

Bull

HACMP 4.4

Administration Guide

AIX



Bull

HACMP 4.4

Administration Guide

AIX

Software

August 2000

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

**ORDER REFERENCE
86 A2 57KX 02**

The following copyright notice protects this book under the Copyright laws of the United States of America and other countries which prohibit such actions as, but not limited to, copying, distributing, modifying, and making derivative works.

Copyright © Bull S.A. 1992, 2000

Printed in France

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

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

Trademarks and Acknowledgements

We acknowledge the right of proprietors of trademarks mentioned in this book.

AIX[®] is a registered trademark of International Business Machines Corporation, and is being used under licence.

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

Year 2000

The product documented in this manual is Year 2000 Ready.

Contents

About This Guide xiii

Chapter 1 **Maintaining an HACMP for AIX System** **1-1**

Overview	1-1
Common Administrative Tasks	1-2
Applying Software Maintenance to an HACMP Cluster ..	1-2
Starting and Stopping Cluster Services	1-2
Monitoring the Cluster	1-3
Maintaining Shared Logical Volume Manager Components	1-3
Maintaining Shared LVM Components in a Concurrent Access Environment	1-3
Changing the Cluster Topology	1-3
Changing Cluster Resources and Resource Groups	1-4
Verifying Cluster Topology and Cluster Resources	1-4
Maintaining Cluster Event Processing	1-4
Additional HACMP for AIX Maintenance Tasks	1-4
Saving and Restoring Cluster Configurations	1-4
Managing Users and Groups in a Cluster	1-5
Related Administrative Tasks	1-5
HACMP for AIX Software	1-6
Daemons	1-6
AIX Files Modified by HACMP for AIX	1-8
Scripts	1-10

Chapter 2 **Starting and Stopping Cluster Services** **2-1**

Overview	2-1
Cluster Services	2-1
Understanding Cluster Service Startup	2-2
Understanding Stopping Cluster Services	2-4
Starting Cluster Services	2-7
Starting Cluster Services on a Single Node	2-7
Starting Cluster Services in C-SPOC Clusters	2-9
Stopping Cluster Services	2-11
Stopping Cluster Services on a Single Cluster Node	2-11
Stopping Cluster Services in a C-SPOC Cluster	2-13
Starting and Stopping Cluster Services on Clients	2-15

Starting Clinfo on a Client	2-15
Stopping Clinfo on a Client	2-15
Maintaining Cluster Information Services	2-16
Maintaining the clhosts File on a Client	2-16
Maintaining the clhosts File on a Cluster Node	2-16
Enabling Clinfo for Asynchronous Event Notification ...	2-16

Chapter 3

Monitoring an HACMP Cluster 3-1

Periodically Monitoring an HACMP Cluster	3-1
Tools for Monitoring an HACMP Cluster	3-1
Monitoring Cluster, Node, and Network Interface Status	3-2
The /usr/sbin/cluster/clstat Utility ASCII Display	3-2
The /usr/sbin/cluster/clstat Utility X Window System Display	3-5
Monitoring a Cluster With HAView	3-7
HAView Installation Considerations	3-8
HAView File Modification Considerations	3-8
NetView Hostname Requirements for HAView	3-9
Starting HAView	3-9
Viewing Clusters and Components	3-11
Obtaining Component Details	3-13
Polling Intervals	3-13
Removing a Cluster from HAView Monitoring	3-14
Using the HAView Cluster Administration Utility	3-15
HAView Browsers	3-16
Monitoring HACMP Clusters with Tivoli	3-18
Consult Your Tivoli Documentation	3-19
Prerequisites and Considerations for Cluster Monitoring with Tivoli	3-19
Installing and Configuring Cluster Monitoring with Tivoli	3-19
Using Tivoli to Monitor the Cluster	3-20
Polling Intervals	3-26
Deinstalling Cluster Monitoring with Tivoli	3-27
Monitoring Cluster Services	3-28
Monitoring Cluster Services on a Node	3-28
Monitoring Cluster Services on a Client	3-29
HACMP for AIX Log Files	3-29
/usr/adm/cluster.log	3-29
/tmp/hacmp.out	3-29
/usr/sbin/cluster/history/cluster.mmdd	3-29
System Error Log	3-29
/tmp/cm.log	3-30
/tmp/cspoc.log	3-30
/tmp/emuhacmp.out	3-30

Chapter 4	Maintaining Shared LVM Components	4-1
	Overview	4-1
	Common Maintenance Tasks	4-1
	Using the C-SPOC Utility to Maintain Shared LVM Components	4-2
	Updating LVM Components in an HACMP Cluster	4-4
	Maintaining Shared Volume Groups	4-4
	TaskGuide for Creating Shared Volume Groups	4-4
	Using AIX Commands to Maintain Shared Volume Groups	4-6
	Using C-SPOC to Maintain Shared Volume Groups	4-13
	Creating a Shared Volume Group with C-SPOC	4-13
	Maintaining Shared Filesystems	4-20
	Creating a Shared Filesystem Using AIX Commands	4-20
	Creating a Shared Filesystem Using C-SPOC	4-23
	Changing a Shared Filesystem in a Cluster	4-25
	Removing a Shared Filesystem	4-27
	Maintaining Logical Volumes	4-31
	Creating a Logical Volume	4-31
	Adding a Logical Volume to a Cluster Using C-SPOC ...	4-33
	Changing the Characteristics of a Shared Logical Volume	4-34
	Removing a Logical Volume	4-37
	Adding Copies (Mirrors) to a Shared Logical Volume ...	4-39
	Setting Characteristics of a Shared Logical Volume Using C-SPOC	4-41
	Maintaining Physical Volumes	4-45
	Adding a Disk Definition to Cluster Nodes Using C-SPOC	4-45
	Removing a Disk Definition on Cluster Nodes Using C-SPOC	4-46
	Migrating Data	4-47
	Lazy Update Processing in an HACMP Cluster	4-49
	Forcing an Update Before Fallover	4-50
 Chapter 5	 Maintaining Shared LVM Components in a Concurrent Access Environment	 5-1
	Overview	5-1
	Concurrent Access and HACMP for AIX Scripts	5-2
	Nodes Join the Cluster	5-2
	Nodes Leave the Cluster	5-2
	Maintaining Concurrent Access Volume Groups	5-2
	Creating a Concurrent Volume Group	5-3
	Activating a Volume Group in Concurrent Access Mode .	5-9
	Determining the Access Mode of a Volume Group	5-10

Replacing a Failed Drive in a Concurrent Access Volume Group	5-11
Maintaining Concurrent Access During a Cluster Upgrade	5-15
Restarting the Concurrent Access Daemon (clvmd)	5-16
Maintaining Concurrent LVM Components Using C-SPOC ..	5-16
Using C-SPOC to Maintain Concurrent Volume Groups ..	5-17
Using C-SPOC to Maintain Concurrent Logical Volumes	5-24
Setting Characteristics of a Concurrent Logical Volume Using C-SPOC	5-26

Chapter 6 Changing the Cluster Topology 6-1

Overview	6-1
Synchronizing Configuration Changes	6-1
Reconfiguring a Cluster Dynamically	6-1
Requirements Before Reconfiguring	6-2
Viewing the Cluster Topology	6-2
Options in SMIT Under Show Cluster Topology	6-3
Changing a Cluster Name or ID	6-4
Changing the Configuration of Cluster Nodes	6-4
Adding a Cluster Node	6-4
Removing a Cluster Node	6-6
Changing the Name of a Cluster Node	6-8
Changing the Configuration of Network Adapters	6-9
Adding a Network Adapter	6-9
Removing a Network Adapter from a Cluster Node	6-12
Changing Network Adapter Attributes	6-12
Swapping a Network Adapter Dynamically	6-13
Changing the Configuration of a Network Module	6-14
Changing the Attributes of a Network Module	6-15
Adding a Network Module	6-17
Removing a Network Module	6-18
Synchronizing the Cluster Topology	6-18
Releasing a Dynamic Reconfiguration Lock	6-19
Processing ODM Data During Dynamic Reconfiguration ..	6-20
Undoing a Dynamic Reconfiguration	6-21
Restoring the ODM Data in the DCD	6-21

Chapter 7 Changing Resources and Resource Groups 7-1

Reconfiguration and Resource Migration—Overview	7-1
Requirements before Reconfiguring Cluster Resources ...	7-1
Cluster Resource Changes You Can Make Dynamically ..	7-2
Emulating Dynamic Reconfiguration Events	7-3
Reconfiguring Application Servers	7-3
Defining an Application Server	7-4

- Removing an Application Server 7-4
- Changing an Application Server 7-5
- Reconfiguring CS/AIX Communication Links 7-6
 - Defining a CS/AIX Communications Link 7-6
 - Removing a CS/AIX Communications Link 7-7
 - Changing a CS/AIX Communications Link 7-8
- Reconfiguring Resource Groups and Resources 7-9
 - Adding a Resource Group 7-9
 - Removing a Resource Group 7-10
 - Changing Resource Groups and Node Relationships 7-10
 - Adding or Removing Resources 7-12
 - Using SMIT to Reconfigure Resources 7-13
- Migrating Resources Dynamically—Overview 7-17
 - 7-17
 - 7-17
 - Resource Migration Types 7-18
- Migrating Resources Dynamically From the Command Line . 7-19
 - Moving or Starting a Resource Group 7-20
 - Stopping a Resource Group 7-20
 - Starting a Resource Group with the Default Keyword 7-21
 - Example: Using cldare to Swap Resource Groups 7-21
- Migrating Resources Dynamically Using SMIT 7-22
 - Bringing a Resource Group Online 7-23
 - Bringing a Resource Group Offline 7-24
 - Moving a Resource Group 7-25
- Checking Resource Group State and Sticky Markers 7-26
 - Checking Resource Group State with the clfindres
Command 7-26
 - Removing Sticky Markers 7-26
- Synchronizing Cluster Resources 7-27
 - Skipping Cluster Verification During Synchronization ... 7-29

Chapter 8

Verifying a Cluster Configuration 8-1

- Overview 8-1
 - Three Ways to Run Cluster Verification 8-1
- The clverify Utility 8-2
 - Specifying Environmental Options for clverify 8-3
- Verifying Software 8-3
- Verifying the Cluster 8-4
- Verifying the Cluster Configuration Using SMIT 8-7
 - Verifying the Cluster using SMIT 8-7
 - Adding a Custom Verification Method 8-8
 - Changing a Custom Verification Method 8-8
 - Removing a Custom Verification Method 8-9

Chapter 9	Maintaining and Emulating Cluster Events	9-1
	Overview	9-1
	Customizing Event Processing	9-1
	Pre- and Post-Event Processing	9-1
	Event Notification	9-1
	Event Recovery and Retry	9-2
	Configuring Custom Cluster Events	9-2
	Adding Customized Cluster Events	9-2
	Changing/Showing Custom Cluster Events	9-2
	Removing Custom Cluster Events	9-3
	Changing Pre- or Post-Event Processing	9-3
	Emulating Cluster Events	9-5
Chapter 10	Additional Tasks: Configuring Run-Time Parameters, Cluster Log Files, and Maintaining NFS	10-1
	Changing Run-Time Parameters	10-1
	Customizing Cluster Log Files	10-2
	Maintaining NFS	10-2
	Creating Shared Volume Groups	10-3
	NFS Mounting and Fallover	10-4
	Reliable NFS Server Capability	10-7
Chapter 11	Saving and Restoring Cluster Configurations	11-1
	Overview	11-1
	What Information Is Saved in a Cluster Snapshot	11-1
	Format of a Cluster Snapshot	11-3
	clconvert_snapshot Utility	11-4
	Defining a Custom Snapshot Method	11-4
	Changing or Removing a Custom Snapshot Method	11-5
	Creating (Adding) a Cluster Snapshot Using SMIT	11-5
	Applying a Cluster Snapshot Using SMIT	11-6
	Undoing an Applied Snapshot	11-8
	Changing a Cluster Snapshot	11-8
	Removing a Cluster Snapshot	11-9
Chapter 12	Managing Users and Groups in a Cluster	12-1
	Overview	12-1
	Managing User Accounts	12-1
	Listing Users On All Cluster Nodes	12-1
	Adding User Accounts on all Cluster Nodes	12-2
	Changing Attributes of Users in a Cluster	12-3
	Removing Users from a Cluster	12-4
	Managing Group Accounts	12-5

	Listing Groups on All Cluster Nodes	12-5
	Adding Groups on Cluster Nodes	12-6
	Changing Characteristics of Groups in a Cluster	12-6
	Removing Groups from the Cluster	12-7
Appendix A	HACMP for AIX Commands	A-1
Appendix B	Script Utilities	B-1
Appendix C	Command Execution Language Guide	C-1
Appendix D	VSM Graphical Configuration Application	D-1
Appendix E	7x24 Maintenance	E-1
	Processor ID Licensing Issues	E-13
Index		X-1

Contents

About This Guide

Administering an HACMP for AIX system involves making *planned* changes to the system after it has been installed and configured. This guide provides information about the tasks you perform to maintain an HACMP for AIX system. System management tasks include modifying a shared volume group, reconfiguring a cluster, and changing the node environment.

Who Should Use This Guide

This guide is intended for the system administrator responsible for maintaining an HACMP for AIX system. To maintain an HACMP for AIX system, you must understand the components that make up the system. Accordingly, this guide assumes that you understand:

- HACMP for AIX software and concepts
- Communications, including the TCP/IP subsystem
- The AIX operating system, including the Logical Volume Manager subsystem
- The hardware and software installed at your site.

Before You Begin

As a prerequisite, read the following documents to familiarize yourself with the facilities and capabilities of the HACMP for AIX software: *HACMP for AIX, Version 4.4 Concepts and Facilities*, *HACMP for AIX, Version 4.4 Planning Guide*, and *HACMP for AIX, Version 4.4 Installation Guide*.

How To Use This Guide

Overview of Contents

This guide contains the following chapters and appendixes:

- Chapter 1, *Maintaining an HACMP for AIX System*, presents an overview of the various tasks involved in maintaining an HACMP for AIX system. Each task is described more fully in a subsequent chapter. The chapter also describes the software components of an HACMP for AIX system.
- Chapter 2, *Starting and Stopping Cluster Services*, describes how to start and stop the HACMP for AIX software on nodes and clients.
- Chapter 3, *Monitoring an HACMP Cluster*, describes how to determine the status of an HACMP cluster.
- Chapter 4, *Maintaining Shared LVM Components*, describes how to maintain volume groups shared by cluster nodes within an HACMP for AIX system.
- Chapter 5, *Maintaining Shared LVM Components in a Concurrent Access Environment*, discusses special considerations for maintaining LVM components in a concurrent access configuration.

- Chapter 6, Changing the Cluster Topology, describes how to change a cluster definition after its initial configuration.
- Chapter 7, Changing Resources and Resource Groups, describes how to change the configuration of cluster resources, such as application servers and resource groups.
- Chapter 8, Verifying a Cluster Configuration, describes how to verify that the cluster is properly configured.
- Chapter 9, Maintaining and Emulating Cluster Events, describes how to define pre- and post-event processing scripts and customize cluster event processing.
- Chapter 10, Additional Tasks: Configuring Run-Time Parameters, Cluster Log Files, and Maintaining NFS, describes how to set values for parameters that control the debug output level for HACMP for AIX scripts, how HACMP interacts with NIS, how to maintain the cluster configuration files used by the cluster verification utility, and how to ensure that NFS works properly on an HACMP cluster.
- Chapter 11, Saving and Restoring Cluster Configurations, describes how to use the cluster snapshot utility to save and restore cluster configurations.
- Chapter 12, Managing Users and Groups in a Cluster, describes how to create and manage user and group accounts in clusters using the C-SPOC utility.
- Appendix A, HACMP for AIX Commands, lists syntax diagrams and examples for common HACMP for AIX informational commands.
- Appendix B, Script Utilities, describes the utilities called by the event and startup scripts supplied with the HACMP for AIX software.
- Appendix C, Command Execution Language Guide, describes how to use the Command Execution Language (CEL) to create commands that act on all nodes in clusters using the C-SPOC utility. All C-SPOC commands are created using this language.
- Appendix D, VSM Graphical Configuration Application, describes the **xhacmpm** graphical interface tool for configuring cluster environments and components.
- Appendix E, 7x24 Maintenance, is a checklist for maintenance procedures that should help you keep a cluster running as close to a 7 X 24 environment as possible.

Highlighting

The following highlighting conventions are used in this guide:

<i>Italic</i>	Identifies variables in command syntax, new terms and concepts, or indicates emphasis.
Bold	Identifies pathnames, commands, subroutines, keywords, files, structures, directories, and other items whose names are predefined by the system. Also identifies graphical objects such as buttons, labels, and icons that the user selects.
Monospace	Identifies examples of specific data values, examples of text similar to what you might see displayed, examples of program code similar to what you might write as a programmer, messages from the system, or information that you should actually type.

ISO 9000

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

Related Publications

The following publications provide additional information about HACMP for AIX:

- *Release Notes* in `/usr/lpp/cluster/doc/release_notes` describe hardware and software requirements
- *HACMP for AIX, Version 4.4: Concepts and Facilities*, order number 86 A2 54KX 02
- *HACMP for AIX, Version 4.4: Planning Guide*, order number 86 A2 55KX 02
- *HACMP for AIX, Version 4.4: Installation Guide*, order number 86 A2 56KX 02
- *HACMP for AIX, Version 4.4: Troubleshooting Guide*, order number 86 A2 58KX 02
- *HACMP for AIX, Version 4.4: Programming Locking Applications*, order number 86 A2 59KX 02
- *HACMP for AIX, Version 4.4: Programming Client Applications*, order number 86 A2 60KX 02
- *HACMP for AIX, Version 4.4: Master Index and Glossary*, order number 86 A2 65KX 02
- *HACMP for AIX, Version 4.4: Enhanced Scalability Installation and Administration Guide, Vol.1*, order number 86 A2 62KX 02
- *HACMP for AIX, Version 4.4: Enhanced Scalability Installation and Administration Guide, Vol.2*, order number 86 A2 89KX 01.

The AIX document set, as well as manuals accompanying machine and disk hardware, also provide relevant information.

Ordering Publications

To order additional copies of this guide, use order number 86 A2 57KX 02.

Chapter 1 Maintaining an HACMP for AIX System

This chapter provides an overview of the tasks you perform to maintain an HACMP for AIX system. Each task is described in detail in a subsequent chapter. The chapter also describes the components of the software distributed with the HACMP for AIX, Version 4.4.

Overview

The HACMP for AIX system provides a highly available environment that ensures that critical applications at your site are available to end users. As system administrator, your job is to make sure that HACMP for AIX itself is stable and operational.

Before undertaking any management task, you should have read carefully the following documents to familiarize yourself with the facilities and capabilities of the HACMP for AIX system: *HACMP for AIX Concepts and Facilities*, *HACMP for AIX Planning Guide*, and *HACMP for AIX Installation Guide*. Even if HACMP for AIX has already been installed at your site, these books contain essential information you must understand in order to maintain an HACMP for AIX system. For a description of (or for help with) common HACMP for AIX problems, see the chapter Solving Common Problems in the *HACMP for AIX Troubleshooting Guide*.

As the system administrator of a highly available environment, you should strive to minimize cluster down time. Whenever possible, use the HACMP for AIX dynamic reconfiguration capability to change an active cluster without requiring down time. For those changes that require the cluster be shut down, schedule the down time for off hours when there is minimal activity on the system. When you plan to take the system down, notify end users in advance and inform them when services will be restored.

Another HACMP for AIX administrative utility that can make cluster administration easier is the Cluster-Single Point of Control (C-SPOC) utility. The C-SPOC utility allows you to perform some administrative tasks on all cluster nodes by executing distributive command scripts which execute across the cluster. For an overview of C-SPOC, see Chapter 6 in the *HACMP for AIX Concepts and Facilities* guide.

Finally, remember that planning is the key to successfully administering an HACMP for AIX system because you must consider the effects of changes on the highly available applications and their use.

Common Administrative Tasks

The following sections briefly describe the types of administrative tasks you may be called upon to perform while maintaining an HACMP for AIX environment.

Applying Software Maintenance to an HACMP Cluster

You can install software maintenance, called Program Temporary Fixes (PTFs), to your HACMP cluster will running HACMP for AIX cluster services on cluster nodes; however, you must stop cluster services on the node on which you are applying a PTF.

To apply software maintenance to your HACMP cluster:

1. Use the **smit clstop** fastpath to stop cluster services on the node on which the PTF is to be applied. If you would like the resources provided by this node to remain available to users, stop cluster with takeover so that the takeover node will continue to provide these resources to users.
2. Apply the software maintenance to this node using the procedure described in the *AIX Installation Guide* and in any documentation distributed with the PTF.
3. Run the **/usr/sbin/cluster/diag/clverify** utility to ensure that no errors exist after installing the PTF. See Chapter 8, Verifying a Cluster Configuration for information about using the **clverify** utility.

Note: If cluster services are active on any cluster node and service adapters are configured with their service versus boot addresses (IPAT is enabled), a message appears indicating that the node is not on its boot address. This message does not reflect a problem with the cluster configuration; instead, it indicates that cluster services are running on the specified node.

4. Reboot the node to reload any HACMP for AIX kernel extensions that may have changed as a result of the PTF being applied.
If an update to the cluster.base.client.lib file set has been applied and you are using Cluster Lock Manager or Clinfo API functions, you may need to relink your applications.
5. Restart the HACMP for AIX software on the node using the **smit clstart** fastpath and verify that the node successfully joined the cluster.
6. Repeat Steps 1 through 5 on the remaining cluster nodes.

Starting and Stopping Cluster Services

Starting and stopping cluster services requires a thorough understanding of how nodes interact and of the tools available to stop and start HACMP for AIX. You must also consider the impact that stopping HACMP for AIX has on end users. Chapter 2, Starting and Stopping Cluster Services describes how to start and stop HACMP for AIX on server nodes and client nodes. This chapter also describes how to use the C-SPOC utility to start or stop cluster services in clusters.

Monitoring the Cluster

By design, HACMP for AIX masks from end users various failures that occur within a cluster. For example, HACMP for AIX masks a network adapter failure by swapping in a standby adapter. As a result, it is possible that a component in the cluster can fail and that you can be unaware that a failure has occurred. The danger here is that, while HACMP for AIX can survive one or possibly several failures, an unnoticed failure threatens a cluster's ability to handle future failures.

To avoid this situation, it is recommended that you customize your HACMP for AIX system by adding event notification to the scripts designated to handle the various cluster events. You can also use the AIX Error Notification facility to provide warnings about hardware errors that do not cause HACMP for AIX events. The Automatic Error Notification function automatically turns Error Notification on or off on all nodes in the cluster for particular devices. (See the *HACMP for AIX Installation Guide*, chapter 16 for more information on the Automatic Error Notification function.) Moreover, as a general practice, you should periodically inspect your cluster to make sure that it works properly. Chapter 3, *Monitoring an HACMP Cluster*, describes how to check the status of an HACMP cluster, the nodes within that cluster, and the daemons that run on the nodes.

Although the combination of HACMP for AIX and the inherent high availability features built into the AIX system keeps single points of failure to a minimum, there are still failures that, although detected, are not handled in any useful way. See the *HACMP for AIX Installation Guide* for suggestions on customizing error notification.

Maintaining Shared Logical Volume Manager Components

Maintaining the volume groups, logical volumes, and filesystems shared by cluster nodes requires that you keep the ODM definitions of these components synchronized across all nodes in the cluster. To do so, you must propagate configuration changes on one node to other affected nodes in the cluster. Chapter 4, *Maintaining Shared LVM Components*, describes a set of procedures for maintaining shared LVM components.

This chapter also describes how to use the C-SPOC utility to manage LVM components in clusters.

Maintaining Shared LVM Components in a Concurrent Access Environment

While maintaining the shared LVM components in a concurrent access environment is similar to a non-concurrent access environment, there are some additional considerations. Chapter 5, *Maintaining Shared LVM Components in a Concurrent Access Environment*, describes these special requirements.

Changing the Cluster Topology

Over time, the composition of a cluster can change. For example, you might add an additional network or remove an adapter. If you add a new component to a cluster, or remove an existing component, you must reconfigure HACMP for AIX to update the cluster configuration. Chapter 6, *Changing the Cluster Topology*, describes how to modify a cluster topology after its initial configuration.

This chapter also describes which cluster topology reconfigurations can be performed dynamically and which changes require you to stop and restart cluster services to make the changed configuration the currently active configuration.

Changing Cluster Resources and Resource Groups

In addition to the cluster topology, the configuration of the cluster nodes and their relationships to cluster resources may also change. As with cluster topology changes, when you change the configuration of a cluster node, you must update all other cluster nodes. Chapter 7, *Changing Resources and Resource Groups*, describes how to perform these tasks. This chapter also includes information about changing application servers.

This chapter also describes which cluster resource reconfigurations can be performed dynamically and which changes require you to stop and restart cluster services to make the changed configuration the currently active configuration.

Verifying Cluster Topology and Cluster Resources

If you change your cluster topology or the configuration of a cluster node, you should perform the cluster verification procedure to make sure that all nodes agree on the cluster topology, network configuration, and ownership of resources. Chapter 8, *Verifying a Cluster Configuration*, describes this procedure and also how the `/usr/sbin/cluster/diag/clverify` utility works.

Maintaining Cluster Event Processing

To configure cluster events, you indicate the script that handles the event and any additional customized processing that should accompany an event. If you add, delete, or change scripts related to the processing of a cluster event, you must update each node in the cluster to reflect the changes. Chapter 9, *Maintaining and Emulating Cluster Events*, describes how to manage cluster event processing.

Additional HACMP for AIX Maintenance Tasks

Additional tasks that you must perform to maintain an HACMP for AIX system include updating the configuration files used by the `/usr/sbin/cluster/diag/clverify` utility, configuring NFS files, and changing the security level of the cluster. It also describes how to set the debug output level of HACMP for AIX scripts, and how to control HACMP's interaction with NIS

Saving and Restoring Cluster Configurations

After you configure the topology and resources of a cluster, you can use the cluster snapshot utility to save in a file the ODM information that defines the cluster. This saved configuration can later be used to restore the configuration. A cluster snapshot can also be applied to an active cluster to dynamically reconfigure the cluster. Chapter 11, *Saving and Restoring Cluster Configurations*, describes these tasks.

Managing Users and Groups in a Cluster

As system administrator of an HACMP for AIX cluster, you may be called upon to create user accounts and groups on the nodes of your cluster. This chapter describes how to use the C-SPOC utility to manage users and groups in clusters. Using C-SPOC, you can create, change, or remove users and groups from all cluster nodes by executing a C-SPOC command on any single cluster node. Chapter 12, Managing Users and Groups in a Cluster, describes these tasks.

Related Administrative Tasks

The tasks below, while not specifically discussed in this book, are essential for effective system administration.

Backing Up Your System

The practice of allocating multiple copies of a logical volume can enhance high availability in a cluster environment, but it should not be considered a replacement for regular system backups. Although HACMP for AIX is designed to survive failures within the cluster, it cannot survive a catastrophic failure where multiple points of failure leave data on disks unavailable. Therefore, it is imperative that you back up your system on a regular basis. You must have a backup procedure in place to ensure data reliability and to protect against catastrophic physical volume failure.

To maintain your HACMP for AIX environment, you must back up the root volume group (which contains the HACMP for AIX software) and the shared volume groups (which contain the data for highly available applications) regularly. HACMP for AIX is like other AIX environments from this perspective. Back up all nodes.

Documenting Your System

As your HACMP for AIX system grows and changes, it differs from its initial cluster configuration. It is your responsibility as system administrator to document all aspects of the HACMP for AIX system unique to your environment. This responsibility includes documenting procedures concerning the highly available applications, recording changes that you make to the configuration scripts distributed with HACMP for AIX, documenting any custom scripts you write, recording the status of backups, maintaining a log of user problems, and maintaining records of all hardware.

Maintaining Highly Available Applications

As system administrator, you should understand the relationship between your applications and HACMP for AIX. Highly available applications are started and stopped by HACMP for AIX in response to cluster events. Understanding when, how, and why this happens is critical to keeping a cluster highly available, as problems can occur that require corrective actions.

Helping Users

As the resident HACMP for AIX expert, you can expect to receive many questions from end users at your site about HACMP for AIX. Again, the more you know about HACMP for AIX, the better you are able to answer these questions. If you cannot answer questions about your HACMP for AIX cluster environment, refer to the *HACMP for AIX Troubleshooting Guide* for more information, or contact your IBM support representative.

HACMP for AIX Software

This section describes the software distributed with the HACMP for AIX software. The HACMP for AIX software includes the following components:

- Daemons
- Files
- Scripts
- Utilities

Daemons

The HACMP for AIX software includes the following daemons:

- Cluster Manager daemon
- Cluster SMUX Peer daemon
- Cluster Information Program daemon
- Cluster Lock Manager daemon

Cluster Manager Daemon

The Cluster Manager (`/usr/sbin/cluster/clstrmgr`) daemon is the core process in the HACMP for AIX system. The Cluster Manager, which runs on each node in the cluster, is essentially a monitoring and notification process. The Cluster Manager monitors cluster objects—node, network interface, and network—for changes in status. When a change occurs, the Cluster Manager generates an event that triggers the execution of an event script.

HACMP for AIX provides a set of predefined events and subevents. Major events are listed below.

node_up
node_up_complete
node_down
node_down_complete
swap_adapter
swap_adapter_complete
network_up
network_up_complete
network_down
network_down_complete

reconfig_resource_acquire
reconfig_resource_release
reconfig_resource_complete
reconfig_topology_start
reconfig_topology_complete

See the *HACMP for AIX Concepts and Facilities Guide* or the **clstrmgr** man page for additional information on the Cluster Manager.

Cluster SMUX Peer Daemon

The Cluster SMUX Peer (**/usr/sbin/cluster/clsmuxpd**) daemon runs on each node in a cluster and provides SNMP support to client applications through **clinfo**. The Cluster SMUX Peer maintains cluster status information in a Management Information Base (MIB). You can use standard SNMP routines to access the cluster information in this MIB. Note that the **/usr/sbin/cluster/clsmuxpd** daemon requires that the **snmpd** daemon also be running on the node.

See the *HACMP for AIX Concepts and Facilities Guide* or the **clsmuxpd** man page for additional information.

Cluster Information Program Daemon

The Cluster Information (**/usr/sbin/cluster/clinfo**) daemon is an SNMP monitor. It can run on RS/6000 clients and on cluster nodes. The **clinfo** daemon queries the Cluster SMUX Peer to receive an updated topology map of the cluster and provides an API for clients to access that data. When the **clinfo** daemon detects a change in the cluster status, it calls the **/usr/sbin/cluster/etc/clinfo.rc** script to respond to that change.

Note: You must maintain the **/usr/sbin/cluster/etc/clhosts** file in order for **clinfo** to operate properly. Also **clinfo** must be stopped and re-started after making changes to your cluster configuration.

See the *HACMP for AIX Concepts and Facilities Guide* or the **clinfo** man page for additional information. Chapter 2, Starting and Stopping Cluster Services, contains information on maintaining the **/usr/sbin/cluster/etc/clhosts** file.

The **clinfo** daemon also provides C and C++ APIs that enable you to develop “cluster-aware” applications that recognize and respond to changes within a cluster. See *HACMP for AIX Programming Client Applications* for more information.

Cluster Lock Manager Daemon

The Cluster Lock Manager (**/usr/sbin/cluster/cllockd**) daemon runs on each node in a cluster configured for concurrent access. It provides an API that applications can use to create a single, unified lock image that is shared among all nodes in the cluster. Cooperating applications running on different nodes in an HACMP for AIX cluster can then share common resources without corrupting those resources. Locking is required in a concurrent access configuration where multiple nodes are simultaneously accessing a shared external disk.

See *HACMP for AIX Programming Locking Applications* and the **cllockd** man page for more information.

AIX Files Modified by HACMP for AIX

The following AIX files are modified to support HACMP for AIX. They are not distributed with HACMP for AIX.

/.rhosts

On each cluster node, you must list the service and boot adapters for each cluster node, and include the following loopback entry: 127.0.0.1. The loopback entry is required to allow for the proper execution of C-SPOC commands initiated on the local node. These entries allow the Global ODM (**godm**) daemon to run so you can configure the cluster nodes from a central location. They also allows the **clruncmd** command to run.

If you have security concerns, add these entries when necessary to use the functionality noted, and remove them afterwards. These entries, however, are required to use C-SPOC commands in a cluster environment. The **/usr/sbin/cluster/clstrmgr** daemon, however, does not depend on **/.rhosts** file entries.

The **/.rhosts** file is not required on SP systems running HACMP for AIX Enhanced Security. See Chapter 6 of the *HACMP for AIX Concepts and Facilities Guide* for a brief description of this feature.

/etc/hosts

The cluster event scripts use the **/etc/hosts** file for name resolution. All cluster node IP interfaces must be added to this file on each node. Note that DNS and NIS are disabled during HACMP-related name resolution. This is why HACMP for AIX IP addresses must be maintained locally.

/etc/inetd.conf

The **/etc/inetd.conf** file is the Internet server configuration file. The install process adds the following entry for the HACMP for AIX Global ODM socket (**godm**) to this file:

```
godm stream tcp nowait root /usr/sbin/cluster/godmd
```

/etc/inittab

The **/etc/inittab** file is modified when HACMP for AIX is configured for IP address takeover or the **Start at system restart** option is chosen on the **SMIT Start Cluster Services** screen.

IP Address Takeover

The following entry is added to the **/etc/inittab** file for HACMP for AIX network startup with IP address takeover.

```
harc:2:wait:/usr/sbin/cluster/etc/harc.net # HACMP network startup
```

When IP address takeover is enabled, the system edits **/etc/inittab** to change the **rc.tcpip** and **inet**-dependent entries from run level 2 (the default multi-user level) to run level a. Entries that have run level a are processed only when the **telinit** command is executed specifying that specific run level.

System Boot

The `/etc/inittab` file is used by the `init` process to control the startup of processes at boot time. The following entry is added to the `/etc/inittab` file if the **Start at system restart** option is chosen on the **SMIT Start Cluster Services** screen:

```
hacmp:2:wait:/usr/sbin/etc/rc.cluster -boot> /dev/console 2>&1 # Bring
up Cluster
```

When the system boots, the `/etc/inittab` file calls the `/usr/sbin/cluster/etc/rc.cluster` script to start HACMP for AIX.

Because the `inet` daemons must not be started until after HACMP-controlled adapters have swapped to their service address, HACMP for AIX also adds the following entry to the end of the `/etc/inittab` file to indicate that `/etc/inittab` processing has completed.

```
clinit:a:wait:touch /usr/sbin/cluster/.telinit # HACMP This must be last
entry in inittab!
```

See Chapter 2, *Starting and Stopping Cluster Services*, for more information about the files involved in starting and stopping HACMP for AIX.

`/etc/rc.net`

The `/etc/rc.net` file is called by `cfgmgr` to configure and start TCP/IP during the boot process. It sets hostname, default gateway and static routes. The following entry is added at the beginning of the file for a node on which IP address takeover is enabled:

```
# HACMP for AIX
# HACMP for AIX These lines added by HACMP for AIX software
[ "$1" = "-boot" ] && shift || { ifconfig lo0 127.0.0.1 up; exit 0; }
#HACMP for AIX
# HACMP for AIX
```

The entry prevents `cfgmgr` from reconfiguring boot and service addresses while HACMP for AIX is running.

`/etc/services`

The `/etc/services` file defines the sockets and protocols used for network services on a system. The ports and protocols used by the HACMP for AIX components are defined here.

```
#clinfo_deadman      6176/tcp
#clm_keepalive       6255/udp
#cllockd             6100/udp
#clm_pts             6200/tcp
#clsmuxpd            6270/tcp
#clm_lkm              6150/tcp
#clm_smux            6175/tcp
#godm                 6177/tcp
```

`/etc/snmpd.conf`

The SNMP daemon reads the `/etc/snmpd.conf` configuration file when it starts up and when a refresh or **kill -1** signal is issued. This file specifies the community names and associated access privileges and views, hosts for trap notification, logging attributes, `snmpd`-specific parameter configurations, and SMUX configurations for the `snmpd`. The HACMP for AIX installation process adds the `clsmuxpd` password to this file. The following entry is added to the end of the file, to include the HACMP for AIX MIB managed by the `clsmuxpd`.

```
smux 1.3.6.1.4.1.2.3.1.2.1.5 "clsmuxpd_password" # HACMP clsmuxpd
```

/etc/snmpd.peers

The **/etc/snmpd.peers** file configures **snmpd** SMUX peers. The HACMP for AIX install process adds the following entry to include the **clsmuxpd**.

```
clsmuxpd 1.3.6.1.4.1.2.3.1.2.1.5 "clsmuxpd_password" # HACMP clsmuxpd
```

/etc/syslog.conf

The **/etc/syslog.conf** file is used to control output of the **syslogd** daemon, which logs system messages. During the install, process HACMP for AIX adds entries to this file that direct the output from HACMP for AIX-related problems to certain files.

```
# example:
# "mail messages, at debug or higher, go to Log file. File must exist."
# "all facilities, at debug and higher, go to console"
# "all facilities, at crit or higher, go to all users"
# mail.debug          /usr/spool/mqueue/syslog
# *.debug             /dev/console
# *.crit              *
# HACMP for AIX Critical Messages from HACMP for AIX
local0.crit /dev/console
# HACMP for AIX Informational Messages from HACMP for AIX
local0.info /usr/adm/cluster.log
# HACMP for AIX Messages from Cluster Scripts
user.notice /usr/adm/cluster.log
```

The **/etc/syslog.conf** file should be identical on all cluster nodes.

/etc/tcp.clean

The **/etc/tcp.clean** file stops TCP/IP daemons, brings down network interfaces, and removes lock files. During the install, HACMP for AIX adds the **clstop** command to this file.

/etc/trcfmt

The **/etc/trcfmt** file is the template file for the system trace logging and report utility, **trcrpt**. The install process adds HACMP for AIX tracing to the trace format file. HACMP for AIX tracing applies to the following daemons: **clstrmgr**, **clinfo**, **clockd**, and **clsmuxpd**.

/var/spool/cron/crontab/root

The **/var/spool/cron/crontab/root** file contains commands needed for basic system control. The install process adds HACMP for AIX logfile rotation to the file.

Scripts

The following scripts are supplied with the HACMP for AIX software.

Startup and Shutdown Scripts

Each of the following scripts is involved in starting and stopping the HACMP for AIX software. See Chapter 2, Starting and Stopping Cluster Services for more information about these scripts.

/usr/sbin/cluster/utilities/clstart

The **/usr/sbin/cluster/utilities/clstart** script, which is called by the **/usr/sbin/cluster/etc/rc.cluster** script, invokes the AIX System Resource Controller (SRC) facility to start the cluster daemons. The **clstart** script starts HACMP for AIX with the options currently specified on the Start Cluster Services screen. See the **clstart** man page for additional information.

/usr/sbin/cluster/utilities/clstop

The **/usr/sbin/cluster/utilities/clstop** script, which is called by the SMIT Stop Cluster Services screen, invokes the SRC facility to stop the cluster daemons with the options specified on the **Stop Cluster Services** screen. See the **clstop** man page for additional information.

There is a corresponding C-SPOC version of this script that stops cluster services on each cluster node. The **/usr/sbin/cluster/utilities/cl_clstop** script calls the HACMP for AIX **clstop** script. See the **cl_clstop** man page for additional information.

/usr/sbin/cluster/utilities/clexit.rc

If the SRC detects that the **clstrmgr** daemon has exited abnormally, it executes the **/usr/sbin/cluster/utilities/clexit.rc** script to halt the system. If the SRC detects that any other HACMP for AIX daemon has exited abnormally, it executes the **clexit.rc** script to stop these processes, but does not halt the node. See the **clexit.rc** man page for additional information.

/usr/sbin/cluster/etc/rc.cluster

If the **Start at system restart** option is chosen on the Start Cluster Services screen, the **/usr/sbin/cluster/etc/rc.cluster** script is called by the **/etc/inittab** file to start HACMP for AIX. The **/usr/sbin/cluster/etc/rc.cluster** script does some necessary initialization and then calls the **/usr/sbin/cluster/utilities/clstart** script to start HACMP for AIX.

The **/usr/sbin/cluster/etc/rc.cluster** script is also used to start the **clinfo** daemon on a client. See the **rc.cluster** man page for additional information.

There is a corresponding C-SPOC version of this script that starts cluster services on each cluster node. The **/usr/sbin/cluster/utilities/cl_rc.cluster** script calls the HACMP for AIX **rc.cluster** script. See the **cl_rc.cluster** man page for additional information.

/etc/rc.net

The **/etc/rc.net** script is called by the **/usr/sbin/cluster/etc/rc.cluster** script to configure and start the TCP/IP interfaces and to set the required network options. The **/etc/rc.net** script is used in the boot process to retrieve interface information from the ODM and to configure all defined interfaces. If IP address takeover is configured, the **/etc/rc.net** script is called from the **/usr/sbin/cluster/etc/rc.cluster** script at cluster startup instead of during the boot process.

Event Scripts

The following event scripts described below are called by the cluster daemons to respond to cluster events. Chapter 9, Maintaining and Emulating Cluster Events, in this book, and Chapter 14, Customizing Cluster Events, in the *HACMP for AIX Installation Guide* for detailed information about these scripts. See the man pages for additional information.

The `/usr/sbin/cluster/events/node_up` Scripts

The `/usr/sbin/cluster/events/node_up` scripts handle node attachment:

- `/usr/sbin/cluster/events/node_up_local`
- `/usr/sbin/cluster/events/node_up_remote`
- `/usr/sbin/cluster/events/node_up_local_complete`
- `/usr/sbin/cluster/events/node_up_remote_complete`.

The `/usr/sbin/cluster/events/node_down` Scripts

The `/usr/sbin/cluster/events/node_down` scripts handle node detachment:

- `/usr/sbin/cluster/events/node_down_local`
- `/usr/sbin/cluster/events/node_down_remote`
- `/usr/sbin/cluster/events/node_down_local_complete`
- `/usr/sbin/cluster/events/node_down_remote_complete`.

The `/usr/sbin/cluster/events/swap_adapters` Script

The `/usr/sbin/cluster/events/swap_adapters` script swaps the service and standby adapters.

The `/usr/sbin/cluster/events/network` Scripts

The `/usr/sbin/cluster/events/network_up` and `/usr/sbin/cluster/events/network_down` scripts handle network failure and recovery. These scripts take no default actions; they must be customized for the site.

The `/usr/sbin/cluster/events/reconfig` Scripts

The `/usr/sbin/cluster/events/reconfig` scripts handle dynamic reconfiguration:

- `/usr/sbin/cluster/events/reconfig_resource_acquire`
- `/usr/sbin/cluster/events/reconfig_resource_release`
- `/usr/sbin/cluster/events/reconfig_resource_complete`
- `/usr/sbin/cluster/events/reconfig_topology_start`
- `/usr/sbin/cluster/events/reconfig_topology_complete`

The `/usr/sbin/cluster/etc/clinfo.rc` Script

The `/usr/sbin/cluster/etc/clinfo.rc` script, which is invoked by the `clinfo` daemon whenever a network or node event occurs, updates the system's ARP cache. You can customize this script for additional processing. There must be a copy of the `/usr/sbin/cluster/etc/clinfo.rc` script on each node in the cluster. See the `clinfo.rc` man page for additional information.

Chapter 2 Starting and Stopping Cluster Services

This chapter explains how to start and stop cluster services on cluster nodes and clients. The chapter also describes how to use the Cluster-Single Point of Control (C-SPOC) utility to start and stop cluster services in clusters environments.

Overview

Starting cluster services refers to the process of starting the HACMP for AIX daemons that enable the coordination required between nodes in a cluster. Starting cluster services on a node also triggers the execution of certain HACMP for AIX scripts that create the cluster. Stopping cluster services refers to stopping these same daemons on a node and may or may not cause the execution of additional HACMP for AIX scripts, depending on the type of shutdown you perform.

Before starting or stopping cluster services, you must have a thorough understanding of the node interactions it causes and the impact on your system's availability. The following sections define the HACMP for AIX cluster services and briefly describe the processing involved in startup or shutdown of these services. Later sections in this chapter describe the step-by-step procedures you must follow to start or stop cluster services on a node. These sections include information about starting or stopping cluster services on several cluster nodes using the C-SPOC utility. The final section of this chapter describes how to start and stop cluster services on HACMP for AIX clients.

Warning: Do not start cluster services on a node that has been dynamically removed from a cluster. When you remove a node from the cluster, the cluster definition remains in the node's ODM. If you start cluster services on the removed node, the node reads this cluster configuration data and attempts to rejoin the cluster from which it had been removed. See the *HACMP for AIX Troubleshooting Guide* for more information.

Cluster Services

The following lists the required and optional HACMP for AIX daemons:

Cluster Manager daemon (clstrmgr)

This daemon maintains the heartbeat protocol between the nodes in the cluster, monitors the status of the nodes and their interfaces, and invokes the appropriate scripts in response to node or network events. All cluster nodes must run the **clstrmgr** daemon.

Cluster SMUX Peer daemon (clsmuxpd)

This daemon maintains status information about cluster objects. This daemon works in conjunction with the Simple Network Management Protocol (**snmpd**) daemon. All cluster nodes must run the **clsmuxpd** daemon.

Note: The **clsmuxpd** daemon cannot be started unless the **snmpd** daemon is running.

Cluster Lock Manager daemon (cllockd)

This daemon provides advisory locking services. The **cllockd** daemon is required on cluster nodes only if those nodes are part of a concurrent access configuration.

Cluster Information Program daemon (clinfo)

This daemon provides status information about the cluster to cluster nodes and clients and invokes the `/usr/sbin/cluster/etc/clinfo.rc` script in response to a cluster event. The **clinfo** daemon is optional on cluster nodes and clients.

The HACMP for AIX daemons are controlled by the AIX System Resource Controller (SRC). The SRC provides a consistent interface for starting, stopping, and monitoring processes by grouping sets of related programs into subsystems and groups. In addition, the SRC provides facilities for logging of abnormal terminations of subsystems or groups and for tracing of one or more subsystems. (For information about tracing HACMP for AIX subsystems, see Appendix B, “HACMP for AIX Tracing,” in the *HACMP for AIX Troubleshooting Guide*.)

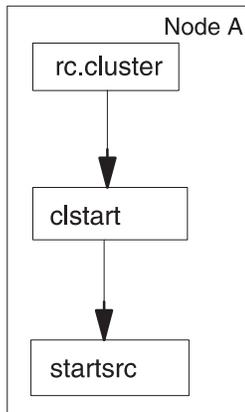
The HACMP for AIX daemons are collected into the following SRC subsystems and groups:

Daemon	Subsystem	Group
<code>/usr/sbin/cluster/clstrmgr</code>	clstrmgr	cluster
<code>/usr/sbin/cluster/clinfo</code>	clinfo	cluster
<code>/usr/sbin/cluster/clsmuxpd</code>	clsmuxpd	cluster
<code>/usr/sbin/cluster/cllockd</code>	cllockd	lock

When using the SRC commands, you can control the **clstrmgr**, **clinfo**, and **clsmuxpd** daemons by specifying the SRC **cluster** group.

Understanding Cluster Service Startup

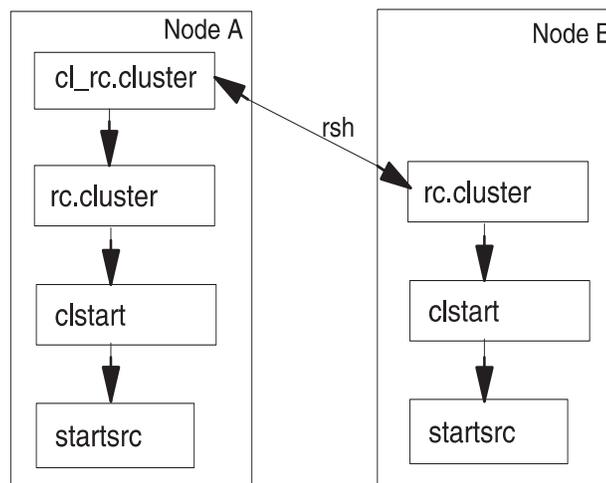
You start cluster services on a node by executing the HACMP for AIX `/usr/sbin/cluster/etc/rc.cluster` script. Use the HACMP for AIX “Start Cluster Services” SMIT screen, described in the section “Starting Cluster Services,” to build and execute this command. The **rc.cluster** script initializes the environment required for HACMP for AIX by setting environment variables and then calls the `/usr/sbin/cluster/utilities/clstart` script to start the HACMP for AIX daemons. The **clstart** script is the HACMP for AIX script that starts all the cluster services. The **clstart** script calls the SRC **startsrc** command to start the specified subsystem or group. The following figure illustrates the major commands and scripts called at cluster startup.



Flow of Commands Used in Cluster Startup

Using the C-SPOC utility, you can start cluster services on one or any number of nodes in a cluster by executing the C-SPOC `/usr/sbin/cluster/utilities/cl_rc.cluster` command on a single cluster node. The C-SPOC `cl_rc.cluster` command calls the `rc.cluster` command to start cluster services on the nodes specified from the originating node. The nodes are started in sequential order, not in parallel. Node startup is based on the node's priority in the resource group, or on the order in which they are listed on the command line. The output of the command runs on the local node and is returned to the originating node. Because the command is executed remotely, there can be a delay before the command output is returned.

The following example shows the major commands and scripts executed on all cluster nodes when cluster services are started in clusters using the C-SPOC utility.



Flow of Commands Used at Cluster Startup by C-SPOC Utility

Note: After the C-SPOC command executes on Node A and the node is stable, the **rsh** command is executed on the next node in sequence.

Automatically Restarting Cluster Services

You can optionally have cluster services start whenever the system is rebooted. If you specify the **-R** flag to the **rc.cluster** command, or specify “restart or both” in the Start Cluster Services SMIT screen, the **rc.cluster** script adds the following line to the **/etc/inittab** file.

```
hacmp:2:wait:/usr/sbin/cluster/etc/rc.cluster -boot> /dev/console 2>&1  
# Bring up Cluster
```

At system boot, this entry causes AIX to execute the **/usr/sbin/cluster/etc/rc.cluster** script to start HACMP for AIX.

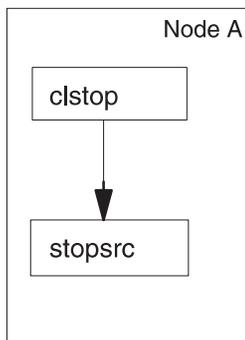
Note: Be aware that if the cluster services are set to restart automatically at boot time, you may face problems with node integration after a power failure and restoration, or you may want to test a node after doing maintenance work before having it rejoin the cluster.

Starting Cluster Services with IP Address Takeover Enabled

If IP address takeover is enabled, the **/etc/rc.net** script is called by the **/usr/sbin/cluster/etc/rc.cluster** script to configure and start the TCP/IP interfaces and to set the required network options.

Understanding Stopping Cluster Services

You stop cluster services on a node by executing the HACMP for AIX **/usr/sbin/cluster/etc/clstop** script. Use the HACMP for AIX Stop Cluster Services SMIT screen, described in the section “Stopping Cluster Services,” to build and execute this command. The **clstop** script stops an HACMP for AIX daemon or daemons. The **clstop** script starts all the cluster services or individual cluster services by calling the SRC command **stopsrc**. The following figure illustrates the major commands and scripts called at cluster shutdown.

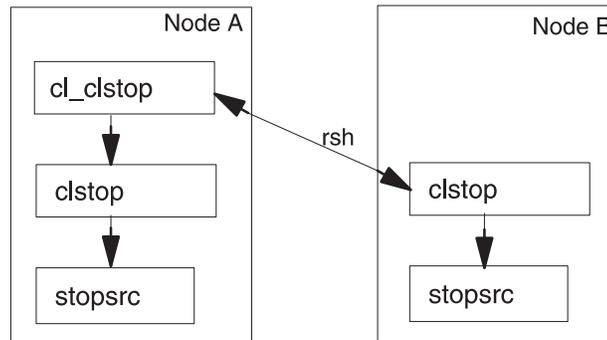


Flow of Commands Used at Cluster Shutdown

Using the C-SPOC utility, you can stop cluster services on a single node, on several nodes, or on all nodes in a cluster by executing the C-SPOC **/usr/sbin/cluster/utilities/cl_clstop** command on a single node. The C-SPOC **cl_clstop** command performs some cluster-wide verification and then calls the **clstop** command to stop cluster services on the specified nodes.

The nodes are stopped in sequential order—not in parallel. The output of the command run on the remote node is returned to the originating node. Because the command is executed remotely, there can be a delay before the command output is returned.

The following figure provides an illustration of how commands and scripts are executed when cluster services are stopped on a single-node in a cluster using the C-SPOC utility. All nodes in the cluster are stopped accordingly, if they are specified from the node originating the command.



Flow of Commands Used at Cluster Shutdown

Note: After the C-SPOC command executes on Node A and the node is stable, the **rsh** command is executed on the next node in sequence.

When to Stop Cluster Services

You typically stop cluster services:

- Before making any hardware or software changes or other scheduled node shutdowns or reboots. Failing to do so may cause unintended cluster events to be triggered on other nodes.
- Before certain reconfiguration activity. Some changes to the cluster information stored in the ODM require stopping and restarting the cluster services on *all* nodes for the changes to become active. For example, if you wish to change the name of the cluster, the name of a node, or the name of an adapter, you must stop and restart cluster services.

For more information about which changes to the cluster require reconfiguration, see Chapter 6, Changing the Cluster Topology and Chapter 7, Changing Resources and Resource Groups.

Types of Stops

When you stop cluster services, you must also decide how to handle the resources that were owned by the node you are removing from the cluster. You have the following options:

- Graceful** In a *graceful* stop, the HACMP for AIX software shuts down its applications and releases its resources. The other nodes *do not* take over the resources of the stopped node.
- Graceful with Takeover** In a *graceful with takeover* stop, the HACMP for AIX software shuts down its applications and releases its resources. The surviving nodes take over these resources. This is also called *intentional fallover*.
- Forced** In a *forced* stop, the HACMP for AIX software shuts down immediately. No HACMP for AIX scripts are run. The other nodes view it as a graceful shutdown and, therefore, do not take over any resources. The node that is shutting down retains control of all its resources. Applications that require no HACMP for AIX daemons continue to run. Typically, you use the forced option so that users and clients are not interrupted by stopping the Cluster Manager.
- Note:** If applications are stopped manually after a *forced* stop, the relevant entry in the `/usr/sbin/cluster/server.status` file, which lists all applications running on the node, must be removed prior to reintegrating. Otherwise, the application entry is not deleted from the `server.status` file. Consequently, HACMP will not restart the applications, since an application entry in the file usually means that the application is already running.

Note: When the AIX operating system is stopped with the **shutdown** command, HACMP is stopped gracefully, without takeover.

Abnormal Termination of a Cluster Daemon

If the SRC detects that any HACMP for AIX daemon has exited abnormally (without being shut down using the **clstop** command), it executes the `/usr/sbin/cluster/utilities/clexit.rc` script to halt the system. This prevents unpredictable behavior from corrupting the data on the shared disks. The **clexit.rc** script stops all processes in the lock group when the **cllockd** daemon exits abnormally. See the **clexit.rc** man page for additional information.

Warning: Never use the **kill -9** command on the **clstrmgr** daemon. Using the **kill** command causes the **clstrmgr** daemon to exit abnormally. This causes the SRC to run the `/usr/sbin/cluster/utilities/clexit.rc` script which halts the system immediately, causing the surviving nodes to initiate fallover.

Starting Cluster Services

The following sections describe how to start cluster services on a single node or on several cluster nodes in a C-SPOC cluster.

Note: If a node has a **tty** console, it must be powered on for HACMP for AIX to be started on that node. If this is not desirable, find the line in the `/usr/sbin/cluster/etc/rc.cluster` script that ends with `2>/dev/console` and change this line to reflect whatever behavior is desired. For example, you can redirect output to another **tty**. Be aware that this redirects the startup messages on the node.

Starting Cluster Services on a Single Node

The steps below describe the procedure for starting cluster services on a cluster node. If you have defined cascading or rotating resources, start the nodes in the logical order for the correct distribution of those resources.

Note: All cluster nodes should be running. For cold boots, let one node fully boot before powering up the other nodes. Starting cluster services also depends on the cluster resource configuration and on the order in which nodes should be started.

To start cluster services on one node, several nodes, or on all cluster nodes:

1. Enter:

```
smit clstart
```

SMIT displays the Start Cluster Services screen.

2. Enter field values as follows:

Start now, on system restart or both Indicate whether you want the **clstrmgr** and **clsmuxpd** daemons to start when you commit the values on this screen (**now**), when the operating system reboots (**restart**), or on **both** occasions. Choosing **on system restart** or **both** means that the cluster services are always brought up automatically after a system reboot.

BROADCAST message at startup? Indicate whether you want to send a broadcast message to users when the cluster services start.

Startup Cluster Lock Services? Indicate whether you want to start the **clockd** daemon. If you have a concurrent access configuration or if your application uses the Cluster Lock Manager, set this field to **yes**. Otherwise, set it to **no**.

**Startup Cluster
Information Daemon?**

Indicate whether you want to start the **clinfo** daemon. If your application uses Clinfo or if you use the **clstat** monitor, set this field to **yes**. Otherwise, set it to **no**.

The values that you enter in the **Startup Cluster Lock Services?** and **Startup Cluster Information Services?** fields work in conjunction with the value you enter in the **Start now, on system restart or both** field. If you set either (or both) of the startup fields to **yes** and the **Start now, on system restart or both** field to **both**, then these daemons are also started whenever the **clstrmgr** and **clsmuxpd** daemons are started.

3. Press Enter. The system starts the cluster services, activating the cluster configuration that you have defined, if you specified that you wanted cluster services started immediately.

After a few seconds, a message appears on the console indicating that the **/usr/sbin/cluster/events/node_up** script has begun. The time that it takes this script to run depends on your configuration (that is, the number of disks, the number of interfaces to configure, the type of configuration, the number of filesystems to mount, and the number of applications being started). When the **node_up** script completes, a message appears on the console indicating that the **/usr/sbin/cluster/events/node_up_complete** is running.

When the scripts complete, HACMP for AIX is up and running on the selected node. SMIT displays a command status window.

If cluster services fail to start, check the **cluster.log** or **hacmp.out** log files for error messages. Consult Appendix A in the *HACMP for AIX Troubleshooting Guide* for an explanation of what caused the problem and suggested actions to correct the problem.

4. Ensure that the HACMP-controlled applications are available.

Access the applications controlled by HACMP for AIX to verify that they behave as expected.

Repeat this procedure on the other cluster nodes. To verify that the nodes are up, use the **/usr/sbin/cluster/clstat** utility, described in Chapter 3, Monitoring an HACMP Cluster.

Reintegrating Nodes

A node reintegrating into a cluster is just like a node starting up except that the Cluster Manager is already running on other cluster nodes. The cluster's fallback configuration determines whether the node joining the cluster takes back shared resources other nodes took over during failover. If the resources are defined as rotating, the node joining the cluster (the original owner) does not take back the shared resources.

When a node that owns a resource group is forced down, the resource group remains with that node. If a second node that participates in the resource group joins the cluster while the owning node is down, the second node attempts to acquire the resource group, causing possible resource contention.

To avoid this situation, integrate the owning node before any other nodes, if possible. If this is not possible, release the resources owned by the downed node before integrating a different node.

Note: See the **Forced** entry under Types of Stops on page 2-6 for reintegration considerations regarding the `/usr/sbin/cluster/server.status` file.

Ensuring that Network Daemons Start as Part of HACMP Start-Up

Clusters that have both rotating and cascading resource groups, where the cascading resource group does not have IP address takeover configured (that is, the cascading resources are not tied to a service label), can experience problems starting the network daemons. At cluster startup, the first node to join the cluster acquires the rotating address and runs the **telinit -a** command. The second node to join the cluster, since it does not migrate from a boot to a service address, does not run the **telinit -a** command. As a result, the network-related daemons (NFS and NIS for example) do not get started. If your cluster requires this combination of resource groups, customize the pre- and post-event processing to issue the **telinit -a** command.

Starting Cluster Services in C-SPOC Clusters

The steps below describe the procedure for starting cluster services on a single node or on any number of nodes by executing the C-SPOC `/usr/sbin/cluster/utilities/cl_rc.cluster` command on one of the cluster nodes.

Note: All cluster nodes should be running. For cold boots, let one node fully boot before powering up the other nodes. Starting cluster services also depends on the cluster resource configuration and on the order in which nodes should be started.

As the root user, perform the following steps to start the cluster services on a node.

5. Enter:

```
smit cl_admin
```

SMIT displays the **Cluster System Management** screen.

6. Select the **HACMP for AIX Cluster Services** option and press Enter.

Note: The C-SPOC SMIT screen for starting cluster services in a cluster is entitled **HACMP for AIX Cluster Services**. This screen lets you start cluster services on a single node or on N nodes, where N is the number of cluster nodes. Starting cluster services through the main HACMP for AIX SMIT interface only lets you start services on the local node.

7. Select the **Start Cluster Services** option and press Enter.

8. Enter field values as follows:

Start now, on system restart or both Indicate whether you want the **clstrmgr** and **clsmuxpd** daemons to start when you commit the values on this screen (**now**), when the operating system reboots (**restart**), or on **both** occasions. Choosing **on system restart** or **both** means that the cluster services are always brought up automatically after a system reboot.

Start on local node, remote node, or both Indicate whether you want to start cluster services on the local node, on a remote node, or on both using the C-SPOC utility. Press the F4 key for a list of options.

BROADCAST message at startup? Indicate whether you want to send a broadcast message to users when the cluster services start. The messages appear on all the nodes being started.

Startup Cluster Lock Services? Indicate whether you want to start the **clockd** daemon. If you have a concurrent access configuration or if your application uses the Cluster Lock Manager, set this field to **yes**. Otherwise, set it to **no**.

Startup Cluster Information Daemon? Indicate whether you want to start the **clinfo** daemon. If your application uses Clinfo or if you use the **clstat** monitor, set this field to **yes**. Otherwise, set it to **no**.

The values that you enter in the **Startup Cluster Lock Services?** and **Startup Cluster Information Services?** fields work in conjunction with the value you enter in the **Start now, on system restart or both** field. If you set either (or both) of the startup fields to **yes** and the **Start now, on system restart or both** field to **both**, then these daemons are also started whenever the **clstrmgr** and **clsmuxpd** daemons are started.

9. Press Enter. The system starts the cluster services on the nodes specified, activating the cluster configuration that you have defined. The time that it takes the commands and scripts to run depends on your configuration (that is, the number of disks, the number of interfaces to configure, the number of filesystems to mount, and the number of applications being started).

When command execution completes and HACMP for AIX cluster services are started on all nodes specified, SMIT displays a command status window.

If cluster services fail to start on any cluster node, check the C-SPOC utility log file, named **/tmp/cspoc.log**, for error messages. This file contains the command execution status of the C-SPOC command executed on each cluster node.

The following example illustrates the status returned in the C-SPOC log file:

```
06/26/96 12:04:38 [===== C_SPOC COMMAND LINE =====]
06/26/96 12:04:39 /usr/sbin/cluster/sbin/cl_lsuser -c -a id home ALL
06/26/96 12:04:48 abalone: success:
/usr/sbin/cluster/utilities/clresactive -c
clresactive:-V:4200::uname:-m
06/26/96 12:04:48 peri: success: /usr/sbin/cluster/utilities/clresactive
-c clr
esactive:-V:4200::uname:-m
06/26/96 12:04:54 abalone: success: lsuser -c -a id home ALL
```

```
06/26/96 12:04:54 peri: success: lsuser -c -a id home ALL
06/26/96 12:13:14 [===== C_SPOC COMMAND LINE =====]
06/26/96 12:13:14 /usr/sbin/cluster/sbin/cl_lsgroup -c ALL
06/26/96 12:13:21 abalone: success:
/usr/sbin/cluster/utilities/clresactive -c
clresactive:-V:4200::uname:-m
06/26/96 12:13:21 peri: success: /usr/sbin/cluster/utilities/clresactive
-c clr
esactive:-V:4200::uname:-m
06/26/96 12:13:27 abalone: success: lsgroup -c ALL
06/26/96 12:13:27 peri: success: lsgroup -c ALL
```

Consult Appendix A in the *HACMP for AIX Troubleshooting Guide* for an explanation of error messages and suggestions about actions to correct the problem.

Stopping Cluster Services

The following sections describe how to stop cluster services on a single cluster node, on several nodes, or on all cluster nodes using the C-SPOC utility.

Note: Minimize activity on the system. If the node you are stopping is currently providing highly available services, notify users of your intentions if their applications will be unavailable. Let them know when services will be restored.

Stopping Cluster Services on a Single Cluster Node

The steps below describe the procedure for stopping HACMP for AIX on a node. If all nodes need to be stopped, follow the same steps on all nodes, but stop them sequentially, not in parallel. Stopping cluster services also depends on the cluster resource configuration and on the order in which nodes should be stopped.

To stop cluster services on a single node, on several nodes, or on all cluster nodes:

1. Enter:

```
smit clstop
```

The **Stop Cluster Services** screen appears.

2. Enter field values as follows:

Stop now, on system restart or both Indicate whether you want the cluster services to stop **now**, at **restart** (when the operating system reboots), or on **both** occasions. If you choose **restart** or **both**, the cluster services will no longer come up automatically after a reboot.

BROADCAST cluster shutdown? Indicate whether you want to send a broadcast message to users before the cluster services stop.

Shutdown mode Indicate the type of shutdown:

- **graceful**—Shutdown after the `/usr/sbin/cluster/events/node_down_complete` script is run on this node to release its resources. The other nodes *do not* take over the resources of the stopped node.
- **graceful with takeover**—Shutdown after the `/usr/sbin/cluster/events/node_down_complete` script runs to release its resources.
- **forced**—Shutdown immediately. No `/usr/sbin/cluster/events/node_down` script is run on this node. This node retains control of all its resources. The forced shutdown is intended to be used when a cluster is unstable or is in reconfiguration. You can also use this option to bring down a node while you make a change to the cluster configuration, such as adding an adapter. The node's applications remain available (but without services of HACMP for AIX daemons). Refer to the section Reintegrating Nodes on page 2-8 for information on reintegrating a node that has been forced down back into the cluster.

3. Press Enter. HACMP for AIX stops the cluster services according to your specifications. SMIT displays a command status window.

Note: When you stop the cluster and restart it immediately, there is a two-minute delay before the `clstrmgr` daemon becomes active.

AIX Shutdown and Cluster Services

When the AIX operating system is stopped with the `shutdown` command, HACMP is stopped gracefully, without takeover.

Using the `stopsrc` Command to Stop an HACMP Daemon

You can also stop an HACMP daemon by using the standard AIX `stopsrc` command with the `-g` flag, specifying the cluster group:

```
stopsrc -g cluster
```

Note: Using the `stopsrc` command to stop an individual HACMP for AIX daemon (by specifying the `-s` flag with the name of the daemon) is not recommended and may cause unexpected behavior.

Stopping a Node Configured for IP Address Takeover

When you stop the Cluster Manager on an HACMP for AIX node with the forced option, no scripts are run. Applications and resources (disk and IP addresses) are unaffected.

If a network adapter fails and a standby adapter is configured as a service adapter, the `cfgmgr` attempts to reconfigure the failed adapter as the service adapter.

To prevent this reconfiguration from happening, the HACMP for AIX software places an identifier in the call to the `/etc/rc.net` file which causes it to exit immediately when called by `cfgmgr`, but to run to completion when called by the Cluster Manager.

For more information about the modifications to the `/etc/rc.net` file, see Appendix A of the *HACMP for AIX Installation Guide*.

Stopping Cluster Services in a C-SPOC Cluster

The steps below describe the procedure for stopping cluster services on a single node or on N nodes, where N is the number of cluster nodes, by executing the C-SPOC `/usr/sbin/cluster/utilities/cl_clstop` command on one of the cluster nodes.

When stopping several nodes, C-SPOC stops them sequentially, not in parallel. If any node specified to be stopped is inactive, the shutdown operation aborts. Stopping cluster services also depends on the cluster resource configuration and on the order in which nodes should be stopped.

To stop cluster services in a C-SPOC cluster:

1. Enter:

```
smit cl_admin
```

SMIT displays the HACMP for AIX Cluster Services screen.

Note: Do not confuse this C-SPOC **HACMP for AIX Cluster Services** screen with the **Cluster Services** screen available as an option on the main HACMP for AIX SMIT menu.

2. Select the **HACMP for AIX Cluster Services** option and press Enter.

Note: The C-SPOC SMIT screen for stopping cluster services in a cluster is entitled **HACMP for AIX Cluster Services**. This screen lets you stop cluster services on a single node or on N nodes, where N is the number of cluster nodes. Stopping cluster services through the main HACMP for AIX SMIT interface only lets you stop services on the local node.

3. Select the **Stop Cluster Services** option and press Enter.

4. Enter field values as follows:

Stop now, on system restart or both Indicate whether you want the cluster services to stop **now**, at **restart** (when the operating system reboots), or on **both** occasions. If you choose **restart** or **both**, the cluster services will no longer come up automatically after a reboot.

Stop on local node, remote node, or both Indicate whether you want to stop cluster services on the local node, on a remote node, or on both using the C-SPOC utility. Press the F4 key for a list of options.

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 each cluster node being stopped.
Shutdown mode	Indicate the type of shutdown:
graceful	Shutdown after the <code>/usr/sbin/cluster/events/node_down_complete</code> script is run on this node to release its resources. The other nodes <i>do not</i> take over the resources of the stopped node.
graceful with takeover	Shutdown after the <code>/usr/sbin/cluster/events/node_down_complete</code> script runs to release its resources. You cannot shut down cluster services on other cluster nodes if you select this shutdown mode.
forced	Shutdown immediately. No <code>/usr/sbin/cluster/events/node_down</code> script is run on the nodes. The nodes retain control of all their resources. The forced shutdown is intended to be used when a cluster is unstable or is in reconfiguration. You can also use this option to bring down a node while you make a change to the cluster configuration, such as adding an adapter. The node's applications remain available (but without services of HACMP for AIX daemons). Refer to the section Reintegrating Nodes on page 2-8 for information on reintegrating a node that has been forced down back into the cluster.

5. Press Enter. The system stops the cluster services on the nodes specified.

SMIT displays a command status window.

If the stop operation fails, check the C-SPOC utility log file, named `/tmp/cspoc.log`, for error messages. This file contains the command execution status of the C-SPOC command executed on each cluster node.

The following example illustrates the status returned in the C-SPOC log file:

```
06/26/96 12:04:38 [===== C_SPOC COMMAND LINE =====]
06/26/96 12:04:39 /usr/sbin/cluster/sbin/cl_lsuser -c -a id home ALL
06/26/96 12:04:48 abalone: success:
/usr/sbin/cluster/utilities/clresactive -c
clresactive:-V:4200::uname:-m
06/26/96 12:04:48 peri: success: /usr/sbin/cluster/utilities/clresactive
-c clr
esactive:-V:4200::uname:-m
06/26/96 12:04:54 abalone: success: lsuser -c -a id home ALL
06/26/96 12:04:54 peri: success: lsuser -c -a id home ALL
06/26/96 12:13:14 [===== C_SPOC COMMAND LINE =====]
06/26/96 12:13:14 /usr/sbin/cluster/sbin/cl_lsgroup -c ALL
```

```
06/26/96 12:13:21 abalone: success:
/usr/sbin/cluster/utilities/clresactive -c
clresactive:-V:4200::uname:-m
06/26/96 12:13:21 peri: success: /usr/sbin/cluster/utilities/clresactive
-c clr
esactive:-V:4200::uname:-m
06/26/96 12:13:27 abalone: success: lsgroup -c ALL
06/26/96 12:13:27 peri: success: lsgroup -c ALL
```

To interpret the messages that may appear in this file, consult the *HACMP for AIX Troubleshooting Guide* for an explanation of what caused the problem and for suggested actions to correct the problem.

Starting and Stopping Cluster Services on Clients

The cluster services on clients consist solely of the **clinfo** daemon, which provides clients with status information about the cluster. This daemon is under the control of the SRC. Although **clinfo** is under the control of the SRC, its abnormal exit does not cause a **halt -q** by default.

Note that the **/etc/inittab** file is modified when the HACMP for AIX software is installed to start the **clinfo** daemon whenever the system is rebooted.

Starting Clinfo on a Client

Use the **/usr/sbin/cluster/etc/rc.cluster** script or the **startsrc** command to start **clinfo** on a client, as shown below:

```
/usr/sbin/cluster/etc/rc.cluster
```

You can also use the standard AIX **startsrc** command:

```
startsrc -s clinfo
```

Stopping Clinfo on a Client

To stop **clinfo** on a client, use either the HACMP for AIX **clstop** command or the AIX **stopsrc** command.

Using the **clstop** Command

Enter the following command to stop **clinfo** on a client:

```
clstop -f -y
```

The **-f** argument specifies a forced shutdown. The option is needed for the command to execute, though it does not apply to client machines. The **-y** argument suppresses the confirmation prompt.

Using the **stopsrc** Command

You can also use the standard AIX **stopsrc** command. Enter either of the following commands to stop **clinfo** on a client machine:

```
stopsrc -g cluster
```

Note: Using the **stopsrc** command with the **-s** flag to stop an individual HACMP for AIX daemon is not recommended and may cause unexpected behavior.

Maintaining Cluster Information Services

In order for the **clinfo** daemon to get the information it needs, you must edit the **/usr/sbin/cluster/etc/clhosts** file. This file should contain IP service and boot labels or addresses of any HACMP for AIX servers to which **clinfo** may communicate, including servers from any clusters accessible through logical connections. The installed version of the **clhosts** file on an HACMP client node differs from that on an HACMP server node.

Maintaining the clhosts File on a Client

As installed, the **clhosts** file on an HACMP for AIX client node contains no hostnames or addresses. HACMP server addresses must be provided by the user at installation time. This file should contain all boot and service names or addresses of HACMP servers from any cluster accessible through logical connections with this client node. Upon startup, **clinfo** uses these names or addresses to attempt communication with a **clsmuxpd** process executing on an HACMP server.

An example list of hostnames/addresses in a **clhosts** file follows:

```
cowrie_en0_cl83 # cowrie service
140.186.91.189 # limpet service
floyd_en0_cl83 # floyd service
squid_en0_cl83 # squid service
```

Maintaining the clhosts File on a Cluster Node

As installed, the **/usr/sbin/cluster/etc/clhosts** file on an HACMP server node contains a loopback address. **clinfo** first attempts to communicate with a **clsmuxpd** process locally. If it succeeds, it then acquires an entire cluster map, including a list of all HACMP for AIX server interface addresses. From then on, it uses this list rather than the provided loopback address to recover from **clsmuxpd** communication loss.

However, if **clinfo** does not succeed in communicating with a **clsmuxpd** process locally, it continues indefinitely to attempt communication periodically only to the local address. For this reason, you should add all HACMP server boot and service addresses accessible through logical connections with this node, just as on an HACMP client node. The loopback address is provided only as a convenience.

Enabling Clinfo for Asynchronous Event Notification

By default, **clinfo** polls the **clsmuxpd** periodically for information. You can optionally choose to have **clinfo** receive notification of events as asynchronous messages (traps).

Note: Only one SNMP application can receive traps. If you are running NetView, for example, you cannot enable **clinfo** to receive traps. See Appendix E in the *HACMP for AIX Installation Guide* for more information.

To enable asynchronous event notification, use the following procedure.

1. Start **clinfo** with the **-a** option, by entering the following:

```
chssys -s clinfo -a "-a".
```

To verify that the SRC has the correct command line arguments for **clinfo**, enter the following:

```
lssrc -Ss clinfo | awk -F: '{print $3}'
```

2. Edit the **snmpd.conf** file on the nodes that will send traps. As installed, traps are directed to the loopback address (**clinfo** receives those traps generated by the **clsmuxpd** on the same node).

- a. Find the trap line at the end of the file. It looks like this:

```
view 1.17.2 system enterprises view
trap public 127.0.0.1 1.2.3 fe # loopback
```

- b. Add trap lines as desired. Multiple **clinfo** processes can receive traps from a single **clsmuxpd** process. Make sure that the "1.2.3 fe" field is unique.

An entry might look like the following example, with two more trap lines added:

```
trap public 127.0.0.1 1.2.3 fe #loopback
trap public 123.456.789.1 #adam
trap public 123.456.789.2 #eve
```

3. Stop and restart the **snmpd** and **clsmuxpd** processes on the hosts where you made the changes in the **snmpd.conf** file:

```
stopsrc -s clsmuxpd
stopsrc -s snmpd
startsrc -s snmpd
startsrc -s clsmuxpd
```

Starting and Stopping Cluster Services
Maintaining Cluster Information Services

Chapter 3 Monitoring an HACMP Cluster

This chapter describes tools you can use to monitor an HACMP cluster.

Note: The default locations of log files are used in this chapter. If you redirected any logs, check the appropriate location.

Periodically Monitoring an HACMP Cluster

By design, HACMP for AIX compensates for various failures that occur within a cluster. For example, HACMP for AIX compensates for a network adapter failure by swapping in a standby adapter. As a result, it is possible that a component in the cluster has failed and that you are unaware of the fact. The danger here is that, while HACMP for AIX can survive one or possibly several failures, *a failure that escapes your notice threatens a cluster's ability to provide a highly available environment.*

To avoid this situation, you should customize your system by adding event notification to the scripts designated to handle the various cluster events. You can specify a command that sends you mail indicating that an event is about to happen (or that an event has just occurred), along with information about the success or failure of the event. The mail notification system enhances the standard event notification methods.

Use the AIX Error Notification facility to add an additional layer of high availability to an HACMP for AIX environment. Although the combination of HACMP for AIX and the inherent high availability features built into the AIX system keeps single points of failure to a minimum, there are still failures that, although detected, are not handled as fully as you would like. See the chapter on supporting AIX Error Notification in the *HACMP for AIX Installation Guide* for suggestions on customizing error notification.

See the chapter on tailoring cluster event processing in the *HACMP for AIX Planning Guide* for detailed information on predefined events and on customizing event handling. Also, be sure to consult your worksheets, to document any changes you make to your system, and to periodically inspect the key cluster components to make sure they are in full working order.

Tools for Monitoring an HACMP Cluster

HACMP for AIX provides the following tools for monitoring an HACMP cluster:

- The `/usr/sbin/cluster/clstat` utility, which reports the status of key cluster components—the cluster itself, the nodes in the cluster, and the network adapters connected to the nodes.
- The HAView utility, which monitors HACMP clusters through the TME 10 NetView for AIX graphical network management interface. HAView enables you to monitor multiple HACMP clusters and cluster components across a network from a single node.
- Capability for cluster monitoring through a Tivoli management interface. You can use Tivoli to monitor the state of your cluster nodes and networks.
- The SMIT Show Cluster Services screen, which shows the status of the HACMP for AIX daemons

Monitoring an HACMP Cluster

Monitoring Cluster, Node, and Network Interface Status

- The following log files: the `/usr/adm/cluster.log` file, which tracks cluster events, the `/tmp/hacmp.out` file, which records the output generated by configuration scripts as they execute, the `/usr/sbin/cluster/history/clustermmdd` log file, which logs the daily cluster history, and the `/tmp/cspoc.log` file, which logs the status of C-SPOC commands executed on cluster nodes.

When you monitor a cluster, use the `clstat` utility to examine the cluster and its components. Also, constantly monitor the `hacmp.out` file. Use the SMIT Show Cluster Services screen to make sure that the necessary HACMP for AIX daemons are running on each node. Finally (if necessary) examine the other cluster log files to get a more in-depth view of the cluster status.

Consult the *HACMP for AIX Troubleshooting Guide* for help if you detect a problem with an HACMP cluster.

Monitoring Cluster, Node, and Network Interface Status

HACMP for AIX provides the `/usr/sbin/cluster/clstat` and HAView utilities for monitoring a cluster and its components. This utility `clstat` is a Clinfo client program that uses the Clinfo API to retrieve information about the cluster. Clinfo must be running on a node for this utility to work properly.

The `/usr/sbin/cluster/clstat` Utility ASCII Display

The `/usr/sbin/cluster/clstat` utility runs on both ASCII and X Window Display clients in either single-cluster or multi-cluster mode. Multi-cluster mode requires that you use the `-i` flag when invoking the `clstat` utility. The client display automatically corresponds to the capability of the system. For example, if you run `clstat` on an X Window client, a graphical display appears; however, you can run an ASCII display on an X-capable machine by specifying the `-a` flag.

The `clstat` utility reports whether the cluster is up, down, or unstable. It also reports whether a node is up, down, joining, leaving, or reconfiguring, and the number of nodes in the cluster. For each node, the utility displays the IP label and address of each network interface attached to the node, and whether that interface is up or down. See the `clstat` man page for additional information about this utility.

Single-Cluster ASCII Display Mode

In single-cluster ASCII display mode, the **clstat** utility displays information about *only* one cluster.

To invoke the **clstat** utility in single-cluster (non-interactive) mode, enter `clstat` to get a screen similar to the following:

```
clstat - HACMP for AIX Cluster Status Monitor
-----
Cluster: ibm_26c          (666)          Thu Jul  9 18:49:37 EDT 1998
          State: DOWN                Nodes: 2
          SubState: UNSTABLE
Node: poseidon           State: DOWN
  Interface: poseidon_enboot (0)      Address: 140.186.70.106
                                       State:  DOWN

Node: venus              State: DOWN
  Interface: venus_enboot (0)        Address: 140.186.70.107
                                       State:  DOWN
***** f/forward, b/back, r/refresh, q/quit*****
```

The **clstat** Single-Cluster ASCII Display Mode

The cluster information displayed shows both a cluster name and an ID. In this example, the cluster is up and has two nodes, both of which are down. Each node has one network adapter. Note that the *forward* and *back* menu options apply when more than one page of information is available to display.

If more than one cluster exists when you run the **clstat** command, the utility notifies you of this fact and requests that you retry the command specifying one of the following options:

```
clstat [-c cluster ID] [ -r seconds] [-i]
```

where:

- c *ID*** Displays information about the cluster with the specified ID if that cluster is active. This option cannot be used with the **-n** option. Displays information about the cluster with the specified ID. If the cluster is not available, the **clstat** utility continues looking for it until it is found or until the program is cancelled. Note that this option cannot be used if the **-i** option (for multi-cluster mode) is used.
- r *seconds*** Updates the cluster status display at the specified number of seconds. The default is 1 second; however, the display is updated only if the cluster state changes.
- a** Causes **clstat** to display in ASCII mode.
- i** Displays information about clusters interactively. Only valid when running **clstat** in ASCII mode.

To see cluster information about a specific cluster, enter:

```
clstat [-c cluster ID]
```

Multi-Cluster ASCII Display Mode

The multi-cluster (interactive) mode lets you monitor all clusters that Clinfo can access from the list of active service IP labels or addresses found in the `/usr/sbin/cluster/etc/clhosts` file. In multi-cluster mode, the **clstat** utility displays this list of recognized clusters and their IDs, allowing you to select a specific cluster to monitor. For example, to invoke the **clstat** utility in multi-cluster mode, enter:

```
clstat -i
```

where the **-i** indicates multi-cluster (interactive) ASCII mode. A screen similar to the following appears.

```
clstat - HACMP for AIX Cluster Status Monitor
-----
Number of clusters active: 1

      ID      Name      State
      666    ibm_26c     DOWN

Select an option:
# - the Cluster ID          x- quit
```

The **clstat** Multi-Cluster Mode Menu

This screen displays the ID, name, and state of each active cluster accessible by the local node. You can either select a cluster for which you want to see detailed information, or quit the **clstat** utility. When you enter a cluster ID, a screen appears similar to that shown in the figure:

```
clstat - HACMP for AIX Cluster Status Monitor
-----
Cluster: ibm_26c      (666)      Thu Jul  9 18:58:33 EDT 1998
      State: DOWN      Nodes: 2
      SubState: UNSTABLE
Node: poseidon      State: DOWN
      Interface: poseidon_enboot (0)      Address: 140.186.70.106
                                          State: DOWN

Node: venus      State: DOWN
      Interface: venus_enboot (0)      Address: 140.186.70.107
                                          State: DOWN

***** f/forward, b/back, r/refresh, q/quit *****
```

The **clstat** Multi-Cluster ASCII Display Mode

After viewing this screen, press “q” to exit the display. The multi-cluster mode returns you to the cluster list so you can select a different cluster. Note that you can use all menu options displayed. The *forward* and *back* options allow you to scroll through displays of active clusters without returning to the previous screen.

The /usr/sbin/cluster/clstat Utility X Window System Display

When you start the `/usr/sbin/cluster/clstat` utility on a node capable of displaying X Window System applications, the `clstat` utility displays its graphical interface, if the client's `DISPLAY` environment variable is set to the value of the X server's node address.

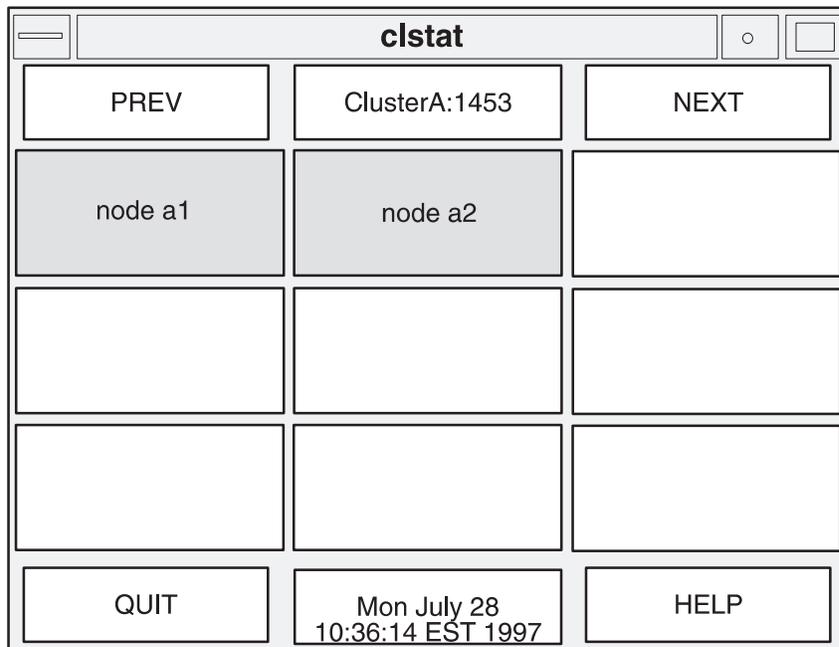
To invoke the `clstat` utility X Window System display, enter the `clstat` command:

```
clstat [-n name] [-c Id] [-r #] [-D debug_level]
```

where:

- n name** The cluster name.
- c ID** Displays information about the cluster with the specified ID if that cluster is active. This option cannot be used with the **-n** option.
- r #** the interval at which the `clstat` utility updates the display. For the graphical interface, this value is interpreted in tenths of seconds. By default, `clstat` updates the display every 0.10 seconds.
- D debug_level** The level of debugging to be performed. The levels range from 1 to 9 in increasing amounts of information. The default (0) turns debugging off.

The `clstat` utility graphical interface uses windows to represent cluster nodes, as shown below.



The `clstat` X Window System Display

Monitoring an HACMP Cluster

Monitoring Cluster, Node, and Network Interface Status

The middle box in the top row indicates the cluster name and ID. If the cluster is stable, this box appears green. If the cluster becomes unstable for any reason, this box changes to red.

Note: While a node reintegrates into a cluster, ignore any flashes of color. The node turns green when it finishes reintegrating.

The large boxes in other rows represent nodes. A node name appears in a box for each active node in the cluster. Up to eight nodes can be displayed per cluster. Nodes that are up are shown in green, nodes that are down are shown in red, nodes that are joining or leaving the cluster are shown in yellow (topology changes), and nodes that are undefined are shown in the background color. Colors are configured in the **Xclstat** X Window resource file in the `/usr/lpp/cluster/samples/clstat/` directory.

On a monochrome display, gray shading represents the colors as follows:

red dark gray

yellow gray

green light gray

Five buttons are available on the **clstat** display:

PREV Displays the previous cluster (loops from end to start).

NEXT Displays the next cluster (loops from start to end).

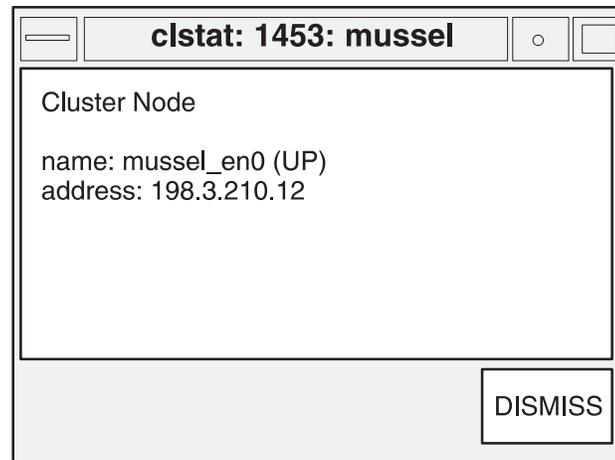
cluster:ID The refresh bar. Pressing this bar updates the status display.

QUIT Cancels the **clstat** utility.

HELP Displays help information.

Viewing Network Interface Information on an X Window System Display

To view information about network interfaces attached to a node, click mouse button 1 on the appropriate node box in the **clstat** display. A pop-up window similar to the following appears. The title in the example shows that you are viewing node *mussel* in cluster *1453*.



The clstat Node Information Display

Press the **DISMISS** button to close the pop-up window and to return to the **clstat** display window. Do *not* use the **Close** option in the pull-down menu in the upper left corner of the window to close this display; it will terminate the **clstat** utility.

Note: **clstat** may exit if you increase and decrease the size of the pop-up window. **clstat** *does* exit if you resize the window, then close it using the **Close** option in the window menu.

Monitoring a Cluster With HAView

HAView is a cluster monitoring utility that allows you to monitor HACMP clusters using TME 10 NetView for AIX. Using NetView, you can monitor clusters and cluster components across a network from a single management station.

HAView creates and modifies NetView objects that represent clusters and cluster components. It also creates submaps that present information about the state of all nodes, networks, and network interfaces associated with a particular cluster. This cluster status and configuration information is accessible through NetView's menu bar.

HAView monitors cluster status using the Simple Network Management Protocol (SNMP). It combines periodic polling and event notification through traps to retrieve cluster topology and state changes from the HACMP management agent, the Cluster SMUX peer daemon (**clsmuxpd**).

You can view cluster event history using the HACMP Event Browser and node event history using the Cluster Event Log. Both browsers can be accessed from the NetView menu bar. The `/usr/sbin/cluster/history/cluster.mmd` file contains more specific event history. This information is helpful for diagnosing and troubleshooting fallover situations. See page 3-29 for more information about this log file.

HAView Installation Considerations

HAView has a client/server architecture. You must install both an HAView server image and an HAView client image, on the same machine or on separate server and client machines. For more information on installation considerations, refer to the *HACMP for AIX Installation Guide*, Chapter 23, Installing HACMP for AIX Software, and in particular, the section Installation Server on page 23-3.

HAView File Modification Considerations

Certain files may need to be modified in order for HAView to monitor your cluster properly. When configuring HAView, check and/or edit the following files:

- **haview_start**
- **clhosts**
- **snmpd.conf**

The haview_start File

The **haview_start** file must be edited by the user so that it includes the name of the node that has the HAView server executable installed. This is how the HAView client knows where the HAView server is located. Regardless of whether the HAView server and client are on the same node or different nodes, you are required to specify the HAView server node in the **haview_start** file.

The **haview_start** file is loaded when the haview client is installed and is stored in `/usr/haview`. Initially, the **haview_start** file contains only the following line:

```
"${HAVIEW_CLIENT:-/usr/haview/haview_client}" $SERVER
```

You must add the following line to the file:

```
SERVER="${SERVER:-<your server name>}"
```

For example, for a cluster with the HAView server installed on *mynode*, the edited **haview_start** file appears as follows:

```
SERVER="${SERVER:-mynode}"  
"${HAVIEW_CLIENT:-/usr/haview/haview_client}" $SERVERa
```

where *mynode* is the node that contains the HAView server executable.

The clhosts File

HAView monitors a cluster's state within a network topology based on cluster-specific information in the `/usr/sbin/cluster/etc/clhosts` configuration file. The **clhosts** file must be present on the management node, or on the workstation running NetView. Make sure this file contains the IP address or IP label of the service and/or boot adapters of the nodes in each cluster that HAView is to monitor.

Note: Make sure the hostname and the service label of your NetView nodes are exactly the same. (If they are not the same, add an alias in the `/etc/hosts` file to resolve the name difference.)



Warning: If an invalid IP address exists in the `clhosts` file, HAView will fail to monitor the cluster. Make sure the IP addresses are valid, and there are no extraneous characters in the file.

The `snmpd.conf` File

The NetView management node must also be configured in the list of trap destinations in the `snmpd.conf` files on the cluster nodes of all clusters you want it to monitor. This makes it possible for HAView to utilize traps in order to reflect cluster state changes in the submap in a timely manner. Also, HAView can discover clusters not specified in the `clhosts` file on the nodes in another cluster.

The format for configuring trap destinations is as follows:

```
trap <community name> <IP address of NetView management node>1.2.3 fe
```

For example, enter:

```
trap          public          140.186.131.121 1.2.3 fe
```

Note the following:

- You can specify the name of the management node instead of the IP address.
- You can include multiple trap lines in the `snmpd.conf` file.

NetView Hostname Requirements for HAView

The following hostname requirements apply to using HAView in a NetView environment. If you change the hostname of an adapter, the NetView database daemons and the default map are affected.

Hostname Effect on the NetView Daemon

The hostname required to start NetView daemons must be associated with a valid interface name; if not, NetView fails to start.

Hostname Effect on the NetView Default Map

If you change the hostname of the NetView client, the new hostname does not match the original hostname referenced in the NetView default map database and NetView will not open the default map. Using the NetView `mapadmin` command, you need to update the default map (or an invalid map) to match the new hostname.

See the *NetView for AIX Administrator's Guide* for more information about updating or deleting an invalid NetView map.

Starting HAView

Once you have installed the HAView client and server, HAView is started and stopped when you start or stop NetView, but there are some conditions you should verify on the management node before starting HAView.

Before starting NetView/HAView, check the management node as follows:

Monitoring an HACMP Cluster

Monitoring a Cluster With HAView

- Make sure both client and server components of HAView are installed. See Chapter 23, *Installing HACMP for AIX Software*, for more information.
- Make sure access control has been granted to remote nodes by running the **xhost** command with the plus sign (+) or with specified nodes.

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>
```

- Make sure the DISPLAY variable has been set to the monitoring node and to a label that can be resolved by and contacted from remote nodes:

```
export DISPLAY=<monitoring node>:0.0
```

These actions allow you to access the HAView Cluster Administration option.

After ensuring these conditions are set, type the following to start NetView:

```
/usr/OV/bin/nv6000.
```

Refer to the *NetView for AIX User's Guide for Beginners* for further instructions about starting NetView.

When NetView starts, HAView creates objects and symbols to represent a cluster and its components. Through submaps, you can view detailed information about these components.

HAView places the Clusters symbol on the NetView map after NetView starts. As shown in the following figure, the Clusters symbol is placed alongside the NetView Collections, Manager Submap, and IP Internet symbols.



HAView Clusters Symbol

Viewing Clusters and Components

To see which clusters HAView is currently monitoring, double-click the Clusters symbol. The Clusters submap appears. You may see one or more symbols that represent specific clusters. Each symbol is identified by a label indicating the cluster's name. Double-click a cluster symbol to display symbols for nodes and networks within that cluster.

Note that the cluster status symbol may remain unknown until the next polling cycle, even though the status of its cluster components is known. See the section Polling Intervals on page 3-13 for more information about the default intervals and how to change them using SMIT.

Note: You can view component details at any time using the shortcut **ctrl-o**. See Obtaining Component Details on page 3-13 for information and instructions.

Read-Write and Read-Only NetView Maps

Normally, you have one master monitoring station for NetView/HAView. This station is supplied with new information as cluster events occur, and its map is updated so it always reflects the current cluster status.

In normal cluster monitoring operations, you will not need to open multiple NetView stations from the same node. If you do, and you want the additional stations to be updated with current cluster status information, you must be sure they use separate maps with different map names. For more information on multiple maps and changing map permissions, see the *NetView for AIX Administrators Guide*.

Interpreting Cluster Topology States

When using HAView to view cluster topology, symbols for clusters and cluster components such as nodes and networks are displayed in various colors depending on the object's state. The following table summarizes colors you may see when monitoring a cluster.)

Status	Meaning	Symbol Color	Connection Color (network submap)
Critical	The object has failed or is not functioning. If the symbol is a node or network, the node or network is DOWN.	Red	Red
Normal	The object is functioning correctly. If the symbol is a node object, the node is UP.	Green	Black
Marginal	Some object functions are working correctly; others are not.	Yellow	Red
Unknown	The object's state cannot be determined. It may not be currently monitored by HAView.	Blue	Blue

Note: You can select **Legend** at any time from the Help pull-down menu to view NetView and HAView symbols and to understand their associative colors.

The Navigation Tree and Submap Windows

In addition to the submap window, the NetView Navigation Tree Window can help you keep track of your current location in the HAView hierarchy. Press the Tree button to see the Navigation Tree Window. In the Navigation Tree, you can see which submap you are in by checking the blue outline.

The Symbols Legend

At any time, you can select **Legend** from the Help pull-down menu to view all NetView and HAView symbols and the meanings of the symbols and their various colors.

The Help Menu

To view help topics, select **Help > Index > Tasks > HA View Topics**.

Viewing Networks

To view the state of the nodes and addresses connected to a network associated with a specific cluster, double-click a network symbol in the specific Cluster submap. A network submap appears displaying symbols for all nodes connected to the network. The symbols appear in a color that indicates the nodes' current state. The vertical line representing a network is called the network connection. Its color indicates the status of the connection between the node and the network.

Viewing Nodes

To view the state of nodes associated with a particular network, double-click a network symbol. A submap appears displaying all nodes connected to the network. Each symbol's color indicates the associated node's current state.

You can also view the state of any individual node associated with a cluster by double-clicking on that node's symbol in the specific cluster submap.

Viewing Addresses

To view the status of addresses serviced by a particular node, double-click a node symbol from either a cluster or network submap. A submap appears displaying symbols for all addresses configured on a node. Each symbol's color indicates the associated address's current state.

Note: When viewing adapters in a node submap from a network submap, all adapters relating to that node are shown, even if they are not related to a particular network.

Obtaining Component Details

NetView dialog boxes allow you to view detailed information about a cluster object. A dialog box can contain information about a cluster, network, node, network adapter, or about cluster events. You can access an object's dialog box using the NetView menu bar or the Object Context menu, or by pressing **ctrl-o** at any time:

To view details about a cluster object using the NetView menu bar:

1. Click on an object in any submap.
2. Select the **Modify/Describe** option from the NetView Edit menu.
3. Select the **Object** option.
An Object Description dialog window appears.
4. Select **HAView for AIX** and click on **View/Modify Object Attributes**.
An Attributes dialog window appears.

Note: You can view dialog boxes for more than one object simultaneously by either clicking the left mouse button and dragging to select multiple objects, or by pressing the **Alt** key and clicking on all object symbols for which you want more information.

To view details about a cluster object using the Object Context menu:

1. Click on an object in any submap.
2. Click on the symbol you have highlighted to display the object context menu, using:
 - Button 3 on a three-button mouse
 - Button 2 on a two-button mouse.
3. Select **Edit** from the object context menu.
4. Select **Modify/Describe** from the Edit cascade menu.
5. Select the **Object** option.
An Object Description dialog window appears.
6. Select **HAView for AIX** and click on **View/Modify Object Attributes**.
An Attributes dialog window appears.

Polling Intervals

To ensure that HAView is optimized for system performance and reporting requirements, you can customize these two parameters:

- The polling interval (in seconds) at which HAView polls the HACMP clusters to determine if cluster configuration or object status has changed. The default is 60 seconds.
- The polling interval (in minutes) at which HAView polls the **clhosts** file to determine if new clusters have been added. The default for Cluster Discovery polling is 120 minutes.

You can change the HAView polling intervals using the SMIT interface as follows:

1. On the HAView server node, open a SMIT screen by typing:

```
smitty haview
```

The Change/Show Server Configuration window opens.

2. Enter the polling interval numbers you want (between 1 and 32000) and press OK.

Note: If the **snmpd.conf** file is not properly configured to include the NetView management node as a trap destination, HAView can detect a trap that occurs as a result of a cluster event, but information about the network topology may not be timely. Refer back to the section HAView File Modification Considerations on page 3-8 for more information on the **snmpd.conf** file.

Removing a Cluster from HAView Monitoring

If a cluster does not respond to status polling, you can use the Remove Cluster option to remove the cluster from the database. To remove a cluster, it must be in an UNKNOWN state, represented by a blue cluster symbol. If the cluster is in any other state, the Remove Cluster option is disabled.

Warning: The **Remove Cluster** option is the only supported way to delete HAView objects from submaps. Do not delete an HAView symbol (cluster or otherwise) through the Delete Object or Delete Symbol menu items. If you use these menu items, HAView continues to poll the cluster.

When you remove a cluster, the following actions occur:

- The cluster name is removed from the NetView object database and HAView stops polling the cluster.
- The symbol for the cluster is deleted.
- The symbols for all child nodes, networks, and addresses specific to that cluster are deleted.

If you are removing the cluster permanently, remember to remove the cluster addresses from the **/usr/sbin/cluster/etc/clhosts** file. If you do not remove the cluster addresses from the **clhosts** file, New Cluster Discovery polling continues to search for the cluster.

To remove a cluster:

1. Click on the cluster symbol you wish to remove. The cluster must be in an UNKNOWN state, represented by a blue cluster symbol.
2. Select **HAView** from the Tools pull-down menu.
3. Select **Remove Cluster** from the HAView cascade menu.

Using the HAView Cluster Administration Utility

HAView allows you to start a **SMIT hacmp** session to perform cluster administration functions from within the NetView session. The administration session is run on an aixterm opened on the chosen node through a remote shell. You can open multiple sessions of SMIT hacmp while in HAView.

Note: You can start an administration session for any node that is in an UP state (the node symbol is green). If you attempt to start an administration session when the state of the node is DOWN or UNKNOWN, no action occurs.

When bringing a node up, the HAView node symbol may show green before all resources are acquired. If you select the node symbol and attempt to open an administration session before all resources are acquired, you may receive an error.

Opening and Closing a Cluster Administration Session

To open a cluster administration session:

1. Click on an available node symbol (one that is green).
2. Select the **Tools > HAView > Cluster Administration**.
3. Proceed with your tasks in SMIT.
4. Select **F10** to exit the Cluster Administration session. The aixterm session will also close when using either of these choices.

HAView Cluster Administration Utility Notes and Requirements

Keep the following considerations in mind when using the Cluster Administration option:

- Be sure you have run the **xhost** command prior to starting NetView, so that a remote node can start an aixterm session on your machine.
- Be sure you have set the DISPLAY variable to a label that can be resolved and contacted from remote nodes.
- For the cluster administration session to proceed properly, the current NetView user (the account that started NetView) must have sufficient permission and be authenticated to perform an rsh to the remote node in the **/.rhosts** file or through Kerberos if you have HACMP on an RS/6000 SP. For more information, see the following sections in the *HACMP for AIX Installation Guide*: Chapter 7 (section on editing the **/.rhosts** file) and Appendix F (section on Kerberos).
- If an IP Address Takeover (IPAT) occurs while a cluster administration session is running, the route between the remote node that the HAView monitoring node may be lost.

HAView Browsers

HAView provides two browsers which allow you to view the event history of a cluster, the Cluster Event Log and the HACMP Event Browser.

Cluster Event Log

Using the Cluster Event Log you can view the event history for a cluster as recorded by a specific node. The Log browser is accessible through the NetView Tools menu, and is only selectable if an active node symbol is highlighted.

For more detailed information on a node's event history, log onto the specific node and check the Cluster Message Log Files. See the Examining Cluster Log Files chapter in the *HACMP for AIX Troubleshooting Guide* for more information on Cluster Message Log Files.

Note: To ensure that the header for the Cluster Event Log displays properly, install all the NetView fonts on your system.

1. Click on the node symbol for which you wish to view a Cluster Event Log.
2. Select **HAView** from the Netview Tools menu.
3. Select the **Cluster Event Log** option.
4. Set the **number of events to view** field. You can use the up and down arrows to change this number or you can enter a number directly into the field. The possible range of values is 1 to 1000 records. The default value is 100.
5. Press the **Issue** button to generate the list of events. The message area at the bottom of the dialog box indicates when the list is done generating.

When the list is done generating, the dialog box displays the following view-only fields:

Event ID	This field displays a numeric identification for each event which occurred on the cluster.
Node Name	The name of the node on which the event occurred.
Time	The date and time the event occurred. This field is in the format MM DD hh:mm:ss.
Description	A description of the event.

6. Press the **Dismiss** button to close the dialog box.

HACMP Event Browser

HAView provides a NetView browser which allows you to view the accumulative event history of a cluster. The browser shows the history of all nodes in the cluster, broadcast through an assigned primary node. If the primary node fails, another node will assume the primary role and continue broadcasting the event history.

The HACMP Event Browser provides information on cluster state events. A filter is used to block all redundant traps.

The HACMP Event Browser is available through the NetView menu bar. The menu item is always active, and when selected will start a NetView browser showing the event history for all active clusters. Note that you can access only one instantiation of the Event Browser at a time.

To view the HACMP Event Browser:

1. Select **HAView** from the Netview Tools menu.
2. Select the **HACMP Event Browser** option.

The HACMP Event Browser appears. Note that only one instantiation of the Event Browser can be accessed at a time. See the *NetView for AIX User's Guide for Beginners* for more information on the NetView browser functions.

3. Select the **Close** option from the File menu of the HACMP Event Browser menu bar to close the browser.

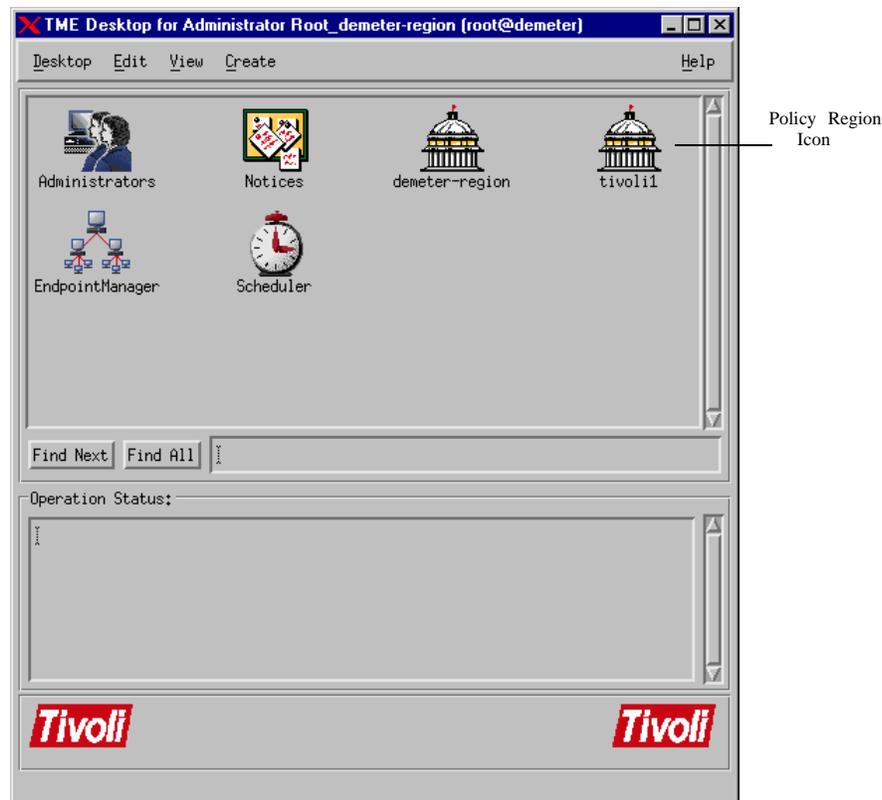
Note: When you exit the Event Browser, the HAView application restarts. At this time, the HACMP cluster icon turns blue, disappears, and then reappears.

Monitoring HACMP Clusters with Tivoli

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 Desktop, you can monitor the following aspects of your cluster:

- Cluster state and substate
- Configured networks and network state
- Participating nodes and node state

The Tivoli Desktop interface is shown here:



The Tivoli Desktop Initial Screen

Tivoli's thermometer icons (see figure on page 3-22) provide a visual indication of whether components are up, down, in transition, or in an unknown or error state. From the window for a selected Policy Region, you can go to a cluster's Indicator Collection window, which displays thermometer icons indicating the state of all cluster components.

The cluster status information shown by the thermometers is updated every three minutes or at another interval you specify. (Further information on changing the default polling interval appears later in this chapter.)

In order to integrate HACMP with Tivoli, you must configure your HACMP cluster nodes as subscriber (client) nodes to the Tivoli server, or Tivoli Management Region (TMR). Each cluster node can then maintain detailed node information in its local Tivoli database, which the TMR accesses for updated node information to display.

For full installation details, see Appendix F, Installing and Configuring Cluster Monitoring with Tivoli, in the *HACMP for AIX Installation Guide*.

Consult Your Tivoli Documentation

The following sections provide information on monitoring an HACMP cluster through the Tivoli interface. Descriptions of Tivoli components and processes are provided here as needed, but for full information on installing, configuring, and using the Tivoli software itself, consult your Tivoli user documentation or Tivoli support.

Prerequisites and Considerations for Cluster Monitoring with Tivoli

As you planned and configured Cluster Monitoring with Tivoli, the following points should have been considered:

- The Tivoli Management Region (TMR) should be located on an AIX node outside the cluster.
- The Tivoli Framework, Distributed Monitoring, and AEF components must be installed on the Tivoli Management Region node and on each cluster node.
- In the Policy Region that contains HACMP cluster monitoring tasks and monitors, the Task Library and Profile Managers must be configured as managed resources.
- To ensure accurate monitoring of IP address takeover, consider using a dedicated network for communication between the TMR and the cluster nodes. If you do not have a separate network dedicated to Tivoli, you must set up your cluster to have one physical network with an extra subnet in order for IP address takeover and related post-events to occur. This subnet is used for the TMR adapter and for the cluster node's alias IP address.
- Also for proper IPAT monitoring, you create an alias for the IP address of each cluster node's standby adapter. You must place this alias in the `/etc/hosts` file and the `/usr/sbin/hativoli/ipaliases.conf` file. It is also recommended that it be placed in the Tivoli `/etc/wlocalhost` file.

For further detail on all prerequisites and installation issues for cluster monitoring with Tivoli, see Appendix F, Installing and Configuring Cluster Monitoring with Tivoli, in the *HACMP for AIX Installation Guide*. for details on installation.

Installing and Configuring Cluster Monitoring with Tivoli

Preparing to monitor a cluster with Tivoli involves properly installing the Tivoli software, defining the necessary Tivoli components to handle cluster monitoring, and then installing the proper HACMP filesets to make Tivoli aware of your HACMP cluster nodes. In addition, if you do not have a dedicated network for communication between the TMR and HACMP cluster nodes, there are some prerequisite tasks you must perform in order to facilitate IP address takeover.

For complete instructions in installing HACMP Tivoli-related filesets and making Tivoli aware of your cluster, see Appendix F, Installing and Configuring Cluster Monitoring with Tivoli, in the *HACMP for AIX Installation Guide*.

Using Tivoli to Monitor the Cluster

Once you have properly installed your **hativoli** files and defined your nodes to Tivoli, you can view information on the status of your HACMP cluster components.

When you monitor your cluster through Tivoli, you can access cluster information in both icon and text form, in a number of different Tivoli windows. The next few sections are meant to orient you to the flow of Tivoli cluster monitoring information.

Note: **When all nodes are down, node status may not be displayed accurately.** You should be aware that in the event that your last remaining cluster node goes down, Tivoli may still indicate that the cluster is up. This can occur when HACMP is unable to contact the MIB for updated information. In this case, the Tivoli display will show information *as of the last successful poll*.

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.

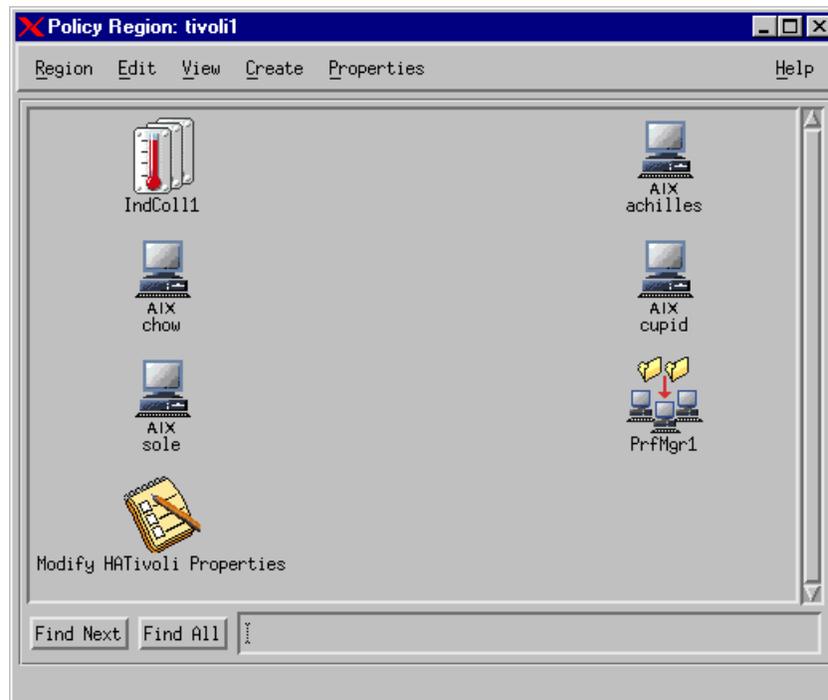
Note that there may be a delay as Tivoli adds the indicators for the cluster.

Tivoli Policy Regions

A Tivoli *Policy Region* groups together all entities related to a specific area of Tivoli monitoring. In this case, that area will be the HACMP cluster or clusters. The HACMP Policy Region will encompass the nodes, clusters, indicator icons, and tasks related to your HACMP configuration.

You create this Policy Region during the process of defining your cluster nodes to Tivoli, as described in Appendix F, Installing and Configuring Cluster Monitoring with Tivoli, in the *HACMP for AIX Installation Guide*.

Policy Region icons appear in the initial Tivoli Desktop window (shown on page 3-18). Clicking on a Policy Region icon opens the Policy Region window, in which you see the thermometer icons of the Indicator Collections for each cluster, as well as icons for the specific profiles and tasks associated with the HACMP Policy Region.



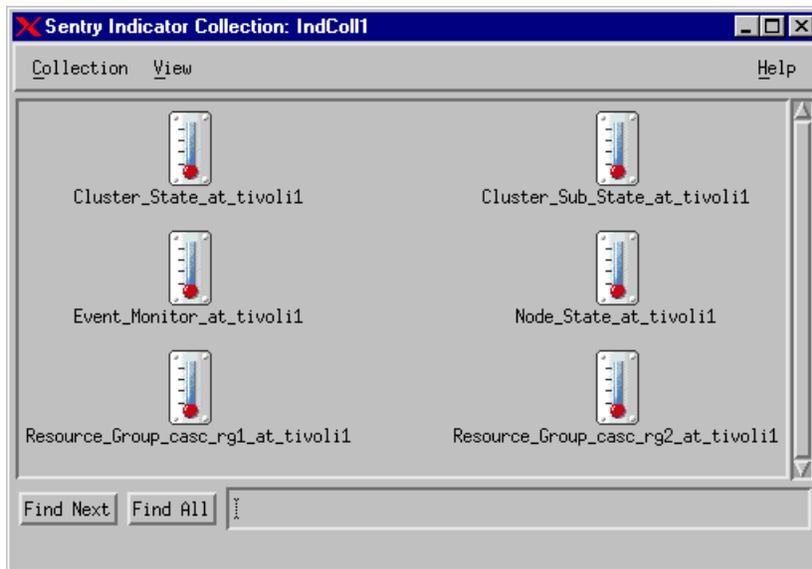
HAMCP Policy Region Window

Tivoli Distributed Monitors and Indicator Collections

For each cluster, a group of Tivoli *distributed monitors* is created. These distributed monitors query the HACMP cluster node at set intervals for information about the various cluster components. The group of monitors is associated with an *indicator collection* that displays the state of the cluster components. If a change is detected in the state of the cluster or one of its components, the distributed monitor takes action, changing an icon in the associated Indicator Collection window. This provides the Tivoli administrator with a visual representation of any changes in the status of the monitored items.

The icon for cluster and cluster component status is a thermometer figure with varying levels of red color depending on the severity of the component status. When you click on a cluster's Indicator Collection icon in the Policy Region window, the Indicator Collection window appears, showing status icons for that cluster's components.

Monitoring an HACMP Cluster
Monitoring HACMP Clusters with Tivoli



Indicator Collection Window

Note: The diagram above includes icons for resource groups. Resource group monitoring is available only in the HACMP/ES (Enhanced Scalability) product feature code. For more information, see the *HACMP for AIX: Enhanced Scalability Installation and Administration Guide*.

Interpreting Indicator Displays for Various Cluster Components

The height of the red color in the thermometer and the color-coded marker alongside it indicate whether a component is up, down, or in various degrees of trouble. The following tables list the indicator displays for various cluster component states:

CLUSTER STATE Indicator Display	Cluster State
Normal	UP
Fatal	DOWN
Severe	UNKNOWN

CLUSTER SUBSTATE Indicator Display	Cluster Substate
Normal	STABLE
Warning	UNSTABLE
Severe	RECONFIG
Critical	ERROR

NODE Indicator Display*	Node State(s)
Normal	All nodes ONLINE
Warning	One or more nodes OFFLINE

*Note that node state is displayed in the Distributed Monitor Indicator Collection as a composite view of all nodes rather than an individual node view.

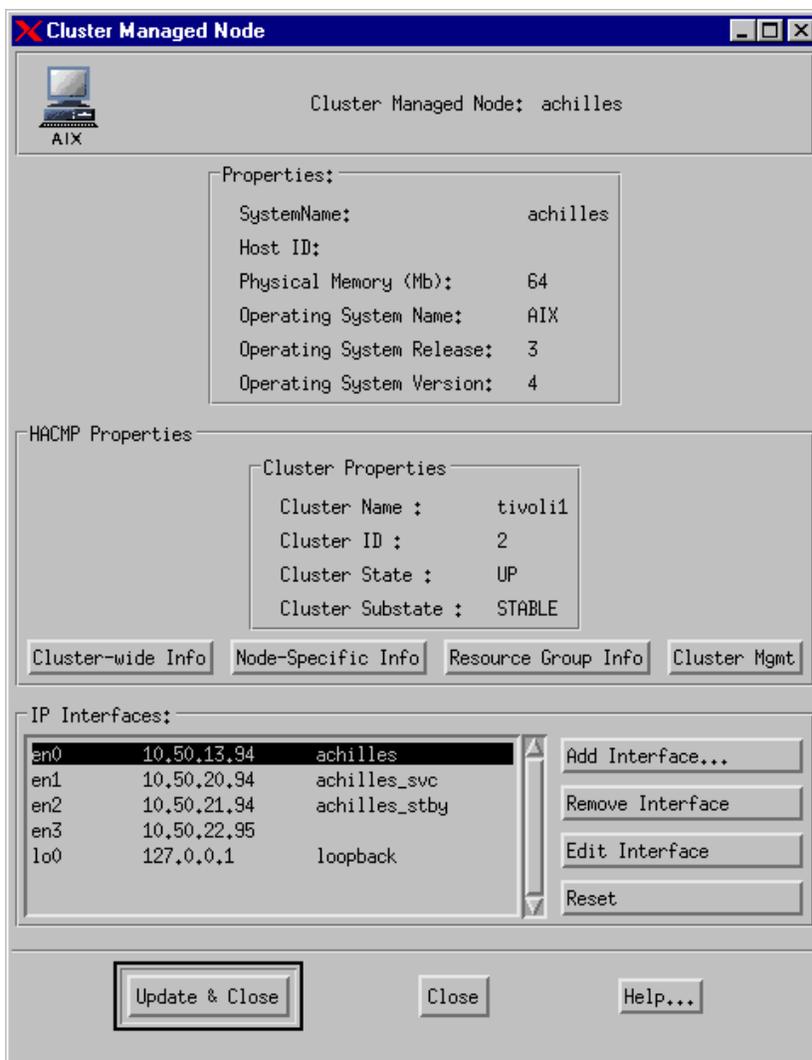
Viewing Cluster Information

The Cluster Managed Node window gives you all information about the current cluster topology and configuration.

The Properties section displays standard system properties information for the managed node, and the IP Interfaces section at the bottom shows standard IP Interface information for the node.

With the addition of the HACMP cluster monitoring feature, the standard Tivoli node icon is extended to capture and display additional information specific to your HACMP cluster. The items that you see in the HACMP Properties portion of the window are detailed in the following sections.

To view the Managed Node window, right-click on a node icon in the Policy Region window.



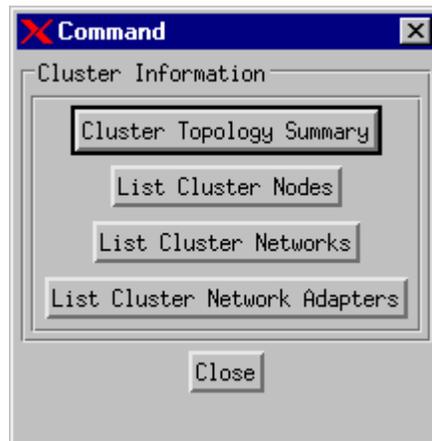
Cluster Managed Node Window

In the HACMP Properties portion of the Managed Node window shown above, you see the following four items of HACMP-specific top-level cluster information:

- Cluster name
- Cluster ID
- Cluster state
- Cluster substate

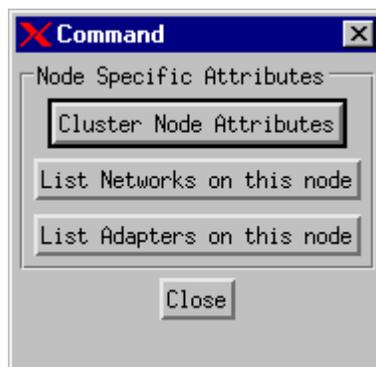
In addition, the HACMP Properties buttons lead you to further cluster and component details. Selecting a button brings up a new popup window with options for retrieving specific information. These buttons and the choices within them are detailed below.

Cluster-wide Information Button



Clicking this button gives you options to view details on the configuration and status information for the cluster as a whole.

Node-Specific Information Button



From the Node Specific Attributes window, you can access further information about the attributes, networks, and adapters associated with the node.

Cluster Management Button



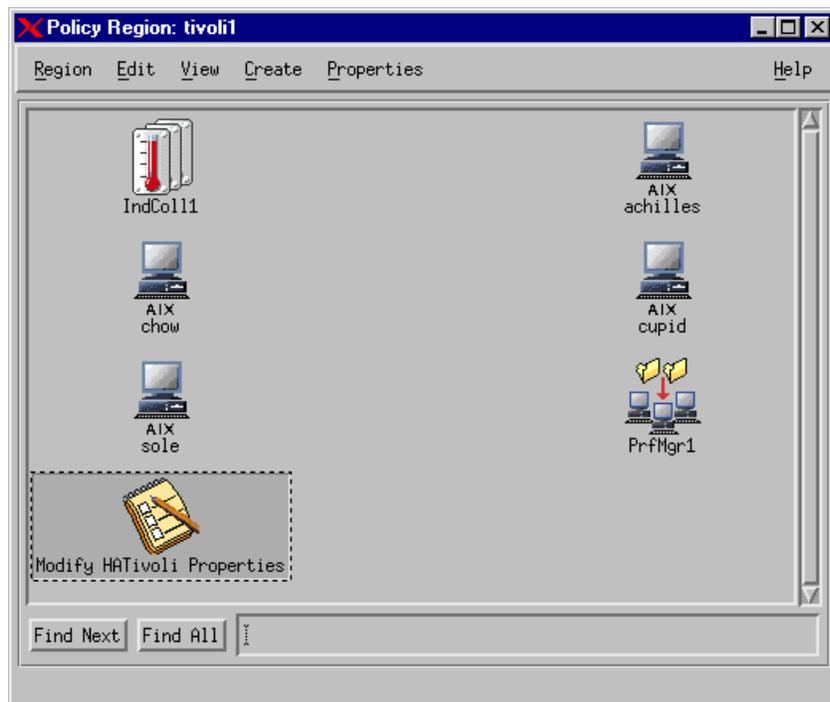
From the Cluster Management window, you can open a SMIT window to perform all of your normal cluster management tasks.

Note: In order to open the SMIT window from within Tivoli, you must be sure to run the command `xhosts + (or xhosts <node name>)` to grant access to remote nodes, as explained in the instructions for starting Tivoli.

The HACMP Task Library

The Task Library contains options to perform cluster management tasks such as configuring, modifying, and deleting various items associated with the cluster, and refreshing the cluster view.

Task categories appear as “notepad” icons in the Policy Region window.



Policy Region Window with Task Icon Highlighted

Polling Intervals

The Distributed Monitors poll the cluster nodes periodically for cluster topology and status changes. The default polling interval is three minutes. If this interval is too short for your particular cluster monitoring needs, you can increase this interval through an HACMP Task found in the Policy Region window. It is not recommended to make the polling interval shorter than the default.

Note: As mentioned earlier, be aware that in the event that your last remaining cluster node goes down, Tivoli may still indicate that the cluster is up. This can occur when HACMP is unable to contact the MIB for updated information. In this case, the Tivoli display will show information *as of the last successful poll*.

Deinstalling Cluster Monitoring with Tivoli

To discontinue cluster monitoring with Tivoli, you must perform the following steps to delete the HACMP-specific information from Tivoli.

1. Run a deinstall through the SMIT interface, deinstalling the three **hativoli** filesets on all cluster nodes and the TMR.
2. If it is not already running, invoke Tivoli on the TMR:
 1. type `./etc/Tivoli/setup_env.sh`
 2. Type **tivoli**
3. In the Policy Region for the cluster, go to HATivoli Properties.
4. Select the Modify Properties task.
A window appears containing task icons.
5. Choose **Edit > Select All** to select all tasks, and then **Edit > Delete** to delete.
The Operations Status window at the left shows the progress of the deletions.
6. Return to the Properties window and delete the Modify Properties task icon.
7. Open the Profile Manager.
8. Choose **Edit > Profiles > Select All** to select all HACMP Indicators.
9. Choose **Edit > Profiles > Delete** to delete the Indicators.
10. Unsubscribe the cluster nodes from the Profile Manager:
 1. In the Profile Manager window, choose Subscribers.
 2. Highlight each HACMP node on the left, and click to move it to the right side.
 3. Click **Set & Close** to unsubscribe the nodes.

Monitoring Cluster Services

After checking cluster, node, and network interface status, check the status of the HACMP for AIX daemons on both nodes and clients.

Monitoring Cluster Services on a Node

Use the SMIT Show Cluster Services screen to check the status of the HACMP for AIX daemons on a node. Check the following daemons:

- Cluster Manager (**clstrmgr**) daemon
- Cluster Lock Manager (**cllockd**) daemon
- Cluster SMUX Peer (**clsmuxpd**) daemon

Note: If the **clsmuxpd** daemon is not active, also check the status of the SNMP (**snmpd**) daemon.

- Clinfo (**clinfo**) daemon (optional on a node).

Use the **smit clshow** fastpath to view cluster services on a node.

1. Enter:

```
smit clshow
```

A screen similar to that shown below appears.

```

                                COMMAND STATUS
Command: OK                      stdout: yes                      stderr: no
Before command completion, additional instructions may appear below.
Subsystem      Group          PID          Status
clstrmgr       cluster        1000         inoperative
clinfo         cluster        1000         inoperative
clsmuxpd       cluster        1000         inoperative
cllockd        lock           1000         inoperative
```

The Command Status Display of Node Information

The screen indicates the:

- HACMP for AIX subsystem
- Group to which the subsystem belongs
- Process ID number of the subsystem, if it is running
- Status of the subsystem.

Monitoring Cluster Services on a Client

The only HACMP for AIX process that can run on a client is the **clinfo** daemon. (Not all clients run this daemon.) You can use the AIX **lssrc** command with either the **-g cluster** or **-s clinfo** arguments to check the status of the **clinfo** daemon on a client. The output looks similar to the following:

Subsystem	Group	PID	Status
clinfo	cluster	9843	active

You can also use the **ps** command and **grep** for “clinfo”. For example:

```
ps -aux | grep clinfo
```

HACMP for AIX Log Files

HACMP for AIX writes the messages it generates to the system console and to several log files. Because each log file contains a different subset of the types of messages generated by HACMP for AIX, you can get different views of cluster status by viewing different log files. HACMP for AIX writes messages into the log files described below. See Chapter 2, Examining Cluster Log Files, in the *HACMP for AIX Troubleshooting Guide* for more information about these files.

/usr/adm/cluster.log

The **/usr/adm/cluster.log** file is the main HACMP for AIX log file. HACMP error messages and messages about HACMP for AIX-related events are appended to this log with the time and date at which they occurred.

/tmp/hacmp.out

The **/tmp/hacmp.out** file records the output generated by the configuration and startup scripts as they execute. This information supplements and expands upon the information in the **/usr/adm/cluster.log** file. To receive verbose output, the Debug Level run-time parameter should be set to *high* (the default). See Chapter 10, Additional Tasks: Configuring Run-Time Parameters, Cluster Log Files, and Maintaining NFS for details on setting run-time parameters.

/usr/sbin/cluster/history/cluster.mmdd

The **/usr/sbin/cluster/history/cluster.mmdd** file contains time stamped, formatted messages generated by HACMP for AIX scripts. The system creates a cluster history file whenever cluster events occur, identifying each file by the file name extension *mmdd*, where *mm* indicates the month and *dd* indicates the day.

While it is more likely that you will use these files during troubleshooting, you should occasionally look at them to get a more detailed picture of the activity within a cluster.

System Error Log

The system error log file contains time stamped, formatted messages from all AIX subsystems, including HACMP for AIX scripts and daemons. Cluster events are logged as operator messages (error id: AA8AB241) in the system error log.

/tmp/cm.log

Contains time-stamped, formatted messages generated by HACMP for AIX **clstrmgr** activity. This file is typically used by IBM support personnel.

/tmp/cspoc.log

Contains time-stamped, formatted messages generated by HACMP for AIX C-SPOC commands. The **/tmp/cspoc.log** file resides on the node that invokes the C-SPOC command.

/tmp/emuhacmp.out

The **/tmp/emuhacmp.out** file records the output generated by the event emulator scripts as they execute. The **/tmp/emuhacmp.out** file resides on the node from which the event emulator is invoked. You can use the environment variable **EMUL_OUTPUT** to specify another name and location for this file, but the format and information remains the same.

Chapter 4 Maintaining Shared LVM Components

This chapter explains how to maintain LVM components shared by nodes in an HACMP cluster. It includes specific procedures for managing volume groups, filesystems, logical volumes, and physical volumes. For most procedures, you can use either AIX commands or the C-SPOC utility. For creating a shared volume group or adding nodes to a volume group, you can use the TaskGuide utility.

For information about LVM components in a concurrent access environment, see Chapter 5, *Maintaining Shared LVM Components in a Concurrent Access Environment*. For specific information on AIX commands and SMIT screens, see your *AIX System Management Guide*.

Overview

A key element of any HACMP for AIX cluster is the data used by the highly available applications. This data is stored on AIX Logical Volume Manager (LVM) entities. HACMP for AIX clusters use the capabilities of the LVM to make this data accessible to multiple nodes.

In an HACMP for AIX cluster, a *shared volume group* is a volume group that resides entirely on the external disks shared by cluster nodes. A *shared physical volume* is a disk that resides in a shared volume group. A *shared logical volume* is a logical volume that resides entirely in a shared volume group. A *shared filesystem* is a filesystem that resides entirely in a shared logical volume.

Common Maintenance Tasks

As a system administrator of an HACMP for AIX cluster, you may be called upon to perform any of the following LVM-related tasks:

- Creating a new shared volume group
- Extending, reducing, changing, or removing an existing volume group
- Importing, mirroring, unmirroring, or synchronizing mirrors of a volume group
- Creating a new shared logical volume
- Extending, reducing, changing, copying, or removing an existing logical volume (or a copy)
- Creating a new shared filesystem
- Extending, changing, or removing an existing filesystem.
- Adding, removing physical volumes

When performing any of these maintenance tasks on shared LVM components, make sure that ownership and permissions are reset when a volume group is exported and then re-imported. After exporting and importing, a volume group is owned by root and accessible by the system group. Applications, such as some database servers, that use raw logical volumes may be affected by this if they change the ownership of the raw logical volume device. You must restore the ownership and permissions back to what is needed after this sequence.

Using the C-SPOC Utility to Maintain Shared LVM Components

HACMP for AIX provides the Cluster-Single Point of Control (C-SPOC) utility to automate repetitive tasks, thus simplifying maintenance of shared LVM components in clusters. C-SPOC commands provide functions in a cluster environment that are comparable to the standard AIX commands that work on a single node. For example, the C-SPOC utility includes a command called **cl_chlv** that provides similar functions to the AIX **chlv** command. (The C-SPOC command calls the AIX command.) By automating repetitive tasks, C-SPOC eliminates a potential source of errors.

You can perform the following LVM maintenance tasks using C-SPOC:

Note: The C-SPOC commands operate on both *shared* and *concurrent* LVM components (where applicable). See Chapter 5, *Maintaining Shared LVM Components in a Concurrent Access Environment*, for information on maintaining shared LVM components in a concurrent access environment.

- Shared volume groups
 - Create a shared volume group
 - Import a volume group
 - Extend a volume group
 - Reduce a volume group
 - Mirror a volume group
 - Unmirror a volume group
 - Synchronize volume group mirrors
 - List all shared volume groups
 - List all active shared volume groups
 - Display characteristics of a shared volume group
- Shared logical volumes
 - Create a logical volume
 - Make a copy of a logical volume
 - Remove a copy of a logical volume
 - List all shared logical volumes by volume group
 - Change (or view) the characteristics of a shared logical volume (name, size)
 - Remove a shared logical volume
- Shared filesystems
 - Create a shared filesystem
 - List all shared filesystems
 - Change (or view) the characteristics of a shared filesystem
 - Remove a shared filesystem.

- Physical volumes
 - Add a definition to cluster nodes
 - Remove a definition from cluster nodes

For more information about performing these tasks, see later sections of this chapter. (See the *HACMP for AIX Troubleshooting Guide* for information about using the list commands.)

When you execute a C-SPOC command, you specify the list of nodes where you want to execute the required commands. You can supply either a list of nodes, or a resource group name. Typically, C-SPOC executes the command on the node that owns the LVM component (has it varied on). However, you can use C-SPOC commands from the command line (not from SMIT) on an LVM component that is not currently activated on any cluster node. These are nodes that can own the LVM component when it is activated, as specified in the HACMP resource group.

Understanding C-SPOC and its Relation to Resource Groups

The C-SPOC commands that modify LVM components, such as **cl_chlv** or **cl_chfs**, require that you specify a resource group name as an argument. The LVM component that is the target of the command *must* be in the resource group specified. C-SPOC uses the resource group information to determine on which nodes it must execute the specified operation.

To illustrate, consider a cluster with two resource groups defined: **resgrp1** and **resgrp2**. The following summarizes the resource group configuration. Filesystems defined as resources and participating nodes in the configuration are listed:

Resource Group	Filesystem	Participating Nodes
resgrp1	/fs1	Nodes A – G
resgrp2	/fs2	Nodes A – G

Note: In a C-SPOC cluster, any number of nodes can be defined as participating in a resource group. If, however, you change a filesystem's name, be sure to update the resource group with the new name.

To change the mount point of filesystem **/fs1**, you could enter the following C-SPOC command:

```
cl_chfs -cspoc "-g resgrp1" -m /fsnew /fs1
```

The example specifies **resgrp1** in which **/fs1** is configured as a resource. However, the command will also succeed if you specify **resgrp2** as the resource group because both resource groups have the same set of participating nodes.

When removing a filesystem or logical volume using the C-SPOC **cl_rmfs** or **cl_rmlv** command, the target filesystem or logical volume must *not* be configured as a resource in the specified resource group. If the filesystem or logical volume is a part of the resource group, you should remove it from the resource group before executing the C-SPOC command to avoid errors that may occur when the takeover node tries to access the removed filesystem or logical volume.

Updating LVM Components in an HACMP Cluster

When you change the definition of a shared LVM component in a cluster, the operation updates the LVM data that describes the component on the local node and in the Volume Group Descriptor Area (VGDA) on the disks in the volume group. LVM enhancements, as provided by AIX version 4.3.1 and higher, allow all nodes in the cluster to be aware of changes to a volume group, logical volume, and filesystem, at the time the changes are made, rather than waiting for the information to be retrieved during a lazy update.

If for some reason the node is not updated via the C-SPOC enhanced utilities, due to an error condition (a node is down, for example), the volume group will be updated and the change will be taken care of during the “lazy update” mechanism or during execution of the **clvaryonvg** command.

If node failure does occur during a C-SPOC operation an error is displayed to the screen and the error messages are recorded in the C-SPOC errorlog (**/tmp/cspoc.log**).

Error Reporting provides detailed information about inconsistency in volume group state across the cluster. If an error occurs, you must take corrective action manually.

Maintaining Shared Volume Groups

The following administrative tasks involve shared volume groups:

- Creating a shared volume group
- Extending a shared volume group
- Importing a shared volume group
- Reducing a shared volume group
- Making a copy of a volume group
- Removing a copy of a volume group
- Mirroring a volume group
- Unmirroring a Volume group
- Removing a shared volume group
- Synchronizing volume group mirrors.

The following section describes how to do these tasks using the standard AIX commands. The next section describes how to do these tasks using C-SPOC.

Using C-SPOC simplifies the steps required for all tasks. Moreover, you do not have to stop and restart cluster services in order to do the tasks.

TaskGuide for Creating Shared Volume Groups

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

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

Note that you may still want to rename and mirror the default JFS log, as discussed in the procedure for creating a shared filesystem that starts on page 4-23.

TaskGuide Requirements

Before starting the TaskGuide, make sure:

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

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

Starting the TaskGuide

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

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

or you can use the SMIT interface as follows:

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

After a pause, the TaskGuide Welcome panel appears.

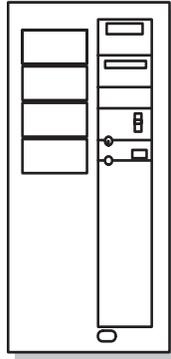
3. Proceed through the panels to create or share a volume group.

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

Using AIX Commands to Maintain Shared Volume Groups

Creating a Shared Volume Group with AIX Commands

The figure below summarizes the steps you must complete on each cluster node to create a shared volume group. Perform these tasks on each destination node, one node at a time.



Source Node

Complete prerequisite tasks
Create shared volume group
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Import volume group
Change volume group to remain dormant at startup
Vary off volume group
Complete follow-up tasks

Creating a Shared Volume Group

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

The following procedure provides more detail about each of the steps.

Note: It is essential that you perform all the steps in the correct order so the data does not become corrupted.

1. Complete prerequisite tasks. On the source node, create the volume group, using the SMIT **mkvg** fastpath.
2. Enter the specific field values and press Enter. For other fields use the defaults or the appropriate entries for your operation:

**Activate volume group
AUTOMATICALLY at
system restart?**

Set to **no** so that the volume group can be activated as appropriate by the cluster event scripts.

**ACTIVATE volume
group after it is created?**

Set this field to **yes**.

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

3. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
4. On each destination node, in turn, import the volume group using the SMIT **importvg** fastpath. When you import a volume group, the system copies data from the Volume Group Description Area (VGDA) on a physical volume in the volume group into kernel data structures.

Enter specific field values and press Enter. 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 **yes**.

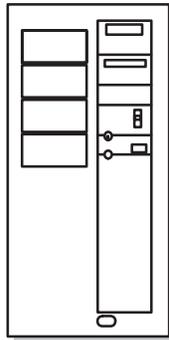
Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be sure to use the same major number on all nodes. Use the **lvlstmajor** command on each node to determine a free major number common to all nodes.

5. On each destination node, in turn, change the volume group to remain dormant at startup, using the **chvg** command. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
6. On each destination node, in turn, vary off the volume group, using the SMIT **varyoffvg** fastpath.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Extending a Shared Volume Group with AIX Commands

The figure below summarizes the steps you must complete on all cluster nodes to extend (add one or more physical volumes to) an existing shared volume group.



Source Node

Complete prerequisite tasks
Vary on volume group
Extend shared volume group
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
Import volume group
Change volume group characteristics
Vary off volume group
Complete follow-up tasks

Extending a Shared Volume Group

Note: The physical volumes (**hdisks**) being added to the volume group must be installed, configured, and available. They must have PVIDs on all nodes that can own the volume group.

The following procedure provides more detail about each step.

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on the nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, extend the volume group, using the SMIT **extendvg** fastpath. SMIT displays the Add a Physical Volume to a Volume Group screen.

Press the F4 key to obtain a list of volume groups from which you can select one. Press the F4 key again to obtain a listing of physical volumes from which you can pick one or more to add to the volume group. Press Enter and AIX adds the physical volumes to the volume group.

4. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
5. On each destination node, in turn, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.

6. On each destination node, in turn, import the volume group to make it known to the system using the SMIT **importvg** fastpath. On the **Import a Volume Group** SMIT screen, you must enter data for the following fields:

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 **yes**.

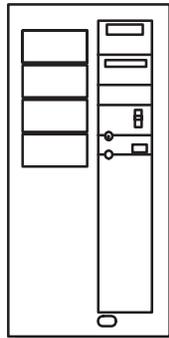
Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be sure to use the same major number on all nodes. Use the **lvlstmajor** command on each node to determine a free major number common to all nodes.

7. On each destination node, in turn, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
8. On each destination node, in turn, vary off the volume group, using the SMIT **varyoffvg** fastpath.
9. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Reducing a Shared Volume Group with AIX Commands

The figure below summarizes the steps you must complete on all cluster nodes to reduce (remove one or more physical volumes from) a shared volume group.



Source Node

Complete prerequisite tasks
Vary on volume group
Remove data from physical volume
Reduce shared volume group
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
Import volume group
Change volume group to remain dormant at startup
Vary off volume group
Complete follow-up tasks

Reducing a Shared Volume Group

The following procedure provides more detail about each step.

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, move data from the physical volume (or volumes) being removed from the volume group, using the SMIT **migratepv** fastpath. Set the **Move only data belonging to this LOGICAL VOLUME** field to **no**.

Warning: If you do not perform this step, data will be lost.

4. On the source node, reduce the size of the volume group, using the SMIT **reducevg** fastpath. SMIT displays the **Reduce a Volume Group** screen.

Select the **Remove a Physical Volume from a Volume Group** option and press Enter. SMIT prompts you to choose the volume group you want to reduce. Press the F4 key to obtain a list of volume groups.

After you select the volume group, SMIT displays the **Remove a Physical Volume from a Volume Group** screen.

Specify the names of the physical volumes you want to remove from the volume group. Set the value of the **FORCE deallocation of all partitions on this physical volume** field to **no**.

5. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
6. On each destination node, in turn, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.
7. On each destination node, in turn, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

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

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

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

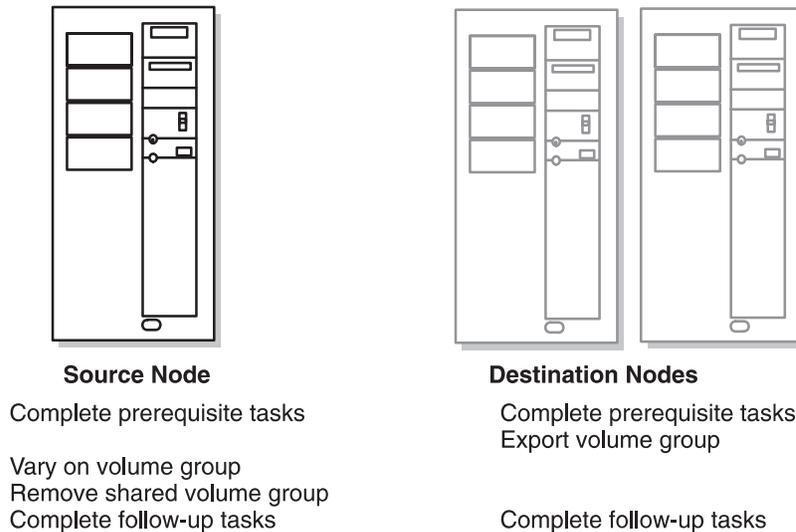
Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be sure to use the same major number on all nodes. Use the **lvlstmajor** command on each node to determine a free major number common to all nodes.

8. On each destination node, in turn, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
9. On each destination node, in turn, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Removing a Shared Volume Group with AIX Commands

The figure below summarizes the steps you must complete on all cluster nodes to remove a shared volume group.



Removing a Shared Volume Group

The following procedure provides more detail about each step.

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On each destination node, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.
3. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
4. On the source node, remove the shared volume group by using the SMIT **reducevg** fastpath. SMIT displays the **Reduce a Volume Group** screen:
Select the **Remove a Volume Group** option and press Enter. In the next screen, you must specify the volume group you want to remove. Press the F4 key to obtain a list of volume groups from which to choose. After making your selection, press Enter and AIX deletes the volume group.
5. Vary off the volume group on the source node.
6. Restart cluster services on all nodes on all cluster nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Using C-SPOC to Maintain Shared Volume Groups

You can use C-SPOC to do all the shared volume maintenance tasks.

Creating a Shared Volume Group with C-SPOC

Before creating a shared volume group for the cluster using C-SPOC, check that:

- All disk devices are properly attached to the cluster nodes
- All disk devices are properly configured on all cluster nodes and the device is listed as available on all nodes
- Disks have a PVID.

Take the following steps to create a shared volume group for a selected list of cluster nodes:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Logical Volume Manager > Shared Volume Groups > Add a Shared Volume Group**.

SMIT displays a list of cluster nodes.

3. Select two or more nodes from the list and press Enter.

The system correlates a list of all free physical disks that are available to all nodes selected. (Free disks are those disks that currently are not part of a volume group and have PVIDs.) SMIT displays the list of free physical disks in a multi-pick list by PVIDs.

4. Select one or more disks from the list and press Enter.

SMIT displays the **Add a Volume Group** screen.

5. Complete the selections as follows and press Enter.

VOLUME GROUP name Enter the name for this volume group.

Physical partition SIZE in megabytes Accept the default.

PHYSICAL VOLUME names Selected physical volumes are listed.

Volume group MAJOR NUMBER The system displays the number C-SPOC has determined to be correct.

Warning: Changing the volume group major number may result in the command's inability to execute on a node that does not have that major number currently available. Please check for a commonly available major number on all nodes before changing this setting.

C-SPOC verifies communication paths and version compatibility and then executes the command on all nodes in selection

Note: If the major number entered on the SMIT panel was not free at the time that the system attempted to make the volume group the command will display an error for the node that did not complete the execution and continue to the other nodes. At the completion of the command the volume group will not be active on any node in cluster.

Extending a Shared Volume Group with C-SPOC

Take the following steps to add a physical volume to a shared volume group using C-SPOC.

1. The physical volumes (**hdisks**) being added to the volume group must be installed, configured, and available on all nodes that can own the volume group. They must have PVIDs on all nodes that can own the volume group.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Shared Volume Groups > Set Characteristics of a Shared Volume Group > Add a Physical Volume to a Shared Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Select the ones you want to add to the shared volume group and Press Enter.
7. SMIT displays the **Add a Physical Volume to a Shared Volume Group** screen, with the following entries filled in.

Resource Group name The cluster resource group to which this shared volume group belongs.

Volume Group name Name of the shared volume group where hdisks are to be added.

Reference node Name of the node where the hdisks are found.

Physical Volume names Names of the hdisks to be added to the volume group.

8. If this screen reflects the correct information, press Enter to add the disks to the shared volume group. All nodes in the cluster receive this updated information.
9. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Importing a Shared Volume Group with C-SPOC

Take the following steps to import a volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.

2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume Manager > Shared Volume Groups > Import a Shared Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the **Import a Shared Volume Group** screen. Values for fields you have selected are displayed. For other fields, use the defaults or the appropriate entries for your operation:

Resource Group name The cluster resource group to which this shared volume group belongs.

VOLUME GROUP name The name of the volume group that you are importing.

PHYSICAL VOLUME name The name of one of the physical volumes that resides in the volume group. This is the hdisk name on the reference node.

Reference node Node from which the physical disk was retrieved.

Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

Make this VG concurrent capable? The default is **no**. For a non-concurrent VG, leave this **no**.

Make default varyon of VG Concurrent? The default is **no**. For a non-concurrent VG, leave this **no**.

8. If this screen reflects the correct information, press Enter to import the shared volume group. All nodes in the cluster receive this updated information.
9. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Removing a Physical Volume from a Shared Volume Group with C-SPOC

Take the following steps to remove a physical volume from a shared volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).

3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Shared Volume Groups > Set Characteristics > Remove a Physical Volume from a Shared Volume Group**.

SMIT displays a list of volume groups. Select the desired volume group and press Enter.

SMIT displays a list of physical volumes. Pick one and Press Enter.

SMIT displays the **Remove a Physical Volume from a Shared Volume Group** screen, with the following entries filled in.:

Resource Group name The cluster resource group to which this shared volume group belongs.

VOLUME GROUP name The name of the volume group that you are reducing.

Reference node Node from which the name of the physical disk was retrieved.

PHYSICAL VOLUME name The name of the physical volume that you want to remove. This is the hdisk name on the reference node.

5. If this screen reflects the correct information, press Enter to reduce the shared volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
6. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Mirroring a Volume Group Using C-SPOC

Take the following steps to mirror a shared volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Shared Volume Groups > Mirror a Shared Volume Group**.

SMIT displays a list of volume groups.

5. Select the desired volume group and press Enter.

SMIT displays a list of physical volumes.

6. Pick the desired physical volume and Press Enter.

SMIT displays the **Mirror a Shared Volume Group** screen, with the selected entries filled in. For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name	The name of the resource group to which this shared volume group belongs is displayed.
VOLUME GROUP name	The name of the volume group that you want to mirror is displayed.
Reference node	Node from which the name of the physical disk was retrieved is displayed.
PHYSICAL VOLUME names	The name of a physical volume on the volume group that you want to mirror. This is the hdisk name on the reference node.
Mirror sync mode	Foreground is the default. Other choices are Background and No Sync .
Number of COPIES of each logical partition	The default is 2 . You can also select 3 .
Keep Quorum Checking On?	The default is no . You can also choose yes .
Create Exact LV Mapping?	The default is no .

7. If this screen reflects the correct information, press Enter to mirror the shared volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
8. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Unmirroring a Volume Group Using C-SPOC

Take the following steps to unmirror a shared volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Shared Volume Groups > Unmirror a Shared Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.

7. SMIT displays the Unmirror a Shared Volume Group screen, with the chosen fields filled in. For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name The name of the resource group to which this shared volume group belongs is displayed.

VOLUME GROUP name The name of the volume group that you want to mirror is displayed.

Reference node Node from which the name of the physical disk was retrieved is displayed.

PHYSICAL VOLUME names The name of a physical volume on the volume group that you want to unmirror. This is the hdisk name on the reference node.

Mirror sync mode **Foreground** is the default. Other choices are **Background** and **No Sync**.

Number of COPIES of each logical partition The default is **2**. You can also select **3**.

8. If this screen reflects the correct information, press Enter to unmirror the shared volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Synchronizing Volume Group Mirrors

Take the following steps to synchronize shared LVM Mirrors by volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Synchronize Shared LVM Mirrors > Synchronize By Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Synchronize LVM Mirrors by Volume Group screen, with the chosen entries filled in. For other fields, use the defaults or the appropriate entries for your operation.

Resource Group Name The name of the resource group to which this shared volume group belongs is displayed.

VOLUME GROUP name The name of the volume group that you want to mirror is displayed.

Reference node The node from which the name of the physical disk was retrieved is displayed.

Number of Partitions to Sync in Parallel Leave empty. Range is 1 to 32.

Synchronize All Partitions The default is **no**.

Delay Writes to VG from other cluster nodes during this Sync .The default is **no**.

8. If this screen reflects the correct information, press Enter to synchronize LVM mirrors by the shared volume group. All nodes in the cluster receive this updated information.
9. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node. Synchronizing a Shared Volume Group Definition

Take the following steps to synchronize a shared volume group definition using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Synchronize a Shared Volume Group Definition** and press Enter.
5. SMIT displays the Synchronize a Shared Volume Group Definition screen. Here you must enter the name of the shared volume group to synchronize. Press F4 for a picklist, select the desired volume group and press Enter.
6. The command is run. All nodes in the cluster receive this updated information.
7. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Maintaining Shared Filesystems

The following administrative tasks involve shared filesystems:

- Creating a shared filesystem
- Extending a shared filesystem
- Changing a shared filesystem
- Removing a shared filesystem.

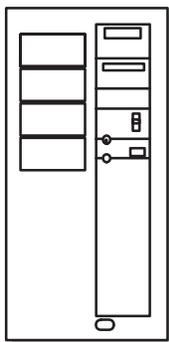
Each operation is described below. The sections also describe how to use the C-SPOC utility to create, change, or remove a shared filesystem in a cluster.

Creating a Shared Filesystem Using AIX Commands

This section describes how to use the **smit crjfs** fastpath to create a shared journaled filesystem. The **crfs** command also creates the logical volume associated with the journaled filesystem. If you prefer, you can use the **mklv** command to create a logical volume and then use the **mkfs** command to create the filesystem for the logical volume.

Note: You must have previously created the shared volume group that will contain the shared filesystem.

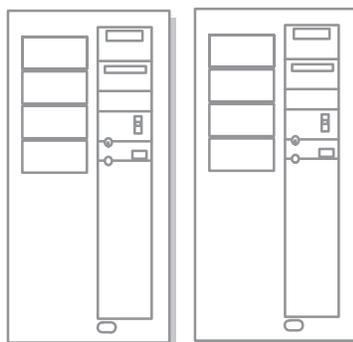
The figure below summarizes the steps you must complete on all cluster nodes to create a shared filesystem.



Source Node

- Complete prerequisite tasks
- Vary on volume group
- Create shared file system
- Check **/etc/filesystems** file
- Rename jfslog and logical volumes
- Check **/etc/filesystems** file
- Mirror logical volumes
- Vary off volume group

Complete follow-up tasks



Destination Nodes

- Complete prerequisite tasks

- Export volume group
- Import volume group
- Change volume group to remain dormant at startup
- Vary off volume group
- Complete follow-up tasks

Creating a Shared Filesystem on a Cluster Larger than Two Nodes

The following procedure provides more detail about each step.

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, create the filesystem, using the **crjfs SMIT** fastpath. SMIT prompts you to specify the volume group on which you want to create the filesystem. After you choose the volume group, SMIT displays the Add a Journalled Filesystem screen.
You must specify the size of the filesystem and the name of the mount point. Make sure the **Mount AUTOMATICALLY at system restart?** field is set to **no**.
4. On the source node, verify that the mount attribute for the filesystem in **/etc/filesystems** file is set to false. Setting the **mount** attribute to *false* prevents AIX from mounting the filesystem at boot time.
5. On the source node, rename the jfslog and logical volumes to ensure that they will be consistent across all cluster nodes. If you do not specify names for the jfslog and logical volume, AIX assigns them a name which could be vary on each node.
Use the **lsvg -l volume_group** command to determine the name of the jfslog and the logical volume associated with the filesystem. In the output from this command, look for the logical volume with the type jfs and the log logical volume that has type jfslog.
Use the SMIT **chlv** fastpath to rename these logical volumes. Use a naming convention that makes the relationship of the logical volumes clear.
6. On the source node, check the **dev** and **log** attributes in the **/etc/filesystems** file to make sure they reflect the change.
7. On the source node, add copies to logical volume, using the SMIT **mklvcopy** fastpath. Add copies to both the **jfslog** log logical volume and the logical volumes in the shared filesystems. To avoid space problems, first mirror the **jfslog** log logical volume and then the shared logical volumes.

Note: This step does not apply to RAID devices using RAID levels 3 or 5. They provide their own data redundancy.

Enter the specific field values as follows and press Enter to create the copies. For other fields use the defaults or the appropriate entries for your operation:

Allocate each logical partition copy on a SEPARATE physical volume? Specify **yes**. Doing so ensures that copies reside on separate disks.

SYNCHRONIZE the data in the new logical partition copies Specify **yes**.

To verify the number of logical volume copies, enter:

```
lsvg -l volume_group_name
```

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

To verify the placement of logical volume copies, enter:

```
lspv -l hdiskx
```

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

8. On the source node, run a consistency check on each filesystem. Enter:

```
fsck /filesystem_name
```

Verify that you can mount the filesystem by entering:

```
mount /filesystem_name
```

Verify that you can unmount the filesystem by entering:

```
umount /filesystem_name
```

9. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. On each destination node, in turn, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.
11. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

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

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

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

Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

12. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
13. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
14. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Creating a Shared Filesystem Using C-SPOC

Before creating a journaled filesystem for the cluster using C-SPOC, check that:

- All disk devices are properly attached to the cluster nodes
- All disk devices are properly configured and available on all cluster nodes
- The volume group that will contain the filesystem must be varied on on at least one cluster node.

You can add a journaled filesystem to either:

- A shared volume group (no previously defined cluster logical volume)
- A previously defined cluster logical volume (on a shared volume group).

To add a filesystem where no logical volume is currently defined:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Logical Volume Manager > Shared Filesystems > Add a Journaled Filesystem**.

SMIT displays a list of filesystem types (Standard, Compressed or Large File Enabled).

3. Select the desired filesystem type from the list.
SMIT generates a list of all volume groups in the cluster.
4. Select the volume group where the filesystem will be added.
SMIT displays the AIX SMIT screen for selecting filesystem attributes.

Node Names	The names of the selected cluster nodes are displayed.
Volume Group Name	The selected volume group name is displayed.
MOUNT POINT	Enter the mount point for the filesystem.
PERMISSIONS	Set as desired.
Mount OPTIONS	Set as desired.
Start Disk Accounting?	Set as desired. Default is no .
Fragment Size (Bytes)	4096 is the default.

Number of Bytes per node 4096 is the default.

Compression algorithm Default is **no**.

5. Select the filesystem attributes and press Enter.
SMIT checks the list of nodes that can own the resource group that contains the volume group, creates the logical volume (on an existing log logical volume if present, otherwise it will create a new log logical volume) and adds the filesystem to the node where the volume group is varied on. All other nodes in the resource group will run an **importvg -L** .

To add a filesystem on a previously defined logical volume:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Logical Volume Manager > Shared Filesystems > Add a Journaled Filesystem to a Previously Defined Logical Volume**.

SMIT displays a list of filesystem types (Standard, Compressed or Large File Enabled).

3. Select the desired filesystem type from the list.
SMIT generates a list of all free logical volumes in the cluster and nodes they are on. SMIT reports a logical volume as free if:
 - the logical volume is part of a parent volume group that is configured as a resource in the cluster
 - the logical volume is varied on prior to and during the system polling the disk for logical volume information
 - the logical volume does not have a filesystem mount point.
4. Select a logical volume where the filesystem will be added.
SMIT displays the AIX SMIT screen for selecting filesystem attributes.

LOGICAL VOLUME name The name of the selected logical volume is displayed.

Node Names The names of the selected cluster nodes are displayed.

MOUNT POINT Enter the mount point for the filesystem.

PERMISSIONS Set as desired.

Mount OPTIONS Set as desired.

Start Disk Accounting? Set as desired. Default is **no**.

Fragment Size (Bytes) 4096 is the default.

Number of Bytes per node 4096 is the default.

Compression algorithm Default is **no**.

5. Select the filesystem attributes and press Enter.

SMIT checks the list of nodes that can own the resource group which contains the volume group where the logical volume is located. It adds the filesystem to the node where the volume group is varied on. All other nodes in the resource group will run an **importvg -L**.

Changing a Shared Filesystem in a Cluster

As system administrator of an HACMP for AIX cluster, you may need to change the characteristics of an existing filesystem. The following sections describe how to perform this task in a cluster using either the C-SPOC utility or using standard AIX commands.

Changing a Filesystem using C-SPOC

Using the C-SPOC utility, you can change the characteristics of a shared filesystem on cluster nodes by executing a command on a single cluster node. The C-SPOC command changes the attributes of the shared filesystem on the node which currently has the shared volume group varied on. The definition of the filesystem on other cluster nodes is not updated immediately; it is updated when the filesystem is activated on these nodes.

1. Vary on the volume group, if needed, by using the **lsvg** command. You can use the C-SPOC utility to change a filesystem even if the volume group on which it is defined is varied off. In this case, specify the **-f** flag. See the **cl_chfs** man page for more information.
2. Use C-SPOC to change attributes of the filesystem. Enter the SMIT fastpath:

```
smit hacmp
```

SMIT displays the main HACMP for AIX menu.

3. Select **Cluster System Management > Cluster Logical Volume Manager > Shared File > Change/Show Characteristics of a Shared Filesystem in the Cluster**.

SMIT displays a pick list of existing filesystems from which you must select one. After you select the filesystem, SMIT displays a screen containing the characteristics of the filesystem you can change using the C-SPOC utility.

4. Enter data in the fields you want to change and press Enter.

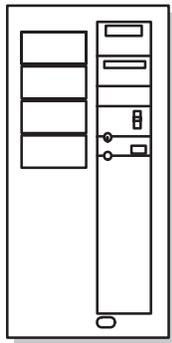
The C-SPOC utility changes the filesystem characteristics on the local node. The filesystem characteristics are not updated on remote nodes until the volume group is activated on each node.

To check the status of the C-SPOC command execution on cluster nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Changing a Shared Filesystem Using AIX Commands

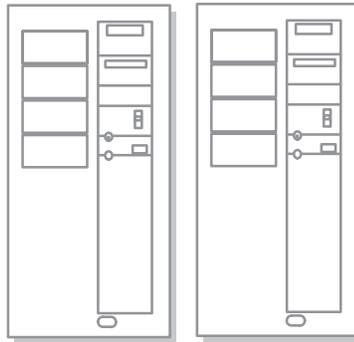
To change the characteristics of a filesystem on a cluster larger than two nodes, you must make the changes on one cluster node, the source node, and then import the changed filesystems on all the other cluster nodes. The figure below summarizes the steps you must complete on all cluster nodes to change a shared filesystem.



Source Node

Complete prerequisite tasks
Vary on volume group
Extend file system
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
Import volume group
Change volume group to remain dormant at startup
Vary off volume group
Complete follow-up tasks

Changing a Shared Filesystem on Cluster Larger Than Two Nodes

As shown in the figure, to change a shared filesystem on a cluster larger than two nodes:

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, change the characteristics of the shared filesystem, using the SMIT **chjfs** fastpath.
Enter data in the fields you want to change and press Enter.
4. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
5. On each destination node, in turn, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.

6. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

VOLUME GROUP name	Enter the name of the volume group that you are importing. Make sure the volume group name is the same name that you used on the source node.
PHYSICAL VOLUME name	Enter the name of one of the physical volumes that resides in the volume group. <i>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.</i> 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 yes .
Volume Group MAJOR NUMBER	If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

7. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
8. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
9. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Removing a Shared Filesystem

As system administrator of an HACMP cluster, you may need to remove a filesystem. You can optionally remove the filesystem's mount point as part of the same operation. The following sections describe how to perform this task on a cluster using the C-SPOC utility, and on any cluster using standard AIX commands.

Removing a Shared Filesystem Using C-SPOC

Using the following procedure, you can remove a shared filesystem on any node in a cluster by executing a C-SPOC command on that node. The C-SPOC command deletes the shared filesystem on the node which currently has the shared volume group varied on. The command removes both the shared logical volume on which the filesystem resides and the associated stanza in the **/etc/filesystems** file. These changes occur on other cluster nodes when the filesystem is activated on each node.

The filesystem being deleted must not be configured as a resource in the resource group specified as an argument in the C-SPOC command line.

1. On the source node, vary on the volume group, if needed, using the SMIT **varyonvg** fastpath. You can use the C-SPOC utility on a volume group that is varied off; however, you must specify the **-f** flag. See the **cl_rmfs** man page for more information.

2. Use C-SPOC to delete the filesystem. Enter the fastpath:

```
smit cl_admin
```

3. Select the **Cluster Logical Volume Manager** option and press Enter.
4. Choose the **Shared Filesystems** option and press Enter.
5. Select the **Remove a Shared Filesystem** option and press Enter.

SMIT displays the Remove a Journaled Filesystem from the Cluster screen.

6. Press the F4 key to obtain a pick list of existing filesystems from which you may select one. Toggle the **Remove Mount Point** option to **yes** if you want to remove the mount point in the same operation. When you finish entering data, press Enter.

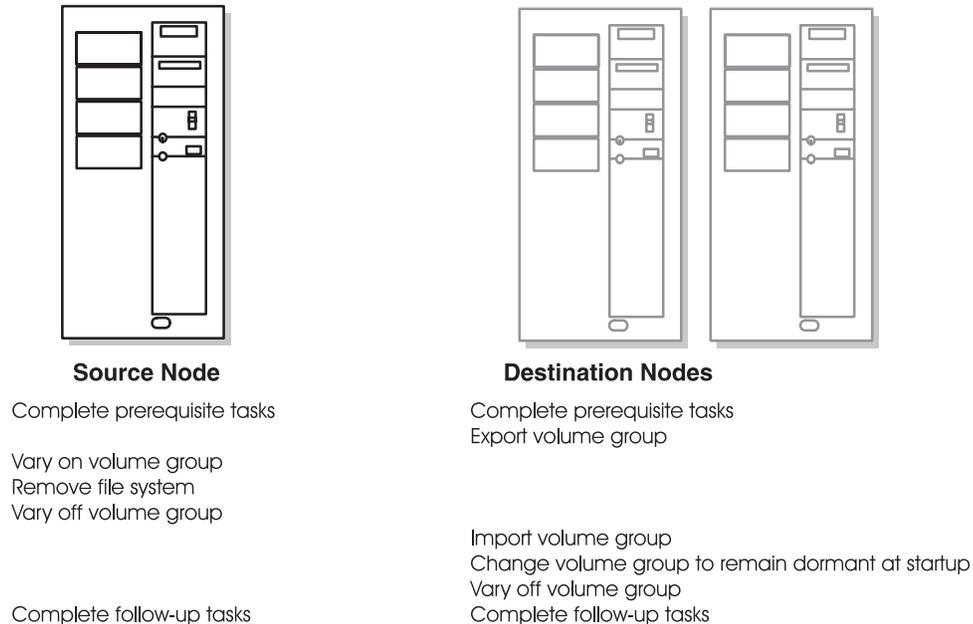
The C-SPOC utility removes the filesystem on the source (local) node. The filesystem is not removed on remote nodes until the volume group on which the filesystem is defined is activated.

To check the status of the C-SPOC command execution on both nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Removing a Shared Filesystem Using AIX Commands

The figure below summarizes the steps you must complete on the source and destination nodes to remove a shared filesystem.



Removing a Shared Filesystem

The following procedure provides more detail about each step.

To remove a shared filesystem using AIX commands:

1. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services for more information.
2. On the destination nodes, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.
3. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
4. On the source node, delete the shared filesystem, using the SMIT **rmjfs** fastpath.
Press the F4 key to obtain a pick list of existing filesystems from which you may select one. Toggle the **Remove Mount Point** option to **yes** if you want to remove the mount point in the same operation. When you finish entering data, press Enter.
5. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.

6. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

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

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

ACTIVATE volume group Set the field to **yes**.

Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

7. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath. Set the **Activate volume group AUTOMATICALLY at system restart?** option to **no**.
8. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
9. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Maintaining Logical Volumes

The following administrative tasks involve shared logical volumes. You can perform all these tasks using AIX commands, or using the C-SPOC utility (except for creating a shared logical volume):

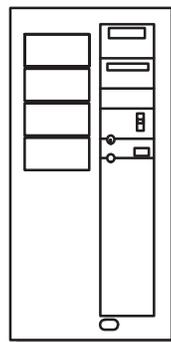
- Creating a shared logical volume
- Changing a shared logical volume (renaming, increasing size, adding or removing copy)
- Adding or removing a shared logical volume

Note: Mirroring, unmirroring, or increasing or decreasing the number of copies (mirrors) of a shared logical volume: *This task does not apply to RAID devices.*

Creating a Logical Volume

The figure below summarizes the steps you must complete on the source and destination nodes to create a logical volume.

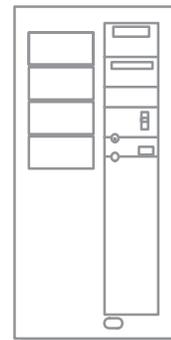
Note: The volume group on which you are creating the logical volume must already exist.



Source Node

Complete prerequisite tasks
Vary on volume group
Create logical volume
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
Import volume group
Change volume group to remain dormant at startup
Vary off volume group
Complete follow-up tasks

Creating a Logical Volume

The following procedure provides more detail about each step.

1. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.

3. On the source node, create the logical volume, using the SMIT **mklv** fastpath.
4. Enter the specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

Logical Volume NAME	Enter the name of the logical volume. Make sure that the name you assign to this logical volume is unique within the cluster.
Number of COPIES of each logical partition	Enter the number of logical volume copies (mirrors). Entering a value of three eliminates the physical volume as a single point of failure. If using a RAID device, set this field to 1. These devices provide their own data redundancy if you use RAID level 1, 3, or 5 (not 0).
Mirror Write Consistency	Set this field to yes , for nonconcurrent environments, or to no for concurrent environments.
Allocate each logical partition copy on a SEPARATE physical volume	Make sure this field is set to yes . (Does not apply to RAID devices.)
ENABLE BAD BLOCK relocation	Set this field to no to disable bad block relocation (applies to RAID devices).

5. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
6. On the destination nodes, export the volume group, using the SMIT **exportvg** fastpath.
7. On the destination nodes, import the volume group, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

VOLUME GROUP name	Enter the name of the volume group that you are importing. Make sure the volume group name is the same name that you used on the source node.
PHYSICAL VOLUME name	Enter the name of one of the physical volumes that resides in the volume group. <i>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.</i> 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 yes .

**Volume Group
 MAJOR NUMBER**

If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

8. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath.
9. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. Restart cluster services on all nodes.

Adding a Logical Volume to a Cluster Using C-SPOC

To add a logical volume to a cluster using C-SPOC, take the following steps:

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can add a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_mklv** command man page for more information.
2. Enter the SMIT C-SPOC fastpath:

```
smit cl_admin
```
3. Select these options: **Cluster Logical Volume Manager > Shared Logical Volumes > Add a Shared Logical Volume**.
4. SMIT displays a pickup list of shared volume groups. Select one and press Enter.
5. SMIT displays a list of physical volumes. Select one and press Enter.

The **Add a Shared Logical Volume** screen appears, with chosen fields filled in as shown in the sample below:

Resource Group name	casg_rg
VOLUME GROUP name	sharedvg
Reference node	a1
* Number of LOGICAL PARTITIONS	[]
PHYSICAL VOLUME names	hdisk16
Logical volume NAME	[]
Logical volume TYPE	[]
POSITION on physical volume	middle
RANGE of physical volumes	minimum

MAXIMUM NUMBER of PHYSICAL VOLUMES to use for allocation	<input type="checkbox"/>
Number of COPIES of each logical partition	1
Mirror Write Consistency?	yes
Allocate each logical partition copy on a SEPARATE physical volume?	yes
RELOCATE the logical volume during reorganization	yes
Logical volume LABEL	<input type="checkbox"/>
MAXIMUM NUMBER of LOGICAL PARTITIONS	[512]
Enable BAD BLOCK relocation?	yes
SCHEDULING POLICY for writing logical partition copies	parallel
Enable WRITE VERIFY?	no
File containing ALLOCATION MAP	<input type="checkbox"/>
Stripe Size?	[Not Striped]

6. The default characteristics are most common. Make changes if necessary for your system and press Enter. Other cluster nodes are updated with this information.

Changing the Characteristics of a Shared Logical Volume

As system administrator of an HACMP for AIX cluster, you may need to change the characteristics of an existing logical volume. The following sections describe how to perform this task on a cluster both by using the C-SPOC utility and by using standard AIX commands.

Changing a Shared Logical Volume on a C-SPOC Cluster

Using the following procedure, you can change the characteristics of a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can change a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_chlv** command man page for more information.
2. Use the C-SPOC utility to change the logical volume. Enter the SMIT fastpath:

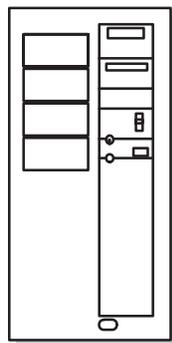
```
smit cl_admin
```

3. Select the **Cluster Logical Volume Manager** option and press Enter.
4. Choose the **Shared Logical Volumes** option and press Enter.
5. Select **Change a Shared Logical Volume > Change a Logical Volume on the Cluster** and press Enter.
6. SMIT displays a picklist of existing logical volumes from which you may select one. When you finish entering data, press Enter. SMIT displays the screen, with the values of the chosen logical volume with current attributes filled in.
7. Enter data in the fields you want to change and press Enter. The C-SPOC utility changes the characteristics on the local node. The logical volume definition is updated on remote nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Changing a Shared Logical Volume Using AIX Commands

To change the characteristics of a shared logical volume on large clusters, you must make the changes on one cluster node, the source node, and then import the changes on all other cluster nodes. The figure below summarizes the steps you must complete on all cluster nodes.



Source Node

Complete prerequisite tasks
Vary on volume group
Change the logical volume
Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
Import volume group
Change volume group to remain dormant at startup
Vary off volume group
Complete follow-up tasks

Changing a Shared Logical Volume on Cluster Larger Than Two Nodes

The following procedure provides more detail about each step.

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.

2. On the source node, vary on the volume group, if necessary, using the SMIT **varyonvg** fastpath.
3. On the source node, change the characteristics of the shared logical volume, using the SMIT **chlv** fastpath.
4. Select the **Change a Logical Volume** option and press Enter. SMIT displays the **Change a Logical Volume on the Cluster** screen.
5. Press the F4 key to obtain a pick list of existing logical volumes from which you may select one. When you finish entering data, press Enter. SMIT displays the screen with the values of the current logical volume attributes filled in.
6. Enter data in the fields you want to change and press Enter. The C-SPOC utility changes the filesystem characteristics on the local node.
7. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

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

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

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

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

8. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath.
9. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. Restart cluster services on all nodes.

Warning: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Removing a Logical Volume

As system administrator of an HACMP cluster, you may need to remove a logical volume. The following sections describe how to perform this task on a cluster using the C-SPOC utility, and on any cluster using standard AIX commands.

Removing a Logical Volume on a C-SPOC Cluster

Using the following procedure, you can remove a logical volume on any node in a cluster by executing a C-SPOC command on one of the nodes.

1. On a cluster node, vary on the volume group, if needed, using the SMIT **varyonvg** fastpath. You can remove a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_rmlv** command man page for more information.

Note: If the logical volume to be removed contains a filesystem, you first must remove the filesystem from any specified resource group before attempting to remove the logical volume. Afterwards, be sure to synchronize cluster resources on all cluster nodes.

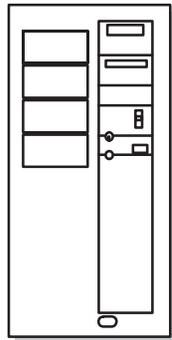
2. Use the C-SPOC utility to change the logical volume. Enter the SMIT fastpath:

```
smit cl_admin
```
3. Select the **Cluster Logical Volume Manager** option and press Enter.
4. Choose the **Shared Logical Volumes** option and press Enter.
5. Select the **Remove a Shared Logical Volume** option and press Enter.
6. C-SPOC provides a list of shared logical volumes, organized by HACMP resource group. Select the logical volume you want to remove and press Enter.

All remote nodes are updated.

Removing a Logical Volume on a Cluster Using AIX Commands

To remove a logical volume, you must remove the logical volume on one cluster node, the source node, and then import the volume group on which the logical volume is defined on all the other cluster nodes. The figure below summarizes the steps you must complete on the source and destination nodes to remove a logical volume.



Source Node

Complete prerequisite tasks

- Vary on volume group
- Remove the logical volume
- Vary off volume group
- Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks.
 Export the volume group

- Import the volume group
- Change it to remain dormant at startup
- Vary off the volume group

1. Complete prerequisite tasks. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the destination nodes, export the volume group to remove the definition from the system, using the SMIT **exportvg** fastpath.
3. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
4. On the source node, remove the shared logical volume, using the SMIT **rmlv** fastpath. You can accept the defaults for all fields.
5. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
6. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

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

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

ACTIVATE volume group Set the field to **yes**.

**Volume Group
MAJOR NUMBER**

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

7. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath.
8. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
9. Restart cluster services on all nodes.

After completing this procedure, verify failover behavior.

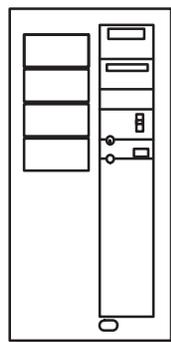
Adding Copies (Mirrors) to a Shared Logical Volume

You should use at least two copies—and preferably three—per logical volume, each on a separate physical volume, to maintain high availability.

Note: This task does not apply to RAID devices, which provide their own data redundancy if you use RAID level 1, 3, or 5 (not 0). Also, you should have created the logical volume to which you are adding copies.

Adding Copies to a Shared Logical Volume Using AIX Commands

The figure below summarizes the steps you must complete on the source and destination nodes to add copies to a logical volume.



Source Node

- Complete prerequisite tasks
- Vary on volume group
- Add copies to logical volume
- Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

- Export volume group
- Import volume group
- Change volume group to remain dormant at startup
- Vary on volume group
- Complete follow-up tasks

Mirroring a Logical Volume

The following procedure provides more detail about each step.

1. Stop HACMP for AIX cluster services on nodes in the resource group chain. See Chapter 2, Starting and Stopping Cluster Services, for more information.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, add the mirror copies to the logical volume, using the SMIT **mklvcopy** fastpath. Enter the specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

NEW TOTAL number of logical partition copies Specify the total number of copies. As many as three copies are allowed, and three is the recommended number.

PHYSICAL VOLUME names Specify the physical disk volumes on which the logical volume copies will reside. Press F4 to list the disks available in the volume group.

The copies, if possible, should also reside on disks that are controlled by different disk adapters and are located in separate drawers or units. Because copies should reside on separate disks, you should specify as many disks as there are copies.

Allocate each logical partition copy on a SEPARATE physical volume? Specify **yes**. Doing so ensures that copies reside on separate disks.

Mirror Write Consistency Set this field to **yes**.

ENABLE BAD BLOCK relocation Set this field to **no** to disable bad block relocation (applies to RAID devices).

SYNCHRONIZE the data in the new logical partition copies Specify **yes**.

4. Verify the creation of the logical volume copies, using the **lsvg -l volume_group** command. In the output from the command, check 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. To check that the copies were created on the correct physical volumes, use the **lspv -l hdiskx** command.

5. Test the filesystem (if the logical volume contains a filesystem) by running a consistency check on each filesystem. Enter:

```
fsck /filesystem_name
```

Verify that you can mount the filesystem by entering:

```
mount /filesystem_name
```

Verify that you can unmount the filesystem by entering:

```
umount /filesystem_name
```

6. On the destination nodes, export the volume group to remove its definition from the system, using the SMIT **exportvg** fastpath.

7. On the destination nodes, import the volume group to make it known to the system, using the SMIT **importvg** fastpath. Enter specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:

VOLUME GROUP name	Enter the name of the volume group that you are importing. Make sure the volume group name is the same name that you used on the source node.
PHYSICAL VOLUME name	Enter the name of one of the physical volumes that resides in the volume group. <i>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.</i> 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 yes .
Volume Group MAJOR NUMBER	If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

8. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath.
9. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. Restart cluster services on all nodes.

After completing the procedure, verify failover behavior.

Setting Characteristics of a Shared Logical Volume Using C-SPOC

You can use C-SPOC to do the following tasks for all cluster nodes from one node:

- Rename a shared logical volume
- Increase the size of a shared logical volume
- Add copies to a shared logical volume
- Remove copies from a shared logical volume.

Renaming a Shared Logical Volume Using C-SPOC

Using the following procedure, you can rename a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can rename a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_chlv** command man page for more information.

2. Use the C-SPOC utility to rename the logical volume. Enter the SMIT fastpath:

```
smit cl_admin
```

3. Select these options: **Cluster Logical Volume Manager > Shared Logical Volumes > Set Characteristics of A Shared Logical Volume > Rename a Shared Logical Volume** and press Enter.

SMIT displays the **Rename a Logical Volume on the Cluster** screen.

4. Press Enter to obtain a pick list of existing logical volumes; select one and press Enter. SMIT displays a screen with the **Resource group name** and **Current logical volume name** filled in.

5. Enter the new name in the **NEW logical volume name** field and press Enter. The C-SPOC utility changes the name on all cluster nodes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Caution: After completing this procedure, confirm your changes by initiating failures and verifying correct failover behavior before resuming normal user operations.

Increasing the Size of a Shared Logical Volume Using C-SPOC

Using the following procedure, you can increase the size of a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can rename a logical volume using the C-SPOC command when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_chlv** command man page for more information.

2. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```

3. Select these options: **Cluster Logical Volume Manager > Shared Logical Volumes > Set Characteristics of a Shared Logical Volume > Increase Size of a Shared Logical Volume** and press Enter.

SMIT displays a list of logical volumes arranged by resource group.

4. Choose the desired logical volume from the pick list and press Enter. SMIT displays a list of physical volumes.
5. Choose a physical volume and press Enter. SMIT displays the **Increase Size of a Shared Logical Volume** screen with the Resource Group, Logical Volume, Reference Node and default fields filled in.

6. Enter the new size in the **Number of ADDITIONAL logical partitions** field and press Enter. The C-SPOC utility changes the size of this logical volume on all cluster nodes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Adding a Copy to a Shared Logical Volume Using C-SPOC

Using the following procedure, you can add a copy to a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can rename a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_chlv** command man page for more information.

2. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```

3. Select these options: **Cluster Logical Volume Manager > Shared Logical Volumes > Set Characteristics of a Shared Logical Volume > Add a Copy to a Shared Logical Volume** and press Enter.

SMIT displays a list of logical volumes arranged by resource group.

4. Choose the desired logical volume from the pick list and press Enter.

SMIT displays a list of physical volumes.

5. Choose a physical volume and press Enter. SMIT displays the **Add a Copy to a Shared Logical Volume** screen with the Resource Group, Logical Volume, Reference Node and default fields filled.

6. Enter the new number of mirrors in the **NEW TOTAL number of logical partitions** field and press Enter. The C-SPOC utility changes the number of copies of this logical volume on all cluster nodes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Removing a Copy from a Shared Logical Volume Using C-SPOC

Using the following procedure, you can remove a copy of a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can rename a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **cspoc -f** flag. See the **cl_chlv** command man page for more information.

2. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```

3. Select these options: **Cluster Logical Volume Manager > Shared Logical Volumes > Set Characteristics of A Shared Logical Volume > Remove a Copy from a Shared Logical Volume** and press Enter.

SMIT displays a list of logical volumes arranged by resource group.

4. Choose the desired logical volume from the pick list and press Enter. SMIT displays a list of physical volumes.

5. Choose the physical volumes from which you want to remove copies and press Enter. SMIT displays the **Remove a Copy from a Shared Logical Volume** screen with the Resource Group, Logical Volume name, Reference Node and Physical Volume names fields filled in.
6. Enter the new number of mirrors in the **NEW maximum number of logical partitions copies** field and check the **PHYSICAL VOLUME name(s) to remove copies from** field to make sure it is correct and press Enter. The C-SPOC utility changes the number of copies of this logical volume on all cluster nodes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Synchronizing LVM Mirrors by Logical Volume

Take the following steps to synchronize shared LVM Mirrors by logical volume using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the shared volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Synchronize Shared LVM Mirrors > Synchronize By Logical Volume**.
5. SMIT displays a list of logical volumes. Select the desired logical volume and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Synchronize LVM Mirrors by Volume Group screen, with the chosen entries filled in. For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name	The name of the resource group to which this logical volume belongs is displayed.
LOGICAL VOLUME name	The name of the logical volume that you want to synchronize is displayed.
Reference node	Node from which the name of the physical disk was retrieved is displayed.
Number of Partitions to Sync in Parallel	Leave empty. The range is 1 to 32.
Synchronize All Partitions	The default is no .
Delay Writes to VG from other cluster nodes during this Sync	The default is no .

8. If this screen reflects the correct information, press Enter to synchronize LVM mirrors by the shared logical volume. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the shared volume group varied on, vary off the volume group on that node.

Maintaining Physical Volumes

Administrative tasks that involve shared physical volumes include:

- Adding a disk definition to cluster nodes
- Removing a disk definition from cluster nodes
- Moving (migrating) the data on one physical volume to other physical volumes within the shared volume group.

Adding a Disk Definition to Cluster Nodes Using C-SPOC

The nodes must be configured as part of the cluster. Only SSA and SCSI disks are currently supported for this utility.

Take the following steps to add a raw disk on selected cluster nodes:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Physical Volume Manager > Add a Disk to the Cluster**.
SMIT displays a list of nodes in the cluster and prompts you to select the nodes where the disk definition should be added
3. Select one or more nodes where you want to have the new disk configured.
The system generates a list of available disk types based on those disk types known to the first node in the list (above).
4. Select the type of disk you want to add to the cluster.
The set of screens that follow depend on the disk type selected. Possible disk types include:
 - SCSI (various types listed)
 - SSA (available types listed)

Adding a SCSI Disk

If you are select a SCSI disk type to define, SMIT displays a multi-select list of parent adapter/node name pairs. You are prompted to select one parent adapter per node.

- 4a. Select a parent adapter for each node connected to the disk and accept the selection.

SMIT displays the AIX SCSI **Add a Disk** screen with all entries filled in except Connection Address.

Disk type	1000mb
Disk interface	scsi
Description	1.0GB SCSI Disk Drive

Parent adapter	nodea,vscsi0 nodeb,vscsi1,...etc.
CONNECTION address	Use F4 to display the list; select the address.
ASSIGN physical volume identifier	no is the default.

4b. Select the connection address and press Enter.

C-SPOC executes the necessary commands to define the disk on all selected nodes.

Adding an SSA Disk

If you select an SSA disk to define, SMIT displays the AIX **Add an SSA Logical Disk** screen.

Disk type	hdisk
Disk interface	ssa
Description	SSA Logical Disk Drive
Parent	ssar
CONNECTION address	Press F4 for a list; add the address.The connection address must be the same on all nodes.
Location Label	(Optional entry). Press F1 for information.
ASSIGN physical volume identifier	no is the default.
RESERVE disk on open	yes is the default.
Queue depth	(Optional entry). Press F3 for information on ranges allowed. Press F1 for help.
Maximum coalesce	(Optional entry). Press F3 for information on ranges allowed. Press F1 for help.

4a. Select the connection address and other attributes as desired and press Enter.

C-SPOC executes the necessary commands to define the disk on all selected nodes.

Removing a Disk Definition on Cluster Nodes Using C-SPOC

Before removing a disk from the cluster using C-SPOC, check that the disk to be removed is not currently part of an existing volume group. If it is, use the C-SPOC **cl_reducevg** command to remove a physical volume from a volume group.

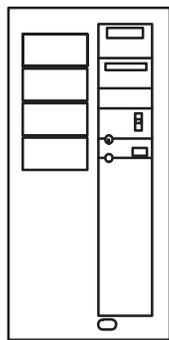
Take the following steps to remove a configured disk on all selected nodes in the cluster:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Physical Volume Manager > Remove a Disk from the Cluster**.
 SMIT displays a list of nodes in the cluster that currently have the disk configured, and prompts you to select the nodes where the disk should be removed.
3. Select the nodes from which you want to remove the disk configuration. (You may have removed the cable from some of the nodes in the cluster and only want the disk configuration removed from those nodes.)
 SMIT displays the AIX **Remove a Disk** screen with the selected disks displayed.
4. For the entry “Keep the disk definition in database” Select **yes** to keep the definition in the database; select **no** to delete the disk from the database. Press Enter.
 C-SPOC sends the **rmdev** command to all nodes listed to remove the selected disk.

Migrating Data

Use AIX commands to migrate data. The figure below summarizes the steps you must complete on the source and destination nodes to migrate data from one physical volume to another physical volume.

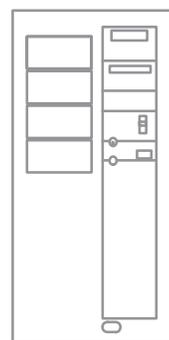
Note: The physical volumes (**hdisks**) should be installed, configured, and available.



Source Node

Complete prerequisite tasks
 Vary on the volume group
 Migrate data
 Vary off volume group

Complete follow-up tasks



Destination Nodes

Complete prerequisite tasks

Export volume group
 Import volume group
 Change volume group to remain dormant at startup
 Vary off volume group
 Complete follow-up tasks

Migrating Data

To migrate data from one physical volume to another, perform the following procedure.

1. Stop HACMP for AIX cluster services on nodes in the resource group chain.
2. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath.
3. On the source node, move data from the physical volume (or volumes) being removed from the volume group, using the SMIT **migratepv** fastpath. Leave the **Move only data belonging to this LOGICAL VOLUME** field blank.

Warning: If you do not perform this step, data will be lost.

4. On the source node, reduce the size of the volume group, using the SMIT **reducevg** fastpath. SMIT displays the **Reduce a Volume Group** screen.

Select the **Remove a Physical Volume from a Volume Group** option and press Enter. SMIT prompts you to choose the volume group you want to reduce. Press the F4 key to obtain a list of volume groups.

5. On the source node, vary off the volume group, using the SMIT **varyoffvg** fastpath.
6. On the destination nodes, export the volume group, using the SMIT **exportvg** fastpath.
7. On the destination nodes, import the volume group, using the SMIT **importvg** fastpath.

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

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

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

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

Volume Group MAJOR NUMBER If you are not using NFS, use the default (which is the next available number in the valid range). If you are using NFS, you must be 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.

8. On the destination nodes, change the volume group to remain dormant at startup, using the SMIT **chvg** fastpath.
9. On the destination nodes, vary off the volume group, using the SMIT **varyoffvg** fastpath.
10. Restart cluster services on all nodes.

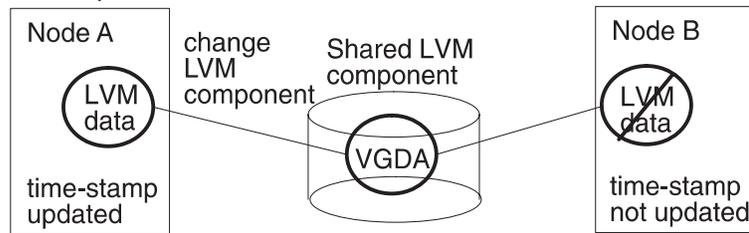
Lazy Update Processing in an HACMP Cluster

Procedures included in this chapter describe how to update the LVM data on the other cluster nodes by exporting the volume group and then re-importing it. (In the import process, data describing the volume group is copied from the VGDA on the disk.)

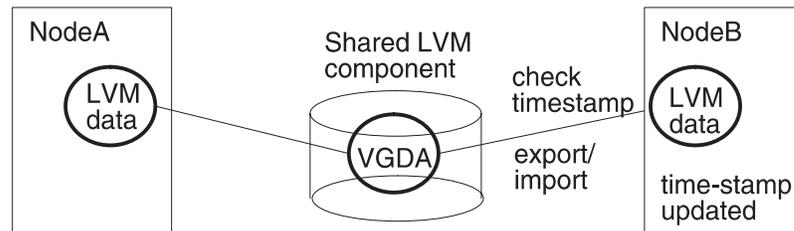
For LVM components under the control of HACMP for AIX, you do not have to explicitly export and import to bring the other cluster nodes up-to-date. Instead, HACMP for AIX can perform the export and import when it activates the volume group during a failover. (In a cluster, HACMP controls when volume groups are activated.) HACMP for AIX implements a function, called *lazy update*, by keeping a copy of the time-stamp from the volume group's VGDA. AIX updates this time-stamp whenever the LVM component is modified. When another cluster node attempts to vary on the volume group, HACMP for AIX compares its copy of the time-stamp with the time-stamp in the VGDA on the disk. If the values are different, the HACMP for AIX software exports and re-imports the volume group before activating it. If the time-stamps are the same, HACMP for AIX activates the volume group without exporting and re-importing.

The following figure illustrates how a lazy update can be achieved in a cluster using the C-SPOC utility. All nodes in the cluster are updated accordingly, if they are specified from the node originating the command.

1. LVM component modified on Node A:



2. LVM Data on Node B after lazy update:



Lazy Update Processing

Forcing an Update Before Fallover

In certain circumstances, you may want to update the LVM definition on remote cluster nodes before a fallover occurs. For example, if you rename a logical volume using C-SPOC, the LVM data describing the component is updated on the local node and is updated in the VGDA on the disks, as previously described. If you attempt to rename the logical volume a second time using C-SPOC; the operation will fail if the LVM data on any other cluster node has not been updated.

To update the LVM data on a remote node, use the C-SPOC **/usr/sbin/cluster/sbin/cl_updatevg** command. This command causes the specified remote node to export and import the LVM data, whether or not the time-stamp associated with the LVM component is the same. For more information about the **cl_updatevg** command, see Appendix A or its man page.

Note: The volume group must be varied off on all nodes accessing the shared LVM component before running the **cl_updatevg** command.

Chapter 5 Maintaining Shared LVM Components in a Concurrent Access Environment

This chapter explains how to maintain shared LVM components in a concurrent access environment.

While the maintenance tasks for shared LVM components in concurrent access environments are similar to those of non-concurrent access environments, described in Chapter 4, *Maintaining Shared LVM Components*, there are some special considerations in a concurrent access environment.

Overview

In a *concurrent access* environment, two or more nodes can have simultaneous (concurrent) access to a volume group that resides on shared external disks.

Note: You cannot define filesystems on a concurrent access volume group.

HACMP for AIX supports concurrent access volume groups and logical volumes defined on any of the following devices:

- IBM 9333 Serial Disk Subsystem
- IBM 7133 and 7131-405 Serial Storage Architecture (SSA) Disk Subsystems
- IBM 7135-110 and IBM 7135-210 RAIDiant Disk Arrays
- IBM 7137 Disk Array
- IBM 2105-B09 and 2105-100 Versatile Storage Servers

The following sections first describe how the HACMP for AIX scripts handle concurrent access LVM components. Then the chapter includes sections that describe how to perform specific administrative tasks for LVM components in a concurrent access environment.

Most maintenance tasks can be performed using the HACMP C-SPOC utility.

Creating a concurrent volume group or adding nodes to an existing volume group can also be done using the TaskGuide utility. For information on using the TaskGuide for creating a shared volume group, see the *HACMP for AIX Concepts and Facilities* guide.

Concurrent Access and HACMP for AIX Scripts

You should seldom, if ever, need to intervene in a concurrent access cluster. In a concurrent access environment, as in a non-concurrent environment, the HACMP for AIX event scripts control the actions taken by a node and coordinate the interactions between the nodes. However, as a system administrator, you should monitor the status of the concurrent access volume groups when HACMP for AIX events occur.

To intervene in a cluster, you must understand how nodes in a concurrent access environment control their interaction with shared LVM components. For example, the HACMP **node_up_local** script may fail before varying on a volume group in concurrent mode. After fixing whatever problem caused the script to fail, you may need to manually vary on the volume group in concurrent access mode. The following sections describe the processing performed by these scripts.

Nodes Join the Cluster

A node joining a cluster calls the **node_up_local** script which calls the **cl_mode3** script to activate the concurrent capable volume group in concurrent access mode.

The **cl_mode3** script calls the **varyonvg** command with the **-c** flag. For more information about this command and its flags, see “Activating a Concurrent Capable Volume Group in Concurrent Access Mode.” If the concurrent capable volume group is defined on a RAID disk array device, the scripts use the **convaryonvg** command to vary on the concurrent volume groups in concurrent mode.

Nodes Leave the Cluster

Nodes leaving the cluster do not affect the concurrent access environment. They simply vary off the volume groups normally. The remaining nodes take no action to change the concurrent mode of the shared volume groups.

When a node leaves the cluster gracefully, it executes the **node_down_local** script which calls the **cl_deactivate_vgs** script. This script uses the **varyoffvg** command to vary off the concurrent volume groups.

Maintaining Concurrent Access Volume Groups

The LVM enables you to create concurrent access volume groups that can be varied on in either concurrent access mode or non-concurrent access mode. Maintaining these volume groups may require you to perform any of the following tasks:

- Creating a concurrent capable volume group in concurrent access mode
- Varying on a concurrent capable volume group in concurrent access mode
- Determining if a volume group is a concurrent capable volume group
- Determining if a volume group is varied on in concurrent access mode
- Replacing a failed drive in a concurrent access volume group
- Maintaining a concurrent access environment while upgrading HACMP for AIX software

- Restarting the LVM concurrent access daemon (**clvmd**).

The following sections describe how to perform these tasks.

The tasks you can perform with C-SPOC are described in the last part of the chapter.

You can also create or add nodes to a concurrent volume group using the TaskGuide utility. For more information on the TaskGuide for creating shared volume groups, see the Administrative Facilities chapter in the *HACMP for AIX Concepts and Facilities* guide.

Creating a Concurrent Volume Group

You can use AIX (commands and SMIT) to create a concurrent volume group. Both methods are explained in this section.

You are likely to plan to create concurrent volume groups on SSA disk subsystems as part of the procedure for replacing a failed drive in an SSA concurrent access volume group. In order to enable this procedure, you must assign unique non-zero node numbers on each node of the cluster.

If you specify the use of SSA disk fencing in your concurrent resource group, HACMP assigns the node numbers when you synchronize the resources.

If you don't specify the use of SSA disk fencing in your concurrent resource group, assign the node numbers with

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

where *x* is the number to assign to that node. Then reboot the system.

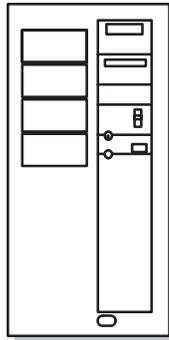
The following sections describe each step of creating a concurrent volume group in detail.

Creating a Concurrent Volume Group Using AIX

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

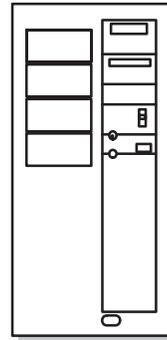
Maintaining Shared LVM Components in a Concurrent Access Environment

Maintaining Concurrent Access Volume Groups



Source Node

1. Complete prerequisite tasks
2. Create concurrent capable volume group
3. Vary on the volume group in non-concurrent mode
4. Create logical volume
5. Vary off volume group
8. Complete follow-up tasks



Destination Nodes

1. Complete prerequisite tasks
6. Import volume group
7. Change volume group to remain dormant at startup
8. Complete follow-up tasks

Creating a Concurrent Access Volume Group

Step 1. Complete Prerequisite Tasks

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

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

The procedure you use to create a concurrent access volume group varies depending on the type of device you are using.

Creating a Concurrent Access Volume Group on Serial Disk Subsystems

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

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

```
mkvg -n -s 4 -c -V 35 -y myvg hdisk1 hdisk2
```

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

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

```
smit mkvg
```

SMIT displays the Add a Volume Group screen. This screen has additional fields in a concurrent access environment.

2. Enter the specific field values as follows:

VOLUME GROUP name	Specify name of volume group.
Physical partition SIZE in megabytes	Accept default.
PHYSICAL VOLUME NAMES	Specify the names of the physical volumes you want included in the volume group.
Activate volume group AUTOMATICALLY at system restart?	Set to no so that the volume group can be activated as appropriate by the cluster event scripts.
ACTIVATE volume group after it is created?	Set this field to no .
Volume Group MAJOR NUMBER	You can accept the default; however, using the lvlstmajor command is recommended to find a free major number.
Create VG concurrent capable?	Set to yes so that the volume group can be activated in concurrent access mode by the HACMP for AIX event scripts.
Auto-varyon concurrent mode?	Set to no so that the volume group can be activated as appropriate by the cluster event scripts.

3. Press Enter. The system asks if you are sure. Press Enter again.
4. After the command completes, press F10 to exit SMIT.

Creating a Concurrent Access Volume Group on RAID Disk Subsystems

On IBM 7135 RAIDiant Disk Arrays and IBM 2105 Versatile Storage Servers, use the **mkvg** command to create concurrent access volume groups. Do not use the **-c** flag, since RAID disks don't use the CLVM.

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

VOLUME GROUP name	The name of the shared volume group should be unique within the cluster.
Activate volume group AUTOMATICALLY at system restart?	Set to no so that the volume group can be activated as appropriate by the cluster event scripts.
ACTIVATE volume group after it is created?	Set to yes .

**Volume Group MAJOR
NUMBER**

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

Step 3. Vary On the Concurrent Access Volume Group in Non-concurrent Mode

Use the `varyonvg` command to activate the volume group in non-concurrent mode. To create logical volumes, the volume group must be varied on in non-concurrent mode. For example, to vary on the concurrent access volume group `myvg` in non-concurrent mode, enter the following command:

```
varyonvg myvg
```

You can also use SMIT to vary on the volume groups, as follows:

1. To vary on a concurrent capable volume group in non-concurrent mode, enter:

```
smit varyonvg
```

SMIT displays the Activate a Volume Group screen; it has additional fields in a concurrent access environment.

2. Enter field values as follows.

VOLUME GROUP name Specify name of volume group.

**RESYNCHRONIZE stale
physical partitions?** Set 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.

3. Press Enter. The system asks if you are sure. Press Enter again.
4. After the command completes, press F10 to exit SMIT.

Step 4. Create Logical Volumes on Concurrent Access 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. For more information about creating logical volumes, see the *AIX System Management Guide*.

1. To create a logical volume, enter:

```
smit mklv
```

SMIT displays the Add a Logical Volume menu.

2. In general, you can supply values for these fields following the guidelines in the *HACMP for AIX Installation Guide*. However, specify the following values for these selected fields for logical volumes created on concurrent access volume groups:

Number of COPIES of each logical partition Specify 1, 2, or 3 mirror copies. **Note:** If you defined the volume group on an IBM 7135-110 or IBM 7135-210 disk array, do not use LVM mirroring.

Mirror Write Consistency Specify the value **no**.

Enable BAD BLOCK relocation? Specify the value **no**.

3. Press Enter. The system asks if you are sure. Press Enter again.
4. After the command completes, press F10 to exit SMIT.

Step 5. Vary Off Volume Group on Source Node

After creating the logical volume, vary off the volume group using the **varyoffvg** command so that it can be varied on by the HACMP for AIX scripts. Enter:

```
varyoffvg volume_group_name
```

Step 6. Import Volume Group Information on Destination Nodes

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

```
importvg -V 35 -y vg_name physical_volume_name
```

Specify the name of any of the disks in the volume group as an argument to the **importvg** command.

Note: You must vary on concurrent capable volume groups defined on serial disk subsystems—they are not varied on automatically upon import by AIX. Concurrent access volume groups defined on RAID disk subsystems are varied on automatically when imported.

With SMIT use the following procedure.

1. To import a concurrent access volume group, enter:

```
smit importvg
```

SMIT displays the Import a Volume Group screen; this field has additional fields in a concurrent access environment.

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 one of the physical volumes that resides in the volume group. <i>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.</i> 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	You can accept the default; however, using the lvlstmajor command to find a free major number is recommended.
Make this VG concurrent capable	Accept the default value (no). Only used when importing a volume group created on an earlier version of AIX. For more information, see <i>Maintaining Concurrent Access During a Cluster Upgrade</i> on page 5-15
Make default varyon of VG concurrent	Accept the default value (no). Only used when importing a volume group created on an earlier version of AIX. For more information, see <i>Maintaining Concurrent Access During a Cluster Upgrade</i> on page 5-15.

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

If your cluster uses SCSI external disks (including RAID devices), and the import of the volume group fails, make sure there is no reserve on any of the disks in the volume group, by executing the following command:

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

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

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

Step 7. Change Volume Group to Remain Dormant at Startup on Destination Nodes

In the HACMP for AIX system, a volume group should be varied on as appropriate by the HACMP for AIX scripts. Therefore, after importing a volume group, you must change the volume group so that it remains dormant at startup.

To change this characteristic of a volume group enter the following:

```
chvg -a n volume_group_name
```

You can also use SMIT by using the following procedure:

1. Use the **smit chvg** fastpath to change the characteristics of a volume group.
2. Enter the specific field values as follows. For other fields use the defaults or the appropriate entries for your operation:
Set the **Activate volume group Automatically at system restart?** field to **no**.
3. Press Enter to commit this change.
4. Press F10 to exit SMIT and return to the command line.

Step 8. Complete Follow-up Tasks

Verify that the HACMP for AIX scripts can activate the concurrent access volume group as a concurrent cluster resource.

Activating a Volume Group in Concurrent Access Mode

As a system administrator, you may, at times, need to activate (vary on) a concurrent capable volume group in concurrent access mode. For example, if an HACMP for AIX script fails before activating a volume group in concurrent access mode, you may want to activate the volume group manually after correcting the failure.

The method used to vary on a volume group in concurrent access mode varies, depending on the type of device. The following sections describe the procedure for IBM 9333 Disk Subsystems, IBM 7133 or 7131-405 SSA Disk Subsystems, IBM 7135 and IBM 2105 Versatile Storage Servers RAID devices.

For concurrent access volume groups, bad block relocation is automatically turned off.

Activating Volume Groups Defined on Serial Disk Subsystems

If the volume group is defined on an IBM 9333 Disk Subsystem or an IBM 7133 or 7131-405 SSA Disk Subsystem, you can activate the volume group using **varyonvg** command, specifying the **-c** flag, as in the following example:

```
varyonvg -c myvg
```

Note: To vary on a volume group in concurrent access mode, the HACMP for AIX software must be present on the node.

You can also use SMIT to activate a volume group in concurrent access mode. Use the following procedure.

1. Enter:

```
smit varyonvg
```

SMIT displays the **Activate a Volume Group** screen; this screen has an additional field in a concurrent access environment.

2. Enter the field values as follows.

VOLUME GROUP name Specify name of volume group.

RESYNCHRONIZE stale physical partitions? Set this field to **no**.

Activate volume group in SYSTEM MANAGEMENT mode? Accept the default (**no**).

FORCE activation of the Volume Group? Accept the default (**no**).

Varyon VG in concurrent mode? Set to **yes**.

3. Press Enter. The system asks if you are sure. Press Enter again.
4. After the command completes, press F10 to exit SMIT.

Activating Volume Groups Defined on RAID Devices

To activate a volume group defined on a 7135, 7137 or 2105 Versatile Storage Server RAID device in concurrent access mode, you must use the `/sbin/cluster/events/utlis/convaryonvg` command. This command first checks to make sure the disks in the volume group are not reserved.

For example, to vary on the `sharedvg1` volume group in concurrent mode, enter:

```
/usr/sbin/cluster/events/utlis/convaryonvg sharedvg1
```

Determining the Access Mode of a Volume Group

To determine if a volume group is a concurrent capable volume group and to determine its current mode, use the `lsvg` command specifying the name of the volume group as an argument. The `lsvg` command displays the following information about the volume group:

```
# lsvg myvg
VOLUME GROUP:   myvg                VG IDENTIFIER:   000025426733bc50
VG STATE:       active                PP SIZE:         16 megabyte(s)
VG PERMISSION:  read/write                TOTAL PPs:       119 (1904 megabytes)
MAX LVs:        256                  FREE PPs:        119 (1904 megabytes)
LVs:            10                   USED PPs:        0 (0 megabytes)
OPEN LVs:       8                    QUORUM:          2
TOTAL PVs:      1                    VG DESCRIPTORS:  2
STALE PVs:      0                    STALE PPs        0
ACTIVE PVs:     1                    AUTO ON:         no
Concurrent:     Capable                Auto-Concurrent: disabled
VG Mode:        Concurrent
```

To determine if the volume group is concurrent capable, check the value of the **Concurrent** field. The volume group in the example was created as a concurrent capable volume group, as indicated by the value of this field. If this volume group was not a concurrent capable volume group, the value of this field would be *Non-Capable*.

To determine if the volume group is activated in concurrent access mode, check the value of the **VG Mode** field. In the example, the volume group is activated in concurrent access mode. If this volume group had not been varied on in concurrent access mode, the value of this field would be *Non-Concurrent*.

The **Auto-Concurrent** field indicates whether the volume group should be varied on in concurrent access mode when the volume group is started automatically at system reboot. The value of this field is determined by the value of the `-x` option to the `mkvg` command when the volume group was created. In an HACMP for AIX environment, this option should always be disabled; HACMP for AIX scripts control when the volume should be varied on.

Replacing a Failed Drive in a Concurrent Access Volume Group

The following five volume group reconfiguration commands allow you to replace a failed disk drive belonging to a concurrent access volume group:

extendvg	Add a physical volume to an existing volume group.
reducevg	Remove a physical volume from an existing volume group.
mklvcopy	Add a mirror copy to an existing logical volume.
rmlvcopy	Remove a mirror copy from an existing logical volume.
syncvg	Brings all copies of physical partitions in a logical volume up to date.

The concurrent capable volume group must be varied on in concurrent access mode *on all nodes* for the changes made by these reconfiguration commands to be made known to all cluster nodes automatically. If a node has the volume group varied off, it must export the volume group and then re-import the volume group before varying it on again.

Note: Make sure that the concurrent access volume group is in a quiescent state (no I/O operations in progress) before executing these reconfiguration commands.

Two procedures follow. The first is for IBM 9333 serial disk subsystems and IBM RAIDiant Disk Arrays; the second is for IBM SSA disk subsystems.

Replacing a Failed Drive in a Concurrent Access Volume Group on an IBM 9333 Serial Disk Subsystem or an IBM RAIDiant Disk Array

The following example illustrates the procedure for replacing a failed drive in a concurrent access volume group on an IBM 9333 serial disk subsystem, IBM 7135 RAIDiant Disk Array, 7137 Disk Array, or IBM 2105 Versatile Storage Server. In the example, three cluster nodes (named A, B, and C) share access to a concurrent access volume group, named *sharedvg*, on which three copies of the logical volume *sharelv* are maintained.

The example assumes that the system error log reports problems with *hdisk4*. The data in the volume group remains accessible because the number of disks in the volume group that are still available constitute a quorum. However, if the disk is not repaired or replaced, high availability is endangered.

To replace a failed drive in a concurrent access volume group:

1. Determine the contents of the failed disk drive by executing the **lsvg** command:

```
lsvg -M -n sharedvg
  hdisk2:1      sharelv:1:1
  hdisk2:2      sharelv:2:1
  hdisk2:3      sharelv:3:1
  hdisk2:4-50
  hdisk3:1      sharelv:1:2
  hdisk3:2      sharelv:2:2
  hdisk3:3      sharelv:3:2
  hdisk3:4-50
  hdisk4:1      sharelv:1:3   stale
  hdisk4:2      sharelv:2:3   stale
  hdisk4:3      sharelv:3:3   stale
```

```
hdisk4:4-50
```

The output indicates that the logical volume mirror copies on *hdisk4* are stale. This is the drive that has failed and must be replaced.

2. Reduce the number of copies of the physical partitions to only those that are currently available, using the **rmlvcopy** command:

```
rmlvcopy sharelv 2 hdisk4
```

This command reduces the number of copies of the logical volume *sharelv* to two and specifies that the copy to be removed is on the failed drive, *hdisk4*.

3. Reduce the size of the volume group to omit the failed drive, using the **reducevg** command:

```
reducevg sharedvg hdisk4
```

This command specifies that the physical volume *hdisk4* should be removed from the volume group *sharedvg*.

4. Execute the **rmdev** command on *each* cluster node to remove the definition of the failed drive from the device configuration database:

```
rmdev -l hdisk4 -d
```

The disk may not be named *hdisk4* on all the nodes. Use the **lspv** command on each node to obtain a list of disk PVIDs to accurately identify the failed disk.

5. Remove the failed drive and replace it with a new drive. Refer to the documentation that accompanied the disk subsystem for information about removing and replacing an individual drive. *Power down only the failed drive. Do not power down the entire disk subsystem.*

Run the **diag** command to format the new drive, if necessary.

6. Configure the new drive on *each* cluster node. Do *not* use the **cfgmgr** command to reconfigure the disk drive; instead, use the following commands. The example configures the adapters and disks on an IBM 9333 disk subsystem:

```
lsdev -Cadapter -tserdasda | awk '{print $1}' | xargs -n1 >  
-i{} /etc/methods/cfgserdasda -l {}  
lsdev -Cadapter -tsserdasda | awk '{print $1}' | xargs -n1 >  
-i{} /etc/methods/cfgsserdasda -l {}  
lsdev -Cdisk -tsserdasdc | awk '{print $1}' | xargs -n1 >  
-i{} mkdev -l {}
```

7. Determine the logical name of the new disk, using the **lsdev** command:

```
lsdev -Cc disk
```

The system displays a list of all currently configured disks. To find out what name the system assigned to the replacement disk, you must know the location code of the drive. This code can be obtained by entering:

```
lsdev -Cdisk -tsserdasdc
```

The following steps assume that the system assigned the same logical name to the new drive (*hdisk4*).

8. Add the new disk to the concurrent access volume group, using the **extendvg** command:

```
extendvg sharedvg hdisk4
```

9. Recreate the mirrors of the logical volume on the new drive, use the **mklvcopy** command:

```
mklvcopy sharelv 3 hdisk4
```

Do not specify the **-k** option which synchronizes the logical volume mirrors as you create the copies.

10. Synchronize the new mirror copies of the logical volume with the current mirrors of the logical volume, using the **syncvg** command:

```
syncvg -v sharevg
```

This completes the procedure for replacing a failed drive in a concurrent access volume group on an IBM 9333 serial disk subsystem or an IBM RAIDiant Disk Array.

Replacing a Failed Drive in a Concurrent Access Volume Group on an IBM SSA Disk Subsystem

The following example illustrates the procedure for replacing a failed SSA drive in a concurrent access volume group. In the example, two cluster nodes (named A and B) share access to a concurrent access volume group named *sharevg*, on which two copies of the logical volume *lvmarket* are maintained. The example assumes that an SSA disk configured as *hdisk6* has failed.

Warning: For this procedure to work, you must first have assigned unique non-zero *node_numbers* through the *ssar* on each cluster node. If you have specified (or are planning to specify) SSA disk fencing in your concurrent resource group, these *node_numbers* are assigned when you synchronize the resources. If, however, you do not specify SSA disk fencing, issue the following command to assign *node_numbers*: `chdev -l ssar -a node_number=x`, where *x* is the number to assign to the node. Reboot the system to make the *node_number* change effective. Changing *node_numbers* is best done after installing HACMP for AIX and after attaching and configuring your SSA devices.

Complete the following steps to replace a failed SSA drive in a concurrent access volume group. You must be the root user.

Perform Steps 1 Through 5 on Node A.

1. Run the **lsvg** command to determine which logical volumes have partitions on the failed disk:

```
# lsvg -M sharevg
  hdisk6:126      lvmarket:72:2    stale
  hdisk6:127      lvmarket:73:2    stale
  hdisk6:128      lvmarket:74:2    stale
  hdisk6:129      lvmarket:75:2    stale
  hdisk6:130      lvmarket:76:2    stale
  hdisk6:131      lvmarket:77:2    stale
  hdisk6:132      lvmarket:78:2    stale
  hdisk6:133      lvmarket:79:2    stale
  hdisk6:134      lvmarket:80:2    stale
```

The output indicates that the logical volume mirror copies on *hdisk6* are stale. This is the drive that has failed and must be replaced.

2. For each logical volume with partitions on the failed disk, use the **rmlvcopy** command to remove the mirror copies from that disk:

```
# rmlvcopy lvmarket 1 hdisk6
```

This command reduces the number of copies of the logical volume *lvmarket* to one and specifies that the copy to be removed is on the failed drive, *hdisk6*.

Note: The **rmlvcopy** command also causes a **syncldvdm** to be run on the other node, which resets the permissions on the raw device to the default values. A permissions change can affect applications using the device since it can cause them to lose access to the device.

3. Run the **reducevg** command to remove the failed disk from the volume group:

```
# reducevg sharedvg hdisk6
```

This command specifies that the physical volume *hdisk6* should be removed from the volume group *sharedvg*.

4. Determine which **pdisk** corresponds to the **hdisk**, if the disk has failed completely, by issuing the following command:

```
lsdev -Ccpdisk -Fconnwhere,name|grep `lsdev -ClhdiskX HR>-Fconnwhere`
```

This command gives the **pdisk** name and the “connwhere” value of the failing **hdisk** and **pdisk**. The serial number printed on the front of the failed SSA disk drive is represented by the fifth through the twelfth characters of the connwhere value.

If the disk has not failed completely, run the **diag** command as follows:

- a. Choose the Service Aids option.
- b. Choose the SSA Service Aids option.
- c. Choose the Configuration Verification option.
- d. Choose the appropriate **hdisk** and press ENTER. The corresponding **pdisk** is displayed with the serial number on the disk drive. Remember the **pdisk** number, for example, **pdisk1**.
- e. Exit the diagnostic service aids.

Use the **pdisk** number in the following step:

5. Remove the failed disk.

If the disk has failed completely, its check light remains on, or all indicators on the device are off. The adjacent devices show flashing ready lights. You can now remove the drive without first setting Service Mode.

If, however, the disk has not failed completely, run the **diag** command as follows:

- a. Choose the Service Aids option.
- b. Choose the SSA Service Aids option.
- c. Choose the Configuration Verification option.
- d. Choose “Set or Reset Service Mode”
The drive’s yellow check light turns on.
- e. Remove this drive from the SSA enclosure.
- f. Exit the diagnostic service aids.

Perform Steps 6 Through 9 on Both Nodes.

6. Run the **rmdev** command to remove the **hdisk** and **pdisk** devices on both nodes. Note that, while the **pdisk** device number is always the same on both nodes, the **hdisk** number for the same disk may differ from node to node, depending on the disk configuration of each node.

You can verify that you have the correct disk using the following commands:

```
# lsdev -Cl hdiskX -Fconnwhere  
# lsdev -Cl pdiskX -Fconnwhere
```

Then remove the hdisk and pdisk devices on both nodes as follows:

```
# rmdev -d -l hdisk6  
# rmdev -d -l pdisk1
```

The resulting character string should be identical for the pdisk and hdisk on node A and node B.

7. Replace the failed disk with a new one.
8. Run the following commands on each node to configure the new device. Do not run the **cfgmgr** command while the cluster is running.

```
# /usr/lib/methods/cfgssar -l ssar >/dev/null  
# lsdev -Cssar -S0 iFname|xargs -n1
```

9. Run the **lspv** command to check the hdisk name of the newly configured disk. It should be named identical to the removed hdisk.

Perform Steps 10 Through 12 on Node A.

10. Run the **extendvg** command to add the disk back into the volume group:

```
# extendvg sharedvg hdisk6
```

11. Run the **mklvcopy** command to recreate the mirror copies onto the new disk:

```
# mklvcopy lvmarket 2 hdisk6
```

12. Run the **syncvg** command to synchronize the new mirror copies with the existing mirror copies for each affected logical volume:

```
# syncvg -l lvmarket
```

This completes the procedure for replacing a failed SSA drive in a concurrent access volume group.

Maintaining Concurrent Access During a Cluster Upgrade

When upgrading a cluster, you typically bring down the cluster nodes one at a time, install the new HACMP for AIX software, and then reintegrate the node into the cluster. As part of this process, you import the shared volume group definitions onto the node that is reintegrating.

When upgrading a cluster to HACMP for AIX, Version 4.2.2, in which a concurrent access volume group is configured, you must specify the **-c** flag with the **importvg** command when importing the shared concurrent access volume groups, as in the following:

```
importvg -y volume_group_name -c physical_disk_name
```

Concurrent access volume groups created with earlier versions of AIX do not identify the volume groups as concurrent capable in the Volume Group Descriptor Area (VGDA). If the volume group is not identified as concurrent capable, the new version of AIX activates the imported volume group automatically in non-concurrent access mode. When one node varies on a concurrent access volume group in non-concurrent mode, the access mode reverts to non-concurrent access for all nodes accessing the volume group, thus locking out the other cluster nodes. By specifying the **-c** flag with the **importvg** command, AIX updates the on-disk copy of the volume group's VGDA to indicate it is concurrent capable.

Note: Once you change a volume group to be concurrent capable, you cannot change it back to a standard volume group.

Restarting the Concurrent Access Daemon (clvmd)

As a system administrator, you may, at times, need to restart the concurrent access daemon (**clvmd**). The **clvmd** daemon normally gets started by the **varyonvg** command when you vary on a volume group in concurrent access mode by specifying the **-c** flag. You can restart the **clvmd** by re-executing the **varyonvg -c** command on the already-varied on concurrent access volume group. You cannot vary on an already-varied on volume group in a different mode, however.

You can also restart the **clvmd** daemon using the following SRC command:

```
startsrc -s clvmd
```

By using the SRC facility command to start the daemon, you ensure that only a single copy of the daemon is running on a node at any one time.

Maintaining Concurrent LVM Components Using C-SPOC

HACMP 4.3.1 C-SPOC uses the AIX 4.3.1 CLVM capabilities that allow changes to concurrent LVM components without stopping and restarting the cluster. See Chapter 4, *Maintaining Shared LVM Components*, for a general explanation of how C-SPOC works.

You can use the C-SPOC utility to do the following tasks:

- Create a concurrent-capable volume group.
- Concurrent volume groups tasks (volume group must be varied on in concurrent mode)
 - List all concurrent volume groups in the cluster
 - Import a concurrent volume group
 - Extend a concurrent volume group
 - Reduce a concurrent volume group
 - Mirror a concurrent volume group
 - Unmirror a concurrent volume group
 - Synchronize concurrent LVM mirrors by volume group.
- Concurrent logical volumes tasks
 - List all concurrent logical volumes by volume group
 - Add a concurrent logical volume to a volume group
 - Make a copy of a concurrent logical volume
 - Remove a copy of a concurrent logical volume
 - Show the characteristics of a concurrent logical volume
 - Remove a concurrent logical volume
 - Synchronize concurrent LVM mirrors by logical volume.

Using C-SPOC to Maintain Concurrent Volume Groups

You can use C-SPOC to do all the concurrent volume maintenance tasks.

Creating a Concurrent Volume Group on Cluster Nodes Using C-SPOC

Using C-SPOC simplifies the procedure for creating a concurrent volume group on selected cluster nodes. Before creating a concurrent volume group for the cluster using C-SPOC, check that:

- All disk devices are properly attached to the cluster nodes
- All disk devices are properly configured on all cluster nodes and listed as available on all nodes
- The cluster concurrent logical volume manager is installed
- All disks that will be part of the volume group must be concurrent capable.

Take the following steps to create a shared volume group for a selected list of cluster nodes:

1. Enter the fastpath **smitty cl_admin**
2. Select **Cluster Concurrent Logical Volume Manager > Concurrent Volume Groups > Add a Concurrent Volume Group**.

SMIT displays a list of cluster nodes.

3. Select two or more nodes from the list of cluster nodes and press Enter.

The system correlates a list of all free concurrent-capable physical disks that are available to all nodes selected. (Free disks are those disks that currently are not part of a volume group and have PVIDs.) SMIT displays the list of free physical disks in a multi-pick list by PVIDs

4. Select one or more disks from the list and press Enter.

SMIT displays the **cl_mkvg** screen with a major number inserted into the Major Number data field. The system determines this free major number; do not change it.

5. Complete the selections as follows

Nodes Names of the selected nodes are displayed.

VOLUME GROUP name Enter a name for this concurrent capable volume group.

Physical partition SIZE in megabytes Accept the default.

*** PHYSICAL VOLUME names** The system displays the selected disks.

**Volume Group MAJOR
NUMBER**

The system displays the number C-SPOC has determined to be correct.

Warning: Changing the volume group major number may result in the command's inability to execute on a node that does not have that major number currently available. Please check for a commonly available major number on all nodes before changing this setting.

C-SPOC verifies communication paths and version compatibility and then executes the command on all the nodes you selected.

Note: If the major number entered on the SMIT panel was not free at the time that the system attempted to make the volume group the command will display an error for the node that did not complete the execution and continue to the other nodes. At the completion of the command the volume group will not be active on any node in cluster.

Listing all Concurrent Volume Groups in the Cluster

To list all concurrent volume groups in the cluster, take the following steps:

1. On the source node, enter `smit hacmp`.
2. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > List all Concurrent Volume Groups** and press Enter.

SMIT displays a message asking if you want to view only active concurrent volume groups.

3. Choose **yes** to see a list of active concurrent volume groups only, or choose **no** to see a list of all concurrent volume groups.

Extending a Concurrent Volume Group with C-SPOC

Take the following steps to add a physical volume to a concurrent volume group using C-SPOC.

1. The physical volumes (**hdisks**) being added to the volume group must be installed, configured, and available.
2. On any cluster node that can own the volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > Concurrent Volume Groups > Set Characteristics of a Concurrent Volume Group > Add a Physical Volume to a Concurrent Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. You can pick one or more to add to the volume group. Select the ones you want to add to the volume group and press Enter.

7. SMIT displays the **Add a Physical Volume to a Concurrent Volume Group** screen, with the following entries filled in.

Resource Group name	The cluster resource group to which this concurrent volume group belongs.
Volume Group name	Name of the volume group where hdisks are to be added.
Reference node	Name of the node where the hdisks are found.
Physical Volume names	Names of the hdisks to be added to the volume group.

8. If this screen reflects the correct information, press Enter to add the disks to the concurrent volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Importing a Concurrent Volume Group with C-SPOC

Take the following steps to import a concurrent volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > Concurrent Volume Groups > Import a Concurrent Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Import a Concurrent Volume Group screen. Values for fields you have selected are displayed. For other fields, use the defaults or the appropriate entries for your operation:

VOLUME GROUP name	The name of the volume group that you are importing.
--------------------------	--

PHYSICAL VOLUME name	The name of one of the physical volumes that resides in the volume group. This is the hdisk name on the reference node.
-----------------------------	---

Reference node	Node from which the physical disk was retrieved.
-----------------------	--

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

Make this VG concurrent capable? The default is **no**. For a non-concurrent VG, leave this **no**.

Make default varyon of VG concurrent? The default is **no**. For a non-concurrent VG, leave this **no**.

8. If this screen reflects the correct information, press Enter to import the concurrent volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Removing a Physical Volume from a Concurrent Volume Group with C-SPOC

Take the following steps to remove a physical volume from a concurrent volume group using the C-SPOC utility.

1. Complete prerequisite tasks. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > Concurrent Volume Groups > Set Characteristics > Remove a Physical Volume from a Concurrent Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Remove a Physical Volume from a Concurrent Volume Group screen, with the following entries filled in.:

VOLUME GROUP name The name of the volume group that you are reducing.

Reference node Node from which the name of the physical disk was retrieved.

PHYSICAL VOLUME name The name of the physical volume that you want to remove. This is the hdisk name on the reference node.

8. If this screen reflects the correct information, press Enter to reduce the concurrent volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Mirroring a Concurrent Volume Group Using C-SPOC

Take the following steps to mirror a concurrent volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > Concurrent Volume Groups > Mirror a Concurrent Volume Group**.

SMIT displays a list of volume groups.

5. Select the desired volume group and press Enter.

SMIT displays a list of physical volumes.

6. Select the physical volume you want and Press Enter.

SMIT displays the Mirror a Concurrent Volume Group screen, with the following entries filled in.

For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name The name of the resource group to which this concurrent volume group belongs is displayed.

VOLUME GROUP name The name of the volume group that you want to mirror is displayed.

Reference node Node from which the name of the physical disk was retrieved is displayed.

PHYSICAL VOLUME names The name of a physical volume on the volume group that you want to mirror. This is the hdisk name on the reference node.

Number of COPIES of each logical partition The default is **2**. You can also select **3**.

Keep Quorum Checking On? The default is **no**. You can also choose **yes**.

Create Exact LV Mapping? The default is **no**.

7. If this screen reflects the correct information, press Enter to mirror the concurrent volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
8. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Unmirroring a Concurrent Volume Group Using C-SPOC

Take the following steps to unmirror a concurrent volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > Concurrent Volume Groups > Unmirror a Concurrent Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Unmirror a Concurrent Volume Group screen, with the chosen fields filled in.

For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name The name of the resource group to which this concurrent volume group belongs is displayed.

VOLUME GROUP name The name of the volume group that you want to mirror is displayed.

Reference node Node from which the name of the physical disk was retrieved is displayed.

PHYSICAL VOLUME names The name of a physical volume on the volume group that you want to unmirror. This is the hdisk name on the reference node.

Number of COPIES of each logical partition The default is **2**. You can also select **3**.

8. If this screen reflects the correct information, press Enter to unmirror the concurrent volume group. All nodes in the cluster receive this updated information.
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Synchronizing Concurrent Volume Group Mirrors

Take the following steps to synchronize concurrent LVM Mirrors by volume group using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter **smit hacmp**.
4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Synchronize Concurrent LVM Mirrors > Synchronize By Volume Group**.
5. SMIT displays a list of volume groups. Select the desired volume group and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Synchronize Concurrent LVM Mirrors by Volume Group screen, with the chosen entries filled in.

For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name	The name of the resource group to which this concurrent volume group belongs is displayed.
VOLUME GROUP name	The name of the volume group that you want to mirror is displayed.
Reference node	Node from which the name of the physical disk was retrieved is displayed.
Number of Partitions to Sync in Parallel	Set range from 1 to 32.
Synchronize All Partitions	The default is no .
Delay Writes to VG from other cluster nodes during this Sync	The default is no .

8. If this screen reflects the correct information, press Enter to synchronize LVM mirrors by the concurrent volume group. All nodes in the cluster receive this updated information immediately (before lazy update).
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Using C-SPOC to Maintain Concurrent Logical Volumes

You can use the C-SPOC utility to do many maintenance tasks on concurrent logical volumes. You cannot change the size of the logical volume. After you complete the procedure using SMIT, the other cluster nodes are updated with the new information.

Listing all Concurrent Logical Volumes in the Cluster

To list all concurrent logical volumes in the cluster, take the following steps:

1. On the source node, enter `smit hacmp`.
2. From the main HACMP menu, choose **Cluster System Management > Cluster Concurrent Logical Volume manager > List all Concurrent Logical Volumes by Volume Groups** and press Enter.
3. SMIT displays a message asking if you want to view only active concurrent volume groups. Choose **yes** to see a list of active concurrent volume groups only, or choose **no** to see a list of all concurrent volume groups.

Adding a Concurrent Logical Volume to a Cluster Using C-SPOC

To add a concurrent logical volume to a cluster using C-SPOC, take the following steps:

1. Enter the SMIT C-SPOC fastpath:
`smit cl_admin`
2. Select these options **Cluster Concurrent Logical Volume Manager > Concurrent Logical Volumes > Add a Concurrent Logical Volume** and press Enter.
3. SMIT displays a popup list of concurrent volume groups. Select one and press Enter.
4. SMIT displays a list of physical volumes. Select one and press Enter. The **Add a Concurrent Logical Volume** screen appears, with chosen fields filled in as shown in the sample below.:

Resource Group name	ccur_rg
VOLUME GROUP name	concurrentvg
Reference node	a1
* Number of LOGICAL PARTITIONS	[]
PHYSICAL VOLUME names	hdisk16
Logical volume NAME]
Logical volume TYPE	[]
POSITION on physical volume	middle
RANGE of physical volumes	minimum
MAXIMUM NUMBER of PHYSICAL VOLUMES to use for allocation	[]

Number of COPIES of each logical partition	1
Mirror Write Consistency?	yes
Allocate each logical partition copy on a SEPARATE physical volume?	yes
RELOCATE the logical volume during reorganization	yes
Logical volume LABEL	[]
MAXIMUM NUMBER of LOGICAL PARTITIONS	[512]
Enable BAD BLOCK relocation?	yes
SCHEDULING POLICY for writing logical partition copies	parallel
Enable WRITE VERIFY?	no
File containing ALLOCATION MAP	[]
Stripe Size?	[Not Striped]

5. The default LV characteristics are most common. Make changes if necessary for your system and press Enter. Other cluster nodes are updated with this information.

Removing a Concurrent Logical Volume on a C-SPOC Cluster

Using the following procedure, you can remove a concurrent logical volume on any node in a cluster by executing a C-SPOC command on one of the nodes.

Note: If the logical volume to be removed contains a filesystem, you first must remove the filesystem from any specified resource group before attempting to remove the logical volume. Afterwards, be sure to synchronize cluster resources on all cluster nodes.

1. Enter the SMIT fastpath:

```
smit cl_admin
```
2. Select the **Cluster Logical Volume Manager** option and press Enter.
3. Choose the **Concurrent Logical Volumes** option and press Enter.
4. Select the **Remove a Concurrent Logical Volume** option and press Enter.
5. C-SPOC provides a list of concurrent logical volumes, organized by HACMP resource group. Select the logical volume you want to remove and press Enter.

To check the status of the C-SPOC command execution on other cluster nodes, view the C-SPOC log file in `/tmp/cspoc.log`.

Setting Characteristics of a Concurrent Logical Volume Using C-SPOC

You can use C-SPOC to do the following tasks:

- Add copies to a concurrent logical volume
- Remove copies from a concurrent logical volume

Adding a Copy to a Concurrent Logical Volume Using C-SPOC

Using the following procedure, you can add a copy to a concurrent logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. On the source node, vary on the volume group, using the SMIT **varyonvg** fastpath. You can rename a logical volume using C-SPOC when the volume group is varied off; however, you must specify the **-f** flag. See the **cl_chlv** command man page for more information.
2. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```
3. Select these options: **Cluster Logical Volume Manager > Concurrent Logical Volumes > Set Characteristics of A Concurrent Logical Volume > Add a Copy to a Concurrent Logical Volume.**
SMIT displays a list of logical volumes arranged by resource group.
4. Choose the desired logical volume from the pick list and press Enter. SMIT displays a list of physical volumes.
5. Choose a physical volume and press Enter. SMIT displays the **Add a Copy to a Concurrent Logical Volume** screen with the Resource Group, Logical Volume, Reference Node and default fields filled.
6. Enter the new number of mirrors in the **NEW TOTAL number of logical partitions** field and press Enter. The C-SPOC utility changes the number of copies of this logical volume on all cluster nodes.

Note: The current version of the AIX **mklvcopy** command ignores the “synchronize copies” option for concurrent logical volumes, Therefore, to synchronize the new copies you must run the **Synchronizing LVM Mirrors by Logical Volume** command after the **mklvcopy** command completes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Removing a Copy from a Concurrent Logical Volume Using C-SPOC

Using the following procedure, you can remove a copy of a concurrent logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

1. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```
2. Select these options: **Cluster Logical Volume Manager > Concurrent Logical Volumes > Set Characteristics of A Concurrent Logical Volume > Remove a Copy from a Concurrent Logical Volume.**

SMIT displays a list of logical volumes arranged by resource group.

3. Choose the desired logical volume from the pick list and press Enter. SMIT displays a list of physical volumes.
4. Choose the physical volumes from which you want to remove copies and press Enter.

SMIT displays the **Remove a Copy from a Concurrent Logical Volume** screen with the Resource Group, Logical Volume name, Reference Node and Physical Volume names fields filled in.

5. Enter the new number of mirrors in the **NEW maximum number of logical partitions copies** field and check the **PHYSICAL VOLUME name(s) to remove copies from** field to make sure it is correct and press Enter.

The C-SPOC utility changes the number of copies of this logical volume on all cluster nodes.

To check the status of the C-SPOC command execution on all nodes, view the C-SPOC log file in **/tmp/cspoc.log**.

Show Characteristics of a Concurrent Logical Volume Using C-SPOC

To show the current characteristics of a concurrent logical volume, take the following steps.

1. On the source node, vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath.
2. Enter the C-SPOC SMIT fastpath:

```
smit cl_admin
```

3. Select these options: **Cluster Logical Volume Manager > Concurrent Logical Volumes > Show Characteristics of A Concurrent Logical Volume**.

SMIT displays a list of logical volumes arranged by resource group.

4. Choose the desired logical volume from the pick list and press Enter.

SMIT displays the Show Characteristics of A Concurrent Logical Volume screen with the **Resource Group**, and **Logical Volume name** fields filled in. The **Option list** field offers three choices for the display. Toggle between **Status**, **Physical Volume map**, and **Logical Partition map**.

Synchronizing LVM Mirrors by Logical Volume

Take the following steps to synchronize concurrent LVM Mirrors by logical volume using the C-SPOC utility.

1. The physical volumes (**hdisks**) in the volume group must be installed, configured, and available.
2. On any cluster node that can own the concurrent volume group (is in the participating nodes list for the resource group), vary on the volume group in concurrent mode, using the SMIT **varyonvg** fastpath (if it is not varied on already).
3. On the source node, enter `smit hacmp`

4. From the main HACMP menu, choose **Cluster System Management > Cluster Logical Volume manager > Synchronize Concurrent LVM Mirrors > Synchronize By Logical Volume**.
5. SMIT displays a list of logical volumes. Select the desired logical volume and press Enter.
6. SMIT displays a list of physical volumes. Pick one and Press Enter.
7. SMIT displays the Synchronize LVM Mirrors by Volume Group screen, with the chosen entries filled in. For other fields, use the defaults or the appropriate entries for your operation:

Resource Group Name The name of the resource group to which this logical volume belongs is displayed.

LOGICAL VOLUME name The name of the logical volume that you want to synchronize is displayed.

Reference node Node from which the name of the physical disk was retrieved is displayed.

Number of Partitions to Sync in Parallel Set the range from 1 to 32.

Synchronize All Partitions The default is **no**.

Delay Writes to VG from other cluster nodes during this Sync The default is **no**.

8. If this screen reflects the correct information, press Enter to synchronize LVM mirrors by the concurrent logical volume. All nodes in the cluster receive this updated information.
9. If you did this task from a cluster node that does not need the concurrent volume group varied on, vary off the volume group on that node.

Chapter 6 Changing the Cluster Topology

The cluster topology refers to the physical layout of the cluster, that is, the number of nodes, their adapters and the networks that connect them. This chapter explains how to view the current cluster topology, how to update the topology definition when changes are made to the cluster, and how to synchronize the cluster topology definition across all cluster nodes.

Note: You can also use the **xhacmpm** application to view and reconfigure a cluster. For more information, see the Administrative Facilities chapter in the *HACMP for AIX Concepts and Facilities Guide*.

Overview

When you configure a cluster, you enter data that describes the cluster topology into the AIX configuration database, called the Object Data Manager (ODM). The ODM is an object-oriented database made up of records, called object classes, that contain key configuration information. When you configure an HACMP for AIX cluster, configuration data is stored in HACMP-specific object classes. (To view these object classes, use the **odmget** command, specifying as an argument one of the HACMP object classes, such as **HACMPcluster**.) The AIX ODM object classes are stored in the default system configuration directory (DCD), **/etc/objrepos**.

Synchronizing Configuration Changes

When you change the topology of a cluster, you update the data stored in the ODM in the DCD. For example, when you add an additional network adapter to a cluster node, you must add the adapter to the cluster definition so that the cluster nodes can recognize and use the adapter. When you change the cluster definition on one cluster node, you must also update the ODMs on the other cluster nodes, a process called synchronization. Synchronization causes the information stored in the DCD on the local cluster node to be copied to the ODM object classes in the DCD on the other cluster nodes.

Note that some topology changes, such as changing node names and adapter names, are propagated through the cluster topology only. You still need to make manual updates in SMIT for the resources and resource groups associated with the changed node or adapter.

Reconfiguring a Cluster Dynamically

While a cluster is running, the HACMP daemons, scripts, and utilities reference the ODM data stored in the active configuration directory (ACD), not the ODM data stored in the DCD. At cluster start-up, HACMP for AIX copies HACMP-specific ODM classes into this directory. If you synchronize the cluster topology definition while the Cluster Manager is running on the local node, it triggers a dynamic reconfiguration event. In a dynamic reconfiguration event, the ODM data in the DCDs on all cluster nodes is updated and the ODM data in the ACD is overwritten with the new configuration data. The HACMP daemons are refreshed so that the new configuration becomes the currently active configuration.

Cluster topology changes that can be made to an active cluster dynamically:

- Adding or removing one or more nodes
- Adding or removing one or more network adapters
- Swapping an adapter card
- Adding, removing, or modifying one or more network modules.

Note: Before changing the cluster topology, you should remove any dependencies on the topology in the cluster resource definitions. For example, a resource group may include in its list of participating nodes, the name of the node you intend to remove. You must reconfigure cluster resources in a separate dynamic reconfiguration operation. The following sections note when corresponding changes to cluster resources are required. See Chapter 7, *Changing Resources and Resource Groups*, for more information about changing cluster resources.

By default, whenever you synchronize the cluster topology, HACMP for AIX verifies the configuration, calling the **clverify** utility. When synchronizing the cluster triggers a dynamic reconfiguration event, HACMP for AIX verifies that both cluster topology and cluster resources are correctly configured, even though you may have only changed an element of one of these. Since a change in topology may invalidate the resource configuration, and vice versa, the software checks both. For more information, see Chapter 8, *Verifying a Cluster Configuration*.

You can skip cluster verification during cluster synchronization by choosing the appropriate value in the **Skip Cluster Verification** field in SMIT or by running the appropriate command at the command line. Cluster verification is optional only when a cluster is *inactive*.

For more information on the option to skip cluster verification, see *Synchronizing the Cluster Topology* on page 6-18.

Requirements Before Reconfiguring

Before making changes to a cluster topology, ensure that:

- HACMP for AIX is installed on all nodes
- All nodes are up and able to communicate with each other
- All nodes include each other's boot and service interface IP labels or addresses in their **.rhosts** file as well as their own.

Viewing the Cluster Topology

Note: When you view the cluster topology, you are viewing the ODM data stored in the DCD, not the data stored in the ACD.

To view the cluster topology, perform the following procedure.

1. Enter the SMIT fastpath:

```
smit hacmp
```

2. Select **Cluster Configuration** from the main menu and press Enter.
3. Select **Cluster Topology** and press Enter.
4. Select **Show Cluster Topology** and press Enter.

A list of options appears, as described in the next section.

Options in SMIT Under Show Cluster Topology

Each option provides a different view of the cluster. The following sections provide more information about each option.

Show Cluster Topology	Select the Show Cluster Topology option when you want to obtain complete information about the cluster topology including the nodes in the cluster, their adapters, and the networks that connect them.
Show Cluster Definitions	Select the Show Cluster Definitions option when you want to obtain the name and ID of all clusters accessible from this node.
Show Topology Information by Node	<p>Select the Show Topology Information by Node option to obtain information about cluster nodes and their adapters. If you choose the Select a Node to Show option, SMIT displays a pick list of nodes.</p> <p>After you select a node and press Enter, SMIT displays the attributes of the node you specified. Or, you can choose the Show All Nodes option.</p>
Show Topology Information by Network Name	<p>Select the Show Topology Information by Network Name option to obtain information about the networks that connect the cluster nodes. If you choose the Select a Network to Show option, SMIT displays a pick list of networks.</p> <p>After you select a network and press Enter, SMIT displays the attributes of the network you specified. Or, you can choose the Show All Networks option.</p>
Show Topology Information by Network Adapter	<p>Select the Show Topology Information by Network Adapter option to obtain information about the network adapters defined in the cluster. If you choose the Select an Adapter to Show option, SMIT displays a pick list of adapters.</p> <p>After you select an adapter and press Enter, SMIT displays the attributes of the adapter you specified. Or, you can choose the Show All Adapters option.</p>

Changing a Cluster Name or ID

Note: You cannot change the name or ID of an active cluster dynamically.

To change a cluster's name or ID:

1. Enter the fastpath `smit hacmp`.
2. Select **Cluster Configuration > Cluster Topology > Configure Cluster > Change/Show Cluster Definition**.

SMIT displays the fields that define the cluster definition with their current values filled in.

3. Enter any changes to the values displayed and press Enter.

A cluster name can include alphabetic and numeric characters and underscores. Use no more than 31 characters.

4. After the command completes, press the F3 key to return to the HACMP for AIX SMIT menus to perform further topology reconfiguration or to synchronize the changes you made. To synchronize the cluster topology, return to the Cluster Topology menu and select the Synchronize Cluster Topology option. See Synchronizing the Cluster Topology on page 6-18 for more information.

Changing the Configuration of Cluster Nodes

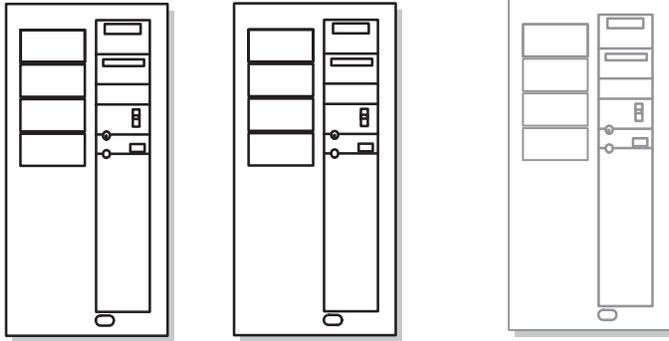
As the system administrator of an HACMP for AIX cluster, you may need to perform any of the following tasks relating to cluster nodes:

- Adding one or more nodes to the cluster
- Removing a node from the cluster
- Changing the attributes of a cluster node.

Adding a Cluster Node

You can add a node to an active cluster dynamically. You do not need to stop and restart cluster services for the node to become part of the cluster.

The following figure illustrates the steps necessary to add a third node to a two-node cluster dynamically.



On any cluster node:

2. Add node to cluster topology
3. Add an adapter for the new node
4. Synchronize cluster topology
5. Add node to resource group
6. Synchronize cluster resources

On node to be added:

1. Make sure that cluster services are inactive
7. Start cluster services on added node

Adding a Third Node to a Two-Node Cluster Dynamically

The following describes each step in more detail.

1. On the node to be added:
 - Ensure that the node to be added is powered on.
 - Ensure that the node to be added has network adapters configured to provide connectivity to the cluster. See the section Changing the Configuration of Network Adapters on page 6-9 for more information.
 - Ensure that cluster services are inactive. For information about how to view the state of cluster services, see Chapter 3, Monitoring an HACMP Cluster.
2. On any other node (called the local node from here on), add the new node to the cluster topology definition.
Enter the fastpath `smit hacmp`
3. From the main HACMP for AIX SMIT screen, select **Cluster Configuration > Cluster Topology > Configure Nodes > Add Cluster Nodes**.
4. Enter the name of the node (or nodes) that you want to add to the cluster.

Note: A node name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. Separate multiple names with spaces. If you specify a duplicate node name, the operation fails.

Press Enter to add the node or nodes to the cluster definition.

5. To add a node to a cluster dynamically, you must also add a service network adapter for the new node to the cluster definition. To be an active cluster member, each cluster node must have at least one network adapter defined as the service adapter. A node can also have multiple standby adapters, depending on the specific configuration. For information about adding network adapters, see the section Adding a Network Adapter on page 6-9.
6. On the local node, if you have completed the topology changes you want to make, synchronize the cluster topology definition. To do so, return to the **Cluster Topology** menu and select the **Synchronize Cluster Topology** option.

If the cluster manager is running on the local node, the synchronization triggers a dynamic reconfiguration event. See Synchronizing the Cluster Topology on page 6-18 for more information.

Note: In a single dynamic reconfiguration, you can change either the cluster topology definition or the definition of cluster resources, but not both.

7. If desired, on the local node, add the newly added node (or nodes) to the list of participating nodes in a resource group. For information about adding a node to a resource group, see Chapter 7, Changing Resources and Resource Groups.

Warning: When adding a node to a cluster that has one or more resource groups with *SSA disk fencing enabled*, add the node to the concurrent resource group immediately. All nodes in an SSA concurrent access cluster must participate in the concurrent access resource group. Include the new node in this resource group immediately to avoid the possibility of unrecoverable data loss.

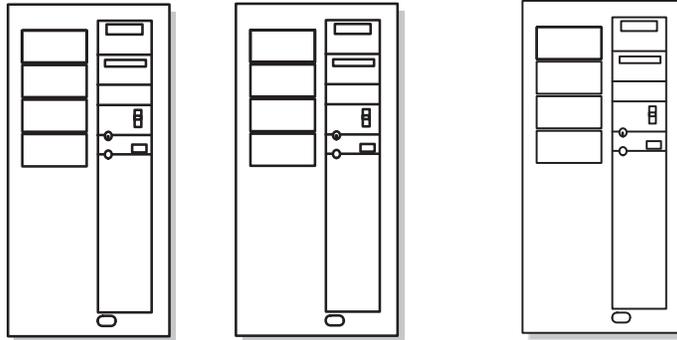
In a cascading resource group, if you give the new node the highest priority, by specifying it first in the list of participating nodes, the newly added node will acquire control of the resource group when you synchronize cluster resources. This can be useful when you want the new node to take over a specific resource. For example, you may be adding a high-powered node to a cluster that runs a heavily used database application and you want this application to run on the newly added node.

8. On the local node, synchronize cluster resources. Return to the **SMIT Cluster Resources** menu and select the **Synchronize Cluster Resources** option. See the section entitled Synchronizing Cluster Resources on page 7-27 for more information.
9. On the newly added node, start cluster services to integrate the new node into the cluster.

Removing a Cluster Node

You can remove a node from an active cluster dynamically.

Before removing a node from the cluster, remove the node from any resource groups it participates in. The following figure illustrates the steps necessary to remove a node from the cluster.



On any cluster node:

2. Remove the node from resource groups.
3. Synchronize cluster resources.
4. Remove the node from the topology.
5. Synchronize the cluster topology.

On node to be removed:

1. Stop cluster services.

The following procedure provides more detail about each step:

1. On the node to be removed, stop cluster services.

Note: For nodes in cascading or rotating resource groups, you do not need to stop cluster services on the node to be removed before removing it from the list of participating nodes. You can stop cluster services after removing the node. However, in clusters with concurrent access resource groups with SSA disk fencing enabled, it is recommended.

2. On any other cluster node (hereafter called the local node), remove the node from the resource groups in which it participates. For information about removing a node from a resource group, see Chapter 7, *Changing Resources and Resource Groups*.

Note: When removing a node from the list of nodes participating in a concurrent access resource group that has *SSA disk fencing enabled*, remove the node from the cluster immediately after removing the node from the concurrent access resource group. All nodes in an SSA concurrent access cluster must participate in the concurrent access resource group. Remove the node from the cluster to avoid the possibility of unrecoverable data loss.

3. On the local node, if you have completed all the topology changes you want to make, synchronize cluster resources. Return to the **SMIT Cluster Resources** menu and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, the synchronization triggers a dynamic reconfiguration event. In a dynamic reconfiguration, the definition of cluster resources is updated on all cluster nodes and the new (changed) configuration becomes the current active configuration. See the section entitled *Synchronizing Cluster Resources* on page 7-27 for more information.

4. On the local node, remove the node from the cluster topology definition.

From the main HACMP for AIX SMIT screen, select the **Cluster Configuration > Cluster Topology > Configure Nodes > Remove a Cluster Node** option and press Enter.

SMIT displays a list of all cluster nodes from which you must select the node to be removed.

5. Select the node you want to remove from the cluster and press Enter. Before executing the command, SMIT displays a popup message asking if you are sure you want to proceed. Press Enter again to remove the node from the cluster topology definition.

Note: When you remove a node, all adapter information associated with the node is also removed.

6. On the local node, synchronize the cluster topology.

Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option.

If the Cluster Manager is running on the local node, the synchronization triggers a dynamic reconfiguration event. See *Synchronizing the Cluster Topology* on page 6-18 for more information. When performing a dynamic reconfiguration, the Command Status window displays “OK” before the operation is actually complete.

When the synchronization completes, the node is removed from the cluster topology definition.

Changing the Name of a Cluster Node

You cannot change the name of a cluster node dynamically. You must stop cluster services, make the change, and then restart cluster services to apply it to the active configuration.

To change the name of a cluster node, perform the following procedure.

1. From the Cluster Topology menu, select the **Configure Nodes** option and press Enter.
2. Select the **Change/Show Cluster Node Name** option and press Enter.

SMIT displays a pick list of cluster nodes.

3. Select the desired node.

SMIT displays the current node name.

4. Enter the new name you want to assign to the node in the **New Node Name** field. A node name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. When you finish entering data, press the Enter key. SMIT makes the changes you specified.
5. After the command completes, press the F3 key to return to the HACMP for AIX SMIT menus to perform further topology reconfiguration or to synchronize the changes you made.
6. To synchronize the cluster topology, return to the Cluster Topology menu and select the **Synchronize Cluster Topology** option. For more information, see the section *Synchronizing the Cluster Topology* on page 6-18 in this chapter.

Manual Updates Required After Node Name Change

Be aware that after changing a node name, the change is propagated through the cluster topology only. Cluster *resources* are not modified, and so will still contain references to the old name. The Participating Node Names field must be updated manually (in the Change/Show a Resource Group SMIT screen) for all resource groups associated with the changed node.

Changing the Configuration of Network Adapters

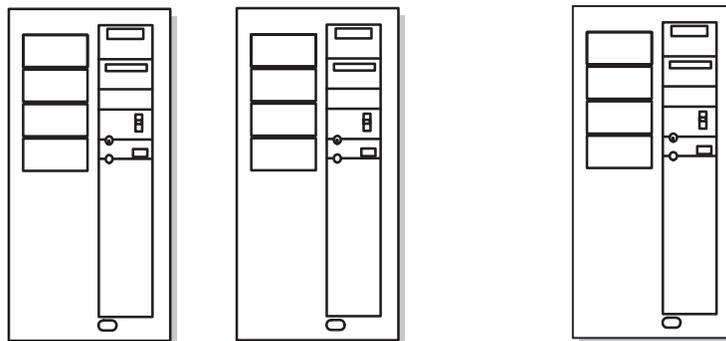
As a system administrator, you may need to perform any of the following tasks relating to cluster network adapters:

- Adding a network adapter
- Removing a network adapter
- Changing a network adapter

Adding a Network Adapter

You can add a network adapter to an active cluster dynamically. You do not need to stop and restart cluster services for the node to become part of the cluster.

The following figure illustrates the steps necessary to add a network adapter to a cluster node.



On any cluster node:

2. Update `/etc/hosts` on all nodes
3. Add the adapter to the cluster topology.
4. Synchronize the cluster topology.

On node getting new adapter:

1. Complete prerequisite tasks.

Adding a Network Adapter

The following procedure provides more detail about each step.

1. On the node getting the new adapter, complete prerequisite tasks. These tasks include:
 - Installing the new adapter.
 - Configuring the new adapter. For information about configuring an adapter, see the *HACMP for AIX Installation Guide*.
2. On all cluster nodes, update the `/etc/hosts` file to include the IP address of the new adapter.
3. On any cluster node (hereafter called the local node), add the network adapter to the cluster topology definition.

From the main HACMP for AIX SMIT screen, select **Cluster Configuration > Cluster Topology > Configure Adapters** option, and press Enter.

SMIT displays the Configure Adapters menu.

4. Select **Add an Adapter** and enter field values as follows:

Adapter IP Label	<p>Type the IP label (the name) of the adapter as defined in the <i>/etc/hosts</i> file.</p> <p>If the cluster uses IP address takeover or rotating resources, each adapter that can have its IP address taken over must have a boot adapter (address) label defined for it. Use a consistent naming convention for boot adapter labels. The label for an RS232 line must end in the characters <i>ttyn</i>, where <i>n</i> is the number of the tty device associated with the serial network (for example, <i>clam_tty1</i>).</p> <p>The label for a target mode SCSI-2 bus must end in the characters <i>tmcsin</i>, where <i>n</i> is the SCSI device number (for example, <i>clam_tmcsi2</i>).</p> <p>The label for a target mode SSA loop must end in the characters <i>tmssan</i> where <i>n</i> is the SSA device number.</p> <p>The label for a SLIP line should end in <i>SLn</i>, where <i>n</i> is the number of the SLIP device associated with the SLIP line (for example <i>clam_SLI</i>).</p>
Network Type	<p>Indicate the type of network, for instance, <i>ethernet</i> or <i>token ring</i>, to which this adapter belongs. Press F4 for a popup pick list of pre-installed network types. The network type for an SP switch must be <i>hps</i>.</p>
Network Name	<p>Enter an ASCII text string that identifies the network connected to this adapter. The network name can include alphabetic and numeric characters and underscores. Use no more than 31 characters. The network name is arbitrary, but must be used consistently. If several adapters share the same logical network or subnet, make sure that you use the same network name for each of these adapters.</p>
Network Attribute	<p>Indicate whether the network is public, private, or serial. Press TAB to toggle the values. Ethernet, Token-Ring, FDDI, and SLIP can be public networks. SOCC, ATM, and SP Switch networks can only be private networks. RS232 lines, target mode SCSI-2 busses, and target mode SSA loops are serial networks.</p>
Adapter Function	<p>Indicate whether the adapter's function is service, standby, or boot. Press the TAB key to toggle the values. A node has a single service adapter for each network. A node can have none, one, or more standby adapters for each public network; however, one or more standby adapters is highly recommended to avoid a single point of failure in your cluster environment. Serial networks and SP switch networks do not have standby adapters.</p>

Adapter Identifier Enter the IP address in dotted decimal format or as a device file name.

IP address information is required for non-serial network adapters only if the node's address cannot be obtained from the local `/etc/hosts` file (using the adapter IP label given).

You must enter device file names for serial network adapters. RS232 serial adapters should have the device file name `/dev/tty n` .

Target mode SCSI serial adapters should have the device file name `/dev/tm s cs n` , where n should match the adapter label number.

Target mode SSA serial adapters should have the device file name `/dev/tm s sa n` , where n should match the adapter label number.

Adapter Hardware Address

This field is optional. Enter a hardware address for the adapter. The hardware address must be unique within the physical network. Enter a value in this field only if:

- You are currently defining a service adapter, *and*
- The service adapter has a boot address, *and*
- You want to use hardware address swapping.

The hardware address is 12 digits for Ethernet, Token-Ring and FDDI; and 14 digits for ATM.

Node Name

This field is optional. Only adapters which have addresses which are *always* associated with a particular node need a node name defined. These are service addresses which are always associated with only a single node, boot addresses, and standby addresses.

For adapters which have addresses which may be rotated or cascaded, no node name should be defined. Instead, the event scripts use the user-defined resource group configuration to associate these service addresses with the proper node.

5. Press Enter. The system adds these values to the HACMP for AIX ODM.
6. On the local node, synchronize the cluster topology.

Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option.

If the Cluster Manager is running on the local node, the synchronization triggers a dynamic reconfiguration event. See Synchronizing the Cluster Topology on page 6-18 below for more information.

When the synchronization completes, the network adapter is added to the cluster topology definition.

Manual Updates Required After Adapter Name Change

Be aware that after changing an adapter name, the change is propagated through the cluster topology only. Cluster *resources* are not modified, and so will still contain references to the old name. The Service IP Label field (in the Change/Show Resources/Attributes SMIT screen) must be updated manually with the new name for any resource group associated with the changed adapter.

Removing a Network Adapter from a Cluster Node

You can remove a network adapter from an active cluster dynamically; you do not need to stop and restart cluster services.

To remove a network adapter from a cluster node, perform the following procedure:

1. From the main HACMP for AIX SMIT screen, select **Cluster Configuration > Cluster Topology > Configure Adapters > Remove an Adapter**.

SMIT displays the list of defined adapter names.

2. Select the adapter you want to remove and press Enter.

When you remove an adapter, all information associated with the adapter is removed from the ODM. A pop-up message asks if you are sure you want to do this operation. Press Enter again *only* if you are sure you want to remove the adapter and its associated information.

3. On the same node, synchronize the cluster topology.

Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option. If the Cluster Manager is running on the local node, the synchronization triggers a dynamic reconfiguration event. See Synchronizing the Cluster Topology on page 6-18 for more information.

When the synchronization completes, the network adapter is removed from the cluster topology definition.

Changing Network Adapter Attributes

You cannot change the attributes of a network adapter dynamically. You must stop and restart cluster services to make the changed configuration the active configuration.

To change the attributes of a network adapter:

1. Stop cluster services on all nodes.
2. On any cluster node, select the **Configure Adapters** option from the Cluster Topology menu and press Enter. SMIT displays the **Configure Adapters** menu.
3. Choose the **Change/Show an Adapter** option and press Enter. SMIT displays a list of adapter names.
4. Select the adapter to change and press Enter. SMIT displays a screen listing all the attributes of the adapter, with their current values filled in.
5. Make any desired changes and press Enter.
6. On the local node, synchronize the cluster topology.

Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option. If the Cluster Manager is running on the local node, the synchronization fails. You cannot reconfigure the attributes of a network adapter dynamically. See *Synchronizing the Cluster Topology* on page 6-18 for more information.

When the synchronization completes, the network adapter is removed from the cluster topology definition.

Swapping a Network Adapter Dynamically

Note: All nodes must be at HACMP for AIX Version 4.3.1 or higher to use this feature.

As a systems administrator, you may at some point experience a problem with a network adapter card on one of the HACMP for AIX cluster nodes. If this occurs, the dynamic adapter swap feature can be used to swap the IP address of an active or boot adapter with the IP address of an active, available standby adapter on the same node and network. You can select the available standby to which the address will be moved. Cluster services do not have to be stopped to perform the swap.

This feature can be used to move an IP address off of an adapter that is behaving erratically, to another standby adapter, without shutting down the node. It can also be used if a hot pluggable adapter device is being replaced on the node. Hot pluggable adapters can be physically removed and replaced without powering off the node.

If hardware address swapping is enabled, the hardware address will be swapped along with the IP address.

Note: This type of dynamic adapter swap can only be performed within a single node. You cannot swap the IP address of a service or boot adapter with the IP address of a standby adapter on a different node with this feature. To move an IP address to another node, move its resource group using the DARE Resource Migration utility.

To dynamically swap a network adapter, perform the following procedure:

1. Make sure that no other HACMP for AIX events are running before swapping an adapter.
2. From the main HACMP for AIX SMIT screen, select **Cluster Systems Management > Swap Network Adapter**.
SMIT displays a list of available service/boot adapters.
3. Select the service/boot adapter you want to remove from cluster use, and press Enter.
SMIT displays a list of available standby adapters.
4. Select the standby adapter you want, and press Enter.
The Swap Network Adapter menu appears.
5. Verify the service/boot IP label, and the standby IP label you have chosen. If these are correct, press Enter.

A pop-up message asks if you are sure you want to do this operation. Press Enter again *only* if you are sure you want to swap the network adapter.

After the adapter swap, the service/boot address becomes an available standby adapter. At this point, you can take action to repair the faulty adapter. If you have a hot pluggable adapter, you can replace the adapter while the node and cluster services are up. Otherwise, you will have to stop cluster services and power down the node to replace the adapter.

If you have a hot pluggable adapter, HACMP for AIX will make the adapter unavailable as a standby when you pull it from the node. When the new adapter card is placed in the node, the adapter is incorporated into the cluster as an available standby again. You can then use the dynamic adapter swap feature again to swap the IP address from the standby back to the original adapter.

If you need to power down the node to replace the faulty adapter, HACMP for AIX will configure the service/boot and standby addresses on their original adapters when cluster services are restarted. You do not need to use the dynamic adapter swap feature again to swap back the adapters. HACMP for AIX does not record the swapped adapter information in the AIX ODM. Therefore, the changes are not persistent across system reboots or cluster restarts.

Changing the Configuration of a Network Module

The HACMP for AIX SMIT interface allows you to add, remove or change an HACMP for AIX network module. You rarely need to add or remove an HACMP for AIX network module. However, you may want to change its failure detection rate.

If HACMP for AIX cannot get enough CPU resources to send heartbeats on IP and serial networks, other nodes in the cluster will assume the node has failed, and initiate takeover of the node's resources. In order to ensure a clean takeover, the deadman switch crashes the busy node if it is not reset within a given time period. The deadman switch uses the following formula:

```
N=((keepalives * missed_keepalives) -1)
Where keepalives and missed_keepalives are for the slowest network in
the cluster.
```

If you have trouble with node failure due to network traffic, you can change the **Failure Detection Rate** for the network modules in your cluster. This field controls the time interval the Cluster Manager uses between keepalives. You can choose between three set values: **Fast**, **Normal**, **Slow**; you can also select **Custom** to adjust the *heartbeat rate* and *failure cycle* ODM fields manually. The normal heartbeat rate is usually optimal. Speeding up or slowing down failure detection is a small, but potentially significant area where you can adjust cluster fallover behavior. However, the amount and type of customization you add to event processing has a much greater impact on the total fallover time. You should test the system for some time before deciding to change the failure detection speed of any network module.

If you decide to change the failure detection rate of a network module, keep the following considerations in mind:

- Tune the AIX I/O pacing and **syncd** parameters for each node before making any change to network module settings
- Failure detection is dependent on the *fastest* network linking two nodes.

- Faster heartbeat rates may lead to false failure detections, particularly on busy networks. For example, bursts of high network traffic may delay heartbeats and this may result in nodes being falsely ejected from the cluster. Faster heartbeat rates also place a greater load on networks.
- If your networks are very busy and you experience false failure detections, you can try changing the failure detection speed on the network modules to **slow** to avoid this problem.

Note: In rare cases, it is necessary to slow the Failure Detection Rate to even longer than the **Slow** option SMIT offers. In this case, use the **Custom** option and then adjust the two fields as needed.

The failure rate of networks varies, depending on their characteristics. For example, for an Ethernet, the normal failure detection rate is two heartbeats per second; fast is about four per second; slow is about one per second. For an SP switch network, because no network traffic is allowed when a node joins the cluster, normal failure detection is 30 seconds; fast is 10 seconds; slow is 60 seconds.

The table below shows the Deadman Switch Timeout for each network. Each of these times is one second less than the time to trigger an HACMP event on that network. Remember that the Deadman Switch is triggered on the slowest network in your cluster.

NETWORK	SLOW	NORMAL	FAST
ATM	63	31	15
Ethernet	11	5	3.5
FDDI	11	5	3,5
Token-Ring	11	5	3.5
RS232	17	11	5
SLIP	17	11	5
SOCC	11	5	3.5
SP Switch	63	15	7
TMSSA	17	11	5
TMSCSI	17	11	5

Deadman Switch Timeouts in Seconds Per Network

Changing the Attributes of a Network Module

To change the attributes of a network module:

1. Stop cluster services on all cluster nodes.
2. Enter `smitty hacmp`
3. Select **Configure Network Modules** from the Cluster Topology menu and press Enter.

4. Select the **Change/Show Network Modules** option and press Enter. SMIT displays a list of defined network modules.
5. Select the network module you want to change and press Enter.
SMIT displays the attributes of the network module, with their current values.

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

New Network Module Name []

Description For example, Ethernet Protocol

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

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

Parameters

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

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

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

6. Make your changes and press Enter. SMIT executes the command to modify the values of these attributes in the ODM.
7. On the local node, synchronize the cluster topology. Return to the SMIT **Cluster Topology** menu and select the **Synchronize Cluster Topology** option.
The configuration data stored in the DCD on each cluster node is updated and the changed configuration becomes the active configuration when cluster services are started. See Synchronizing the Cluster Topology on page 6-18 for more information.

Adding a Network Module

Note: You cannot add a network module to an active cluster dynamically.

If you want to add a new type of network to the cluster (not originally supported by the HACMP for AIX software), you must add a network module for the new type of network. After installing this new software on each node in the cluster, use the following procedure to define the new network module to HACMP cluster.

1. Select **Configure Network Modules** from the Cluster Topology menu and press Enter.
2. Select **Add a Network Module** and press Enter.
3. Enter field values as follows:

Network Module Name	Network module names cannot exceed 31 characters. They can include alphabetic and numeric characters and underscores.
Description	Enter a descriptive comment about the network module. Usually this describes the type of network the module monitors.
Address type	Specify either IP address (address , the default) or device file (dev). Press TAB to toggle the field.
Path	Enter the path where the executable for the network module resides.
Parameters	<i>This field is optional.</i> Any additional parameters.
Failure Detection Rate	Switch for adjusting time between keepalives. Press TAB to toggle between Fast , Normal , Slow , Custom . The failure cycle and the heartbeat interval determine how soon a failure can be detected. The time needed to detect a failure can be calculated using this formula: (heartbeat interval) * (failure cycle).
Failure Cycle	The current setting is the default for the network module selected. (Default for ether is 2.) This is the number of successive heartbeats that can be missed before the interface is considered to have failed. If you select Custom in the Failure Detection Rate field, you can enter a number from 1 to 21474.
Heartbeat Rate	The current setting is the default for the network module selected. This parameter tunes the interval between heartbeats (in tenths of a second) for the selected network module. If you select Custom in the Failure Detection Rate field, you can enter a number from 1 to 21474.

4. After the command completes, you must update the other cluster nodes. For the HACMP cluster to function properly, the HACMP for AIX ODM entries must be the same on all cluster nodes. You must synchronize all cluster nodes *whenever* you make a change to the cluster's definition.

To synchronize the cluster topology, press the F3 key until you return to the **Cluster Topology** menu and then select the **Synchronize Cluster Topology** option. See Synchronizing the Cluster Topology on page 6-18 for more information.

Removing a Network Module

Note: You cannot remove a network module from an active cluster dynamically.

It is highly unlikely that you will ever need to remove a network module. However, if you must, use the following procedure.

1. Select **Configure a Network Module** from the Cluster Topology menu and press Enter.
2. Select **Remove a Network Module** and press Enter.

SMIT displays a list of defined network modules.

3. Select the name of the network module you want to remove and press Enter.

A popup message appears, warning you that all the information associated with the network module will be lost. Press Enter again, if you are completely certain you understand the consequences of this operation.

4. After the command completes, you must update the other cluster nodes. For the HACMP cluster to function properly, the HACMP ODM entries must be the same on all cluster nodes. You must synchronize all cluster nodes *whenever* you make a change to the cluster's definition.
5. To synchronize the cluster topology, press the F3 key until you return to the **Cluster Topology** menu and then select the **Synchronize Cluster Topology** option. See Synchronizing the Cluster Topology on page 6-18 for more information.

Synchronizing the Cluster Topology

Whenever you modify the cluster topology definition in the ODM on one node, you must synchronize the change with the ODM data on all cluster nodes. You perform a synchronization by choosing the **Synchronize Cluster Topology** option from the **Cluster Topology** SMIT screen.

Note: Before synchronizing a cluster configuration, ensure that all nodes are powered on, that the HACMP for AIX software is installed, that if the cluster security mode is **Standard** the **/etc/hosts** and **/.rhosts** files on all nodes include all HACMP for AIX boot and service IP labels, and that if the cluster security mode is **Enhanced**, the **/.klogin** files on the control workstation and all nodes include all service principals configured for Kerberos. See the *HACMP Installation Guide* for a discussion of configuring for Kerberos.

The processing performed in synchronization varies depending on whether the Cluster Manager is active on the local node:

- If the Cluster Manager is not active on the local node, the ODM data in the DCD on the local node is copied to the ODMs stored in the DCDs on all cluster nodes.
- If the Cluster Manager is active on the local node, a dynamic reconfiguration event is triggered. In dynamic reconfiguration, the DCDs on all cluster nodes are synchronized and the configuration data stored in the ACD on each cluster node is replaced by the new configuration data. Finally, the dynamic reconfiguration causes the daemons to refresh, making the new configuration the active configuration. In the HACMP for AIX log file, a **reconfig_topology_start** event marks the beginning of a dynamic reconfiguration and a **reconfig_topology_complete** event marks the finish of the dynamic reconfiguration.
- If the Cluster Manager is active on some cluster nodes but not on the local node, the synchronization operation is aborted.

To synchronize changes to the cluster topology across all cluster nodes:

1. From the **Cluster Configuration** menu, select **Cluster Topology > Synchronize Cluster Topology** and press Enter.

The following fields appear:

Ignore Cluster Verification Errors	By choosing yes , the result of the cluster verification is ignored and the configuration is synchronized even if verification fails. By choosing no , the synchronization process terminates; view the error messages in the system error log to determine the configuration problem.
---	---

Emulate or Actual	If you set this field to Emulate , the synchronization is an emulation and does not affect the Cluster Manager. If you set this field to Actual , the synchronization actually occurs, and any subsequent changes affect the Cluster Manager. Actual is the default value.
--------------------------	---

Skip Cluster Verification	By default, this field is set to no and the cluster topology verification program is run. To save time in the cluster synchronization process, you can toggle this entry field to yes .
----------------------------------	---

2. Specify the desired option in the **Ignore Cluster Verification Errors** field. Set the **Emulate or Actual** field to **Actual** to run cluster synchronization. Set the **Skip Cluster Verification** field to the desired option.

Press enter

The cluster definition (including all node, adapter, and network method information) is copied to the other cluster nodes.

3. Press F10 to exit SMIT.

Releasing a Dynamic Reconfiguration Lock

As described in Chapter 3 in the *HACMP for AIX Concepts and Facilities Guide*, during a dynamic reconfiguration, HACMP for AIX creates a temporary copy of the HACMP-specific ODM classes and stores them in the Staging Configuration Directory (SCD). This allows you to modify the cluster configuration while a dynamic reconfiguration is in progress. You cannot, however, synchronize the new reconfiguration until the first is finished. The presence of an SCD on any cluster node prevents dynamic reconfiguration. If, because of a node failure or

other reason, an SCD remains on a node after a dynamic reconfiguration is finished, it will prevent any further dynamic reconfiguration. Before you can perform further reconfiguration, you must remove this lock.

To remove a dynamic reconfiguration lock, perform the following procedure.

1. Enter the SMIT fastpath:

```
smit hacmp
```
2. Select **Cluster Recovery Aids** and press Enter.
3. Select the **Release Locks Set By Dynamic Automatic Reconfiguration Event** option and press Enter.

SMIT displays a screen asking if you want to proceed. If you want to remove the SCD, press Enter.

Processing ODM Data During Dynamic Reconfiguration

When you synchronize the cluster topology or cluster resources, the processing performed by HACMP for AIX varies depending on the status of the Cluster Manager.

Note: Whenever you synchronize the cluster topology or cluster resources, HACMP for AIX verifies both the cluster topology and the cluster resource configurations. Verifying both the topology and resources, even though only one of the two has changed, is necessary because a change to one may invalidate the other.

The following variations may occur when synchronizing the topology or resources:

Cluster Manager Status	Synchronization results
The Cluster Manager is not running on any cluster node	If the Cluster Manager is not running on any cluster node, which is typically the case when a cluster is first configured, synchronizing the topology or resources updates the configuration data stored on each node that is reachable from the local node.
The Cluster Manager is running on the local node	If the Cluster Manager is running on the local node, synchronizing the topology or resources triggers a dynamic reconfiguration event. In processing this event, HACMP for AIX updates the configuration data stored on each cluster node that is reachable. Further processing makes the new configuration the currently active configuration.
The Cluster Manager is running on some cluster nodes but not on the local node	If the Cluster Manager is running on some cluster nodes but not on the local node, synchronizing the topology or resources updates the configuration data stored on each node that is reachable from the local node. However, the processing performed during a dynamic reconfiguration to make the new configuration the active configuration is <i>not</i> performed.

Undoing a Dynamic Reconfiguration

Before HACMP for AIX overwrites the configuration defined in the ACD, it saves a record of the configuration in a cluster snapshot. Only the **.odm** portion of a cluster snapshot is created; the **.info** file is not created. (For more information about cluster snapshots, see Chapter 11, Saving and Restoring Cluster Configurations.) If you want to undo the dynamic reconfiguration, you can use this cluster snapshot to restore the previous configuration.

HACMP for AIX saves snapshots of the last ten configurations in the default cluster snapshot directory, **/usr/sbin/cluster/snapshots**, with the name **active.x.odm**, where *x* is a digit between 0 and 9, with 0 being the most recent.

Restoring the ODM Data in the DCD

If a dynamic reconfiguration operation fails or is interrupted, you may want to restore the configuration in the DCD with the current active configuration, which is stored in the ACD. HACMP for AIX allows you to save in a snapshot the changes you made to the configuration in the DCD before you overwrite it.

To replace the ODM data stored in the DCD with the ODM data in the ACD, perform the following procedure.

1. Enter the SMIT fastpath:

```
smit hacmp
```

2. Select **Cluster Configuration** and press Enter.
3. Select the **Restore System Default Configuration from Active Configuration** option and press Enter.
4. Enter field values as follows:

**Cluster Snapshot Name
of System Default
HACMP ODMs**

In this field, you specify the name you want assigned to the cluster snapshot HACMP for AIX creates before it overwrites the ODM data stored in the DCD with the ODM data from the ACD. You can use this snapshot to save the configuration changes you made.

**Cluster Snapshot Description
of System Default
HACMP ODMs**

Enter any text string you want stored at the beginning of the snapshot.

Changing the Cluster Topology
Synchronizing the Cluster Topology

Chapter 7 Changing Resources and Resource Groups

This chapter describes how to reconfigure the resources and resource groups in your cluster. It describes changing application server resources in detail. It also covers the DARE Resource Migration utility that allows you to change the status and location of resource groups dynamically using the **cldare** command or the SMIT interface.

Reconfiguration and Resource Migration—Overview

Over time, you may need to change the attributes of existing resource groups or change the configuration of cluster resources such as application servers. To do this, you update the HACMP resource and resource group ODM information on any of the cluster nodes, using SMIT or the **xhacmpm** application. You then synchronize the changes so that the ODM data that defines the cluster, stored in the system default configuration directory (DCD), is updated on all the cluster nodes.

If the cluster is active when you perform the synchronization, the synchronization triggers a dynamic reconfiguration (DARE) event. In dynamic reconfiguration, HACMP replaces the currently active configuration, stored in the active configuration directory (ACD), with the changed configuration, and then refreshes the cluster daemons so that the new configuration becomes the currently active configuration.

Note: When dynamically reconfiguring a cluster, you can make changes either to the cluster topology or to cluster resources in a single operation, or you can migrate resources dynamically in a single operation. You cannot do more than one of these three activities at a time.

You can also dynamically migrate resource groups to other cluster nodes, at the command line using the **cldare** command and its associated flags, or by using the SMIT interface. DARE resource migration allows you to move, stop, or start a resource group without bringing the node down. This makes it possible to perform selective maintenance on a node without rebooting the cluster or losing accesses to the node's resources. See the section Migrating Resources Dynamically—Overview on page 7-17 in this chapter for more information about DARE resource migration.

Requirements before Reconfiguring Cluster Resources

Before making changes to cluster resources, ensure that:

- HACMP for AIX is installed on all nodes
- All nodes are up and running HACMP and able to communicate with each other
- Either all nodes include each other's boot and service interface IP labels (addresses) in their **/.rhosts** file as well as their own, or they are properly configured for Kerberos security.

Before starting, review your planning worksheets. See Chapter 1, Planning an HACMP Cluster, in the *HACMP for AIX Planning Guide* for more information about resource group names and types.

Cluster Resource Changes You Can Make Dynamically

The following lists the types of changes to cluster resources that you can make dynamically to an active cluster:

Resource/attribute	Dynamic Change Options
Resources	<ul style="list-style-type: none">• Add to resource group• Remove from resource group• Change start/stop scripts of application servers
Resource Groups	<ul style="list-style-type: none">• Add or remove a resource group• Bring a resource group up/online• Bring a resource group down/offline• Migrate a resource group to a different node location• Toggle the state of resource group attributes:<ul style="list-style-type: none">• Inactive Takeover• 9333 Disk Fencing• SSA Disk Fencing• Filesystems Mounted Before IP Configured
Node Relationship	<ul style="list-style-type: none">• Add nodes (already in cluster) to a resource group• Remove participating nodes from a resource group• Change the order (priority) of participating nodes• Change the resource policy, <i>except to/from rotating</i> (see notes on changing to/from rotating configuration on page 7-11.)

Note: Other configuration changes, such as changing the *name* of an application server or resource group, require that you stop cluster services before they become active. You can include such a change in a dynamic reconfiguration; however, HACMP for AIX interprets these changes, particularly name change, as defining a new cluster component rather than as changing an existing component. Such a change causes HACMP to stop the active component before starting the new component, causing an interruption in service.

Dynamic reconfiguration is not supported during a cluster migration.

Remember, you can perform only one type of dynamic reconfiguration change at a time. In other words, you can perform cluster configuration changes OR topology changes OR resource migration, but not more than one of these at a time.

Note: When reconfiguring a cluster dynamically, remove any dependencies that would be adversely affected by the change. For example, if you want to remove a node from the cluster, you must remove the node from all resource groups in which it participates. Similarly, if you want to remove a service adapter, you must remove the resource group that is using the associated service address or remove the service address from the resource group.

Emulating Dynamic Reconfiguration Events

You can emulate any dynamic reconfiguration or resource migration before actually performing it. This option is offered in the SMIT screens for synchronizing cluster resources or topology, and for performing DARE migrations. Emulation can also be specified at the command line with the **cldare** command.

The Event Emulator will not change the configuration of a cluster device. Therefore, if your configuration contains a process that makes changes to the Cluster Manager (disk fencing, for example), the Event Emulator will not show these changes. This could lead to a different output, especially if the hardware devices cause a failover.

You should add a cluster snapshot before running an emulation of a DARE event. Instructions for adding cluster snapshots are in Chapter 11, Saving and Restoring Cluster Configurations.

After you run the emulation, if you do not wish to run an actual DARE, you can restore the original configuration using the SMIT screen option **Restore System Default Configuration from Active Configuration**.

Reconfiguring Application Servers

An *application server* is a cluster resource used to control an application that must be kept highly available. Configuring or reconfiguring an application server does the following:

- Associates a meaningful name with the server application. For example, you could give the Image Cataloger demo supplied with the HACMP for AIX software a name such as *imserv*. You then use this name to refer to the application server when you define it as a resource during resource group configuration.
- Points the cluster event scripts to the scripts or programs that they call to *start* the server. These may be user-defined scripts specifically tailored to initialize the system before starting the actual application, or just the application itself with a set of command line options.
- Points the cluster event scripts to the scripts or programs used to *stop* the server. These may be user-defined scripts specifically tailored to bring down the server in a pre-defined order, or a simple kill signal.

Note: This section does not discuss writing the start and stop scripts. See the vendor documentation for specific product information on starting and stopping an application. Also see the appendix on applications and HACMP in the *HACMP for AIX: Planning Guide*.

Defining an Application Server

You can define an application server dynamically. You can define an application server and add it to a resource group in a single dynamic reconfiguration operation.

Complete the following steps to create an application server on any cluster node.

1. Type:

```
smit hacmp
```

2. When the SMIT main menu appears, select **Cluster Configuration > Cluster Resources > Define Application Servers > Add an Application Server**.

3. Enter field values as follows:

Server Name Enter a text string that identifies the server. You will use this name to refer to the application server when you define resources during node configuration. The server name can include alphabetic and numeric characters and underscores. Use no more than 31 characters.

Start Script Enter the pathname of the script (followed by arguments) called by the cluster event scripts to start the application server. This script must be in the same location on each cluster node that might start the server. The contents of the script, however, may differ.

Stop Script Enter the pathname of the script called by the cluster event scripts to stop the server. This script must be in the same location on each cluster node that may start the server. The contents of the script, however, may differ.

4. Press Enter to add this information to the ODM on the local node.
5. Press the F3 key to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes to the cluster you just made. To synchronize the cluster definition, go to the Cluster Resources SMIT screen and select the Synchronize Cluster Resources option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27.

Removing an Application Server

You can remove an application server from an active cluster dynamically. Before removing an application server, you must remove it from any resource group where it is included as a resource. For more information, see the section Adding or Removing Resources on page 7-12.

To remove an application server, complete the following steps:

1. In SMIT, go to the **Define Application Servers** screen, as described on page 7-4. From this menu, the select the **Remove an Application Server** option and press Enter.

2. Select the server you want to remove and press Enter.

A message appears asking if you are sure you want to remove the server.

3. Press Enter again to confirm the removal.

The server is removed from the ODM object classes stored in the DCD on the local node.

4. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see Synchronizing Cluster Resources on page 7-27 in this chapter.

Changing an Application Server

When you change the start or stop scripts associated with an application server, the ODM data is updated but the application server is not configured or unconfigured dynamically; thus the application controlled by the application server is not stopped and restarted. The next time the application is stopped, HACMP for AIX will call the new stop script—not the stop script that was defined when the application server was originally started.

To change (or view) an application server, complete the following steps:

1. In SMIT, go to the **Define Application Server** screen, as described on page 7-4. From this menu, select the **Change/Show an Application Server** option and press Enter.

SMIT displays a list of application servers.

2. Select the application server you want to change and press Enter. The **Change Application Server** screen appears, with the server name filled in.
3. Make the desired changes on this screen.

New Server Name Enter the new name that will identify the server you want to change. This field initially contains the name of the application server you selected. You may use alphabetic or numeric characters and underscores in the name. Use no more than 31 characters.

Note: Changing the name dynamically causes the application server to be stopped and restarted.

Start Script Enter the pathname of the script called by the cluster event scripts to start the application server. This script must be in the same location on each cluster node that might start the server. The contents of the script, however, may differ. Note that this field initially contains the path to the start script for the application server you selected.

Stop Script

Enter the pathname of the script called by the cluster event scripts to stop the server. This script must be in the same location on each cluster node that may start the server. The contents of the script, however, may differ. Note that this field initially contains the path to the stop script for the application server you selected.

4. Press Enter to add this information to the ODM stored in the DCD on the local node.
5. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section *Synchronizing Cluster Resources* on page 7-27.

Reconfiguring CS/AIX Communication Links

A CS/AIX communications link contains CS/AIX configuration information for a specific node and network adapter. This information enables an RS/6000 computer to participate in an SNA network that includes mainframes, PCs and other workstations.

Changes to a CS/AIX communications link may involve: changing the DLC name; changing the ports or link stations; or changing the service application start script. You can reconfigure CS/AIX communication links using the SMIT interface.

Defining a CS/AIX Communications Link

You can define a CS/AIX communications link dynamically. You can define a CS/AIX communications link and add it to a resource group in a single dynamic reconfiguration operation.

Complete the following steps to create a highly available CS/AIX communications link on any cluster node:

1. Type:

```
smit hacmp
```
2. Select **Cluster Configuration > Cluster Resources > Define Highly Available Communications Links > Define Communication Links**. This brings you to the main AIX menu for CS/AIX system configuration. Use the F1 help key if you need help with these SMIT entries. A valid CS/AIX configuration must exist before a CS/AIX DLC profile can be made highly available.
3. Press F3 to return to the **Define Highly Available Communications Links** screen. Select **Make Communications Links Highly Available > Add a Highly Available Communications Link**

4. Enter field values as follows:

DLC Name	Identify the CS/AIX DLC profile to be made highly available. Pick F4 to see a list of the DLC names.
Port	Enter ASCII text strings for the names of any CS/AIX ports to be started automatically.
Link Station	Enter ASCII text string for the names of any CS/AIX link stations. This field is only available if you are using CS/AIX Version 5.0.
Service	Enter the full pathname of any application start script. This start script starts any application layer processes that use the communications link. This field is optional.

5. Press Enter to add this information to the ODM on the local node.
6. Once you have defined a highly available CS/AIX communications link, you must configure it as a resource in a resource group. See the section, Reconfiguring Resource Groups and Resources in this chapter for information on adding a CS/AIX communications link to a resource group.
7. Press F3 key to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes to the cluster you just made. To synchronize the cluster definition, go to the Cluster Resources SMIT screen and select the Synchronize Cluster Resources option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27.

Removing a CS/AIX Communications Link

You can remove a CS/AIX communications link from an active cluster dynamically. Before removing a CS/AIX communications link, you must remove it from any resource group where it is included as a resource. For more information, see the section Adding or Removing Resources on page 7-12.

To remove a CS/AIX communications link, complete the following steps:

1. In SMIT, go to the **Make Communications Links Highly Available** screen, as described on page 7-6. From this menu, select the **Remove a Highly Available Communications Link** option and press Enter.
2. Select the communications link you want to remove and press Enter.
A message appears asking if you are sure you want to remove the communications link.
3. Press Enter again to confirm the removal.

The server is removed from the ODM object classes stored in the DCD on the local node.

4. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see Synchronizing Cluster Resources on page 7-27 in this chapter.

Changing a CS/AIX Communications Link

To change (or view) a CS/AIX communications link, complete the following steps:

1. In SMIT, go to the **Make Communications Links Highly Available** screen, as described on page 7-6. From this menu, select the **Change/Show a Highly Available Communications Link** option and press Enter.

SMIT displays a list of communications links.

2. Select the communications link you want to change and press Enter. The **Change Communications Link** screen appears, with the server name filled in.
3. Make the desired changes on this screen.

DLC Name	Make any changes to the CS/AIX communications link name.
Port	Make any changes to the CS/AIX ports to be started automatically.
Link Station	Make any changes to the CS/AIX link stations (CS/AIX version 5.0 only).
Service	Make any changes to the application start script. This is the full pathname of the start script that starts any application layer processes that use the communications link.

4. Press Enter to add this information to the ODM stored in the DCD on the local node.
5. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27.

Reconfiguring Resource Groups and Resources

In HACMP for AIX, you define each resource as part of a resource group. This allows you to combine related resources into a single logical entity for easier configuration and management. You then configure each resource group to have a particular kind of relationship with a set of nodes. Depending on this relationship, resources can be defined as one of three types: cascading, rotating, or concurrent access. You also assign a *priority* to each participating node in a cascading resource group chain.

To change the nodes associated with a given resource group or to change the priorities assigned to the nodes in a resource group chain, you must redefine the resource group. You must also redefine the resource group if you add or change a resource assigned to the group.

This section describes how to view, change, add, and delete a resource group. For more information about the initial configuration of cluster resources, see the chapter on configuring cluster resources in the *HACMP for AIX Installation Guide*.

Adding a Resource Group

You can add a resource group to an active cluster. You do not need to stop and then restart cluster services for the resource group to become part of the current cluster configuration.

To add a resource group, perform the following procedure:

1. From the Cluster Resources menu, select **Define Resource Groups** and press Enter.
2. Select **Add a Resource Group** and press Enter.
3. Enter field values as follows:

Resource Group Name	Enter a text string that identifies the resource group. You can use alphabetic or numeric characters and underscores. Use no more than 31 characters. If you specify a resource group name already in use, the command fails.
Node Relationship	Toggle the entry field to select among Cascading , Concurrent , and Rotating .
Participating Node Names	Enter the names of the nodes that you want to be members of the resource chain for this resource group. Enter the node names in order from highest to lowest priority (left to right). Leave a space between node names. Priority is ignored for concurrent resource groups.

4. Press Enter to add the resource group information to the ODM on the local node.
5. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27.

Removing a Resource Group

You can remove a resource group from an active cluster. You do not need to stop and then restart cluster services for the resource group to be removed from the current cluster configuration.

To remove a resource group, perform the following procedure.

1. From the Cluster Resources menu, select **Define Resource Groups** and press Enter.
2. Select the **Remove a Resource Group** option. SMIT displays a screen listing the defined resource groups.
3. Select the resource group you want to remove and press Enter.
SMIT displays a popup warning, reminding you that all information about the resource group will be lost.
4. Press Enter again to confirm your action.
5. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section *Synchronizing Cluster Resources* on page 7-27.

Changing Resource Groups and Node Relationships

Using the SMIT **Change/Show Cluster Resource** screen, you can perform any of the following tasks:

- Change the name of the resource group
- Add or remove nodes from the list of participating nodes
- Change the priority of participating nodes (by changing their position in the list of participating nodes)
- Change the relationship of the nodes to the resource group.

You can change many of the attributes of a resource group in an active cluster without having to stop and then restart cluster services. However, to change the name of a resource group, you must stop and then restart the cluster to make the change part of the current cluster configuration.

1. Enter the following SMIT fastpath:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Resources > Define Resource Groups > Change/Show a Resource Group**.

SMIT displays a list of the currently defined resource groups.

3. Select a resource group and press the Enter key.

SMIT displays a screen listing the present attributes of the resource group and their values.

4. Enter field values in the screen as follows:

Resource Group Name	This field is filled in from the previous menu selection.
New Resource Group Name	Enter the new name you want assigned to this resource group. Use no more than 31 characters. You can use alphabetic or numeric characters and underscores. Duplicate entries are not allowed. Note: HACMP interprets a name change as the inclusion of a new component, rather than as a change to an existing component, which will cause the resource group to be unconfigured.
Node Relationship	Toggle the entry field to select among Cascading, Concurrent, and Rotating. Note: Do not change a resource group from cascading to rotating in an active cluster. See the next section for changing a resource group to or from rotating.
Participating Node Names	Use this field to enter the names of the nodes that you want to add the resource group or remove from the resource group. You can also change the priority of the nodes in the resource group chain by changing the order in which they are specified. You enter node names in order from highest to lowest priority (left to right). Leave a space between node names. Priority is ignored for concurrent resource groups.

5. Press Enter to change the resource group information stored in the ODM.
6. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27 in this chapter.

Changing a Resource Group to or from Rotating

You cannot change a resource group to or from rotating in a single dynamic reconfiguration event. To make the change, do the following:

1. Remove the resource group and synchronize the cluster resource configuration, triggering a dynamic reconfiguration event. This should stop any application server configured in the resource group.
2. Make the appropriate changes to the adapter configuration and synchronize the cluster topology, triggering a dynamic reconfiguration event. You must at least change the node field for the service adapters associated with the resource group. You may also need to add

or remove boot adapters for the nodes participating in the rotating resource group. For more information about configuring a rotating resource group, see the *HACMP for AIX Installation Guide*.

3. Add a resource group with the proper node relationship, application servers, and other resources and synchronize the cluster resource configuration, triggering a dynamic reconfiguration event.

Adding or Removing Resources

You can add a resource to or remove a resource from a resource group in an active cluster without having to stop and restart cluster services to apply the change to the current configuration. You can add or remove resources from a resource groups even if a node is powered down; however, when a node is powered up, you can obtain a list of possible shared resources for each field by pressing the F4 key.

Resource groups can contain many different types of cluster resources, including IP labels, filesystems, volume groups and application servers. You can change the mix of resources in a resource group and the settings of other cluster resource attributes by using the SMIT **Change/Show Resources for a Resource Group** screen.

Converting from AIX Connections to AIX Fast Connect

If you previously configured the AIX Connections application as a highly available resource, and you now wish to switch to AIX Fast Connect, you should take care to examine your AIX Connections planning and configuration information before removing it from the resource group. Remember that you cannot have both of these applications configured at the same time in the same resource group, so you must unconfigure all AIX Connections realm/service pairs *before* configuring Fast Connect resources.

Keep in mind that AIX Fast Connect does not handle the AppleTalk and NetWare protocols that AIX Connections is able to handle. Fast Connect is primarily for connecting with clients running Windows operating systems. Fast Connect uses NetBIOS over TCP/IP.

Follow these steps when converting from AIX Connections to Fast Connect:

1. Refer to your original planning worksheet for AIX Connections, where you listed the participating nodes and the realm/service pairs you planned to configure. Compare this information to your Fast Connect planning worksheet so you can be sure you are not leaving anything out.

If you do not have your planning sheet, note the information you see in the AIX Connections Services field when you go into SMIT to remove the AIX Connections realm/service pairs from the resource group.
2. Start the Fast Connect server on each node and verify that you can connect to the shared directories and files on each node in turn.
3. In SMIT, go to the Change/Show Resource Groups screen, as described in the section below.
4. Select the AIX Connections Resources field and remove all specified realm/service pairs.
5. Select Fast Connect Services and specify the resources you wish to configure in the resource group. If you are specifying Fast Connect fileshares, make sure you have defined their filesystems in the Filesystems field earlier in the SMIT screen.

6. Synchronize the cluster as usual after you have made all changes. You'll find instructions for synchronization are at the end of this chapter, beginning on page 7-27.

Using SMIT to Reconfigure Resources

To change the resources in a resource group, perform the following procedure.

1. From the Cluster Resources menu, select the **Change/Show Resources/Attributes for a Resource Group** option and press Enter. SMIT displays a pick list of resource groups.
2. Select the resource group you want to change and press Enter.

SMIT displays a screen that lists all the types of resources that can be added to a resource group, with their current values.

3. Enter the field values as follows:

Resource Group Name	These fields are already filled in and are not editable in this screen.
Node Relationship	
Participating Node Names	
Service IP Label	If IP address takeover is being used, list the IP label to be taken over when this resource group is taken over. Press F4 to see a list of valid IP labels. These include addresses which rotate or may be taken over. Note: You must use IP address takeover if you configure a cascading resource group with an NFS mount point. You must also set the field Filesystems Mounted Before IP Configured (see below) to true.
Filesystems	Identify the filesystems to include in this resource group. Press F4 to see a list of the filesystems. When you enter a filesystem in this field, the HACMP for AIX system determines the correct values for the Volume Groups and Disks fields.
Filesystems Consistency Check	Identify the method of checking consistency of filesystems, fsck (default) or logredo (for fast recovery).
Filesystems Recovery Method	Identify the recovery method for the filesystems, parallel (for fast recovery) or sequential (default).
Filesystems/Directories to Export	Identify the filesystems or directories to be exported. These should be a subset of the filesystems listed above. Press F4 for a list.
Filesystems/Directories to NFS Mount	Identify the filesystems or directories to NFS mount. All nodes in the resource chain will attempt to NFS mount these filesystems or directories while the owner node is active in the cluster.

Network for NFS Mount

(This field is optional.)

Choose a previously defined IP network where you want to NFS mount the filesystems. The F4 key lists valid networks.

This field is relevant only if you have filled in the previous field. The **Service IP Label** field should contain a service label which is on the network you choose.

Note: You can specify more than one service label in the **Service IP Label** field. It is highly recommended that at least one entry be an IP label on the network chosen here.

If the network you have specified is unavailable when the node is attempting to NFS mount, it will seek other defined, available IP networks in the cluster on which to establish the NFS mount.

Volume Groups

Identify the shared volume groups that should be varied on when this resource group is acquired or taken over. Press F4 to see a list of shared volume groups. If you have previously entered values in the Filesystems field, the appropriate volume groups are already known to the HACMP for AIX system.

If you are using raw logical volumes in non-concurrent mode, you only need to specify the volume group in which the raw logical volume resides in order to include the raw logical volumes in the resource group.

Concurrent Volume Groups

Identify the shared volume groups that can be accessed simultaneously by multiple nodes. Press F4 to see a list of shared volume groups.

Raw Disk PVIDs

Press F4 for a listing of the PVIDs and associated hdisk device names. If you have previously entered values in the Filesystems or Volume groups fields, the appropriate disks are already known to the HACMP for AIX system.

If you are using an application that directly accesses raw disks, list the raw disks here.

**AIX Connections
Services**

If you have AIX Connections software installed, press F4 to choose from a list of all realm/service pairs that are common to all nodes in the resource group. You can also type in realm/service pairs. Use % as a divider between service name and service type; do not use a colon. *Note that you cannot configure both AIX Connections and AIX Fast Connect in the same resource group.*

**AIX Fast Connect
Resources**

If you have AIX Fast Connect software installed, press F4 to choose from a list of Fast Connect resources common to all nodes in the resource group. If you are adding Fast Connect fileshares, make sure you have defined their filesystems in the resource group. *Note that you cannot configure both AIX Connections and AIX Fast Connect in the same resource group.*

Note: If you are changing from AIX Connections to AIX Fast Connect, see the instructions on page 7-12.

Application Servers

Indicate the application servers to include in the resource group. Press F4 to see a list of application servers. See Defining an Application Server on page 7-4 for information on defining application servers.

**Highly Available
Communications Link**

Indicate the communications links to include in the resource group. Press F4 to see a list of communications links. See Defining a CS/AIX Communications Link on page 7-6 for information on defining communications links.

Miscellaneous Data

A text string you want to place into the environment along with the resource group information. This string is accessible by the HACMP for AIX scripts.

**Inactive Takeover
Activated**

Set this variable to control the *initial acquisition* of a resource group by a node when the node/resource relationship is cascading. This variable does not apply to rotating or concurrent resource groups.

If Inactive Takeover is **true**, then the first node in the resource group to join the cluster acquires the resource group, regardless of the node's designated priority.

If Inactive Takeover is **false**, the first node to join the cluster acquires only those resource groups for which it has been designated the highest priority node.

The default is **false**.

Cascading without Fallback	<p>Set this variable to determine the fallback behavior of a cascading resource group.</p> <p>When the CWOFF variable is set to false, a cascading resource group will fallback as a node of higher priority joins or reintegrates into the cluster.</p> <p>When CWOFF is true, a cascading resource group will not fallback as a node of higher priority joins or reintegrates into the cluster. It migrates from its owner node only if the owner node fails. It will not fallback to the owner node when it reintegrates into the cluster.</p> <p>Note: You may find it useful to review the <i>HACMP for AIX Concepts Guide</i> section, Fallover vs. Fallback on page 1-10.</p> <p>The default for CWOFF is false.</p>
9333 Disk Fencing Activated	<p>Once set, the values in a fence register can typically only be changed by power-cycling the 9333 unit. The fence register is immune to all other “reset” conditions. This can cause the IBM 9333 disk subsystem to be fenced out from the node (becoming) responsible for managing them. If you need to clear the fence registers for that disk to allow the node access to disk resources, use the command provided on the Cluster Recovery Aids SMIT screen. Clear these fence registers <i>in extraordinary circumstances only</i>.</p>
SSA Disk Fencing Activated	<p>If this field is set to true, the disk fencing of the IBM SSA disk subsystem is enabled. By default, this field is set to false. Disk fencing on SSA disk subsystems is only applicable for concurrent access configurations.</p> <p>Once set, the values in an SSA fence register cannot be changed by power-cycling the SSA unit. To clear the fence registers, use the Cluster Recovery Aids SMIT screen. For more information on SSA disk fencing, see the chapter on Planning Shared Disk Devices in the <i>HACMP for AIX Planning Guide</i>.</p>
Filesystems Mounted Before IP Configured	<p>This field specifies whether, on takeover, HACMP takes over volume groups and mounts a failed node’s filesystems before or after taking over the failed node’s IP address or addresses.</p> <p>The default is false, meaning the IP address is taken over first. Similarly, upon reintegration of a node, the IP address is acquired before the filesystems.</p> <p>Set this field to true if the resource group contains filesystems to export. This is so that the filesystems will be available once NFS requests are received on the service address.</p>

4. After entering data in the fields you want to change, press the Enter key.
5. Press F10 to exit SMIT or press F3 to return to previous SMIT screens to perform other configuration tasks or to synchronize the changes you just made. To synchronize the cluster definition, go to the **Cluster Resources** SMIT screen and select the **Synchronize Cluster Resources** option.

If the Cluster Manager is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event. For more information, see the section Synchronizing Cluster Resources on page 7-27.

Migrating Resources Dynamically—Overview

The Dynamic Reconfiguration (DARE) Resource Migration utility allows you to change the status or location of a resource group without stopping cluster services. This utility provides improved cluster management by allowing you to:

- Bring a resource group up/online (This option also restores a resource group to its default setting, removing any sticky markers. Sticky and non-sticky designations are discussed later in this chapter.)
- Bring a resource group down/offline
- Move a resource group to a new location
- Perform maintenance on a node without losing access to the node's resources
- Relocate resource groups for enhanced performance.

Using **cldare** from the command line you can change the status or location of multiple resource groups on multiple nodes at the same time; with the SMIT interface, you can perform one migration at a time. You can also disable resource groups dynamically, preventing them from being acquired during a reintegration. This disabling option allows a “swap” of resources in certain situations. (See page 7-21 for an example of such a situation.)

The rest of this chapter covers:

- Performing two types of migrations: sticky (permanent) and non-sticky.
- Using the **cldare** command and its associated keywords and arguments to perform dynamic migrations from the command line
- Using the SMIT interface to perform dynamic migrations
- Using **clfindres** to check resource group state and presence of sticky markers
- Removing sticky markers with the cluster up or down
- Synchronizing cluster resources.

Resource Migration Types

Before performing a resource migration, you must choose to declare the migration *sticky* or *non-sticky*.

Sticky Resource Migration

The sticky designation supersedes the defined node hierarchy of a resource group. A sticky migration to a specified node (“assigning the sticky attribute to a node”) moves the resource group to that node and gives it highest ownership priority of the resource group. Only one sticky attribute can be set in each resource group. The sticky attribute persists until a new resource group migration is done. Thus, it persists even after node failures and a cluster stop. In a next migration request, the user can reassign the sticky designation to another node, or when a non-sticky migration is issued, the sticky attribute will be removed, thus reinstating the default node hierarchy for that resource group.

A sticky attribute in a resource group affects the failover and fallback behavior, as well as the initial acquisition of a resource group during the start of the cluster, by changing the default node hierarchy. Despite this new node hierarchy, the resource group behavior follows the rules specified by CWOFF and Inactive Takeover flags. For example, in a cascading resource group with CWOFF set to **false**, after the reintegration of a node with a sticky attribute, the resource group will fallback onto that node. Furthermore, when the cluster starts and the Inactive Takeover flag is set to **false**, the first node up will acquire the resource group only if it is the sticky node.

In a rotating resource group, the sticky attribute does not alter the initial acquisition of a resource group. As always, the first node which joins the cluster will acquire it. The resource group will be acquired by the sticky node if the node which owns it fails.

If the sticky node fails, the resource group will fall over to the next highest priority node, which is the default node.

Non-Sticky Resource Migration

A non-sticky resource group migration is the default case when no sticky attribute is specified, and is allowed for rotating and Cascading without Fallback resource groups. Non-sticky resource groups are temporarily placed on the specified node until the next failover or reintegration occurs and all resource group location policies are reevaluated.

A cascading resource group with the CWOFF flag set to **false** cannot be moved by a non-sticky migration since it always resides on the highest priority node which is active. Thus, non-sticky migrations are supported only for cascading resource groups with CWOFF = **true**. A rotating resource group immediately resumes a normal rotating resource group failover policy after a non-sticky migration, but from the new location.

Note: You may find it helpful to perform a non-sticky DARE migration when doing maintenance on a CWOFF owner node. Should the default node fail, the CWOFF resource group will return to the owner node if it is available.

Note: You cannot mix node locations, default and stop keywords within a single DARE resource migration.

Migrating Resources Dynamically From the Command Line

This section explains how to perform DARE migrations using **cldare** at the command line. Instructions on using the SMIT interface to do this begin on page 7-22. Before performing either method of DARE migration, you should read the preceding overview sections.

The general syntax for migrating, starting, or stopping resource groups dynamically from the command line is as follows:

```
cldare -M <resource group name>:[location|keyword][:sticky] ...
```

-M	Specifies a migration action.
Resource Group Name	The name of the resource group to move or bring online.
Location (Node Name)	<p>To move a resource group or to bring a resource group online on a specific node, you enter a node name in the <i>location</i> field to specify the target node where the resource group will be relocated or started.</p> <p>The node name must be in the resource group's participating node list and the node must be available.</p>
Default Keyword	<p>If you use the <i>default</i> keyword to move or start a resource group, the DARE Resource Migration utility removes all previous stickiness for the resource group and returns the resource group to its default fallover behavior where normal configured node priorities apply.</p> <ul style="list-style-type: none">• A cascading resource group will migrate to the highest priority node currently up.• A rotating resource group will be released from wherever it resides and the highest priority node with a boot address will reacquire the resources. <p>If you include neither of the location keywords (<i>location</i>, <i>default</i>) in the field after the first colon, the DARE Resource Migration utility performs a <i>default</i> migration, again making the resources available for reacquisition. You can use this action to start a cascading resource group that has <code>INACTIVE_TAKEOVER</code> set to false and that has not yet started because its primary node is down.</p>
Stop Keyword	Using the keyword <i>stop</i> deactivates the resource group. If you also use the <i>sticky</i> keyword, the resources remain unavailable for reacquisition even after a fallover or reintegration. See Stopping a Resource Group on page 7-20 for more information about using the <i>stop</i> keyword.

sticky When moving or stopping a resource group, you can make the group *sticky* or non-sticky. Non-sticky is the default. You must specify *sticky* if you want this type.

Repeat this syntax on the command line for each resource group you want to migrate. Do not include spaces between arguments.

For full information on the **cldare** command and all of its associated flags, see Appendix B, Script Utilities, and the **cldare** man page.

Note: The **cldare** command attempts to perform all requested migrations simultaneously. If, for some reason, the command cannot simultaneously cause all specified resources to be released and reacquired at the new locations, it fails and no migrations occur.

Moving or Starting a Resource Group

Use the **cldare** command with the **-M** flag and a node location specified after the first colon to move a resource group from one node to another or start the resource group if it was previously inactive. The syntax for moving or starting a resource group is as follows:

```
cldare -M <resgroup name>:<target node> [:sticky]
```

You can specify multiple move commands on the command line. The **cldare** utility first parses and checks the command line and verifies the cluster configuration; then the resources are released and reacquired by the specified cluster node. The resource migration process first releases all specified resource groups (wherever they reside in the cluster), then causes the newly specified node to acquire the resource groups.

Attaching the *sticky* designation causes the resource group to attempt to return to the specified node after failover and reintegration, superseding the normal resource policy.

Stopping a Resource Group

The DARE facility provides a command to stop a resource group; that is, to bring that resource group offline. You can stop a resource group in a sticky and in a non-sticky way. The former is referred to as a DARE sticky stop, the latter as a DARE stop. A resource group which has been brought offline via a DARE stop remains offline for as long as no cluster event occurs. When a cluster event does occur, the state of the resource group is undefined. It may have been brought online or it may stay offline. If you want to bring a resource group permanently offline, use a DARE sticky stop. A sticky-stopped resource group will stay offline until a new DARE migration request for that resource group is issued, which brings it online.

To stop a resource group, use the keyword *stop* instead of an actual nodename in the location field, as follows:

```
cldare -M <resgroup name>:stop [:sticky]
```

The DARE Resource Migration utility brings the resource group offline, which includes taking down any service label, unmounting filesystems, etc.

As with sticky *locations*, sticky *stop* requests are in effect until they are:

- superseded by new sticky migration requests for the same resource group

- removed by *default*, non-sticky migration requests for the same resource group.

Note: The sticky designation is optional, but usually you will want to use it with the *stop* request. Be careful when using a non-sticky *stop* request, since the resource group will likely be restarted at the next major cluster event. As a result, all non-sticky requests produce warning messages. A non-sticky stop could be used to halt a cascading resource group that has `INACTIVE_TAKEOVER` set to `false` during periods in which its primary node is down.

Special Considerations when Stopping a Resource Group

After DARE stopping resource group, you should not assume that a joining or rejoining node will bring that resource group online. The following are instances when a resource group must be brought back online with a DARE command.

1. If you use **cldare stop** to bring down a cascading resource group (with `CWOF = false`) which is assigned an Inactive Takeover value of **false**, and which resides on the highest priority node, it will remain in an inactive state. You must manually bring the resource group online through DARE migration.
2. In a cascading resource group with Cascading without Fallback set to **true**, the possibility exists that while the highest priority node is up, the resource group is down. This situation can occur if the resource group does not reside on its highest priority node, and you bring it down by either a graceful shutdown or by a **cldare stop** command. Unless you bring the resource group up manually, it will remain in an inactive state.

For more information on this issue, see Miscellaneous Issues on page 48-25 of the HACMP for AIX Troubleshooting Guide.

Starting a Resource Group with the Default Keyword

To start a resource group using the **cldare** command with the *default* keyword after the first colon, enter the command as follows:

```
cldare -M <resgroup name>:default [:sticky]
```

You need not specify a target node; the resource group is activated on the node that has been designated as its highest priority node.

You can use the *default* keyword when you want to restore a resource group to its original state, for example to remove an earlier sticky designation.

You can attach the *sticky* keyword to this start command if you want the resource group to stay on that node after a failure or reintegration of another cluster node.

Note: When using the SMIT interface to bring a resource group online, *sticky* is not an option. The resource group will failover and fallback to other nodes just as the resource policy dictates.

Example: Using cldare to Swap Resource Groups

In the three-node cluster indicated here, each node—Node1, Node2, and Node3—has a service adapter and a standby adapter.

There are three cascading resource groups with node priority lists as follows:

RG1 - Node1, Node3 CrucialRG - Node2, Node3 RG3 - Node3, Node1

Each node is up and possesses a resource group as follows:

Node1 - UP (RG1) Node2 - UP (CrucialRG) Node3 - UP (RG3)

Suppose Node2's resources—contained in CrucialRG—are of particular importance to your operation. A situation occurs in which two cluster nodes fail. Node1 fails first; its resources fall over to Node3, since Node3 is in RG1's priority list. Then Node2 fails. In this case, Node2's crucial resources remain down; they have nowhere to go, since Node3's only standby adapter has already been taken. The cluster now looks like this:

Node1 - DOWN Node2 - DOWN Node3 - UP (RG3, RG1)

The crucial resource group is unavailable. HACMP is able to take care of only one failure, because there are no more standby adapters, so it handles the first failure, Node1, but not the second. However, if you need CrucialRG's resources more than you need RG1's, you can use the DARE resource migration utility to “swap” the resource groups so you can access CrucialRG instead of RG1.

You do this by issuing a *stop* command to bring RG1 offline, then issue a second command to bring CrucialRG online on Node3, as follows:

```
cldare -M rg1:stop:sticky -M crucialrg:node3
```

Note that both commands can be typed on the same line with no space between them.

After these DARE migration commands are completed, access to CrucialRG is restored, and the cluster looks like this:

Node1 - DOWN Node2 - DOWN Node3 - UP (RG3, CrucialRG)

Migrating Resources Dynamically Using SMIT

You can also access the DARE migration functions using the SMIT interface.

From the main hacmp menu, choose **Cluster System Management > Cluster Resource Group Management**.

Here SMIT presents the following options for dynamic resource migration.

Bring a Resource Group Online

This is equivalent to using the **cldare default** keyword to start a resource group. Choose this to activate all resources in a specified resource group on the node designated as highest priority for that resource group. You can also use this option to return a resource group to its default setting, removing any previously set sticky designation.

Bring a Resource Group Offline

This is equivalent to the **cldare stop** keyword. Use this option to deactivate all resources in a specified resource group. You can choose sticky or normal migration.

Move a Resource Group This option allows you to deactivate a resource group on one node and then activate it on another node. You can choose sticky or normal migration.

For each migration option, SMIT offers the choice to emulate the migration rather than actually perform it.

Bringing a Resource Group Online

Bringing a resource group online is the equivalent of using the *default* keyword with the **cldare** command. This command activates (starts) the specified resource group on the node identified as its highest priority node, or moves the resource group back to its initially designated node, removing any previously assigned stickiness.

1. From the main hacmp menu, choose **Cluster System Management > Cluster Resource Group Management > Bring a Resource Group Online**.
2. Select the resource group from the picklist and press Enter.
3. The Bring a Resource Group Online screen appears. Press the Tab key to toggle through valid entries, or press F4 to view and choose from a picklist of entries. You cannot add entries manually in any of the fields.

Note: When using SMIT to bring a resource group online, *sticky* is not an option. The resource group will fallover and fallback to other nodes just as the resource policy dictates.

Enter field values as follows:

Resource Group to Bring Online	Shows the name of the resource group that you have chosen to activate. You cannot edit this field here.
Emulate or Actual?	Choose Actual (the default) to bring the resource group online. Choosing Emulate to run the command in a mode that does not make any changes to your cluster, but still displays the results/messages as if the action was being carried out.
Perform Verification First?	Specify whether you want to run the clverify utility to verify configuration before performing the migration. The default is Yes . It is recommended that any prospective configuration changes be verified before execution. However, if you are certain that the new configuration is appropriate, the verification step may be skipped to quicken the process.
Ignore Cluster Verification Errors?	You can choose to have the DARE action continue even if configuration errors are detected during verification. The default is No . It is recommended that any prospective configuration changes be verified before execution. However, there may be instances when it is appropriate to proceed despite verification failure. If this is the case, choose Yes .

To return to the Cluster Resource Group Management menu, press F3 at any time.

Bringing a Resource Group Offline

Bringing a resource group offline is equivalent to using the **clbare** command and the *stop* keyword in the location field. This action deactivates, or stops, a resource group.

Take the following steps to bring a resource group offline.

1. From the main hacmp menu, choose **Cluster System Management > Cluster Resource Group Management > Bring a Resource Group Offline**.
2. Select a resource group from the picklist.
The Bring a Resource Group Offline screen appears.
3. To bring the resource group offline, enter field values as follows:

Resource Group to Bring Offline

Indicates the Resource Group you have selected. This field cannot be edited here.

Use Sticky Migration?

Specify sticky or normal (nonsticky) migration. The default is **No**, causing a nonsticky migration. If you choose **yes**, the DARE special migration location *stop* becomes the highest-priority location for that resource group (until the sticky migration is turned off). The resource group attempts to remain offline during a cluster fallover or reintegration. If you choose **No**, the resource group will be stopped, but its highest-priority location will remain unchanged, and the resource group may be restarted at the time of the next cluster event.

Emulate or Actual?

Choose **Actual** to bring the resource group online. Choosing **Emulate** to run the command in a mode that does not make any changes to your cluster, but still displays the results/messages as if the action was being carried out.

Perform Verification First?

Specify whether you want to run the **clverify** utility before performing the migration. The default is **Yes**. It is recommended that any prospective configuration changes be verified before execution. However, if you are certain that the new configuration is appropriate, the verification step may be skipped to quicken the process.

Ignore Cluster Verification Errors?

You can choose to have the DARE action continue even if errors are detected during verification. The default is **No**. It is recommended that any prospective configuration changes should be verified before execution. However, there may be instances when it is appropriate to proceed despite verification failure. If this is the case, choose **Yes**.

Press F3 at any time to go back to the Cluster Resource Group Management menu.

Moving a Resource Group

Take the following steps to move a resource group.

1. From the main hacmp menu, choose **Cluster System Management > Cluster Resource Group Management > Move a Resource Group**.
2. Select a resource group from the picklist. The **Move a Resource Group** screen appears.
3. Enter field values as follows:

Resource Group To Be Moved	Indicates the Resource Group you have selected. You cannot edit this field here.
Move Resource Group To Which Node?	<p>Enter the name of the target node to which you are moving the resource group; press F4 for a picklist of the cluster nodes participating in that resource group to which the resource group can migrate—the list does not include the node the resource group is presently on. The target node you specify acquires the resource group.</p> <p>If you leave this field blank or enter an invalid node name in this field, an error message is displayed.</p>
Use Sticky Migration?	Specify sticky or normal (non-sticky) migration. The default is No , causing a nonsticky migration. If you choose Yes , the specified target node becomes the highest-priority node for that resource group (until the sticky migration is turned off). The resource group will attempt to remain on that target node—regardless of that node’s original priority for the resource group—during a cluster failover or reintegration.
Emulate or Actual?	Choose Actual to bring the resource group online. Choosing Emulate to run the command in a mode that does not make any changes to your cluster, but still displays the results/messages as if the action was being carried out.
Perform Verification First?	Specify whether you want to run the clverify utility before performing the migration. The default is Yes . It is recommended that any prospective configuration changes be verified before execution. However, if you are certain that the new configuration is appropriate, the verification step may be skipped to quicken the process.
Ignore Cluster Verification Errors?	You can choose to have the DARE action continue even if errors are detected during verification. The default is No . It is recommended that any prospective configuration changes should be verified before execution. However, there may be instances when it is appropriate to proceed despite verification failure. If this is the case, choose Yes .

Press F3 at any time to go back to the Cluster Resource Group Management menu.

Checking Resource Group State and Sticky Markers

The following information will be of use when you wish to check the state of your cluster resource groups at times, identify if they have sticky markers persisting from earlier resource migrations, and remove the sticky markers.

Checking Resource Group State with the `clfindres` Command

To help you check the location and state of resources placed on a specific node, the DARE Resource Migration utility includes a command, `clfindres`, that displays the state and location of specified resource groups. It also indicates whether a resource group has a sticky location, and it identifies that location.

See Appendix A, HACMP for AIX Commands for the complete syntax and typical output of the `clfindres` command.

Be aware that persistent sticky location markers are saved and restored in cluster snapshots (discussed in Chapter 11, Saving and Restoring Cluster Configurations), and can cause confusion if a snapshot is restored and the user does not realize the sticky markers are there.

Removing Sticky Markers

When the Cluster is Running

Sticky location markers are stored in the HACMPresource class in the HACMP ODM and are a persistent cluster attribute. While the cluster is *up*, you can only remove these locations by performing a subsequent non-sticky migration on the same resource group, using the *default* special location keyword or specifying no location.

When the Cluster is Down

If you want to remove sticky location markers while the cluster is *down*, do not use the *default* keyword, since it implies activating the resource. Instead, when the cluster is down, you use a non-sticky **stop** request, as in this example:

```
cldare -v -M <resgroup name>:stop
```

(The optional `-v` flag indicates to skip verification.)

Synchronizing Cluster Resources

Whenever you modify the configuration of cluster resources in the ODM on one node, you must synchronize the change across all cluster nodes. You perform a synchronization by choosing the **Synchronize Cluster Resources** option from the Cluster Resources SMIT screen.

Note: In HACMP for AIX, the event customization information stored in the ODM is synchronized across all cluster nodes when the cluster resources are synchronized. Thus, pre, post, notify, and recovery event script names must be the same on all nodes although the actual processing done by these scripts can be different.

The processing performed in synchronization varies depending on whether the Cluster Manager is active on the local node:

- If the cluster manager is *not* active on the local node when you select this option, the ODM data in the DCD on the local node is copied to the ODMs stored in the DCDs on all cluster nodes.
- If the Cluster Manager is active on the local node, synchronization triggers a cluster-wide, dynamic reconfiguration event. In dynamic reconfiguration, the configuration data stored in the DCD is updated on each cluster node and, in addition, the new ODM data replaces the ODM data stored in the ACD on each cluster node. The cluster daemons are refreshed and the new configuration becomes the active configuration. In the HACMP for AIX log file, **reconfig_resource_release**, **reconfig_resource_acquire**, and **reconfig_resource_complete** events mark the progress of the dynamic reconfiguration.
- If the Cluster Manager is active on some cluster nodes but not on the local node, the synchronization is aborted.

For more information on these variations see the section Processing ODM Data During Dynamic Reconfiguration in Chapter 6, Changing the Cluster Topology.

To synchronize changes to the cluster resources across all cluster nodes, perform the following procedure.

1. From the **Cluster Configuration** menu, select **Cluster Resources > Synchronize Cluster Resources**.
2. Enter the following values:

Ignore Cluster Verification Errors?

If this field is set to **no**, synchronization aborts if verification of the new configuration fails. As part of dynamic reconfiguration processing, the new configuration is verified before it is made the active configuration. If you want synchronization to proceed even if verification fails, set this value to **yes**.

The default is **no**.

- Un/Configure Cluster Resources?** If you set this field to **yes**, HACMP for AIX changes the definition of the resource in the ODM and it performs any configuration triggered by the resource change. For example, if you remove a filesystem, HACMP for AIX removes the filesystem from the ODM and also unmounts the filesystem.
- If you set this field to **no**, HACMP for AIX changes the definition of the resource in the ODM but does not perform any configuration processing that the change may require. For example, a filesystem would be removed from the HACMP for AIX cluster definition but would not be unmounted. This processing is left to be performed by HACMP for AIX during a fallover.
- The default is **yes**. HACMP for AIX attempts to limit the impact on the resource group when a component resource is changed. For example, if you add a filesystem to the resource group that already includes the underlying volume group as a resource, HACMP for AIX does not require any processing of the volume group. Other modifications made to the contents of a resource group may cause the entire resource group to be unconfigured and reconfigured during the dynamic reconfiguration. Cluster clients will experience an interruption in related services while the dynamic reconfiguration is in progress.
- Emulate or Actual** If you set this field to **Emulate**, the synchronization is an emulation and does not affect the Cluster Manager. If you set this field to **Actual**, the synchronization actually occurs, and any subsequent changes affect the Cluster Manager. **Actual** is the default value.
- Skip Cluster Verification** By default, this field is set to **no** and the verification of cluster resources program is run.
- If the cluster is inactive, you can toggle this entry field to **yes** to save time in the synchronization process. By doing so, cluster verification will be skipped.
- Note:** In some cases, the verification uncovers errors that do not cause the synchronization to fail. HACMP reports the errors in the SMIT command status window so that you are aware that something about the configuration may be a problem. You should investigate any error reports, even when they don't interfere with the synchronization.

Skipping Cluster Verification During Synchronization

To save time during synchronization of cluster resources, you can skip cluster verification. This option is available only when a cluster is *inactive*. Even if one node is active, cluster verification will be run.

You can skip cluster verification using SMIT or at the command line, using the **cldare** flag **-x**.

To skip verification of cluster resources using the SMIT interface, choose **YES** at the **Skip Cluster Verification** field.

Changing Resources and Resource Groups
Synchronizing Cluster Resources

Chapter 8 Verifying a Cluster Configuration

This chapter describes how to verify the HACMP modifications to AIX system files and your cluster configuration (topology and resources). It covers the **clverify** utility and how to run cluster verification in three ways: interactively, from the command line, or by using SMIT.

The final sections of the chapter cover custom verification methods.

Overview

Verifying the cluster configuration assures you that all resources used by the HACMP for AIX software are configured correctly, and that ownership and takeover of those resources are defined and in agreement across all nodes. You should verify the cluster configuration after making changes to your hardware or software. You may want to include the **clverify** utility in a cron job to perform cluster verification nightly.

To ensure that an HACMP cluster works properly, you must verify that the HACMP-specific modifications to AIX system files are made correctly, and that the cluster topology and resources are configured correctly, before starting the system. The *software verification* procedure notifies you if a problem exists with the changes HACMP has made to AIX system files. The *cluster verification* procedure ensures that all nodes are in agreement on cluster topology, network configuration, and ownership and takeover of resources.

Three Ways to Run Cluster Verification

You can run cluster verification in three ways:

**Interactively
(using a
command menu)**

To run clverify “interactively,” you start with a simple command, **/usr/sbin/cluster/diag/clverify**, and select options from a series of menus until you get to the specific option you want to run.

This option includes a help facility. To use the help facility, type **help** followed by one of the listed options you need help with. For example, if you type **help config**, you see a screen that contains a brief message explaining the purpose of the option, and returns you to the previous prompt to make your next selection.

**Directly from the
command line**

If you already know the option you want to run, and know the complete syntax, you can type the full command and the appropriate options and arguments to run clverify immediately.

**Using the SMIT
interface**

You can use the SMIT interface to run all clverify options except software.

In the following sections on software verification and cluster verification, the interactive (command menu) method is explained first, followed by the full syntax for running clverify options directly from the command line.

A separate section, beginning on page 8-7 covers how to perform cluster verification using SMIT.

The clverify Utility

The **clverify** utility supports several nested options.

- **software**—verifies that the HACMP-specific modifications to AIX system files exist and are correct. It has one option, **lpp**.
- **cluster**—allows you to verify the configuration agreement using the **topology** and **config** options:
 - **topology**—verifies that all nodes agree on the cluster configuration. This program has two options, **check** and **sync**:
 - **check**—tells you if the nodes are in agreement. For example, checks for invalid characters in cluster names, node names, network names, adapter names and resource group names.
 - **sync**—synchronizes the cluster topology, if necessary, so that all nodes agree.
 - **config**—verifies that networks are configured correctly, and that all nodes agree on the ownership and takeover of defined resources. This program has three options, **networks**, **resources**, or **all**:
 - **networks**—checks for the valid configuration of network adapters and **tty** lines on all cluster nodes. Also checks each cluster node to determine whether multiple RS232 serial networks exist on the same tty device.
 - The **resources** option
 - checks for agreement among all nodes on the ownership of defined resources (filesystems, volume groups, disks, application servers, events, AIX Connections, and CS/AIX communications links) and on the distribution of resources in case of a takeover
 - checks, more specifically, for the existence and defined ownership of filesystems to be taken over
 - checks the volume group and disks where the filesystems reside to verify that the takeover information matches the owned resources information
 - checks the major device numbers for NFS-exported directories
 - checks to ensure that application servers are configured correctly
 - prints out diagnostic information about custom snapshot methods, custom verification methods, custom pre/post events, and redirection of cluster log files.
 - **all**—checks the network topology and resources, and runs all custom-defined verification methods. (For more information about defining custom verification methods, see Adding a Custom Verification Method on page 8-8.)

If you have configured Kerberos on your system, the **clverify** utility also verifies that:

- All IP labels listed in the configuration have the appropriate service principals in the **.klogin** file on each node in the cluster.
- All nodes have the proper service principals.
- Kerberos is installed on all nodes in the cluster.
- All nodes have the same security mode setting.

Specifying Environmental Options for clverify

You can use either or both of the following arguments to qualify the environment for the clverify program options:

-e *num* Allows you to specify the maximum number of errors clverify can encounter before aborting the check, where *num* is the number of errors you specify. For example:

```
clverify cluster topology check -e 3
```

If you do not use the -e flag, then by default, clverify continues to run all checks until complete, no matter how many errors it finds.

-R *file* Specifies the name of a file into which output from the verification procedure should be saved, where *file* is the name of the file. For example, to store the information from software verification in a file called warnings.out, you type:

```
clverify lpp -R warnings.out
```

Verifying Software

To ensure that an HACMP cluster works properly, you must verify that the correct HACMP-specific modifications to the AIX system files exist. The clverify utility software option performs this task for you.

Note: You should run the software verification procedure before verifying your cluster configuration.

Interactive (command menu) method

Use this method if you wish to be shown menus of valid options to choose from. To quit the interactive mode at any time, type CTRL-C or **quit**.

To run the software verification procedure interactively:

1. Type:

```
/usr/sbin/cluster/diag/clverify
```

The command returns a list of command options, and the clverify prompt at the bottom of the screen. The screen looks like this:

```
-----
To get help on a specific option, type: help <option>
To return to previous menu, type: back
To quit the program, type: quit
-----Valid
Options are:
software
cluster
clverify>
```

2. Type:

```
software
```

The following option and an updated prompt appear:

```
Valid Options are:  
lpp  
clverify.software>
```

3. To verify the HACMP for AIX modifications to AIX system files, and to log the results in a file called **verify_hacmp**, use the **-R** argument. Type:

```
lpp -R verify_hacmp
```

When the program completes, read the **verify_hacmp** file. If no problems exist, no messages are logged. If the software modifications are not correct, you receive a message similar to the following that informs you of errors.

```
inetd is not configured for HACMP.  
The /.rhost file does not exist.
```

Note: If you receive messages about configuration errors but have not yet configured the cluster, ignore them.

4. Type CTRL-C or **quit** to return to the system prompt.

Command Line Method

To run the `clverify` software option directly from the command line, you must type the entire command and desired arguments at the system prompt.

The full syntax for verifying software from the command line is as follows:

```
/usr/sbin/cluster/diag/clverify software {lpp} [-e num] [-R file]
```

Verifying the Cluster

After verifying the software, verify your cluster. The `clverify` cluster options check that all resources are configured correctly and that ownership and takeover of resources are defined and consistent across all nodes.

Interactive (command menu) method

When you run the **clverify** utility in the interactive mode, you step through the valid programs and their options until you get to the specific option you want to run. If you are unsure of which option to run, the interactive mode includes a help facility to provide information on a specified option.

Note: If you already know which option you want to run and you know the complete syntax the utility needs to verify a given feature, you can skip this section and run **clverify** immediately by typing the command and its required options at the system prompt, as described in the Command Line Method section starting on page 8-6.

To run the `clverify` cluster option interactively:

1. Type:

```
/usr/sbin/cluster/diag/clverify
```

The command returns a list of command options and the **clverify** prompt:

```
-----  
To get help on a specific option, type: help <option>  
To return to previous menu, type: back  
To quit the program, type: quit  
-----
```

```
Valid Options are:  
software  
cluster  
clverify>
```

2. Type the **cluster** subcommand:

```
cluster
```

The following options and an updated prompt appear:

```
...  
Valid Options are:  
topology  
config
```

```
clverify.cluster>
```

3. To see the options for the **topology** program, type:

```
topology
```

The following options and the updated prompt appear:

```
Valid Options are:  
check  
sync
```

```
clverify.cluster.topology>
```

4. To verify that all nodes agree on the cluster topology, type:

```
check
```

When the program finishes, review the output. If a cluster topology problem exists, you receive a message similar to the following:

```
ERROR: Could not read local configuration  
ERROR: Local Cluster ID XXX different from Remote Cluster ID XXX.  
ERROR: Nodes have different numbers of networks
```

5. (Optional) To synchronize the cluster topology on all cluster nodes as it is defined on the node where you are running the **clverify** program, type the following command at the `clverify.cluster.topology` prompt:

```
sync
```

You should run this program only if the **topology check** program reveals disagreement among nodes on the cluster topology and if you are sure you want the configuration as it is defined on the local node.

6. To verify that all nodes agree on the ownership of all defined resources, and that networks and interfaces are properly configured, type **back** until you return to the `clverify.cluster` prompt, then type:

```
config
```

The system lists three options to the **config** program. If you choose **all**, the system runs the **networks** program followed by the **resources** program followed by any custom verification methods. The **networks** and **resources** programs can be run separately.

7. To verify that all networks and interfaces are configured correctly and that results are logged in a file called **verify_nw**, type:

```
networks -R verify_nw
```

When the program completes, review the **verify_nw** file. If no problems exist, no messages are logged. If a problem with any node's adapters or **tty** lines exists, you receive a message similar to the following:

```
ERROR: The serial device XXX does not exist on node XXX  
ERROR: Service adapter XXX is improperly configured on node XXX
```

8. To verify that cluster nodes agree on the ownership of defined resources, that all takeover resources are assigned (and are in agreement), and that results are logged in a file called **verify_own**, type:

```
resources -R verify_own
```

The program checks information on the cluster configured resources (filesystems, NFS filesystems, volume groups, disks, and application servers) stored in the global ODM. When the program completes, review the **verify_own** file. If no problems exist, no messages are logged. If any resource is configured incorrectly, you receive a message similar to the following:

Checking:

```
Owned Resources Verification.  
clver: filesystem /sand is improperly configured on node seaweed  
clver: filesystem /home is configured to auto-mount on node seaweed  
clver: volume group sandvg is improperly configured on node seaweed  
clver: volume group homevg is configured to auto-varyon on node  
seaweed  
clver: volume group sand2vg is improperly configured on node seashell  
clver: application clmarket's start-file  
(/usr/sbin/cluster/demos/clmarket) does not exist on node seaweed  
clver: application clmarket's stop-file  
(/usr/sbin/cluster/demos/clmarket) does not exist on node seaweed  
clver: physical disk hdisk3, used in owned volume group homevg is not  
owned on node seaweed
```

Command Line Method

Once you have run **clverify** several times, you may be familiar with the utility's options. If so, you can type the complete syntax in the command line to run a specified **clverify** option. The complete syntax for the **clverify** command is as follows:

```
clverify cluster {topology check | topology sync |  
config networks | config resources | config all}[-e num][-R file]
```

Include [-e *num*] to abort the program after *num* errors.

Include [-R *file*] to redirect output to a specified file

For example, to verify the cluster networks configuration and to log the results in a file called **verify_nw**, type;

```
/usr/sbin/cluster/diag/clverify cluster config networks -R verify_nw
```

If you type an incomplete command, the utility lists the options for the level typed and puts you in interactive mode to complete the command. For example, if you type:

```
clverify cluster
```

You get the list of valid options:

```
topology  
config
```

```
clverify.cluster>
```

To quit the program at any time, type CTRL-C or **quit**.

Verifying the Cluster Configuration Using SMIT

After reconfiguring or updating a cluster, run the cluster verification utility (**clverify**) on one node to check that all resources used by the HACMP for AIX software are configured correctly, and that ownership and takeover of those resources are defined (and are in agreement) across all nodes. You can do this through the SMIT interface as well as the command line.

Through SMIT, you can also add new custom verification methods to the clverify utility that check specific components within your cluster configuration. You can change or remove these methods from the verification process depending on the level of cluster verification you want to perform. (See the sections on custom verification methods beginning on page 8-8.)

Verifying the Cluster using SMIT

To verify the cluster network topology and resource configurations using the SMIT interface, perform the following steps.

1. To reach the HACMP SMIT menu, type:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Verification > Verify Cluster**.

Fill in the fields as follows:

Base HACMP Verification Methods	By default, both the cluster topology and resources verification programs are run. You can toggle this entry field to run either program, or you can select none to specify a custom-defined verification method in the Define Custom Verification Method field.
Define Custom Verification Method	Enter the name of a custom-defined verification method. You can also press F4 for a list of previously defined verification methods. By default, if no methods are selected, the clverify utility also will not check the base verification methods, and it generates an error message. The order in which verification methods are listed determines the sequence in which selected methods are run. This sequence remains the same for subsequent verifications until different methods are selected.
Error Count	By default, the program will run to the end, even if it finds many errors. To cancel the program after a specific number of errors, specify the number in this field.

Log File to store output Enter the name of an output file in which to store verification output. By default, verification output is stored the **smit.log** file.

3. Press Enter.

SMIT runs the **clverify** utility. The output from the verification is displayed in the SMIT Command Status window. If you receive error messages, make the necessary changes and run the verification procedure again.

4. To synchronize all cluster nodes, use the **Synchronize Cluster Topology** option on the **Cluster Topology** SMIT menu, and **Synchronize Cluster Resources** under the SMIT **Cluster Resources** menu.

If you choose, you can skip verification of cluster topology and cluster resources in order to save time during synchronization.

For more information, see the sections Synchronizing the Cluster Topology on page 6-18 and Synchronizing Cluster Resources on page 7-27.

Adding a Custom Verification Method

The HACMP for AIX software provides an ODM class, **HACMPcustom**, that allows you to store custom-defined verification methods. The new ODM class includes the method name, full pathname to the method, a method description and type definition.

To add a custom verification method:

1. Select **Cluster Configuration > Cluster Verification > Define Custom Verification Method > Add a Custom Verification Method**.
2. Fill in the fields as follows:

Verification Method Name Enter a name for the verification method. Method names can be 32 bytes long and can contain alphanumeric characters. Do not use the word “all,” as this is a keyword indicating that all custom verification methods are to be run.

Verification Method Description Enter a description of the verification method. Method descriptions can be 1024 bytes.

Verification Method Enter the full path executable filename of the verification method. The method name can be different from the filename and can be 1024 bytes.

3. Press Enter.

SMIT runs the **clverify** utility. The output from the verification is displayed in the SMIT Command Status window.

Changing a Custom Verification Method

To change a custom verification method:

1. Select **Cluster Verification > Define Custom Verification Method > Change/Show a Custom Verification Method**.

SMIT displays a picklist of custom verification methods.

2. Select a specific verification method and press Enter.

SMIT displays the current information.

3. Enter any new information (name, description, filename) for the verification method.
4. Press Enter.

SMIT runs the **clverify** utility. The output from the verification is displayed in the SMIT Command Status window.

Removing a Custom Verification Method

To remove a custom verification method:

1. Select **Cluster Configuration > Cluster Verification > Define Custom Verification Method > Remove a Custom Verification Method**.

SMIT displays a picklist of custom verification methods.

2. Select the verification method you want to remove and press Enter.

SMIT prompts you to confirm that you want to remove the specified verification method.

3. Press Enter to remove the verification method, or press F3 to cancel.

Verifying a Cluster Configuration
Verifying the Cluster Configuration Using SMIT

Chapter 9 Maintaining and Emulating Cluster Events

This chapter explains how to configure and test event processing within an HACMP cluster.

Overview

The HACMP for AIX system is event-driven. An event is a change of status within a cluster. When the Cluster Manager detects a change in cluster status, it executes the designated script to handle the event and initiates any user-defined customized processing.

See the chapter on customizing cluster events, in the *HACMP for AIX Installation Guide*, for detailed information on the predefined events and subevents that the Cluster Manager processes. Also, be sure to consult your worksheets and to document any changes you make to your system.

To configure cluster events, you indicate the script that handles the event and any additional processing that should accompany an event, as described below.

Customizing Event Processing

The HACMP for AIX system provides a default script for each pre-defined cluster event and subevent. You can tailor event processing for your environment by adding multiple pre- or post-event processing of your own design.

HACMP versions 4.3.1 and higher also allow you to add, change, and remove customized cluster events.

Pre- and Post-Event Processing

You can specify commands that execute before and after a specific event is generated by the Cluster Manager. For pre-processing, for example, you may want to send a message to specific users, informing them to please stand by because a fallover is occurring. For post-processing, you may want to disable login for a specific group of users if a particular network fails.

Event Notification

You can specify a command that sends mail indicating that an event is about to happen or that an event has just occurred, along with information about the success or failure of the event.

Event Recovery and Retry

You can specify a command that attempts to recover from an event command failure. If the retry count is greater than zero, and the recovery command succeeds, the event script command is rerun.

You can specify the number of times to attempt to execute the recovery command.

Note: HACMP for AIX does not call the notify, recovery, and pre-and post-event processing scripts and commands when the force option of the node_down event is specified.

Configuring Custom Cluster Events

To add, change, or remove customized cluster events, take the following steps.

1. To start system management for HACMP for AIX, enter:

```
smit hacmp
```
2. Select **Cluster Configuration > Cluster Custom Modification > Define Custom Cluster Events**.

SMIT displays the menu choices for adding, changing, or removing a custom event.

Adding Customized Cluster Events

To add a customized event:

1. Select **Add a Custom Cluster Event** from the menu.
2. Enter the field values as follows:

Cluster Event Name	The name can have a maximum of 32 characters.
Cluster Event Description	Enter a short description of the event.
Cluster Event Script Filename	Enter the full pathname of the script to execute.
3. Press Enter to add the information to the local ODM.
4. Synchronize your changes across all cluster nodes by selecting the **Synchronize Cluster Resources** option off the **Cluster Resources** SMIT screen. Press F10 to exit SMIT.

Note: Synchronizing does not propagate the actual new or changed scripts; you must add these to each node manually.

Changing/Showing Custom Cluster Events

To change or show a customized event:

1. Select **Change/Show a Custom Cluster Event** from the menu.
SMIT displays a picklist of available Cluster Event Methods.

2. Select one entry by highlighting it and press Enter.
After you select an event, SMIT displays the Change/Show a Custom Cluster Event screen.

Enter field values as follows:

Cluster Event Name	The current event name is displayed.
New Cluster Event Name	Enter the new name for the event.
Cluster Event Description	Enter a short description of the event.
Cluster Event Script Filename	Enter the full pathname of the script to execute.

Removing Custom Cluster Events

To remove a custom cluster event, take the following steps:

1. Select **Remove a Custom Cluster Event** from the **Define Custom Cluster Events** menu.
SMIT displays a picklist of customized events that exist in HACMPcustom.
2. Choose a single value. Pressing Enter gives an “Are you sure?” pop-up message. After removal SMIT returns an OK and control returns to the **Define Custom Cluster Events** SMIT screen

Changing Pre- or Post-Event Processing

Complete the following steps to change the processing for an event. The changes you can make include pointing the Cluster Manager to a different script to process the event, or using the event customization facility to specify pre- or post- processing event scripts. You only need to complete these steps on a single node and then synchronize the ODM data on the other cluster nodes. The HACMP for AIX system propagates the information to the other nodes.

1. To start system management for HACMP for AIX, enter:

```
smit hacmp
```
2. From the main menu, select **Cluster Configuration > Cluster Resources > Cluster Events > Change/Show Cluster Events**.
SMIT displays the list of events.
3. Select a specific event or subevent that you want to configure and press Enter.
SMIT displays the node name, event name, description, and default event command.
4. Enter field values as follows:

Event Command	Enter the name of the command that processes the event. HACMP for AIX provides a default script. If additional functionality is required, make any changes by adding pre- or post-event processing scripts or commands of your own design, rather than modifying the default scripts or writing new ones.
----------------------	---

Notify Command	<p><i>This field is optional.</i> Enter the full pathname of a user-supplied script to run both before and after a cluster event. This script can notify the system administrator that an event has occurred. The arguments passed to the command are: The event name, one keyword (either start or complete), the exit status of the event (if the keyword was complete), and the same arguments passed to the event command.</p>
Pre-Event Command	<p><i>This field is optional.</i> The field has a picklist if custom events have been defined. You can enter more than one custom event name. Use the F7 key to get an alphabetized list of custom event names, or enter the custom event names in the desired order, delimited by commas. The names must correspond to a custom event name already defined in the HACMPcustom ODM.</p> <p>Enter the full pathname of a user-supplied script to run before the cluster event command executes. This command provides pre-processing before a cluster event occurs. The arguments passed to this command are the event name and the arguments passed to the event command.</p>
Post-Event Command	<p><i>This field is optional.</i> The field has a picklist if custom events have been defined. You can enter more than one custom event name. Use the F7 key to get an alphabetized list of event names, or enter the custom event names in the desired order, delimited by commas. The event names must correspond to a custom event name already defined in the HACMPcustom ODM.</p> <p>Enter the full pathname of the user-supplied script to run after the cluster event command executes. This command provides post-processing after a cluster event. The arguments passed to this command are the event name, event exit status, and the arguments passed to the event command.</p>
Recovery Command	<p><i>This field is optional.</i> Enter the full pathname of a user-supplied script or AIX command to execute to attempt to recover from a cluster event command failure. If the recovery command succeeds and the retry count is greater than zero, the cluster event command is rerun. The arguments passed to this command are the event name and the arguments passed to the event command.</p>
Recovery Counter	<p><i>This field is optional.</i> Enter the number of times to run the recovery command. Set this field to zero if no recovery command is specified, and to at least one if a recovery command is specified.</p>

5. Press Enter to add this information to the ODM on the local node.
6. Synchronize your changes across all cluster nodes by selecting the **Synchronize Cluster Resources** option off the **Cluster Resources** SMIT screen. Press F10 to exit SMIT.

Note: Synchronizing does not propagate the actual new or changed scripts; you must add these to each node manually.

Emulating Cluster Events

HACMP for AIX provides a utility that lets you emulate cluster events by running event scripts that produce output but do not affect the cluster configuration status. This allows you to predict a cluster's reaction to an event as though the event has actually occurred.

The Event Emulator Function

The Event Emulator follows the same procedure used by the Cluster Manager given a particular event, but does not execute any commands that would change the status of the Cluster Manager. For descriptions of cluster events and how the Cluster Manager processes these events, see the chapter on customizing cluster events in the *HACMP for AIX Installation Guide*.

The event emulator runs the events scripts on every active node of a stable cluster. Output from each node is stored in an output file on the node from which you invoked the emulation. You can specify the name and location of the output file using the environment variable `EMUL_OUTPUT` or use the default output file, `/tmp/emuhacmp.out`.

Event Emulator Considerations

Keep the following cautions in mind when using the Event Emulator:

- You can only run one instance of the Event Emulator at a time. If you attempt to start a new emulation in a cluster while an emulation is already running, the integrity of the results cannot be guaranteed. Each emulation is a stand-alone process; one emulation cannot be based on the results of a previous emulation.
- **clinfo** must be running.
- You should add a cluster snapshot before running an emulation, just in case uncontrolled events occur during the emulation. (See Chapter 11, *Saving and Restoring Cluster Configurations* for more information.)
- The Event Emulator can run only event scripts that comply with the currently active configuration. For example:
 - The Emulator expects to see the same environmental arguments used by the Cluster Manager; if you define arbitrary arguments, the event scripts will run, but error reports will result.
 - In the case of `swap_adapter`, you must enter the `ip_label` supplied for service and standby adapters in the correct order, as specified in the usage statement. Both adapters must be located on the same node at emulation time. Both must be configured as part of the same HACMP logical network.

For other events, the same types of restrictions apply. If errors occur during emulation, recheck your configuration to ensure that the cluster state supports the event to be emulated.

- The Event Emulator runs customized scripts (pre- and post-event scripts) associated with an event, but does not execute commands within these scripts. Therefore, if these customized scripts change the cluster configuration when actually run, the outcome may differ from the outcome of an emulation.
- When emulating an event which contains a customized script, the Event Emulator uses the **ksh** flags **-n** and **-v**. The **-n** flag reads commands and checks them for syntax errors, but does not execute them. The **-v** flag indicates verbose mode. When writing customized

scripts that may be accessed during an emulation, be aware that the other **ksh** flags may not be compatible with the **-n** flag and may cause unpredictable results during the emulation. See the **ksh** man page for flag descriptions.

You can run the event emulator through SMIT or the command line. See the **cl_emulate** man page for information on emulating Cluster Events from the command line. Complete the following steps to emulate a cluster event from SMIT.

1. To start system management for HACMP for AIX, enter:

```
smit hacmp
```

2. Select **RAS Support > Event Emulator**.

SMIT displays the list of events. Each option provides a different cluster event to emulate. The following sections provide more information about each option.

Emulating a Node Up Event

To emulate a Node Up event:

1. Select **Node Up Event** from the **Select Event to Emulate** screen, and press enter.
2. Enter the name of the node to use in the emulation in the **Node Name** field.
3. Press Enter to start the emulation.

Emulating a Node Down Event

To emulate a Node Down event:

1. Select **Node Down Event** from the **Select Event to Emulate** screen, and press Enter.
2. Enter field data as follows:

Node Name Enter the node to use in the emulation.

Shutdown Mode Indicate the type of shutdown to emulate:

graceful: Shutdown after the /usr/sbin/cluster/events/node_down_complete script is run on this node to release its resources. The other nodes do not take over the resources of the stopped node.

graceful with takeover: Shutdown after the /usr/sbin/cluster/events/node_down_complete script runs to release its resources. The other nodes do take over the resources of the stopped node.

forced Shutdown immediately. No /usr/sbin/cluster/events/node_down script is run on this node. This node retains control of all its resources.

3. Press Enter to start the emulation.

Emulating a Network Up Event

To emulate a Network Up event:

1. From the **Select Event to Emulate** screen, select **Network Up Event** and press enter.
2. Enter the **Network Name** to use in the emulation.
3. (Optional) Enter the **Node Name**.
4. Press Enter to start the emulation.

Emulating a Network Down Event

To emulate a Network Down event:

1. From the **Select Event to Emulate** screen, select **Network Down Event** and press enter.
2. Enter the **Network Name** to use in the emulation.
3. (Optional) Enter the **Node Name**.
4. Press Enter to start the emulation.

Emulating a Fail Standby Event

To emulate a Fail Standby event:

1. From the **Select Event to Emulate** screen, select **Fail Standby Event** and press enter.
2. Enter field data as follows:

Node Name	Enter the node to use in the emulation.
IP Label	Enter the IP label to use in the emulation.
3. Press Enter to start the emulation.

Note: The following messages are displayed on all active cluster nodes when emulating the Fail Standby and Join Standby events:

```
Adapter $ADDR is no longer available for use as a standby, due to  
either a standby adapter failure or IP address takeover.
```

```
Standby adapter $ADDR is now available.
```

Emulating a Join Standby Event

To emulate a Join Standby event:

1. From the **Select Event to Emulate** screen, select **Join Standby Event** and press enter.
2. Enter field data as follows:

Node Name	Enter the node to use in the emulation.
IP Label	Enter the IP label to use in the emulation.
3. Press Enter to start the emulation.

Note: The following messages are displayed on all active cluster nodes when emulating the Fail Standby and Join Standby events:

```
Adapter $ADDR is no longer available for use as a standby, due to  
either a standby adapter failure or IP address takeover.  
  
Standby adapter $ADDR is now available.
```

Emulating a Swap Adapter Event

To emulate a Swap Adapter event:

1. From the **Select Event to Emulate** screen, select **Swap Adapter Event** and press enter.
2. Enter field data as follows:

Node Name	Enter the node to use in the emulation.
Network Name	Enter the network to use in the emulation.
IP Label (Standby Label used to Swap)	The name of the Standby Adapter to which you want to switch.
IP Label (Service Label to Fail)	The name of the Service Adapter that failed.

3. Press Enter to start the emulation.

Chapter 10 Additional Tasks: Configuring Run-Time Parameters, Cluster Log Files, and Maintaining NFS

This chapter describes how to:

- Set the run-time parameters for a node
- Redirect cluster log files
- Ensure that NFS works properly on an HACMP cluster

Changing Run-Time Parameters

Each cluster node supports two parameters that determine the level of debug information output by the HACMP for AIX scripts and that identify if the node is running Network Information Services (NIS) or name serving. By default, HACMP for AIX sets the debug information parameter to high, which produces detailed output from script execution. If the node is running NIS, HACMP for AIX must disable these services before resolving addresses and re-enable them after name resolution completes.

To change the run-time parameters for a node:

1. Enter:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Resources > Change/Show Run Time Parameters**.

SMIT prompts you to pick a node.

3. Select a node name and press Enter.

SMIT displays the Configure Run Time Parameters screen, showing the node name.

4. Enter field values as follows:

Debug Level

Cluster event scripts have two levels of logging. The low level only logs events and errors encountered while the script executes. The high (default) level logs all commands performed by the script and is strongly recommended. The high level provides the level of script tracing needed to resolve many cluster problems.

Host uses NIS or Name Server

If the cluster uses Network Information Services (NIS) or name serving, set this field to **true**. HACMP for AIX then disables these services before resolving addresses and reenables them after name resolution completes. The default is **false**.

Customizing Cluster Log Files

During normal operation HACMP produces several output log files that you can use to monitor and debug your systems. A cluster log redirection functionality allows you to store a cluster log in a location other than its default directory if you so choose, as well as rename the destination directory.

Should you redirect a log file to a directory of your choice, keep in mind that the requisite (upper limit) disk space for most cluster logs is 2MB. 14MB is recommended for **hacmp.out**. As part of the installation and configuration of the software, you can use SMIT to change the default pathnames for the following log file.

Log Names	Descriptions
cluster.mmdd	Cluster history files generated daily
cm.log	Generated by clstrmgr activity
cspoc.log	Generated by C-SPOC commands
dms_loads.out	Generated by deadman switch activity
emuhacmp.out	Generated by event emulator scripts
hacmp.out	Generated by event scripts and utilities
cluster.log	Generated by cluster scripts and daemons

Note: You cannot change the name of a log file; only its storage destination directory.

For more information on the cluster log redirection functionality, see the HACMP Installation Guide, Chapter 14.

Maintaining NFS

In order for NFS to work as expected on an HACMP cluster, you must be aware of certain configuration issues:

- Creating shared volume groups
- Exporting NFS filesystems
- Mounting issues
- Fallover/takeover issues.

The HACMP for AIX scripts have only minimal NFS support. You may need to modify them to handle your particular configuration. The following sections contain some suggestions for handling a variety of issues.

Creating Shared Volume Groups

When creating shared volume groups, normally you can leave the **Major Number** field blank and let the system provide a default for you. However, if you are using NFS, all nodes in your cluster must have identical major numbers, as HACMP uses the major number as part of the file handle to uniquely identify a Network Filesystem.

In the event of node failure, NFS clients attached to an HACMP cluster operate exactly the way they do when a standard NFS server fails and reboots. If the major numbers are not the same, when another cluster node takes over the filesystem and re-exports the filesystem, the client application will not recover, since the filesystem exported by the node will appear to be different from the one exported by the failed node.

Useful Commands for Handling Volume Groups

To see the major number of existing volume groups, from the `/dev` directory, use the command `ls -l *vg*`. Each line of output looks like the following example. The first number on each line after the owner and group names (“system” is the group name in the example) is the major number of that device.

```
crw-rw---- 1 root system 10, 0 Jun 23 1995 IPL_rootvg
crw-rw---- 1 root system 28, 0 Apr 12 15:15 chowvg
```

To find a free common major number, you can do the command shown above on each node, and compare the numbers. Or, use the command `lvlstmajor` on each node. The output is the next free major number. Compare the numbers returned, then select a common free major number.

To create the volume group `sharedvg` on `hdiskx` and `hdisky` with a major number 24:

```
mkvg -f -n -y sharedvg -s 4 -V 24 hdiskx hdisky
```

To import the volume group `sharedvg` with the major number 24 from `hdiskx`:

```
importvg -y sharedvg -V24 hdiskx
```

To complete this process you must set the `autovaryon = NO`. Do this by running the following command:

```
chvg -a n sharedvg
```

NFS Exporting Filesystems and Directories

The process of NFS-exporting filesystems and directories in HACMP for AIX is different from that in AIX. Remember the following when NFS-exporting in HACMP:

Specifying Filesystems and Directories to NFS Export

While in AIX, you list filesystems and directories to NFS-export in the `/etc/exports` file, in HACMP for AIX, you must put these in a resource group.

Specifying Export Options for NFS Exported Filesystems and Directories

If you want to specify special options in for NFS-exporting in HACMP, you can create a `/usr/sbin/cluster/etc/exports` file. This file has the same format as the regular `/etc/exports` file used in AIX.

Note: Use of this alternate exports file is optional. HACMP checks the `/usr/sbin/cluster/etc/exports` file when NFS-exporting a filesystem or directory. If there is an entry for the filesystem or directory in this file, HACMP will use the options listed. If the filesystem or directory for NFS-export is not listed in the file, or, if the user has not created the `/usr/sbin/cluster/etc/exports` file, the filesystem or directory will be NFS-exported with the default option of root access for all cluster nodes.

Configuring the Optional `/usr/sbin/cluster/etc/exports` File

In this step you add the directories of the shared filesystems to the exports file. Complete the following steps for each filesystem you want to add to the exports file. Refer to your NFS-Exported Filesystem Worksheet.

Note: Remember that this alternate exports file does not specify *what* will be exported, only *how* it will be exported. To specify what to export, you must put it in a resource group.

1. Enter the **smit mknfsexp** fastpath to display the **Add a Directory to Exports List** screen.
2. In the **EXPORT directory now, system restart or both** field, enter **restart**.
3. In the **PATHNAME of alternate Exports file** field, enter `/usr/sbin/cluster/etc/exports`. This step will create the alternate exports file which will list the special NFS export options.
4. Add values for the other fields as appropriate for your site, and press Enter. Use this information to update the `/usr/sbin/cluster/etc/exports` file.
5. Press F3 to return to the **Add a Directory to Exports List** screen, or F10 to exit SMIT.
6. Repeat steps 1 through 4 for each filesystem or directory listed in the **FileSystems/Directories to Export** field on your planning worksheets.

NFS Mounting and Fallover

For HACMP/ES and NFS to work properly together, you must be aware of the following mount issues.

- To NFS mount, a resource group must be configured with IPAT.
- If you want to use the Reliable NFS Server capability that preserves NFS locks and the dupcache in two-node clusters, the IPAT adapter for the resource group must be configured to use Hardware Address Takeover.

Cascading Takeover with Cross Mounted NFS Filesystems

This section describes how to set up cascading resource groups with cross mounted NFS filesystems.

Note: Only cascading resource groups support automatic NFS mounting across servers during fallover. Rotating resource groups do not provide this support. Instead, you must use additional post events or perform NFS mounting using normal AIX routines.

Creating NFS Mount Points on Clients

A mount point is required in order to mount a filesystem via NFS. In a cascading resource group, all the nodes in the resource group will NFS mount the filesystem; thus you must create a mount point on each node in the resource group. On each of these nodes, create a mount point by executing the following command:

```
mkdir /mountpoint
```

where *mountpoint* is the name of the local mountpoint over which the remote filesystem will be mounted.

Setting Up NFS Mount Point Different from Local Mount Point

HACMP handles NFS mounting in cascading resource groups as follows: The node that currently owns the resource group will mount the filesystem over the filesystem's local mount point, and this node will NFS export the filesystem. All the nodes in the resource group (including the current owner of the group) will NFS mount the filesystem over a different mount point. Therefore the owner of the group will have the filesystem mounted twice - once as a local mount, and once as an NFS mount.

Since IPAT is used in resource groups that have NFS mounted filesystems, the nodes will not unmount and remount NFS filesystems in the event of a fallover. When the resource group falls over to a new node, the acquiring node will locally mount the filesystem and NFS-export it. (The NFS mounted filesystem will be temporarily unavailable to cluster nodes during fallover.) As soon as the new node acquires the IPAT label, access to the NFS filesystem is restored.

All applications must reference the filesystem through the NFS mount. If the applications used are dependent upon always referencing the filesystem by the same mount point name, you can change the mount point for the local filesystem mount (for example, change it to mount point_local, and use the previous local mount point as the new NFS mount point.

In the **Change/Show Resources/Attributes for a Resource Group** SMIT screen, the **Filesystem to NFS Mount** field must specify both mount points. Put 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
```

If there are nested mount points, the nfs mount points should be nested in the same manner as the local mount points so that they match up properly. When cross mounting NFS filesystems, you must also set the "*filesystems mounted before IP configured*" field of the Resource Group to **true**.

Server-to-Server NFS Cross Mounting: Example

HACMP/ES allows you to configure a cluster so that servers can NFS-mount each other's filesystems. Configuring cascading resource groups allows the Cluster Manager to decide which node should take over a failed resource, based on priority and node availability.

Ensure that the shared volume groups have the same major number on the server nodes. This allows the clients to re-establish the NFS-mount transparently after the takeover.

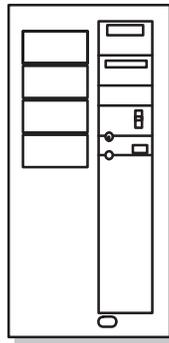
In the example cluster shown below, you have two resource groups, NodeA_rg and NodeB_rg. These resource groups are defined in SMIT as follows:

Resource Group	<i>NodeA_rg</i>
Participating node names	Node A Node B
Filesystems	<i>/afs</i> (filesystems to be locally mounted by node currently owning the resource group)
Filesystems to export	<i>/afs</i> (Filesystem to NFS-export by node currently owning resource group. Filesystem is subset of filesystem listed above.)
Filesystems to NFS mount	<i>/mountpointa;/afs</i> (Filesystems/directories to be NFS-mounted by all nodes in the resource group. First value is NFS mount point; second value is local mount point)
Resource Group	<i>NodeB_rg</i>
Participating node names	Node B Node A
Filesystems	<i>/bfs</i>
Filesystems to export	<i>/bfs</i>
Filesystems to NFS mount	<i>/mountpointb;/bfs</i>

The filesystem you want the local node (Node A) in this resource group to locally mount and export is */afs*, on Node A. You want the remote node (Node B) in this resource group to NFS-mount */afs*, from Node A.

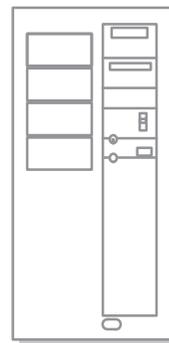
Setting up your cascading resource groups like this ensures the expected default server-to-server NFS behavior described above. On reintegration, */afs* is passed back to Node A, locally mounted and exported. Node B mounts it via NFS again.

When the cluster as originally defined is up and running on both nodes, the filesystems are mounted as shown:



Node A

/afs locally mounted
/afs NFS exported
a_svc:/afs NFS mounted over /mountpointa
b_svc:/bfs NFS mounted over /mountpointb



Node B

/bfs locally mounted
/bfs NFS exported
b_svc:/bfs NFS mounted over /mountpointb
a_svc:/afs NFS mounted over /mountpointa

Cross-Mounted Nodes, Normal Operation

When Node A fails, Node B uses the **cl_nfskill** utility to close open files in Node A:/afs, unmounts it, mounts it locally, and re-exports it to waiting clients.

After takeover, Node B has:

- **/bfs** locally mounted
- **/bfs** NFS-exported
- **/afs** locally mounted
- **/afs** NFS-exported
- **a_svc:/afs** NFS mounted over **/mountpointa**
- **b_svc:/bfs** NFS mounted over **/mountpointb**

See the man page in **/usr/sbin/cluster/events/utlils** for information about the usage and syntax for the **cl_nfskill** command.

Caveats about Node Names and NFS

In the configuration described above the node name is used as the NFS hostname for the mount. This can fail if the node name is not a legitimate TCP/IP adapter label.

To avoid this problem do one of the following:

- Ensure that node name and the service adapter label are the same on each node in the cluster
or
- Alias the node name to the service adapter label in the **/etc/hosts** file.

Reliable NFS Server Capability

An HACMP two-node cluster now takes advantage of AIX extensions to the standard NFS functionality that enable it to handle duplicate requests correctly and restore lock state during NFS server failover and reintegration. This support was previously only available in the HANFS feature. More detail can be found in the **/usr/lpp/cluster/doc/release_notes**.

Note on Reliable NFS Server Functionality

While HACMP is exporting or unexporting filesystems during event processing, the NFS daemons stop temporarily. This affects all NFS mounts on a node, causing the NFS clients to be suspended during the stop, as they cannot reach the server. Any clients holding locks are notified that they must reclaim these once the NFS daemons are restarted.

Chapter 11 Saving and Restoring Cluster Configurations

This chapter explains how to use the cluster snapshot utility to save and restore cluster configurations.

Overview

The cluster snapshot utility allows you to save in a file a record of all the data that defines a particular cluster configuration. This facility gives you the ability to recreate a particular cluster configuration, a process called *applying a snapshot*, provided the cluster is configured with the requisite hardware and software to support the configuration.

Note: You can perform many of the cluster snapshot utility operations, such as saving a configuration and applying a saved configuration, using the HACMP for AIX VSM application (**xhacmpm**). For more information, see the *HACMP for AIX, Administration Guide*, Appendix D, VSM Graphical Configuration Application, and the online help information available with the application.

In addition, a snapshot can provide useful information for troubleshooting cluster problems. Because the snapshots are simple ASCII files that can be sent via e-mail, they can make remote problem determination easier. For information about this cluster snapshot utility, see the *HACMP for AIX Troubleshooting Guide*.

You can also add your own custom snapshot methods to store additional user-specified cluster and system information in your snapshots. The output from these user-defined custom methods is reported along with the conventional snapshot information.

You cannot use the cluster snapshot facility in a cluster concurrently running different versions of HACMP for AIX.

What Information Is Saved in a Cluster Snapshot

The primary information saved in a cluster snapshot is the data stored in the HACMP for AIX ODM classes. This is the information used to recreate the cluster configuration when a cluster snapshot is applied. The cluster snapshot utility saves the following HACMP for AIX ODM classes in the cluster snapshot.

HACMPcluster	Cluster configuration information including cluster ID number, cluster name, names of participating nodes, and information about their IDs.
HACMPnode	Node information including name, and ID number.
HACMPnetwork	Network information including the name, attribute, and cluster ID.

HACMPnim	Network interface information including name, description, and pathname of the module.
HACMPadapter	Adapter information including the type, IP label, and function.
HACMPgroup	Resource group information including name, type and participating nodes.
HACMPresource	Resource group information including the values of the Inactive takeover attribute and the disk fencing attribute.
HACMPserver	Information about application servers.
HACMPevent	Event information including the name, description, and names of pre- and post-processing scripts.
HACMPcommand	Data needed by certain HACMP for AIX commands.
HACMPfence	Settings of the IBM 9333 fence registers. This data, which is device-dependent, is saved but not restored. The objects in this class are regenerated as part of applying a snapshot.
HACMPdaemons	HACMP for AIX daemon start-up and stop parameters which may or may not be node-specific. This node-specific information is preserved during restoration; however, the object class itself (including the object data) is the same on all nodes. This class also contains the Cluster Lock Manager Resource Allocation parameters.
HACMPsp2	Data HACMP for AIX requires to support the IBM Scalable POWERparallel (SP) system. This information is device-specific and is not used during a snapshot restoration.
HACMPcustom	Custom verification, custom event, and custom snapshot method information: method name, full pathname to the method, method description and type definition.

The cluster snapshot does not save any user-customized scripts, applications, or other non-HACMP configuration parameters. For example, the name of an application server and the location of its start and stop scripts are stored in the HACMPserver ODM object class. However, the scripts themselves as well as any applications they may call are not saved.

The cluster snapshot also does not save any device- or configuration-specific data which is outside the scope of HACMP for AIX. For instance, the facility saves the names of shared filesystems and volume groups; however, other details, such as NFS options or LVM mirroring configuration are not saved.

Note: Be aware that sticky location markers specified during earlier dynamic reconfigurations may be present in the snapshot. For information on locating and removing these markers while the cluster is down, see *Removing Sticky Markers* on page 7-26 in Chapter 7, *Changing Resources and Resource Groups*.

Format of a Cluster Snapshot

The cluster snapshot utility stores the data it saves in two separate files

ODM Data File	This file contains all the data stored in the HACMP for AIX ODM object classes for the cluster. This file is given a user-defined basename with the .odm file extension. Because the ODM information must be largely the same on every cluster node, the cluster snapshot saves the values from only one node.
Cluster State Information File	This file contains the output from standard AIX and HACMP for AIX system management commands. This file is given the same user-defined basename with the .info file extension. Output from any custom snapshot methods is appended to the .info file.

The following describes the contents of the **.odm** file. For more information about the **.info** file, see the *HACMP for AIX Troubleshooting Guide*.

Cluster Snapshot ODM Data File

The cluster snapshot ODM data file is an ASCII text file divided into three delimited sections:

Version section	This section identifies the version of the cluster snapshot. The characters <VER identify the start of this section; the characters </VER identify the end of this section. The version number is set by the cluster snapshot software.
Description section	This section contains user-defined text that describes the cluster snapshot. You can specify up to 255 characters of descriptive text. The characters <DSC identify the start of this section; the characters </DSC identify the end of this section.
ODM data section	This section contains the HACMP for AIX ODM object classes in generic AIX ODM stanza format. The characters <ODM identify the start of this section; the characters </ODM identify the end of this section.

The following is an excerpt from a sample cluster snapshot ODM data file showing the various ODM stanzas that are saved.

```
<VER
1.0
</VER

<DSC
My Cluster Snapshot
</DSC
<ODM

HACMPcluster:
    id = 97531
    name = "Breezel"
    nodename = "mynode"

HACMPnode:
    name = "mynode"
    object = "VERBOSE_LOGGING"
    value = "high"
.
.
.
</ODM
```

clconvert_snapshot Utility

You can run **clconvert_snapshot** to upgrade cluster snapshots from previous versions of HACMP to the most recent version. The **clconvert_snapshot** is not run automatically during installation, and must always be run from the command line. Each time you run the **clconvert_snapshot** command, conversion progress is logged to the **/tmp/clconvert.log** file.

Note: Root user privilege is required to run **clconvert_snapshot**. You must know the HACMP version from which you are converting in order to run this utility.

For more information on the **clconvert_snapshot** utility, refer to *HACMP for AIX Administration Guide*, Appendix A, HACMP for AIX Common Commands on page A-5, or to the **clconvert_snapshot** man page.

Defining a Custom Snapshot Method

If you want additional, customized system and cluster information to be appended to the **.info** file, you should define custom snapshot methods to be executed when you create your cluster snapshot.

To define a custom snapshot method, perform the following steps. Notice that there are two paths to the correct screen.

1. To reach the main HACMP SMIT menu, enter:

```
smit hacmp
```
2. Select **Cluster Configuration**.

3. From Cluster Configuration, select *either* **Cluster Custom Modification** or **Cluster Snapshot**.
4. Select **Define Custom Snapshot Method > Add a Custom Snapshot Method**.
5. Enter information as follows:

Custom Snapshot Method Name Enter a name for this custom snapshot method.

Custom Snapshot Method Description Add any descriptive information about the custom method.

Custom Snapshot Script Filename Add the full pathname to the custom snapshot scriptfile.

Once you have defined one or more custom snapshot methods, when you create a cluster snapshot you are asked to specify which custom method(s) you wish to be executed in addition to the conventional snapshot.

Changing or Removing a Custom Snapshot Method

After you have defined a custom snapshot method, you can change or delete it using the other menu items in the Define Custom Snapshot Method SMIT screen: Change/Show a Custom Snapshot Method and Remove a Custom Snapshot Method.

When you select one of these menus, a picklist of existing custom snapshot methods appears. Choose the one you wish to change or remove and fill in the appropriate fields, or answer the prompt to confirm deletion.

Creating (Adding) a Cluster Snapshot Using SMIT

You can initiate cluster snapshot creation from any cluster node. You can create a cluster snapshot on a running cluster and multiple snapshots may be created. The cluster snapshot facility retrieves information from each node in the cluster. Accessibility to all nodes is required. Because of the large amount of data which must be retrieved when creating the cluster snapshot, the time and memory consumed may be substantial, especially when the number of cluster nodes is high. Cluster snapshot files typically require approximately 10 Kb per node.

Note: To get an accurate snapshot of a system that has been configured with Kerberos security, you must set up *all* Kerberos service principals before taking the snapshot. For more information about configuring your system with Kerberos, see the *HACMP for AIX Installation Guide*.

To create a cluster snapshot:

1. To view the main HACMP for AIX SMIT menu, enter:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Snapshots > Add a Cluster Snapshot**.
3. Enter field values as follows:

Cluster Snapshot Name	Enter the name you want for the basename for the cluster snapshot files. The default directory path for storage and retrieval of the snapshot is <code>/usr/sbin/cluster/snapshots</code> . You can specify an alternate path using the <code>SNAPSHOTPATH</code> environment variable.
Custom Defined Snapshot Methods	Specify one or more custom snapshot methods to be executed if desired; press F4 for a picklist of custom methods on this node. If you select All , the custom methods will be executed in alphabetical order on each node.
Cluster Snapshot Description?	Enter any descriptive text you want inserted into the cluster snapshot. You can specify any text string up to 255 characters in length.

Applying a Cluster Snapshot Using SMIT

Applying a cluster snapshot overwrites the data in the existing HACMP for AIX ODM classes on all nodes in the cluster with the new ODM data contained in the snapshot. You can apply a cluster snapshot from any cluster node.

If cluster services are inactive on all cluster nodes, applying the snapshot changes the ODM data stored in the system default configuration directory (DCD). If cluster services are active on the local node, applying a snapshot triggers a cluster-wide dynamic reconfiguration event. In dynamic reconfiguration, in addition to synchronizing the ODM data stored in the DCDs on each node, HACMP for AIX replaces the current configuration data stored in the active configuration directory (ACD) with the changed configuration data in the DCD. The snapshot becomes the currently active configuration. For more information about dynamic reconfiguration of a cluster, see Chapter 6, Changing the Cluster Topology.

Note: A cluster snapshot used for dynamic reconfiguration may contain changes to either the cluster topology OR to cluster resources, but not both. You cannot change both the cluster topology and cluster resources in a single dynamic reconfiguration event.

Note: Applying a cluster snapshot may affect both AIX and HACMP for AIX ODM objects and system files as well as user-defined files.

To apply a cluster snapshot using SMIT, perform the following steps.

1. To view the main HACMP for AIX SMIT menu, enter:

```
smit hacmp
```
2. Select **Cluster Configuration > Cluster Snapshots > Apply a Cluster Snapshot**.

SMIT displays the Cluster Snapshots screen containing a list of all the cluster snapshots that exist in the directory specified by the `SNAPSHOTPATH` environment variable.

3. Select the cluster snapshot that you want to apply and press Enter.

SMIT displays the **Apply a Cluster Snapshot** screen.

4. Enter field values as follows:

Cluster Snapshot Name	Displays the current basename of the cluster snapshot.
Cluster Snapshot Description?	Displays the text stored in the description section of the snapshot files.
Un/Configure Cluster Resources?	<p>If you set this field to yes, HACMP for AIX changes the definition of the resource in the ODM and it performs any configuration triggered by the resource change. For example, if you remove a filesystem, HACMP for AIX removes the filesystem from the ODM and also unmounts the filesystem.</p> <p>By default, this field is set to yes.</p> <p>If you set this field to no, HACMP for AIX changes the definition of the resource in the ODM but does not perform any configuration processing that the change may require. For example, a filesystem would be removed from the HACMP for AIX cluster definition but would not be unmounted. This processing is left to be performed by HACMP for AIX during a fallover.</p> <p>HACMP for AIX attempts to limit the impact on the resource group when a component resource is changed. For example, if you add a filesystem to the resource group that already includes the underlying volume group as a resource, HACMP for AIX does not require any processing of the volume group. Other modifications made to the contents of a resource group may cause the entire resource group to be unconfigured and reconfigured during the dynamic reconfiguration. Cluster clients will experience an interruption in related services while the dynamic reconfiguration is in progress.</p>
Force apply if verify fails?	<p>If this field is set to no, synchronization aborts if verification of the new configuration fails. As part of dynamic reconfiguration processing, the new configuration is verified before it is made the active configuration.</p> <p>If you want synchronization to proceed even if verification fails, set this value to yes.</p> <p>By default, this field is set to no.</p>

Note: In some cases, the verification uncovers errors that do not cause the synchronization to fail. HACMP reports the errors in the SMIT command status window so that you are aware of an area of the configuration that may be a problem.

You should investigate any error reports, even when they don't interfere with the synchronization.

Undoing an Applied Snapshot

Before the new configuration is applied, the cluster snapshot facility saves the current configuration in a snapshot called `~snapshot.n.odm`, where `n` is either 1, 2, or 3. The saved snapshots are cycled so that only three generations of snapshots exist. If the apply process fails, you can re-apply the previous configuration. These saved snapshots are stored in the directory specified by the `SNAPSHOTPATH` environment variable.

Changing a Cluster Snapshot

After creating a cluster snapshot, you can change the basename assigned to cluster snapshot files and the description contained in these files. Note that you must use the SMIT interface to perform this task.

To change a cluster snapshot, perform the following steps.

1. To view the main HACMP for AIX SMIT menu, enter:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Snapshots > Change/Show a Cluster Snapshot**.

SMIT generates and displays a list of all the cluster snapshots that exist in the directory specified by `SNAPSHOTPATH`.

3. Select the cluster snapshot that you want to change and press Enter.
4. Enter field values as follows:

Cluster Snapshot Name	Displays the current basename of the cluster snapshot.
New Cluster Snapshot Name	Enter the new name you want assigned as the basename of the cluster snapshot files.
Cluster Snapshot Description?	SMIT displays the current description. You can edit the text using up to 255 characters.

Removing a Cluster Snapshot

Removing a cluster snapshot deletes both of the ASCII files that define the snapshot from the snapshots directory. (The directory in which the snapshots are stored is defined in the SNAPSHOTPATH environment variable.) You must use SMIT to remove a cluster snapshot.

To remove a cluster snapshot using the SMIT interface, perform the following steps.

1. To view the main HACMP for AIX SMIT menu, enter:

```
smit hacmp
```

2. Select **Cluster Configuration > Cluster Snapshots > Remove a Cluster Snapshot**.

SMIT generates and displays a list of all the cluster snapshots that exist in the directory specified by the SNAPSHOTPATH environment variable.

3. Select the cluster snapshot that you want to remove and press Enter.

The cluster snapshot facility deletes the files in the snapshot directory that are associated with that snapshot.

Saving and Restoring Cluster Configurations

Removing a Cluster Snapshot

Chapter 12 Managing Users and Groups in a Cluster

This chapter explains how to use the C-SPOC utility to manage user accounts and groups on all nodes in a cluster by executing a C-SPOC command on a single node.

Overview

One of the basic tasks any system administrator must perform is setting up user accounts and groups. All users require accounts to gain access to the system. Every user account must belong to a group. Groups provide an additional level of security and allow system administrators to manipulate a group of users as a single entity.

For users of an HACMP for AIX cluster, system administrators must create duplicate accounts on each cluster node. The user account information stored in the `/etc/passwd` file and in other files stored in the `/etc/security` directory should be consistent on all cluster nodes. For example, if a cluster 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.

System administrators typically keep user accounts synchronized across cluster nodes by copying the key system account and security files to all cluster nodes whenever a new account is created or an existing account is changed. For C-SPOC clusters, the C-SPOC utility simplifies the cluster-wide synchronization of user accounts by propagating the new account or changes to an existing account across all cluster nodes automatically.

The following sections describe how to perform many common administrative tasks involved with managing user accounts and groups in a cluster using the C-SPOC utility.

Managing User Accounts

As system administrator of an HACMP for AIX cluster, you may be called upon to perform the following tasks:

- Listing all user accounts on all cluster nodes
- Adding users to all cluster nodes
- Changing characteristics of a user account on all cluster nodes
- Removing a user account from all cluster nodes.

The following sections describe how to accomplish these tasks on all nodes in a cluster using the C-SPOC utility.

Listing Users On All Cluster Nodes

To obtain information about all user accounts on cluster nodes (or about a particular user account), use the C-SPOC `cl_lsuser` command or the C-SPOC SMIT **List all the Users on the Cluster** screen. The `cl_lsuser` command executes the AIX `lsuser` command on each node. To obtain a listing of all user accounts in the cluster, you must specify the **ALL** argument.

If you specify a user name that does not exist on one of the cluster nodes, the **cl_lsuser** command outputs a warning message but continues execution of the command on other cluster nodes. For more information about the **cl_lsuser** command, see its man page.

Warning: If you have a Network Information Service (NIS) database installed on any cluster node, some user information may not appear when you use the **cl_lsuser** command.

To list all user accounts on all cluster nodes using the C-SPOC utility:

1. Enter:

```
smit cl_admin (or cl_lsuser)
```

2. Select **Cluster Users & Groups > Users > List All Users in the Cluster**

SMIT executes the **cl_lsuser** command and displays a listing of user accounts.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**, to obtain the status of the command on each cluster node.

Adding User Accounts on all Cluster Nodes

Adding a user to the cluster involves three steps:

- Add an entry for the new user to the **/etc/passwd** file and other system security files
- Create a home directory for the new user
- Add the user to a group file

On AIX systems, you use the **mkuser** command to perform these tasks. This command adds entries for the new user to various system security files, including **/etc/passwd** and **/etc/security/passwd**, adds the new user to a group, and creates a home directory for the new user. Every user account has a number of attributes associated with it. When you create a user, the **mkuser** command fills in values for these attributes from the system default file **/usr/lib/security/mkuser.default** file. You can override these default values by specifying an attribute and a value on the **mkuser** command line.

To add a user on one or more nodes in a C-SPOC cluster, use the C-SPOC **cl_mkuser** command or the Add a User to the Cluster SMIT screen. The **cl_mkuser** command calls the AIX **mkuser** command to create the user account on each cluster node you specify. The **cl_mkuser** command creates a home directory for the new account on each cluster node.

If a user with the same name already exists on one of the cluster nodes, the operation fails, returning the message “*user-name* already exists on node *node-name*.” Optionally, you can specify that the **cl_mkuser** command continue processing even if the user name already exists on one of the cluster nodes by specifying the C-SPOC **-f** flag. For more information, see the **cl_mkuser** command man page.

Note: User accounts do not become active until the root user assigns a password for the account on each node, using the **/etc/passwd** command. The **cl_mkuser** command, like the AIX **mkuser** command, does not create passwords. Until the initial password is set with the **passwd** command or the **pwdadm** command, the password field in the **/etc/passwd** file is set to * (an asterisk) to indicate that there is no valid password. The new user account is disabled until authentication information is added to the **/etc/security/passwd** file.

To add a user to all nodes in a cluster using the C-SPOC utility, perform the following procedure on any cluster node.

1. Enter the SMIT fastpath:

```
smit cl_mkuser
```

2. Press F4 to generate a resource group picklist.
3. Select a resource group and press Enter.
4. Specify a user name and other attributes as needed.

Enter data in the entry fields to set up the account. AIX provides help screens that describe each attribute. The **User Name** field is the only required field.

Note: You should specify a value in the **User ID** field so that the account's user ID will be the same on all cluster nodes. If you do not specify this value, AIX could assign different user IDs on each node. A mismatch of user IDs for an account could prevent a user from logging on to another cluster node in the event of a fallover.

5. After entering in user data, press Enter. The **cl_mkuser** command executes, creating the user account specified on both cluster nodes.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file contains the status of the **cl_mkuser** command on each cluster node.

Changing Attributes of Users in a Cluster

On AIX systems, you can change any of the attributes associated with an existing user account by using the **chuser** command. Using the **chuser** command you specify the name of the user account you want to change and then specify the attributes with their new values. If you use the SMIT Change User Attributes screen, the complete list of user attributes is displayed and you can supply new values for any attributes. The **chuser** command modifies the user information stored in the **/etc/passwd** file and the files in the **/etc/security** directory.

To change the attributes of a user account on one or more cluster nodes, you use the C-SPOC **cl_chuser** command or the C-SPOC Change User Attributes SMIT screen. The **cl_chuser** command executes the AIX **chuser** command on each cluster node.

Warning: Do not use the **cl_chuser** command if you have a Network Information Service (NIS) database installed on any node in your cluster.

Both cluster nodes must be active and a user with the specified name must exist on both the nodes for the change operation to proceed. Optionally, you can specify that **cl_chuser** command continue processing if the specified user name exists on any of the cluster nodes. See the **cl_chuser** command man page for more information.

To change the characteristics of a user account on all cluster nodes using the C-SPOC utility:

1. Enter the SMIT fastpath:

```
smit cl_chuser
```

SMIT displays the Change/Show Characteristics of a User in the Cluster screen.

2. Specify the name of the user account you want to change and press Enter. Press F4 to obtain a listing of users from which to choose. SMIT displays a complete listing of the user account attributes with their current values filled in.
3. Enter the new values for attributes you want to change and press Enter. AIX provides help screens that explain each attribute. SMIT executes the C-SPOC **cl_chuser** command to change the attributes of the user account on all cluster nodes.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file contains the status of the command on each cluster node.

Removing Users from a Cluster

On AIX systems, you remove a user account by using the **rmuser** command or the SMIT Remove a User From the System screen. Using the **rmuser** command you specify the name of the user account you want to remove and specify whether you want the user password and other authentication information removed from the **/etc/security/passwd** file.

To remove a user account from one or more cluster nodes, use the C-SPOC **cl_rmuser** command or the C-SPOC Remove a User from the Cluster SMIT screen. The **cl_rmuser** command executes the AIX **rmuser** command on all cluster nodes.

Note: The system removes the user account but does not remove the home directory or any files owned by the user. These files are only accessible to users with root authority or by the group in which the user was a member.

To remove a user from all cluster nodes using the C-SPOC utility:

1. Enter the following SMIT fastpath:

```
smit cl_rmuser
```

2. Enter field data as follows:

User Name	This is the only required field. You must enter a user name for the new account. The name can be up to 8 characters long.
Remove Authentication information?	If you specify yes , password and other authentication information is deleted from system security files.

3. After entering in user data, press Enter.

SMIT removes the user account specified from both nodes.

If the command fails, check the C-SPOC log file, `/tmp/cspoc.log`. This file contains the status of the command's execution on each cluster node.

Managing Group Accounts

All users must belong to a group. Groups add a level of security. As the system administrator of an HACMP for AIX cluster, you may be called upon to perform the following tasks:

- Listing all groups defined on all cluster nodes
- Adding groups to all cluster nodes
- Changing the characteristics of a group on all cluster nodes
- Removing a group from all cluster nodes.

Using the C-SPOC facility, you can perform these tasks cluster-wide from any node in a cluster.

Listing Groups on All Cluster Nodes

To obtain information about all the groups defined on an AIX system, or about a particular group, you use the **lsgroup** command. Each group has associated attributes that include the names of the users in the group, the user name of the administrator of the group, and the group ID.

To obtain information about the groups defined on all cluster nodes, use the C-SPOC **cl_lsgroup** command, specifying the **ALL** argument, or by using the C-SPOC SMIT List all the Groups on the Cluster screen. The **cl_lsgroup** command executes the **lsgroup** command on each cluster node. The output from the **lsgroup** command for all nodes is displayed on the node on which the command was executed.

If you specify a group name that does not exist on a cluster node, the **cl_lsgroup** command outputs a warning message but continues execution of the command on all other cluster nodes. For more information about the **cl_lsgroup** command, see its man page.

Note: If you have a Network Information Service (NIS) database installed on any cluster node, some user information may not appear when you use the **cl_lsgroup** command.

To list all the groups defined on each cluster node using the C-SPOC utility's SMIT interface:

1. Enter:

```
smit cl_admin
```

(You can also use the `cl_lsgroup` SMIT fastpath to bypass the initial screens in the SMIT hierarchy.)

2. From the main HACMP menu, select **Cluster Users & Groups > Groups > List all Groups on the Cluster**.

SMIT displays the command status window.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file contains the status of the command execution on each cluster node.

Adding Groups on Cluster Nodes

To define a new group on AIX systems, you use the **mkgroup** command. This command adds an entry for the new group to various system security files, including **/etc/group** and **/etc/security/group**. Every group has a number of attributes associated with it. When you create a group, you must specify the name of the group. You can optionally specify values for other group attributes.

To define a new group on all cluster nodes, use the C-SPOC **cl_mkgroup** command or the C-SPOC **Add a Group to the Cluster** SMIT screen. The **cl_mkgroup** command performs some verification and then calls the AIX **mkgroup** command on each cluster node to create the group you specify.

If a group with the same name already exists on a node, the operation is aborted. By default, the **cl_mkgroup** command requires that all nodes in the HACMP cluster must be powered up and accessible over the network; otherwise, the **cl_mkgroup** command fails with an error. Optionally, if you specify the C-SPOC **-f** flag, the **cl_mkgroup** command continues processing even if it encounters these errors on one of the cluster nodes. See the **cl_mkgroup** command man page for more information.

To define a new group on cluster nodes using the C-SPOC utility:

1. Enter the SMIT fastpath:

```
smit cl_mkgroup
```

Enter data in entry fields to create the group account. The **Group Name** is the only required field. Note, however, that you should specify the **Group ID** field as well.

2. After you finish filling in the SMIT fields, press Enter. The C-SPOC **cl_mkgroup** command executes, creating the new group on all cluster nodes.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file provides more information about the execution of the command on each cluster node.

Changing Characteristics of Groups in a Cluster

To change the attributes of a group on an AIX system, you use the **chgroup** command. Using the **chgroup** command you specify the name of the group you want to change and the attributes with their new values. If you use the SMIT Change Group Attributes screen, SMIT displays the complete list of attributes associated with a group which you can modify. The **chgroup** command modifies the user information stored in the **/etc/group** and the **/etc/security/group** files.

To change the attributes of a group on all cluster nodes, use the C-SPOC **cl_chgroup** command or the C-SPOC Change Group Attributes SMIT screen. The **cl_chgroup** command executes the AIX **chgroup** command on each cluster node.

Warning: Do not use the **cl_chgroup** command if you have a Network Information Service (NIS) database installed on either node in your cluster. Using the command in this environment could cause serious system database inconsistencies.

For the C-SPOC command to succeed, all cluster nodes must be accessible and a group with the specified name must exist on all cluster nodes. Optionally, if you specify the C-SPOC **-f** flag, the **cl_chgroup** command continues processing even if it encounters an error on one of the cluster nodes. See the **cl_chgroup** command man page for more information.

To change the attributes of a group on all cluster nodes using the C-SPOC utility:

1. Enter the SMIT fastpath:

```
smit cl_chgroup
```

2. Specify the name of the group you want to change and press Enter. Press F4 to obtain a listing of groups from which to choose. SMIT displays a complete listing of the attributes of the group specified, with their current values filled in.
3. Change the value of any group attribute and press Enter. The **cl_chgroup** command executes, writing the new attribute value in the appropriate system security files on all cluster nodes.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file contains the status of the command execution on each cluster node.

Removing Groups from the Cluster

To delete a group on an AIX system, you use the **rmgroup** command. This command removes the entry for the group from the **/etc/group** and **/etc/security/group** files. Users that are members of the group are not deleted.

If the group is the primary group for any user, the remove operation fails unless you redefine the user's primary group with the **chuser** command. (For more information about using the **chuser** command, see Changing Attributes of Users in a Cluster on page 12-3.) Only the root user can remove an administrative group or a group with administrative users as members.

To remove a group from any number of cluster nodes, use the C-SPOC **cl_rmgroup** command or the C-SPOC Remove a Group from the Cluster SMIT screen. The **cl_rmgroup** command performs some cluster-wide verification checks and then calls the AIX **rmgroup** command to remove the group on each node.

If a group with the name specified does not exist on one of the cluster nodes, the **cl_rmgroup** command reports a warning message but continues the operation on the other cluster node. By default, the **cl_rmgroup** command requires that all cluster nodes be powered up and accessible over the network; otherwise, the **cl_rmgroup** command fails with an error. Optionally, if you specify the C-SPOC **-f** flag, the **cl_rmgroup** command continues processing even if it encounters an error on one of the cluster nodes. See the **cl_rmgroup** command man page for more information.

Managing Users and Groups in a Cluster

Managing Group Accounts

To remove a group from cluster nodes using the C-SPOC utility:

1. Enter:

```
smit cl_rmgroup
```

2. Enter the name of the group you want to remove. Press the F4 key to get a listing of available groups from which to choose. After specifying the group name, press Enter. The **cl_rmgroup** command executes, removing the group from all cluster nodes.

If the command fails, check the C-SPOC log file, **/tmp/cspoc.log**. This file contains the status of the command execution on each cluster node.

Appendix A HACMP for AIX Commands

This appendix provides a quick reference to commands commonly used to obtain information about the cluster environment or to execute a specific function. The chapter lists syntax diagrams and provides examples for using each command.

Overview of Contents

As system administrator, you often must obtain information about your cluster to determine if it is operating correctly. The commands you need to obtain this information are listed in alphabetical order in this chapter.

Highlighting

The following highlighting conventions are used in this appendix:

Bold	Identifies command words, keywords, files, directories, and other items whose actual names are predefined by the system.
<i>Italics</i>	Identifies parameters whose actual names or values are supplied by the user.
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 you should actually type.

Reading Syntax Diagrams

Usually, a command follows this syntax:

[]	Material within brackets is optional.
{ }	Material within braces is required.
	Indicates an alternative. Only one of the options can be chosen.
...	Indicates that one or more of the kinds of parameters or objects preceding the ellipsis can be entered.

Note: Flags listed in syntax diagrams throughout this appendix are those recommended for use with the HACMP for AIX software. Flags used internally by SMIT are not listed.

Related Information

For complete information on a command's capabilities and restrictions, see the online man page and the relevant chapter in this guide. Man pages for HACMP for AIX, Version 4.4 commands and utilities are installed in the `/usr/share/man/cat1` directory. Use the following syntax to read man page information:

```
man [command-name]
```

where *command-name* is the actual name of the HACMP command or script. For example, type **man clstart** to obtain information about the HACMP cluster startup command.

HACMP for AIX Common Commands

The following commands can be run from the command line to obtain information about your HACMP for AIX cluster environment. **For complete syntax and descriptions for these commands, see the section beginning on HACMP for AIX Common Commands on page A-2.**

cl_convert	Converts ODM of previous HACMP release to ODM of current release. Run from the command line only if installation fails. Otherwise, it runs automatically with installation.
clconvert_snapshot	Upgrades cluster snapshots.
cldiag debug	Enables Cluster Manager and lock resource debugging. - logs – Allows selected viewing of script or syslog output - vgs – Checks for consistencies of volume groups - error – Allows parsing of system error log - trace – Allows tracing HACMP for AIX daemons
clfindres	Finds a given resource group in the cluster configuration.
clgetaddr	Returns IP address for the specified node name.
clgetgrp	Retrieves and displays HACMPgroup class objects by format.
clgetif	Prints name/netmask of IP address/label.
clgodmget	Lists ODM class remotely.
cllscf	Lists cluster topology information.
cllslstr	Shows cluster name and ID.
cllsdisk	Lists PVIDs of accessible disks in a specified resource chain.
cllsfs	List filesystems accessible in a specified resource chain.
cllslv	List the names of filesystems accessible by nodes in a specified resource chain.
cllsgrp	Lists all resource groups.

cclsif	Shows adapter information.
cclsnim	Lists contents of HACMPnetwork interface module ODM class.
cclsnode	Shows node information.
cclsnw	Shows network information.
cclsparam	Lists run-time parameters.
cclsrres	Lists ODM resource data by name and arguments.
cclsserv	Lists application servers by name.
cclsvg	List volume groups accessible in a specified resource chain.
clnodename	Lists, adds, changes, or removes node name.
clshowres	Shows node environment resources.
clstat	Monitors status of cluster.
clverify cluster	Verifies cluster configuration.
clverify software	Verifies HACMP-specific modifications to AIX system files.
get_local_nodename	Retrieves the name of the local node.
clgetactivenodes	Retrieves the names of all active cluster nodes.
clresactive	Retrieves the names of all active resources.
clm_stats	Queries the Cluster Lock Manager for status information.

HACMP for AIX C-SPOC Commands

The following C-SPOC commands function in cluster environments and can be run from the command line to manage the cluster. **For complete syntax and description for these commands, see the section beginning on page A-23.**

cl_chfs	Changes shared filesystem attributes.
cl_chgroup	Changes group attributes on all cluster nodes.
cl_chlv	Changes a shared logical volume's characteristics.
cl_chuser	Changes user attributes for a user across cluster nodes.
cl_crfs	Creates a shared filesystem on a logical volume already in the cluster.
cl_crlvfs	Creates a logical volume and a shared filesystem on a shared volume group already in the cluster.

cl_clstop	Stops cluster services on nodes running C-SPOC.
cl_extendvg	Allows a physical volume to be added to a cluster volume group.
cl_importvg	Imports a volume group to cluster node.
cl_lsfs	Displays shared filesystem attributes for all cluster nodes.
cl_lsgroup	Displays group attributes for all cluster nodes.
cl_lslv	Displays shared logical volume attributes for cluster nodes.
cl_lsuser	Displays user account attributes for all nodes.
cl_lsvg	Displays shared volume group attributes for cluster nodes.
cl_mirrorvg	Allows logical volumes in a cluster volume group to be mirrored.
cl_mkdev	Configures a disk device on specified cluster nodes.
cl_mkvg	Creates a shared (or concurrent, using -c) volume group on specified cluster nodes.
cl_mkgroup	Creates a new group on all cluster nodes.
cl_mklv	Creates a logical volume.
cl_mklvcopy	Creates a copy of a logical volume.
cl_mkuser	Creates a new user account on all cluster nodes.
cl_nodcmd	Runs a given command in parallel on a given set of nodes.
cl_reducevg	Removes a physical volume from a volume group.
cl_rc.cluster	Sets the operating environment and starts cluster daemons.
cl_rmdev	Removes a previously configured disk from the cluster.
cl_rmfs	Removes a shared filesystem from all cluster nodes.
cl_rmggroup	Removes a group from all cluster nodes.
cl_rmlv	Removes a logical volume from cluster nodes.
cl_rmlvcopy	Removes a copy of a logical volume from cluster nodes.
cl_rmuser	Removes a user account from cluster nodes.
cl_syncvg	Synchronizes any non-current logical volume copies.
cl_unmirror	Removes the mirrors that exist in a volume group.
cl_updatevg	Exports/imports the shared volume group definition.

AIX Connections Commands

If you have AIX Connections configured, the following commands may apply. **For complete syntax and descriptions of these commands, see the section beginning on page A-60.**

cllsres	Lists the NAME=VALUE pairs when doing a colon list
clshowres	Lists the NAME=VALUE pairs
clsetenvres	Initializes the AIX_CONNECTIONS_SERVICES NAME=VALUE pairs
cllsaconnrs	Retrieves and lists AIX Connections realms and services
clisaconn	Sets the exit condition
cllsaconngrs	Retrieves a list of AIX Connections realm/service pairs accessible to all nodes in a resource group

CS/AIX Communication Links Commands

If you have highly available CS/AIX communications links configured, the following commands may apply. **For complete syntax and descriptions of these commands, see the section beginning on page A-61 .**

clshowres	Lists resources for a resource group
cllssnaobj	List CS/AIX resources

HACMP for AIX Common Commands

cl_convert [-C | -N | -E] -v [release] [-F] [-s<simulation file>] [-i]

Upgrading HACMP software to the newest version involves converting the ODM from a previous release to that of the current release. When you install HACMP, **cl_convert** is run automatically. However, if installation fails, you must run **cl_convert** from the command line. Root user privilege is required to run **cl_convert**.

If **-C**, **-N** or **-E** is not provided, **cl_convert** will assume that the previously installed product (HACMP for AIX, HACMP ES or HANFS) is the same as that which is being installed.

-C	Classic Conversion Flag. This flag specifies conversion from HACMP “classic”. Cannot be used with -N or -E flag.
-N	HANFS Conversion Flag. Specifies conversion from HANFS. Cannot be used with -C flag or -E flag.
-E	HACMP ES Conversion Flag. Specifies conversion from HACMP ES. Cannot be used with -C or -N flag.

- v** Release Version Flag. Specifies the source version of HACMP. (“Source version” is the previously installed version.)
- F** Force Flag. Forces an unconditional overwrite of the existing configuration.
- s** Simulation Flag. Specifies the file to which the ODMs should be output. If this flag is provided, only the file is created and the actual ODMs are not updated.
- i** Ignore Copy Flag. Should not be used from the command line. Used by HACMP to copy the old ODM stanzas to the new ODM stanzas.

Note: The AIX environmental variable ODMDIR must be set to the directory you wish to convert.

Examples

Example 1

```
cl_convert -v 4.3.1 -F
```

Converts ODM from HACMP 4.3.1 to HACMP 4.4 (assuming that HACMP 4.4 is currently installed on the node).

Example 2

```
cl_convert -C -v 4.4 -s /tmp/convert.out -F
```

Converts ODM from HACMP 4.4 to HACMP ES 4.4 (assuming HACMP/ES is currently installed on the node) and outputs the results to a file named **/tmp/convert.out**. The actual ODMs are not changed. (The **-s** flag prevents the actual ODM in **/etc/objrepos** from being changed.)

clconvert_snapshot [-C | -N | -E] -v [release] -s <cluster snapshot file>

You can run **clconvert_snapshot** to upgrade cluster snapshots from previous versions of HACMP to the most recent version of HACMP for AIX Enhanced Scalability. The **clconvert_snapshot** is not run automatically during installation, and must always be run from the command line.

If **-C**, **-N** or **-E** is not provided, **clconvert_snapshot** will assume that the previously installed product is the same as that which is being installed.

- C** Classic Conversion Flag. This flag specifies conversion from HACMP “classic”. Cannot be used with **-N** or **-E** flag.
- N** Currently not used.
- E** Currently not used.
- v** Release Version Flag. Specifies the source version of HACMP. (“Source version” is the previously installed version.)

-s Snapshot File Flag. Specifies the snapshot file to convert. If you do not specify a path for the snapshot file, the command uses the path specified in the \$SNAPSHOTPATH variable. The default is `/usr/[es]/sbin/cluster/snapshots`.

Examples

Example 1

```
clconvert_snapshot -v 4.3.1 -s mysnapshot.odm
```

Note: This example is applicable if HACMP 4.4 is currently installed on the node where **clconvert_snapshot** will be run and **mysnapshot.odm** is an HACMP 4.3.1 snapshot ODM data file. (Refer to the section Format of a Cluster Snapshot on page 11-3 for a definition of a snapshot ODM data file.)

In this example, **clconvert_snapshot** will look for **mysnapshot.odm** in the directory specified by the \$SNAPSHOTPATH environmental variable. If a \$SNAPSHOTPATH environmental variable is not provided, **clconvert_snapshot** will look in `/usr/[es]/sbin/cluster/snapshots`.

Since HACMP 4.4 is currently installed on the node, **clconvert_snapshot** will assume **mysnapshot.odm** is an HACMP 4.3.1 snapshot ODM data file. When **clconvert_snapshot** completes, **mysnapshot.odm** will be upgraded to a HACMP 4.4 snapshot ODM data file. Also, a new file, `/tmp/mysnapshot.odm.old`, will be created, which is a copy of the original HACMP 4.3.1 snapshot ODM data file.

Example 2

```
clconvert_snapshot -C -v 4.4 -s /tmp/mysnapshot
```

Note: This example is applicable if HACMP ES 4.4 is currently installed on the node where **clconvert_snapshot** will be run and **mysnapshot.odm** is an HACMP 4.4 snapshot ODM data file. (Refer to the section Format of a Cluster Snapshot on page 11-3 for a definition of a snapshot ODM data file.)

In this example, the `.odm` extension was not specified, therefore, **clconvert_snapshot** will look for a snapshot ODM data file called `/tmp/mysnapshot.odm`.

Since the `-C` flag was specified, **clconvert_snapshot** will assume that `/tmp/mysnapshot.odm` is an HACMP 4.4 snapshot ODM data file. When **clconvert_snapshot** completes, `/tmp/mysnapshot.odm` will be upgraded to a HACMP ES 4.4 snapshot ODM data file. Also, a new file, `/tmp/mysnapshot.odm.old`, will be created. This is a copy of the original HACMP 4.4 snapshot ODM data file.

Example 1

```
clconvert_snapshot -C -v 4.3.1 -s mysnapshot
```

Converts an HACMP 4.3.1 snapshot to an HACMP Enhanced Scalability 4.4 snapshot named "mysnapshot."

“Mysnapshot” is in turn placed in the directory specified by the \$SNAPSHOTPATH environmental variable. If a \$SNAPSHOTPATH variable is not specified, the file is put in **/usr/[es]/sbin/cluster/snapshots**.

Example 2

```
clconvert_snapshot -C -v 4.3.1 -s /tmp/mysnapshot
```

Converts an HACMP 4.3.1 snapshot to an HACMP ES 4.4 snapshot named **/tmp/mysnapshot**.

In this case, since a directory is specified, **clconvert_snapshot** will ignore the \$SNAPSHOTPATH environmental variable.

The snapshot will appear as **/tmp/mysnapshot**.

cldiag debug { clstrmgr [-l *level*] | cllockd } [-R *file*]

Enables real-time debugging of the Cluster Manager or dumping of the Lock Manager’s lock resource table.

clstrmgr	Allows viewing of Cluster Manager debug information.
cllockd	Gets a dump of the Cluster Lock Manager’s lock resource table.
-l <i>level</i>	The level of debugging to be performed. The levels range from 1 to 9 in increasing amounts of information. The default (0) turns debugging off.
-R <i>file</i>	Saves output in the specified file.

Example

```
cldiag debug clstrmgr -l2 -R foo
```

Enables Cluster Manager debugging at debug level 2 and saves output to the file named *foo*.

**cldiag logs {scripts [-s] [-f] [event...] | syslog [-e] [-w] }
[-h *hostname*] [-d #_of_days] [-R *file*]**

Allows for selected viewing and parsing of HACMP for AIX process and script output files.

-h <i>hostname</i>	Hostname of the system from which to gather data.
-s	Captures all “Start” and “Complete” events (scripts option).
-f	Captures all “Fail” events (scripts option).
-d #_of_days	The number of days preceding the present day from which information will be gathered.
event...	Captures all lines containing the cluster event (such as node_up_local). This can be a list of events (scripts option).
-e	Captures all “Error” events (syslog option).
-w	Captures all “Warning” events (syslog option).
-R <i>file</i>	Saves output in the specified file.

Example

```
cldiag logs scripts -f -d3
```

Captures all failed script events that occurred within the last three days.

cldiag vgs -h *hostnames* [-v *vgnames*]

Checks for consistencies among volume groups on various hosts, ODMs, and disks.

-h <i>hostnames</i>	List of 2 – 8 hostnames, separated by commas (no space).
-v <i>vgnames</i>	List of 2 – 8 volume group names, separated by commas (no space).

Examples

Example 1

```
cldiag vgs -h jaws,kelp
```

Checks all common volume groups for hosts *jaws* and *kelp*.

Example 2

```
cldiag vgs -h limpet,cowrie -v vgck1
```

Checks volume group *vgck1* for hosts *limpet* and *cowrie*.

Note: Occasionally, using the **vgs** option causes the utility to exit prematurely. If you want to check the consistency of volume group, logical volume, and filesystem information among nodes, and you encounter this problem, run the **clverify** routine instead, using SMIT or the command line. For more information about running **clverify**, see Chapter 8, Verifying a Cluster Configuration, and the **clverify** syntax section on page A-20 of this appendix.

cldiag error { [short | long | cluster] } [-h *hostname*] [-R *file*]

Allows for parsing the system error log for errors that occurred in a cluster.

short	Short error report.
long	Long error report.
cluster	clstrmgr and HACMP for AIX error report.
-h <i>hostname</i>	Hostname of the system from which to gather data.
-R <i>file</i>	Saves output in the specified file.

Example

```
cldiag error long -h steamer -R err_rep
```

Generates a long error report for host steamer and sends it to the file **err_rep**.

cldiag trace [-t *time*] [-R *file*] [-l] *daemon ...*

Allows for tracing HACMP daemons (**clstrmgr**, **cllockd**, **clsmuxpd**, **clinfo**).

-t <i>time</i>	Number of seconds to perform the trace.
-R <i>file</i>	Redirects output to this file.
-l	Chooses more detailed trace daemon.
<i>daemon</i>	List of cluster daemons to trace.

Example

```
cldiag trace -t 45 clinfo
```

Traces the **clinfo** daemon for 45 seconds; writing to **stdout**.

clfindres [-s] [resgroup1] [resgroup2] ...

Finds a given resource group or groups in a cluster configuration. This command is intended to be used when trying to locate migrated resource groups.

-s	Requests abbreviated (location only) output.
-----------	--

Example 1: Normal Output

If no arguments are specified, the **clfindres** command lists all resources. As shown in the following output, the output lists resource groups, and defines the resource group type, state, location, and migration type/location. Note that in the type column, the order is always *cascading, rotating, concurrent*.

Typical normal output looks something like the following:

Name	Type	State	Location	Sticky Loc
myres	cascading	UP	crusty	crusty
ourres	cascading	DOWN	N/A	
yourres	rotating	UP	oyster	STOP
hisres	rotating	????	oyster	
themoney	concurrent	UP	[crusty oyster]	
zztop	concurrent	DOWN	N/A	

Note: There are five possible states tracked by **clfindres**: UP, DOWN, ACQUIRING, RELEASING, and ERROR. Resources in any of the latter three states are indicated as being in “????” state. Check the **hacmp.out** file to see if an error exists. If an error condition is the cause of this indication, stop the resource group using the **clbare** command with the **stop** keyword instead of a location, as follows:

```
clbare -M <resgroup name>:stop
```

For more information, see the **clbare** man page and the section Migrating Resources Dynamically—Overview on page 7-17 of this guide.

Example 2: Abbreviated Output

Abbreviated output, when you specify the **-s** flag, shows only the name and location:

```
% clfindres -s myres themoney yourres
myres:crusty
themoney:[crusty oyster]
ourres:**down**
```

clgetaddr [-o *odmdir*] *nodename*

Returns a PINGable address for the specified node name.

-o Specifies an alternate ODM directory.

Example

To get a PINGable address for the node seaweed, enter:

```
clgetaddr seaweed
```

The following address is returned: 2361059035

clgetgrp { -g *group* } [-o] [-c | -h -f *field*]

Retrieves and displays HACMPgroup class objects.

-g <i>group</i>	Specifies name of resource group to list.
-o <i>objdir</i>	Specifies an alternate ODM directory to /etc/objrepos .
-c	Specifies a colon output format.
-h	Specifies to print a header.
-f <i>field</i>	Specifies the field in object to list.

Example

```
clgetgrp -g grp3
```

Lists information for resource group *grp3*.

clgetif { [-a | -n | -d] *IPlabel* | *IPaddress* }

Prints the interface name and/or netmask associated with a specified IP label or address.

-a	Prints interface name (for example, <i>en0</i>).
-n	Prints interface netmask in dotted decimal.
-d	Prints interface device name.
<i>IPlabel</i>	Specifies the name of the adapter to show.
<i>IPaddress</i>	Specifies the address of the adapter to show.

Examples

Example 1

```
clgetif -a clam
```

Prints interface name for adapter with IP label *clam*.

Example 2

```
clgetif -n 1.1.1.22
```

Prints netmask for adapter name with IP address *1.1.1.22*.

Example 3

```
clgetif -an clam_svc clam_stby
```

Prints interface names and netmasks for adapters with IP labels *clam_svc* and *clam_stby*.

clgodmget [-c] [-q *criteria*] { -n *nodename class* }

Lists ODM class remotely.

-c	Specifies a colon output format.
-q <i>criteria</i>	Specifies search criteria for ODM retrieve. See the <code>odmget</code> man page for information on search criteria. If no criteria are entered, all objects in the class are retrieved.
-n <i>nodename</i>	Searches the ODM from the specified node.
<i>class</i>	ODM object class.

Example

```
clgodmget -n abalone HACMPadapter
```

Prints adapter information for the node *abalone*.

cllscf

Lists complete cluster topology information.

cllscfstr [-i *id*]

Shows cluster name and ID in the cluster configuration ODM object class. If no cluster ID is included, shows the information for the cluster where the local node is configured.

-i <i>id</i>	Cluster ID to show.
---------------------	---------------------

Example

```
cllscfstr -i 2
```

Shows the cluster name and ID as configured for cluster 2.

cllsdisk { -g *resource_group* }

Lists PVIDs of accessible disks in a specified resource chain.

-g *resource_group*

Specifies name of resource group to list.

Example

```
cllsdisk -g grp3
```

Lists PVIDs of disks accessible in resource group *grp3*.

cllsfs { -g *resource_group* } [-n]

Lists shared filesystems contained in a resource group.

- g *resource_group*** Specifies name of resource group for which to list filesystems.
- n** Lists the nodes that share the filesystem in the resource group.

Note: Do not run the cllsfs command from the command line. Use the SMIT interface to retrieve filesystem information, as explained in Chapter 4, Maintaining Shared LVM Components.

cllslv [-g *resource_group*] [-n] [-v]

Lists the names of logical volumes accessible by nodes in a specified resource chain.

- g *resource_group*** Specifies name of resource group for which to list logical volumes.
- n** Lists the nodes in the resource group.
- v** Lists only those logical volumes that belong to volume groups that are currently varied on.

Example 1

```
cllslv -g grp2
```

Lists all shared logical volumes contained in resource group *grp2*.

Example 2

```
cllslv -g grp2 -n -v
```

Displays the nodes and those logical volumes that belong to currently varied-on volume groups in resource group *grp2*.

cllsgrp

Lists names of all resource groups configured in the cluster.

cllsif [-n *Ilabel*] [-i *nodename*] [-S] [-c] [-x]

Shows adapter information. If no options are used, shows information for all adapters.

- n *Ilabel*** Name of the adapter to show.
- i *nodename*** Node name. All adapters configured to HACMP on this node are shown.
- S** Specifies that the output appear in column format without colons. Also lists node names associated with this interface.
- c** Specifies that the output appear in colon-separated format. When used with the **-S** options, node names are not listed.
- x** Shows only adapter labels (no additional adapter attributes).

Examples

Example 1

```
cllsif -n foo
```

Shows information about the adapter labeled *foo*.

Example 2

```
cllsif -i seaweed
```

Shows information about all adapters configured on node *seaweed*.

Example 3

```
cllsif -i seaweed -x
```

Lists the labels of adapters on node *seaweed*.

cllsnim [-d *odmdir*] [-c] [-n *nimname*]

Lists contents of HACMPnetwork interface module ODM class.

- | | |
|--------------------------|---|
| -d <i>odmdir</i> | Specifies an alternate ODM directory to <i>/etc/objrepos</i> . |
| -c | Specifies a colon output format. |
| -n <i>nimname</i> | Name of the network interface module for which to list information. |

Examples

Example 1

```
cllsnim
```

Shows information for all configured network modules.

Example 2

```
cllsnim -n ether
```

Shows information for all configured Ethernet network modules.

cllsnode [-i *nodename*]

Shows node information retrieved from adapter and network ODM object classes.

- | | |
|---------------------------|---------------------------|
| -i <i>nodename</i> | Node name of target node. |
|---------------------------|---------------------------|

Examples

Example 1

```
cllsnode
```

Shows information for all configured nodes.

Example 2

```
cllsnode -i seaweed
```

Shows information for node *seaweed*.

cllsnw [-n *name*]

Shows network information defined in the adapter and network configuration ODMs.

-n *name* Name of network for information to display.

Examples

Example 1

```
cllsnw
```

Shows all networks.

Example 2

```
cllsnw -n ether
```

Shows network named *ether*.

cllsparam { -n *nodename* } [-c] [-s] [-d *odmdir*]

Lists run-time parameters.

-n *nodename* Specifies a node for which to list the information.

-c Specifies a colon output format.

-s Used along with the **-c** flag, specifies native language instead of English.

-d *odmdir* Specifies an alternate ODM directory.

Example

```
cllsparam -n abalone
```

Shows run-time parameters for node *abalone*.

cllsres [-g *group*] [-c] [-s] [-f *field*] [-d *odmdir*] [-q *query*]

Sorts HACMP for AIX ODM resource data by name and arguments.

-g *group* Specifies name of resource group to list.

-c Specifies a colon output format.

-s Used with the **-c** flag, specifies native language instead of English.

-f *field* Specifies the field in object to list. Fields are the resource group name, the type of resource group (cascading, concurrent, or rotating), and participating nodes.

-d *odmdir* Specifies an alternate ODM directory.

-q *query* Specifies search criteria for ODM retrieve. See the **odmget** man page for information on search criteria.

Examples

Example 1

```
cllsres
```

Lists resource data for all resource groups.

Example 2

```
cllsres -g grp1
```

Lists resource data for resource group *grp1*.

Example 3

```
cllsres -g grp1 -q"name = FILESYSTEM"
```

Lists filesystem resource data for resource group *grp1*.

cllsserv [-c] [-h] [-n *name*] [-d *odmdir*]

Lists application servers by name.

- | | |
|-------------------------|---|
| -c | Specifies a colon output format. |
| -h | Specifies to print a header. |
| -n <i>name</i> | Specifies an application server for which to check information. |
| -d <i>odmdir</i> | Specifies an alternate ODM directory. |

Examples

Example 1

```
cllsserv
```

Lists all application servers.

Example 2

```
cllsres -c -n test1
```

Lists information in colon format for application server *test1*.

cllsvg { -g *resource_group* } [-n]

List shared volume groups in a specified resource chain.

- | | |
|---------------------------------|---|
| -g <i>resource_group</i> | Specifies name of resource group for which to list volume groups. |
| -n <i>nodes</i> | Specifies all nodes participating in a resource group. |

Example

```
cllsvg -g grp1
```

Lists all shared volume groups in resource group *grp1*.

**clnodename [-d *odmdir*] [-o *oldname* -n *newname*]
[-V *high/low*][-N *true/false*]**

clnodename [-d *odmdir*] [-V *high/low*] [-N *true/false*] -a *name*

clnodename [-d *odmdir*] -r *name*]

Lists, adds, changes, or removes a cluster node name.

-d <i>odmdir</i>	Specifies an alternate ODM directory other than /etc/objrepos .
-o <i>oldname</i>	Specifies the old name when changing a node name.
-n <i>newname</i>	Specifies the new name when changing a node name.
-a <i>name</i>	Specifies the node name or names to be added to the cluster.
-r <i>name</i>	Specifies the node name to be removed
-V <i>high/low</i>	Verbose logging on high or low.
-N <i>true/false</i>	Nameserving is used: true or false.

Example

To list all node names configured for a cluster, enter:

```
clnodename
```

The command returns the list of node names.

clshowres [-g *group*] [-n *nodename*] [-d *odmdir*]

Shows resource group information for a cluster or a node.

-g <i>group</i>	Name of resource group to show.
-n <i>nodename</i>	Searches the resources ODM from the specified node.
-d <i>odmdir</i>	Specifies <i>odmdir</i> as the ODM object repository directory instead of the default /etc/objrepos .

Examples

Example 1

```
clshowres
```

Lists all the resource group information for the cluster.

Example 2

```
clshowres -n clam
```

Lists the resource group information for node *clam*.

clstat [-c *id* | -i] [-r *seconds*] [-a]

Cluster Status Monitor (ASCII mode).

- c *id*** Displays cluster information only about the cluster with the specified ID. If the specified cluster is not available, **clstat** continues looking for the cluster until the cluster is found or the program is cancelled. May not be specified if the **-i** option is used.
- i** Runs ASCII **clstat** in interactive mode. Initially displays a list of all clusters accessible to the system. The user must select the cluster for which to display the detailed information. A number of functions are available from the detailed display.
- r *seconds*** Updates the cluster status display at the specified number of seconds. The default is 1 second; however, the display is updated only if the cluster state changes.
- a** Causes **clstat** to display in ASCII mode.

clstat [-a] [-c *id*] [-r *tenths-of-seconds*]

Cluster Status Monitor (X Windows mode).

- a** Runs **clstat** in ASCII mode.
- c *id*** Displays cluster information only about the cluster with the specified ID. If the specified cluster is not available, **clstat** continues looking for the cluster until the cluster is found or the program is cancelled. May not be specified if the **-n** option is used.
- r *tenths-of-seconds*** The interval at which the **clstat** utility updates the display. For the graphical interface, this value is interpreted in tenths of seconds. By default, **clstat** updates the display every 0.10 seconds.

Examples

Example 1

```
clstat -c 10
```

Displays the cluster information about the cluster whose ID is *10*.

Example 2

```
clstat -i
```

Runs ASCII **clstat** in interactive mode, allowing multi-cluster monitoring.

Example 3

```
clstat -n waves
```

Displays information about the cluster named *waves*.

Buttons on X Window System Display

Prev	Displays previous cluster.
Next	Displays next cluster.
Name:Id	Refresh bar, pressing bar causes clstat to refresh immediately.
Quit	Exits application.
Help	Pop-up help window shows the clstat manual page.

clverify cluster { topology check | topology sync | config networks | config resources | config all} [-e *num*] [-R *file*]

Verifies cluster installation and configuration.

topology check	Checks that all nodes agree on cluster topology.
topology sync	Forces all nodes to agree on cluster topology as defined on the local node.
config resources	Verifies ownership of all disks and other resources.
config networks	Verifies configuration of network adapters.
config all	Runs resources and networks programs and all custom verification methods.
-e <i>num</i>	Aborts the program after <i>num</i> errors (not available for use with topology options).
-R <i>file</i>	Redirects output to a file (not available for use with topology options).

Example

```
clverify cluster config networks -R verify_nw
```

Verifies the cluster networks configuration and logs the results in a file called **verify_nw**.

clverify software {lpp} [-R *file*]

Verifies the modifications to AIX system files.

lpp Verifies that HACMP-specific modifications to AIX system files are correct.

-R *file* Redirects output to a file.

Example

```
clverify software lpp -R verify_lpp
```

Verifies that the proper HACMP-specific modifications exist and are correct, and logs the output in a file called **verify_lpp**.

get_local_nodename

Returns the name of the local node.

clgetactivenodes [-n *nodename*] [-o *odmdir*] [-t *timeout*] [-v *verbose*]

Retrieves the names of all cluster nodes.

-n *nodename* Determines if the specified node is active.

-o *odmdir* Specifies *odmdir* as the ODM object repository directory instead of the default **/etc/objrepos**.

-t *timeout* Specifies a maximum time interval for receiving information about active nodes.

-v *verbose* Specifies that information about active nodes be displayed as verbose output.

Example

```
clgetactivenodes -n java
```

Verifies that node java is active.

**clresactive { -v *volume*group | -l *logical*volume | -f *filesystem*
|-u *user* |-g *group* |-V *HACMP* version |-c [::*cmd*]}**

Retrieves the names of all active cluster resources.

-v <i>volume</i>group	Specifies the status of a volume group.
-l <i>logical</i>volume	Specifies the status of a logical volume.
-f <i>filesystem</i>	Specifies the status of a filesystems.
-u <i>user</i>	Specifies a user account.
-g <i>group</i>	Specifies a user group.
-V <i>HACMP</i> version	Specifies the current HACMP for AIX version.
-c [::<i>cmd</i>]	Specifies several commands to be executed simultaneously.

Example

```
clresactive -g finance
```

clm_stats [-t *Seconds*]

Queries the HACMP Cluster Lock Manager (CLM) for current information about the number of locks, resources, and memory usage.

-t <i>Seconds</i>	Specifies the interval in seconds between each line of output.
--------------------------	--

Example

To specify that the **clm_stats** utility runs continuously, producing output every five seconds, enter:

```
clm_stats -t 5
```

Information displayed will list the following:

- Number of locks in use, allocated, and refused
- Number of resources in use or allocated
- Number of megabytes (MB) of memory available, in use, percentage used, and the percentage limit.

For more information, see the **clm_stats** manpage.

HACMP for AIX C-SPOC Commands

The following C-SPOC commands can be executed from the command line and through SMIT. Error messages and warnings returned by the commands are based on the underlying AIX-related commands.

Note: While the AIX commands, underlying the C-SPOC commands, allow you to specify flags in any order, even flags that require arguments, the C-SPOC commands require that arguments to command flags must immediately follow the flag. See the `cl_lsuser` command for an example.

cl_chfs -cspoc “[-f] [-g *ResourceGroup* | -n *Nodelist*]”
[-m *newmntpt*] [-u *mtgrp*] [-A { *yes* | *no* }] [-p { *ro* | *rw* }]
[-a *attr1=val1*] [-d *attr*] [-t {*yes* | *no*}] **FileSystem**

Changes the attributes of a shared filesystem on the node that currently has the filesystem’s underlying shared volume group varied on. Not supported for concurrent volume groups.

Note: Arguments associated with a particular flag must be specified immediately following the flag.

- | | |
|-----------------------------|--|
| -cspoc | Argument used to specify one of the following C-SPOC options:
-f – Forces <code>cl_chfs</code> to check the accessibility of the participating nodes in the specified resource group.
-g <i>ResourceGroup</i> – Generates the list of nodes participating in the resource group where the command will be executed.
-n <i>Nodelist</i> – Specifies this list of nodes on which to execute the command. If more than one node, separate nodes listed by commas. |
| -a <i>attr1=val1</i> | Specifies the attribute/value pairs. In relation to the journaled filesystem, the <code>size=Value</code> parameter indicates the size of the journaled filesystem in 512-byte blocks. If <code>Value</code> begins with a <code>+</code> , it is interpreted as a request to increase the filesystem size by the specified amount. If the specified size is not evenly divisible by the physical partition size, it is rounded up to the closest evenly divisible number. Also specific to the journaled filesystem is the <code>logname=LVName</code> parameter, which specifies the logical volume name of the existing log to be used. |
| -A | Specifies the attributes for Auto-mount: yes =Filesystem is automatically mounted at system restart; no =Filesystem is not mounted at system restart. |
| -m <i>newmntpt</i> | Specifies the new mount point. |
| -p | Sets the permissions for the filesystem: ro specifies read-only permissions; rw specifies read-write permissions. |

- t** Sets an accounting entry in the `/etc/filesystems` file: **yes**=The accounting=entry is set to true; **no**=The accounting=entry is set to false.
- u *mtgrp*** Specifies the mount group. Mount groups are used to group related mounts, so they can be mounted as one instead of mounting each individually. For example, if several scratch filesystems always need to be mounted together when performing certain tests, they can each be placed in the test mount group. They can then all be mounted with a single command, such as the **mount -t test** command.

Examples

Example 1

To change the filesystem size of the `/test` shared filesystem, enter:

```
cl_chfs -cspoc "-g resource_grp1" -a size=24576 /test
```

This command changes the size of the `/test` journaled filesystem to 24576 512-byte blocks, or 12 MB on the node in the given resource group.

Example 2

To increase the size of the `/test` shared filesystem, enter:

```
cl_chfs -cspoc "-g resource_grp1" -a size+=8192 /test
```

This increases the size of the `/test` journaled filesystem by 8192 512-byte blocks, or 4 MB on the node in the given resource group.

Example 3

To change the mount point of a filesystem, enter:

```
cl_chfs -cspoc "-g resource_grp1" -m /test2 /test
```

This changes the mount point of a filesystem from `/test` to `/test2` on the node in the given resource group.

cl_chgroup [-cspoc “[-f] [-g *ResourceGroup*|-n *nodelist*]]” Attribute=Value ... *Group*

Changes group attributes on all nodes in an HACMP cluster.

Warning: Do not use the `cl_chgroup` command if you have a Network Information Service (NIS) database installed on any cluster node; using the command in this environment can cause database inconsistencies.

Note: Arguments associated with a particular flag must be specified immediately following the flag.

-cspoc Argument used to specify the following C-SPOC option:

- f** – Forces **cl_chgroup** to skip default attribute verification. If this flag is set and a cluster node is not accessible, **cl_chgroup** reports a warning and continues execution on the other cluster nodes.
- g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
- n nodelist** – Specifies the list of nodes where the command should execute. If more than one node, separate nodes listed by commas.

Examples

Example 1

To add Sam and Carol to the finance group on all cluster nodes that currently only have Frank as a member, enter:

```
cl_chgroup users=sam,carol,finance
```

Example 2

To remove Frank as a user in the finance group on all cluster nodes and retain Sam and Carol, enter:

```
cl_chgroup users=sam,carol
```

In the second example, two attribute values were changed. The name Frank was omitted from the list of members, and the value of the adms attribute was left blank.

cl_chlv -cspoc “[-f] [-g ResourceGroup / -n nodelist]” [-a Position] [-b BadBlocks] [-d Schedule] [-e Range] [-L label] [-p Permission][-r Relocate] [-s Strict] [-t Type] [-u Upperbound] [-v Verify] [-w MirrorWriteConsistency] [-x Maximum] LogicalVolume ...

Changes an attribute of a shared logical volume.

cl_chlv -cspoc “[-f] [-g ResourceGroup | -n nodelist]” -n NewLogicalVolume LogicalVolume

Change the name of a shared logical volume:

Note: Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – At least one node in the resource group is accessible. If no node has the volume group varied on, the system chooses a node in the resource group, varies on the volume group there, executes the command, and varies off the volume group.
 - g ResourceGroup** – Generates the list of nodes participating in this resource group on which the command will be executed.
 - n nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- a Position** Sets the intra-physical volume allocation policy (the position of the logical partitions on the physical volume). The **Position** variable is represented by one of the following:
- m** – Allocates logical partitions in the outer middle section of each physical volume. This is the default position.
 - c** – Allocates logical partitions in the center section of each physical volume.
 - e** – Allocates logical partitions in the outer edge section of each physical volume.
 - ie** – Allocates logical partitions in the inner edge section of each physical volume.
 - im** – Allocates logical partitions in the inner middle section of each physical volume.
- b BadBlocks** Sets the bad-block relocation policy. The **BadBlocks** variable is represented by one of the following:
- y** – Causes bad-block relocation to occur.
 - n** – Prevents bad-block relocation from occurring.
- d Schedule** Sets the scheduling policy when more than one logical partition is written. The **Schedule** variable is represented by one of the following:
- p** – Establishes a parallel scheduling policy.
 - s** – Establishes a sequential scheduling policy.

- e Range** Sets the inter-physical volume allocation policy (the number of physical volumes to extend across, using the volumes that provide the best allocation). The value of the **Range** variable is limited by the **Upperbound** variable, set with the **-u** flag, and is represented by one of the following:
- x** – Allocates logical partitions across the maximum number of physical volumes.
 - m** – Allocates logical partitions across the minimum number of physical volumes.
- L Label** Sets the logical volume label. The maximum size of the **Label** variable is 127 characters.
- n NewLogicalVolume** Changes the name of the logical volume to that specified by the **NewLogicalVolume** variable. Logical volume names must be unique and system-wide, and can range from 1 to 15 characters.
- p Permission** Sets the access permission to read-write or read-only. The **Permission** variable is represented by one of the following:
- w** – Sets the access permission to read-write.
 - r** – Sets the access permission to read-only.
- r Relocate** Sets the reorganization flag to allow or prevent the relocation of the logical volume during reorganization. The **Relocate** variable is represented by one of the following:
- y** – Allows the logical volume to be relocated during reorganization.
 - n** – Prevents the logical volume from being relocated during reorganization.
- s Strict** Determines the strict allocation policy. Copies of a logical partition can be allocated to share or not to share the same physical volume. The **Strict** variable is represented by one of the following:
- y** – Sets a strict allocation policy, so copies of a logical partition cannot share the same physical volume.
 - n** – Does not set a strict allocation policy, so copies of a logical partition can share the same physical volume.
- t Type** Sets the logical volume type. The maximum size is 31 characters.
- u Upperbound** Sets the maximum number of physical volumes for new allocation. The value of the **Upperbound** variable should be between one and the total number of physical volumes.

- v Verify** Sets the write-verify state for the logical volume. Causes all writes to the logical volume either to be verified with a follow-up read or not to be verified with a follow-up read. The **Verify** variable is represented by one of the following:
- y** – Causes all writes to the logical volume to be verified with a follow-up read.
 - n** – Causes all writes to the logical volume not to be verified with a follow-up read.
- MirrorWriteConsistency y** – Turns on mirror write consistency which insures data consistency among mirrored copies of a logical volume during normal I/O processing.
- n** – No mirror write consistency. See the **-f** flag of the **syncvg** command.
- x Maximum** – Sets the maximum number of logical partitions that can be allocated to the logical volume.

Examples

Example 1

To change the inter-physical volume allocation policy of logical volume lv01, enter:

```
cl_chlv -cspoc "-g resource_grp1" -e m lv01
```

The inter-physical volume allocation policy is set to minimum.

Example 2

To change the type of logical volume lv03, enter:

```
cl_chlv -cspoc "-g resource_grp1" -t copy lv03
```

Example 3

To change the permission of logical volume lv03 to read-only, enter:

```
cl_chlv -cspoc "-g resource_grp1" -p r lv03
```

Logical volume lv03 now has read-only permission.

Example 4

To change the type to paging and the maximum number of physical volumes for logical volume lv03, enter:

```
cl_chlv -cspoc "-g resource_grp1" -t paging -u 10 lv03
```

The change in the type of logical volume takes effect immediately, but the change in the maximum number of physical volumes does not take effect until a new allocation is made.

Example 5

To change the allocation characteristics of logical volume lv07, enter:

```
cl_chlv -cspoc "-g resource_grp1" -a e -e x -r y -s n -u 5 lv07
```

cl_chuser [-cspoc “[-f] [-g ResourceGroup | -n nodelist]”] Attribute=Value ... Name

Changes user attributes for a user on all nodes in an HACMP cluster.

Warning: Do not use the **cl_chuser** command if you have a Network Information Service (NIS) database installed on any cluster node; using the command in this environment can cause database inconsistencies.

Note: Arguments associated with a particular flag must be specified immediately following the flag.

-cspoc Argument used to specify the following C-SPOC option:

- f** – Forces **cl_chuser** to skip default verification. If the flag is set and a cluster node is not accessible, **cl_chuser** reports a warning and continues execution on the other cluster nodes. If the flag is set and the user does not exist on a node, **cl_chuser** reports a warning and continues execution on the other cluster nodes.
- g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
- n nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

Examples

Example 1

To enable user smith to access every node in the cluster remotely, enter:

```
cl_chuser rlogin=true smith
```

Example 2

To change the expiration date for the davis user account to 8 a.m., 1 May, 1999 on every cluster node, enter:

```
cl_chuser expires=0501080099 davis
```

Example 3

To add davis to the groups finance and accounting on every cluster node, enter:

```
cl_chuser groups=finance,accounting davis
```

cl_crfs [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” -v VfsType { -g VolumeGroup | -d Device } [-l LogPartitions]-m MountPoint [-n NodeName] [-u MountGroup] [-A { yes | no }] [-p {ro | rw }] [-a Attribute=Value ...] [-t {yes | no }]

Creates a filesystem on specified list of cluster nodes.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
 - g *ResourceGroup*** – Generates the list of nodes participating in the resource group where the command will be executed.
 - n *Nodelist*** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

All other flags listed belong to the AIX **crfs** command and operate as outlined in the **crfs** man pages.

Example

To create a filesystem on nodes alpha and beta, members of resource group foo, enter:

```
cl_crfs -cspoc "-f -g foo" -d hdisk1 -m/home2 -a frag=4096 -a nbpi=4096  
-a ag=8
```

cl_crlvfs [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” -v VfsType { -g VolumeGroup | -d Device } [-l LogPartitions]-m MountPoint [-n NodeName] [-u MountGroup] [-A { yes| no }] [-p {ro | rw }] [-a Attribute=Value ...] [-t {yes | no }]

Creates a logical volume and journaled filesystem on specified list of cluster nodes.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
 - g *ResourceGroup*** – Generates the list of nodes participating in the resource group where the command will be executed.
 - n *Nodelist*** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

All other flags listed belong to the AIX **crfs** command and operate as outlined in the **crfs** man pages.

Example

To create a logical volume and filesystem on nodes alpha and beta, members of resource group foo, enter:

```
cl_crlvfs -cspoc "-f -g foo" -g vgl -m/home2 -a frag=4096 -a nbpi=4096
-a ag=8
```

```
cl_clstop [ -cspoc "[ -f ] [-g ResourceGroup | -n NodeList ] "" -f
cl_clstop [ -cspoc "[ -f ] [-g ResourceGroup | -n NodeList ] "" -g [ -s ] [ -y ]
[ -N | -R | -B ]
cl_clstop [ -cspoc "[ -f ] [-g ResourceGroup | -n NodeList ] "" -gr [ -s ] [ -y ]
[ -N | -R | -B ]
```

Stops Cluster daemons using the System Resource Controller (SRC) facility.

Note: Arguments associated with a particular flag must be specified immediately following the flag.

- | | |
|---------------|---|
| -cspoc | Argument used to specify one of the following C-SPOC options:
-f – Forces cl_stop to skip default verification. If this flag is set and a cluster node is not accessible, cl_clstop reports a warning and continues execution on the other nodes.
-g ResourceGroup – Generates the list of nodes participating in the resource group where the command will be executed.
-n NodeList – Shuts down cluster services on the nodes specified in the node list. |
| -f | Forced shutdown. Cluster daemons should terminate without running any local events. Resources are not released. |
| -g | Graceful shutdown with no takeover. Resources are not released. |
| -gr | Graceful shutdown with the resources being released by this node and taken over by another node, if configured. The daemons should terminate gracefully, and the node should release its resources, which will then be taken over. A node list must be specified for graceful shutdown with takeover. |
| -s | Silent shutdown, specifies not to broadcast a shutdown message through /bin/wall . The default is to broadcast. |
| -y | Do not asks operator for confirmation before shutting down. |
| -N | Shut down now. |
| -R | Stops on subsequent system restart (removes the inittab entry). |
| -B | Stop now. |

Examples

Example 1

To shut down the cluster node gracefully on node1 (releasing the resources) with no warning broadcast to users before the cluster processes are stopped and resources are released, enter:

```
cl_clstop -cspoc "-n node1" -gr -s -y
```

Example 2

To forcefully and immediately shut down the cluster on all cluster nodes (resources not released) with warning broadcast to users before the cluster processes are stopped, enter:

```
cl_clstop -f -y
```

Example 3

To shut down the cluster node gracefully on all cluster nodes, broadcasting a warning to users before the cluster processes are stopped, enter:

```
cl_clstop -g -y
```

cl_extendvg -cspoc “[-c|-s] [-f][-g ResourceGroup | -n NodeList]” -R RefNode [-f] VolumeGroup PhysicalVolume

The **cl_extendvg** command allows a physical volume to be added to a volume group.

Note: If you do not specify either the **-g** or **-n** option, the default action occurs on all cluster nodes.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached. If the VG is defined as concurrent, the command does not run and the user is notified.
 - g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
 - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- R RefNode** Node from which to retrieve hdisk names.
- f** Forces the physical volume to be added to the volume group unless it is a member of another volume group in the Configuration Database or of another active volume group.

Example

To add physical volumes hdisk3 and hdisk8 residing on nodeA to volume group vg3 (member of rsgrp1), enter:

```
cl_extendvg -cspoc "-g rsgrp1" -R nodeA vg3 hdisk3 hdisk8
```

**cl_extendlv [-cspoc “[-f] [-g ResourceGroup | -n NodeList]”][-a Position]
[-e Range] [-u Upperbound] [-s Strict] [-R ReferenceNode]
LogicalVolume Partitions [PhysicalVolume ...]**

Increases the size of a logical volume by adding available unallocated physical partitions from within the volume group.

**cl_extendlv [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” -m MapFile
LogicalVolume Partitions**

Increases the size of a logical volume by adding specific unallocated physical partitions from within the volume group.

Note: This command is not allowed if the volume group is varied on in concurrent mode.

Note: The -e, -m, -s, and -u flags are not valid with a striped logical volume.

-cspoc

Argument used to specify one of the following C-SPOC options:

-f – If no node has the underlying VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached. If the VG is defined as concurrent, the command does not run and the user is notified.

-g ResourceGroup – Generates the list of nodes participating in the resource group where the command will be executed.

-n Nodelist – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

-R RefNode

Node from which to retrieve hdisk names.

-a position

Sets the intraphysical volume allocation policy (the position of the logical partitions on the physical volume). The Position variable can be one of the following:

m Allocates logical partitions in the outer middle section of each physical volume. This is the default position.

c Allocates logical partitions in the center section of each physical volume.

e Allocates logical partitions in the outer edge section of each physical volume.

ie Allocates logical partitions in the inner edge section of each physical volume.

im Allocates logical partitions in the inner middle section of each physical volume.

- e range** Sets the interphysical volume allocation policy (the number of physical volumes to extend across, using the volumes that provide the best allocation). The value of the Range variable is limited by the Upperbound variable (set with the **-u** flag) and can be one of the following:
- x Allocates logical partitions across the maximum number of physical volumes.
 - m Allocates logical partitions across the minimum number of physical volumes.
- m MapFile** Specifies the exact physical partitions to allocate. Partitions are used in the order given in the **MapFile** parameter. Used partitions in the **MapFile** parameter are skipped. All physical partitions belonging to a copy are allocated before allocating for the next copy of the logical volume. The **MapFile** parameter format is: PVname:PPnum1[-PPnum2]. In this example, PVname is a physical volume name (for example,hdisk0). It is one record per physical partition or a range of consecutive physical partitions. PPnum is the physical partition number, which can range from 1 to 1016.
- s Strict** Determines the strict allocation policy. Copies for a logical partition can be allocated to share or not to share the same physical volume. The Strict variable can be one of the following:
- y Sets a strict allocation policy, so copies for a logical partition cannot share the same physical volume.
 - n Does not set a strict allocation policy, so copies for a logical partition can share the same physical volume.
- u Upperbound** Sets the maximum number of physical volumes for new allocation. The value should be between one and the total number of physical volumes in the volume group.

Example

To increase the size of the logical volume represented by the lv05 directory by three logical partitions, enter:

```
cl_extendlv lv05 3
```

cl_importvg -cspoc “[-f] [-g ResourceGroup | -n NodeList]” -R RefNode [-V MajorNumber] [-c] [-x] -y VolumeGroup PhysicalVolume

The **cl_importvg** command imports a volume group. The list of nodes provided must contain at least one node that has knowledge of the volume group. This ensures that the physical volume provided can be verified to belong to the volume group in question. The underlying **importvg** operates on nodes in the list that have no knowledge of the volume group, or on nodes that are aware of the volume group, but do not have the volume group varied on, using the appropriate flags.

Note: When you issue the **cl_importvg** command to a previously defined volume group, the QUORUM value is set to default and the AUTO ON value is set to NO. You should verify the parameters of the newly imported volume group with the **clsvg** command and change any values with the **cl_chvg** command.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** –If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
 - g ResourceGroup** – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.
 - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- R RefNode** Node from which to retrieve hdisk names.
- V MajorNumber** Major number of imported volume group
- c** Make volume group concurrent-capable.
- x** If used with **-c**; make concurrent-capable volume group autovaryon in concurrent mode.
- In order for this auto-varyon into concurrency of the volume group to take effect, you must enter the following line into the **/etc/inittab** file:
- ```
rc_clvmv:2:wait:/usr/sbin/clvm_cfg 2>&1
```
- This entry must be added after the entry used to initiate **srcmstr**.
- y VolumeGroup** Name of new volume group. If this flag is not used, the system automatically generates a new name. The volume group name can only contain the following characters: “A” through “Z,” “a” through “z,” “0” through “9,” or “\_” (the underscore), “-” (the minus sign), or “.” (the period). All other characters are considered invalid.
- Physical Volume** One physical volume to identify volume group. Get name from **-R RefNode**.

## Example

To import the volume group **bkgv**, member of resource group **resgrp1**, from physical volume **hdisk7** residing on **NodeA**, enter:

```
cl_importvg -cspoc "-g rsgpr1" -R NodeA -y bkgv hdisk7
```

The volume group **bkgv** is made known to the system.

## **cl\_lsfs [-cspoc "[-f] [-g ResourceGroup | -n Nodelist]" [-q] [-c | -l] FileSystem] ...**

Displays the characteristics of shared filesystems.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- |               |                                                                                                                                                                                                                                                                                                                                                                                                            |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b> | Argument used to specify one of the following C-SPOC options:<br><b>-f</b> – This option has no effect when used with the <b>cl_lsfs</b> command.<br><b>-g ResourceGroup</b> – Generates the list of nodes participating in the resource group where the command will be executed.<br><b>-n nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-c</b>     | Specifies a different search pattern to determine if the underlying AIX <b>lsfs</b> command returned data or not.                                                                                                                                                                                                                                                                                          |
| <b>-l</b>     | Specifies that the output should be in list format.                                                                                                                                                                                                                                                                                                                                                        |
| <b>-q</b>     | Queries the logical volume manager (LVM) for the logical volume size (in 512-byte blocks) and queries the JFS superblock for the filesystem size, the fragment size, the compression algorithm (if any), and the number of bytes per i-node (nbpi). This information is displayed in addition to other filesystem characteristics reported by the <b>lsfs</b> command.                                     |

### Example

#### Example 1

To display characteristics about all shared filesystems in the cluster, enter:

```
cl_lsfs
```

#### Example 2

Display characteristics about the filesystems shared amongst the participating nodes in **resource\_grp1**.

```
cl_lsfs -cspoc "-g resource_grp1"
```

## **cl\_lsgroup [-cspoc "[-f] -g ResourceGroup | -n Nodelist]" [-c|-f] [-a | -a List] {ALL | Group [ ,Group] ... }**

Displays attributes of groups that exist on an HACMP cluster.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc**                   Argument used to specify the following C-SPOC option:
- f** – This option has no effect when used with the **cl\_lsgroup** command.
  - g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
  - n nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- a List**                   Specifies the attributes to display. The **List** parameter can include any attribute defined in the **chgroup** command, and requires a blank space between attributes. If you specify an empty list using only the **-a** flag, only the group names are listed.
- c**                         Displays the attributes for each group in colon-separated records, as follows:
- ```
# name: attribute1: attribute2: ...
Group: value1:      value2:      ...
```
- f** Displays the group attributes in stanzas. Each stanza is identified by a group name. Each Attribute=Value pair is listed on a separate line:
- ```
group:
attribute1=value
attribute2=value
attribute3=value
```

## Examples

### Example 1

To display the attributes of the finance group from all cluster nodes enter:

```
cl_lsgroup finance
```

### Example 2

To display in stanza format the ID, members (users), and administrators (adms) of the finance group from all cluster nodes, enter:

```
cl_lsgroup -f -a id users adms finance
```

### Example 3

To display the attributes of all the groups from all the cluster nodes in colon-separated format, enter:

```
cl_lsgroup -c ALL
```

All the attribute information appears, with each attribute separated by a blank space.

## **cl\_lslv [-cspoc “[-f] [-g ResourceGroup | -n Nodelist”] [-l | -m] LogicalVolume**

Displays shared logical volume attributes.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – This option has no effect when used with the **cl\_lsfs** command.
  - g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
  - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- l** Lists information for each physical volume in the shared logical volume. Refer to the **lslv** command for information about the fields displayed.
- m** Lists information for each logical partition. Refer to the **lslv** command for information about the fields displayed. If no flags are specified, information about the shared logical volume and its underlying shared volume group is displayed. Refer to the **lslv** command for the information about the fields displayed.

## Examples

### Example 1

To display information about the shared logical volume lv03, enter:

```
cl_lslv -cspoc -g resource_grp1 lv03
```

Information about logical volume lv03, its logical and physical partitions, and the volume group to which it belongs is displayed.

### Example 2

To display information about a specific logical volume, using the identifier, enter:

```
cl_lslv -g resource_grp1 00000256a81634bc.2
```

All available characteristics and status of this logical volume are displayed.

**cl\_lsuser [-cspoc “[-f] [-g ResourceGroup | -n Nodelist]”]  
[-c | -f] [-a List] {ALL | Name [ ,Name ] ...}**

Displays user account attributes for users that exist on an HACMP cluster.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc**                      Argument used to specify the following C-SPOC option:
  - f** – This option has no effect when used with the **cl\_lsuser** command.
  - g *ResourceGroup*** – Generates the list of nodes participating in the resource group where the command will be executed.
  - n *Nodelist*** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
  
- a *Lists***                      Specifies the attributes to display. The List variable can include any attribute defined in the **chuser** command and requires a blank space between attributes. If you specify an empty list, only the user names are displayed.
  
- c**                                  Displays the user attributes in colon-separated records, as follows:
 

```
name: attribute1: attribute2: ...
User: value1: value2: ...
```
  
- f**                                  Displays the output in stanzas, with each stanza identified by a user name. Each Attribute=Value pair is listed on a separate line:
 

```
user:
attribute1=value
attribute2=value
attribute3=value
```

## Examples

### Example 1

To display in stanza format the user ID and group-related information about the smith account from all cluster nodes, enter:

```
cl_lsuser -fa id pgrp groups admgroups smith
```

### Example 2

To display all attributes of user smith in the default format from all cluster nodes, enter:

```
cl_lsuser smith
```

All attribute information appears, with each attribute separated by a blank space.

### Example 3

To display all attributes of all the users on the cluster, enter:

```
cl_lsuser ALL
```

All attribute information appears, with each attribute separated by a blank space.

## **cl\_lsvg [-cspoc “[-f] [-g *ResourceGroup* / n- *Nodelist*]” [-o] | [-l | -M | -p] *VolumeGroup* ...**

Displays information about shared volume groups.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – This option has no effect when used with the **cl\_lsvg** command.
  - g ResourceGroup** – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.
  - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- p** Lists the following information for each physical volume within the group specified by the **VolumeGroup** parameter:
- Physical volume: A physical volume within the group.
  - PVstate: State of the physical volume.
  - Total PPs: Total number of physical partitions on the physical volume.
  - Free PPs: Number of free physical partitions on the physical volume.
  - Distribution: The number of physical partitions allocated within each section of the physical volume: outer edge, outer middle, center, inner middle, and inner edge of the physical volume.
- l** Lists the following information for each logical volume within the group specified by the **VolumeGroup** parameter:
- LV: A logical volume within the volume group.
  - Type: Logical volume type.
  - LPs: Number of logical partitions in the logical volume.
  - PPs: Number of physical partitions used by the logical volume.
  - PVs: Number of physical volumes used by the logical volume.
- M** Lists the following fields for each logical volume on the physical volume:
- PVname: **PPnum [LVname: LPnum [:Copynum] [PPstate]]**
  - PVname: Name of the physical volume as specified by the system.
  - PPnum: Physical partition number. Physical partition numbers can range from 1 to 1016.
- o** Lists only the active volume groups (those that are varied on). An active volume group is one that is available for use. Refer to the **lsvg** command for the information displayed if no flags are specified.

## Examples

### Example 1

To display the names of all shared volume groups in the cluster, enter:

```
cl_lsvg
nodeA: testvg

nodeB: testvg
```

### Example 2

To display the names of all active shared volume groups in the cluster, enter:

```
cl_lsvg -o
nodeA: testvg
```

The **cl\_lsvg** command lists only the node on which the volume group is varied on.

### Example 3

To display information about the shared volume group vg02, enter:

```
cl_lsvg -cspoc testvg
```

The **cl\_lsvg** command displays the same data as the **lsvg** command, prefixing each line of output with the name of the node on which the volume group is varied on.

## **cl\_mirrorvg -cspoc “[-f][-g ResourceGroup | -n NodeList]” [-R RefNode ] [-S | -s] [-Q] [-c Copies] [-m VolumeGroup [PhysicalVolume...]**

The **cl\_mirrorvg** command allows all logical volumes in a volume group to be mirrored. (Not supported on RAID devices.)

**Note:** When mirroring a concurrent volume group, the volume group must be varied on in concurrent mode on at least one node in the cluster.

- |                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b>     | Argument used to specify one of the following C-SPOC options:<br><b>f</b> – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.<br><b>-g ResourceGroup</b> – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.<br><b>-n Nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-R RefNode</b> | Node from which to retrieve hdisk names.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>-S</b>         | Background Sync.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>-s</b>         | Disable sync.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>-Q</b>         | Keep quorum                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <b>-c Copies</b>  | Specifies minimum number of copies of each logical volume.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

**-m** exact mapping.

## Example

To triply mirror a volume group, that is part of resource group resgrp1, enter:

```
cl_mirrorvg -cspoc"-g resgrp1" -c 3 workvg
```

**cl\_mkdev [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” -c disk -t <Disk Type> -s <Disk Type Technology> -m <NodeName:Parent Adapter .....>**

Creates a disk definition on specified cluster nodes.

**-cspoc** Argument used to specify one of the following C-SPOC options:

- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
- g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
- n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

**SCSI Disk Technology** -w <Connection Address>

**SSA Disk Technology** -a <location>|| -a pv = [yes|no] || -a reserve\_lock=[yes|no] ||  
-a queue\_depth= [0...200(+1)]  
-a max\_coalesce= [0x20000...0x200000 (+0x1000)]

**-m** Supply a comma separated list of node/disk parent adapter pairs. The node and its disk parent adapter will be separated by a colon.

All other flags listed belong to the AIX **mkdev** command and operate as outlined in the **mkdev** man pages.

## Examples

To create hdisk2 on nodes alpha and beta, members of resource group foo, enter:

```
SCSI: cl_mkdev -cspoc "-f -n alpha,beta" -c disk -t 1000mb -s scsi
 -p scsi0 -w 1,2
SSA: cl_mkdev -cspoc "-f -n alpha,beta" -c disk -c ssar -t hdisk
 -p ssar -w 1,2
```

**cl\_mkgroup [-cspoc “[-f] -g ResourceGroup | -n Nodelist]” [-a] [-A] [Attribute=Value ...] Group**

Creates a new group on all nodes in an HACMP cluster. If the cluster is running on SP machines and the **usermgt\_config** flag is set to true, the command exits.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc**                      Argument used to specify the following C-SPOC option:
  - f** – Forces **cl\_mkggroup** to skip default verification. If this flag is set and a cluster node is not accessible, **cl\_mkggroup** reports a warning and continues execution on the other cluster nodes. If this flag is set and the group already exists on a node, **cl\_mkggroup** reports a warning and continues execution on the other cluster nodes
  - .-g ResourceGroup** – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.
  - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

If neither **-g** nor **-n** specified, the command executes on all nodes.
- a**                              Creates an administrative group. Only a user with root permissions can use this flag.
- A**                              Sets the group administrator to the person who invoked the **cl\_mkggroup** command.
- Attribute=Value**            Initializes a group with a specific attribute. See the **chgroup** command for more information about the group attributes.

## Examples

### Example 1

To create a new group account called finance on all nodes in the cluster, enter:

```
cl_mkggroup -cspoc "-f" finance
```

### Example

To create a new administrative group account called payroll on all cluster nodes, enter:

```
cl_mkggroup -cspoc "-f" -a payroll
```

**Note:** Only a root user can issue this command.

**cl\_mklv -cspoc “[-c|-s] [-f] [-g ResourceGroup | -n NodeList]” [-a IntraPolicy][ -b BadBlocks] [-c Copies] [-d Schedule][ -e InterPolicy] [-i] [-L Label] [-m MapFile] [-r Relocate] [-s Strict] [-t Type] [-u UpperBound][ -v Verify] [-w MWC] [-x MaxLPs] [-y LVName][ -Y Prefix] [-S StripeSize] -V VolumeGroup NumberOfLPs [PhysicalVolume...]**

The **cl\_mklv** command creates a logical volume on a volume group in the HACMP cluster.

**Note:** When creating a logical volume on a concurrent volume group, the volume group must be varied on in concurrent mode on at least one node in the cluster.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
  - g *ResourceGroup*** – Generates the list of nodes participating in this resource group on which the command will be executed.
  - n *nodelist*** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.
- a *Position*** Sets the intra-physical volume allocation policy (the position of the logical partitions on the physical volume). The ***Position*** variable is represented by one of the following:
- m** – Allocates logical partitions in the outer middle section of each physical volume. This is the default position.
  - c** – Allocates logical partitions in the center section of each physical volume.
  - e** – Allocates logical partitions in the outer edge section of each physical volume.
  - ie** – Allocates logical partitions in the inner edge section of each physical volume.
  - im** – Allocates logical partitions in the inner middle section of each physical volume.
- b *BadBlocks*** Sets the bad-block relocation policy. The ***BadBlocks*** variable is represented by one of the following:
- y** – Causes bad-block relocation to occur.
  - n** – Prevents bad-block relocation from occurring.
- d *Schedule*** Sets the scheduling policy when more than one logical partition is written. The ***Schedule*** variable is represented by one of the following:
- p** – Establishes a parallel scheduling policy.
  - s** – Establishes a sequential scheduling policy.
- e *Range*** Sets the inter-physical volume allocation policy (the number of physical volumes to extend across, using the volumes that provide the best allocation). The value of the ***Range*** variable is limited by the ***Upperbound*** variable, set with the **-u** flag, and is represented by one of the following:
- x** – Allocates logical partitions across the maximum number of physical volumes.
  - m** – Allocates logical partitions across the minimum number of physical volumes.

- L *Label*** Sets the logical volume label. The maximum size of the *Label* variable is 127 characters.
- n *NewLogicalVolume*** Changes the name of the logical volume to that specified by the *NewLogicalVolume* variable. Logical volume names must be unique and system-wide, and can range from 1 to 15 characters.
- p *Permission*** Sets the access permission to read-write or read-only. The *Permission* variable is represented by one of the following:  
**w** – Sets the access permission to read-write.  
**r** – Sets the access permission to read-only.
- r *Relocate*** Sets the reorganization flag to allow or prevent the relocation of the logical volume during reorganization. The *Relocate* variable is represented by one of the following:  
**y** – Allows the logical volume to be relocated during reorganization.  
**n** – Prevents the logical volume from being relocated during reorganization.
- s *Strict*** Determines the strict allocation policy. Copies of a logical partition can be allocated to share or not to share the same physical volume. The *Strict* variable is represented by one of the following:  
**y** – Sets a strict allocation policy, so copies of a logical partition cannot share the same physical volume.  
**n** – Does not set a strict allocation policy, so copies of a logical partition can share the same physical volume.
- t *Type*** Sets the logical volume type. The maximum size is 31 characters.
- u *Upperbound*** Sets the maximum number of physical volumes for new allocation. The value of the *Upperbound* variable should be between one and the total number of physical volumes.
- v *Verify*** Sets the write-verify state for the logical volume. Causes all writes to the logical volume either to be verified with a follow-up read or not to be verified with a follow-up read. The *Verify* variable is represented by one of the following:  
**y** – Causes all writes to the logical volume to be verified with a follow-up read.  
**n** – Causes all writes to the logical volume not to be verified with a follow-up read.

**MirrorWriteConsistency y** – Turns on mirror write consistency which insures data consistency among mirrored copies of a logical volume during normal I/O processing.

**n** – No mirror write consistency. See the **-f** flag of the **syncvg** command.

**-x Maximum** – Sets the maximum number of logical partitions that can be allocated to the logical volume.

## Examples

1. To make a logical volume in volume group vg02 with one logical partition and a total of two copies of the data, enter:

```
cl_mklv -cspoc "-g resgrp1" -c 2 vg02 1
```

2. To make a logical volume in volume group vg03 with nine logical partitions and a total of three copies spread across a maximum of two physical volumes, and whose allocation policy is not strict, enter:

```
cl_mklv -cspoc "-g resgrp1" -c 3 -u 2 -s n vg03 9
```

**cl\_mklvcopy -cspoc “[-f][-g ResourceGroup | -n NodeList]” [-R RefNode] [-a IntraPolicy][-e InterPolicy] [-k] [-m MapFile] [-u UpperBound][-s Strict] -V LogicalVolume Copies [PhysicalVolume...]**

The **cl\_mklvcopy** command copies a logical volume. (Not supported on RAID devices.)

**Note:** When creating a copy of a logical volume on a concurrent volume group, the volume group must be varied on in concurrent mode on at least one node in the cluster.

### **-cspoc**

Argument used to specify one of the following C-SPOC options:

**-f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.

**-g ResourceGroup** – Generates the list of nodes participating in this resource group on which the command will be executed.

**-n nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

### **-R RefNode**

Node from which to retrieve hdisk names (used with Physical Volume option).

- a *Position*** Sets the intra-physical volume allocation policy (the position of the logical partitions on the physical volume). The ***Position*** variable is represented by one of the following:
- m** – Allocates logical partitions in the outer middle section of each physical volume. This is the default position.
  - c** – Allocates logical partitions in the center section of each physical volume.
  - e** – Allocates logical partitions in the outer edge section of each physical volume.
  - ie** – Allocates logical partitions in the inner edge section of each physical volume.
  - im** – Allocates logical partitions in the inner middle section of each physical volume.
- b *BadBlocks*** Sets the bad-block relocation policy. The ***BadBlocks*** variable is represented by one of the following:
- y** – Causes bad-block relocation to occur.
  - n** – Prevents bad-block relocation from occurring.
- d *Schedule*** Sets the scheduling policy when more than one logical partition is written. The ***Schedule*** variable is represented by one of the following:
- p** – Establishes a parallel scheduling policy.
  - s** – Establishes a sequential scheduling policy.
- e *Range*** Sets the inter-physical volume allocation policy (the number of physical volumes to extend across, using the volumes that provide the best allocation). The value of the ***Range*** variable is limited by the ***Upperbound*** variable, set with the **-u** flag, and is represented by one of the following:
- x** – Allocates logical partitions across the maximum number of physical volumes.
  - m** – Allocates logical partitions across the minimum number of physical volumes.
- L *Label*** Sets the logical volume label. The maximum size of the ***Label*** variable is 127 characters.
- n *NewLogicalVolume*** Changes the name of the logical volume to that specified by the ***NewLogicalVolume*** variable. Logical volume names must be unique and system-wide, and can range from 1 to 15 characters.
- p *Permission*** Sets the access permission to read-write or read-only. The ***Permission*** variable is represented by one of the following:
- w** – Sets the access permission to read-write.
  - r** – Sets the access permission to read-only.

- r Relocate** Sets the reorganization flag to allow or prevent the relocation of the logical volume during reorganization. The **Relocate** variable is represented by one of the following:
- y** – Allows the logical volume to be relocated during reorganization.
  - n** – Prevents the logical volume from being relocated during reorganization.
- s Strict** Determines the strict allocation policy. Copies of a logical partition can be allocated to share or not to share the same physical volume. The **Strict** variable is represented by one of the following:
- y** – Sets a strict allocation policy, so copies of a logical partition cannot share the same physical volume.
  - n** – Does not set a strict allocation policy, so copies of a logical partition can share the same physical volume.
- t Type** Sets the logical volume type. The maximum size is 31 characters.
- u Upperbound** Sets the maximum number of physical volumes for new allocation. The value of the **Upperbound** variable should be between one and the total number of physical volumes.
- v Verify** Sets the write-verify state for the logical volume. Causes all writes to the logical volume either to be verified with a follow-up read or not to be verified with a follow-up read. The **Verify** variable is represented by one of the following:
- y** – Causes all writes to the logical volume to be verified with a follow-up read.
  - n** – Causes all writes to the logical volume not to be verified with a follow-up read.
- MirrorWriteConsistency** **y** – Turns on mirror write consistency which insures data consistency among mirrored copies of a logical volume during normal I/O processing.
- n** – No mirror write consistency. See the **-f** flag of the **syncvg** command.
- x Maximum** – Sets the maximum number of logical partitions that can be allocated to the logical volume.

## Example

To add physical partitions to the logical partitions in the logical volume lv01 , so that a total of three copies exists for each logical partition, enter:

```
cl_mklvcopy -cspoc "-g resgrp1" lv01 3
```

## **cl\_mkuser [-cspoc “[-f] [-g *ResourceGroup* | -n *Nodelist*]”] [-a] [Attribute=Value ...] *Name***

Creates a new user account on all nodes in an HACMP cluster. If the cluster is running on SP machines and the **usermgt\_config** flag is set to true, the command exits.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

|                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b>          | Argument used to specify the following C-SPOC option:<br><b>-f</b> – Forces <b>cl_mkuser</b> to skip default verification. If the flag is set and a cluster node is not accessible, <b>cl_mkuser</b> reports a warning and continues execution on the other cluster nodes. If the flag is set and the user already exists on a node, <b>cl_mkuser</b> reports a warning and continues execution on the other cluster nodes.<br><b>-g <i>ResourceGroup</i></b> – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.<br><b>-n <i>Nodelist</i></b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-a</b>              | Specifies that the user is an administrator. Only the root user can use this flag or alter the attributes of an administrative user.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Attribute=Value</b> | Initializes a user attribute. Refer to the <b>chuser</b> command for the valid attributes and values.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

## **Examples**

### **Example 1**

To create the davis user account with the default values in the **/usr/lib/security/mkuser.default** file, enter:

```
cl_mkuser -cspoc "-f" davis
```

### **Example 2**

To create the davis account with davis as an administrator, enter:

```
cl_mkuser -cspoc "-f" -a davis
```

You must be the root user to create davis as an administrative user.

### **Example 3**

To create the davis user account and set the su attribute to a value of false, enter:

```
cl_mkuser -cspoc "-f" su=false davis
```

**cl\_mkvg [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” [-d maxPVs] [-B] [-G] [-f] [-c] [-x] [-i] [-s PPsize] [-n] [-m MaxSize | -t factor] [-V MajorNumber] [-y Vgname] PVname ...**

Creates a shared volume group for a specified list of cluster nodes

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

**-cspoc** Argument used to specify one of the following C-SPOC options:

- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
- g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
- n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

The remaining flags are AIX 4.3.2 supported flags for the **mkvg** command.

## Example

To create a shared volume group on hdisk2 for nodes alpha and beta, members of resource group foo, enter:

```
cl_mkvg -cspoc "-f -n alpha,beta" -f -y foo hdisk2 hdisk3
```

**cl\_mkvg [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” [-d maxPVs] [-B] [-G] [-f] [-c] [-x] [-i] [-s PPsize] [-n] [-m MaxSize | -t factor] [-V MajorNumber] [-y Vgname] PVname ...**

Creates a concurrent-capable shared volume group for a specified list of cluster nodes.

**-cspoc** Argument used to specify one of the following C-SPOC options:

- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
- g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
- n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

**-c** tells the system to make this volume group concurrent capable.

All other flags listed belong to the AIX **mkvg** command and operate as outlined in the **mkvg** man pages.

## Example

To create a concurrent volume group on hdisk2 for nodes alpha and beta, members of resource group foo, enter:

```
cl_mkvg -cspoc "-f -n alpha,beta" -c hdisk2
```

## **cl\_nodcmd [-q] [-cspoc "[-f] [-n *nodelist* | -g *resource group*]" ] *command* *args***

Runs a given command in parallel on a given set of nodes.

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-q</b>             | Specifies quiet mode. All standard output is suppressed.                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>-cspoc</b>         | Argument used to specify one of the following C-SPOC options:<br><b>-f</b> – Forces <b>cl_nodcmd</b> to skip HACMP version compatibility checks and node accessibility verification.<br><b>-g <i>ResourceGroup</i></b> – Generates the list of nodes participating in the resource group where the command will be executed.<br><b>-n <i>nodelist</i></b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b><i>command</i></b> | Specifies the command to be run on all nodes in the nodelist.                                                                                                                                                                                                                                                                                                                                                                                               |
| <b><i>args</i></b>    | Specifies any arguments required for use with the <b>cl_nodcmd</b> command.                                                                                                                                                                                                                                                                                                                                                                                 |

## Examples

### Example 1

```
cl_nodcmd lspv
```

Runs the **lspv** command on all cluster nodes.

### Example 2

```
cl_nodcmd -cspoc "-n beaver,dam" lsvg rootvg
```

Runs the **lsvg rootvg** command on nodes beaver and dam, suppressing standard output.

**cl\_rc.cluster [ -cspoc “[ -f ] [-g *ResourceGroup* | -n *NodeList* ]”] [ -boot ] [ -l ] [ -b ] [-i ] [-N | -R | -B]**

Sets up the operating system environment and starts the cluster daemons across cluster nodes.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify the following C-SPOC option:
- f** – Forces **cl\_rc.cluster** to skip HACMP version compatibility checks and node accessibility verification.
  - g *ResourceGroup*** – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.
  - n *NodeList*** – Executes underlying AIX commands across nodes in the node list.
- boot** Configures the service adapter to be on its boot address if IPAT is enabled.
- l** Starts the Cluster Lock Manager daemon (**clockd**) with its default option.
- i** Starts the Cluster Information (**clinfo**) daemon with its default options.
- b** Broadcasts the startup.
- N** Starts the daemons immediately (no **inittab** change).
- R** Starts the HACMP daemons on system restart only (the HACMP start up command is added to **inittab** file).
- B** Starts the daemons immediately and adds the HACMP entry to the **inittab** file.
- f** Forced startup. Cluster daemons should initialize running local procedures.

## Examples

### Example 1

To start the cluster with cluster lock services and to broadcast the event, enter:

```
cl_rc.cluster -cspoc "-n node1,node2" -boot -l -b
```

### Example 2

To start the cluster with **clinfo** running on all the cluster nodes, enter:

```
cl_rc.cluster -boot -i
```

## **cl\_reducevg -cspoc “[-f][-g ResourceGroup | -n NodeList]” -R RefNode [-d] [-f] VolumeGroup PhysicalVolume**

The `cl_reducevg` command removes a physical volume from a volume group.

**Note:** You can use the command while the volume group is in concurrent mode. However, if you run this command while the volume group is in concurrent mode and the end result is the deletion of the volume group, then the command will fail.

|                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b>     | <p><b>-f</b> – If no node with access has the volume group varied on, an appropriate node is chosen, the volume group is varied on and the command executes on that node. The volume group is varied off upon successful completion. Warnings are sent about nodes that cannot be reached.</p> <p><b>-g ResourceGroup</b> – Specifies the name of the resource group whose participating nodes share the volume group. The command executes on these nodes.</p> <p><b>-n Nodelist</b> – Executes underlying AIX commands across nodes in the node list.</p> <p>Note: If you do not specify either the <b>-g</b> or <b>-n</b> options, the default action occurs on all cluster nodes.</p> |
| <b>-R RefNode</b> | Name of node from which to retrieve hdisk names.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>-d</b>         | Deallocates existing logical volume partitions and then deletes resultant empty logical volumes from the specified physical volumes. User confirmation required due to danger of jeopardizing integrity of entire logical volume.                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>-f</b>         | Removes requirement for user confirmation.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

### Examples

- To remove physical volume `hdisk1` from volume group `vg01` where `vg01` is a member of `resgrp1` and `NodeA` is the reference node, enter:

```
cl_reducevg -cspoc "-g resgrp1" -R NodeA vg01 hdisk1
```

- To remove physical volume `hdisk1` and all residing logical volumes from volume group `vg01` without user confirmation, enter the following command. Attention: The `-d` flag automatically deletes all logical volume data before removing the physical volume.

```
cl_reducevg -cspoc "-resgrp1" -R NodeA -d -f vg01 hdisk1
```

The physical volume `hdisk1` and all residing logical volumes are removed.

## **cl\_rmdev [-cspoc “[-f] [-g ResourceGroup | -n NodeList]” -l <hdisk#> [-d]**

Removes a previously configured disk from a list of cluster nodes.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
  - g ResourceGroup** – Generates the list of nodes participating in the resource group where the command will be executed.
  - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

All other flags listed belong to the AIX **rmdev** command and operate as outlined in the **rmdev** man pages.

### **Example**

To remove hdisk2 on nodes alpha and beta, members of resource group foo, enter:

```
cl_rmdev -cspoc“-f -n alpha,beta -f” -l hdisk2 -d
```

## **cl\_rmfs -cspoc “[-f] [-g ResourceGroup | -n Nodelist]” FileSystem**

Removes a shared filesystem.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – Forces **cl\_rmfs** to check the accessibility of only the participating nodes in the specified resource group.
  - g ResourceGroup** – Specifies the name of the resource group whose participating nodes share the filesystem. The command executes on these nodes.
  - n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

**FileSystem** Specifies the shared filesystem to be removed.

## Examples

To remove the */test* filesystem, enter:

```
cl_rmfs -cspoc "-g resource_grp1" /test
```

This removes the */test* filesystem, its entry in the */etc/filesystems* file, and the underlying logical volume from the cluster node in *resource\_grp1* that currently has the */test* filesystem's volume group varied on.

## **cl\_rmggroup [-cspoc “[-f] [-g ResourceGroup | n- Nodelist]”] Name**

Removes a group from all cluster nodes.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

**-cspoc** Argument used to specify one of the following C-SPOC options:

- f** – Forces **cl\_rmggroup** to skip node accessibility verification. If this flag is set and a cluster node is not accessible, **cl\_rmggroup** reports a warning and continues execution on the other cluster nodes.
- g ResourceGroup** – Generates a list of nodes participating in the given resource group. The command executes on these nodes.
- n Nodelist** – Executes the command on nodes in the node list.

## Example

To remove the finance group from all cluster nodes, enter:

```
cl_rmggroup finance
```

## **cl\_rmlv -cspoc “[-f] [-g ResourceGroup | -n Nodelist]” [-f] [-p Physical Volume] LogicalVolume ...**

Removes a shared logical volume from a volume group.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

**-cspoc** Argument used to specify one of the following C-SPOC options:

- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
- g ResourceGroup** – Specifies the name of the resource group whose participating nodes share the logical volume. The command executes on these nodes.
- n Nodelist** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

**-f** Removes the logical volumes without requesting confirmation.

**LogicalVolume** Specifies the shared logical volume to be removed.

## Examples

To remove logical volume lv05 without requiring user confirmation, enter the following command.

```
cl_rmlv -cspoc "-g resource_grp1" -f lv05
```

**Warning:** This command destroys all data in the logical volumes, and the logical volume is removed from the volume group.

## **cl\_rmlvcopy -cspoc “[-f][-g ResourceGroup | -n NodeList]” [-R RefNode] LogicalVolume Copies [PhysicalVolume...]**

The **cl\_rmlvcopy** command removes a copy of a logical volume.

|                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b>          | Argument used to specify one of the following C-SPOC options:<br><b>-f</b> – If no node in accessibility list has VG varied on, choose one from list and vary it on, execute command, and vary off.<br><b>-g ResourceGroup</b> – Specifies the name of the resource group whose participating nodes share the logical volume. The command executes on these nodes.<br><b>-n Nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-R RefNode</b>      | Node from which to retrieve hdisk names.                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Logical Volume</b>  | Logical volume of which to remove copy. Name or ID.                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Copies</b>          | Maximum number of copies to leave.                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>Physical Volume</b> | Name or ID of physical volume. Copies only removed from this PV.                                                                                                                                                                                                                                                                                                                                                                                                                           |

## **cl\_rmuser [-cspoc “[-f] [-g ResourceGroup | -n Nodelist]” [-p] Name**

Removes a user account from all HACMP cluster nodes.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

|               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b> | Argument used to specify one of the following C-SPOC options:<br><b>-f</b> – Forces <b>cl_rmuser</b> to skip node accessibility verification. If this flag is set and a cluster node is not accessible, <b>cl_rmuser</b> reports a warning and continues execution on the other cluster nodes.<br><b>-g ResourceGroup</b> – Generates the list of nodes participating in the resource group where the command will be executed.<br><b>-n Nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.<br><b>-p</b> – Removes user password information from the <b>/etc/security/passwd</b> file. |
| <b>Name</b>   | The specified user name.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

### Examples

#### Example 1

To remove the user davis account and its attributes from all cluster nodes, enter:

```
cl_rmuser davis
```

#### Example 2

To remove the user davis account and all its attributes, including passwords and other user authentication information in the **/etc/security/passwd** file from all cluster nodes, enter:

```
cl_rmuser -p davis
```

## **cl\_syncvg -cspoc “[-f] [-g ResourceGroup | -n NodeList]” [-f] [-i] [-H] [-P NumParallelLps] {-l|-p|-v} Name...**

The **cl\_syncvg** command synchronizes any non-current logical volume copies.

|               |                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b> | <b>-f</b> – If no node in resource group has VG varied on, chooses one node in list, varies it on, executes command, varies it off. Issues warning for nodes that cannot be reached.<br><b>-g ResourceGroup</b> – Generates the list of nodes participating in the resource group where the command will be executed.<br><b>-n Nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-f</b>     | Specifies good physical copy is chosen and propagated to all, whether stale or not.                                                                                                                                                                                                                                                                                                                                                           |

|               |                                                                                              |
|---------------|----------------------------------------------------------------------------------------------|
| <b>-i</b>     | reads names from standard input.                                                             |
| <b>-H</b>     | Postpone writes for this volume group on other active concurrent nodes until sync completes. |
| <b>-P Num</b> | Number of logical partitions to sync in parallel. Valid range is 1 - 32.                     |
| <b>Name</b>   | -l logical volume name, -p physical volume name, -v volume group name.                       |

### Example

To synchronize the copies on volume groups vg04 and vg05, enter:

```
cl_syncvg -cspoc "-g resgrp1" -v vg04 vg05
```

### **cl\_unmirrorvg -cspoc “[-f][-g ResourceGroup |-n NodeList]” [-R RefNode][-c Copies] VolumeGroup [PhysicalVolume...]**

The **cl\_unmirrorvg** command removes the mirrors that exist in a volume group.

**Note:** When unmirroring a concurrent volume group, the volume group must be varied on in concurrent mode on at least one node in the cluster.

|                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-cspoc</b>          | Argument used to specify one of the following C-SPOC options:<br><b>-f</b> – If no node in accessibility list has VG varied on, choose one from list and vary it on, execute command, and vary off. Issues warning for nodes that cannot be reached.<br><b>-g ResourceGroup</b> – Specifies the name of the resource group whose participating nodes share the logical volume. The command executes on these nodes.<br><b>-n Nodelist</b> – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas. |
| <b>-R RefNode</b>      | Node from which to retrieve hdisk names.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Copies</b>          | Maximum number of copies to leave. Default is 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Physical Volume</b> | Name or ID of physical volume. Copies only removed from this PV.                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

### Example

To unmirror a triply mirrored volume group and leave two copies, enter:

```
cl_unmirrorvg -c 2 workvg
```

The logical partitions in the logical volumes held on workvg now have 2 copies.

## **cl\_updatevg -cspoc “[-f] [-g *ResourceGroup* | -n *Nodelist*]” *VolumeGroup***

The **cl\_updatevg** command exports and imports the shared volume group definition onto the nodes participating in the given resource group.

**Note:** Arguments associated with a particular flag must be specified immediately following the flag.

- cspoc** Argument used to specify one of the following C-SPOC options:
- f** – If no node has the VG varied on, the system chooses a node in the resource group, varies on the VG and executes the command there, and varies off the VG. A warning is issued for nodes that cannot be reached.
  - g *ResourceGroup*** – Generates the list of nodes participating in the resource group where the command will be executed.
  - n *Nodelist*** – Runs the command on this list of nodes. If more than one node, separate nodes listed by commas.

### **Examples**

To export and re-import a shared volume group definition on a local or remote node, enter:

```
cl_updatevg -cspoc "-f -g resource_grp1" /test
```

---

## **C-SPOC Shell Commands/Utilities**

These utilities perform operations on remote nodes in the cluster.

### **cllsconvg *VolumeGroup***

Determines if *VolumeGroup* is shared or concurrent.

### **climportvg [-V *MajorNumber*] [-y *VolumeGroup*] [-f] [-c] [-x] *PhysicalVolume***

Imports a volume group on the remote node. This script determines which type of import operation needs to be performed. It also updates the volume group timestamp in **/usr/sbin/cluster/etc/vg**.

- V *MajorNumber*** Major number of imported volume group.
- y *Volume Group*** Name for new volume group.
- c** Imports volume group and creates it as concurrent-capable.
- x** When used with **-c** flag, sets the volume group to be varied on in concurrent mode. Must have **rc\_clvmv** entry in **/etc/inittab** set.

### **clupdatevg *VolumeGroup PhysicalVolumeId***

Performs **importvg -L *VolumeGroup PhysicalVolume*** operation on the remote node. The *PhysicalVolumeId* is translated into the *hdiskx* name on the remote machine. This script also updates the volume group timestamp in **/usr/sbin/cluster/etc/vg**.

### **clsynsvg *VolumeGroup***

Performs the **synsvg** operation on the remote node, in addition to updating the volume group timestamp in the **/usr/sbin/cluster/etc/vg** directory.

### **clsynclvodm *VolumeGroup***

Performs the **synclvodm** operation on the remote node, in addition to updating the volume group timestamp in the **/usr/sbin/cluster/etc/vg** directory. This script is only used when performing an **extendvg** or **reducevg** operation, or when performing an LVM operation on a concurrent RAID device.

---

## **AIX Connections Shell Commands/Utilities**

The following HACMP utilities and commands handle information about AIX Connections:

### **cclsres**

Lists the NAME=VALUE pairs when doing a colon list.

### **clshowres**

Lists the NAME=VALUE pairs.

### **clsetenvres**

Initializes the AIX\_CONNECTIONS\_SERVICES NAME=VALUE pairs.

### **cclsacnrrs**

Retrieves all AIX Connections realms and services lists them in a realm/service format. In the service portion this script replaces the colon (:) with a percent sign (%).

### **clisacnncn -r realm -s service [-v volume\_ref | -p print\_ref | -j attach\_point]**

Sets the exit condition to 0 if service, volume reference, printer reference, or attach point is defined in AIX Connections. Otherwise it returns a non-zero value. **clisacnncn** returns no output. The HACMP verification tool runs this script remotely.

### **cclsacnncnrs [-g resource\_group] [-n]**

Retrieves a list of AIX Connections realm/service pairs shared by nodes in a resource group. A realm/service pair is considered shared if it is accessible by all participating nodes in a configured resource group. Note that a listed realm/service pair may or may not be configured as a resource in the resource group.

## CS/AIX Communications Links Commands/Utilities

The following HACMP utilities and commands handle information about CS/AIX Communications Links:

### **clshowres**

Lists the NAME=VALUE pairs.

### **cllssnaobj -n CS/AIX DLC name or 'All' [-p (port) ] [-l(linkstation)]**

List CS/AIX resources for a specified CS/AIX communications link.



---

## Appendix B Script Utilities

This appendix describes the utilities called by the event and startup scripts supplied with HACMP for AIX. These utilities are general-purpose tools that can be called from any script, or from the AIX command line. The examples assume they are called in a script.

---

### Utilities

The script utilities are stored in the `/usr/sbin/cluster/events/utls` directory. The utilities described in this chapter are grouped in the following categories:

- Disk utilities
- RS/6000 SP system utilities
- Filesystem and volume group utilities
- Logging utilities
- Network utilities
- Startup and stop utilities
- DARE utilities
- Emulation utilities
- Security utilities.

---

### Disk Utilities

#### `cl_9333clear`

**Syntax**

```
cl_9333clear pvid
```

**Description**

Clears the fence register for the IBM 9333 disk identified by the PVID, restoring it to its power-on state where any connected node has access to the disk. Use this utility only when all the cluster nodes that access the disk are down (that is, the HACMP for AIX software is not running).

**Warning:** This utility allows hosts not running the HACMP for AIX software to modify the data on the IBM 9333 disk, therefore exposing the integrity of that data to risk. It requires adequate operator controls and warnings and should not be included within any script.

### Parameters

**pvid** PVID of the disk whose fence register is to be cleared.

### Return Values

- 0** Success.
- 1** Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file.
- 2** Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.

### Example

```
cl_9333clear 0000045683922922
echo $?
0
```

## cl\_9333diskreset

### Syntax

```
cl_9333diskreset /dev/controller_special_file disk_number
```

### Description

Issues a reset (to break reservation) to each IBM 9333 disk named as an argument. Disk numbers are 0, 1, 2, or 3.

### Parameters

**controller\_special\_file** Controller to which disk is attached.

**disk\_number** Number assigned to the disk to be reset.

### Return Values

- 0** All specified disks have been reset.
- 1** No disks have been reset.

### Example

```
lsdev -Cc disk
hdisk Available 00-02-00-03 857MB F Serial Link Disk Drive
serdasdc0 Available 00-02-00 Serial Link Disk Controller
cl_9333diskreset /dev/serdasdc0 3
```

## cl\_9333fence

### Syntax

```
cl_9333fence -r|x|s pvid
```

### Description

The **cl\_9333fence** utility reads or sets the value of the fence register for the 9333 disk identified by the PVID.

If the **-r** flag is specified, the contents of the fence register for the disk identified by the PVID are written to stdout.

If the **-s** flag is specified, the contents of the fence register are evaluated and reflected in the exit code. If the fence register is in its default state (clear), then a 0 is returned. Otherwise, a 1 is returned. If an execution error occurs, the utility exits with a 2.

If the **-x** flag is specified, the fence register is updated from the existing membership to a new membership. The existing membership is determined by the logical AND of the nodes allowed access to the PVID by the HACMPfence entry in the ODM and the nodes listed in the PRE\_EVENT\_MEMBERSHIP environment variable. The new membership is determined by the logical AND of the nodes allowed access to the PVID by the HACMPfence entry in the ODM and the nodes listed in the POST\_EVENT\_MEMBERSHIP environment variable.

**Note:** For this command to succeed, PRE\_EVENT\_MEMBERSHIP or POST\_EVENT\_MEMBERSHIP must match the current fence register contents.

### Parameters

**pvid** PVID of the disk whose fence register is to be read or updated.

### Return Values

- 0** Success (for the **-r** and **-x** flags). Fence register is in clear (default) state (for the **-s** flag).
- 1** Failure (for the **-r** and **-x** flags). A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. Fence register is in non-default state (for the **-s** flag).
- 2** Failure. Invalid number of arguments (for the **-r** and **-x** flags). A message describing the problem is written to stderr and to the cluster log file. A problem occurred during execution (for the **-s** flag).

### Examples

```
echo $PRE_EVENT_MEMBERSHIP
caviar whale anchovy kippers

echo $POST_EVENT_MEMBERSHIP
caviar whale

cl_9333fence -r 0000045683922922
caviar,whale,anchovy,kippers
echo $?
0
cl_9333fence -x 0000045683922922
echo $?
0
```

## cl\_9333genodm

### Syntax

```
cl_9333genodm
```

### Description

This utility creates the HACMPfence ODM object used by the **cl\_9333fence** utility. The **cl\_9333genodm** utility is only executed when the cluster is configured or a change is made to the configuration. The utility is executed when the user synchronizes the node environment.

The **cl\_9333genodm** utility first searches the ODM on all nodes for all PVIDs associated with 9333 disks. Then, for each node, a request is sent to get the bit position that node occupies for each PVID it can access. Once this information is obtained for all nodes, an ODM entry object is created to hold this information. During script execution, only objects participating in fencing are updated.

Note that the command sends a message about all the 9333 disks it finds, including those not using fencing.

### Parameters

None.

### Return Values

|          |                                                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------|
| <b>0</b> | Success.                                                                                                                         |
| <b>1</b> | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |
| <b>2</b> | Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.         |

### Example

```
cl_9333genodm
echo $?
0
odmget -q "pvid = 00000275175134fa" HACMPfence
caviar:0,whale:3,anchovy:5,kippers:2
```

## cl\_9333getbit

### Syntax

```
cl_9333getbit pvid
```

### Description

This utility is only used when synchronizing the node environment. **cl\_9333getbit** takes a PVID for a physical volume installed on an IBM 9333 serial disk and writes to stdout the bit position the node it is running on occupies in the fence register for that disk. This utility is only called by **cl\_9333genodm** during cluster configuration.

### Parameters

**pvid** PVID of the disk.

### Return Values

- |          |                                                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------|
| <b>0</b> | Success.                                                                                                                         |
| <b>1</b> | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |
| <b>2</b> | Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.         |

### Example

```
cl_9333getbit 0000045683922922
3
echo $?
0
```

## cl\_array\_mode3

### Syntax

```
cl_array_mode3 volume_group ...
```

### Description

Varies on a concurrent RAID volume group and handles calls to **convaryonvg** for RAID devices.

### Parameters

**Volume\_group** List of one or more RAID volume groups.

### Return Values

- |          |                                                                   |
|----------|-------------------------------------------------------------------|
| <b>0</b> | All volume groups are varied on. Modes have changed.              |
| <b>1</b> | The <b>varyonvg</b> command failed for at least one volume group. |
| <b>2</b> | Incorrect or bad arguments were used.                             |

## cl\_disk\_available

### Syntax

```
cl_disk_available diskname ...
```

### Description

Checks to see if a disk named as an argument is currently available to the system, and, if not, makes the disk available.

### Parameters

**diskname** List of one or more disks to be made available; for example, *hdisk1*.

### Return Values

**0** Successfully made the specified disk available.

**1** Failed to make the specified disk available.

**2** Incorrect or bad arguments were used.

## cl\_fs2disk

### Syntax

```
cl_fs2disk [-lvip] mount_point
```

or

```
cl_fs2disk -g volume_group
```

where **-l** identifies and returns the logical volume, **-v** returns the volume group, **-i** returns the physical volume ID, **-p** returns the physical volume, and **-g** is the mount point of a filesystem (given a volume group).

### Description

Checks the ODM for the specified logical volume, volume group, physical volume ID, and physical volume information.

### Parameters

**mount point** Mount point of filesystem to check.

**volume group** Volume group to check.

### Return Values

**0** Successfully retrieved filesystem information.

**1** Failed to retrieve filesystem information.

## cl\_get\_disk\_vg\_fs\_pvids

### Syntax

```
cl_get_disk_vg_fs_pvids [filesystem_list volumegroup_list]
```

### Description

Given filesystems and/or volume groups, the function returns a list of the associated PVIDs.

### Parameters

|                          |                             |
|--------------------------|-----------------------------|
| <b>filesystem_list</b>   | The filesystems to check.   |
| <b>volume_group_list</b> | The volume groups to check. |

### Return Values

|          |                    |
|----------|--------------------|
| <b>0</b> | Success.           |
| <b>1</b> | Failure.           |
| <b>2</b> | Invalid arguments. |

## cl\_is\_array

### Syntax

```
cl_is_array diskname
```

### Description

Checks to see if a disk is a READI disk array.

### Parameters

|                 |                                                   |
|-----------------|---------------------------------------------------|
| <b>diskname</b> | Single disk to test; for example, <i>hdisk1</i> . |
|-----------------|---------------------------------------------------|

### Return Values

|          |                                 |
|----------|---------------------------------|
| <b>0</b> | Disk is a READI disk array.     |
| <b>1</b> | Disk is not a READI disk array. |
| <b>2</b> | An error occurred.              |

## cl\_is\_scsidisk

### Syntax

```
cl_is_scsidisk diskname
```

### Description

Determines if a disk is a SCSI disk.

### Parameters

**diskname** Single disk to test; for example, *hdisk1*.

### Return Values

**0** Disk is a SCSI disk.  
**1** Disk is not a SCSI disk.  
**2** An error occurred.

## cl\_raid\_vg

### Syntax

```
cl_raid_vg volume_group
```

### Description

Checks to see if the volume group is comprised of RAID disk arrays.

### Parameters

**volume\_group** Single volume group to check.

### Return Values

**0** Successfully identified a RAID volume group.  
**1** Could not identify a RAID volume group; volume group must be 9333.  
**2** An error occurred. Mixed volume group identified.

## cl\_scdiskreset

### Syntax

```
cl_scdiskreset /dev/diskname ...
```

### Description

Issues a reset (SCSI ioctl) to each SCSI disk named as an argument.

### Parameters

**/dev/diskname** List of one or more SCSI disks.

### Return Values

**0** All specified disks have been reset.

**-1** No disks have been reset.

**n** Number of disks successfully reset.

## cl\_scdiskrsrv

### Syntax

```
cl_scsidiskrsrv /dev/diskname ...
```

### Description

Reserves the specified SCSI disk.

### Parameters

**/dev/diskname** List of one or more SCSI disks.

### Return Values

**0** All specified disks have been reserved.

**-1** No disks have been reserved.

**n** Number of disks successfully reserved.

## cl\_sync\_vgs

### Syntax

```
cl_sync_vgs -b|f volume_group ...
```

### Description

Attempts to synchronize a volume group by calling **syncvg** for the specified volume group.

### Parameters

|                     |                    |
|---------------------|--------------------|
| <b>volume_group</b> | Volume group list. |
| <b>-b</b>           | Background sync.   |
| <b>-f</b>           | Foreground sync.   |

### Return Values

|          |                                                                          |
|----------|--------------------------------------------------------------------------|
| <b>0</b> | Successfully started <b>syncvg</b> for all specified volume groups.      |
| <b>1</b> | The <b>syncvg</b> of at least one of the specified volume groups failed. |
| <b>2</b> | No arguments were passed.                                                |

## convaryonvg

### Syntax

```
convaryonvg volume_group
```

### Description

Varies on a volume group in concurrent mode. This service applies to RAID disk arrays being used in a concurrent access configuration. An initial check of the volume group checks to see if it is reserved.

### Parameters

|                     |                                               |
|---------------------|-----------------------------------------------|
| <b>volume_group</b> | Single volume group comprised of RAID arrays. |
|---------------------|-----------------------------------------------|

### Return Values

|          |                                                                          |
|----------|--------------------------------------------------------------------------|
| <b>0</b> | Successfully varied on the specified volume group for concurrent access. |
| <b>1</b> | Could not vary on the specified volume group for concurrent access.      |

## scdiskutil

### Syntax

```
scdiskutil -t /dev/diskname ...
```

### Description

Tests and clears any pending SCSI disk status.

### Parameters

|                      |                                  |
|----------------------|----------------------------------|
| <b>-t</b>            | Tests to see if a unit is ready. |
| <b>/dev/diskname</b> | Single SCSI disk.                |

### Return Values

|              |                           |
|--------------|---------------------------|
| <b>-1</b>    | An error occurred.        |
| <b>0</b>     | The disk is not reserved. |
| <b>&gt;0</b> | The disk is reserved.     |
| <b>2</b>     | No arguments were passed. |

## ssa\_fence

### Syntax

```
ssa_fence -e event pvid
```

### Description

Fences a node in or out.

Additionally, this command also relies on environment variables; the first node up fences out all other nodes of the cluster regardless of their participation in the resource group.

If it is not the first node up, then the remote nodes fence in the node coming up. The node joining the cluster will not do anything.

If it is a **node\_down** event, the remote nodes will fence out the node that is leaving. The node leaving the cluster will not do anything.

The last node going down clears the fence register.

### Environment Variables

|                       |                         |
|-----------------------|-------------------------|
| PRE_EVENT_MEMBERSHIP  | Set by cluster manager. |
| POST_EVENT_MEMBERSHIP | Set by cluster manager. |
| EVENT_ON_NODE         | Set by calling script.  |

### Parameters

|                 |                                                 |
|-----------------|-------------------------------------------------|
| <b>-e event</b> | 1=up; 2=down.                                   |
| <b>pvid</b>     | Physical volume ID on which fencing will occur. |

### Return Values

|          |                                                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------|
| <b>0</b> | Success.                                                                                                                         |
| <b>1</b> | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |
| <b>2</b> | Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.         |

## ssa\_clear

### Syntax

```
ssa_clear -x | -d pvid
```

### Description

Clears or displays the contents of the fence register. If **-d** is used, a list of fenced out nodes will be displayed. If **-x** is used, the fence register will be cleared.

**Note:** This command exposes data integrity of a disk, by unconditionally clearing its fencing register. It requires adequate operator controls and warnings, and should *not* be included within any takeover script.

### Return Values

|          |                                                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------|
| <b>0</b> | Success.                                                                                                                         |
| <b>1</b> | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |
| <b>2</b> | Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.         |

## ssa\_clear\_all

### Syntax

```
ssa_clear_all pvid1, pvid2 ... pvidn
```

### Description

Clears the fence register on multiple physical volumes.

### Return Values

- |   |                                                                                                                                  |
|---|----------------------------------------------------------------------------------------------------------------------------------|
| 0 | Success.                                                                                                                         |
| 1 | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |
| 2 | Failure. Invalid number of arguments. A message describing the problem is written to stderr and to the cluster log file.         |

## ssa\_configure

### Syntax

```
ssa_configure
```

### Description

Assigns unique node IDs to all the nodes of the cluster. Then it configures and unconfigures all SSA pdisks and hdisks on *all* nodes thus activating SSA fencing. This command is called from the SMIT screen during the sync of a node environment. If this command fails for any reason, that node should be rebooted.

### Return Values

- |   |                                                                                                                                  |
|---|----------------------------------------------------------------------------------------------------------------------------------|
| 0 | Success.                                                                                                                         |
| 1 | Failure. A problem occurred during execution. A message describing the problem is written to stderr and to the cluster log file. |

---

## RS/6000 SP Utilities

### cl\_HPS\_Eprimary

#### Syntax

```
cl_HPS_Eprimary manage | unmanage
```

or

```
cl_HPS_Eprimary event [nodename]
```

#### Description

This script has two distinct uses.

In its first form, you can use this utility to specify that you want HACMP for AIX to manage the Eprimary node in an IBM RS/6000 SP system. The script sets the value of HPS\_CONTROL entry in the HACMPsp2 ODM database to true. When used in this way, the script must be called manually before HACMP for AIX is started.

In its second form, this utility is called in HACMP scripts to update local and remote HACMPsp2 ODM entries with Eprimary data. Assigns unique node IDs to all the nodes of the cluster. Then it configures and unconfigures all SSA pdisks and hdisks on *all* nodes thus activating SSA fencing. This command is called from the SMIT screen during the sync of a node environment. If this command fails for any reason, that node should be rebooted.

In addition, you can optionally use this script to change the Eprimary node in the SP.

### Parameters

|                 |                                                                                                                                   |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------|
| <b>manage</b>   | Configures the HACMP cluster for Eprimary management.                                                                             |
| <b>unmanage</b> | Unconfigures the HACMP cluster for Eprimary management.                                                                           |
| <b>event</b>    | One of the following HACMP events: <b>node_up_local</b> ;<br><b>node_down_remote_complete</b> ; <b>node_down_local_complete</b> . |
| <b>nodename</b> | Optionally, specifies the name of the new EPrimary node                                                                           |

### Return Values

|          |                 |
|----------|-----------------|
| <b>0</b> | Success.        |
| <b>1</b> | Invalid Syntax. |

## cl\_Eprimary\_app

### Syntax

```
cl_Eprimary_app -d | -u
```

### Description

This script is called as part of HACMP for AIX processing in response to a `node_up_local_complete` event (with the `-u` flag specified) or a `node_down_remote_complete` event (with the `-d` flag specified).

When a remote node fails, the HACMP for AIX event script that runs in response to this event calls this script to determine if the failed node was the Eprimary node. When a node comes up, the HACMP for AIX event script that runs in response to this event must determine if it is the Eprimary node or if there is already an Eprimary node.

### Parameters

|           |                                                              |
|-----------|--------------------------------------------------------------|
| <b>-d</b> | Specifies <b>node_down_remote_complete</b> event processing. |
| <b>-u</b> | Specifies <b>node_up_local_complete</b> event processing.    |

### Return Values

|          |                 |
|----------|-----------------|
| <b>1</b> | Failure.        |
| <b>2</b> | Invalid syntax. |

## cl\_HPS\_init

### Syntax

```
cl_HPS_init -s | -t
```

### Description

This script has two distinct uses.

This script calls the **Estart** command to start the SP Switch network. This script should be called as part of HACMP for AIX processing when a node comes up for the first time or when a failed node reintegrates after a failover.

### Parameters

- s** Specifies when called as part of **node\_up\_local\_complete** event processing.
- t** Specified when called as part of **node\_down\_remote\_complete** event processing.

### Return Values

- 0** Success.
- 1** Failure.
- 2** Invalid syntax.

## cl\_reassign\_Eprimary

### Syntax

```
cl_reassign_Eprimary service_address
```

### Description

This script ensures that the SP Switch Eprimary node is reassigned in the System Data Repository (SDR). (The SP configuration information is stored in the SDR.) It runs in the background, once a minute, until it successfully retrieves information from the SDR.

### Parameters

- service\_address** The address of the node you want to make Eprimary.

### Return Values

- 0** Success.
- 1** Failure.
- 2** Invalid syntax.

## cl\_swap\_HPS\_IP\_address

### Syntax

```
cl_swap_HPS_IP_address interface address netmask [delete]
```

### Description

This script is used to specify an alias address to an SP Switch interface, or remove an alias address, during IP address takeover. Note that adapter swapping does not make sense for the SP Switch since all addresses are alias addresses on the same adapter.

### Parameters

|                  |                                                         |
|------------------|---------------------------------------------------------|
| <b>interface</b> | The name of the interface.                              |
| <b>alias</b>     | The alias address.                                      |
| <b>netmask</b>   | Netmask.                                                |
| <b>delete</b>    | When this parameter is specified, the alias is removed. |

### Return Values

|          |                                                                                                   |
|----------|---------------------------------------------------------------------------------------------------|
| <b>0</b> | Success.                                                                                          |
| <b>1</b> | The adapter could not be configured (using the <b>ifconfig</b> command) at the specified address. |
| <b>2</b> | Invalid syntax.                                                                                   |

### Examples

The following example specifies the alias *1.1.1.1* for the *css0* interface.

```
cl_swap_HPS_IP_address css0 1.1.1.1 255.255.255.128
```

The following example removes the alias *1.1.1.1* from the *css0* interface.

```
cl_swap_HPS_IP_address css0 1.1.1.1 255.255.255.128 delete
```

---

## Filesystem and Volume Group Utilities

### cl\_activate\_fs

#### Syntax

```
cl_activate_fs /filesystem_mount_point
```

#### Description

Mounts the filesystems passed as arguments.

#### Parameters

**/filesystem\_mount\_point** A list of one or more filesystems to mount.

#### Return Values

|          |                                                                                              |
|----------|----------------------------------------------------------------------------------------------|
| <b>0</b> | All filesystems named as arguments were either already mounted or were successfully mounted. |
| <b>1</b> | One or more filesystems failed to <b>mount</b> .                                             |
| <b>2</b> | No arguments were passed.                                                                    |

### cl\_activate\_nfs

#### Syntax

```
cl_activate_nfs retry host /filesystem_mount_point
```

### Description

NFS-mounts the filesystems passed as arguments. The routine backgrounds the mounts rather than specifying the **bg** option to the **mount** command because the standard **mount** command fails if the server host is not ready (that is, has not exported the filesystem). This command will retry in the background. Note that this route assumes the filesystem is already mounted if any mounted filesystem has a matching name.

### Parameters

**retry** Number of attempts. Sleeps 15 seconds between attempts.

**host** NFS server host.

**/filesystem\_mount\_point** List of one or more filesystems to activate.

### Return Values

**0** All filesystems passed were either mounted or mounts were scheduled in the background.

**1** One or more filesystems failed to mount.

**2** No arguments were passed.

## cl\_activate\_vgs

### Syntax

```
cl_activate_vgs [-n] volume_group_to_activate ...
```

### Description

Initiates a **varyonvg** of the volume groups passed as arguments.

### Parameters

**-n** Do not sync the volume group when **varyon** is called.

**volume\_group\_to\_activate** List of one of more volume groups to activate.

### Return Values

**0** All of the volume groups are successfully varied on.

**1** The **varyonvg** of at least one volume group failed.

**2** No arguments were passed.

## cl\_deactivate\_fs

### Syntax

```
cl_deactivate_fs /filesystem_mount_point
```

### Description

Attempts to **unmount** any filesystem passed as an argument that is currently mounted.

### Parameters

**/filesystem\_mount\_point** List of one or more filesystems to unmount.

### Return Values

**0** All filesystems were successfully unmounted.  
**1** One or more filesystems failed to unmount.  
**2** No arguments were passed.

## **cl\_deactivate\_nfs**

### Syntax

```
cl_deactivate_nfs file_system_to_deactivate ...
```

### Description

Attempts to **unmount -f** any filesystem passed as an argument that is currently mounted.

### Parameters

**file\_system\_to\_deactivate** List of one or more NFS-mounted filesystems to unmount.

### Return Values

**0** Successfully unmounted a specified filesystem.  
**2** No arguments were passed.

## **cl\_deactivate\_vgs**

### Syntax

```
cl_deactivate_vgs volume_group_to_deactivate ...
```

### Description

Initiates a **varyoffvg** of any volume group that is currently varied on and that was passed as an argument.

### Parameters

**volume\_group\_to\_deactivate** List of one or more volume groups to vary off.

### Return Values

**0** All of the volume groups are successfully varied off.  
**1** The **varyoffvg** of at least one volume group failed.  
**2** No arguments were passed.

## **cl\_export\_fs**

### Syntax

```
cl_export_fs hostname file_system_to_export ...
```

### Description

NFS-exports the filesystems given as arguments so that NFS clients can continue to work.

### Parameters

**hostname** Hostname of host given root access.

**filesystem\_to\_export** List of one or more filesystems to NFS-export.

### Return Values

**0** Successfully exported all filesystems specified.

**1** A runtime error occurred: unable to export or unable to startsrc failures.

**2** No arguments were passed.

## cl\_nfskill

### Syntax

```
cl_nfskill [-k] [-t] [-u] directory ...
```

### Description

Lists the process numbers of local processes using the specified NFS directory.

Find and kill processes that are executables fetched from the NFS-mounted filesystem. Only the root user can kill a process of another user.

If you specify the **-t** flag, all processes that have certain NFS module names within their stack will be killed.

**Warning:** When using the **-t** flag it is not possible to tell which NFS filesystem the process is related to. This could result in killing processes which belong to NFS-mounted filesystems other than those which are cross-mounted from another HACMP node and under HACMP control. This could also mean that the processes found could be related to filesystems under HACMP control but not part of the current resources being taken. This flag should therefore be used with caution and only if you know you have a specific problem with unmounting the NFS filesystems.

To help to control this, the **cl\_deactivate\_nfs** script contains the normal calls to **cl\_nfskill** with the **-k** and **-u** flags and commented calls using the **-t** flag as well. If you choose to use the **-t** flag, you should uncomment those calls and comment the original calls.

### Parameters

|                  |                                                                                      |
|------------------|--------------------------------------------------------------------------------------|
| <b>-k</b>        | Sends the SIGKILL signal to each local process,                                      |
| <b>-u</b>        | Provides the login name for local processes in parentheses after the process number. |
| <b>-t</b>        | Finds and kills processes that are just opening on NFS filesystems.                  |
| <i>directory</i> | Lists of one or more NFS directories to check.                                       |

### Return Values

None.

---

## Logging Utilities

### cl\_log

#### Syntax

`cl_log message_id, default_message, variables`

#### Description

Logs messages to **syslog** and standard error.

#### Parameters

|                        |                                             |
|------------------------|---------------------------------------------|
| <b>message_id</b>      | Message ID for the messages to be logged.   |
| <b>default_message</b> | Default message to be logged.               |
| <b>variables</b>       | List of one or more variables to be logged. |

#### Return Values

|          |                                                                   |
|----------|-------------------------------------------------------------------|
| <b>0</b> | Successfully logged messages to <b>syslog</b> and standard error. |
| <b>2</b> | No arguments were passed.                                         |

---

## Network Utilities

### cl\_echo

#### Syntax

`cl_echo message_id, default_message, variables`

#### Description

Logs messages to standard error.

### Parameters

|                        |                                                |
|------------------------|------------------------------------------------|
| <b>message_id</b>      | Message ID for the messages to be displayed.   |
| <b>default_message</b> | Default message to be displayed.               |
| <b>variables</b>       | List of one or more variables to be displayed. |

### Return Values

|          |                                            |
|----------|--------------------------------------------|
| <b>0</b> | Successfully displayed messages to stdout. |
| <b>2</b> | No arguments were passed.                  |

## cl\_nm\_nis\_off

### Syntax

```
cl_nm_nis_off
```

### Description

Turns off name serving and NIS.

### Parameters

None

### Return Values

|          |                                                               |
|----------|---------------------------------------------------------------|
| <b>0</b> | Successfully turned off name serving and NIS client services. |
| <b>1</b> | Could not turn off name serving, or could not turn off NIS.   |

## cl\_nm\_nis\_on

### Syntax

```
cl_nm_nis_on
```

### Description

Turns on name serving and NIS.

### Parameters

None.

### Return Values

|          |                                                              |
|----------|--------------------------------------------------------------|
| <b>0</b> | Successfully turned on name serving and NIS client services. |
| <b>1</b> | Could not turn on name serving, NIS, or both.                |

## cl\_swap\_HW\_address

### Syntax

```
cl_swap_HW_address address interface
```

### Description

Checks to see if an alternate hardware address is specified for the address passed as the first argument. If so, it assigns the hardware address specified to the new adapter.

### Parameters

|                  |                                                          |
|------------------|----------------------------------------------------------|
| <b>address</b>   | Interface address or IP label.                           |
| <b>interface</b> | Interface name (for example, <i>en0</i> or <i>tr0</i> ). |

### Return Values

|          |                                                                     |
|----------|---------------------------------------------------------------------|
| <b>0</b> | Successfully assigned the specified hardware address to an adapter. |
| <b>1</b> | Could not assign the specified hardware address to an adapter.      |
| <b>2</b> | Wrong number of arguments were passed.                              |

**Note:** This utility is used during adapter swap and IP address takeover.

## cl\_swap\_IP\_address

### Syntax

```
cl_swap_IP_address interface address netmask
cl_swap_IP_address interface1 address1 interface2 address2 netmask
```

### Description

This routine is used during adapter swap and IP address takeover.

In the first form, the routine sets the specified interface to the specified address:

```
cl_swap_IP_address en0 1.1.1.1 255.255.255.128
```

In the second form, the routine sets two interfaces in a single call. This is sometimes necessary due to AIX routing quirks where an existing route gets spuriously deleted. An example where this is required is the case of swapping two interfaces:

```
cl_swap_IP_address en0 1.1.1.1 en1 2.2.2.2 255.255.255.128
```

### Parameters

|                  |                                          |
|------------------|------------------------------------------|
| <b>interface</b> | Interface name.                          |
| <b>address</b>   | IP address.                              |
| <b>netmask</b>   | Network mask. Must be in decimal format. |

### Return Values

|          |                                         |
|----------|-----------------------------------------|
| <b>0</b> | Successfully swapped IP addresses.      |
| <b>1</b> | <b>ifconfig</b> failed.                 |
| <b>2</b> | Wrong or incorrect number of arguments. |

**Note:** This utility is used for swapping the IP address of either a standby adapter with a local service adapter (called adapter swapping), or a standby adapter with a remote service adapter (called masquerading). For masquerading, the **cl\_swap\_IP\_address** routine should sometimes be called *before* processes are stopped, and sometimes *after* processes are stopped. This is application dependent. Some applications respond better if they shutdown *before* the network connection is broken, and some respond better if the network connection is closed first.

## cl\_unswap\_HW\_address

### Syntax

```
cl_unswap_HW_address interface
```

### Description

Script used during adapter swap and IP address takeover. It restores an adapter to its boot address.

### Parameters

|                  |                                                          |
|------------------|----------------------------------------------------------|
| <b>interface</b> | Interface name (for example, <i>en0</i> or <i>tr0</i> ). |
|------------------|----------------------------------------------------------|

### Return Values

|          |                     |
|----------|---------------------|
| <b>0</b> | Success.            |
| <b>1</b> | Failure.            |
| <b>2</b> | Invalid parameters. |

## Startup and Stop Utilities

### start\_clmarketdemo

**Syntax**

```
start_clmarketdemo
```

**Description**

Starts the Cluster Market Server (CLMarket) demo.

**Parameters**

**-d filename** Database file.

**-l filename** Log file.

**Return Values**

- |          |                                                                            |
|----------|----------------------------------------------------------------------------|
| <b>0</b> | Successfully started the CLMarket demo, or it is already running.          |
| <b>1</b> | Could not start the CLMarket demo because the lock manager is not running. |
| <b>2</b> | The required file to start CLMarket does not exist.                        |

### start\_imagedemo

**Syntax**

```
start_imagedemo [image_directory]
```

**Description**

Starts the Image Cataloger (Image Server) demo.

**Parameters**

**image\_directory** Optional image file directory.

**Return Values**

- |          |                                                                                   |
|----------|-----------------------------------------------------------------------------------|
| <b>0</b> | Successfully started the Image Cataloger demo, or it is already running.          |
| <b>1</b> | Could not start the Image Cataloger demo because the lock manager is not running. |
| <b>2</b> | The given image path does not exist.                                              |

### stop\_clmarketdemo

**Syntax**

```
stop_clmarketdemo
```

**Description**

Stops the CLMarket demo.

#### Parameters

None.

#### Return Values

**0** Successfully stopped the CLMarket demo, or the **market server** process cannot be found.

## stop\_imagedemo

#### Syntax

```
stop_imagedemo
```

#### Description

Stops the Image Cataloger demo.

#### Parameters

None.

#### Return Values

**0** Successfully stopped the Image Cataloger demo, or the **imserv** process cannot be found.

---

## DARE Utilities

### cldare

#### Syntax

```
cldare [-u]
```

```
cldare [-n] [-i] [-r] [-t] [-v] [-f] [-x][-M resgroup:[location|
[stop|default]] [:sticky] ...
```

Note that the -u flag cannot be run with any other flag or command.

#### Description

Updates the HACMP Cluster Manager and associated daemons with new configuration information, provides an emulation feature to emulate cluster resource and cluster topology dynamic reconfiguration, allows resource group migration to other cluster nodes while the cluster is running.

**Note:** When run in emulation mode (using the -f flag), cldare makes no changes to the configuration information.

## Parameters

- n** If this flag is specified, cluster resources removed from the configuration will keep their configuration as part of the DARE event. The default behavior is for the resources not to keep their configuration.
- u** If this flag is specified, the DARE will remove the stage ODM lock on all nodes. This may be used in case of a failure to remove the lock. This flag is exclusive—it cannot be run with any other flags or commands in `cl dare`.
- i** If this flag is specified, the DARE event will be performed whether or not the cluster verification finds configuration errors.
- r** If this flag is specified, Cluster Resources will be synchronized across cluster nodes.
- t** If this flag is specified, Cluster Topology will be synchronized across cluster nodes.
- v** If this flag is specified, verification will not run.  
Use this flag when the cluster is active or not active.
- f** If this flag is specified, the **cl dare** command will run as an emulation only and no changes will be made to the Cluster Manager. You must specify either the **-r** flag to indicate the emulation of Synchronizing Cluster Resources, or the **-t** flag to indicate the emulation of Synchronizing Cluster Topology. The **-u** flag and the **-v** flag cannot be used if you specify emulation mode.
- x** If this flag is specified, cluster verification will not be performed during synchronization of cluster resources or topology.  
Use this flag only when the cluster is not active.
- M resgroup: [location | [stop | default]] [:sticky]** Permits one or more pairs of resource groups and nodes to be specified on the command line to perform a dynamic resource group migration to another node. The **stop** keyword in place of a node location causes the utility to attempt to stop the resource group, taking down any service label, unmounting filesystems, etc. The use of the **default** keyword instead of a location causes the resource group to be started/moved to a cluster node based on the node priority.  
  
The keyword **sticky** to specify migration type is optional. For more information on sticky and non-sticky migration, see *Migrating Resources Dynamically—Overview* on page 7-17 of this book.

---

## Emulation Utilities

### cl\_emulate

#### Syntax

```
cl_emulate -e node_up -n nodename
cl_emulate -e node_down -n nodename {f|g|t}
cl_emulate -e network_up -w networkname -n nodename
cl_emulate -e network_down -w networkname -n nodename
cl_emulate -e join_standby -n nodename -a ip_label
cl_emulate -e fail_standby -n nodename -a ip_label
cl_emulate -e swap_adapter -n nodename -w network -a ip_label
-d ip_label
```

#### Description

Emulates a specific cluster event and outputs the result of the emulation. The output is shown on the screen as the emulation runs, and is saved to an output file on the node from which the emulation was executed.

The Event Emulation utility does not run customized scripts such as pre- and post- event scripts. In the output file the script is echoed and the syntax is checked, so you can predict possible errors in the script. However, if customized scripts exist, the outcome of running the actual event may differ from the outcome of the emulation

When emulating an event which contains a customized script, the Event Emulator uses the **ksh** flags **-n** and **-v**. The **-n** flag reads commands and checks them for syntax errors, but does not execute them. The **-v** flag indicates verbose mode. When writing customized scripts that may be accessed during an emulation, be aware that the other **ksh** flags may not be compatible with the **-n** flag and may cause unpredictable results during the emulation. See the **ksh** man page for flag descriptions.

You can run only one instance of an event emulation at a time. If you attempt to start an emulation while an emulation is already running on a cluster, the integrity of the output cannot be guaranteed.

## Parameters

|                              |                                                                                                                                                                                   |
|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-e</b> <i>eventname</i>   | The name of the event to emulate: <b>node_up</b> , <b>node_down</b> , <b>network_up</b> , <b>network_down</b> , <b>join standby</b> , <b>fail standby</b> , <b>swap adapter</b> . |
| <b>-n</b> <i>nodename</i>    | The node name used in the emulation.                                                                                                                                              |
| <b>-f</b>                    | Forced shut down emulation. Cluster daemons terminate without running any local procedures.                                                                                       |
| <b>-g</b>                    | Graceful shutdown with no takeover emulation.                                                                                                                                     |
| <b>-t</b>                    | Graceful shutdown emulation with the resources being released by this node and taken over by another node.                                                                        |
| <b>-w</b> <i>networkname</i> | The network name used in the emulation.                                                                                                                                           |
| <b>-a</b> <i>ip_label</i>    | The standby adapter address with which to switch.                                                                                                                                 |
| <b>-d</b> <i>ip_label</i>    | The service adapter to fail.                                                                                                                                                      |

**Note:** The **cldare** command also provides an emulation feature for dynamic reconfiguration events. See DARE Utilities on page B-25 for more information on **cldare**.

---

## Security Utilities

To simplify and automate the process of creating a secure system, two scripts for setting up Kerberos service principals are provided with HACMP for AIX and HACMP/ES versions 4.3.1 and higher:

- **cl\_setup\_kerberos**—Extracts the HACMP adapter labels from an already configured node and creates a file, **cl\_krb\_service**, that contains all of the HACMP adapter labels and additional format information required by the **add\_principal** Kerberos setup utility. Also creates the **cl\_adapters** file that contains a list of the adapters required to extract the service principals from the authentication database.
- **cl\_ext\_krb**—Prompts the user to enter the Kerberos password to be used for the new principals, and uses this password to update the **cl\_krb\_service** file. Checks for a valid **.k** file and alerts the user if one doesn't exist. Once a valid **.k** file is found, the **cl\_ext\_krb** script runs the **add\_principal** utility to add all the adapter labels from the **cl\_krb\_service** file into the authentication database; extracts the service principals and places them in a new Kerberos services file, **cl\_krb-srvtab**; creates the **cl\_klogin** file that contains additional entries required by the **.klogin** file; updates the **.klogin** file on the control workstation and all nodes in the cluster; and concatenates the **cl\_krb-srvtab** file to each node's **/etc/krb-srvtab** file.

---

## Appendix C Command Execution Language Guide

This appendix describes how you can use the HACMP for AIX Command Execution Language (CEL) to create additional Cluster Single Point of Control (C-SPOC) commands for your cluster. It also explains CEL constructs and how to convert a command's "execution plan" written in CEL into a Korn shell (C-SPOC) script.

You should be familiar with Korn shell programming (and with programming concepts in general) before attempting to create C-SPOC commands.

---

### Overview

CEL is a programming language that lets you integrate the dsh's distributed functionality into each C-SPOC script the CEL preprocessor (**celpp**) generates. When you invoke a C-SPOC script from a single cluster node to perform an administrative task, the script is automatically executed on all nodes in the cluster. Without C-SPOC's distributed functionality, you must execute each administrative task separately on each cluster node, which can lead to inconsistent node states within the cluster.

Appendix A, HACMP for AIX Commands, provides a list of all C-SPOC commands provided with the HACMP for AIX software.

---

### Creating C-SPOC Commands

C-SPOC commands are written as execution plans in CEL. Each plan contains constructs to handle one or more underlying AIX tasks (a command, executable, or script) with a minimum of user input. An execution plan becomes a C-SPOC command when the **/usr/sbin/cluster/utilities/celpp** utility converts it into a *cluster aware* **ksh** script, meaning the script uses the C-SPOC distributed mechanism - the C-SPOC Execution Engine - to execute the underlying AIX commands on cluster nodes to complete the defined tasks.

C-SPOC commands provide a means for controlling the execution of specific tasks on cluster nodes, for collecting status and log information in the **/tmp/cspoc.log** file, and for responding to errors generated by its script. Each command corresponds to an existing, underlying AIX system administration task that lets you maintain user accounts, maintain shared Logical Volume Manager (LVM) components, or control HACMP cluster services on a cluster-wide basis.

To create a C-SPOC command:

1. Write an execution plan for the command.
2. Run **celpp** to convert the plan (**.cel** file) into a **ksh** script and make the script executable.
3. Store the script in a user-defined directory.

The following sections describe each step.

---

## Writing an Execution Plan

In a C-SPOC cluster environment, various commands let you manage user and group accounts, logical volumes, and cluster services. You can expand this command set by creating new commands for specific tasks that meet your administrative needs.

To create a new command, you first must write an execution plan for the command. The execution plan contains **ksh** and CEL constructs (statements and clauses) that together describe the tasks the script will perform. CEL constructs are shown in bold in the following pseudo execution plan example. Execution plans you create should be similar in format and content to this example.

See the section “CEL Constructs” below for an explanation of each construct and to see an actual execution plan for the **cl\_chuser** command.

```

#Other CEL scripts may be included. This one defines routines used
#by all C-SPOC scripts (e.g. initialization, verification, logging,
#etc.)
%include cl_init.cel
#The following variables must be defined when including cl_init.cel in
an execution plan:

_CSPOC_OPT_STR="" # Define valid C-SPOC option flags
_OPT_STR="" # Define valid option flags for the generated
script
_USAGE="" # Define the usage for the generated script
%define ERROR_CODE 2
#An error that is not handled within a try statement using %except
#can be handled using a global %others statement.

%others
 print "ERROR: Command failed!"
%end
#Each plan must contain a try statement to execute commands across
#nodes in the cluster. A %try_parallel executes commands on all
#nodes simultaneously. Note that only a single line ksh statement
#is allowed after a %try_parallel or %try_serial.

%try_parallel
 ls -l $*
 %except ERROR_CODE
 print "ERROR: Unable to open: $*"
 %end
%end
#####
```

For a description of option string variables used in the above example, refer to the **cl\_init.cel** file in the **/usr/lpp/cluster/samples/cspoc** directory. The **cl\_init.cel** file provides examples of functionality required in any execution plan you create; it should be included in each **.cel** file.

The initialization and verification routines in the **cl\_init.cel** file provide the following functionality:

- Gets a list of nodes in the cluster
- Gets a list of target nodes for command execution
- Determines nodes associated with any resource groups specified
- Processes & implements the standard C-SPOC flags (**-f**, **-n**, & **-g**)

- Validates option strings. Requires several environment variables to be set within the plan. (`_OPT_STR`, `CSPOC_OPT_STR`, & `_USAGE`)
- Saves log file entries upon command termination
- Performs C-SPOC verification as follows:
  - Ensures **cdsh** is available in `$PATH`
  - Checks HACMP version on all nodes.

## Encoding and Decoding Command-Line Arguments

If the command execution plans you create require that you pass additional arguments (or perform additional processing to the existing command-line arguments), you must use the C-SPOC **clencodearg** and **cldecodearg** programs located in `/usr/sbin/cluster/cspoc` directory to code and decode a list of arguments to ensure that the CEL execution engine handles the embedded spacing and quotes within arguments across multiple invocations of the shell.

For example, in the following command, the argument “John Smith” contains an embedded space:

```
chuser gecos="John Smith" jsmith
```

Thus, you should encode all command-line arguments to C-SPOC commands unless they begin with a dash. Arguments that begin with a dash are generally command-line flags and do not contain spaces or quotes. If a string that begins with a dash is passed to the **clencodearg** or **cldecodearg** program, the string is passed through without being encoded or decoded.

The following script fragments shows how to encode and decode a list of arguments.

To encode a list of args:

```
ENCODED_ARGS=""
for ARG in "$@"
do
 ENCODED_ARGS="$ENCODED_ARGS $(print ARG | clencodearg)"
done
```

To decode a list of args:

```
UNENCODED_ARGS=""
for ARG in $ENCODED_ARGS
do
 UNENCODED_ARGS="$UNENCODED_ARGS $(print $ARG | cldecodearg)"
done
```

**Warning:** You must use the **clencodearg** or **cldecodearg** program to process command-line arguments that are to be passed to any commands contained within a **%try\_serial** or **%try\_parallel** statement because the C-SPOC Execution Engine (**cdsh**) tries to decode all command-line arguments before executing the command. See the section “CEL Constructs” for more information about **%try\_serial** and **%try\_parallel** statements.

**Warning:** Any arguments obtained from the environment variables set by routines within `cl_init.cel`, such as `_getopts()`, will have been encoded. If a command contained within a `%try_serial` or `%try_parallel` statement includes arguments generated within the execution plan, then they must first be encoded. For example, most execution plans pass command-line arguments as follows:

```
%try_parallel
 chuser $_CMD_ARGS
%end
```

The following CEL plan uses the `clencodearg` and `cldecodearg` programs.

### Example CEL Plan

```
#Initialize C-SPOC variables required by cl_init.cel

_CMD_NAME=$(basename $0) # Specify the name of this script.
_CSPOC_OPT_STR="d:f" # Specify valid C-SPOC option flags.
_OPT_STR="+2" # Specify valid AIX command option flags.

#Specify a Usage statement for this script.

_USAGE="USAGE: cl_chuser [-cspoc -f] Attr=Value...Username"
#Initialize variables local to this script. The default return code
#is '0' for success.

RETCODE=0
#Include the required C-SPOC Initialization and Verification
#Routines

%include cl_init.cel
#This define makes the exception clause below more readable.

%define USER_NOT_FOUND 2
#Get the username, which is the last command line arg and perform
#some crude checks for validity. Note: This is an example of
#decoding args within a script.

user=${_CMD_ARGS##*[]} # The []'s contain a tab & a space!
case $user in
 -*|"")
 print "$_USAGE"
 exit 2
 ;;
esac
#Since cl_init.cel has encoded all the args we must decode it to
use it.

Duser=$(print $user | cldecodearg)
#Construct a Gecos field based on the 411 entry for the username.
#Note: 411 is a local script that prints the following
tab-separated fields: username Firstname Lastname Work_phone
Home_phone

#This plan cuts the "Firstname Lastname" to put into the Gecos
#field of /etc/passwd

GECOS=$(/usr/local/bin/411 $Duser | cut -d' ' -f2)
#Construct a new set of command line args. Note the following:
1) We put the 'gecos=' arg just before the username so that it
will supercede any that may have been specified on the
command line.
2) We must encode the 'gecos=' arg explicitly.
```

```
3) This is an example of encoding args specified inside a script.

NEW_CMD_ARGS=${_CMD_ARGS%[]*} # []'s contain a tab & a space
NEW_CMD_ARGS="$NEW_CMD_ARGS $(print gecos="$GECOS" | HR>
clencodearg) $user"

#Perform a check that the username exists on all nodes. The check
#is not performed when the C-SPOC force flag is specified.

 if [[-z "${_SPOC_FORCE}"]]
 then

#Check if user exists across all cluster nodes

%try_parallel _NODE _TARGET_NODES silent_err silent_out
 lsuser $user
%except USER_NOT_FOUND
 print "${_CMD_NAME}: User ${Duser} does "
 print "not exist on node ${_NODE}" >&2"
 RETCODE=1
%end
%end
fi

#If the username does not exist on a node then exit immediately

 if [[${RETCODE} -ne 0]]
 then
 exit ${RETCODE}
 fi

#Execute the chuser command in parallel on all nodes in the #cluster.

%try_parallel _NODE _TARGET_NODES
 chuser $NEW_CMD_ARGS
%others
 # If chuser returned an error on any node
 # set the return code to '1' to indicate
 # that one or more nodes failed.
 RETCODE=1
%end
%end

#Exit with the appropriate value.

 exit ${RETCODE}
```

## CEL Constructs

You can use the following CEL constructs (statements or clauses) in a command's execution plan. All C-SPOC commands must contain the **%include** statement, a **%try\_parallel** or **%try\_serial** statement, and an **%end** statement.

The **%include** statement is used to access **ksh** libraries within the **cl\_init.cel** file. These libraries make C-SPOC commands "cluster aware." The **%except** and **%others** clauses are used typically for error handling in response to a command's execution on one or more cluster nodes.

### **%define statement:**

```
%define key value
```

For improved readability, the **%define** statement (keyword) is used to provide descriptive names for error\_id values in **%except** clauses. The error\_id given in the define statement is inserted in place of the given id in any subsequent **%except** clauses.

**%include statement:**

```
%include filename
```

The **%include** statement allows a copy of a file to be included in an execution plan, which means that common code can be shared among execution plans. The **%include** statement also allows C-SPOC commands with similar algorithms to share the same plans. Using **%include**, CEL statements can be used within library routines included in any execution plan.

**%try\_parallel statement:**

```
%try_parallel node nodelist [silent_err] [silent_out]
 ksh statement
 [%except clause ...]
[%end | %end_p]
```

The **%try\_parallel** statement executes the enclosed **ksh** statement simultaneously across all nodes in the node list. It waits until a command completes its execution on all cluster nodes before checking for errors.

The **%try\_parallel** statement is equivalent to the following pseudo-code:

```
for each node in nodelist
 execute ksh statement in the background on each node
end
wait for all background execution to complete on all nodes
for each node in nodelist
 check status for each node and execute %except
 clause(s)
end
```

**%try\_serial statement:**

```
%try_serial node nodelist [silent_err] [silent_out]
 ksh statement
 [%except clause ...]
 [%others clause]
[%end | %end_s]
```

The **%try\_serial** statement executes the enclosed **ksh** statement consecutively on each node specified in the node list. The command must complete before **%try\_serial** continues its execution on the next node. All **%try\_serial** statements sequentially check for errors after a command completes on each node in the list.

The **%try\_serial** statement is equivalent to the following pseudo-code:

```
for each node in nodelist
 execute ksh statement in the on each node
end
for each node in nodelist
 check status for each node and execute %except
 clause(s)
end
```

**Note:** Any non-flag arguments (those not preceded by a dash) used inside a **try** statement must be encoded using the **/usr/sbin/cluster/cspoc/clencodearg** utility; otherwise, arguments will be decoded incorrectly.

**%except clause:**

```
%except error_id
 ksh code
[%end | %end_e]
```

The **%except** clause is used for error checking on a per-node basis, and for the execution of the **ksh** statement contained within a **%try\_parallel** or **%try\_serial** statement. If the **ksh** statement fails or times out, the **%except** clauses define actions to take for each return code value.

If an `error_id` is defined using the **%define** statement, the return code of the **ksh** statement (the status code set by a command's execution on a particular node) is compared to the `error_id` value. If a match is found, the **ksh** code within the **%except** statement is executed.

If the **%others** clause is used within the **%except** clause, the **ksh** code within the **%except** clause is executed only if the return code value did not match an `error_id` defined in previous **%except** clauses.

An **%except** clause can be defined within or outside a **try** statement. If it is defined within a **try** statement, the clause's scope is local to the **%try\_parallel** or **%try\_serial** statement.

If the **%except** clause is defined outside a **try** statement, it is used in any subsequent **%try\_serial** or **%try\_parallel** statements that do not already provide one for the specified `error_id` or **%others** clause. In this case, the **%except** clause's scope is global until an **%unexcept** statement is found. Global **%except** clauses are used to simplify and standardize repetitive types of error handling.

**%stop\_try clause:**

```
%except error_id
 ksh code
 [%stop_try clause]
 ksh code
[%end | %end_e]
```

The **%stop\_try** clause can be used within an **%except** clause; it causes an exit from the enclosing **try** statement. This statement has the effect of a `break` statement in other languages.

The **%stop\_try** clause has a different effect that depends on whether it is defined with a **%try\_parallel** versus a **%try\_serial** statement.

In a **%try\_serial** statement, defining a **%stop\_try** clause prevents further execution of the **ksh** statement on all cluster nodes; it also prevents further error checking defined by the **%except** clause.

In a **%try\_parallel** statement, defining a **%stop\_try** clause prevents error checking on all cluster nodes, since execution of the **try** statement happens simultaneously on all nodes. The commands within the **ksh** statement will have completed on other nodes by the time the **%except** clauses are evaluated.

**%others clause:**

```
%others error_id
 ksh code
 [%stop_try clause]
 ksh code
[%end | %end_o]
```

The **%others** clause is the default error action performed if a command's return code does not match the return code of a specific **%except** clause. The **%others** clause can be used in a **try** statement, or globally within an execution plan.

**%end statement:**

`%end`

The **%end** statement is used with all CEL constructs. Ensure that your **.cel** file includes an **%end** statement with each statement or clause. A construct like **%try\_parallel**, for example, can be ended with **%end** or with **%end\_p**, where "p" represents parallel.

**Actual Execution Plan**

The following example shows an actual execution plan, **cl\_chuser.cel**, written in CEL that uses the required constructs to create the **cl\_chuser** command. The CEL constructs are shown in bold. Other examples of execution plans are in the **/usr/lpp/cluster/samples/cspoc** directory.

```
#####
Name:
cl_chuser.cel
#
Description:
The cl_chuser command changes user attributes for a user on all # nodes in an
HACMP cluster.
#
Usage: cl_chuser [-cspoc "[-f]"] Attribute=Value ... Name
#
Arguments:
The cl_chuser command arguments include all options and
arguments that are valid for the chuser command as well as
C-SPOC specific arguments. The C-SPOC specific arguments are as #follows:
-f C-SPOC force flag
#
Return Values:
0 success
1 failure
#
#####
Initialize variables
_CMD_NAME=`basename $0`
_CSPOC_OPT_STR="d:f"
_OPT_STR=""
_USAGE="Usage: cl_chuser [-cspoc #7f>[-f]] Attribute=Value ... Name"
_MSET=12
_RETCODE=0
Default exception handling for a COULD_NOT_CONNECT error
%except 1000
 nls_msg -l $cspoc_tmp_log ${_MSET} CDSH_UNABLE_TO_CONNECT "${_CMD_NAME}":
Unable to
 connect to node ${_NODE}" >& 2
 if [[-z "${_SPOC_FORCE}"]]
 then
 exit 1
 fi
%end
C-SPOC Initialization and Verification
%include cl_init.cel
%define USER_NOT_FOUND 2
user=${_CMD_ARGS### }
if [[-z "${_SPOC_FORCE}"]]
then
 #
 # Check if user exists across all cluster nodes
```

```

#
%try_parallel _NODE _TARGET_NODES silent_err silent_out
 lsuser $user
 %except USER_NOT_FOUND
 nls_msg -l $cspoc_tmp_log ${_MSET} 1 "${_CMD_NAME}: User ${user} does not
exist on node ${_NODE}" ${_CMD_NAME} ${user} ${_NODE} >& 2
 _RETCODE=1
 %end
%end
fi
If user does not exist on a node, exit 1
if [[${_RETCODE} -ne 0]]
then
 exit ${_RETCODE}
fi
Run chuser across the cluster nodes
%try_parallel _NODE _TARGET_NODES
 chuser $_CMD_ARGS
 %others
 # If chuser returned an error on any node, exit 1
 _RETCODE=1
 %end
%end
exit ${_RETCODE}

```

---

## Converting Execution Plans to ksh Scripts

After writing an execution plan for a command, you must use the CEL preprocessor (**celpp**) to convert it into a **ksh** script; then you must execute the **chmod** command on the script to make it executable. An executable **ksh** script includes all code required to implement the command across cluster nodes. The preprocessor generates much of the code you would otherwise have to write.

To convert an execution plan into a **ksh** script:

1. Change directories to the one containing the execution plan (**.cel** file) for the command you want to convert.
2. Enter the following command to run **celpp**, specifying the **.cel** file as the input file and the command to be generated as the output file:

```
celpp [-i inputfile] [-o outputfile] [-I IncludeDirectory]
```

where the elements of this command line are:

- i inputfile**            Uses *inputfile* for input, or **stdin** by default.
- o outputfile**            Uses *outputfile* for output, or **stdout** by default.
- I IncludeDirectory**      Uses the specified directory name to locate **.cel** files (execution plans). You can specify multiple **-I** options.

For example, enter:

```
celpp [-i cl_chuser.cel] [-o cl_chuser] HR>[-I /usr/sbin/cluster/samples/cspoc]
```

This command converts the **cl\_chuser.cel** file into a **ksh** script.

3. Enter the following command to make the generated **ksh** script executable:

```
chmod +x inputfile
```

For example, enter:

```
chmod +x cl_chuser
```

You can now invoke the C-SPOC command. If you change a command's execution plan after converting it, repeat the above steps to generate a new **ksh** script. Then make the script executable. Note that the above steps can be included in a **Makefile**.

---

## Storing C-SPOC Commands

The HACMP for AIX software must be installed in your cluster to use C-SPOC commands. You can store C-SPOC commands you create in any directory.

---

## Handling Command Output

When you are ready to use a C-SPOC command, be aware that all error output (messages) resulting from the command's execution across the cluster is returned to **stdout** and **stderr**. By default, **stdout** and **stderr** remain directed to your display. You can override this action by using the **silent\_err** or **silent\_out** arguments with the **try** statements in a command's execution plan. See "Writing an Execution Plan" for an example of how you can use these arguments.

However you redirect error output, all error messages are prefixed with the node name on which the command was executed. When the command completes on each node, error messages are logged in the **/tmp/cspoc.log** file. See the *HACMP for AIX Troubleshooting Guide* for more information about C-SPOC-related log entries.

---

## Appendix D VSM Graphical Configuration Application

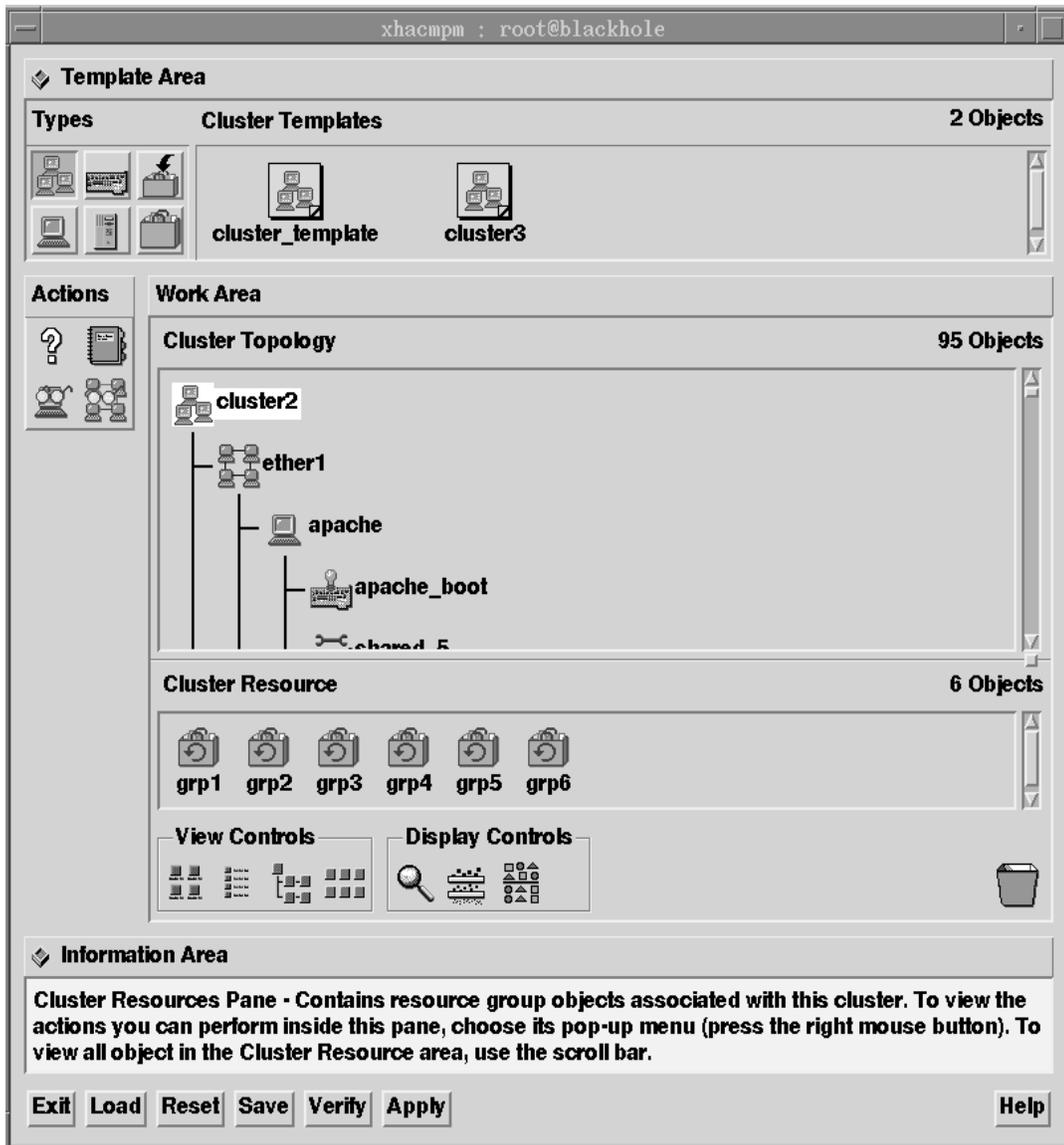
This appendix describes the **xhacmpm** graphical interface tool for configuring cluster environments and components.

---

### Overview

Through the graphical interface **xhacmpm**, the HACMP for AIX Visual System Management (VSM) application, you can perform many configuration tasks. Using **xhacmpm**, a Common Desktop Environment (CDE) application, you can define attributes for clusters and cluster objects like nodes, resource groups, or application servers, and simultaneously and dynamically configure your cluster environment. You also can save snapshots of cluster configurations and restore them when needed. By clicking, or by dragging and dropping objects onto other cluster objects (or into a Work Area pane) in VSM, you can configure cluster components and the cluster environment.

For example, to configure a cluster, click on the cluster icon in the Types window and then drag and drop a cluster\_template into the Cluster Topology pane. A dialogue window appears to let you define the cluster name and ID. Drag and drop to add nodes, adapters, resource groups, and other components to the cluster. The following figure illustrates a cluster configuration represented in the **xhacmpm** window.



The HACMP for AIX xhacmpm Application

---

## Starting xhacmpm

To start **xhacmpm**, enter the following command:

```
/usr/lpp/cluster/samples/xhacmpm &
```

When started, **xhacmpm** displays its default configuration. If you want to cancel a configuration you are working on but have not applied, click the Reset button to return **xhacmpm** to its startup display, canceling all changes. To save a snapshot of the current configuration, click the Save button. You can restore the snapshot by clicking the Load button and by specifying a PATH and file name in the dialog window.

## The xhacmpm Window

The **xhacmpm** window contains several distinct areas:

### Template Area

In the Template Area, at the top of the window, you define characteristics of cluster objects.

The **xhacmpm** application includes templates for:

- Clusters
- Nodes
- Adapters
- Application servers
- Resources
- Resource groups.

Each template defines characteristics of a cluster object. For example, the cluster template allows you to specify the cluster name and ID. To define the characteristics of a cluster object, drag its icon from the Template Area into the Work Area.

### Work Area

In the Work Area, the middle section of the window, you create a cluster configuration by combining cluster object templates. For example, to configure a cluster to include nodes, drag and drop a node\_template onto a cluster object in the Work Area's Cluster Topology pane. The **xhacmpm** application displays a dialog box where you can specify each node's attributes. After configuring a node object, you can keep a permanent record of its definition by dragging the icon up into the Node Templates pane in the Template Area. The node object becomes a template object with the same name.

### Information Area

The Information Area at the bottom of the window displays brief descriptions of whatever part of the **xhacmpm** interface is under the mouse pointer. For example, when you move the pointer over the wastebasket icon in the lower-right corner of the window, the Information Area displays a description of how to use it.

The **xhacmpm** interface includes other buttons that allow you to customize the view of the cluster you are configuring, search for particular cluster objects, and determine the current values of their characteristics.

## Using xhacmpm Online Help

For more information about using **xhacmpm** to configure an HACMP cluster, start the application and view the online help that comes with it. To access help, drag the question mark (?) Action icon onto any icon representing a cluster object and **xhacmpm** displays a window providing information about the object and any related configuration tasks.

You must perform the cluster configuration steps in the order specified in the online help:

1. Define the cluster topology
2. Configure application servers
3. Configure cluster resources
4. Verify a cluster topology
5. Save a cluster configuration
6. Activate the configuration.

**Note:** You can customize many aspects of **xhacmpm**'s window appearance, such as the background color. Specify customizations using command-line arguments, or by defining resources in the **.Xdefaults** file in your \$HOME directory to override the defaults in the **xhacmpm** resource file.

---

## Appendix E 7x24 Maintenance

The goal of high availability is to keep systems up and running, allowing continuous access to critical applications. In many enterprises it has become necessary to keep applications running seven days a week, 24 hours a day. With proper planning, customizing, and monitoring, an HACMP cluster can provide nearly continuous availability, interrupted only by scheduled, necessary maintenance.

This appendix is a collection of information describing the issues and procedures involved in keeping a cluster running on as close to a 7 X 24 basis as possible.

---

### Overview

Throughout all stages of cluster administration—planning, configuration, maintenance, troubleshooting, and upgrading—here are tasks you can do and systems you can put in place to help ensure your cluster’s nearly continuous availability.

Once you have configured the cluster and brought it online, it is very important to do maintenance tasks in as non-disruptive a way as possible. The HACMP cluster is a distributed operating system environment. Therefore maintaining an HACMP cluster requires attention to some issues that have different ramifications in the cluster environment compared to maintaining a single-server system.

Making changes to a cluster must be thoroughly planned, since changes to one component may have cascading effects. Changes on one node affect other nodes, but this is not often apparent until failover occurs (or cannot occur due to a non-synchronized change to the cluster). Some of the “do’s and don’ts” of cluster maintenance are explained in this appendix.

Finally, setting up and following regular preventive maintenance procedures helps alert you to any potential problems before they occur. Then you can take timely action or plan failovers or cluster downtime at your convenience as necessary to deal with any impending issues.

The appendix is organized as follows:

- **Planning for 7 X 24 Maintenance.** This section reemphasizes the importance of careful planning and customization of the original installation of the cluster.
- **Run-Time Maintenance.** This section offers reminders and tips to help you avoid actions that endanger a stable, running cluster.
- **Hardware Maintenance.** This section contains procedures for changing or replacing certain hardware
- **Practicing Preventive Maintenance.** This sections reviews tools you can use to avoid problems or catch them early.

---

## Planning for 7 X 24 Maintenance

Planning the original installation of your cluster carefully goes a long way toward making cluster maintenance easier. A well-configured and customized cluster is the first step for good preventive maintenance. Proper cluster configuration also makes it less likely you will have to make changes that affect cluster performance while users are accessing their applications.

Planning the cluster starts with a Single Point of Failure analysis. See the *HACMP Planning Guide* for a detailed list of issues to consider. Once the cluster is installed and running, you need to handle any failures as quickly and/or as automatically as possible. Planning for run-time failure recovery will help ensure that HACMP for AIX does all that it is capable of doing to keep your critical resources online.

This section includes information on the following topics:

- Customizing the cluster, including setting up error notification to improve monitoring and management of the cluster
- Tuning the communications system—network and nameserving issues
- Planning disk and volume group layout
- General planning for hardware and software maintenance

### Customizing the Cluster

Customizing the cluster enhances your ability to monitor the cluster and keep it running. You can define a pre-event, a post-event, and a notification method for every cluster event. Notification of events is critical to maintain service for any HACMP cluster. Although HACMP writes messages to the **hacmp.out** and **cluster.log** log files, it is very useful to include notifications to the console or mail to the system administrator when an event occurs that demands immediate attention.

You can include automatic recovery actions as well as notification in the cluster customization.

Use the HACMP and AIX tools available to customize some or all of the following:

- Hardware error notification
- Hardware failure
- Cluster event notification
- Pre- and post-event recovery actions
- Network failure escalation
- ARP cache refresh
- Application server scripts.

It is highly recommended that you maintain a test cluster as well as your production cluster. Thus before you make any major change to the production cluster, you can test the procedure on the test cluster. HACMP supplies event emulation utilities to aid in testing.

## Customizing AIX Error Notification of Hardware Errors

Customizing notification when you configure the cluster is a good preventive measure. See Chapter 14, Tailoring Cluster Event Processing in the *HACMP Planning Guide* and Chapter 29, Customizing Cluster Events and Log Files in the *HACMP Installation Guide* for complete information on customizing and setting up notification of cluster events.

See Chapter 31, Supporting AIX Error Notification, in the *HACMP Installation Guide* for information on supporting AIX Error Notification, and for information on setting up automatic notification for hardware errors that do not cause cluster events.

Using the HACMP Automatic Error Notification SMIT screens, you can turn on automatic error notification for selected hard, non-recoverable error types: disk, disk adapter, and SP switch adapter errors. All disks defined as HACMP resources, and disks in the rootvg and HACMP volume groups and filesystems are included.

You may want to set up error notification for certain media or temporary errors. You may also want to customize the error notification for some devices rather than using one of the two automatic error notification methods.

Chapter 31, Supporting AIX Error Notification in the *HACMP Installation Guide* includes an example of how to promote a hardware error to a node failure, using the AIX Error Notification facility.

### List of Hardware Errors to Monitor

The following list of hardware errors gives you a good idea of types of errors to monitor. The first list shows which errors are handled by the HACMP automatic error notification utility. The following lists show other types of errors you may want to address. You can decide whether you want to escalate any of these to node failure. For each device monitored, you can determine an additional action other than notification, such as:

- Initiate a graceful takeover
- Initiate a custom recovery action such as reconfiguration for a failed device using an alternative device.

| <b>Hardware Errors Handled by HACMP Auto-Error Notification</b> |                                                               |
|-----------------------------------------------------------------|---------------------------------------------------------------|
| <b>DISK_ERR2</b>                                                | Permanent physical disk error (known error)                   |
| <b>DISK_ERR3</b>                                                | Permanent physical disk error, adapter detected (known error) |
| <b>SCSI_ERR1</b>                                                | Permanent SCSI adapter HW error (known error)                 |
| <b>SCSI_ERR3</b>                                                | Permanent SCSI adapter microcode error (known error)          |
| <b>SCSI_ERR5</b>                                                | Temporary SCSI bus error                                      |
| <b>SCSI_ERR7</b>                                                | Permanent unknown system error                                |
| <b>SCSI_ERR9</b>                                                | Potential Data loss condition                                 |
| <b>SDA_ERR1</b>                                                 | Adapter hardware error condition                              |

|                         |                                               |
|-------------------------|-----------------------------------------------|
| <b>SDA_ERR3</b>         | Permanent unknown system error                |
| <b>SDC_ERR1</b>         | Controller/DASD link error                    |
| <b>SDC_ERR2</b>         | Controller hardware error                     |
| <b>SSA_HDW_ERR</b>      | SSA hardware error condition                  |
| <b>SSA_DISK_ERR1</b>    | Permanant microcode program error             |
| <b>SSA_DISK_ERR4</b>    | Permanent disk operation error                |
| <b>DISK_ARRAY_ERR2</b>  | Permanent disk operation error (disk failure) |
| <b>DISK_ARRAY_ERR3</b>  | Permanent disk operation error (disk failure) |
| <b>DISK_ARRAY_ERR5</b>  | Permanent disk operation error (disk failure) |
| <b>SCSI_ARRAY_ERR2</b>  | SCSI hardware error                           |
| <b>HPS_FAULT4_ER</b>    | SP switch error                               |
| <b>TB3_HARDWARE_ER</b>  | TB3 type switch hardware error                |
| <b>TB3_MICROCODE_ER</b> | TB3 switch microcode error                    |
| <b>TB3_SLIH_ER</b>      |                                               |

| <b>Disk and Adapter Errors Not Covered by HACMP Auto-Error Notification</b> |                                                 |
|-----------------------------------------------------------------------------|-------------------------------------------------|
| <b>SSA_HDW_RECOVERED</b>                                                    | Temporary adapter error                         |
| <b>SSA_DISK_ERR3</b>                                                        | Temporary disk operation error                  |
| <b>SSA_DEGRADED_ERROR</b>                                                   | Adapter performance degraded                    |
| <b>SSA_LOGGING_ERROR</b>                                                    | Permanent unable to log an error against a disk |
| <b>SSA_LINK_OPEN</b>                                                        | Permanent adapter detected open serial link     |
| <b>SSA_SOFTWARE_ERROR</b>                                                   | Permanent software program error                |
| <b>SSA_LINK_ERROR</b>                                                       | Temporary link error                            |
| <b>SSA_DETECTED_ERROR</b>                                                   | Permanent loss of redundant power/cooling       |
| <b>LVM_MISSPVADDED</b>                                                      | PV defined as missing (unknown error)           |
| <b>LVM_SA_WRT</b>                                                           | PV defined as missing (unknown error)           |
| <b>LVM_SA_PVMISS</b>                                                        | Failed to write VGSA (unknown error)            |

| Disk Array Errors not Covered by HACMP Auto-Error Notification |                                                            |
|----------------------------------------------------------------|------------------------------------------------------------|
| DISK_ARRAY_ERR4                                                | Temporary disk operation error (disk media failing)        |
| DISK_ARRAY_ERR6                                                | Permanent array subsystem degradation (disk media failure) |
| DISK_ARRAY_ERR7                                                | Permanent array subsystem degradation (controller)         |
| DISK_ARRAY_ERR8                                                | Permanent array active controller switch (controller)      |
| DISK_ARRAY_ERR9                                                | Permanent array controller switch failure                  |

| Failed 64-port Adapter (tty device driver) |                                         |
|--------------------------------------------|-----------------------------------------|
| COM_PERM_PIO                               | PIO exception, possible adapter failure |

You may have additional devices critical to your operation that are not supported by HACMP for AIX. For example, you may have X.25 components that need monitoring. You can set up AIX error notification to monitor X.25 microcode errors or adapter time-outs.

## Customizing Cluster Events

Customizing cluster events to send notification or to take recovery actions is another method you can use to help maintain the cluster running as smoothly as possible.

See Chapter 14, Tailoring Cluster Event Processing in the *HACMP Planning Guide* and Chapter 29, Customizing Cluster Events and Log Files in the *HACMP Installation Guide* for complete information on customizing and setting up notification of cluster events.

See the section Sample Custom Scripts on page 29-14 in the *HACMP Installation Guide* for tips on writing scripts to make **cron** jobs and print queues highly available.

## Customizing Application Server Scripts

See Appendix D, Applications and HACMP in the *HACMP Planning Guide* for tips on handling applications.

Some key things to keep in mind:

- Define an HACMP application server for each node that supports applications requiring recovery.
- Applications must be started up and shutdown in an orderly fashion. Some situations exist where the timing and control of starting and stopping applications needs to be handled based on pre/post event process. You may need to take into account the order in which applications assigned to the same node are started.
- Check for dependencies between nodes. For example, a process on node1 cannot start until a process that runs on node2 is up. Include a check for remote node/application availability before issuing the local startup command.
- You may need to perform some checks to make sure the application is not running and to clean up logs or roll back files before starting the application process.

See the section Sample Custom Scripts on page 29-14 in the *HACMP Installation Guide* for tips on writing scripts to make **cron** jobs and print queues highly available.

## Network Configuration and Nameserving

Setting up and maintaining clear communication paths for the Cluster Manager is a key element for efficient cluster operation.

### Setting up Serial Networks

It is crucial to have at least one serial network configured for the cluster. Without a serial network, you run the risk of a partitioned cluster if TCP/IP networks fail, since the nodes will be unable to maintain heartbeat communication.

If there is a problem with the serial network (or there is none configured and defined) you may also see the DGSP message in the AIX error log. (DGSP = Diagnostic Group Shutdown Partition.) This message is sent if a takeover is initiated after a node\_down, and the downed node revives and sends heartbeats. The takeover nodes kill the reviving node to ensure non-contention over resource ownership within the cluster. This prevents that node from being able to rejoin the cluster normally.

### Integrating HACMP with Network Services

HACMP requires IP address to name resolution. The three most commonly used methods include:

- Domain Name Service
- Network Information Service
- Flat file name resolution (**/etc/hosts**)

By default, a name request will look first for the DNS (**/etc/resolv.conf**), second for NIS, and last for **/etc/hosts** to resolve the name. Since DNS and NIS both require certain hosts as designated servers, it is necessary to maintain the **/etc/hosts** file in case the DNS or NIS name server is unavailable, and to identify hosts that are not known to the name server. It is required to have all cluster adapter IP labels in all cluster nodes' **/etc/hosts** tables.

To ensure the most rapid name resolution of cluster nodes, it is recommended that you change the default order for name serving so that **/etc/hosts** is used first (at least for cluster nodes).

To do this, edit the **/etc/netsvc.conf** file so that this line appears as follows:

```
Hosts=local,nis
```

Putting the local option first tells the system to use **/etc/hosts** first, then NIS.

You can also change the order for name resolution by changing the environment variable NSORDER so it looks like this:

```
NSORDER=local,bind
```

**Note:** If you are using NIS, it is recommended to have the NIS master server outside the cluster, and have the cluster nodes run as NIS slave servers.

If you are using DNS or NIS, this fact must be configured in the run-time parameters of all nodes in the cluster. The Cluster Manager then knows to stop and restart the **named** when it is changing adapter IP addresses (in an adapter\_swap event).

See Using HACMP with NIS and DNS on page 9-18 in the *HACMP Planning Guide* for more information on possible NIS and DNS issues such as changing a host name. See Chapter 22, Additional AIX Administrative Tasks in the *HACMP Installation Guide* for information on editing the `/etc/hosts` file, and also for notes on NIS and **cron** considerations.

**Warning:** You cannot use DHCP to allocate IP addresses to HACMP cluster nodes. Clients may use this method, but cluster nodes cannot.

## Tuning Networks for Best Performance

HACMP 4.4 provides easier and greater control over several tuning parameters that affect the cluster's performance. Setting these tuning parameters correctly to ensure throughput and adjusting the HACMP failure detection rate can help avoid "failures" caused by heavy network traffic.

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

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

Setting these tuning parameters correctly may avoid some of the performance problems noted above.

See Planning for Cluster Performance on page 9-21 in the *HACMP Planning Guide* for information on setting AIX and HACMP parameters for best performance.

If you are running a cluster on an SP, also consult your SP manual set for instructions on tuning SP switch networks.

## Planning Disks and Volume Groups

Planning the disk layout is crucial for the protection of your critical data in an HACMP cluster. Follow the guidelines carefully, and keep in mind these issues:

- All operating system files should reside in the root volume group (**rootvg**) and all user data should reside outside that group. This makes updating or reinstalling the operating system and backing up data more manageable.
- A node whose resources are not designed to be taken over should not own critical volume groups.
- When using copies, each physical volume using a mirror copy should get its power from a UPS system.
- Volume groups that contain at least three physical volumes provide the maximum availability when implementing mirroring (one mirrored copy for each physical volume).
- **auto-varyon** must be set to **false**. HACMP will be managing the disks and varying them on and off as needed to handle cluster events.

## Quorum Issues

Setting up quorum correctly when laying out a volume group is very important. Quorum *must* be enabled on concurrent volume groups. With quorum enabled, a two-disk non-concurrent volume group puts you at risk for losing quorum and data access. The failure of a single adapter or cable would cause half the disks to be inaccessible.

Either build three-disk volume groups or disable quorum on non-concurrent volume groups. You can also use a quorum buster disk that contains no critical data to ensure maintenance of quorum.

Maintaining volume groups, logical volumes, and filesystems shared by cluster nodes require that you keep the ODM definitions of these components synchronized across all cluster nodes. It is recommended to use the C-SPOC commands HACMP provides for maintaining these components, thus updating configuration changes from one node to all the others.

See the detailed section on Quorum in Chapter 6 of the *HACMP Planning Guide*.

## Planning Hardware Maintenance

Good maintenance practice in general dictates that you

- Check cluster power supplies periodically.
- Check the **errlog** and/or any other logs where you have redirected information of interest and attend to all notifications in a timely manner.
- Be prepared to replace any failed or outdated cluster hardware.

If possible, you should have replacement parts readily available. If the cluster has no single points of failure, it will continue to function even though a part has failed. However, now a single point of failure may exist. If you have set up notification for hardware errors, you have an early warning system in place.

This Guide contains procedures detailing how to replace the following cluster components while keeping the cluster running:

- Network
- Adapter
- Disk
- Node.

See Hardware Maintenance on page E-13 for more information.

## Planning Software Maintenance

Planning for software maintenance includes:

- Customizing notification of software problems
- Periodically checking and cleaning up log files
- Taking cluster snapshots when making any change to the cluster configuration
- Preparing for upgrading AIX, applications, and HACMP for AIX.

See Preventive Maintenance on page E-19 in this Appendix for more details.

---

## Run-Time Maintenance

Once you have configured the cluster and brought it online, it is very important to do maintenance tasks in as non-disruptive a way as possible. Maintaining an HACMP cluster requires attention to some issues that have different ramifications in the cluster environment compared to maintaining a single system.

This section discusses the following issues:

- Tasks that require stopping the cluster
- Warnings about the cascading effects caused by making certain types of changes to a stable, running cluster

### Tasks That Require Stopping the Cluster

HACMP allows you to do many tasks without stopping the cluster; you can do many tasks dynamically using the DARE and C-SPOC utilities. However, in order to do the following tasks, you must stop the cluster:

- Change the name of a cluster component: network module, cluster node, network adapter, application server, or resource group. Once you configure the cluster, you should not need to change these names,
- Change the cluster ID. This is also a rare occurrence.
- Change auto error notification.
- Change SSA fence registers.

### Changing the Cluster Configuration—Cascading Effects on Cluster Behavior

Installing HACMP makes changes to several AIX files (see Chapter 1, Maintaining an HACMP for AIX System). All the components of the cluster are under HACMP control once you configure, synchronize, and run the cluster software. Using AIX to change any cluster component, instead of using the HACMP menus and synchronizing the topology and/or the cluster resources, will interfere with the proper behavior of the HACMP cluster software and thus affect critical cluster services.

This section contains warnings about actions that will endanger the proper behavior of an HACMP cluster. It also includes some reminders about proper maintenance procedures.

### Stopping and Starting Cluster Services

Do not start or stop daemons or services that are running under the control of HACMP. Any such action will affect cluster communication and behavior. You can choose to run certain daemons (Clinfo) but others are required to run under HACMP control.

Most important, never use the **kill – 9** command to stop the Cluster Manager. This causes an abnormal exit. SRC will run the **clexit.rc** script and halt the system immediately. This causes the other nodes to initiate a failover.

TCP/IP services are required for cluster communication. Do not stop this service on a cluster node. If you need to stop HACMP or TCP/IP to maintain a node, use the proper procedure to move the node's resources to another cluster node, then bring down the node. Follow the instructions in Chapters 6 and 7 of this Guide to make changes to cluster topology or resources.

## **Node, Network and Network Adapter Issues**

The HACMP configuration of the nodes and IP addresses of network adapters is crucial to the communication system of the cluster. Any change in the definitions of these elements must be updated in the cluster configuration and resynchronized.

Do not change the configuration of a cluster node, network, or adapter using AIX SMIT menus or commands, individually on a cluster node, outside of HACMP. See Chapter 6, Changing the Cluster Topology, for instructions on changing the configuration dynamically following the proper HACMP cluster procedures.

Do not start or stop daemons or services that are running under the control of HACMP. Any such action will affect cluster communication and behavior.

Be sure to follow proper procedures for the following types of changes:

- Changing boot, service, or standby IP address of any adapter defined to HACMP. Changes to IP addresses must be updated in the HACMP cluster definition and the cluster must then be resynchronized. Any change to network adapter attributes normally requires stopping cluster services, making the change, and restarting cluster services.

Note that in some circumstances you can use the HACMP facility to swap a network service or boot adapter IP address dynamically, to an active standby adapter on the same node and network, without shutting down cluster services on the node. See *Swapping a Network Adapter Dynamically* on page 6-13 for more information.

- Changing netmasks of network adapters. Service and standby adapters on the same network must have the same netmask on all cluster nodes. Changes made outside the cluster definition will affect the ability of the Cluster Manager to send heartbeat messages across the network.
- Enabling an alternate Hardware Address for a service adapter using AIX SMIT.
- Taking down adapters: Do not take down all adapters on the same network if `network_down` is set up to perform a graceful down with takeover. If the cluster is customized to perform a graceful down with takeover when all communications on a specific network fail, and you take down all adapters, this will force the graceful down with takeover event to occur whether you intend it or not.
- Taking down adapters: Do not bring all adapters down on the same network if there is only one network and no serial point-to-point network is defined. Doing this will cause system contention between cluster nodes and failover attempts made by each node. On an HACMP system, you must force down all the nodes and then bring them back up. The cluster is unavailable while you fix the situation.

### **Making Changes to Network Interfaces**

In some circumstances, you can use the HACMP facility to swap a network service or boot adapter IP address dynamically, to an active standby adapter on the same node and network, without shutting down cluster services on the node. See *Swapping a Network Adapter Dynamically* on page 6-13 for more information.

You must normally stop the cluster to make any change to network interfaces. If you must change the IP address of an adapter, or if you change the adapter label, make sure to make the changes to both DNS or NIS *and* the `/etc/hosts` file. If DNS or NIS and `/etc/hosts` are not updated, you will be unable to synchronize the cluster nodes or do any DARE operations. If DNS or NIS services are interrupted, the `/etc/hosts` file is used for name resolution. You must also fix the `/.rhosts` file or redo `cl_setup kerberos` if you are using enhanced security.

### Handling Network Load/Error rates

Dropped packets due to network loads may cause false fallovers. Also, high throughput may cause the deadman switch to time-out. If either of these conditions occurs, check the AIX network options and the Failure Detection Rate you have set for the cluster. These parameters are contained in the Advanced Performance Tuning option of SMIT.

See Changing the Configuration of a Network Module on page 6-14 for information on tuning the Failure Detection Rate.

### Network Maintenance/Reconfiguration

Moving Ethernet ports on a running cluster will result in adapter swap or node failure. Even a brief outage will result in a cluster event.

## Shared Disk, Volume Group, and Filesystem Issues

Do not change the configuration of a shared volume group or filesystem using AIX, outside of HACMP. Any such action will affect cluster behavior. The Cluster Manager and the cluster event scripts assume the shared volume groups and filesystems are under HACMP control. If you change the environment, the event scripts will not be able to complete properly and you will get unexpected results. HACMP is very sensitive to filesystem mounting and the way volume groups are varied on. (HACMP is very sensitive to all configuration information.)

### Disk Issues

Disks should always be mirrored (or use a disk array), to protect against loss of data. Once they are defined and configured within the HACMP cluster, you should always use the HACMP C-SPOC utility (`smitty cl_admin`) to add or remove disks from a volume group with the cluster running. The cluster needs to be made aware of disks being added to or removed from a shared volume group. If you add or remove disks using the conventional method, the cluster will not be aware that these changes have occurred.

### Volume Group and Filesystems Issues

Use the C-SPOC utility (`smitty cl_admin`) for common maintenance tasks like creating, extending, changing, or removing a shared filesystem. See Chapter 4, Maintaining Shared LVM Components, in this Guide.

See Maintaining NFS on page 10-2 for information on NFS and HACMP.

- Do not set filesystems to automount; HACMP handles the mounts at startup and during cluster events.
- Do not set volume groups to autovaryon; HACMP executes the varying on and off as needed.

- If you are testing something when the cluster is not running and you vary on a volume group or mount a filesystem, remember to unmount the filesystem and vary off the volume group before you start HACMP.
- Do not have any processes running that would point to a shared filesystem when a graceful down occurs on the node that currently owns that filesystem. If a graceful down occurs and the application stop script fails to terminate the processes that are using the filesystem, that filesystem will be unable to unmount and the fallover will not occur. The cluster will go into a “reconfiguration too long” condition.

One of the more common reasons for a filesystem to fail being unmounted during a graceful down is because the filesystem is busy. In order to successfully unmount a filesystem, no processes or users can be accessing it at the time. If a user or process is holding it, the filesystem will be “busy” and will not unmount.

This is easy to overlook when you write stop application scripts. The script to stop an application should also include a check to make sure that the shared filesystems are not in use. You can do this by using the **fuser** command. The script should use the **fuser** command to see what processes or users are accessing the filesystems in question. The pids of these processes can then be acquired and killed. This will free the filesystem so it can be unmounted.

Refer to the AIX man pages for complete information on this command.

### **Expanding Filesystems**

Use C-SPOC to increase the size of a filesystem.

- Log in as root.
- Enter `smitty cl_fs`
- Select the option to change a cluster filesystem
- Select the filesystem to change
- Enter the new size for the filesystem.
- Synchronize the new definition to all cluster nodes via Sync Cluster Resources.

### **General Filesystems Issues**

**/etc/filesystem** concerns:

- Changes are made to stanzas in various files during startups, fallover, and other cluster events. You can set up a **crontab** job to monitor filesystem size to help avoid trouble (for example, the **hacmp.out** file can get quite large).
- Shared filesystems must have the *mount* option set to false, so that HACMP can mount and unmount them as needed to handle cluster events.
- Full filesystems in the root volume group may cause cluster events to fail. You should monitor this volume group and clean it up periodically.

**NFS Filesystem** concerns:

HACMP 4.4 includes improvements to the way NFS filesystems are handled. See Chapter 10, Additional Tasks: Configuring Run-Time Parameters, Cluster Log Files, and Maintaining NFS in this Guide.

## Application Issues

Appendix D, Applications and HACMP in the *HACMP Planning Guide* gives many pointers on planning and maintaining applications in the HACMP environment. You can monitor applications if you are running HACMP/ES 4.4. Some key points to remember:

- Application maintenance will require downtime for resource groups due to binaries residing on a shared disk.
- Upgrades should be tested prior to implementation to anticipate effects on the production cluster.
- Changes to application start and stop procedures should be thoroughly tested prior to going into production.
- Don't have shared applications already running when you start the cluster. A second attempt at starting already running applications may cause a problem.
- Don't manually execute the application stop script for any reason on a running cluster without starting the application back up again. Problems may occur if an attempt is made to stop the application that is already down. This could potentially cause a failover attempt to be unsuccessful.

---

## Hardware Maintenance

Hardware failures must be dealt with promptly to maintain no single point of failure in the cluster. If you have carefully set up error notification and event customization as recommended, you receive quick notification via email of any problems. You should also periodically do error log analysis. See Chapter 2, Examining Cluster Log Files in the *HACMP Troubleshooting Guide* for details on error log analysis.

Some issues to be aware of in a high availability environment include:

- Shared disks connect to both systems, thus open loops and failed disks can result in fragmented SSA loops and the loss of access to one mirror set.
- Set up mirroring so that the mirrored disk copy is accessible by a different controller. This will prevent loss of data access in the event of a disk controller failure. In the event that a disk controller fails, the mirror disk will be accessible through the other controller.

## Processor ID Licensing Issues

The Concurrent Resource Manager is licensed to the processor ID of a cluster node. Many of the **clvm** or concurrent access commands validate the processor ID against the license file. A mismatch will cause the command to fail, with an error message indicating the lack of a license.

Restoring a system image from a **mksysb** tape created on a different node or replacing the planar board on a node will cause this problem. In such cases, you must recreate the license file by removing and reinstalling the **cluster.clvm** component of the current release from the original installation images.

## Replacing Topology Hardware

Nodes, networks, and network adapters comprise the topology hardware. Changes to the cluster topology often involves down time on one or more nodes if changes to cabling or adding/removing adapters is involved. In most situations, you can use the DARE utilities to add a topology resource without down time.

The following sections indicate the conditions under which you can use DARE and the conditions under which you must plan cluster downtime.

See Chapter 6, Changing the Cluster Topology for full documentation on all procedures.

## Replacing Nodes

Using the DARE utility, you can add or remove a node while the cluster is running

### Replacing a Node or Node Component

If you are replacing a cluster node keep this list in mind:

- The new node must typically have the same amount of RAM (or more) as the original cluster node
- The new node must typically be same type of system if your applications are optimized for a particular processor
- The new node's slot capacity typically must be the same or better than the old node
- Adapter placement is important – use the same slots as originally assigned
- If you have a concurrent environment, you must reinstall the CRM software. This is also a consideration if you are cloning nodes.
- Get the new license key from the application vendor for the new CPU ID if necessary.

If you are replacing a component of the node:

- Be aware of CPU ID issues
- For SCSI adapter replacement – reset external bus SCSI ID to original SCSI ID
- Adapter replacement – use the same slots as originally assigned.

### Procedure for Adding or Removing a Node

See Changing the Configuration of Cluster Nodes on page 6-4 for more complete information.

The basic procedure is outlined here:

- Install AIX, HACMP and LPPs on new node
- Connect networks and SSA cabling and test
- Configure TCP/IP
- Import volume group definitions
- Connect serial network and test
- Change HACMP ODM configuration
- Synchronize and verify.

## Replacing Networks and Network Adapters

HACMP does not provide highly available networks. You can only protect your applications from downtime due to a network failure if you configure more than one IP network. You should also have a serial network. If no backup network is configured, the cluster will be inaccessible to all but directly connected clients.

You can replace network cabling without taking HACMP off line. You can also replace hubs, routers, and bridges while HACMP is running. Be sure to use the correct IP addresses when reconfiguring a router.

You can use the DARE `swap_adapter` function to swap the IP address from a boot or service adapter to an active standby adapter on the same node and network. Then you can service the failed adapter without stopping the node.

### Procedure for Replacing a LAN adapter

If the hardware supports hot-pluggable network adapters, no cluster downtime is required for this procedure.

A general procedure, if you cannot use the `swap_adapter` function:

- Move resource groups to another node using DARE
- Stop and detach the device driver for the failed adapter
- Remove the device from the ODM
- Shut down the node containing the failed adapter
- Boot into Diagnostics and test the failed adapter – the problem may be a faulty cable, for example.
- Replace the adapter if necessary (reconnecting drop cables)
- Reboot in normal mode (`cfgmgr` runs)
- Start the device driver for the adapter
- Assign IP addresses and netmasks (using `smitty chinnet`) for interfaces if they were undefined
- Test IP communications.

## Handling Disk Failures

Handling shared disk failures differs depending on the type of disk and whether it is a concurrent access configuration or not.

- SCSI non-RAID - you will have to shut down the nodes sharing the disks
- SCSI RAID - perhaps no downtime, depends on the capabilities of the array
- SSA non-RAID – requires manual intervention – you can replace disks with no system downtime
- SSA RAID – no downtime necessary.

See Replacing a Failed Drive in a Concurrent Access Volume Group on page 5-11 for information on that procedure.

## Replacing a Failed SSA non-RAID Disk

**Note:** AIX 4.3.3 has a new utility, **replacepv**. See the AIX manpage to see if you can use this utility to replace your failed disk.

This section describes the process of replacing a mirrored disk drive, using C-SPOC commands. The sample cluster is set up as follows:

- Two nodes, NodeB and NodeA
- Two cascading resource groups, each node has top priority for one of the resource groups.
  - NodeArg: NodeA, NodeB
  - NodeBrg: NodeB, NodeA.
- One shared Volume Group: vg1 (includes hdisk1 and hdisk2)
  - vg1 is included in the resource group NodeArg
  - vg1 has one mirrored logical volume (lv01) and one filesystem: (/fs1)
- The logical volume is mirrored on the two disks in the volume group:

```
>lslv -m lv01
lv01:/fs1
LP PP1 PV1 PP2 PV2 PP3 PV3
0001 0111 hdisk2 0056 hdisk1
```

- Initial disk configuration on both nodes:

| <u>hdisk#</u> | <u>pdisk#</u> | <u>vg defined?</u> |
|---------------|---------------|--------------------|
| hdisk1        | pdisk0        | vg1                |
| hdisk2        | pdisk1        | vg1                |
| hdisk3        | pdisk2        | None               |

When HACMP starts on the cluster, NodeA varies on vg1

```
>lsvg -o
NodeB:
NodeA: vg1
```

### Procedure to Replace Failed Disk

1. Unmirror a shared vg
  - a. From NodeB, execute:

```
smitty cl_admin / Cluster LVM/ Shared vg / Unmirror a shared vg
```
  - b. Select: NodeArg vg1  
the choices:

```
None
NodeA hdisk1
NodeA hdisk2
```
  - c. Select:

```
NodeA hdisk2
```
  - d. lslv on NodeA shows logical volume is no longer mirrored:

```
>lslv -m lv01
lv01:/fs1
LP PP1 PV1 PP2 PV2 PP3 PV3
0001 0056 hdisk1
```
2. Remove a physical volume from a shared vg

a. From NodeB execute:

```
smitty cl_admin /Cluster LVM/Shared vg /set vg chars/remove a pv from vg
```

b. Select: NodeArg vg1

the choices:

```
NodeA hdisk1
NodeA hdisk2
```

c. Select:

```
NodeA hdisk2
```

d. Force deallocation of physical partitions = yes

e. On both nodes, lspv shows hdisk2 is removed from vg1

```
hdisk1 vg1
hdisk2 none
hdisk3 none
```

3. Remove a disk from the cluster

a. From NodeB execute:

```
smitty cl_admin / Cluster Physical VM/ remove a disk from the cluster
```

b. Select:

```
NodeA
NodeB
```

the choices:

```
0cc6
0578
```

c. Select:

```
0578 (hdisk2)
```

d. Do not keep definition in database

e. On both nodes, lspv shows hdisk1 is removed.

```
hdisk1 vg1
hdisk3 none
```

f. On both nodes, ssaxlate shows pdisk1 is still present and unpaired:

```
hdisk1: pdisk0
hdisk3: pdisk2

pdisk0: hdisk1
pdisk1:
pdisk2: hdisk3
```

4. Pull failed disk and replace with a new disk.

5. Add a disk to the cluster:

a. From NodeB, execute:

```
smitty cl_admin / Cluster Physical VM/ add a disk to the cluster
```

b. Select:

```
NodeA
NodeB
```

c. Select:

```
hdisk ssar SSA logical disk drive
connection address: (from F4)
0004ac7c036500d
```

d. On both nodes, lspv shows new disk added as hdisk2

```
hdisk1 vg1
hdisk2 none
hdisk3 none
```

e. On both nodes, ssaxlate shows hdisk2 is unpaired with a pdisk.

```
hdisk1: pdisk0
hdisk2:
hdisk3: pdisk2
```

```
pdisk0: hdisk1
pdisk1:
pdisk2: hdisk3
```

6. Add a disk to a shared vg

a. Execute:

```
smitty cl_admin /Cluster LVM/shared vg/set chars of vg/add pv to vg
```

b. select:

```
NodeArg vg1
```

the choices:

```
NodeA hdisk2
NodeA hdisk3
```

c. select:

```
NodeA hdisk2
```

d. All data on physical volume will be destroyed. Continue? yes

e. On both nodes, lspv shows hdisk2 added to vg1

```
hdisk1 vg1
hdisk2 vg1
hdisk3 none
```

7. Mirror a shared vg:

a. Execute:

```
smitty cl_admin / Cluster LVM/ shared vg/ mirror a shared vg:
```

b. select:

```
NodeArg vg1
```

the choice:

```
None
NodeA hdisk1
NodeA hdisk2
```

c. select:

```
NodeA hdisk2
```

d. On NodeA (where vg is varied on) lslv shows lv is mirrored

```
lslv -m lv01
lv01: /fs1
LP PP1 PV1 PP2 PV2 PP3 PV3
0001 0056 hdisk1 0056 hdisk2
```

---

## Preventive Maintenance

It is highly recommended that you maintain a test cluster as well as your production cluster. Thus before you make any major change to the production cluster, you can test the procedure on the test cluster.

HACMP also supplies event emulation utilities to aid in testing.

### Cluster Snapshots

It is highly recommended to take periodic snapshots of the cluster in case you need to reapply a configuration. You should take a snapshot any time you change the configuration. Keep a copy of the snapshot on another system, off the cluster, as protection against loss of the cluster. You can use the snapshot to rebuild the cluster quickly in case of an emergency. See Chapter 11, Saving and Restoring Cluster Configurations in this Guide for complete information on cluster snapshots. You might want to consider setting up a **cron** job to do this on a regular basis.

### Backups

HACMP does not provide tools for backing up the system. You should plan for periodic backups just as you do for a single system. You should do backups of **rootvg** and shared volume groups.

Backups of shared volume groups should be done more frequently.

Some applications have their own online backup methods.

You can use any of the following:

- **mksysb** backups
- Online backups (**sysback**, **splitlvcopy**)

### Using mksysb

You should do a **mksysb** on each node prior to and following any changes to the node environment. Such changes include:

- Applying PTFs
- Upgrading AIX or HACMP software
- Adding new applications
- Adding new device drivers
- Changing TCP/IP configuration
- Changing cluster topology or resources
- Changing LVM components of **rootvg** (paging space, filesystem sizes)
- Changing AIX parameters (including the tuning parameters: I/O pacing, **syncd**)

## Shared Volume Group Backup—Using `splitlvcopy`

You can use the **splitlvcopy** method on raw logical volumes and filesystems to do a backup while the application is still running. This method is only possible for LVM mirrored logical volumes.

By taking advantage of the LVM's mirroring capability, you can stop the application briefly to split off a copy of the data using the AIX **splitlvcopy** command. Stopping the application gives the application its checkpoint. Then restart the application so it continues processing while you do a backup of the copy.

You can do the backup using **tar**, **cpio**, or any other AIX backup command that operates on a logical volume or a filesystem. Using **cron**, you can automate this type of backup.

## Using `cron`

Use the AIX **cron** utility to automate scheduled maintenance and to monitor the system.

### Using `cron` To Automate Maintenance Of Log Files

Use this utility to automate some of the administrative functions that need to be done on a regular basis. Some of the HACMP log files need **cron** jobs to ensure that they do not use up too much space.

Use **crontab -e** to edit `/var/spool/cron/crontabs/root`.

**Cron** will recognize the change without need for rebooting.

You might establish a policy for each log, depending how long you want to keep the log, and what size you will allow it to grow. **hacmp.out** is already set to expire every seven days.

The RSCT logs are stored in the `/var/ha/log` directory. These logs are trimmed regularly. If you want to save information for a longer period of time you can either redirect the logging to a different directory, or change the maximum size file parameter (using SMIT).

### Using `cron` To Set Up An Early Warning System

You can set up **clverify** to run as a **cron** job, sending email notification to the system administrator if **clverify** fails. The output from **clverify** then serves to notify where you might have configuration problems.

Use **cron** to set up jobs to proactively check out the system.

- Run a custom verification daily and send a report to the system administrator
- Check for full filesystems (and take action if necessary)
- Check that certain processes are running
- Run event emulation and send a report to the system administrator

## Do Regular Testing

It is recommended to regularly schedule a testing window where a failure is conducted in a controlled environment. That way you can evaluate a fallover before anything happens in your production cluster and people are anxiously awaiting the fix. It should include fallovers of all nodes and full verification of tested protected apps. This is strongly encouraged if you are changing or evolving your cluster environment.

## Upgrading Software (both AIX and HACMP)

Take a cluster snapshot and save it somewhere off the cluster.

Back up the operating system and data before performing any upgrade. Prepare a backout plan in case you encounter problems with the upgrade.

Whenever possible, plan and do an initial run through on a test cluster.

AIX patches need to be applied according to the HACMP operations guide:

- Apply patch to standby node.
- Fallover (graceful shutdown with takeover) to standby machine.
- Apply patch to primary node.

Follow this same general rule for patches to the application; follow specific instructions for the application.

See the install and upgrade procedures in the *HACMP Installation Guide* for complete instructions for upgrades, including node-by-node migration from HANFS 4.3.1 to HACMP 4.4.

**7x24 Maintenance**  
Preventive Maintenance

---

# Index

adding 8-8  
changing 8-9

## +-\*/\*

/.rhosts file 1-8  
/etc/hosts file 1-8, E-6, E-11  
/etc/inetd.conf file 1-8  
/etc/inittab  
    defined 1-8  
    entry for HACMP 2-4  
    IP address takeover 1-8  
/etc/rc.net script 1-9  
    for cluster startup 1-11  
/etc/resolv.conf file E-6  
/etc/services file 1-9  
/etc/snmp.peers file 1-10  
/etc/snmpd.conf file 1-9  
/etc/syslog.conf file 1-10  
/etc/tcp.clean file 1-10  
/etc/trcfmt file 1-10  
/tmp/hacmp.out file 3-29  
/usr/sbin/cluster/clinfo daemon  
    Clinfo 2-8, 2-10, 2-16  
    clinfo daemon 2-2, 2-8, 2-10  
/usr/sbin/cluster/clsmuxpd daemon 2-2  
    clsmuxpd daemon 2-8, 2-10  
    Cluster SMUX Peer 2-8, 2-10  
/usr/sbin/cluster/clstrmgr daemon 2-1  
/usr/sbin/cluster/etc/clhosts file 1-7  
    maintaining on clients 2-16  
    maintaining on nodes 2-16  
/usr/sbin/cluster/etc/rc.cluster script 1-11  
    starting clients 2-15  
/usr/sbin/cluster/events/network scripts 1-12  
/usr/sbin/cluster/events/node scripts 1-12  
/usr/sbin/cluster/events/swap\_adapters script 1-12  
/usr/sbin/cluster/events/utills directory B-1  
/usr/sbin/cluster/events/utills/cl\_deactivate\_nfs utility  
    10-7  
/usr/sbin/cluster/events/utills/cl\_nfskill command 10-7  
/usr/sbin/cluster/events/utills/convaryonvg command  
    5-9  
/usr/sbin/cluster/server.status 2-6  
/usr/sbin/cluster/snapshots/active.x.odm file  
    dynamic reconfiguration backup file 6-21  
/usr/sbin/cluster/utilities/cl\_clstop script 1-11  
/usr/sbin/cluster/utilities/cl\_rc.cluster script 1-11

/usr/sbin/cluster/utilities/clexit.rc script 1-11, 2-6  
/usr/sbin/cluster/utilities/clstart script 1-11  
/usr/sbin/cluster/utilities/clstop script 1-11  
/usr/share/man/cat1  
    HACMP for AIX man pages A-2  
/var/spool/cron/crontab/root file 1-10

## 0,1,2...

7 X 24 environment xiv, E-1

## A

active.n.odm file  
    dynamic reconfiguration backup file 6-21  
adapters  
    swapping dynamically 6-13  
adding  
    concurrent logical volume  
        using C-SPOC 5-24  
    disk definition to cluster nodes 4-45  
    JFS  
        using C-SPOC 4-23  
    network adapters 6-9  
    resource groups 7-9  
    user accounts 12-2  
AIX  
    files modified by HACMP 1-8  
AIX Connections  
    commands A-5, A-60  
AIX Fast Connect  
    converting from AIX Connections 7-12  
application servers  
    changing 7-5  
    defining 7-4  
    removing 7-4  
applications  
    customizing scripts E-5  
applying  
    cluster snapshots 11-6

## B

backing up  
    HACMP for AIX system 1-5  
backups E-19

## C

- cascading resource groups
  - and cldare stop command 7-21
  - changing priority of nodes dynamically 6-6
  - inactive takeover 7-15
  - NFS cross mounting issues 10-4
- CEL guide
  - CEL constructs C-5
  - writing execution plans C-1
- changing
  - cluster configuration 7-9
  - cluster environment properly E-11
  - cluster event processing 9-1
  - cluster ID or name 6-4
  - cluster topology 6-1
  - IP address properly E-11
  - network modules 6-15
  - nodename
    - clnodename command A-18
  - resource groups 7-10
  - user accounts 12-3, 12-4, 12-5
- changing cluster configuration
  - effects on components E-9
- cl\_9333clear utility B-1
- cl\_9333diskreset utility B-2
- cl\_9333fence utility B-2
- cl\_9333genodm utility B-4
- cl\_9333getbit utility B-4
- cl\_activate\_fs utility B-16
- cl\_activate\_nfs utility B-16
- cl\_activate\_vgs utility B-17
- cl\_array\_mode3 utility B-5
- cl\_chfs command A-23
- cl\_chgroup command A-24
- cl\_chlv command A-25
- cl\_chuser command A-29
- cl\_clstop command A-31
- cl\_convert utility A-5
- cl\_crfs command A-29
- cl\_crlvfs command A-30
- cl\_deactivate\_fs utility B-17
- cl\_deactivate\_nfs utility B-18
- cl\_deactivate\_vgs utility B-18
- cl\_disk\_available utility B-5
- cl\_echo utility B-20
- cl\_Eprimary\_app utility B-14
- cl\_export\_fs utility B-18
- cl\_extendvg command A-32
- cl\_fs2disk utility B-6
- cl\_get\_disk\_vg\_fs\_pvids utility B-6
- cl\_HPS\_Eprimary utility B-13
- cl\_HPS\_init utility B-14
- cl\_importvg command A-34
- cl\_is\_array utility B-7
- cl\_is\_scsidisk utility B-7
- cl\_log utility B-20
- cl\_lsfs command A-36
- cl\_lsgroup command A-36
- cl\_lslv command A-37
- cl\_lsuser command A-38
  - using 12-1
- cl\_lsvg command A-39
- cl\_mkgroup command A-42
- cl\_mkuser command A-49
  - using 12-2
- cl\_mvkg command
  - creating concurrent volume group 5-17
- cl\_nfskill command B-19
- cl\_nm\_nis\_off utility B-21
- cl\_nm\_nis\_on utility B-21
- cl\_nodecmd command A-51
- cl\_raid\_vg utility B-8
- cl\_rc.cluster command A-52
- cl\_reassign\_Eprimary utility B-15
- cl\_rmfs command A-54
- cl\_rmgroup command A-55
- cl\_rmlv command A-55
- cl\_rmlvcopy command A-56
- cl\_rmuser command A-57
- cl\_scsidiskreset utility B-8
- cl\_scsidiskrsrv utility B-9
- cl\_swap\_HPS\_IP\_address utility B-15
- cl\_swap\_HW\_address utility B-21
- cl\_swap\_IP\_address utility B-22
- cl\_sync\_vgs utility B-9
- cl\_synevg command A-57
- cl\_unmirrorvg command A-58
- cl\_unswap\_HW\_address B-23
- cl\_updatevg command A-59
  - updating ODM data on remote nodes 4-50
- clconvert\_snapshot utility A-6
- cldare command
  - syntax and flags B-25
- cldiag utility A-8
- clexit.rc script
  - /usr/sbin/cluster/clexit.rc script 2-6
- clfindres command 7-26, A-10
- clgetactivenodes command A-21
- clgetaddr command A-11
- clgetgrp command A-12
- clgetif command A-12
- clgodmget command A-13
- Clinfo
  - enabling traps 2-16
  - starting on clients 2-15
  - stopping on clients 2-15
- clinfo daemon 1-7, 2-2
- clinfo.rc script 1-12
- cllockd daemon 1-7, 2-2
  - starting 2-7, 2-10
- cllscf command A-13
- cllslstr command A-13
- cllsdisk command A-13

- cllsfs command A-14
- cllsgrp command A-14
- cllsif command A-14
- clslv command A-14
- cllsnim command A-15
- cllsnode command A-15
- cllsnw command A-16
- cllsparam command A-16
- cllsres command A-16
- cllsserv command A-17
- cllsvg command A-17
- clm\_stats command A-22
- clnodename command A-18
- clresactive command A-22
- clshowres command A-18
- clsmuxpd daemon 1-7, 2-2
- clsnapshot utility 11-1
- clstat utility
  - command syntax A-19
  - defined 3-2
  - multi-cluster mode 3-4
  - single-cluster mode 3-3
  - X Window display 3-5
- clstrmgr daemon 1-6, 2-1
  - starting 2-7, 2-10
- cluster
  - changing IDs 6-4
  - changing names 6-4
  - event processing 9-1
  - managing resources 7-9
  - monitoring 3-1
    - overview 3-1
  - reconfiguration 6-1
  - topology information 6-3
  - verifying configuration 8-1
- cluster configuration
  - custom verification methods
    - adding 8-8
    - changing 8-8
    - removing 8-9
  - verifying 8-2
- cluster events
  - emulating 9-5
  - pre- and post-event processing 9-1, 9-3
- Cluster Manager 1-6
- cluster monitoring
  - clstat utility 3-5
  - with Tivoli 3-18
    - deinstalling 3-27
    - prerequisites 3-19
- cluster services
  - defined 2-1
  - starting
    - on a single node 2-7
    - on clients 2-15
  - stopping
    - on clients 2-15
    - on nodes 2-11
- cluster snapshot
  - applying 11-6
  - backup files 11-8
  - changing 11-8
  - contents 11-1
  - creating 11-5
  - cron jobs E-19
  - custom methods
    - defining 11-4, 11-6
  - description section 11-6
  - format 11-3
  - memory requirement 11-5
  - naming 11-6
  - removing 11-9
  - reverting to previous configuration 11-8
- cluster topology
  - synchronizing 6-19
    - skipping clverify during 6-19
  - verifying 8-2, 8-4
  - viewing 6-2
- cluster verification
  - custom methods
    - adding 8-8
    - changing 8-8
    - removing 8-9
  - overview 8-1
  - skip during synchronization 7-29
  - ways to run clverify 8-1
- cluster.mmdd file 3-29
- clverify utility
  - cluster verification 8-4, 8-5, 8-6
  - command syntax A-20
  - custom methods
    - adding, changing 8-8
    - removing 8-9
  - overview 8-1
  - run as cron job E-20
  - software verification 8-3, 8-4
  - ways to run 8-1
- Command Execution Language
  - CEL guide C-1
- command-line arguments
  - encoding/decoding C-SPOC commands C-3
- commands
  - AIX Connections A-5, A-60
  - C-SPOC A-3, A-23
  - HACMP A-1, A-5
- communication system E-10
- concurrent access mode
  - convaryonvg command 5-9
  - devices 5-1
  - HACMP for AIX scripts 5-2
  - maintaining disk subsystems 5-1
  - maintaining RAID devices 5-1
  - varyonvg command 5-2



- custom verification methods
  - adding 8-8
  - changing 8-8
  - removing 8-9
- customizing
  - 7 X 24 maintenance E-2

## D

- daemons
  - checking status 2-9
  - clinfo 2-2
  - clockd 2-2
  - clsmuxpd 2-2
  - clstrmgr 2-1
  - Cluster Manager 1-6
  - HACMP for AIX (list) 1-6
  - handling properly under HACMP E-9
  - monitoring 3-28
    - on clients 3-29
    - on cluster nodes 3-28
  - snmpd 2-2
- DARE Resource Migration
  - example 7-21
  - overview 7-1, 7-17
  - using SMIT 7-22
  - using the command line 7-19
- DARE utility
  - use for hardware maintenance E-14
- DCD
  - restoring from ACD 6-21
- Deadman Switch
  - formula 6-14
  - timeouts per network 6-15
- debug levels
  - setting 10-1
- default location keyword 7-19
- defining
  - CS/AIX communications links 7-6
- deleting
  - application server 7-4
  - cluster nodes 6-7
    - dynamically 6-6
  - CS/AIX communications links 7-7
  - resource groups 7-10
- DGSP message
  - avoiding E-6
- DHCP
  - allocating IP addresses E-7
- disk drive
  - replacing E-16
- disk failures
  - handling E-15
- disk fencing
  - in dynamic reconfiguration 6-6
- disk layout
  - planning issues E-7

- disks
  - 9333 serial disks 5-1
  - defining to cluster using C-SPOC 4-45
  - planning issues E-7
  - RAID devices
    - concurrent access 5-1
  - SSA disk subsystems
    - concurrent access mode 5-1

## DNS

- integrating with HACMP E-6
- dynamic adapter swap 6-13
- dynamic reconfiguration
  - adding a cluster node 6-5
  - effect of disk fencing 6-6
  - effect on resources 7-28, 11-7
  - of cluster resources 7-2
  - of cluster topology 6-1
  - overview 1-1
  - releasing the SCD lock 6-19
  - restoring the DCD from the ACD 6-21
  - scripts 1-12
  - triggered by applied snapshot 11-6
  - triggering 6-19
  - undoing 6-21

## E

- editing
  - snmpd.conf file 2-16
- emulating
  - cluster events 9-1, 9-5
  - dynamic reconfiguration events 7-3
- enabling
  - asynchronous event notification 2-16
- Eprimary node
  - SP utilities B-13
- error notification
  - customizing E-3
- events
  - changing custom events processing 9-2, 9-3
  - emulating 9-1, 9-5
  - maintaining scripts 9-1
  - predefined (list) 1-6
- extending
  - shared volume groups 4-8

## F

- failure cycle
  - tuning 6-14
- failure detection rate
  - changing 6-14, 6-15
- fallover
  - intentional 2-6
- Fast Connect
  - converting from AIX Connections 7-12

## Index

### G – L

- files
  - log 3-29
- filesystems
  - changing
    - using AIX commands 4-25
  - mount failures E-12
  - removing
    - using AIX commands 4-29
    - using C-SPOC 4-27
  - shared
    - maintaining 4-20
- forced stops
  - on nodes 2-6
- fuser command
  - using in scripts E-12

## G

- get\_local\_nodename command A-21
- graceful stops
  - on nodes 2-6
  - with takeover 2-6
- groups
  - adding 12-6
  - changing 12-6
  - listing 12-5
  - managing 12-5
  - removing 12-7
  - using C-SPOC to manage 12-1

## H

- HACMP for AIX
  - backing up 1-5
  - commands
    - commonly used commands A-1
    - syntax conventions A-1
  - C-SPOC commands A-23
  - log files 3-29
  - loopback address on local node 1-8
  - scripts 1-10
    - configuration 1-12
    - shutdown 1-10
    - startup 1-10
  - software daemons (list) 1-6
  - starting
    - on nodes 2-1
  - starting cluster services
    - on clients 2-15
    - overview 2-1
  - stopping
    - on clients 2-15
    - on nodes 2-11
  - stopping cluster services
    - overview 2-1
  - viewing man pages A-2

- HACMP log files
  - using cron to maintain E-20
- harc.net entry
  - for IP address takeover 1-8
- hardware
  - guidelines for maintenance E-8
- hardware errors
  - list of errors to monitor E-3
- hardware maintenance
  - proper procedures E-13
- HAView
  - browsers 3-16
  - changing polling intervals 3-13
  - monitoring a cluster 3-7
  - read-only NetView maps 3-11
  - starting 3-9
  - symbols for components 3-11
- heartbeat rate
  - tuning 6-14

## I

- importing volume groups
  - using C-SPOC 4-14
- importvg command
  - with shared LVM components 5-7
- intentional fallover 2-6
- IP address
  - changing properly E-11
- IP address takeover
  - /etc/inittab file 1-8

## J

- jfslog
  - renaming 4-21

## K

- keepalives
  - tuning 6-15
- Kerberos
  - verifying with clverify 8-2
- kill – 9 command
  - warning E-9

## L

- LAN adapter
  - replacing E-15
- lazy update
  - overview 4-4, 4-49
- listing
  - concurrent volume groups 5-18
- local node
  - loopback address required in /rhosts 1-8

- log files 3-29
  - recommended use 3-29
  - redirecting 10-2
  - system error log 3-29
- log logical volume
  - renaming 4-21
- logical volumes
  - adding copies 4-21
  - adding mirror copies to 4-39
  - changing
    - using AIX commands 4-35
    - using C-SPOC 4-34
  - creating 4-31
  - maintaining 4-31
  - removing
    - using AIX commands 4-37
    - using C-SPOC 4-37, 5-25
  - renaming 4-21
- loopback address
  - required to run C-SPOC commands 1-8
- LVM
  - forcing an update of ODM data on remote nodes 4-50
  - maintaining shared components 4-1
    - using C-SPOC 4-2
  - updating ODM definitions on remote nodes 4-4, 4-49

## M

- maintaining
  - 7 X 24 cluster E-1
  - an HACMP for AIX system
    - common tasks 1-2
    - overview 1-1
  - clhost file
    - on clients 2-16
  - cluster event scripts 9-1
  - concurrent access environments 5-1
  - NFS 10-2
  - shared LVM components 4-1
    - using C-SPOC 4-2
- man pages
  - stored in /usr/share/man/cat1 directory A-2
  - using the man command A-2
- Management Information Base (MIB)
  - and clsmuxpd 1-7
- managing
  - cluster resources 7-9
- migrating resources
  - default location keyword 7-19
  - example using cldare 7-21
  - non-sticky migration 7-18
  - sticky migration 7-18
  - stop location keyword 7-19
  - using SMIT 7-22
  - using the cldare command 7-20

- mirroring
  - jfslog 4-21
  - logical volumes 4-21
- mksysb backups E-19
- modifying
  - LVM
    - HACMP environment 4-1
- monitoring
  - cluster 3-1, 3-2
  - cluster daemons
    - on nodes and clients 3-28
  - cluster services 3-28
  - daemons 3-28
  - network interfaces 3-2
  - node status 3-2
  - nodes and network interfaces 3-2
  - overview 3-1
- mounting
  - NFS 10-4

## N

- name resolution
  - integrating with HACMP E-6
- NetView
  - dialog boxes 3-13
  - using HAView 3-7
- network adapters
  - adding 6-9
  - monitoring 3-2, 3-7
  - showing cluster topology information by 6-3
- Network Information Services (NIS) 10-1
- network interfaces
  - making changes E-10
- network loads
  - handling problems E-11
- network modules
  - adding to cluster 6-17
  - changing or showing 6-15
  - configuring 6-14
  - removing from cluster 6-18
- network services
  - integrating with HACMP E-6
- networks
  - NFS mounting filesystems and directories 7-14
  - replacing hardware E-15
- NFS
  - creating shared volume groups 10-3
  - maintaining 10-2
  - mount issues 10-4
  - mounting filesystems 10-4
  - nested mount points 10-5
  - node name issues 10-7
  - reliable server functionality 10-7, 10-8
  - setting up mount points for cascading groups 10-5
  - takeover issues 10-4

## Index

O – R

### NIS

- integrating with HACMP E-6
- setting run-time parameter 10-1

### node\_down\_remote

- concurrent access environment 5-2

### node\_up\_local event

- concurrent access environment 5-2

### nodes

- adding to cluster 6-8
- adding to cluster dynamically 6-5
- changing run-time parameters 10-1
- deleting dynamically 6-6
- monitoring 3-2
- naming 6-8
- priority in cascading configurations 7-9
- procedure for replacing hardware E-14
- removing from cluster 6-7
- verifying configuration 8-1

### non-sticky migration 7-18

## O

### ODM

- restoring the DCD 6-21

### Online backups E-19

## P

### physical volume

- shared
  - removing using C-SPOC 4-46
  - adding to cluster using C-SPOC 4-45
  - maintaining 4-45
  - migrating data 4-47

### planning

- for 7 X 24 maintenance E-2
- for software maintenance E-8

### polling interval

- HAView 3-13

### post-processing

- cluster events 9-1

### pre-processing

- cluster events 9-1

### preventive maintenance E-19

### priorities

- nodes in cascading resource configurations 7-9

### procedure

- replace failed SSA disk E-16

## R

### RAID devices

- concurrent access environment 5-1
- importing volume groups 5-8

### rc.cluster script

- starting clients 2-15

### reconfiguring

- cluster event processing 9-1
- cluster topology 6-1
- resources for a resource group 7-13

### reintegration

- of cluster nodes 2-8

### reliable NFS server 10-7

### removing

- application servers 7-4
- cluster nodes
  - dynamically 6-6
- cluster snapshot 11-9
- concurrent logical volume
  - using C-SPOC 5-25
- CS/AIX communications links 7-7
- filesystems
  - using C-SPOC utility 4-3
- logical volumes
  - using C-SPOC utility 4-3
- physical volume from concurrent volume group
  - using C-SPOC 5-20
- resource groups 7-10
- shared filesystem 4-27

### renaming

- cluster log files 10-2
- logical volumes 4-21

### replacing

- cluster node E-14
- LAN adapter E-15
- mirrored disk drive E-16
- network hardware E-15
- topology hardware E-14

### resource groups

- adding 7-9
- changing
  - changing to or from rotating 7-11
  - resources in 7-12
- dynamic reconfiguration of 7-2
- relation to C-SPOC utility 4-3
- removing 7-10

### resources

- changes you can make dynamically 7-2
- changing in resource group 7-13
- effect of dynamic reconfiguration 7-28, 11-7
- managing 7-9
- migrating dynamically 7-20
- synchronizing 7-27, 7-29

### rootvg

- planning volume groups E-7

### RS/6000 SP B-13

### run-time maintenance E-9

## S

- SCD
  - removing 6-19
- scripts
  - HACMP for AIX 1-10
    - configuration 1-12
    - shutdown 1-10
    - startup 1-10
    - node events 9-1
    - setting debug levels 10-1
  - scsidiskutil utility B-10
  - server.status file (see /usr/sbin/cluster/server.status) 2-6
- shared
  - filesystem
    - changing 4-25
    - creating 4-20
    - maintaining 4-20
  - logical volume
    - creating 4-31
    - maintaining 4-31
  - LVM components
    - maintaining 4-1
    - maintaining with C-SPOC 4-2
  - physical volume
    - maintaining 4-45
    - migrating data 4-47
  - volume groups
    - creating 4-6
    - deleting 4-12
    - extending 4-8
    - maintaining 4-4
    - NFS issues 10-3
    - reducing 4-10
- shared volume groups
  - creating with TaskGuide 4-4
- showing
  - cluster definitions 6-3
  - detailed topology information 6-3
  - network adapters 6-12
- smit clshow fastpath 3-28
- smit clstart fastpath 2-7
- smit clstop fastpath 2-11
- SNA network
  - reconfiguring CS/AIX communication links 7-6
- snapshot *See* cluster snapshot
- SNAPSHOTPATH environment variable 11-9
- snmpd daemon 2-2
- snmpd.conf file
  - editing 2-16
- software maintenance
  - planning E-8
- software verification 8-3, 8-4
- SP Switch B-13
- SP Utilities B-13
- splitlvcopy utility E-19
- SRC 1-11
  - System Resource Controller 2-2
- SSA Disk Subsystems
  - in concurrent access environment 5-1
- SSA disks
  - procedure for replacing E-16
- ssa\_clear utility B-12
- ssa\_clear\_all utility B-12
- ssa\_configure utility B-13
- ssa\_fence utility B-11
- start\_clmarketdemo utility B-24
- start\_imagedemo utility B-24
- starting
  - cluster services 2-2
    - on clients 2-15
    - using C-SPOC 2-3, 2-9
  - daemons E-9
  - NetView/HAView 3-9
- sticky migration 7-18
  - removing sticky markers 7-26
- stop location keyword 7-19
- stop\_clmarketdemo utility B-24
- stop\_imagedemo utility B-25
- stopping
  - cluster services
    - on clients 2-15
    - on nodes 2-6
    - on single node 2-11
    - understanding 2-4
    - using C-SPOC 2-5, 2-13
  - daemons E-9
  - the cluster E-9
- stopsrc command
  - stop cluster services 2-12
  - stopping Clinfo on clients 2-15
  - stopping cluster services 2-12
- swap address feature 6-13
- synchronizing
  - cluster topology 6-19
  - concurrent LVM mirrors 5-27
  - concurrent volume group mirrors 5-23
  - prevented by SCD lock 6-20
  - resources 7-27, 7-29
  - shared LVM Mirrors 4-18
  - skipping cluster verification during 6-19
- syntax conventions
  - HACMP for AIX commands A-1
- sysback utility E-19
- system error log file 3-29
- System Resource Controller (SRC) 1-11, 2-2

## Index

T – U

### T

- takeover
  - NFS issues 10-4
- TaskGuide for creating shared volume groups 4-4
- TCP/IP services
  - proper procedure for stopping E-10
- timeout
  - Deadman Switch 6-15
- Tivoli, cluster monitoring
  - cluster monitoring with Tivoli 3-18
  - deinstalling 3-27
  - prerequisites 3-19
- topology
  - reconfiguring dynamically 6-1
  - verifying with clverify 8-2
- topology hardware
  - replacing E-14
- traps
  - Clinfo 2-16
- tuning
  - failure detection rate 6-15
  - network module 6-15

### U

- user accounts
  - adding 12-2
  - changing 12-3
  - creating 12-1
  - removing 12-4
  - using C-SPOC to manage 12-1

### utilities

- 9333fence B-2
- cl\_9333clear B-1
- cl\_9333diskreset B-2
- cl\_9333genodm B-4
- cl\_9333getbit B-4
- cl\_activate\_fs B-16
- cl\_activate\_nfs B-16
- cl\_activate\_vgs B-17
- cl\_array\_mode3 B-5
- cl\_deactivate\_fs B-17
- cl\_deactivate\_nfs B-18
- cl\_deactivate\_vgs B-18
- cl\_disk\_available B-5
- cl\_echo B-20
- cl\_Eprimary\_app B-14
- cl\_export\_fs B-18
- cl\_fs2disk B-6
- cl\_getdisk\_vg\_fs\_pvids B-6
- cl\_HPS\_Eprimary B-13
- cl\_HPS\_init B-14
- cl\_is\_array B-7
- cl\_is\_scsidisk B-7
- cl\_log B-20
- cl\_nfskill B-19
- cl\_nm\_nis\_off B-21
- cl\_nm\_nis\_on B-21
- cl\_nodectmd A-51
- cl\_raid\_vg B-8
- cl\_reassign\_Eprimary B-15
- cl\_scdiskreset B-8
- cl\_scsidiskrsrv B-9
- cl\_swap\_HPS\_IP\_address B-15
- cl\_swap\_HW\_address B-21
- cl\_swap\_IP\_address B-22
- cl\_sync\_vgs B-9
- cl\_unswap\_HW\_address B-23
- clhare B-25
- cllsvg A-17
- clm\_stats A-22
- convaryonvg B-10
- script B-1
- scsidiskutil B-10
- ssa\_clear B-12
- ssa\_clear\_all B-12
- ssa\_configure B-13
- ssa\_fence B-11
- start\_clmarketdemo B-24
- start\_imagedemo B-24
- stop\_clmarketdemo B-24
- stop\_imagedemo B-25

## **V**

- varyonvg command
  - in concurrent access mode 5-2
- verification
  - errors ignored during synchronization 6-19, 7-27, 11-7
- verifying
  - cluster configuration
    - overview 8-1
    - using SMIT 8-7
  - cluster topology
    - using SMIT 8-7
  - resources 8-6
  - software 8-3
- Versatile Storage Servers
  - concurrent access mode 5-1
  - concurrent access volume group 5-5
- viewing
  - details about cluster
    - HAView 3-13
- volume groups
  - activating
    - concurrent access mode 5-9
  - determining access mode 5-10
  - dormant at startup 5-8
  - planning issues E-7
  - shared
    - creating 4-6
    - creating with TaskGuide 4-4
    - deleting 4-12
    - extending 4-8
    - maintaining 4-4
    - reducing 4-10
- VSM (Visual System Management)
  - overview D-1

## **XYZ**

- X Window System display
  - using clstat 3-5
- xclstat 3-5
- xhacmpm application
  - overview D-1



## Vos remarques sur ce document / Technical publication remark form

**Titre / Title :** Bull HACMP 4.4 Administration Guide

**N° Référence / Reference N° :** 86 A2 57KX 02

**Daté / Dated :** August 2000

### ERREURS DETECTEES / ERRORS IN PUBLICATION

### AMELIORATIONS SUGGEREES / SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION

Vos remarques et suggestions seront examinées attentivement.

Si vous désirez une réponse écrite, veuillez indiquer ci-après votre adresse postale complète.

Your comments will be promptly investigated by qualified technical personnel and action will be taken as required.

If you require a written reply, please furnish your complete mailing address below.

NOM / NAME : \_\_\_\_\_ Date : \_\_\_\_\_

SOCIETE / COMPANY : \_\_\_\_\_

ADRESSE / ADDRESS : \_\_\_\_\_

Remettez cet imprimé à un responsable BULL ou envoyez-le directement à :

Please give this technical publication remark form to your BULL representative or mail to:

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

# Technical Publications Ordering Form

## Bon de Commande de Documents Techniques

To order additional publications, please fill up a copy of this form and send it via mail to:

Pour commander des documents techniques, remplissez une copie de ce formulaire et envoyez-la à :

**BULL CEDOC**  
**ATTN / MME DUMOULIN**  
**357 AVENUE PATTON**  
**B.P.20845**  
**49008 ANGERS CEDEX 01**  
**FRANCE**

**Managers / Gestionnaires :**  
**Mrs. / Mme :** C. DUMOULIN +33 (0) 2 41 73 76 65  
**Mr. / M :** L. CHERUBIN +33 (0) 2 41 73 63 96  
**FAX :** +33 (0) 2 41 73 60 19  
**E-Mail / Courrier Electronique :** srv.Cedoc@franp.bull.fr

Or visit our web site at: / Ou visitez notre site web à:

<http://www-frec.bull.com> (PUBLICATIONS, Technical Literature, Ordering Form)

| CEDOC Reference #<br>N° Référence CEDOC | Qty<br>Qté | CEDOC Reference #<br>N° Référence CEDOC | Qty<br>Qté | CEDOC Reference #<br>N° Référence CEDOC | Qty<br>Qté |
|-----------------------------------------|------------|-----------------------------------------|------------|-----------------------------------------|------------|
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |
| ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            | ___ - ___ - ___ - [__]                  |            |

[\_\_]: no revision number means latest revision / pas de numéro de révision signifie révision la plus récente

NOM / NAME : \_\_\_\_\_ Date : \_\_\_\_\_

SOCIETE / COMPANY : \_\_\_\_\_

ADRESSE / ADDRESS : \_\_\_\_\_

PHONE / TELEPHONE : \_\_\_\_\_ FAX : \_\_\_\_\_

E-MAIL : \_\_\_\_\_

**For Bull Subsidiaries / Pour les Filiales Bull :**

Identification: \_\_\_\_\_

**For Bull Affiliated Customers / Pour les Clients Affiliés Bull :**

**Customer Code / Code Client :** \_\_\_\_\_

**For Bull Internal Customers / Pour les Clients Internes Bull :**

**Budgetary Section / Section Budgétaire :** \_\_\_\_\_

**For Others / Pour les Autres :**

**Please ask your Bull representative. / Merci de demander à votre contact Bull.**



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

**ORDER REFERENCE  
86 A2 57KX 02**

PLACE BAR CODE IN LOWER  
LEFT CORNER



Utiliser les marques de découpe pour obtenir les étiquettes.  
Use the cut marks to get the labels.

