- Each OEM volume is composed of one or more physical disks (equivalent to `hdisks` in AIX) and the system can determine the names of the physical disks from the name of the volume.
- OEM disks, volume groups, and file systems must have operations and sequences of operations comparable to the functions of AIX LVM, although the command names, syntax, and arguments may differ.
- For Veritas volume groups, HACMP automatically uses them once they are configured in HACMP with the SMIT panels described below. For other OEM volume groups, HACMP uses the custom methods that you specify in SMIT.

Limitations

HACMP identifies OEM volume groups and file systems and performs all major functions with them. See OEM Volume Group Functions Performed by HACMP and OEM File Systems Functions Performed by HACMP.

However, some of the HACMP functions have limitations for OEM volume groups and file systems:

- You cannot use C-SPOC (Cluster Single Point of Control) to manage OEM disk, volume, and file systems operations. In particular, if you use C-SPOC for your other operations, HACMP does *not* include OEM disks, volume groups and file systems in picklists for your selection in SMIT.
- HACMP does *not* automatically discover OEM volume groups and file systems and does *not* list them for your selections in picklists in SMIT.
- HACMP does *not* use NFS (Network File System) functions for OEM file systems.
- HACMP does *not* provide a workaround for any limitations of OEM disks, volume groups, and file systems that exist outside of HACMP.
- In addition to listing, varying on and off, and verifying volume groups and file systems created with the AIX LVM, HACMP supports other extended functions for its “own” volume groups and file systems, such as enhanced concurrent mode, active and passive varyon process, heartbeating over disk, selective fallover upon volume group loss and other.

  These functions of HACMP utilize the AIX LVM capabilities, and since OEM volume groups and file systems do *not* use AIX LVM, HACMP does *not* support these functions for OEM volume groups.
- HACMP’s mechanism for fast disk takeover cannot be utilized for mixed physical volumes, that is, for disks comprised of both IBM and non-IBM disks. The LVM enhanced concurrent mode capabilities of AIX are used to enable fast disk takeover—only LVM supported disks can be used with enhanced concurrent mode and the Fast Disk Takeover feature.
- The Automatic Error Notification methods available in AIX cannot be configured in HACMP for OEM volume groups and file systems.

- You can have sites defined in the cluster, but you cannot configure and use HACMP/XD replicated resources (with the use of PPRC or Geographic LVM) in the cluster that utilizes OEM volume groups. The inter-site management policy for resource groups should be set to Ignore.
- HACMP does *not* support the LVM cross-site mirroring function for resource groups with OEM volume groups and file systems.

Software Requirements

HACMP supports Veritas volume groups in AIX v.5.2 (depending on the version level supported by VFS 4.0).

Supporting Veritas Volume Groups

Among other OEM volume groups and file systems, HACMP 5.4.1 supports volume groups and file systems created with VxVM in Veritas Foundation Suite v.4.0.

To make it easier for you to accommodate Veritas volume groups in the HACMP cluster, the methods for Veritas volume groups support are predefined in HACMP and are used automatically. After you add Veritas volume groups to HACMP resource groups, you can select the methods for the volume groups from the picklists in HACMP SMIT menus for OEM volume groups support.

Note: After you configure Veritas volume groups or file systems in HACMP cluster, the HACMP cluster verification utility runs the associated Veritas verification check. This verification check is effective only for the volume groups or file systems that are currently *online* on the node on which the verification is being run.

Configuring OEM Volume Groups Methods in SMIT

You can add, change, and remove custom volume groups processing methods for a specific OEM volume group using SMIT. You can select existing custom volume group methods that are supported by HACMP, or you can use your own custom methods.

You can add, change, and remove the methods that HACMP will use to manage OEM volume groups.

See these sections:

- Adding Custom Volume Group Methods
- Changing Custom Volume Group Methods
- Removing Custom Volume Group Methods.

Adding Custom Volume Group Methods

To add a custom volume group method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Volume Group Methods > Add Custom Volume Group Methods** and press Enter.

Note: If you are configuring Veritas volume groups, perform this operation in SMIT on the same node on which the volume group is currently active.

3. Enter field values that define the volume group processing methods you want to specify for the particular OEM volume group:

Volume Group Type	Enter the identifier for the particular volume group type, or use your own custom volume group methods. For Veritas volume groups, HACMP uses the method to call this type. By default, it is the value of the **PdDvLn** field of the **CuDv** entry for the volume. This value can be retrieved from the **CuDv** class in the HACMP configuration database (ODM) with the following command: `odmget -q "name = <volume name>" CuDv`
Method to List Volume Group Names	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to list volume group names for the specified volume group. As a default method, HACMP uses the AIX **lsvg** command. The custom method that you use must generate a list of the volume group names as an output. HACMP expects that the method returns zero for success and non-zero for failure. HACMP also expects that if a method passed to it is a list, the list should have one item per line.
Method to determine physical disks (hdisks) comprising the volume group	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to determine which physical disks must be made available on the node before HACMP brings the specified volume group online. As a default method, HACMP uses LSPV which calls the AIX **lspv** command. The custom method that you use must generate a list of the `hdisk` names as an output. HACMP expects that the method returns zero for success and non-zero for failure.

Method for bringing the volume group online	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to bring the volume group online. As a default method, HACMP uses VARYONVG which calls the AIX **varyonvg** command. HACMP expects that the method returns zero for success and non-zero for failure.
Method for determining volume group status	Select the method from the picklist, or enter the full path name of a custom method that HACMP should use to determine the volume group status. As a default method, HACMP uses LSACTIVEVG which calls the AIX **lsvg -o** command. HACMP takes as an input the volume group name and expects that the method returns 0 for offline, 1 for online and 2 for command failure. HACMP also expects that if a method passed to it is a list, the list should have one item per line.
Method for bringing the volume group offline	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to take the volume group offline. As a default method, HACMP uses VARYOFFVG which calls the AIX **varyoffvg** command. HACMP expects that the method returns zero for success and non-zero for failure.
Method for verifying volume configuration	Enter the full path name of a custom method that HACMP should use to verify the volume configuration on each cluster node. **Note:** By default, HACMP verifies the configuration for AIX volume groups and file systems. It also uses the Veritas verification method for Veritas volume groups and file systems. However, the Veritas verification is effective *only* for those volume groups and file systems that are currently active on the nodes on which the verification is being run. HACMP expects that the method returns zero for success and non-zero for failure.
Directory containing log information (optional)	Empty by default. Enter one or more space-separated directories in this field. You can use the AIX **snap -e** command to view this information.

Note: The custom volume group processing method that you specify for a particular OEM volume group is added to the local node only. This information is *not* propagated to other nodes; you must copy this custom volume group processing method to each node manually. Alternatively, you can use the HACMP File Collections facility to make the disk, volume, and file system methods available on all nodes.

Once you have made the selections, the information is applied to the volume groups on the local node.

4. Configure the same custom volume group type on other nodes in the cluster and synchronize the cluster resources. You may configure it manually or use the HACMP File Collections facility. The cluster verification process ensures that the type that you configured exists and is executable on all nodes. HACMP also verifies the root user permissions for the type and methods you specified, and the fact that the methods do *not* reside on a shared physical volume. The synchronization process ensures that the entries in the HACMP configuration database are the same on all nodes and synchronizes the methods named in the HACMP File Collections. HACMP runs the specific verification checks for the OEM-type volume groups only if they are configured in the cluster.

   Note: After you configure Veritas volume groups or file systems in HACMP cluster, the HACMP cluster verification utility runs the associated Veritas verification check. This verification check is effective *only* for the volume groups or file systems that are currently online on the node on which the verification is being run.

Changing Custom Volume Group Methods

To change a custom volume group method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Volume Group Methods > Change/Show Custom Volume Group Methods** and press Enter.

   SMIT displays a picklist containing the names of the specified volume group processing types (or methods).
3. Select a name of a particular volume group type and press Enter. SMIT displays the current information.
4. Enter new information in the fields you want to change and press Enter.

   Note: Unless you use the HACMP File Collections facility, the method's characteristics are changed on the local node, but they are *not* updated on remote nodes. Manually update the custom volume group method on other nodes in the cluster.

Removing Custom Volume Group Methods

To remove a custom volume method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Volume Group Methods > Remove Custom Volume Group Methods**.

   SMIT displays a picklist with the names of the specified volume group types (or processing methods).
3. Select a type or method you want to remove and press Enter. SMIT prompts you to confirm your selection. If you choose to continue, HACMP will delete the corresponding entry.

   Note: Unless you use the HACMP File Collections facility, HACMP deletes the characteristics of the volume group types or methods on the local node but does *not* update them on remote nodes. Manually delete custom types or methods on other nodes in the cluster.

Integrating OEM File Systems in an HACMP Cluster

You can configure OEM file systems in AIX and use HACMP as an IBM High Availability solution to manage such volume groups, their corresponding file systems, and application servers. (Applications servers are defined as start and stop scripts for the applications supported by the volume groups).

This section contains the following topics:

- Overview
- Supporting Veritas File Systems
- Configuring OEM File Systems Methods in SMIT.

Overview

Depending on the type of OEM volume, custom methods in HACMP allow you (or an OEM vendor) to tell HACMP that a file system unknown to AIX LVM should be treated the same way as a known and supported file system to specify the custom methods that provide the file systems processing functions supported by HACMP.

You can either use custom methods for file systems offered by HACMP, or create your own custom methods to use for the non-IBM file systems in the cluster. These functions are performed within the normal HACMP event processing.

In particular, HACMP 5.4.1 automatically detects and provides the methods for file systems created with Veritas File System (VxFS) using Veritas Foundation Suite v. 4.0.

OEM File Systems Functions Performed by HACMP

When HACMP identifies OEM file systems of a particular type, it provides the following processing functions for them or lets you specify your own custom methods for any one of these functions:

- Determine a list of file systems belonging to a specified type of volume group.
- List volume groups hosting a specified file system.
- Bring the file system online and offline.
- Determine the file system status.
- Verify the file systems configuration.
- Provide the pathnames to the file systems log files, for troubleshooting purposes.

OEM Volume Groups and File Systems as Resources in an HACMP Resource Group

You can include OEM disks, volume groups, and file systems in an HACMP resource group. HACMP recognizes OEM volume groups and file systems and handles them as the AIX LVM volume groups and file systems by listing, activating, checking and taking them offline when it is necessary for the cluster event processing tasks.

You can mix the OEM disks, volume groups and file systems with AIX volume groups and file systems in a single HACMP resource group.

When you include OEM or Veritas volume groups or file systems into HACMP resource groups, **Automatically Import Volume Groups** field should be set to **false**. Also OEM volume groups should be set to *not* automatically varyon when the node is restarted; OEM file systems should be set to *not* automatically mount when the node is restarted.

To view OEM disks, volume groups and file systems used in the HACMP cluster, you can use the HACMP SMIT interface.

Prerequisites

In order for HACMP to handle OEM volume groups and file systems in your configuration, the volume groups and file systems must adhere to these conditions:

- Each OEM volume is composed of one or more physical disks (equivalent to `hdisks` in AIX) and the system can determine the names of the physical disks from the name of the volume.
- OEM disks, volume groups, and file systems must have operations and sequences of operations comparable to the functions of AIX LVM, although the command names, syntax, and arguments may differ.

Limitations

HACMP identifies OEM volume groups and file systems and performs all major functions with them. See OEM Volume Group Functions Performed by HACMP and OEM File Systems Functions Performed by HACMP.

However, some of the HACMP functions have limitations for OEM volume groups and file systems:

- You cannot use C-SPOC (Cluster Single Point of Control) to manage OEM disk, volume, and file systems operations.

- HACMP does *not* automatically discover OEM volume groups and file systems and does *not* list them for your selections in picklists in SMIT.
- HACMP does *not* use NFS (Network File System) functions for OEM file systems.
- HACMP does *not* provide a workaround for any limitations of OEM disks, volume groups, and file systems that exist outside of HACMP.
- In addition to listing, varying on and off, and verifying volume groups and file systems created with the AIX LVM, HACMP supports other extended functions for its "own" volume groups and file systems, such as enhanced concurrent mode, active and passive varyon process, disk heartbeating and selective fallover upon volume group loss and other.

  These functions of HACMP utilize the AIX LVM capabilities, and since OEM volume groups and file systems do *not* use AIX LVM, HACMP does *not* support these functions for OEM volume groups.
- The Automatic Error Notification methods available in AIX cannot be configured in HACMP for OEM volume groups and file systems.
- You can have sites defined in the cluster, but you cannot configure and use HACMP/XD replicated resources (with the use of PPRC or Geographic LVM) in the cluster that utilizes OEM volume groups. The inter-site management policy for resource groups should be set to Ignore.
- HACMP does *not* support the LVM cross-site mirroring function for resource groups with OEM volume groups and file systems.
- The HACMP's mechanism for fast disk takeover cannot be utilized for mixed physical volumes, that is, for disks comprised of both IBM and non-IBM disks. The LVM enhanced concurrent mode capabilities of AIX are used to enable fast disk takeover—only LVM supported disks can be used with enhanced concurrent mode and the Fast Disk Takeover feature.
- HACMP uses its serial method of processing resource groups if they contain OEM volume groups or file systems.

Software Requirements

HACMP 5.4.1 supports Veritas volume groups in the following software environment:

- AIX v.5.2 (depending on the version level supported by VFS 4.0).

Supporting Veritas File Systems

Among other OEM volume groups and file systems, HACMP 5.4.1 supports volume groups and file systems created with VxVM in Veritas Foundation Suite v.4.0.

To make it easier for you to accommodate Veritas file systems in the HACMP cluster, the methods for Veritas file systems support are predefined in HACMP. After you add Veritas file systems to HACMP resource groups, you can select the methods for the file systems from the picklists in HACMP SMIT menus for OEM file systems support.

Note: After you configure Veritas volume groups or file systems in HACMP cluster, the HACMP cluster verification utility runs the associated Veritas verification check. This verification check is effective *only* for the volume groups or file systems that are currently online on the node on which the verification is being run.

Configuring OEM File Systems Methods in SMIT

You can add, change, and remove custom file systems processing methods for a specific OEM file system using SMIT. You can select existing custom file systems methods that are supported by HACMP, or you can use your own custom methods.

See these sections:

- Adding Custom File Systems Methods
- Changing Custom File Systems Methods
- Removing Custom File Systems Methods.

Adding Custom File Systems Methods

To add a custom file system method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Filesystems Methods > Add Custom Filesystem Methods** and press Enter.

   Note: If you are configuring Veritas file systems, perform this operation in SMIT on the same node on which the file system is currently active.

3. Enter field values that define the file system processing methods you want to specify for the particular OEM file system type:

Filesystem Type	Enter the identifier for the particular file system type for which you want to configure existing methods in HACMP or use your own custom methods. By default, it is the value of the **VFS** field of the **/etc/filesystems** file for the volume. Use the **lsfs** command to obtain this value from the **/etc/filesystems** file.
Method for listing filesystem names	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to list file system names for the specified volume type. As a default method, HACMP uses LSFS which calls the AIX **lsfs** command. The custom method that you use must generate a list of the file system names as an output. HACMP expects that the method returns zero for success and non-zero for failure. HACMP also expects that if a method passed to it is a list, the list should have one item per line.

Method for listing volume groups hosting a specified filesystem	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to list volume groups that host the file system. As a default method, HACMP compares the AIX logical volume associated with the file system with the AIX logical volume that is part of the volume group. By default, HACMP uses ODM, which calls the **/usr/es/sbin/cluster/events/utils/cl_fs2disk** command. HACMP expects that the method returns zero for success and non-zero for failure. HACMP also expects that if a method passed to it is a list, the list should have one item per line.
Method for bringing the filesystem online	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to activate the file system. As a default method, HACMP uses MOUNT which calls the AIX **mount** command. HACMP expects that the method returns zero for success and non-zero for failure.
Method for bringing the filesystem offline	Select a method from the picklist, or enter the full path name of a custom method that HACMP should use to take the file system offline. As a default method, HACMP uses UMOUNT which calls the AIX **umount** command. HACMP expects that the method returns zero for success and non-zero for failure.
Method for determining the filesystem status	Select the method from the picklist, or enter the full path name of a custom method that HACMP should use to determine the file system status. As a default method, HACMP uses LSACTIVEFS which calls the **mount 2>/dev/null | awk $3 ~ /jfs2*$/ {print $2}** command. HACMP takes as an input the file system name and expects that the method returns 0 if the file system is *not* mounted, 1 for the file system that is mounted, and 2 for command failure.

Method for verifying filesystem configuration	Enter the full path name of a custom method that HACMP should use to verify the file system configuration on each cluster node. **Note:** By default, HACMP verifies the configuration for AIX volume groups and file systems. It also uses the Veritas verification method for Veritas volume groups and file systems. However, the Veritas verification is effective *only* for those volume groups and file systems that are currently active on the nodes on which the verification is being run. HACMP expects that the method returns zero for success and non-zero for failure.
Directories containing log information (optional)	Enter one or more space-separated pathnames to the log files in this field, or leave the field blank. If HACMP finds the files in this field, you can use the **snap -e** command to troubleshoot the log files.

Note: The custom file system processing method or custom file system type is added to the local node only. This information is *not* propagated to other nodes; you copy this method or type to each node manually. Alternatively, you can use the HACMP File Collections facility to make the disk, volume, and file system methods available on all nodes.

Once you have made the selections, the information is applied to the file systems on the local node.

4. Configure the same custom file systems processing method on other nodes in the cluster and synchronize the cluster resources. The cluster verification process ensures that the method (or the file system type) that you configured exists and is executable on all nodes. HACMP also verifies the root user permissions for the methods you specified, and the fact that the methods do *not* reside on a shared physical volume. The synchronization process ensures that the entries in the HACMP configuration database are the same on all nodes and synchronizes the methods named in the HACMP File Collections.

Note: After you configure Veritas volume groups or file systems in HACMP cluster, the HACMP cluster verification utility runs the associated Veritas verification check. This verification check is effective *only* for the volume groups or file systems that are currently online on the node on which the verification is being run.

Changing Custom File Systems Methods

To change a custom file system method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Filesystem Methods > Change/Show Custom Filesystems Methods** and press Enter.

SMIT displays a picklist with the names of the specified file systems processing methods (or types).

3. Select a name of a particular file system method (or type) and press Enter. SMIT displays the current information.
4. Enter new information in the fields you want to change and press Enter.

   Note: Unless you use the HACMP File Collections facility, the type or method characteristics are changed on the local node but they are *not* updated on remote nodes. Manually update the custom methods or types on other nodes in the cluster.

Removing Custom File Systems Methods

To remove a custom file system method:

1. Enter `smit hacmp`
2. In SMIT, select **Extended Configuration > Extended Resource Configuration > HACMP Extended Resources Configuration > Configure Custom Filesystem Methods > Remove Custom Filesystem Methods** and press Enter.

   SMIT displays a picklist with the names of the specified file system types or processing methods.
3. Select the file system type or method you want to remove, and press Enter. SMIT prompts you to confirm your selection. If you choose to continue, HACMP will delete the corresponding entry.

   Note: Unless you use the HACMP File Collections facility, HACMP deletes the characteristics of the types or methods on the local node but does *not* update them on remote nodes. Manually delete custom types or methods on other nodes in the cluster.

Appendix C: GPFS Cluster Configuration

This appendix describes how to configure, manage, and remove an IBM General Parallel File System (GPFS) version 2.3 cluster within an HACMP cluster environment, using the HACMP SMIT interface.

You can also configure and manage GPFS clusters using the native GPFS command set (**mm***xxx*). *Do not mix these two methods of configuring and managing a GPFS cluster*; choose one or the other. For complete documentation on GPFS including restrictions and how to use the native GPFS command set, see the following URL:

http://www.ibm.com/servers/eserver/clusters/software/gpfs.html

Before configuring GPFS, the hardware must be set up as required by GPFS and all restrictions imposed by GPFS apply. This appendix contains these sections:

- Overview
- Planning for a GPFS Cluster
- Installing GPFS
- Configuring a GPFS Cluster Using HACMP SMIT
- HACMP Cluster Operations and GPFS
- Troubleshooting the GPFS Cluster.

Overview

GPFS 2.3 provides concurrent high speed file access to applications executing on multiple systems that are part of an HACMP cluster. It offers:

- High-performance file system for IBM UNIX® clusters capable of supporting multi-terabytes of storage within a single file system
- Shared-disk file system where every GPFS cluster node can have concurrent read/write access to the same file
- High availability through automatic recovery from node and disk failures.

GPFS in an HACMP Cluster Environment

HACMP and AIX provide the operating and administrative environment for GPFS 2.3 in a cluster environment. GPFS uses some part of the CLVM component of AIX and also has its own data replication method.

GPFS uses the topology framework of an HACMP cluster: Nodes, networks, and network interfaces used by a GPFS cluster are configured as part of the HACMP cluster. Only one GPFS cluster can be configured per HACMP cluster.

GPFS Nodes

Each node in the GPFS cluster can be defined as belonging to a GPFS nodeset. A GPFS cluster can contain multiple nodesets. All nodes within a GPFS nodeset have access to the same GPFS file systems and must have access to all the directly-attached disks that form part of the GPFS cluster.

Note: When using the HACMP SMIT panels to configure a GPFS cluster, all the cluster nodes are included in one GPFS nodeset.

GPFS cluster nodes can host HACMP cluster resources as well as GPFS.

GPFS Network and Network Interfaces

One IP network is defined for communication between GPFS cluster nodes. This network should *not* be used for service by HACMP. You can either use node-bound service IP interfaces on a network or base interfaces on networks using IPAT via IP Aliases.

Note: When you are configuring GPFS on SP nodes running HACMP using the HACMP SMIT method, you will get an error if you configure the SP Switch network for use by GPFS. To use GPFS on SP nodes, configure another IP network in HACMP that either has node-bound service addresses defined on it, or uses IPAT via IP Aliases.

GPFS Disk Devices

The disk devices are *not* defined to the HACMP cluster as HACMP resources; they are defined only for use by GPFS. If you include the GPFS disk, volume group, or file system in an HACMP resource group, verification reports this as an error.

Planning for a GPFS Cluster

See the GPFS 2.3 manuals for extensive information on planning for GPFS in HACMP clusters.

Basic Configuration Considerations for GPFS Clusters

Basic considerations for planning the GPFS set up:

- GPFS uses an IP network to connect all of the nodes. This is typically a LAN with sufficient bandwidth (minimum 100MB/sec bandwidth) available for GPFS control traffic. The IP network and network interfaces are configured within the HACMP cluster.

  GPFS requires invariant network connections; network interfaces supporting TCP/IP socket connections for GPFS cannot be configured for Hardware Address Takeover (HWAT). Use a single service network interface per network, or base interfaces on networks using IPAT via IP Aliases.
- GPFS requires that the SSA or Fibre Channel disk devices be configured and directly attached on all of the nodes that will mount the file systems. The disks, volume groups, and file systems *must not* be configured within the HACMP cluster resource groups.

Tuning the HACMP Cluster for GPFS

Tune the cluster before you configure GPFS. Components to consider include:

Communications I/O

Set the network option *ipqmaxlen* to 512 (the default is 128). This parameter controls the number of incoming packets that can exist on the IP interrupt queue.

```
no -o ipmaxlen=512
```

Since this option must be modified at every reboot, it is suggested to place it at the end of one of the system startup files, such as the **/etc/rc.net** shell script.

Disk I/O

Set the SSA disk I/O option *max_coalesce* parameter to allow the SSA device driver to coalesce requests that have been broken up to satisfy LVM requirements:

```
chdev -1 hdiskX - a max_coalesce=0x40000
```

Security

Ensure you have an ~**/.rhosts** file on every node in the GPFS cluster. You must be the root user.

Installing GPFS

Before you install GPFS, install the correct levels of the following software:

- AIX
- HACMP
- RSCT.

You can configure a GPFS Cluster using HACMP SMIT. To use this capability, be sure to install the **cluster.es.cfs.rte** fileset.

For GPFS installation instructions, see the GPFS version 2.3 documentation.

Configuring a GPFS Cluster Using HACMP SMIT

A **Configure GPFS** entry is available on the **System Management (C-SPOC)** SMIT panel if you have installed the GPFS fileset (**cluster.es.cfs.rte**). Selecting this item leads to the following menu of actions available for configuring and manipulating a GPFS cluster and the GPFS file systems:

- Configure a GPFS Cluster - create and configure a GPFS cluster
- List All GPFS Filesystems - List all configured GPFS file systems
- Add GPFS Filesystem - Add a GPFS file system to an existing GPFS cluster
- Remove GPFS Filesystem - Remove existing GPFS file system(s)
- Remove a GPFS Cluster - Remove a GPFS cluster and all file systems.

Note: The use of the HACMP SMIT panels is limited to GPFS clusters created and configured using these panels. You cannot use the Add or Remove function on any previously configured GPFS cluster or GPFS file system. However, the List option will list all GPFS file systems, no matter the method used to configure them.

Prerequisites and Limitations

Using the **Configure GPFS** panels simplifies the process of creating and removing the GPFS cluster components. However, certain conditions apply:

- You can have one GPFS cluster per HACMP cluster. No previously configured GPFS cluster may exist in the HACMP cluster. *All nodes in the HACMP cluster will be included in the GPFS cluster created by HACMP SMIT.*
- If you are using Fibre Channel connected disks, you must have at least three nodes in the cluster (for GPFS quorum). This is a GPFS limitation.

  You must have configured and synchronized an HACMP cluster with topology that includes one IP-based network that includes all nodes, with node-bound service IP label/addresses (for exclusive use of GPFS) or non-service interfaces on networks using IPAT via IP Aliases.
- All nodes must be up and running HACMP.

Step 1: Create the HACMP and GPFS Cluster Topology

Use the HACMP SMIT panels to configure the cluster name, nodes, networks and network interfaces that will be used by the GPFS cluster. For information about how to configure the topology from HACMP SMIT, see Chapter 4: Configuring HACMP Cluster Topology and Resources (Extended) in the *Administration Guide*.

You must include at least one IP-based network that includes all nodes, with node-bound service IP label/addresses (for exclusive use of GPFS) or non-service interfaces on networks using IPAT via IP Aliases.

Step 2: Create the GPFS Cluster Using SMIT

From the main HACMP menu, select **System Management (C-SPOC) > Configure GPFS > Configure a GPFS Cluster**.

The software performs the following steps to create and configure the GPFS cluster:

- Check that no GPFS cluster is currently configured

  Check that an HACMP network with a node-bound service IP label/address on each node is available for the GPFS network. If none, check for non-service interfaces on networks using IPAT via IP Aliases.
- Select the first node listed as the primary server
- Create a nodeset with the name HAgpfs01 that includes all the HACMP cluster nodes
- Run the **mmcrcluster** command to create the GPFS cluster
- Verify and synchronize the GPFS cluster configuration.

Notes on the GPFS Configuration Created Using SMIT

If you have configured more than one network that meets the criteria for the GPFS cluster, the networks are sorted by name and the first one in the list is picked as the GPFS network. (HACMP checks for a network with node-bound service IP labels/addresses first.)

The GPFS network interface information is stored along with the GPFS Cluster ID. HACMP uses this information to recognize the presence of a GPFS cluster in the HACMP environment.

The GPFS cluster created by using the HACMP SMIT panels will always contain all the HACMP cluster nodes. The nodeset is named **HAgpfs01**.

The first node listed is selected as a primary GPFS cluster data server node.

Step 3: Add the GPFS File System

To add the GPFS file system:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Configure GPFS** and press Enter.
3. Select the **Add a GPFS Filesystem** SMIT panel and press Enter.
4. Fill in the fields as follows to set up the GPFS file system:

GPFS Filesystem Name	The name of the file system to be created. Use no more than eight alphanumeric characters and underscores.
GPFS Filesystem Mount Point	The mount point for the file system.
hdisks	The list of hdisks to be used for the GPFS descriptor file. You can select more than one hdisk in the picklist.
Force?	**true** = Force creation of the GPFS file system; overwrite any existing GPFS file system from a previous configuration. The default is **false**.

5. Press Enter. The GPFS daemon is started if it is *not* already running. The network shared disk is created on the chosen hdisk. The file system is created and mounted on each cluster node.

If there are any problems, error messages are printed to the screen.

HACMP Cluster Operations and GPFS

As administrator of the HACMP cluster with GPFS, be aware of the effect on GPFS when you make any changes to the cluster.

GPFS assumes that IP addresses, node names, and hostnames remain constant. With this in mind, if you make HACMP cluster topology configuration changes, take care that GPFS is *not* affected. For example:

- Before changing a cluster node's name, IP address, or hostname, remove the node from the cluster, make the changes, then add it back to the cluster. You can use a DARE operation to remove the node and add it back in.
- If a network interface fails, quorum may be affected if the gpfs daemon stops on a node. GPFS may *not* be available. GPFS does *not* support the use of backup network interfaces; therefore HACMP cannot take corrective action.

Using the HACMP SMIT panels to configure and manipulate the GPFS cluster ensures that these actions are done properly and issues a warning if quorum is affected. The verification utility reports errors if the GPFS network attributes are changed.

GPFS Cluster Dynamic Reconfiguration Operations

If you configured the GPFS cluster using the HACMP SMIT method, you can make the following changes to HACMP cluster topology while GPFS is active (following the HACMP DARE rules):

- You can add or remove a cluster node as part of a dynamic reconfiguration. These operations will cause the GPFS cluster to be reconfigured to reflect the change. During reconfiguration the GPFS cluster is stopped temporarily, then restarted after the node is added or deleted. If you delete the GPFS primary node, the software will reassign the primary.

  Adding a node dynamically to the HACMP configuration also adds it to the GPFS cluster configuration. However, you then have to mount the GPFS file system on the new node the first time when cluster services are started.
- Swap a GPFS service network interface using the **Swap Network Adapter** option on the **Cluster Communications Adapter Management** SMIT panel if there is a non-service network interface configured. The GPFS file system may be unavailable for a short time.
- Use the **PCI Hot Plug Replace Network Adapter** option on the **Cluster Communications Adapter Management** SMIT panel to replace a GPFS service network interface.

The verification utility reports errors if you have performed an operation that affects the GPFS cluster.

Verifying the Modified GPFS Cluster

You do *not* need to run verification after you first create the GPFS cluster. However, you should verify and synchronize after making *any* change to HACMP configuration.

Running the verification utility on a GPFS cluster:

- Checks that networks, network interfaces, and disk access are all correctly configured for GPFS

- Prints a message reminding you that the GPFS file system must be manually mounted on the new node(s) after starting the HACMP cluster services.
- If you are adding node(s) to a cluster that previously had two nodes, the GPFS daemon must be stopped on all nodes and started with multi-node quorum. An onscreen message indicates that the GPFS file system will be unavailable during this time.
- Prints an informational message that the removed node will be automatically deleted from the GPFS cluster configuration and GPFS file systems will *not* be available on that node
- Prints a WARNING message if you deleted a node and the number of remaining nodes in the GPFS cluster causes a quorum issue.
- (On dynamic reconfiguration only) If there are fewer nodes in the GPFS cluster than in the HACMP node, a forward check is made to ensure that the new node(s) contains a network interface of the same kind, on the same subnet and network that is configured for use by GPFS.

Listing GPFS File Systems Using HACMP SMIT

To list GPFS file systems:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Configure GPFS > List All GPFS Filesystems**. The system lists the information for all the GPFS file systems files found in **/etc/filesystems**, similar to the following sample output:

```
============================================================
Filesystem Name: gpfs1
Filesystem Mount Dir: /gpfs1
Filesystem Details:
flag value            description
---- --------------
------------------------------------------------------------
 -s  roundRobin       Stripe method
 -f  8192             Minimum fragment size in bytes
 -i  512              Inode size in bytes
 -I  16384            Indirect block size in bytes
 -m 1                 Default number of metadata replicas
 -M  1                Maximum number of metadata replicas
 -r  1                Default number of data replicas
 -R  1                Maximum number of data replicas
 -a  1048576          Estimated average file size
 -n  32               Estimated number of nodes that will mount
                      file system
 -B  262144           Block size
 -Q  none             Quotas enforced
     none             Default quotas enabled
 -F  33792            Maximum number of inodes
 -V  5.01             File system version. Highest supported
                      version: 5.01
 -z  no               Is DMAPI enabled?
 -d  gpfs0lv;gpfs1lv  Disks in file system
 -A  yes              Automatic mount option
 -C  gpfsns1          GPFS nodeset identifier
 -E  no               Exact mtime default mount option
 -S  no               Suppress atime default mount option
```

Removing a GPFS File System Using HACMP SMIT

If you configured the GPFS file system using HACMP SMIT, you can remove it with the GPFS Cluster **Remove GPFS Filesystem** option.

To remove a GPFS file system:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Configure GPFS** and press Enter.
3. Select the **Remove GPFS Filesystem** option.
4. Enter the following information and press Enter:

*** GPFS Filesystem Name**	You can select more than one GPFS file system name from the list. The selected file systems will be deleted.
*** Force?**	Selecting **True** will force the removal of the file system, even if the disks are damaged. The default is **False**.

Note: Although the file systems are deleted, the network shared disks and volume groups used by the GPFS file systems are left untouched. Remove them manually if you do *not* want them. See the GPFS documentation.

Removing a GPFS Cluster Using HACMP SMIT

If you configured the GPFS cluster using HACMP SMIT, you can remove it with the **Remove a GPFS Cluster** option. First remove all GPFS file systems. You can force-delete the GPFS cluster even when there are active file systems, by setting the **Remove All GPFS Filesystems** option to TRUE.

To force-delete a GPFS cluster:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Configure GPFS**.
3. Select the **Remove a GPFS Cluster** option.
4. Choose whether to force-delete the active file systems or *not* and press Enter.

   A message appears warning you that this action will remove all GPFS cluster information and all GPFS file systems. SMIT prompts you to confirm your selections.
5. To complete the removal, press Enter again.

If a cluster node is *not* available when the **mmdelnode** command runs, GPFS cannot remove its configuration files. In this case, you will have to manually remove the files from that node. See the GPFS documentation for the procedure.

Removing the HACMP Cluster Definition of a GPFS Cluster

Remove the GPFS cluster definition *before* removing the HACMP cluster definition. If you remove the HACMP cluster definition first, the GPFS cluster definition still exists and you will have to manually delete it from the configuration database.

Troubleshooting the GPFS Cluster

The **/usr/sbin/cluster/utilities/clgetesdbginfo** log file includes GPFS logs for debugging purposes.

If you use the HACMP SMIT panels to create the GPFS cluster, then when you take a cluster snapshot the GPFS configuration file **/var/mmfs/gen/mmsdrfs** is saved in the HACMP cluster snapshot **.info** file. However, restoring and applying snapshots does *not* create a GPFS cluster. You must recreate it following the procedures outlined here after restoring the HACMP cluster.

GPFS ODM (HACMPgpfs) Data

If the HACMP GPFS fileset is installed, the HACMPgpfs ODM is added to the HACMP cluster configuration. This ODM is used to store the configuration data that GPFS uses for its operation. The ODM stores one data entry per one HACMP node. This ODM consists of the following fields:

HACMPgpfs:

```
{
gpfs_clusterid [20];
gpfs_version [25];
nodeset_name [10]
gpfs_network [33];
identifier [32];
interfacename [33];
nodename [33];
}
```

References

- *Sizing and Tuning GPFS: SG24-5610*
- *GPFS in an RS/6000 Cluster: SG24-6035*
- *IBM GPFS Concepts, Planning, and Installation Guide: GA22-7453*
- *IBM GPFS for AIX Administration and Programming Reference: SA22-7452*
- *AIX Performance Management Guide*

Appendix D: HACMP and SNMP Utilities

This appendix discusses the Simple Network Management Protocol (SNMP) and describes the relationship between the HACMP 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 a 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 standard for defining 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. The Cluster Manager 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 software, NetView for AIX, and Systems Monitor for AIX all include the following SMUX daemons: **clstrmgr**, **trapgend**, and **sysinfod**, respectively. You must be aware of possible conflicts between these daemons.

HACMP SNMP Components

The HACMP software provides an enterprise-specific (generic type 6) MIB. The source file is **hacmp.my**. The **mosy** command compiles **hacmp.my** (with other standard MIBs) to generate the **hacmp.defs** file.

The HACMP MIB, is associated with and maintained by the Cluster Manager. The HACMP software also provides two cluster monitor programs, the Cluster Information Program (Clinfo) and **clstat**.

Cluster Information Program (Clinfo)

Clinfo is a cluster monitoring program. It requests information about the current cluster state from the Cluster Manager. The Cluster Manger updates data using internal, dynamically allocated data structures that are accessible to Clinfo clients—applications that use Clinfo API functions.

By default, Clinfo receives information from the Cluster Manager by polling. The time between polling is set by an argument to Clinfo, which defaults to 15. Clinfo can also receive information asynchronously through traps. In response to traps, Clinfo sends a request for more information to the Cluster Manager. It does *not* parse the trap message data itself; instead, Clinfo employs a trap-directed polling policy.

To enable Clinfo to receive traps, call it using the **-a** option. Since Clinfo is started through the System Resource Controller (SRC), the best way to do this is by entering:

```
chssys -s clinfoES -a "-a"
```

Then use the **lssrc** command to ensure this change occurred. Enter:

```
lssrc -Ss clinfoES | awk -F: '{print $3}'
```

Traps provide more timely information to Clinfo clients. The trap function is completely transparent to these clients—they simply register to receive various events. Clinfo notifies the traps 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** switch by entering:

```
chssys -s clinfoES -a "-c abcdef"
```

where `abcdef` is an SNMP community name defined as such in the **snmpd.conf** file.

Then use the **lssrc** command to ensure this change occurred. Enter:

```
lssrc -Ss clinfoES | awk -F: '{print $3}'
```

HACMP now supports a SNMP Community Name other than “public”. If the default SNMP Community Name has been changed in **/etc/snmpd.conf** to something different from the default of “public” HACMP will function correctly. The SNMP Community Name used by HACMP will be the first name found that is *not* “private” or “system” using the **lssrc -ls snmpd** command. The Clinfo service will also get the SNMP Community Name in the same manner.

The Clinfo service still supports the **-c** option for specifying SNMP Community Name but its use is *not* required. The use of the **-c** option is considered a security risk because doing a **ps** command could find the SNMP Community Name. If it is important to keep the SNMP Community Name protected, change permissions on **/tmp/hacmp.out, /etc/snmpd.conf, /smit.log** and **/usr/tmp/snmpd.log** to *not* be world readable.

Important Notes on snmpdv3.conf File

AIX **snmpdv3** has three functions or parts: One is the SNMP v3 agent, one is the SMUX server, and the last is the DPI2 agent. The DPI2 agent has to use community “public” to get a port number from DPI2 subagents (**hostmibd**, **snmpmibd**, **aixmibd**) to communicate with them. For this reason you should still keep community name “public” and give the “public” a view of

only this *dpiPortForTCP.0 (1.3.6.1.4.1.2.2.1.1.1.0)* MIB variable so that the DPI2 agent can get the port number from subagents. See the example **snmpdv3.conf** file below. Also, refer to the documentation at:

http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/commadmn/snmp_v3architecture.htm

Sample snmpdv3.conf File with Non-Public Community Name

```
VACM_GROUP group1 SNMPv1  YourLongNameHere  -
VACM_GROUP group2 SNMPv1  public  -

VACM_VIEW defaultView        internet                  - included -
VACM_VIEW dpi2view        1.3.6.1.4.1.2.2.1.1.1        - included -

VACM_VIEW defaultView        snmpModules               - excluded -
VACM_VIEW defaultView        1.3.6.1.6.3.1.1.4         - included -
VACM_VIEW defaultView        1.3.6.1.6.3.1.1.5         - included -

VACM_VIEW defaultView        1.3.6.1.4.1.2.6.191       - excluded -

VACM_ACCESS  group1 - - noAuthNoPriv SNMPv1  defaultView - defaultView -
VACM_ACCESS  group2 - - noAuthNoPriv SNMPv1  dpi2view - - -

NOTIFY notify1 traptag trap -

TARGET_ADDRESS Target1 UDP 127.0.0.1       traptag trapparms1 - - -

TARGET_PARAMETERS trapparms1 SNMPv1  SNMPv1  YourLongNameHere
noAuthNoPriv -

COMMUNITY YourLongNameHere   YourLongNameHere    noAuthNoPriv 0.0.0.0
0.0.0.0         -
COMMUNITY public    public     noAuthNoPriv 0.0.0.0     0.0.0.0
-

DEFAULT_SECURITY no-access - -

logging         file=/usr/tmp/snmpdv3.log       enabled
logging         size=4194304                    level=0

smux            1.3.6.1.4.1.2.3.1.2.1.2         gated_password  # gated

smux 1.3.6.1.4.1.2.3.1.2.3.1.1 muxatmd_password #muxatmd
smux     1.3.6.1.4.1.2.3.1.2.1.5      clsmuxpd_password # HACMP/ES for
AIX clsmuxpd
```

Note: See the AIX documentation for full information on the **snmpd.conf** file. Version 3 (default for AIX 5.2 and up) has some differences from Version 1.

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. However, if you want to run an ASCII display on an X-capable machine, you can do so by specifying the **-a** option.

In addition, you can set up **clstat** to display in a web browser, if you set up a web server on a node that has **clinfo** running. The browser display makes it easier to view multiple clusters without having to select the clusters one at a time.

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/sbin/cluster/samples/clstat.**

NetView for AIX

NetView for AIX is a network manager that 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 System p™ processors running AIX, 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 (**errdaemon**) 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 that monitors the following characteristics:

- Machine name, type, and processor ID
- Devices installed on the machine
- Operating system configuration
- Status of subsystems, paging devices, and file systems
- 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.

For more information about using this product, see the Systems Monitor for AIX documentation.

Systems Monitor Startup Options for HACMP Compatibility

If you are using the Systems Monitor for AIX along with HACMP on your system, start the **sysinfod** with the **-H** option. This option allows the HACMP 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. However, only one SNMP manager (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 is set in the **/usr/adm/sm6000/config/install.config** configuration file. 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. 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 the SNMP process (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.

Notices for HACMP Installation Guide

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

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

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

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Dept. LRAS / Bldg. 003
11400 Burnet Road
Austin, TX 78758-3493
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

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

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

Index

+-*/

/etc/cluster/conraid.dat file 159
/etc/cluster/disktype.lst file 159
/etc/cluster/lunreset.lst file 159
/etc/filesystems file 106
/etc/hosts file 119, 136
/etc/rc.net script
  modifying for use with GPFS 179
/etc/resolv.conf file 119
/etc/snmpd.conf
  notes on upgrading AIX 49
/lpp/cluster/objrepos/HACMPnpp 42
/lpp/cluster/objrepos/HACMPude 42
/usr/es/sbin/cluster/etc/clhosts.client 70
/usr/es/sbin/cluster/etc/clinfo.rc script 71
/usr/es/sbin/cluster/etc/harc.net file (automounter daemon) 120
/usr/es/sbin/cluster/wsm/README file
  Web SMIT installation information 40
/usr/lpp/save.config directory 42

A

AIX
  error notification 124
  setting
    syncd frequency 118
  user and group IDs 118
AIX error notification 124
  automatic 125
ARP 71
  enabling on SP switch 130
ARP cache
  ATM classic IP configuration 79
  clinfo.rc script 71
Asynchronous Transfer Mode. See ATM.
ATM
  configuring 79
    Classic IP clients for HACMP 80
  LAN emulation 84
automatic error notification 125
  deleting methods assigned 127
automatically importing volume groups 109
automounter daemon, enabling 120

B

boot addresses
  in nameserver configuration 119
boot non-service IP labels 74

C

cascading resource groups
  migrating to HACMP 5.4.1 26
changing GPFS cluster topology 182
checking
  installed hardware 73
  SCSI devices installation 88
classic IP and ATM protocol 79
clhosts file 70
Clinfo
  polling 188
  relation to SNMP 188
  setting up 69
    files and scripts 69
clinfo.rc script 71
  customizing 71
clstat utility 189
cluster
  GPFS 177
  planning
    shared LVM components 15
cluster monitoring
  with Tivoli
    defining managed nodes 146
    installation prerequisites 144
    IPAT considerations 148
    required Tivoli software 51, 146
    subnet requirements 145
    uninstalling 149
cluster.es.server.cfs
  GPFS cluster fileset 179
commands
  Eclock 131
  Estart 131
  importvg 112
  restrictions during upgrade 31
concurrent access mode
  migration issues 28
  shared LVM components 110
Concurrent Resource Manager 47
config_too_long
  handling during migration 45
configuration database. See HACMP Configuration Database. 46
Configuration_Files
  default file collection 121

configuring
ATM adapters 79
automatic error notification 126
basic cluster 133
custom disk methods 159
GPFS with HACMP SMIT 179
HACMP 19, 66
network adapters in AIX 73
target mode SCSI 92
target mode SSA 95, 96
creating
shared file systems 105
shared volume groups
concurrent access 111
cron utility and NIS 119
custom
disk
methods 157
processing 159
file systems
processing 172
volume groups
processing 165
custom resource groups
migrating to HACMP 5.4.1 26

D

daemons
snmpd 187
sysinfod 187, 191
trapd 190
trapgend 187, 190
default file collections 121
defining
shared LVM components 103
concurrent access 110
tty device 76, 93
deleting automatic error methods 127
destination nodes, importing volume group 107
discovering volume group configuration 107
disk reserve 152
disks
adapters
checking SCSI installation 88
installing SCSI 89
discovery 107
ghost 157
OEM 151
supported by HACMP 153
distribution
policy during migration 27
documentation 14
dynamic node priority
migration of 5.1 fallback policy 30

E

Eclock command 131
editing
/etc/hosts file 119
/usr/es/sbin/cluster/etc/clinfo.rc script 71
clinfo.rc on client 71
emulating error log entries 129
enabling
target mode SSA interface 96
enhanced concurrent volume groups
mksysb 49
enhanced security mode 29
Enterprise Storage System (ESS)
automatic error notification 125, 128
Subsystem Device Driver 125
error
emulation 129
notification 124
automatic 125
for volume group failures 124
for volume group loss 127
Estart command 131

F

fallback
policy for Two-Node Cluster Configuration Assistant 134
fallover for resource groups
policy for Two-Node Cluster Configuration Assistant 134
fibre tape drive, installing 100
file collection
default file collections 121
naming 123
files
/etc/cluster/conraid.dat 159
/etc/cluster/disktype.lst 159
/etc/cluster/lunreset.lst 159
/etc/filesystems 106
/etc/hosts 119, 136
/etc/rc.net 179
/etc/resolv.conf 119
/lpp/cluster/objrepos/HACMPnpp 42
/lpp/cluster/objrepos/HACMPude 42
/usr/es/sbin/cluster/etc/harc.net 120
clhosts 70
hacmp.my 187
target mode 94, 96
filesets
ibm2105.rte 88
forced stop 38
forced stop, or stop with "unmanage resource groups" 38

G

ghost disks 157
GPFS
/etc/rc.net script 179
adding file system using HACMP SMIT 181
cluster topology changes 182
configuring with HACMP SMIT 179
defining cluster topology 180
disk requirements 178
installing 179
network requirements 178
removing
GPFS cluster using HACMP SMIT 184
GPFS file system using HACMP SMIT 184
tuning cluster components 179

H

HACMP
configuring 19, 66
installing 47
uninstalling 35, 61
verifying installation of software 42, 60
HACMP 5.2
upgrading to HACMP 5.4.1
rolling migration 32
HACMP 5.3
upgrading to HACMP 5.4.1
snapshot 34
HACMP Configuration Database
converting for upgrade 46
restoring 45
security changes 27
HACMP installing HACMP
Concurrent Resource Manager, notes on installing 47
HACMP Smart Assists
WebSphere, DB2 or Oracle 14
hacmp.my source file 187
HACMP_Files
default file collection 121
HACMPgpfs ODM 185
hard disk, installing HACMP from 56
hardware
checking installation 73
HAView installation notes 50
high water mark, setting 117
HPS switch, upgrading to SP switch 78

I

I/O pacing 118
IBM SCSI 91
ibm2105.rte fileset 88
importing non-concurrent volume groups 107
importvg command, with shared LVM components 112
installation overview 47
installing
GPFS 179
HACMP Concurrent Resource Manager 47
shared fibre tape drive 100
shared tape drives 98
SSA serial disk subsystem 97
WebSMIT 40
installing HACMP
Concurrent Resource Manager
entry in /etc/inittab 51
from
a hard disk 56
installation medium 58
installation server 55
installing 19, 66
overview 47
prerequisites 47
problem resolution 61
using a saved snapshot 35
interface cards. See NIC.
IP address
aliasing
for monitoring with Tivoli 144, 148
IP labels/addresses
Two-Node Cluster Configuration Assistant 136
IPAT
IPAT via IP Replacement 74

J

jfslog
renaming 105

K

Kerberos security 29

L

LAN Emulation
ATM protocol 79
configuring in AIX for ATM 84
LANG variable setting for F1 help 51
lock manager (not supported after HACMP 5.1) 28
log logical volume, renaming 105
logging for Two-Node Cluster Configuration Assistant 139
Logical Device Name
fibre tape drives 101
SCSI tape drives 100
logical volumes
adding copies 106
as shared LVM component 15
renaming 105
low-water mark 117
LVM
mirroring 104

LVM_SA_QUORCLOSE 128

M

message catalogs, setting LANG variable 51
MIB
and SNMP 187
MIB-2 (Internet standard MIB) 187
migration
ODM security changes 27
PSSP File Collections issue 27
reversing 46
supported upgrade paths 23
troubleshooting 45
mirroring
jfslog 106
logical partitions 106
shared LVM components 104
super strict disk allocation policy 16
mksysb
disabling restore, during an upgrade of AIX 49
modifying
/etc/rc.net or use with GPFS 179
mounting NFS 16
Multi-Initiator RAID adapters for TMSSA 95

N

nameserving
configuration 119
naming file collection 123
NetView
daemons 187
traps 190
Network Interface Card. See NIC.
networks
network manager, NetView for AIX 190
options for customizing AIX 118
types 25
NFS
mount options 16
nested mount points 16
NFS Version 4
configuring HACMP to use 17
NIC
configuring in AIX 73
interface name 74
NIS
and cron 119
nodes
changing node number for TMSSA 95
nondisruptive upgrade 23
notification methods, testing 130

O

OEM disks, customizing 159
OEM file systems, customizing 172
OEM volume groups, customizing 165

P

permissions
on HACMP ODM files 27
planning
shared LVM components 15
logical volumes 15
polling with Clinfo 188
prerequisites
Two-Node Cluster Configuration Assistant 135
protocol, SNMP 187
PSSP File Collections
migration issue 27

R

RAID devices
importing volume groups 113
shared LVM components 104
concurrent access 110
recovery
SP Switch failures 131
removing
GPFS cluster using HACMP SMIT 184
GPFS file system using HACMP SMIT 184
HACMP 35, 61
previous version of HACMP 61
renaming
log logical volume 105
logical volumes 105
resetting HACMP tunable values 41
resource groups
distribution policy during migration 27
migrating 26
ring configuration for RS232 serial line 75
rotating resource groups
migrating to HACMP 5.4.1 26
routerevalidate 119
routerevalidate network option, changing 119
RS232 serial lines
checking connections 74
configuring 75
testing 75, 77
RSCT
installation images 53

S

scripts
/etc/rc.net 179
/usr/es/sbin/cluster/etc/clinfo.rc 71
clinfo.rc 71
SCSI devices
installing shared disks 88, 89
installing tape drive 98
status of 92

security
ODM changes that may affect upgrading 27
serial networks
configuring 74
testing
RS232 line 75, 77
target mode SCSI 94
service IP labels 74
setting
Clinfo scripts 69
I/O pacing 118
shared LVM components
creating file systems 105
creating volume groups 104
defining volume groups 103
logical volumes 15
planning 15
volume groups
concurrent access 110
Shared SCSI disks 89
Smart Assists
WebSphere, DB2 or Oracle 14
snapshot conversion
upgrade HACMP using snapshot 24
SNMP
utilities 187
version compatibility for upgrade of AIX 49
snmpd daemon 187
source node, creating shared volume groups 104
SP Switch
checking configuration 78
enabling ARP 130
specifying AIX error notification 131
SSA
disks
verifying installation 97
stalled upgrade
troubleshooting 36
startup policy
Two-Node Cluster Configuration Assistant 134
stopping cluster services with moving groups into unmanaged state 38
Subsystem Device Driver, ESS 125, 128
super strict
disk allocation for logical volumes 15
supported OEM disks 151
syncd, setting frequency for flushing buffers 118
sysinfod daemon 187, 191
Systems Monitor for AIX 191
daemon 187

T

tape drives
confirming installation 101
installing fibre 100
installing SCSI 98
target mode files 94, 96
target mode SCSI
configuring serial network 92
enabling interface 93
testing connections 94
target mode SSA
configuring connections 95
enabling interfaces 96
testing connections 96
TCP/IP standard MIB 187
testing
RS232 serial line 75
SCSI target mode connection 94
serial connection 77
target mode SSA connection 96
Tivoli, cluster monitoring
defining managed nodes 146
IPAT considerations 148
overview of installation instructions 145
required Tivoli software 51, 146
subnet requirements 145
uninstalling 149
trapd daemon 190
trapgend daemon 187, 190
traps
and SMUX peer daemons 191
Clinfo 188
troubleshooting
after an upgrade to HACMP 5.4.1 45
tty device
defining 76
tunable values
resetting 41
tuning GPFS cluster components 179
two-node cluster 133
planning 136
Two-Node Cluster Configuration Assistant 133
application configuration 135
logging 139
prerequisites 135
privileges to run 138
running 138
service IP labels/addresses 136
SMIT interface 139
standalone application 139
start and stop scripts 135
volume groups 135
types of networks 25

U

uninstalling
HACMP 35, 61

upgrading
commands restricted 31
current HACMP cluster to HACMP 5.4.1 21
on an offline cluster 37
terminology 22
upgrading HACMP
concurrent access migration issues 28
user privileges
Two-Node Cluster Configuration Assistant 138

V

verifying
GPFS cluster configuration 182
install of HACMP software 42, 60
Veritas 162
disk group 162
disk groups 162
volumes and file systems in an HACMP cluster 151
volumes, accommodating in an HACMP cluster 164
volume groups
changing startup status 110
creating for concurrent access 111
dormant at startup 115
importing
automatically 108
concurrent access 112
non-concurrent access 107

W

WebSMIT
install and configure 40, 62