IBM Administering

Administering

Contents

Highlighting ... 7
Case-sensitivity in AIX® ... 7
ISO 9000 ... 7
Related information ... 7

Administering PowerHA® SystemMirror® ... 8

What's new in Administering PowerHA® SystemMirror® ... 8
How to see what's new or changed ... 8
July 2025 ... 8
December 2024 ... 8
December 2023 ... 8
December 2022 ... 8
June 2022 ... 9
December 2021 ... 9
December 2020 ... 9
December 2019 ... 9
November 2019 ... 10
September 2019 ... 10
June 2019 ... 10
April 2019 ... 10
December 2018 ... 10
Administering a PowerHA® SystemMirror® cluster ... 11
Options for configuring a PowerHA® SystemMirror® cluster ... 11
Configuration tasks ... 11
Maintaining a PowerHA® SystemMirror® cluster ... 13
Monitoring the cluster ... 14
AIX® files modified by PowerHA® SystemMirror® ... 15
/etc/syslog.conf and /etc/rsyslog.conf files ... 19
Changing script behavior for unrecoverable errors in PowerHA® SystemMirror® ... 21
PowerHA® SystemMirror® and AIX® commands ... 21
Configuring a PowerHA® SystemMirror® cluster ... 21
Overview of configuring a cluster ... 21
Configuring a cluster using Smart Assists ... 24
Defining PowerHA® SystemMirror® cluster topology ... 25
Configuring PowerHA® SystemMirror® resources ... 25
Configuring PowerHA® SystemMirror® resource groups ... 30
Configuring resources in resource groups ... 31
Verifying and synchronizing the standard configuration ... 33
Viewing the PowerHA® SystemMirror® configuration ... 34
Configuring split and merge policies ... 35
Configuring a quarantine policy ... 39
Additional cluster configuration ... 40
Understanding custom cluster configuration options ... 40
Discovering PowerHA® SystemMirror®-related information ... 40
Cluster, nodes, and networks ... 40
Configuring PowerHA® SystemMirror® resources ... 50
Configuring PowerHA® SystemMirror® resource groups ... 70
Configuring resource groups ... 71
Limitations and prerequisites for configuring resource groups ... 71
Configuring resource groups by using SMIT ... 71
Dynamic node priority policies ... 74
Configuring resource group run-time policies ... 75
Configuring dependencies between resource groups ... 75
Adding resources and attributes to resource groups ... 89
Reliable NFS function ... 94
Forcing a varyon of volume groups ... 96
Running a resource group in an AIX® WPAR ... 97
Testing your configuration ... 99
Configuring cluster events ... 100
Considerations for pre-event and post-event scripts ... 100
Configuring pre-event and post-event commands ... 101
Configuring pre-event and post-event processing ... 102
Tuning event duration time until warning ... 103
Configuring a custom remote notification method ... 105
Verifying and synchronizing a PowerHA® SystemMirror® cluster ... 109
Running cluster verification ... 110
Automatic verification and synchronization ... 110
Verifying the PowerHA® SystemMirror® configuration using SMIT ... 113
Inactive components report ... 120

Managing PowerHA® SystemMirror® file collections....121
Adding a custom verification method....127
List of reserved words....128
Verifying two different clusters by using the CCV utility....129
Requirements for using the CCV utility....129
Data inputs to the CCV utility....130
Comparing configuration attributes by using the CCV utility....130
Output from the CCV utility....130
Usage scenarios....130
Testing a PowerHA® SystemMirror® cluster....131
Overview for testing a cluster....131
Running automated tests....133
Understanding automated testing....134
Setting up custom cluster testing....137
Description of tests....139
Running custom test procedures....152
Evaluating results....153
Recovering the control node after cluster manager stops....154
Error logging....154
Fixing problems when running cluster tests....159
Starting and stopping cluster services....161
Starting cluster services....162
Stopping cluster services....166
Maintaining cluster information services....170
Monitoring a PowerHA® SystemMirror® cluster....172
Periodically monitoring a PowerHA® SystemMirror® cluster....172
Monitoring clusters with clstat....173
Monitoring applications....180
Displaying an application-centric cluster view....181
Measuring application availability....182
Using the cldisp command....185
Using PowerHA® SystemMirror® topology information commands....187
Monitoring cluster services....187
PowerHA® SystemMirror® log files....188
Modifying the log file size by using SMIT....193
Managing shared LVM components....193
Shared LVM overview....193
Understanding C-SPOC....194
Maintaining shared volume groups....196
Maintaining logical volumes....212
Maintaining shared file systems....219
Maintaining raw volumes....222
Maintaining physical volumes....223
Configuring LVM split-site mirroring....229
Managing shared LVM components in a concurrent access environment....231
Understanding concurrent access and PowerHA® SystemMirror® scripts....232
Maintaining concurrent volume groups with C-SPOC....232
Maintaining concurrent access volume groups....234
Managing the cluster topology....235
Reconfiguring a cluster dynamically....235
Viewing the cluster topology....236
Managing communication interfaces in PowerHA® SystemMirror®....237
Adding PowerHA® SystemMirror® site definitions....242
Changing the host name for a cluster node in PowerHA® SystemMirror® 7.1.2, or earlier....243
Changing the host name for a cluster node in PowerHA® SystemMirror®....244
Changing how PowerHA® SystemMirror® 7.1.3, or later, responds to host name change....244
Changing the IP address for a PowerHA® SystemMirror® cluster node....245
Changing a cluster name....246
Changing the configuration of cluster nodes....246
Changing the configuration of a PowerHA® SystemMirror® network....247
Changing the configuration of communication interfaces....249
Managing persistent node IP labels....250
Synchronizing the cluster configuration....251
Dynamic reconfiguration issues and synchronization....251
Managing the cluster resources....253
Reconfiguring a cluster dynamically....253
Requirements before reconfiguring....254
Reconfiguring application controllers....254
Changing or removing application monitors....255
Reconfiguring service IP labels as resources in resource groups....257
Reconfiguring tape drive resources....259
Using NFS with PowerHA® SystemMirror®....260
Reconfiguring resources in clusters with dependent resource groups....260
Synchronizing cluster resources....261
Managing resource groups in a cluster....262
Changing a resource group....262
Moving resource groups....273
Bringing a resource group online....275

Taking a resource group offline 276
Checking resource group state 277
Special considerations when stopping a resource group 277
Example: Using clRGmove to swap resource groups 277
Managing users and groups 278
Overview for AIX® and LDAP users and groups 278
Managing AIX® and LDAP user accounts across a cluster 279
Managing password changes for users 282
Changing the password for your own user account 285
Managing AIX® and LDAP group accounts 286
Managing cluster security 289
Configuring cluster security 289
Configuring PowerHA® SystemMirror® with IP security filter rules 289
Standard security mode 291
Configuring message authentication and encryption 292
PowerHA® SystemMirror® federated security 298
Planning for federated security 298
Installing federated security 299
Configuring federated security 299
Managing PowerHA® SystemMirror® federated security 302
Removing PowerHA® SystemMirror® federated security 302
Troubleshooting PowerHA® SystemMirror® federated security 303
Saving and restoring cluster configurations 304
Information saved in a cluster snapshot 304
Format of a cluster snapshot 305
clconvert_snapshot utility 306
Defining a custom snapshot method 306
Changing or removing a custom snapshot method 306
Creating a snapshot of the cluster configuration 307
Restoring the cluster configuration from a snapshot 307
Changing a snapshot of the cluster configuration 309
Removing a snapshot of the cluster configuration 310
7x24 maintenance 310
Planning for 7x24 maintenance 310
Runtime maintenance 316
Hardware maintenance 319
Preventive maintenance 321
Resource group behavior during cluster events 323
Resource group event handling and recovery 324
Selective fallover for handling resource groups 327
Handling of resource group acquisition failures 330
Recovering resource groups when nodes join the cluster 332
Handling of resource groups configured with IPAT via IP aliases 332
Examples of location dependency and resource group behavior 333
Using DLPAR and CoD in a PowerHA® SystemMirror® cluster 347
Overview of DLPAR and CoD 347
PowerHA® SystemMirror® integration with the CoD function 349
Types of CoD licenses 351
Resource Optimized High Availability in PowerHA® SystemMirror® 353
Application provisioning in PowerHA® SystemMirror® 366
Using pre-event and post-event scripts 371
Resource Optimized High Availability in Cloud 371
Planning for ROHA in Cloud 372
Configuring Cloud instances for cluster nodes to work with ROHA 373
Changing hardware provisioning for an application controller in ROHA in Cloud configuration 375
Troubleshooting ROHA in Cloud 376
SAP high availability management with PowerHA® SystemMirror® 377
SAP high availability infrastructure 377
SAP liveCache Hot Standby with PowerHA® SystemMirror® 377
PowerHA® SystemMirror® SAP liveCache Hot Standby wizard 377
Live Partition Mobility 384
Configuring SAN communication with LPM 385
Performing LPM operation on a cluster node and modifying LPM variables 386
Backing up and restoring an application data in PowerHA® SystemMirror® by using cloud storage 387
Planning for backing up data by using cloud storage in PowerHA® SystemMirror® 387
Configuring PowerHA® SystemMirror® for cloud storage 389
Configuring backup profiles 391
Restoring backup profiles 393
Understanding network instability 394

Notices 395

Privacy policy considerations 396
Trademarks 396

Index 398

Note

Before using this information and the product it supports, read the information in “Notices” on page 395.

This edition applies to IBM® PowerHA® SystemMirror® 7.2 Standard Edition for AIX® and to all subsequent releases and modifications until otherwise indicated in new editions.

About this document

This document provides information about how you can configure, maintain, and monitor clusters with PowerHA® SystemMirror® for AIX®.

Highlighting

The following highlighting conventions are used in this document:

Bold	Identifies 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.
Italics	Identifies parameters whose actual names or values are to be supplied by the user.
`Monospace`	Identifies examples of specific data values, examples of text similar to what you might see displayed, examples of portions of program code similar to what you might write as a programmer, messages from the system, or information you should actually type.

Case-sensitivity in AIX®

Everything in the AIX® operating system is case-sensitive, which means that it distinguishes between uppercase and lowercase letters. For example, you can use the **ls** command to list files. If you type `LS`, the system responds that the command `is not found`. Likewise, **FILEA**, **FiLea**, and **filea** are three distinct file names, even if they reside in the same directory. To avoid causing undesirable actions to be performed, always ensure that you use the correct case.

ISO 9000

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

Related information

- The PowerHA® SystemMirror® Version 7.2 for AIX® PDF documents are available in the PowerHA SystemMirror 7.2 PDFs topic.
- The PowerHA® SystemMirror® Version 7.2 for AIX® release notes are available in the PowerHA SystemMirror 7.2 release notes topic.

Administering PowerHA® SystemMirror®

Use this information to configure, manage, and troubleshoot PowerHA® SystemMirror®.

What's new in Administering PowerHA® SystemMirror®

Read about new or significantly changed information for the Administering PowerHA® SystemMirror® topic collection.

How to see what's new or changed

In this PDF file, you might see revision tags (>| and |<) surrounding new and changed information.

July 2025

The following information is a summary of the updates that were made to the PowerHA SystemMirror Version 7.2.9 Service Pack 1:

- Updated the following topics related to the SAP NetWeaver NFS resource improvements:
  - “Removing a shared file system using C-SPOC” on page 222
  - “Adding the file system to a PowerHA SystemMirror cluster logical volume” on page 221
  - “Creating shared file systems with C-SPOC” on page 220

December 2024

The following information is a summary of the updates that were made to this topic collection:

- Added notes about configuring any global filesystem, either Network File System version 4 (NFSv4) or GPFS in the “SAP high availability management with PowerHA SystemMirror” on page 377 topic.
- Added new topics, Enabling or disabling encryption for physical volumes using C-SPOC and Listing encrypted physical volumes with C-SPOC under “Maintaining physical volumes” on page 223.
- Added new topics Removing a raw disk definition on cluster nodes using C-SPOC, Enabling or disabling encryption for raw volumes using C-SPOC and Renaming raw volumes with C-SPOC about raw volume monitor features under “Maintaining raw volumes” on page 222.

December 2023

The following information is a summary of the updates that were made to this topic collection:

- Added a note about HMC ports. For more information, see. Configuring HMC or PowerVM® NovaLink to work with Resource Optimized High Availability.
- Added new Volume Group fields to the Creating a shared volume group by using C-SPOC topic.
- Updated Split and Merge Management Policy fields in the Configuring split and merge policies topic.
- Added new topic and subtopics, Enabling Safeguard function for a volume group using C-SPOC under Maintaining shared volume groups
- Added notes about Remote notification method and system events. For more information, see. Defining a remote notification method
- Updated command to reach Configure Remote Notification Methods panel. For more information see. Changing a remote notification method
- Added note about using RPM fileset to install Python 3.x.x. and related Boto packages. For more information see. “Planning for backing up data by using cloud storage in PowerHA SystemMirror” on page 387

December 2022

The following information is a summary of the updates that were made to this topic collection:

- Cluster Aware AIX® (CAA) supports a new tunable PowerVM® Watchdog Timer in IBM® AIX® 7.3 with the Technology Level 1, or later. The PowerVM® Watchdog Timer tunable is supported in PowerHA® SystemMirror® Version 7.2.7, or later. For more information, see. "Configuring cluster heartbeat settings" on page 41.
- The following topics were updated for Cloud Backup Management (CBM) and Cloud Backup Restoration features:
  - "Backing up and restoring an application data in PowerHA SystemMirror by using cloud storage" on page 387
  - "Restoring backup profiles" on page 393
  - "Troubleshooting ROHA in Cloud" on page 376
  - "Resource Optimized High Availability in Cloud" on page 371
- Starting with PowerHA® SystemMirror® Version 7.2.7, or later, Resource Optimization High availability (ROHA) in Cloud support in IBM® Power® Virtual Server (PowerVS) is available. To know more, see "Resource Optimized High Availability in Cloud" on page 371.

June 2022

The following information is a summary of the updates that were made to this topic collection:

- Added a function Resource Optimized High Availability in Cloud in PowerHA® SystemMirror® that automatically and dynamically manages Dynamic Logical Partitions (DLPARs) resources.

December 2021

The following information is a summary of the updates that were made to this topic collection:

- Added the Changing a logical volume to a cluster using C-SPOC topic that explains how to change a logical volume to a cluster by using the C-SPOC utility.
- Updated the following topics:
  - "Adding a logical volume to a cluster using C-SPOC" on page 212
  - "Changing or displaying the characteristics of a volume group" on page 203
  - "Creating a shared volume group by using C-SPOC" on page 200

December 2020

The following information is a summary of the updates that were made to this topic collection:

- Added the topic "Configuring the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable" on page 48 that explains the configuration of CRIT_DAEMON_RESTART_GRACE_PERIOD tunable.
- Updated the Split and Merge Management Policy field in the "Configuring split and merge policies" on page 35 topic.
- Updated the following topics:
  - "Starting cluster services on a node with a resource group in the UNMANAGED state" on page 165
  - "Live Partition Mobility" on page 384

December 2019

- Added the topic Verifying two different clusters by using the CCV utility that explains the requirements and data inputs for using the cross-cluster verification (CCV) utility, and how to use the CCV utility to compare two different clusters.
- Federated security is enhanced to enable PowerHA® SystemMirror® commands that are not already supported for role-based access control (RBAC). The following topic collection is updated for this enhancement:
  - "PowerHA SystemMirror federated security" on page 298
  - "Planning for federated security" on page 298
  - "Installing federated security" on page 299
  - "Configuring LDAP clients" on page 301

- “Removing PowerHA SystemMirror federated security” on page 302

November 2019

The following information is a summary of the updates that were made to this topic collection:

- Added a note about steps that must be completed before you update to a new AIX® Technology Level in the “AIX Live Update for PowerHA SystemMirror nodes” on page 41 topic.
- Updated the supported files and directories formats for file collections in the “Creating a PowerHA SystemMirror file collection” on page 123 topic.

September 2019

- Added resource allocation order and methods to verify and synchronize the cluster in the “Changing the default cluster tunables” on page 363 topic.
- Added a note about the methods to verify and synchronize the cluster in the “Configuring HMC or PowerVM NovaLink to work with Resource Optimized High Availability” on page 357 topic.

June 2019

The following cloud backup management topics were updated:

- “Configuring backup profiles” on page 391
- “Configuring PowerHA SystemMirror for cloud storage” on page 389
- “Planning for backing up data by using cloud storage in PowerHA SystemMirror” on page 387

April 2019

The following information is a summary of the updates that were made to this topic collection:

- Updated information about changing the boot IP address in the “Changing the IP address for a PowerHA SystemMirror cluster node” on page 245 topic.
- Added a note about the`Long-running`monitoring mode in the “Steps for configuring a custom application monitor” on page 63 topic.
- Updated information about placing resource groups in an unmanaged state in the “Stopping PowerHA SystemMirror cluster services without stopping applications” on page 168 topic.
- Added information about changing the startup policy for a resource group in the “Limitations and prerequisites for configuring resource groups” on page 71 topic.
- Updated information about target resource groups in the “Configuring a start after dependency between resource groups” on page 78 and “Configuring a stop after dependency between resource groups” on page 79 topics.

December 2018

The following information is a summary of the updates that were made to this topic collection:

- Added information about backing up application data by using a cloud storage solution in the following topics:
  - “Backing up and restoring an application data in PowerHA SystemMirror by using cloud storage” on page 387
  - “Planning for backing up data by using cloud storage in PowerHA SystemMirror” on page 387
  - “Configuring PowerHA SystemMirror for cloud storage” on page 389
  - “Configuring backup profiles” on page 391
- Added information about the `/etc/rsyslog.conf` file in the “/etc/syslog.conf and /etc/rsyslog.conf files” on page 19 topic.
- Added information about the **Enable CPU usage statistics** and **CPU usage monitor interval** SMIT options in the “Configuring application controllers” on page 25 topic.
- Added information about the **Instability Threshold** and **Instability Period** SMIT options in the “Configuring networks” on page 45 topic.

- Added information about changing volume group characteristics in the “Changing or displaying the characteristics of a volume group” on page 203 topic.
- Added information about getting notifications for pre-event and post-event scripts in the “Considerations for pre-event and post-event scripts” on page 100 topic.
- Added information about the **Mirror pool name** and **Storage Location** SMIT options in the “Creating a shared volume group by using C-SPOC” on page 200 topic.
- Updated information about pre-event and post-event commands in the “Configuring pre-event and post-event commands” on page 101 topic.

Administering a PowerHA® SystemMirror® cluster

These topics provide a list of the tasks you perform to configure, maintain, monitor, and troubleshoot a PowerHA® SystemMirror® system, related administrative tasks, and a list of AIX® files modified by PowerHA® SystemMirror®.

Options for configuring a PowerHA® SystemMirror® cluster

In PowerHA® SystemMirror®, you can configure a cluster using one of the several different PowerHA® SystemMirror® tools.

These tools include:

- PowerHA® SystemMirror® SMIT user interface. You can also use the SMIT menus under the **Cluster Nodes and Networks › Initial Cluster Setup (Typical)** path to configure a typical cluster. You can alternately use the menus under **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Setup (Custom)** to create a custom configuration.
- *Cluster Snapshot Utility*: If you have a snapshot of the PowerHA® SystemMirror® cluster configuration taken from a prior release, you can use the Cluster Snapshot utility to perform the initial configuration.

Configuration tasks

The PowerHA® SystemMirror® configuration tasks are described in these topics. You can reach configuration tasks related to the cluster, nodes, and networks from the Cluster Nodes and Networks SMIT menu. You can reach tasks related to resources and applications from the Cluster Applications and Resources SMIT menu.

The major steps in the process are:

1. Configure the cluster topology using the Cluster Nodes and Networks SMIT menu path.
2. Configure the cluster applications and resources using the Cluster Applications and Resources SMIT menu path.
3. Verify and synchronize your cluster configuration.
4. (Optional) Perform custom configuration of your cluster, such as configuring pre-events and post-events, remote notification, file collections, and other optional settings. You must verify and synchronize if you make any additional changes to the cluster configuration.
5. Test the cluster.

Initial cluster setup

Using the options under the Cluster Nodes and Networks SMIT menu, you can add the basic components of the cluster to the PowerHA® SystemMirror® Configuration Database in a few steps. This configuration path automates the discovery and selection of configuration information and chooses default behaviors.

The following are the prerequisites and default settings for configuring a cluster:

- PowerHA® SystemMirror® uses unicast communications by default for heartbeat and messaging between nodes in the cluster. You can choose to use multicast communication. If you use multicast communication, you must verify that the network devices are configured for multicast communication. When you create a

cluster that uses multicast communication, PowerHA® SystemMirror® uses a default multicast IP address for your environment or you can specify a multicast IP address.

- Connectivity for communication must already be established between all cluster nodes. Automatic discovery of cluster information runs by default when you use the Initial Cluster Setup (Typical) menus, found under the SMIT menu **Cluster Nodes and Networks**. Once you have specified the nodes to add and their established communication paths, PowerHA® SystemMirror® automatically collects cluster-related information and configures the cluster nodes and networks based on physical connectivity. All discovered networks are added to the cluster configuration.
- The cluster configuration is stored in a central repository disk, and PowerHA® SystemMirror® assumes that all nodes in the cluster have common access to at least one physical volume or disk. This common disk cannot be used for any other purpose, such as hosting application data. You can specify this dedicated shared disk when you initially configure the cluster.

Cluster and application configuration options

Use the following information to configure cluster and application components after the initial cluster setup.

Configuring topology and resources

You can configure the cluster, nodes, networks, network interfaces, and the cluster repository disk and IP address using the SMIT menus under **Cluster Nodes and Networks**.

Once the cluster is created, manage the cluster, nodes, networks, and network interfaces using the **Manage** menus under **Cluster Nodes and Networks** in SMIT.

You can add resources and resource groups to support clustered applications under the SMIT menu **Cluster Applications and Resources**.

There are some options for configuring cluster topology and resource that are not required, but might be desired in some configurations. For example, by default all network interfaces discovered on cluster nodes at the time of cluster creation are included in the cluster topology configuration and are used for cluster communications, monitoring, and to keep application IP addresses highly available. It is possible to exclude some interfaces from the cluster configuration when desired.

Configuring dynamic LPAR and Capacity Upgrade on Demand resources

Using Dynamic LPAR (DLPAR) and Capacity on Demand (CoD) in a PowerHA® SystemMirror® cluster describes how to plan, integrate, configure, and troubleshoot application provisioning for PowerHA® SystemMirror® through the use of DLPAR CoD functions available on some IBM® Power Systems™ servers. It also includes examples and recommendations about customizing your existing pre- and post-event scripts.

Custom cluster configuration

Use the Custom Cluster Configuration menus to access less typical configuration options for the cluster and applications. To access this menu, enter `smit sysmirror` and select **Custom Cluster Configuration**.

Cluster Nodes and Networks

Use the **Initial Cluster Setup (Custom)** option to perform a custom setup by manually creating the cluster and adding nodes, networks, and network interfaces. Use the **Manage the Cluster** option to customize the cluster startup settings to start cluster services automatically when the system starts, customize cluster heartbeat settings, and reset the cluster tunable parameters to the default values.

Sites

Add sites, remove nodes from a site, and change site settings.

Resources

These menus are available for custom resource configuration and include the following submenus: Define custom disk, volume group, and file system methods, Configure user defined resources and types, and Customize resource recovery.

Resource Groups

Configure resource groups for PowerHA® SystemMirror®, and assign resources to resource groups.

Events

Customize cluster events by adding pre and post-event commands, notify commands, recovery commands, user-defined events, and remote notification methods by using the **Cluster Event** menu. This menu also has the options to allow you to change the default time until warning values. If you want to customize the responses to system events use the **System Events** menu.

Verify and Synchronize Cluster Configuration (Advanced)

Verify cluster topology, verify cluster resources, and identify custom verification methods.

Configuring cluster events

The PowerHA® SystemMirror® 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.

To configure customized cluster events, you indicate the script that handles the event and any additional processing that should accompany an event. Configuring cluster events describes the procedures for customization of event handling in PowerHA® SystemMirror®.

Configuring remote notification for cluster events

The remote notification function allows you to direct SMS text-message notifications to any address including your cell phone.

In addition to sending e-mails or SMS messages to your cell phone, remote notification methods can be used to send numeric or alphanumeric pages through a dialer modem, which uses the standard Telocator Alphanumeric Protocol (TAP) protocol.

Verifying and synchronizing the configuration

Verifying the cluster configuration assures you that all resources used by PowerHA® SystemMirror® are properly configured, and that ownership and takeover of those resources are defined and are in agreement across all nodes. By default, if the verification is successful, the configuration is automatically synchronized.

You should verify the configuration after making changes to a cluster or node. The Verifying and synchronizing a PowerHA® SystemMirror® cluster section describes the SMIT menus for verification, explains the contents and uses of the **clverify.log** file, and describes how to verify your cluster.

Verifying and synchronizing a PowerHA® SystemMirror® cluster also explains how to create and maintain PowerHA® SystemMirror® File Collections. Using the PowerHA® SystemMirror® File Collections utility, you can request that a list of files is automatically kept synchronized across the cluster. You no longer have to manually copy an updated file to every cluster node, verify that the file is properly copied, and confirm that each node has the same version of it. If you use the PowerHA® SystemMirror® File Collections utility, PowerHA® SystemMirror® can detect and warn you if one or more files in a collection is deleted or has a zero value on one or more cluster nodes during cluster verifications.

Testing the cluster

PowerHA® SystemMirror® includes the Cluster Test Tool to help you test the recovery procedures for a new cluster before the cluster becomes part of your production environment.

You can also use the tool to test configuration changes in an existing cluster, when the cluster services are not running. Testing a PowerHA® SystemMirror® cluster explains how to use the Cluster Test Tool.

Maintaining a PowerHA® SystemMirror® cluster

PowerHA® SystemMirror® systems have different maintenance tasks.

Starting and stopping cluster services

Various methods for starting and stopping cluster services are available.

Maintaining shared logical volume manager components

Any changes to logical volume components must be synchronized across all nodes in the cluster. Using C-SPOC (the Cluster Single Point of Control) to configure the cluster components on one node and then synchronize the cluster saves you time and effort.

Managing the cluster topology

Any changes to cluster configuration must be propagated across all nodes. Managing the cluster topology describes how to modify cluster topology after the initial configuration. You can make most changes on one node and then synchronize the cluster.

Managing cluster resources

Any changes to cluster resources require updating the cluster across all nodes. You can make most changes on one node and then synchronize the cluster.

Managing cluster resource groups

The Managing resource groups in a cluster section describes how to modify cluster resource groups after the initial configuration. You can add or delete resources and change the runtime policies of resource groups.
You can dynamically migrate resource groups to other nodes and take them online or offline by using the Resource Group Management utility (clRGmove) from the command line or by using SMIT.

Managing users and groups in a cluster

PowerHA® SystemMirror® allows you to manage user accounts for a cluster from a Single Point of Control (C-SPOC). Use the C-SPOC SMIT panels on any node to create, change, or remove users and groups from all cluster nodes by executing a C-SPOC command on any single cluster node.

Managing cluster security and inter-node communications

You can protect access to your PowerHA® SystemMirror® cluster by setting up security for cluster communications between nodes.

Understanding the /etc/cluster/rhosts file

The /etc/cluster/rhosts file
A Cluster Communications daemon (**clcomd**) runs on each PowerHA® SystemMirror® node to transparently manage inter-node communications for PowerHA® SystemMirror®.

In other words, PowerHA® SystemMirror® manages connections for you automatically:

- Populate the **/etc/cluster/rhosts** file with the host names or IP addresses of each node that will form the cluster.
- The **clcomd** command validates the addresses of the incoming connections to ensure that they are received from a node in the cluster. The rules for validation are based on the presence and contents of the **/etc/cluster/rhosts** file.
- If the **/etc/cluster/rhosts** file is not present, **clcomd** rejects all connections

After you synchronize the cluster, you can empty the **/etc/cluster/rhosts** file (but not remove it), because the information present in the PowerHA® SystemMirror® Configuration Database would be sufficient for all future connections.

The ~/.rhosts File

PowerHA® SystemMirror® does not use native AIX® remote execution (rsh) so you do not need to configure a **~/.rhosts** file unless you intend to use Workload Partitions (WPAR) which have their own requirements on this file.

Saving and restoring PowerHA® SystemMirror® cluster configurations

After you configure the topology and resources of a cluster, you can save the cluster configuration by taking a cluster snapshot. This saved configuration can later be used to restore the configuration if this is needed by applying the cluster snapshot. A cluster snapshot can also be applied to an active cluster to dynamically reconfigure the cluster.

Monitoring the cluster

By design, failures of components in the cluster are handled automatically, but you need to be aware of all such events.

Monitoring a PowerHA® SystemMirror® cluster describes various tools you can use to check the status of a PowerHA® SystemMirror® cluster, the nodes, networks, and resource groups within that cluster, and the daemons that run on the nodes.

The PowerHA® SystemMirror® software includes the Cluster Information Program (Clinfo), based on SNMP. The PowerHA® SystemMirror® for AIX® software provides the PowerHA® SystemMirror® for AIX® MIB, associated with and maintained by PowerHA® SystemMirror®. Clinfo retrieves this information from the PowerHA® SystemMirror® for AIX® Management Information Base (MIB).

The Cluster Manager gathers information relative to cluster state changes of nodes and interfaces. The Cluster Information Program (Clinfo) gets this information from the Cluster Manager and allows clients communicating with Clinfo to be aware of a cluster's state changes. This cluster state information is stored in the PowerHA® SystemMirror® MIB.

Clinfo runs on cluster server nodes and on PowerHA® SystemMirror® client machines. It makes information about the state of a PowerHA® SystemMirror® cluster and its components available to clients and applications via an application programming interface (API). Clinfo and its associated APIs enable you to write applications that recognize and respond to changes within a cluster.

Although the combination of PowerHA® SystemMirror® and the high availability features built into the AIX® system keeps single points of failure to a minimum, there are still failures that, although detected, can cause other problems.

AIX® files modified by PowerHA® SystemMirror®

The following AIX® files are modified to support PowerHA® SystemMirror®. They are not distributed with PowerHA® SystemMirror®.

- /etc/hosts
- /etc/inittab
- /etc/services
- /etc/snmpdv3.conf file
- /etc/snmpd.conf file
- /etc/syslog.conf and /etc/rsyslog.conf files
- /var/spool/cron/crontabs/root

/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.
PowerHA® SystemMirror® might modify this file to ensure that all nodes have the necessary information in their `/etc/hosts` file, for proper PowerHA® SystemMirror® operations.

If you delete service IP labels from the cluster configuration by using SMIT, you must also them from the `/etc/hosts` file. Removing service IP label from the `/etc/hosts` file reduces the possibility of having conflicting entries if the labels are reused with different addresses in a future configuration.

DNS and NIS are not available during PowerHA® SystemMirror®-related name resolution. Hence, PowerHA® SystemMirror® IP addresses must be maintained locally.

/etc/inittab

The **/etc/inittab** file is modified in the following cases:

- PowerHA® SystemMirror® is configured for IP address takeover.
- The **Start at System Restart** option is chosen on the SMIT **System Management (C-SPOC) › PowerHA SystemMirror Start Cluster Services** panel.
- The `/etc/inittab` file has the following entry in the `/user/es/sbin/cluster/etc/rc.init`:

  ```
  hacmp:2:once:/usr/es/sbin/cluster/etc/rc.init
  ```

  This entry starts the PowerHA® SystemMirror® communications daemons, **clcomd**, and the **clstrmgr** subsystem.

Modifications to the /etc/inittab file due to IP address takeover

The following entry is added to the `/etc/inittab` file for PowerHA® SystemMirror® network startup with IP address takeover:

```
harc:2:wait:/usr/es/sbin/cluster/etc/harc.net # PowerHA SystemMirror network startup
```

Modifications to the /etc/inittab file due to system restart

The /etc/inittab file is used by the **init** process to initialize processes during system restart.

When the system restarts, the /etc/inittab file calls the /usr/es/sbin/cluster/etc/rc.cluster script to start PowerHA® SystemMirror®. The entry is added to the /etc/inittab file if the **Start at system restart** option is selected in the SMIT **System Management (C-SPOC) > PowerHA® SystemMirror® Services > Start Cluster Services** panel or when the system restarts.

The following entry starts the PowerHA® SystemMirror® communications daemons, **clcomd** and **clstrmgr**.

```
hacmp:2:once:/usr/es/sbin/cluster/etc/rc.init
```

Because some of the daemons that are started through the rc.tcpip script are required during the system restart, PowerHA® SystemMirror® adds an **inittab** entry for the **harc.net** script with a run level of 2. The **harc.net** script starts the following daemons when you restart the system:

syslogd
portmap
inetd

The **harc.net** script also has code to start the following NFS-related daemons:

nfsd
rpc.mountd
rpc.statd
rpc.lockd

The code to start the NFS-related daemons is commented out, and you can uncomment if needed.

The **syslogd**, **portmap**, and **inetd** daemons are common to both the **rc.tcpip** script and the **harc.net** script. However, there is always the possibility that the NFS-related daemons are added to the **rc.tcpip** script by the user.

For more information about the files that are involved in starting and stopping PowerHA® SystemMirror®, see Starting and stopping cluster services.

/etc/services

The **/etc/services** file defines the sockets and protocols that are used for network services on a system. The ports and protocols that are used by the PowerHA® SystemMirror® components are defined here.

```
clinfo_deadman 6176/tcp
clinfo_client  6174/tcp
clsmuxpd       6270/tcp
clm_lkm        6150/tcp
clm_smux       6175/tcp
godm           6177/tcp
topsvcs        6178/udp
grpsvcs        6179/udp
emsvcs         6180/udp
clcomd         6191/tcp
```

/etc/snmpdv3.conf file

The Simple Network Management Protocol (SNMP) version 3 is the default version that is used by the AIX® operating system. You can configure SNMP version 3 with the /etc/snmpdv3.conf file.
The SNMP daemon reads the /etc/snmpdv3.conf file when it starts and when a **refresh** command or a **kill** command is used. The /etc/snmpdv3.conf file specifies the community names and associated access privileges and views, hosts for trap notification, logging attributes, parameter configurations, and SMUX configurations for the **snmpd** daemon.

The PowerHA® SystemMirror® installation process adds a clsmuxpd password at the /etc/snmpdv3.conf file. The following line is added to the end of the /etc/snmpdv3.conf file to include the Management Information Base (MIB) variables for PowerHA® SystemMirror® that are supervised by the Cluster Manager:

```
smux  1.3.6.1.4.1.2.3.1.2.1.5 clsmuxpd_password # PowerHA SystemMirror clsmuxpd
```

The /usr/es/sbin/cluster/clstat utility and the /usr/es/sbin/cluster/utilities/cldump utility do not work if the internet MIB tree is not enabled in the /etc/snmpdv3.conf file. These utilities use the risc6000clsmuxpd (1.3.6.1.4.1.2.3.1.2.1.5) MIB subtree.

To enable the risc6000clsmuxpd (1.3.6.1.4.1.2.3.1.2.1.5) MIB subtree in the /etc/snmpdv3.conf file, complete the following steps:

1. From the command line, enter `vi /etc/snmpdv3.conf`.
2. On a new line in the /etc/snmpdv3.conf file, add the following entry: `VACM_VIEW defaultView 1.3.6.1.4.1.2.3.1.2.1.5 - included -`
3. To stop the **snmpd** daemon on the hosts where you changed the /etc/snmpdv3.conf file, enter `stopsrc -s snmpd` from the command line.
4. To start the **snmpd** daemon on the hosts where you changed the /etc/snmpdv3.conf file, enter `startsrc -s snmpd` from the command line.

If your system is running SNMP version 3, the community name is found in the VACM_GROUP entry in the /etc/snmpdv3.conf file.

The **clinfo** daemon also receives the SNMP community name by using the same process. You can use the **-c** flag with the **clinfo** daemon to specify an SNMP community name.

> **Note:** If you want to protect the SNMP community name, do not use the **-c** flag with the **clinfo** daemon. If you use the **-c** flag, an unauthorized user might use the **ps** command to identify the SNMP community name. To protect the SNMP community name, change the permissions on the following files so that these files cannot be read by unauthorized users:
>
> - /etc/snmpd.conf
> - /smit.log
> - /usr/tmp/snmpd.log
> - /var/hacmp/log/hacmp.out
> - /etc/snmpd.peers

/etc/snmpd.conf file

Simple Network Management Protocol (SNMP) version 3 is the default version that is used for the AIX® operating system. However, you can use SNMP version 1 and configure it with the /etc/snmpd.conf file. You can switch from SNMP version 3 to SNMP version 1 by using the **snmpv3_ssw** command.

The SNMP daemon reads the /etc/snmpdv3.conf file when it starts and when a **refresh** command or a **kill** command is used. The /etc/snmpd.conf file specifies the community names and associated access privileges and views, hosts for trap notification, logging attributes, parameter configurations, and SMUX configurations for the **snmpd** daemon.

The PowerHA® SystemMirror® installation process adds a `clsmuxpd` password to the /etc/snmpd.conf file. The following line is added at the end of the /etc/snmpd.conf file to include the PowerHA® SystemMirror® Management Information Base (MIB) variables that are supervised by the Cluster Manager:

```
smux  1.3.6.1.4.1.2.3.1.2.1.5 clsmuxpd_password # PowerHA SystemMirror clsmuxpd
```

The SNMP version 1 community name is the first name found that is not `private` or `system` in the output of the **lssrc -ls snmpd** command.

The **clinfo** daemon also receives the SNMP community name by using the same process. You can use the **-c** flag with the **clinfo** daemon to specify an SNMP community name.

> **Note:** If you want to protect the SNMP community name, do not use the **-c** flag with the **clinfo** daemon. If you use the **-c** flag, an unauthorized user might use the **ps** command to identify the SNMP community name. To protect the SNMP community name, change the permissions on the following files so that these files cannot be read by unauthorized users:

- /etc/snmpd.conf
- /smit.log
- /usr/tmp/snmpd.log
- /var/hacmp/log/hacmp.out
- /etc/snmpd.peers

/etc/snmpd.peers

The /etc/snmpd.peers file configures **snmpd** SMUX peers.
During installation, PowerHA® SystemMirror® adds the following entry to include the **clsmuxpd** password to this file:

```
clsmuxpd  1.3.6.1.4.1.2.3.1.2.1.5  "clsmuxpd_password" # PowerHA SystemMirror/ES for AIX clsmuxpd
```

/etc/syslog.conf and /etc/rsyslog.conf files

During the installation process, PowerHA® SystemMirror® reads the subsystem. Depending on the subsystem, following entries are added to the file that directs the output from PowerHA® SystemMirror® 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                       *
#  *.debug              /tmp/syslog.out     rotate size 100k files 4
#  *.crit               /tmp/syslog.out     rotate time 1d
local0.crit /dev/console
local0.info /var/hacmp/adm/cluster.log
user.notice /var/hacmp/adm/cluster.log
daemon.notice /var/hacmp/adm/cluster.log
```

If you want to use the /rtc/rsyslogd.conf file after installing PowerHA® SystemMirror®, you can run the following command to convert the existing /etc/syslog.conf file into the /etc/rsyslog.conf file on all the cluster nodes.

```
/usr/sbin/syslog_ssw -c /etc/syslog.conf /etc/rsyslog.conf
```

Then, convert the *syslogd* daemon by running the following command:

```
syslog_ssw -r
```

An output that is similar to the following example is displayed:

```


0513-077 Subsystem has been changed.
Start daemon: syslogd
0513-059 The syslogd Subsystem has been started. Subsystem PID is 4456860.
```

You can view the following entries in the /etc/rsyslog.conf file and you might also view more entries based on your /etc/rsyslog.conf configuration file.

```


aso.notice /var/log/aso/aso.log
aso.info /var/log/aso/aso_process.log
aso.debug /var/log/aso/aso_debug.log
caa.debug;caa. /var/adm/ras/syslog.caa .info /var/adm/ras/syslog.txt
local0.info;user.notice;daemon.notice /var/hacmp/adm/cluster.log
```

Irrespective of the *syslogd* daemon type, the **lssrc** command always displays the state of the subsystem as *syslog*. To determine the type of *syslogd* daemon that is enabled, run the following commands:

- **ps -ef | grep syslog**

An output similar to the following example is displayed:

```
root 26869770 4128770 0 04:19:37 - 0:00 /usr/sbin/rsyslogd
```

- **odmget -q "subsysname = 'syslogd'" SRCsubsys**
  An output similar to the following example is displayed:

```
SRCsubsys:
subsysname = "syslogd"
synonym = ""
cmdargs = ""
path = "/usr/sbin/rsyslogd" <<<< rsyslogd is enabled
uid = 0
auditid = 0
standin = "/dev/console"
standout = "/dev/console"
standerr = "/dev/console"
action = 1
multi = 1
contact = 3
svrkey = 0
svrmtype = 0
priority = 20
signorm = 0
sigforce = 0
display = 1
waittime = 20
grpname = "ras"
```

Note:

- The `/etc/rsyslog.conf` file must be identical on all cluster nodes.
- PowerHA® SystemMirror® does not support rotation of log files when the *rsyslogd* daemon is enabled because the *rsyslogd* daemon requires special mechanism to rotate the log files.

/var/spool/cron/crontabs/root

The `/var/spool/cron/crontabs/root` file contains commands that are needed for basic system control. The installation process adds PowerHA® SystemMirror® log file rotation to the file.
During the installation process, PowerHA® SystemMirror® adds entries to this file that direct the output from PowerHA® SystemMirror® related problems to certain files.

```
0 0 * * * /usr/es/sbin/cluster/utilities/clcycle 1>/dev/null 2>/dev/null # PowerHA SystemMirror for AIX Logfile rotation
```

/etc/syslog.conf and /etc/rsyslog.conf files

The `/etc/syslog.conf` and `/etc/rsyslog.conf` files are used to control the output of the *syslogd* daemon log files, which Cluster Aware AIX uses to log the debug information and PowerHA® SystemMirror® uses to log the non-critical information.

During the installation process, PowerHA® SystemMirror® reads the subsystem. Depending on the subsystem, following entries are added to the file that directs the output from PowerHA® SystemMirror®-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                       *
#  *.debug              /tmp/syslog.out     rotate size 100k files 4
#  *.crit               /tmp/syslog.out     rotate time 1d
local0.crit /dev/console
local0.info /var/hacmp/adm/cluster.log
user.notice /var/hacmp/adm/cluster.log
daemon.notice /var/hacmp/adm/cluster.log
```

If you want to use the `/rtc/rsyslogd.conf file` after installing the PowerHA® SystemMirror®, you can run the following command to convert the existing `/etc/syslog.conf file` into the `/etc/rsyslog.conf` file on all the cluster nodes.

/usr/sbin/syslog_ssw -c /etc/syslog.conf /etc/rsyslog.conf

Then, convert the *syslogd* daemon by running the following command:

syslog_ssw -r

An output similar following example is displayed:

```
0513-077 Subsystem has been changed.
Start daemon: syslogd
0513-059 The syslogd Subsystem has been started. Subsystem PID is 4456860.
```

You can view the following entries in the `/etc/rsyslog.conf` file and you might also view additional entries based on your `/etc/rsyslog.conf` configuration file.

```
aso.notice /var/log/aso/aso.log
aso.info /var/log/aso/aso_process.log
aso.debug /var/log/aso/aso_debug.log
caa.debug;caa. /var/adm/ras/syslog.caa .info /var/adm/ras/syslog.txt
local0.info;user.notice;daemon.notice /var/hacmp/adm/cluster.log
```

Note:
Irrespective of type of the *syslogd* daemon type, the **lssrc** command always shows the state of the subsystem as *syslog*.

To determine which *syslogd* daemon is enabled, run the following commands:

- **ps -ef | grep syslog**
  An output similar to the following example is displayed:

  ```
  root 26869770 4128770 0 04:19:37 - 0:00 /usr/sbin/rsyslogd
  ```

- **odmget -q** “subsysname = 'syslogd'” **SRCsubsys**
  An output similar to the following example is displayed:

  ```
  SRCsubsys:
  subsysname = "syslogd"
  synonym = ""
  cmdargs = ""
  path = "/usr/sbin/rsyslogd" <<<< rsyslogd is enabled
  uid = 0
  auditid = 0
  standin = "/dev/console"
  standout = "/dev/console"
  standerr = "/dev/console"
  action = 1
  multi = 1
  contact = 3
  svrkey = 0
  svrmtype = 0
  priority = 20
  signorm = 0
  sigforce = 0
  display = 1
  waittime = 20
  grpname = "ras"
  ```

Note:

- The **/etc/rsyslog.conf** files must be identical on all cluster nodes.

- PowerHA® SystemMirror® does not support rotation of log files when the *rsyslogd* daemon is enabled because the *rsyslogd* daemon requires special mechanism to rotate the log files.

Changing script behavior for unrecoverable errors in PowerHA® SystemMirror®

If the System Resource Controller (SRC) detects that the **clstrmgr** daemon has exited abnormally, it runs the **/usr/es/sbin/cluster/utilities/clexit.rc** script to halt the system. If the SRC detects that any other PowerHA® SystemMirror® daemon has exited abnormally, it runs the **clexit.rc** script to stop these processes, but does not halt the system.

You can change the default behavior of the **clexit.rc** script by configuring the **/usr/es/sbin/cluster/etc/hacmp.term** file, which is called when the PowerHA® SystemMirror® cluster services end abnormally. Depending on how you customize the **hacmp.term** file, PowerHA® SystemMirror® operates specific to your installation.

PowerHA® SystemMirror® and AIX® commands

PowerHA® SystemMirror® provides a comprehensive set of functions for managing shared volume groups, logical volumes, and file systems.

These functions are available from the System Management (C-SPOC) menu in the System Management Interface Tool (SMIT). If you must configure scripts for these functions, you can use the **`clmgr`** command or a set of the **`cli`** commands in the `/usr/es/sbin/cluster/cspoc/` directory.

Use these PowerHA® SystemMirror® functions to manage shared storage instead of using basic AIX® commands. If you incorrectly use the AIX® commands, you can cause various problems such as data corruption on a shared storage disk.

Configuring a PowerHA® SystemMirror® cluster

These topics describe how to configure a PowerHA® SystemMirror® cluster using the SMIT **Cluster Nodes and Networks** path.

Overview of configuring a cluster

Using the options under the SMIT menu **Cluster Nodes and Networks › Initial Cluster Setup (Typical)** , you can configure the basic components of a cluster. This configuration path significantly automates the discovery and selection of configuration information and chooses default behaviors.

You can also use the General Configuration Smart Assist to quickly set up your application.

Prerequisite tasks for configuring a cluster

Before configuring the cluster, PowerHA® SystemMirror® must be installed on all the nodes, and connectivity must exist between the node where you are performing the configuration and all other nodes to be included in the cluster.

Network interfaces must be both physically and logically configured to the AIX® operating system so that communication occurs from one node to each of the other nodes. The PowerHA® SystemMirror® discovery process runs on all server nodes, not just the local node.

All node IP addresses and host names must be added to the `/etc/cluster/rhosts` file before configuring the cluster to verify that information is gathered from the systems that belong to the cluster.

After you have configured and powered on all disks and configured communication paths to other nodes in the AIX® operating system, PowerHA® SystemMirror® automatically collects information about the physical and logical configuration. It displays this information in corresponding SMIT picklists.

PowerHA® SystemMirror® uses all configured interfaces on the cluster nodes for cluster communication and monitoring. All configured interfaces are used to keep cluster IP addresses highly available.

After you install PowerHA® SystemMirror®, verify that the following Cluster Aware AIX® (CAA) entry is commented out in the `/etc/inetd.conf` file before creating PowerHA® SystemMirror® cluster.

```
caa_cfg stream tcp6 nowait root /usr/sbin/clusterconf
clusterconf >>/var/adm/ras/clusterconf.log 2>&1
```

Starting with AIX® 7.3, CAA entry in the `/etc/inetd.conf` file is already commented. During PowerHA® SystemMirror® installation, this CAA entry in the `/etc/inetd.conf` file is commented out for the following releases:

- PowerHA® SystemMirror® Version 7.2.3 SP6, or later
- PowerHA® SystemMirror® Version 7.2.4 SP4, or later
- PowerHA® SystemMirror® Version 7.2.5 SP2, or later
- PowerHA® SystemMirror® Version 7.2.6

If AIX® 7.3 is installed with an earlier version of PowerHA® SystemMirror® than the versions listed previously, the CAA entry is not commented out automatically. You must manually comment out the CAA entry in the `/etc/inetd.conf` file.

To uncomment the CAA entry, run the following command on all the nodes of the cluster before creating the cluster:

```
chsubserver -a -r inetd -v caa_cfg -p tcp6
```

Assumptions and defaults for configuring a cluster

PowerHA® SystemMirror® makes some assumptions regarding the environment, such as assuming all network interfaces on a physical network belong to the same PowerHA® SystemMirror® network. Using these assumptions, PowerHA® SystemMirror® supplies or automatically configures intelligent and default parameters to its configuration process in SMIT. This helps to minimize the number of steps it takes to configure the cluster.

PowerHA® SystemMirror® makes the following basic assumptions:

- The cluster configuration is stored in a central repository disk, and PowerHA® SystemMirror® assumes that all nodes in the cluster have common access to at least one physical volume or disk. This common disk cannot be used for any other purpose, such as hosting application data. You can specify this dedicated shared disk when you initially configure the cluster.
- PowerHA® SystemMirror® uses unicast communications by default for heartbeat and messaging between nodes in the cluster. You can choose to use multicast communication. If you use multicast communication, you must verify that the network devices are configured for multicast communication. When you create a cluster that uses multicast communication, PowerHA® SystemMirror® uses a default multicast IP address for your environment or you can specify a multicast IP address.
- If you use the **Cluster Nodes and Networks** System Management Interface Tool (SMIT) path to create a cluster, the host name is used as the PowerHA® SystemMirror® node name. After the cluster is created, you can change the node name. You can use the **Custom Cluster Configuration** SMIT path to specify the node name when you create the cluster.
- PowerHA® SystemMirror® uses IP aliasing for binding a service IP label/address to a network interface.

Steps for configuring a cluster

Here are the steps to configure the typical cluster components.

What You Do	Description
Step 1: Configure a basic cluster, or a cluster with an application	Use the SMIT menus under **Initial Cluster Setup (Typical)** to configure a basic cluster with default options and discovered network components. You can use one of the Smart Assists to configure a cluster with an application.
Step 2: Configure additional topology components	You can choose to configure the cluster with the SMIT menus under **Initial Cluster Setup (Custom)** if you want to create it piece by piece. This may be because you want to name the networks or nodes with something other than the default names, or because

What You Do	Description
	you want to choose specific network interfaces to support clustered applications (by default all interfaces will be used). Regardless of how you initially setup the cluster, you can add or remove components to the initial configuration using the **Manage** menus under the **Cluster Nodes and Networks** SMIT menu.
Step 3: Configure the cluster resources	Configure the resources to be made highly available. Use the menus under the SMIT path **Cluster Applications and Resources › Resources** to configure resources that are to be shared among the nodes in the cluster. You can configure these resources: • Application IP addresses and IP labels • Application controllers (start and stop scripts for the applications) • Volume groups • Logical volumes and file systems • Raw disks (for concurrently accessed data) • Tape resources • NFS exports and cross-mounts • Custom user-defined resources
Step 4: Configure the resource groups	Use the menus under the SMIT path **Cluster Applications and Resources › Resources Groups** to create the resource groups you have planned for each set of related resources.
Step 5: Put the resources to be managed together into their respective resource groups	To assign resources to each resource group use the following SMIT menu, **Cluster Applications and Resources › Resource Groups - Change/Show Resources and Attributes for a Resource Group**.
Step 6: Adjust log viewing and management	*(Optional)* Adjust log viewing and management using the SMIT dialogs under **Problem Determination Tools › PowerHA SystemMirror Logs**.
Step 7: Verify and synchronize the cluster configuration	Use the **Verify and Synchronize Cluster Configuration** dialog to verify that the desired configuration is valid and to ensure that all nodes in the cluster have the same view of the configuration.
Step 8: Display the cluster configuration	*(Optional)* Use the **PowerHA SystemMirror Configuration** dialog, found under SMIT menu **Cluster Nodes and Networks › Manage the Cluster** to view the cluster topology and resources configuration.
Step 9: Make further additions or adjustments to the cluster configuration	*(Optional)* You may want to perform some optional cluster configuration for your application environment needs. Such additions or adjustments include, for example: • Configuring a distribution preference for service IP aliases • Configuring resource group runtime policies, including Workload Manager • Adding resource group timers for startup and fallback • Configuring dependencies between resource groups • Adding application monitors • Configuring File Collections • Configuring cluster users or security

What You Do	Description
	• Customizing remote notifications (pager, SMS messages, and e-mail) • Customizing cluster events
Step 10: Test the cluster before it goes into the production environment	*(Recommended)* Use the **Cluster Test Tool**, found under SMIT menu **Problem Determination Tools**, to test recovery procedures for the cluster. **Note:** The Cluster Test Tool is not fully supported by the latest PowerHA releases.

Configuring a cluster using Smart Assists

You can configure a basic cluster with just a few configuration steps by using the **Initial Cluster Setup (Typical)** menus from SMIT.

If you are configuring a WebSphere®, DB2® UDB or Oracle application, see the corresponding PowerHA® SystemMirror® Smart Assists guide.

The initial requirements for using Smart Assists are:

- The application must be verified that it can run on all cluster nodes. You also need to verify that application response is common while running on all cluster nodes.
- The Smart Assists must be installed on all cluster nodes that run the application.

To configure your installed application (other than DB2®, WebSphere®, or Oracle), complete the following steps:

1. On a local node, enter `smitty sysmirror`.

   If you use a Smart Assist, the cluster topology components and the application will both be configured and no further steps are necessary.

2. Select **Cluster Applications and Resources › Configuration Assistants › Make Applications Highly Available (Use Smart Assists) › Add an Application to the PowerHA® SystemMirror® Configuration** and press Enter.
3. If the cluster is not yet configured, you are directed to a window that prompts you to **Enter Communication Path to Nodes**. You need to list the communication paths to all nodes in the cluster.
4. If the cluster is configured, SMIT displays a list of installed Smart Assists common in the cluster node. Select **Other Applications**, and press Enter.
5. Select **General Application Smart Assist**, and press Enter.
6. Enter values for the following fields in the **Add an Application to PowerHA® SystemMirror®** window:
   - **Application Controller Name**
   - **Primary Node**
   - **Takeover Nodes**
   - **Application Controller Start Script**
   - **Application Controller Stop Script**
   - **Service IP Label**
7. Press Enter after you completed all fields. The configuration is synchronized and verified automatically.
8. Return to the window **Make Applications Highly Available** and select **Test Your Application for Availability**. Press Enter.
   The Cluster Test Tool runs and displays the results. If you receive error messages, make the necessary corrections.

Defining PowerHA® SystemMirror® cluster topology

When you verify and synchronize the cluster topology, its definition is copied to the other nodes.

To configure the cluster topology:

1. From the command line, enter `smit sysmirror`
2. Depending on how you want to deploy your cluster, use one of the following paths in SMIT:
   - For a cluster that does not have sites, select **Cluster Nodes and Networks › Standard Cluster Deployment › Setup a Cluster, Nodes and Networks**, and press Enter.
   - For a cluster that has sites, select **Cluster Nodes and Networks › Multi Site Cluster Deployment › Setup a Cluster, Sites, Nodes and Networks**, and press Enter.
3. In SMIT, select **Cluster Nodes and Networks › Initial Cluster Configuration (Typical) › Setup a Cluster, Nodes and Networks** and press Enter.
4. Complete the fields and press Enter.
5. Press F3 to return to the previous SMIT panel and select **Define Repository Disk and Cluster IP Address**.
6. Complete the fields and press Enter.

> The default value for the **Heartbeat Mechanism** field is **unicast**. If you select **multicast** for the **Heartbeat Mechanism** field, you must verify that your network devices are configured for multicast communication.

Configuring PowerHA® SystemMirror® resources

Using the SMIT path **Cluster Applications and Resources › Resources**, you can configure resources that are required by your clustered applications.

You must first define resources that will be kept highly available by PowerHA® SystemMirror® for an application, then group them together in a resource group. You can add all the resources at once or separately.

This section explains how to configure the following types of resources in a cluster:

- Application controllers (the scripts used to start and stop the application).
- PowerHA® SystemMirror® service IP labels/addresses. The service IP label/address is the IP label/address over which services are provided and which is kept highly available by PowerHA® SystemMirror®.
- Volume groups, logical volumes, and file systems.

Configuring application controllers

A PowerHA® SystemMirror® application controller is a cluster resource used to control an application that must be highly available. It contains application start and stop scripts.

When you configure an application controller the following occurs:

- Associates a meaningful name with the application. For example, the application you are using with PowerHA® SystemMirror® is named *dbinst1*. You use this name to refer to the application controller when you define it as a resource. When you set up the resource group that contains this resource, you define an application controller as a resource.
- Points the cluster event scripts to the scripts that they call to start and stop the application.
- Allows you to configure application monitoring for that application. You can configure multiple application monitors for one application. For more information, see Steps for configuring multiple application monitors.

Review the vendor documentation for specific product information about starting and stopping a particular application.

Verify that the scripts exist on all nodes that participate as possible owners of the resource group where this application controller is defined.

To configure an application controller on any cluster node:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Application Controller Scripts › Add Application Controller Scripts**, and press Enter.
3. Enter the field values as follows:

Add Application Controller Scripts fields	
Field name	**Value**
Application controller name	Enter an ASCII text string that identifies the application controller. You use this name to refer to the application controller when you add it to the resource group. The controller name can include alphanumeric characters and underscores. Use a maximum of 64 characters.
Start Script	Enter the full path name of the script followed by arguments that are called by the cluster event scripts to start the application. This field can have a maximum of 256 characters. Although, this script must have the same name and location on every node, the content and function of the script can be different. You can use the same script and runtime conditions to modify a node's runtime behavior.
Stop Script	Enter the full path name of the script that is called by the cluster event scripts to stop the application. This field can have a maximum of 256 characters. This script must be in the same location on each cluster node that can start the application. Although, this script must have the same name and location on every node, the content and function of the script can be different. You can use the same script and runtime conditions to modify a node's runtime behavior.
Resource Group Name	Specify the resource group to contain this resource. Use F4 to view a picklist. If you have not configured your resource groups yet, you can leave this field blank and add the resource to the group later.
Startup Mode	Specify how the application controller startup script is called. Select **background**, the default value, if you want the start script called as a background process and event processing continues even if the start script has not completed. Select **foreground** if you want the event to suspend processing until the start script exits. This field is only available in PowerHA® SystemMirror® 7.1.1, or later.
Enable CPU usage statistics	Specify **No** to disable CPU monitoring. Specify **Yes** to enable CPU monitoring. The default value is **No**.
Process to monitor CPU usage	Specify the full path name of the application binary that you want to monitor. You can use this field to determine the associated process-id (PID) to monitor.
CPU usage monitor interval	Specify the time interval in minutes to wait after each CPU and Memory usage reading. Valid values are between 1 and 120 minutes inclusively.

4. Press Enter to add the application controller as a cluster resource. Remember that you must verify and synchronize this change to add it to the cluster definition on all nodes in the cluster.

Configuring PowerHA® SystemMirror® service IP labels and IP addresses

Service IP labels and IP addresses are used to establish communication between client nodes and the server node. Services, such as a database application, are provided using the connection made over the service IP label.

The **/etc/hosts** file on all nodes must contain all IP labels and associated IP addresses that you will define to the cluster, including service IP labels and addresses.

To define service IP labels for your cluster, complete the following steps:

1. From the command line, enter `smit sysmirror`

2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Service IP Labels/ Addresses › Add Service IP Label/Address** and press Enter.
3. Fill in field values as follows:

Add Service IP Label/Address fields	
Field	**Value**
IP Label/IP Address	Select from the picklist or enter the service IP label/address to be kept highly available. The name of the service IP label/address must be unique within the cluster and distinct from the volume group and resource group names; it should relate to the application it serves, as well as to any corresponding device, such as `websphere_service_address`.
Network Name	Enter the symbolic name of the PowerHA® SystemMirror® network on which this Service IP label/address will be configured. If you leave the field blank, PowerHA® SystemMirror® fills in this field automatically with the network type plus a number appended, starting with 1, for example, `netether1`.
Netmask(IPv4)/ Prefix Length(IPv6)	For the configuration of the IP version 4 service interface, enter the network mask for the address. For the configuration of the IP version 6 service interface, enter the prefix length for the address. This field is not a required field. If you do not enter a value, the prefix length or netmask of the underlying network is used. If a prefix length value or netmask value is specified, it is checked for compatibility with the underlying network.

4. Press Enter after filling in all required fields. PowerHA® SystemMirror® checks the validity of the IP interface configuration.
5. Repeat the previous steps until you have configured all service IP labels for each network, as needed.

Configuring volume groups, logical volumes, and files systems

You can configure volume groups, logical volumes, files systems, and user defined resources.

Configuring volume groups, logical volumes, and file systems as cluster shared resources

You must define and properly configure volume groups, logical volumes and file systems to the AIX® operating system, before using them as shared resources in a PowerHA® SystemMirror® cluster.

Configuring concurrent volume groups, logical volumes, and file systems

These components must be defined to the AIX® operating system and properly configured, for use as shared resources.

Configuring a user-defined resource type

PowerHA® SystemMirror® allows users to add their own resource type and to specify management scripts to configure how and where PowerHA® SystemMirror® processes the resource type. You can then configure a user-defined resource instance for use in a resource group.

A user-defined resource type is one that you can define a customized resource that you can add to a resource group. A user-defined resource type contains several attributes that describe the properties of the instances of the resource type.

Ensure that the user-defined resource type management scripts exist on all nodes that participate as possible owners of the resource group where the user-defined resource resides.

To configure a user-defined resource group type, complete these steps:

1. From the command line, enter `smit sysmirror`

2. In SMIT, select **Custom Cluster Configuration › Resources › Configure User Defined Resources and Types › Add a User Defined Resource Type**, and press Enter.
3. Enter the field values as follows:

Add a User Defined Resource Type fields	
Field	**Value**
Resource Type Name	Enter an ASCII text string that identifies the resource type. You use this name to refer to the resource type when you define the resource configuration. The resource type name can include alphanumeric characters and underscores. Use a maximum of 64 characters.
Process At/After	Specify the Processing order that you want to use to process the user-defined resources. Use F4 to view a pick list of all existing resource types and chose one from the list. PowerHA® SystemMirror® processes the user-defined resources at the very beginning of the resource acquisition order if you choose FIRST. If you chose any other value, for example, VOLUME_GROUP, the user-defined resources will be acquired after varying on the volume groups and they will be released after varying off the volume groups.
Verification Method	Specify a verification method to be invoked by the cluster verification process. You will need to provide the verification checks so that before starting the cluster services, user-defined resources will be verified to avoid failures during cluster operation.
Verification Type	Specify the type of verification method to be used. The verification method can be either a script or a library. If you chose Library, it should be written as per the guidelines mentioned in *Writing custom verification libraries*.
Start Method	Enter the name of the script and its full path name (followed by arguments) to be called by the cluster event scripts to start the user-defined resource. Use a maximum of 256 characters. 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 Method	Enter the full path name of the script to be called by the cluster event scripts to stop the user-defined resource. Use a maximum of 256 characters. This script must be in the same location on each cluster node that might stop the resource. The contents of the script, however, may differ.
Monitor Method	Enter the full path name of the script to be called by the cluster event scripts to monitor the user-defined resource. Use a maximum of 256 characters. This script must be in the same location on each cluster node that might monitor the monitor. The contents of the script, however, may differ.
Cleanup Method	(Optional) Specify a resource cleanup script to be called when a failed user-defined resource is detected, before calling the restart method. The default for the cleanup script is the stop script defined when the user-defined resource type was set up. If you are changing the monitor mode to be used only in the startup monitoring mode, the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field. With resource monitoring, since the resource is already stopped when this script is called, the resource stop script might fail.
Restart Method	The default restart method is the resource start script defined previously. You can specify a different method here, if desired. If you are changing the monitor mode to be used only in the startup monitoring mode, the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field.
Failure Notification Method	Define a notify method to run when the user-defined resource fails. This custom method runs during the restart process and during notify activity. If you are changing the monitor mode to be used only in the startup monitoring mode,

Field	Value
	the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field.
Required Attributes	Specify a list of attribute names, with each name separated by a comma. These attributes must be assigned with values when creating the user-defined resource, for example, `Rattr1,Rattr2`. The purpose of the attributes is to store resource-specific attributes which can be used in the different methods specified in the resource type configuration.
Optional Attributes	Specify a list of attribute names, with each name separated by a comma. These attributes might or might not be assigned with values when creating the user-defined resource, for example, `Oattr1, Oattr2`. The purpose of the attributes is to store resource-specific attributes which can be used in the different methods specified in the resource type configuration.
Description	Provide a description of the user-defined resource type.

4. Press Enter to add this information to the PowerHA® SystemMirror® Configuration Database on the local node. Return to previous PowerHA® SystemMirror® SMIT panels to perform other configuration tasks.

Configuring user-defined resources

PowerHA® SystemMirror® allows you to configure user-defined resource instances for an already configured user-defined resource type.

User-defined resource configuration does the following:

- Creates an instance of the chosen resource type with default attribute values.
- All resource type management scripts/methods can access the attributes associated with the user-defined resource by using the `cludres -q` command.
- A method is called in a specific format, for example, if the start method is `/opt/udrmethods/start_resource.sh`, then the format to call the method is `/opt/udrmethods/start_resource.sh <resourcename>`.
- If a monitor method is specified in the user-defined resource type configuration, a custom resource monitor is added for the current resource with a name in the format of `cludrm_<RESOURCENAME>`. The custom monitor is added with the following default values for monitor attributes, which you can change by using the **Change/Show User Defined Resource Monitor** option in SMIT:
  - INVOCATION = longrunning
  - MONITOR_INTERVAL = 60
  - HUNG_MONITOR_SIGNAL = 9
  - STABILIZATION_INTERVAL = 15
  - RESTART_COUNT = 3
  - FAILURE_ACTION = fallover
  - RESTART_INTERVAL = 15
  - MONITOR_METHOD = monitor method defined in resource type configuration
  - CLEANUP_METHOD = cleanup method defined in resource type configuration, otherwise, this is the stop script
  - FAILURE_NOTIFY_METHOD = failure notification method defined in resource type configuration
  - RESTART_METHOD = restart method defined in resource type configuration otherwise start method
- Enables the ability to add the user defined resource type to the resource group in the **User defined resources** field of the resource group.

To configure a user-defined resource, complete these steps:

1. From the command line, enter `smit sysmirror`

2. In SMIT, select **Custom Cluster Configuration › Resources › Configure User Defined Resources and Types › Add a User Defined Resource**, and press Enter.
3. SMIT displays the **Select User Defined Resource type** panel, which lists all user-defined resource types.

   If you have not previously configured a resource type, none are listed and you cannot complete this task.

4. SMIT displays the **Add a User Defined Resource** panel. Enter the field values as follows:

Add a User Defined Resource fields	
Field	**Value**
Resource Type Name	Displays the name of the selected user-defined resource type.
Resource Name	Enter an ASCII text string that identifies the resource. You use this name to refer to the resource when you define resources during node configuration. The resource name can include alphanumeric characters and underscores. You can enter maximum of 64 characters. The resource name must be unique across the cluster. When you define a volume group as a user-defined resource for a Peer-to-Peer Remote Copy (PPRC) configuration or a HyperSwap® configuration, the resource name must match the volume group.
Attribute Data	Specify a list of attributes and values in the form of attribute=value, with each pair separated by a space; for example, `Rattr1="value1" Rattr2="value2" Oattr1="value3"`.

5. Press Enter to add this information to the PowerHA® SystemMirror® Configuration Database on the local node. Return to previous PowerHA® SystemMirror® SMIT panels to perform other configuration tasks.

   You can also import user-defined resource configuration from an xml file by using the **Import User Defined Resource Types and Resources Definition from XML File** option in SMIT. Before using this option, you must first create an xml file with all necessary information. You can do this by using a the `/usr/es/sbin/cluster/etc/udrt_sample.xml` as a template.

Configuring PowerHA® SystemMirror® resource groups

You can configure resource groups that use different startup, fallover, and fallback policies.

Configuring a resource group involves two phases:

- Configuring the resource group name, startup, fallover and fallback policies, and the nodes that can own it (nodelist for a resource group).
- Adding the resources and additional attributes to the resource group.

To create a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Add a Resource Group**.
3. Enter information in the following fields:

Add a Resource Group fields	
Field	**Value**
Resource Group Name	Enter the name for this group. The name of the resource group must be unique within the cluster and distinct from the service IP label and volume group name. It is helpful to create the name related to the application it serves, as well as to any corresponding device, such as `websphere_service_address`. Use no more than 64 alphanumeric characters or underscores; do not use a leading numeric. Do not use reserved words. See List of reserved words. Duplicate entries are not allowed.
Participating Node Names	Enter the names of the nodes that can own or take over this resource group. Enter the node with the highest priority for ownership first, followed by the nodes with the less priorities, in the desired order. Leave a space between node names, for example, `NodeA NodeB NodeX`.
Startup Policy	Select a value from the picklist that defines the startup policy of the resource group: **ONLINE ON HOME NODE ONLY**. The resource group should be brought online *only* on its home (highest priority) node during the resource group startup. This requires the highest priority node to be available. **ONLINE ON FIRST AVAILABLE NODE**. The resource group activates on the first node that becomes available. **ONLINE USING NODE DISTRIBUTION POLICY**. If you select the node distribution policy, only one resource group is brought online on a node during startup. **ONLINE ON ALL AVAILABLE NODES**. The resource group is brought online on *all* nodes. This is equivalent to concurrent resource group behavior. If you select this option for the resource group, ensure that resources in this group can be brought online on multiple nodes simultaneously.
Fallover policy	Select a value from the list that defines the fallover policy of the resource group: **FALLOVER TO NEXT PRIORITY NODE IN THE LIST**. In the case of fallover, the resource group that is online on only one node at a time follows the default node priority order specified in the resource group's nodelist (it moves to the highest priority node currently available). **FALLOVER USING DYNAMIC NODE PRIORITY**. If you select this option for the resource group (and Online on Home Node startup policy), you can choose either one of the three predefined dynamic node priority policies or one of the two user defined policies. See Configuring dependencies between resource groups.
Fallback policy	Select a value from the list that defines the fallback policy of the resource group: **NEVER FALLBACK**. A resource group does not fall back when a higher priority node joins the cluster. **FALLBACK TO HIGHER PRIORITY NODE IN THE LIST**. A resource group falls back when a higher priority node joins the cluster.

4. Press Enter.
5. Return to the **Add a Resource Group** panel to continue adding all the resource groups you have planned for the PowerHA® SystemMirror® cluster.

Configuring resources in resource groups

After you have defined a resource group, you will add resources to it. SMIT can list possible shared resources for the node if the node is powered on (helping you avoid configuration errors).

When you are adding or changing resources in a resource group, PowerHA® SystemMirror® displays only valid choices for resources, based on the resource group management policies that you have selected.

Keep the following in mind as you prepare to define the resources in your resource group:

- You cannot configure a resource group until you have completed the information on the **Add a Resource Group** panel. If you need to do this, refer back to the instructions under Configuring PowerHA® SystemMirror® resource groups .
- A resource group may include multiple service IP addresses. When a resource group is moved, all service labels in the resource group are moved as aliases to the available interfaces, according to the resource group management policies in PowerHA® SystemMirror®.
  Also, you can specify the distribution preference for service IP labels. For more information, see Steps to configure distribution preference for service IP label aliases.

  For information on how PowerHA® SystemMirror® handles the resource groups, see Resource group behavior during cluster events.
- When you define a service IP label/address on a cluster node, the service label can be used in any non-concurrent resource group.
- IPAT function does not apply to concurrent resource groups (those with the startup policy Online on All Available Nodes).

To assign the resources for a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group** and press Enter to display a list of defined resource groups.
3. Select the resource group you want to configure and press Enter. SMIT displays the panel that matches the type of resource group you selected, with the Resource Group Name, and Participating Node Names (Default Node Priority) fields filled in.
   Note: SMIT displays only valid choices for resources, depending on the type of resource group that you selected.

   If the participating nodes are powered on, you can press F4 to list the shared resources in the picklists. If a resource group/node relationship has not been defined, or if a node is not powered on, pressing F4 causes PowerHA® SystemMirror® SMIT to display the appropriate warnings.
4. Enter the field values as follows:

Resource group fields	
Field	**Value**
Service IP Label/IP Addresses	This option appears only if you are adding resources to a non-concurrent resource group. List the service IP labels to be taken over when this resource group is taken over. See a picklist of valid IP labels. These include addresses that rotate or may be taken over.
File systems (empty is All for specified VGs)	This option appears only if you are adding resources to a non-concurrent resource group. If you leave the File systems (**empty is All for specified VGs**) field blank *and* specify the shared volume groups in the Volume Groups field here, all file systems will be mounted in the volume group. If you leave the File systems field blank and do not specify the volume groups in the field here, no file systems will be mounted. You may also select individual file systems to include in the resource group. Press F4 to see a list of the file systems. In this case only the specified file systems will be mounted when the resource group is brought online.
Volume Groups	This option appears only if you are adding resources to a non-concurrent resource group. Identify the shared volume groups that should be varied on when this resource group is acquired or taken over. Select the volume

Field	Value
	groups from the picklist or enter desired volume groups names in this field.
Volume Groups (continued)	Press F4 for a list of all shared volume groups in the resource group *and* the volume groups that are currently available for import onto the resource group nodes. Specify the shared volume groups in this field if you want to leave the field Filesystems **(empty is All for specified VGs)** blank *and* to mount all file systems in the volume group. If you specify more than one volume group in this field, all file systems in all specified volume groups are mounted. You cannot choose to mount all file systems in one volume group and not to mount them in another. For example, in a resource group with two volume groups, vg1 and vg2, if the **Filesystems (empty is All for specified VGs)** is kept blank, all the file systems in vg1 and vg2 are mounted when the resource group is brought up. However, if the **Filesystems (empty is All for specified VGs)** has only file systems that are part of the vg1 volume group, none of the file systems in vg2 are mounted, because they were not entered in the **Filesystems (empty is All for specified VGs)** field along with the file systems from vg1. If you have previously entered values in the Filesystems field, the appropriate volume groups are already known to PowerHA® SystemMirror®.
Concurrent Volume Groups	This option appears only if you are adding resources to a concurrent resource group. Identify the shared volume groups that can be accessed simultaneously by multiple nodes. Select the volume groups from the picklist, or enter desired volume groups names in this field. If you previously requested that PowerHA® SystemMirror® collect information about the appropriate volume groups, then the picklist displays a list of all existing concurrent capable volume groups that are currently available in the resource group, *and* concurrent capable volume groups available to be imported onto the nodes in the resource group. Disk fencing is turned **on** by default.
Application Controllers	Specify the application controllers to include in the resource group. The picklist displays a list of application controllers.
User defined resource	Specify the user defined resource to include in the resource group. The picklist displays a list of user defined resources configured.

If you are configuring a resource group with the startup policy of Online on Home Node and the fallover policy Fallover Using Dynamic Node Priority, this SMIT panel displays the field where you can select one of the three predefined dynamic node priority policies or one of the two user defined policies you want to use.

5. Press Enter to add the values to the PowerHA® SystemMirror® Configuration Database.

Verifying and synchronizing the standard configuration

After all resource groups have been configured, verify the cluster configuration on all nodes to ensure compatibility. If no errors are found, the configuration is then copied (synchronized) to each node of the cluster. If you synchronize from a node where Cluster Services are running, one or more resources may change state when the configuration changes take effect.

At the beginning of verification, before PowerHA® SystemMirror® verifies the cluster topology, the Cluster Topology Summary is displayed listing any nodes, networks, network interfaces, and resource groups that are "unavailable" at the time that cluster verification is run. "Unavailable" refers to those that have failed and are considered offline by the Cluster Manager. These components are also listed in the **/var/hacmp/clverify/clverify.log** file.

PowerHA® SystemMirror® displays informational messages during the verification process. The first phase of verification involves collecting data from all the nodes in the cluster. Messages are displayed as collection completes on each node, and if a node is slow to respond, the amount of time passed since collection started is displayed.

The second phase of the process is the verification of the collected data. PowerHA® SystemMirror® display the progress of the verification checks in increments of 10 percent.

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.

The output may take one of the following forms:

- You may see warnings if the configuration has a limitation on its availability, for example, if only one interface per node per network is configured.
- Although no summary will be displayed to the user when no cluster topology components have failed, the **clverify.log** file displays the following:

  ```
  <DATE/TIME>  Verification detected that all cluster topology components are available.
  ```

- If cluster components are unavailable, the utility providing the list of failed components puts similar information in the log file.

To verify and synchronize the cluster topology and resources configuration:

1. From the command line, enter `smit sysmirror`
2. You can access the **Verify and Synchronize Cluster Configuration** dialog from many menu paths. This dialog is accessible from main menus that contain dialogs used to change the cluster configuration, such as under the **Cluster Nodes and Networks** menu or the **Cluster Applications and Resources** menu. The **Verify and Synchronize Cluster Configuration (Advanced)** dialog is located under the **Custom Cluster Configuration** menus.
3. To use default options for verification and synchronization, press Enter.
   SMIT runs the **verification** utility.

Viewing the PowerHA® SystemMirror® configuration

Once you have configured, verified, and synchronized the PowerHA® SystemMirror® configuration, you can display the PowerHA® SystemMirror® cluster.

To display the PowerHA® SystemMirror® cluster:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster and Nodes and Networks › Manage the Cluster › PowerHA® SystemMirror® Configuration** and press Enter.
   SMIT displays the current topology and resource information.

Result
After you have finished configuring and synchronizing the cluster configuration, consider further customizing the cluster. For example, you can:

- Refine the distribution of service IP aliases placement on the nodes. For more information, see Steps to configure distribution preference for service IP label aliases.
- Configure dependencies between resource groups. Consider this step if you are planning to include multi-tiered applications in the cluster, where the startup of one application depends on the successful startup of another application.
- Refine resource groups behavior by specifying the delayed fallback timer, the settling time, and the node distribution policy.
- Configure multiple monitors for an application controller, to monitor the health of your applications.

- Change runtime parameters and redirect log files for a node.
- Customize cluster events.
- Customize and configure different types of remote notification, such as pager, SMS messages, and email.
- Configure PowerHA® SystemMirror® File Collections.
- Enable the cluster verification to run corrective actions.

Configuring split and merge policies

You can use the SMIT interface to configure split and merge policies.

The possible split policy and merge policy configuration depends on the version of the AIX® operating system running in your environment.

The following table displays the possible split and merge policy combinations for different cluster types in AIX® Version 7.2.0, or earlier:
Displays the cluster types and possible split and merge policy options.

AIX® Version 7.2.0, or earlier, split and merge policy combinations		
Cluster type	**Split policy option**	**Merge policy option**
Standard cluster	Not supported	Majority
Stretched cluster	None	Majority
	Tie breaker	Tie breaker
Linked cluster	None	Majority
	Tie breaker	Tie breaker
	Manual	Manual

In AIX® Version 7.1.5, or later, and AIX® Version 7.2.1, or later, all standard, stretched, and linked clusters support the same split policy and merge policy combinations.

The following table displays the possible split and merge policy combinations in AIX® Version 7.2.1, or later:
Split and merge policy combinations

AIX® Version 7.2.1, or later, split and merge policy combinations	
Split policy options	**Merge policy options**
None	Majority
None	None
Tie breaker	Tie breaker
Manual	Manual
Cloud	Cloud

To configure a split and merge policy in PowerHA® SystemMirror®, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. In the SMIT interface, select **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Setup (Custom) › Configure Cluster Split and Merge Policy › Split and Merge Management Policy**, and press Enter.
3. Complete the following fields, and press Enter.

   > Some of the following fields are displayed only when certain options are specified in the **Split Handling Policy** and **Merge Handling Policy** fields.

   Split and Merge Management Policy fields

Split and Merge Management Policy fields	
Field	**Description**
Split Handling Policy	You can select the following options for this field: **None** This option is the default setting. If you select this option, the partitions operate independently of each other after the cluster split occurs. If you select this option, the **Merge Handling Policy** field is set to **Majority**. **Tie Breaker** You can select the following options for the tie breaker: **Disk** Select this option to specify a disk as the tie breaker. When the cluster split occurs, one site wins the SCSI reservation on the tie breaker disk. The site that losses the SCSI reservation uses the recovery action that is specified in the policy setting. If you select this option, the **Merge Handling Policy** field is set to **Tie breaker**. **NFS** Select this option to specify an NFS file as the tie breaker. During the cluster split, a predefined NFS file is used to decide the winning partition. The partition that losses the NFS file reservation uses the recovery action that is specified in the policy setting. If you select this option, the **Merge Handling Policy** field is set to **NFS**. **Cloud** Select this option when the nodes are configured for cloud communication. When a cluster split event occurs, each partition attempts to acquire the lock by uploading a file to the configured Cloud service. The partition that successfully uploads the file to the configured Cloud service continues to function. The partition that cannot upload the file to the configured Cloud service is rebooted or the cluster services are restarted if it is the specified action in the policy setting. If you select this option, the **Merge handling policy** field is set to **Cloud**. Before you configure the split and merge policies as **Cloud**, you must enable Cloud communication on all nodes of the cluster and you must have access to create a Cloud Storage bucket for the Cloud service. During the verification and synchronization process, PowerHA® SystemMirror® verifies the communication to the Cloud and displays a warning for any failure in the cloud communication when the **Split handling policy** or **Merge handling policy** field is configured as **Cloud**. Also, you must install Python and Boto3 Software Development Kit (SDK) on all nodes of the cluster. For more information about the Boto3 Software Development Kit (SDK), see the Boto3 SDK Documentation website. **Manual** Select this option to wait for manual intervention when a cluster split occurs. PowerHA® SystemMirror® does not perform any actions on the cluster until you specify how to recover from the cluster split. If you select this option, the **Merge handling policy** field is set to **Manual**.
Merge Handling Policy	Do not change the options that are specified in this field. The options that are specified in this field are determined by the options you selected in the **Split Handling Policy** panel.
Split and Merge Action Plan	You can select the following options for this field: **Reboot**

Field	Description
	Select this option to reboot nodes on the losing partition when a cluster split event occurs. **Disable Applications Auto-Start and Reboot** Select this option to reboot nodes on the losing partition when a cluster split event occurs. If you select this option, the resource groups are not brought online automatically after the system reboots. This option is available only if your environment is running AIX® Version 7.2.1, or later. **Disable Cluster Services Auto-Start and Reboot** Select this option to reboot nodes on the losing partition when a cluster split event occurs. If you select this option, Cluster Aware AIX® (CAA) is not started. The resource groups are not brought online automatically. After the cluster split event is resolved, in SMIT, you must select **Problem Determination Tools › Start CAA on Merged Node** to restore the cluster. If you specified **None** in the **Merge Handling Policy** and **Split Handling Policy** fields, the action plan is not implemented and a reboot does not occur after the cluster split event. This option is available only if your environment is running AIX® Version 7.2.1, or later.
Select tie breaker	This field is available if you select **Tie Breaker - Disk** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Select an iSCSI disk or a SCSI disk that you want to use as the tie breaker disk.
Select Site Priority	This field is available if you select **Tie Breaker - Disk** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Select any cluster site in PowerHA® SystemMirror® that enables the tie breaker win during cluster split operation. Site Priority configuration is functional only with Reliable Scalable Cluster Technology (RSCT) Version 3.3.2.0, or later.
Time Delay for Lower Priority Site	This field is available if you select **Tie Breaker - Disk** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. To reserve the tie breaker, specify the time in seconds to delay the lower priority site. You can enter a value in the range 0 – 300 seconds. The default value is 30 seconds only when **Site Priority** option is selected.
NFS Export Server	This field is available if you specify **Tie Breaker - NFS** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Specify the fully qualified domain name of the NFS server that is used for the NFS tie-breaker. The NFS server must be accessible from each node in the cluster by using the NFS server IP address. To export NFS from a Linux® operating system, you must disable the dir_index option by running the **`tune2fs -O ^dir_index <filesystem>`** command. Where, **`<filesystem>`** is the NFS directory.

Field	Description
Local Mount Directory	This field is available if you specify **Tie Breaker - NFS** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Specify the absolute path of the NFS mount point that is used for the NFS tie-breaker. The NFS mount point must be mounted on all nodes in the cluster. The default NFS mount options are `vers=4,fg,soft,retry=1,timeo=10`. Modifying the default values might lead to failure in acquiring the NFS lock.
NFS Export Directory	This field is available if you specify **Tie Breaker - NFS** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Specify the absolute path of the NFSv4 exported directory that is used for the NFS tie-breaker. The NFS exported directory must be accessible from all nodes in the cluster that use NFSv4. You must verify that the following services are active in the NFS server: • biod • nfsd • nfsgryd • portmap • rpc.lockd • rpc.mountd • rpc.statd • TCP You must verify that the following services are active in the NFS client on all cluster nodes: • biod • nfsd • rpc.mountd • rpc.statd • TCP • portmap
Notify Method	This field is available if you specify **Manual** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Specify the method that informs you to select which partition remains online after a cluster split or merge event. The method is specified as a path name that is followed by optional parameters. When this method is called, the last parameter is either **`split`** or **`merge`** to indicate the type of cluster event.
Notify Interval	This field is available if you specify **Manual** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Enter the seconds between notification messages that are sent asking you to select which partition remains online after a cluster split or merge event. You can enter a value in the range 10 - 3600.
Maximum Notifications	This field is available if you specify **Manual** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Enter the number of times you want a notification message to be sent to select which partition remains online after a cluster split or merge event. You can enter a

Field	Description
	value in the range 3 - 1000. You must enter a value for this field when the surviving (winning) site is specified.
Default Surviving Site	This field is available for only linked clusters and if you specified **Manual** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. Enter the site that remains operational if you do not manually specify the surviving site. The other site that you do not specify in this field uses the action that is specified in the **Split and Merge Action Plan** field.
Apply to Storage Replication Recovery	This field is available for only linked clusters and if you have specified **Manual** in both the **Split Handling Policy** and the **Merge Handling Policy** fields. This field determines if the manual response when a cluster split event occurs also applies to storage replication recovery mechanisms. If you select **Yes**, the partition that was selected to remain online after a cluster split event proceeds with the takeover of the storage replication recovery. You cannot specify this option if you are using IBM® XIV® Storage Systems or IBM® System Storage® DS8000®, or later.

4. Verify that all fields are correct and press Enter.
5. Verify and synchronize the changes across the cluster.
STEP RESULT:

Other:

- Split and merge policy are effective only in an active cluster.
- The Deadman mode sets to an Assert mode when the cluster services start.

Configuring a quarantine policy

You can configure a quarantine policy with PowerHA® SystemMirror® to isolate the previously active node that was hosting a critical resource group after a cluster split event or node failure occurs. The quarantine policy ensures that your application data is not corrupted or lost.

To configure a quarantine policy, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In the SMIT interface, select **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Setup (Custom) › Configure Cluster Split and Merge Policy › Quarantine Policy**, and press Enter.
3. Select one of the following options:
Active Node Halt Policy

   Select this option to stop the node that is unresponsive before PowerHA® SystemMirror® acquires the critical resource group. To use this option, you must identify the critical resource group. The critical resource group is the first resource group that is processed when a cluster split occurs, and the **Active Node Halt Policy** is applied to your cluster.

   Disk Fencing

   The disks that are related to the critical resource group are fenced off from the previously active node. The disks in the volume groups that are configured for all resource groups on that node are fenced off. This option ensures that the node that contains the critical resource group cannot access the disks in case a cluster split event occurs. To use this option, you must identify the critical resource group and the storage subsystem must support SCSI-3 persistent reservation and ODM reserve_policy of PR_SHARED. This policy is applied to all disks that are part of the volume group and resource group.
4. Verify and synchronize the changes across the cluster.

Additional cluster configuration

You can configure additional cluster components after initial cluster configuration.

Understanding custom cluster configuration options

Custom cluster configuration may be desired in some environments. These are less common configuration tasks that are not required in most cases.

Most options for configuration cluster components are found under the **Cluster Nodes and Networks** or **Cluster Applications and Resources** menus. Some options which are not required in most typical configurations can be found under the **Custom Cluster Configuration** menu. This includes dialogs to configure custom disk, volume group and file system methods for cluster resources, options to customize resource recovery and service IP label distribution policy, and options for event customization. You can also use the **Initial Cluster Setup (Custom)** menus in this path to create the initial cluster one piece at a time, giving you complete control over which components are added to the cluster and how they are named.

Discovering PowerHA® SystemMirror®-related information

You can use the SMIT interface to discovery network and storage devices.

After you have configured and powered on all disks, created shared volume groups, and configured communication paths to other nodes, PowerHA® SystemMirror® can automatically collect this information and display it in corresponding SMIT picklists, to help you make accurate selections of existing components. When you add new disks, network interfaces, or volume groups to the cluster, you can run the discovery task again.

> **Note:** The discovery process runs on *all* nodes, not just on the local node.

To run the PowerHA® SystemMirror® cluster discovery process, take the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Discover Network Interfaces and Disks** and press Enter.
3. The software executes the discovery process.

Cluster, nodes, and networks

Review the custom topology configuration options that you might want to use for specific cases.

Options for configuring cluster topology components includes:

Configuring a PowerHA® SystemMirror® cluster

You can use the **Custom Cluster Configuration** path and the menus under **Initial Cluster Setup (Custom)** to configure the cluster a piece at a time. You can start by giving the cluster a name.

To assign a cluster name, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Setup (Custom) › Cluster › Add/Change/Show a Cluster** and press Enter.
3. Enter field values as follows:

Add/Change/Show a Cluster field	
Field	**Value**
Cluster Name	Enter an ASCII text string that identifies the cluster. The cluster name can include alphanumeric characters and underscores, but cannot have a leading

Field	Value
	numeric. Use no more than 64 characters. Do not use reserved names. For a list of reserved names see List of reserved words.

4. Press Enter.
5. Return to the **Return to Initial Cluster Setup (Custom)** SMIT panel.

Configuring cluster heartbeat settings

The heartbeat function is configured to use specific paths between nodes. These paths allow heartbeats to monitor the health of all PowerHA® SystemMirror® networks, network interfaces, and nodes in the cluster.

To configure heartbeat settings for a cluster, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Custom Cluster Configuration › Cluster Nodes and Networks › Manage the Cluster › Cluster heartbeat settings** and press Enter.
3. Enter the settings for the following fields:

   Node Failure Detection Timeout

   The time in seconds for the health management layer to wait before sending a message that the node has failed. The valid values are in the range 10 - 600.

   Node Failure Detection Grace Time

   The time in seconds for the node to wait before declaring that a node has actually failed. The valid values are in the range 5- 600. This function starts after being contacted by the health management layer, which is specified in the **Node Failure Detection Timeout** field.

   Node Failure Detection Timeout during LPM

   The time out value, in seconds, that is used during Live Partition Mobility (LPM) instead of the **Node Failure Detection Timeout** value. You can set this option to be greater than the LPM freeze duration to avoid any unwanted cluster events during the LPM process. The valid values are in the range 10 - 600. If a value is not specified for this field, the value of the **Node Failure Detection Timeout** is used.

   LPM Node Policy

   Select **unmanage** with Unmanage Resource Group option in SMIT to stop cluster services during the LPM process. Select **manage** for PowerHA® SystemMirror® to continue to monitor the resource groups and application availability during the LPM process. The default value is **manage**.

   Link Failure Detection Timeout

   The time in seconds for the health management layer to wait before sending a message that the links between the sites have failed. A link failure can cause your cluster to fallover to another link and continue to function. If all the links in the cluster are not responding, a message is sent identifying that site is in an offline state.

   PowerVM® Watchdog Timer

   Cluster Aware AIX® (CAA) supports a new tunable **PowerVM Watchdog Timer** in IBM® AIX® 7.3 with the Technology Level 1, or later. The **PowerVM Watchdog Timer** tunable is supported in PowerHA® SystemMirror® 7.2.7, or later. When you enable this tunable, CAA resets the Power® hypervisor (PHYP) timer of every cluster node. If CAA fails to reset the PHYP timer, for example, if the logical partition (LPAR) fails, PHYP manages LPAR based on the option set by the user. The valid values for the **PowerVM Watchdog Timer** are **DISABLE**, **DUMP_RESTART**, **HARD_RESET**, and **HARD_POWER_OFF**. The default value is **DISABLE**.
4. Verify and synchronize the cluster.

AIX® Live Update for PowerHA® SystemMirror® nodes

You can use the AIX® Live Update function to apply an interim fix for the AIX® operating system without restarting the system. The workloads on a system are not stopped during the update process of the system that uses the Live Update function.

To use the Live Update function, you must have the following software installed:

- PowerHA® SystemMirror® Version 7.2.0, or later
- AIX® Version 7.2.0, or later

PowerHA® SystemMirror® can support the Live Update function on any nodes in the cluster. However, you can use the Live Update function on only one node at a time. PowerHA® SystemMirror® uses the Live Update framework to verify that disruptions do not occur to any resource groups during the Live Update process. When you are using the Live Update process and commands, PowerHA® SystemMirror® provides automation that is performed internally and no additional steps are required (except for an asynchronous GLVM environment). If you want to use the Live Update function and your environment is using asynchronous GLVM, you must convert it to a synchronous GLVM during the Live Update process. After the update process completes, you can switch your environment back to an asynchronous GLVM.

PowerHA® SystemMirror® supports the Live Update function only when the cluster that is specified for the update is in an unmanaged state. During the Live Update process, PowerHA® SystemMirror® workloads continue to run with all storage devices available.

During the Live Update process, PowerHA® SystemMirror® completes the following tasks on a node:

- Verifies that any active Geographic Logical Volume Manager (GLVM) mirror pools are synchronous, and that all peer GLVM partitions belong to the same cluster.
- Suspends all GLVM network traffic. When the Live Update process completes, PowerHA® SystemMirror® resumes GLVM network traffic.
- Verifies that a Live Update process is not currently in progress on other nodes in the cluster. You can complete only one Live Update process at a time.
- Stops cluster services by using the unmanage resource group option at the beginning of the Live Update process. When the Live Update process completes, PowerHA® SystemMirror® restarts cluster services.

> **Note:** When the cluster is in an unmanaged state, PowerHA® SystemMirror® does not monitor any applications.

The Live Update process does not support asynchronous GLVM mirroring (volume groups that contain asynchronous mirror pools). If you attempt to use the Live Update process with asynchronous GLVM mirroring, the process fails and an error message is logged in the AIX® system error log. To perform a Live Update process with a GLVM configuration that uses asynchronous mirroring, complete the following steps:

1. Convert all the asynchronous mirror pools to synchronous mirror pools by running the `chmp -S -m <mirror_pool> <glvm_vg>` command.
2. Perform the Live Update process.
3. Convert the synchronous mirror pools back to asynchronous mirror pools by running the `chmp -A -m <mirror_pool> <glvm_vg>` command.

If you upgrade to AIX® Version 7.2.0, or later, you must complete the following steps to enable the Live Update function in PowerHA® SystemMirror® Version 7.2.0, or later:

1. From the command line, enter `smit sysmirror`.
2. In the SMIT interface, select **Cluster Nodes and Networks › Manage Nodes › Change/Show a Node ›**
3. Select the node from the list that uses Live Update.
4. From the **Enable AIX Live Update operation** field, select **Yes**.
5. Verify and synchronize the cluster.

> Before you update to a new AIX® Technology Level, you must complete steps 1 - 4 that is mentioned in the Updating software on a cluster topic.

Resetting cluster tunables

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

Resetting cluster tunables is useful when an administrative change does not produce ideal results, and you want to return to the default values. While this change might not produce the optimum configuration, it is likely to produce a working and stable configuration.

Use this option to reset all the tunables (customizations) made to the cluster. Using this option returns all tunable values to their default values but does not change the cluster configuration. PowerHA® SystemMirror® takes a snapshot file prior to resetting and informs you about the name and location of the snapshot file. You can choose to have PowerHA® SystemMirror® synchronize the cluster when this operation is complete.

Adding a node to a PowerHA® SystemMirror® cluster

You can use the SMIT interface to add a node when you are initially setting up the cluster or you can add a node to an existing cluster.

You can add a node to an active cluster dynamically. You do not need to stop and restart cluster services on the already-participating cluster for the new node to become part of the cluster.

To add a node to a cluster, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. You can add a node to a new cluster or to an existing cluster.

   > If you want to add more than one node at a time, you must verify and synchronize the cluster after you add each new node to the cluster.

   a. To add a node to a new cluster from the SMIT interface, select **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Setup (Custom) › Nodes › Add a Node**, and press Enter.

   b. To add a node to an existing cluster from the SMIT interface, select **Cluster Nodes and Networks › Manage Nodes › Add a Node**, and press Enter.
3. Enter field values as follows:

Add a Node fields	
Field	**Value**
Node name	Enter a unique node name for the node. The name can be up to 64 characters in length. The node name does not have to be the same as the host name of the node. You can enter one node name at a time, with up to 16 nodes in the cluster.

Communication Path to Node	Enter (or add) one resolvable IP Label (host name), IP address, or fully qualified domain name for each new node in the cluster. Each entry must be separated by a space. PowerHA® SystemMirror® uses this path to initiate communication with the node.
	Example 1:
	`10.11.12.13 NodeC.ibm.com`
	Example 2:
	`NodeA NodeB`
	You can also press F4 to select IP labels and IP addresses from a list that are added to the **/etc/hosts** file, but are not configured in PowerHA® SystemMirror®.

> The AIX® Live Update operation is automatically enabled when you add a node with PowerHA® SystemMirror® 7.2.0, or later, on a system that is running the AIX® Version 7.2, or later, operating system. You can use the AIX® Live Update process to apply a kernel interim fix (iFix) without rebooting the system. To disable the Live Update operation, use the **Change / Show a Node** menu in the SMIT interface.

4. Verify that all fields are correct, and press Enter.
5. Verify and synchronize the cluster. After the verification and synchronization processes are complete, you can start cluster services on the newly added node to integrate it into the cluster.

Adding a PowerVM® NovaLink definition

You can use the SMIT interface to add a PowerVM® NovaLink definition that is used by the PowerHA® SystemMirror® cluster.

To add a PowerVM® NovaLink definition, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optmized High Availability › NovaLink Configuration › Add NovaLink Definition**, and press Enter.
3. Complete the following fields, and press Enter.

   NovaLink name

   If the PowerVM® NovaLink name or IP address is not discovered automatically, enter the name or IP address of PowerVM® NovaLink.

   DLPAR operations timeout

   Enter a timeout value, in minutes, for the Dynamic Logical Partitioning (DLPAR) commands that are run on a PowerVM® NovaLink. If you do not specify a value, the default values, which are specified in the **Change/Show Default NovaLink Tunables** SMIT panel, are used. The default values are:

   ```
   DLPAR operations timeout (in minutes)                    10
   Number of retries                                        5
   Delay between retries (in seconds)                       10
   Connection Type                                          SSH
   ```

   Number of retries

   Enter the number of times you want any command to try before the PowerVM® NovaLink is considered as not responding. The next PowerVM® NovaLink in the list is used after the number of failed retries that you entered. If you do not specify a value, the default values, which are specified in the **Change/Show Default NovaLink Tunables** SMIT panel, are used.

   Delay between retries

   Enter the number of seconds you want to delay before attempting to resend a PowerVM® NovaLink command. If you do not specify a value, the default values, which are specified in the **Change/Show Default NovaLink Tunables** SMIT panel, are used.

   User name

   Specify a user name that can be used with the Representational State Transfer (REST) application programming interface (API) to establish a secure connection with the PowerVM® NovaLink or SSH. You must specify the associated password for the user name when the connection type is REST API or when connection type is SSH, you can verify whether the user name can be used without a password.

   The following limitations applys to PowerVM® NovaLink:

- With the PowerHA® SystemMirror® 7.2.2, and later, the PowerVM® NovaLink does not support the Representational State Transfer (REST) application programming interface (API) connection type.
- PowerVM® NovaLink cannot be used for the Enterprise Pool CoD or On/Off CoD operations. You can use an Hardware Management Console (HMC) for the Enterprise Pool CoD or On/Off CoD operations. If the HMC is not defined in the PowerHA® SystemMirror® cluster, the Enterprise Pool CoD or On/Off CoD operations are not supported.
- To run the Enterprise Pool or On/Off CoD operations on HMC, the HMC must be set as controller by using the `setmaster` option. For both these operations, the `setmaster` option is handled automatically by PowerHA® SystemMirror®.
- If multiple clusters are running on the same CEC, ensure that each ROHA operation that is running does not interfere with all the other ROHA operations that are running simultaneously.
- If an HMC is controller, the PowerVM® NovaLink cannot run operations on any operating system. Any requests to run the operations results in an error that can impact the other nodes in a managed system.
- If the managed system is managed by both the HMC and PowerVM® NovaLink, the PowerVM® NovaLink must be defined in PowerHA® SystemMirror® cluster. If the PowerVM® NovaLink is not defined in the PowerHA® SystemMirror® cluster and HMC is not set as controller, then all DLPAR operations might be impacted.

Configuring PowerHA® SystemMirror® networks

You can configure multiple networks in PowerHA® SystemMirror® to control application traffic over the cluster network interfaces. Use the **Manage Networks and Network Interfaces › Networks** path in SMIT to add, change, show, or remove networks from the cluster.

To speed up the configuration process, run discovery before configuring networks.

Configuring networks

You can use the System Management Interface Tool (SMIT) to configure IP-based networks to use PowerHA® SystemMirror®.

To configure networks, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. When initially setting up the cluster, if you are using the custom configuration path, you can use the **Networks** menu under **Initial Cluster Configuration (Custom › Cluster Nodes and Networks** to add networks to the new cluster configuration.
   After initial configuration of the cluster is complete and you want to add additional networks to an existing cluster, you can use the menus found directly under **Cluster Nodes and Networks › Manage Networks and Network Interfaces › Networks** from the main PowerHA® SystemMirror® SMIT menu.
3. Select the type of network to configure.
4. Enter the information as follows:

Network fields	
Field	**Value**
Network Name	If you do not enter a name, PowerHA® SystemMirror® will give the network a default network name made up of the type of network with a number appended (for example, `net_ether_01`). If you change the name for this network, use no more than 128 alphanumeric characters and underscores.
Network Type	This field is filled in depending on the type of network you selected.
Netmask(IPv4)/Prefix Length(IPv6)	For the configuration of the IP version 4 service interface, enter the network mask for the address. For the configuration of the IP version 6 service interface, enter the prefix length for the address.

Field	Value
	This field is not a required field. If you do not enter a value, the prefix length or netmask of the underlying network is used. If a prefix length value or netmask value is specified, it is checked for compatibility with the underlying network.
Instability Threshold	The network is considered unstable when more than this number of state changes are detected within the Instability Period.
Instability Period	Used with the Instability Threshold to determine when a network is unstable. The network is considered unstable when more than the Threshold number of state changes occur within this time period. If the network was unstable, and this period of time passes without any new state changes, the network is then considered stable. Period is specified in seconds.

5. Press Enter to configure this network.
6. Repeat the operation to configure more networks.

Configuring an application service interface

If you already have an application that is active and using a particular IP Address as a base address on network interface, you can configure this service IP label in PowerHA® SystemMirror® without disrupting your application.

If you are configuring the cluster when the application is not active, you do not need to follow this procedure.

The following steps guide you through configuring your application service IP label in PowerHA® SystemMirror® in order not to disrupt your application:

1. Configure a PowerHA® SystemMirror® cluster
2. Configure PowerHA® SystemMirror® nodes
3. Configure PowerHA® SystemMirror® networks
4. Run Discovery.
5. Configure PowerHA® SystemMirror® network interfaces.
6. Run verification and synchronization to propagate your configuration to all the nodes.
7. For each node that has an application using a particular IP Address:
   a. For the network interface currently hosting the application IP address, determine a new address to use as the interfaces base address. This address will be configured on the interface when the system boots and is referred to as the Boot_IP_Address. The cluster manager will alias the application IP address onto the interface when bringing the application online during normal cluster operation, but you will run the following command to perform this step manually initially to avoid disrupting your application.
   b. Run the sample utility **clchipdev**:

```
/usr/es/sbin/cluster/samples/appsvclabel/clchipdev
```

Result
The utility **clchipdev** helps configure an application service interface correctly in PowerHA® SystemMirror® when you have an active application that is using a particular IP Address as a base address on network interface before starting PowerHA® SystemMirror®.

```
clchdev -n NODE -w network_name -a 'App_IP_Address=Boot_IP_Address'
```

Where:

- NODE is the nodename.
- network_name is the name of the network that contains this service interface.
- App_IP_Address is the IP Address currently in use by the application (and currently configured in the CuAt as the base address for the given interface).

- Boot_IP_Address is the IP Address that is to be used as the new base (boot) address.

For example, if NodeA has an IP Address 10.10.10.1 that is being used to make an application highly available, you would use the following steps:

1. Run the utility **clchipdev**.
   `clchdev -n NodeA -w net_ip -a '10.10.10.1=192.3.42.1'.`

   The sample utility performs the following:

   - Performs **rsh** to NodeA and determines the network interface for which 10.10.10.1 is currently configured as the base address.
   - Determines the network interface to be en0.
   - Determines the network type as defined in PowerHA® SystemMirror®network ODM, using the network name.
   - Runs: `chdev -l en0 -a netaddr=192.3.42.1 -P`
     This changes the CuAt on that node to use the new Boot_IP_Address as the base address.
   - Replaces 10.10.10.1 in PowerHA® SystemMirror®adapter ODM with 192.3.42.1.
   - Configures to PowerHA® SystemMirror® the IP Address 10.10.10.1 as a service IP address.
2. Add this service IP label to a resource group.
3. Run verification and synchronization.

Configuring network interfaces to PowerHA® SystemMirror®

You can define which network interfaces are used to host clustered application IP traffic.

When you are configuring network interfaces to PowerHA® SystemMirror® components, you can have the following scenarios:

- Network interfaces are already configured to AIX®, and you have run the PowerHA® SystemMirror® discovery process to add them to PowerHA® SystemMirror® picklists to aid in the PowerHA® SystemMirror® configuration process.
- Network interfaces are already configured to AIX®, and need to be configured to PowerHA® SystemMirror® (no discovery was run).
- Network interfaces need to be defined to AIX® before you can configure them in PowerHA® SystemMirror®. In this case, you should use the AIX® SMIT menus to define new network interfaces with base IP addresses prior to adding them to the PowerHA® SystemMirror® cluster.
  To configure network interfaces to the AIX® operating system without leaving PowerHA® SystemMirror® SMIT, use the **System Management (C-SPOC) › Communication Interfaces** SMIT path.

Configuring PowerHA® SystemMirror® persistent node IP labels/addresses

A persistent node IP label is an IP alias that can be assigned to a network for a specified node.

A persistent node IP label is a label that:

- Always stays on the same node (is node-bound)
- Co-exists with other IP labels present on an interface
- Does not require installing an additional physical interface on that node
- Always stays on an interface of the PowerHA® SystemMirror® network and will not move to an interface.
- *Is not* part of any resource group.

Assigning a persistent node IP label for a network on a node allows you to have a node-bound address on a cluster network that you can use for administrative purposes to access a specific node in the cluster.

To add persistent node IP labels:

1. From the command line, enter `smit sysmirror`

2. In SMIT, select **Cluster Nodes and Networks › Manage Nodes › Configure Persistent Node IP Labels/ Addresses › Add a Persistent Node IP Label** and press Enter.
3. Enter the field values as follows:

Add a Persistent Node IP Label fields	
Field	**Value**
Node Name	The name of the node on which the IP label/address will be bound.
Network Name	The name of the network on which the IP label/address will be bound.
Node IP Label/Address	The IP label/address to keep bound to the specified node.

4. Press Enter.

Result
If you are using persistent node IP Labels/Addresses, note the following issues:

- You can define only one persistent IP label on each node per cluster network.
- Persistent IP labels become available at a node's boot time.
- Once a persistent IP label is configured for a network interface on a particular network on a particular node, it becomes available on that node on a boot interface at operating system boot time and remains configured on that network when PowerHA® SystemMirror® is shut down on that node.
- You can remove a persistent IP label from the cluster configuration using the **Remove a Persistent Node IP Label/Address** SMIT panel. However, after the persistent IP label has been removed from the cluster configuration, it is not automatically deleted from the interface on which it was aliased. In order to completely remove the persistent IP label from the node, you should manually remove the alias with the **ifconfig delete** command or reboot the cluster node.
- Configure persistent node IP labels individually on each node. You cannot use the PowerHA® SystemMirror® discovery process for this task.
- To change or show persistent node IP labels, use the **Change/Show a Persistent Node IP label** SMIT menu.

Managing a backup repository disk with SMIT

You can use the System Management Interface Tool (SMIT) to add backup repository disk, remove a backup repository disk that is not active, and view all configured repository disks.

To manage a backup repository disk with SMIT, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Nodes and Networs › Manage Repository Disks**, and press Enter.
3. Select one of the following options:
   Add a Repository Disk

   Select this option to add an extra backup repository disk to the cluster.

   Remove a Repository Disk

   Select this option to remove a backup repository disk from the cluster. You cannot remove an active repository disk. However, you can replace an active repository disk with an existing repository disk.

   Show Repository Disks

   Select this option to view all repository disks that are configured in the cluster.
4. If you add or remove a backup repository disk, you must verify and synchronize the cluster.

Configuring the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable

The Reliable Scalable Cluster Technology (RSCT) suite consists of multiple subsystems that provide various clustering tools and solutions for several products on AIX. The traditional RSCT Peer Domain (RPD), when configured without Cluster-Aware AIX (CAA), consists of several critical subsystems that monitor each other,

and failures or excessive blockages detected between them triggers a node failure (often a panic or reboot), if necessary to protect the integrity of resources shared between multiple nodes in a cluster.

However, in the CAA environment used by PowerHA® SystemMirror®, the Group Services subsystem`cthags`is the only critical element on which PowerHA® SystemMirror® depends directly. If other critical subsystems are allowed to recycle without impairing the functionality provided by the RSCT layer, their failure and restart are allowed by PowerHA® SystemMirror®.

The capability to allow certain subsystem restarts was introduced as part of the AIX version 7.2, Technology Level 4 (and was also included in the base of AIX version 7.3.0). However, it was not enabled by default.

The option is implemented in seconds, which indicates the time to allow problematic subsystems to restart before exiting the node. For example, if you use a value of 30, and one of the critical RSCT subsystems other than`cthags`fails, the subsystem has 30 seconds grace period to restart. If the subsystem restarts successfully and returns to normal operation within 30 seconds, the node continues to run. However, if the restart fails, the node is exited without this option and exited by normal methods.

There are two ways to implement this change in PowerHA® SystemMirror®, depending on what code level of the product is in use.

Customers with PowerHA® SystemMirror® version 7.2.5 or higher should set this tunable through the **`clmgr`** command, either at the cluster or node level:

To set the value of the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable by using the `/etc/ctfile.cfg` RSCT configuration file, enter the following syntax:

```
CritDaemonRestartGracePeriod=<value in secs>
```

To set the value of the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable at a node level or cluster level by using the **`clmgr`** command:

```
clmgr add cluster <clustername> CRIT_DAEMON_RESTART_GRACE_PERIOD=<value in secs>
clmgr modify cluster CRIT_DAEMON_RESTART_GRACE_PERIOD=<value in secs>
clmgr add node <nodename> CRIT_DAEMON_RESTART_GRACE_PERIOD=<value in secs>
clmgr modify node <nodename> CRIT_DAEMON_RESTART_GRACE_PERIOD=<value in secs>
```

> **Note:** For PowerHA® SystemMirror®, the supported range of the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable is 0 - 240 seconds.

The CRIT_DAEMON_RESTART_GRACE_PERIOD tunable supports the following values:

0

The node is halted when the RMC subsystems or IBMConfigRm daemon fails. This is the default value.

>0

RSCT will wait for the specified grace period and if the RMC subsystems or IBMConfigRm daemon are not restarted within the specified grace period, the node will be halted.

A value of **0** means that no grace is allowed and all subsystem failures trigger an immediate reaction. This requires the process to abort a potentially hung subsystem by an external signal from another process, write the core file to the disk, restart the system with SRC, and finally reconnect to the`cthags`subsystem as a new client. It is not advisable to have a safe grade period of less than 10 seconds as there are other system activity that may continue, and a heavy workload or I/O issues may impact process core dumping to the disk.

You can use the cluster level setting for the same grace timeout on all nodes. Setting a value at the node level supersedes the cluster setting on that node only. This is required if you have certain nodes with very high workload or other performance factors, different from other nodes in the cluster. You can use PowerHA® SystemMirror® Version 7.2.4 or lower, with AIX Version 7.2 TL4 to set the grace period directly at the RSCT layer. For more information, see https://www.ibm.com/support/pages/apar/IJ29784.

> **Note:** If you use the **`clmgr`** command in PowerHA® SystemMirror® Version 7.2.5, setting the node-specific timeout supersedes any cluster-wide setting.

In PowerHA® SystemMirror®, the value of the CRIT_DAEMON_RESTART_GRACE_PERIOD tunable set by the `/etc/ctfile.cfg` RSCT configuration file is given highest priority followed by node level and then the cluster level settings. If all three options are enabled, the value set in the `/etc/ctfile.cfg` RSCT configuration file value is consider as the grace period.

After you configure the cluster level or node level tunable values, you must run the verify and synchronization operation of the cluster to update the RSCT CritDaemonRestartGracePeriod tunable.

Based on the configuration settings specified by a user, few warning messages are displayed during configuration or verify and synchronization operation of the cluster.

Scenario 1

When you configure both cluster level and node level tunable for any specific node, the following warning message is displayed during verify and synchronization operation of the cluster:

> **Warning:** Node level critical daemon restart grace period is configured for the following nodes: node1 node2. Node level configuration has highest priority then cluster level so the cluster level attribute is ignored for the specified nodes.

Scenario 2

During verify and synchronization operation of the cluster, the preferred value of the CRIT_DAEMON_RESTART_GRACE_PERIOD is displayed. If you sets the CRIT_DAEMON_RESTART_GRACE_PERIOD value for node level or cluster level to more than 120, the following warning message is displayed during verify and synchronization operation of the cluster:

> **Warning:** Critical daemon restart grace period at node level is configured for the following node: node1 with 130 seconds. The preferred value is less than 120 seconds.

Replacing a repository disk with SMIT

Cluster Aware AIX® (CAA) detects when a repository disk failure occurs and generates a notification message. You will continue to receive notification messages until you replace the failed repository disk with a new repository disk.

The cluster operates in a restricted mode until you replace the failed repository disk. You cannot change the cluster configuration or rejoin nodes to the cluster until you replace the failed repository disk.

To replace a repository disk with a new disk, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Problem Determination Tools › Replace the Primary Repository Disk**, and press Enter.
3. In the **Repository Disk** field, press F4 (List) to select an available disk from all nodes in the cluster, or enter the name of the disk if a backup repository disk has not been configured.

   > When you first access the **Replace the Primary Repository Disk** window, the **Repository Disk** field displays the current repository disk.

4. Press Enter to set up the selected disk as the new repository disk for the cluster.
5. After synchronizing the configuration, you can verify that the new repository disk is working by running the **`/usr/sbin/lscluster -d`** command.

Configuring PowerHA® SystemMirror® resources

After you have configured the cluster topology, continue setting up your cluster by configuring the resources that are for the resource groups. Use the SMIT interface to configure resources to support highly available applications.

In SMIT, use the **Cluster Applications and Resources › Resources** path, to configure the following resources:

- Application controller
- Service IP label
- Shared volume group
- File system
- Application monitors
- Tape drive
- User-defined resources

Configuring service IP labels as PowerHA® SystemMirror® resources

Before you start configuring service IP labels as PowerHA® SystemMirror® resources, you must understand how the network for your environment is configured.

For the initial configuration, follow the procedures described in this section.

Discovering IP network information (optional)

You can choose to run the PowerHA® SystemMirror® cluster information discovery process. If you choose to run discovery, all communication paths must be configured first. Then PowerHA® SystemMirror® will discover nodes, networks, and communication interfaces and devices for you and show them in the SMIT picklists. If you choose not to run discovery, PowerHA® SystemMirror® will only include in the picklist network information that is predefined in AIX®.

To run cluster discovery, complete the following steps:

1. Enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Discover Network Interfaces and Disks** and press Enter. PowerHA® SystemMirror® retrieves current AIX® configuration information from all cluster nodes. This information is displayed in picklists to help you make accurate selections of existing components. PowerHA® SystemMirror® informs you about which components have been discovered by the system. Predefined components (those that are supported but are not discovered) are also made available as selections in picklists.

Configuring service IP labels and addresses

This topic discusses configuring service IP labels and addresses.

To add service IP labels/addresses as resources to the resource group in your cluster:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Service IP Labels/ Addresses › Add a Service IP Label/Address** and press Enter.
3. Fill in field values as follows:

Add a Service IP Label/Address fields	
Field	**Value**
IP Label/Address	Enter, or select from the picklist the IP label/address to be kept highly available.
Network Name	Enter the symbolic name of the PowerHA® SystemMirror® network on which this Service IP label/ address will be configured.

4. Press Enter after completing all required fields. PowerHA® SystemMirror® now checks the validity of the IP label/address configuration.
5. Repeat the previous steps until you have configured all service IP labels/addresses for each network, as needed.

Distribution preference for service IP label aliases

You can configure a distribution preference for the service IP labels that are placed under PowerHA® SystemMirror® control.

A distribution preference for service IP label aliases is a network-wide attribute used to control the placement of the service IP label aliases on the physical network interface cards on the nodes in the cluster.

Configuring a distribution preference for service IP label aliases does the following:

- Lets you customize the load balancing for service IP labels in the cluster.
- Enables PowerHA® SystemMirror® to redistribute the alias service IP labels according to the preference you specify.
- Allows you to configure the type of distribution preference suitable for the VPN firewall external connectivity requirements.
- The distribution preference is exercised as long as there are acceptable network interfaces available. PowerHA® SystemMirror® always keeps service IP labels active, even if the preference cannot be satisfied.

Rules for the distribution preference for service IP label aliases

The following rules apply to the distribution preference:

- If you do not specify any preference, PowerHA® SystemMirror® by default distributes all service IP label aliases across all available boot interfaces on a network using the IPAT via IP Aliasing function. For more information on how the default method for service IP label distribution works, see Resource group behavior during cluster events.
- If there are insufficient network interface cards available to satisfy the preference that you have specified, PowerHA® SystemMirror® allocates service IP label aliases to an active network interface card that may be hosting other IP labels.
- You can change the IP labels distribution preference dynamically: The new selection becomes active during subsequent cluster events. (PowerHA® SystemMirror® does not require the currently active service IP labels to conform to the newly changed preference.)
- If you did not configure persistent labels, PowerHA® SystemMirror® lets you select the Collocation with Persistent and Anti-Collocation with Persistent distribution preferences, but it issues a warning and uses the regular collocation or anti-collocation preferences by default.
- When a service IP label fails and another one is available on the same node, PowerHA® SystemMirror® recovers the service IP label aliases by moving them to another NIC on the same node. During this event, the distribution preference that you specified remains in effect.
- You can view the distribution preference per network using the **cltopinfo** or the **cllsnw** commands.

Types of distribution for service IP label aliases

You can specify in SMIT the distribution preferences for the placement of service IP label aliases

These preference include:

Type of distribution preference	Description
Anti-collocation	This is the default. PowerHA® SystemMirror® distributes all service IP label aliases across all boot IP labels using a "least loaded" selection process.
Anti-collocation with source	Service labels are mapped using the Anti-Collocation preference. If there are not enough adapters, more than one service label can be placed on one adapter. This choice will allow one label to be chosen as source address for outgoing communication.
Collocation	PowerHA® SystemMirror® allocates all service IP label aliases on the same network interface card (NIC).

Type of distribution preference	Description
Collocation with source	Service labels are mapped using Collocation preference. This choice will allow to choose one service label as source for outgoing communication. The service label chosen in the next field is source address.
Anti-collocation with persistent	PowerHA® SystemMirror® distributes all service IP label aliases across all active physical interfaces that are *not* hosting the persistent IP label. PowerHA® SystemMirror® will place the service IP label alias on the interface that is hosting the persistent label only if no other network interface is available. If you did not configure persistent IP labels, PowerHA® SystemMirror® lets you select the Anti-Collocation with Persistent distribution preference, but it issues a warning and uses the regular anti-collocation preference by default.
Anti-collocation with persistent label and source	Service labels will be mapped using the Anti-Collocation with Persistent preference. One service address can be chosen as a source address for the case when there are more service addresses than the boot adapters.
Collocation with persistent	All service IP label aliases are allocated on the same NIC that is hosting the persistent IP label. This option may be useful in VPN firewall configurations where only one interface is granted external connectivity and all IP labels (persistent and service) must be allocated on the same interface card. If you did not configure persistent IP labels, PowerHA® SystemMirror® lets you select the Collocation with Persistent distribution preference, but it issues a warning and uses the regular collocation preference by default.

Steps to configure distribution preference for service IP label aliases

This topic describes the procedure to configure distribution preference for service IP label aliases on any cluster node.

To configure a distribution preference:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Service IP Labels/ Addresses › Configure Service IP Labels/Addresses Distribution Preferences** and press Enter.
   A list of available networks is displayed.
3. Select a network for which you want to specify the distribution preference.
4. SMIT displays the **Configure Resource Distribution Preferences** screen. Enter field values as follows:

Configure Resource Distribution Preferences	
Field	**Value**
Network Name	The field is filled in with the network for which you want to specify or change the distribution preference for service IP label aliases.
Distribution Preference	From the picklist, select the distribution preference as follows: • **Anti-collocation**. This is the default. PowerHA® SystemMirror® distributes all service IP label aliases across all boot IP labels using a least loaded selection process. • **Anti-collocation with source**. Service labels are mapped using the Anti-Collocation preference. If there are not enough adapters, more than one service label can be placed on one adapter. This choice will allow one label to be chosen as source address for outgoing communication. • **Collocation**. PowerHA® SystemMirror® allocates all service IP label aliases on the same network interface card (NIC).

Field	Value
	• **Collocation with source**. Service labels are mapped using Collocation preference. This choice will allow to choose one service label as source for outgoing communication. The service label chosen in the next field is source address. • **Anti-collocation with persistent**. PowerHA® SystemMirror® distributes all service IP label aliases across all active physical interfaces that are *not* hosting the persistent IP label. PowerHA® SystemMirror® allocates the service IP label alias on the interface that is hosting the persistent label *only* if no other interface is available. • **Anti-Collocation with persistent label and source**. Service labels will be mapped using the Anti-Collocation with Persistent preference. One service address can be chosen as a source address for the case when there are more service addresses than the boot adapters. • **Collocation with persistent**. All service IP label aliases are allocated on the same NIC that is hosting the persistent IP label. This option may be useful in firewall configurations where only one interface is granted external connectivity and all IP labels (persistent and service) must be allocated on the same interface card.
Source IP Label for outgoing packets	This field allows to choose a Service or persistent address to be used as a source address on the selected network. All the service labels and persistent labels will be shown as choices.

If you did not configure persistent IP labels, PowerHA® SystemMirror® lets you select the **Collocation with Persistent** and **Anti-Collocation with Persistent** distribution preferences but issues a warning and uses the regular collocation or anti-collocation preferences by default.

5. Press Enter to add this information to the PowerHA® SystemMirror® Configuration Database on the local node. Return to previous PowerHA® SystemMirror® SMIT screens to perform other configuration tasks.
6. Verify and synchronize the changes to the cluster configuration. If the Cluster Manager is running on the local node, synchronizing the cluster resources triggers a dynamic reconfiguration event.

Configuring PowerHA® SystemMirror® application controller scripts

An *application controller* is a cluster component that is included in the resource group as a cluster resource, and that is used to control an application that must be highly available. An application controller consists of application start and stop scripts.

Configuring an application controller does the following:

- Associates a meaningful name with the application. For example, you could give the tax software a name such as taxes. You then use this name to refer to the application controller when you define it as a resource. When you set up the resource group, you add an application controller as a resource.
- Points the cluster event scripts to the scripts that they call to start and stop the application.
- Allows you to then configure application monitoring for that application.

Note: This section does not discuss how to write the start and stop scripts. See the vendor documentation for specific product information on starting and stopping a particular application.

Configuring volume groups, logical volumes, and file systems as resources

You define volume groups, logical volumes, and file systems in the AIX® operating system and then configure them as resources for PowerHA® SystemMirror®.

Configuring concurrent volume groups, logical volumes, and file systems as resources

Concurrent volume groups, logical volumes, and file systems must be defined in AIX® and then configured as resources for PowerHA® SystemMirror®.

Configuring multiple application monitors

PowerHA® SystemMirror® can monitor specified applications using application monitors.

These application monitors can:

- Check if an application is running before PowerHA® SystemMirror® starts it.
- Watch for the successful startup of the application.
- Check that the application runs successfully after the stabilization interval has passed.
- Monitor both the startup and the long-running process.
- Automatically take action to restart applications upon detecting process termination or other application failures.

You can configure multiple application monitors and associate them with one or more application controllers.

By supporting multiple monitors per application, PowerHA® SystemMirror® can support more complex configurations. For example, you can configure one monitor for each instance of an Oracle parallel server in use. Or, you can configure a custom monitor to check the health of the database along with a process termination monitor to instantly detect termination of the database process.

> **Note:** If a monitored application is under control of the system resource controller, ensure that *action:multi* are **-O** and **-Q**. The **-O** specifies that the subsystem is not restarted if it stops abnormally. The **-Q** specifies that multiple instances of the subsystem are not allowed to run at the same time. These values can be checked using the following command:

```
lssrc -Ss Subsystem | cut -d : -f 10,11
```

If the values are not **-O** and **-Q**, change them using the **chssys** command.

Process and custom monitoring

You can select either process application monitoring or custom application monitoring method.

- *Process application monitoring* detects the termination of one or more processes of an application.
- *Custom application monitoring* checks the health of an application with a custom monitor method at user-specified polling intervals.

Process monitoring is easier to set up, as it uses the built-in monitoring capability provided by the operating system and requires no custom scripts. However, process monitoring may not be an appropriate option for all applications. Custom monitoring can monitor more subtle aspects of an application's performance and is more customizable, but it takes more planning, as you must create the custom scripts.

Using shell environment variables in custom PowerHA® scripts

You can use shell environment variables in custom controller scripts, custom monitoring scripts, cluster event scripts, or other custom scripts used in PowerHA® SystemMirror®.

When you write your controller scripts, custom monitoring scripts, cluster event scripts, or other custom scripts, none of the shell environment variables that are defined in the `/etc/environment` file are available to your program. If you need to use any of these variables, you must explicitly source them by including the following line in your script:

```
. /etc/environment
```

Fallover and notify actions

In both process and custom monitoring methods, when the monitor detects a problem, PowerHA® SystemMirror® attempts to restart the application on the current node and continues the attempts until the specified restart count has been exhausted.

When an application cannot be restarted within the restart count, PowerHA® SystemMirror® takes one of two actions, which you specified when configuring the application monitor:

- Choosing **fallover** causes the resource group containing the application to fall over to the node with the next highest priority according to the nodelist. (See Application monitoring prerequisites and considerations for more information.)
- Choosing **notify** causes PowerHA® SystemMirror® to generate a **server_down** event, which informs the cluster of the failure.

Monitor modes

When you configure process monitors and custom monitors for the application controller, you can also specify the mode in which the application monitor is used.

You can specify the following modes for an application monitor:

Startup Monitoring Mode

In this mode, the application monitor runs when cluster services are starting. The monitor checks whether the application is already running on the node. If the monitor indicates that the application is running, cluster services does not run the application controller start script during the startup operation process. This function is useful if you have previously stopped cluster services by using the **unmanage** option. The **unmanage** option keeps the application and other resources active on the node. When cluster services are restarting, the startup monitor can indicate whether the application is still running and avoid creating a second instance by running the controller start script.

The startup monitor is stopped after a specific amount of time that is specified by the stabilization interval for the application controller. If the monitor returns within the stabilization interval, a zero return code indicates that the application is already running and the controller start script is not run. If the monitor returns a nonzero code within the stabilization interval, or if it does not return at all, the application is not running and the controller start script begins to run.

Startup monitors are important if you are using parent/child resource dependencies or start after resource group dependencies. A startup monitor also verifies that the application in the parent resource group or target resource group is running before the dependent resource group is started.

Long-Running Mode

In this mode, the application monitor is used after the application is started and the application monitor checks whether that the application is still running. The monitoring starts after the stabilization interval expires. For custom monitors, the user supplied monitor is called at regular intervals. In SMIT, the value for the interval is specified in the **Monitor Interval** field. For process monitors, the monitoring relies on the Reliable Scalable Cluster Technology (RSCT) subsystem that can detect the end of a specified process.

You can configure a monitor in this mode for any application controller, and you can specify multiple monitors that indicate the health of the application.

For example, you can configure a process application monitor to immediately indicate the end of a critical database process. You can also configure a custom monitor that periodically sends transactions to the

database. The custom monitor verifies that the database is actively responding to requests, while the process monitor immediately detects the end of a critical database process.

> **Note:** In this mode, the application monitors do not run at the start of the application, and do not check the status of the running application. If you start the cluster services operation after you unmanage the node, the start script handles the running status of the application.

Both

In this mode, the same application monitor is used for both the startup monitor mode and the long-running mode. When you are writing and testing a custom monitor to be used in this mode, the same monitor must operate correctly when started under different circumstances. When the monitor is called during startup, the monitor is run only for the amount of time specified by the stabilization interval. If the monitor does not return within the amount of time that is specified by stabilization interval, the monitor is stopped and the application start script is run.

The interval that you specified for the **Monitor Interval** field in SMIT, determines how often the custom long-running monitor is run after the application is started and is stable.

If you use the same application monitor for both startup and long-running monitoring, you must verify that the monitor can reliably determine the application state and return an appropriate indication within the associated time period.

Retry count and restart interval

The restart behavior depends on two parameters, the *retry count* and the *restart interval*, that you configure in SMIT.

- *Retry count*. The retry count specifies how many times PowerHA® SystemMirror® should try restarting before considering the application failed and taking subsequent fallover or notify action.
- *Restart interval*. The restart interval dictates the number of seconds that the restarted application must remain stable before the retry count is reset to zero, thus completing the monitor activity until the next failure occurs.

> **Note:** Do not specify both of these parameters if you are creating an application monitor that will only be used as in a startup monitoring mode.

If the application successfully starts up before the retry count is exhausted, the restart interval comes into play. By resetting the restart count, it prevents unnecessary fallover action that could occur when applications fail several times over an extended time period. For example, a monitored application with a restart count set to three (the default) could fail to restart twice, and then successfully start and run cleanly for a week before failing again. This third failure should be counted as a new failure with three new restart attempts before invoking the fallover policy. The restart interval, set properly, would ensure the correct behavior: it would have reset the count to zero when the application was successfully started and found in a stable state after the earlier failure.

Be careful not to set the restart interval for a too short period of time. If the time period is too short, the count could be reset to zero too soon, before the immediate next failure, and the fallover or notify activity will never occur.

Application monitoring prerequisites and considerations

This topic discusses some prerequisites and considerations for planning and configuring application monitoring.

Keep the following in mind:

- Any application to be monitored must be defined to an application controller in an existing cluster resource group.
- If you have configured dependent resource groups, we recommend to configure multiple monitors:
  - For applications included in parent resource groups, and for applications in child resource groups.

- For applications included in target resource groups, and for applications in source resource groups in start after dependency and stopafter dependency

  For example, a monitor for a parent resource group can monitor the successful startup of the application, and a monitor for a child resource group can monitor the process for an application. For more information, see Monitor modes.

- Multiple monitors can be configured for the same application controller. Each monitor can be assigned a unique name in SMIT.
- The monitors that you configure must conform to existing configuration rules. For more information, see Configuring a process application monitor and Configuring a custom application monitor.
- It is recommend that you first configure an application controller, and then configure the monitor(s) that you can associate with the application controller. Before configuring an application monitor, configure all your application controllers. Then configure the monitors and associate them with the controllers. You can go back at any time and change the association of monitors to controllers.
- You can configure no more than 128 monitors per cluster. No limit exists on the number of monitors per application controller, as long as the total number of all monitors in the cluster is less than 128.
- When multiple monitors are configured that use different fallover policies, each monitor specifies a failure action of either "notify" or "fallover". PowerHA® SystemMirror® processes actions in the order in which the monitors indicate an error. For example, if two monitors are configured for an application controller and one monitor uses the "notify" method and the other uses the "fallover" method, the following occurs:
  - If a monitor with "fallover" action indicates an error first, PowerHA® SystemMirror® moves the resource group to another node, and the remaining monitor(s) are shut down and restarted on another node. PowerHA® SystemMirror® takes no actions specified in any other monitor.
  - If a monitor with "notify" action indicates an error first, PowerHA® SystemMirror® runs the "notify" method and shuts down that monitor, but any remaining monitors continue to operate as before. You can manually restart the "notify" monitor on that node using the **Suspend/Resume Application Monitoring** SMIT panel.
- If multiple monitors are used, PowerHA® SystemMirror® does not use a particular order for the monitors startup or shutdown. All monitors for an application controller are started at the same time. If two monitors are configured with different fallover policies, and they fail at precisely the same time, PowerHA® SystemMirror® does not guarantee it processes methods specified for one monitor before methods for the other.
- The same monitor can be associated with multiple application controllers using the **Application Monitor(s)** field in the **Change/Show an Application Controller** SMIT panel. You can select a monitor from the picklist.
- If you remove an application monitor, PowerHA® SystemMirror® removes it from the definition for all application controllers that were using the monitor, and indicates which application controllers are no longer using the monitor.
- If you remove an application controller, PowerHA® SystemMirror® removes it from the definition of all application monitors that were configured to monitor the application. PowerHA® SystemMirror® also sends a message about which monitor will no longer be used for the application. If you remove the last application controller in use for any particular monitor, that is, if the monitor will no longer be used for any application, verification issues a warning that the monitor will no longer be used.
- If you configure an application monitor for an application controller, PowerHA® SystemMirror® starts the monitor when the application is being brought online to determine the state of the application. PowerHA® SystemMirror® event processing is suspended for the number of seconds specified in the stabilization interval for the monitoring script . If you did not configure an application monitor, PowerHA® SystemMirror® event processing is suspended for 10 seconds to allow the application to start.

Steps for configuring multiple application monitors

These topic outline the procedures for configuring multiple application monitors.

To define multiple application monitors for an application:

- Define one or more application controllers. For instructions, see Configuring application controller scripts.

- Add the monitors to PowerHA® SystemMirror®. The monitors can be added using the following path in SMIT, **Configure Applications and Resources › Resources › Configure Resources › Configure User Applications (Scripts and Monitors) > Application Monitors**.

Configuring a process application monitor

You can configure multiple application monitors and associate them with one or more application controllers. By supporting multiple monitors per application, PowerHA® SystemMirror® can support more complex configurations.

Process application monitoring detects the termination of a process and generates an event. This section describes how to configure process application monitoring, in which you specify one or more processes of a single application to be monitored.

> **Note:** Process monitoring may not be the appropriate solution for all applications. For instance, you cannot monitor a shell script with a process application monitor. If you want to monitor a shell script, configure a custom monitor.

Identifying correct process names

For process monitoring, it is important that you list the correct process names in the SMIT **Add Process Application Monitor** panel. You must use the processes that are listed in response to the **ps -e** command and not use the **-f** flag.

Any process that is started through a #!<path name> in the script must use the processes that are listed in response to the **ps -e** command. For example, the **bsh** command and the **csh** command.

To identify correct process names in your process list, complete the following steps:

1. Enter the following command:
   ```
   ps -e | awk '{print $4}' | sort -u >/tmp/list1
   ```
2. Run the application controller start script.
3. Enter the following command:
   ```
   ps -e | awk '{print $4}' | sort -u >/tmp/list2
   ```
4. Compare the two lists by entering:
   ```
   diff list1 list2 | grep \>
   ```

Result
The result is a complete and accurate list of possible processes to monitor. You might choose not to include all of them in your process list.

Steps for configuring a process application monitor

An application must have been defined to an application controller before you set up the monitor.

To configure a process application monitor (in any of the three running modes: startup mode, long-running mode or both):

1. From the command line, enter `smit sysmirror`
2. In SMIT, select**Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Application Monitors › Configure Process Application Monitors › Add Process Application Monitor** and press Enter. A list of previously defined application controllers appears.
3. Select the application controller to which you want to add a process monitor.
4. In the **Add a Process Application Monitor** panel, fill in the field values as follows:

Add a Process Application Monitor	
Field	**Value**
Monitor Name	Enter the name of the application monitor. Each monitor can have a unique name that does not have to be the same name as the application controller name.
Monitor Mode	Select the mode in which the application monitor monitors the application: • **startup monitoring**. In this mode the application monitor checks that the application controller has successfully started *within* the specified stabilization interval. The monitor in this mode may run multiple times, as long as it is being run *within* the stabilization interval that you specify. If the monitor in this mode returns a zero code, this means that the application had started successfully. If a non-zero code is returned, this means that the application did not start within the stabilization interval. Select this mode if you are configuring an application monitor for an application that is included in a parent resource group (in addition to other monitors that you may need for dependent resource groups). • **long-running monitoring**. In this mode, the application monitor periodically checks that the application controller is running. The monitor is run multiple times based on the monitoring interval that you specify. If the monitor returns a zero code, it means that the application is running successfully. A non-zero return code indicates that the application has failed. The checking starts *after* the specified stabilization interval has passed. This mode is the default. • **both**. In this mode, the application monitor checks that within the stabilization interval the application controller has started successfully, *and* periodically monitors that the application controller is running after the stabilization interval has passed. If the same monitor is used in the "both" mode, PowerHA® SystemMirror® interprets the return codes differently, according to which type of monitoring is used (see the description of modes).
Processes to Monitor	Specify the process(es) to monitor. You can type more than one process name. Use spaces to separate the names. To be sure you are using correct process names, use the names as they appear from the **ps -el** command (not **ps -f**), as explained in Identifying correct process names.
Process Owner	Specify the user ID of the owner of the processes specified above, for example *root*. Note that the process owner must own all processes to be monitored.
Instance Count	Specify how many instances of the application to monitor. The default is **1** instance. The number of instances must exactly match the number of processes to monitor. If you put one instance, and another instance of the application starts, you will receive an application monitor error. This number must be more than **1** if you have specified more than one process to monitor (1 instance for each process).
Stabilization Interval	Specify the time (in seconds). PowerHA® SystemMirror® uses the stabilization period for the monitor in different ways, depending on which monitor mode is selected in this SMIT panel:

Field	Value
	• If you select the **startup monitoring** mode, the stabilization interval is the period *within* which PowerHA® SystemMirror® runs the monitor to check that the application has successfully started. When the specified time expires, PowerHA® SystemMirror® terminates the monitoring of the application startup and continues event processing. If the application fails to start within the stabilization interval, the resource group's acquisition fails on the node, and PowerHA® SystemMirror® launches resource group recovery actions to acquire a resource group on another node. The number of seconds you specify should be approximately equal to the period of time it takes for the application to start. This depends on the application you are using. • If you select the **long-running** mode for the monitor, the stabilization interval is the period during which PowerHA® SystemMirror® waits for the application to stabilize, before beginning to monitor that the application is running successfully. For instance, with a database application, you may wish to delay monitoring until after the start script and initial database search have been completed. You may need to experiment with this value to balance performance with reliability. • If you select **both** as a monitoring mode, the application monitor uses the stabilization interval to wait for the application to start successfully. It uses the same interval to wait until it starts checking periodically that the application is successfully running on the node. In most circumstances, this value should *not* be zero.
Restart Count	Specify the number of times to try restarting the application before taking any other actions. The default is **3** . If you are configuring a monitor that is going to be used only in the startup monitoring mode, restart count does not apply, and PowerHA® SystemMirror® ignores values entered in this field. Make sure you enter a Restart Method if your Restart Count is any non-zero value.
Restart Interval	Specify the interval (in seconds) that the application must remain stable before resetting the restart count. Do not set this to be shorter than *(Restart Count) x (Stabilization Interval)* . The default is 10% longer than that value. If the restart interval is too short, the restart count will be reset too soon and the desired fallover or notify action may not occur when it should. If you are configuring a monitor that is going to be used only in the startup monitoring mode, restart interval does not apply, and PowerHA® SystemMirror® ignores values entered in this field.
Action on Application Failure	Specify the action to be taken if the application cannot be restarted within the restart count. You can keep the default choice **notify**, which runs an event to inform the cluster of the failure, or select **fallover** , in which case PowerHA® SystemMirror® recovers the resource group containing the failed application on the cluster node with the next highest priority for that resource group. If you are configuring a monitor that is going to be used only in the startup monitoring mode, the action specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field.

Field	Value
	See Application monitoring prerequisites and considerations for more information.
Notify Method	(*Optional*) Define a notify method that will run when the application fails. This custom method runs during the restart process and during notify activity. If you are configuring a monitor that is going to be used only in the startup monitoring mode, the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field.
Cleanup Method	(*Optional*) Specify an application cleanup script to be called when a failed application is detected, before calling the restart method. The default is the application controller stop script defined when the application controller was set up (if you have only one application controller defined. If you have multiple application controllers, enter the stop script in this field that is used for the associated application controller). If you are configuring a monitor that is going to be used only in the startup monitoring mode, the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field. With application monitoring, since the application is already stopped when this script is called, the server stop script may fail.
Restart Method	(Required if **Restart Count** is not zero.) The default restart method is the application controller start script defined previously, if only one application controller was set up. This field is empty if multiple servers are defined. You can specify a different method here if desired. If you are configuring a monitor that is going to be used only in the startup monitoring mode, the method specified in this field does not apply, and PowerHA® SystemMirror® ignores values entered in this field.

5. Press Enter.
   SMIT checks the values for consistency and enters them into the PowerHA® SystemMirror® Configuration Database. When the resource group is brought online, the application monitor in the long-running mode starts (if it is defined). Note that the application monitor in the startup monitoring mode starts before the resource group is brought online.

   When you synchronize the cluster, verification ensures that all methods you have specified exist and are executable on all nodes.

Configuring a custom application monitor

You can configure multiple application monitors and associate them with one or more application controllers. By supporting multiple monitors per application, PowerHA® SystemMirror® can support more complex configurations.

Custom application monitoring allows you to write a monitor method to test for conditions other than process termination. For example, if an application sometimes becomes unresponsive while still running, a custom monitor method could test the application at defined intervals and report when the application's response is too slow. Also, some applications (`shell` scripts, for example) cannot be registered with RSCT, so process monitoring cannot be configured for them. A custom application monitor method can monitor these types of applications.

For instructions on defining a process application monitor, which requires no custom monitor method, refer to Application monitoring prerequisites and considerations.

Defining a monitor method

Unlike process monitoring, custom application monitoring requires you to provide a script to test the health of the application. You must also decide on a suitable polling interval.

When devising your custom monitor method, keep the following points in mind:

- The monitor method must be an executable program (it can be a `shell` script) that tests the application and exits, returning an integer value that indicates the application's status. The return value must be zero if the application is healthy, and must be a non zero value if the application has failed.
- The method can log messages by printing them to the standard output **stdout** file. For long running monitors, the output is stored in the `/var/hacmp/log/clappmond.`*`application monitor name.resource group name.monitor`*`.log` file. For startup monitors, this output is stored in the `/var/hacmp/log/clappmond.`*`application controller name.resource group name.monitor`*`.log` file. In PowerHA® SystemMirror® Version 7.1.1, or earlier, there is a single log file that is overwritten each time the application monitor is restarted. In PowerHA® SystemMirror® Version 7.1.2, or later, a new log file is created each time the application monitor is restarted.
- Since the monitor method is set up to terminate if it does not return within the specified polling interval, do not make the method overly complicated.

Steps for configuring a custom application monitor

This topic explains the steps for configuring a custom application monitor.

To set up a custom application monitoring method, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource › Configure User Applications (Scripts and Monitors) › Application Monitors › Configure Custom Application Monitors › Add a Custom Application Monitor** and press Enter.
   A list of defined application controllers is displayed.
3. Select the application controller for which you want to add a monitoring method.
4. In the **Add Custom Application Monitor** panel, fill in field values as follows. The **Monitor Method** and **Monitor Interval** fields require you to supply your own scripts and specify your own preference for the polling interval:

Add Custom Application Monitor fields	
Field	**Value**
Application Controller Name	Select the application controller from the picklist.
Monitor Mode	Select the mode in which the application monitor monitors the application: • **Startup monitoring**. In this mode the application monitor checks that the application controller has successfully started *within* the specified stabilization period. If you are configuring a monitor for an application that is included in a parent resource group, select this mode (in addition to other monitors that you may need for dependent resource groups). • **Long-running monitoring**. In this mode, the application monitor periodically checks that the application controller is running. The checking starts *after* the specified stabilization interval has passed. This is the default. • **Both**. In this mode, the application monitor checks that within the stabilization interval the application controller has started successfully, and periodically monitors that the application controller is running after the stabilization interval have passed. If the **Monitor Mode** field is set to the `Long-running monitoring`mode, the application monitor does not run when the cluster services are started. The application controller runs the start script each time to start your application irrespective of whether the application is already running. When the resource group that contains the

Field	Value
	application is online, the application monitor is used to monitor the application. You can also specify the startup application monitor to run when you start the cluster services. When the startup application monitor runs, it determines whether the application is already active. The application controller does not run the start script when the application monitor indicates that the application is up and running. You can use the same application monitor for both the startup monitoring mode and long-running monitoring mode, by selectingBothas the value for the **Monitor Mode** field. The mode that you select for your application monitor determines whether the application monitor is run when the cluster services are started or when the resource group is online. You must carefully select these modes when you write your monitoring scripts so that proper checks are performed and actions are taken when the application monitor runs. For example, when the cluster services are started for the first time, the other resources on which the application relies on might not be fully active. In this case, you might consider specifying the startup application monitor to run when you start the cluster services.
Monitor Method	Enter a script or executable for custom monitoring of the health of the specified application. Do not leave this field blank. Note that the method must return a zero value if the application is healthy and a non-zero value if a problem is detected. The method can log messages by printing them to the standard output **stdout** file. For long running monitors, the output is stored in the /var/hacmp/log/clappmond.*application monitor name.resource group name.monitor*.log file. For startup monitors, this output is stored in the /var/hacmp/log/clappmond.*application controller name.resource group name.monitor*.log file. In PowerHA® SystemMirror® Version 7.1.1, or earlier, there is a single log file that is overwritten each time the application monitor is restarted. In PowerHA® SystemMirror® Version 7.1.2, or later, a new log file is created each time the application monitor is restarted.
Monitor Interval	Enter the polling interval (in seconds) for checking the health of the application. If the monitor does not respond within this interval, it is considered hung.
Monitor Retry Count	Specifies the number of times PowerHA® SystemMirror® tries to restart the custom application monitor before performing any other actions. The default value is 0. This field is related to the **Restart Count.**
Hung Monitor Signal	The signal the system should send to stop the **Monitor Method** script if it does not return within the time specified for the **Monitor Interval**. The default is SIGKILL(9).
Stabilization Interval	Specify the time (in seconds). PowerHA® SystemMirror® uses the stabilization period for the monitor in different ways, depending on which monitor mode is selected in this SMIT panel: • If you select the **startup monitoring** mode, the stabilization interval is the period within which PowerHA® SystemMirror® monitors that the application has successfully started. When the specified time expires, PowerHA® SystemMirror® terminates the monitoring of the application startup, and

Field	Value
	continues event processing. If the application fails to start within the stabilization interval, the resource group's acquisition fails on the node, and PowerHA® SystemMirror® launches resource group recovery actions to acquire a resource group on another node. The number of seconds you specify should be approximately equal to the period of time it takes for the application to start. This depends on the application you are using. • If you select the **long-running** mode for the monitor, the stabilization interval is the period during which PowerHA® SystemMirror® waits for the application to stabilize, before beginning to monitor that the application is running successfully. For instance, with a database application, you may wish to delay monitoring until after the start script and initial database search have been completed. You may need to experiment with this value to balance performance with reliability. • If you select **both** as a monitoring mode, the application monitor uses the stabilization interval to wait for the application to start successfully. It uses the same interval to wait until it starts checking periodically that the application is successfully running on the node. In most circumstances, this value should *not* be zero.
Restart Count	Specify the number of times to try restarting the application before taking any other actions. The default is **3** .
Restart Interval	Specify the interval (in seconds) that the application must remain stable before resetting the restart count. Do not set this to be shorter than (Restart Count) x (Stabilization Interval + Monitor Interval). The default is 10% longer than that value. If the restart interval is too short, the restart count will be reset too soon and the desired failure response action may not occur when it should.
Action on Application Failure	Specify the action to be taken if the application cannot be restarted within the restart count. You can keep the default choice **notify**, which runs an event to inform the cluster of the failure, or select **fallover**, in which case the resource group containing the failed application moves over to the cluster node with the next highest priority for that resource group.
Notify Method	(*Optional*) The full pathname of a user defined method to perform notification when a monitored application fails. This method will execute each time an application is restarted, fails completely, or falls over to the next node in the cluster. Configuring this method is strongly recommended.
Cleanup Method	(*Optional*) Specify an application cleanup script to be invoked when a failed application is detected, before invoking the restart method. The default is the application controller stop script defined when the application controller was set up. With application monitoring, since the application may be already stopped when this script is called, the server stop script may fail.
Restart Method	(Required if **Restart Count** is not zero.) The default restart method is the application controller start script defined previously, when the application controller was set up. You can specify a different method here if desired.

5. Press Enter.

Result
SMIT checks the values for consistency and enters them into the PowerHA® SystemMirror® Configuration Database. When the resource group comes online, the application monitor in the long-running mode starts. The application startup monitor starts before the resource group is brought online.

When you synchronize the cluster, verification ensures that all methods you have specified exist and are executable on all nodes.

Suspending, changing, and removing application monitors

You can temporarily suspend an application monitor in order to perform cluster maintenance. You should not change the application monitor configuration while it is in a suspended state.

If you have multiple application monitors configured, and choose to temporarily suspend an application monitor, all monitors configured for a specified server are suspended.

Configuring tape drives as PowerHA® SystemMirror® resources

The PowerHA® SystemMirror® SMIT panels enable certain actions for configuring tape drives.

These actions include:

- Add tape drives as PowerHA® SystemMirror® resources
  - Specify synchronous or asynchronous tape operations
  - Specify appropriate error recovery procedures
- Change or show tape drive resources
- Remove tape drive resources
- Add tape drives to PowerHA® SystemMirror® resource groups
- Remove tape drives from PowerHA® SystemMirror® resource groups.

Adding a tape resource

This topic describes how to add a tape drive as a cluster resource.

To add a tape drive:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Tape Resources › Add a Tape Resource** and press Enter.
3. Enter the field values as follows:

Add a Tape Resource fields	
Field	**Value**
Tape Resource Name	The symbolic name for the tape resource. This is a required field, and must be unique within the cluster. The name can have up to 64 alphanumeric characters and underscores.
Description	Description of the tape resource.
Tape Device Name	The name of the special file for the tape drive, for example, **/dev/rmt0**. This is a required field.
Start Script	Enter the full pathname of the script called by the cluster event scripts to start the application controller. You can use a maximum of 256 characters. This script must be in the same location on each cluster node that can start the server. The contents of the script, however, can be different.

Field	Value
	Passing arguments to this script is not permitted.
Start Processing Synchronous?	If yes, then tape start processing is synchronous. If no, it is asynchronous. The default is synchronous operation.
Stop Script	Enter the full pathname of the script called by the cluster event scripts to stop the server. You can use a maximum of 256 characters. This script must be in the same location on each cluster node that can start the server. The contents of the script, however, can be different. Passing arguments to this script is not permitted.
Stop Processing Synchronous?	If yes, then tape stop processing is synchronous. If no, it is asynchronous. The default is synchronous operation.

Result
Sample scripts are available in the **/usr/es/sbin/cluster/samples/tape** directory. The sample scripts rewind the tape drive explicitly.

To change or show the current configuration of a tape drive resource, see Reconfiguring tape drive resources.

Adding a tape resource to a resource group

This topic describes how to add a tape drive resource to a resource group.

To add a tape drive resource:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group** and press Enter.
   SMIT displays the list of resource groups.
3. Select the resource group to which you want to add the tape resource.
   SMIT displays the **Change/Show all Resources/Attributes for a <selected type of> Resource Group** panel.
4. Enter the field value for tape resource.

Result
Type in the resource name or press F4 to display a picklist of defined tape resources. Select the desired resource. If there are no tape resources defined, SMIT displays an error message.

Verifying and synchronizing tape drive configuration

After adding a resource to a resource group, verify that the configuration is correct and then synchronize shared tape resources to all nodes in the cluster.

Verification ensures the following:

- Validity of the specified tape special file (is it a tape drive?)
- Accessibility of the tape drive (does a device on the specified SCSI LUN exist?)
- Consistency of the configuration (does the device have the same LUN on the nodes sharing the tape drive?)
- Validity of the user defined start and stop scripts (do the scripts exist and are they executable?)

Dynamic reconfiguration of tape resources

When a tape drive is added to a resource group, or when a new resource group is created with tape resources, DARE will reserve the tape and invoke the user-provided tape start script.

When a tape drive is removed from a resource group, or when a resource group with tape resources is removed, DARE invokes the user-provided tape stop script and releases the tape drive.

Adding a user-defined resource to a resource group

After you create a user-defined resource based on a user-defined resource type, you can add the resource to a resource group.

To add a user-defined resource to a resource group, complete these steps:

1. From the command line, enter `smit sysmirror`
2. Select the resource group to which you want to add the user-defined resource.
3. SMIT displays the **Change/Show all Resources and Attributes for a <selected type of> Resource Group** panel.
4. Enter the field value for the user-defined resource.
5. Enter the resource name or press F4 to display a picklist of user-defined resources and select the resource that you want to add.
   If there are no user-defined resources configured, SMIT displays an error message. Refer to *Configuring a user-defined resource type* for more information about this type of resource.

Dynamic reconfiguration of a user-defined resource type and resources

When a user-defined resource type is added or when a user-defined resource is added to a resource group, or when a new resource group is created a with user-defined resource, DARE will start the user-defined resource according to the order specified in the resource type.

When a user-defined resource type is removed or when a user-defined resource type is removed from a resource group, or when a resource group with user-defined resources is removed, DARE invokes the user-provided stop script and releases the user-defined resource.

Customizing resource recovery

PowerHA® SystemMirror® monitors system resources and initiates recovery when a failure is detected. Recovery involves moving a set of resources (grouped together in a resource group) to another node. PowerHA® SystemMirror® uses *selective fallover* function when it can. Selective fallover enables PowerHA® SystemMirror® to recover only those resource groups that are affected by the failure of a specific resource.

PowerHA® SystemMirror® uses selective fallover in the following cases.

- Loss of a volume group
- Local network failure
- Resource group acquisition failure
- Application failure
- User-defined resource failure

You can customize recovery for two types of resources where PowerHA® SystemMirror® uses selective fallover.

- *Service IP labels.* By default, for a local network failure, PowerHA® SystemMirror® responds by scanning the configuration for any service labels on that network and moving only the resource group containing the failed service IP label to another available node.

  > **Note:** You cannot customize recovery for service IP labels for the secondary instance of a replicated resource group.

- *Volume groups.* For volume groups where the recovery is triggered by a loss of quorum for a volume group, PowerHA® SystemMirror® moves the resource group to a takeover node.

  > **Note:** Customizing volume group recovery (disabling selective fallover) in a cluster with this type of resource in a replicated resource group applies to both the primary and the secondary instances of the resource group.

However, selective fallover may not be the behavior you want when one of these resources fails. After upgrading from a previous release, if you have custom pre-event and post-events to handle these situations, these may act in unexpected ways when combined with the selective fallover behavior. PowerHA® SystemMirror® includes the **Customize Resource Recovery** option for changing the behavior of the selective fallover action for these resources. You can select to have the fallover occur, or to simply receive a notification.

Take the following steps to customize resource recovery for service label and volume group resources (especially if you have your own custom pre-event and post-event scripts).

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Resource › Customize Resource Recovery** and press Enter.
3. Select the resource to customize from the list.
4. Enter field values as follows.

Customize Resource Recovery fields	
Field	**Value**
Name	The resource you selected. Resources with duplicate names will not be listed because this function only supports resources with unique names. XD resources (GMD, PPRC, ERCMF, SVCPPRC and GMVG) are not supported.
Action on Resource failure	Select either **fallover** or **notify**. **Fallover** is the default.

5. **Fallover** initiates an **rg_move** event to move the affected resource group to another node.
6. **Notify** causes a **server_down** event that calls out the specific failed resource but takes no recovery action.

   > In the **Notify Method** field, enter the full pathname of your own method to perform notification when this resource fails. This method will be called by the **server_down** event. Passing arguments to this method is not permitted.

7. Press Enter to apply the customized resource recovery action.
8. If you use the **Notify Method**, make sure it is on all nodes in the resource group nodelist.
9. Verify and synchronize the cluster.

Fallover option and resource group availability

Be aware that if you select the **fallover** option of customized resource recovery - which could cause a resource group to migrate from its original node - the possibility exists that while the highest priority node is up, the resource group remains down.

This situation occurs when an **rg_move** event moves a resource group from its highest priority node to a lesser priority node, and then you stop the cluster services on the lower priority node with an option to bring the resource groups offline. Unless you bring the resource group up manually, it remains in an inactive state.

For more information on resource group availability, see Selective fallover for handling resource groups.

Testing customized resource recovery

Once you have configured the options and synchronized your cluster successfully, you are ready to test that the new options provide the desired behavior.

Testing the fallover action on resource failure

This is the default behavior. When a resource failure occurs (local_network_down or volume group quorum loss), an **rg_move** event will be run for the affected resource group. You can test this behavior by inducing a local_network_down (fail all interfaces on that network on a single node) or by inducing the LVM_SA_QUORCLOSE error (power off a disk while writes are occurring such that quorum is lost for that volume group).

Testing the notify action on resource failure

Induce the same failures mentioned previously, selecting **notify** but no **Notify Method**. Instead of an **rg_move** event, a **server_down** event should run. Check the output in **hacmp.out**.

Testing the notify method

Configure a resource and resource group and specify the **notify** option for that resource, with a **Notify Method**. Induce one of the failures to trigger the **server_down** event. The **server_down** event will call the **Notify Method** and any output from that method will be logged in **hacmp.out**.

Configuring PowerHA® SystemMirror® resource groups

Use the following SMIT menu path, **Configure Applications and Resources › Resource Groups** to configure resource groups in a cluster.

The Configure Resource Groups menu path can be used to add, change, show or remove a resource group, as well as to configure resource group run-time policies.

- **Configure Resource Group Run-Time Policies**:
  Use this set of menus to manage the following:
  - Dependencies between resource groups
  - Workload Manager parameters
  - Resource group processing order
  - Delayed Fallback Timer
  - Settling Time
- **Manage resource group configuration**:
  From the Configure Resource Groups menu you can complete the following tasks:
  - Add a Resource group
  - Change or show resource group nodes and policies
  - Change or show the resources included in a resource group
  - Remove a resource group
  - Show all resources by node or resource group

Configuring resource groups

Use these topics to find out how to configure resource groups with different combinations of startup, fallover and fallback policies, and run-time policies.

You can add resource groups with different startup, fallover and fallback policies. Prior to configuring resource groups, you should read the planning information.

> **Note:** You can use SMIT to configure and manage the cluster and view interactive cluster status.

Limitations and prerequisites for configuring resource groups

When configuring a resource group, certain limitations and conditions apply.

These conditions include:

- By default, PowerHA® SystemMirror® processes resource groups in parallel. You may include a resource group in a list of resource groups that are processed serially. However, if you do not include a resource group in a serially-processed list, but specify a settling time or a delayed fallback timer for a resource group, the acquisition of this resource group is delayed. For complete information, see Configuring processing order for resource groups.
- Clocks on all nodes must be synchronized for the settings for fallover and fallback of a resource group to work as expected.
- To view the information about resource groups and for troubleshooting purposes, use the **clRGinfo** command. Also, for troubleshooting purposes, you can use the **Show All Resources by Node or Resource Group** SMIT option.
- If you want to change the startup policy for a resource group from **Online On All Available Nodes** (OAAN) to some other policy, you must first take the resource group offline from all nodes where the resource group is active. Alternately, you can stop cluster services on all nodes. Change the startup policy as needed, then verify and synchronize your cluster configuration for the changes to take effect.

Configuring resource groups by using SMIT

The System Management Interface Tool (SMIT) fields that you use to configure resource groups depend on whether you configured sites for the cluster.

To configure a resource group by using SMIT, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Add a Resource Group**, and press Enter.
3. Enter the field values as follows:

Configuring resource group fields

Field	Description
Resource Group Name	The name of the resource group must be unique within the cluster and distinct from the volume group and service IP label. You cannot use more than 64 alphanumeric characters and underscores. You cannot start the resource group name with a number. Duplicate entries and reserved words are not allowed.
Inter-site Management Policy	This field is only available if you configured sites. The default setting is Ignore. You can select the following options for this field: **Ignore** The resource group does not have ONLINE SECONDARY instances. Use this option if you use cross-site LVM mirroring.

Field	Description
	Prefer Primary Site The primary instance of the resource group is brought ONLINE on the primary site at startup, the secondary instance is started on the other site. The primary instance falls back when the primary site rejoins the cluster. **Online on Either Site** During startup, the primary instance of the resource group is brought ONLINE on the first node that meets the node policy criteria (either site). The secondary instance is started on the other site. The primary instance does not fall back when the original site rejoins the cluster. **Online on Both Sites** During startup, the resource group (node policy must be defined as Online on All Available Nodes) is brought ONLINE on both sites. There is no fallover policy or fallback policy. If no nodes or conditions exist that bring or keep the resource group online on the site, the resource group moves to another site. The site that owns the active resource group is called the primary site.
Participating Nodes (Default Node Priority)	Enter the names of the nodes that can own or take over this resource group. Enter the node with the highest priority first, followed by the other nodes in priority order. Leave a space between node names. If you configured sites for the cluster, this field is not available.
Participating Nodes from Primary Site	This field is only available if you configured sites. Select the nodes that belong to the primary site for the resource group. Resource group policies that refer to the preferred primary site belong to the nodes in this list. The nodes in this list are from the same site in the cluster. If the Inter-Site Management Policy field is set to **Ignore**, there is no practical distinction between the primary site and secondary site of the resource group.
Participating Nodes from Secondary Site	This field is only available if you configured sites. Select the nodes that belong to the secondary site for the resource group. Resource group policies that refer to the preferred secondary site belong to the nodes in this list. The nodes in this list are from the same site in the cluster. If the Inter-Site Management Policy field is set to **Ignore**, there is no practical distinction between the primary site and secondary site of the resource group.
Startup Policy	Select the following option that defines the startup policy of the resource group: **Online On Home Node Only** The resource group is brought online only on its home node, the highest priority node, during the resource group startup. This function requires the highest priority node to be available. **Online On First Available Node**

Field	Description
	The resource group is activated on the first participating node that becomes available. If you configured the settling time for resource groups, it is only used for the startup policy of this resource group. **Online Using Node Distribution Policy** The resource group is brought online according to the node-based distribution policy. This policy allows only one resource group to be brought online on a node during startup. **Online On All Available Nodes** The resource group is brought online on all nodes. If you select this option, you must verify that the resources in this group can be brought online on multiple nodes simultaneously.
Fallover Policy	Select the following option that defines the fallover policy of the resource group: **Fallover To Next Priority Node In The List** The resource group that is online on only one node at a time follows the default node priority order that is specified in the resource group's node list. **Fallover Using Dynamic Node Priority** You can use a predefined dynamic-node priority policy or one of the two user-defined policies. **Bring Offline (On Error Node Only)** The resource group is brought offline on a node during an error condition. Select this option if you want to ensure that when a particular node fails, the resource group goes offline only on the specified node but remains online on other nodes. Selecting this option as the fallover preference when the startup preference is not set to **Online On All Available Nodes**, might allow resources to become unavailable during error conditions. In this case, PowerHA® SystemMirror® issues an error message.
Fallback Policy	Select the following option that defines the fallback policy of the resource group: **Fallback To Higher Priority Node In The List** The resource group falls back when a higher priority node joins the cluster. Select this option if you configured the delayed fallback timer setting. If you did not configure the delayed fallback timer setting, the resource group falls back immediately when a higher priority node joins the cluster. **Never Fallback** The resource group does not fall back when a higher priority node joins the cluster.

4. Press Enter to add the resource group information to the PowerHA® SystemMirror® configuration database.
   During the configuration of resource groups, if you selected an option that prevents high availability of a resource group, PowerHA® SystemMirror® issues a warning message. Thus, PowerHA® SystemMirror® prevents invalid or incompatible resource group configurations.

Dynamic node priority policies

The default node priority policy is the order in the participating nodelist. However, can have a takeover node selected dynamically, according to the value of a specific system property at the time of failure.

Dynamic Node Priority entails the selection of a node that acquires the resource group based on values of system attributes calculated at run time. These values are obtained by querying the RMC subsystem. In particular, one of the following attributes can be chosen for Dynamic Node Priority:

- `cl_highest_free_mem` - select the node with the highest percentage of free memory
- `cl_highest_idle_cpu` - select the node with the most available processor time
- `cl_lowest_disk_busy` - select the disk that is least busy

The PowerHA® SystemMirror® cluster manager queries the RMC subsystem every three minutes to obtain the current value of attributes on each node and distributes them cluster wide. The interval at which the queries of the RMC subsystem are performed, 3 minutes, is not user configurable. During a fallover event of a resource group with Dynamic Node Priority configured, the most recently collected values are used in the determination of the best node to acquire the resource group.

Collected values		
PowerHA® SystemMirror®	**RMC Resource Manager**	**Attribute**
cl_highest_free_mem	IBM®.Host	PgSpFree
cl_highest_idle_cpu	IBM®.Host	PctTotalTimeIdle
cl_lowest_disk_busy	IBM®.PhysicalVolume	PvPctBusy

The RMC resource monitor on a node may be queried to obtain the current values of these attributes:

```
lsrsrc -Ad IBM.Host
lsrsrc -Ad IBM.PhysicalVolume
```

Note: If you have defined a resource group over multiple sites (using the PowerHA® SystemMirror® Enterprise Edition software) and a dynamic node priority policy is configured for the group, you will receive this warning when **verification** runs:

```
"Warning:
Dynamic Node Priority is configured in a resource group
with nodes in more than one site. The priority calculation may
fail due to slow communication, in which case the default node
priority will be used instead."
```

You can choose Dynamic node priority based on the user defined property by selecting one of the following attributes:

```
cl_highest_udscript_rc
cl_lowest_nonzero_udscript_rc
```

When you select one of the these criteria, you must also provide values for the **DNP script path** and **DNP timeout** attributes for a resource group. When the **DNP script path** attribute is specified, the given script is invoked on all nodes and return values are collected from all nodes. The failover node decision is made by using these values and the specified criteria. If you choose the **cl_highest_udscript_rc** attribute, collected values are sorted and the node which returned the highest value is selected as a candidate node to failover. Similarly, if you choose the **cl_lowest_nonzero_udscript_rc** attribute, collected values are sorted and the node which returned lowest nonzero positive value is selected as a candidate node to failover. If the return value of the script from all nodes are same or zero, the default node priority will be considered. PowerHA® verifies the script existence and the execution permissions during verification.

When you select a timeout value, ensure that it is within the time period for running and completing a script. If you do not specify a timeout value, a default value that is equal to the **config_too_long** time is specified. If you provide a timeout value that is greater than the default allowed timeout value, PowerHA® sets the value to the default timeout value and creates the following warning message:

```
warning: The parameter "SDNP_SCRIPT_TIMEOUT" value specified is greater than the Maximum allowed timeout value. will use " 360."
```

> **Note:** In the preceding warning message, 360 seconds is the **config_too_long** attribute time that is currently set in the cluster.

Do the following to specify these values:

1. From the command line, enter `smit sysmirror`.
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group**
3. Select the resource group with the failover policy of `Failover Using Dynamic Node Priority`.

The following reminders apply when writing DNP scripts:

- Script return values are considered for DNP calculation.
- A return value of zero on a node indicates that the node is not a candidate node to failover.
- A return value greater than zero on a node indicates that the node can be one of the candidate node.

Configuring resource group run-time policies

Review the following for information on resource group runtime policies.

Resource Group run-time policies include:

- Dependencies between resource groups. See Configuring dependencies between resource groups
- Resource group processing order See Configuring processing order for resource groups
- Workload Manager. See Configuring Workload Manager
- Settling Time for resource groups. See Configuring a settling time for resource groups
- Delayed Fallback Timer for resource groups. See Configuring delayed fallback timers in SMIT
- Node distribution policy. See Using the node distribution startup policy

Configuring dependencies between resource groups

You can set up more complex clusters by specifying dependencies between resource groups.

Business configurations that use multi-tiered applications can utilize parent/child dependent resource groups. For example, the back end database must be online before the application controller. In this case, if the database goes down and is moved to a different node, the resource group containing the application controller must be brought down and back up on any node in the cluster.

Business configurations that require different applications to run on the same node, or on different nodes can use location dependency runtime policies. See Examples of location dependency and resource group behavior for more information.

The dependencies that you configure are:

- Explicitly specified using the SMIT interface
- Established cluster-wide, not just on the local node
- Guaranteed to occur in the cluster, that is, they are not affected by the current cluster conditions.

You can configure six types of dependencies between resource groups:

- Parent/child dependency
- Start After dependency
- Stop After dependency
- Online On Same Node Location dependency

- Online On Different Nodes Location dependency
- Online On Same Site Location dependency

Considerations for dependencies between resource groups

Look here for additional considerations you may need to keep in mind when configuring resource group dependencies. These include interaction with sites, use of pre-and post-event scripts, and information about the **clRGinfo** command.

- To obtain more granular control over the resource group movements, use the **clRGinfo -a** command to view what resource groups are going to be moved during the current cluster event. Also, use the output in the **hacmp.out** file. For more information, see Using resource groups information commands.
- Dependencies between resource groups offer a predictable and reliable way of building clusters with multi-tiered applications. However, **node_up** processing in clusters with dependencies could take more time than in the clusters where the processing of resource groups upon **node_up** is done in parallel. A resource group that is dependent on other resource groups cannot be started until others have been started first. The **config_too_long** warning timer for **node_up** should be adjusted large enough to allow for this.
- During verification, PowerHA® SystemMirror® verifies that your configuration is valid and that application monitoring is configured.
- You can configure resource group dependencies in PowerHA® SystemMirror® Enterprise Edition clusters that use replicated resources for disaster recovery. However, you cannot have the combination of any non-concurrent startup policy and concurrent (Online on Both Sites) inter-site management policy. You can have a concurrent startup policy combined with a non-concurrent inter-site management policy.

The high-level steps required to specify resource group dependencies are described in the following sections.

Steps to configure dependencies between resource groups

This section provides a high-level outline of the steps required to configure a dependency between resource groups.

These steps include:

1. For each application that is going to be included in dependent resource groups, configure application controllers and application monitors.
2. Create resource groups and include application controllers as resources. For instructions, see Configuring resource groups and Adding resources and attributes to resource groups using the extended path.
3. Specify a dependency between resource groups. For instructions, see Configuring resource groups with dependencies.
4. Use the SMIT **Verify and Synchronize Cluster Configuration** option to guarantee the desired configuration is feasible given the dependencies specified, and ensure that all nodes in the cluster have the same view of the configuration.

To ensure that the applications in the dependent resource groups start successfully, you should configure multiple application monitors.

In general, we recommend that you configure the following monitors:

- A monitor that will check the running process for an application in the child resource group, and a monitor that will check the running process for an application in the parent resource group.
- A monitor that will check the running process for an application in the source resource group in a startafter dependency, and a monitor that will check the running process for an application in the target resource group in a startafter dependency.

For a parent resource group, it is also advisable to configure a monitor in a *startup monitoring mode* to watch the application startup. This ensures that after the parent resource group is acquired, the child resource group(s) also can be acquired successfully. Similarly, for a target resource group in a startafter dependency, it also is advisable to configure a monitor in a *startup monitoring mode* to watch the application startup. This ensures that after the target resource group is acquired, the source resource group(s) also can be acquired successfully.

For information on monitor modes that you can specify (long-running mode, startup monitoring mode, and both), see Monitor modes.

For instructions on configuring application monitoring, see Configuring multiple application monitors.

Configuring resource groups with dependencies

You can configure six types of dependencies between resource groups.

These dependencies include:

- Parent/child dependency
- Start After dependency
- Stop After dependency
- Online On Same Node Location Dependency
- Online On Different Nodes Location Dependency
- Online On Same Site Location Dependency.

The following limitations apply to configurations that combine dependencies:

- Only one resource group can belong to a Same Node Dependency and a Different Node Dependency at the same time
- If a resource group belongs to *both* a Same Node Dependency and a Different Node Dependency, all nodes in the Same Node Dependency set have the same Priority as the shared resource group.
- Only resource groups with the same Priority within a Different Node Dependency can participate in a Same Site Dependency.

Configuring a parent and child dependency between resource groups

In this type of dependency, the parent resource group must be online on any node in the cluster before a child (dependent) resource group can be activated on a node.

These are the guidelines and limitations:

- A resource group can serve as both a parent and a child resource group, depending on which end of a given dependency link it is placed.
- You can specify three levels of dependencies for resource groups.
- You cannot specify circular dependencies between resource groups.
- A child resource group cannot be acquired on a node until its parent resource group is fully functional. If the parent node does not become fully functional, the child resource group goes into an ERROR state. If you notice that a resource group is in this state, you may need to troubleshoot which resources might need to be brought online manually to resolve the resource group dependency.
- When a resource group in a parent role falls over from one node to another, the resource groups that depend on it are stopped before the parent resource group falls over, and restarted again once the parent resource group is stable again.
- For information on dynamic reconfiguration (DARE), see the section Reconfiguring resources in clusters with dependent resource groups.

To configure a parent/child dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between resource groups › Configure Parent/Child Dependency › Add Parent/Child Dependency between resource groups** and press Enter.
3. Fill in the fields as follows:

Add Parent/Child Dependency between resource groups fields	
Field	**Value**
Parent Resource Group	Select the parent resource group from the list. The parent resource group provides services upon which another resource group depends. During resource group acquisition, PowerHA® SystemMirror® acquires the parent resource group on a node *before* the child resource group is acquired.
Child Resource Group	Select the child resource group from the list and press Enter. PowerHA® SystemMirror® prevents you from specifying circular dependencies. The child resource group depends on services another resource group provides. During resource group acquisition, PowerHA® SystemMirror® acquires the parent resource group on a node *before* the child resource group is acquired. During release, PowerHA® SystemMirror® releases the child resource group *before* the parent resource group is released.

4. Press Enter and verify the cluster.

Configuring a start after dependency between resource groups

In this type of dependency, the source resource group on a node can be bought online only when the target resource groups are online. There is no dependency when releasing resource groups and the groups are released in parallel.

These are the guidelines and limitations:

- A resource group can serve as both a target and a source resource group, depending on which end of a given dependency link it is placed. However, a resource group cannot serve as both a target resource group and a source resource group within the same dependency.
- You can specify three levels of dependencies for resource groups.
- You cannot specify circular dependencies between resource groups.
- This dependency applies only at the time of resource group acquisition. There is no dependency between these resource groups during resource group release.
- A source resource group cannot be acquired on a node until the target resource groups are fully functional. If the target resource groups do not become fully functional, the source resource group goes into an OFFLINE DUE TO TARGET OFFLINE state. If you notice that a resource group is in this state, you might need to troubleshoot which resources might need to be brought online manually to resolve the resource group dependency.
- When a resource group in a target role falls over from one node to another, there will be no effect on the resource groups that depend on it.
- After the source resource group is online, any operation (bring offline, move resource group) on the target resource groups will not affect the source resource group.
- A manual resource group move or bring resource group online on the source resource group is not allowed if the target resource groups are offline.
- For information on dynamic reconfiguration (DARE), see the section Reconfiguring resources in clusters with dependent resource groups.

To configure a Start After dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Start After Dependency › Add Start After Dependency Between Resource Groups** and press Enter.
3. Fill in the fields as follows:

Field name	Description
Source Resource Group	Select a source resource group from the list and press Enter. PowerHA® SystemMirror® prevents you from specifying circular dependencies. The source resource group depends on services that another resource group provides. During resource group acquisition, PowerHA® SystemMirror® acquires the target resource groups on a node *before* the source resource group is acquired.
Target Resource Groups	Select the target resource groups from the list and press Enter. PowerHA® SystemMirror® prevents you from specifying circular dependencies. The target resource groups provide services, which another resource group depends on. During resource group acquisition, PowerHA® SystemMirror® acquires the target resource groups on a node *before* the source resource group is acquired. There is no dependency between the source resource group and the target resource groups during release.

4. Press Enter and verify the cluster.

Configuring a stop after dependency between resource groups

In this type of dependency, the source resource group on a node can be bought offline only when the target resource groups are offline. There is no dependency when acquiring resource groups and the groups are acquired in parallel.

These are the guidelines and limitations:

- A resource group can serve as both a target and a source resource group, depending on which end of a given dependency link it is placed. However, a resource group cannot serve as both a target resource group and a source resource group within the same dependency.
- You can specify three levels of dependencies for resource groups.
- You cannot specify circular dependencies between resource groups.
- This dependency applies only at the time of resource group release. There is no dependency between these resource groups during resource group acquisition.
- A source resource group cannot be released on a node until the target resource groups are offline.
- When a resource group in a source role falls over from one node to another, first the target resource groups are released and then the source resource group will be released. After that, both resource groups will be acquired in parallel, assuming that there is no start after or parent/child dependency between these resource groups.
- A manual resource group move or bring resource group offline on the source resource group is not allowed if the target resource groups are online.
- For information on dynamic reconfiguration (DARE), see the section Reconfiguring resources in clusters with dependent resource groups.

To configure a Stop After dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Stop After Dependency › Add Stop After Dependency Between Resource Groups** and press Enter.
3. Complete the fields as follows:

Add Stop After Dependency Between Resource Groups fields	
Field	**Value**
Source Resource Group	Select a source resource group from the list and press Enter. PowerHA® SystemMirror® prevents you from specifying circular

Field	Value
	dependencies. The source resource group will be stopped only after the target resource groups are offline. During the resource group release process, PowerHA® SystemMirror® releases the target resource groups on a node *before* releasing the source resource group. There is no dependency between source and target resource groups during acquisition.
Target Resource Groups	Select the target resource groups from the list and press Enter. PowerHA® SystemMirror® prevents you from specifying circular dependencies. The target resource groups provide services on which another resource group depends on. During the resource group release process, PowerHA® SystemMirror® releases the target resource groups on a node before releasing the source resource group. There is no dependency between the source resource group and the target resource groups during acquisition.

4. Press Enter and verify the cluster.

Configuring Online On Same Node Dependency for resource groups

When you configure two or more resource groups to establish a location dependency between them, they belong to a set for that particular dependency. This topic discusses the Online On Same Node Dependency for resource groups.

The following rules and restrictions apply to the Online On Same Node Dependency set of resource groups:

- All resource groups configured as part of a given Same Node Dependency set must have the same nodelist (the same nodes in the same order).
- All non-concurrent resource groups in the Same Node Dependency set must have the same Startup/ Fallover/Fallback Policies.
  - Online Using Node Distribution Policy is not allowed for Startup.
  - If a Dynamic Node Priority Policy is chosen as Fallover Policy, then all resource groups in the set must have the same policy.
  - If one resource group in the set has a fallback timer, it applies to the set.
  - All resource groups in the set must have the same setting for fallback timers.
- Both concurrent and non-concurrent resource groups are allowed.
- You can have more than one Same Node Dependency set in the cluster.
- All resource groups in the Same Node Dependency set that are active (ONLINE) are required to be ONLINE on the same node, even though some resource groups in the set may be OFFLINE or in the ERROR state.
- If one or more resource groups in the Same Node Dependency set fails, PowerHA® SystemMirror® tries to place all resource groups in the set on the node that can host all resource groups that are currently ONLINE (the ones that are still active) plus one or more failed resource groups.

To configure an Online on Same Node dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between resource groups › Configure Online on Same Node Dependency › Add Online on Same Node Dependency between resource groups** and press Enter.
3. Fill in the field as follows:

Add Online on Same Node Dependency between resource groups fields	
Field	**Value**
Resource groups to be Online on the same node	Select the resource groups from the list to be in this set of resource groups to be acquired and

Field	Value
	brought ONLINE on the same node (according to the startup policy and the availability of the node required). On fallback and fallover, the resource groups are processed simultaneously and brought ONLINE on the same target node (using the fallover and fallback policy defined for these groups).

4. Press Enter.
5. Verify the configuration.

Configuring Online on Different Nodes Dependency for resource groups

When you configure two or more resource groups to establish a location dependency between them, they belong to a *set* for that particular dependency. This topic discusses the Online On Different Nodes Dependency set of resource groups.

The following rules and restrictions apply to the Online On Different Nodes Dependency set of resource groups:

- Only one Online On Different Nodes Dependency set is allowed per cluster.
- Each resource group in the set should have a different home node for startup.
- When you configure resource groups in the Online On Different Nodes Dependency set you assign priorities to each resource group in case there is contention for a given node at any point in time. You can assign High, Intermediate, and Low priority. Higher priority resource groups take precedence over lower priority groups at startup, fallover, and fallback:
  - If a resource group with High Priority is ONLINE on a node, then no other resource group in the Different Nodes Dependency set can come ONLINE on that node.
  - If a resource group in this set is ONLINE on a node, but a resource group with a higher priority falls over or falls back to this node, the resource group with the higher priority will come ONLINE and the one with the lower priority will be taken OFFLINE and moved to another node if this is possible.
  - resource groups with the same priority cannot come ONLINE (startup) on the same node. Priority of a resource group for a node *within the same Priority Level* is determined by alphabetical order of the groups.
  - resource groups with the same priority do not cause one another to be moved from the node after a fallover or fallback.
  - If a parent/child dependency is specified, then the child cannot have a higher priority than its parent.
  - If a Start after dependency is specified, then the source cannot have a higher priority than its target.

To configure an Online On Different Nodes dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Online on Different Node Dependency › Add Online on Different Node Dependency between Resource Groups** and press Enter.
3. Complete the following fields and press Enter.

Add Online on Different Node Dependency between Resource Groups fields	
Field	**Value**
High Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) before lower priority resource groups. On fallback and fallover, these resource groups are processed simultaneously and brought ONLINE on different target nodes before any other groups are

Field	Value
	processed. If different target nodes are unavailable for fallover or fallback, these groups (same priority level) can remain on the same node. The highest relative priority within this list is the group listed first (on the left), as for the nodelist.
Intermediate Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) after the high priority groups and before low priority resource groups are brought ONLINE. On fallback and fallover, these resource groups are processed simultaneously and brought ONLINE on different target nodes after the high priority groups and before low priority resource groups are processed. If different target nodes are unavailable for fallover or fallback, these groups (same priority level) can remain on the same node. The highest relative priority within this list is the group listed first (on the left), as for the nodelist.
Low Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) after the higher priority resource groups are brought ONLINE. On fallback and fallover, these resource groups are brought ONLINE on different target nodes after the higher priority groups are processed. Higher priority groups moving to a node may cause these groups to be moved or taken OFFLINE.

4. Continue configuring runtime policies for other resource groups or verify the cluster.

Configuring processing order for resource groups

By default, PowerHA® SystemMirror® acquires and releases resource groups in parallel.

Any pair of resource groups that do not have the following attributes might be processed in any order, even if one of the resource groups of the pair has a relationship (serial order or dependency) with another resource group.

The following resource group attributes affect the acquisition order:

- Resource groups that are configured with a serial acquisition order are acquired in the specified order.
- Resource groups that are configured with dependencies with other resource groups are acquired in phases. Parent resource groups are acquired before child resource groups. Resource groups that are configured as the target for a Start After dependency are acquired before resource groups configured as the Source for the dependency.
- Resource groups that include NFS mounts are acquired before resource groups that do not include NFS mounts.

You can configure your application controller start scripts to run in the foreground or as a background process. The default setting for the application controller start script is to run as a background process. When the application controller start script runs as a background process the rest of the resource group processing continues, possibly starting other resource or groups, without waiting for the start script to complete. To change the application controller to foreground, see the Configuring application controllers topic.

The following resource group attributes affect the releasing order:

- Resource groups that are configured with a serial release order are released in the specified order.

- Resource groups that are configured with dependencies with other resource groups are released in phases. Child resource groups are released before parent resource groups. Resource groups that are configured as the target for Stop After dependency are released before resource groups configured as the Source for the dependency.
- Resource groups that include NFS mounts are released after resource groups that do not include NFS mounts.

Resource groups processing order and timers

PowerHA® SystemMirror® acquires resource groups in parallel, but if the settling time or the delayed fallback timer policy is configured for a particular resource group, PowerHA® SystemMirror® delays its acquisition for the duration specified in the timer policy.

Settling and delayed fallback timers do not affect the release process.

Prerequisites and notes for resource group ordering

These sections detail limitations of the resource group ordering.

Serial processing notes

When you configure individual resource groups that depend on other resource groups, you can customize to use the serial processing order that will dictate the order of processing on the local node. If you specify dependencies between resource groups, the order in which PowerHA® SystemMirror® processes resource groups cluster-wide is dictated by the dependency.

- Specify the same customized serial processing order on all nodes in the cluster. To do this, you specify the order on one node and synchronize cluster resources to propagate the change to the other nodes in the cluster. Also, since resource group dependencies also override any serial processing order, make sure that the serial order you specify does not contradict the dependencies. If it does, it will be ignored.
- If you have specified serial processing order for resource groups, and if in some of the resource groups only the NFS cross-mounting takes place during the acquisition (**node_up** event), or release (**node_down** event), then PowerHA® SystemMirror® automatically processes these resource groups after other resource groups in the list.
- If you remove a resource group that has been included in the customized serial ordering list from the cluster, then the name of that resource group is automatically removed from the processing order list. If you change a name of a resource group, the list is updated appropriately.

Parallel processing notes

In clusters where some groups have dependencies defined, these resource groups are processed in parallel using event phasing.

Error handling

If an error occurs during the acquisition of a resource group, recovery procedures are run after the processing of all other resource groups is complete.

If an error occurs during the release of a resource group, the resource group goes offline temporarily while PowerHA® SystemMirror® tries to recover it. If it moves to the ERROR state, you should take care of it manually.

Steps for changing resource group processing order

This topic discusses the steps for viewing or changing resource group processing order.

To view or change the current resource group processing order in SMIT:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-time Policies › Configure Resource Group Processing Ordering** and press Enter.
   SMIT displays the current processing order for the resource groups.

3. Enter field values as follows:

Configure Resource Group Processing Ordering	
Field	**Value**
Resource groups Acquired in Parallel	The current list of resource groups which are acquired in parallel by PowerHA® SystemMirror® on this node.
Serial Acquisition Order	The current serial order in which PowerHA® SystemMirror® serially acquires the specified resource groups on this node.
New Serial Acquisition Order	Enter the new list of resource group names. This list is the new serial order in which you want PowerHA® SystemMirror® to acquire the specified resource groups on this cluster node. The resource groups that are not included in this list are acquired in parallel by default.
Resource groups Released in Parallel	The current list of resource groups which are released in parallel by PowerHA® SystemMirror® on this node.
Serial Release Order	The current serial order in which PowerHA® SystemMirror® releases these resource groups on this node.
New Serial Release Order	Enter the new list of resource group names. This list is the new serial order in which you want PowerHA® SystemMirror® to release the specified resource groups on this cluster node.The resource groups that are not included in this list are released in parallel by default.

4. Press Enter to accept the changes. PowerHA® SystemMirror® checks that a resource group name is entered only once on a list and that all specified resource groups are configured in the cluster. Then it stores the changes in the PowerHA® SystemMirror® Configuration Database.
5. Synchronize the cluster in order for the changes to take effect across the cluster.
6. You can determine whether or not the resource groups are being processed in the expected order based on the content of the event summaries.

Configuring Workload Manager

IBM® offers AIX® Workload Manager (WLM) as a system administration resource included with AIX®. WLM allows users to set targets for and limits on CPU time, physical memory usage, and disk I/O bandwidth for different processes and applications. This provides better control over the usage of critical system resources at peak loads.

PowerHA® SystemMirror® allows you to configure WLM classes in PowerHA® SystemMirror® resource groups so that the starting, stopping, and active configuration of WLM can be under cluster control.

Steps for configuring WLM in PowerHA® SystemMirror®

Follow these basic steps for configuring WLM classes in PowerHA® SystemMirror®.

These steps include:

1. Configure WLM classes and rules, using the appropriate AIX® SMIT panels, as described here.
2. If you select a configuration other than the default (“PowerHA® SystemMirror®_WLM_config”), specify the WLM configuration to be used in PowerHA® SystemMirror®, as described here.
3. Assign the classes for this configuration to a resource group, selecting from a picklist of the classes associated with the default WLM configuration or the configuration you specified in Step 2. For instructions on adding resources to resource groups, see Adding resources and attributes to resource groups using the extended path.
4. After adding the WLM classes to the resource group - or after all resource group configuration is complete - verify and synchronize the configuration.

Once WLM is configured in PowerHA® SystemMirror®, PowerHA® SystemMirror® starts and stops WLM. If WLM is already running when PowerHA® SystemMirror® is started, PowerHA® SystemMirror® restarts it with a new configuration file. Therefore, only the WLM rules associated with classes in a resource group that can be acquired on a given node will be active on that node. Once PowerHA® SystemMirror® is stopped, WLM will be switched back to the configuration it was using when it was started.

Creating a new Workload Manager configuration

You can create a new WLM set of classes and rules

To set up WLM classes and rules, use the AIX® SMIT panels.

1. In AIX® SMIT, select **Performance & Resource Scheduling** > **Workload Management** > **Work on alternate configurations > Create a configuration.** (You can also get to the "alternate configurations" panel by typing *smitty wlm*.)
2. Enter the new name for the configuration in the **New configuration name** field. It is recommended to use the default name that PowerHA® SystemMirror® supplies: PowerHA® SystemMirror®_WLM_config.
3. Define classes and rules for the PowerHA® SystemMirror® configuration.

Defining a non-default Workload Manager configuration in PowerHA® SystemMirror®

You may have a non-default Workload Manager configuration. In this case, make this configuration known to PowerHA® SystemMirror®, so that it is managed.

To ensure that a non-default Workload Manager Configuration is managed by PowerHA® SystemMirror®:

1. Change the WLM runtime parameters to specify the PowerHA® SystemMirror® configuration.
2. From the main PowerHA® SystemMirror® SMIT panel, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-time Policies › Configure Workload Manager Parameters** and press Enter.
   This field indicates the WLM configuration to be managed by PowerHA® SystemMirror®. By default, the configuration name is set to **PowerHA® SystemMirror®_WLM_config**.
3. Specify a different configuration name if needed.

Verification of the Workload Manager configuration

After adding WLM classes to resource groups, or after you have finished configuring all your resource groups, verify that the configuration is correct.

Verification checks for the following conditions:

- For each resource group with which a WLM class is associated, an application controller is associated with this resource group. It is not required that an application controller exists in the resource group, but it is expected. PowerHA® SystemMirror® issues a warning if no application controller is found.
- Each WLM class defined to a PowerHA® SystemMirror® resource group exists in the specified PowerHA® SystemMirror® WLM configuration directory.
- A non-concurrent resource group (that does not have the Online Using Node Distribution Policy startup policy) does not contain a secondary WLM class without a primary class.
- A resource group with the startup policy Online on All Available Nodes has only a primary WLM class.
- A resource group with the startup policy Online Using Node Distribution Policy has only a primary WLM class.

Note: The **verification** utility cannot check class assignment rules to verify that the correct assignment will take place, since **PowerHA® SystemMirror®** has no way of determining the eventual gid, uid and

pathname of the user application. The user is entirely responsible for assigning user applications to the WLM classes when configuring WLM class assignment rules.

Cluster verification looks only for obvious problems and cannot verify all aspects of your WLM configuration; for proper integration of WLM with PowerHA® SystemMirror®, you should take the time to plan your WLM configuration carefully in advance.

Reconfiguration, startup, and shutdown of WLM by PowerHA® SystemMirror®

This section describes the way WLM is reconfigured or started or stopped once you have placed WLM under the control of PowerHA® SystemMirror®.

Workload Manager reconfiguration

When WLM classes are added to a PowerHA® SystemMirror® resource group, then at the time of cluster synchronization on the node, PowerHA® SystemMirror® reconfigures WLM so that it will use the rules required by the classes associated with the node.

In the event of dynamic resource reconfiguration on the node, WLM will be reconfigured in accordance with any changes made to WLM classes associated with a resource group.

Workload Manager startup

WLM startup occurs either when the node joins the cluster or when a dynamic reconfiguration of the WLM configuration takes place.

The configuration is node-specific and depends upon the resource groups in which the node participates. If the node cannot acquire any resource groups associated with WLM classes, WLM will not be started.

For a non-concurrent resource group with the startup policy other than Online Using Node Distribution Policy, the startup script will determine whether the resource group is running on a primary or on a secondary node and will add the corresponding WLM class assignment rules to the WLM configuration.

For each concurrent access resource group, and for each non-concurrent resource group with the startup policy Online Using Node Distribution Policy that the node can acquire, the primary WLM class associated with the resource group will be placed in the WLM configuration; the corresponding rules will be put into the rules table.

Finally, if WLM is currently running and was not started by PowerHA® SystemMirror®, the startup script restarts WLM from the user-specified configuration, saving the prior configuration. When PowerHA® SystemMirror® is stopped, it returns WLM back to its prior configuration.

Failure to start up WLM generates an error message logged in the hacmp.out log file, but node startup and/or the resource reconfiguration will proceed normally.

Workload Manager shutdown

WLM shutdown occurs either when the node leaves the cluster or on dynamic cluster reconfiguration.

If WLM is currently running, the shutdown script checks if the WLM was running prior to being started by the PowerHA® SystemMirror® and what configuration it was using. It then either does nothing (if WLM is not currently running), or stops WLM (if it was not running prior to PowerHA® SystemMirror® startup), or stops it and restarts it in the previous configuration (if WLM was Configuring Resources in a Resource Group

Once you have defined a resource group, you assign resources to it. SMIT cannot list possible shared resources for the node (making configuration errors more likely) if the node is powered off.

Configuring a settling time for resource groups

The settling time specifies how long PowerHA® SystemMirror® waits for a higher priority node (to join the cluster) to activate a resource group that is currently offline on that node. If you set the settling time, PowerHA® SystemMirror® waits for the duration of the settling time interval to see if a higher priority node may join the cluster, rather than simply activating the resource group on the first possible node that reintegrates into the cluster.

To configure a settling time for resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Settling Time for Resource Group** and press Enter.
   The **Configure Settling Time** panel appears.
3. Enter field values as follows:

Configure Settling Time fields	
Field	**Value**
Settling Time (in Seconds)	Enter any positive integer number in this field. The default is zero. In this case the resource group does not wait before attempting to start on a joining higher priority node. If you set the settling time, then if the currently available node that reintegrated into the cluster is not the highest priority node, the resource group waits for the duration of the settling time interval. When the settling time expires, the resource group is acquired on the node which has the highest priority among the list of nodes that joined the cluster during the settling time interval. If no nodes joined the cluster, the resource group remains offline. The settling time is only valid for resource groups that have the **Online on First Available Node** startup policy.

4. Press Enter to commit the changes and synchronize the cluster. This settling time is assigned to all resource groups with the **Online on First Available Node** startup policy.

Result
You can change, show or delete a previously configured settling time using the same SMIT path as described for configuring a settling time.

Defining delayed fallback timers

A delayed fallback timer lets a resource group fall back to its higher priority node at a specified time. This lets you plan for outages for maintenance associated with this resource group.

You can specify a recurring time at which a resource group will be scheduled to fall back, or a specific time and date when you want to schedule a fallback to occur.

You can specify the following types of delayed fallback timers for a resource group:

- Daily
- Weekly
- Monthly
- Yearly
- On a specific date.

> **Note:** It is assumed that the delayed timer is configured so that the fallback time is valid. If the configured time occurs in the past or is not valid, you receive a warning and the delayed fallback policy is ignored. If you use a specific date, the fallback attempt is made only once, at the specified time.

To make a resource group use a delayed fallback policy, follow these steps:

1. Configure a delayed fallback timer that you want to use. After you have configured the delayed fallback timer, you can use it in one or several resource groups as the default fallback policy. For instructions, see Configuring delayed fallback timers in SMIT.
2. Select the **Fallback to Higher Priority Node** option from the picklist of fallback policies for your resource group. You can do so when configuring a resource group.
   For instructions, see Steps for configuring resource groups in SMIT.

3. Assign a fallback timer to a resource group, by adding it as an attribute to the resource group.
   If the **delayed fallback timer** entry does not show up in the list of attributes/resources that you can add to a resource group, this indicates that you did not follow the instructions in steps 1 and 2, because PowerHA® SystemMirror® only displays attributes and resources that are valid in each particular case.

   For instructions, see Assigning a delayed fallback policy to a resource group.

Configuring delayed fallback timers in SMIT

Configure a delayed fallback timer that you want to use. After you have configured the delayed fallback timer, you can use it in one or several resource groups as the default fallback policy.

To configure a delayed fallback timer:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Delayed Fallback Timer Policies › Add a Delayed Fallback Timer Policy** and press Enter.
   A picklist **Recurrence for Fallback Timer** displays. It lists **Daily, Weekly, Monthly**, **Yearly** and **Specific Date** policies.
3. Select the timer policy from the picklist and press Enter. Depending on which option you select, a corresponding SMIT panel displays that lets you configure this type of a fallback policy.

Assigning a delayed fallback policy to a resource group

You must define the delayed fallback policies before you can assign them as attributes to resource groups.

To assign a delayed fallback policy to a resource group:

1. In PowerHA® SystemMirror® SMIT, create a resource group, or select an existing resource group.
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resource and Attributes for a Resource Group** and press Enter. SMIT displays a list of resource groups.
3. Select the resource group for which you want to assign a delayed fallback policy. The following panel appears. (The SMIT panel is abbreviated here. All valid options for the resource group are displayed based on the startup, fallover and fallback preferences that you have specified for the resource group.)
4. Enter field values as follows:

Resource group fields	
Field	**Value**
Resource Group Name	The name of the selected resource group displays here.
Participating Node Names (Default Node Priority)	The names of the nodes that can own or take over this resource group. The node with the highest priority is listed first, followed by the nodes with the lesser priorities.
Dynamic Node Priority (Overrides default)	The default is blank (the ordered nodelist). The preconfigured policies are listed. Note that this SMIT option displays only if you have previously selected **Fallover Using Dynamic Node Priority** as a fallover policy for this resource group.
Fallback Timer Policy (empty is immediate)	The default is blank (the resource group falls back immediately after a higher priority node joins). All configured fallback timer policies are listed in the picklist. Note that this SMIT option displays only if you have previously selected **Fallback to Higher Priority Node in the List** as a fallback policy for this resource group.

5. Press the F4 key to see the picklist in the **Fallback Timer Policy** field and select the fallback timer policy you want to use for this resource group.
6. Press Enter to commit the changes. The configuration is checked before populating the PowerHA® SystemMirror® Configuration Database. You can assign the same fallback timer policy to other resource groups.
7. Assign fallback timer policies to other resource groups and synchronize the cluster when you are done.

Using the node distribution startup policy

For each resource group in the cluster, you can specify a startup policy to be Online Using Node Distribution Policy.

This resource group policy is a cluster-wide attribute that causes the resource groups to distribute themselves in a way that only one resource group is acquired on a node during startup. Using this policy ensures that you distribute your CPU-intensive applications on different nodes.

The following statements apply:

- The only distribution policy supported in PowerHA® SystemMirror® is node-based distribution. You can use this policy whether or not you have sites configured in the cluster.
- If two or more resource groups with this startup policy are offline at the time when a particular node joins, the node acquires the resource group that has the least number of nodes in its nodelist. After considering the number of nodes, PowerHA® SystemMirror® sorts the list of resource groups alphabetically.
- If one of the resource groups with this startup policy is a parent resource group (it has a dependent resource group), PowerHA® gives preference to the parent resource group.
- If one of the resource groups with this startup policy is a startafter target resource group (it has a dependent resource group), PowerHA® gives preference to the target resource group.
- If you upgrade from a previous release that allowed network-based distribution, that configuration is automatically changed to node-based distribution.

When configuring the node distribution startup policy, take into consideration the following:

- If the number of resource groups is larger than the number of cluster nodes, PowerHA® SystemMirror® issues a warning. It is recommended that all resource groups that use node-based distribution have potential nodes on which they could be brought online during the cluster startup.
- resource groups configured for distribution during startup cannot have the fallover policy set to Bring Offline (on Error Node Only). If you select this combination of policies, PowerHA® SystemMirror® issues an error.
- resource groups configured for distribution during startup must use the Never Fallback policy. This is the only fallback policy PowerHA® SystemMirror® allows for such resource groups.
- If you configure multiple resource groups to use the Online Using Node Distribution startup policy, and you select the Prefer Primary Site inter-site management policy for all groups, the node-based distribution policy ensures that the primary site hosts *one group per node*. Whether the resource group will fall back to the primary site depends on the availability of nodes on that site.

PowerHA® SystemMirror® allows only valid startup, fallover and fallback policy combinations and prevents you from configuring invalid combinations.

Adding resources and attributes to resource groups

You can add, change or show resources and attributes for resource groups.

Keep the following in mind as you prepare to define the resources in your resource group:

- If you are configuring a resource group, first you configure timers (optional), startup, fallover, and fallback policies for a resource group, and then add specific resources to it. For information on configuring resource groups, see Configuring resource groups.
- You cannot change a resource group's policies once it contains resources. If you have added resources, you need to remove them prior to changing the resource group's policies.

- If you configure a nonconcurrent resource group (with the Online on Home Node startup policy) with an NFS mount point, you must also configure the resource to use IP Address Takeover. If you do not do this, takeover results are unpredictable. You should also set the field value **Filesystems Mounted Before IP Configured** to true so that the takeover process proceeds correctly.
- A resource group may include multiple service IP addresses. When a resource group configured with IPAT via IP Aliasing is moved, all service labels in the resource group are moved as aliases to the available interfaces, according to the resource group management policies in PowerHA® SystemMirror®.
- IPAT functionality is not applicable to concurrent resource groups.
- If you configure application monitoring, remember that PowerHA® SystemMirror® can monitor only one application in a given resource group, so you should put applications you intend to have PowerHA® SystemMirror® monitor in separate resource groups.
- If you plan to request PowerHA® SystemMirror® to use a forced varyon option to activate volume groups in case a normal varyon operation fails due to a loss of quorum, the logical volumes should be mirrored. It is recommended to use the **super strict** disk allocation policy for the logical volumes in AIX®.

To configure resources and attributes for a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group** and press Enter.
   SMIT displays a list of defined resource groups.
3. Select the resource group you want to configure and press Enter. SMIT returns the panel that matches the type of resource group you selected, with the **Resource Group Name**, **Inter-site Management Policy**, and **Participating Node Names (Default Node Priority)** fields filled in.
   SMIT displays only valid choices for resources, depending on the resource group startup, fallover, and fallback policies that you selected.

   Be aware that once you add resources to a resource group, its startup, fallover, and fallback policies cannot be changed unless you remove the resources. You can only change the resource group policies in a resource group that does not contain any resources yet. Plan your resource group's policies in advance, before adding resources to it.

   If the participating nodes are powered on, press F4 to list the shared resources. If a resource group/node relationship has not been defined, or if a node is not powered on, the picklist displays the appropriate warnings.
4. Enter the field values as follows (non-concurrent resource group is shown):

Nonconcurrent resource group fields	
Field	**Value**
Dynamic Node Priority (Overrides default)	Select the dynamic node priority policy.The default is blank (the ordered nodelist). The preconfigured dynamic node priority policies are listed.
DNP Script Path	If the Dynamic Node Priority (DNP) policy is either `cl_highest_udscript_rc` or `cl_lowest_nonzero_udscript_rc`, you must enter the full path and file name of a script. There is no default value.
DNP Script Timeout Value	If the Dynamic Node Priority (DNP) policy is either `cl_highest_udscript_rc` or `cl_lowest_nonzero_udscript_rc`, you must enter a maximum time for PowerHA® SystemMirror® to wait for the script to exit. The default value is `config_too_long`.
Service IP Labels/Addresses	Enter the IP labels/addresses to be taken over when this resource group is taken over, or select an IP label/address from the picklist. The picklist includes IP labels/addresses which rotate or may be taken over.
Application Controller	Enter or select from the picklist the application controllers to include in the resource group.
Volume Groups	Identify the shared volume groups that should be varied on when this resource group is acquired or taken over. Select the

Field	Value
	volume groups from the picklist or enter desired volume groups names in this field. If you previously requested that PowerHA® SystemMirror® collect information about the appropriate volume groups, the picklist displays list of all shared volume groups in the resource group and the volume groups that are currently available for import onto the resource group nodes. Specify the shared volume groups in this field if you want to leave the field file systems (empty is All for specified VGs) blank and to mount all file systems in the volume group. If you specify more than one volume group in this field, then all file systems in all specified volume groups will be mounted; you cannot choose to mount all file systems in one volume group and not to mount them in another. For example, in a resource group with two volume groups, vg1 and vg2, if the **file systems** (empty is All for specified VGs) is left blank, then all the file systems in vg1 and vg2 will be mounted when the resource group is brought up. However, if the **file systems** (empty is All for specified VGs) has only file systems that are part of the vg1 volume group, then none of the file systems in vg2 will be mounted, because they were not entered in the **file systems** (empty is All for specified VGs) field along with the file systems from vg1. If you have previously entered values in the file systems field, the appropriate volume groups are already known to the PowerHA® SystemMirror® software.
Use Forced Varyon of Volume Groups, if Necessary	The default is **false**. If this flag is set to **true**, PowerHA® SystemMirror® uses a forced varyon to bring each volume group that belongs to this resource group online in the event that a normal varyon for the volume group fails due to a lack of quorum, and if PowerHA® SystemMirror® finds at least one complete copy of every logical partition in every logical volume available for this volume group. Use this option only for volume groups in which every logical volume is mirrored. It is recommended to use the **super strict** disk allocation policy; the forced varyon operation is unlikely to be successful for other choices of logical volume configuration.
File Systems (empty is All for specified VGs)	Leave this field blank, if you want all file systems in the specified volume groups to be mounted by default when the resource group, containing this volume group, is brought online. If you leave the file systems (empty is All for specified VGs) field blank and specify the shared volume groups in the Volume Groups field below, all file systems will be mounted in the volume group. If you leave the file systems field blank and do not specify the volume groups in the field below, no file systems will be mounted. You may also select individual file systems to include in the resource group. Press F4 to see a list of the file systems. In this case, only the specified file systems will be mounted when the resource group is brought online. file systems (empty is All for specified VGs) is a valid option only for non-concurrent resource groups.
File systems Consistency Check	Identify the method to check consistency of the fsck (default) or logredo (for fast recovery) file systems. If you choose logredo and the logredo function fails, then the fsck runs in its place.

Field	Value
File systems Recovery Method	Identify the recovery method for the file systems, parallel (for fast recovery) or sequential (default). Do not set this field to parallel if you have shared, nested file systems. These must be recovered sequentially. (Note that the cluster verification process does not report file system and fast recovery inconsistencies.)
File systems Mounted Before IP Configured	Specify whether, on takeover, PowerHA® SystemMirror® takes over volume groups and mounts a failed node's file systems before or after taking over the failed node's IP address or addresses. The default is **false**, meaning the IP address is taken over first. Similarly, upon reintegration of a node, the IP address is acquired before the file systems. Set this field to **true** if the resource group contains file systems to export. This is so that the file systems will be available once NFS requests are received on the service address.
File systems/Directories to Export (NFSv2/3)	Formerly named file systems/Directories to Export. This field contains a space-separated list of local file systems managed by the resource group that should be exported using the NFSv2/3 protocol. • If the field is left blank (default), then no file systems are exported using the NFSv2/3 protocol. • A file system can be listed for both NFSv2/3 and NFSv4 exporting by listing it in both fields of this SMIT panel. • This field must be left blank if the resource group contains more than two nodes. • For backwards compatibility, if File systems/Directories to Export (NFSv4) is left blank, then PowerHA® SystemMirror® uses the protocol versions specified in the /usr/es/sbin/cluster/etc/exports file. If File systems/Directories to Export (NFSv4) is not blank, then PowerHA® SystemMirror® ignores the protocol versions specified in the exports file.
File systems/Directories to Export (NFSv4)	This field contains a space separated list of local file systems managed by the resource group that should be exported using the NFSv4 protocol. • If the field is left blank (default), then PowerHA® SystemMirror® uses the protocol versions specified in the /usr/es/sbin/cluster/etc/exports file. • A file system can be listed for both NFSv2/3 and NFSv4 exporting by listing it in both fields of this SMIT panel. • This field is present only if the cluster.es.nfs.rte fileset is installed.
File systems/Directories to NFS Mount	Identify the file systems or directories to NFS mount. All nodes in the resource chain will attempt to NFS mount these file systems or directories while the owner node is active in the cluster.Example: Using `/fs1` from the previous entry, you enter the remote mount, then the local mount: `/rfs1;/fs1` .

Field	Value
Network for NFS Mount	(*Optional* .) Select the network where you want to NFS mount the file systems from a picklist of a previously defined IP networks. This field is relevant only if you have filled in the **File systems/ Directories to NFS Mount** field. The **Service IP Labels/IP Addresses** field should contain a service label which is on the network you select. You can specify more than one service label in the **Service IP Labels/IP Addresses 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.
NFSv4 Stable Storage Path	This field contains the path where NFSv4 stable storage is stored. • The path should belong to a file system managed by the resource group. • The path does not have to be an existing directory. PowerHA® SystemMirror® creates the path automatically. • If this field contains a non-empty value and the File systems/Directories to Export (NFSv4) field is blank, then the contents of this field are ignored and a warning is printed. • This field is present only if the cluster.es.nfs.rte file set is installed.
Raw Disk PVIDs	Press F4 for a listing of the PVIDs and associated hdisk device names. If you have previously entered values in the File systems or Volume groups fields, the appropriate disks are already known to the PowerHA® SystemMirror® software. If you are using an application that directly accesses raw disks, list the raw disks here.
Tape resources	Enter or select from the picklist the tape resources that you want started on the resource group. The picklist displays a list of resources previously defined in the **Define tape resources** panel.
Miscellaneous Data	Miscellaneous Data is a string placed into the MISC_DATA environment variable. The MISC_DATA environment variable is accessible by scripts, for example pre-event and post-event scripts and application controller start and stop scripts. You can use this field to provide a description to your resource group.
Primary Workload Manager Class	Select from the picklist of Workload Manager (WLM) classes associated with the PowerHA® SystemMirror® WLM configuration specified. • For non-concurrent resource groups with the startup policy of Online on Home Node Only, or Online on First Available Node, if no secondary WLM class is specified, all nodes use the primary WLM class. If a secondary class is

Field	Value
	specified, only the primary node uses the primary WLM class. • For non-concurrent resource groups with the startup policy of Online Using a Distribution Policy, all nodes in the resource group will use the primary WLM class. • For concurrent resource groups, all nodes in the resource group will use the primary WLM class.
Secondary Workload Manager Class	(*Optional*) Press F4 and select from the picklist of Workload Manager class associated with this resource group. Only non-concurrent resource groups with the startup policy of either Online On Home Node Only, or Online on First Available node are allowed to use secondary WLM classes. If no secondary WLM class is specified, all nodes in the resource group use the primary WLM class. If you specify a secondary class here, the primary node uses the primary WLM class and all other nodes use the secondary WLM class.
Automatically Import Volume Groups	Specifies whether PowerHA® SystemMirror® should *automatically* import those volume groups that are defined in the **Volume Groups** or **Concurrent Volume Groups** fields. By default, **Automatically Import Volume Groups** flag is set to false. If **Automatically Import Volume Groups** is set to false, then selected volume groups will not be imported automatically. In this case, when you add volume groups to the resource group, make sure that the selected volume groups have already been imported to each of the nodes using the importvg command or C-SPOC. If **Automatically Import Volume Groups** is set to true, then when you press Enter, PowerHA® SystemMirror® determines whether the volume group that you entered or selected in the **Volume Groups** or **Concurrent Volume Groups** fields needs to be imported to any of the nodes in the resource group, and automatically imports it, if needed.
Fallback Timer Policy (empty is immediate)	This field displays only if you have previously selected **Fallback to Higher Priority Node in the List** as a fallback policy. The default is blank (the resource group falls back immediately after a higher priority node joins). The picklist contains all configured fallback timer policies.
WPAR Name (empty is WPAR-disabled)	Setting this field to be same as the resource group name will WPAR-enable this resource group. A pick-list is provided with the expected name for the WPAR. For more information, see Running a resource group in an AIX® WPAR.

5. Press Enter to add the values to the PowerHA® SystemMirror® Configuration Database.
6. Synchronize the cluster.

Reliable NFS function

You can configure NFS in all non-concurrent resource groups.

As you configured resources, you can specify the following items related to NFS:

- Use the Reliable NFS server capability that preserves locks and dupcache. (This functionality is restricted to two-node resource groups if it contains NFSv2 amd NFSv3 exports. If all the exports in the resource group are only NFSv4, then up to 16-node Resource Group configurations are supported.)

- Specify a Stable Storage Location if the Resource Group has NFSv4 exports.
- Specify a network for NFS mounting.
- Define NFS exports and mounts at the directory level.
- Specify export options for NFS-exported directories and file systems.

Managing control over NFS file systems in a PowerHA® SystemMirror® cluster

Once NFS file systems become part of resource groups that belong to an active PowerHA® SystemMirror® cluster, PowerHA® SystemMirror® takes care of exporting, unexporting, cross-mounting and unmounting the file systems, during cluster events, such as fallover of a resource group containing the file system to another node in the cluster.

If for some reason you stop the cluster services and must manage the NFS file systems manually, the file systems must be unmounted before you restart the cluster services. This enables management of NFS file systems by PowerHA® SystemMirror® once the nodes join the cluster.

NFS exporting file systems and directories

The process of NFS-exporting file systems and directories in PowerHA® SystemMirror® differs from that in AIX®.

Specifying file systems and directories to NFS export

In AIX®, you list file systems and directories to be NFS-exported in the **/etc/exports** file; in PowerHA® SystemMirror®, you must put these in a resource group.

You can configure NFS in all non-concurrent resource groups.

Specifying export options for NFS exported file systems and directories

If you want to specify special options for NFS-exporting in PowerHA® SystemMirror®, you can create a **/usr/es/sbin/cluster/etc/exports** file.

This file has the same format as the regular **/etc/exports** file used in AIX®.

Use of this alternate exports file is optional. PowerHA® SystemMirror® checks the **/usr/es/sbin/cluster/etc/exports** file when NFS-exporting a file system or directory. If there is an entry for the file system or directory in this file, PowerHA® SystemMirror® will use the options listed except that PowerHA® SystemMirror® will ignore the version option as described in the section Adding resources and attributes to resource groups using the extended path. If the file system or directory for NFS-export is not listed in the file, or, if the user has not created the **/usr/es/sbin/cluster/etc/exports** file, the file system or directory will be NFS-exported with the default option of root access for all cluster nodes.

Configuring the optional /usr/es/sbin/cluster/etc/exports file

In this step, you add the directories of the shared file systems to the exports file.

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.

To add a directory to Exports List, complete the following steps:

1. In SMIT, enter the fastpath `smit mknfsexp`.
   The system displays the **Add a Directory to Exports List** panel.
2. In the **EXPORT directory now**, **system restart or both** field, enter **restart**.
3. In the **PATHNAME of alternate Exports file** field, enter `/usr/es/sbin/cluster/etc/exports`. This step creates 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/es/sbin/cluster/etc/exports** file.
5. Return to the **Add a Directory to Exports List** panel, or exit SMIT if you are finished.

6. Repeat steps 1 through 4 for each file system or directory.

Forcing a varyon of volume groups

Forcing a varyon of volume groups is an option that you should use only with understanding of its consequences. This section describes the conditions under which you can safely attempt to forcefully bring a volume group online on the node, in the case when a normal varyon operation fails due to a loss of quorum.

We recommend to specify the super strict disk allocation policy for the logical volumes in volume groups for which forced varyon is specified. Configuring the super strict disk allocation policy for volume groups that may be forced on does the following:

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

> **Note:** You should apply the **super strict** disk allocation policy for disk enclosures in the cluster. You specify the **super strict** policy under the **Allocate each logical partition copy on a separate physical volume?** option in the **Add a Logical Volume**, or **Change/Show a Logical Volume** SMIT panels in AIX®. Also, if you are using the super strict disk allocation policy, specify the correct number of physical volumes for this logical volume and do not accept the default setting of 32 physical volumes.

Use independent disk enclosures that use logical volume mirroring; place logical volume mirror copies on separate disks that rely on separate power supplies, and use separate physical network interfaces to ensure access. This ensures that no disk is a single point of failure for your cluster.

You can specify a forced varyon attribute for:

- Volume groups on SCSI disks that use LVM mirroring where you want to NFS mount the file systems
- Volume groups that are mirrored between separate RAID or ESS devices.

> **Note:** Be aware that when the forced varyon facility is used successfully and the volume group is brought online on the node (using the one complete copy of the data that was found), the data that you recover by forcing a volume group to go online is guaranteed to be consistent, but not necessarily the latest.
> During run time, for large volume groups (those with more than 256 disks), checking logical partition maps may take extra processing time. However, since this time delay occurs only when you select a forced varyon for a large volume group in the case when a normal varyon failed due to a lack of quorum, enduring a slow varyon process that enables data recovery is preferable to having no chance at all to activate the volume group.

When PowerHA® SystemMirror® attempts a forced varyon

For troubleshooting purposes, it is helpful to know under what conditions or cluster events PowerHA® SystemMirror® attempts a forced varyon, when this is configured. In general, PowerHA® SystemMirror® attempts a forced varyon in the event of a cluster failure.

The following list contains examples of cluster event failures that can trigger a forced varyon:

- Cluster startup, normal varyon fails due to a loss of quorum on one of the disks.
- Nodes joining the cluster, normal varyon fails due to a loss of quorum on one of the disks.
- Node reintegration, normal varyon fails for concurrent resource groups.
- Selective fallover caused by an application or a node failure moves a resource group to a takeover node.
- Selective fallover caused by a loss of quorum for a volume group moves a resource group to a takeover node.
  When PowerHA® SystemMirror® selectively moves a resource group for which a loss of quorum for a volume group error has occurred, it tries to bring the volume groups online on the takeover node. If a

normal varyon process for volume groups fails at this point, and, if you have specified a forced varyon for the volume groups in this resource group, then, since quorum is lost, PowerHA® SystemMirror® attempts a forced varyon operation.

To summarize, for the cases where PowerHA® SystemMirror® uses selective fallover to move the resource groups, the sequence of events would be the following:

- If, after an **rg_move** event, a forced varyon is launched and is successful, the resource group remains online on the node to which it has been moved.
- If, after an **rg_move** event, a forced varyon is launched and fails, selective fallover continues to move the resource group down the node chain.
  If a resource failure occurs in a concurrent resource group, PowerHA® SystemMirror® takes this resource group offline on a particular node. In this case, use the **clRGmove** utility to manually bring the resource group online on the node.

Avoiding a partitioned cluster

The forced option to activate a volume group must be used with care.

Should the cluster become partitioned, each partition might force on the volume group and continue to run. In this case, two unequal copies of the data will be active at the same time. This situation can cause data divergence and does not allow a clean recovery. Were this to happen with a concurrent volume group, the consequences would be even worse, as the two sides of the cluster would have made uncoordinated updates.

PowerHA® SystemMirror® automatically uses monitoring of all available storage and network paths to avoid a partitioned cluster

Verification checks for forced varyon

If you specified a forced varyon attribute for a resource group, and PowerHA® SystemMirror® detects that the logical volumes are not being mirrored with the **super strict** disk allocation policy, PowerHA® SystemMirror® a warns upon verification of cluster resources. In this case, a forced varyon operation may not succeed.

As part of the process, PowerHA® SystemMirror® checks the logical partitions on each disk for each volume group:

- If it cannot find a complete copy of every logical volume for a volume group, an error message:“ "Unable to vary on volume group <vg name> because logical volume <logical volume name> is incomplete"” displays in the **hacmp.out** file. In this case, a forced varyon operation fails and you will see an event error.
- If PowerHA® SystemMirror® can find a complete copy for every logical volume for all volume groups in this resource group that require a forced varyon, it varies on the volume groups on the node in the cluster.

Running a resource group in an AIX® WPAR

AIX® Workload Partitions (WPAR) are software created virtualized operating system environments within a single instance of the AIX® operating system. To most applications, the workload partition appears to be a separate instance of AIX® because applications and workload partitions have a private execution environment. Applications are isolated in terms of process, signal and file system space. Workload partitions have their own unique users and groups. Workload partitions have dedicated network addresses and inter-process communication is restricted to processes executing in the same workload partition.

For example, within a workload partition:

- A ps command shows only those processes that are in the workload partition.
- Signals can be sent (for example, with the kill command), only to processes in the workload partition.
- In general, files are unique to the workload partition.
- An id command reports the user and groups created and assigned within the workload partition.
- Network servers running within the workload partition only receive requests targeted for the IP addresses assigned to that partition. Network servers are unaware of requests destined for other IP addresses configured on the system. These IP addresses can be assigned to other workload partitions.

- Interprocess communication facilities for messages, shared memory and semaphores can only be employed to communicate with processes in the same workload partition. The ipcs command only reports objects created by processes in the workload partition.

Most applications are unaware of the software creating the workload partition and run unmodified in the workload partition. Workload partitions are also integrated with AIX® resource controls and it is possible to assign processor and/or memory shares to a workload partition, as well as to establish limits on threads and processes.

When AIX® is installed and started, a special partition is created. This partition, termed the global partition, is where the administrator first logs in. All subsequent workload partitions are created from the global partition and many workload partition administrative tasks can only be performed from the global environment. Many commands also function differently when they are run in the global partition.

PowerHA® SystemMirror® support for running a resource group in an AIX® WPAR

When a WPAR-enabled resource group is brought online, all its associated resources are activated within the corresponding WPAR. The WPAR-enabled resource group is associated with a WPAR based on their common name. If a resource group called `test_resource_group` is WPAR-enabled, then it is associated with a WPAR with the name `test_resource_group`.

When a resource group is made WPAR-enabled, all the user-defined scripts (such as application start, stop, and monitoring scripts) must be accessible within the WPAR, at the paths that are specified in the PowerHA® SystemMirror® configuration.

Support for mixed-type nodes

A WPAR-enabled resource group can consist of some nodes that are not WPAR capable. To use the WPAR functions, you do not need to upgrade all their nodes of the resource group to the latest version of the AIX® operating system.

When a WPAR-enabled resource group comes online on a WPAR-incapable node, it behaves as if the WPAR property for the resource group was not set. You must ensure that all the user-defined scripts are accessible at the same path that was previously specified in the PowerHA® SystemMirror® configuration.

Enable/disable WPAR property of a resource group

If the WPAR property of a resource group is changed by using DARE (when the resource group is online), then the WPAR property takes effect only when the resource group is brought online the next time.

Resource assignment to a WPAR-enabled resource group

PowerHA® SystemMirror® automatically assigns and unassigns resources to a WPAR as the corresponding WPAR-enabled resources comes online or goes offline. You must not assign any PowerHA® SystemMirror® resources to a WPAR.

Restrictions for running a resource group in a WPAR

PowerHA® SystemMirror® has the following restrictions for WPAR support:

- Only the following resource types are supported to run in a WPAR: Service Label, Application Controllers, and File Systems.
- Every resource group that is to be run in a WPAR should have at least one service address that is associated with it.
- PowerHA® SystemMirror® Smart Assist scripts are not supported for a WPAR enabled resource group. Thus, any application controller or application monitoring script that uses the PowerHA® SystemMirror® Smart Assist scripts cannot be configured as a part of a WPAR enabled resource group.
- For all applications that are associated with a WPAR-enabled resource group, you must ensure that the application start, stop, and other user-defined scripts (such as application monitoring scripts), are accessible within the WPAR.

> **Note:** The PowerHA® SystemMirror® configuration verification is not checking for the access permissions for the application scripts that are a part of a WPAR-enabled resource group.

- When a resource group is enabled to use WPAR, PowerHA® SystemMirror® manages the WPAR by using the resource group management function during cluster events. You must not manage this WPAR as you would a stand-alone WPAR. For example, you must not log in to the WPAR and stop it, and you must not stop the WPAR from the global environment. PowerHA® SystemMirror® does not monitor the operational state of the WPAR. If you manage the WPAR outside of PowerHA® SystemMirror® it might cause the WPAR enabled resource group to go into the Error state.
- Process application monitoring is not supported for WPAR-enabled resource groups.
- For every WPAR-capable node that is a part of a WPAR-enabled resource group and contains a WPAR for a WPAR-enabled resource group, at least one of the service labels (of the WPAR-enabled resource group) must be accessible from the corresponding global WPAR.

> **Note:** When a WPAR-enabled resource group is brought online on a WPAR capable node, PowerHA® SystemMirror® (which runs in the global WPAR), automatically sets up rsh access to the corresponding WPAR to manage various resources associated with the resource group.

Other operations on a WPAR-enabled resource group

This topic discusses other operations that you can perform on a WPAR-enabled resource group.

Deleting a WPAR-enabled resource group

When a WPAR-enabled resource group is deleted, the corresponding WPAR on the nodes of the resource group are unaffected (that is, the corresponding WPAR is not deleted).

Changing the Name of a WPAR-Enabled resource group

If the name of a WPAR-enabled resource group is changed, then you must ensure that there is a WPAR with the corresponding "new" name on each of the WPAR-capable nodes in which the resource group is run.

DARE Operations on a WPAR-Enabled resource group

All the supported resource types that are supported for a WPAR-enabled resource group can be DARE added and removed from a WPAR-enabled resource group.

If the WPAR property of a resource group is changed through DARE (when the resource group is online), then its affect takes place when the resource group is brought online next time.

PowerHA® SystemMirror® configuration verification and corrective actions

PowerHA® SystemMirror® configuration verification checks that all WPAR-capable nodes of a WPAR-enabled RG, have a WPAR configured for the resource group (that is, a WPAR with the same name as the resource group). If the PowerHA® SystemMirror® configuration verification is run with corrective action enabled, then you are prompted to fix the WPAR related verification errors through PowerHA® SystemMirror® corrective action.

Testing your configuration

After you configure a cluster, you should test it before making it available in a production environment.

For information about using Cluster Test Tool to test your cluster, see Testing a PowerHA® SystemMirror® cluster.

Configuring cluster events

The PowerHA® SystemMirror® 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.

To configure cluster events, you indicate the script that handles the event and any additional processing that should accompany an event as described here. You can define multiple customized pre- and post-event scripts (for a particular cluster event). The environment variable EVENT_STAGE will be set to the appropriate value of *pre, post, notify,* or *recovery* when the corresponding event command is run.

Considerations for pre-event and post-event scripts

Take into account the following information when planning your pre-event and post-event scripts.

Using shell environment variables in pre-event and post-event scripts

When writing your pre-event or post-event script, none of the shell environment variables defined in **/etc/environment** are available to your program. If you need to use any of these variables, for example, PATH and NLSPATH, you must explicitly source them by including this line in your script:

```
. /etc/environment
```

event_error now indicates failure on a remote node

All cluster nodes run the **event_error** event if any node has an unrecoverable error. All nodes log the error and call out the failing node name in the **hacmp.out** log file. If you have added pre-event or post-event scripts for the **event_error** event, be aware that they are called on each node, not just on the failing node.

An environment variable that indicates the node where the event script failed, EVENT_FAILED_NODE, is set to the name of the node where the event occurred. Use this variable in your pre-event or post-event scripts to locate the failure.

The variable LOCALNODENAME identifies the local node; if LOCALNODENAME is not the same as EVENT_FAILED_NODE, then the failure occurred on a remote node.

Parallel processing of resource groups affects event processing

When resource groups are processed in parallel, fewer cluster events occur in the cluster. In particular, only **node_up** and **node_down** events take place, and events such as **node_up_local** or **get_disk_vg_fs** do not occur. This is because PowerHA® SystemMirror® uses other methods to process resources in parallel. As a result, the use of parallel processing reduces the number of particular cluster events for which you can create customized pre-event or post-event scripts. If you start using parallel processing for some of the resource groups in your configuration, be aware that your existing event scripts may not work for the resource groups.

Dependent resource groups and the use of pre-event and post-event scripts

If you are using pre-event and post-event scripts or other methods, such as customized serial resource group processing to establish dependencies between applications that are supported by your cluster, then these methods may no longer be needed or can be significantly simplified. Instead, you can specify dependencies between resource groups in a cluster. For more information on how to configure resource group dependencies, see Configuring dependencies between resource groups.

If you still want to customize behavior for some applications, consider adding a pre-event or post-event script to the **resource_state_change** event.

How notification, pre and post events affect execution of cluster events?

Notification script(s) are invoked as background processes and do not affect the rest of the calling sequence, in other words, the return code from a notification script is ignored.

Pre and post scripts or custom events are invoked in the foreground. By default the return code from a pre or post event is ignored, however, you can change this behavior such that a non-zero return code is treated as an event failure which will terminate any further event processing.

You can configure this option using smit:

`Custom Cluster Configuration->Events->Cluster Events->Change/Show Pre-Defined Events`

Change the `Fail event if pre or post event fails?` option to Yes to have the failure of your pre or post event treated as an event failure.

You can also change this option using the clmgr **`change event`** command by setting the `PREPOSTFAILS` option to true.

Refer:

The default sequence of invocation for pre, post and notify scripts is as follows:

- Notify script(s) are invoked as background processes
- Pre event script(s) are called in the foreground
- The “main” or predefined event script is called in the foreground
- Post event script(s) are called in the foreground
- Notify script(s) are invoked as background processes

If you select the option to fail the event if a pre or post event fails, a failure of any of thoe scripts will be treated the same as a failure of the predefined event and will cause the cluster to enter the RP_FAILED state. The notify script(s) will be run, but no further event processing will occur.

For example, if you select the option to fail the event if a pre event fails, and your pre event script returns non-zero, the sequence of invocations will be:

- Notify script(s) are invoked as background processes
- Pre event script(s) is called in the foreground and returns non-zero
- Notify script(s) are invoked as background processes

In this case neither the main or predefined event script nor any post event script(s) are run.

If you select the option to fail the event if a post event fails, and your post event script returns non-zero, then all scripts will be run in the same sequence as the default case, but the non-zero return code will be treated as an error and the cluster will go to the RP_FAILED state.

If the cluster goes to the RP_FAILED state, you will need to manually recover cluster services as described here:

Configuring pre-event and post-event commands

You can use SMIT to define your customized cluster event script.

Pre and post event scripts can be specified directly by entering the full path to your script, or you can define a custom event as described below. You can specify multiple scripts or custom event names separated by commas.

If you do not want to configure custom events and simply specify a full path instead, you can proceed directly to Configuring pre-event and post-event processing.

To define your customized cluster event scripts:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Pre/Post-Event Commands › Add a Custom Event Command** and press Enter.
3. Enter the field values as follows:

Add a Custom Event Command fields	
Field	**Value**
Cluster Event Name	Enter a name for the command. The name can have a maximum of 64 characters.
Cluster Event Description	Enter a short description of the event.
Cluster Event Script Filename	Enter the full pathname of the user-defined script to execute.

4. Press Enter to add the information to PowerHA® SystemMirror®custom class in the local PowerHA® SystemMirror® Configuration Database (ODM).
5. Go back to the **Events** menu and select **Verify and Synchronize Cluster Configuration (Advanced)** to synchronize your changes across all cluster nodes.

> Synchronizing does not propagate the actual new or changed scripts; you must add these to each node manually.

Configuring pre-event and post-event processing

Complete the following steps to set up or change the processing for an event. In this step you indicate to the Cluster Manager to use your customized pre-event or post-event commands. You only need to complete these steps on a single node. The PowerHA® SystemMirror® software propagates the information to the other nodes when you verify and synchronize the nodes.

To configure pre-event and post-events for customized event processing, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. Select **Custom Cluster Configuration › Events › Cluster Events › Change/Show Pre-defined Events** to display a list of cluster events and subevents.
3. Select an event or subevent that you want to configure and press Enter. SMIT displays the panel with the event name, description, and default event command shown in their respective fields.
4. Enter field values as follows:

Field name	**Description**
Event Name	The name of the cluster event to be customize.
Description	A brief description of the event's function. This information cannot be changed.
Event Command	Indicates the full path of the script that will be called by PowerHA® SystemMirror®. You cannot edit this field.
Notify Command	(*Optional*) Enter the full pathname of a user-supplied script to run before and after a cluster event. This script can notify the system administrator that an event is about to occur or 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 trailing arguments passed to the event command.
Pre-Event Command	(*Optional*) If you have defined custom cluster events, press F4 for the list. Or, enter the name of a custom-defined event to run before the PowerHA® SystemMirror® cluster event command runs. This command is run before the "event command" script is run. The arguments passed to this command are the event name and the trailing arguments passed to the event command. Enter the full path name of a user-supplied script or the name of a custom cluster event. Remember that the Cluster Manager will not process the event until this pre-event script or command has completed.

Field name	Description
	Verify that the pre-event command or script returns with an exit value of 0, otherwise the event fails with an error. You can specify multiple script names or custom event names separated by commas.
Post-Event Command	(*Optional*) If you have defined custom cluster events, press F4 for the list. Or, enter the name of the custom event to run after the PowerHA® SystemMirror® cluster event command executes successfully. This script provides post-processing after a cluster event. Enter the full path name of a user-supplied script or the name of a custom cluster event. The arguments passed to this command are the event name, event exit status, and the trailing arguments passed to the event command. Verify that the post-event command or script returns with an exit value of 0, otherwise the event fails with an error. You can specify multiple script names or custom event names separated by commas.

5. Press Enter to add this information to the PowerHA® SystemMirror® Configuration Database.
6. Return to the **Events** menu and synchronize your event customization by selecting the **Verify and Synchronize Cluster Configuration (Advanced)** option. Note that all PowerHA® SystemMirror® event scripts are maintained in the `/usr/es/sbin/cluster/events` directory. The parameters passed to a script are listed in the script's header.

> You or a third-party system administrator can reset the PowerHA® SystemMirror® tunable values, such as cluster event customizations, to their installation-time defaults.

Tuning event duration time until warning

Depending on cluster configuration, the speed of cluster nodes and the number and types of resources that need to move during cluster events, certain events may take different times to complete. Cluster events run asynchronously and usually call AIX® system commands. Since PowerHA® SystemMirror® has no means to detect whether the event script is actually performing useful work at a given period of time, it runs a **config_too_long** event (which sends messages to the console and to the **hacmp.out** file) each time the processing of the event exceeds a certain amount of time. For such events, you may want to customize the time period PowerHA® SystemMirror® waits for an event to complete before issuing the **config_too_long** warning message.

> **Note:** The **config_too_long** warning timer for **node_up** should be adjusted to allow for longer time to process **node_up** events with dependent resource groups. **node_up** processing in clusters with dependencies could take more time than in the clusters without dependent resource groups.

Event duration overview

Before tuning event duration time, read these prerequisites and notes.

The following are important to keep in mind when you are working with event duration:

- The total duration time is calculated differently for "slow" and "fast" cluster events.
  "Fast" events are those that do not include acquiring or releasing resources and normally take a shorter time to complete.

  For "fast" events, the total duration time during which PowerHA® SystemMirror® waits before issuing a warning is equal to Event Duration Time.

"Slow" cluster events are those that involve acquiring and releasing resources or use application controller start and stop scripts. "Slow" events may take a longer time to complete. Customizing event duration time for "slow" events lets you avoid getting unnecessary system warnings during normal cluster operation.

For "slow" events, the total duration time before receiving a **config_too_long** warning message is set to the sum of **Event-only Duration Time** and **Resource Group Processing Time**.

- Remember, you can customize event duration time before receiving a warning for cluster events, not for nodes or specific resource groups in your cluster. Once the **Total Event Duration Time** is specified, the system waits for the specified period of time and sends a **config_too_long** message to the node which was affected by this event.
  For example, you have a cluster with five resource groups. A **node_down** event (a "slow" event) occurs on Node A, which owns some of the resource groups. And, you have previously specified the **Event-only Duration Time** to be 120 seconds, and the **Resource Group Processing Time** to be 400 seconds.

  When a **node_down** event occurs on Node A, a **config_too_long** message is sent to Node A according to this formula:

  ```
  Event Duration Time (120 seconds) + Resource Group Processing
  Time (400 seconds) = 520 seconds (Total Event Duration Time).
  ```

  A **config_too_long** message appears on Node A after 520 seconds.
- During dynamic reconfiguration events, the Cluster Manager uses the previously specified values of the event duration time until warning. After dynamic reconfiguration is complete and the new values of event duration time get synchronized, the Cluster Manager uses the newly specified values.

To configure event duration, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. From the SMIT interface, select **Custom Cluster Configuration › Events › Cluster Events › Change/Show Time Until Warning**, and press Enter.
3. Modify any of the fields press Enter.

Changing event duration time until warning

This topic discusses how to change the total event duration time before receiving a **config_too_long** warning message.

Perform the following procedure on any cluster node:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Change/Show Time Until Warning** and press Enter.
3. Enter data in the fields as follows:

Change/Show Time Until Warning fields	
Field	**Value**
Max. Event-only Duration (in seconds)	Enter any positive integer. This is the maximum time (in seconds) it takes to execute a cluster event. The default **Max. Event-only Duration** is 180 seconds. For "fast" cluster events, such as events that do not involve acquiring or releasing resource groups, the total event duration time before PowerHA® SystemMirror® issues a warning is equal to the **Max. Event-only Duration.**
Max. Resource Group Processing Time (in seconds)	Enter any positive integer or zero. This is the maximum time (in seconds) it takes to acquire or release a resource group. The default **Max. Resource Group Processing Time** time is 180 seconds. Note that if you have several resource groups that have to be acquired or released during a cluster event, the value in

Field	Value
	this field should be longer than the maximum acquisition or release time for any of the cluster resource groups. For "slow" cluster events, such as events that include acquiring or releasing resources, the total event duration time (before PowerHA® SystemMirror® issues a warning) is equal to the *sum* of **Max. Resource Group Processing Time** and **Max. Event-only Duration**.
Total time to process a Resource Group Event before a warning is displayed	The total time for the Cluster Manager to wait before running the **config_too_long** script. The default is 6 minutes and 0 seconds. This field is the sum of the two other fields and is not editable.

4. Press Enter to change the field values. PowerHA® SystemMirror® changes these values in the PowerHA® SystemMirror® Configuration Database.
5. Synchronize the cluster to propagate the data to the other cluster nodes. PowerHA® SystemMirror® uses the specified total event duration time before issuing **config_too_long** warning messages.

Configuring a custom remote notification method

These topics describe how to configure custom remote notification methods to respond to an event, how cluster verification confirms the remote notification configuration, and how node failure affects the remote notification method.

You can configure a remote notification method through SMIT to issue a customized numeric or alphanumeric page in response to a specified cluster event. You can also send SMS text message notifications to any address, including a cell phone SMS address or mail to an email address. The pager message is sent through the attached dialer modem. Cell phone text messages are sent through email using the TCP/IP connection or an attached GSM wireless modem.

You can send the following custom remote notifications:

- Numeric and alphanumeric page
- SMS text message to any address including a cell phone or mail to an email address.
- SMS text message using a GSM modem to transmit the notification through a wireless connection.

The PowerHA® SystemMirror® remote notification functionality requirements follow:

- A tty port used for paging cannot also be used for heartbeat traffic.
- Any tty port specified must be defined to AIX® and must be available.
- Each node that might send a page or text messages must have an appropriate modem installed and enabled.

  Note: PowerHA® SystemMirror® checks the availability of the tty port when the notification method is configured and before a page is issued. Modem status is not checked.
  To send an SMS text message over the dialer modem, your pager provider must offer this service.

- Each node that might send email messages from the SMIT panel using the AIX® operating system mail must have a TCP/IP connection to the Internet.
- Each node that might send text messages to a cell phone must have an appropriate Hayes-compatible dialer modem installed and enabled.
- Each node that might transmit an SMS message wirelessly must have a Falcom-compatible GSM modem installed in the RS232 port with the password disabled. Ensure that the modem connects to the cell phone system.

Creating a remote notification message file

Before you can issue a message to a pager or cell phone, you must create a file that contains the message text. PowerHA® SystemMirror® provides a template to help you create this file. The template contains default text and instructions for an alphanumeric page or cell phone message.

The template is in:

```
/usr/es/sbin/cluster/samples/pager/sample.txt
```

By default, the message contains the following information: the event, the node on which it occurred, the time and date, and the name of the object affected by the event. This default message is sent if no message file is found at the time a custom alphanumeric page or cell phone message is triggered.

For numeric pages, the provided sample text is not appropriate; your numeric page file should contain only digits. If no message file is found when a numeric page is triggered, the default message sent is "888."

The sample.txt file contains comments that relate to an alphanumeric pager or cell phone message. A numeric page does not use this file. Shown here is the sample.txt file; there is no need to alter the file unless you want to add additional recipients.

> **Note:** Save the sample.txt file with a new name before modifying it. However, if you do alter the file when you migrate to a new version of PowerHA® SystemMirror®, the customized file is preserved, even though a new default sample.txt file is installed.
> Place a separate copy of each message file on each node listed in the notification method definition. PowerHA® SystemMirror® does *not* automatically distribute this file to other nodes during cluster synchronization.

The following lists the contents of the sample.txt file:

```
# sample file for alphanumeric paging
# you can use the following notations in your message
# %d - current time&date
# %n - node that sends the message
# %e - eventname
# '#' is used to comment the line
# for example "Node %n: Event %e occured at %d"
# if nodename=bazilio, event=node_up
# and current date&time = Thu Sep 28 19:41:25 CDT 2006
# will result in sending the message
# "Node bazilio: Event node_up occured at Thu Sep 28 19:41:25 CDT 2006"
```

Defining a remote notification method

You can define a remote notification method using the SMIT interface.

When you synchronize or perform cluster verification, PowerHA® SystemMirror® checks the configuration of your remote notification method.

PowerHA® SystemMirror® sends an error message if the specified pager or cell phone message file is missing from the node. The message can still be sent, but it contains the text that is supplied in the original `sample.txt` file.

To define a remote notification method, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Remote Notification Methods › Add a Custom Remote Notification Method** and press Enter.
3. Fill in field values as follows:

Add a Custom Remote Notification Method fields	
Field	**Value**
Method Name	Assign a name to the notification method. This could also indicate who would get the message.
Description	Add a description, if desired, of the notification method.
Nodename(s)	Enter the name(s) of one or more nodes that you want to issue this or cell phone message. Press F4 to get a list of node names. Each node must have been defined previously in the **Define Port/Node Pairs** SMIT panel. Separate multiple nodes with a space. The sequence of nodes in this SMIT field determines their priority for sending pages or cell phone messages. See Remote notification and node failure for more information about node priority for remote notification.
Number to Dial or Cell Phone Address	Indicate the telephone number to dial to reach the pager or the address of the cell phone. The number-to-dial string can contain any characters or sequences supported by a Hayes-compatible modem using the standard Telocator Alphanumeric Protocol (TAP) - your provider must support this service.

4. Depending on the type of pager, you will need to enter either the number of the pager alone, or the number of the paging company followed by the pager number:
   If you are using a numeric pager, use the form: **18007650102,,,,** The commas create pauses in the dialing sequence. The trailing commas are required because there is always some delay between dialing and the actual sending of the page.

   If the pager is alphanumeric the input should take the form: **180007654321;2119999** where `18007654321` is the paging company number and `2119999` is the actual pager number.

5. For cell phone text messaging using email, enter the address of the cell phone. This is in the format: `phone_number@provider_address`. Consult your provider for the specific provider_address format. It may look like `180007654321@provider.net`. Multiple space-separated addresses can be used. Test this by sending an email. To send email to multiple addresses, separate the addresses using a space.

6. You can send a text message wirelessly to a cell phone, if a GSM modem is used instead of the dialer modem. The format is `<cell phone number>#`. For example, it may look like `7564321#`. The SIM providers may support international calls.

> You can define a remote notification method for shutdown system event.

Phone number fields	
Field	**Value**
Filename	Specify the path of the text file containing the pager message or cell phone message. **Note:** Make sure the path refers to the correct location of the message file on each node specified in the **Node Name(s)** field.
Cluster Event(s)	Specify the event(s) that activate this notification method. Press F4 to get a list of event names. Separate multiple events with a space. This list also includes a system event, **aix_shutdown**.

Field	Value
Retry Counter	Specify how many times to reissue the page or cell phone message if it fails. The default is **3** times.
TIMEOUT	Specify how many seconds to wait before considering a page or cell phone message attempt failed. The default is **45** seconds.

7. When you finish entering values in all fields, press Enter.
8. Synchronize the cluster to propagate the configuration information to the other nodes.

> The configuration information can be entered on one node and propagated to the others during synchronization, but you must manually make sure that the correct page or cell phone message text files reside in the correct locations on each node in the nodelist.

Sending a test remote notification message

You can send a test page or cell phone message to make sure everything is configured correctly, and that the expected notification will be issued for a given event, just as if the event actually occurred.

Before sending the test remote message, you must have a notification method already configured. The test remote message must be sent from a node that is configured for the selected method.

To configure a remote notification message:

1. From the **Configure Custom Remote Notification Method** menu, select **Send a Test Remote Message.**
2. Select a remote notification method to use for the test.
3. In the **Send a Test Remote Message** panel, fill in field values as follows:

Send a Test Remote Message fields	
Field	**Value**
Method Name	The configured method that you selected for the test page.
Event name	Press F4 to get a picklist of events configured for the selected method, and select the event for which you would like to send a test page.

4. Press Enter. The Command Status window then reports the remote message was successful, or errors occurred.

Result
The test remote message will be the message file you specified when you configured the notification method. If an object name is included, for the test remote message, it will appear as a pseudo-name such as node_1, adapter_1, and network_1. If the message file cannot be located, a default message will be sent to the pager or cell phone and an error will be displayed. For alphanumeric pages or cell phone messages, the default message is the sample text; for numeric pages, the default message is "888."

Remote notification and node failure

If a node fails and triggers a page or cell phone message, the remote notification is sent from the node with the next highest priority.

(A node's order of priority is determined by the order in which you listed node names when you defined the method.) If the next highest priority node is up but unable to send the remote notification for some other reason (for instance, the modem is not connected), the system attempts to resend the remote notification message for the number of times specified in the **Retry Counter**. If the remote notification still cannot be sent, it fails. The remote notification is *not* passed on to be issued from another node.

Changing or removing a custom remote notification method

You can change or remove a notification method through SMIT to issue a customized remote notification in response to a specified cluster event.

Changing a remote notification method

This topic discusses changing the configuration of a custom remote notification method.

To change the configuration of a custom remote notification method:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Remote Notification Methods › Change/Show Custom Remote Notification Method** and press Enter.
   You can also reach the **Configure Remote Notification Methods** panel by typing `smit cm_def_cus_pager_menu`.
3. Select the method you want to change.
4. Make your changes.
5. Press Enter.

Deleting a remote notification method

You can delete a custom remote notification method.

To delete a custom remote notification method:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Remote Notification Methods › Remove Custom Remote Notification Method** and press Enter.
3. Specify the name of the method you want to delete.
4. Press Enter to delete the method.

Verifying and synchronizing a PowerHA® SystemMirror® cluster

Verifying and synchronizing your PowerHA® SystemMirror® cluster assures you that all resources used by PowerHA® SystemMirror® are configured appropriately and that rules regarding resource ownership and resource takeover are in agreement across all nodes. You should verify and synchronize your cluster configuration after making any change within a cluster. For example, any change to the hardware operating system, node configuration, or cluster configuration.

Whenever you configure, reconfigure, or update a cluster, run the cluster verification procedure to ensure that all nodes agree on the cluster topology, network configuration, and the ownership and takeover of PowerHA® SystemMirror® resources. If the verification succeeds, the configuration can be synchronized. Synchronization takes effect immediately on an active cluster. A dynamic reconfiguration event is run and the changes are committed to the active cluster.

Note: If you are using the standard configuration paths, synchronization automatically follows a successful verification. If you are using the **Custom Cluster Configuration** path, you have more options for types of verification. If you are using the **Problem Determination Tools** path, you can choose whether to synchronize or not.

The messages output from verification indicate where the error occurred (for example, the node, device, or command). The utility uses verbose logging to write to the **/var/hacmp/clverify/clverify.log** file.

> **Note:** Verification is not supported on a mixed-version PowerHA® SystemMirror® cluster.

Error conditions result when information is not properly configured on all cluster nodes. This information may be important for the operation of PowerHA® SystemMirror®, but not part of the PowerHA® SystemMirror® software itself; for example, AIX® volumes do not exist in the cluster. In some of these situations, you can authorize a

corrective action before verification continues. When verification detects certain conditions, such as mismatched PowerHA® SystemMirror® shared volume group time stamps or a node is missing required entries in **/etc/services**, PowerHA® SystemMirror® fixes the problem.

On the node where you run the utility, detailed information is collected into log files, which contain a record of all data collected and the tasks performed.

You can add your own custom verification methods to ensure that specific components within your cluster are properly configured. You can change or remove these methods from the verification process depending on the level of cluster verification you want.

> **Note:** Verification requires 4 MB of disk space in the **/var** file system in order to run; 18 MB of disk space is recommended for a four-node cluster. Typically, the **/var/hacmp/clverify/clverify.log** files require 1-2 MB of disk space.

Running cluster verification

After making a change to the cluster, you can perform cluster verification in several ways.

These methods include:

- **Automatic verification**. You can automatically verify your cluster:
  - Each time you start cluster services on a node
  - Each time a node rejoins the cluster
  - Every 24 hours.

  By default, automatic verification is enabled to run at midnight.

  For detailed instructions, see Automatic verification and synchronization.
- **Manual verification**. Using the SMIT interface, you can either verify the complete configuration, or only the changes made since the last time the utility was run.
  Typically, you should run verification whenever you add or change anything in your cluster configuration.
  For detailed instructions, see Verifying the PowerHA® SystemMirror® configuration using SMIT.

Automatic verification and synchronization

During *automatic verification and synchronization*, PowerHA® SystemMirror® discovers and corrects several common configuration issues prior to starting cluster services.

This automatic behavior ensures that if you had not manually verified and synchronized your cluster prior to starting cluster services, PowerHA® SystemMirror® will do so. Throughout this section, automatic verification and synchronization is often simply referred to as *verification*.

Understanding the PowerHA® SystemMirror® cluster verification process

By default, verification runs automatically without any configuration required.

Verification occurs on both active and inactive clusters. In order for automatic verification to work, more than one node must exist in the cluster, since PowerHA® SystemMirror® compares the configuration of one node against the configuration of another node.

Verification ensures an error-free cluster startup and poses a negligible impact on performance, which is directly related to the number of nodes, volume groups, and file systems in the cluster configuration.

The phases of the verification and synchronization process are as follows:

1. Verification
2. Snapshot (optional)
3. Synchronization.

For details on these phases, see Understanding the detailed phases of verification. After verification, cluster services start.

Cluster verification during a dynamic cluster reconfiguration event

If a node is down during a dynamic reconfiguration event and later it attempts to join the cluster, cluster verification and synchronization run prior to starting services on the joining node, and the joining node receives its configuration updates from an active cluster node.

If verification fails on the joining node, the node will *not* start cluster services. Likewise, if a node is dynamically removed from the active cluster, the node will *not* be allowed to join the cluster or cluster services.

Parameters automatically corrected

Automatic verification and synchronization ensure that typical configuration inconsistencies are automatically corrected.

These types of inconsistencies are automatically corrected:

- IP addresses (that RSCT expects) are configured on the network interfaces.
- Shared volume groups are *not* set to automatically varyon.
- File systems are *not* set to automatically mount.
- Logical Volumes which are not present in other nodes

Verifying service IP address aliases

At cluster startup, RSCT expects the IP address label to be defined on the interfaces with the same value that has been defined in the PowerHA® SystemMirror® configuration database. The PowerHA® SystemMirror® automatic verification and synchronization process ensures nodes *not* currently running cluster services are verified and corrected; nodes currently running cluster services are *not* automatically corrected.

> **Note:** Only aliased IP interfaces that are used by PowerHA® SystemMirror® are verified and corrected.

If a node has an interface that is *not* defined as it appears in the PowerHA® SystemMirror® configuration database, automatic verification detects this and issues an error message.

Verifying shared volume groups

Shared volume groups that are configured as part of a PowerHA® SystemMirror® resource group must have their automatic varyon attribute set to No. If the verification phase determines that the automatic varyon attribute is set to Yes, verification notifies you about nodes on which the error occurs and prompts you to correct the situation.

Verifying file systems

Any file systems participating in a resource group with AIX® attributes that allow the file system to be automatically mounted at system restart will raise errors. This includes standard journaled file systems (JFS) and enhanced journaled file systems (JFS2). If the file system has been set to mount automatically at boot time, verification displays an error.

Verifying Logical Volume

Any Logical Volume participating in a resource group with AIX® attributes that allow the Logical Volume to be automatically sync to other nodes.

> **Note:** User must run verify and sync from the node where logical volume is present.

Understanding the detailed phases of verification

After verification, cluster services start up. If cluster services do not start, it is because PowerHA® SystemMirror® has discovered errors. You can resolve these errors by correcting inconsistencies.

Phase one: Verification

During the verification process the default system configuration directory (DCD) is compared with the active configuration. On an inactive cluster node, the verification process compares the local DCD across all nodes. On an active cluster node, verification propagates a copy of the active configuration to the joining nodes.

If a node that was once previously synchronized has a DCD that does not match the ACD of an already active cluster node, the ACD of an active node is propagated to the joining node. This new information does not replace the DCD of the joining nodes; it is stored in a temporary directory for the purpose of running verification against it.

PowerHA® SystemMirror® displays progress indicators as the verification is performed.

> **Note:** When you attempt to start a node that has an invalid cluster configuration, PowerHA® SystemMirror® transfers a valid configuration database data structure to it, which may consume 1-2 MB of disk space.

If the verification phase fails, cluster services will not start.

Phase two: (Optional) Snapshot

A snapshot is only taken if a node request to start requires an updated configuration. During the snapshot phase of verification, PowerHA® SystemMirror® records the current cluster configuration to a snapshot file - for backup purposes. PowerHA® SystemMirror® names this snapshot file according to the date of the snapshot and the name of the cluster. Only one snapshot is created per day. If a snapshot file exists and its filename contains the current date, it will not be overwritten.

This snapshot is written to the **/usr/es/sbin/cluster/snapshots/** directory.

The snapshot filename uses the syntax `MM-DD-YYYY-`*`ClusterName`*` -autosnap.odm`. For example, a snapshot taken on April 2, 2006 on a cluster hacluster01 would be named `usr/es/sbin/cluster/snapshots/04-02-06hacluster01-autosnap.odm`.

Phase three: Synchronization

During the synchronization phase of verification, PowerHA® SystemMirror® propagates information to all cluster nodes. For an inactive cluster node, the DCD is propagated to the DCD of the other nodes. For an active cluster node, the ACD is propagated to the DCD.

If the process succeeds, all nodes are synchronized and cluster services start. If synchronization fails, cluster services do not start and PowerHA® SystemMirror® issues an error.

Monitoring verification and resolving configuration inconsistencies

You can monitor the automatic verification and synchronization progress as it occurs by tracking messages as they appear on the SMIT console.

In addition, you can examine any prior processes by reviewing the **smit.log** file or **/var/hacmp/clverify/clverify.log** file.

Verification completion

When cluster verification completes on the selected cluster node, this node supplies the information to the other cluster nodes.

This information includes:

- Name of the node where verification had been run

- Date and time of the last verification
- Results of the verification.

This information is stored on every available cluster node in the **/var/hacmp/clverify/clverify.log** file. If the selected node became unavailable or could *not* complete cluster verification, you can detect this by the lack of a report in the /**var/hacmp/clverify/clverify.log** file. If the log file does *not* indicate a specific node, then the error applies to all nodes and cluster services do *not* start.

If cluster verification completes and detects some configuration errors, you are notified about the potential problems:

- The exit status of verification is published across the cluster along with the information about cluster verification process completion.
- Broadcast messages are sent across the cluster and displayed on the console. These messages inform you of any detected configuration errors.
- A **cluster_notify** event runs on the cluster and is logged in **hacmp.out** (if cluster services are running).
- Information about the node where you ran the cluster verification is written to the **/var/hacmp/clverify/clverify.log** file. If a failure occurs during processing, error messages and warnings indicate the node affected and reasons for the verification failure.
- A configuration snapshot is written to the /usr/es/sbin/cluster/snapshots/ directory.

Ongoing automatic verification

Once a valid configuration is defined, the verification process runs once every 24 hours.

By default, the first node in alphabetical order runs the verification at midnight; however, you can change these defaults by selecting a node and a time that suits your needs. If the selected node is unavailable (powered off), automatic verification does not run.

Verifying the PowerHA® SystemMirror® configuration using SMIT

After reconfiguring or updating a cluster, run the cluster verification procedure.

> **Note:** If you are investigating a problem with the cluster and want to run verification procedures without synchronizing the cluster, use the cluster verification SMIT panels found under the **Problem Determination Tools** menu.

Verifying and synchronizing a cluster configuration

Verification occurs at different points in the configuration process. Verification is done automatically during the start of cluster services and every 24 hours.

Verification is done before synchronization to ensure the validity of the configuration. Synchronization is required whenever a change is made to the cluster definition.

Verifying the topology configuration

Verification ensures that all nodes agree on the topology of the cluster. For example, it checks for invalid characters in cluster names, node names, network names, network interface names, and resource group names. It checks to ensure that interfaces are properly configured, nodes are reachable, and networks have the required number of interfaces.

It also checks for the reserved words used as cluster names, node names, network names, network interface names and resource group names. These names are listed in the **/usr/es/sbin/cluster/etc/reserved_words** file.

Verifying the network configuration

Verification ensures that the networks are configured correctly and that all nodes agree on the ownership of all defined resources, such as the following:

- Configuration of network information, such as addresses on all nodes in the cluster.
- No network interfaces configured on unsupported network types (for example, IP, socc, slip and fcs).

Verifying disk and file system configuration

Verification ensures that disks and file systems are in agreement and configured according to the following:

- Agreement among all nodes on the ownership of defined resources (for example, file systems, volume groups, disks, and application controllers). The verification utility checks for the existence and defined ownership of the file systems to be taken over, and then checks the volume group and disks where the file systems reside.
- Agreement among nodes on the major and minor device numbers for NFS-exported file systems.
- If disk fencing is enabled, verification sends an error if all nodes are not included in the concurrent access resource group.

Verifying resource group information

Verification ensures that the resource group information supplied is in agreement and configured according to the following:

- Verification issues warnings in cases when the startup, fallover or fallback preferences that you choose for resource groups may put the high availability of resources at risk in the case of a cluster failure.
- The verification utility checks that the choices for distribution of resources in case of a takeover (node priorities) so that the takeover information matches the owned resources information.

Verifying individual resources

Verification checks individual resources, such as the following:

- Event customization.
- Application controller start and stop scripts exists and that they are executable.

Verifying automatic error notification methods

Verification ensures that automatic error notification (AEN) methods exist and are properly configured for the following:

- Root volume groups
- PowerHA® SystemMirror®-defined volume groups or PowerHA® SystemMirror®-defined disks
- PowerHA® SystemMirror®-defined file systems (the underlying disks that support the file system)

Verifying custom configurations

The verification function checks for the existence and consistency of any configured custom cluster snapshot methods

Verifying PowerHA® SystemMirror® Enterprise Edition configurations

If you are using PowerHA® SystemMirror® Enterprise Edition configurations, verification confirms that the PowerHA® SystemMirror® Enterprise Edition cluster configuration for sites and its replicated resources are consistent with your PowerHA® SystemMirror® cluster configuration.

Verifying service IP labels

If a service IP label is configured on the interface instead of the boot label, verification issues an error reminding you to run the sample utility **clchipdev** before starting cluster services. If that service IP label is an alias, verification has a correct action to reverse it. The sample utility **clchipdev** helps configure the application service interface correctly in PowerHA® SystemMirror®.

Verifying the cluster using the standard configuration paths

If you use the standard configuration paths, when you select the option **Verify and Synchronize Cluster Configuration**, the command executes immediately. Messages appear in the SMIT command status screen as the configuration is checked.

Verifying the cluster using the Custom Cluster Configuration path

If you use the **Custom Cluster Configuration** path, you can set parameters for the command before it runs. These parameters differ depending on whether or not the cluster is active.

To verify and synchronize the PowerHA® SystemMirror® cluster configuration:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Verify and Synchronize Cluster Configuration (Advanced)** and press Enter.
   The software checks whether cluster services are running on any cluster node and displays one of the following screens:

   If the cluster is active, the following options appear.

Verify and Synchronize Cluster Configuration (Advanced) fields	
Field	**Value**
Verify changes only?	**No** is the default. (Run the full check on resource and topology configuration.) Select **Yes** to verify only resource or topology configurations that have changed since the last time the cluster was verified. If you have changed the AIX® configuration, do not use this mode; it only applies to PowerHA® SystemMirror® configuration changes.
Logging	**Standard** is the default. You can also select **Verbose**. Verification messages are logged to **/var/hacmp/clverify/clverify.log**.

   If the cluster is inactive, the following options appear:

Inactive cluster fields	
Field	**Value**
Verify Synchronize or Both	**Both** is the default. You can also select **Verify** only or **Synchronize** only.
Automatically correct errors found during verification?	**No** is the default. PowerHA® SystemMirror® will not perform corrective actions. If you select **Interactively**, during verification you will be prompted when it finds a problem it can correct, for example: • Importing a volume group • Re-importing shared volume groups (mount points and file systems issues).

Field	Value
	You then choose to have the action taken or not. For more information, see Conditions that can trigger a corrective action.
Force synchronization if verification fails?	**No** is the default. If you select **Yes**, cluster verification runs but verification errors are ignored and the cluster is synchronized. Use the **Yes** option with caution. Correct functioning of the cluster at runtime cannot be guaranteed if you synchronize without verification. Cluster topology errors may lead to an abnormal exit of the Cluster Manager. Resource configuration errors may lead to resource group acquisition errors.
Verify changes only?	**No** is the default. (Run the full check on resource and topology configuration.) **Yes** opts to verify only resource or topology configurations that have changed since the last time the cluster was verified. If you have changed the AIX® configuration, do not use this mode; it only applies to PowerHA® SystemMirror® configuration changes.
Logging	**Standard** is the default. You can also select **Verbose**. All verification messages (including Verbose messages) are logged to **/var/hacmp/clverify/clverify.log**.

3. Press Enter and SMIT starts the verification process. The verification output appears in the SMIT Command Status window.
4. If any error messages appear, make the necessary changes and run the verification procedure again. You may see Warnings if the configuration has a limitation on its availability; for example, only one interface per node per network is configured, or Workload Manager is configured but there is no application controller assigned to use it.

Running corrective actions during verification

You can run automatic corrective actions during cluster verification on an inactive cluster. By default, automatic corrective action is enabled for the standard configuration paths and disabled for custom configuration path.

Automatic corrective actions can be disabled for the advanced verification dialog (from the **System Management (C-SPOC) › Manage Services** menu) but it cannot be disabled for the standard paths. You can run verification with corrective actions in one of two modes:

- *Interactively*. If you select **Interactively**, when verification detects a correctable condition related to importing a volume group or to re-importing mount points and file systems, you are prompted to authorize a corrective action before verification continues.
- *Automatically*. If you select **Yes**, when verification detects that any of the error conditions exists, it takes the corrective action automatically without a prompt.

If an error discovered during verification has a corrective action, the item is corrected and the run continues. For situations when the correction involves importing a shared volume group, re-importing a shared volume group, or updating the **/etc/hosts** file, the utility runs all verification checks again after it corrects one of the previous conditions. If the same error condition is triggered again, the associated corrective action is not executed. The error is logged and verification fails. If the original condition is a warning, verification succeeds.

PowerHA® SystemMirror® detects active service IP labels and active volume groups on nodes regardless of whether or not nodes are running cluster services. PowerHA® SystemMirror® looks at the resource group state instead of cluster services. When verification detects active resources on a node that does not have the resource group in ONLINE state or does not have the resource group in an UNMANAGED state, verification gives you the option to bring these resources OFFLINE according the following table.

Verification cannot tell which node will actually acquire the resource group that has active resources. Thus, the warning messages mentioned in the following table are printed every time active resources are found on a node

that is or is not stopped and the state of the resource group to which active resources belong is UNMANAGED, OFFLINE, or ERROR.

	Resource Group Attribute: Manage Resource Group Automatically	**Resource Group Attribute: Manage Resource Group Manually**
Interactively correct errors	Display message with option to bring resources offline.	Display message with option to bring resources offline.
Automatically correct errors	Reset the startup attribute Managed Resource group to: Manually. Display warning message.	Print reminder/warning and steps to take.
No corrective actions	Print reminder/warning and steps to take	Print reminder/warning and steps to take
Cluster Services are running	N/A	N/A

Conditions that can trigger a corrective action

This topic discusses conditions that can trigger a corrective action.

PowerHA® SystemMirror® shared volume group time stamps are not up-to-date on a node

If the shared volume group time stamp file does *not* exist on a node, or the time stamp files do not match on all nodes, the corrective action ensures that all nodes have the latest up-to-date VGDA time stamp for the volume group and imports the volume group on all cluster nodes where the shared volume group was out of sync with the latest volume group changes. The corrective action ensures that volume groups whose definitions have changed will be properly imported on a node that does *not* have the latest definition.

The /etc/hosts file on a node does not contain all PowerHA® SystemMirror®-managed IP addresses

If an IP label is missing, the corrective action modifies the file to add the entry and saves a copy of the old version to **/etc/hosts.date**. If a backup file already exists for that day, no additional backups are made for that day.

Verification does the following:

- If the **/etc/hosts** entry exists but is commented out, verification adds a new entry; comment lines are ignored.
- If the label specified in the PowerHA® SystemMirror® Configuration does not exist in **/etc/hosts** , but the IP address is defined in **/etc/hosts,** the label is added to the existing **/etc/hosts** entry. If the label is different between **/etc/hosts** and the PowerHA® SystemMirror® configuration, then verification reports a different error message; no corrective action is taken.
- If the entry does not exist, meaning both the IP address and the label are missing from **/etc/hosts,** then the entry is added. This corrective action takes place on a node-by-node basis. If different nodes report different IP labels for the same IP address, verification catches these cases and reports an error. However, this error is unrelated to this corrective action. Inconsistent definitions of an IP label defined to PowerHA® SystemMirror® are not corrected.

A file system is not created on a node, although disks are available

If a file system has *not* been created on one of the cluster nodes, but the volume group is available, the corrective action creates the mount point and file system. The file system must be part of a resource group for this action to take place. In addition, the following conditions must be met:

- This is a shared volume group.
- The volume group must already exist on at least one node.

- One or more node(s) that participate in the resource group where the file system is defined must already have the file system created.
- The file system must already exist within the logical volume on the volume group in such a way that simply re-importing that volume group would acquire the necessary file system information.
- The mount point directory must already exist on the node where the file system does not exist.

The corrective action handles only those mount points that are on a shared volume group, such that exporting and re-importing of the volume group will acquire the missing file systems available on that volume group. The volume group is varied off on the remote node(s), or the cluster is down and the volume group is then varied off if it is currently varied on, prior to executing this corrective action.

If **Mount All File Systems** is specified in the resource group, the node with the latest time stamp is used to compare the list of file systems that exists on that node with other nodes in the cluster. If any node is missing a file system, then PowerHA® SystemMirror® imports the file system.

Disks are available, but the volume group has not been imported to a node

If the disks are available but the volume group has not been imported to a node that participates in a resource group where the volume group is defined, then the corrective action imports the volume group.

The corrective action gets the information regarding the disks and the volume group major number from a node that already has the volume group available. If the major number is unavailable on a node, the next available number is used.

The corrective action is only performed under the following conditions:

- The cluster is down.
- The volume group is varied off if it is currently varied on.
- The volume group is defined as a resource in a resource group.
- The major number and associated PVIDS for the disks can be acquired from a cluster node that participates in the resource group where the volume group is defined.

Note: This functionality will *not* turn off the **auto varyon** flag if the volume group has the attribute set. A separate corrective action handles auto varyon.

Shared volume groups configured as part of a PowerHA® SystemMirror® resource group have their automatic varyon attribute set to Yes.

If verification finds that a shared volume group inadvertently has the auto varyon attribute set to Yes on any node, the corrective action automatically sets the attribute to **No** on that node.

Required /etc/services entries are missing on a node.

If a required entry is commented out, missing, or invalid in **/etc/services** on a node, the corrective action adds it. Required entries are:

Name	Port	Protocol
clcomd_caa	16191	tcp
clinfo_client	6174	tcp
clinfo_deadman	6176	tcp
clsmuxpd	6270	tcp
clm_smux	6175	tcp
NULL	0	NULL

Required PowerHA® SystemMirror® snmpd entries are missing on a node

If a required entry is commented out, missing, or invalid on a node, the corrective action adds it.

Note: The default version of the **snmpd.conf** file for AIX® is **snmpdv3.conf**.

In **/etc/snmpdv3.conf** or **/etc/snmpd.conf**, the required PowerHA® SystemMirror® snmpd entry is:

```
smux   1.3.6.1.4.1.2.3.1.2.1.5   clsmuxpd_password # PowerHA SystemMirror/ES for AIX clsmuxpd
```

In **/etc snmpd.peers**, the required PowerHA® SystemMirror® snmpd entry is:

```
clsmuxpd   1.3.6.1.4.1.2.3.1.2.1.5 "clsmuxpd_password" # PowerHA SystemMirror/ES for AIX clsmuxpd
```

If changes are required to the **/etc/snmpd.peers** or **snmpd[v3].conf** file, PowerHA® SystemMirror® creates a backup of the original file. A copy of the pre-existing version is saved prior to making modifications in the file **/etc/snmpd.**{**peers | conf**}**.date**. If a backup has already been made of the original file, then no additional backups are made.

PowerHA® SystemMirror® makes one backup per day for each **snmpd** configuration file. As a result, running verification a number of times in one day only produces one backup file for each file modified. If no configuration files are changed, PowerHA® SystemMirror® does not make a backup.

Required PowerHA® SystemMirror® network options setting

The corrective action ensures that the value of each of the following network options is consistent across all nodes in a running cluster (out-of-sync setting on any node is corrected):

- **tcp_pmtu_discover**
- **udp_pmtu_discover**
- **ipignoreredirects**

Required routerevalidate network option setting

Changing hardware and IP addresses within PowerHA® SystemMirror® changes and deletes routes. Because AIX® caches routes, setting the **routerevalidate** network option is required as follows:

```
no -o routerevalidate=1
```

This setting ensures the maintenance of communication between cluster nodes. Verification run with corrective action automatically adjusts this setting for nodes in a running cluster.

Note: No corrective actions take place during a dynamic reconfiguration event.

Corrective actions when using IPv6

If you configure an IPv6 address, the verification process can perform 2 more corrective actions:

- **Neighbor discovery (ND)**. Network interfaces must support this protocol which is specific to IPv6. The underlying network interface card is checked for compatibility with ND and the ND related daemons will be started.
- **Configuration of Link Local addresses (LL)**. A special link local (LL) address is required for every network interface that will be used with IPv6 addresses. If a LL address is not present the `autoconf6` program will be run to configure one.

clverify.log file

During verification, PowerHA® SystemMirror® collects configuration data from all the nodes as it runs through a series of checks.

The verbose output is saved to the **/var/hacmp/clverify/clverify.log** file. The log file is rotated; this helps you and IBM® Support obtain a history of what configuration changes have been made when you need to determine the root cause of a problem.

Ten copies of the log are saved, as follows:

```
drwxr-xr-x   3 root     system1024 Mar 13 00:02 .
drwxr-xr-x   6 root     system 512 Mar 11 10:03 ..
-rw-------   1 root     system165229 Mar 13 00:02 clverify.log
-rw-------   1 root     system165261 Mar 12 17:31 clverify.log.1
-rw-------   1 root     system165515 Mar 12 15:22 clverify.log.2
-rw-------   1 root     system163883 Mar 12 15:04 clverify.log.3
-rw-------   1 root     system164781 Mar 12 14:54 clverify.log.4
-rw-------   1 root     system164459 Mar 12 14:36 clverify.log.5
-rw-------   1 root     system160194 Mar 12 09:27 clverify.log.6
-rw-------   1 root     system160410 Mar 12 09:20 clverify.log.7
-rw-------   1 root     system160427 Mar 12 09:16 clverify.log.8
-rw-------   1 root     system160211 Mar 12 09:06 clverify.log.9
```

You can redirect the **clverify.log** file to write to a different location using the standard PowerHA® SystemMirror® logfile redirection mechanism. If the **clverify.log** file is redirected to a different location, the location of all the data saved in the subdirectories in the path **/var/hacmp/clverify** moves along with it. However, pre-existing data under **/var/hacmp/clverify** is not automatically moved if the **clverify.log** is redirected.

Archived configuration databases

All verification checks use PowerHA® SystemMirror® Configuration Database data supplied by the common communication infrastructure, which is designed to provide efficient access to configuration databases from the other nodes.

When the verification runs, it stores copies of the following:

- All PowerHA® SystemMirror® Configuration Databases (ODMs) used during verification
- All AIX® ODMs (Custom Attributes, Device Definitions, and so forth) collected from the remote nodes

The verification utility manages these files by storing the copies in various directories depending on the success or failure of the verification.

Inactive components report

Cluster verification reports the list of inactive cluster components that could be in that state due to errors or failures. The information in this report is valid for active clusters only.

Inactive components are:

- Node that is not running Cluster Services. Resources for such nodes are not listed.
- Interfaces in "down" state
- Networks in "down" state
- resource groups in "ERROR" or "UNMANAGED" state

In the following example, one node has experienced a failure in a boot interface on a network that does not allow aliasing. The other, which was managing two resource groups, has been shutdown unmanaged and restarted with manual resource group management.

```
Node: node1
  Resource Group: rg1                                  State: UNMANAGED
  Resource Group: rg4                                  State: UNMANAGED
Node: node2
  Network: net_ether_01
    Label: node2_stby            Address: 198.168.20.21  State: DOWN
  Resource Group: rg1                                  State: UNMANAGED
  Resource Group: rg4                                  State: UNMANAGED
```

Note:

- Boot interfaces are only displayed for networks that are using aliasing.
- The unmanaged resource groups appear with the **UNMANAGED** status on every node that could potentially host them.

Managing PowerHA® SystemMirror® file collections

PowerHA® SystemMirror® requires that event scripts, application scripts, AIX® files, and PowerHA® SystemMirror® configuration files must be identical on each cluster node.

The PowerHA® SystemMirror® File Collections facility automatically synchronizes these files among cluster nodes and warns you if there are any unexpected results (for example, if one or more files in a collection has been deleted or has a length of zero on one or more cluster nodes).

PowerHA® SystemMirror® has always provided the facilities to keep its own configuration information in sync across the cluster. The value of the PowerHA® SystemMirror® file collections function is that it allows you to easily keep application-specific configuration information in sync on all the nodes in the cluster. It is known that when the application-specific configuration information is allowed to drift out of sync in the cluster - when a change made one node is *not* made on others - this creates problems during application fallovers. You can use file collections to keep your application configuration identical on all cluster nodes.

Default PowerHA® SystemMirror® file collections

When you install PowerHA® SystemMirror®, it sets up the default file collections.

PowerHA® SystemMirror® Configuration_Files collection

PowerHA® SystemMirror® Configuration_Files collection is a container for certain essential system files.

The PowerHA® SystemMirror® Configuration_Files collection contains the following files:

- /etc/hosts
- /etc/services
- /etc/snmpd.conf
- /etc/snmpdv3.conf
- /etc/rc.net
- /etc/inetd.conf
- /etc/cluster/rhosts
- /usr/es/sbin/cluster/etc/clhosts
- /usr/es/sbin/cluster/etc/clinfo.rc
- /usr/es/sbin/cluster/netmon.cf

HACMP_Files collection

HACMP_Files is a container for user-configurable files in the PowerHA® SystemMirror® configuration. PowerHA® SystemMirror® uses this file collection to reference all of the user-configurable files in the PowerHA® SystemMirror® Configuration Database classes.

The **HACMP_Files** collection automatically includes:

- Any pre, post or notification events you have used to customize cluster events.
- The start and stop scripts specified for any application controllers.
- Scripts specified for application monitoring including any monitoring, notification, cleanup and restart scripts.
- Custom pager text messages.

- Scripts for Tape support
- Any custom snapshot methods
- User defined event recovery programs

> **Note:** Do not modify or rename the PowerHA® SystemMirror® event script files. Also, do not include PowerHA® SystemMirror® event scripts in any PowerHA® SystemMirror® file collection.
> When copying a file to a remote node, the local node's owner, group, modification time stamp, and permission settings are maintained on the remote node. That is, the remote node inherits these settings from the local node.

Permissions for all files in the **HACMP_Files** collection are set to execute, which helps to prevent problems if you have not yet set execute permission for scripts on all nodes. (This is often the cause of an event failure.)

You cannot rename or delete the **HACMP_Files** collection. You cannot add or remove files from the collection.

You can add a file that is already included in the **HACMP_Files** collection (for example, an application start script) to another file collection. However, in any other case, a file can only be included in one file collection and you receive the following error message, where XXX _Files is the name of the previously defined collection:

```
This file is already included in the <XXX_Files> collection).
```

You can add and remove files or delete the **Configuration_Files** collection.

Neither of these file collections is enabled by default. If you prefer to include some user-configurable files in another collection instead of propagating all of them, leave the **HACMP_Files** collection disabled.

Options for propagating a PowerHA® SystemMirror® file collection

Propagating a file collection copies the files in a file collection from the current node to the other cluster nodes.

Use one of the following methods to propagate a PowerHA® SystemMirror® file collection:

- Propagate the file collection at any time manually. You can propagate files in a file collection from the PowerHA® SystemMirror® File Collection SMIT menu on the local node (the node that has the files you want to propagate).
- Set the option to propagate the file collection whenever cluster verification and synchronization is executed. The node from which verification is run is the propagation node. (This is set to **No** by default.)
- Set the option to propagate the file collection automatically after a change to one of the files in the collection. PowerHA® SystemMirror® checks the file collection status on each node (every 10 minutes by default) and propagates any changes. (This is set to **No** by default.)
  One timer is set for all file collections. You can change the timer. The maximum is 1440 minutes (24 hours) and the minimum is 10 minutes.

You can set up and change file collections on a running cluster. However, note that if you add a node dynamically, the file collection on that node may have files that are *not* in sync with the files on the other cluster nodes. If the file collection on the node being added is set for automatic propagation upon cluster verification and synchronization, the files on the node just added are updated properly. If this flag is *not* set, you must manually run the file collection propagation from one of the other nodes.

Backup files and error handling

During file propagation, before PowerHA® SystemMirror® copies a file to a remote node, the remote node makes a backup copy of the original file if it exists and its size is greater than zero, with the original time stamp.

The copy is kept in the `/var/hacmp/filebackup/` directory.

Only the most recent backup is kept for each file that is overwritten. When another propagation replaces the file, the new backup overwrites the old one. You cannot customize these backups. If you need to use a backup file, you must manually copy the file back to its original location.

If the local (propagation) node has a zero-length or non-existent file in a file collection, then an error message is logged and the file is not copied during the propagation process. The zero-length or non-existent file remains until you run a manual propagation from another node, or when an automatic propagation from another node sees a change to the file and propagates it.

All errors during file propagation are logged to SMIT if the propagation happens during a cluster verification or synchronization or manual propagation. Errors are also written to the **/var/hacmp/log/clutils.log** file.

You must verify that the file on the local (propagation) node is the latest copy and is not corrupt. PowerHA® SystemMirror® only checks for the existence and length of the file on this node.

File propagation does not occur if the file has a time stamp that is earlier than last time file propagation occurred. For example, if the file is restored from a three month old backup file it has a time stamp earlier than the last propagated file that has a time stamp from one week ago. Therefore, in this example file propagation does not occur.

If you want to use the file collection that has a time stamp older than the latest file collection, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. From the SMIT interface, select **System Management (C-SPOC) › PowerHA® SystemMirror® File Collection Management › Propagate Files in File Collections**, and press Enter.
3. Select the file collection and press Enter.

Tracking PowerHA® SystemMirror® file collection operations

Whenever the PowerHA® SystemMirror® File Collections utility replaces a file on a node, information about it is saved in the **/var/hacmp/log/clutils.log** file.

This information includes:

- Date and time of replacement
- Propagation type
- File name and file collection name
- Name of the remote and local nodes.

For example:

```
Wed Jan 07 11:08:55 2006: clfileprop: Manual file collection propagation
called.
Wed Jan 07 11:08:55 2006: clfileprop: The following file collections
will be processed:
Wed Jan 07 11:08:55 2006: clfileprop: Test_Files Wed Jan 07 11:08:55
2004: clfileprop:
Wed Jan 07 11:08:55 2006: clfileprop: Starting file propagation to
remote node riga.
Wed Jan 07 11:08:55 2006: clfileprop: Successfully propagated file
/tmp/kris to node riga.
Wed Jan 07 11:08:55 2006: clfileprop: Successfully propagated file
/tmp/k2 to node riga.
Wed Jan 07 11:08:55 2006: clfileprop: Total number of files propagated
to node riga: 2
```

Using SMIT to manage PowerHA® SystemMirror® file collections

The SMIT interface enables you to perform the certain actions.

Creating a PowerHA® SystemMirror® file collection

To create a PowerHA® SystemMirror® File Collection, at least one working IP communications path that is defined to PowerHA® SystemMirror® must exist between the node that is running the file propagation and each remote node that is defined to the cluster. The **clcomd** daemon must be running on all nodes.

To create a PowerHA® SystemMirror® file collection:

1. From the command line, enter `smit sysmirror`

2. In SMIT, select **System Management (C-SPOC) › PowerHA® SystemMirror® › File Collection Management › File Collections › Add a File Collection** and press Enter.
3. Enter field values as follows:

Add a File Collection fields	
Field	**Value**
File Collection name	The name can include alphanumeric characters and underscores. Use no more than 64 characters. Do not use reserved names. For a list of reserved names, see List of reserved words.
File Collection Description	A description of the file collection. Use no more than 100 characters.
Propagate files during cluster synchronization?	**No** is the default. If you select **Yes**, PowerHA® SystemMirror® propagates all changed files that are listed in the current collection before every cluster verification and synchronization process.
Propagate changes to files automatically?	**No** is the default. If you select **Yes**, PowerHA® SystemMirror® propagates files that are listed in the current collection across the cluster when a change is detected on any file in the collection. PowerHA® SystemMirror® checks for changes every ten minutes by default. You can adjust the timer on the **Manage File Collections** pane.

4. In SMIT, select **System Management (C-SPOC) › PowerHA® SystemMirror® › File Collection Management › Manage Files in File Collections › Add Files to a File Collection** and press Enter.
5. Select the File Collection where you want to add the files.
6. Enter the file names in the **New Files** field:

New Files fields	
Field	**Value**
File Collection name	The name of the selected file collection is displayed.
File Collection Description	The current description is displayed.
Propagate files during cluster synchronization?	The current choice is displayed.
Propagate changes to files automatically?	The current choice is displayed.
Collection Files	Any files already in the collection are displayed.
New Files	Add the full path name of the new file. The name must begin with a forward slash. A file cannot be a symbolic link, a pipe, a socket, or any file in **/dev** or **/aha** or **/proc** or ***/objrepos**. The file cannot be in another file collection, except for **PowerHA® SystemMirror®_Files**. Adding a directory to a file collection causes all non-empty files in that directory and its subdirectories to be propagated. You cannot add whole **/tmp** or **/var** directory to a file collection. You can add individual files under the **/tmp** or **/var** directory to the file collection.

7. When you finish creating the file collections, you must synchronize the cluster by selecting the following path from the SMIT interface, **Cluster Nodes and Networks › Verify and Synchronize Cluster Configuration**.

Setting the automatic timer for file collections

The default timer for automatic checks on file collections is ten minutes. You can change the amount of time as needed.

> **Note:** The periodic check for changes to a file in a file collection runs on each node. However, these checks are *not* coordinated to run simultaneously on every node. Make changes to a file *only on one node* within the general time limit.

To customize the file collection time interval:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) › PowerHA® SystemMirror® › File Collections › Manage › Change/Show Automatic Update Time** and press Enter.
3. Enter the amount of time (in minutes) that you want PowerHA® SystemMirror® to pause before performing file collection synchronization. The maximum is 1440 minutes (24 hours) and the minimum is 10 minutes. Press Enter.
4. Synchronize the cluster using SMIT.

Changing a file collection

You can modify a file collection in several different ways.

You can modify a file collection as follows:

- Change the attributes of a file collection (name, description, propagation parameters).
- Add or remove files in the collection.
- Remove a file collection.
- Change the automatic timer for all file collections, as described in Setting the automatic timer for file collections.

To change an attribute of a particular file collection:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) › PowerHA® SystemMirror® › File Collection Management › File Collections › Manage › Change/Show Automatic Update Time** and press Enter.
3. Select the file collection.
4. Change the name, description, and synchronization parameters on this panel:

File Collection fields	
Field	**Value**
File Collection name	The current name appears here.
New File Collection name	Enter the new name.
Propagate files during cluster synchronization?	**No** is the default. If you select **Yes**, PowerHA® SystemMirror® propagates all changed files that are listed in the current collection before every cluster verification and synchronization process.
Propagate changes to files automatically?	**No** is the default. If you select **Yes**, PowerHA® SystemMirror® propagates files listed in the current collection across the cluster automatically when a change is detected on any file in the collection. PowerHA® SystemMirror® checks for changes every ten minutes by default. You can adjust the timer on the **Manage File Collections** panel.

Field	Value
Collection Files	Any files already in the collection are displayed. Press F4 to see the list. You cannot change this field.

5. Synchronize the cluster.

Removing files from a file collection

You can use SMIT to remove files from a file collection.

To remove files from a file collection:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) > PowerHA® SystemMirror® File Collection Management > Manage Files in File Collections** > **Remove Files from a File Collection** and press Enter.
3. Select the File Collection from which you want to remove the files.
4. Select one or more files to remove from the file collection and press Enter.
5. Synchronize the cluster to update Configuration Databases.

Removing a file collection

You can remove a file collection from the PowerHA® SystemMirror® configuration using SMIT.

To remove a file collection from the PowerHA® SystemMirror® configuration:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) › PowerHA® SystemMirror® › File Collection Management › File Collections › Remove a File Collection** and press Enter.
3. Select the file collection to remove and press Enter.
4. SMIT displays

   ```
   Are you sure?
   ```

   Press Enter again.
5. Synchronize the cluster.

Verifying and synchronizing file collections

If file collections exist, PowerHA® SystemMirror® checks and propagates the file collections with the flag set to **yes** for “propagate during verify and synchronize” before running the rest of the cluster verification and synchronization process.

Before the files in each collection are propagated to all the cluster nodes, PowerHA® SystemMirror® performs the following verification checks:

- Verifies that no files are listed twice in any file collection. If a file is listed twice, a warning displays and verification continues.
- Verifies that each file listed in each collection is a real file on the local node (the node from which cluster synchronization is being run). A file can*not* be a symbolic link, a directory, a pipe, a socket, or any file in **/dev** or **/proc**. It can*not* begin with **/etc/objrepos/*** or **/etc/es/objrepos/***. If a file in a file collection is one of these, PowerHA® SystemMirror® displays an error and verification fails.
- Verifies that each file exists on the local node and has a file size greater than zero. If a file does *not* exist on the local node or has a size of zero, PowerHA® SystemMirror® displays an error and verification fails.
- Verifies that each file has a full path name that begins with a forward slash. If a file's pathname does *not* begin with a forward slash, PowerHA® SystemMirror® displays an error and verification fails.

Adding a custom verification method

You may want to add a custom verification method to check for a particular issue on your cluster. For example, you could add a script to check for the version of an application. You could include an error message for display and to write to the **clverify.log** file.

> **Note:** During node startup, automatic verification and synchronization does not include any custom verification methods.

To add a custom verification method, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools › PowerHA SystemMirror Verification › Configure Custom Verification Method › Add a Custom Verification Method** and press Enter.
3. Enter the field values as follows:

Add a Custom Verification Method fields	
Field	**Value**
Verification Method Name	Enter a name for the verification method. Method names can be up to 64 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 short description of the verification method.
Verification Type	Select **Script** if the type of verification method you want to use is a script. Select **Library** if the type of verification method you want to use is a library that was built with an API.
Verification Script Filename	Enter the file name for the verification method that is executable.

4. Press Enter. The method is added to the list of verification methods you can use when you select the PowerHA® SystemMirror® Verification option under the **Problem Determination Tools** menu.

Changing a custom verification method

You can use the SMIT interface to change a custom verification method.

To change a custom verification method, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In the SMIT interface, select **Problem Determination Tools › PowerHA SystemMirror Verification › Configure Custom Verification Method › Change/Show a Custom Verification Method** and press Enter.
3. Select the verification method you want to change or show and press Enter.
4. Enter a new name, new verification method description, and/or new filename as desired for the verification method and press Enter.

Removing a custom verification method

You can use the SMIT interface to remove a custom verification method.

To remove a custom verification method, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools › PowerHA SystemMirror Verification › Configure Custom Verification Method › Remove a Custom Verification Method** and press Enter.
3. 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.

4. Press Enter to remove the verification method.

List of reserved words

This topic includes all of the reserved words that you cannot use a names in cluster.

Do not use the following words as names in a cluster:

> **Note:** You can use the following words when combined with numerals or another word (for example, `my_network` or `rs232_02`).

- adapter
- alias
- atm
- BO
- cluster
- command
- custom
- daemon
- disk
- diskhb
- diskhbmulti
- ether
- event
- FBHPN
- fcs
- fddi
- FNPN
- fscsi
- FUDNP
- grep
- group
- hps
- ip
- IP
- ipv6
- IPv6
- IPV6
- IW
- name
- network
- NFB

- nim
- node
- nodename
- OAAN
- OFAN
- OHN
- OUDP
- private
- public
- resource
- rs232
- serial
- slip
- socc
- subnet
- tmscsi
- tmssa
- token
- tty
- volume
- vpath
- vscsi

You can find the most current list of reserved words in the `/usr/es/sbin/cluster/etc/reserved_words` file.

Verifying two different clusters by using the CCV utility

The cross-cluster verification (CCV) utility compares specific attributes of two different cluster configurations. The CCV utility compares data that is collected from different clusters, cluster snapshots, active configuration directory, or the default configuration directory of a local cluster. Learn about the requirements and data inputs for using the CCV utility, and how to use the CCV utility to compare two different clusters. Review the scenarios that illustrate how the CCV utility can be used.

Requirements for using the CCV utility

Because the CCV utility compares the configuration of different clusters, you must collect data from multiple clusters. The data must be gathered on a node where PowerHA® SystemMirror® is installed so that the data can be used by the CCV utility.

To gather data from different clusters, use any of the following methods:

- You can generate data on one node from each cluster, and then use File Transfer Protocol (FTP), Secure Copy Protocol (SCP), smart card or remote execution facilities to manually gather that data on a single node from where the CCV utility can use that data.
- If you configure the remote execution facilities such as Secure Shell (SSH) and SCP between different clusters, you can use the **`clmgr`** command or the **`clccv`** command to automate the generation, collection, and analysis of data from different clusters.

When you generate the cluster configuration data and collect it on a node, you must have adequate disk space to store data. The CCV utility can receive multiple data formats as an input, and the size of data can vary for each cluster configuration. You cannot determine specific disk space requirements in advance. Hence, you must ensure that adequate disk space is available for data input for the CCV utility and the resulting output.

Data inputs to the CCV utility

The CCV utility receives specific configuration information about two different clusters as an input for comparison. You can get the cluster configuration information from the following sources:

- You can use data that is generated by the CCV **cl_ccv_collector** utility.
- You can use the cluster snapshot utility to capture cluster configuration information about a cluster into a file, and then specify that file as an input to the CCV utility.
- You can run the CCV utility from a node of a PowerHA® SystemMirror® cluster. The local cluster configuration information is used as one of the input sources for comparison. The Active configuration directory (ACD) is available after you start cluster services.

You can specify any combination of inputs to the CCV utility for comparison. For example, you can:

- Compare two clusters by creating a snapshot from each cluster and then specifying those snapshots as inputs to the CCV utility.
- Compare the local cluster configuration with a remote cluster snapshot.
- Compare the active cluster configuration with data collected by the CCV utility on a remote node.

Comparing configuration attributes by using the CCV utility

The CCV utility compares a subset of configuration attributes from clusters. Each cluster is uniquely identified by the cluster name and ID. For more information about the list of entities that can be compared by using the CCV utility, see the **cl_ccv_comparator**.

The CCV utility uses the name of the entity as a unique identifier for comparison. Any two entities with the same name are compared and the difference in values are listed. All configuration attributes of all entities are not compared because of expected difference in values. For example, if a resource group with the same name exists on two different clusters, the service labels of those resources groups are expected to be different because the service labels must be unique on a network.

Therefore, when two entities with the same name are compared, only the difference in values are reported. For example, if a resource group with the same name exists on two different clusters, and if one resource group on a cluster contains three service labels and the resource group on another cluster contains only two service labels, this difference in values is reported by the CCV utility.

Output from the CCV utility

The CCV utility generates output files in the JavaScript Object Notation (JSON) format, which can be viewed by using any text editor or any graphic tool. You can also convert the output files to other data formats and view or store the output files by using other tools.

Usage scenarios

The following scenarios illustrate how the CCV utility can be used:

Using the CCV utility for the first time

Review the scenarios to become familiar with the technical and configuration details.

When you are using the CCV utility for the first time, familiarize yourself with the features of the CCV utility. You can run the **cl_ccv_collector** command on the target node in each cluster and collect the resulting data on a cluster node where you can run the **cl_ccv_comparator** command. The output that is generated from the **cl_ccv_collector** command is written to the `/var/hacmp/log/ccv_collector_data.json` file.

You can use file transfer method such as file transfer protocol (FTP) or secure copy (SCP) to retrieve the output that is generated from the **cl_ccv_collector** command to a cluster node where you can run the **cl_ccv_comparator** command. The node that is used to run the **cl_ccv_collector** command need not be a part of either clusters that are being compared. Use the **cl_ccv_comparator** command to compare clusters and to generate the output file, `/var/hacmp/log/ccv_comparison_data.json`.

You can view the output file by using any text editor or you can use a JSON display tool to view the comparison results graphically. You can convert the output file to other data formats such as comma-separated values (CSV) format, which can be viewed by using a spreadsheet tool.

Setting up the CCV utility for recurring use

If your enterprise allows the configuration of the SSH and SCP on your cluster nodes, you can automate the data collection and the cluster comparison process by using the **clccv** command. The **clccv** command takes the names of the cluster nodes as an input. The **clccv** command uses SSH to run the **cl_ccv_collector** command, and uses SCP to retrieve the output generated. If you want to use CCV utility regularly, but your enterprise does not allow you to enable of SSH or SCP continuously, you might enable them for a short duration. For example, you can enable them during a maintenance window. You can disable SSH and SCP after you use the CCV utility. You can retain the SSH or SCP configuration after you use the CCV utility to avoid configuring SSH or the SCP when you want to use the CCV utility the next time.

If you want to use the CCV utility regularly, and your enterprise allows password-less SSH authentication and SCP to be enabled continuously, you can set up the **clccv** command to start the data collection and the cluster comparison process on a recurring basis by using the **at command**. If your enterprise does not allow password-less SSH authentication and SCP to be enabled, you must use the **clccv** command interactively and provide a password when you are prompted. The results of the cluster comparison are written to the `/var/hacmp/clccv/clccv.<first cluster name>.<second cluster name>.<date>.json` file.

> **Note:** The **clmgr compare cluster** command provides the equivalent function of the **clccv** command.

Testing a PowerHA® SystemMirror® cluster

These topics describe how to use the Cluster Test Tool to test the recovery capabilities of a PowerHA® SystemMirror® cluster.

The Cluster Test Tool is available for you to test a new cluster before it becomes part of your production environment, and to test configuration changes to an existing cluster, when the cluster is *not* in service.

The Cluster Test Tool runs only on a cluster that has:

- If the cluster is migrated from an earlier version, cluster migration must be complete.
- The cluster configuration verified and synchronized.
  Before you run the tool on a cluster node, ensure that:
- The node has PowerHA® SystemMirror® installed and is part of the PowerHA® SystemMirror® cluster to be tested.
- The node has network connectivity to all of the other nodes in the PowerHA® SystemMirror® cluster.
- You have root permissions.

Because log file entries include time stamps, consider synchronizing the clocks on the cluster nodes to make it easier to review log file entries produced by test processing.

Overview for testing a cluster

The Cluster Test Tool utility allows you to test a PowerHA® SystemMirror® cluster configuration to evaluate how a cluster operates under a set of specified circumstances, such as when cluster services on a node fail or when a node loses connectivity to a cluster network.

You can start a test, let it run unattended, and return later to evaluate the results of your testing. You should run the tool under both low load and high load conditions to observe how system load affects your PowerHA® SystemMirror® cluster.

You run the Cluster Test Tool from SMIT on one node in a PowerHA® SystemMirror® cluster. For testing purposes, this node is referred to as the *control node*. From the control node, the tool runs a series of specified tests - some on other cluster nodes, gathers information about the success or failure of the tests processed, and stores this information in the Cluster Test Tool log file for evaluation or future reference.

The Cluster Test Tool allows you to test a PowerHA® SystemMirror® cluster in two ways, by running:

- Automated testing (also known as Automated Test Tool). In this mode, the Cluster Test Tool runs a series of predefined sets of tests on the cluster.
- Custom testing (also known as Test Plan). In this mode, you can create your own test plan, or a custom testing routine, that will include different tests available in the Cluster Test Tool library.

Automated testing

Use the automated test procedure (a predefined set of tests) supplied with the tool to perform basic cluster testing on any cluster.

No setup is required. You simply run the test from SMIT and view test results from SMIT and the Cluster Test Tool log file.

The automated test procedure runs a predefined set of tests on a node that the tool randomly selects. The tool ensures that the node selected for testing varies from one test to another. For information about automated testing, see Running automated tests.

Custom testing

If you are an experienced PowerHA® SystemMirror® administrator and want to tailor cluster testing to your environment, you can create custom tests that can be run from SMIT.

You create a custom test plan (a file that lists a series of tests to be run), to meet requirements specific to your environment and apply that test plan to any number of clusters. You specify the order in which tests run and the specific components to be tested. After you set up your custom test environment, you run the test procedure from SMIT and view test results in SMIT and in the Cluster Test Tool log file. For information about customized testing, see Setting up custom cluster testing.

Test duration

Running automated testing on a basic two-node cluster that has a simple cluster configuration takes approximately 30 to 60 minutes to complete.

Individual tests can take around three minutes to run. The following conditions affect the length of time to run the tests:

- Cluster complexity
  Testing in complex environments takes considerably longer.
- Latency on the network
  Cluster testing relies on network communication between the nodes. Any degradation in network performance slows the performance of the Cluster Test Tool.
- Use of verbose logging for the tool
  If you customize verbose logging to run additional commands from which to capture output, testing takes longer to complete. In general, the more commands you add for verbose logging, the longer a test procedure takes to complete.
- Manual intervention on the control node
  At some points in the test, you may need to intervene. See Recovering the control node after cluster manager stops for ways to avoid this situation.
- Running custom tests
  If you run a custom test plan, the number of tests run also affects the time required to run the test procedure. If you run a long list of tests, or if any of the tests require a substantial amount of time to complete, then the time to process the test plan increases.

Security during testing

The Cluster Test Tool uses the PowerHA® SystemMirror® Cluster Communications daemon to communicate between cluster nodes to protect the security of your PowerHA® SystemMirror® cluster.

Limitations of Cluster Test Tool

The Cluster Test Tool has some limitations.

It does not support testing of the following PowerHA® SystemMirror® cluster-related components:

- Network adapters that have FQDN
- RGs with dependencies
- Dynamic cluster reconfiguration.
  You cannot run dynamic reconfiguration while the tool is running.
- Pre-events and post-events.
  Pre-events and post-events run in the usual way, but the tool does not verify that the events were run or that the correct action was taken.

  In addition, the Cluster Test Tool may not recover from the following situations:
- A node that fails unexpectedly, that is a failure not initiated by testing
- The cluster does not stabilize.

> **Note:** The Cluster Test Tool uses the terminology for stopping cluster services that was used in PowerHA® SystemMirror® prior to v.5.4 (graceful stop, graceful with takeover and forced stop).

Running automated tests

You can run the automated test procedure on any PowerHA® SystemMirror® cluster that is not currently in service.

The Cluster Test Tool runs a specified set of tests and randomly selects the nodes, networks, resource groups, and so forth for testing. The tool tests different cluster components during the course of the testing. For a list of the tests that are run, see Understanding automated testing.

Before you start running an automated test:

- Ensure that the cluster is not in service in a production environment
- Stop PowerHA® SystemMirror® cluster services, this is recommended but optional. Note that if the Cluster Manager is running, some of the tests will be irrational for your configuration, but the Test Tool will continue to run.
- Cluster nodes are attached to two IP networks.

One network is used to test a network becoming unavailable then available. The second network provides network connectivity for the Cluster Test Tool. Both networks are tested, one at a time.

Launching the cluster test tool

You can use the cluster test tool to run an automated test procedure.

To run the automated test procedure:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools › Cluster Test Tool › Execute Automated Test Procedure** and press Enter.
   The system displays:

   ```
   Are you sure
   ```

   If you press Enter again, the automated test plan runs.
3. Evaluate the test results.

Result
For information about evaluating test results, see Evaluating results.

Modifying logging and stopping processing in the cluster test tool

You can modify several different features in the cluster test tool.

You can also modify processing for automated test procedure to:

- Turn off verbose logging
- Turn off cycling of log files for the tool
- Stop processing tests after the first test fails

To modify processing for an automated test:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools**.
   Then select **PowerHA® SystemMirror® Cluster Test Tool.**
3. In the **PowerHA® SystemMirror® Cluster Test Tool** panel, select **Execute Automated Test Procedure** .
4. In the **Execute Automated Test Procedure** panel, enter field values as follows:

Execute Automated Test Procedure fields	
Field	**Value**
Verbose Logging	When set to **yes**, includes additional information in the log file. This information may help to judge the success or failure of some tests. For more information about verbose logging and how to modify it for your testing, see Error logging. Select **no** to decrease the amount of information logged by the Cluster Test Tool. The default is **yes**.
Cycle Log File	When set to **yes**, uses a new log file to store output from the Cluster Test Tool. Select **no** to append messages to the current log file. The default is **yes**. For more information about cycling the log file, see Error logging.
Abort on Error	When set to **no**, the Cluster Test Tool continues to run tests after some of the tests being run fail. This may cause subsequent tests to fail because the cluster state is different from the one expected by one of those tests. Select **yes** to stop processing after the first test fails. For information about the conditions under which the Cluster Test Tool stops running, see Cluster test tool stops running. The default is **no**. The tool stops running and issues an error if a test fails and **Abort on Error** is selected.

5. Press Enter to start running the automated tests.
6. Evaluate the test results.

Understanding automated testing

These topics list the sequence that the Cluster Test Tool uses for the automated testing, and describes the syntax of the tests run during automated testing.

The automated test procedure performs sets of predefined tests in the following order:

1. General topology tests
2. Resource group tests on non-concurrent resource groups
3. Resource group tests on concurrent resource groups
4. Tests for each network
5. Volume group tests for each resource group
6. Catastrophic failure test.

The Cluster Test Tool discovers information about the cluster configuration, and randomly selects cluster components, such as nodes and networks, to be used in the testing.

Which nodes are used in testing varies from one test to another. The Cluster Test Tool may select some node(s) for the initial battery of tests, and then, for subsequent tests, it may intentionally select the same node(s), or, choose from nodes on which no tests were run previously. In general, the logic in the automated test sequence ensures that all components are sufficiently tested in all necessary combinations.

The testing follows these rules:

- Tests operation of a concurrent resource group on one randomly selected node - not all nodes in the resource group.
- Tests only those resource groups that include monitored application controllers or volume groups.
- Requires at least two active IP networks in the cluster to test non-concurrent resource groups.

The automated test procedure runs a **node_up** event at the beginning of the test to make sure that all cluster nodes are up and available for testing.

These sections list the tests in each group. For more information about a test, including the criteria to determine the success or failure of a test, see Description of tests. The automated test procedure uses variables for parameters, with values drawn from the PowerHA® SystemMirror® cluster configuration.

The examples in the following sections use variables for node, resource group, application controller, stop script, and network names. For information about the parameters specified for a test, see Description of tests.

General topology tests

The Cluster Test Tool runs the general topology tests in a certain order.

The order is as follows:

1. Bring a node up and start cluster services on all available nodes
2. Stop cluster services on a node and bring resource groups offline.
3. Restart cluster services on the node that was stopped
4. Stop cluster services and move resource groups to another node
5. Restart cluster services on the node that was stopped
6. Stop cluster services on another node and place resource groups in an UNMANAGED state.
7. Restart cluster services on the node that was stopped.

The Cluster Test Tool uses the terminology for stopping cluster services that was used in PowerHA® SystemMirror® in releases prior to v.5.4. For information on how the methods for stopping cluster services map to the terminology used in v.5.4, see Starting and stopping cluster services.

When the automated test procedure starts, the tool runs each of the following tests in the order shown:

1. `NODE_UP, ALL, Start cluster services on all available nodes`
2. `NODE_DOWN_GRACEFUL, node1, Stop cluster services gracefully on a node`
3. `NODE_UP, node1, Restart cluster services on the node that was stopped`
4. `NODE_DOWN_TAKEOVER, node2, Stop cluster services with takeover on a node`
5. `NODE_UP, node2, Restart cluster services on the node that was stopped`

6. `NODE_DOWN_FORCED, node3, Stop cluster services forced on a node`
7. `NODE_UP, node3, Restart cluster services on the node that was stopped`

Resource group tests

There are two groups of resource group tests that can be run. Which group of tests run depends on the startup policy for the resource group: non-concurrent and concurrent resource groups. If a resource of the specified type does not exist in the resource group, the tool logs an error in the Cluster Test Tool log file.

Resource group starts on a specified node

The following tests run if the cluster includes one or more resource groups that have a startup management policy *other than* **Online on All Available Nodes**, that is, the cluster includes one or more non-concurrent resource groups.

The Cluster Test Tool runs each of the following tests in the order shown for each resource group:

1. Bring a resource group offline and online on a node.

   ```
   RG_OFFLINE, RG_ONLINE
   ```

2. Bring a local network down on a node to produce a resource group fallover.

   ```
   NETWORK_DOWN_LOCAL, rg_owner, svc1_net, Selective fallover on local
   network down
   ```

3. Recover the previously failed network.

   ```
   NETWORK_UP_LOCAL, prev_rg_owner, svc1_net, Recover previously failed
   network
   ```

4. Move a resource group to another node. RG_MOVE
5. Bring an application controller down and recover from the application failure.

   ```
   SERVER_DOWN, ANY, app1, /app/stop/script, Recover from
   application failure
   ```

Resource group starts on all available nodes

If the cluster includes one or more resource groups that have a startup management policy of **Online on All Available Nodes**, that is, the cluster has concurrent resource groups, the tool runs one test that brings an application controller down and recovers from the application failure.

The tool runs the following test:

```
RG_OFFLINE, RG_ONLINE
SERVER_DOWN, ANY, app1, /app/stop/script, Recover from
application failure
```

Network tests

The tool runs tests for defined networks.

For each network, the tool runs these tests:

- Bring a network down and up.

  ```
  NETWORK_DOWN_GLOBAL, NETWORK_UP_GLOBAL
  ```

- Fail a network interface, join a network interface. This test is run for the service interface on the network. If no service interface is configured, the test uses a random interface defined on the network.

```
FAIL_LABEL, JOIN_LABEL
```

Volume group tests

The tool runs tests for volume groups.

For each resource group in the cluster, the tool runs tests that fail a volume group in the resource group:

```
VG_DOWN
```

Catastrophic failure test

As a final test, the tool stops the Cluster Manager on a randomly selected node that currently has at least one active resource group.

```
CLSTRMGR_KILL, node1, Kill the cluster manager on a node
```

If the tool terminates the Cluster Manager on the control node, you may need to reboot this node.

Setting up custom cluster testing

If you want to extend cluster testing beyond the scope of the automated testing and you are an experienced PowerHA® SystemMirror® administrator who has experience planning, implementing, and troubleshooting clusters, you can create a custom test procedure to test the PowerHA® SystemMirror® clusters in your environment.

You can specify the tests specific to your clusters, and use variables to specify parameters specific to each cluster. Using variables lets you extend a single custom test procedure to run on a number of different clusters. You then run the custom test procedure from SMIT.

> **Important:** If you uninstall PowerHA® SystemMirror®, the program removes any files you may have customized for the Cluster Test Tool. If you want to retain these files, make a copy of these files before you uninstall PowerHA® SystemMirror®.

Planning a test procedure

Before you create a test procedure, make sure that you are familiar with the PowerHA® SystemMirror® clusters on which you plan to run the test.

List the following components in your cluster and have this list available when setting up a test:

- Nodes
- Networks
- Volume groups
- Resource groups
- Application controllers

Your test procedure should bring each component offline then online, or cause a resource group fallover, to ensure that the cluster recovers from each failure.

Start your test by running a **node_up** event on each cluster node to ensure that all cluster nodes are up and available for testing.

Creating a custom test procedure

This topic describes the high level task of creating a custom test procedure.

To create a custom test procedure:

1. Create a Test Plan, a file that lists the tests to be run.
   For information about creating a Test Plan, see Creating a test plan.
2. Set values for test parameters.

For information about specifying parameters, see Specifying parameters for tests.

Creating a test plan

A test plan is a text file that lists cluster tests to be run in the order in which they are listed in the file. In a test plan, specify one test per line. You can set values for test parameters in the test plan or use variables to set parameter values.

The tool supports the following tests:

Test plans	
Test plan	**Description**
FAIL_LABEL	Brings the interface associated with the specified label down on the specified node.
JOIN_LABEL	Brings the interface associated with the specified label up on the specified node.
NETWORK_UP_GLOBAL	Brings a specified network up (IP network or non-IP network) on all nodes that have interfaces on the network.
NETWORK_DOWN_GLOBAL	Brings a specified network down (IP network or non-IP network) on all nodes that have interfaces on the network.
NETWORK_UP_LOCAL	Brings a network interface on a node up.
NETWORK_DOWN_LOCAL	Brings a network interface on a node down.
NETWORK_UP_NONIP	Brings a non-IP network on a node up.
NETWORK_DOWN_NONIP	Brings a non-IP network on a node down.
NODE_UP	Starts cluster services on the specified node.
NODE_DOWN_GRACEFUL	Stops cluster services and brings the resource groups offline on the specified node.
NODE_DOWN_TAKEOVER	Stops cluster services with the resources acquired by another node.
NODE_DOWN_FORCED	Stops cluster services on the specified node with the Unmanage Resource Group option.
CLSTRMGR_KILL	Terminates the Cluster Manager on the specified node
RG_MOVE	Moves a resource group that is already online to a specific node
RG_MOVE_SITE	Moves a resource group that is already online to an available node at a specific site.
RG_OFFLINE	Brings a resource group offline that is already online
RG_ONLINE	Brings a resource group online that is already offline
SERVER_DOWN	Brings a monitored application controller down
VG_DOWN	Emulates an error condition for a specified disk that contains a volume group in a resource group.
WAIT	Generates a wait period for the Cluster Test Tool.

For a full description of these tests, see Description of tests.

Specifying parameters for tests

You can specify parameters for the tests in the test plan.

Specify parameters by doing one of the following:

- Using a variables file. A variables file defines values for variables assigned to parameters in a test plan.
- Setting values for test parameters as environment variables.
- Identifying values for parameters in the test plan.

When the Cluster Test Tool starts, it uses a variables file if you specified the location of one in SMIT. If it does not locate a variables file, it uses values set in an environment variable. If a value is not specified in an environment variable, it uses the value in the test plan. If the value set in the test plan is not valid, the tool displays an error message.

Using a variables file

The variables file is a text file that defines the values for test parameters. By setting parameter values in a separate variables file, you can use your test plan to test more than one cluster.

The entries in the file have this syntax:

***parameter_name* = value**

For example, to specify a node as **node_waltham**:

```
node=node_waltham
```

To provide more flexibility, you can:

1. Set the name for a parameter in the test plan.
2. Assign the name to another value in the variables file.
   For example, you could specify the value for *node* as **node1** in the test plan:

   NODE_UP,node1, Bring up node1

   In the variables file, you can then set the value of **node1** to **node_waltham**:

   ```
   node1=node_waltham
   ```

   The following example shows a sample variables file:

   ```
   node1=node_waltham
   node2=node_belmont
   node3=node_watertown
   node4=node_lexington
   ```

Using environment variables

If you do not want to use a variables file, you can assign parameter values by setting environment variables for the parameter values. If a variable file is not specified, but there are ***parameter_name* = values** in the cluster environment that match the values in the test plan, the Cluster Test Tool will use the values from the cluster environment.

Using the test plan

If you want to run a test plan on only one cluster, you can define test parameters in the test plan. The associated test can be run only on the cluster that includes those cluster attributes specified. For information about the syntax for parameters for tests, see Description of tests.

Description of tests

The Test Plan supports the tests listed in this section. The description of each test includes information about the test parameters and the success indicators for a test.

Note: One of the success indicators for each test is that the cluster becomes stable. The definition of cluster stability takes a number of factors into account, beyond the state of the Cluster Manager. The **clstat** utility, by comparison, uses only the state of the Cluster Manager to assess stability. For information about the factors used to determine cluster stability for the Cluster Test Tool, see Evaluating results.

Test syntax

This topic describes the syntax for a test.

The syntax for a test is:

TEST_NAME, *parameter1*, *parametern*|*PARAMETER*, *comments*

where:

- The test name is in uppercase letters.
- Parameters follow the test name.
- Italic text indicates parameters expressed as variables.
- Commas separate the test name from the parameters and the parameters from each other. The PowerHA® SystemMirror® Cluster Test Tool supports spaces around commas.
  The example syntax line shows parameters as *parameter1* and *parametern* with *n* representing the next parameter. Tests typically have from two to four parameters.
- A pipe (**|**) indicates parameters that are mutually exclusive alternatives.
  Select one of these parameter options.
- (*Optional*) Comments (user-defined text) appear at the end of the line. The Cluster Test Tool displays the text string when the Cluster Test Tool runs.

In the test plan, the tool ignores:

- Lines that start with a pound sign (#)
- Blank lines.

Node tests

The node tests start and stop cluster services on specified nodes.

NODE_UP, *node* | ALL, *comments*

Starts cluster services on a specified node that is offline or on all nodes that are offline.

node

The name of a node on which cluster services start.

ALL

Any nodes that are offline have cluster services start.

comments

User-defined text to describe the configured test.

Example

```
NODE_UP, node1, Bring up node1
```

Entrance criteria

Any node to be started is inactive.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- The cluster services successfully start on all specified nodes
- No resource group enters the error state
- No resource group moves from online to offline.

NODE_DOWN_GRACEFUL, *node* | ALL, *comments*

Stops cluster services on a specified node and brings resource groups offline.

node

The name of a node on which cluster services stop

ALL

All nodes are to have cluster services stop. If you specify **ALL**, at least one node in the cluster must be online for this test to run.

comments

User-defined text to describe the configured test.

Example

```
NODE_DOWN_GRACEFUL, node3, Bring down node3 gracefully
```

Entrance criteria

Any node to be stopped is active.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services stop on the specified node(s)
- Cluster services continue to run on other nodes if **ALL** is *not* specified
- resource groups on the specified node go offline, and do *not* move to other nodes
- resource groups on other nodes remain in the same state.

NODE_DOWN_TAKEOVER, *node*, *comments*

Stops cluster services on a specified node with a resource group acquired by another node as configured, depending on resource availability.

node

The name of a node on which cluster services stop.

comments

User-defined text to describe the configured test.

Example

```
NODE_DOWN_TAKEOVER, node4, Bring down node4 gracefully with takeover
```

Entrance criteria

The specified node is active.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services stop on the specified node
- Cluster services continue to run on other nodes
- All resource groups remain in the same state.

NODE_DOWN_FORCED, *node* , *comments*

Stops cluster services on a specified node and places resource groups in an UNMANAGED state. Resources on the node remain online, that is they are *not* released.

node

The name of a node on which to stop cluster services

comments

User-defined text to describe the configured test.

Example

```
NODE_DOWN_FORCED, node2, Bring down node2 forced
```

Entrance criteria

Cluster services on another node have *not* already been stopped with its resource groups placed in an UNMANAGED state. The specified node is active.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- The resource groups on the node change to UNMANAGED state
- Cluster services stop on the specified node
- Cluster services continue to run on other nodes
- All resource groups remain in the same state.

Network tests

This section lists tests that bring network interfaces up or down on an IP network.

The Cluster Test Tool requires two IP networks to run any of the tests described in this section. The second network provides network connectivity for the tool to run. The Cluster Test Tool verifies that two IP networks are configured before running the test.

NETWORK_UP_LOCAL, *node* , *network* , *comments*

Brings a specified network up on a specified node by running the **ifconfig up** command on the node.

node

The name of the node on which to run the **ifconfig down** command

network

The name of the network to which the interface is connected

comments

User-defined text to describe the configured test.

Example

```
NETWORK_UP_LOCAL, node6, hanet1, Start hanet1 on node 6
```

Entrance criteria

The specified node is active and has at least one inactive interface on the specified network.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services continue to run on the cluster nodes where they were active before the test
- Resource groups that are in the ERROR state on the specified node and that have a service IP label available on the network can go online, but should not enter the ERROR state
- Resource groups on other nodes remain in the same state.

NETWORK_DOWN_LOCAL, *node*, *network*, *comments*

Brings a specified network down on a specified node by running the **ifconfig down** command.

> **Note:** If one IP network is already unavailable on a node, the cluster may become partitioned. The Cluster Test Tool does not take this into account when determining the success or failure of a test.

node

The name of the node on which to run the **ifconfig down** command

network

The name of the network to which the interface is connected

comments

User-defined text to describe the configured test.

Example

```
NETWORK_DOWN_LOCAL, node8, hanet2, Bring down hanet2 on node 8
```

Entrance criteria

The specified node is active and has at least one active interface on the specified network.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services continue to run on the cluster nodes where they were active before the test

- Resource groups on other nodes remain in the same state; however, some may be hosted on a different node
- If the node hosts a resource group for which the recovery method is set to notify, the resource group does *not* move.

NETWORK_UP_GLOBAL, *network, comments*

Brings specified network up on all nodes that have interfaces on the network. The network specified may be an IP network or a serial network.

network

The name of the network to which the interface is connected

comments

User-defined text to describe the configured test.

Example

```
NETWORK_UP_GLOBAL, hanet1, Start hanet1 on node 6
```

Entrance criteria

Specified network is active on at least one node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services continue to run on the cluster nodes where they were active before the test
- resource groups that are in the ERROR state on the specified node and that have a service IP label available on the network can go online, but should not enter the ERROR state
- resource groups on other nodes remain in the same state.

NETWORK_DOWN_GLOBAL, *network, comments*

Brings the specified network down on all nodes that have interfaces on the network. The network specified may be an IP network or a serial network.

> **Note:** If one IP network is already unavailable on a node, the cluster may become partitioned. The Cluster Test Tool does *not* take this into account when determining the success or failure of a test.

network

The name of the network to which the interface is connected

comments

User-defined text to describe the configured test.

Example

```
NETWORK_DOWN_GLOBAL, hanet1, Bring down hanet1 on node 6
```

Entrance criteria

Specified network is inactive on at least one node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services continue to run on the cluster nodes where they were active before the test
- resource groups on other nodes remain in the same state.

Network interface tests

This section lists tests that bring network interfaces up or down on an IP network.

JOIN_LABEL *iplabel, comments*

Brings up a network interface associated with the specified IP label on a specified node by running the **ifconfig up** command.

> **Note:** You specify the IP label as the parameter. The interface that is currently hosting the IP label is used as the argument to the **ifconfig** command. The IP label can be a service label or a boot label. If it is a service label, then that service label must be hosted on some interface, for example, when the resource group is actually online. You cannot specify a service label that is not already hosted on an interface.

The only time you could have a resource group online and the service label hosted on an inactive interface would be when the service interface fails but there was no place to move the resource group, in which case it stays online.

iplabel

The IP label of the interface.

comments

User-defined text to describe the configured test.

Example

```
JOIN_LABEL, app_serv_address, Start app_serv_address on node 2
```

Entrance criteria

Specified interface is currently active on the specified node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Specified interface comes up on specified node
- Cluster services continue to run on the cluster nodes where they were active before the test
- Resource groups that are in the ERROR state on the specified node and that have a service IP label available on the network can go online, but should not enter the ERROR state
- Resource groups on other nodes remain in the same state.

FAIL_LABEL, *iplabel, comments*

Brings down a network interface associated with a specified label on a specified node by running the **ifconfig down** command.

> **Note:** You specify the IP label as the parameter. The interface that is currently hosting the IP label is used as the argument to the **ifconfig** command The IP label can be a service label or a boot label.

iplabel

The IP label of the interface.

comments

User-defined text to describe the configured test.

Example

```
FAIL_LABEL, app_serv_label, Bring down app_serv_label, on node 2
```

Entrance criteria

The specified interface is currently inactive on the specified node

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Any service labels that were hosted by the interface are recovered
- Resource groups that are in the ERROR state on the specified node and that have a service IP label available on the network can go online, but should not enter the ERROR state
- Resource groups remain in the same state; however, the resource group may be hosted by another node.

Resource group tests

This section lists tests for resource groups.

RG_ONLINE, *rg, node | ALL | ANY | RESTORE, comments*

Brings a resource group online in a running cluster.

Paramaters

rg

The name of the resource group to bring online.

node

The name of the node where the resource group will come online.

ALL

Use **ALL** for concurrent resource groups only. When **ALL** is specified, the resource group will be brought online on all nodes in the resource group. If you use **ALL** for non-concurrent groups, the Test Tool interprets it as **ANY**.

ANY

Use **ANY** for non-concurrent resource groups to pick a node where the resource group is offline. For concurrent resource groups, use **ANY** to pick a random node where the resource group will be brought online.

RESTORE

Use **RESTORE** for non-concurrent resource groups to bring the resource groups online on the highest priority available node. For concurrent resource groups, the resource group will be brought online on all nodes in the nodelist.

comments

User-defined text to describe the configured test.

Example

```
RG_ONLINE, rg_1, node2, Bring rg_1 online on node 2.
```

Entrance criteria

The specified resource group is offline, there are available resources, and can meet all dependencies.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- The resource group is brought online successfully on the specified node
- No resource groups go offline or into ERROR state.

RG_OFFLINE, *rg, node* | ALL | ANY*, comments*

Brings a resource group offline that is already online in a running cluster.

Paramaters

rg

The name of the resource group to bring offline.

node

The name of the node on which the resource group will be taken offline.

ALL

Use **ALL** for concurrent resource groups to bring the resource group offline on all nodes where the resource group is hosted. You can also use **ALL** for non-concurrent resource groups to bring the group offline on the node where it is online.

ANY

Use ANY for non-concurrent resource groups to bring the resource group offline on the node where it is online. You can use **ANY** for concurrent resource groups to select a random node where the resource group is online.

comments

User-defined text to describe the configured test.

Example

```
RG_OFFLINE, rg_1, node2, Bring rg_1 offline from node2
```

Entrance criteria

The specified resource group is online on the specified node

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Resource group, which was online on the specified node, is brought offline successfully
- Other resource groups remain in the same state.

RG_MOVE*, rg, node* | ANY | RESTORE*, comments*

Moves a resource group that is already online in a running cluster to a specific or any available node.

Paramaters

rg

The name of the resource group to bring offline.

node

The target node; the name of the node to which the resource group will move.

ANY

Use **ANY** to let the Cluster Test Tool pick a random available node to which to move the resource group.

RESTORE

Enable the resource group to move to the highest priority node available.

comments

User-defined text to describe the configured test.

Example

```
RG_MOVE, rg_1, ANY, Move rg_1 to any available node.
```

Entrance criteria

The specified resource group must be non-concurrent and must be online on a node other *than* the target node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Resource group is moved to the target node successfully
- Other resource groups remain in the same state.

Volume group tests

This section lists tests for volume groups.

VG_DOWN*, vg, node* | ALL | ANY, *comments*

Forces an error for a disk that contains a volume group in a resource group.

Paramaters

vg

The volume group on the disk of which to fail.

node

The name of the node where the resource group that contains the specified volume group is currently online.

ALL

Use **ALL** for concurrent resource groups. When **ALL** is specified, the Cluster Test Tool will fail the volume group on all nodes in the resource group where the resource group is online. If ALL is used for non-concurrent resource groups, the Tool performs this test for any resource group.

ANY

Use **ANY** to have the Cluster Test Tool will select the node as follows:

- For a non-concurrent resource group, the Cluster Test Tool will select the node where the resource group is currently online.
- For a concurrent resource group, the Cluster Test Tool will select a random node from the concurrent resource group node list, where the resource group is online

comments

User-defined text to describe the configured test.

Example

```
VG_DOWN, sharedvg, ANY, Fail the disk where sharedvg resides
```

Entrance criteria

The resource group containing the specified volume groups is online on the specified node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Resource group containing the specified volume group successfully moves to another node, or if it is a concurrent resource groups, it goes into an ERROR state
- Resource groups may change state to meet dependencies.

General tests

This section lists tests for general.

The other tests available to use in PowerHA® SystemMirror® cluster testing are:

- Bring an application controller down
- Terminate the Cluster Manager on a node
- Add a wait time for test processing.

SERVER_DOWN, *node* | ANY, *appserv*, *command*, *comments*

Runs the specified command to stop an application controller. This test is useful when testing application availability.

In the automated test, the test uses the stop script to turn off the application.

Paramaters

node

The name of a node on which the specified application controller is to become unavailable.

ANY

Any available node that participates in this resource group can have the application controller become unavailable
The Cluster Test Tool tries to simulate server failure on any available cluster node. This test is equivalent to failure on the node that currently owns the resource group, *if the server is in a resource group that has policies other than the following ones*:

- Startup: Online on all available nodes
- Fallover: Bring offline (on error node only)

appserv

The name of the application controller associated with the specified node.

command

The command to be run to stop the application controller.

comments

User-defined text to describe the configured test.

Example

```
SERVER_DOWN,node1,db_app /apps/stop_db.pl, Kill the db app
```

Entrance criteria

The resource group is online on the specified node.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster nodes remain in the same state
- The resource group that contains the application controller is online; however, the resource group may be hosted by another node, unless it is a concurrent resource group, in which case the group goes into ERROR state.

CLSTRMGR_KILL command

Purpose

Runs the **kill** command to terminate the Cluster Manager on a specified node.

Syntax

```
CLSTRMGR_KILL, node , comments
```

Description

If the **CLSTRMGR_KILL** command is run on the local node, you may need to reboot the node. On startup, the Cluster Test Tool automatically starts again. For information about how to avoid manually rebooting the node, see the section (Stops).

For the Cluster Test Tool to accurately assess the success or failure of a **CLSTRMGR_KILL** test, do not perform other activities in the cluster while the Cluster Test Tool is running.

Paramaters

node

The name of the node on which to terminate the Cluster Manager

comments

User-defined text to describe the configured test.

Example

```
CLSTRMGR_KILL, node5, Bring down node5 hard
```

Entrance criteria

The specified node is active.

Success indicators

The following conditions indicate success for this test:

- The cluster becomes stable
- Cluster services stop on the specified node
- Cluster services continue to run on other nodes
- Resource groups that were online on the node where the Cluster Manager fails move to other nodes
- All resource groups on other nodes remain in the same state.

For information about potential conditions caused by a **CLSTRMGR_KILL** test running on the control node, see Recovering the control node after cluster manager stops.

WAIT, *seconds*, *comments*

Generates a wait period for the Cluster Test Tool for a specified number of seconds.

Paramaters

seconds

The number of seconds that the Cluster Test Tool waits before proceeding with processing.

comments

User-defined text to describe the configured test.

Example

```
WAIT, 300, We need to wait for five minutes before the next test
```

Entrance criteria

Not applicable.

Success indicators

Not applicable.

Example test plan

This section includes test examples.

The following excerpt from a sample Test Plan includes the tests:

- **NODE_UP**
- **NODE_DOWN_GRACEFUL**

It also includes a WAIT interval. The comment text at the end of the line describes the action to be taken by the test.

```
NODE_UP,ALL,starts cluster services on all nodes
NODE_DOWN_GRACEFUL,waltham,stops cluster services gracefully on node waltham
WAIT,20
NODE_UP,waltham,starts cluster services on node waltham
```

Running custom test procedures

This topic discusses the processes for launching a custom test procedure.

Before you start running custom tests, ensure that:

- Your Test Plan is configured correctly.
  For information about setting up a Test Plan, see Creating a test plan.
- You have specified values for test parameters.
  For information about parameter values, see Specifying parameters for tests.
- You have logging for the tool configured to capture the information that you want to examine for your cluster.
  For information about customizing verbose logging for the Cluster Test Tool, see Error logging.
- The cluster is *not* in service in a production environment.

To run custom testing:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools**.
   Then select **Cluster Test Tool**.
3. In the **PowerHA® SystemMirror® Cluster Test Tool** panel, select **Execute Custom Test Procedure**.
4. In the **Execute Custom Test Procedure** panel, enter field values as follows:

Execute Custom Test Procedure fields	
Field	**Value**
Test Plan	(*Required*) The full path to the Test Plan for the Cluster Test Tool. This file specifies the tests for the tool to execute.
Variable File	(*Using a variables file is optional but recommended.*) The full path to the variables file for the Cluster Test Tool. This file specifies the variable definitions used in processing the Test Plan.
Verbose Logging	When set to **yes**, includes additional information in the log file that may help to judge the success or failure of some tests. For more information about verbose logging, see Running automated tests. The default is **yes**. Select **no** to decrease the amount of information logged by the Cluster Test Tool.
Cycle Log File	When set to **yes,** uses a new log file to store output from the Cluster Test Tool. The default is **yes**. Select **no** to append messages to the current log file. For more information about cycling the log file, see Log files.

Field	Value
Abort on Error	When set to **no**, the Cluster Test Tool continues to run tests after some of the tests being run fail. This may cause subsequent tests to fail because the cluster state is different from the one expected by one of those tests. The default is **no**. Select **yes** to stop processing after the first test fails. For information about the conditions under which the Cluster Test Tool stops running, see Cluster test tool stops running. The tool stops running and issues an error if a test fails and **Abort on Error** is select.

5. Press Enter to start running the custom tests.
6. Evaluate the test results.

Result
For information about evaluating test results, see Evaluating results.

Evaluating results

You evaluate test results by reviewing the contents of the log file created by the Cluster Test Tool.

When you run the Cluster Test Tool from SMIT, it displays status messages to the screen and stores output from the tests in the file **/var/hacmp/log/cl_testtool.log**. Messages indicate when a test starts and finishes and provide additional status information. More detailed information, especially when verbose logging is enabled, is stored in the log file that appears on the screen. Information is also logged to the **hacmp.out** file.

The following criteria determine the success or failure of cluster tests:

- Did the cluster stabilize?
  For the Cluster Test Tool, a cluster is considered stable when:
  - The Cluster Manager has a status of stable on each node, or is not running.
  - Nodes that should be online are online.
    If a node is stopped and that node is the last node in the cluster, the cluster is considered stable when the Cluster Manager is inoperative on all nodes.
  - No events are in the event queue for PowerHA® SystemMirror®.
    The Cluster Test Tool also monitors PowerHA® SystemMirror® timers that may be active. The tool waits for some of these timers to complete before determining cluster stability. For more information about how the Cluster Test Tool interacts with PowerHA® SystemMirror® timers, see Working with timer settings.
- Has an appropriate recovery event for the test run?
- Is a specific node online or offline as specified?
- Are all expected resource groups still online within the cluster?
- Did a test that was expected to run actually run?

Every test checks to see if it makes sense to be run; this is called a check for "rationality". A test returning a `NOT RATIONAL` status indicates the test could not be run because the entrance criteria could not be met; for example, trying to run the NODE_UP test on a node that is already up. A warning message will be issued along with the exit status to explain why the test was not run. Irrational tests do not cause the Cluster Test Tool to abort.

The `NOT RATIONAL` status indicates the test was not appropriate for your cluster. When performing automated testing, it is important to understand why the test did not run. For Custom Cluster tests, check the sequences of events and modify the test plan to ensure the test runs. Consider the order of the tests and the state of the cluster before running the test plan. For more information, refer to Setting up custom cluster testing.

The tool targets availability as being of primary importance when reporting success or failure for a test. For example, if the resource groups that are expected to be available are available, the test passes.

Keep in mind that the Cluster Test Tool is testing the cluster configuration, not testing PowerHA® SystemMirror®. In some cases the configuration may generate an error that causes a test to fail, even though the error is the expected behavior. For example, if a resource group enters the error state and there is no node to acquire the resource group, the test fails.

> **Note:** If a test generates an error, the Cluster Test Tool interprets the error as a test failure. For information about how the Cluster Test Tool determines the success or failure of a test, see the Success Indicators subsections for each test in Description of tests.

Recovering the control node after cluster manager stops

If a **CLSTRMGR_KILL** test runs on the control node and stops the control node, reboot the control node. No action is taken to recover from the failure. After the node reboots, the testing continues.

To monitor testing after the Cluster Test Tool starts again, review output in the **/var/hacmp/log/cl_testtool.log** file. To determine whether a test procedure completes, run the **tail -f** command on **/var/hacmp/log/cl_testtool.log** file.

You can avoid manual intervention to reboot the control node during testing by:

- Editing the **/etc/cluster/hacmp.term** file to change the default action after an abnormal exit. The clexit.rc script checks for the presence of this file and, if the file is executable, the script calls it instead of halting the system automatically.
- Configuring the node to auto-Initial Program Load (IPL) before running the Cluster Test Tool.

Error logging

The Cluster Test Tool has several useful functions that enable you to work with logs.

Log files

If a test fails, the Cluster Test Tool collects information in the automatically created log files. To collect logs, the Cluster Test Tool creates the directory `/var/hacmp/cl_testtool` if it doesn't exist. PowerHA® SystemMirror® never deletes the files in this directory. You evaluate the success or failure of tests by reviewing the contents of the Cluster Test Tool log file, `/var/hacmp/cl_testtool`.

For each test plan that has any failures, the tool creates a new directory under **/var/hacmp/log/**. If the test plan has no failures, the tool does not create a log directory. The directory name is unique and consists of the name of the Cluster Test Tool plan file, and the time stamp when the test plan was run.

Log file rotation

The Cluster Test Tool saves up to three log files and numbers them so that you can compare the results of different cluster tests. The tool also rotates the files with the oldest file being overwritten. The following list shows the three files saved:

`/var/hacmp/cl_testtool`

`/var/hacmp/cl_testtool`

`/var/hacmp/cl_testtool`

If you do not want the tool to rotate the log files, you can disable this feature from SMIT. For information about turning off this feature, see Running automated tests or Setting up custom cluster testing.

Log file entries

The entries in the log file are in the format:

```
DD/MM/YYYY_hh:mm:ss Message text . . .
```

where `DD/MM/YYYY_hh:mm:ss` indicates day/month/year_hour/minutes/seconds.

The following example shows the type of output stored in the log file:

```
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Initializing Variable Table
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Using Variable File: /tmp/sample_variables
04/02/2006/_13:21:55:  data line: node1=waltham
04/02/2006/_13:21:55:  key: node1 - val: waltham
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Reading Static Configuration Data
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Cluster Name: Test_Cluster
04/02/2006/_13:21:55:   Cluster Version: 7
04/02/2006/_13:21:55:   Local Node Name: waltham
04/02/2006/_13:21:55:   Cluster Nodes: waltham belmont
04/02/2006/_13:21:55:   Found 1 Cluster Networks
04/02/2006/_13:21:55:   Found 4 Cluster Interfaces/Device/Labels
04/02/2006/_13:21:55:   Found 0 Cluster resource groups
04/02/2006/_13:21:55:   Found 0 Cluster Resources
04/02/2006/_13:21:55:   Event Timeout Value: 720
04/02/2006/_13:21:55:   Maximum Timeout Value: 2880
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Building Test Queue
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Test Plan: /tmp/sample_event
04/02/2006/_13:21:55:  Event 1: NODE_UP: NODE_UP,ALL,starts cluster services on all nodes
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Validate NODE_UP
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Event node: ALL
04/02/2006/_13:21:55:   Configured nodes: waltham belmont
04/02/2006/_13:21:55:  Event 2: NODE_DOWN_GRACEFUL:
NODE_DOWN_GRACEFUL,node1,stops cluster services gracefully on node1
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Validate NODE_DOWN_GRACEFUL
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Event node: waltham
04/02/2006/_13:21:55:   Configured nodes: waltham belmont
04/02/2006/_13:21:55:  Event 3: WAIT: WAIT,20
04/02/2006/_13:21:55:  Event 4: NODE_UP: NODE_UP,node1,starts cluster services on node1
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55: | Validate NODE_UP
04/02/2006/_13:21:55: ------------------------------------------------------
04/02/2006/_13:21:55:   Event node: waltham
04/02/2006/_13:21:55:   Configured nodes: waltham belmont
04/02/2006/_13:21:55:
.
.
.
```

Log file example

This topic discusses the log file in depth.

If a test fails, you will see output similar to the following:

```
======================================================================
Test 1 Complete - NETWORK_DOWN_LOCAL: fail service network

Test Completion Status: FAILED

======================================================================

Copying log files hacmp.out and clstrmgr.debug from all nodes to
directory /var/hacmp/cl_testtool/rg_fallover_plan.1144942311
on node prodnode1.
```

After that, you can examine the directory **/var/hacmp/cl_testtool/rg_fallover_plan.1144942311** on node **prodnode1**.

In the log directory, the tool creates separate files for each test. The names for the specific log files stored in the directory have this structure:

`<testnum>.<testname>.<node>.<logfile>`

where:

- `testnum` is the order in which the test appears in the test plan file
- `testname` is the name of the test that failed
- `node` is the node from which the log was collected

- `logfile` the source of the logging information, either the **hacmp.out** or **clstrmgr.debug** files

For example, if the NETWORK_DOWN_LOCAL test fails and it is the first test that was run, and later in the test plan the fourth test, named RG_MOVE also fails, you will see the following files in the **/var/hacmp/cl_testtool/rg_fallover_plan.1144942311** directory:

```
1.NETWORK_DOWN_LOCAL.prodnode1.clstrmgr.debug
1.NETWORK_DOWN_LOCAL.prodnode1.hacmp.out
1.NETWORK_DOWN_LOCAL.prodnode2.clstrmgr.debug
1.NETWORK_DOWN_LOCAL.prodnode2.hacmp.out
4.RG_MOVE.prodnode1.clstrmgr.debug
4.RG_MOVE.prodnode1.hacmp.out
4.RG_MOVE.prodnode2.clstrmgr.debug
4.RG_MOVE.prodnode2.hacmp.out
```

The hacmp.out file

The **hacmp.out** file also logs the start of each test that the Cluster Test Tool runs on each cluster node.

This log entry has the following format:

TestName: *datetimestring1*: *datetimestring2*

where

TestName

The name of the test being processed.

datetimestring1

The date and time on the control node when the Cluster Test Tool starts to run the test. The value of *datetimestring* has the format MMDDHHmmYY (month day hour minute year).

datetimestring2

The date and time on the node on which the test runs. The value of *datetimestring* has the format MMDDHHmmYY (month day hour minute year).

Note: The Cluster Test Tool uses the date and time strings to query the AIX® error log when necessary.

Verbose logging

By default, the Cluster Test Tool uses verbose logging to provide a wealth of information about the results of cluster testing. You can customize the type of information that the tool gathers and stores in the Cluster Test Tool log file.

Note: The Cluster Snapshot utility does not include the Cluster Test Tool log file because this file is specific to PowerHA® SystemMirror® cluster testing at a specific point in time - not an indication of *ongoing* cluster status.

With verbose logging enabled, the Cluster Test Tool:

- Provides detailed information for each test run
- Runs the following utilities on the control node between the processing of one test and the next test in the list:

Utility	Type of Information Collected
clRGinfo	The location and status of resource groups
errpt	Errors stored in the system error log file

- Processes each line in the following files to identify additional information to be included in the Cluster Test Tool log file. The utilities included are run on each node in the cluster after a test finishes running.

File	Type of Information Specified
cl_testtool_log_cmds	A list of utilities to be run to collect additional status information See Customizing the types of information to collect.
cl_testtool_search_strings	Text strings that may be in the **hacmp.out** file. The Cluster Test Tool searches for these strings and inserts any lines that match into the Cluster Test Tool log file. See Adding data from hacmp.out to the cluster test tool log file.

If you want to gather only basic information about the results of cluster testing, you can disable verbose logging for the tool. For information about disabling verbose logging for the Cluster Test Tool, see Running automated tests or Setting up custom cluster testing.

Customizing the types of information to collect

You can customize the types of logging information to be gathered during testing.

When verbose logging is enabled for the Cluster Test Tool, it runs the utilities listed in the **/usr/es/sbin/cluster/etc/cl_testtool_log_cmds** file, and collects status information that the specified commands generate. The Cluster Test Tool runs each of the commands listed in **cl_testtool_log_cmds** file after each test completes, gathers output for each node in the cluster, and stores this information in the Cluster Test Tool log file.

You can collect information specific to a node by adding or removing utilities from the list. For example, if you have an application controller running on two of the nodes in a four-node cluster, you could add application-specific commands to the list on the nodes running the application controllers.

If you want all of the cluster nodes to use the same **cl_testtool_log_cmds** file, you can add it to a file collection. For information about including files in a file collection, see Verifying and synchronizing a PowerHA® SystemMirror® cluster.

By default, the **cl_testtool_log_cmds** file includes the following utilities:

Utility	Type of Information Collected
/usr/es/sbin/cluster/utilities/cldump	A snapshot of the status of key cluster components - the cluster itself, the nodes in the cluster, the network interfaces connected to the nodes, and the resource groups on each node
lssrc -ls clstrmgrES	The status of the Cluster Manager, including a list of any nodes that have been stopped with their resource groups placed in an UNMANAGED state.

The file also contains entries for the following utilities, but they are commented out and *not* run. If you want to run any of these utilities between each test, open the file and remove the comment character from the beginning of the command line for the utility.

Utility	Type of Information Collected
snmpinfo -m dump -v -o /usr/es/sbin/cluster/hacmp.defs cluster	Information on MIB cluster status
snmpinfo -m dump -v -o /usr/sbin/cluster/hacmp.defs resGroupNodeState	Information on MIB resource group state
LANG=C lssrc -a | grep -vw "inoperative$"	The status of all subsystems for each host
svmon -C clstrmgr	Memory usage statistics for the Cluster Manager
/usr/sbin/rsct/bin/hatsdmsinfo	Information about the deadman switch timer
netstat -i ; netstat -r	Information about configured interfaces and routes
lssrc -ls gsclvmd	Information about **gsclvmd** - the access daemon for enhanced concurrent mode volume groups
ps auxw	Process information
lsvg -o	Information about active volume groups (those that are varied on and accessible)
lspv	Information about the physical volumes in a volume group

Utility	Type of Information Collected
vmstat; vmstat -s	System resource utilization information that includes statistics for virtual memory, kernel, disks, traps, and CPU activity

You can also add and remove commands from the **cl_testtool_log_cmds** file.

> **Note:** Enter only one command on each line of the file. The tool executes one command per line.

Adding data from hacmp.out to the cluster test tool log file

You can add messages that include specified text in the **hacmp.out** file to the Cluster Test Tool log file.

With verbose logging enabled, the tool uses the **/usr/es/sbin/cluster/etc/cl_testtool/cl_testtool_search_strings** file to identify text strings to search for in **hacmp.out** . For any text string that you specify on a separate line in the **cl_testtool_search_strings** file, the tool:

- Searches the **hacmp.out** file for a matching string
- Logs the line containing that string, accompanied by the line number from the **hacmp.out** file, to the Cluster Test Tool log file
  You can use the line number to locate the line in the **hacmp.out** file and then review that line within the context of other messages in the file.

By default, the file contains the following lines:

```
!!!!!!!!!!   ERROR   !!!!!!!!!!
 EVENT FAILED
```

You can edit the **cl_testtool_search_strings** file on each node to specify a search string specific to a node. This way, the **cl_testtool_search_strings** file is different on different nodes.

If you want all of the cluster nodes to use the same **cl_testtool_search_strings** file, you can add it to a file collection and synchronize the cluster. For information about including files in a file collection, see Verifying and synchronizing a PowerHA® SystemMirror® cluster.

> **Note:** Cluster synchronization does not propagate a **cl_testtool_search_strings** file to other nodes in a cluster unless the file is part of a file collection.

To edit the **cl_testtool_search_strings** file:

- On each line of the file, specify a single text string that you want the tool to locate in the **hacmp.out** file.

Redirecting all log files to a single directory

This SMIT panel allows you to specify a directory in which to move current logs and redirect future log content.

Before any action is taken, the free disk space is validated within current tolerances. Select **System Management (C-SPOC) > PowerHA® SystemMirror® Logs > Change all Cluster Logs Directory.**

First failure data capture

To prevent loss of critical diagnostic data after a software or node failure, the cluster startup sequence has been enhanced to capture diagnostic data in the `/tmp/ibmsupt/hacmp/ffdc.<timestamp>` directory when a previous failure is being recovered. Only one of these FFDC data captures are retained in case of multiple failures.

Event failures or configuration updates that result in a timeout being reported will save off the event logs in the `/tmp/ibmsupt/hacmp/eventlogs.<date timestamp>` directory on each node in the cluster. A maximum of five of these data collections are retained.

An appropriate message is printed to the log when either of these actions have occurred.

Note: You can disable specific FFDC actions by setting the FFDC_COLLECTION environment variable on each node in the cluster. To disable the FFDC_COLLECTION environment variable on a node, add the following line to the `/etc/environment` file:

```
FFDC_COLLECTION=disable
```

Fixing problems when running cluster tests

This section discusses the following issues that you may encounter when testing a cluster.

Cluster test tool stops running

The Cluster Test Tool can stop running under certain conditions.

These conditions include:

- The Cluster Test Tool fails to initialize
- A test fails and **Abort on Error** is set to **yes** for the test procedure
- The tool times out waiting for cluster stabilization, or the cluster fails to stabilize after a test.
  See Working with timer settings.
- An error that prohibits the Cluster Test Tool from running a test, such as a configuration in AIX® or a script that is missing
- A cluster recovery event fails and requires user intervention.

Control node becomes unavailable

If the control node experiences an unexpected failure while the Cluster Test Tool is running, the testing stops. No action is taken to recover from the failure.

To recover from the failure:

1. Bring the node back online and start cluster services in the usual manner.
   You may need to reboot the control node.
2. Stabilize the cluster.
3. Run the test again.

   Note: The failure of the control node may invalidate the testing that occurred prior to the failure.

If a **CLSTRMGR_KILL** test runs on the control node, the node and cluster services need to restart. For information about handling this situation, see Recovering the control node after cluster manager stops.

Cluster does not return to a stable state

The Cluster Test Tool stops running tests after a timeout if the cluster does not return to a stable state either while a test is running or as a result of a test being processed.

The timeout is based on ongoing cluster activity and the cluster-wide event-duration time until warning values. If the Cluster Test Tool stops running, an error appears on the screen and is logged to the Cluster Test Tool log file before the tool stops running.

After the cluster returns to a stable state, it is possible that the cluster components, such as resource groups, networks, and nodes, are not in a state consistent with the specifications of the list of tests. If the tool cannot run a test due to the state of the cluster, the tool generates an error. The Cluster Test Tool continues to process tests.

If the cluster state does not let you continue a test, you can:

1. Reboot cluster nodes and restart the Cluster Manager.
2. Inspect the Cluster Test Tool log file and the **hacmp.out** file to get more information about what may have happened when the test stopped.
3. Review the timer settings for the following cluster timers, and make sure that the settings are appropriate to your cluster:
    - Time until warning
    - Stabilization interval
    - Monitor interval.

For information about timers in the Cluster Test tool, and about how application monitor timers can affect whether the tool times out, see Working with timer settings.

Working with timer settings

The Cluster Test Tool requires a stable PowerHA® SystemMirror® cluster for testing.

If the cluster becomes unstable, the time that the tool waits for the cluster to stabilize depends on the activity in the cluster:

- No activity.
  The tool waits for twice the time until event duration time until warning (also referred to as **config_too_long)** interval, then times out.
- Activity present.
  The tool calculates a timeout value based on the number of nodes in the cluster and the setting for the time until warning interval.

If the time until warning interval is too short for your cluster, testing may time out. To review or change the setting for the time until warning interval, in PowerHA® SystemMirror® SMIT, select **Custom Cluster Configuration › Events › Cluster Events › Change/Show Time Until Warning** and press Enter.

For complete information on tuning event duration time, see Tuning event duration time until warning.

The settings for the following timers configured for an application monitor can also affect whether testing times out:

- Stabilization interval
- Monitor interval

The settling time for resource groups does not affect whether or not the tool times out.

Stabilization interval for an application monitor

If this timer is active, the Cluster Test Tool does not time out when waiting for cluster stability. If the monitor fails, however, and recovery actions are underway, the Cluster Test Tool may time out before the cluster stabilizes.

Make sure the stabilization interval configured in PowerHA® SystemMirror® is appropriate for the application being monitored.

For information about setting the stabilization interval for an application, see Configuring PowerHA® SystemMirror® cluster topology and resources (extended).

Monitor interval for a custom application monitor

When the Cluster Test Tool runs a **server_down** test, it waits for the length of time specified by the monitor interval before the tool checks for cluster stability. The monitor interval defines how often to poll the application to make sure that the application is running.

The monitor interval should be long enough to allow recovery from a failure. If the monitor interval is too short, the Cluster Test Tool may time out when a recovery is in process.

Testing does not progress as expected

If the Cluster Test Tool is *not* processing tests and recording results as expected, use the Cluster Test Tool log file to try to resolve the problem:

1. Ensure that verbose logging for the tool is enabled.
   For information about verbose logging for the Cluster test Tool, see Error logging.
2. View logging information from the Cluster Test Tool log file **/var/hacmp/log/cl_testtool.log**. The tool directs more information to the log file than to the screen.
3. Add other tools to the **cl_testtool_log_cmds** file to gather additional debugging information. This way you can view this information within the context of the larger log file.

For information about adding commands to the **cl_testtool_log_cmds** file, see Customizing the types of information to collect.

Unexpected test results

The basic measure of success for a test is availability. In some instances, you may consider that a test has passed, when the tool indicates that the test failed. Be sure that you are familiar with the criteria that determines whether a test passes or fails.

For information about the criteria for a test passing or failing, see Evaluating results.

Also ensure that:

- Settings for cluster timers are appropriate to your cluster. See Cluster does not return to a stable state.
- Verbose logging is enabled and customized to investigate an issue. See Testing does not progress as expected.

Starting and stopping cluster services

These topics explain how to start and stop cluster services on cluster nodes and clients.

Starting and stopping cluster services includes these features:

- *Start cluster services*. When you start the cluster services, PowerHA® SystemMirror® by default automatically activates the resources according to how you defined them, taking into consideration application dependencies, application start and stop scripts, dynamic attributes and other parameters. That is, PowerHA® SystemMirror® automatically manages (and activates, if needed) resource groups and applications in them.
  You can also start PowerHA® SystemMirror® with the option to manage resource groups manually. This tells PowerHA® SystemMirror® *not to acquire any resource groups* (and applications) automatically for you. During migration from one version of PowerHA® SystemMirror® to another version you can start and stop cluster services, but you cannot use the manual option when starting cluster services. Manual startup is disabled when migration begins and cannot be used again until the migration is complete.

  You can start PowerHA® SystemMirror® cluster services on the node(s) without stopping your applications, by selecting an option from SMIT (**System Management (C-SPOC) > PowerHA® SystemMirror® Services > Start Cluster Services**).

  PowerHA® SystemMirror® relies on the application monitor and application startup script to verify whether it needs to start the application for you or whether the application is already running. (PowerHA® SystemMirror® attempts *not* to start a second instance of the application.)
- *Stopping cluster services*. When you stop cluster services, you may select one of the following three actions for the resource groups:
  - Bring resource groups Offline.
  - Move resource groups to other node(s).
  - Unmanage resource groups.

  For more information on resource group states, see the section Resource group behavior during cluster events.

Starting cluster services

You can allow your applications that run outside of PowerHA® SystemMirror® to continue running during installation of PowerHA® SystemMirror® and when starting PowerHA® SystemMirror®.

There is no need to stop, restart or reboot the system or applications.

Application monitors

PowerHA® SystemMirror® checks for running applications by using the configured application monitor.

If the monitor indicates that the application is already running, PowerHA® SystemMirror® will not start the second instance of the application. If the application monitors are not configured to PowerHA® SystemMirror®, then you may write an application start script that checks the state of the application before starting it.

Application monitors, configurable in PowerHA® SystemMirror®, are a critical piece of the PowerHA® SystemMirror® cluster configuration; they enable PowerHA® SystemMirror® to keep applications highly available. When PowerHA® SystemMirror® starts an application on a node, it also periodically monitors the application (using the monitor that you configure) to make sure that the application is up and running.

An erroneous application monitor may not detect a failed application. As a result, PowerHA® SystemMirror® would not recover it or may erroneously detect an application as failed, which may cause PowerHA® SystemMirror® to move the application to a takeover node, resulting in unnecessary downtime. To summarize, we highly recommend properly configured and tested application monitors for all applications that you want to keep highly available with the use of PowerHA® SystemMirror®. Use them as follows:

- Use a process monitor if the intent is to monitor whether the process(es) exist on the UNIX™ system.
- Use a custom monitor if the intent is to check the health of the application, for example, whether the database is still functioning by querying a database table.
- Use both process and custom monitors when needed.

During verification, PowerHA® SystemMirror® issues a warning if an application monitor is not configured.

For information on configuring an application monitor, see the section Configuring multiple application monitors.

Procedure for starting cluster services

You can start PowerHA® SystemMirror® cluster services.

To start PowerHA® SystemMirror® cluster services, as the root user, perform the following steps:

> **Note:** Perform the following only after configuring and synchronizing the cluster. For more information, see Configuring a PowerHA® SystemMirror® cluster (standard).

1. Enter `smit cl_admin`
2. In SMIT, select **PowerHA® SystemMirror® Services** > **Start Cluster Services** and press Enter.
3. Enter field values as follows:

Start Cluster Services fields	
Field	**Value**
Start now, on system restart or both	Indicate how you want to start cluster services when you commit the values on this panel by pressing Enter (**now**), when the operating system reboots by selecting **on system 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.

Field	Value
	When you start the PowerHA® SystemMirror® cluster services with the **Manage Resource Group** option set to **Manually**, and select the option **both**, the timing of a power loss or rebooting the node may affect whether the node is in the OFFLINE or UNMANAGED state after the system reboot.
Start Cluster Services on these nodes	Enter the name(s) of one or more nodes on which you want to start cluster services. Alternatively, you can select nodes from a picklist. Separate multiple nodes with a comma.
Manage resource groups	**Automatically** (default). PowerHA® SystemMirror® brings resource group(s) online according to the resource groups' configuration settings and the current cluster state and starts managing the resource group(s) and applications for availability. When you start PowerHA® SystemMirror® cluster services and set the Manage Resource Group option to Automatically, PowerHA® SystemMirror® automatically activates resource groups on the node(s) according to their policies and locations and also starts applications. PowerHA® SystemMirror® may not necessarily start the application on the same node on which it is currently being run, if the application is already running. That is, when this option is selected, PowerHA® SystemMirror® determines on which node to bring a resource group online based on the configured resource group policies, resource group dependency configuration and available resources on the node. If you select this option while starting the cluster services, it is suggested to stop the applications and resources so that PowerHA® SystemMirror® can start them on the appropriate node. See also Running corrective actions during verification.
	Manually. PowerHA® SystemMirror® does not activate resource groups while the cluster services on the selected node are started. After you start cluster services, you can bring any resource groups online or offline, as needed, using the **PowerHA® SystemMirror® Resource Group and Application Management** SMIT menu (**clRGmove**).
BROADCAST message at startup?	Indicate whether you want to send a broadcast message to all nodes when the cluster services start. The default is **true**.
Startup Cluster Information Daemon?	Indicate whether you want to start the **clinfoES** daemon. For example, if your application uses the Cluster Information daemon, if you use the **clstat** monitor, set this field to **true.** Otherwise, set it to **false**. The value that you enter in the **Startup Cluster Information Services?** field works 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 true and the Start now, on system restart or both field to both , then the **clinfoES** daemon is also started whenever the cluster services are started.
Ignore Verification Errors?	Set this value to false (the default) to stop all selected nodes from starting cluster services if verification finds errors on any node.

Field	Value
	Set this value to **true** to start cluster services even if verification finds errors on the specified nodes or in the cluster in general. This setting should be used with caution.
Automatically correct errors found during cluster start?	This field is available only if the automatic verification and synchronization option has been enabled. For more information, see Modifying the startup of cluster services. • Select **Interactively** to receive prompts to correct certain errors as they are found during verification. • Select **No** if you do not want PowerHA® SystemMirror® to correct any verification errors automatically. If you select **No**, you must correct errors, if any, manually. • Select **Yes** if you want PowerHA® SystemMirror® to correct cluster verification errors automatically without first prompting you. Not all verification errors are automatically corrected; some must be corrected manually. For more information, see Automatic verification and synchronization.

4. Press Enter.
   The system performs verification and synchronization as needed, and then 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 (for example, the number of disks, the number of interfaces to configure, the number of file systems to mount, and the number of applications being started).

   SMIT displays a command status window. Note that when the SMIT panel indicates the completion of the cluster startup, PowerHA® SystemMirror® processing of the resource groups in most cases has not yet completed. To verify that the processing has completed, use **/usr/es/sbin/cluster/clstat**, described in Monitoring a PowerHA® SystemMirror® cluster.

Starting PowerHA® SystemMirror® cluster services with manually managed resource groups

Set the cluster services **Manage Resource Group** startup option to **Manually** when you want more control over the node on which an application should run. This method ensures that the services provided by the application controller are not interrupted.

When you choose this option to start the PowerHA® SystemMirror® cluster services, the resource groups on the node remain in the OFFLINE or UNMANAGED state, depending on whether this is a cold start up or a start after the node was stopped and resource groups placed in an UNMANAGED state.

Note: Please be advised that if a resource group is in the UNMANAGED state, it does *not* mean that, from PowerHA® SystemMirror's point of view, the actual resources in the resource group are *not* running. To PowerHA® SystemMirror®, it means that PowerHA® SystemMirror® is *not* managing the resources (and the applications) of the resource group for availability.

Note that either you must have an application monitor configured that PowerHA® SystemMirror® uses to check the application or your application start scripts should be intelligent enough not to start the application if it is already running.

If you want to activate resource groups that are *not* brought online automatically, use the Resource Group Management utility (**clRGmove**) to bring the OFFLINE state resource groups to the ONLINE state.

Consider the following example: If an application is running on a node that is *not* the primary node, and during the startup process you know that PowerHA® SystemMirror® will move the resource group with the application to another node (according to the resource group policy specified), starting PowerHA® SystemMirror® cluster services with the **Manage Resource Group** option set to **Manually** tells PowerHA® SystemMirror® *not* to start the resource groups during startup. You can later use the user-requested **rg-move** to bring the resource group to the ONLINE state on the same node where your application is already running.

To start cluster services on a resource group that is manually managed:

1. Enter `smitty sysmirror`
2. **System Management (C-SPOC) › Resource Group and Applications › Bring Resource Group Online**.
3. Select the node where your application is running.
4. Press Enter.

Starting cluster services on a node with a resource group in the UNMANAGED state

Resource groups may be in the UNMANAGED state on a node if cluster services on that node have been stopped using the **Unmanage resource groups** option.

This **Unmanage resource groups** option causes PowerHA® SystemMirror® to stop providing high availability services to the resource group; that is, the resource groups will not fall over due to resource failures. This option is intended for temporary situations, such as when you want to upgrade PowerHA® SystemMirror® or perform maintenance without bringing your applications offline.

You can start cluster services on the node after it had been stopped with the resource group option set to UNMANAGED. It puts any resource group that is in the UNMANAGED state on that node back to the state in which it was prior to being UNMANAGED. While bringing the resource group ONLINE from the UNMANAGED state, PowerHA® SystemMirror® checks every resource in the resource group to see whether it is active and activates it if it is found inactive. Thus, it is critical to configure the application monitors with `startup` or `both` monitor modes so that the PowerHA® SystemMirror® can correctly detect a running application and so PowerHA® SystemMirror® does not try to start a second instance.

If you start cluster services on the node where the parent resource group is in an UNMANAGED state in a parent-and-child resource-group configuration that has different home nodes, the corresponding child resource group is not released and reacquired.

If you are using PowerHA® SystemMirror® Version 7.2.5, or later, after you restart the cluster services that is stopped by using the **Unmanaged resource groups** option, you can skip cluster resource processing by using the `SKIP_EVENT_PROCESSING_MANAGE_MODE` tunable attribute. This can be an option if you have no application monitors configured or there are only application monitors with long-running monitor mode available. By default, the `SKIP_EVENT_PROCESSING_MANAGE_MODE` tunable attribute is disabled and the initial configuration of cluster services that was set during cluster startup is preserved. If you enable the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable attribute, it restores the state of the cluster resource group to the previous state without disturbing the current state during the cluster restart operation. To enable the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable attribute, you must modify the cluster definition by using the **`clmgr`** command. After you enable the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable attribute in the `hacmp.out` log file, the new **forced** option of the node_up events directs the node_up event to skip processing of the cluster resources.

Other:

- You can use the `SKIP_EVENT_PROCESSING_MANAGE_MODE` tunable attribute to restart the operation of the cluster services, after it is stopped by using the **Unmanaged resource groups** option.
- If you use the `SKIP_EVENT_PROCESSING_MANAGE_MODE` tunable attribute, any cluster resource processing operation is ignored after the cluster services are restarted. After you restart the node, if the cluster resource processing is ignored, you must ensure that the current state of the cluster resource group is not changed when the node is stopped by using the **Unmanage resource groups** option. If the current state of the cluster resource group is changed when cluster services are in the UNMANAGED state, you must manually restore the cluster resources to its previous state once the cluster services are started.

The purpose of the following steps is to give you the option to bring the resource group online on a different node if it is in the UNMANAGED state on a specified node because the current node might be down for prolonged maintenance. Regardless of the state of the node, the cluster manager might be in a FORCED DOWN state on that node, or the system might be down, or might have been rebooted. Bringing the resource group to the OFFLINE state on this node does not affect the state of its resources. If the resource group is online on this node, you need to manually bring it to the OFFLINE state.

In cases where you want to bring a resource group from an UNMANAGED state to an ONLINE state on a different node (because the node that was stopped by using UNMANAGED option is unavailable), you should do the following:

1. Bring the resource groups to the OFFLINE state using a user-requested **rg-move** SMIT panel. Note that during this operation, PowerHA® SystemMirror® will not stop any resources as the node that originally hosted the resource group is no longer available.
2. Ensure that all the resources that are configured in the resource group are OFFLINE, including the application, if any.
3. Bring the resource groups from their OFFLINE state to the ONLINE state, just as was necessary in previous releases using the resource group migration utility **clRGmove** or the SMIT option.

Modifying the startup of cluster services

Typically, you should use the default cluster services startup settings - especially the verification setting, which is automatically enabled to ensure a safe startup. However, you can modify these settings by following the procedure described in the following section.

To modify the startup of cluster services:

1. From the command line, enter `smit sysmirror`
2. Select **Custom Cluster Configuration › Cluster Nodes and Networks › Manage the Cluster › Cluster Startup Settings** and press Enter.
3. Enter field values in the SMIT panel as follows:

Cluster Startup Settings fields	
Field	**Value**
Start PowerHA® SystemMirror® at system restart	False is the default. This removes the entry from the /etc/inittab file and will *not* automatically start cluster services at system restart. True starts the daemons after a system reboot by adding an entry to the /etc/inittab file.
BROADCAST message at startup	True is the default. This broadcasts a message to the console, indicating that cluster services are starting.
Startup Cluster Information Daemon?	False is the default. True starts the clinfo daemon, which allows clstat and xclstat (or any third-party application written against the clinfo API) to read changes in the cluster state.

Stopping cluster services

These topics describe the process of stopping 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 Configuration Database 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 a network interface, you must stop and restart cluster services on that node or on all nodes, depending on the cluster setup.

For more information about which changes to the cluster require PowerHA® SystemMirror® reconfiguration, see 7x24 maintenance.

When stopping cluster services, 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 users know when services will be restored.

Procedure for stopping cluster services

This topic describes the procedure for stopping cluster services on a single node or on all nodes in a cluster by using the C-SPOC utility on one of the cluster nodes.

To stop cluster services:

1. Enter the fastpath `smit cl_admin`.
2. Select **PowerHA® SystemMirror® Services > Stop Cluster Services** and press Enter.
3. Enter field values in the SMIT panel as follows:

Stop Cluster Services fields	
Field	**Value**
Select an Action on resource groups	Indicate the type of shutdown: • **Bring resource groups Offline**. PowerHA® SystemMirror® stops all managed resources currently ONLINE on the node being stopped. PowerHA® SystemMirror® will not activate these resources on any other nodes, that is, no fallover. This option is equivalent to the option to stopping cluster services gracefully in previous releases. After successfully stopping all managed resources, PowerHA® SystemMirror® stops RSCT services and goes into ST_INIT state. • **Move resource groups**. PowerHA® SystemMirror® stops all managed resources currently ONLINE on the node being stopped. The resource groups will be moved to a takeover node according to the configured resource group policies (if defined), dependency configurations (if defined) and available resources. This option is equivalent to the graceful with takeover option in previous releases. After successfully stopping all managed resources PowerHA® SystemMirror®, stops RSCT services and the Cluster Manager daemon goes into ST_INIT state. • **Unmanage resource groups**. The cluster services are stopped immediately. Resources that are online on the node are not stopped. Applications continue to run. This option is equivalent to the forced down option in previous releases. For more information, see Stopping PowerHA® SystemMirror® cluster services without stopping applications. PowerHA® SystemMirror® will not stop the managed resources; applications remain functional. PowerHA® SystemMirror® does not manage the resources on these nodes. PowerHA® SystemMirror® continues to run and RSCT remains functional. On a node that has Enhanced concurrent (ECM) volume groups, cluster services can be stopped with the resource groups

Field	Value
	placed in an unmanaged state. RSCT services will continue to run so that ECM remains functional. If you stop cluster services with this option, the resource groups that are active on this node go into unmanaged state. Once the resource group is in the unmanaged state, PowerHA® SystemMirror® does not process any resource failures. This applies to hardware resources such as disks and adapters as well as any managed applications. Refer to Procedure for starting cluster services for information on reintegrating a node on which the cluster services were stopped back into the cluster.
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 select **restart** or **both**, the entry in the **/etc/inittab** file that starts cluster services is removed. 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. If you specify **true**, a message is broadcast on all cluster nodes.

4. Press Enter. The system stops the cluster services on the nodes specified.
   If the stop operation fails, check the **/var/hacmp/log/cspoc.log** file for error messages. This file contains the command execution status of the C-SPOC command executed on each cluster node.

> After stopping cluster services, you must wait a minimum of two minutes for the RSCT to quiesce before starting cluster services.

Stopping PowerHA® SystemMirror® cluster services without stopping applications

You can stop cluster services without stopping services and applications.

To stop cluster services without stopping your applications, complete the following steps:

1. From the command line, enter `smit cspoc`.
2. In C-SPOC, select **PowerHA® SystemMirror® Services › Stop Cluster Services** and press Enter.
3. Complete the required fields and press Enter.

Result
No matter what type of resource group you have, if you stop cluster services on the node on which this group is active and do not stop the application that belongs to the resource group, PowerHA® SystemMirror® puts the group into an UNMANAGED state and keeps the application running according to your request.

The resource group that contains the application remains in the UNMANAGED state (until you tell PowerHA® SystemMirror® to start managing it again) and the application continues to run. While in this condition, PowerHA® SystemMirror® and the RSCT services continue to run, providing services to ECM VGs that the application controllers might be using.

You can tell PowerHA® SystemMirror® to start managing it again either by restarting Cluster Services on the node, or by using SMIT to move the resource group to a node that is actively managing its resource groups. See Starting PowerHA® SystemMirror® cluster services with manually managed resource groups for more information.

If you have instances of replicated resource groups using the Extended Distance capabilities of the PowerHA® SystemMirror® Enterprise Edition product, the UNMANAGED SECONDARY state is used for resource groups that were previously in the ONLINE SECONDARY state.

You can view the new states of the resource groups using the cluster utilities **clstat** and **clRGinfo**.

You cannot dynamically reconfigure (DARE) the cluster configuration while some cluster nodes have resource groups in the unmanaged state.

Warning about placing resource groups in an unmanaged state

When you stop cluster services by using the **Unmanage Resource Groups** option, the resources and resource groups on that node are neither taken offline, nor they are monitored for failures.

The state of the resource groups on the unmanaged node is maintained by other active nodes in the cluster, even if the stopped node is rebooted.

If you stop the cluster services on all nodes by using the **Unmanage Resource Groups** option, the cluster state information might be lost. Before you restart the cluster services, you must stop the cluster services on all nodes by using the **Resource Groups Offline** option.

Abnormal termination of Cluster Manager daemon

The AIX® resource controller subsystem monitors the cluster manager daemon process. If the controller detects that Cluster Manager daemon has exited abnormally (without being shut down using the **clstop** command), it executes the **/usr/es/sbin/cluster/utilities/clexit.rc** script to halt the system. This prevents unpredictable behavior from corrupting the data on the shared disks.

See the clexit.rc man page for additional information.

The clexit.rc script creates an AIX® error log entry. Here is an example showing the long output:

```
LABEL: OPMSG
IDENTIFIER:     AA8AB241

Date/Time:Fri Jan  7 10:44:46
Sequence Number: 626
Machine Id:      000001331000
Node Id:ppstest8
Class:  O
Type:   TEMP
Resource Name:   OPERATOR

Description
OPERATOR NOTIFICATION

User Causes
ERRLOGGER COMMAND

 Recommended Actions
 REVIEW DETAILED DATA

Detail Data
MESSAGE FROM ERRLOGGER COMMAND
clexit.rc : Unexpected termination of clstrmgrES
```

The clexit.rc error message in short form looks like this:

```
AA8AB241   0107104400 T O OPERATOROPERATOR NOTIFICATION
```

Important: Never use the kill -9 command on the clstrmgr daemon. Using the kill command causes the clstrmgr daemon to exit abnormally. This causes the System Resource Controller (SRC) facility to run the script **/usr/es/sbin/cluster/utilities/clexit.rc** , which halts the system immediately and causes the surviving nodes to initiate fallover.

You can modify the file /etc/cluster/hacmp.term to change the default action after an abnormal exit. The clexit.rc script checks for the presence of this file, and if you have made it executable, the instructions there will be followed instead of the automatic halt called by **clexit.rc** . Please read the caveats contained in the /etc/cluster/hacmp.term file, however, before making any modifications.

Shutting down the AIX operating system in a PowerHA® SystemMirror® node

In PowerHA® SystemMirror®, before you stop or restart the AIX® operating system, ensure that all critical applications are either quiesced or moved onto a backup node.

If you do not quiesce or move the critical application onto a backup node, PowerHA® SystemMirror® might not be able to take all system resources offline. To avoid system failure due to loss of system resources, PowerHA® SystemMirror® might initiate a restart or a failover operation. Also, PowerHA® SystemMirror® might not release a shared system resource like a Network File System (NFS) mount point that might prevent or delay other cluster nodes from acquiring that shared system resource. Delay in acquiring important system resources by the cluster nodes affects critical application availability and causes data corruption.

When the AIX® operating system is shut down on a node where the PowerHA® SystemMirror® cluster services are active, the Cluster Manager daemon (clstrmgr) either recovers the resource groups on a takeover node or leaves them in an offline state depending on the flags that are passed to the **shutdown** or **restart** command.

If you use the AIX® **shutdown** command with the `-F` or `-r` flag or a combination thereof, PowerHA® SystemMirror® does not attempt to recover the resource groups on a backup node. Hence, when the AIX® operating system is started on a node where the PowerHA® SystemMirror® cluster services are active, the resource groups are also started on the same node. If you do not use `-F` or `-r` flag or a combination thereof, PowerHA® SystemMirror® attempts to bring the cluster resources offline. When the cluster resources are offline, AIX® stops the critical resources and moves the resource groups onto a takeover node.

In a complex configuration with many resource groups, volume groups, and file systems, stopping the AIX® operating system might delete the **cthags** subsystem when PowerHA® SystemMirror® attempts to release the cluster resources. If the **cthags** subsystem is deleted when the file systems are mounted, the Logical Volume Manager (LVM) subsystem to moves all the shared volume groups to an offline state. This might corrupt the file systems.

If you use the AIX® **restart** command while the PowerHA® SystemMirror® cluster services are active, the remote cluster nodes interprets it as a system failure and initiates the recovery process.

Stopping PowerHA® SystemMirror® cluster services and RSCT

PowerHA® SystemMirror® manages the RSCT services automatically.

When users stop cluster services using the Move Resource Group option, the RSCT services are stopped after all the resources and applications on the node are released. When users select the Unmanage Resource Group option to stop the cluster services, the Cluster Manager puts the resource groups into the UNMANAGED state but continues to run under the covers thus leaving the RSCT services up and running under this condition

One of the reasons that PowerHA® SystemMirror® does *not* stop the RSCT services from running when you stop cluster services is because *not* only PowerHA® SystemMirror® but also the Enhanced Concurrent Mode (ECM) volume groups use RSCT services. Stopping RSCT services would vary off the ECM volume group and would affect the application that is using it.

There could be rare cases when you need to stop RSCT, for example, to perform an RSCT upgrade. If you need to upgrade RSCT, you can stop and restart it by using SMIT options under the **Problem Determination Tools** menu.

Maintaining cluster information services

The cluster services on clients consist solely of the clinfoES daemon, which provides clients with status information about the cluster.

Note that the /etc/inittab file is modified when the PowerHA® SystemMirror® software is installed to start the clinfoES daemon whenever the system is rebooted.

The Cluster Information Daemon (**clinfo**) retrieves information about the cluster configuration and the state of the cluster, topology and resources from the Management Information Base (MIB) and the Cluster Manager on local or remote nodes. The Cluster Manager updates the MIB with this information.

The **clinfo** daemon populates internal, dynamically allocated data structures with information for each cluster. The cluster(s) can be any combination of local or remote. The **clinfo** daemon calls the **clinfo.rc** script in response to cluster changes.

Starting Clinfo on a client

Use the /usr/es/sbin/cluster/etc/rc.cluster script or the startsrc command to start clinfo on a client.

See the following example:

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

You can also use the standard AIX® startsrc command:

```
startsrc -s clinfoES
```

Stopping Clinfo on a client

Use the standard AIX® stopsrc command to stop clinfo on a client machine.

See the following example:

```
stopsrc -s clinfoES
```

Enabling clinfo for asynchronous event notification

In PowerHA® SystemMirror®, the **clinfo** daemon obtains data only from Simple Network Management Protocol (SNMP). You can configure PowerHA® SystemMirror® to use the **clinfo** daemon to receive notification of events as asynchronous messages (otherwise known as traps).

Only one SNMP application can receive traps. If your system is running the NetView® for AIX® licensed product, you cannot enable the **clinfo** daemon to receive traps.

To enable asynchronous event notification, complete the following steps:

1. To start the **clinfo** daemon, enter `chssys -s clinfoES -a "-a"` from the command line.
2. To verify that the System Resource Controller (SRC) has the correct command-line arguments for the **clinfo** daemon, enter `lssrc -Ss clinfoES` from the command line.
3. Edit the `/etc/snmpdv3.conf` file on the nodes that sends traps. As installed, traps are directed to the loopback address. The **clinfo** daemon receives traps that are generated by the Cluster Manager on the same node. Review the comments at the beginning of the `/etc/snmpdv3.conf` file for a description of all fields.

   > SNMP version 3 is the default version that is used by the AIX® operating system.

   a. Find the trap line at the end of the file. The following code is an example of the trap line:

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

   b. Add trap lines as wanted. Multiple **clinfo** daemon processes can receive traps from the Cluster Manager. Make sure that the `1.2.3 fe` field is unique.
   The following is an 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
   ```

   c. Stop the **snmpd** daemon on the hosts where you changed the `/etc/snmpdv3.conf` file, enter `stopsrc -s snmpd` from the command line.

   d. Start the **snmpd** daemon on the hosts where you changed the `/etc/snmpdv3.conf` file, enter `startsrc -s snmpd` from the command line.

Monitoring a PowerHA® SystemMirror® cluster

These topics describe tools you can use to monitor a PowerHA® SystemMirror® cluster.

You can use SMIT to configure and manage the cluster and view interactive cluster status.

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

Periodically monitoring a PowerHA® SystemMirror® cluster

PowerHA® SystemMirror® provides recovery for various failures that occur within a cluster. For example, PowerHA® SystemMirror® can compensate for a network interface failure by swapping in a boot interface. 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 PowerHA® SystemMirror® can survive one or possibly several failures, each failure that escapes your notice threatens a cluster's ability to provide a highly available environment, as the redundancy of cluster components is diminished.

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.

In addition, PowerHA® SystemMirror® offers application monitoring capability that you can configure and customize in order to monitor the health of specific applications and processes.

Use the AIX® Error Notification facility to add an additional layer of high availability to a PowerHA® SystemMirror® environment. You can add notification for failures of resources for which PowerHA® SystemMirror® does *not* provide recovery by default. The combination of PowerHA® SystemMirror® and the high availability features built into the AIX® system keeps single points of failure to a minimum; the Error Notification facility can further enhance the availability of your particular environment.

Automatic cluster configuration monitoring

Verification automatically runs on one user-selectable PowerHA® SystemMirror® cluster node once every 24 hours.

By default, the first node in alphabetical order runs the verification at midnight. If **verification** finds errors, it warns about recent configuration issues that might cause problems at some point in the future. PowerHA® SystemMirror® stores the results of the automatic monitoring on every available cluster node in the **/var/hacmp/log/clutils.log** file.

If cluster verification detects some configuration errors, you are notified about the potential problems:

- The exit status of **verification** is published across the cluster along with the information about cluster verification process completion.
- Broadcast messages are sent across the cluster and displayed on **stdout**. These messages inform you about detected configuration errors.
- A **cluster_notify** event runs on the cluster and is logged in **hacmp.out** (if cluster services is running).

More detailed information is available on the node that completes cluster verification in **/var/hacmp/clverify/clverify.log** file. If a failure occurs during processing, error messages and warnings clearly indicate the node and reasons for the **verification** failure.

Tools for monitoring a PowerHA® SystemMirror® cluster

PowerHA® SystemMirror® supplies tools for monitoring a cluster.

These are described in subsequent sections:

- **Cluster Monitoring with Tivoli®** allows you to monitor clusters and cluster components and perform cluster administration tasks through your Tivoli® Framework console.

- **clstat** (the **/usr/es/sbin/cluster/clstat** utility) reports the status of key cluster components - the cluster itself, the nodes in the cluster, the network interfaces connected to the nodes, the service labels, and the resource groups on each node.
- **Application Monitoring** allows you to monitor specific applications and processes and define action to take upon detection of process death or other application failures. Application monitors can watch for the successful startup of the application, check that the application runs successfully after the stabilization interval has passed, or monitor both the startup and the long-running process.
- SMIT provides you information on the cluster.
  You have the ability to see the cluster from an application-centric point of view.
  - The **Resource Group and Applications** menu under C-SPOC (`smit cl_admin`) in SMIT has an option to **Show the Current State of Applications and Resource Groups**. You can also reach this panel from **Cluster Applications and Resources › Resource Groups › Show All Resources by Node or Resource Group** (from `smit sysmirror`).

  The **System Management (C-SPOC) > PowerHA® SystemMirror® Services > Show Cluster Services** SMIT panel shows the status of the PowerHA® SystemMirror® daemons.
- The **Application Availability Analysis** tool measures uptime statistics for applications with application controllers defined to PowerHA® SystemMirror®.
- The **clRGinfo** and **cltopinfo** commands display useful information on resource group configuration and status and topology configuration, respectively. For more information, see (Groups Information Commands).
- **Log files** allow you to track cluster events and history: The **/var/hacmp/adm/cluster.log** file tracks cluster events; the **/var/hacmp/log/hacmp.out** file records the output generated by configuration scripts as they execute; the **/var/hacmp/adm/history/cluster**. *mmddyyyy* log file logs the daily cluster history; the **/var/hacmp/log/cspoc.log** file logs the status of C-SPOC commands executed on cluster nodes. You should also check the RSCT log files.

In addition to these cluster monitoring tools, you can use the **Custom Remote Notification** utility allows you to define a notification method through the SMIT interface to issue a customized page in response to a cluster event. You can also send text messaging notification to any address including a cell phone.

Monitoring clusters with clstat

PowerHA® SystemMirror® provides the /usr/es/sbin/cluster/clstat utility for monitoring a cluster and its components. The **clinfo** daemon must be running on the local node for this utility to work properly.

The clstat utility reports on the cluster components as follows:

- Cluster: cluster number (system-assigned); cluster state (up or down); cluster substate (stable, or unstable).
- Nodes: How many, and the state of each node (up, down, joining, leaving, or reconfiguring).
  For each node, clstat displays the IP label and IP address of each network interface attached to each node, and whether that interface is up or down. clstat does *not* display multiple IP labels on one network interface, as in networks with aliases.

  For each node, **clstat** displays service IP labels for serial networks and whether they are up or down.

  > **Note:** By default, **clstat** does *not* display whether the service IP labels for serial networks are down. Use **clstat -s** to display service IP labels on serial networks that are currently down.

  For each node, clstat displays the states of any resource groups (per node): online or offline.

See the clstat man page for additional information.

The /usr/es/sbin/cluster/clstat utility runs on both ASCII and X Window Display clients in either single-cluster or multi-cluster mode. 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.

Viewing clstat in ASCII display mode

In ASCII display mode, you have the option of viewing status for a single cluster or multiple clusters.

You can also use the **-o** option to save a single snapshot of the **clstat** output in a **cron** job.

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:

```
/usr/es/sbin/cluster/clstat
```

A panel similar to the following appears:

```
clstat - PowerHA SystemMirror Cluster Status Monitor
      -------------------------------------

Cluster: myctestcluster   (1044370190)
Tue Mar 11 14:19:50 EST 2004
      State: UP     Nodes: 2
      SubState: STABLE

 Node: holmes  State: UP
 Interface: holmes_en1svc (0)  Address: 192.168.90.40
 State:   UP
 Resource Group: econrg1  State:  online

 Node: u853    State: UP
 Interface: u853_en1svc (0) Address: 192.168.90.50
 State:   UP
 Resource Group: econrg1  State:  online
****************** f/forward, b/back, r/refresh, q/quit ***************
```

clstat single-cluster ASCII display mode

The cluster information displayed shows the cluster ID and name. (Note that PowerHA® SystemMirror® assigns the cluster ID number; this is *not* user-defined.) In this example, the cluster is up and has two nodes, both of which are up. Each node has one network interface. 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:

```
usage: clstat [-c cluster ID] [-n cluster name][ -r seconds] [-i] [-a] [-o] [-s]
```

where:

clstat flags	
Flag	**Description**
-c *cluster ID*	Displays information about the cluster with the specified ID if that cluster is active (PowerHA® SystemMirror® generates this number). This option can*not* be used with the -n option. If the cluster is *not* available, the clstat utility continues looking for it until it is found or until the program is canceled. Note that this option can*not* be used if the -i option (for multi-cluster mode) is used.
-n name	The cluster name. This option can*not* be used with the -c option
-r seconds	Updates the cluster status display at the specified number of seconds. The default is 1 second;

Flag	Description
	however, the display is updated only if the cluster state changes.
-i	Displays information about clusters interactively. Only valid when running clstat in ASCII mode.
-a	Causes clstat to display in ASCII mode.
-o	(once) Provides a single snapshot of the cluster state and exits. This flag can be used to run **clstat** out of a **cron** job. Must be run with the **-a** option; ignores **-i** or **-r** flags.
-s	Displays service labels for serial networks and their state (up or down).

To see cluster information about a specific cluster, enter:

```
clstat [-n name]
```

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/es/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. Multi-cluster mode requires that you use the -i flag when invoking the clstat utility. To invoke the clstat utility in multi-cluster mode, enter:

```
/use/es/sbin/cluster/clstat -i
```

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

```
    clstat - PowerHA SystemMirror for AIX Cluster Status Monitor
    ------------------------------------------------

Number of clusters active: 1

      ID      Name  State

    777 ibm_26c  UP

Select an option:

    # - the Cluster ID  x- quit
```

clstat multi-cluster mode menu

This panel displays the ID, name, and state of each active cluster accessible by the local node. You can either select a cluster to see detailed information, or quit the clstat utility.

When you enter a cluster name, a panel appears similar to the one that follows.

```
    clstat - PowerHA SystemMirror for AIX Cluster Status Monitor
      ------------------------------------------------

Cluster: ibm_26c (777) Thu Jul  9 18:35:46 EDT 2002
      State: UP Nodes: 2
      SubState: STABLE
 Node: poseidonState: UP
 Interface: poseidon-enboot (0)Address: 140.186.70.106
 State:   UP

 Node: venus   State: UP
 Interface: venus-enboot (0)Address: 140.186.70.107
 State:   UP

   Resource Group: rotState: online
```

```
   Resource Gropu: rg1State: online
****************** f/forward, b/back, r/refresh, q/quit *************
```

clstat multi-cluster ASCII display mode

After viewing this panel, 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 panel.

Viewing clstat in X Window System display mode

When you start the /usr/es/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:

```
/usr/es/sbin/cluster/clstat [-n name][-c Id][ -r #][-D debug_level][-s]
```

where:

Flag name	Description
-n name	The cluster name. This option can*not* be used with the -c option.
-c ID	Displays information about the cluster with the specified ID if that cluster is active. This option can*not* 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 10 in increasing amounts of information. The default (0) turns debugging off.
-s	Displays service labels for serial networks and their state (up or down).

The clstat utility graphical interface uses windows to represent cluster nodes, as in the figure shown here:

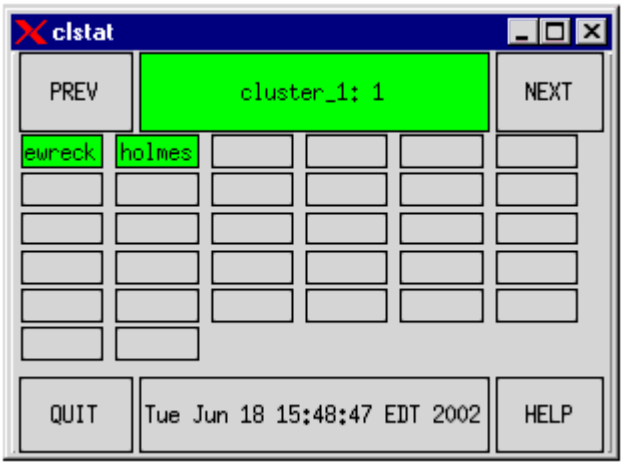


Figure 1: clstat X Window System display

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

The large boxes in other rows represent nodes. A node name appears in a box for each active node in the cluster. You can see up to sixteen nodes 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/es/sbin/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 and resource group information in an X Window display

To view information about network interfaces and resource groups for 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 *holmes* in *cluster_1*.

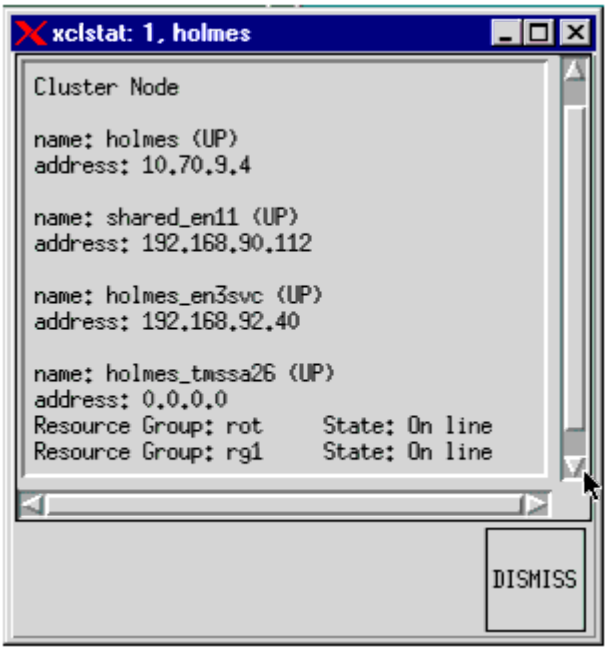


Figure 2: clstat node information display

clstat displays only the state (online or offline) of resource groups.

Click on 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 of the window to close this display; it terminates the clstat utility.

Viewing clstat with a Web browser

With an appropriately configured Web server, you can view **clstat** in a Web browser on any machine that can connect to the cluster node (a node with both a Web server and Clinfo running on it).

Viewing **clstat** through a Web browser allows you to see status for all of your clusters on one panel, using hyperlinks or the scroll bar to view details for each cluster.

When you install PowerHA® SystemMirror®, an executable file called **clstat.cgi** is installed in the same directory (**/usr/es/sbin/cluster/**) as the **clstat** and **xclstat** files . When run, **clstat.cgi** provides a CGI interface that allows cluster status output to be formatted in HTML and viewed in a Web browser.

This feature supports the following browsers:

- Mozilla 1.7.3 for AIX® and FireFox 1.0.6
- Internet Explorer, version 6.0.

Browser display

The **clstat** PowerHA® SystemMirror® Cluster Status Monitor displays the **clstat** output for all clusters from the list of active service IP labels or addresses found in the /usr/es/sbin/cluster/etc/clhosts file.

The following example shows clstat monitoring two clusters, *cluster_1* and *cluster_222* . The browser window displays the status information for one of the clusters, *cluster_1* . To display the other cluster, click the hyperlink for *cluster_222* at the top of the display or scroll down to find it.

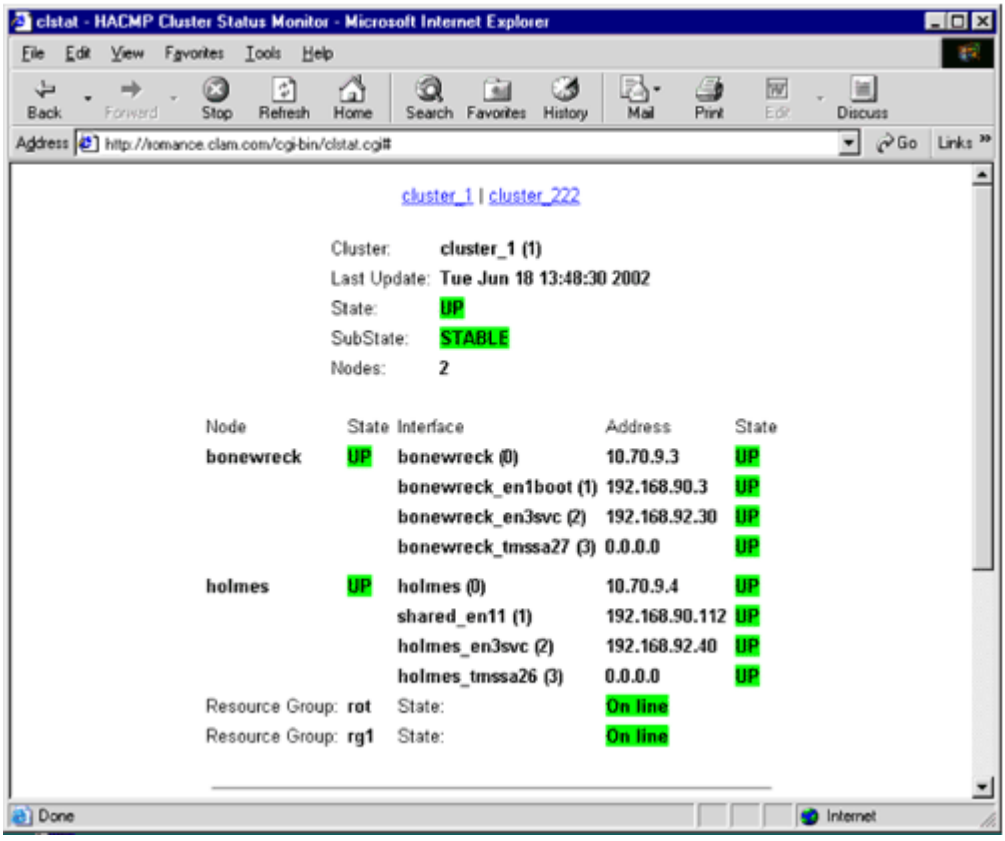


Figure 3: clstat web browser display

The web browser display contains the same types of cluster status information as the ASCII or X Window displays, reorganized and color-coded for easier viewing.

The view automatically refreshes every 30 seconds to display current cluster status.

> **Note:** After an automatic or manual refresh, the view should be retained; that is, the browser window should continue to display the cluster that was last clicked on before the refresh. In Internet Explorer 5.5 only, however, the refresh action causes a return to the top of the display.

In the following example, one of the resource groups is coming online and the cluster is therefore in a reconfiguration substate:

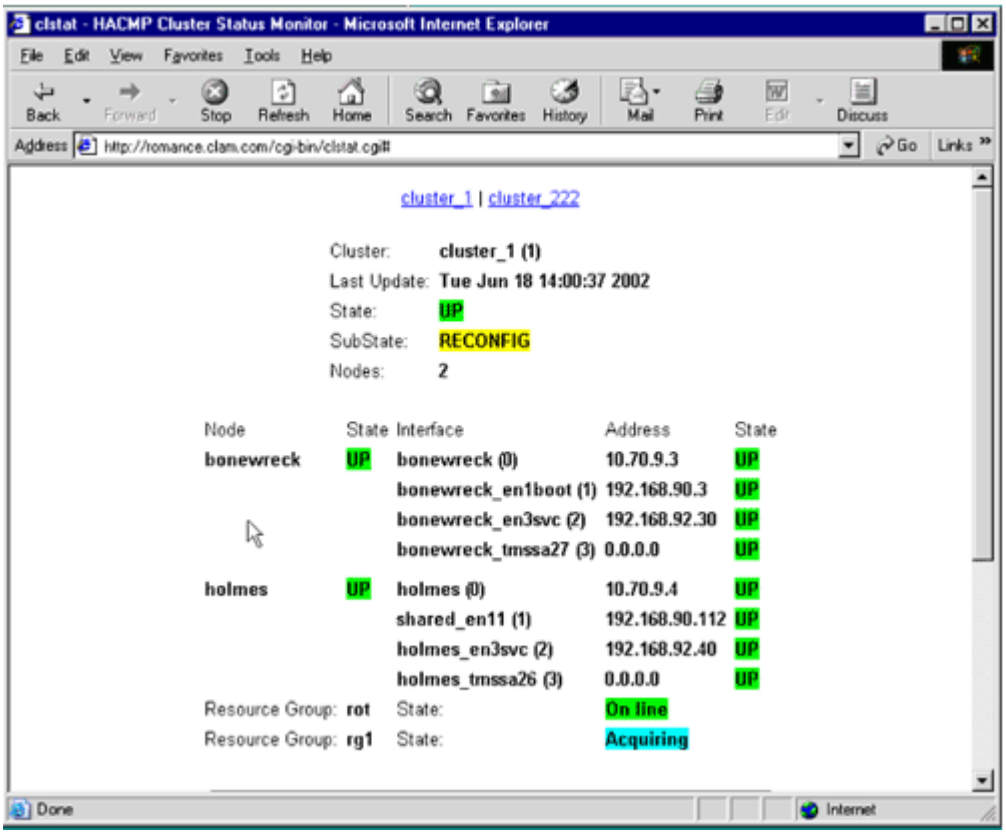


Figure 4: clstat browser display showing a resource group in acquiring state

> **Note:** When a cluster resource group goes offline, it can no longer be displayed by clstat . No information about that resource group appears until it is being reacquired or online.

Configuring Web server access to clstat.cgi

To view the **clstat** display through a web browser, you must have a web server installed on a machine where Clinfo is running and able to gather cluster information. This could be a client node as well as a server node. The clstat.cgi program works with any web server that supports the CGI standard, which includes most currently available web servers for AIX®. For instance, you might use the IBM® HTTP Server, which is included on the Expansion Pack CD for AIX®.

Full instructions for installing and configuring a web server are *not* included here. Please refer to the web server documentation or consult your web administrator if you need additional help.

The following steps complete the configuration of web server access to clstat.cgi using the IBM® HTTP Server with its default configuration. The directories and URL you use for your server and configuration may vary.

1. Move or copy clstat.cgi to the cgi-bin or script directory of the web server, for instance the default HTTP Server directory /usr/HTTPserver/cgi-bin.
2. Verify that the clstat.cgi file still has appropriate permissions (that is, the file is executable by the user *nobody*).
3. You can now view cluster status using a web browser by typing in a URL of the following format:

```
http://<host name or IP label of the web server node>
/cgi-bin/clstat.cgi
```

Result

> **Note:** Although you can change the name of the CGI directory, do *not* rename the clstat.cgi file.

Changing the clstat.cgi refresh interval

You can change the default clstat.cgi refresh interval by specifying the CLSTAT_CGI_REFRESH environment variable in the **/etc/environment** file on the node serving the web page.

Setting the CLSTAT_CGI_REFRESH environment variable (in seconds) overrides the default setting.

For example, to change the refresh interval to 15 seconds from the default setting, add the following to the **/etc/environment** file:

```
# change the clstat.cgi refresh interval to 15 seconds; 30 seconds is the default

CLSTAT_CGI_REFRESH=15
```

clstat and security

Because clstat.cgi is *not* run as root, there should be no immediate security threat of users gaining unauthorized access to PowerHA® SystemMirror® by accessing clstat.cgi from the web server.

Some administrators may wish to restrict access to clstat.cgi from the web server and can use methods built in to the web server to prevent access, such as password authentication or IP address blocking. PowerHA® SystemMirror® does *not* provide any specific means of access restriction to clstat.cgi .

Monitoring applications

PowerHA® SystemMirror® uses monitors to check if the application is running before starting the application, avoiding startup of an undesired second instance of the application.

PowerHA® SystemMirror® also monitors specified applications and attempts to restart them upon detecting process death or application failure.

Application monitoring works in one of two ways:

- *Process application monitoring* detects the termination of one or more processes of an application, using RSCT Resource Monitoring and Control (RMC).
- *Custom application monitoring* checks the health of an application with a custom monitor method at user-specified polling intervals.

PowerHA® SystemMirror® uses monitors to check if the application is running before starting the application. You can configure multiple application monitors and associate them with one or more application controllers. You can assign each monitor a unique name in SMIT.

By supporting multiple monitors per application, PowerHA® SystemMirror® can support more complex configurations. For example, you can configure one monitor for each instance of an Oracle parallel server in use. Or, you can configure a custom monitor to check the health of the database along with a process termination monitor to instantly detect termination of the database process.

Process monitoring is easier to set up, as it uses the built-in monitoring capability provided by RSCT and requires no custom scripts; however, it may not be an appropriate option for all applications. User-defined monitoring can monitor more subtle aspects of an application's performance and is more customizable, but it takes more planning, as you must create the custom scripts.

In either case, when a problem is detected by the monitor, PowerHA® SystemMirror® attempts to restart the application on the current node and continues the attempts until a specified restart count is exhausted. When an application cannot be restarted within this restart count, PowerHA® SystemMirror® takes one of two actions, which you specify when configuring the application monitor:

- Choosing **fallover** causes the resource group containing the application to fall over to the node with the next highest priority according to the resource policy.
- Choosing **notify** causes PowerHA® SystemMirror® to generate a server_down event to inform the cluster of the failure.

When you configure an application monitor, you use the SMIT interface to specify which application is to be monitored and then define various parameters such as time intervals, restart counts, and action to be taken in the event the application can*not* be restarted. You control the application restart process through the Notify

Method, Cleanup Method, and Restart Method SMIT fields, and by adding pre-event and post-event scripts to any of the failure action or restart events you select.

You can temporarily suspend and then resume an application monitor in order to perform cluster maintenance.

When an application monitor is defined, each node's Configuration Database contains the names of monitored applications and their configuration data. This data is propagated to all nodes during cluster synchronization, and is backed up when a cluster snapshot is created. The cluster verification ensures that any user-specified methods exist and are executable on all nodes.

> **Note:** If you specify the **fallover** option, which may cause a resource group to migrate from its original node, even when the highest priority node is up, the resource group may remain offline. Unless you bring the resource group online manually, it could remain in an inactive state.

A note on Application monitors

Application monitors configurable in PowerHA® SystemMirror® are a critical piece of the PowerHA® SystemMirror® cluster configuration; they enable PowerHA® SystemMirror® to keep applications highly available. When PowerHA® SystemMirror® starts an application controller on a node, it uses a monitor that you configure to check if an application is already running to avoid starting two instances of the application. PowerHA® SystemMirror® also periodically manages the application using the monitor that you configure to make sure that the application is up and running.

An erroneous application monitor may not detect a failed application. As a result, PowerHA® SystemMirror® would not recover it or may erroneously detect an application as failed, which may cause PowerHA® SystemMirror® to move the application to a takeover node, resulting in unnecessary downtime. For example, a custom monitor that uses an **sql** command to query a database to detect whether it is functional may not respond that the database process is running on the local node so this is not sufficient for use with PowerHA® SystemMirror®.

If you plan on starting the cluster services with an option of **Manage Resources > Manually**, or stopping the cluster services without stopping the applications, PowerHA® SystemMirror® relies on configured application monitors to determine whether to start the application on the node or not.

When cluster services are stopped using the unmanage option, the long-running application monitors are not brought down. As long as the **clstrmgr** daemon is active, it is aware that there is already a monitor running and a second instance will not be started when PowerHA® SystemMirror® restarts. If the monitor indicates a failure then events are not generated in response. Therefore, no cleanup or restart methods are running during this time. If your application monitor attempts a recovery or restart on its own, PowerHA® SystemMirror® will not be able to react. It is important to separate recovery actions from the monitor itself.

To summarize, we highly recommend properly configured and tested application monitors for all applications that you want to keep highly available with the use of PowerHA® SystemMirror®. During verification, PowerHA® SystemMirror® issues a warning if an application monitor is not configured.

Displaying an application-centric cluster view

You can use the ASCII version of SMIT to view a cluster application.

To show a cluster application in SMIT:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Show Cluster Applications** and press Enter.
   SMIT displays the list of applications.
3. Select the application to show from the list.
   SMIT displays the application with its related components.

Result
To show current resource group and application state, select **Cluster Applications and Resources › Resource Groups › Show All Resources by Node or Resource Group › Show the Current State of Applications and Resource Groups**. This panel displays the current states of applications and resource groups for each resource group.

- For non-concurrent groups, PowerHA® SystemMirror® shows only the node on which they are online and the applications state on this node
- For concurrent groups, PowerHA® SystemMirror® shows ALL nodes on which they are online and the applications states on the nodes
- For groups that are offline on all nodes, only the application states are displayed, node names are not listed.

Measuring application availability

You can use the Application Availability Analysis Tool to measure the amount of time that any of your applications (with defined application controller) is available.

The PowerHA® SystemMirror® software collects, time stamps, and logs the following information:

- An application starts, stops, or fails.
- A node fails or is shut down, or comes up.
- A resource group is taken offline or moved.
- Application monitoring is suspended or resumed.

Using SMIT, you can select a time period and the tool will display uptime and downtime statistics for a given application during that period. The tool displays:

- Percentage of uptime
- Amount of uptime
- Longest period of uptime
- Percentage of downtime
- Amount of downtime
- Longest period of downtime.

All nodes must be available when you run the tool to display the uptime and downtime statistics. Clocks on all nodes must be synchronized in order to get accurate readings.

The Application Availability Analysis tool treats an application that is part of a concurrent resource group as available as long as the application is running on any of the nodes in the cluster. Only when the application has gone offline on all nodes in the cluster will the Application Availability Analysis tool consider the application as unavailable.

The Application Availability Analysis tool reports application availability from the PowerHA® SystemMirror® cluster infrastructure's point of view. It can analyze only those applications that have been properly configured so they will be managed by the PowerHA® SystemMirror® software.

When using the Application Availability Analysis tool, keep in mind that the statistics shown in the report reflect the availability of the PowerHA® SystemMirror® application conroller, resource group, and (if configured) the application monitor that represent your application to PowerHA® SystemMirror®.

The Application Availability Analysis tool cannot detect availability from an end user's point of view. For example, assume that you have configured a client-server application so that PowerHA® SystemMirror® manages the server, and, after the server was brought online, a network outage severed the connection between the end user clients and the server. The end users would view this as an application outage because their client software could not connect to the server, but PowerHA® SystemMirror® would not detect it, because the server it was managing did not go offline. As a result, the Application Availability Analysis tool would not report a period of downtime in this scenario.

Planning and configuring for measuring application availability

If you have application controllers defined, the Application Availability Analysis Tool automatically keeps the statistics for those applications.

In addition to using the Application Availability Analysis Tool, you can also configure Application Monitoring to monitor each application controller’s status. You can define either a Process Application Monitor or a Custom Application Monitor.

If you configure Application Monitoring solely for the purpose of checking on uptime status and do not want the Application Monitoring feature to automatically restart or move applications, you should set the **Action on Application Failure** parameter to just **Notify** and set the **Restart Count** to zero. (The default is three.)

Ensure that there is adequate space for the **clavan.log** file on the file system on which it is being written. Disk storage usage is a function of node and application stability (not availability), that is, of the number (not duration) of node or application failures in a given time period. Roughly speaking, the application availability analysis tool will use 150 bytes of disk storage per outage. For example, on a node that fails once per week and has one application running on it, where that application never fails on its own, this feature uses about 150 bytes of disk storage usage per week.

Whenever **verification** runs, it determines whether there is enough space for the log on all nodes in the cluster.

Configuring and using the application availability analysis tool

You can use SMIT to check on a given application over a certain time period.

Follow these steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) › Resource Group and Applications › Application Availability Analysis** and press Enter.
3. Select an application. Press F4 to see the list of configured applications.
4. Fill in the fields as follows:

Application Availability Analysis fields	
Field	**Value**
Application Name	Application you selected to monitor.
• **Begin analysis on year (1970-2038)** • **month (01-12)** • **day (1-31)**	
• **Begin analysis at hour (00-23)** • **minutes (00-59)** • **seconds (00-59)**	
• **End analysis on year (1970-2038)** • **month (01-12)** • **day (1-31)**	
• **End analysis at hour (00-23)** • **minutes (00-59)** • **seconds (00-59)**	

5. Press Enter. The application availability report is displayed as shown in the following sample.

Result

```
     COMMAND STATUS
Command: OK    stdout: yes    stderr: no
Before command completion, additional instructions may appear as follows.

Application: myapp

Analysis begins: Monday, 1-May-2002, 14:30
Analysis ends:  Friday, 5-May-2002, 14:30

Total time:    5 days, 0 hours, 0 minutes, 0 seconds
```

```
Uptime:
Amount: 4 days, 23 hours, 0 minutes, 0 seconds
Percentage:    99.16 %
Longest period:     4 days, 23 hours, 0 minutes, 0 seconds

Downtime:
Amount: 0 days, 0 hours, 45 minutes, 0 seconds
Percentage:    00.62 %
Longest period:     0 days, 0 hours, 45 minutes, 0 seconds
```

If the utility encounters an error in gathering or analyzing the data, it displays one or more error messages in a **Command Status** panel.

Reading the clavan.log file

The application availability analysis log records are stored in the clavan.log file.

The default directory for this log file is `/var/hacmp/log`. You can change the directory by using the **System Management C-SPOC > PowerHA® SystemMirror® Logs > Change/Show a Cluster Log Directory** SMIT panel. Each node has its own instance of the file. You can look at the logs at any time to get the uptime information for your applications.

> **Note:** If you redirect the log, remember it is a cumulative file. Its usefulness for statistical information and analysis will be affected if you do *not* keep the information in one place.

clavan.log file format

The **clavan.log** file format is described here.

```
Purpose
Records the state transitions of applications managed by PowerHA SystemMirror.

Description
The clavan.log file keeps track of when each application that is managed<
by PowerHA SystemMirror is started or stopped and when the node stops on which
an application is running.  By collecting the records in the
clavan.log file from every node in the cluster, a utility program
can determine how long each application has been up, as well as
compute other statistics describing application availability time.

Each record in the clavan.log file consists of a single line.
Each line contains a fixed portion and a variable portion:

AAA: Ddd Mmm DD hh:mm:ss:YYYY: mnemonic:[data]:[data]: <variable portion>

Where:  is:
------    ----
AAA   a keyword
Ddd  the 3-letter abbreviation for the day of the week
YYYY    the 4-digit year
Mmm  The 3-letter abbreviation for month
DD   the 2-digit day of the month (01...31)
hh   the 2-digit hour of the day (00...23)
mm   the 2-digit minute within the hour (00...59)
ss   the 2-digit second within the minute (00...59)

variable portion: one of the following, as appropriate (note that umt
stands for Uptime Measurement Tool, the original name of this tool):
```

Mnemonic	Description	As used in clavan.log file
umtmonstart	monitor started	umtmonstart:monitor_name:node:
umtmonstop	monitor stopped	umtmonstop:monitor_name:node:
umtmonfail	monitor failed	umtmonfail:monitor_name:node:
umtmonsus	monitor suspended	umtmonsus:monitor_name:node:
umtmonres	monitor resumed	umtmonres:monitor_name:node:

Mnemonic	Description	As used in clavan.log file
umtappstart	application controller started	umtappstart:app_server:node:
umtappstop	application controller stopped	umtappstop:app_server:node:
umtrgonln	resource group online	umtrgonln:group:node:
umtrgoffln	resource group offline	umtrgoffln:group:node:
umtlastmod	file last modified	umtlastmod:date:node:
umtnodefail	node failed	umtnodefail:node:
umteventstart	cluster event started	umteventstart:event
[arguments]:		
umteventcomplete	cluster event completed	umteventcomplete:event
[arguments]:		

```
Implementation Specifics

None.

Files

/var/hacmp/log/clavan.log
This is the default file spec for this log file.
The directory can be changed with the "Change/Show a
PowerHA SystemMirror Log Directory" SMIT panel (fast path =
"clusterlog_redir_menu")

Related Information

None.
```

clvan.log file examples

This example shows output for various types of information captured by the tool.

```
AAA: Thu Feb 21 15:27:59 2002: umteventstart:reconfig_resource_release:
Cluster event reconfig_resource_release started
AAA: Thu Feb 21 15:28:02 2002:
umteventcomplete:reconfig_resource_release: Cluster event
reconfig_resource_release completed
AAA: Thu Feb 21 15:28:15 2002: umteventstart:reconfig_resource_acquire:
Cluster event reconfig_resource_acquire started
AAA: Thu Feb 21 15:30:17 2002:
umteventcomplete:reconfig_resource_acquire: Cluster event
reconfig_resource_acquire completed
AAA: Thu Feb 21 15:30:17 2002: umteventstart:reconfig_resource_complete:
Cluster event reconfig_resource_complete started
AAA: Thu Feb 21 15:30:19 2002: umtappstart:umtappa2:titan: Application
umtappa2 started on node titan
AAA: Thu Feb 21 15:30:19 2002: umtrgonln:rota2:titan: Resource group
rota2 online on node titan
```

> **Note: clavan.log** file records are designed to be human-readable but also easily parsed. This means you can write your own analysis programs. The Application Availability Analysis tool is written in Perl and can be used as a reference for writing your own analysis program. The pathname of the tool is **/usr/es/sbin/cluster/utilities/clavan**.

Using the cldisp command

The **/usr/es/sbin/cluster/utilities/cldisp** command provides the application-centric view of the cluster configuration. This utility can be used to display resource groups and their startup, fallover, and fallback policies.

To show cluster applications:

1. From the command line, enter `smit sysmirror`.
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Show Cluster Applications** and press Enter.
   SMIT displays the information as shown in the example:

```
################
APPLICATIONS
################

 Cluster Test_Cluster_Cities provides the following applications:
 Application_Server_1 Application_Server_NFS_10
      Application: Application_Server_1 State: {online}

Application 'Application_Server_NFS_10' belongs to a resource group
which is configured to run on all its nodes simultaneously. No
fallover will occur.
 This application is part of resource group 'Resource_Group_03'.
  The resource group policies:
     Startup: on all available nodes
    Fallover: bring offline on error node
     Fallback: never
  Nodes configured to provide Application_Server_1: Node_Kiev_1{up} Node_
Minsk_2{up} Node_Moscow_3{up}
  Nodes currently providing Application_Server_1: Node_Kiev_1{up} Node
_Minsk_2{up} Node_Moscow_3{up}
  Application_Server_1 is started by /usr/user1/hacmp/local/ghn_start_4
  Application_Server_1 is stopped by /usr/user1/hacmp/local/ghn_stop_4
  Resources associated with Application_Server_1:
  Concurrent Volume Groups:
Volume_Group_03
  No application monitors are configured for
Application_Server_1.

      Application: Application_Server_NFS_10 State: {online}
  This application is part of resource group
'Resource_Group_01'.
  The resource group policies:
     Startup: on home node only
     Fallover: to next priority node in the list
     Fallback: if higher priority node becomes available
  Nodes configured to provide Application_Server_NFS_10: Node_Kiev_1{up}...
```

Here is an example of the text output from the **cldisp** command:

```
app1{online}
This application belongs to the resource group rg1.
Nodes configured to provide app1: unberto{up} lakin{up}
The node currently providing app1 is: unberto {up}
The node that will provide app1 if unberto fails is: lakin
app1 is started by /home/user1/bin/app1_start
app1 is stopped by /home/user1/bin/app1_stop
Resources associated with app1:
srv1(10.10.11.1){online}
Interfaces are configured to provide srv1:
lcl_unberto (en1-10.10.10.1) on unberto{up}
lcl_lakin (en2-10.10.10.2) on lakin{up}
Shared Volume Groups: NONE
Concurrent Volume Groups: NONE
Filesystems: NONE
AIX Fast Connect Services: NONE
Application monitor of app1: app1
Monitor: app1
Type: custom
Monitor method: /home/user1/bin/app1_monitor
Monitor interval: 30 seconds
Hung monitor signal: 9
Stabilization interval: 30 seconds
Retry count: 3 tries
Restart interval: 198 seconds
Failure action: notify
Notify method: /home/user1/bin/app1_monitor_notify
Cleanup method: /home/user1/bin/app1_stop
Restart method: /home/user1/bin/app1_start
```

Using PowerHA® SystemMirror® topology information commands

You can see the complete topology configuration using the **/usr/es/sbin/cluster/utilities/cltopinfo** command.

See PowerHA® SystemMirror® for AIX® commands for the complete syntax and examples with various flags. The following example uses the basic command:

```
$ /usr/es/sbin/cluster/utilities/cltopinfo
Cluster Description of Cluster: FVT_mycluster
Cluster Security Level: Standard
There are 2 node(s) and 1 network(s) defined

NODE holmes:
 Network ether_ipat
    sherlock_en3svc_a1  192.168.97.50
      holmes_en1svc_a1192.168.95.40
      holmes_en1svc     192.168.90.40

NODE sherlock:
 Network ether_ipat
    sherlock_en3svc_a1  192.168.97.50
      holmes_en1svc_a1 192.168.95.40
    sherlock_en1svc     192.168.90.50

Resource Group econrg1
 Behaviorconcurrent
 Participating Nodes      holmes sherlock
```

Monitoring cluster services

After checking cluster, node, and network interface status, check the status of the PowerHA® SystemMirror® and RSCT daemons on both nodes and clients.

Monitoring cluster services on a node

Depending on what you need to know, you may access the following for information:

- View Management Information Base (MIB)
- Look for cluster events and errors in the hacmp.out file.
- Use SMIT to check the status of the following PowerHA® SystemMirror® subsystems on a node:
  - Cluster Manager (clstrmgrES) subsystem
  - SNMP (snmpd) daemon.
  - Clinfo (clinfoES) Cluster Information subsystem.
  - To view cluster services on a node, enter the fastpath `smit clshow`
    A panel similar to following appears.

```
      COMMAND STATUS

Command: OK  stdout: yes stderr: no

Before command completion, additional instructions may appear below.

Subsystem     Group  PID     Status
clstrmgrES    cluster18524   active
clinfoES      cluster15024   active
```

Monitoring cluster services on a client

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

```
Subsystem   Group   PID   Status
clinfoES    cluster 9843  active
```

You can also use the ps command and grep for "clinfo." For example:

```
ps -aux | grep clinfoES
```

PowerHA® SystemMirror® log files

PowerHA® SystemMirror® 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 PowerHA® SystemMirror®, you can get different views of cluster status by viewing different log files.

PowerHA® SystemMirror® writes messages into the log files described here.

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

> **Note:** If you redirect logs, they should be redirected to local file systems and not to shared or NFS file systems. Having logs on shared or NFS file systems may cause problems if the file system needs to unmount during a fallover event. Redirecting logs to shared or NFS file systems may also prevent cluster services from starting during node reintegration.

Size of /var file system may need to be increased

For each node in your cluster, **verification** requires from 500K to 4 MB of free space in the **/var** file system.

PowerHA® SystemMirror® stores, at most, four different copies of a node's verification data on a disk at a time:

- **/var/hacmp/clverify/current/<nodename>/*** contains logs from a current execution of cluster verification
- **/var/hacmp/clverify/pass/<nodename/*** contains logs from the last time verification passed
- **/var/hacmp/clverify/pass.prev/<nodename/*** contains logs from the second to last time verification passed
- /**var/hacmp/clverify/fail/<nodename>/*** contains information from the last time verification failed.

The **/var/hacmp/clverify/clverify.log[0-9]** log files typically consume 25 MB of disk space. The **/var/hacmp/log/hacmp.out[0-9]** log files consume a maximum of 14 MB of disk space. Other log files consume a maximum of 32 MB of disk space.

In addition, the standard security mechanism that runs the **clcomd** utility has the following requirements for the free space in the **/var** file system:

1. 60 MB, where:
   - **/var/hacmp/clcomd/clcomd.log** requires 8 MB
   - **/var/hacmp/clcomd/clcomddiag.log** requires 80 MB.
2. 1 MB x n, per node (where n is the number of nodes in the cluster) in the file **/var/hacmp/odmcache**.

To summarize, for a four-node cluster it is recommended to have at least 179 MB of free space in the **/var** file system, where:

- 25 MB for writing **clverify.log[0-9]** log files
- 16 MB (4 MB per node) for writing the verification data from the nodes
- 88 MB for writing **clcomd** log information
- 14 MB for writing **hacmp.out[0-9]** files
- 30 MB for writing other log files
- 4 MB (1 MB per node) for writing ODMcache data

Description of log files

This topic contains a listing of log files.

/var/hacmp/adm/cluster.log file

The **cluster.log** file is the main PowerHA® SystemMirror® log file. PowerHA® SystemMirror® error messages and messages about PowerHA® SystemMirror®-related events are appended to this log with the time and date at which they occurred.

/var/hacmp/adm/history/cluster.mmddyyyy file

The cluster.**mmddyyyy** file contains time-stamped, formatted messages generated by PowerHA® SystemMirror® scripts. The system creates a cluster history file whenever cluster events occur, identifying each file by the file name extension *mmddyyyy*, where *mm* indicates the month, *dd* indicates the day, and *yyyy* indicates the year.

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

/var/hacmp/clcomd/clcomd.log file

The **clcomd.log** file contains time-stamped, formatted messages generated by the PowerHA® SystemMirror® Cluster Communication Daemon. This log file contains an entry for every connect request made to another node and the return status of the request.

For information on space requirements for this file and for the file described as follows, see the section Size of /var file system may need to be increased.

/var/hacmp/clcomd/clcomddiag.log file

The **clcomddiag.log** file contains time-stamped, formatted messages generated by the PowerHA® SystemMirror® Communication daemon when tracing is turned on. This log file is typically used by IBM® support personnel for troubleshooting.

/var/hacmp/clverify/clverify.log file

The **clverify.log** file contains verbose messages, output during **verification**. Cluster verification consists of a series of checks performed against various PowerHA® SystemMirror® configurations. Each check attempts to detect either a cluster consistency issue or an error. The verification messages follow a common, standardized format, where feasible, indicating such information as the node(s), devices, and command in which the error occurred. See Verifying and synchronizing a PowerHA® SystemMirror® cluster for complete information.

For information about space requirements for this file, see the section Size of /var file system may need to be increased.

/var/hacmp/log/autoverify.log file

The **autoverify.log** file contains any warnings or errors that occur during Automatic Cluster Verification.

/var/hacmp/log/clavan.log file

The **clavan.log** file keeps track of when each application that is managed by PowerHA® SystemMirror® is started or stopped and when the node stops on which an application is running. By collecting the records in the **clavan.log** file from every node in the cluster, a utility program can determine how long each application has been up, as well as compute other statistics describing application availability time.

/var/hacmp/log/clinfo.log /var/hacmp/log/clinfo.log.n, n=1,..,7 file

Clinfo is typically installed on both client and server systems. Client systems do not have the infrastructure to support log file cycling or redirection.

The **clinfo.log** file records the activity of the **clinfo** daemon.

/var/hacmp/log/cl_testtool.log file

When you run the Cluster Test Tool from SMIT, it displays status messages to the screen and stores output from the tests in the **/var/hacmp/log/cl_testtool.log file**.

/var/hacmp/log/clconfigassist.log file

The **clconfigassist.log** file is the log file for the Cluster Configuration Assistant.

/var/hacmp/log/clstrmgr.debug /var/hacmp/log/clstrmgr.debug.n, n=1,..,7 file

The **clstrmgr.debug** log file contains time-stamped, formatted messages generated by Cluster Manager activity. This file is typically used only by IBM® support personnel.

/var/hacmp/log/clstrmgr.debug.long /var/hacmp/log/clstrmgr.debug.long.n, n=1,..,7 file

The **clstrmgr.debug.long** file contains high-level logging of cluster manager activity, in particular its interaction with other components of PowerHA® SystemMirror® and with RSCT, which event is currently being run, and information about resource groups (for example, their state and actions to be performed, such as acquiring or releasing them during an event).

/var/hacmp/log/clutils.log file

The **clutils.log** file contains the results of the automatic **verification** that runs on one user-selectable PowerHA® SystemMirror® cluster node once every 24 hours. When cluster verification completes on the selected cluster node, this node notifies the other cluster nodes with the following information:

- The name of the node where **verification** had been run.
- The date and time of the last **verification**.
- Results of the **verification**.

The **clutils.log** file also contains messages about any errors found and actions taken by PowerHA® SystemMirror® for the following utilities:

- The PowerHA® SystemMirror® File Collections utility
- The Two-Node Cluster Configuration Assistant
- The Cluster Test Tool
- The script used to manage the Live Partition Mobility (LPM) operations

/var/hacmp/log/cspoc.log file

The **cspoc.log** file contains logging of the execution of C-SPOC commands on the local node with ksh option xtrace enabled (set -x).

/var/hacmp/log/cspoc.log.long file

The **cspoc.log.long** file contains a high-level of logging for the C-SPOC utility - commands and utilities that have been invoked by C-SPOC on specified nodes and their return status.

/var/hacmp/log/cspoc.log.remote file

The **cspoc.log.remote** file contains logging of the execution of C-SPOC commands on remote nodes with ksh option xtrace enabled (set -x).

/var/hacmp/log/hacmp.out /var/hacmp/log/hacmp.out.n n=1,..,7 file

The **hacmp.out** file records the output generated by the event scripts as they execute. This information supplements and expands upon the information in the /var/hacmp/adm/cluster.log file. To receive verbose output, the **debug level** runtime parameter should be set to *high* (the default).

Reported resource group acquisition failures (failures indicated by a non-zero exit code returned by a command) are tracked in **hacmp.out**, and a summary is written near the end of the **hacmp.out** listing for a main event.

Checking this log is important, since the **config_too_long** console message is *not* evident in every case where a problem exists. Event summaries make it easier for you to check the **hacmp.out** file for errors.

/var/hacmp/log/lvupdate_orig.log /var/hacmp/log/lvupdate_orig.log.n, n=1,...,7 file

The **lvupdate_orig.log** file contains time-stamped, formatted messages generated by the PowerHA® SystemMirror® script that is used to manage the AIX® Live Update operations on the original node. This log file provides information about the actions and the execution status for the operations performed on the original cluster node before moving the workload to the surrogate cluster node.

/var/hacmp/log/lvupdate_surr.log /var/hacmp/log/lvupdate_surr.log.n, n=1,...,7 file

The **lvupdate_surr.log** file contains time-stamped, formatted messages generated by the PowerHA® SystemMirror® script that is used to manage the AIX® Live Update operations on the surrogate cluster node. This log file provides information about the actions and execution status for the operations performed on the surrogate cluster.

/var/hacmp/log/migration.log file

The **migration.log** file contains a high level of logging of cluster activity while the cluster manager on the local node operates in a migration state. All actions pertaining to the cluster manager follow the internal migration protocol.

/var/hacmp/log/oraclesa.log file

The **oraclesa.log** file contains information about any Oracle specific errors that occur when using this Smart Assist and is used by the Oracle Smart Assist.

/var/hacmp/log/sa.log file

The **sa.log** file contains information about any general errors that occur when using the Smart Assists and is used by the Smart Assist infrastructure.

Redirecting a cluster log file

You can use the SMIT interface to redirect a cluster log from its default directory to another destination.

To redirect a cluster log file, complete the following steps:

1. Enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > PowerHA® SystemMirror® Logs > Change/Show a Cluster Log Directory** . SMIT displays a picklist of cluster log files with a short description of each:

Change/Show a Cluster Log Directory fields	
Log file	**Description**
autoverify.log	Generated by Auto Verify and Synchronize
clavan.log	Generated by the Application Availability Analysis tool
clconfigassist.log	Generated by the Two-Node Cluster Configuration Assistant

Log file	Description
clinfo.log	Generated by the clinfo daemon
clstrmgr.debug	Generated by the clstrmgr daemon, verbose logging
clstrmgr.debug.long	Generated by the clstrmgr daemon
cl_testtool.log	Generated by the Cluster Test Tool
cluster.log	Generated by cluster scripts and daemons
cluster.mmddyyyy	Cluster history files generated daily
clutils.log	Generated by the cluster utilities and file propagation.
clverify.log	Generated by the cluster verification utility.
cspoc.log	Generated by C-SPOC commands
cspoc.log.long	Generated by the C-SPOC utility, verbose logging
cspoc.log.remote	Generated by the C-SPOC utility
hacmp.out	Generated by event scripts and utilities
migration.log	Generated by the clstrmgr daemon during cluster upgrade
oraclesa.log	Oracle Smart Assist log
sax.log	Smart Assist infrastructure log

3. Select a log that you want to redirect.
SMIT displays a panel with the selected log's name, description, default pathname, and current directory pathname. The current directory pathname will be the same as the default pathname if you do *not* change it. This panel also asks you to specify whether to allow this log on a remote file system (mounted locally using AFS®, DFS, or NFS).The default value is **false**.

> Use of a non-local file system for PowerHA® SystemMirror® logs will prevent log information from being collected if the file system becomes unavailable. To ensure that cluster services are started during node reintegration, log files should be redirected to local file systems, and *not* to NFS file systems.

The example here shows the **cluster.mmddyyyy** log file panel. Edit the fourth field to change the default pathname.

Field	Description
Cluster Log Name	**cluster.mmddyyyy**
Cluster Log Description	Cluster history files generated daily
Default Log Destination Directory	**/usr/es/sbin/cluster/history**
Log Destination Directory	The default directory name appears here. To change the default, enter the desired directory pathname.
Allow Logs on Remote Filesystems	**false**

4. Press Enter to add the values to the PowerHA® SystemMirror® for AIX® Configuration Database.
5. Return to the panel to select another log to redirect, or return to the Cluster System Management panel to proceed to the panel for synchronizing cluster resources.
6. After you change a log directory, a prompt appears reminding you to synchronize cluster resources from this node (Cluster log Configuration Databases must be identical across the cluster). The cluster log destination directories as stored on this node will be synchronized to all nodes in the cluster.

Result
Log destination directory changes will take effect when you synchronize cluster resources.

Note: Existing log files will *not* be moved to the new location.

Modifying the log file size by using SMIT

The Log Analytics function analyzes existing PowerHA® SystemMirror® log files and provides detailed information about different types of errors such as disk failures or interface failures. You can modify the file size of log files by using the SMIT interface.

To modify the file size of the log file, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Problem Determination Tools › PowerHA SystemMirror log Viewing and Management › Change/Show PowerHA Log File Size**, and press Enter.
3. Specify the maximum size for the log file in mega bytes. If the size of the log file is set to a small number, it might discard log file details that had critical information for problem determination. During verification process, the PowerHA® SystemMirror® shows different log file sizes to choose based on specific configuration of the user.

.

- The modified value of the log file size can be changed in the range 1 - 32767.
- The log file size must not be more than available free space in the file system.

Managing shared LVM components

These topics explain how to maintain AIX® Logical Volume Manager (LVM) components shared by nodes in a PowerHA® SystemMirror® cluster and provides procedures for managing volume groups, file systems, logical volumes, and physical volumes using the PowerHA® SystemMirror® Cluster-Single Point of Control (C-SPOC) utility.

The C-SPOC utility simplifies maintenance of shared LVM components in clusters of up to 16 nodes. C-SPOC commands provide comparable functions in a cluster environment to the standard AIX® commands that work on a single node. By automating repetitive tasks, C-SPOC eliminates a potential source of errors, and speeds up the cluster maintenance process.

In SMIT, you access C-SPOC using the **System Management (C-SPOC)** menu. C-SPOC operations can be accessed by entering the fastpath, `smit cspoc`.

Although you can also use AIX® on each node to do these procedures, using the C-SPOC utility ensures that all commands are executed in the proper order.

Shared LVM overview

A key element of any PowerHA® SystemMirror® cluster is the data used by the highly available applications. This data is stored on AIX® LVM entities. PowerHA® SystemMirror® clusters use the capabilities of the LVM to make this data accessible to multiple nodes.

In a PowerHA® SystemMirror® cluster, the following definitions are used:

- 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 file system* is a file system that resides entirely in a shared logical volume.

As a system administrator of a PowerHA® SystemMirror® 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
- Creating a new shared logical volume
- Extending, reducing, changing, or removing an existing logical volume
- Creating a new shared file system
- Extending, changing, or removing an existing file system
- Adding, removing physical volumes.

When performing any of these maintenance tasks on shared LVM components, make sure that ownership and permissions are reset (on logical volumes) 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 to what is needed after this sequence.

Understanding C-SPOC

The C-SPOC commands operate on both shared and concurrent LVM components that are defined as part of a PowerHA® SystemMirror® resource group. When you use C-SPOC, it executes the command on the node that has the LVM component varied on. If no node has the LVM component varied on, then the component is temporarily varied on for the operation.

A resource, such as a volume group, a physical disk, a file system, or an IP address, can be reliably controlled by only one resource manager at a time. For resources defined as part of PowerHA® SystemMirror® resource groups, PowerHA® SystemMirror® should be the only resource manager controlling the resource. You should refrain from using AIX® commands to modify such resources, and only use PowerHA® SystemMirror® operations on the resources. Only use C-SPOC operations on shared volume groups, physical disks, and file systems when the cluster is active. Using AIX® commands on shared volume groups while the cluster is active can result in the volume group becoming inaccessible, and has the potential for introducing data corruption.

Understanding C-SPOC and its relation to resource groups

The C-SPOC commands that modify LVM components use a resource group or a list of node names as an argument. If a resource group is given, it is used to determine the list of nodes. Use the C-SPOC SMIT panels to select LVM objects in picklists. You do not need to enter resource group names or node lists.

Removing a file system or logical volume

When removing a file system or logical volume using C-SPOC, the target file system or logical volume must not be configured as a resource in the resource group specified. You must remove the configuration for it from the resource group before removing the file system or logical volume.

Migrating a resource group

You can use the Resource Group Management utility, under the **System Management Tools (C-SPOC) › Resource Groups and Applications** menu in SMIT, to perform resource group maintenance tasks. This utility enhances failure recovery capabilities of PowerHA® SystemMirror® and allows you to change the status or the location of any type of resource group (along with its resources - IP addresses, applications, and disks), without stopping cluster services. For instance, you can use this utility to free a given node of any resource groups in order to perform system maintenance on that cluster node.

You can complete the following resource group management tasks using the Resource Group and Applications utility:

- Dynamically move a specified non-concurrent resource group from a node it currently resides on to the destination node that you have specified.
- Take a non-concurrent resource group online or offline on one or all nodes in the cluster.

Updating LVM components in a PowerHA® SystemMirror® 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. AIX® LVM enhancements allow all nodes in the cluster to be aware of changes to a volume group, logical volume, and file system, at the time the changes are made, rather than waiting for the information to be retrieved during a *lazy update*.

> **Note:** See Lazy update processing in a PowerHA® SystemMirror® cluster for a full explanation of this process.

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 execution of the **`clvaryonvg`** command.

If node failure does occur during a C-SPOC operation, an error is displayed to the panel and the error messages are recorded in the C-SPOC log. (**/var/hacmp/log/cspoc.log** is the default location of this log.) Other C-SPOC failures are also logged to the **cspoc.log** but are *not* displayed. You should check this log when any C-SPOC problems occur.

Error reporting provides detailed information about inconsistency in volume group state across the cluster. If this happens, you must take manual corrective action. For example, if the file system changes are not updated on all nodes, update the nodes manually with this information.

Lazy update processing in a PowerHA® SystemMirror® cluster

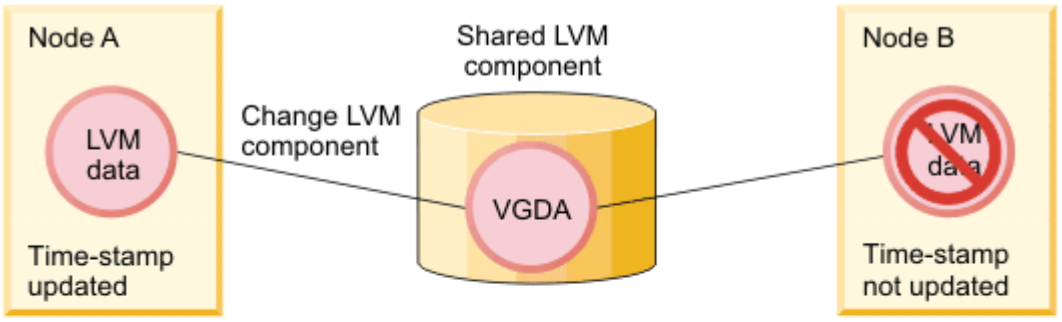
For LVM components under the control of PowerHA® SystemMirror®, you do *not* have to explicitly do anything to bring the other cluster nodes up to date. Instead, PowerHA® SystemMirror® can update LVM information on a node when it activates the volume group during a fallover.

In a cluster, PowerHA® SystemMirror® controls when volume groups are activated. PowerHA® SystemMirror® implements a function called *lazy update*. This function examines the volume group time stamp, which is maintained in both the volume group's VGDA, and the local ODM. AIX® updates both these time stamps whenever a change is made to the volume group. When PowerHA® SystemMirror® is going to vary on a volume group, it compares the copy of the time stamp in the local ODM with that in the VGDA. If the values differ, PowerHA® SystemMirror® will cause the local ODM information on the volume group to be refreshed from the information in the VGDA.

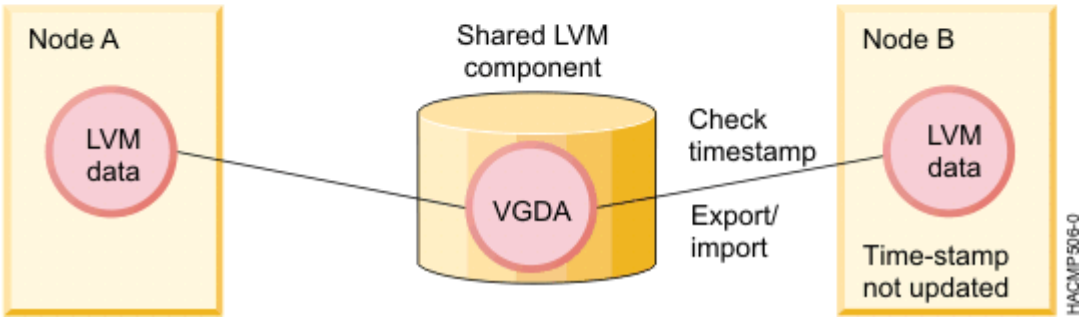
What this means for a PowerHA® SystemMirror® cluster administrator is that if a volume group under PowerHA® SystemMirror® control is updated directly (that is, without going through the C-SPOC facility of PowerHA® SystemMirror®), other nodes' information on that volume group will be updated when PowerHA® SystemMirror® has to bring the volume group online on those nodes, but not before.

The following figure illustrates how a lazy update works in a cluster. While it shows the AIX® export and import functions being used, the actual operations performed will depend on the state of the volume group. As stated previuosly, this happens when PowerHA® SystemMirror® has to bring the volume group online on node B. If the C-SPOC facility of PowerHA® SystemMirror® is used to make changes to the volume group, an equivalent operation is performed automatically at the time of the change.
Figure 5: Lazy update processing example

1. LVM component modified on Node A:



2. LVM data on Node B after lazy update:



Note: PowerHA® SystemMirror® does not require lazy update processing for enhanced concurrent volume groups, as it keeps all cluster nodes updated with the LVM information.

Forcing an update before fallover

To verify that the LVM definition of a volume group is the same on all cluster nodes before a fallover occurs, select the following from the SMIT menu **System Management (C-SPOC) › Storage › Volume Groups › Synchronize a Volume Group Definition**.

Then specify or select the volume group name. When you press Enter, PowerHA® SystemMirror® will cause all the nodes in the cluster to update their local ODM information based on the contents of the volume group.

Maintaining shared volume groups

While maintaining the PowerHA® SystemMirror® cluster, you may need to perform administrative tasks with shared volume groups.

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

A resource, such as a volume group, a physical disk, a file system, or an IP address, can be reliably controlled by only one resource manager at a time. For resources defined as part of PowerHA® SystemMirror® resource groups, PowerHA® SystemMirror® should be the only resource manager controlling the resource. You should refrain from using AIX® commands to modify such resources, and only use PowerHA® SystemMirror® operations on the resources. Only use C-SPOC operations on shared volume groups, physical disks, and file systems when the cluster is active. Using AIX® commands on shared volume groups while the cluster is active can result in the volume group becoming inaccessible, and has the potential for introducing data corruption.

Understanding active and passive varyon in enhanced concurrent mode

An enhanced concurrent volume group can be made active on the node, or varied on, in two states: active or passive.

Note that active or passive state varyons are done automatically by PowerHA® SystemMirror® upon detection of the enhanced concurrent mode volume group, based on the state of the volume group and current cluster configuration.

> **Important:** All nodes in the cluster must be available before making any LVM changes. This ensures that all nodes have an accurate view of the state of the volume group. For more information about safely performing a forced varyon operation, and on instructions how to configure it in SMIT, see Forcing a varyon of volume groups.

Active state varyon

Active state varyon behaves as ordinary varyon, and makes the logical volumes normally available.

When an enhanced concurrent volume group is varied on in active state on a node, it allows the following operations:

- Operations on file systems, such as file system mounts
- Operations on applications
- Operations on logical volumes, such as creating logical volumes
- Synchronizing volume groups.

Passive state varyon

When an enhanced concurrent volume group is varied on in passive state, the LVM provides an equivalent of fencing for the volume group at the LVM level.

Passive state varyon allows only a limited number of read-only operations on the volume group:

- LVM read-only access to the volume group's special file
- LVM read-only access to the first 4k of all logical volumes that are owned by the volume group.

The following operations are *not* allowed when a volume group is varied on in passive state:

- Operations on file systems, such as file systems mounting
- Any operations on logical volumes, such as having logical volumes open
- Synchronizing volume groups.

Using active or passive state varyon in PowerHA® SystemMirror®

PowerHA® SystemMirror® detects when a volume group included in a shared resource group is converted to or defined as an enhanced concurrent mode volume group, and notifies the LVM which node currently owns the volume group.

Based on this information, the LVM activates the volume group in the appropriate active or passive state depending on the node on which this operation takes place:

- Upon cluster startup, if the volume group resides currently on the node that owns the resource group, PowerHA® SystemMirror® activates the volume group on this node in active state. PowerHA® SystemMirror® activates the volume group in passive state on all other nodes in the cluster. Note that PowerHA® SystemMirror® will activate a volume group in active state only on one node at a time.
- Upon fallover, if a node releases a resource group, or, if the resource group is being moved to another node for any other reason, PowerHA® SystemMirror® switches the varyon state for the volume group from active to passive on the node that releases the resource group (if cluster services are still running), and activates the volume group in active state on the node that acquires the resource group. The volume group remains in passive state on all other nodes in the cluster.
- Upon node reintegration, this procedure is repeated. PowerHA® SystemMirror® changes the varyon state of the volume group from active to passive on the node that releases the resource group and varies on the

volume group in active state on the joining node. While activating, the volume group remains passive on all other nodes in the cluster.

> **Note:** The switch between active and passive states is necessary to prevent mounting file systems on more than one node at a time.

Verification checks for shared volume groups defined for auto varyon

Shared volume groups listed within a resource group must have the **auto-varyon** attribute in the AIX® ODM set to **No**. C-SPOC does not allow you to define a volume group with **auto-varyon** set to yes. However, after the volume group is defined you can change the **auto-varyon** attribute from yes to no.

The PowerHA® SystemMirror® **verification** checks that the volume group **auto-varyon** flag is set to **No**. If you use the interactive mode for **verification** , you will be prompted to set the **auto-varyon** flag to **No** on all the cluster nodes listed in the resource group.

Checking the status of a volume group

As with regular cluster takeover operations, you can debug and trace the cluster activity for fast disk takeover using the information logged in the **hacmp.out** file. You may check the status of the volume group by issuing the **lsvg** command.

Depending on your configuration, the **lsvg** command returns the following settings:

- `VG STATE` will be `active` if it is varied on either actively or passively.
- `VG PERMISSION` will be `read/write` if it is actively varied on the node, or `passive-only`, if it is passively varied on.
- `CONCURRENT` will either be `Capable` or `Enhanced-Capable (for concurrent volume groups)`.

Here is an example of **lsvg** output:

```
# lsvg vg1

VOLUME GROUP:   vg1  VG IDENTIFIER:  00020adf00004c00000000f329382713
VG STATE:active  PP SIZE: 16 megabyte(s)
VG PERMISSION:  passive-only   TOTAL PPs:      542 (8672 megabytes)
MAX LVs: 256  FREE PPs:521 (8336 megabytes)
LVs:  3    USED PPs:21 (336 megabytes)
OPEN LVs:0    QUORUM:  2
TOTAL PVs:      1    VG DESCRIPTORS: 2
STALE PVs:      0    STALE PPs:      0
ACTIVE PVs:     1    AUTO ON: no
Concurrent:     Enhanced-Capable  Auto-Concurrent: Disabled
VG Mode: Concurrent
Node ID: 2    Active Nodes:   1 4
MAX PPs per PV: 1016 MAX PVs: 32
LTG size:128 kilobyte(s)AUTO SYNC:      no
HOT SPARE:      no   BB POLICY:      relocatable
```

Avoiding a partitioned cluster

When configuring enhanced concurrent volume groups in shared resource groups, ensure that multiple networks exist for communication between the nodes in the cluster, to avoid cluster partitioning.

When fast disk takeover is used, the normal SCSI reserve is *not* set to prevent multiple nodes from accessing the volume group.

In a partitioned cluster, it is possible that nodes in each partition could accidentally vary on the volume group in active state. Because active state varyon of the volume group allows file system mounts and changes to the physical volumes, this state can potentially lead to different copies of the same volume group. Make sure that you configure multiple communication paths between the nodes in the cluster.

Disk heart beat is strongly recommended for enhanced concurrent mode volume groups.

Restoring fast disk takeover

If PowerHA® SystemMirror® has placed a volume group in the passive-only state and the volume group has then been varied off manually, fast disk takeover has effectively been disabled for that volume group. A fallover of the owning resource group to this node will be forced to use normal disk takeover, which will require additional time.

A volume group can be varied off incorrectly as a result of manual cleanup of resources or any other time when an administrator is manually varying off volume groups.

Restore the fast disk takeover capability for a volume group on a given node by using one of the following methods:

- Stopping and restarting cluster services on the node
- Moving the owning resource group to the node. The owning resource group can then be moved back to its original location, if desired.
- Manually vary on the volume group in passive mode with the command:

```
varyonvg -n -C -P <volume group name>
```

These methods will restore the volume group to a state where fast disk takeover is again available for the given node.

Collecting information on current volume group configuration

PowerHA® SystemMirror® can collect information about all shared volume groups that are currently available on the nodes in the cluster, *and* volume groups that can be imported to the other nodes in the resource group. PowerHA® SystemMirror® filters out volume groups that are already included in any of the resource groups.

You can use this information to import discovered volume groups onto other nodes in the resource group that do *not* have these volume groups.

To collect the information about volume group configuration:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Discover Network Interfaces and Disks** and press Enter.

Result
The information on current volume group configuration is collected and displayed.

Importing a shared volume group with C-SPOC

This topic describes how to import a shared volume group using the C-SPOC utility.

To import a 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 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 the fastpath `smit cl_admin`
4. In SMIT, select **Storage › Volume Groups › Import a Volume Group** and press Enter.
   A list of volume groups appears. (Enhanced concurrent volume groups are also included as choices in picklists for non-concurrent resource groups.)
5. Select a volume group and press Enter.
   A list of physical volumes appears.
6. Select a physical volume and Press Enter.
   SMIT displays the **Import a Shared Volume Group** panel. Values for fields you have selected are displayed.
7. Enter values for other fields as follows:

Import a Shared Volume Group fields	
Field	**Value**
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	The node from which the physical disk was retrieved.
Volume Group MAJOR NUMBER	It is recommended you use the default setting, which is the next available number in the valid range. If you wish to choose your own number, and are using NFS, you must be sure to have 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?	For a non-concurrent volume group, set this field to **no**. The default is **no** .
Make default varyon of VG Concurrent?	For a non-concurrent volume group, set this field to **no**. The default is **no** .

8. If this panel reflects the correct information, press Enter to import the shared volume group. All nodes in the cluster receive this updated information.

Result
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.

Creating a shared volume group by using C-SPOC

You can create a shared volume group using the C-SPOC utility. IBM Cloud®

Consider the following prerequisites before creating a shared volume group for the cluster using the C-SPOC utility:

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

> **Note:** If you add a VPATH disk to a volume group that consist of up of physical volumes (hdisks), the volume group will be converted to VPATHs disks on all nodes.

To create a shared volume group for a selected list of cluster nodes:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Storage › Volume Groups › Create a Volume Group** and press Enter.
   SMIT displays a list of cluster nodes.
3. Select two or more nodes from the list and press Enter.
   SMIT displays a list of volume group types. For more information about volume group types, see the **mkvg** command documentation.
4. Select a volume group type from the list and press Enter.
   SMIT displays the **Add a Volume Group** panel.

5. Complete the selections as follows and press Enter.

Add a Volume Group fields	
Field	**Value**
Node Names	Displays the name of nodes you selected.
PVID	Enter the PVID of the selected disk.
Resource Group	Enter the name of an existing resource group to which the volume group must be added, or the name of a new resource group. This resource group is created in which the volume group will be added.
VOLUME GROUP name	The name of the volume group must be unique within the cluster and distinct from the service IP label or address and resource group names; it should relate to the application it serves, and to any corresponding device. For example, *websphere_service_VG*. If you do not provide a name, a unique name will be generated.
Physical partition SIZE in megabytes	Accept the default value.
Volume group MAJOR NUMBER	The system displays the major number that the C-SPOC utility has determined to be correct. If you change the major number of the volume group the command might not run on a node that the major number is not available. Check commonly available major number on all nodes before changing this setting.
Enable Fast Disk Takeover or Concurrent Access	If Enable Fast Disk Takeover is selected, an enhanced concurrent mode volume group is created. If a resource group was created, it is created with the policies *online on the highest priority node* and *never fall back*. If Concurrent Access is selected, an enhanced concurrent mode volume group is created. If a resource group was created, it is created with the policies *online on all available nodes* and *never fall back*.
Volume Group Type	Displays the type of volume group. You cannot change this field.
CRITICAL volume group	Select **yes** to identify this volume group as a critical volume group. When a volume work is identified as a critical volume group, you can configure the volume group by using the **Manage Critical Volume Groups** option to set how PowerHA® SystemMirror® responds when access to the volume group is lost.
Mirror pool name	Mirror pools provide you a method to group disks within a volume group. This option is available only for scalable volume groups
Storage Location	Select an option from the list. • Site name - Select a corresponding site for the mirror pool.

Field	Value
	• Flash storage - Choose if the associated physical volumes belong to Flash storage.
Enable LVM encryption	Select **yes** to enable Logical Volume Manager (LVM) encryption. LVM Encryption is not supported for legacy volume groups. The default value for this field is **yes**.
Enable PV encryption	Select **yes** to enable Physical Volume (PV) encryption. The default value for this field is **no**. To avoid any verification error, you must add a authentication method and the authentication method name when you enable PV encryption.
Auth method	Select this option to add a authentication method for the PV encryption. The supported value for this field is **pks**.
Auth method name	This option is an alias name that is used to differentiate between multiple authentication methods that are configured for a physical volume. The authentication method name can range from 1 to 15 characters and can contain only the following characters: • 0-9 • A-z • a-z • _ (underscore character) • - (minus character) • . (period character) All other characters are considered invalid. The authentication name must not start with a - (minus character).

Depending on the type of volume group being created, additional information might be required in the configuration panel.

The C-SPOC utility verifies communication paths and version compatibility, and then executes the command on all nodes that you selected. If cross-site LVM mirroring is enabled, the associated configuration is verified.

If the major number that you entered on the SMIT panel was not free when the system attempted to create the volume group, PowerHA® SystemMirror® displays an error for the node that did not complete the command, and continues to the other nodes. After running the command, the volume group will not be active on any node in the cluster.

6. The discovery process runs automatically so that the new volume group is included in picklists for further actions.

Setting characteristics of a shared volume group

You can change the volume group's characteristics by adding or removing a volume group from the shared volume group.

Adding or removing a volume group from a shared volume group

You can add or remove a volume group to or from a shared volume group.

To add or remove a volume group to or from a shared volume group:

1. Enter the fastpath `smit cspoc`
2. To add a volume group in SMIT, select **Storage › Volume Groups › Set Characteristics of a Volume Group › Add a Volume to a Volume Group** and press Enter.
   To remove a volume group in SMIT, select **Storage › Volume Groups › Set Characteristics of a Volume Group › Remove a Volume from a Volume Group** and press Enter.
   SMIT displays a list of shared volume groups and its owning resource group, if any, and the node list.
3. Select the volume group and press Enter.
4. Select the volume group to add or remove from the list and press Enter.
5. If the volume group that you added is part of the mirror pool, specify the mirror pool name and storage location.
   The values of the **Storage Location** field includes Flash storage, site1, or site2.
   - Site name - Select a corresponding site for the mirror pool.
   - Flash storage - Choose if the associated physical volumes belong to Flash storage.

Changing or displaying the characteristics of a volume group

You can change or display the characteristics of a volume group.

To change or display a shared volume group:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Storage › Volume Groups › Set Characteristics of a Volume Group › Change/Show characteristics of a Volume Group** and press Enter.
3. Select Volume group. SMIT displays the **Change/Show characteristics of a Volume Group** panel.
4. Complete the selections as follows and press Enter

Change or show a volume group fields	
Field	**Value**
Volume group name	The volume group selected from the prior pick list of volume groups known across the cluster
Resource Group Name	The resource group (if any) that owns the selected volume group.
Node Names	The list of nodes on which the volume group is known.
Activate volume group automatically at system restart?	Specifies whether the volume group is automatically activated during the system startup. If a volume group is infrequently used, do not active the volume group at system startup because it uses kernel resources (memory).
A QUORUM of disks required to keep the volume group on-line ?	Determines whether the volume group is automatically varied off after losing its quorum of physical volumes. The default value is **yes**, which

Field	Value
	means that the volume group is automatically varied off after losing its quorum of physical volumes (51% of descriptor areas). The value **no** indicates that the volume group stays active until it looses all of its physical volumes.
Enable Fast Disk Takeover or Concurrent Access	Change the volume group to Enhanced Concurrent Mode. Enhanced concurrent mode can be used for fast disk takeover, or it can be used in a resource group that is online on all available nodes (OAAN).
Change to big VG format?	Changes the volume group to Big VG format. This can accommodate up to 128 physical volumes and 512 logical volumes. For more information about the Big VG format, see the **mkvg** command.
Change to scalable VG format?	Changes the volume group to Scalable VG format. This format can accommodate up to 1024 physical volumes and 4096 logical volumes. For more information about the Scalable VG format, see the **mkvg** command.
LTG Size in kbytes	Sets the group size of the logical track, in number of kilobytes, for the volume group. The specified value must be 128, 256, 512, or 1024. The number that you specify must be less than or equal to the maximum transfer size of all disks in the volume group. The default size is 128 kilobytes.
Set hotspare characteristics	Sets the sparing characteristics for the volume group specified by the Volume Group parameter. Either allows or prohibits the automatic migration of failed disks. For more information about hotspare characteristics, see the **mkvg** command.
Max PPs per VG in units of 1024	Increases the number of physical partitions a volume group can accommodate. The Physical Partitions variable is represented in units of 1024 partitions. Valid values are 64, 128, 256, 512, 768, 1024, and 2048. The value must be greater than the current value or no action is taken. This option is only valid with Scalable-type volume groups.
Max Logical Volumes	Increases the number of logical volumes that can be created. Valid values are 512, 1024, 2048, and 4096. The value must be greater than the current value, otherwise no action is taken. This option is valid only with scalable-type volume groups.
Mirror Pool Strictness	Enables mirror pool strictness for the volume group. **Note:** • If you specify **yes**, mirror pools must be used on each logical volume in the volume group. • If you specify **superstrict**, super-strict mirror pools are enforced on the selected volume group.

Field	Value
LVM Preferred Read	LVM preferred read options follows: • Roundrobin - Default policy for LVM Preferrred read copy. LVM determines which mirror copy to read. • Favour copy - Select this option if you want to read from the Flash storage irrespective of the resource group location. • Site Affinity - Select this option if you want to read from the local storage path that is based on the resource group location.
Enable LVM Encryption	Use this option to enable or disable the Logical Volume Manager (LVM) encryption. LVM encryption is not supported on legacy volume groups.

5. The C-SPOC utility verifies communication paths and version compatibility, and then executes the command on all nodes that you selected. If cross-site LVM mirroring is enabled, the associated configuration is verified.

Mirroring a volume group using C-SPOC

This topic describes mirroring a shared volume group using the C-SPOC utility.

To mirror a shared 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 the any cluster node, enter `smit cspoc`
3. In SMIT, select **Storage › Volume Groups › Mirror a Volume Group**.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
4. Select a volume group and press Enter.
5. Select entries from the list of nodes and physical volumes (hdisks) and Press Enter.
   SMIT displays the **Mirror a Volume Group** panel, with the selected entries filled in.
6. Enter values for other fields as follows:

Mirror a Volume Group fields	
Field	**Value**
Resource Group Name	SMIT displays the name of the resource group to which this shared volume group belongs.
VOLUME GROUP name	SMIT displays the name of the volume group that you selected to mirror.
Node List	The nodes on which this volume group is known.
Reference node	SMIT displays the node from which the name of the physical disk was retrieved.
VOLUME names	SMIT displays the name of a physical volume on the volume group that you selected to unmirror. This is the hdisk name on the reference node.
FORCE deallocation of all partitions on this physical volume?	The default is **no**.
Mirror sync mode	Select **Foreground**, **Background**, or **No Sync**. **Foreground** is the default.
Number of COPIES of each logical partition	Select **2** or **3** . The default is **2**.

Field	Value
Keep Quorum Checking On?	You can also select **yes** or **no** .The default is **no**.
Create Exact LV Mapping?	The default is **no**.

7. If this panel reflects the correct information, press Enter to mirror the shared volume group. All nodes in the cluster receive this updated information.

Unmirroring a volume group using C-SPOC

This topic describes unmirroring a shared volume group using the C-SPOC utility.

To unmirror a shared 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, enter the fastpath `smit cspoc`.
3. In SMIT, select **Storage › Volume Groups › Unmirror a Volume Group** and press Enter.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
4. Select a volume group and press Enter.
5. Select entries from the list of nodes and physical volumes (**hdisks**) and Press Enter.
   SMIT displays the **Unmirror a Volume Group** panel, with the chosen fields filled in.
6. Enter values for other fields as follows:

Unmirror a Volume Group fields	
Field	**Value**
Resource Group Name	SMIT displays the name of the resource group to which this shared volume group belongs.
VOLUME GROUP name	SMIT displays the name of the volume group that you selected to mirror.
Node List	The nodes on which this volume group is known.
Reference node	SMIT displays the node from which the name of the physical disk was retrieved.
VOLUME names	SMIT displays the name of a physical volume on the volume group that you selected to unmirror. This is the hdisk name on the reference node.
Number of COPIES of each logical partition	Select **2** or **3** . The default is **2**.
Mirror Pool name	Select a mirror pool to unmirror a volume group.

7. If this panel reflects the correct information, press Enter to unmirror the shared volume group. All nodes in the cluster receive this updated information.

Synchronizing volume group mirrors

You can use the C-SPOC utility to synchronize shared LVM Mirrors by volume group.

To synchronize shared LVM Mirrors by volume group using the C-SPOC utility, complete the following steps:

1. Physical volumes (hdisks) in the volume group must be installed, configured, and all nodes must be available and running the **clcomd** daemon.
2. From the SMIT interface, select **System Management C-SPOC › Storage › Volume Groups › Synchronize LVM Mirrors › Synchronize By Volume Group** and press Enter.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
3. Select a volume group and press Enter.
   SMIT displays the **Synchronize Mirrors by Volume Group** panel, with the chosen entries filled in.
4. Enter values for other fields as follows:

Synchronize Mirrors by Volume Group fields	
Field	**Value**
Resource Group Name	SMIT displays the name of the resource group to which this shared volume group belongs.
VOLUME GROUP name	SMIT displays the name of the volume group that you selected to mirror.
Node List	The nodes on which this volume groups is known
Reference node	SMIT displays the node from which the name of the physical disk was retrieved.
Number of Partitions to Sync in Parallel	Leave empty.
Synchronize All Partitions	**The default is no**.
Delay Writes to VG from other cluster nodes during this Sync	The default is **no**.

5. If this panel reflects the correct information, press Enter to synchronize LVM mirrors by the shared volume group. All nodes in the cluster receive this updated information.

Synchronizing a shared volume group definition

This topic describes synchronizing a shared volume group definition using the C-SPOC utility.

To synchronize a shared volume group definition 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, enter the fastpath `smit cspoc`.
3. In SMIT, select **Storage › Volume Groups › Synchronize a Volume Group Definition** and press Enter.
   A list of volume groups known across the cluster is displayed. For each volume group, a list of nodes and the owning resource group, if any, are also displayed.
4. Select a volume group and press Enter.
   The command runs. All nodes in the cluster receive this updated information.

Creating critical volume groups

Critical volume groups are volume groups that you want to monitor for continuous access. You can configure critical volume groups in PowerHA® SystemMirror® 7.1.0, or later.

Critical volume groups contain volume groups that are critical to and application. You can configure how PowerHA® SystemMirror® 7.1.0, or later, responds when access to the application data in the volume group is lost. For example, in an environment that is using Oracle Real Application Clusters (RAC) 11gR2, you can specify the volume group that contains the RAC voting disk as a critical volume group.

To create a critical volume group in an existing cluster with nodes, complete the following steps:

1. From the command line, enter `smit cl_vg` and select **Create a Volume Group**.
2. From the list of available nodes, select the node that you want to create the critical volume group on and press Enter.
3. From the list of available disks on the node, select the ones that you want to include in the critical volume group and press Enter.
4. From the **Volume Group Type** menu, select **Scalable** and press Enter.
5. From the **Create a Scalable Volume Group** menu, enter the following values:
   Create a Scalable Volume Group fields

Create a Scalable Volume Group fields	
Field name	**Description**
Resource Group name	Press F4 to select an available resource group from a list.
Volume Group name	Enter the name for the volume group. The name must be unique across all nodes in the cluster.
Enable Fast Disk Takeover or Concurrent Access	Press F4 and select **Concurrent Access**.
Critical volume group?	Press F4 and select **yes**.

You can use the default values for all other fields.

6. Verify and synchronize the cluster.

AFTER COMPLETING THE TASK
If access to a critical volume group is lost, you can configure PowerHA® SystemMirror® to respond in the following ways:

- Run a notification method
- Halt all node processes
- Fence the node away from the disks so that the node remains online, but it cannot access the disk
- Shutdown cluster services and bring all resource groups offline.

You can make these configuration changes from the following menu in SMIT, **System Management (C-SPOC) › Storage › Manage Critical Volume Groups › Configure failure actions for Critical Volume Groups**.

With IBM® AIX® 7 with Technology Level 3, and later, the **CriticalVG** attribute provides protection against disk failures and volume group failures. The **CriticalVG** attribute can be set as follows:

- If the **CriticalVG** attribute is set for a rootvg volume group, the node gets rebooted when a rootvg volume group failure is detected.
- If the **CriticalVG** attribute is set for a normal non-rootvg volume group, when an I/O error occurs in the logical volume manager (LVM), the LVM checks the metadata area of the LVM. If the metadata checks fails, the LVM declares the volume group as offline, and then PowerHA® SystemMirror® relocates the resource group that depends on the failed volume group.

Note: The AIX® **CriticalVG** attribute is different from the PowerHA® **CriticalVG** attribute. PowerHA® SystemMirror® sets the **CriticalVG** attribute for all the disks that are associated with the resource groups that are managed by PowerHA® SystemMirror®. The **CriticalVG** attribute enables PowerHA® SystemMirror® to detect storage failures quickly and to resolve it.

Migrating an existing Oracle RAC cluster to PowerHA® SystemMirror®

You can migrate only an Oracle Real Application Clusters (RAC) 11gR2 Version 11.2.0.1, or later, cluster to PowerHA® SystemMirror® Version 7.1.0, or later.

If you are using PowerHA® SystemMirror® 6.1, or earlier, you must first upgrade to Oracle RAC 11gR2 before you upgrade PowerHA® SystemMirror®.
To migrate an existing Oracle RAC 11gR2, or later, cluster to be used with PowerHA® SystemMirror®, complete the following steps:

1. From the command line, enter `smit cl_manage_critical_vgs`.
2. From the SMIT interface, select **Mark a Volume Group as Critical** and press Enter.
3. From the list of volume groups, select the volume group that contains the Oracle voting disk and press Enter.
4. Verify and synchronize the cluster.

Enabling Safeguard function for a volume group using C-SPOC

The Safeguarded Copy function supports the ability to create cuber-resilient, point-in-time copies of volumes that cannot be changed or deleted by user errors, malicious actions, or ransomware attacks.

The PowerHA® SystemMirror® 7.2.8 integrates with IBM® Copy Services Manager (CSM) to provide automated backup copies and data recovery. The PowerHA® SystemMirror® 7.2.8 supports IBM® Copy Services Manager as an external scheduling application. IBM® Copy Services Manager coordinates and automates Safeguarded Copy function across multiple systems. For more information see, Safeguarded Copy function - IBM® Documentation.

> **Note:** The Safeguarded Copy function is available with PowerHA storage 8.4.2 or later, and is not supported for the IBM®® FlashSystem 5015, IBM FlashSystem® 5035, FlashSystem 5010, FlashSystem 5030, Storwize® V5010E, Storwize® V5030E, Storwize® V7000 Gen2, and Storwize® V7000 Gen2+ models.

Starting with PowerHA® SystemMirror® 7.2.8, support for management of Safeguard Copy functions is provided. By using PowerHA® SystemMirror®, you can enable an AIX® volume group for safeguard snapshots function by assigning a safeguard policy.

Before you enable the safeguard function for an AIX volume group, you must complete the following requirements.

- Configure a PowerHA® SystemMirror® cluster and create a volume group by using the C-SPOC utility.
- Ensure that the disks of volume group are mapped to a SVC storage device that supports the safeguard function.
- Configure a safeguard child pool inside the storage device as a safeguard backup location for the safeguard copies or snapshots.
- Configure a user in the SVC storage and enable password-less communication with the hosts to perform safeguard operations by using the CLI commands.

To enable the safeguard function for a volume group, complete the following steps by using the SMIT interface:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Storage › Volume Groups › Safeguard a Volume Group** and press Enter. **The Safeguard a Volume Group** panel is displayed.

Limitations
In PowerHA® SystemMirror®, the functionalities are not implemented as it requires additional user privileges. You must perform these functions manually by using storage Security admin privileges.

- Disable safeguard
- Restoring safeguard copy
- Unassign a policy
- Extend VG and Shrink VG

Configuring storage

You can configure a storage device by adding or updating the SAN Volume Controller (SVC) storage.

- From the command line, enter `smit cl_admin`.
- In SMIT, select **Storage › Volume Groups › Safeguard a Volume Group › Configure Storage** and press Enter.

  Adding SVC storage

  To add SVC storage, complete the following steps.

  - To configure the SVC storage, enter the values for the cluster name, type, SVC cluster IP address, and user name fields as shown in Table 1.
  - A user do not require a password to communicate with the storage device. Also, the user should be able to ping the configured IP address.

Configuring SVC cluster	
Field	**Value**
Cluster names	Enter the same name that is used by SVC. The name can contain up to 20 alphanumeric characters and underscores.
Type	Select type as **SVC**.
SVC cluster IP address	Enter the IP address of SVC storage cluster.
SVC cluster second IP address	You can enter a second IP address for SVC storage cluster. If the IP address that is specified in the SVC cluster IP Address field is not available, the IP address specified in this field is used.
User name	Enter the user name that is used to access the IBM SVC cluster by using the Secured Shell (SSH) protocol. The user name must be configured to use an SSH key-based authentication.

Changing or modifying an SVC cluster

To modify the configured SVC storage information, update the following field values.

Modifying SVC cluster	
Field	**Value**
SVC cluster name	Enter the current name of the SVC cluster.
New SVC cluster name	Enter the name of the SVC cluster. You can use a maximum of 20 characters.
SVC cluster IP address	Enter the IP address of SVC cluster.
User name	Enter the user name that is used to access the IBM SVC cluster by using the Secured Shell (SSH) protocol. The user name must be configured to use an SSH key-based authentication.

Removing SVC cluster

To remove the SVC storage configuration with a cluster name, complete the following steps.

a. Select the SVC cluster you want to remove from the selected list and press Enter.

b. Press Enter to confirm that you want to remove this SVC cluster definition.

Configuring safeguard snapshot policy

You can configure a safeguard snapshot policy by creating a new policy.

- From the command line, enter `smit cl_admin`.
- In SMIT, select **Storage › Volume Groups › Safeguard a Volume Group › Configure Snapshot Policy** and press Enter.

  Creating safeguard snapshot policy

  To create a safeguard snapshot policy, complete the following steps.

  - Create safeguard snapshot policy with Policy name, Snapshot backup unit, Backup Interval, Retention days, and Policy start time as shown in Table1.
  - A user do not require a password to communicate.

Creating safeguard snapshot policy	
Field	**Value**
Policy name	Enter the safeguard snapshot policy name to be created.

Field	Value
Snapshot backup unit	Select backup unit in the following format: `minute/hour/day/week/month`.
Backup interval	Enter the backup interval for the safeguard snapshot policy.
Retention days	Enter number of days the safeguard snapshot created at the time of snapshot creation, should remain available.
Policy start time	Policy start time is the start time from when the snapshots are created for the volume group. The format of start time is `yymmddHHMM`.

Removing safeguard snapshot policy

To remove a safeguard snapshot policy, complete the following steps.

a. Select the name of the snapshot policy you want to remove from the selected list and press Enter.

b. Press Enter to confirm that you want to remove this safeguard snapshot policy.

Other:

a. Delete the safeguard snapshot policy based on the policy name.

b. A snapshot policy cannot be deleted if it is associated with a volume group.

c. Predefined policies cannot be deleted.

List safeguard snapshot policy

Safeguard snapshot policies are displayed on selecting the option. It lists all the snapshot policies along with the predefined policies. The values of backup unit, backup interval, policy name, retention days, and policy start time are displayed.

Enabling safeguard snapshot function

You can enable the safeguard snapshots function for a volume group.

You can enable the safeguard snapshots function for a volume group whose volumes mapped to a SAN Volume Controller (SVC) storage device. The SVC device supports the safeguard function by using the safeguard snapshot policy and start time. This creates a volume group in the storage device and adds the volumes that are mapped to the hard disks. This also assigns the snapshot policy to a volume group to automate snapshots in the safeguarded child pool.

Safeguard a volume group	
Field	**Value**
Select safeguard volume group	Select the volume group to be safeguarded.
Select safeguard policy ID	Select the safeguard snapshot policy ID.
Start time(optional)	Safeguard snapshot for volume group start time. If the start time is in past, the snapshot is created based on the frequency of the snapshot policy. If the start time is in future, the snapshot is created from the specified start time. If the start time is not specified, the snapshot policy start time is considered internally. The start time format should be in `yymmddHHMM`.

List safeguarded volume group

You can generate a list of the names of the safeguarded volume group.

Lists the names of the safeguarded volume groups for which the safeguard snapshot policy is assigned and for which the snapshots is created in the safeguarded location.

Maintaining logical volumes

These topics describe the administrative tasks involve shared logical volumes. You can perform all these tasks using the C-SPOC utility.

Adding a logical volume to a cluster using C-SPOC

This topic describes adding a logical volume to a cluster using the C-SPOC utility.

To add a logical volume to a cluster using C-SPOC:

1. Enter the C-SPOC fastpath, `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Add a Logical Volume** and press Enter.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
3. Select a resource group-volume group combination and press Enter.
   SMIT displays a list of physical volumes and the Auto-Select option.
4. Select a physical volume and press Enter.
5. Select Auto-Select to allow the AIX® Logical Volume Manager (LVM) to place the logical volume anywhere in the volume group. Complete the following fields:

Add a Shared Logical Volume fields	
Field	**Value**
Resource Group name	SMIT displays the name of the resource group to which this volume group and logical volume belong
VOLUME GROUP name	SMIT displays the name of the volume group you selected to hold this logical volume
Node List	The nodes on which this volume groups is known
Reference node	SMIT displays the node from which the name of the physical disks were retrieved.
Number of LOGICAL PARTITIONS	Determines the size of the logical volume
PHYSICAL VOLUME names	SMIT displays the names of the physical disks you selected to hold this logical volume
Logical volume NAME	Enter the name of your choice, unique across the cluster, for the logical volume, or leave blank to have C-SPOC provide an appropriate name.
Logical volume TYPE	Based on the expected used of the logical volume
POSITION on physical volume	Middle is the default. Any valid value can be used.
RANGE of physical volumes	Minimum is the default. Any valid value can be used.
Mirror Write Consistency?	Specifies the data consistency among mirrored copies of a logical volume during the normal I/O processing for enhanced concurrent or non concurrent mode volume groups, such as VGs in the active or passive state. **Active** is the default. Other options are **passive** and **off**.
Allocate each logical partition copy on a SEPARATE physical volume?	**yes** is the default. If you specify a forced varyon attribute in SMIT for volume groups in a resource group, it is recommended to set this field to **super strict**.
RELOCATE the logical volume during reorganization	Yes is the default. Either yes or no can be used.

Field	Value
Logical volume LABEL	Leave blank, if the logical volume will hold a file system.
MAXIMUM NUMBER of LOGICAL PARTITIONS	512 is the default. The value must be greater than or equal to the number parameter.
Enable BAD BLOCK relocation?	Cannot be specified for enhanced concurrent mode volume groups
SCHEDULING POLICY for reading/writing logical partition copies	Parallel is the default. Any valid value can be used
Enable WRITE VERIFY?	No is the default. Either no or yes can be used
File containing ALLOCATION MAP	Enter the name of any mapfile you have created to define the layout of the logical volume
Stripe Size?	Specify any valid value to create a striped logical volume
Serialize I/O?	No is the default. This is appropriate for file systems and data bases
Make first block available for applications?	Enter Yes for data bases that are using raw logical volumes.
Enable LVM Encryption	Use this option to enable logical volume (LV) level encryption. To use LV level encryption, you must first enable VG level encryption. To avoid verification error, you must add the authentication method when you enable LVM encryption.
Auth Method	Use this option to add either key server or PKS authentication method.
Method Details	Use this option to specify the key server ID, which is configured as a part of key server. This option is applicable only for key server authentication method.
Auth Method Name	This option is an alias name that is used to differentiate multiple authentication methods that are configured to a logical volume.

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

Changing a logical volume to a cluster using C-SPOC

This topic describes changing a logical volume to a cluster using the C-SPOC utility.

To change a logical volume to a cluster using C-SPOC:

1. Enter the C-SPOC fastpath, `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Change a Logical Volume** and press Enter.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
3. Select a resource group-volume group combination and press Enter.
   SMIT displays a list of physical volumes and the Auto-Select option.
4. Select a logical volume and press Enter.
5. Select Auto-Select to allow the AIX® Logical Volume Manager (LVM) to place the logical volume anywhere in the volume group. Complete the following fields:

Change a Shared Logical Volume fields	
Field	**Value**
Resource Group name	SMIT displays the name of the resource group to which this volume group and logical volume belong
VOLUME GROUP name	SMIT displays the name of the volume group you selected to hold this logical volume
MAXIMUM NUMBER OF PHYSICAL VOLUMES to use for allocation	Sets the maximum number of physical volumes for new allocation. The set value should be between one and the total number of physical volumes in the volume group.
Logical volume NAME	Enter the name of your choice, unique across the cluster, for the logical volume, or leave blank to have C-SPOC provide an appropriate name.
Logical volume TYPE	Based on the expected used of the logical volume
POSITION on physical volume	Middle is the default. Any valid value can be used.
RANGE of physical volumes	Minimum is the default. Any valid value can be used.
Mirror Write Consistency?	Specifies the data consistency among mirrored copies of a logical volume during the normal I/O processing for enhanced concurrent or non concurrent mode volume groups, such as VGs in the active or passive state. **Active** is the default. Other options are **passive** and **off**.
Allocate each logical partition copy on a SEPARATE physical volume?	**yes** is the default. If you specify a forced varyon attribute in SMIT for volume groups in a resource group, it is recommended to set this field to **super strict**.
RELOCATE the logical volume during reorganization	**Yes** is the default. Either **Yes** or **No** can be used.
Logical volume LABEL	Leave blank, if the logical volume will hold a file system.
MAXIMUM NUMBER of LOGICAL PARTITIONS	512 is the default. The value must be greater than or equal to the number parameter.
Enable BAD BLOCK relocation?	Cannot be specified for enhanced concurrent mode volume groups
SCHEDULING POLICY for reading/writing logical partition copies	Parallel is the default. Any valid value can be used
Enable WRITE VERIFY?	**No** is the default. Either **No** or **Yes** can be used
Permissions	Sets the access permission to read-write or read-only
Serialize I/O?	**No** is the default. This field is required for file system and databases. Enter **Yes** for the databases that use raw logical volume.
Enable LVM Encryption	Use this option to enable the logical volume (LV) encryption. To use LV encryption, you must first enable the volume group (VG) encryption. To avoid errors generated during cluster verification, you must add the authentication method when you enable the LVM encryption.
Auth Method	Use this option to add either the key server or PKS authentication method.

Field	Value
Method Details	Use this option to specify the key server ID, which is configured as a part of key server. This option is applicable only for the key server authentication method.
Auth Method Name	This option is an alias name that is used to differentiate multiple authentication methods that are configured to a logical volume.
Remove Auth Method	Use this option to remove authentication method name. Do not select any other LV level encryption fields while selecting **Remove Auth Method**

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

Changing a logical volume to a cluster using C-SPOC

This topic describes changing a logical volume to a cluster using the C-SPOC utility.

To change a logical volume to a cluster using C-SPOC:

1. Enter the C-SPOC fastpath, `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Change a Logical Volume** and press Enter.
   SMIT displays a list of shared volume groups and their owning resource group, if any, and the node list.
3. Select a resource group-volume group combination and press Enter.
   SMIT displays a list of physical volumes and the Auto-Select option.
4. Select a logical volume and press Enter.
5. Select Auto-Select to allow the AIX® Logical Volume Manager (LVM) to place the logical volume anywhere in the volume group. Complete the following fields:

Change a Shared Logical Volume fields	
Field	**Value**
Resource Group name	SMIT displays the name of the resource group to which this volume group and logical volume belong
VOLUME GROUP name	SMIT displays the name of the volume group you selected to hold this logical volume
MAXIMUM NUMBER OF PHYSICAL VOLUMES to use for allocation	Sets the maximum number of physical volumes for new allocation. The set value should be between one and the total number of physical volumes in the volume group.
Logical volume NAME	Enter the name of your choice, unique across the cluster, for the logical volume, or leave blank to have C-SPOC provide an appropriate name.
Logical volume TYPE	Based on the expected used of the logical volume
POSITION on physical volume	Middle is the default. Any valid value can be used.
RANGE of physical volumes	Minimum is the default. Any valid value can be used.
Mirror Write Consistency?	Specifies the data consistency among mirrored copies of a logical volume during the normal I/O processing for enhanced concurrent or non concurrent mode volume groups, such as VGs in the active or passive state. **Active** is the default. Other options are **passive** and **off**.
Allocate each logical partition copy on a SEPARATE physical volume?	**yes** is the default.

Field	Value
	If you specify a forced varyon attribute in SMIT for volume groups in a resource group, it is recommended to set this field to **super strict**.
RELOCATE the logical volume during reorganization	**Yes** is the default. Either **Yes** or **No** can be used.
Logical volume LABEL	Leave blank, if the logical volume will hold a file system.
MAXIMUM NUMBER of LOGICAL PARTITIONS	512 is the default. The value must be greater than or equal to the number parameter.
Enable BAD BLOCK relocation?	Cannot be specified for enhanced concurrent mode volume groups
SCHEDULING POLICY for reading/writing logical partition copies	Parallel is the default. Any valid value can be used
Enable WRITE VERIFY?	**No** is the default. Either **No** or **Yes** can be used
Permissions	Sets the access permission to read-write or read-only
Serialize I/O?	**No** is the default. This field is required for file system and databases. Enter **Yes** for the databases that use raw logical volume.
Enable LVM Encryption	Use this option to enable the logical volume (LV) encryption. To use LV encryption, you must first enable the volume group (VG) encryption. To avoid errors generated during cluster verification, you must add the authentication method when you enable the LVM encryption.
Auth Method	Use this option to add either the key server or PKS authentication method.
Method Details	Use this option to specify the key server ID, which is configured as a part of key server. This option is applicable only for the key server authentication method.
Auth Method Name	This option is an alias name that is used to differentiate multiple authentication methods that are configured to a logical volume.
Remove Auth Method	Use this option to remove authentication method name. Do not select any other LV level encryption fields while selecting **Remove Auth Method**

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

Setting characteristics of a shared logical volume using C-SPOC

These topics contain instructions for tasks that you can do for all cluster nodes from one node with the C-SPOC utility.

Renaming a shared logical volume using C-SPOC

This topic describes renaming a shared logical volume on all nodes in a cluster by executing a C-SPOC command on any node.

To rename a shared logical volume, complete the following steps:

1. Enter the fastpat, `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Set Characteristics of a Logical Volume › Rename a Logical Volume** and press Enter. When SMIT displays the Rename a Logical Volume on the Cluster panel press Enter.
   SMIT displays a picklist of shared volume groups and their owning resource group, if any, and the list of nodes on which the volume group is known.
3. Select a logical volume and press Enter. SMIT displays a panel with the **Resource group name**, **Volume Group name**, **Node list**, and **Current® logical volume name** fields completed.
4. 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.

Result

> **Note:** After completing this procedure, confirm your changes by initiating failures and verifying correct fallover behavior before resuming normal cluster operations.

Increasing the size of a shared logical volume using C-SPOC

This topic describes increasing the size of a shared logical volume on all nodes in a cluster using the C-SPOC utility.

To increase the size of a shared logical volume on all nodes in a cluster:

1. On any node, enter the SMIT fastpath: `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Set Characteristics of A Logical Volume › Increase the Size of a Logical Volume** and press Enter. SMIT displays a list of logical volumes arranged by resource group.
3. SMIT displays a list of logical volumes in the previously selected volume group.
4. Select a logical volume from the picklist and press Enter.
   SMIT displays a list of physical volumes.
5. Select a physical volume and press Enter. SMIT displays the **Increase Size of a Logical Volume** panel with the **Resource Group**, **Logical Volume**, **Reference Node** and default fields filled.
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.

Adding a copy to a shared logical volume using C-SPOC

This topic describes adding a copy to a shared logical volume on all nodes in a cluster using the C-SPOC utility.

To add a copy to a shared logical volume on all nodes in a cluster:

1. On any node, enter the fastpath `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Set Characteristics of A Logical Volume › Add a Copy from to Logical Volume** and press Enter. SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
3. Select a volume group from the picklist and press Enter. SMIT displays a list of logical volumes in the selected volume group.
4. Select a logical volume from the picklist and press Enter. SMIT displays a list of physical volume and the Auto-Select option.
5. Select a physical volume or Auto-Select and press Enter. Selecting Auto-Select allows the AIX® Logical Volume Manager to place the logical volume any where in the volume group. SMIT displays the **Add a Copy to a Logical Volume** panel with the **Resource Group**, **Logical Volume**, **Node list**, **Reference Node**, and the default fields filled.
6. Enter the new number of mirrors in the **NEW TOTAL number of logical partition copies** field and press Enter. The C-SPOC utility changes the number of copies of this logical volume on all cluster nodes.

Removing a copy from a shared logical volume using C-SPOC

This topic describes removing a copy from a shared logical volume on all nodes in a cluster using the C-SPOC utility.

To remove a copy of a shared logical volume on all nodes in a cluster:

1. On any node, enter the fastpath `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Set Characteristics of A Logical Volume › Remove a Copy from a Logical Volume** and press Enter. SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
3. Select a volume group from the picklist and press Enter. SMIT displays a list of logical volumes in the selected volume group.
4. Select the logical volume from the picklist and press Enter. SMIT displays a list of nodes and physical volumes.
5. Select the physical volumes from which you want to remove copies and press Enter. SMIT displays the **Remove a Copy from a Logical Volume** panel 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.
7. Enter mirror pool names in the **Mirror Pool Name** files to remove the copies of the logical volume.

Changing a shared logical volume

This topic describes changing the characteristics of a shared logical volume on all nodes in a cluster.

To change the characteristics of a shared logical volume:

1. On any node, enter the fastpath `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Change a Logical Volume** and press Enter. SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
3. Select a volume group from the picklist and press Enter. SMIT displays a list of logical volumes in the selected volume group.
4. Select the logical volume. SMIT displays the panel, with the values of the selected logical volume attributes filled in.
5. 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.

Removing a logical volume using C-SPOC

This topic describes removing a logical volume on any node in a cluster using the C-SPOC utility.

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

To remove a logical volume on any node in a cluster:

1. On any node, enter the fastpath `smit cspoc`.
2. In SMIT, select **Storage › Logical Volumes › Remove a Logical Volume** and press Enter. SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
3. Select the logical volume you want to remove and press Enter. The remote nodes are updated.

Synchronizing LVM mirrors by logical volume

You can use the C-SPOC utility to synchronize shared LVM mirrors by logical volumes.

Before you complete the following task, all nodes must be available and running the **clcomd** daemon.
To synchronize shared LVM mirrors, complete the following steps:

1. On any cluster node, from the command line enter `smit cspoc`.
2. In the C-SPOC menu, select **Storage › Volume Groups › Synchronize LVM Mirrors › Synchronize by Logical Volume** and press Enter. SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
3. Select a volume group from the picklist and press Enter. SMIT displays a list of logical volumes in the selected volume group.
4. Select a logical volume and press Enter. SMIT displays the **Synchronize LVM Mirrors by Volume Group** panel, with the chosen entries filled in.
5. Enter values for other fields as follows:

Field name	Description
Resource Group Name	SMIT displays the name of the resource group to which this logical volume belongs.
LOGICAL VOLUME name	SMIT displays the name of the logical volume that you selected to synchronize.
Node List	Can be empty, or set to a larger value to improve performance.
Synchronize All Partitions	This value is needed only when the integrity of a mirror is in question.
Delay Writes to VG from other cluster nodes during this Sync	Appropriate for volume groups in concurrent access configurations.

6. If this panel reflects the correct information, press Enter to synchronize LVM mirrors by the shared logical volume. All nodes in the cluster receive this updated information.

Maintaining shared file systems

These topics describe administrative tasks that involve shared file systems. These topics also describe how to use the C-SPOC utility to create, change or remove a shared file system in a cluster.

Journaled file system and enhanced journaled file system

Enhanced Journaled File System (JFS2) provides the capability to store much larger files than the Journaled File System (JFS). Additionally, it is the default file system for the 64-bit kernel. You can choose to implement either JFS, which is the recommended file system for 32-bit environments, or JFS2, which offers 64-bit functionality.

> **Note:** Unlike the JFS file system, the JFS2 file system will *not* allow the **link()** API to be used on files of type **directory**. This limitation may cause some applications that operate correctly on a JFS file system to fail on a JFS2 file system.

The SMIT paths shown in the following topic collections use the Journaled File System; similar paths exist for the Enhanced Journaled File System.

Reliable NFS server and enhanced journaled file system

You can use either JFS or JFS2 file systems with the Reliable NFS Server functionality of PowerHA® SystemMirror®.

Creating shared file systems with C-SPOC

This topic describes adding a shared file system where no logical volume is current defined using the C-SPOC utility.

Before creating a journaled file system 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 owning volume group does not need to be varied on to create a filesystem with C-SPOC.

Avoid using a period (.) in the file system name.

You can add a journaled file system or enhanced journaled file system to either of the following volumes:

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

To add a file system where no logical volume is currently defined, complete the following steps:

1. Enter the fastpath `smit cspoc`.
2. In the C-SPOC interface, select **Storage › File System › Add a File System** and press Enter. SMIT displays a list of the filesystem types, Standard, Enhanced, Compressed, or Large File Enabled.
3. Select the volume group where the file system will be added.
4. Select a file system type from the list.
5. Enter field values as follows:

Filesystem attribute fields	
Field	**Value**
Resource Group	Displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
Node Names	Displays the names of the cluster nodes on which the volume group is known.
Volume Group Name	Displays the selected volume group name.
SIZE of filesystem	Set the value as required. The size can be specified in 512-byte blocks, megabytes or gigabytes.
MOUNT POINT	Enter the mount point for the file system.
PERMISSIONS	Set the value as required.
Mount OPTIONS	Set the value as required. Select from the predefined mount options.
Extended Mount OPTIONS	Set the value as required. This field provides supported mount options that are not listed in the predefined **Mount OPTIONS** field. Enter one or more values, separated by a comma. Changes in the **Extended Mount Options** field is reflected only when you restart the cluster services.
Start Disk Accounting?	Set the value as required. The default value is **no**.
Fragment Size (Bytes)	4096 is the default.
Number of Bytes per inode	4096 is the default.
Allocation Group Size (MBytes)	8 is the default.
Logical Volume for Log	Enter the logical volume that is used as the logging device for the new file system.

6. Select the file system attributes and press Enter.
   When you specify values in the **Extended Mount OPTIONS** field, an entry is added to the `/usr/es/sbin/cluster/etc/mntopt` mapping file. This new file is automatically created if it does not exist and included in the PowerHA® SystemMirror® file collection under the `mountoptions` label. It helps ensure that the file stays synchronized and consistent across all the cluster nodes.

Result
SMIT checks the nodelist for the resource group that contains the volume group, creates the logical volume and uses an existing log logical volume if present, otherwise a new log logical volume is created. Next, the filesystem is created by using the logical volume on the node where the volume group is varied-on.

Adding the file system to a PowerHA® SystemMirror® cluster logical volume

You can add a file system to a previously defined cluster logical volume.

To add a file system:

1. Enter the fast path `smit cspoc`.
2. In SMIT, select **Storage › file system › Add a file system** and press Enter. SMIT displays a list of the following file system types: Standard, Enhanced, Compressed, or Large File Enabled.
3. Select the file system type from the list. SMIT generates a list of all free logical volumes in the cluster, the owning volume group and resource group, if any, and the nodes the free logical volumes are on. SMIT reports a logical volume as free in case the logical volume does not have a mount point for the file system.
4. Select a logical volume where the file systems are added. SMIT displays the AIX® **SMIT** window for selecting file system attributes.
5. Enter the field values as follows:

File system attribute fields	
Field	**Value**
Resource Group	SMIT displays a list of shared volume groups, their owning resource group, if any, and the list of nodes on which the volume groups are known.
Volume Group Name	SMIT displays the selected volume group name.
Node Names	SMIT displays the names of the selected cluster nodes.
LOGICAL VOLUME name	SMIT displays the name of the selected logical volume.
MOUNT POINT	Enter the mount point for the file system.
PERMISSIONS	Set the value as required.
Mount OPTIONS	Set the value as required. Select from the predefined mount options.
Extended Mount OPTIONS	Set the value as required. This field provides supported mount options that are not listed in the predefined **Mount OPTIONS** field. Enter one or more values, separated by a comma.
Start Disk Accounting?	Set the value as required. Default is no.
Fragment Size (Bytes)	4096 is the default.
Number of Bytes per inode	4096 is the default.
Allocation Group Size (MBytes)	8 is the default.

6. Select the file system attributes and press Enter. SMIT checks the node list for the resource group that contains the volume group where the logical volume is located. It adds the file system to the node where the volume group is varied on. All other nodes in the resource group are notified about the new file system.
   When you specify values in the **Extended Mount OPTIONS** field, an entry is added to the `/usr/es/sbin/cluster/etc/mntopt` mapping file. This new file is automatically created if it does not exist and included in the PowerHA® SystemMirror® file collection under the `mountoptions` label. It helps ensure that the file stays synchronized and consistent across all the cluster nodes.

Changing a shared file system in PowerHA® SystemMirror® using C-SPOC

As system administrator of a PowerHA® SystemMirror® cluster, you may need to change the characteristics of an existing file system. Using the C-SPOC utility, you can change the characteristics of a shared file system on cluster nodes by executing a command on a single cluster node. The C-SPOC command changes the attributes of the shared file system on all the nodes in the resource group.

To change the characteristics of a shared file system:

1. Enter the fastpath `smit cspoc`.
2. In SMIT, select the **Storage › File System › Change/Show Characteristics of a File System** and press Enter.
   SMIT displays a picklist of existing file systems.
3. Select the file system to change.
   SMIT displays a panel containing the characteristics of the file system.
4. Enter data in the fields to change and press Enter. The C-SPOC utility changes the file system characteristics on all nodes in the resource group.

Removing a shared file system using C-SPOC

As system administrator of a PowerHA® SystemMirror® cluster, you may need to remove a file system. You can optionally remove the file system's mount point as part of the same operation. Using this procedure, you can remove a shared file system on any node in a cluster.

C-SPOC deletes the shared file system on the node that currently has the shared volume group varied on. It removes both the shared logical volume on which the file system resides and the associated stanza in the /etc/filesystems file.

To remove a shared file system:

1. Enter the fastpath `smit cspoc`.
2. In SMIT, select **Storage › File System › Remove a File System** and press Enter.
3. Press the F4 key to obtain a picklist of existing file systems from which you can select one. Select Yes for the Remove Mount Point option if you want to remove the mount point. When you finish entering data, press Enter.

   > When a file system that is configured with extended mount options is removed, the corresponding entry in the `/usr/es/sbin/cluster/etc/mntopt` mapping file is automatically deleted. The `mountoptions` file collection ensures that this change is distributed to all cluster nodes.

   The C-SPOC utility removes the file system on the local node. All other nodes in the resource group are informed of the new filesystem.

Maintaining raw volumes

You can use the C-SPOC utility to complete the administrative tasks that involve raw volumes in a cluster.

You can use the C-SPOC utility in PowerHA® SystemMirror® to perform the administrative tasks related to the raw volumes of a configured disk on all the selected nodes in the cluster.

Removing a raw disk definition on cluster nodes using C-SPOC

This topic describes removing a raw disk on all selected nodes in the cluster using the C-SPOC utility.

Before removing a raw 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.

To remove a raw volume on configured disk on all selected nodes in the cluster:

1. Enter the fastpath `smitty cl_admin`

2. In SMIT, select **Storage › Raw Volumes › Remove a Raw Disk From the Cluster** and press Enter.
   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 one or more node names 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** panel 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.

Enabling or disabling encryption for raw volumes using C-SPOC

You can encrypt or decrypt raw volumes with the C-SPOC utility.

You can encrypt or decrypt raw volumes of a configured disk on all selected nodes in the cluster with the C-SPOC utility. You can encrypt or decrypt a disk if it is not part of a volume group. Before you encrypt or decrypt a disk in the cluster, verify that the disk is not part of an existing volume group. If the disk is part of an existing volume group, enter the **`smitty cl_vgsc`** command and select the **Remove a Volume from a Volume Group** option.

To encrypt or decrypt a raw volume with C-SPOC, complete the following steps:

1. Enter `smitty  cl_raw_disk_man`.
2. Select **Enable/Disable Encryption for a Raw Volume**.
3. Select the`enable/disable`for the physical volume that you selected and press Enter.
4. Select the disk which you want to encrypt or decrypt.
5. Select the`enable/disable`option for the selected raw volume in the disk.
6. Enter the authorisation name for the physical volume in the **Auth Method Name**field.
7. Press Enter.

Result
C-SPOC will encrypt or decrypt the raw volume of a disk from the cluster nodes.

Renaming raw volumes with C-SPOC

You can rename a configured disk on all selected nodes in the cluster with the C-SPOC utility.

AIX® does not guarantee that a given physical volume will have the same name on all nodes that have access to it. Therefore, it is ideal to rename a disk so that it has a common name on all nodes.

A disk can only be renamed if it is not part of a volume group. Before you rename a disk in the cluster, verify that the disk is not part of an existing volume group. If the disk is part of an existing volume group, enter `smitty cl_vgsc` from the command line and select **Remove a Volume from a Volume Group**.

To rename a raw volume with C-SPOC, complete the following steps:

1. From the command line, enter `smitty cl_raw_disk_man`.
2. Select **Rename a Raw Physical Volume**.
3. From the list, select the disk for which you want to rename.
4. Enter the new name for the disk. You cannot use a name that is already in use by another disk, volume group, or logical volume.
5. Press Enter.

Maintaining physical volumes

You can use C-SPOC utility to complete administrative tasks that involve shared physical volumes.

By default, the C-SPOC utility in PowerHA® SystemMirror® automatically assigns a PVID to a disk if a PVID is not already defined. To disable the automatic PVID assignment function, add `CL_PVID_ASSIGNMENT=0` to the `/etc/environment` file on every node in the cluster. The changes take effect immediately.

Note: If you disable the automatic PVID assignment function, you must assign a PVID to the disk outside of PowerHA® SystemMirror®. To use the C-SPOC utility for disk management, you must assign a PVID to the disk.

Removing a disk definition on cluster nodes using C-SPOC

This topic describes remoing a configured disk on all selected nodes in the cluster using the C-SPOC utility.

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.

To remove a configured disk on all selected nodes in the cluster:

1. Enter the fastpath `smitty cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Remove a Disk from the Cluster** and press Enter.
   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 one or more node names 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** panel 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.

Renaming physical volumes with C-SPOC

You can rename a configured disk on all selected nodes in the cluster with the C-SPOC utility.

AIX® does not guarantee that a given physical volume will have the same name on all nodes that have access to it. Therefore, it is ideal to rename a disk so that it has a common name on all nodes.

A disk can only be renamed if it is not part of a volume group. Before you rename a disk in the cluster, verify that the disk is not part of an existing volume group. If the disk is part of an existing volume group, enter `smitty cl_vgsc` from the command line and select **Remove a Volume from a Volume Group**.

To rename a physical volume with C-SPOC, complete the following steps:

1. From the command line, enter `smitty cl_disk_man`.
2. Select **Rename a Physical Volume**.
3. From the list, select the disk for which you want to rename.
4. Enter the new name for the disk. You cannot use a name that is already in use by another disk, volume group, or logical volume.
5. It is possible that the disk you want to rename has a different name on different nodes in the cluster. To rename all instances of the disk to the new name, in the **Change all Physical Volumes with this PVID?** field specify **Yes**.
6. Press Enter.

Enabling or disabling encryption for physical volumes using C-SPOC

You can encrypt or decrypt the physical volumes with the C-SPOC utility.

You can encrypt or decrypt the physical volumes of a configured disk on all selected nodes in the cluster with the C-SPOC utility. You can encrypt or decrypt a disk if it is not part of a volume group. Before you encrypt or decrypt a disk in the cluster, verify that the disk is not part of an existing volume group. If the disk is part of an existing volume group, enter the **`smitty cl_vgsc`** command and select **Remove a Volume from a Volume Group**.

To encrypt or decrypt a physical volume with C-SPOC, complete the following steps:

1. From the command line, enter `smitty cl_disk_man`.
2. Select **Enable/Disable Encryption for a Physical Volume**.
3. Select the disk which you want to encrypt or decrypt.
4. Select the`enable/disable`option for the selected physical volume in the disk.
5. Enter the authorisation name for the physical volume in the **Auth Method Name**field.
6. Press Enter.

Result
C-SPOC will encrypt or decrypt the physical volume of a disk from the cluster nodes.

Listing encrypted physical volumes with C-SPOC

You can display the list of encrypted physical volumes with the C-SPOC utility.

You can display the list of the physical volumes that have been encrypted in a configured disk with the C-SPOC utility.

To display the list of the encrypted physical volumes:

1. From the command line, enter `fastpath smit cl_admin`.
2. In SMIT, select **Storage › Physical Volumes › List All Encrypted Physical Volumes** and press Enter.
   SMIT displays a list of encrypted physical volumes in the configuration:

```
Disk               PVID              Auth Method           Auth Method Name
         hdisk3       00c1cfc835243d82           PKS                   pks_hdisk3
         hdisk12      00c1cfc82af5154e           PKS                   pks_hdisk12
```

Result
C-SPOC will display the list of encrypted physical volumes in a configured disk.

Using SMIT to replace a cluster disk

The SMIT interface simplifies the process of replacing a failed disk by using C-SPOC commands.

> **Note:** If you have VPATH devices configured, the procedure for replacing a cluster disk using C-SPOC requires additional steps.

Before you replace a disk, ensure that:

- You have root user privilege to perform the disk replacement.
- You have a replacement disk with an assigned PVID configured on all nodes in the resource group to which the volume group belongs. If you do not have the PVID assigned, run **chdev** on all nodes in the resource group.
- To add a new disk, remove the old disk and put the new one in its place.

To replace a disk in the cluster:

1. Locate the failed disk. Make note of the PVID volume group.
2. Enter `smitty cspoc`.
3. In SMIT, select **Storage › Physical Volumes › Cluster Disk Replacement** and press Enter.
   SMIT displays a list of disks that are members of volume groups contained in cluster resource groups. There must be two or more disks in the volume group where the failed disk is located. The list includes the volume group, the hdisk, the disk PVID, and the reference cluster node. (This node is usually the cluster node that has the volume group varied on.)

   > The new disk that is available for replacement must have a PVID assigned to in on all nodes in the cluster. Use the **chdev** command to assign a PVID to the disk.

4. Select the disk for disk replacement (**source disk**) and press Enter.
   SMIT displays a list of those available shared disk candidates that have a PVID assigned to them, to use for replacement. (Only a disk that is of the same capacity or larger than the failed disk is suitable to replace the failed disk.)
5. Select the replacement disk (**destination disk**) and press Enter.
   SMIT displays your selections from the two previous panels.
6. Press Enter to continue or Cancel to terminate the disk replacement process.
   SMIT warns you that continuing will delete any information you may have stored on the destination disk.
7. Press Enter to continue or Cancel to terminate.
   SMIT displays a command status panel, and informs you of the replacepv recovery directory.

   If disk configuration fails and you want to proceed with disk replacement, you must manually configure the destination disk. If you terminate the procedure at this point, be aware that the destination disk may be configured on more than one node in the cluster.

   The **replacepv** utility updates the volume group in use in the disk replacement process (on the reference node only).

   > SMIT displays the name of the recovery directory to use should **replacepv** fail. Make note of this information, as it is required in the recovery process.

   Configuration of the destination disk on all nodes in the resource group takes place.
8. If a node in the resource group fails to import the updated volume group, you must do this manually.

Result
C-SPOC will not remove the failed disk information from the cluster nodes, hdisk, and pdisk. You must do this manually.

Managing data path devices with C-SPOC

All VPATH disk operations currently supported on AIX® are now supported by C-SPOC. You can define and configure VPATH devices, add paths, configure defined VPATHs, and remove VPATH devices. You can also display VPATH device and adapter configuration and status.

You must have SDD 1.6.2.0, or later, or SDDPCM 2.1.1.0, or later, installed.

Displaying data path device configuration

This topic describes displaying data path device configuration.

To display data path device configuration:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Display Data Path Device Configuration** and press Enter.

   SMIT displays the node picklist.
3. Select a node and press Enter.

   SMIT displays the configuration as shown in the following example for node`herbert`:

```
PVID: 000240bfd57e0746
herbert: vpath9 (Avail pv shvg1) 10612027 = hdisk59 (Avail ) hdisk65 (Avail )
PVID: 000240ffd5691fba
herbert: vpath12 (Avail ) 10C12027 = hdisk62 (Avail pv ) hdisk68 (Avail pv )
PVID: 000240ffd5693251
herbert: vpath14 (Avail pv ) 10E12027 = hdisk64 (Avail ) hdisk70 (Avail )
PVID: 000240ffd56957ce
herbert: vpath11 (Avail ) 10812027 = hdisk67 (Avail pv ) hdisk71 (Avail pv )
PVID: 0002413fef72a8f0
herbert: vpath13 (Avail pv ) 10D12027 = hdisk63 (Avail ) hdisk69 (Avail )
PVID: 0002413fef73d477
herbert: vpath10 (Avail pv ) 10712027 = hdisk60 (Avail ) hdisk66 (Avail )
```

Displaying data path device status

This topic describes displaying data path device status.

To display data path device status:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Display Data Path Device Status** and press Enter.

   SMIT displays the node picklist.
3. Select a node and press Enter.
4. SMIT displays the status as shown in the following example for node `herbert`:

```
[TOP]
herbert: Total Devices : 6

PVID 000240bfd57e0746

herbert:
DEV#: 0 DEVICE NAME: vpath9 TYPE: 2105F20 SERIAL: 10612027
POLICY: Optimized
============================================================
Path# Adapter/Hard Disk State Mode Select Errors
0 fscsi1/hdisk59 OPEN NORMAL 1696 0
1 fscsi0/hdisk65 OPEN NORMAL 1677 0

PVID 000240ffd5691fba
[MORE...57]
```

Displaying data path device adapter status

This topic describes displaying data path device adapter status.

To display data path device adapter status:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Display Data Path Device Adapter Status** and press Enter.

   SMIT displays the node picklist.
3. Select a node and press Enter.
4. SMIT displays the status as shown in the following example for node `herbert`:

```
herbert:

Active Adapters :2

Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0 fscsi1  NORMAL   ACTIVE22040      61
    1 fscsi0  NORMAL   ACTIVE22130      61
```

Defining and configuring all data path devices

This topic describes defining and configuring all data path devices.

To define and configure data path devices:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Define and Configure all Data Path Devices** and press Enter.

Result
The command runs and the command status displays on the panel.

Adding paths to available data path devices

This topic describes adding paths to available data path devices.

To add paths to available data path devices:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Add Paths to Available Data Path Devices** and press Enter.
   SMIT displays the list of node names.
3. Select one or more nodes, and press Enter.

Result
The command runs and the command status is displayed on the panel.

Configuring a defined data path device

This topic describes configuring a defined data path devices.

To configure a defined data path device:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Configure a Defined Data Path Device** and press Enter.
   SMIT displays the list of node names.
3. Select one or more nodes, and press Enter.
   SMIT displays the list of defined VPATHs by PVID.
4. Select a PVID and press Enter.

Result
The command runs and the command status is displayed on the panel.

Removing a data path device

This topic describes removing a data path devices.

To remove a data path device:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Remove a Data Path Device** and press Enter.
   SMIT displays the list of node names.
3. Select a node and press Enter.
4. Keep the definition in Data Base Selector.
   SMIT displays the list of devices.
5. Select one or more devices and press Enter.

Result
The command runs and the command status is displayed on the panel.

Converting ESS hdisk device volume group to an SDD VPATH device volume group

This topic discusses converting ESS hdisk device volume group to an SDD VPATH device volume group.

To convert ESS hdisk volume groups to SDD VPATH device volume groups:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Convert ESS hdisk Device Volume Group to an SDD VPATH Device Volume Group** and press Enter.
   SMIT displays the picklist of ESS hdisk volume groups.
3. Select the ESS hdisk volume group to convert and press Enter.
   SMIT displays the current resource group and volume group names.

4. Press Enter.

Result
The command runs and the command status is displayed on the panel.

Converting SDD VPATH device volume group to an ESS hdisk device volume group

This topic discusses converting SDD VPATH device volume group to a hdisk device volume group.

To convert SDD VPATH device volume groups to ESS hdisk device volume groups:

1. Enter the fastpath `smit cl_admin`
2. In SMIT, select **Storage › Physical Volumes › Cluster Data Path Device Management › Convert SDD VPATH Device Volume Group to an ESS hdisk Device Volume Group** and press Enter.
   SMIT displays the picklist of SDD VPATH volume groups.
3. Select the one to convert and press Enter.
   SMIT displays the current resource group and volume group names.
4. Press Enter.

Result
The command runs and the command status is displayed on the panel.

Replacing a cluster disk with a VPATH device

If you need to replace a cluster disk that has a VPATH device configured, before you use C-SPOC, move the PVID of the VPATH devices to the corresponding hdisks. This is done by converting the volume group from VPATH devices to hdisks. After converting, use the C-SPOC procedure to replace a disk.

> **Note:** The C-SPOC disk replacement utility does not recognize VPATH devices. If you do not convert the volume group from VPATH to hdisk, then during the C-SPOC disk replacement procedure, PowerHA® SystemMirror® returns a “no free disks” message, although unused VPATH devices are available for replacement.

To replace a cluster disk that has a VPATH device configured, complete the following steps:

1. Convert the volume group from VPATHs to hdisks.
2. Use the C-SPOC procedure to replace a cluster disk.
3. Convert the volume group back to VPATH devices.

Configuring LVM split-site mirroring

LVM split-site mirroring is a mechanism that replicates data between the disk subsystems located and configured at two remote sites to ensure disaster recovery. You can use SMIT and C-SPOC to configure mirrored pools that belong to a volume group for LVM split-site mirroring.

Before you configure LVM split-site mirroring, you must complete the following tasks:

- Plan for implementing LVM split-site mirroring in your environment.
- Run the PowerHA® SystemMirror® discovery process.
- Configure all nodes and resource groups.
- Determine the disk names that are at each location.
- Verify that forced varyon is set to **Yes** for all resource groups that contain a volume group for LVM split-site mirroring.

Configuring LVM split-site mirroring for a new volume group

You can use SMIT and C-SPOC to configure mirrored pools for LVM split-site mirroring for a new volume group. The cluster services can be active or inactive when you configure mirrored pools.

To configure mirror pools for LVM split-site mirroring for a new volume group, complete the following steps:

1. Verify that the disks are visible at all locations by running the **cfgmgr** command and the **chdev** command.
2. From the command line, enter `smit cl_vg`.
3. From the SMIT interface, select **Create a Volume Group**.
4. Specify the nodes at all locations that will have access to the volume group.
5. Select all the disks at one location that will be in the volume group.
6. Select **Scalable** for the volume group type.
7. For the **Enable Strict Mirror Pools** field, enter **Superstrict**.

   > You can specify other volume group parameters in any of the available fields.

8. For the **Mirror Pool name** field, specify a name for the mirror poo,l and press Enter to create the volume group.
9. Navigate to the Volume Group panel and select **Set Characteristics of a Volume Group**.
10. Select **Add a Volume to a Volume Group**.
11. Select the volume group you created in step 8.
12. Select all disks at the other location that will be in the volume group.
13. Specify a name for the mirror pool at this location and press Enter.

    > If you have a third location for LVM split-site mirroring repeat step 8 - 13.

Configuring LVM split-site mirroring for an existing volume group

The steps for configuring mirror pools for LVM split-site mirroring for an existing volume group depend on the properties of the volume group and the locations of the disks in the volume group.

To configure mirror pools for LVM split-site mirroring for an existing volume group, complete the following steps:

1. Verify that the disks are visible at all locations by running the **cfgmgr** command and the **chdev** command.
2. Identify the type of the volume group.
   - If the volume group is not a scalable volume group, go to step 3.
   - If the volume group is a scalable volume group, go to step 7.
3. From the command line, enter `smit cl_vg`.
4. From the SMIT interface, select **Set Characteristics of a Volume Group › Change/Show characteristics for a Volume Group**.
5. Select the volume group, and press Enter.
6. Change the **Mirror Pool Strictness** field to **Superstrict**, and press Enter.

   > You can specify other volume group parameters in any of the available fields.

7. Identify the disks at each location and place them in location-specific mirror pools.

8. If you want to add disks to the volume group, continue to step 9; otherwise, the configuration is complete.
9. Navigate to the Volume Group panel, and select **Set Characteristics of a Volume Group**.
10. Select **Add a Volume to a Volume Group**.
11. Select the volume group for which you want to configure mirror pools for LVM split-site mirroring.
12. Select all disks at the other location that will be in the volume group.
13. Specify a name for the mirror pool at this location and press Enter.

> If you have a third location for LVM split-site mirroring, repeat step 9-13.

Configuring mirror pools

Using the SMIT interface you can display mirror pools, change characteristics of a mirror pool, remove a mirror pool, and add disks to mirror pools.

From the command line, enter `smit cl_mirrorpool_mgt` to display the following configuration options:

```
Show all Mirror Pools
Show Mirror Pools for a Volume Group
Change/Show Characteristics of a Mirror Pool
Add Disks to a Mirror Pool
Remove Disks from a Mirror Pool
Rename a Mirror Pool
Remove a Mirror Pool
```

To verify the placement of disks in the mirror pools for a volume group that is used for LVM split-site mirroring, select **Show Mirror Pools for a Volume Group**. If a disk is placed in the wrong mirror pool, you can remove the disk and add it to the correct mirror pool. If you complete this process after the logical volumes or the file systems are created on the volume group, it can affect the mirroring of the volume groups data.

Extending a volume group that uses LVM split-site mirroring

When mirror pools are used for LVM split-site mirroring, you need to have the same amount of disk space in each mirror pool. Otherwise, the expansion or creation of file systems and logical volumes is blocked before all the disk space is used.

If all disks that are being used in a cluster are the same size, the same number of disks can be added to each mirror pool.

If the disks in a volume group with LVM split-site mirroring are not the same size, you must ensure that an equal amount of space is added to each mirror pool. To determine the size of a disk, run the **`bootinfo -s`** command.

Managing shared LVM components in a concurrent access environment

There are a few different steps for managing a shared LVM components in a concurrent access environment using the C-SPOC facility compared to managing a non-concurrent access environment. However, most of the steps are done in exactly the same order and using exactly the same SMIT panels as a non-concurrent configuration.

You can define concurrent access volume groups and logical volumes on all the disk devices supported by the PowerHA® SystemMirror® software.

> **Note:** You cannot define file systems on a concurrent access volume group unless it is an enhanced concurrent mode volume group used as a serial resource.

Most maintenance tasks can be performed using the PowerHA® SystemMirror® C-SPOC utility. All operations for maintaining concurrent volume groups and logical volumes are done in exactly the same order and using exactly the same SMIT panels as non-concurrent configuration.

Understanding concurrent access and PowerHA® SystemMirror® 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 PowerHA® SystemMirror® 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 PowerHA® SystemMirror® events occur.

When intervening in a cluster, you must understand how nodes in a concurrent access environment control their interaction with shared LVM components. For example, the PowerHA® SystemMirror® **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. If resource groups are processed in parallel, **process_resources** calls **cl_mode3**.

The **cl_mode3** script calls the **varyonvg** command with the -c flag. For more information about this command and its flags, see Activating a 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 has cluster services stopped with resource groups brought offline, 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 volume groups with C-SPOC

C-SPOC uses the AIX® CLVM capabilities that allow changes to concurrent LVM components without stopping and restarting the cluster.

You can use the C-SPOC utility to do the following concurrent volume groups tasks:

- Create a concurrent volume group on selected cluster nodes (using hdisks or data path devices)
- Convert RAID concurrent volume groups to enhanced concurrent mode

All other operations on concurrent access volume groups are performed using the same SMIT panels and C-SPOC operations as non-concurrent volume groups.

To perform concurrent resource group maintenance tasks, use the following SMIT menu: **System Management (C-SPOC) › Resource Groups and Applications**.

This utility allows you to take a concurrent resource group online or offline (along with its resources - IP addresses, applications, and disks) - without stopping cluster services. For more information on Resource Group Migration, see Resource group migration.

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.

For creating a concurrent volume path on VPATH disks, see Managing data path devices with C-SPOC. If you add a VPATH disk to a volume group made up of hdisks, the volume group will be converted to VPATHs on all nodes.

- 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 are concurrent capable.

To create a concurrent volume group for a selected list of cluster nodes:

1. From the command line, enter `smit cspoc`.
2. In SMIT, select **Storage › Volume Groups › Create a Volume Group (or Create a Volume Group with Data Path Devices)** and press Enter.
   SMIT displays a list of cluster nodes.
3. Select one 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-picklist by PVIDs. If you are creating a volume group with data path devices, only disks capable of hosting them will be listed.
4. Select one or more PVIDs from the list and press Enter.
   SMIT displays the **cl_mkvg** panel with a major number inserted into the **Major Number** data field. The system determines this free major number; do not change it.
5. Enter field values as follows:

PVID fields

Field	Value
Node Names	Names of the selected nodes are displayed.
PVID	PVID of the selected disk.
VOLUME GROUP name	The name of the volume group must be unique within the cluster and distinct from the service IP address and resource group names; it should relate to the application it serves, as well as to any corresponding device. For example, `websphere_service_VG`. If you do not provide a name, a unique name will be generated.
Physical partition SIZE in megabytes	Accept the default.
Volume Group MAJOR NUMBER	The system displays the number C-SPOC has determined to be correct. If you change the volume group major number, the command might not be able to run on a node that does not have that major number currently available. Check for a commonly available major number on all nodes before changing this setting.
Enable Fast Disk Takeover or Concurrent Access	Select **Concurrent Access**. An enhanced concurrent mode volume group is created. If a resource group was created, it is created with the policies **online on all available nodes** and **never fall back**.
Volume Group Type	Displays the type of volume group. You cannot change this field.
CRITICAL volume group	Select **yes** to identify this volume group as a critical volume group. When a volume work is identified as a critical volume group, you can configure the volume group by using the **Manage Critical Volume Groups** option. This setting determines how PowerHA® SystemMirror® responds when it cannot access the volume group.

Result

C-SPOC verifies communication paths and version compatibility and then runs 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.

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. Enhanced concurrent mode uses the **gsclvmd** daemon, which starts when PowerHA® SystemMirror® services are started.

Activating a volume group in concurrent access mode

As a system administrator, you may, at times, need to bring a resource group online. After correcting the failure, bring the resource group online

Follow these steps:

1. Enter `smitty cl_admin`
2. In SMIT, select **Resource Group and Applications › Bring a Resource Group Online**
3. Select the resource group to bring online and press Enter.

Activating concurrent access volume groups

This topic describes activating a volume group in concurrent access mode.

To activate a volume group:

1. Enter `smit varyonvg`
   The Activate a Volume Group SMIT panel appears; it has an additional field in a concurrent access environment.
2. Enter the field values as follows:

Activate a Volume Group fields	
Field	**Value**
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 prompts you to confirm. Press Enter again.

Determining the access mode of a volume group

To determine whether 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 information about the volume group, as in the following example:

```
# lsvg db2_vg
VOLUME GROUP:       db2_vg                  VG IDENTIFIER:
```

```
 00c3a28e00004c000000014184437f98
VG STATE:           active                   PP SIZE:        4 megabyte(s)
VG PERMISSION:      read/write               TOTAL PPs:      988 (3952 megabytes)
MAX LVs:            256                      FREE PPs:       983 (3932 megabytes)
LVs:                2                        USED PPs:       5 (20 megabytes)
OPEN LVs:           0                        QUORUM:         2 (Enabled)
TOTAL PVs:          2                        VG DESCRIPTORS: 3
STALE PVs:          0                        STALE PPs:      0
ACTIVE PVs:         2                        AUTO ON:        no
Concurrent:         Enhanced-Capable         Auto-Concurrent: Disabled
VG Mode:            Non-Concurrent
MAX PPs per VG:     32768                    MAX PVs:        1024
LTG size (Dynamic): 512 kilobyte(s)          AUTO SYNC:      no
HOT SPARE:          no                       BB POLICY:      relocatable
MIRROR POOL STRICT: super
PV RESTRICTION:     none                     INFINITE RETRY: no
DISK BLOCK SIZE:    512
```

To determine whether the volume group is concurrent capable, check the value of the Concurrent field. The volume group in the example was created as an Enhanced-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 or the Concurrent field is not present.

To determine whether 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 a PowerHA® SystemMirror® environment, this option should always be disabled; PowerHA® SystemMirror® scripts control when the volume should be varied on.

Verifying a concurrent volume group

On all nodes participating in a resource group that have the volume group defined, a volume group consistency check that PowerHA® SystemMirror® runs during the verification process.

This check ensures the following:

- The **concurrent** attribute setting for the volume group is consistent across all related cluster nodes
- The list of PVIDs for this volume group is identical on all related cluster nodes
- An automatic corrective action of the cluster verification utility updates volume group definitions on all related cluster nodes
- Any problems detected are reported as errors.

Managing the cluster topology

These topics describe how to reconfigure the cluster topology.

Reconfiguring a cluster dynamically

When you configure a PowerHA® SystemMirror® cluster, configuration data is stored in PowerHA® SystemMirror®-specific object classes in the Configuration Database (ODM). The AIX® ODM object classes are stored in the default system configuration directory (DCD), **/etc/es/objrepos**.

You can make certain changes to both the cluster topology and to the cluster resources while the cluster is running. This is called a dynamic reconfiguration (DARE). You can make a combination of resource and topology changes via one dynamic reconfiguration operation.

If you have dependent resource groups in the cluster, see Reconfiguring resources in clusters with dependent resource groups for information on making dynamic reconfiguration changes to the cluster topology.

At cluster startup, PowerHA® SystemMirror® copies PowerHA® SystemMirror®-specific ODM classes into a separate directory called the Active Configuration Directory (ACD). While a cluster is running, the PowerHA®

SystemMirror® daemons, scripts, and utilities reference the Configuration Database data stored in the active configuration directory (ACD) in the PowerHA® SystemMirror® Configuration Database.

If you synchronize the cluster topology or cluster resources definition while the Cluster Manager is running on the local node, this action triggers a dynamic reconfiguration event. In a dynamic reconfiguration event, the PowerHA® SystemMirror® Configuration Database data in the Default Configuration Directories (DCDs) on all cluster nodes is updated and the PowerHA® SystemMirror® Configuration Database data in the ACD is overwritten with the new configuration data. The PowerHA® SystemMirror® daemons are refreshed so that the new configuration becomes the currently active configuration.

The dynamic reconfiguration operation (that changes both resources and topology) progresses in the following order:

- Releases any resources affected by the reconfiguration
- Reconfigures the topology
- Acquires and reacquires any resources affected by the reconfiguration operation.

Before making changes to a cluster definition, ensure that:

- All nodes are up and running the AIX® operating system and able to communicate with each other.
- Any change you make to a cluster definition must be done on an active node.
- The cluster is stable; no recent event errors or config_too_long messages exist.

Synchronizing configuration changes

When you change the cluster configuration, you update the data stored in the PowerHA® SystemMirror® Configuration Database in the DCD. For example, when you add an additional network interface to a cluster node, you must add the interface to the cluster definition so that the cluster nodes can recognize and use it.

When you change the cluster definition on one cluster node, you must also update the PowerHA® SystemMirror® Configuration Databases 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 PowerHA® SystemMirror® Configuration Database object classes in the DCD on the other cluster nodes.

When synchronizing the cluster triggers a dynamic reconfiguration event, PowerHA® SystemMirror® 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.

Dynamic cluster topology changes

DARE (Dynamic Reconfiguration) supports resource and topology changes done in one operation.

You can make the following changes to the cluster topology in an active cluster, dynamically:

- Adding or removing nodes
- Adding or removing network interfaces
- Adding or removing a PowerHA® SystemMirror® network
- Swapping a network interface card

To avoid unnecessary processing of resources, use **clRGmove** to move resource groups that will be affected by the change before you make the change. When dynamically reconfiguring a cluster, PowerHA® SystemMirror® will release resource groups if this is found to be necessary, and they will be reacquired later. For example, PowerHA® SystemMirror® will release and reacquire the resource group that is using the associated service IP address on a network interface affected by the change to topology.

Viewing the cluster topology

When you view the cluster topology, you are viewing the PowerHA® SystemMirror® Configuration Database data stored in the DCD, not the data stored in the ACD.

Before making changes to a cluster topology, view the current configuration.

To view the cluster topology:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks > Manage the Cluster > PowerHA® SystemMirror® Configuration** and press Enter.
   This will display the current cluster configuration.

Result
You can also use the **/usr/es/sbin/cluster/utilities/cltopinfo** command to view the cluster topology configuration. The command shows all the topology information and you can choose to see it organized by node, network, or interface.

Managing communication interfaces in PowerHA® SystemMirror®

This section describes the options under the **System Management (C-SPOC) › Communication Interfaces** SMIT menu.

Configuring network interfaces to the operating system on a node

You can configure network interfaces to AIX® without leaving PowerHA® SystemMirror® SMIT, by using the **System Management (C-SPOC) › Communication Interfaces** SMIT path.

To configure communication interfaces/devices to the operating system on a node:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) › Communication Interfaces › Configure Communication Interfaces/Devices to the Operating System on a Node** and press Enter.
   A picklist with node names appears.
3. Select a node on which to configure a network interface or device from the picklist.
4. Select a communication interface or a device type and press Enter:

Interface fields	
Fieldo	**Description**
Network Interfaces	This option leads to the AIX® configuration SMIT menus for a particular node. Each network interface must be defined to the operating system before it can be used by PowerHA® SystemMirror®. It is equivalent to running `smitty mktcpip`.
Physical Disk Devices	This option leads to the AIX® configuration SMIT menus for a particular node. Each physical disk device must be defined to the operating system before it can be used by PowerHA® SystemMirror®.

5. To finish configuring the network interface on a node, fill in the fields in the corresponding AIX® SMIT panel that will open.

Updating PowerHA® SystemMirror® network interfaces with AIX® settings

When you define network interfaces by entering or selecting a PowerHA® SystemMirror® IP label or device, PowerHA® SystemMirror® discovers the associated AIX® network interface name. PowerHA® SystemMirror® expects this relationship to remain unchanged. If you change the IP Label/Address associated with the AIX® network interface after configuring and synchronizing the cluster, PowerHA® SystemMirror® will *not* function correctly.

If this problem occurs, you can reset the network interface IP Label/Address with the AIX settings using the below procedure.

Use the following steps to update PowerHA® SystemMirror® after you make any changes to the underlying AIX® configuration of the mapping of a network interface to an IP label or address. For example, you should update PowerHA® SystemMirror® after modifying the **nameserver** or **/etc/hosts.** options.

You can make these changes dynamically. The following steps describe how to update the IP address from one network interface to another.

To update PowerHA® SystemMirror® with new AIX® settings:

1. To transfer a boot and service IPs, set the **Action on Resource Failure** to **notify** using the `Customizing Resource Recovery`option (`smitty _cm_change_show_resource_action_select`) for the relevant resource.

   > This setting is applied temporarily to avoid a failover during the boot IP transfer. After completing the procedure, you can revert the setting back to the **fallover** state..

2. Run the cluster verification and synchronization.

   > You can use the en1 and en2 as examples to transfer IP address from en1 to en2.

3. Run the following command to access the interface where you want to update the IP address.

   ```
   chdev -l Interface -a state=up
    Ex: chdev -l en2 -a state=up
   ```

4. Run the following command to configure a new network interface and transfer configuration from the previous interface.

   ```
   ifconfig <source_interface> <ip_address> netmask <netmask> transfer <target_interface>
    Ex: ifconfig en1 11.xx.xx.xx netmask 255.255.255.0 transfer en2
   ```

   > Repeat the step for all the affected IP addresses. Transfer the IP addresses in the same order as they are configured on the source interface.

5. Run the following command to remove the IP address and netmask from the older network interface.

   ```
   chdev -l interface -a netaddr='' -a netmask='' -a state=detach
    Ex: chdev -l en1 -a netaddr='' -a netmask='' -a state=detachs
   ```

6. Run the following command to assign IP address to a new network interface.

   ```
   chdev -l en2 -a netaddr=<IP_ADDRESS> -a netmask=<NETMASK> -P
   Ex: chdev -l en2 -a netaddr=11.xx.xx.xx -a netmask=255.255.255.0 -P
   ```

7. Run the cluster verification and synchronization.
8. Set the **Action on Resource Failure** option to **fallover** by using the `smitty _cm_change_show_resource_action_select` flag.
9. Run the cluster verification and synchronization.

   > You can perform these steps on a cluster running on the RSCT version 3.2.6.6. The steps are applicable up to the RSCT version 3.3.2.X. For the RSCT version 3.3.3, you require a fix for `APAR IJ55422 BOOT IP CHANGES IN POWERHA UNDETECTED BY CTHAGS` cluster.

Swapping IP addresses between network interfaces dynamically

As a systems administrator, you may at some point experience a problem with a network interface card on one of the PowerHA® SystemMirror® cluster nodes. If this occurs, you can use the dynamic communications interface swap feature to swap the IP address of an active service network interface with the IP address of

another active, available network interface on the same node and network. Cluster services do not have to be stopped to perform the swap.

You can use this feature to move an IP address off of a NIC that is behaving erratically without shutting down the node. It can also be used if a hot pluggable communication device is being replaced on the node. Hot pluggable NICs can be physically removed and replaced without powering off the node.

This feature can also be used to move the persistent IP label to another network interface.

Make sure that no other PowerHA® SystemMirror® events are running before swapping a network interface.

To dynamically swap an IP address between communication interfaces:

1. Enter `smit cspoc`
2. In SMIT, select **Communication Interfaces › Swap IP Addresses Between Communication Interfaces** and press Enter.
   SMIT displays a list of available service interfaces. It also displays those interfaces that have persistent labels placed on them, but are *not* hosting service IP labels. This allows you to move the persistent label to another interface.
3. Select the service communication interface you want to remove from cluster use, and press Enter.
   SMIT displays a list of available boot interfaces.
4. Select a boot interface and press Enter.
   The **Swap IP Addresses Between Communication Interfaces** menu appears.
5. Verify the service IP label, and the boot IP label you have chosen. If this is correct, press Enter.
   SMIT prompts you to confirm that you want to do this operation.
6. Press Enter only if you are sure you want to swap the communication interface.

Result
After the swapping of IP addresses between communication interfaces, the service address becomes an available boot interface. At this point, you can take action to repair the faulty network interface card. If you have a hot pluggable network interface card, you can replace it while the node and cluster services are up. Otherwise, you will have to stop cluster services and power off the node to replace it.

If you have a hot pluggable network interface card, PowerHA® SystemMirror® makes the interface unavailable when you pull it from the node. When the new card is placed in the node, the network interface is incorporated into the cluster as an available boot IP label again. You can then use the dynamic network interface swap feature again to swap the IP address back to the original network interface.

If you need to power off the node to replace the faulty network interface card, PowerHA® SystemMirror® will configure the service and boot addresses on their original communication interfaces when cluster services are restarted. You do not need to use the dynamic network interface swap feature again to swap the interfaces. PowerHA® SystemMirror® does not record the swapped interface information in the AIX® Configuration Database (ODM). Therefore, the changes are not persistent across system reboots or cluster restarts.

Note the following restrictions:

- The dynamic IP address swap can only be performed within a single node. To move an IP address to another node, move its resource group using the **clRGmove** Resource Group Management utility.

Replacing a PCI hot-pluggable network interface card

This topic takes you through the process of replacing a PCI hot plug network interface card.

Keep the following in mind before you replace a hot-pluggable PCI network interface card:

- Be aware of the following consideration: If a network interface you are hot-replacing is the only available keepalive path on the node where it resides, *you must shut down PowerHA® SystemMirror® on this node in order to prevent a partitioned cluster* **while the interface is being replaced**.
- SMIT gives you the option of stopping cluster services on this node with resource groups brought offline. From this point, you can manually hot-replace the network interface card.
- Hot-replacement of Ethernet network interface cards is supported.
- You should manually record the IP address settings of the network interface being replaced to prepare for unplanned failures.
- You should *not* attempt to change any configuration settings while the hot replacement is in progress.

- To avoid a network failure when using multiple dual-port Ethernet adapter cards on the same node for a particular network, you must configure the interfaces on different physical dual-port Ethernet adapter cards.

Note: Hot-replacement of the dual-port Ethernet adapter used to configure two interfaces for one PowerHA® SystemMirror® IP network is currently *not* supported

Hot-replacing a PCI network interface card

The SMIT interface simplifies the process of replacing a hot-pluggable PCI network interface card. PowerHA® SystemMirror® supports only one PCI hot plug network interface card replacement via SMIT at one time per node.

Note: If the network interface was alive before the replacement process began, then between the initiation and completion of the hot-replacement, the interface being replaced is in a maintenance mode. During this time, network connectivity monitoring is suspended on the interface for the duration of the replacement process.

Scenario 1 (Live NICs only)

This scenario discusses hot-replacing a live PCI network service or boot interface.

Complete the listed steps in the procedure when hot-replacing the following:

- A live PCI network service interface in a resource group and with an available boot interface
- A live PCI network service interface not in a resource group and with an available boot interface
- A live PCI network boot interface with an available boot interface.

1. On the node for which you want to replace a hot-pluggable PCI network interface card, enter `smit sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › Communication Interfaces › PCI Hot Plug Replace a Network Interface Card** and press Enter.
   SMIT displays a list of available PCI network interfaces that are hot-pluggable.
3. Select the network interface you want to hot-replace. Press Enter. The service address of the PCI interface is moved to the available boot interface.
4. SMIT prompts you to physically replace the network interface card. After you have replaced the card, you are asked to confirm that replacement has occurred.
   If you select **yes**, the service address will be moved back to the network interface that has been hot-replaced. On aliased networks, the service address will not move back to the original network interface, but will remain as an alias on the same network interface. The hot-replacement is complete.

   If you select **no**, you must manually reconfigure the interface settings to their original values:

   a. Run the **drslot** command to take the PCI slot out of the removed state.
   b. Run the **mkdev** command on the physical interface.
   c. Use the **ifconfig** command manually as opposed to the `smit chinet` command, the **cfgmgr** command, or the **mkdev** command in order to avoid configuring duplicate IP addresses or an unwanted boot address.

Scenario 2 (Live NICs only)

This scenario discusses hot-replacing a live PCI network service interface on a resource group but with no available boot interface.

Complete the following procedure:

1. On the node for which you want to replace a hot-pluggable PCI network interface card, enter `smit sysmirror`.

2. Select **System Management (C-SPOC) › Communication Interfaces › PCI Hot Plug Replace a Network Interface Card** and press Enter.
   SMIT displays a list of available PCI network interfaces that are hot-pluggable.
3. Select the network interface you wish to hot-replace and press Enter.
   SMIT prompts you to choose whether to move the resource group to another node during the replacement process in order to ensure its availability.
4. If you choose to do this, SMIT gives you the option of moving the resource group back to the node on which the hot-replacement took place after completing the replacement process.
   If you do *not* move the resource group to another node, it will be offline for the duration of the replacement process.
5. SMIT prompts you to physically replace the card. After you have replaced the network interface card, you are asked to confirm that replacement has occurred.
   If you select **Yes**, the hot-replacement is complete.

   If you select **No**, you must manually reconfigure the interface settings to their original values:

   a. Run the **drslot** command to take the PCI slot out of the removed state.
   b. Run the **mkdev** command on the physical interface.
   c. Use the **ifconfig** command manually as opposed to the `smit chinet` command, the **cfgmgr** command, or the **mkdev** command in order to avoid configuring duplicate IP addresses or an unwanted boot address.
   d. Move the resource group back to the node from which you moved it in step 5.

Scenario 3 (Inactive NICs only)

This scenario discusses hot-replacing non-alive PCI network service and boot interface.

Follow the procedure here when hot-replacing the following:

- A inactive PCI network service interface in a resource group and with an available boot interface
- A inactive PCI network service interface not in a resource group and with an available boot interface
- A inactive PCI network boot interface with an available boot interface.

1. On the node for which you want to replace a hot-pluggable PCI network interface card, enter `smit sysmirror`.
2. Select **System Management (C-SPOC) › Communication Interfaces › PCI Hot Plug Replace a Network Interface Card** and press Enter.
   SMIT displays a list of available PCI network interfaces that are hot-pluggable.
3. Select the network interface you want to hot-replace. Press Enter.
   SMIT prompts you to physically replace the network interface card.
4. After you have replaced it, SMIT prompts you to confirm that replacement has occurred.
   If you select **yes**, the hot-replacement is complete.

   If you select **no**, you must manually reconfigure the interface settings to their original values:

   a. Run the **drslot** command to take the PCI slot out of the removed state.
   b. Run **mkdev** command on the physical interface.
   c. Use **ifconfig** command manually as opposed to the smit chinet command, the **cfgmgr** command, or the **mkdev** command in order to avoid configuring duplicate IP addresses or an unwanted boot address.

Service interface failure during hot-replacement

While an interface is unavailable during its replacement, PowerHA® SystemMirror® continues processing events that occur during this time.

Consider, for example, where a node in a cluster has a service interface (Interface A) and an available boot interface (Interface B) on the same network. If you want to hot-replace Interface A, the service network address will first be swapped to Interface B.

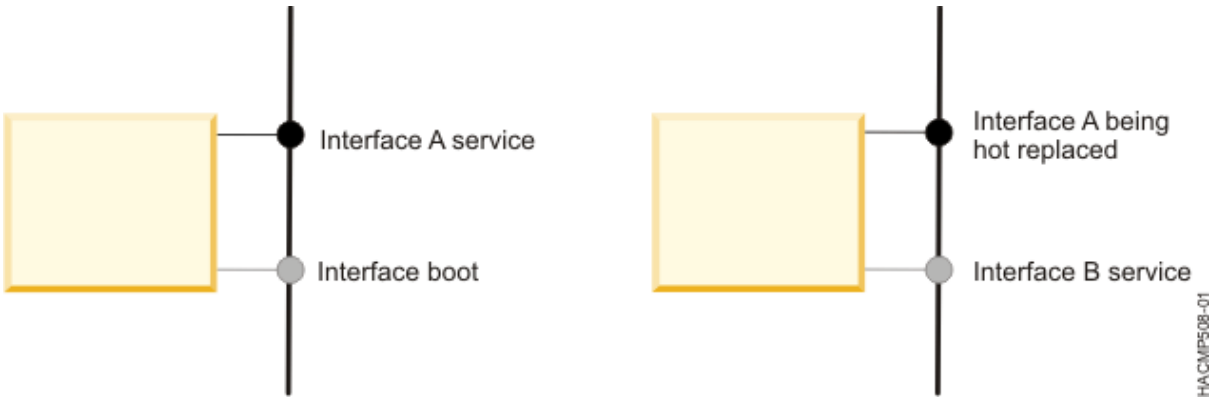


Figure 6: Behavior of interface B while interface a being hot-replaced

Now consider that Interface B (now the service interface) fails while the hot-replace of Interface A is in progress. If there is another available boot interface (C), PowerHA® SystemMirror® does a swap of Interface B to Interface C. When the hot-replacement is finished, the service network settings are swapped from Interface C back to Interface A (the newly replaced interface), and Interface C is reconfigured to boot settings.

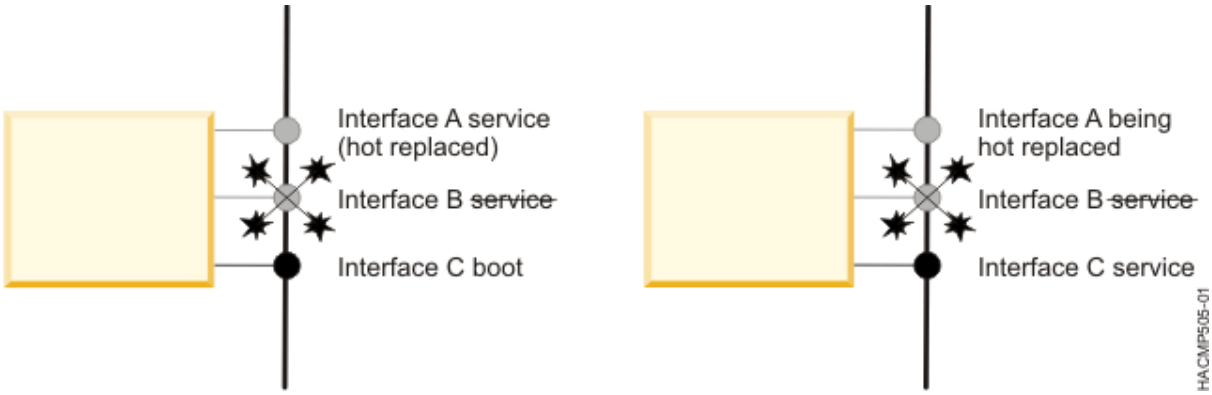


Figure 7: Behavior of interface C while interface a being hot-replaced and interface B fails

If there are no extra available boot interfaces, then between the time Interface B (the service interface) fails and the replacement of Interface A is complete, the node has no network connectivity on that network. In this case, if there are no other sufficient network paths alive on the node for keepalive traffic, a partitioned cluster results. If there are sufficient other network paths alive for keepalive traffic, then a local network failure event is generated for the network to which Interfaces A and B belong.

Any resource group dependent on a service interface in that same network moves to another node, thus the service address moves with the resource group. Following hot plug replacement, Interface A (the newly replaced interface) is reconfigured to a boot address not currently used on that node and network.

Recovering from PCI hot plug network interface card failure

If an unrecoverable error causes the hot-replacement process to fail, PowerHA® SystemMirror® may be left in a state where your network interface is unconfigured and still in maintenance mode.

To recover from this, manually fix the script, then run *smit clruncmd* to remove any maintenance modes that are still set. You can also use **ifconfig** to reconfigure the network settings of the interface.

Adding PowerHA® SystemMirror® site definitions

You can define sites in PowerHA® SystemMirror® Standard Edition or PowerHA® SystemMirror® Enterprise Edition. You must define sites to enable PowerHA® SystemMirror® Enterprise Edition storage replication support, including Geographic Logical Volume Manager (GLVM) and Metro Mirror.

When you associate nodes and storage devices with a site, you can use PowerHA® SystemMirror® to help implement a split-site LVM mirroring configuration. PowerHA® SystemMirror® identifies appropriate selections that are based on site information and verifies the consistency of the mirroring configuration at a site level.

Before you add a site definition, you must decide whether you want to use a stretched cluster or a linked cluster.
A *linked cluster* contains nodes from sites that are located at different geographical locations. Linked clusters do not require sites to share a repository disk or support multicast communication.

A *stretched cluster* contains nodes from sites that are located at the same geographical locations. Stretched clusters share at least one repository disk between sites.

If you define sites to be used in some other way, appropriate methods or customization must be provided to handle site operations. If sites are defined, site events run during **node_up** and **node_down** events.

If you are configuring sites, two sites must be configured and all nodes must belong to one of the two sites.

To add a site definition to a PowerHA® SystemMirror® cluster, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Nodes and Network › Manage Sites › Add a Site**, and press Enter.
3. Complete the following fields.
   Adding site fields.

Adding site fields	
Field	**Description**
Site Name	Enter a name for this site. The site name can be up to 64 alphanumeric characters.
Site Nodes	Press F4 (List) to select the available nodes that you want to be part of the site.
Cluster Type	Press F4 (List) to select stretched cluster or linked cluster.

4. Press enter to add the site definition to PowerHA® SystemMirror®.
5. Repeat the steps 1 - 4 to add definitions for the second site.

Changing the host name for a cluster node in PowerHA® SystemMirror® 7.1.2, or earlier

You cannot change the host name of a cluster node after the cluster is configured. To change the host name of a cluster node, you must first remove the Cluster Aware AIX® (CAA) cluster definition, update PowerHA® SystemMirror® and the AIX® operating system configurations, and then synchronize the changes to re-create the CAA cluster with the new host name.

To change the host name for a cluster node in PowerHA® SystemMirror® 7.1.2, or earlier, complete the following steps:

1. Stop cluster services by using the Bring resource group offline option.
2. To remove the CAA cluster, run the `rmcluster -f -n` *`clustername`* command, where *`clustername`* is the name of the CAA cluster, on a node in the cluster.

   > You can run the `lscluster -i` command to display the name of the CAA cluster.

3. To change the host name, complete the following steps:
   a. From the command line, run `smit hostname` on the cluster node where you want to change the host name. In the SMIT interface, select **Set Hostname** and enter the new host name.
   b. To change the host name for the *`COMMUNCATION_PATH`* variable, complete the following steps:
      i. Enter the following command:

      ```
      odmget -q "object = COMMUNICATION_PATH " HACMPnode > tmp1
      ```

      ii. Edit the `tmp1` file and change the `value` for the corresponding node to the new *`COMMUNCATION_PATH`* name.
      iii. Enter the following command to update the ODM:

      ```
      odmchange -o HACMPnode -q "object = COMMUNICATION_PATH" tmp1
      ```

c. Change the `/etc/hosts` file for each node in the cluster with the new host name. If your environment is using a domain name system (DNS), you must update the DNS with the new host name.

d. Change the `/etc/cluster/rhosts` file and run the `refresh -s clcomd` command on all cluster nodes.

e. To change any of PowerHA® SystemMirror® configuration settings, such as the node name, run `smit sysmirror` from the command line.

4. Verify and synchronize the cluster. This process creates the CAA cluster configuration with the updated host name.
5. Start cluster services.

Changing the host name for a cluster node in PowerHA® SystemMirror®

You can use PowerHA® SystemMirror® to change the host name for a node while cluster services are active.

Cluster Aware AIX® (CAA) uses the host name as a point of reference for cluster communication. PowerHA® SystemMirror® stores the IP address that is associated with a host name in a configuration database. Therefore, you must consider the effect of each type of host name change (permenant or temporary) and whether the PowerHA® SystemMirror® configuration requires an update to reflect a host name change.

With AIX® interfaces, you can change a host name temporarily or permanently. If you change a host name by using either method, the host name must resolve to a TCP/IP address.

If you change the host name temporarily, the host name is updated in the AIX® kernel but not in the AIX® Object Data Manager (ODM). If you use the temporary host name change method, the host name is not saved when the system reboots.

If you change the host name permanently, the host name is updated in the AIX® kernel and in the AIX® Object Data Manager (ODM). If you use the permanent host name change method, the new host name is used after the system reboots.

Review the following information before you change the host name for a cluster node in PowerHA® SystemMirror®:

- Do not change the host name of the cluster nodes during initial cluster configuration.
- Do not change the host name of the cluster nodes when you are migrating from a prior version of PowerHA® SystemMirror®.
- Do not change the host name on multiple nodes in the cluster at the same time. If you change the host name for multiple nodes, change it on one node at a time and synchronize the cluster configuration after each change.

You can use one of the following methods to change a host name for a cluster node in PowerHA® SystemMirror®:

Temporary host name change

You can use the **hostname** command to change the host name for the node. By default, PowerHA® SystemMirror® and CAA ignore this type of change and the PowerHA® SystemMirror® configuration database is not updated.

Permanent host name change

From the command line, enter **smitty hostname** and use the SMIT interface to change the host name. When you change the host name with this method, PowerHA® SystemMirror® responds by updating the configuration database with new host name information. After you complete the changes, you must verify and synchronize the cluster.

Changing how PowerHA® SystemMirror® 7.1.3, or later, responds to host name change

Cluster Aware AIX® (CAA) uses the host name of a cluster node as a point of contact for cluster communications. PowerHA® SystemMirror® stores the host name information for each node in the configuration database. If the host name of a cluster node changes, you must update the host name information in the PowerHA® SystemMirror® configuration database.

If you change a host name with the **hostname** command (temporary host name change method), PowerHA® SystemMirror® and CAA, by default, ignore the host name change and the PowerHA® SystemMirror® configuration database is not updated.

If you change the host name with the SMIT interface (permanent host name change method), PowerHA® SystemMirror® updates the PowerHA® SystemMirror® configuration database. Whenever you change the host name with the permanent host name change method, you must verify and synchronize the cluster. If you do not verify and synchronize the cluster, your changes are not applied to the cluster.

Any change to a host name that uses the temporary host name change method are not saved after a system reboots. You can change the host name to the service label as part of the workload startup script and then change the host name back to the actual host name of the LPAR while you stop the workload. This configuration ensures that there are not two LPARs with the same host name at any point in time.

Changing the IP address for a PowerHA® SystemMirror® cluster node

When you change the boot IP address, a cluster outage might occur. Therefore, you need to plan your downtime.

To change the boot IP address of an LPAR, you can use the **clmgr** command, which is located in `/usr/es/sbin/cluster/utilities/clmgr` directory. For information about the options for the **clmgr** command, run the following command:

```
clmgr -h
```

1. To stop cluster services, run the following command:

   ```
   clmgr stop cluster STOP_CAA=yes
   ```

   This command stops the PowerHA® cluster services along with Cluster Aware AIX® (CAA) repository disk. All applications that are running on the cluster stops.
2. Change the boot IP of the LPAR by using SMITTY **tcpip** command.
3. Clean the RSCT peer domains on both nodes by running the following command:

   ```
   /usr/sbin/rsct/install/bin/recfgct -s
   ```

   > The **recfgct** command might prompt you to remove the node from the RSCT peer domain `rpdomain`.

   You can delete the RSCT peer domain **rpdomain** by using following command and then run the **recfgct** command.

   ```
   rmrpdomain -f <rpdomain_name>
   ```

4. To remove the CAA cluster, run the following command:

   ```
   CAA_FORCE_ENABLED=yes /usr/sbin/rmcluster -fr hdiskX
   ```

   where, **hdiskX** is the CAA repository disk. If you do not know the location of the CAA repository disk, run the following command:

   ```
   clmgr query repository
   ```

5. Replace the entries of the old boot IP addresses in the `/etc/hosts` and `/etc/cluster/rhosts` files on all the cluster nodes.
6. Refresh the **clcomd** command by using the following command:

   ```
   refresh -s clcomd
   ```

   This command allows the **clcomd** command to use the new address.

7. On one node in the cluster, complete the following steps:
    a. Remove any one of the available network interfaces on the PowerHA® network. For instructions, see Removing a network interface from a cluster node.
    b. Run discover network interface and disks. For instructions, see Discovering IP network information.
    c. Add a newly discovered interface on the PowerHA® network. In SMIT, select **Cluster Nodes and Networks › Add Discovered Network Interfaces to the Configuration**.
8. To verify and synchronize the configuration, run the following command:

```
clmgr sync cluster
```

9. When you are ready to start the cluster services, run the following command:

```
clmgr start cluster start_CAA=yes
```

Changing a cluster name

Changing the name of a cluster is allowed only prior to initial synchronization of the cluster configuration in PowerHA® SystemMirror®. After you have synchronized the initial cluster configuration, you cannot change the cluster name. Instead, you must remove the cluster completely and recreate it.

If you have not yet synchronized the initial cluster configuration, you can change the cluster name by completing these steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Custom Cluster Configuration › Cluster Nodes and Networks › Initial Cluster Configuration (Custom) › Cluster › Add/Change/Show a Cluster** and press Enter.
   SMIT displays the cluster definition with the current value for the cluster name.
3. Enter the name change. A cluster name can include alphanumeric characters and underscores; it cannot begin with a numeric. Use no more than 64 characters.

Changing the configuration of cluster nodes

As the system administrator of a PowerHA® SystemMirror® cluster, you may need to perform any of several tasks relating to cluster nodes.

Adding nodes to a resource group

Once you have added the new node to the cluster topology, you can continue by adding the new node (or nodes) to the list of participating nodes in a resource group.

In a non-concurrent resource group with the startup policy of either Online on Home Node Only or Online on First Available Node, 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 start up cluster services on this node. 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.

When you are finished adding the node to a resource group:

1. Synchronize the cluster.
2. When you press Enter, the cluster resources are dynamically reconfigured.

Removing a cluster node from the PowerHA® SystemMirror® configuration

You can remove a node from an active cluster dynamically. However, before removing a node from the cluster, you must remove the node from any resource groups it participates in and synchronize resources.

To remove a cluster node:

1. Stop cluster services on the node to be removed (usually this is done by stopping cluster services with the **Move Resource Groups** option).
2. On another active node, enter `smit sysmirror`
3. In SMIT, select **Cluster Nodes and Networks › Manage Nodes › Remove a Node**. SMIT displays a list of all cluster nodes.
4. Select the node you want to remove and press Enter. SMIT prompts you to confirm that you want to proceed. Press Enter again to remove the node from the cluster.

   > When you remove a node from the cluster topology, all communication path information associated with the node is also removed, its resources are released and reacquired, and the node is removed from the resource configuration.

5. Synchronize the change on the local node. When the synchronization completes, the node is removed from the cluster definition.

Changing the name of a cluster node

When changing the name of a cluster node, 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:

1. From the command line, enter `smit sysmirror`
2. Select **Cluster Nodes and Networks › Manage Nodes › Network Interfaces › Change/Show a Node** and press Enter. SMIT displays a picklist of cluster nodes.
3. Make your selection and press Enter. SMIT displays the current node name.
4. Enter the new name for the node in the New Node Name field. A node name can include alphanumeric characters and underscores, but cannot have a leading numeric. Use no more than 64 characters. When you finish entering data, press Enter. SMIT makes the changes you specified.
5. After the command completes, return to the PowerHA® SystemMirror® SMIT menus to perform further topology reconfiguration or to synchronize the changes you made.

Result
The change is propagated through both the cluster topology and resource configuration.

Changing the configuration of a PowerHA® SystemMirror® network

You can change network attributes, but you cannot change them dynamically.

Changing the network attribute to private for Oracle inter-node communication

Changing the network attribute to **private** makes the network Oracle-compatible by changing all interfaces to service (as well as changing the attribute in PowerHA® SystemMirror®network ODM).

ORACLE uses the **private** network attribute setting to select networks for Oracle inter-node communications. This attribute is not used by PowerHA® SystemMirror® and will not affect PowerHA® SystemMirror® in any way. The default attribute is **public**.

To configure private networks for use by Oracle:

1. Configure the network and add all interfaces. You cannot change the attribute if the network has no interfaces.
2. Change the network attribute to **private**.
3. Private networks must have either all boot or all service interfaces. If the network has all boot interfaces (the default when using discovery) PowerHA® SystemMirror® converts these interfaces to service. (Oracle only looks at service interfaces.)
4. Synchronize the cluster after changing the attribute.

Changing attributes of an IP-based network

You can use SMIT to change the network name or an attribute of an IP-based PowerHA® SystemMirror® network.

To change the name or an attribute of an IP-based network, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Nodes and Networks › Manage Networks and Network Interfaces › Networks › Change/Show a Network** and press Enter.
3. From the list, select the network that you want to change.
4. Enter the changes for the following fields:

Change/Show a Network fields	
Field	**Value**
Network Name	Displays the current network name. You cannot change this field.
New Network Name	Enter the new name for this network. The name cannot start with a number and use more than 64 alphanumeric characters and underscores.
Network Type	Select **XD_data**, **XD_ip**, or **ether** for this network.
Netmask (IPv4) / Prefix Length (IPv6)	Enter the netmask for an IP version 4 network or the prefix length for an IP version 6 network.
Network attribute	Select the default value of **public** to allow all types of cluster communication on this network. Select **private** to prevent heartbeat communication and cluster communication from using this network.

5. Verify that the changes you made are correct, and press Enter.
6. Verify and synchronize the cluster configuration.

Removing a PowerHA® SystemMirror® network

You can remove a network from a PowerHA® SystemMirror® cluster definition.

> **Note:** Deleting all network interfaces associated with a network deletes the network definition from PowerHA® SystemMirror®.

To remove a network:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Networks and Network Interfaces › Networks › Remove a Network** and press Enter.
   SMIT displays the **Select a Network to Remove** panel.
3. Select the network to remove.
   SMIT displays

   ```
   Are you sure?
   ```

4. Press Enter to remove the network. All of this network's subnets and their interfaces are removed from the PowerHA® SystemMirror® configuration.
5. On the same node, synchronize the cluster configuration.

Result
If the Cluster Services is running on the local node, the synchronization triggers a dynamic reconfiguration event.

Establishing default and static routes on aliased networks

If you require the default route, and possibly, other static routes to be established on the IP aliased service subnet, these routes will fail to be established automatically when the **rc.net** file runs at boot time. This is because there is no address on that subnet in the Configuration Database.

To ensure that these routes are established at boot time, you can configure a persistent address on that subnet. After you configure the persistent address, PowerHA® SystemMirror® configures the routes. If you have a single network adapter that is configured in your environment, you can have boot services and persistent addresses on the same subnet. In this case, boot services can serve the same purposes of the persistent address. Therefore, you do not need to configure the persistent address.

If you do not configure persistent addresses, then you should use your own scripts that configure routes on aliased service subnets.

Controlling distribution preferences for service IP label aliases

To control the placement of the service IP label aliases on the cluster node physical network interface cards, you can configure a distribution preference for the aliases of the service IP labels that are placed under PowerHA® SystemMirror® control.

Changing the configuration of communication interfaces

As a system administrator, you may need to perform several different tasks relating to cluster network interfaces.

Adding PowerHA® SystemMirror® network interfaces

You can add a network interface to an active cluster dynamically. You do not need to stop and restart cluster services for the network interface to become part of the cluster.

1. On the node getting the new network interface card, complete the prerequisite tasks:
   - Install the new network interface card.
   - Configure the new logical network interface to AIX®.
2. On all cluster nodes, update the **/etc/hosts** file to include the IP address of the new network interface.
3. On any cluster node, add the PowerHA® SystemMirror® communication interface to the cluster topology definition.
4. Synchronize the cluster.

Changing network interface attributes

You cannot change the attributes of a network interface dynamically. You must stop and restart cluster services for the changes to take effect in the configuration.

To change a network interface for a cluster, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Networks and Network Interfaces › Network Interfaces › Change/Show a Network Interface**, and press Enter.
3. Select the network interface from the picklist.
4. Enter the field values as follows:

Network Interface fields	
Field	**Value**
Node Name	The name of the node on which this network interface physically exists.

Field	Value
Network Interface	The network interface that is associated with the communication interface.
IP Label/Address	The IP label or IP address that is associated with this communication interface that is configured on the network interface when the node starts up. The picklist filters out IP labels or IP addresses already configured to PowerHA® SystemMirror®.
Network Type	Displays the network type for the selected network interface. You cannot edit this field.
Network Name	Enter a unique name for this network interface.

5. Verify your changes are correct, and press Enter. PowerHA® SystemMirror® now checks the validity of the configuration. If a node cannot be reached, you might receive a warning
6. Synchronize the cluster.
7. Restart cluster services.

Removing a network interface from a cluster node

You can remove a PowerHA® SystemMirror® network interface from an active cluster dynamically; you do not need to stop and restart cluster services.

> **Note:** Deleting all network interfaces associated with a network deletes the network from PowerHA® SystemMirror®.

To remove a network interface from a cluster node:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Networks and Network Interfaces › Network Interfaces › Remove a Network Interface** and press Enter.
3. Select the network interface(s) or the serial device(s) from the picklist and press Enter.
   When you remove a network interface, all information associated with the interface is removed from the Configuration Database. SMIT prompts you to confirm that you want to do this operation. Press Enter again only if you are sure you want to remove the interface and its associated information.
4. On the same node, synchronize the cluster. If the Cluster Manager is running on the local node, the synchronization triggers a dynamic reconfiguration event.
   When the synchronization completes, the selected network interfaces are removed from the cluster topology definition.

Managing persistent node IP labels

This topic describes different tasks for managing persistent node IP labels.

Configuring persistent node IP labels/addresses

This topic describes configuring a persistent node IP label/address on a specified node.

To configure persistent node IP labels on a specified node:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Nodes › Configure Persistent Node IP Labels/ Addresses › Add a Persistent Node IP Label/Address** and press Enter.
3. Select a cluster node.
4. Enter the field values as follows:

Cluster node fields	
Field	**Value**
Node Name	The name of the node on which the IP Label/Address will be bound.
Network Name	The name of the network on which the IP Label/Address will be bound.
Node IP Label/Address	The IP Label/Address to keep bound to the specified node.

5. Press Enter. The resulting SMIT panel displays the current node name and persistent node IP labels defined on IP networks on that node.

Changing persistent node IP labels

This topic describes changing or viewing persistent node IP labels configured on a specified node.

To change or view persistent node IP labels:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Nodes › Configure Persistent Node IP Labels/ Addresses › Change/Show a Persistent Node IP Label/Address** and press Enter
3. Enter field values as follows:

Change/Show a Persistent Node IP Label/Address fields	
Field	**Value**
Node Name	The name of the node on which the IP Label/Address will be bound.
New Node Name	The new node name for binding the IP Label/Address.
Network Name	The name of the network on which the IP Label/Address will be bound.
Node IP Label/Address	The IP Label/Address to keep bound to the specified node.
New Node IP Label/ Address	The new IP Label/Address to be bound to the specified node.

4. Press Enter. The resulting SMIT panel displays the current node name and persistent node IP labels defined on IP networks on that node.

Deleting persistent node IP labels

This topic describes deleting persistent node IP labels configuring on a specified node.

To delete persistent node IP labels:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage Nodes › Configure Persistent Node IP Labels/ Addresses › Remove a Persistent Node IP Label/Address**.
3. Press Enter.
   PowerHA® SystemMirror® deletes the persistent node IP label from the node.

Synchronizing the cluster configuration

Whenever you modify the cluster definition in the Configuration Database on one node, you must synchronize the change with the Configuration Database data on all cluster nodes.

You perform a synchronization by choosing the **Verification and Synchronize Cluster Configuration** option from either the standard configuration paths, the **Custom Cluster Configuration** path, or from the **Problem Determination Tools** menu.

Dynamic reconfiguration issues and synchronization

These topics are relevant for dynamic reconfiguration of both topology and resources.

Releasing a dynamic reconfiguration lock

During a dynamic reconfiguration, PowerHA® SystemMirror® creates a temporary copy of the PowerHA® SystemMirror®-specific Configuration Database 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 configuration 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:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools** and press Enter.
3. Select the **Release Locks Set By Dynamic Reconfiguration** option and press Enter. SMIT displays a panel asking if you want to proceed. If you want to remove the SCD, press Enter.

Processing configuration database data during dynamic reconfiguration

When you synchronize the cluster topology, the processing performed by PowerHA® SystemMirror® varies depending on the status of the Cluster Services.

The following describe the variations that may occur:

Cluster Services is not running on any cluster node

If the Cluster Services is *not* running on any cluster node (typically the case when a cluster is first configured), synchronizing the topology causes the configuration data stored on each node reachable from the local node to be updated.

Cluster Services is running on the local node

If the Cluster Services is running on the local node, synchronizing the topology triggers a dynamic reconfiguration event. While processing this event, PowerHA® SystemMirror® updates the configuration data stored on each cluster node that is reachable. Further processing makes the new configuration the currently active configuration.

Cluster Services is running on some cluster nodes but not on the local node

If the Cluster Services is running on some cluster nodes but *not* on the local node, synchronizing the topology causes the configuration data stored on each node that is reachable from the local node to be updated. However, the processing performed during a dynamic reconfiguration to make the new configuration the active configuration is *not* performed.

Undoing a dynamic reconfiguration

Before PowerHA® SystemMirror® 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. If you want to undo the dynamic reconfiguration, you can use this cluster snapshot to restore the previous configuration.

PowerHA® SystemMirror® saves snapshots of the last ten configurations in the default cluster snapshot directory, **/usr/es/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 configuration database 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. PowerHA® SystemMirror® allows you to save in a snapshot the changes you made to the configuration in the DCD before you overwrite it.

To replace the Configuration Database data stored in the DCD with the Configuration Database data in the ACD:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Problem Determination Tools** and press Enter.
3. Select **Restore PowerHA® SystemMirror® Configuration Database from Active Configuration** and press Enter.
4. Enter field values as follows:

Restore PowerHA® SystemMirror® Configuration Database from Active Configuration fields	
Field	**Value**
Cluster Snapshot Name of System Default PowerHA® SystemMirror® ODMs	In this field, specify the name you want assigned to the cluster snapshot PowerHA® SystemMirror® 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 PowerHA® SystemMirror® ODMs	Enter any text string you want stored at the beginning of the snapshot.

5. Press Enter. SMIT displays the results.

Managing the cluster resources

Use these topics to manage the resources in your cluster. The first part describes the dynamic reconfiguration process. The second part describes procedures for making changes to individual cluster resources.

When you configure a PowerHA® SystemMirror® cluster, configuration data is stored in specific PowerHA® SystemMirror® object classes in the ODM. The AIX® ODM object classes are stored in the default system configuration directory (DCD), **/etc/objrepos**.

You can make certain changes to both the cluster topology and to the cluster resources while the cluster is running (dynamic reconfiguration, or DARE). You can make a combination of resource and topology changes via one dynamic reconfiguration operation making the whole operation faster, especially for complex configuration changes.

> **Note:** No automatic corrective actions take place during a DARE.

Reconfiguring a cluster dynamically

At cluster startup, PowerHA® SystemMirror® copies PowerHA® SystemMirror®-specific ODM classes into a separate directory called the Active Configuration Directory (ACD). While a cluster is running, the PowerHA® SystemMirror® daemons, scripts, and utilities reference the ODM data stored in the active configuration directory (ACD) in the ODM.

> **Important:** Do not make configuration changes or perform any action that affects a resource if any node in the cluster has cluster services stopped and its resource groups placed in an UNMANAGED state.

If you synchronize the cluster topology and cluster resources definition while the Cluster Manager is running on the local node, this action triggers a dynamic reconfiguration (DARE) event. In a dynamic reconfiguration event, the ODM data in the Default Configuration Directory (DCD) on all cluster nodes is updated and the ODM data in the ACD is overwritten with the new configuration data. The PowerHA® SystemMirror® daemons are refreshed so that the new configuration becomes the currently active configuration.

The dynamic reconfiguration operation (that changes both resources and topology) progresses in the following order that ensures proper handling of resources:

- Releases any resources affected by the reconfiguration
- Reconfigures the topology

- Acquires and reacquires any resources affected by the reconfiguration operation.

Requirements before reconfiguring

There are requirements you must verify before you change a cluster definition. These requirements help ensure that the verification process can effectively analyze the configuration and that the synchronization process can distribute changes to all nodes in the cluster.

Before you change a cluster definition, verify the following settings:

- The same version of PowerHA® SystemMirror® is installed on all nodes.
- All nodes are online and running the AIX® operating system and are able to communicate with each other using the **clcomd** subsystem.
- Resource groups are not in the UNMANAGED state.
- The cluster is stable; the **hacmp.out** file does not contain recent event errors or **config_too_long** events.

Reconfiguring application controllers

An *application controller* is a cluster resource used to control an application that must be kept highly available. It includes start and stop scripts.

Note that this section does *not* discuss how to write the start and stop scripts. See the vendor documentation for specific product information on starting and stopping a particular application.

If you intend to add an application controller dynamically, it is very important to test the server scripts beforehand, as they will take effect during the dynamic reconfiguration operation.

Changing an application controller

When you specify new start or stop scripts to be associated with an application controller, the PowerHA® SystemMirror® configuration database is updated but the application controller is not configured or unconfigured dynamically; thus the application controlled by the application controller is not stopped and restarted. The next time the application is stopped, PowerHA® SystemMirror® calls the new stop script - not the stop script that was defined when the application was originally started.

> **Note:** Changes to application controller information are not automatically communicated to the application monitor configuration. Only the name of the application controller is updated on the SMIT panel for changing monitors. If you change an application controller that has an application monitor defined, you must make the change separately to the application monitor as well.

To change an application controller, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Application Controller Scripts** and press Enter.
3. From this menu, select the **Change/Show an Application Controller** option and press Enter. SMIT displays the application controllers.
4. Select the application controller to change and press Enter. The **Change/Show an Application Controller** panel appears, with the application controller name filled in.
5. You can change the application name and/or the start and stop scripts.
6. Press Enter to add this information to the PowerHA® SystemMirror® configuration database on the local node.
7. (*Optional*) Return to previous SMIT panels to perform other configuration tasks.
8. Verify and synchronize the cluster configuration after you make any changes.

Removing an application controller

You can remove an application controller from an active cluster dynamically. Before removing an application controller, you must remove it from any resource group where it has been included as a resource.

> **Note:** Note: If you remove an application controller, PowerHA® SystemMirror® checks all application monitors for that server, and if only this controller (and no other controllers) use the associated monitors, it also removes the monitors. PowerHA® SystemMirror® sends a message if monitors have been removed or preserved as a result of removing an application controller.

To remove an application controller, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Application Controller Scripts › Remove an Application Controller** and press Enter. SMIT displays the list of application controllers.
3. Select the application controller to remove and press Enter. PowerHA® SystemMirror® asks if you are sure you want to remove it.
4. Press Enter again to confirm the removal. The application controller is removed from the PowerHA® SystemMirror® configuration database on the local node.
5. (*Optional*) Return to previous SMIT panels to perform other configuration tasks.
6. Synchronize the cluster definition. If the Cluster Manager is running on the local node, synchronizing the cluster resources triggers a dynamic reconfiguration event.

Changing or removing application monitors

If you have configured application monitoring, you may want to suspend or remove the monitor at some point. You can also change some aspect of the monitoring you have set up (for instance, the processes to be monitored, the scripts to run, or the notify, cleanup, or restart methods).

This section discusses changing an existing application monitor. For information about adding a new application monitor, see the section Configuring PowerHA® SystemMirror® cluster topology and resources.

Suspending and resuming application monitoring

You can suspend and resume the monitoring of an application while that application is online. When you suspend the monitoring of an application, the suspension is temporary and the monitoring can be resumed automatically under certain conditions. For example, if a cluster events results in moving the affected resource group to a different node, or the resource group is manually moved to another node, the application monitoring is resumed automatically on the new node. Similarly, if a node has a resource group that is brought offline and if the node is rebooted, the application monitoring is resumed automatically on the node after the cluster services are started.

> **Note:** If you have multiple monitors configured for one application, and if a monitor with **notify** action is launched first, PowerHA® SystemMirror® runs the notification methods for that monitor, and the remaining monitor(s) are shut down on that node. PowerHA® SystemMirror® takes no actions specified in any other monitor. You can restart the **fallover** monitor using the **Suspend/Resume Application Monitoring** SMIT panel.

To permanently stop monitoring of an application, see the section Removing an application monitor.

To temporarily suspend application monitoring:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Management (C-SPOC) › Resource Groups and Applications › Suspend/Resume Application Monitoring › Suspend Application Monitoring** and press Enter.

Result
You are prompted to select the application controller for which this monitor is configured. If you have multiple application monitors, they are all suspended until you choose to resume them or until a cluster event occurs to resume them automatically, as explained earlier.

To resume monitoring after suspending it:

1. From the command line, enter `smit sysmirror`.
2. In SMIT, select **Cluster Management (C-SPOC) › Resource Groups and Applications › Suspend/Resume Application Monitoring › Resume Application Monitoring** and press Enter.
   PowerHA® SystemMirror® prompts you to select the application controller that is associated with the suspended application monitor you want to resume.
3. Select the controller. All monitors resume, configured as they were prior to suspending them.

> **Note:** Do not make changes to the application monitor(s) configurations while they are in a suspended state.

Changing the configuration of an application monitor

You can change the configuration details of an application monitor by editing the SMIT fields you defined when you configured the monitor initially.

> **Note:** When you configured application monitors originally, the Restart Method and Cleanup Method fields had default values. If you changed those fields, and now want to change back to the defaults, you must enter the information manually (by copying the scripts from the **Change/Show an Application Controller** SMIT panel).

To alter a defined application monitor:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Application Monitors** and press Enter.
3. Depending on which type of monitor you are altering, select either:
   Configure Process Application Monitor > Change/Show Process Application Monitor

   or

   Configure Custom Application Monitor > Change/Show Custom Application Monitor.
4. From the list of monitors, select the previously defined application monitor you want to change.
5. Make changes in the SMIT panel fields and press Enter. Remember that default values are not restored automatically.

Result
Verify and synchronize the cluster configuration after you make any changes.

Removing an application monitor

You can permanently remove an application monitor.

To permanently remove an application monitor:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select**Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Configure Application Monitoring** and press Enter.
3. Depending on which type of monitor you are altering, select either:
   Configure Process Application Monitor > Remove a Process Application Monitor

   or

Configure Custom Application Monitor > Remove a Custom Application Monitor.

4. Select the monitor to remove.
5. Press Enter. The selected monitor is deleted.

Result
If the monitor is currently running, it is *not* stopped until the next dynamic reconfiguration or synchronization occurs.

> **Note:** If you remove an application monitor, PowerHA® SystemMirror® removes it from the definition for all application controllers that were using the monitor, and sends a message about the servers that will no longer use the monitor.

If you remove an application controller, PowerHA® SystemMirror® removes it from the definition of all application monitors that were configured to monitor the application. PowerHA® SystemMirror® also sends a message about which monitor will no longer be used for the application. If you remove the last application controller in use for any particular monitor; that is, if the monitor will no longer be used for any application, **verification** issues a warning that the monitor will no longer be used.

Reconfiguring service IP labels as resources in resource groups

You must stop cluster services to change service IP labels/address resources that are already included in resource group.

Remember to add any new service IP labels/addresses to the **/etc/hosts** file before using them. If you intend to change the names of existing labels, first create the new names and add them to the **etc/hosts** file. Then make the name change in SMIT.

Do *not* remove the previously used service IP label/address from the **/etc/hosts** file until after you have made the change in the cluster configuration. Once you make the change in the configuration and in the **/etc/hosts** file on the local node, make the change in the **/etc/hosts** files of the other nodes before you synchronize and restart the cluster.

Steps for changing the service IP labels/addresses definitions

This topic describes changing the service IP labels/addresses definitions.

To change a service IP label/address definition:

1. Stop cluster services on all nodes.
2. On any cluster node, enter `smit sysmirror`
3. Select **Cluster Applications and Resources › Resources › Configure Service IP Labels/ Addresses › Change/Show a Service IP Label/Address**.
4. In the **IP Label/Address to Change** panel, select the IP Label/Address you want to change. The **Change/ Show a Service IP Label/Address** panel appears.
5. Make changes in the field values as needed.
6. Press Enter after filling in all required fields. PowerHA® SystemMirror® now checks the validity of the new configuration. You may receive warnings if a node cannot be reached, or if network interfaces are found to *not* actually be on the same physical network.
7. On the local node, verify and synchronize the cluster.
8. Restart Cluster Services.

Deleting service IP labels

You can use the SMIT interface to remove IP labels and an IP addresses.

To delete an IP label or an IP address, complete the following steps:

1. From the command line on any node in the cluster, enter `smit sysmirror`.

2. In the SMIT interface, select **Cluster Applications and Resources › Resources › Configure Service IP Labels/Addresses › Remove Service IP Label(s)/Address(es)**.
3. Select one or more labels that you want to delete from the list and press Enter.
4. Verify and synchronize the cluster.
5. For maintenance purposes, delete the labels/addresses from the **/etc/hosts** file.

Result
After you delete service ip labels from the cluster configuration by using smit, remove the deleted service ip labels from the **/etc/hosts** file to avoid conflicting entries in future.

> **Note:** If you delete the service ip labels when the associated resource group is online, an application outage may occur.

Changing AIX® network interface names

You can use SMIT to change or reset the PowerHA® SystemMirror® network interface.

When you define network interfaces by entering or selecting a PowerHA® SystemMirror® IP label/address, PowerHA® SystemMirror® discovers the associated AIX® network interface name. PowerHA® SystemMirror® expects this relationship to remain unchanged. If you change the name of the AIX® network interface name after configuring and synchronizing the cluster, PowerHA® SystemMirror® will not function correctly.

If this problem occurs, you can reset the network interface name from the SMIT PowerHA® SystemMirror® **Cluster System Management (C-SPOC**) menu.

To reset the PowerHA® SystemMirror® communication interface:

1. From the command line, enter `smit cspoc`.
2. In SMIT, select **Communication Interfaces › UpdatePowerHA® SystemMirror® Communication Interface with Operating System Settings** and press Enter.
3. Select the network interface that you want to reset from the list.
4. Press Enter to complete the reset operation.
5. On the local node, verify and synchronize the cluster.

Changing distribution preference for service IP label aliases

You can configure a distribution preference for the service IP labels that are placed under PowerHA® SystemMirror® control. PowerHA® SystemMirror® lets you specify the distribution preference for the service IP label aliases.

When you specify the new distribution preference to be associated with a network, the PowerHA® SystemMirror® configuration database is updated but the preference is *not* changed dynamically; that is, PowerHA® SystemMirror® does *not* interrupt the processing by relocating service IP labels at the time the preference is changed. Instead, the next time a cluster event, such as a fallover takes place for a resource group that has service IP labels on the network, PowerHA® SystemMirror® uses the new distribution preference when it allocates the service IP label alias on the network interface on the backup node.

For information on types of distribution preferences, see Types of distribution for service IP label aliases.

To change a defined distribution preference for service IP labels:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Service IP Labels/ Addresses › Configure Service IP labels/addresses Distribution Preferences** and press Enter. PowerHA® SystemMirror® displays a list of networks.
3. From the list of networks, select the network for which you want to change the distribution preference and press Enter.
4. Change the distribution preference and press Enter. Remember that default values are not restored automatically.

Result

The changes you enter take effect the next time the resource group containing the service IP label is restarted.

Viewing distribution preference for service IP label aliases

Use the **cltopinfo** command to display the service IP label distribution preference specified for a particular network.

Example output:

```
Network net_ether_02
        NODE Ora_app_1:
                App_svc1          1.1.1.1
                App1_boot        192.9.201.129
        NODE Ora_app_2
                App2_boot        192.9.201.131
```

Network net_ether_02 is using the following distribution preference for service labels: Collocation with persistent. Service label(s) will be mapped to the same interface as the persistent label.

Reconfiguring tape drive resources

Using PowerHA® SystemMirror® SMIT panels, you can reconfigure tape drives in several different ways.

Take the following actions to reconfigure tape drives:

- Add tape drives as PowerHA® SystemMirror® resources
  - Specify synchronous or asynchronous tape operations
  - Specify appropriate error recovery procedures
- Change/Show tape drive resources
- Remove tape drive resources
- Add or remove tape drives to/from PowerHA® SystemMirror® resource groups.

To add tape drive resources, see Configuring a PowerHA® SystemMirror® cluster.

Changing a tape resource

This topic describes how to change or show the current configuration of a tape drive resource.

To change or show the current configuration:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Tape Resources › Change/Show a Tape Resource** and press Enter.
   SMIT returns a picklist of the configured tape drive resources.
3. Select the tape resource you want to see or change.
   SMIT displays the current configuration for the chosen tape device.
4. Change the field values as necessary.
5. Press Enter.

Removing a tape device resource

Use the information in this topic to remove a tape device resource.

To remove a tape device resource:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resources › Configure Tape Resources › Remove a Tape Resource** and press Enter.
   SMIT returns a picklist of the configured tape drive resources.
3. Select the tape resource you want to remove.

SMIT displays the message

```
Are You Sure?
```

Using NFS with PowerHA® SystemMirror®

You can use NFS with PowerHA® SystemMirror®.

PowerHA® SystemMirror® includes the following functionality:

- NFS Configuration Assist to relinquish NFS exports and cross mounts into PowerHA® SystemMirror® cluster.
- Pre-configured application monitor (clam_nfsv4) to monitor NFSv4 exports and the health of the NFS daemons.
- Ability to configure NFSv2/v3 and/or NFSv4 exports of the same file system/directory.
- Reliable NFS server capability that allows a backup processor to recover current NFS activity should the primary NFS server fail, preserving the locks on NFS file systems and dupcache. This functionality is restricted to two-node Resource Groups if it contains NFSv2/v3 exports. If all the exports in the resource group are NFSv4 only, then up to 16-node Resource Group configurations are supported.
- Ability to specify a Stable Storage location to maintain the NFSv4 state across all the nodes of the Resource Group.
- Ability to specify a network for NFS mounting.
- Ability to define NFS exports and mounts at the directory level.
- Ability to specify export options for NFS-exported directories and file systems.

> **Note:** The NFSv4 support for PowerHA® SystemMirror® needs the cluster.es.nfs fileset installed.

Reconfiguring resources in clusters with dependent resource groups

These topics describe the conditions under which PowerHA® SystemMirror® performs dynamic reconfigurations in clusters with dependent resource groups.

If you have configured dependent resources in the cluster, the dynamic reconfiguration (DARE) allows you to do the following:

- Make changes to the cluster resources
- Make changes to the cluster topology
- Dynamically add or remove resource groups from the cluster configuration.

When reconfiguring resources dynamically, PowerHA® SystemMirror® ensures the availability of applications in resource groups. For resource groups that have dependencies between them, it means that PowerHA® SystemMirror® only allows changing resources when it is safe to do so.

Reconfiguring resources and topology dynamically

Consider a cluster where resource group A (child) depends on resource group B (parent). In turn, resource group B depends on resource group C. Note that resource group B serves both as a parent for resource group A and a child for resource group C.

The following rules for DARE apply:

- You can make changes to the cluster topology and cluster resources dynamically for a child resource group and for a parent resource group.
- For a child resource group, if this resource group has no other groups that depend on it, PowerHA® SystemMirror® runs the reconfiguration events and performs the requested changes. PowerHA®

SystemMirror® performs a dynamic reconfiguration of a child resource group without taking any other resource groups offline and online.

- For a parent resource group, before proceeding with dynamic reconfiguration events, you must manually take offline *all* child resource groups that depend on the parent resource group. After the dynamic reconfiguration is complete, you can bring the child resource groups back online.

For instance, in a A>B>C dependency, where A is a child resource group that depends on B, and B is a child resource group that depends on C, to make changes to the resource group C, you must first take offline resource group A, then resource group B, and then perform a dynamic reconfiguration for resource group C. Once PowerHA® SystemMirror® completes the event, you can bring online resource group B and then resource group A.

If you attempt a dynamic reconfiguration event and PowerHA® SystemMirror® detects that the resource group has dependent resource groups, the DARE operation fails and PowerHA® SystemMirror® displays a message prompting you to take the child resource groups offline, before attempting to dynamically change resources or make topology changes in the parent resource group.

Making dynamic changes to dependent resource groups

If you have dependent resource groups configured, there are some rules that apply.

These rules include:

- If you dynamically add a resource group to the cluster, PowerHA® SystemMirror® processes this event without taking any resource groups offline or online.
- If you dynamically remove a resource group from the cluster configuration and the resource group is included in a dependency with one or more resource groups, then:
- If a resource group that you remove dynamically is a parent resource group, then before processing the dynamic reconfiguration event to remove the group, PowerHA® SystemMirror® temporarily takes offline dependent (child) resource group(s). After the DARE event is complete, PowerHA® SystemMirror® reacquires child resource groups.

For instance, consider the following resource group dependency: A >B>C, where A (child) depends on B, and B depends on C (parent). B is a child to resource group C and is a parent to resource group A.

In this case, if you dynamically remove resource group C from the cluster configuration, PowerHA® SystemMirror® takes resource group A offline, then it takes resource group B offline, removes resource group C, and reacquires first resource group B and then resource group A.

Cluster processing during DARE in clusters with dependent resource groups

As with cluster processing for other events, if you have dependencies configured in the cluster, cluster processing for dynamic reconfiguration is done in a different way than in clusters without dependencies between resource groups. As a result, the sequence of events in the **hacmp.out** file shows a series of **rg_move** events.

Synchronizing cluster resources

Whenever you modify the configuration of cluster resources in the Configuration Database on one node, you must synchronize the change across all cluster nodes. You perform a synchronization by choosing the Verification and Synchronization option from either the Cluster Nodes and Networks or Cluster Applications and Resources SMIT panel.

> **Note:** If the cluster is running, make sure no node has been stopped with its resource groups placed in UNMANAGED state when performing a synchronization.

The processing performed in synchronization varies depending on whether the Cluster Manager is active on the local node:

- If the Cluster Services is not active on the local node when you select this option, the Configuration Database data in the DCD on the local node is copied to the Configuration Databases stored in the DCDs on all cluster nodes.
- If the Cluster Services 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 Configuration Database data replaces the Configuration Database data stored in the ACD on each cluster node. The cluster daemons are refreshed and the new configuration becomes the active configuration. In the PowerHA® SystemMirror® log file, **reconfig_resource_release**, **reconfig_resource_acquire**, and **reconfig_resource_complete** events mark the progress of the dynamic reconfiguration.

In some cases, the verification uncovers errors that do not cause the synchronization to fail. PowerHA® SystemMirror® 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 do not interfere with the synchronization.

Log files that are no longer stored in a default directory, but a user-specified directory instead, are verified by the cluster **verification** utility, which checks that each log file has the same pathname on every node in the cluster and reports an error if this is not the case.

Managing resource groups in a cluster

These topics describe how to reconfigure the cluster resource groups. It describes adding and removing resource groups, and changing resource group attributes and processing order.

It also covers the **Resource Group Management** utility that allows you to change the status and location of resource groups dynamically using the SMIT interface, or the **clRGmove** command. This utility lets you move resource groups to other cluster nodes, for instance to perform system maintenance on a particular cluster node.

If you have dependent resource groups in the cluster, see Reconfiguring resources in clusters with dependent resource groups for information on making dynamic reconfiguration changes to the cluster resources.

Changing a resource group

These topics describe changes that you can make to resource groups.

Reconfiguring cluster resources and resource groups

You can view, change, add, and delete a resource group.

When you initially configured your PowerHA® SystemMirror® system, you defined 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 configured each resource group to have a particular kind of relationship with a set of nodes. You also assigned a priority to each participating node for some non-concurrent resource groups.

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.

You can also redefine the order in which PowerHA® SystemMirror® attempts to acquire and release the resource groups in your cluster. In general, PowerHA® SystemMirror® processes all individual resource groups configured in your cluster in parallel unless you define a specific serial order upon which certain resource groups should be acquired or released, using the Change/Show Resource Group Processing Order panel in SMIT.

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.

If Cluster Services is running on the local node, synchronizing the cluster triggers a dynamic reconfiguration event.

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:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Remove a Resource Group** and press Enter.
   SMIT displays a panel 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.

   > If you have the following parent/child resource group dependency chain configured: A > B > C, and remove the resource group B, PowerHA® SystemMirror® sends a warning that the dependency links between A and B, and between B and C are also removed.

4. Press Enter again to confirm your action.
5. Return to previous SMIT panels to perform other configuration tasks.
6. Synchronize the cluster configuration.

Result
If Cluster Services is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event.

Changing resource group processing order

By default, PowerHA® SystemMirror® acquires and releases resource groups in parallel. You can view or change the current order in which PowerHA® SystemMirror® processes resource groups in your cluster.

Use the **Change/Show Resource Group Processing Order** panel in SMIT to change or view the current order (smit cm_processing_order fastpath).

Resource group ordering during DARE

In general, PowerHA® SystemMirror® processes all individual resource groups configured in your cluster in parallel unless you define a specific serial order upon which certain resource groups should be acquired or released. Handling of any dependencies between resource groups takes priority over any serial processing you specify.

If you need to control the actual processing order during dynamic reconfiguration (DARE), make the changes to only one resource group at a time. Otherwise, the order in which resource groups are acquired and released may be unpredictable.

During the dynamic reconfiguration process, you could have two scenarios:

- Prior to dynamically changing any of the resource groups:
  - The processing order for *all* the resource groups was parallel
    and
  - You did *not* change it during dynamic reconfiguration (DARE).
- In this case, during the dynamic reconfiguration process, PowerHA® SystemMirror® processes the resource groups according to an alphabetically-sorted order, and *not* in parallel. If you made the changes to particular resource groups in the cluster, these changes may affect the order in which these resources will be actually released and acquired.
- Prior to dynamically changing any of the resource groups:
  - The processing order for some of the resource groups was parallel
    and
  - Some of the resource groups were included in the list for serial processing.

In this case, if during DARE, you change the serial order in which some of the resource groups are acquired or released on the nodes, then the newly specified order becomes valid during the reconfiguration process. PowerHA® SystemMirror® uses the new order during the same cluster reconfiguration cycle.

After reconfiguration is complete, PowerHA® SystemMirror® returns to the usual order of processing, as described below.

Resource group acquisition in PowerHA® SystemMirror® occurs in the following order:

1. Resource groups for which the customized order is specified are acquired in the customized serial order.
2. If some of the resource groups in the cluster have dependencies between them, these resource groups are processed in phases. For example, parent resource groups are acquired before child resource groups are acquired.
3. Resource groups that mount NFS only are processed in the specified order.
4. Resource groups that are *not* included in the customized ordering lists are acquired in parallel.

Resource group release in PowerHA® SystemMirror® occurs in the following order:

1. Resource groups for which no customized order has been specified are released in parallel.
2. PowerHA® SystemMirror® releases resource groups that are included in the customized release ordering list.
3. If some of the resource groups in the cluster have dependencies between them, these resource groups are processed in phases. For example, the child resource groups are released before the parent resource groups are released.
4. Resource groups that must unmount NFS are processed in the specified order.

However, if you made changes to particular resource groups in the cluster, these changes may affect the order in which these resource groups are released and acquired. As a result, during the dynamic reconfiguration process, the actual order in which resource groups are acquired and released is unpredictable.

This order is dependent on the changes you make to the order during DARE, and on the types of dynamic changes you make to the resource groups themselves. For instance, due to the changes you made to a particular resource group, this resource group may need to be released before others in the list, even though the alphabetically-sorted order is used for the remaining resource groups.

Changing the configuration of a resource group

You can change some of the configuration attributes of a resource group.

You can change the following attributes:

- Name of the resource group
- Nodes in the list of participating nodes
- Priority of participating nodes (by changing their position in the list of participating nodes)
- Startup, fallover and fallback policies for resource groups
- Attributes of the resource group.

You can change most 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.

Changing the basic configuration of a resource group

You can change the basic configuration of a resource group:

To change the basic configuration of a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Nodes and Policies for a Resource Group**. SMIT displays a list of the currently defined resource groups.
3. Select the resource group to change and press Enter.

> PowerHA® SystemMirror® shows *only* the valid choices for the specified resource group.

4. Enter field values as necessary.
5. Press Enter to change the resource group information stored in the PowerHA® SystemMirror® Configuration Database (ODM).
6. Return to previous SMIT panels to perform other configuration tasks or to synchronize the changes you just made.

Result
If Cluster Services is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event.

Changing resource group attributes

You can change the attributes of a resource group.

To change the attributes of a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group**. SMIT displays a list of the currently defined resource groups.
3. Select the resource group you want to change and press Enter.
   SMIT displays a list of resource group attributes and the values set.
4. Change field values as needed.
5. Press Enter to change the resource group information stored in the PowerHA® SystemMirror® Configuration Database.
6. Return to previous SMIT panels to perform other configuration tasks.
7. Synchronize changes to the configuration.

Result
If Cluster Services is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event.

Changing a dynamic node priority policy

You can use SMIT to change or show a dynamic node priority policy.

To show or change the dynamic node priority policy for a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources and Attributes for a Resource Group** and press Enter.
3. Select the resource group.
   You can change the dynamic node priority policy on the next panel, if you have one configured previously.
4. Select the policy you want and press Enter.

Changing a delayed fallback timer policy

You can use SMIT to change or show a delayed fallback timer policy.

To change or show a previously configured fallback policy, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Delayed Fallback Timer Policies › Change/Show a Delayed Fallback Timer Policy** and press Enter.
3. Select the fallback timer policy to change.

4. Change the fallback timer policy on the next panel.

Result
The new value for the timer will come into effect after synchronizing the cluster and after the resource group is released and restarted (on a different node or on the same node) due to either a cluster event or if you move the group to another node.

Note that you can change the parameters, but you cannot change the type of recurrence for the specific fallback timer. However, you can configure another fallback timer policy that uses a different predefined recurrence, and assign it to a resource group.

Removing a delayed fallback timer policy for a resource group

You can delete a previously configured delayed fallback timer policy.

You cannot remove a delayed fallback timer if any resource groups are configured to use it. First, change or remove the delayed fallback timer included as an attribute to any resource groups configured to use the unwanted timer, then proceed to remove it, as described in the following procedure.

To delete a previously configured delayed fallback timer policy, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Delayed Fallback Timer Policies › Remove a Delayed Fallback Timer Policy** and press Enter.
3. Select the fallback timer policy to remove and press Enter.
4. Press Enter.

Showing, changing, or deleting a settling time policy

You can change, show, or delete previously configured setting time policies.

Use the **Cluster Applications and Resource › Resource Groups › Configure Resource Group Run-Time Policies › Configure Settling Time for Resource Groups** SMIT path.

Changing a location dependency between resource groups

Location dependencies between resource groups come in three types: Online On Same Node, Online On Different Nodes, and Online On Same Site. You can change location dependencies between resource groups.

Changing an Online on Same Node dependency

You can change an **Online on Same Node** location dependency between resource groups.

To change an **Online on Same Node** location dependency:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Online on Same Site Dependency › Change/Show Online on Same Node Dependency between Resource Groups** and press Enter.
   PowerHA® SystemMirror® displays a list of resource groups configured with this location dependency.
3. Select the **Online on Same Node** dependency set of resource groups to show.
4. Add a resource group to the selected **Online on Same Node** dependency set of resource groups:

Online on Same Node Resource Group fields	
Field	**Value**
Resource Groups to be Online on the same node	PowerHA® SystemMirror® displays the resource groups listed in the selected set.

Field	Value
New Resource Groups to be Online on the same node	Press F4 to display the list of available resource groups. Select the resource groups from the list to be in this set of resource groups to be acquired and brought ONLINE on the same node (according to the startup policy and the availability of the node required). On fallback and fallover, the resource groups are processed simultaneously and brought ONLINE on the same target node (using the fallover and fallback policy defined for these groups).

5. Press Enter.
6. Verify and synchronize the cluster.

Changing an Online on Different Nodes dependency

Follow these instructions to change an **Online on Different Nodes** location dependency between resource groups.

To change an **Online on Different Nodes** location dependency:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resource › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Online on Different Nodes Dependency › Change/Show Online on Different Nodes Dependency between Resource Groups** and press Enter.
3. Select the **Online on Different Nodes** dependency set of resource groups to show.
4. Make changes as required and then press Enter.

Online on Different Nodes fields	
Field	**Value**
High Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) before less priority resource groups. On fallback and fallover, these resource groups are processed simultaneously and brought ONLINE on different target nodes before any other groups are processed. If different target nodes are unavailable for fallover or fallback, these groups (same priority level) can remain on the same node. Relative priority within this list is alphabetical by group names.
Intermediate Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) after the high priority groups and before low priority resource groups. are brought ONLINE. On fallback and fallover, these resource groups are processed simultaneously and brought ONLINE on different target nodes after the high priority groups and before low priority resource groups are processed. If different target nodes are unavailable for fallover or fallback, these groups (same priority level) can remain on the same node. Relative priority within this list is alphabetical by group names.
Low Priority Resource Group(s)	Select the resource groups to be in this set of resource groups to be acquired and brought ONLINE (according to the startup policy and the availability of the node required) after the higher priority resource groups are brought ONLINE. On fallback and fallover, these resource groups are processed and brought ONLINE on different target nodes after the higher priority groups are processed.

Field	Value
	Higher priority groups moving to a node may cause these groups to be moved or taken OFFLINE. Relative priority within this list is alphabetical by group names.

5. Verify and synchronize the cluster.

Changing a parent/child dependency between resource groups

You can change a parent/child dependency between resource groups.

To change a parent/child dependency:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Parent/Child Dependency › Change/Show Parent/Child Dependency between Resource Groups** and press Enter.
   A list of child-parent resource group pairs appears.
3. Select a pair from the list and press Enter. A screen appears where you can change the parent resource group or the child resource group.
4. Change the resource groups as required and press Enter. Note that you cannot change the **Dependency Type**.

Changing a start after dependency between resource groups

You can change a Start After dependency between resource groups

To change a Start After dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Start After Dependency › Change/Show Start After Dependency Between Resource Groups** and press Enter.
   A list of Source-Target resource group pairs appears.
3. Select a pair from the list and press Enter.
   A panel appears from which you can change the Source resource group or the Target resource group.
4. Change the resource groups as required and press Enter.
   You cannot change the Dependency Type.

Changing a stop after dependency between resource groups

You can change a Stop After dependency between resource groups

To change a Stop After dependency between resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Stop After Dependency › Change/Show Stop After Dependency Between Resource Groups** and press Enter.
   A list of Source-Target resource group pairs appears.
3. Select a pair from the list and press Enter.
   A panel appears from which you can change the Source resource group or the Target resource group.
4. Change the resource groups as required and press Enter.
   You cannot change the Dependency Type.

Displaying a parent/child dependency between resource groups

You can display parent/child dependencies between resource groups.

To display a dependency between parent/child resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Parent/Child Dependency › Display All Parent/Child Resource Group Dependencies** and press Enter.
   A selector screen appears.
3. Select **Display per Child** or **Display per Parent** to display all resource group dependencies for a child resource group, or for a parent resource group. Press Enter.
4. PowerHA® SystemMirror® displays a list similar to one of the following:

   ```
   Resource Group (RG_b) has the following parent resource groups:

   RG_a

   RG_e
   ```

   Or:

   ```
   Resource Group (RG_a) has the following child resource groups:

   RG_b

   RG_c

   RG_d

   Resource Group (RG_e) has the following child resource groups:

   RG_b

   RG_c

   RG_d
   ```

Displaying a start after dependency between resource groups

You can display start after dependencies between resource groups.

> **Note:** You can use ASCII SMIT to display Start After dependencies between resource groups.

To display a dependency between start after resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Start After Dependency › Display All Start After Resource Group Dependencies** and press Enter.
   A selector screen appears.
3. Select **Display per Source** or **Display per Target** to display all resource group dependencies for a source resource group, or for a target resource group. Press Enter.
4. PowerHA® SystemMirror® displays a list similar to one of the following:

   ```
   Resource Group (RG_2) has the following target resource groups:

   RG_3

   RG_4
   Resource Group (RG_1) has the following target resource groups:

   RG_2

   RG_3
   ```

```
RG_4
```

Or:

```
Resource Group (RG_2) has the following source resource groups:
RG_1
Resource Group (RG_3) has the following source resource groups:
RG_2
RG_1
Resource Group (RG_4) has the following source resource groups:
RG_2
RG_1
```

Displaying a stop after dependency between resource groups

You can display stop after dependencies between resource groups.

> **Note:** You can use ASCII SMIT to display Stop After dependencies between resource groups.

To display a dependency between stop after resource groups:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Stop After Dependency › Display All Stop After Resource Group Dependencies** and press Enter.
   A selector screen appears.
3. Select **Display per Source** or **Display per Target** to display all resource group dependencies for a source resource group, or for a target resource group. Press Enter.
4. PowerHA® SystemMirror® displays a list similar to one of the following:

```
Resource Group (RG_2) has the following target resource groups:
RG_3
RG_4
Resource Group (RG_1) has the following target resource groups:
RG_2
RG_3
RG_4
```

Or:

```
Resource Group (RG_2) has the following source resource groups:
RG_1
Resource Group (RG_3) has the following source resource groups:
RG_2
RG_1
Resource Group (RG_4) has the following source resource groups:
RG_2
```

```
RG_1
```

Removing a dependency between resource groups

You can remove any of the four types of dependencies between resource groups.

Deleting a parent/child dependency between resource groups

You can delete a parent/child dependency between resource groups.

To delete a parent/child dependency, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Parent/Child Dependency › Remove a Dependency between Parent/Child Resource Groups** and press Enter. PowerHA® SystemMirror® displays a list of parent/child resource group pairs.
3. Select a pair from the list to delete and press Enter. Deleting a dependency between resource groups does not delete the resource groups themselves.

Result

> **Note:** If you have the following dependency chain configured: A > B > C, and remove the resource group B, PowerHA® SystemMirror® sends a warning that the dependency links between A and B, and between B and C are also removed.

Deleting a start after dependency between resource groups

You can delete a start after dependency between resource groups.

To delete a start after dependency, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Start After Dependency › Remove a Dependency Between Start After Resource Groups** and press Enter. PowerHA® SystemMirror® displays a list of start after resource group pairs.
3. Select a pair from the list to delete and press Enter. Deleting a dependency between resource groups does not delete the resource groups themselves.

Deleting a stop after dependency between resource groups

You can delete a stop after dependency between resource groups.

To delete a stop after dependency, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies Between Resource Groups › Configure Stop After Dependency › Remove a Dependency Between Stop After Resource Groups** and press Enter. PowerHA® SystemMirror® displays a list of stop after resource group pairs.
3. Select a pair from the list to delete and press Enter. Deleting a dependency between resource groups does not delete the resource groups themselves.

Deleting a location dependency between resource groups

You can delete a location dependency between resource groups.

To delete a location dependency, complete the following steps:

1. In SMIT, select the path for configuring the location dependency that you want to remove.
   This example shows the path for Online on Same Node Dependency: **Cluster Applications and Resources › Resource Groups › Configure Resource Group Run-Time Policies › Configure Dependencies between Resource Groups › Configure Online on same node Dependency › Remove Online on Same Node Dependency between Resource Groups**and press Enter.

   PowerHA® SystemMirror® displays a list of resource groups with this location dependency.
2. Select the **Online on same node** dependency to remove and press Enter.
   Deleting a dependency between resource groups does not delete the resource groups themselves. The resource groups are now handled individually according to their startup, fallover, and fallback policies.

Adding or removing individual 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 resource groups even if another node in the cluster is inactive. However, it is more convenient to have nodes active, so you can obtain a list of possible shared resources for each field by pressing the F4 key when you are in the SMIT **Change/Show Resources/Attributes for a Resource Group** panel.

Resource groups can contain many different types of cluster resources, including IP labels/addresses, file systems, volume groups and application controllers. 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/Attributes for a Resource Group** panel. See the following section.

Reconfiguring resources in a resource group

You can change the resources in a resource group.

To change the resources in a resource group:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Applications and Resources › Resource Groups › Change/Show Resources/ Attributes for a Resource Group**and press Enter. SMIT displays a picklist of configured resource groups.
3. Select the resource group you want to change and press Enter. SMIT displays a panel that lists all the types of resources that can be added to the type of selected resource group, with their current values.

   > If you specify file systems to NFS mount in a non-concurrent resource group with the startup policy of either Online on Home Node Only, or Online on First Available Node, you must also configure the resource to use IP Address Takeover. If you do *not* do this, takeover results are unpredictable. You should also set the field value **Filesystems Mounted Before IP Configured** to true so that the takeover process proceeds correctly.

4. Enter the field values you want to change, and press Enter.
5. Return to previous SMIT panels to perform other configuration tasks or to synchronize the changes you just made.

Result
If Cluster Services is running on the local node, synchronizing cluster resources triggers a dynamic reconfiguration event.

Forcing a varyon of a volume group

You can force a varyon of a volume group either by specifying an attribute in SMIT, or by entering a command at the command line.

It is recommended that you use SMIT to force a varyon because PowerHA® SystemMirror® does the following before attempting to activate a volume group on a node:

- Checks whether the LVM mirroring is used for the disks

- Verifies that at least one copy of every logical volume for this volume group can be found.

It is recommended to specify the **super strict** allocation policy for the logical volumes in volume groups for which forced varyon is specified.

As with regular volume group operations, you can determine the final status of the volume group using the messages logged by PowerHA® SystemMirror® during the verification process and the information logged in the **hacmp.out** file. You can also use the **lsvg -o** command to verify whether a volume group is offline or online, and the **lsvg -l** command to check the volume group status and attributes.

If after checking partition maps, PowerHA® SystemMirror® cannot find a complete copy of every logical volume for a volume group, an error message: `"Unable to vary on volume group <vg name> because logical volume <logical volume name> is incomplete"` displays in the **hacmp.out** file and the volume group remains offline.

Forcing a varyon of a volume group from SMIT

Use the following procedure to ensure that you always have access to your data if there is a copy available, and that you receive notification if you lose either a copy of your data, or all copies.

Note that you specify a forced varyon attribute for all volume groups that belong to a resource group. For instructions on setting a forced varyon attribute using SMIT, see Forcing a varyon of volume groups.

With this attribute, if a normal **varyonvg** fails, a check is made to ensure that there is at least one complete copy of all data available in the volume group. If there is, it runs **varyonvg -f**; otherwise, the volume group remains offline. Specifying the forced varyon attribute for a volume group eliminates the need for a quorum buster disk or special scripts to force a varyon, although you can continue to use these methods.

To use PowerHA® SystemMirror® forced varyon and error notification:

1. Disable quorum on your volume group. This ensures that it does not vary off if you still have access to a copy of your data.
2. Use the SMIT forced varyon option to vary on your volume group if your data is available.
3. Set up error notification to inform you if a file system or logical volume becomes unavailable.

Forcing a varyon of a volume group from the command line

Issue the **varyonvg -f** command for a specific volume group on a node in the cluster.

If you use this method, PowerHA® SystemMirror® does not verify that the disks are LVM mirrored, and does not check the logical partitions to verify that at least one complete copy of every logical volume can be found for this volume group. You should use this command with caution to avoid forcing a varyon of a volume group in a partitioned cluster.

> **Important:** Forcing a varyon with non-mirrored logical volumes and missing disk resources can cause unpredictable results (both conditions must be present to cause problems.) Forcing a varyon should only be performed with a complete understanding of the risks involved. Also, refer to the AIX® documentation.

Avoiding a partitioned cluster

Use care when using forced varyon to activate a volume group. If the cluster becomes partitioned, each partition might force the volume group to vary on and continue to run. In this case, two different copies of the data are active at the same time.

This situation is referred to as *data divergence* , and does *not* allow a clean recovery. If this happens in a concurrent volume group, the two sides of the cluster have made uncoordinated updates.

Moving resource groups

The Resource Group Management utility (clRGmove) allows you to perform maintenance on a node without losing access to the node's resources. You are not required to synchronize cluster resources or stop cluster services.

The Resource Group Management utility provides improved cluster management by allowing you to:

- Bring a resource group online or offline.
- Move a resource group to a new location. This location can be a node in the same site or a node in the other site.

If you have requested PowerHA® SystemMirror® to move, activate, or stop a particular resource group, then no additional operations on any additional groups will run until this operation is completed.

Specific considerations related to moving a resource group are:

- PowerHA® SystemMirror® attempts to recover resource groups in the ERROR state upon **node_up** events. However, if you moved a group to Node A, the group remains on Node A (even in the error state).
- When node B joins the cluster, it does not acquire any resource groups that are currently in the ERROR state on node A. To recover such resource groups, manually bring them online or move them to other nodes.

> **Note:** When you request PowerHA® SystemMirror® to move a resource group, it uses the **clRGmove** utility, which moves resource groups by calling an **rg_move** event. It is important to distinguish between an **rg_move** event that is triggered automatically by PowerHA® SystemMirror®, and an **rg_move** event that occurs when you explicitly request PowerHA® SystemMirror® to manage resource groups for you. To track and identify the causes of operations performed on the resource groups in the cluster, look for the command output in SMIT and for the information in the **hacmp.out** file.

Before attempting to explicitly move a resource group from one node to another, or to take a resource group online or offline, ensure that:

- Cluster Services is running on the node that releases the resource group and on the node that acquires it.
- The cluster is stable. If the cluster is not stable, the operation that you request with the resource group terminates and you receive an error message.

To move a resource group, complete the following steps:

1. From the command line, enter `smit cspoc`.
2. In SMIT, select **Resource Group and Applications** and press Enter.
3. Depending on where you want to move the resource group, select one of the following options:
   - **Move Resource Group to Another Node**
   - **Move Resource Group to Another Site**
4. Select the resource that you want to move from the list and press Enter.
5. Complete all required fields and press Enter.

Moving resource groups with dependencies

PowerHA® SystemMirror® prevents you from moving resource groups online or to another node under certain conditions

These conditions include:

- If you took the parent resource groups offline with the Resource Group Management utility, **clRGmove**, PowerHA® SystemMirror® rejects manual attempts to bring the resource groups that depend on these resource groups online. The error message lists the parent resource groups that you must activate first to satisfy the resource group dependency.
- If you took the target resource groups offline with the Resource Group Management utility, **clRGmove**, PowerHA® SystemMirror® rejects manual attempts to bring the resource groups that depend on these resource groups online. The error message lists the target resource groups that you must activate first to satisfy the resource group dependency.
- If you have a parent and a child resource group online, and want to move the parent resource group to another node or take it offline, PowerHA® SystemMirror® prevents you from doing so before a child

resource group is taken offline. However, if both parent and child are in the same **Same Node** or **Same Site** location dependency set, you can move them both as you move the whole set.

- If you have a source and a target resource group online (with stopafter dependency), and want to move the stopafter-source resource group to another node or take it offline, PowerHA® SystemMirror® prevents you from doing so before a stopafter-target resource group is taken offline. However, if both source and target are in the same **Same Node** or **Same Site** location dependency set, you can move them both as you move the whole set.
- You can move **Same Node** dependency or **Same Site** dependency sets of resource groups. You must specify that all groups are moved in the same operation.
- The rules for location dependencies may not allow all combinations of resource groups moves. You need to consider your needs for manual control of resource groups when configuring dependencies.

No automatic recovery for resource groups that fail to migrate

If you request PowerHA® SystemMirror® to move a resource group to a node and during this operation the destination node fails to acquire the group, the resource group is put into an ERROR state. If you try to move a resource group that has a dependency (parent/child, Startafter, Stopafter, or location) that prohibits the move, the resource group will be in the DEPENDENCY_ERROR state.

Similarly, if you request PowerHA® SystemMirror® to activate the resource group on a particular node, and this node fails to bring the resource group online, the resource group is put into an ERROR state.

In either case, PowerHA® SystemMirror® does *not* attempt to acquire or activate the resource group on any other node in the cluster. The error messages in these cases indicate that your intervention is required to move the resource group to another node.

If you request PowerHA® SystemMirror® to migrate a resource group to another node, but the node that owns it fails to release it, or if you request to bring a resource group offline on a particular node, but the node fails to release it, an error message indicates that your intervention is required to stabilize the cluster.

Moving resource groups by using the command line

You can use the **clRGmove** command to move resource groups.

The **clRGmove** utility lets you manually control the location and the state of resource groups by calling the **rg_move** event. With this command, you can bring a specified resource group offline or online, or move a resource group to a different node. This utility provides the command line interface to the Resource Group Migration functionality, which can be accessed through SMIT. You can also use this command from the command line, or include it in the pre-event and post-event scripts.

For a resource groups that do not have the **Online On All Available Nodes** startup policy (non-concurrent resource groups), you can complete the following tasks:

- Take the resource group offline from an online node
- Bring the resource group online to a specific node
- Move the resource group from its current hosting node to a new location.

For a resource groups that have the **Online On All Available Nodes** startup policy (concurrent resource groups), you can complete the following tasks:

- Take the resource group offline from all nodes in the group's nodelist
- Take the resource group offline from one node in the group's nodelist
- Bring the resource group online on all nodes in the group's nodelist
- Bring the resource group online on one node in the group's nodelist.

Bringing a resource group online

A resource group must be offline or in an error state to bring the resource group online. You can use the SMIT interface to bring a resource group online.

To bring a resource group online, complete the following steps:

1. From the command line, enter `smit cspoc`.
2. In SMIT, select **Resource Groups and Applications › Bring a Resource Group Online** and press Enter.
3. Select the resource group from the list and press Enter.
4. Select a destination node from the list and press Enter. If an originally configured highest priority node for the resource group is now available to host the group, an asterisk (*) is displayed next to the resource group.

   > The list displays only nodes that are running cluster services, participate in the resource group nodelist, and have enough available resources to host the resource group. The nodes in this list appear in the same order of priority as in the resource group nodelist.

5. From the **Bring a Resource Group Online** menu, complete the following field:

Bring a Resource Group Online fields	
Field	**Value**
Resource Group to Bring Online	Resource group to be activated.
Destination Node	Destination node that you selected.

6. Confirm your selections and press Enter to start the execution of the **rg_move** event and bring the resource group online. You do not need to synchronize the cluster.

Result
If the event completes successfully, PowerHA® SystemMirror® displays a message and the status and location of the resource group that was successfully brought online on the specified node.

If you requested PowerHA® SystemMirror® to activate the resource group on a particular node, and this node fails to bring the resource group online, the resource group is put into the ERROR state. In this case, PowerHA® SystemMirror® does not attempt to activate the resource group on any other node in the cluster without your intervention. The error message in this case indicates that your intervention is required to activate the resource group on another node and stabilize the cluster.

Taking a resource group offline

A resource group must be online or in an error state to take the resource group offline. You can use the SMIT interface to take a resource group offline.

To take a resource group offline, complete the following steps:

1. From the command line, enter `smit cspoc`.
2. In SMIT, select **Resource Groups and Applications › Bring a Resource Group Offline** and press Enter.
3. Select the resource group from the list and press Enter.
4. Select a destination node from the list. The list displays only nodes that are running cluster services. The destination node that you select is temporarily set as the highest priority node for this resource group.
5. From the **Bring a Resource Group Offline** menu, complete the following fields:

Bring a Resource Group Offline fields	
Field	**Value**
Resource Group to Bring Offline	Resource group that is to be stopped or brought offline.
Destination Node	Node on which the resource group is stopped.

6. Confirm your selections and press Enter to start the execution of the **rg_move** event and bring the resource group offline. You do not need to synchronize the cluster.

Result
If the event completes successfully, PowerHA® SystemMirror® displays a message and the status and location of the resource group that was successfully stopped on the specified node.

If you requested to bring a resource group offline on a particular node, and the resource group fails to release from the node on which it is online, an error message indicates that your intervention is required to stabilize the cluster.

Checking resource group state

As with regular cluster events, you can debug the status of resource groups using the messages logged by PowerHA® SystemMirror® in the **hacmp.out** file.

In addition, you can use **clRGinfo** to view the resource group location and status. See Using the clRGinfo command for an example of the command output. Use **clRGinfo** to view the node that is temporarily the highest priority node.

Special considerations when stopping a resource group

After taking a resource group offline, 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 using the Resource Group and Application Management utility.

- If you use **clRGmove -d** to bring down a resource group with Online on Home Node startup policy, Fallover to Next Priority Node in the List fallover policy and Fallback to Higher Priority Node in the List fallback policy, and which resides on the highest priority node, it will remain in an inactive state. You must manually bring the resource group online through resource group management.
- If you specify the **fallover** option of application monitoring for a resource group using the **Customize Resource Recovery** SMIT panel, which may cause resource groups to migrate from their original owner node, the possibility exists that while the highest priority node is up, the resource group remains down. Unless you bring the resource group up manually, it will remain in an inactive state.
- If your resource group was placed in an UNMANAGED state, due to stopping cluster services without stopping the applications, you may need to bring this resource group online manually.

Example: Using clRGmove to swap resource groups

In the three-node cluster indicated here, each node (Node1, Node2, and Node3) has a service label and a boot label.

The three nonconcurrent resource groups have the following policies:

- Startup: **Online on Home Node Only**
- Fallover: **Fallover to Next Priority Node in the List**
- Fallback: **Fallback to Higher Priority Node in the List** .

These resource groups have 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)

Node2's resources - contained in **CrucialRG** - are of particular importance to your operation. A situation could occur 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 boot label has already been taken. The cluster now looks like this:

Node1

DOWN

Node2

DOWN

Node3

UP (RG3, RG1)

The crucial resource group is unavailable. PowerHA® SystemMirror® is able to take care of only one failure, because there are no more boot label, 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 Resource Group Management utility to "swap" the resource groups so you can access **CrucialRG** instead of `RG1`.

You do this by issuing the following commands:

`clRGmove -g RG1 -n node3 -d` to bring `RG1` offline on `Node3`, and `clRGmove -g CrucialRG -n node3 -u` to bring `CrucialRG` online on `Node3`.
After these resource group migration commands are completed, access to `CrucialRG` is restored, and the cluster looks like this:

Node1

DOWN

Node2

DOWN

Node3

UP (RG3, CrucialRG)

> **Note:** You can move one or more resource groups at once with the **clRGmove** command to another node.

Managing users and groups

These topics describe how to use the SMIT Cluster Management (C-SPOC) utility to manage user accounts and groups, this applies to LDAP as well, on all nodes in a cluster by making configuration changes on a single node, and on LDAP from any node in a cluster.

Overview for AIX® and LDAP users and groups

PowerHA® SystemMirror® allows you to manage AIX® and LDAP user and group accounts across a PowerHA® SystemMirror® cluster. Groups provide an additional level of security and enable system administrators to manipulate a group of users as a single entity. In addition, PowerHA® SystemMirror® provides a utility that lets you authorize specified users to change their own password across nodes in a PowerHA® SystemMirror® cluster.

Requirements for managing user accounts in a PowerHA® SystemMirror® cluster

AIX® files that store user account information should be consistent across cluster nodes. These files are:

- The system **/etc/passwd** file
- Other system files in the **/etc/security** directory.

This way, if a cluster node fails, users can log on to the surviving nodes without experiencing problems caused by mismatched user or group IDs.

As the system administrator of a PowerHA® SystemMirror® cluster, you can use the C-SPOC utility to manage user and group accounts from any node in a cluster. C-SPOC propagates new and updated information to all of the other nodes in the cluster.

Note: Managing user accounts through C-SPOC requires that the Cluster Communications daemon is running and that all cluster nodes are active.

Important: If you manage user accounts with a utility such as Network Information Service (NIS) or Distributed Computing Environment (DCE) Manager, do not use PowerHA® SystemMirror® user management. Using PowerHA® SystemMirror® user management in this environment might cause serious system inconsistencies in the database.

Requirements for managing LDAP user accounts

You can use the C-SPOC utility to mange users and group accounts from any node in a cluster. If you create a user name that already exists on any node in the cluster, the operation might fail.

Note: The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

User account configuration

Make sure user accounts are the same on all nodes in the cluster. Run verification after you make changes to user accounts.

If a node in the cluster has fewer password restrictions than the other nodes, a user could make changes from the node with fewer restrictions and degrade cluster security.

Status of C-SPOC actions

If an action initiated by the C-SPOC utility fails, check the C-SPOC log file, **/tmp/cspoc.log**, to obtain the status of the command on each cluster node.

Note: The default location of this log file is **/tmp/cspoc.log**. If you redirected this log, check the appropriate location.

Managing AIX® and LDAP user accounts across a cluster

You can authorize users to change their own password and have C-SPOC propagate that password across cluster nodes.

Listing AIX® and LDAP users on all cluster nodes

To obtain information about all user accounts on cluster nodes or on the nodes in a specified resource group, you can use the following procedures or run the **`cl_lsuser`** command.

To list all user accounts on all cluster nodes by using the C-SPOC utility, complete the following steps.

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › List Users in the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. In the **List Users in the Cluster** window, leave the selection for a resource group blank to display information about all users, and press Enter.

A listing of user accounts similar to the following is displayed.

```
      COMMAND STATUS

Command: OK  stdout: yes stderr: no

Before command completion, additional instructions may appear.

[TOP]
sigmund  root   0/
sigmund  daemon 1/etc
sigmund  bin    2/bin
sigmund  sys    3/usr/sys
sigmund  adm    4/var/adm
sigmund  uucp   5/usr/lib/uucp
sigmund  guest  100     /home/guest
sigmund  nobody -2      /
sigmund  lpd    9/
sigmund  nuucp  6/var/spool/uucppublic
orion    root   0/
orion    daemon 1/etc
orion    bin    2/bin
[MORE...18]
```

AFTER COMPLETING THE TASK

Note: The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

To list all user accounts for LDAP by using the C-SPOC utility, complete the following steps.

1. Enter the fast path `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › List Users in the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the LDAP mode and press Enter.

SMIT lists user accounts similar to the following output:

```
      COMMAND STATUS

Command: OK  stdout: yes stderr: no

Before command completion, additional instructions may appear.

[TOP]
daemon 1/etc
bin    2/bin
sys    3/usr/sys
adm    4/var/adm
uucp   5/usr/lib/uucp
guest  100     /home/guest
nobody -2      /
lpd    9/
nuucp  6/var/spool/uucppublic
 [MORE...18]
```

Adding AIX® and LDAP user accounts on all cluster nodes

In the AIX® operating system, you can add user accounts by using the **mkuser** command or the **smit mkuser** command.

Note: The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

The user account information is stored in the **/etc/passwd** file. The data files are stored in the **/etc/security** directory. For more information about the **mkuser** command, see its man page.

To add an LDAP user or an AIX® operating system user to all nodes in a cluster using the C-SPOC utility, complete the following steps on any cluster node:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › Add a User to the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. Enter data in the applicable fields to set up the account, and press Enter.
   AIX® provides help information that describes each attribute. The **User Name** field is the only required field.

   > You can specify a value in the **User ID** field so that the account user ID is the same on all cluster nodes. If you do not specify this value, AIX® could assign a different user ID 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. If you are adding an LDAP user account, you must select a role that is specific to PowerHA® SystemMirror® in the **Roles** field.

5. The user account is created on all cluster nodes. If you are adding an LDAP user, the user is created in LDAP.

Result
The C-SPOC utility creates the AIX® user account and home directory for the new account on each remote cluster node that you specify.

If a user with the same name exists on one of the AIX® cluster nodes, the operation fails, returning this message:

```
user-name already exists on node nodename
```

You can specify that the command continue processing, even if the user name exists on one of the AIX® cluster nodes, by specifying the **force** option.

If you are adding an LDAP user, that user name cannot exist on any nodes in the cluster. Also a home directory is created automatically on all the nodes in the cluster.

Changing attributes of AIX® and LDAP user accounts in a cluster

Using the AIX® operating system, you can change any of the attributes that are associated with an existing user account by using the **chuser** command or SMIT interface.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

The **chuser** command changes the user information that is stored in the **/etc/passwd** file and the files in the **/etc/security** directory.

You can also change attributes associated with an existing user account from C-SPOC, as described in the following procedure. This procedure runs the AIX® **chuser** command on each cluster node. All cluster nodes must be active, the Cluster Communications daemon must be running, and a user with the specified name must exist on all the nodes for the change operation to proceed.

To change the characteristics of an LDAP user account or an AIX® user account on all cluster nodes by using the C-SPOC utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
3. Specify the name of the user account you want to change and press Enter. Press F4 to obtain a list of users from which to select. If you are changing LDAP attribute,s pressing F4 displays LDAP users. SMIT displays a list of the user account attributes and their current values.
4. Enter the new values for attributes you want to change, and press Enter. AIX® provides help information that explain each attribute. SMIT runs the C-SPOC command to change the attributes of the AIX® user

account on all cluster nodes. This process does not occur if you are changing attributes of an LDAP user account.

Removing AIX® and LDAP user accounts from a cluster

Using the AIX® operating system, you can remove a user account by using the **rmuser** command or the fastpath **smit cl_rmuser**.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

You can also remove a user account from cluster nodes from C-SPOC, as described in the following procedure. This procedure runs 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 accessible only to users with root permissions or by the group in which the user was a member.

To remove an LDAP user account or an AIX® user account from all cluster nodes by using the C-SPOC utility, complete the following steps.

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › User in a PowerHA® SystemMirror® cluster › Remove a User from the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. Complete the following fields, and press Enter.

Authentication and registry mode fields	
Field	**Value**
User Name	Enter the user name for the account you want to remove. The user name can be up to 8 characters in length. Press F4 to display a list of LDAP users that you can remove.
Remove Authentication information?	Specify **Yes** to delete the password and other authentication information from system security files. For an LDAP user account, the directory structure for the selected user account is removed.

Managing password changes for users

You can manage user passwords from any node in a cluster.

You can allow specified users to change their password on multiple nodes in the cluster by changing their password on one node.

a PowerHA® SystemMirror® user, that is a user who has an AIX® user account on each node in a cluster, can use the C-SPOC utility to change their own password across nodes in the cluster.

> **Important:** If you manage user accounts with a utility such as Network Information Service (NIS) or Distributed Computing Environment (DCE) Manager, do not use PowerHA® SystemMirror® user management. Using PowerHA® SystemMirror® user management in this environment might cause serious system inconsistencies in the database.

Before you authorize users to change their password or change a user's passwords, ensure that:

- The cluster topology is configured properly.

- The user's account exists on every cluster node in a specified resource group, and if no resource group is specified, in the entire cluster.
- The user's account exists on the local node. (The password changes on the local node, even if that node is not in the selected resource group.)
- All cluster nodes are powered up and accessible.

Note: These conditions should also be met before a user changes their own password. As a user may not have this information, the utility displays messages to a user should their attempt to change their password fail.

Allowing users to change their own passwords

System administrators can enable the new Cluster Password (**clpasswd)** utility.

This utility, when enabled, links to the AIX® system password utility to:

- Let system administrators authorize specified users to change their password across cluster nodes
- Let authorized users change their own password across a resource group or cluster (as configured), rather than having to change their password on each node in the cluster.
  This means that the user's AIX® system password is the same on the set of nodes specified.

  Note: The security of the password propagated to other nodes is only as secure as the network used to distribute the password.

Depending on the configuration of the Cluster Password utility, it lets users change their password using either:

- C-SPOC
- **clpasswd** command.

Both of these call the AIX® **passwd** command. The **clpasswd** command uses the same arguments as the **passwd** command. For more information about the **clpasswd** command, see its man page.

The following table shows where a user's password is changed based on the user's authorization, the password utility that is active, and the command executed:

Displays where a user's password is changed based on the user's authorization, the password utility that is active, and the command executed.

User password options			
User authorization	**When the system password utility is linked to clpasswd and the AIX® passwd command is run**	**When the system password utility is active (not linked to clpasswd)**	
		The AIX® passwd command is run	**The PowerHA® SystemMirror® clpasswd command is run**
The user authorized to change password across cluster	The password is changed on all cluster nodes.	The password is changed only on the local node.	The password is changed on all cluster nodes.
The user is not authorized to change password across cluster	The password is changed only on the local node.	The password is changed only on the local node.	The password is not changed.

Configuring the cluster password utility

With the SMIT interface, you can configure the cluster password utility.

To enable the Cluster Password utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.

2. From the SMIT interface, select **Security and Users › Passwords in an PowerHA® SystemMirror® cluster › Modify System Password Utility** and press Enter.
3. Complete the following fields:
   /bin/passwd utility is

   Select **Link to Cluster Password Utility** to link the Cluster Password Utility to the AIX® password utility. This option enables the Cluster Password utility. Select **Original AIX® System Command** to remove the link from the Cluster Password utility to the AIX® password utility. This option disables the Cluster Password utility.

   Select nodes by resource group

   Select one or more resource groups to enable the Cluster Password utility on the nodes in the specified groups. Leave the field blank to enable the Cluster Password utility on all cluster nodes.

   When the Cluster Password utility is linked to the AIX® password utility, PowerHA® SystemMirror® creates a `/usr/es/sbin/cluster/etc/clpasswd/usr_bin_passwd.orig` file to store the AIX® passwd utility. If you disable the Cluster Password utility, PowerHA® SystemMirror® removes the link between the two files, and the `usr_bin_passwd.orig` file is moved to `/bin/passwd` file.

Configuring authorization for users

After the Cluster Password utility is linked to the AIX® system password utility (**passwd**), you can specify and update which users have permission to change their passwords across a cluster.

To specify which users can change their own password, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. From the SMIT interface, select **Security and Users › Passwords in an PowerHA® SystemMirror® cluster › Manage List of Users Allowed to Change Password** and press Enter.
3. Press F4 to select a user from a list that you want to allow to change their passwords across the cluster. Select **ALL_USERS** to allow all users the authority to change their password across the cluster.
4. Verify that the user name is correct, and press Enter.

> You can view the list of users that are allowed to change their password across a cluster. You can also remove a user from the list. The `/usr/es/sbin/cluster/etc/clpasswd/cl_passwd_users` file stores the list of users that are allowed to change their password across a cluster.

Changing passwords for user accounts

You must have root authority to use C-SPOC to change a user's password or to specify that a user can change their password during the next login. You can configure this setting to change passwords on all cluster nodes.

If you use C-SPOC to change a user password for all nodes that belong to a resource group, make sure you complete this operation on a node that is included in the resource group. If you run this C-SPOC command from a node that is not part of the resource group, the password changes on that node also.

To use SMIT to change a user's password on a list of nodes in the cluster or in LDAP, complete the following steps.

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Passwords in a PowerHA® SystemMirror® cluster › Change a User's Password in the Cluster** and press Enter.
3. Select the corresponding authentication and registry mode for the user account for which you want to change the password and press Enter.
4. Select the node that contains the user account and press Enter.

> If you leave the field blank, all nodes in the cluster are selected.

5. Complete the Enter field values as follows:

Change a user's password in the cluster fields	
Field	**Value**
User Name	Select the name of the user whose password you want to change. If you are changing the password for an LDAP user account, press F4 to select from a list of LDAP users accounts.
User must change Password on first login?	Values that are entered in this field do not impact the cluster. This field was removed in the following versions of PowerHA® SystemMirror®: • PowerHA® SystemMirror® Version 7.1.3 Service Pack 6 • PowerHA® SystemMirror® Version 7.2.0 Service Pack 2 • PowerHA® SystemMirror® Version 7.2.1 • PowerHA® SystemMirror® Version 7.2.2

> PowerHA® SystemMirror® does not support the **ADMCHG** flag for changing passwords.

6. Press Enter to change the password.

Changing the password for your own user account

As an individual user, you can change your password on all cluster nodes, or on nodes within a specified resource group, if the Cluster Password utility is enabled on each cluster node and the administrator (who has root privileges) has given you permission to change your password on nodes across a cluster.

> **Note:** The password you are changing is your AIX® password on the specified nodes.

If you are unsure whether or not you are authorized to change your password, or if you try to change your password and receive an error message, contact your system administrator.

To change your password on cluster nodes or in LDAP, complete the following steps.

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Passwords in a PowerHA® SystemMirror® cluster › Change Current User's Password** and press Enter.
3. Complete the following fields.

Change Current® User's Password fields	
Field	**Value**
Select nodes by Resource Group	Select the resource groups that contains the nodes where you want to change your password. If you are changing the password for AIX® cluster nodes, you can leave this field blank to select all nodes in the cluster.
User Name	Verify that this field displays your user name. If it displays another name, contact your system administrator.

4. Press Enter.
5. Change your password on the panel that appears.

Result
If C-SPOC can distribute your new password to all cluster nodes or the nodes in a specified resource group, it changes your password across the nodes. Messages advise you of the progress of the password change and display the nodes on which the change takes place.

If C-SPOC cannot communicate with all cluster nodes, it does not change your password, and it displays a message to that affect.

> **Note:** If your password is changed on some, but not all, of the cluster nodes, a message appears that directs you to contact your system administrator. Be sure to talk with your system administrator because your password might be inconsistent among nodes in the specified resource groups or cluster.

You can also use the **clpasswd** command to change your cluster password. If you have not been authorized to change your password on cluster nodes, the **clpasswd** command does not let you change your password on any node, including the one you are currently logged in to.

Managing AIX® and LDAP group accounts

All users must belong to an AIX® or LDAP group. AIX® and LDAP groups add a level of security.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

If you manage user accounts with a utility such, as Network Information Service (NIS) or Distributed Computing Environment (DCE) Manager, do not use PowerHA® SystemMirror® user management. Using PowerHA® SystemMirror® user management in this environment might cause serious system inconsistencies in the database.

Listing AIX® and LDAP groups on all cluster nodes

Each AIX® and LDAP 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. In the AIX® operating system, you obtain information about all the groups defined on an AIX® system by running the **lsgroup** command.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

You can obtain information about the groups defined on all cluster nodes from C-SPOC, as described in the following procedure, or by running the C-SPOC **cl_lsgroup** command, specifying the ALL argument. Both C-SPOC and the **cl_lsgroup** command run 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 run.

If you specify a group name that does not exist on a cluster node, the **cl_lsgroup** command displays a warning message but continues running the command on all of the other cluster nodes.

To list all the groups defined in LDAP or on each AIX® cluster node by using the C-SPOC utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › List all groups in the Cluster**, and press Enter.
3. On the **Select an Authentication and registry mode** panel, select the mode and press Enter.
   a. If you select **LOCAL** for the mode, SMIT displays the following command status window.

```
      COMMAND STATUS

Command: OK  stdout: yes stderr: no

Before command completion, additional instructions may appear as follows.
```

```
[TOP]
cav      system 0true    root
cav      staff  1false   daemo
cav      bin    2true    root,bin
cav      sys    3true    root,bin,sys
cav      adm    4true    bin,adm
cav      uucp   5true    nuucp,uucp
cav      mail   6true
cav      security7true    root
cav      cron   8true    root
cav      printq 9true
cav      audit  10      true    root
cav      ecs    28      true
cav      nobody -2      false   nobody,lpd
[MORE...56]
```

b. If you select LDAP for the mode, SMIT displays the following command status window.

```
      COMMAND STATUS

Command: OK  stdout: yes stderr: no

Before command completion, additional instructions may appear as follows.

[TOP]
system 0true    root
taff  1false   daemo
bin    2true    root,bin
ys    3true    root,bin,sys
adm    4true    bin,adm
uucp   5true    nuucp,uucp
mail   6true
security7true    root
cron   8true    root
printq 9true
audit  10      true    root
ecs    28      true
nobody -2      false   nobody,lpd
[MORE...56]
```

Adding AIX® and LDAP groups on cluster nodes

To define a new group on AIX® systems, the **mkgroup** command. This command adds an entry for the new group to various system security files, including /etc/group and /etc/security/group.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

You can also define a new group on all AIX® cluster nodes and in LDAP from C-SPOC as described in the following procedure. The C-SPOC command performs some verification and then calls the AIX® **mkgroup** command on each cluster node to create the group you specify. If you want to add an LDAP group, use the **mkgroup -R LDAP** command.

If a group with the same name exists on a cluster node, the operation ends. By default, the C-SPOC command requires that the nodes in the PowerHA® SystemMirror® cluster must be powered on and accessible over the network; otherwise, the command is not run successfully and produces an an error.

To define a new LDAP or AIX® group on cluster nodes using the C-SPOC utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › Add a group to the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. Enter data in applicable fields to create the group account. The **Group Name** is a required field. You can also specify the group ID .

   > If you are adding a group to LDAP, you cannot edit all fields.

5. Press Enter. The C-SPOC command runs, creating a new group on all AIX® cluster nodes or in LDAP, depending on the mode you selected in step 4.

Changing characteristics of AIX® and LDAP groups in a cluster

In the AIX® operating system, you can change the attributes of a group by using the **chgroup** command or the SMIT interface.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

The **chgroup** command changes the user information stored in the **/etc/group** file and the **/etc/security/group** file.

You can change the attributes of a group on all cluster nodes from C-SPOC as described in the following procedure. This procedure runs the AIX® **chgroup** command on each cluster node.

To change group characteristics by using C-SPOC, you must meet the following requirements:

- All cluster nodes are accessible.
- The Cluster Communications daemon is running.
- A group with the name specified exists on all cluster nodes.

Optionally, you can force the C-SPOC command to continue processing even if it encounters an error on one of the cluster nodes.

To change the attributes of a group in LDAP or on all AIX® cluster nodes by using the C-SPOC utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › Change/Show Characteristics of a Group in the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. Specify the name of the group you want to change, and press Enter.
   Press F4 to obtain a list of groups from which to select. SMIT displays a list of the attributes of the specified group and their current values.
5. Change the value of any group attribute, and press Enter.

Result
The command runs, writing the new attribute value in the appropriate system security files on all cluster nodes or in LDAP.

Removing AIX® and LDAP groups from the cluster

To delete a group on an AIX® system, you must use the **rmgroup** command. This command removes the entry for the group from the /etc/group file and the /etc/security/group file. Users that are members of the group are not deleted.

> **Note:** The LDAP function is only available for PowerHA® SystemMirror® 7.1.1, or later.

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. Only the root user can remove an administrative group or a group with administrative users as members.

To remove a group from all cluster nodes, complete the steps in the following procedure. C-SPOC performs some clusterwide verification checks and then calls the AIX® **rmgroup** command to remove the group on each cluster node.

If a group with the name specified does not exist on one of the cluster nodes, the command reports a warning message but continues the operation on the other cluster nodes. By default, the command requires that all cluster nodes are powered up and accessible over the network; otherwise, the command fails with an error. Optionally, you can force the command to continue processing even if it encounters an error on one of the cluster nodes.

To remove a group from LDAP or AIX® cluster nodes by using the C-SPOC utility, complete the following steps:

1. From the command line, enter `smit cl_admin`.
2. In SMIT, select **Security and Users › Users in a PowerHA® SystemMirror® cluster › Remove a group to the Cluster**, and press Enter.
3. In the **Select an Authentication and registry mode** window, select the mode and press Enter.
4. Enter the name of the group you want to remove. Press the F4 to list available groups from which to select. After specifying the group name, press Enter.

Managing cluster security

These topics describe how to configure security options to protect your PowerHA® SystemMirror® cluster.

You can protect access to your PowerHA® SystemMirror® cluster by setting up security for cluster communications between nodes. PowerHA® SystemMirror® provides security for connections between nodes, with higher levels of security for inter-node communications provided through virtual private networks. In addition, you can configure authentication and encryption of the messages sent between nodes.

Configuring cluster security

PowerHA® SystemMirror® secures communications between cluster nodes for PowerHA® SystemMirror® operations in several different ways.

These methods include providing:

- Connection authentication for each new connection request
- (*Optional*) Message authentication
  Messages are signed on the sending node, and this signature is verified on the receiving node.
- (*Optional*) Message encryption.

Messages are encrypted on the sending node and decrypted on the receiving node, using a common, shared (symmetric) key.

A Cluster Communications daemon (**clcomd**) runs on each PowerHA® SystemMirror® node to transparently manage inter-node communications for PowerHA® SystemMirror®. This daemon consolidates communication mechanisms in PowerHA® SystemMirror® and decreases management traffic on the network. This communication infrastructure requires only one common communication path, rather than multiple TCP connections, between each pair of nodes.

The Cluster Communications daemon logs information about all attempted connections (those accepted and those rejected) to **clcomd.log**.

Although most components communicate through the Cluster Communications daemon, the following PowerHA® SystemMirror® components use a different mechanism for inter-node communications:

Component	Communication Method
Cluster Manager	RSCT
Cluster Information Program (Clinfo)	SNMP

Configuring PowerHA® SystemMirror® with IP security filter rules

You must enable specific ports for cluster commands and for cluster services to work correctly.

If you manually configure IP security filter rules, or if you use a tool such as AIX® Security Expert, which creates filter rules, you must ensure that those rules do not affect the ports that are used by PowerHA® SystemMirror®, Cluster Aware AIX®, and Reliable Scalable Cluster Technology (RSCT).

To use the IP security filter rules for cluster services, which you configured manually, complete the following steps:

1. From the command line, enter **smitty tcpip**.
2. In SMIT, select **Configure IP Security › Advanced IP Security Configuration › Configure IP Security Filter Rules › Add an IP Security Filter Rule** and press Enter.
3. From the **Add an IP Security Filter Rule** menu, enter the values for a single port according to the following table.

Valid port numbers and values for the Add an IP security filter rule menu in SMIT

Source port number / ICMP type	Rule action	Protocol	Source port / ICMP type operation	Description
0	permit	icmp	any	The clcomd daemon uses ICMP to identify a working IP address to connect to a node.
512	deny	all	le	Blocks all port numbers that are less than 512.
1023	permit	all	le	Opens all port numbers that are less than 1024.
6174	permit	all	eq	The **clinfo_client** daemon uses this port number for the clstat utility and other clinfo applications.
6175	permit	all	eq	The **clm_smux** daemon uses this port number for Simple Network Management Protocol (SNMP) smux peer operations.
6176	permit	all	eq	The **clinfo_deadman** daemon uses this port number for clinfo monitoring operations.
6180	permit	all	eq	The **emsvcs** command uses this port number for RSCT events.
6270	permit	all	eq	The **clsmuxpd** daemon uses this port number for SNMP operations.
12348	permit	all	eq	The **cthags** command uses this port number

Source port number / ICMP type	Rule action	Protocol	Source port / ICMP type operation	Description
				for RSCT group services.
16191	permit	all	eq	The **clcomd** daemon uses this port number during the migration process from a prior release of PowerHA® SystemMirror®

4. Repeat steps 1-3 for each port that is listed in the "Valid port numbers and values for the Add an IP security filter rule menu in SMIT" on page 290 table.

Standard security mode

In standard security mode, PowerHA® SystemMirror® authenticates requests for incoming connections by checking the source IP address, the port number, and user privilege.

Remote command execution for commands in **/usr/es/sbin/cluster** uses the principle of *least privileged*. This ensures that no arbitrary command can run on a remote node with root privilege. A select set of PowerHA® SystemMirror® commands are considered *trusted* and allowed to run as root; all other commands run as user *nobody*.

The dependency on **rsh** and the **~/.rhosts** file to configure host access has been eliminated. Although this file is optional, some commands external to PowerHA® SystemMirror® - for example user-defined event scripts and user programs - may still require an **~/.rhosts** file. PowerHA® SystemMirror® now relies on an internal PowerHA® SystemMirror® trusted host file, **/etc/cluster/rhosts** to authenticate PowerHA® SystemMirror® communications.

> **Note:** PowerHA® SystemMirror® does not use remote native AIX® remote execution (rsh) so you do not need to configure a ~/.rhosts file unless you intend to use Workload Partitions (WPAR) which have their own requirements on this file.

To manage inter-node communications, the Cluster Communications daemon requires a list of valid cluster IP labels or addresses to use. There are two ways to provide this information:

- Automatic node configuration
- Individual node configuration (more secure).

> **Note:** During discovery, each node that receives a connect request checks the **/etc/cluster/rhosts** file to ensure that the request is from a legitimate cluster node. The **smit.log** file indicates whether this file is missing or has incorrect entries.

Manually configuring /etc/cluster/rhosts file on individual nodes

For a more secure initial configuration, manually configure a **/etc/cluster/rhosts** file for PowerHA® SystemMirror® on each node before configuration.

The PowerHA® SystemMirror® installation creates this empty file with read-write permissions for root only. Ensure that each IP address/label is valid for the cluster. Otherwise, an error is logged in **smit.log** and **clcomd.log** .

To manually set up the **/etc/cluster/rhosts** file:

1. As root, open the file **/etc/cluster/rhosts** on a node.

2. Edit the file to add all possible network interface IP labels or addresses for each node. Put only one IP label or address on each line. Do *not* add any other characters or comments. The format of this file does *not* allow to have comments, additional lines, or characters in it, besides the IP labels.

Troubleshooting the Cluster Communications daemon

In some cases, if you change or remove IP addresses in the AIX® adapter configuration, and this takes place *after* the cluster has been synchronized, the Cluster Communications daemon cannot validate these addresses against the **/etc/cluster/rhosts** file or against the entries in the PowerHA® SystemMirror®'s Configuration Database, and PowerHA® SystemMirror® issues an error.

Or, you may obtain an error during the cluster synchronization.

In this case, you must update the information that is saved in the **/etc/cluster/rhosts** file on all cluster nodes, and refresh **clcomd** to make it aware of the changes. When you synchronize and verify the cluster again, **clcomd** starts using IP addresses added to PowerHA® SystemMirror® Configuration Database.

To refresh the Cluster Communications daemon, use:

```
refresh -s clcomd
```

Also, configure the **/etc/cluster/rhosts** file to contain all the addresses currently used by PowerHA® SystemMirror® for inter-node communication, and then copy this file to all cluster nodes.

Configuring message authentication and encryption

In addition to connection authentication, you can secure the messages sent through the Cluster Communications Daemon between cluster nodes by authenticating and encrypting those messages. You can use message encryption with message authentication, but you cannot use message encryption alone. Message authentication and encryption are disabled by default.

Both message authentication and message encryption rely on *secret key* technology. For authentication, the message is signed and the signature is encrypted by a key when sent, and the signature is decrypted and verified when received. For encryption, the encryption algorithm uses the key to make data unreadable. The message is encrypted when sent and decrypted when received.

Message authentication and encryption rely on Cluster Security (CtSec) Services in AIX®, and use the encryption keys available from Cluster Security Services. PowerHA® SystemMirror® message authentication uses message digest version 5 (MD5) to create the digital signatures for the message digest. Message authentication uses the following types of keys to encrypt and decrypt signatures and messages (if selected):

- Data encryption standard (DES)
- Triple DES
- Advanced encryption standard (AES).

The message authentication mode is based on the encryption algorithm. Your selection of a message authentication mode depends on the security requirements for your PowerHA® SystemMirror® cluster.

Authenticating and encrypting messages increases the overhead required to process messages and may impact PowerHA® SystemMirror® performance. Processing more sophisticated encryption algorithms may take more time than less complex algorithms. Message authentication and encryption are disabled by default.

The PowerHA® SystemMirror® product does not include encryption libraries. Message authentication and encryption rely on Cluster Security (CtSec) Services in AIX®, and use the encryption keys available from Cluster Security Services. Before you can use message authentication and encryption, the following AIX® filesets must be installed on each cluster node:

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

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

If you install the AIX® encryption filesets after you have PowerHA® SystemMirror® running, restart the Cluster Communications daemon to enable PowerHA® SystemMirror® to use these filesets. To restart the Cluster Communications daemon:

```
stopsrc -s clcomd
startsrc -s clcomd
```

If your configuration includes persistent labels, make sure that this configuration is synchronized before proceeding.

PowerHA® SystemMirror® provides a SMIT interface to configure message authentication and encryption.

To open the SMIT interface, enter:

```
smit cspoc
```

Select **Security and Users** > **PowerHA SystemMirror Cluster Security**.

> **Important:** Do not perform other cluster configuration activities while you are configuring message authentication and encryption for a cluster. Doing so may cause communication problems between nodes. Make sure that security configuration is complete and the cluster synchronized before performing other configuration tasks.

For more information about the available authentication and encryption options, see the CAA documentation at Configuring cluster security.

Managing keys

PowerHA® SystemMirror® cluster security uses a shared common (symmetric) key. This means that each node must have a copy of the *same* key for inter-node communications to be successful. You control when keys change and how keys are distributed.

You can allow PowerHA® SystemMirror® to distribute a key for you, or you can manually copy a key to each node in a cluster. Copying a key to each cluster node can provide a higher level of security than having PowerHA® SystemMirror® distribute the key, depending on the method you use to copy the key to the cluster nodes.

Cluster synchronization does *not* update keys and does *not* distribute keys among nodes.

Location of keys

On each node, a key is stored in the **/etc/cluster/security** directory. The name of the key identifies the encryption type selected:

- key_md5_des
- key_md5_3des
- key_md5_aes

When to generate and distribute a key

Generate and distribute a key after:

- Enabling message authentication
- Changing the configuration for message authentication.

Also, change the key in accordance with the security policies for your organization.

> **Note:** Communication between cluster nodes requires that all nodes have active copies of the same key. You activate a new key after you distribute the key to each node in the cluster.

Configuring message authentication and encryption using Automatic Key Distribution

Make sure that the cluster is synchronized before you start to configure message authentication and encryption. This ensures that cluster nodes can communicate with each other.

Step 1: Enable Automatic Key Distribution on each node

The first step is to enable Automatic Key Distribution on each node.

To make sure that you can distribute a new key through PowerHA® SystemMirror®, enable **Automatic Key Distribution** on each node in the cluster before:

- You change the message authentication mode
- You try to automatically distribute a key to cluster nodes.

To enable key distribution on each cluster node:

1. Enter `smit cspoc`
2. In SMIT, select **Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Enable/Disable Automatic Key Distribution** and press Enter.
   The **Enable/Disable Automatic Key Distribution** panel appears.
3. For **Enable Key Distribution**, select **Yes**.
4. Repeat step 1 through step 3 on the other nodes in the cluster.

Step 2: Enable or change message authentication

Step 2 is to enable or change message authentication and encryption *from one cluster node*.

To enable or change message authentication:

1. Enter `smit cspoc`
2. In SMIT, select **Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Configure Message Authentication Mode** and press Enter.
   The **Configure Message Authentication Mode** panel appears.
3. Enter field values as follows:

Configure Message Authentication Mode	
Field	**Value**
Message Authentication Mode	Select one of the following modes: **MD5_DES** The MD5 algorithm is used for message digest (signature) and the DES algorithm is used for signature encryption. **MD5_3DES** The MD5 algorithm is used for message digest (signature) and the triple DES algorithm is used for signature encryption. **MD5_AES** The MD5 algorithm is used for message digest (signature) and the AES algorithm is used for signature encryption. **None** This indicates that neither message authentication nor message encryption is being used.
Enable Encryption	Select **Yes** to *enable* message encryption for messages sent between PowerHA® SystemMirror® nodes. Select **No** to *disable* message encryption for messages sent between PowerHA® SystemMirror® nodes.

4. Press Enter.

Step 3: Generate and distribute a key from one node

Step 3 is to generate and distribute a key from one node.

If you are enabling or changing message authentication and encryption, complete this procedure on the same node where you completed Step 2: Enable or change message authentication.

To generate a new key and distribute it through PowerHA® SystemMirror®:

1. From the **System Management (C-SPOC)** menu, select **Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Generate/ Distribute a Key** and press Enter.
   The **Generate/Distribute a Key** panel appears.
2. Enter field values as follows:

Generate/Distribute a Key fields	
Field	**Value**
Type of Key to Generate	Lists the active authentication mode
Distribute a Key	**Yes**

3. When prompted, confirm that you want PowerHA® SystemMirror® to distribute a key. This information is written to the **/var/hacmp/clcomd/clcomd.log** file.

Result

> **Note:** If for some reason SMIT cannot copy the key to cluster nodes, copy the key file to diskette and copy it to the node.

Step 4: Activate the key on each node

After you distribute a new key to each node in the cluster, on the node from which you distributed the key, active it for *all cluster nodes*. This action makes it possible for cluster nodes to communicate with each other.

To activate a new key:

1. In SMIT, select **System Management (C-SPOC) > Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Activate the New Key on All Cluster Nodes** and press Enter.
   SMIT displays Are you sure?
2. Press Enter to activate the key on all cluster nodes.

Result
The **Command Status** panel lists the nodes on which the key is active.

Step 5: Synchronize the cluster

Synchronize the cluster configuration.

For information about synchronizing the cluster, see Verifying and synchronizing a PowerHA® SystemMirror® cluster.

Step 6: Disable Automatic Key Distribution on each node

After you distribute a key to cluster nodes through PowerHA® SystemMirror® and activate the key, disable **Automatic Key Distribution** *on each node* in the cluster.

Important: Do *not* leave **Automatic Key Distribution** enabled. Doing so might allow an unwelcome user to distribute a spurious key to cluster nodes, which would compromise cluster security.

To disable automation key distribution from each cluster node:

1. Enter `smit cspoc`
2. In SMIT, select **Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Enable/Disable Automatic Key Distribution** and press Enter.
   The **Enable/Disable Automatic Key Distribution** panel appears.
3. For **Enable Key Distribution**, select **No**.

Configuring message authentication and encryption using Manual Key Distribution

Synchronize the cluster before you start to configure message authentication and encryption. This ensures that cluster nodes can communicate with each other.

Step 1: Enable or change message authentication and encryption

You enable message authentication and encryption from one cluster node.

To enable or change message authentication:

1. Enter `smit cspoc`
2. In SMIT, select **Security and Users > PowerHA® SystemMirror® Cluster Configuration > Configure Message Authentication Mode and Key Management > Configure Message Authentication Mode** and press Enter.
   The **Configure Message Authentication Mode** panel appears.
3. Enter field values as follows:

Configure Message Authentication Mode fields	
Field	**Value**
Message Authentication Mode	Select one of the following modes: **MD5_DES** The MD5 algorithm is used for message digest (signature) and the DES algorithm is used for signature encryption. **MD5_3DES** The MD5 algorithm is used for message digest (signature) and the triple DES algorithm is used for signature encryption. **MD5_AES** The MD5 algorithm is used for message digest (signature) and the AES algorithm is used for signature encryption. **None** This indicates that neither message authentication nor message encryption is being used.
Enable Encryption	Select **Yes** to *enable* message encryption for messages sent between PowerHA® SystemMirror® nodes. Select **No** to *disable* message encryption for messages sent between PowerHA® SystemMirror® nodes.

4. Press Enter.

Step 2: Distribute a new key by copying it to cluster nodes

Ensure that you distribute the same encryption key to each cluster node; otherwise, PowerHA® SystemMirror® cannot communicate between cluster nodes.

To generate a new key and copy it to other cluster nodes:

1. On the node where you want to create a key, enter `smit hacmp`
2. In SMIT, select **System Management (C-SPOC) > Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Generate/ Distribute a Key** and press Enter.
   The **Generate/Distribute a Key** panel appears.
3. Enter field values as follows:

Generate/Distribute a Key fields	
Field	**Value**
Type of Key to Generate	Lists the active authentication mode
Distribute a Key	**No**

4. Copy the key file from the node where the key was generated to each node in the PowerHA® SystemMirror® cluster.
   On each node, a key is stored in the **/usr/es/sbin/cluster/etc** directory. The name of the key identifies the encryption type selected:
   - key_md5_des
   - key_md5_3des
   - key_md5_aes

Result
You can copy the file to diskette and then go to each node and copy the key file to the appropriate directory, or you can use a remote copy command such as **ftp** or **rcp** .

> **Important:** A key may already be present on each node, make sure that you copy the key to each node. The new key overwrites the older one if the keys are of the same type, for example if the key is for 3DES. *If the keys on the nodes do not match, PowerHA® SystemMirror® does not function.*

Step 3: Activate the key on each node

After you distribute a new key to each node in the cluster, from one node you activate the key *on all cluster nodes* to make it possible for cluster nodes to communicate with each other. If you enabled or changed the message authentication mode, you should activate the key from the cluster node where you made that configuration change.

To activate a new key:

1. Enter `smit cspoc`
2. In SMIT, select **Security and Users > PowerHA® SystemMirror® Cluster Security > Configure Message Authentication Mode and Key Management > Activate the New Key on All Cluster Nodes** and press Enter.
   SMIT displays Are you sure?
3. Press Enter to activate the key on all cluster nodes.

Result
The **Command Status** panel lists the nodes on which the key is active.

Step 4: Synchronize the cluster

Synchronize the cluster configuration.

For information about synchronizing the cluster, see Verifying and synchronizing a PowerHA® SystemMirror® cluster.

PowerHA® SystemMirror® federated security

To successfully implement PowerHA® SystemMirror® federated security, you must use role-based access control (RBAC) and Encrypted File System (EFS) with PowerHA® SystemMirror® using Lightweight Directory Access Protocol (LDAP), as a centralized information base for clusters.

> **Note:** The PowerHA® SystemMirror® federated security features are available in PowerHA® SystemMirror® Version 7.2.4, and later.

With federated security, you can complete the following tasks.

- Configure and manage an IBM® or non-IBM® LDAP server as a centralized information base.
- Configure and manage a peer-to-peer IBM® LDAP server.
- Configure and manage the LDAP client for all the nodes of the cluster.
- Create and manage a highly available EFS file system.
- Create and manage Role Based Access Control (RBAC) roles for users and groups. You can use the following roles to control which commands can be run by different sets of users of PowerHA® SystemMirror®.
  - `ha_admin` (for administrator): A user with this role is a PowerHA® SystemMirror® administrator, who can perform all PowerHA® SystemMirror® operations, such as configuring clusters, verifying and synchronizing clusters, and starting and stopping cluster services.
  - `ha_op` (for operations): A user with the `ha_op` role can perform all PowerHA® SystemMirror® operations except the configuration changes. The PowerHA® SystemMirror® operations include starting and stopping cluster services, resource group movements and verification.
  - `ha_mon` (for viewer): A user with the `ha_mon` role can perform only querying and listing related operations.
  - `ha_view` (for monitor): A user with the `ha_view` role cannot perform any PowerHA® SystemMirror® related operations. The user can only view the PowerHA® SystemMirror® log files.

> **Other:**
> - In PowerHA® SystemMirror® Version 7.2.4 or later, a user must belong to the `ldapha` group to use PowerHA® SystemMirror® commands.
> - If RBAC is enabled for the PowerHA® SystemMirror® cluster, you must remove the LDAP server and client configuration from PowerHA® SystemMirror® before you migrate PowerHA® SystemMirror® to a later version. You must configure the LDAP server and client after completing the migration operation.

Planning for federated security

Before you can use the features of federated security, you must plan for its implementation in your environment.

To use the features of federated security, your environment must meet the following requirements:

- A cluster must be configured before using LDAP and RBAC.
- PowerHA® SystemMirror® services must be started before you can use the EFS feature for shared file system mode.

- If you want to use a non-IBM® LDAP server you must load the schema. You can configure the rsh service to automatically load the schema, which is how it is loaded for the AIX® operating system. If the rsh service is not configured, you must manually load the schema. For more information on how manually load schemas, see Extending non-IBM® LDAP servers to support full AIX® functionality.

System requirements

Your environment must have the following hardware and software to implement the federated security features.

- The AIX® operating system must be at one of the following technology levels:
  - IBM® AIX® 7.1 with Technology Level 1, or later
  - IBM® AIX® 7.2, or later
- Your environment must be running PowerHA® SystemMirror® Version 7.2.4, or later
- Your environment must be running IBM® LDAP 6.4, or later
- Your environment must be running one of the following versions of Microsoft™ Windows™ Server:
  - Microsoft™ Windows™ Server 2003 Active Directory
  - Microsoft™ Windows™ Server 2008 Active Directory
  - Microsoft™ Windows™ Server 2008 R2 Active Directory
- Your environment must be running Services for UNIX® (SFU) 3.5, or later, or the Subsystem for UNIX®-based Applications (SUA)

Installing federated security

You must install the file sets before you can use the federated security features.

To install all the federated security features, complete the following steps:

1. Install PowerHA® SystemMirror® Version 7.2.4, or later.
2. Install the LDAP client file sets.
3. Install GSkit file sets on all nodes in the cluster. Verify that the installation completed successfully. These file sets come with the LDAP packages.

   > **Note:** If you are using an IBM® LDAP server, you must install the LDAP server file sets, DB2® file sets, and GSkit file sets that come with the LDAP server package. For more information, see Setting up an LDAP client.

4. Verify that the `clic.rte` file set is installed on all the nodes in the cluster.
5. Verify that the `expect.base` file set is installed on all the nodes in the cluster.

Configuring federated security

After you have installed the required file sets, you can configure the federated security features.

Configuring LDAP servers

The Lightweight Directory Access Protocol (LDAP) defines a standard method for accessing and updating information in a directory, either locally or remotely in a client/server model.

To configure an existing LDAP server, complete the following steps:

1. From the command line enter, `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › LDAP › LDAP server configuration › Add an existing LDAP Server**, and press Enter.
3. Complete the following fields.

Fields for adding an existing LDAP server	
Fields	**Value**
LDAP servers	Enter the host name of LDAP servers that are currently running. If you have more than one host name, you must separate each host name with a comma.
Bind DN	Enter the LDAP administrator domain name (DN).
Bind password	Enter the LDAP administrator password.
Suffix/Base DN	Enter the base DN that will be the root for all other DNs that store information in the LDAP directory for the cluster.
Server port number	Enter the server port number.
SSL key path	Enter the SSL key path for the client.
SSL password	Enter the SSL key password for the client.

4. Verify all fields are correct, and press Enter.

> Verify that you have correctly configured the SSL keys between the server and the client. Also verify that the Microsoft™ Windows™ Server Active Directory is communicating with PowerHA® SystemMirror®.

Configuring peer-to-peer LDAP servers

If a prior LDAP server configuration doesn't exist, you can configure a new peer-to-peer LDAP server to replicate the configuration. Only LDAP servers based on the AIX® operating system are created.

In a peer-to-peer replication, several servers act as primary servers for directory information. Each primary server is responsible for updating other primary servers and replication servers. This process is known as peer replication and can help improve performance, availability, and reliability.

To configure a peer-to-peer LDAP server, complete the following steps:

1. From the command line enter, `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › LDAP › LDAP server configuration › Configure a new peer-to-peer LDAP server**, and press Enter.
3. Complete the following fields.

Fields for configuring a new peer-to-peer LDAP server	
Fields	**Value**
Host names	Press **F4** to select one or more host names from the list that you want to configure. You can select 2 - 6 nodes.
LDAP administrator DN	Enter the LDAP administrator domain name (DN).
LDAP administrator password	Enter the LDAP administrator password.
Schema type	The default value is `rfc2307aix`. You cannot edit this field.
Suffix/Base DN	Enter the base DN, that is the root for all other DNs that store information in the LDAP directory for the cluster.
Server port number	Enter the server port number.
SSL key path	Enter the SSL key path for the server.
SSL password	Enter the SSL key password for the server.

Fields	Value
Version	Displays the LDAP version. You cannot edit this field.
DB2® instance password	Enter a password for the DB2® instance that is created by the directory instance.
Encryption seed to generate key stash file	Enter a minimum of 12 alphanumeric characters to generate key stash files for LDAP.

4. Verify that all fields are correct, and press Enter.

> You can manually configure the SSL keys, or you can use PowerHA® SystemMirror® to configure the SSL keys.

Configuring LDAP clients

You must set up your LDAP client before you can configure it.

Ensure that all cluster nodes are up and reachable.
To configure an LDAP client, complete the following steps.

1. From the command line enter, `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › LDAP › LDAP client configuration › Configure LDAP client**, and press Enter.
3. Complete the following fields:

Fields for configuring LDAP client	
Fields	**Value**
LDAP servers	Press **F4** to select the LDAP servers that are configured in your environment.
Bind DN	Displays the bind DN. You cannot edit this field.
Bind password	Enter the bind DN password.
Authentication type	Press **F4** to select the authentication type. The default value is **ldap_auth**.
Suffix/Base DN	Enter the base DN, which is the root for all other DNs that store information in the LDAP directory for the cluster.
Server port number	Enter the server port number.
SSL key path	Enter the SSL key path to store the client key.
SSL password	Enter the SSL key password for the client.

4. Verify that all fields are correct, and press Enter. The configuration process can take a long time to complete.

> The permissions of the `root` user and user groups are preserved. The execute permission for all PowerHA® SystemMirror® commands on all nodes in the cluster is removed for other users.

Creating an Encrypted File System

The Encrypted Files System (EFS) enables individual users on the system to encrypt their data on the J2 file system through their individual key stores.

To create a highly available EFS, complete the following steps:

1. From the command line enter, `smitty sysmirror`.

2. In SMIT, select **System Management (C-SPOC) › Security and Users › EFS management › Enable EFS Keystore**, and press Enter.
3. Complete the following fields.

Fields for enabling EFS Keystore	
Fields	**Value**
EFS keystore mode	Press **F4** to select either LDAP or a shared file system from a list.
EFS admin password	Enter the EFS administrator password.
Volume group for keystore	Press **F4** to select a concurrent volume group in the cluster from a list. This field is disabled if you selected **LDAP** in the **EFS keystore** field.
Service IP	Press **F4** to select the service IP in the cluster from a list. This field is disabled if you selected **LDAP** in the **EFS keystore** field.

4. Verify that all fields are correct, and press Enter.

Managing PowerHA® SystemMirror® federated security

You can use PowerHA® SystemMirror® to manage the LDAP server and the Encrypted Files System (EFS).

Managing LDAP servers

To manage LDAP servers, you can use AIX® commands and LDAP management commands.

To change LDAP server settings, complete the following steps:

1. Remove the LDAP client.
2. Remove the LDAP server.
3. Create a new LDAP server with the changed parameters.
4. Configure the LDAP client.
5. Run verification and synchronization on the LDAP client.

Note: Before you change LDAP server parameters, you must disable all federated security features.

Managing EFS

To manage EFS, complete the following steps:

1. From the command line, enter `smitty sysmirror`
2. In SMIT, select **System Management (C-SPOC) › Security and Users › EFS management in cluster › Change / Show EFS characteristic**, and press Enter.
3. Change the applicable fields. You cannot change the **Password** field.
4. Verify that the modifications you made are correct, and press Enter.

Removing PowerHA® SystemMirror® federated security

You can use PowerHA® SystemMirror® to remove LDAP servers, LDAP clients, and EFS from the cluster.

Removing LDAP servers

> **Note:** When you are removing any of the following federated security features, read any warning or error messages carefully and verify that the removal will not cause problems for your cluster environment.

To remove LDAP servers from the cluster, complete the following steps:

1. From the command line, enter `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › Configure LDAP › LDAP server configuration for cluster › Delete the LDAP server from the cluster**, and press Enter.

   > **Note:** Completing this task removes the entries from the PowerHA® SystemMirror® ODM. The data will still be available on the LDAP server if you want to configure it again in the future.

Removing LDAP clients

To remove LDAP clients from the cluster, complete the following steps:

1. From the command line, enter `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › Configure LDAP › LDAP server configuration for cluster › Delete the LDAP clients from the cluster**, and press Enter.

> **Note:** The permissions of the `root` user and user groups are preserved. The execute permission for all PowerHA® SystemMirror® commands on all nodes in the cluster is removed for other users.

Removing EFS

To remove EFS management from the cluster, complete the following steps.

> **Note:** Before you remove EFS, you can create a backup of the EFS keystore in case you want to reuse it in the future.

1. From the command line, enter `smitty sysmirror`.
2. In SMIT, select **System Management (C-SPOC) › Security and Users › EFS management in cluster › Delete EFS keystore**, and press Enter.

   > **Note:** You must also remove the EFS from the resource group and the file system.

Troubleshooting PowerHA® SystemMirror® federated security

You can view failure messages from log files in the `/var/hacmp/log/fsec` directory.

Use the following table to help troubleshoot any problems with federated security.

Troubleshooting federated security	
Problem	**Solution**
You are configuring a non-IBM LDAP server, and the SMIT commands are not working. The following error message is displayed: `Not able to communicate with server`.	Verify that the rsh function and the rcp function are working correctly from the source system to the target system.

Problem	Solution
The **Security and Users** SMIT panel for LDAP configuration is not working correctly. The following error message is displayed: `Could not complete the command`.	Try the following solution: • Verify that the LDAP client is able to communicate with the LDAP server. • Correct any errors in the **cspoc.log** file. • Run verification and synchronization on the cluster.

Saving and restoring cluster configurations

You can use the cluster snapshot utility to save and restore cluster configurations. The cluster snapshot utility allows you to save to 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, provided the cluster is configured with the requisite hardware and software to support the configuration.

This process is called applying a snapshot.

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.

PowerHA® SystemMirror® captures all the Cluster Aware AIX (CAA) tunables and customer security preferences in the cluster snapshot database and restores it at a later time.

> **Note:** You can *not* use the cluster snapshot facility in a cluster concurrently running different versions of PowerHA® SystemMirror®.

By default, PowerHA® SystemMirror® does *not* collect cluster log files when you create the cluster snapshot. Cluster snapshots are used for recording the cluster configuration information, whereas cluster logs only record the operation of the cluster and *not* the configuration information. Skipping the log collection reduces the size of the snapshot and speeds up running the snapshot utility. The size of the cluster snapshot depends on the configuration. For instance, a basic two-node configuration requires roughly 40 KB.

> **Note:** You can change the default to collect cluster log files using SMIT if you need logs for problem reporting. This option is available under the SMIT menu **Problem Determination Tools › Log Viewing and Management › Collect Cluster log files for Problem reporting**. It is recommended to use this option only if IBM® support personnel request logs.

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.

Information saved in a cluster snapshot

The primary information saved in a cluster snapshot is the data stored in the PowerHA® SystemMirror® Configuration Database classes (such as PowerHA® SystemMirror®cluster, PowerHA® SystemMirror®node, PowerHA® SystemMirror®network, and PowerHA® SystemMirror®daemons). This information is used to re-create the cluster configuration when a cluster snapshot is applied to nodes installed with PowerHA® SystemMirror®.

The cluster snapshot does not save any user-customized scripts, applications, or other non-PowerHA® SystemMirror® configuration parameters. For example, the names of application controllers and the locations of their start and stop scripts are stored in the PowerHA® SystemMirror®server Configuration Database 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 that is outside the scope of PowerHA® SystemMirror®. For instance, the facility saves the names of shared file systems and volume groups; however, other details, such as NFS options or LVM mirroring configuration are not saved.

If you moved resource groups using the Resource Group Management utility **clRGmove**, once you apply a snapshot, the resource groups return to behaviors specified by their default nodelists.

To investigate a cluster after a snapshot has been applied, run **clRGinfo** to view the locations and states of resource groups.

> **Note:** You can reset cluster tunable values using the SMIT interface. PowerHA® SystemMirror® creates a cluster snapshot, prior to resetting. After the values have been reset to their defaults, you can apply the snapshot and return to customized cluster settings, if needed.

Format of a cluster snapshot

The cluster snapshot utility stores the data it saves in two separate files created in the directory **/usr/es/sbin/cluster/snapshots**: the ODM data file and the Cluster state information file.

ODM Data File (.odm)

This file contains all the data stored in the PowerHA® SystemMirror® Configuration Database object classes for the cluster. This file is given a user-defined basename with the .odm file extension. Because the Configuration Database information is largely the same on every cluster node, the cluster snapshot saves the values from only one node.

Cluster State Information File (.info)

This file contains the output from standard AIX® and PowerHA® SystemMirror® system management commands. This file is given the same user-defined basename file with the **.info** file extension. Output from custom snapshot methods is appended to this file.

Cluster snapshot ODM data file

The cluster snapshot Configuration Database 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 PowerHA® SystemMirror® Configuration Database 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 Configuration Database data file showing some of the ODM stanzas that are saved.

```
<VER
1.0
</VER

<DSC
My Cluster Snapshot
</DSC

<ODM

PowerHA SystemMirror cluster:
id = 97531
name = "Breeze1"
nodename = "mynode"
sec_level = "Standard"
last_node_ids = "2,3"
```

```
highest_node_id = 3
last_network_ids = "3,6"
highest_network_id = 6
last_site_ides = " "
highest_site_id = 0
handle = 3
cluster_version = 5
reserved1 = 0
reserved2 = 0
wlm_subdir = " "

PowerHA SystemMiror node:
name = "mynode"
object = "VERBOSE_LOGGING"
value = "high"
.
.
.
</ODM
```

clconvert_snapshot utility

You can run **clconvert_snapshot** to convert cluster snapshots from a release supported for upgrade to a recent PowerHA® SystemMirror® release. The **clconvert_snapshot** is not run automatically during installation, and you must always run it 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 PowerHA® SystemMirror® version from which you are converting in order to run this utility.

For more information on the **clconvert_snapshot** utility, refer 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.

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Configure Custom Snapshot Method › Add a Custom Snapshot Method** and press Enter.
3. Enter field values as follows:

Add a Custom Snapshot Method fields	
Field	**Value**
Custom Snapshot Method Name	A name for the custom snapshot method you would like to create.
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.

Result
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 run 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 SMIT interface.

To change or remove a custom snapshot method, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. In the SMIT interface, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Configure Custom Snapshot Method** and press Enter.
3. Depending on the task you want to complete, select **Change/Show a Custom Snapshot Method** or **Remove a Custom Snapshot Method**.
4. Complete all required fields and press Enter.

Creating a snapshot of the cluster configuration

You can initiate cluster snapshot creation from any cluster node. You can create a cluster snapshot on a running cluster. The cluster snapshot facility retrieves information from each node in the cluster. Accessibility to all nodes is required and the snapshot is stored on the local node.

To create a cluster snapshot, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Create a Snapshot of the Cluster Configuration** and press Enter.
3. Enter field values as follows:

Create a Cluster Snapshot of the Cluster Configuration fields

Field	Value
Cluster Snapshot Name	The name you want for the basename for the cluster snapshot files. The default directory path for storage and retrieval of the snapshot is **/usr/es/sbin/cluster/snapshots**. You can specify an alternate path using the SNAPSHOTPATH 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.
Save Cluster Log Files in a Snapshot	The default is **No**. If you select **Yes**, PowerHA® SystemMirror® collects cluster log files from all nodes and saves them in the snapshot. Saving log files can significantly increase the size of the snapshot.
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.

Restoring the cluster configuration from a snapshot

Restoring a cluster snapshot overwrites the data in the existing PowerHA® SystemMirror® Configuration Database classes on all nodes in the cluster with the new Configuration Database data contained in the snapshot. You can restore a cluster snapshot from any cluster node.

You cannot restore a cluster from a snapshot if cluster services are active and you are using PowerHA® SystemMirror® 7.1.2, or later.

> **Note:** Only the information in the **.odm** file is applied. The **.info** file is not needed to restore a snapshot.

Restoring a cluster snapshot may affect PowerHA® SystemMirror® Configuration Database objects and system files as well as user-defined files.

- If cluster services are inactive on all cluster nodes, restoring the snapshot changes the Configuration Database data stored in the system default configuration directory (DCD).

- If cluster services are active on the local node, restoring a snapshot triggers a cluster-wide dynamic reconfiguration event.

During dynamic reconfiguration, in addition to synchronizing the Configuration Database data stored in the DCDs on each node, PowerHA® SystemMirror® replaces the current configuration data stored in the active configuration directory (ACD) with the updated configuration data in the DCD. The snapshot becomes the active configuration.

> **Note:** A cluster snapshot used for dynamic reconfiguration might contain changes to the cluster topology and to the cluster resources. You can change both the cluster topology and cluster resources in a single dynamic reconfiguration event.

The snapshot restoration process for a cluster might fail if the repository disk changed since the snapshot was created. The following conditions cause the snapshot restoration process to fail:

- The repository disk is corrupted or fails and must be physically replaced.
- The snapshot of one cluster is used to restore a node that has a different repository disk.

To restore a cluster by using a snapshot, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT interface, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Restore the Cluster Configuration From a Snapshot** and press Enter.
   SMIT displays the **Cluster Snapshot to Apply** panel 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 restore and press Enter. SMIT displays the panel.
4. From **Restore the Cluster Configuration From a Snapshot** panel, complete the following fields:

Restore the Cluster Configuration From a Snapshot fields

Field	Description
Cluster Snapshot Name	Displays the current basename of the cluster snapshot. This field is not editable.
Cluster Snapshot Description	Displays the text stored in the description section of the snapshot files. This field is not editable.
Un/Configure Cluster Resources?	If cluster services are running and you are using PowerHA® SystemMirror® 7.1.2, or later, changes you make to this field are not applied. If you set this field to **Yes**, PowerHA® SystemMirror® changes the definition of the resource in the Configuration Database and performs any configuration triggered by the resource change. For example, if you remove a file system, PowerHA® SystemMirror® removes the file system from the Configuration Database and also unmounts the file system. By default, this field is set to **Yes**. If you set this field to **No**, PowerHA® SystemMirror® changes the definition of the resource in the Configuration Database but does not perform any configuration processing that the change may require. For example, a file system would be removed from the PowerHA® SystemMirror® cluster definition but would not be unmounted. This processing must be performed by PowerHA® SystemMirror® during a fallover. PowerHA® SystemMirror® attempts to limit the impact on the resource group when a component resource is changed. For example, if you add a file system to the resource group that already includes the underlying volume group as a resource, PowerHA® SystemMirror® 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.

Field	Description
	Cluster clients experience an interruption in related services while the dynamic reconfiguration is in progress.
Force apply if verify fails?	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. By default, this field is set to **No**. If you want synchronization to proceed even if verification fails, set this value to **Yes**.

5. Verify that all fields are correct and press Enter.

Result

> **Note:** In some cases, the verification uncovers errors that do not cause the synchronization to fail. PowerHA® SystemMirror® 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 do not interfere with the synchronization.

If the restore process fails or you want to go back to the previous configuration for any reason, you can reapply an automatically saved configuration.

If you create a cluster snapshot and make a dynamic automatic reconfiguration (DARE) change to a working cluster, such as removing and then re-adding a network, the snapshot may fail due to naming issues. For example, the following steps would make a snapshot fail:

1. Start the cluster.
2. Create a snapshot.
3. Remove a network dynamically.
4. Add a network dynamically using the same name as the one that was removed in step 3.
5. Attempt to apply snapshot from step 2.

However, if you use a different network name in step 4 than the network that was removed, you can apply the snapshot successfully. The problem is that a different network ID is used when the network is added back into the cluster.

AFTER COMPLETING THE TASK
Before the new configuration is applied, the cluster snapshot facility automatically saves the current configuration in a snapshot called **~snapshot.** *n* **.odm**, where **n** is 1 (the most recent), 2, or 3. The saved snapshots are cycled so that only three generations of snapshots exist. If the restore process fails or you want to go back to the previous configuration for any reason, you can reapply the saved configuration. The saved snapshot is stored in the directory specified by the SNAPSHOTPATH environment variable.

Changing a snapshot of the cluster configuration

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:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Change/Show a Snapshot of the Cluster Configuration** and press Enter.
   SMIT displays the **Change/Show a Snapshot of the Cluster Configuration** panel with a list of all the cluster snapshots that exist in the directory specified by SNAPSHOTPATH.
3. Select the cluster snapshot to change and press Enter.
4. Enter field values as follows:

Cluster snapshot fields	
Field	**Value**
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 snapshot of the cluster configuration

Removing a cluster snapshot deletes both of the ASCII files (**.odm** and **.info**) 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. From the command line, enter `smit sysmirror`
2. In SMIT, select **Cluster Nodes and Networks › Manage the Cluster › Snapshot Configuration › Remove a Snapshot of the Cluster Configuration** and press Enter.
   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 to remove and press Enter.

Result
The cluster snapshot facility deletes the files in the snapshot directory that are associated with that snapshot.

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, a PowerHA® SystemMirror® cluster can provide nearly continuous availability, interrupted only by scheduled, necessary maintenance.

These topics are a collection of information describing the issues and procedures involved in keeping a cluster running as closely as possible to a 7 X 24 basis.

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 PowerHA® SystemMirror® cluster is a distributed operating system environment. Therefore, maintaining a PowerHA® SystemMirror® 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 may affect other nodes, but this may not be apparent until fallover 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 topic.

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 fallovers or cluster downtime at your convenience as necessary to deal with any impending issues.

Planning for 7x24 maintenance

Carefully planning the original installation of your cluster goes a long way toward making cluster maintenance easier. A well-configured and customized cluster is the first step to 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. Once the cluster is installed and running, you need to handle any failures as quickly and as automatically as possible. Planning for runtime failure recovery helps ensure that PowerHA® SystemMirror® for AIX® does all that it is capable of doing to keep your critical resources online.

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 PowerHA® SystemMirror® cluster. Although PowerHA® SystemMirror® 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 PowerHA® SystemMirror® and AIX® tools available to customize some or all of the following:

- Hardware error notification
- Hardware failure notification
- Cluster event notification
- Pre-event and post-event recovery actions
- Network failure escalation
- ARP cache refresh
- Pager notification
- Application controller scripts.

It is highly recommended that you maintain a test cluster as well as your production cluster. Before you make any major change to the production cluster, you can test the procedure on the test cluster.

Customizing AIX® error notification of hardware errors

Customizing notification when you configure the cluster is a good preventive measure.

Using the PowerHA® SystemMirror® Automatic Error Notification SMIT panels, you can turn on automatic error notification for selected hard, non-recoverable error types, such as disk, disk adapter. All disks defined as PowerHA® SystemMirror® resources, and disks in the rootvg and PowerHA® SystemMirror® volume groups and file systems 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.

> **Note:** Most of these errors return Notification only.

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 PowerHA® SystemMirror® automatic error notification utility. The following lists show other types of errors you may want to address. For each device monitored, you can determine an additional action other than notification, such as:

- Stop cluster services and move resource groups to another node.
- Initiate® a custom recovery action such as reconfiguration for a failed device using an alternative device.

Hardware Errors Handled by PowerHA® SystemMirror® Auto-Error Notification	
Error	**Description**
DISK_ERR2	Permanent physical disk error (known error)
DISK_ERR3	Permanent physical disk error, adapter detected (known error)

Error	Description
SCSI_ERR1	Permanent SCSI adapter hardware error (known error)
SCSI_ERR3	Permanent SCSI adapter microcode error (known error)
SCSI_ERR5	Temporary SCSI bus error
SCSI_ERR7	Permanent unknown system error
SCSI_ERR9	Potential Data loss condition
SDA_ERR1	Adapter hardware error condition
SDA_ERR3	Permanent unknown system error
SDC_ERR1	Controller/DASD link error
SDC_ERR2	Controller hardware error
DISK_ARRAY_ERR2	Permanent disk operation error (disk failure)
DISK_ARRAY_ERR3	Permanent disk operation error (disk failure)
DISK_ARRAY_ERR5	Permanent disk operation error (disk failure)
SCSI_ARRAY_ERR2	SCSI hardware error

Disk and Adapter Errors Not Covered by PowerHA® SystemMirror® Auto-Error Notification	
Error	**Description**
LVM_MISSPVADDED	PV defined as missing (unknown error)
LVM_SA_WRT	PV defined as missing (unknown error)
LVM_SA_PVMISS	Failed to write VGSA (unknown error)

Disk Array Errors Not Covered by PowerHA® SystemMirror® Auto-Error Notification	
Error	**Description**
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

You may have additional devices critical to your operation that are not supported by PowerHA® SystemMirror® for AIX®. You can set up AIX® error notification to monitor microcode errors for those devices 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.

Customizing Application controller scripts

When customizing Application controller scripts, there are several key ideas to keep in mind.

These topics include:

- Define a PowerHA® SystemMirror® application controller for each node that supports applications requiring recovery.
- Applications must be started and stopped 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.

Optionally, you can also include applications in different resource groups and establish dependencies between resource groups. For more information, see Adding resources and attributes to resource groups.

- Check for dependencies between nodes. For example, a process on node1 may not 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.

There is also a plug-in for print queues in the **/usr/es/sbin/cluster/plugins/printserver** directory.

Application monitoring

You can monitor a set of applications that you define through the SMIT interface.

You can configure multiple application monitors and associate them with one or more application controllers. By supporting multiple monitors per application, PowerHA® SystemMirror® can support more complex configurations. For example, you can configure one monitor for each instance of an Oracle parallel server in use. Or, you can configure a custom monitor to check the health of the database, and a process termination monitor to instantly detect termination of the database process.

You assign each monitor a unique name in SMIT.

It is possible to configure either a process monitor or a custom monitor. For example, you can supply a customized script to SystemMirror that sends a request to a database to check that it is functioning. A non-zero exit from the script indicates a failure of the monitored application, and PowerHA® SystemMirror® responds by trying to recover the resource group that contains the application.

With each monitor configured, when a problem is detected, PowerHA® SystemMirror® attempts to restart the application, and continues up to a specified restart count. You select one of the following responses for PowerHA® SystemMirror® to take when an application cannot be restarted within the restart count:

- The fallover option causes the resource group containing the application to fall over to the node with the next-highest priority according to the resource policy.
- The notify option causes PowerHA® SystemMirror® to generate a **server_down** event, to inform the cluster of the failure.

You can customize the restart process through the Notify Method, Cleanup Method, and Restart Method for the application monitor.

Note: If the System Resource Controller (SRC) is configured to restart the application, this can interfere with actions taken by application monitoring. Disable the SRC restart for the application (application start and stop scripts should *not* use the SRC unless the application is not restartable). For the case of a custom monitor, the script is responsible for the correct operation. The action taken by application monitoring is supported based on the script return.
If a monitored application is under control of the system resource controller, check to be certain that `action:multi` are **-O** and **-Q**. The **-O** Specifies that the subsystem is not restarted if it stops abnormally. The **-Q** Specifies that multiple instances of the subsystem are not allowed to run at the same time. These values can be checked using the following command:

```
lssrc -Ss <Subsystem> | cut -d : -f 10,11
```

If the values are not **-O** and **-Q,** they must be changed using the **chssys** command.

Measuring application availability

You can use the Application Availability Analysis Tool to measure the amount of time that any of your applications (with defined application controller) is available.

The PowerHA® SystemMirror® software collects, time stamps, and logs the following information:

- An application starts, stops, or fails.

- A node fails or is shut down, or comes up.
- A resource group is taken offline or moved.
- Application monitoring is suspended or resumed.

Using SMIT, you can select a time period and the tool will display uptime and downtime statistics for a given application during that period. The tool displays:

- Percentage of uptime
- Amount of uptime
- Longest period of uptime
- Percentage of downtime
- Amount of downtime
- Longest period of downtime.

All nodes must be available when you run the tool to display the uptime and downtime statistics. Clocks on all nodes must be synchronized in order to get accurate readings.

The Application Availability Analysis tool treats an application that is part of a concurrent resource group as available as long as the application is running on any of the nodes in the cluster. Only when the application has gone offline on all nodes in the cluster will the Application Availability Analysis tool consider the application as unavailable.

The Application Availability Analysis tool reports application availability from the PowerHA® SystemMirror® cluster infrastructure's point of view. It can analyze only those applications that have been properly configured so they will be managed by the PowerHA® SystemMirror® software.

When using the Application Availability Analysis tool, keep in mind that the statistics shown in the report reflect the availability of the PowerHA® SystemMirror® application conroller, resource group, and (if configured) the application monitor that represent your application to PowerHA® SystemMirror®.

The Application Availability Analysis tool cannot detect availability from an end user's point of view. For example, assume that you have configured a client-server application so that PowerHA® SystemMirror® manages the server, and, after the server was brought online, a network outage severed the connection between the end user clients and the server. The end users would view this as an application outage because their client software could not connect to the server, but PowerHA® SystemMirror® would not detect it, because the server it was managing did not go offline. As a result, the Application Availability Analysis tool would not report a period of downtime in this scenario.

Network configuration and name serving

Setting up and maintaining clear communication paths for the Cluster Manager is a key element for efficient cluster operation.

Integrating PowerHA® SystemMirror® with network services

PowerHA® SystemMirror® requires IP address to name resolution during the configuration process. 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 PowerHA® SystemMirror® IP labels in all cluster nodes' **/etc/hosts** tables.

To ensure the most rapid name resolution of cluster nodes, 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,bind
```

Putting the local option first tells the system to use **/etc/hosts** first. If your installation uses NIS you can also add `nis`. For example,

```
hosts=local,bind,nis
```

You can also change the order for name resolution by changing the environment variable NSORDER as follows:

```
NSORDER=local,bind,nis
```

Note: By default, during the process of IP address swapping, to ensure that the external name service does not cause AIX® to map the service IP address to the wrong network interface, PowerHA® SystemMirror® automatically disables NIS or DNS by temporarily setting the AIX® environment variable *NSORDER=local* within the event scripts.
If you are using NIS, have the NIS primary server outside the cluster, and have the cluster nodes run as NIS backup servers. At a minimum, every PowerHA® SystemMirror® node must be able to access NIS primary or backup servers on a local subnet, and not via a router.

Important: You cannot use DHCP to allocate IP addresses to PowerHA® SystemMirror® cluster nodes. Clients may use this method, but cluster nodes cannot.

Planning disks and volume groups

Planning the disk layout is crucial for the protection of your critical data in a PowerHA® SystemMirror® 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**. PowerHA® SystemMirror® will be managing the disks and varying them on and off as needed to handle cluster events.
- One disk must be available for dedicated use by the cluster repository. This disk is ideally a LUN defined on a SAN for all nodes in the cluster.

Quorum issues

Setting up quorum correctly when laying out a volume group is very important. Quorum must be enabled on concurrently accessed 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. PowerHA® SystemMirror® provides some protections to avoid the failure, but planning is still important.

Either build three-disk volume groups or disable quorum on non-concurrent volume groups. You can also use the **forced varyon** option to work around quorum issues.

PowerHA® SystemMirror® selectively provides recovery for resource groups that are affected by failures of individual resources. PowerHA® SystemMirror® automatically reacts to a "loss of quorum" LVM_SA_QUORCLOSE error associated with a volume group going offline on a cluster node. If quorum is lost for a volume group that belongs to a resource group on a cluster node, the system checks whether the

LVM_SA_QUORCLOSE error appeared in the node's AIX® error log file and informs the Cluster Manager to selectively move the affected resource group.

> **Note:** When the AIX® error log buffer is full, new entries are discarded until space becomes available in the buffer and adds an error log entry to inform you of this problem.
> PowerHA® SystemMirror® launches selective fallover and moves the affected resource group in the case of the LVM_SA_QUORCLOSE error. Be aware that this error can occur only if you use mirrored volume groups with quorum enabled. PowerHA® SystemMirror® monitors for LVM_IO_FAIL errors for all volume groups. When this error is reported, it is determined if quorum has been lost for the affected volume group. If quorum is lost, LVM_SA_QUORCLOSE is issued and selective fallover occurs.

Planning hardware maintenance

You should plan for hardware maintenance.

Good maintenance practice in general dictates that you:

- Check cluster power supplies periodically
- Check the **errlog** and 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
- Network interface card
- Disk
- Node.

Planning software maintenance

Planning for software maintenance includes several different actions.

These actions include:

- 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 PowerHA® SystemMirror® for AIX®.

Runtime 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 a PowerHA® SystemMirror® cluster requires attention to some issues that have different ramifications in the cluster environment compared to maintaining a single system.

Tasks that require stopping the cluster

PowerHA® SystemMirror® allows you to do many tasks without stopping the cluster; you can do many tasks dynamically using the DARE and C-SPOC utilities. However, there are some tasks that require that you stop the cluster. For example, renaming the cluster or a cluster node requires restarting cluster services.

Changing the cluster configuration and cluster behavior

Changing the cluster configuration can have cascading effects on cluster behavior. This topic contains warnings about actions that endanger the proper behavior of a PowerHA® SystemMirror® cluster. It also includes some reminders about proper maintenance procedures.

Installing PowerHA® SystemMirror® makes changes to several AIX® files. All the components of the cluster are under PowerHA® SystemMirror® control once you configure, synchronize, and run the cluster software. Using AIX® to change any cluster component, instead of using the PowerHA® SystemMirror® menus and synchronizing the topology and/or the cluster resources, interferes with the proper behavior of the PowerHA® SystemMirror® cluster software and thus affect critical cluster services.

Stopping and starting cluster services

Do not directly start or stop daemons or services that are running under the control of PowerHA® SystemMirror®. Any such action affects cluster communication and behavior. You can choose to run certain daemons (Clinfo) but others are required to run under PowerHA® SystemMirror® control.

Most important, never use the **kill - 9** command to stop the Cluster Manager or any RSCT daemons. This causes an abnormal exit. SRC runs the **clexit.rc** script and halts the system immediately. This causes the other nodes to initiate a fallover.

TCP/IP services are required for cluster communication. Do not stop this service on a cluster node. If you need to stop PowerHA® SystemMirror® or TCP/IP to maintain a node, use the proper procedure to move the node's resources to another cluster node, then stop cluster services on this node.

Node, network, and network interface issues

The PowerHA® SystemMirror® configuration of the nodes and IP addresses 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 network interface using AIX® SMIT menus or commands, individually on a cluster node, outside of PowerHA® SystemMirror®.

Do not start or stop daemons or services that are running under the control of PowerHA® SystemMirror®. This action will affect cluster communication and behavior.

Be sure to follow proper procedures for the following types of changes:

- Changing the IP label/address of any network interface defined to PowerHA® SystemMirror®. Changes to IP addresses must be updated in the PowerHA® SystemMirror® cluster definition and the cluster must then be resynchronized. Any change to network interface attributes normally requires stopping cluster services, making the change, and restarting cluster services.
  Note that in some circumstances you can use the PowerHA® SystemMirror® facility to swap a network service IP address dynamically, to another active network interface on the same node and network, without shutting down cluster services on the node.
- Changing netmasks of network interfaces. Service and other network interfaces 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.
  It is important to configure the correct interface name for network interfaces.
- Taking down network interface cards. Do not take down all cards on the same network if a local network failure event is set up to stop cluster services and move resource groups to another node. If the cluster is customized to stop cluster services and move resource groups to another node when all communications on a specific network fail, and you take down all network interfaces, this will force the resource groups to move to another node whether you intend it or not.
- Taking down network interfaces. Do not bring all network interfaces down on the same network if there is only one network and no point-to-point network is defined. Doing this will cause system contention between cluster nodes and fallover attempts made by each node. A Group Services domain merge message is issued when a node has been out of communication with the cluster and then attempts to reestablish communication. One or more nodes may be halted until the Group Services domain merge occurs.

Making changes to network interfaces

In some circumstances, you can use the PowerHA® SystemMirror® facility to swap a network service IP address dynamically, to an active boot interface on the same node and network, without shutting down cluster services on the node.

Typically, stop the cluster to make any change to network interfaces. If you must change the IP address of an network interface, or if you change the IP label/address, 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.

Maintaining and reconfiguring networks

Moving Ethernet ports on a running cluster results in network interface swap or node failure. Even a brief outage results in a cluster event.

Shared disk, volume group, and file system issues

You should not change the configuration of a PowerHA® SystemMirror® shared volume group or file system using AIX®, outside of PowerHA® SystemMirror®. Any such action could affect cluster behavior if the change is not properly propagated to all the nodes. The Cluster Manager and the cluster event scripts assume the shared volume groups and file systems are under PowerHA® SystemMirror® control. If you change the environment, the event scripts will not be able to complete properly and you could get unexpected results.

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 PowerHA® SystemMirror® cluster, you should always use the PowerHA® SystemMirror® C-SPOC utility (*smit 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 file systems issues

Do not change the configuration of a PowerHA® SystemMirror® shared volume group or file system using AIX®, outside of PowerHA® SystemMirror®. Any such action will affect cluster behavior. The Cluster Manager and the cluster event scripts assume the shared volume groups and file systems are under PowerHA® SystemMirror® control. If you change the environment, the event scripts will not be able to complete properly and you will get unexpected results.

Use the C-SPOC utility (`smit cl_admin`) for common maintenance tasks like creating, extending, changing, or removing a shared file system.

When configuring volume groups and file systems:

- Do not set file systems to automount; PowerHA® SystemMirror® handles the mounts at startup and during cluster events.
- Do not set volume groups to autovaryon; PowerHA® SystemMirror® executes the varying on and off as needed.
- If you are testing something when the cluster is not running and you varyon a volume group or mount a file system, remember to unmount the file system and vary off the volume group before you start PowerHA® SystemMirror®.
- Do not have any processes running that would point to a shared file system when cluster services are stopped with resource groups brought offline on the node that currently owns that file system. If cluster services are stopped with resource groups brought offline and the application stop script fails to terminate the processes that are using the file system, that file system will be unable to unmount and the fallover will not occur. The cluster will go into a **config_too_long** condition.

One of the more common reasons for a file system to fail being unmounted when cluster services are stopped with resource groups brought offline is because the file system is busy. To unmount a file system successfully,

no processes or users can be accessing it at the time. If a user or process is holding it, the file system will be "busy" and will not unmount. The same issue may result if a file has been deleted but is still open.

This is easy to overlook when you write application stop scripts. The script to stop an application should also include a check to make sure that the shared file systems 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 file systems in question. These processes can then be killed. This will free the file system so it can be unmounted.

Refer to the AIX® man pages for complete information on this command.

General file systems issues

There are some general file system issues that you should be aware of.

The following are some more general file systems concerns:

- Full file systems in the root volume group may cause cluster events to fail. You should monitor this volume group and clean it up periodically. You can set up a cron job to monitor file system size to help avoid filling a critical file system (for example, the **hacmp.out** file can get quite large).
- Shared file systems must have the *mount* option set to false, so that PowerHA® SystemMirror® can mount and unmount them as needed to handle cluster events.
- Be aware of the way NFS file systems are handled.

Expanding file systems

You can use C-SPOC to increase the size of a file system.

Follow this procedure:

1. Enter `smit cl_admin`
2. Go to **System Management (C-SPOC) › Storage › File Systems** and press Enter.
3. Select the option to change a cluster file system.
4. Select the file system to change.
5. Enter the new size for the file system.
6. Synchronize the new definition to all cluster nodes via the **Synchronize a Shared Volume Group Definition** dialog.

Application issues

There are some key points to remember when planning and maintaining applications.

These points include:

- Application maintenance will require downtime for resource groups if binaries reside 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.
- Do not have shared applications already running when you start the cluster. A second attempt at starting already running applications may cause a problem.
- Do not 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 fallover attempt to be unsuccessful.

Hardware maintenance

Hardware failures must be dealt with promptly, as they may create single points 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.

Some issues to be aware of in a high availability environment include:

- Shared disks connect to both systems.
- Set up mirroring so that the mirrored disk copy is accessible by a different controller. This prevents loss of data access when a disk controller fails. When a disk controller fails, the mirror disk is accessible through the other controller.

Replacing topology hardware

There are certain conditions under which you can use DARE and certain conditions under which you must plan cluster downtime.

Nodes, networks, and network interfaces and devices comprise the topology hardware. Changes to the cluster topology often involves downtime on one or more nodes if changes to cabling or adding/removing network interfaces is involved. In most situations, you can use the DARE utilities to add a topology resource without downtime.

> **Note:** No automatic corrective actions take place during a DARE.

Replacing a node or node component

Using the DARE utility, you can add or remove a node while the cluster is running.

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 the 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.
- NIC physical placement is important - use the same slots as originally assigned.
- 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.
- For NIC replacement - use the same slots as originally assigned.

Removing a node

You can add or remove a node.

The basic procedure for adding or removing a node:

1. Install AIX®, PowerHA® SystemMirror® and LPPs on new node and apply PTFs to match the levels of the previous node.
2. Connect networks and test.
3. Configure TCP/IP.
4. Import volume group definitions.
5. Change the Configuration Database configuration on one of the existing nodes.
6. Synchronize and verify from the node where you made the changes.

Replacing networks and network interfaces

You can only protect your applications from downtime due to a network failure if you configure more than one IP network. If no backup network is configured, the cluster will be inaccessible to all but directly connected clients.

> **Note:** It is important to configure the correct interface name for network interfaces.

You can replace network cabling without taking PowerHA® SystemMirror® off line. You can also replace hubs, routers, and bridges while PowerHA® SystemMirror® 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 on the same node and network. Then, you can service the failed network interface card without stopping the node.

If the hardware supports hot-pluggable network interfaces, no cluster downtime is required for this procedure.

If you cannot use the **swap_adapter** function, use this procedure:

1. Move resource groups to another node using the Resource Group Management utility.
2. Use the hotplug mechanism to replace the card.
3. Assign IP addresses and netmasks for interfaces if they were undefined.
4. Test IP communications.

Preventive maintenance

If you have a complex and/or very critical cluster, 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.

Cluster snapshots

Periodically take 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 configuration. You can use the snapshot to rebuild the cluster quickly in case of an emergency. You might want to consider setting up a cron job to do this on a regular basis.

Backups

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 frequently.

Some applications have their own online backup methods.

You can use any of the following:

- **mksysb** backups
- **sysback, splitlvcopy** online backups.

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 PowerHA® SystemMirror® software

- Adding new applications
- Adding new device drivers
- Changing TCP/IP configuration
- Changing cluster topology or resources
- Changing LVM components of **rootvg** (paging space, file system sizes)
- Changing AIX® parameters (including the tuning parameters: I/O pacing, **syncd**).

Using splitlvcopy

You can use the splitlvcopy method on raw logical volumes and file systems 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 file system. 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 PowerHA® SystemMirror® 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 should 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 after it cycles more than 7 times.

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

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 file systems (and take action if necessary).
- Check that certain processes are running.

Regular testing

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.

It should include fallovers of all nodes and full verification of tested protected applications. This is strongly encouraged if you are changing or evolving your cluster environment.

Updating software on a cluster

Before you upgrade software on a highly available system, you must consider how the upgrade affects other software and applications on all nodes in the cluster.

When cluster services are active, the cluster software is interfacing with the system software and external components such as Cluster Aware AIX® (CAA) and Reliable Scalable Cluster Technology (RSCT). The cluster software might also be monitoring applications.

Installing or updating the software for any of these components might cause a disruption. This disruption might be interpreted as a failure by the cluster software and can trigger a failover or corrupt the cluster software.

Before you upgrade software, consider the following information:

- Assess the impact to the application and software that you are upgrading. For example, upgrading to CAA and RSCT might affect PowerHA® SystemMirror®.
- Before you update the active node, you must stop cluster services. If the update includes changes to the RSCT filesets, you must stop cluster services by using the **Bring Resource Groups Offline** option. Otherwise, you can stop cluster services by using the **Unmanage Resource Groups** option.
- To update the software, if you must restart the system, you can stop cluster services and move any resource groups to a standby node before the updates are applied.
- If you are updating AIX® or PowerHA® SystemMirror® software, take a cluster configuration snapshot and save it in a directory outside the cluster.
- Back up the operating system and any critical data. Prepare a backout plan in case you encounter problems with the upgrade.
- Test the update process on a test cluster before you update the production cluster.
- Use the alternate disk migration installation process.
- Maintain the installed software at the same version and fix levels on all nodes in the cluster.

To update to a new AIX® Technology Level or to apply a PowerHA® SystemMirror® Service Pack, complete the following steps:

1. Stop cluster services. If the update includes changes to the RSCT filesets, you must stop cluster services by using the **Bring Resource Groups Offline** option. Otherwise, you can stop cluster services by using the **Unmanage Resource Groups** option.
2. Before you upgrade AIX® or PowerHA® SystemMirror®, you must disable the **cthags** options by running the following commands:

   ```
   /usr/sbin/rsct/bin/hags_disable_client_kill -s cthags
   /usr/sbin/rsct/bin/hags_stopdms -s cthags
   ```

3. After the update is complete, you must enable the **cthags** options by running the following commands:

   ```
   /usr/sbin/rsct/bin/hags_enable_client_kill -s cthags
   /usr/sbin/rsct/bin/hags_startdms -s cthags
   ```

4. Start cluster services and resume normal cluster operations.

Resource group behavior during cluster events

Look here for an overview of resource group events and describe when PowerHA® SystemMirror® moves resource groups in the cluster, how the resource groups are placed on the nodes, and how to identify the causes of the underlying cluster events.

This information may be especially useful for experienced PowerHA® SystemMirror® users who are familiar with previous resource group handling by PowerHA® SystemMirror® but may *not* be aware of changes made in recent releases.

It is assumed that the reader is familiar with basic resource group fallover policies.

Once you have planned and defined the cluster topology and resources, PowerHA® SystemMirror® monitors the working cluster and takes actions to recover when failures are detected. PowerHA® SystemMirror® monitors resources and launches events to handle resource groups. You do not need to specify these actions in the cluster; they are initiated automatically.

PowerHA® SystemMirror® manages resource groups by:

- Moving only the resource groups that are affected by a failure of an individual resource to another node in the cluster.
- Taking recovery actions when it fails to acquire resource groups on a node. This can happen when PowerHA® SystemMirror® attempts to move a resource group to another node in the cluster, but fails to acquire it on that node. You can disable automatic recovery, if needed.

Resource group event handling and recovery

PowerHA® SystemMirror® tracks the state of all cluster resources and manages the recovery according to the available backup resources.

When multiple backup resources are available, PowerHA® SystemMirror® can be configured to dynamically select the backup resources to use based on the current performance statistics (using a dynamic node priority policy). Event logging includes a detailed summary for each high-level event to help you understand exactly what actions were taken for each resource group during the handling of failures.

Events and resource groups with dependencies

If either parent/child or location dependencies between any resource groupsare configured in the cluster, PowerHA® SystemMirror® processes all events related to resource groups in the cluster with the use of **resource_state_change** trigger events that are launched for all resource groups for events where resource groups are affected.

The Cluster Manager then takes into account all configured runtime policies, especially the configuration of dependencies for resource groups, and the current distribution and state of resource groups on all nodes to properly handle any acquiring, releasing, bringing online or taking offline of resource groups.

When events handle resource groups with dependencies, a preamble is written to the **hacmp.out** log file listing the plan of sub_events for handling the resource groups.

resource_state_change

This trigger event is used for resource group recovery if resource group parent/child or location dependencies are configured in the cluster. This action indicates that the Cluster Manager needs to change the state of one or more resource groups, or there is a change in the state of a resource managed by the Cluster Manager. This event runs on all nodes if one of the following occurs:

- Application monitoring failure
- Selective fallover for loss of volume group
- Local network down
- Resource Group Acquisition Failure
- Resource Group Recovery on IP Interface Availability
- Expiration of settling timer for a resource group
- Expiration of fallback timer for a resource group.

While the event runs, the state of the resource group is changed to TEMP_ERROR or SECONDARY_TEMP_ERROR. This is broadcast to all nodes.

> **Note:** This is a place where you can add pre-event or post-events for specific resources if needed.

resource_state_change_ complete

This event runs on all nodes when the **resource_state_change** event has successfully completed. (Necessary recovery actions including release and acquire events have completed.)

Events for moving resource groups

PowerHA® SystemMirror® may move resource groups as a result of recovery actions taken during the processing of events such as **node_down** and especially for **resource_state_change**.

rg_move

This event moves a specified resource group from one node to another.

rg_move_complete

This action indicates that the **rg_move** event has successfully completed.

Resource group subevents and states

Handling of individual resources during the processing of an event may include the following actions or resource group states. For example, when a file system is in the process of being unmounted and mounted it is taken offline and then released by one node. Then, if there is an available backup node the file system will be acquired and brought online.

Resource group states	
Resource group state	**Description**
RELEASING	A resource group is being released either to be brought offline, or to be acquired on another node.
ACQUIRING	A resource group is being acquired on a node.
ONLINE	The resource group is online.
OFFLINE	The resource group is offline.
ERROR	The resource group is in an error state.
TEMPORARY ERROR	The resource group is in a temporary error state. It occurs, for instance, due to a local network failure or an application failure. This state informs the Cluster Manager to initiate an **rg_move** event for this resource group. Resource groups should not be in this state when the cluster is stable.
UNKNOWN	The state of the resource group is unknown.
UNMANAGED	You have stopped the cluster services without stopping the running applications. In this case: • PowerHA® SystemMirror® is not managing the resources. • The previous state of the group was ONLINE. • The application and other resources may continue to be running on the node.

After the completion of an event, PowerHA® SystemMirror® is aware of the state of resources and resource groups involved in the event. PowerHA® SystemMirror® then analyzes the resource group information that it maintains internally and determines whether recovery events need to be queued for any of the resource groups. PowerHA® SystemMirror® also uses status of individual resources in resource groups to print out a comprehensive event summary to the **hacmp.out** log file.

For each resource group, PowerHA® SystemMirror® keeps track of the nodes on which the resource group has tried to come online and failed. This information is updated when recovery events are processed. PowerHA® SystemMirror® resets the nodelist for a resource group as soon as the resource group moves to the online or error states.

When a resource group is in the process of being moved, application monitoring is suspended and resumed appropriately. The Application Monitor sees that the application is in recovery state while the event is being processed.

resume_appmon

This action is used by the Application Monitor to resume monitoring of an application.

suspend_appmon

This action is used by the Application Monitor to suspend monitoring of an application.

Note: If you change the node list for a resource group while cluster services are active, the resource group does not move. This function avoids any interruptions for the application in an active cluster environment. You can dynamically move resource groups to other nodes and take them online or offline using the Resource Group Management utility (clRGmove) from the command line or through SMIT. Using this function might cause an outage in the application during the move.

Cluster event processing

The resource group handling features add steps to the overall processing for an event.

These steps include:

1. The Cluster Manager communicates with RSCT Group Services to obtain information about topology events, and consolidates information about events related to resource groups.
2. The Cluster Manager performs event rollup and determines the actual cluster event to run.
3. The Group Services protocol is run to get all cluster nodes to agree on the event (voting).
4. The Cluster Manager starts up event scripts on the cluster nodes.
5. Event scripts get information about resource groups to process for the event:
    - Get information from the PowerHA® SystemMirror® Configuration Database and the Cluster Manager and determine resource groups to process for the event.
    - Get information about the nodes already tried and exclude these nodes from the default nodelist.
    - Exclude nodes with insufficient network interfaces (for resource groups that require network interfaces).
6. Check the node priority policy to prioritize the list of target nodes for a resource group.
7. Event scripts process resource groups. (Bring them online/offline, etc.)
8. The Cluster Manager internally marks resource groups as recoverable if failures are encountered during the "acquire" phase.
9. Event scripts complete.

    Note: For more information on steps 5-9, and for information on which attributes and policies take precedence when the Cluster Manager determines which resource group to process first, see the section (See General Resource Group Event Processing Logic).

10. The Cluster Manager gets the return code from scripts.
11. If the return code is **0**, the event has completed (it may or may *not* have been successful); otherwise, return **event_error**. (User intervention may be required to get the cluster back to a stable state.)
12. The Cluster Manager notes the resource groups affected by a local network failure event and marks affected resource groups as recoverable.
13. The Cluster Manager notes the resource groups affected by a local network failure event (13) or by an acquiring error (8) and enqueues recovery events for each resource group in the recoverable state.
14. End of event.

General resource group event processing logic

You can specify various policies for resource groups in the cluster that affect the order in which resource group events take place.

Such policies may include:

- Requesting PowerHA® SystemMirror® to move a resource group to a particular destination node or state.
- Setting a dynamic node priority.

- Specifying parent/child or location dependencies between resource groups
- Customizing resource recovery for service IP labels and volume group resources (specifying fallover or notify)

This section presents a high-level view of what actions take precedence in the resource group event processing process. The Cluster Manager examines the following variables to determine the order in which to handle events for resource groups. If all variables are equally true for two or more groups, groups are sorted according to the processing order specified on the nodes in the cluster.

1. Resource group states (online, offline, error, unmanaged) determine which policies will be considered and, therefore, which resource groups will be taken for processing first. For instance, if the resource group is offline, the dynamic node priority set for this resource group is not considered. If the resource group is online, the Cluster Manager does not have to move it and does not perform look-ahead process to find an appropriate node.
   Also, in this step, resource groups' dependencies are considered to determine which resource groups must be processed before other resource groups can be processed.
   - Look-ahead for resource availability is considered. Nodes with insufficient network interfaces (for resource groups that require network interfaces) are excluded, and this affects the order in which events for resource groups will be handled.
   - The node distribution policy for resource groups is considered.
   - Dynamic node priority is considered.
   - Participating nodelist for the particular resource group is considered.
   - Startup, fallover and fallback settings of resource groups are considered: These settings include any previously configured delayed fallback timers and settling times for resource groups.
   - Once the Cluster Manager decides which resource group to process, it considers the resource group dependencies, processing order settings and inter-site management policies. At this step, the Cluster Manager chooses the path for processing.

Resource groups in clusters with dependencies or sites are processed in phases since more variables must be considered.

Selective fallover for handling resource groups

Selective fallover is a function of PowerHA® SystemMirror® that attempts to selectively move only a resource group that has been affected by an individual resource failure, rather than moving all resource groups, to another node in the cluster. Selective fallover provides recovery for individual resource groups that are affected by failures of specific resources.

Resources for which selective fallover Is used

PowerHA® SystemMirror® utilizes selective fallover in the case of a failure of several types of resources that can belong to a resource group.

These types include:

- Service IP labels
- Applications
- Volume groups.

You can customize the default selective fallover behavior to use a notification instead of a fallover for the following types of resources:

- Service IP labels
- Volume groups. This is the only resource type where customization also affects the secondary instance (if the resource group is replicated).

Look-ahead for moving resource groups and choosing takeover nodes

PowerHA® SystemMirror® determines the highest priority node using the resource group nodelist, dynamic node priority, and persistent migration requests, as well as the availability of backup resources.

For example, if a group contains a service IP label, PowerHA® SystemMirror® looks at the status of the available interfaces on the backup nodes. If the resource group is part of a resource group parent/child, startafter, stopafter, or location dependency set, PowerHA® SystemMirror® takes this into account also.

PowerHA® SystemMirror® does not move a resource group when there are no available backup resources. Instead, it simply takes the group offline from the current node. This result is clearly indicated in the event summary in the **hacmp.out** log file.

Selective fallover caused by network interface failures

When a network interface with a PowerHA® SystemMirror® service IP label fails and there are no other network interfaces available on the node on the same PowerHA® SystemMirror® network, the affected applications on that node cannot run. If the service network interface is the last one available on the node, the network interface failure triggers a network failure event.

PowerHA® SystemMirror® distinguishes between two types of network failure, *local* and *global*. A local network failure occurs when a node can no longer communicate over a particular network, but the network is still in use by other nodes. A global network failure occurs when all nodes lose the ability to communicate over a network.

PowerHA® SystemMirror® uses the following formats for local and global network failure events:

Local Network Failure Event

```
network_down <node_name> <network_name>
```

Global Network Failure Event

```
network_down -1 <network_name>
```

In the case of a local network failure, you may create a post-event to trigger a **node_down** event. While this has the desired effect of moving the resource group with the failed resource to another node, it has the undesired effect of moving all of the resource groups on the node to other nodes.

Selective fallover uses this infrastructure to better handle network interface failures. You do not have to create a post-event to promote a local network failure to a node failure in this case. See the section below for more information on how PowerHA® SystemMirror® handles network interface failures.

You should not promote global network failures to **node_down** events as the global network event applies to all nodes and would result in a node down for all nodes.

Actions taken for network interface failures

PowerHA® SystemMirror® takes the following actions in cases of network interface failures:

- When a network interface with a service IP label fails, and there are no network interfaces available on the same node (therefore, a **swap_adapter** is not possible), it moves only the resource group associated with the failed service network interface to another node.
- When a network interface fails and this can result in launching an **rg_move** for the affected resource group, a check for available network interfaces is made. The highest priority node with an available network interface attempts to acquire the resource group.
- PowerHA® SystemMirror® checks that a network interface is available on the node joining the cluster before releasing the resource group. If no network interfaces are available, the resource group is not released.

The above actions assume available nodes in the resource group definitions.

The **hacmp.out** file contains messages informing you about cluster activity that results from selective fallover actions.

Selective fallover caused by local network failures

When a local network failure event occurs, the Cluster Manager takes selective recovery actions for resource groups containing a service IP label connected to that network. The Cluster Manager tries to move only the resource groups affected by the local network failure event, rather than all resource groups on a particular node.

> **Note:** You do not need to create a post-event to promote a local network failure to a node failure in cases of network interface failures.

For example, if you have two resource groups:

```
RG1 - service label on network net_ether_01
RG2 - service label on network net_ether_02
```

If network net_ether_02 fails, the Cluster Manager will move RG2. RG2 will not touch RG1.

Selective fallover caused by application failures

When an application that is being monitored by Application Monitoring fails, PowerHA® SystemMirror® attempts to move the resource group containing that application to another node. Only the affected resource group is moved.

PowerHA® SystemMirror® can monitor your applications to keep the application software itself highly available. There are two types of application monitors you can use:

- Process monitors.
  With process monitoring you specify key attributes about a process and the number of instances of that process that should be active. If the number of instances of that process falls is less than the number you specify, PowerHA® SystemMirror® will trigger a recovery or failover action. Process monitoring is effective for applications with key processes though monitoring the process does not necessarily indicate it is capable of doing real work.
- Custom monitors.
  To use custom monitoring you must supply a script which PowerHA® SystemMirror® will execute to determine the health of the application. The script can perform any operation you want, such as sending a dummy transaction to a database to verify that it is responding. Custom monitors are effective for detecting an application that is hung or otherwise unresponsive.

You can combine both types of monitoring for a single application. This way you can instantly detect failure of key processes as well as hang conditions. Both types of monitors support optional scripts which let you retry or restart the application, as well as special notification scripts which can alert a system administrator to respond to the problem.

Selective fallover caused by a volume group loss

Selective fallover can also be triggered when PowerHA® SystemMirror® detects a volume group failure on a node containing that resource group. In other words, PowerHA® SystemMirror® automatically reacts to a "loss of quorum" error associated with a volume group going offline on a cluster node.

If a volume group in the resource group has dropped offline due to a loss of quorum error for the volume group on that node, PowerHA® SystemMirror® selectively moves the resource group to another node.

PowerHA® SystemMirror® uses the selective fallover for volume group loss functionality under the following conditions:

- PowerHA® SystemMirror® monitors all volume groups that are included in the resource group, and all volume groups on which file systems that are part of the resource group depend.
- PowerHA® SystemMirror® moves only resource groups that contain volume groups with volume groups for which the LVM_SA_QUORCLOSE error has been logged by the error daemon in the AIX® **errpt** on that node.

Note: PowerHA® SystemMirror® does not react to any other type of volume group errors automatically. In these cases, you still need to configure customized error notification methods, or use AIX® Automatic Error Notification methods to react to volume group failures.

PowerHA® SystemMirror® uses an Error Notification method to inform the Cluster Manager about the failure of a volume group. When using this error notification method:

- Do *not* modify this error notification method. PowerHA® SystemMirror® issues a warning and takes no action if you attempt to customize this notification method, or to use it to protect against the failure of other types of resources.
- Synchronize the cluster after making changes to the cluster configuration. A notification script used for a volume group failure should correspond to the current configuration of cluster resources, otherwise PowerHA® SystemMirror® issues a warning during verification and takes no action to selectively move the affected resource group.
- Besides the **errnotify** entries created by PowerHA® SystemMirror® for selective fallover, the **errnotify** ODM may also contain other entries related to the same AIX® error labels and resources. However, selective fallover provides the most effective recovery mechanism to protect a resource group from the failure of a single resource.
- The notification method that is run in the case of a volume group failure provides the following information in the **hacmp.out** and **clstrmgr.debug** log files:
  - AIX® error label and ID
  - The name of the affected resource group
  - The node's name on which the error occurred.
- You can test the error notification methods generated by the selective fallover facility by emulating an error for each volume group in SMIT.

To test error notification:

1. From the command line, enter `smit sysmirror`.
2. In SMIT, select **Problem Determination Tools > PowerHA® SystemMirror® Error Notification > Emulate Error Log Entry**, and press Enter.
3. Select from the picklist the error notification object that was generated by the selective fallover facility for each volume group.

Handling of resource group acquisition failures

PowerHA® SystemMirror® uses event scripts to move resources around the PowerHA® SystemMirror® cluster. PowerHA® SystemMirror® differentiates certain types of failures in the event script. There are still fatal type errors where an error in the script logic or environment causes the script to fail, but now PowerHA® SystemMirror® traps recoverable errors related to the processing of the resources. This allows PowerHA® SystemMirror® to continue event processing and to try to bring the group online on the next available node.

Attempts by PowerHA® SystemMirror® to start or move a resource group may fail for a variety of reasons, such as busy or unavailable devices, or lack of disk space. PowerHA® SystemMirror® may react to such failures by attempting to move the resource group to another node.

In the event of a resource group acquisition failure on a particular node:

- Not all resource group acquisition failures require immediate manual intervention. In some cases, a resource group will be successfully brought online on another node. However, the fact that a resource group acquisition failure occurred indicates a system problem that needs attention.
- The Cluster Manager logs error messages when a node cannot acquire a resource group, and continues processing events so the cluster resources remain available.
  PowerHA® SystemMirror® *automatically* attempts to activate resource groups in the ERROR state on the node during a **node_up** event. You cannot disable this functionality. If an attempt to recover a resource group in the ERROR state on a joining node is made but the resource group acquisition on the node fails,

the non-concurrent resource group falls over to the next node in the nodelist, if one is available. If the concurrent resource group acquisition fails, the resource group remains in the ERROR state.

- PowerHA® SystemMirror® logs reported resource group acquisition failures (failures indicated by a non-zero exit code returned by a command) in **hacmp.out**. The information appears in an event summary that follows each main event's details.
  The event summary makes it easier for you to check the **hacmp.out** file for errors. Checking this log becomes more important, since the **config_too_long** console message will not be evident in every case where a problem exists.

  The **config_too_long** event runs whenever a cluster event takes too long to complete. When the **config_too_long** event runs, it indicates that an error has occurred or that a recovery operation might have stopped. By configuring a notification for the **config_too_long** event, the operator is alerted to take the appropriate action.

How PowerHA® SystemMirror® processes resource group acquisition failures

If a node fails in an attempt to acquire a resource group during a fallover, PowerHA® SystemMirror® marks the resource group "recoverable" and triggers an **rg_move** event to try to bring up the resource group on some other node.

Note that a failure may occur during the acquisition phase of an **rg_move**, and this may cause a queue of **rg_move** events. The software goes through the queue until the resource group is successfully brought online, or until all possible owners have failed to acquire it, in which case the resource group is kept in the ERROR state.

Resource group recovery when the network or interface is up

When a local network failure occurs, PowerHA® SystemMirror® determines whether any resource groups are affected by the failure. The affected resource groups are those that contain service labels defined on the failed network.

In this case, PowerHA® SystemMirror® checks whether such resource groups are online on a node, and attempts to move each affected resource group to another node by initiating an **rg_move** event for it.

The **rg_move** event attempts to bring the resource group online on another node. If PowerHA® SystemMirror® does *not* find available resources on any nodes to bring the resource group online, then the **rg_move** event leaves the resource group in an error state and the resource group goes offline and becomes unavailable.

PowerHA® SystemMirror® tries to bring resource groups in ERROR state online whenever a network interface becomes available. When bringing the affected resource groups online, PowerHA® SystemMirror® does the following:

1. If it finds resource groups that went to an error state due to a resource failure and contain a service IP label of a network interface that became available again, then it moves such resource groups to the **rg_temp_error_state**.
2. Prior to running an **rg_move** for an affected resource group, PowerHA® SystemMirror® determines possible candidate nodes for bringing the resource group online. If it cannot find any candidate nodes, then the resource group remains offline and unavailable.
3. If PowerHA® SystemMirror® finds candidate nodes, it initiates an **rg_move** event to attempt to bring the resource groups online.

Disabling automatic recovery of resource groups

When a local network failure occurs, you may need to replace the failed resource first, before letting PowerHA® SystemMirror® automatically bring the affected resource group online. For instance, you may need to replace and test the network interface before bringing the affected resource group online.

To avoid automatic recovery of a resource group if the resource group goes into an error state, take these steps:

1. From the command line, enter `smit sysmirror`
2. In SMIT, select **System Management (C-SPOC) > Resource Group and Application Management > Bring a Resource Group Offline** and press Enter.
3. Specify that this resource group must remain offline on a node.

Result

Remember to bring the resource group back online manually when needed.

Recovering resource groups when nodes join the cluster

An attempt is made to automatically bring online the resource groups that are currently in the ERROR state. This further increases the chances of bringing the applications back online. When a node that is included in the nodelist for the resource group starts up, if the resource group is in the ERROR state on any node in the cluster, this node attempts to acquire the resource group. The node must be included in the nodelist for the resource group.

The resource group recovery on node startup is different for non-concurrent and concurrent resource groups:

- If the starting node fails to activate a *non-concurrent resource group* that is in the ERROR state, the resource group continues to selectively fall over to another node in the nodelist, if a node is available. In this case, PowerHA® SystemMirror® uses selective fallover. The fallover action continues until all available nodes in the nodelist have been tried.
- If the starting node fails to activate a *concurrent resource group* that is in the ERROR state, the concurrent resource group is not attended in the ERROR state.

> **Note:** PowerHA® SystemMirror® *automatically* attempts to activate resource groups in the ERROR state on the node during a **node_up** event. You cannot disable this functionality. If an attempt to recover a resource group in the ERROR state on a joining node is made but the resource group acquisition on the node fails, the non-concurrent resource groups falls over to the next node in the nodelist, if one is available. If the concurrent resource group acquisition fails, the resource group remains in the ERROR state.

Handling of resource groups configured with IPAT via IP aliases

When you configure your PowerHA® SystemMirror® cluster, you define certain IP labels/IP addresses (service addresses) to be kept highly available. These service addresses are typically the IP address used by clients to access the server application. PowerHA® SystemMirror® keeps the IP address available to clients by moving the address between different network interfaces.

IP aliasing is a function of the TCP/IP stack where multiple IP addresses can be added to the same physical interface. PowerHA® SystemMirror® uses IP aliases for recovery, so the base address for the interface does not change. PowerHA® SystemMirror® recovers the service address by adding it as a second, or alias address, on the same interface. A single physical interface can host or back up multiple service addresses. This greatly improves the configuration flexibility and fallover options by requiring fewer physical resources to serve as backup. IPAT via IP Aliases is also faster as there are fewer commands required when moving addresses.

To control the placement of the service IP label aliases on the cluster node physical network interface cards, you can configure a distribution preference for the aliases of the service IP labels that are placed under PowerHA® SystemMirror® control.

Resource group behavior when using IPAT via IP aliases

With IPAT via IP Aliases, the service address is added as an alias address on an available boot interface. This applies to the node where the resource group is first acquired, as well as to the node(s) that might acquire it later. When the resource group is released, the service address is removed from the interface, but this does not alter the base or boot address on the interface.

The mechanics of IP address takeover on a network using aliases works the same way for all non-concurrent resource groups. While the mechanics are identical, IPAT via IP aliases does affect the initial startup and fallover placement of the resource group.

The aliased service IP labels are distributed across all available boot interfaces. To facilitate even distribution of labels across all available IP interface cards, PowerHA® SystemMirror® sorts all available interfaces by state and then by the number of aliased addresses already placed on the interface, and places the aliased labels accordingly. Note that this distribution is only done at fallover time, PowerHA® SystemMirror® makes no attempt to redistribute the labels later, if another interface becomes active.

Note: If you want PowerHA® SystemMirror® to activate only a specific resource group on a node during startup among multiple resource groups that could potentially be acquired on this node, we recommend that you use the startup policy Online Using Node Distribution Policy.

Resource group placement on cluster startup

The presence of a service IP label in the resource group does not change the placement policy for the resource group on initial cluster startup. Therefore, on initial cluster startup, a non-concurrent resource group is placed according to the defined startup policy.

On the subsequent cluster startup, PowerHA® SystemMirror® moves the resource group containing the service IP label onto a node with a boot interface that:

- Is up
- Has a different subnet than the IP label that is being moved.

In addition, PowerHA® SystemMirror® follows these rules:

- If multiple boot interfaces are found that are up and have different subnets, then PowerHA® SystemMirror® moves the resource group onto the one that comes first in the alphabetically-sorted list of network interfaces configured on the node.
- If the resource group uses Online Using Node Distribution Policy startup, the resource group is placed on a node that does not host another resource group.

Resource group placement on fallover

On fallover, if you have configured resource groups that contain aliased service IP labels, this allows having more than one non-concurrent resource group on the same node. Therefore, more than one resource group can be serviced by a node with a single physical interface.

On fallover, PowerHA® SystemMirror® moves the resource group containing the service IP label onto a node with a boot interface which:

- Is up
- Has a different subnet
  - Is preferably not hosting another service label (if available).
  - Comes first in the alphabetically-sorted list of network interfaces in the network configuration.

The key advantage of IPAT via IP Aliases is that, on fallover, more than one resource group can be serviced by a node with a single physical interface.

Examples of location dependency and resource group behavior

Look here for scenarios that illustrate how location dependent resource groups are processed at startup and also how they are processed for various failure scenarios.

Publishing model with same node and different nodes dependencies

The XYZ Publishing company follows a business continuity model that involves prioritizing the different platforms used to develop the web content. XYZ uses location dependency policies to keep some resource groups strictly on separate nodes and others together on the same node.

The Production database (PDB) and Production application (Papp) are hosted on the same node to facilitate maintenance (and perhaps the highest priority node for these resource groups has the most memory or faster processor). It also makes sense to set up a parent/child relation between them, since the application depends on the database. The database must be online for the application to function.The same conditions are true for the System Database (SDB) and the System application (Sapp) and for the QA Database (QADB) and the QA application (QAapp).

Since keeping the production database and application running is the highest priority, it makes sense to configure the cluster so that the three database resource groups stay on different nodes (make them an Online On Different Nodes dependency set), and assign the PDB resource group with the **high** priority. The SDB is the **Intermediate** priority and the QADB is the **low** priority.

The databases and their related applications are each configured to belong to an Online On Same Node dependency set.

PowerHA® SystemMirror® handles these groups somewhat differently depending on how you configure startup, fallover, and fallback policies. It makes sense to have the participating nodelists differ for each database and application set to facilitate keeping these resource groups on the preferred nodes.

The following figure shows the basic configuration of the three nodes and six resource groups.

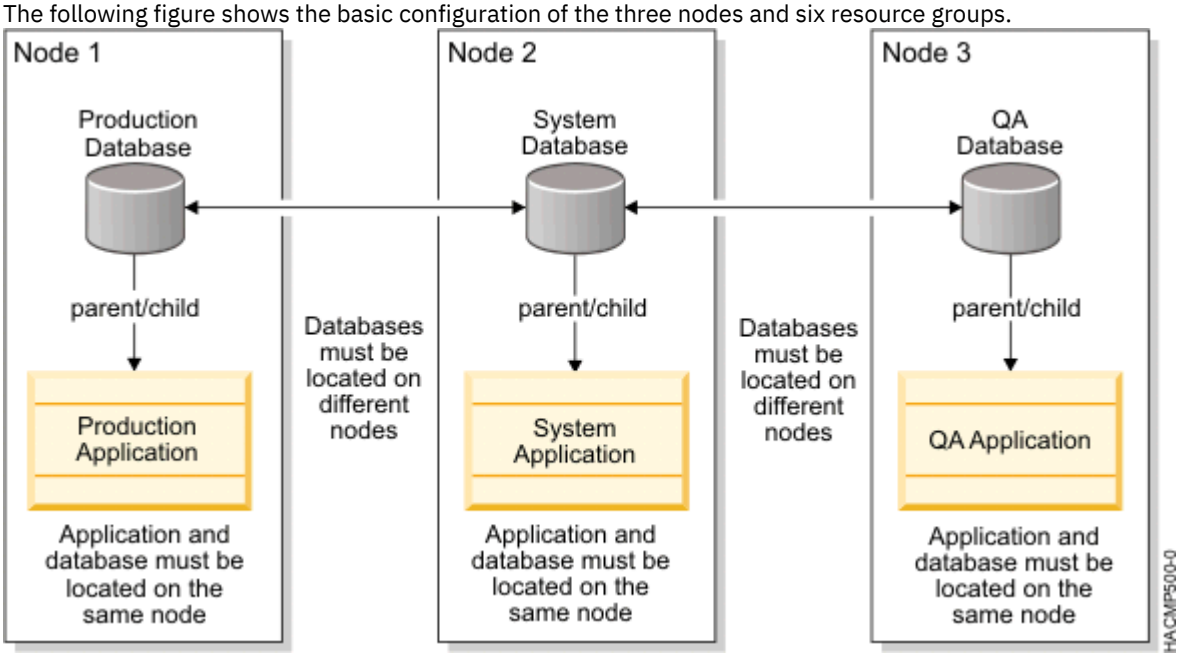


Figure 8: Publishing model with parent and child and location dependencies

Resource group policies: Online on first available node

For the following use case discussions, all six resource groups have the following policies:

- Startup Policy: Online On First Available Node
- Fallover Policy: Fallover to Next Priority Node
- Fallback Policy: Never Fallback

Resource group policies		
Participating Nodes	**Location Dependency**	**Parent/Child Dependency**
• PApp: 1, 2, 3 • PDB: 1, 2, 3 • SApp: 2, 3 • SDB: 2, 3 • QAApp: 3 • QADB: 3	Online On The Same Node Dependent Groups: • PApp with PDB • SApp with SDB • QAApp with QADB Online On Different Nodes Dependent set: [PDB SDB QADB] Priority: PDB > SDB > QADB	• PApp (child) depends on PDB (parent) • SApp (child) depends on SDB (parent) • QAApp (child) depends on QADB (parent)

Use case 1: Start nodes in numerical order (Node1 first)

Starting the nodes in numerical order we expect the Production resource groups to come online on Node 1, the System resource groups to come online on Node 2, and the QA resource groups to come online on Node 3. There is no contention.

Node 1 is the highest priority node for resource groups PDB and PApp. The parent/child dependency dictates that PDB must be brought online prior to processing PApp. Therefore, PowerHA® SystemMirror® processes the **rg_move** event to acquire PDB first and then it acquires PApp.

Node 1 is *not* in the nodelist for any other groups. Even if it were, the Online on Different Nodes dependency would disallow any less priority groups from coming online on this node.

Consolidated view of start node sequence: 1, 2, 3

Consolidated view of start node sequence			
Step	**Node 1**	**Node 2**	**Node 3**
Start Node 1	PApp: ONLINE PDB: ONLINE	PApp: PDB: SApp: SDB:	PApp: PDB: SApp: SDB: QAApp: QADB:
Start Node 2	PApp: ONLINE PDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE	PApp: PDB: SApp: SDB: QAApp: QADB:
Start Node 3	PApp: ONLINE PDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

Use Case 2: Start nodes out of order (Node 3)

> **Note:** Resource groups are offline, all nodes are offline

Start nodes out of order (Node 3)				
Step/ Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**
1	Start Node 3			
2				Acquire PDB
3				Acquire PApp
Post-condition/ Resource group states		PDB PApp	PDB: Papp	PApp: ONLINE PDB: ONLINE

Step/ Qualifier	Action	Node 1	Node 2	Node 3
			SDB SApp	SApp: ERROR SDB: ERROR QAApp: ERROR QADB: ERROR

Node 3 is the lowest priority node for PDB and PApp, as well as for SDB and SApp. Node 3 is the highest priority node for QADB and QAApp. However, the PDB/PApp pair has the highest priority due to the Online On Different Nodes Dependency. Therefore, PowerHA® SystemMirror® will acquire and start PDB on Node 3 and then process its child PApp. The other resource groups will go to the ERROR state based on the rule—these resource groups could have been brought online on Node 3 but were *not* acquired due to the Online On Different Nodes Dependency policy.

Use Case 2 Continued: Start Nodes Out of Order (Node 2)

> **Note:** Node 3 is up; cluster and group states as at the end of the previous table.

Start Nodes Out of Order (Node 2)				
Step/Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**
1	Start Node 2			
2				Release PApp
3				Release PDB
4			Acquire PDB	Acquire SDB
5			Acquire PApp	Acquire SApp
Post-condition/ Resource group states		PApp: PDB:	PApp: ONLINE PDB: ONLINE SApp: OFFLINE SDB: OFFLINE	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE QAApp: ERROR QADB: ERROR

Node 2 is the highest priority node for the SDB and SApp resource groups. However, the higher priority resource group in the Online On Different Nodes Dependency set is PDB. Therefore, PDB will fall over to this joining node while SDB and SApp will be acquired and started on Node 3. PowerHA® SystemMirror® moves Papp to the same node with PDB because these two resource groups belong to an Online on Same Node dependency set. QA PDB is less priority than SDB, so it stays in the ERROR state along with QAapp.

When Node 1 comes up, PDB and Papp will fall over to Node 1, SDB and Sapp will fall over to Node 2, and the QA resource groups will be acquired and started on Node 3.

Consolidated view of start nodes out of order: Sequence 3, 2, 1

Consolidated view of start nodes out of order: Sequence 3, 2, 1			
Step	**Node 1**	**Node 2**	**Node 3**
Start Node 3	PApp: PDB:	PApp: PDB: SApp: SDB:	PApp: ONLINE PDB: ONLINE SApp: ERROR SDB: ERROR

Step	Node 1	Node 2	Node 3
			QAApp: ERROR QADB: ERROR
Start Node 2	PApp: PDB:	PApp: ONLINE PDB: ONLINE SApp: OFFLINE SDB: OFFLINE	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE QAApp: ERROR QADB: ERROR
Start Node 1	PApp: ONLINE PDB: ONLINE	PApp: OFFLINE PDB: OFFLINE App: ONLINES SDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

Use Case 3: Fallover of resource groups due to node failure

Note: All nodes are ONLINE. Resource groups PDB and PApp are online on Node 1, SDB and SApp are online on Node 2, QAApp and QADB are online on Node 3

Fallover of resource groups due to node failure					
Step/ Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**	**Comments**
1	Node 1 crash.		Release SApp	Release QAApp	
2			Release SDB	Release QADB	QAApp and QDB go to ERROR state.
3			Acquire PDB	Acquire SDB	
4			Acquire PApp	Acquire SApp	
Post-condition/ Resource Group states		PApp: PDB:	PApp: ONLINE PDB: ONLINE SApp: ERROR SDB: ERROR	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE QAApp: ERROR QADB: ERROR	

When Node 1 fails, PowerHA® SystemMirror® releases SApp and SDB and QADB and QAapp and moves the highest priority resource group PDB and its Same Node dependency partner and child PApp to Node 2 and likewise moves the System groups to Node 3. The QA groups have nowhere to go; they go into ERROR state).

Use Case 4: Fallover of resource groups: Network goes down during fallover

Note: All nodes are ONLINE. Resource groups PDB and PApp are online on Node 1, SDB and SApp are online on Node 2, QAApp and QADB are online on Node 3. All applications use the app_network.

Fallover of resource groups: Network goes down during fallover					
Step/ Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**	**Comments**
1	Node 1 crash.		Release SApp	Release QAApp	
2			Release SDB	Release QADB	
3			Acquire PDB	Acquire SDB	QADB goes into ERROR state
4	app_network down Node 2				app_network failure.
5			Acquire PApp	Acquire SApp	PApp and SApp go to the ERROR state (network *not* available)
6			resource_state _ change event	resource_state _ change event	Triggers **rg_move** events
7			Release PDB	Release SApp	
8				Release SDB	
9			Acquire SDB	Acquire PDB	
10			Acquire SApp	Acquire PApp	
Post-condition/ Resource group states		PApp: PDB:	PApp: OFFLINE PDB: OFFLINE SApp: ERROR SDB: ONLINE	PApp: ONLINE PDB: ONLINE SApp: ERROR SDB: ERROR QAApp: ERROR QADB: ERROR	

In step 5, PApp goes to the ERROR state directly, instead of going through the acquisition phase since the Cluster Manager knows that the network required for PApp on Node 2 is currently down. This is in contrast to an acquisition failure.

In step 6, the event queue gets a resource_state_change event on the queue, which gets voted and queues additional ACQUIRE/RELEASE events.

In steps 7 & 8: SApp goes to the ERROR state due to network failure.

Publishing model: Alternate configuration

This model consists of three pairs of parent and child resource groups (total of six resource groups) and three nodes in the PowerHA® SystemMirror® cluster. The applications (PApp, SApp and QAApp) are Online On The Same Node with their corresponding databases (PDB, SDB and QADB.) All databases (which are parent resource groups also) are Online On Different Nodes from the other databases.

The only difference between the original Publishing Model configuration and this alternate Publishing Model is the resource group's startup preferences. This section uses the **Online On Home Node Only** startup policy whereas the original Publishing Model configuration uses **Online On First Available Node** as the startup policy for the resource groups.

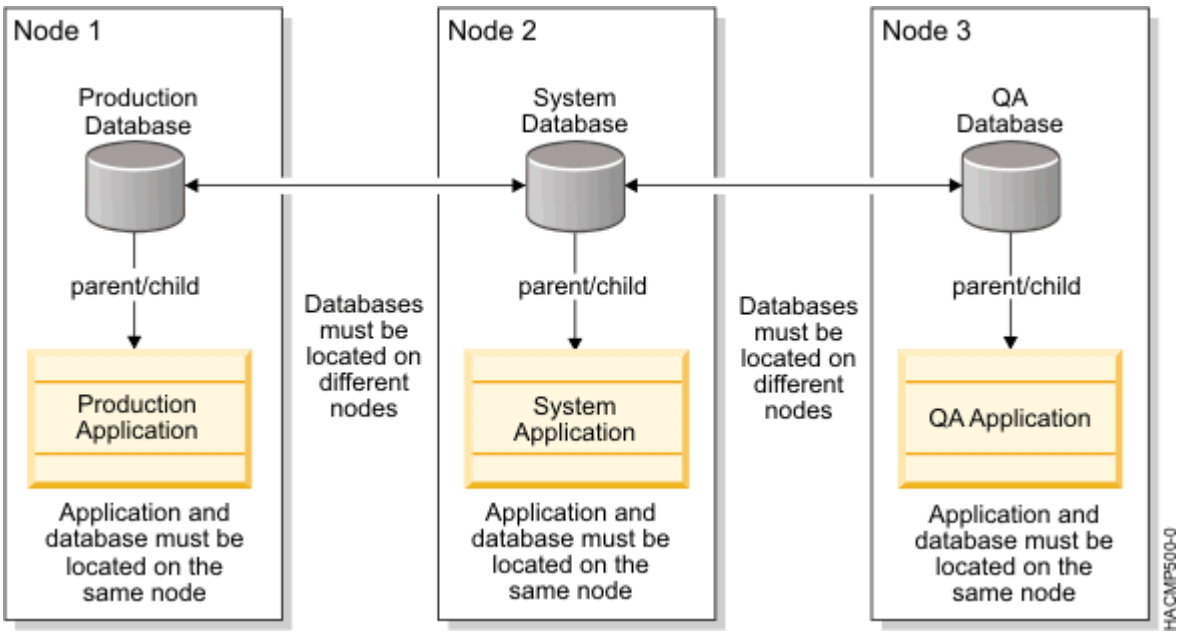


Figure 9: Alternate publishing model: Startup on home node only

Resource group policies: Online on home node only

All six resource groups have the following policies:

- Startup Policy: Online On Home Node Only - this is different from the previous set of use cases.
- Fallover Policy: Fallover to Next priority node
- Fallback Policy: Never Fallback

Table has 3 columns. The table displays the nodes, location dependency, parent and child dependency.

Resource group policies: Online on home node only		
Participating Nodes	**Location Dependency**	**Parent/Child Dependency**
PApp: 1, 2, 3 PDB: 1, 2, 3 SApp: 2, 3 SDB: 2, 3 QAApp: 3 QADB: 3	Online On The Same Node Dependent Groups: • PApp is with PDB • SApp is with SDB • QAApp is with QADB Online On Different Nodes Dependent set: [PDB SDB QADB] Priority: PDB>SDB>QADB	PApp (child) depends on PDB (parent) SApp (child) depends on SDB (parent) QAApp (child) depends on QADB (parent)

Use case 1: Start lowest priority node (Node 3)

> **Note:** All resource groups are offline, all nodes are offline.

Table show the step number and the corresponding actions and states on different nodes.

Use case 1: Start lowest priority node (Node 3)				
Step/ Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**
1	Start Node 3			
2				Acquire QADB

Step/ Qualifier	Action	Node 1	Node 2	Node 3
3				Acquire QAApp
Post-condition/ Resource group states		PApp: PDB:	Papp PDB: SApp: SDB:	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

Node 3 is the home node for resource groups QAApp and QADB. Although PDB and PApp have higher priority as defined by the Online On Different Nodes Dependency set, during cluster startup the startup policy allows only the QAApp and QADB resource groups to come online on Node 3. Therefore, the higher priority resource groups remain in the OFFLINE state at startup.

Use case 2: Start second node (Node 2)

> **Note:** Node 3 is up; cluster and group states as at the end of the previous use case.

Table show the step number and the corresponding actions and states on different nodes.

Use case 2: Start second node (Node 2)				
Step/Qualifier	**Action**	**Node 1**	**Node 2**	**Node 3**
1	Start Node 2			
2			Acquire SDB	
3			Acquire SApp	
Post-condition/ Resource group states		PApp: PDB:	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

Node 2 is the highest priority node for SDB and SApp resource groups. Since the startup policy for the resource groups is Online On Home Node, these resource groups will be started even though PDB and PApp are the highest priority resource groups.

Consolidated view of start node sequence: 3, 2, 1

Table show the start node sequence for different start nodes. It shows what happens for the node that starts and the corresponding nodes.

Consolidated view of start node sequence: 3, 2, 1			
Step	**Node 1**	**Node 2**	**Node 3**
Start Node 3	PApp: PDB:	PApp: PDB: SApp: SDB:	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

Step	Node 1	Node 2	Node 3
Start Node 2	PApp: PDB:	PApp: ONLINE PDB: ONLINE SApp: OFFLINE SDB: OFFLINE	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE
Start Node 1	PApp: ONLINE PDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: ONLINE SDB: ONLINE	PApp: OFFLINE PDB: OFFLINE SApp: OFFLINE SDB: OFFLINE QAApp: ONLINE QADB: ONLINE

WAS/DB2 cluster model and use cases

This model contains a DB2® database, a WebSphere® Application Server application that depends on DB2® and four WebSphere® Applications. The parent/child dependency for this model is that DB2® should be available prior to activating the WAS and WebSphere® (WS#) applications depend on the availability of WAS.

The location dependency of the resource groups is that DB2® and WAS should *not* be activated on the same node and WAS is Online On The Same Node Dependent with WS4 (see figure) and DB2® is Online On The Same Node with WS1, WS2 and WS3. The location dependency for the WS# is purely artificial in this example. However, this is a configuration where one of the nodes is fine-tuned for DB2® (hence will be the highest priority node for DB2®) and the other one is fine-tuned for WAS. They both have a common backup node, which can host only one of the two groups at a time.

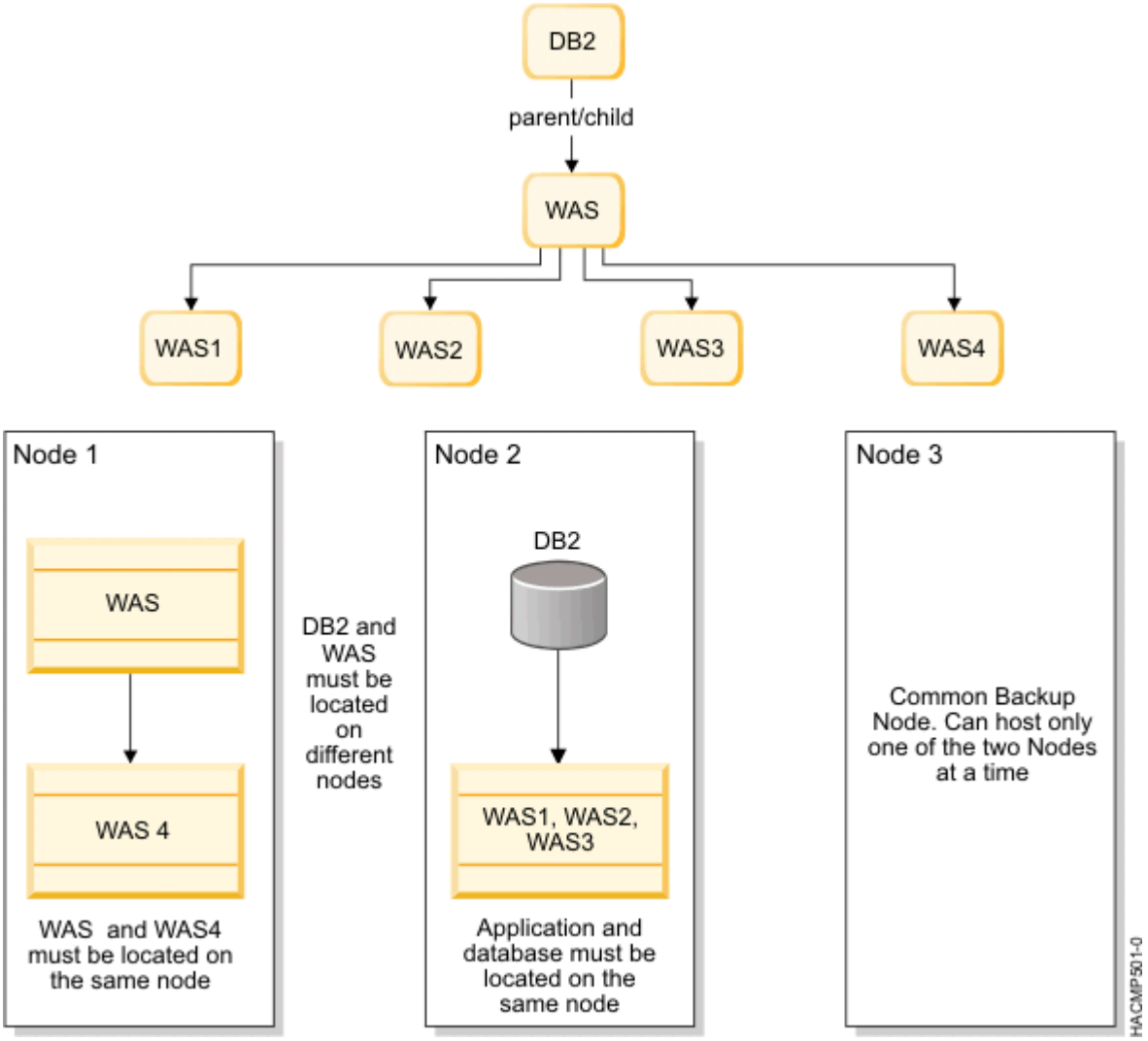


Figure 10: A WAS cluster and a DB2® cluster with location and parent and child dependencies

Resource group policies

All resource groups have the following policies:

- Startup Policy: Online On First Available Node
- Fallover Policy: Fallover to Next priority node
- Fallback Policy: Never Fallback

Participating Nodes	Location Dependency	Parent/Child Dependency
• DB2® [2, 3] • WS1 [2, 3] • WS2 [2, 3] • WS3 [2, 3] • WS4 [1, 3] • WAS [1, 3]	Online On The Same Node Dependent Groups: 1. DB2®, WS1, WS2, WS3 2. WAS, WS4 Online On Different Nodes Dependent Groups: • DB2®, WAS	1. WS1, WS2, WS3 and WS4 (children) depend on WAS (parent) 2. WAS (child) depends on DB2® (parent)

Use case 1: Start first node (Node 1)

> **Note:** All resource groups are offline, all nodes are offline.

All resource groups are offline, all nodes are offline.

Step/ Qualifier	Action	Node 1	Node 2	Node 3
1	Start Node 1			a Parent/child dependency *not* met.
2		WAS: ERROR		
3		WS4: ERROR		
Post-condition/ Resource group states		WAS: ERROR WS4: ERROR	• DB2®: ◦ WS1: ◦ WS2: ◦ WS3:	WAS: DB2®: • WS1: • WS2: • WS3: • WS4

WAS and WS4 could have started on Node 1, but the parent resource group DB2® is still in the offline state. Therefore, WAS and WS4 are put in the ERROR state.

Use case 2: Start second node (Node 2)

> **Note:** Cluster state as in the post-condition from the use case 2.

Step/ Qualifier	Action	Node 1	Node 2	Node 3
1	Start Node 2			
2			Acquire DB2®	
3		Acquire WAS		
4		Acquire WS4	Acquire WS1, WS2, WS3	
Post-condition/ Resource group states		WAS:ONLINE WS4: ONLINE	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	WAS: DB2®: • WS1: • WS2: • WS3: • WS4:

Node 2 starts DB2® (the parent RG), which in turn triggers processing of WAS (child of DB2®). Finally all the grandchildren are started on their respective nodes.

Consolidated view of start node sequence 1, 2, 3

Step	Node 1	Node 2	Node 3
Start node 1	WAS: ERROR WS4: ERROR	DB2®: • WS1: • WS2:	WAS: DB2®: • WS1:

Step	Node 1	Node 2	Node 3
		• WS3:	• WS2: • WS3: • WS4:
Start node 2	WAS: ONLINE WS4: ONLINE	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	WAS: DB2®: • WS1: • WS2: • WS3: • WS4:
Start node 3	WAS: ONLINE WS4: ONLINE	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: OFFLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: OFFLINE

Use case 3: Start nodes out of order (Node 3)

Note: All cluster nodes and resource groups are in the offline state.

Step/ Qualifier	Action	Node 1	Node 2	Node 3	Comments
1	Start Node 3				
2				Acquire DB2®	
Post-condition/ Resource group states		WAS: DB2®: WS4	WS1: • WS2: • WS3:	• WAS: ERROR • DB2®: ONLINE • WS1: ERROR • WS2: ERROR • WS3: ERROR • WS4: ERRORli	

Node 3 is a participating node for all the resource groups. However, WAS and DB2® cannot coexist on the same node. DB2® - being a parent - is started on Node 3, which means that WAS cannot be started on the same node. Since WAS is *not* online none of the children of WAS can come online on Node 3.

Use case 4: Start second node out of order (Node 2)

> **Note:** Cluster and RG states as at the end of the previous use case.

Step/ Qualifier	Action	Node 1	Node 2	Node 3
1	Start Node 2			
2				Release DB2®
3			Acquire DB2®	
4				Acquire WAS
			Acquire WS1, WS2, WS3	Acquire WS4
Post-condition/ Resource group states		WAS: WS4	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: ONLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: ONLINE

Node 2 is the higher priority node for DB2®. Therefore DB2® falls back to Node 2 and WAS (Online On Different Nodes Dependency set) can now be acquired on Node 3.

Use case 5: Start third node (Node) 1

> **Note:** Cluster and RG states as at the end of the previous use case.

Step/ Qualifier	Action	Node 1	Node 2	Node 3
1	Start Node 1			
2			Release WS1, WS2, and WS3	Release WS4
3		Acquire WAS		
4		Acquire WS4		
5			Acquire WS1, WS2 and WS3	
Post-condition/ Resource group states		WAS: ONLINE WS4: ONLINE	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: OFFLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: OFFLINE

All groups are now online.

Consolidated view of start node sequence: 3, 2, 1

Step	Node 1	Node 2	Node 3
Start Node 3	WAS: WS4:	DB2®: • WS1: • WS2: • WS3:	• WAS: ERROR • DB2®: ONLINE • WS1: ERROR • WS2: ERROR • WS3: ERROR • WS4: ERROR
Start Node 2	WAS: WS4:	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: ONLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: ONLINE
Start Node 1	WAS: ONLINE WS4: ONLINE	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: OFFLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: OFFLINE

Use case 6: Acquisition failure example

> **Note:** Node 1 is offline and all resource groups are ONLINE on Nodes 2 and 3.

Step / Qualifier	Action	Node 1	Node 2	Node 3	Comments
		WAS: WS4:	• DB2®: ONLINE • WS1: ONLINE • WS2: ONLINE • WS3:ONLINE	• WAS: ONLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: ONLINE	
1	Node_up 1				

Step / Qualifier	Action	Node 1	Node 2	Node 3	Comments
2			Release WS1 WS2 WS3 Release	Release WS4	
3				Release WAS	
4		Acquire WAS			Acquisition Failure for WAS
5		rg_move WAS			Normal rg_move event
6				Acquire WAS	
7			Acquire WS1 WS2 WS3 Acquire	Acquire WS4	
Post condition/ Resource group states		WAS:OFFLINE WS4: OFFLINE	DB2®: ONLINE WS1: ONLINE • WS2: ONLINE • WS3: ONLINE	• WAS: ONLINE • DB2®: OFFLINE • WS1: OFFLINE • WS2: OFFLINE • WS3: OFFLINE • WS4: ONLINE	

As Node 1 joins the cluster, WAS attempts to fallback but gets the acquisition failure. The acquisition failure launches a **resource_state_change** event; this triggers an **rg_move** event, which moves WAS to its original node.

Using DLPAR and CoD in a PowerHA® SystemMirror® cluster

You can configure PowerHA® SystemMirror® in a hardware and software configuration to use Dynamic Logical Partitions (DLPARs) and the Capacity on Demand (CoD) function.

The Resource Optimization High Availability feature in PowerHA® SystemMirror® manages DLPAR and CoD functions. CoD resources for PowerHA® SystemMirror® are composed of On/Off CoD resources and Enterprise Pool CoD resources.

Overview of DLPAR and CoD

You can use IBM® Power Systems™ to configure multiple Logical Partitions (LPARs) on a single physical frame, where each of the LPARs behaves as a standalone IBM® Power Systems™ processor. Using this configuration, you can install and run multiple applications on different LPARs that use a single physical hardware component.

The applications running on LPARs are completely isolated from each other at the software level. Each LPAR can be optimally tuned for a particular application that runs on it.

In addition, Dynamic Logical Partitioning (DLPAR) allows you to dynamically allocate additional resources (such as memory and CPUs) to each logical partition, if needed, without stopping the application. These additional resources must be physically present on the frame that uses logical partitions.

PowerHA® SystemMirror® can dynamically use the Capacity on Demand (CoD) resources (On/Off CoD resources or Enterprise Pool CoD resources) and activate or allocate these resources. This process allows the frame to receive more configurable resources that can be allocated to the LPAR through DLPAR operations. The Resource Optimized High Availability (ROHA) function in PowerHA® SystemMirror® manages DLPAR and CoD operations.

LPAR, DLPAR, and CoD terminology

Logical Partition (LPAR)

The division of a computer's processors, memory, and hardware resources into multiple environments so that each environment can be operated independently with its own operating system and applications. The number of logical partitions that can be created depends on the system. Typically, partitions are used for different purposes, such as database operation, client/server operations, web server operations, test environments, and production environments. Each partition can communicate with the other partitions as if each partition is a separate machine.

Dynamic Logical Partitioning (DLPAR)

A facility in some IBM® Power Systems™ processors that provide the ability to logically attach and detach a managed system's resources to and from a logical partition's operating system without rebooting the system. The following features are available in a DLPAR:

Capacity on Demand (CoD)

A feature of IBM® Power Systems™ servers that you can use to activate preinstalled but inactive processors when resource requirements change.

Dynamic Processor Deallocation

A feature of IBM® Power Systems™ servers and some SMP models. The processor is taken offline dynamically when an internal threshold of recoverable errors is exceeded. DLPAR allows substitution of the inactive processor, for the processor that is suspected of being defective. This online switch does not affect applications and kernel extensions. This function is not supported by PowerHA® SystemMirror®.

Cross-partition workload management

A feature that is used to manage system resources across partitions. This function is not supported by PowerHA® SystemMirror®.

Capacity on Demand (CoD)

A function in some IBM® Power Systems™ processors you can use to acquire, but not pay for a fully configured system. The additional CPUs and memory, while physically present, are not used until you decide that the additional capacity you need is worth the cost. This provides you with a fast and easy upgrade in capacity to meet peak or unexpected loads. CoD is composed of the following resources:

> **Note:** The following resources are managed through the ROHA function in PowerHA® SystemMirror®.

On/Off CoD

Resources that are preinstalled in your system, but you have not paid for the resources or activated the resources. You can use this type of CoD license to temporarily activate resources.

Trial CoD

Resources that are available to be used for a limited number of days. You do not have to pay for this type of CoD license.

Enterprise Pool CoD (EPCoD)

Resources that can move between systems in the same EPCoD pool to where the resources are needed. Physical resources such as CPU and memory are not moved between systems, but rights to access the physical resources are moved between systems. The rights to use resources are shared across systems. You can allocate the resources where they are required.

Hardware Management Console (HMC)

An interface that you can use to collect CoD system profile information and enter activation codes for CoD. You must manually enter the activation codes for CoD in the HMC.

The HMC also manages all DLPAR, On/Off CoD, Trial CoD, and EPCoD operations for the LPARs created on the CEC frame. PowerHA® SystemMirror® automatically performs all DLPAR, On/Off CoD, Trial CoD, and EPCoD operations for starting, stopping, and moving resource groups.

For integration with PowerHA® SystemMirror®, HMC must have a TCP/IP connection to the LPAR and a configured IP label through which a connection is established. Also, the Secure Shell (SSH) link must be established between all LPARs and the HMC. The **lshmc** command displays the HMC configuration.

ROHA function uses the AIX® **`ping`** command to verify connectivity with HMC.

Managed System

Am IBM® Power Systems™ that is LPAR-capable and that is managed by an HMC.

CoD Vital Product Data (VPD)

A collection of system profile information that describes the hardware configuration and identification numbers. In this document, VPD refers to CoD VPD.

Activation Code (or License Key)

A password that is used to activate a processor that is inactive (standby) or to activate memory in CoD. Each activation code is uniquely created for a system and requires the system Vital Product Data (VPD) to ensure correctness.

> **Note:** In the PowerHA® SystemMirror® SMIT interface and in the PowerHA® SystemMirror® documentation, the activation code is also referred to as the license key.

PowerHA® SystemMirror® integration with the CoD function

By integrating with DLPAR and CoD, PowerHA® SystemMirror® ensures that each node can support the application with reasonable performance at a minimum cost. You can use the On/Off CoD function to upgrade the capacity of the logical partition when your application requires more resources, without having to pay for idle capacity until you need it. You can use Enterprise Pool CoD (EPCoD) to share resources across systems in the same mirror pool.

You can configure cluster resources so that the logical partition with minimally allocated resources serves as a standby node, and the application resides on another LPAR node that has more resources than the standby node. This way, you do not use any additional resources that the frames have until the resources are required by the application.

When it is necessary to run the application on the standby node, PowerHA® SystemMirror® ensures that the node has sufficient resources to successfully run the application. The resources can be dynamically allocated from the free pool. The DLPAR function provides the resources to the standby node, by allocating the resources that are available in the free pool on the system.

If there are not enough available resources in the free pool that can be allocated through DLPAR to the standby node, PowerHA® SystemMirror® can dynamically provision resources from the EPCoD resource pool or the On/Off resource pool. When the free pool contains enough resources, these resources are allocated by PowerHA® SystemMirror® through a DLPAR operation for the standby node.

The On/Off CoD resources are temporary resources that you pay for on a per consumption basis. EPCoD resources are permanent resources that are a onetime payment and has unlimited use. To reduce costs, PowerHA® SystemMirror® always gets resources from the EPCoD resource pool before getting resources from the On/Off resource pool. PowerHA® SystemMirror® always releases resources to the On/Off resource pool before releasing resources to the EPCoD resource pool.

The following table displays all available types of CoD for PowerHA® SystemMirror®.

CoD type	PowerHA® SystemMirror® Version 7.1	PowerHA® SystemMirror® Version 7.2
On/Off CoD	• CPU: Yes • Memory: No	• CPU: Yes • Memory: Yes
Trial CoD	Yes	Yes
Enterprise Pool CoD	No	Yes

You can configure a PowerHA® SystemMirror® cluster within one or more IBM® Power Systems™ servers, using two or more logical partitions. You can also configure a cluster on a subset of LPARs within one frame. Or, the cluster can use partitions from two or more frames, where the nodes can be defined as a subset of LPARs from one frame and a subset of LPARs from another frame, all connected to one or more HMCs. The following figure illustrates a typical two-frame configuration:

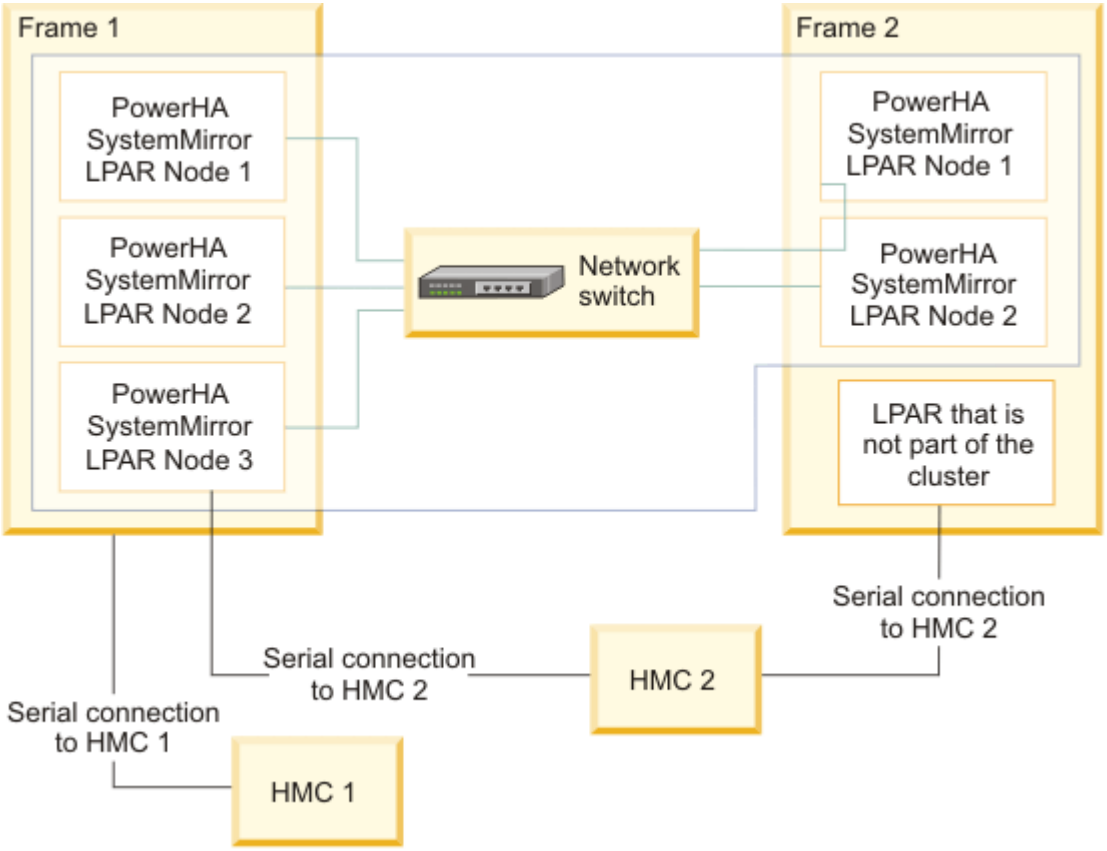


Figure 11: PowerHA® SystemMirror® cluster configuration with LPARs

Terminology for resource types and memory allocation

The following terms can help you to distinguish between different types of resources allocation that can occur in a PowerHA® SystemMirror® cluster that uses DLPAR and CoD functions:

Installed amount of resources

The number of CPUs and the amount of memory that is physically present on the frame. Although these resources are physically present, they are only usable if they are configurable resources.

Configurable amount of resources

The number of CPUs and the amount of memory that is physically available for use by all LPARs on the frame. This amount includes all permanent resources, paid for resources, and CoD resources that are active on the frame (On/Off or EPCoD).

Inactive amount of resources

The difference between the installed amount of resources and the configurable amount of resources. Also called unlicensed resources. You must have an inactive amount of resources to be able to use CoD resources.

Free pool amount of resources

The number of CPUs and the amount of memory that can be dynamically allocated by PowerHA® SystemMirror® through the HMC to an LPAR that requires more resources. The free pool amount of resources is the difference between the configurable amount of resources on the frame minus the resources that are current being used by the LPAR.
The free pool includes resources on a particular frame only. For example, if a cluster is configured with LPARs that reside on frames A and B, PowerHA® SystemMirror® does not request resources from a free pool on frame B for an LPAR that resides on frame A.

Enterprise Pool CoD (EPCoD) amount of resources

The number of CPUs and the amount of memory that is available in the EPCoD resource pool that can be allocated by PowerHA® SystemMirror® if the frame requires more resources. When EPCoD resources are allocated to a frame, the amount of configurable resources on the frame is increased. The increase of configurable resources makes it possible for a DLPAR operation to provide more resources to an LPAR.

You can use the **lscodpool** command on the HMC to view the characteristics on the EPCoD resource pool. You can also use the **clmgr view report roha** command on a PowerHA® SystemMirror® node to view general information that is related to the Resource Optimized High Availability (ROHA) data.

> **Note:** The EPCoD resource pool includes resources for multiple frames.

On/Off CoD amount of resources

The number of CPUs and the amount of memory that is available in the On/Off resource pool that can be allocated by PowerHA® SystemMirror® if the frame requires more resources. When On/Off CoD resources are allocated to a frame, the amount of configurable resources on the frame is increased. The increase of configurable resources makes it possible for a DLPAR operation to provide more resources to an LPAR. You can use the **lscod** command on the HMC to view the characteristics on the On/Off CoD resource pool. You can also use the **clmgr view report roha** command on a PowerHA® SystemMirror® node to view general information that is related to the ROHA data.

> **Note:** The On/Off CoD resource pool includes resources for a single frame.

LPAR minimum amount

The minimum amount (or quantity) of a resource, such as CPU or memory, that an LPAR requires to be brought online or started. The LPAR does not start unless it meets the specified LPAR minimum. When DLPAR operations are performed between LPARs, the amount of resources that are removed from an LPAR cannot go less than this value. This value is set on the HMC on the LPAR profile and is not modified by PowerHA® SystemMirror®. All allocations that are performed by PowerHA® SystemMirror® are computed starting for this minimum value.

LPAR desired amount

The desired amount of a resource that an LPAR acquires when it starts, if the resources are available. This value is set on the HMC on the LPAR profile and is not modified by PowerHA® SystemMirror®.

LPAR maximum amount

The maximum amount (or quantity) of a resource that an LPAR can acquire. When DLPAR operations are performed, the amount of resources added to an LPAR cannot go more than this value. This value is set on the HMC on the LPAR profile and is not modified by PowerHA® SystemMirror®. Use the **lshwres** command on the HMC to verify the minimal, desired, and maximum values. You can also use the **clmgr view report roha** command on a PowerHA® SystemMirror® node to view general information that is related to the ROHA data.

Types of CoD licenses

There are different types of Capacity on Demand (CoD) licenses that are available for PowerHA® SystemMirror®.

The following table indicates the type of CoD licenses and whether PowerHA® SystemMirror® allows the use of a particular license.

The table describes the different types of CoD licenses and if they are supported by PowerHA® SystemMirror®.

Types of CoD licenses

License type	Description	Supported by PowerHA® SystemMirror®	Comments
On/Off CoD	**CPU and Memory:** When active these resources can be used as temporary resources until they expire.	**CPU:** Yes **Memory:** Yes	PowerHA® SystemMirror® does not manage this license. You must enter the On/Off CoD license in the Hardware Management Console (HMC) before PowerHA® SystemMirror® can use

License type	Description	Supported by PowerHA® SystemMirror®	Comments
	A user pays for these resources on a consumption basis. With this license, your system can use a processor or memory for a predetermined number of days. For example, if you purchase 300 processor days, you can use either 30 processors for 10 days or 10 processors for 30 days.		this type of licenses. PowerHA® SystemMirror® can dynamically activate (On) or deactivate (Off) CoD resources. When resources are activated, PowerHA® SystemMirror® can allocate the resources to a logical partition through a DLPAR operation. By default, On/Off CoD resources are activated for 30 days. You can change the number of days On/Off CoD is activated in SMIT by running the **smitty cm_cfg_def_cl_tun** command.
Trial CoD	**CPU and Memory:** The Trial CoD resources are activated for a single period of 30 consecutive days. If your system was configured with Trial CoD features and if the features are not activated, you can turn on the features for a trial period. With the Trial CoD capability, you can gauge how much capacity you might need in the future.	**CPU:** Yes **Memory:** Yes	PowerHA® SystemMirror® does not manage this license. You must enter the Trial CoD license in the (HMC) before PowerHA® SystemMirror® can use this type of licenses. Trial CoD resources are temporary resources. PowerHA® SystemMirror® does not dynamically put temporary resources in an On or Off state. When the Trial CoD license is entered into HMC, the associated resources are put in an On state. The configured Trial CoD resources are included in the calculation of available DLPAR resources.
Enterprise Pool CoD (EPCoD)	You pay for the EPCoD resources once and you can use these resource without any limits.	**CPU:** Yes **Memory:** Yes	PowerHA® SystemMirror® does not manage this license. You must enter the EPCoD license in the (HMC) before PowerHA® SystemMirror® can use this type of license. EPCoD resources can move between systems in the same EPCoD pool to where the resources are needed. Physical resources such as CPU

License type	Description	Supported by PowerHA® SystemMirror®	Comments
			and memory are not moved between systems, but rights to access the physical resources are moved between systems. The rights to use resources are shared across systems, but only one system can use it at a time. To use the EPCoD license, your system must be using HMC 7.8, or later. **Note:** PowerHA® SystemMirror® does not support a blank EPCoD pool name or an EPCoD pool name that has space. For example, if you use the EPCoD pool name as **POOL NAME**, an error message is displayed. To rename the EPCoD pool name from **POOL NAME** to **POOL_NAME**, enter the following command: `chcodpool -o update -p "POOL NAME" -a new_name="POOL_NAME"`

Resource Optimized High Availability in PowerHA® SystemMirror®

Resource Optimized High Availability (ROHA) is a function in PowerHA® SystemMirror® that automatically and dynamically manages DLPAR, Enterprise Pool CoD (EPCoD), and On/Off CoD resources. You can configure ROHA with the Hardware Management Console (HMC), hardware resource provisioning, and cluster tunable configurations.

Before you use ROHA, you must determine which HMC manages a specific LPAR and any LPARs that you plan to use in the future. You must also plan for the necessary resources for your applications, identify your workloads, and your requirements for physical resources (CPU cores, virtual CPUs, and memory). After you identify all these requirements, you must configure ROHA.

Before using the ROHA function for the first time, you must complete the following steps:

1. Create an HMC for each LPAR that you identified by running the **`clmgr add hmc`** command.
2. Create hardware resource provisioning for each application controller that you identified for your workloads by running the **`clmgr add roha`** command.

> **Note:** When you provision resources for the first time, you must agree or not agree to use On/Off CoD function. If you agree to use the On/Off CoD function, you are billed for any extra costs. The On/Off CoD agreement is displayed again only if you did not previously accept the agreement. However, you can use the ROHA function without accepting the On/Off CoD agreement. In this case, ROHA uses only the DLPAR and EPCoD operations.

Depending on your environment configuration, you might have to complete the following optional tasks:

- Define HMCs at the site level or node level if your topology requires them.
- Change the default HMC tunable values such as retry count, retry delay, or timeout on DLPAR operations.
- Verify the HMC list that is used by the cluster and by each node.
- Change the clusters ROHA tunables for DLPAR, Enterprise Pool CoD, and On/Off CoD acquisition and release operations.

The following figure describes the high-level details for configuring ROHA:

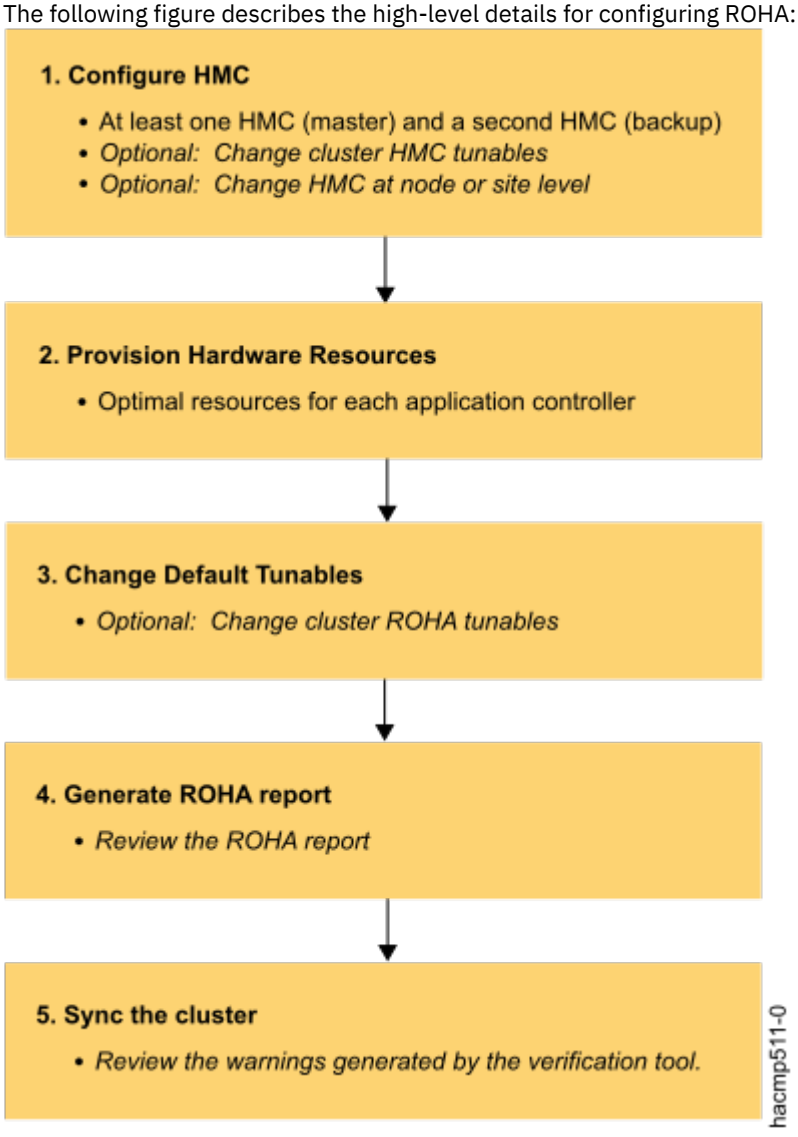


Figure 12: ROHA configuration steps

Planning for Resource Optimized High Availability

If you plan to use the Resource Optimized High Availability (ROHA) function in a PowerHA® SystemMirror® cluster, you must plan and allocate resources to the LPARs through the HMC. You must also be familiar with the types of Capacity on Demand (CoD) licenses that are available.

Review the following planning information about ROHA in a PowerHA® SystemMirror® cluster:

- Obtain the LPAR resources information and resource group policies information:
  - How much memory and resources the applications that are supported by your cluster require when they run on their regular hosting nodes. Under normal running conditions, check how much memory and what number of CPUs each application uses to run with optimum performance on the LPAR node on which its resource group resides normally (home node for the resource group).
  - Determine the startup, fallover, and fallback policies of the resource group that contains the application controller by using the **clRGinfo** command. This identifies the LPAR node to which the resource group will fall over, in case of a failure.
  - How much memory and what number of CPUs are allocated to the LPAR node on which the resource group will fall over, in case of a failure. This LPAR node is referred to as a standby node. With these numbers in mind, consider whether the application's performance will be impaired on the standby node, if the application is running with fewer resources.
  - Check the existing values for the LPAR minimums, LPAR maximums, and LPAR desired amounts (resources and memory) specified by using the **lshwres** or **pvmctl** command on the standby node.
- Estimate the resources that are required for the application:
  - For each standby node that can host a resource group, you must estimate the optimal amount of resources (CPU and memory) that this node requires for the application to run successfully. The optimal amount of resources that you identify are specified in PowerHA® SystemMirror®. PowerHA® SystemMirror® verifies that the optimal amount is contained within the boundaries of the LPAR maximums that are configured outside of PowerHA® SystemMirror® for each LPAR.
  - When you specify for an application to use resources through the DLPAR operation, PowerHA® SystemMirror® dynamically activates CoD resources from either Enterprise Pool CoD or On/Off CoD resource pools. When the application no longer requires these extra resources, they are returned to the corresponding free pool.
- Revise existing pre-event and post-event scripts that were used to allocated DLPAR resources.
- The following limitations are applicable for ROHA multi-cluster deployments. A multi-cluster deployment is an environment with multiple clusters deployed across two different systems. In a multi-cluster deployment environment, each cluster has one node on an active system, and another node on the standby system.
  - You can have up to 45 clusters in a ROHA multi-cluster deployment. For example, you can have 45 LPARs hosting nodes on the active system and 45 LPARs hosting nodes on the standby system. In this example, if a failure occurs on the active system, the 45 LPARs on the standby system independently contact the PowerVM® NovaLink or HMC simultaneously to obtain resources.
  - Enterprise Pool CoD is supported, but you must verify that your system has enough available resources to meet resource requirements for all clusters and possible resource conflicts. You should plan to allocate an additional 15% amount of the total resource requirement for all clusters on the system.
  - On/Off CoD is not supported in a multi-cluster deployment.
- ROHA does not support resource groups that have a startup policy of `Online on All Available Nodes`.
- In PowerHA® SystemMirror® Version 7.2.1 SP1, or later, ROHA supports dynamic automatic reconfiguration (DARE). With DARE, you can change CoD resource requirements for application controllers without bringing down a workload. The resource changes that you specify are synchronized across the cluster.
- In PowerHA® SystemMirror® Version 7.2.1 SP1, or later, ROHA supports Live Partition Mobility (LPM) to migrate partitions from one system to another. Before you implement LPM, you must verify the following information about the target system, target HMC, and target PowerVM® NovaLink:
  - Before you start the LPM process, you must verify and synchronize all nodes in the target system and the source system.
  - The source system and the target system must be part of the same Enterprise Pool CoD.
  - If the target system is running HMC version 8.40, or earlier, you must establish connections for the primary HMC and the backup HMC. If the target system is running HMC version 8.50, or later, only a connection to the primary HMC is required.

- The resources on the source system and the target system must be similar. If the target system has less resources than the source system, a failover might occur. Before the LPM process starts, you must verify that the target system has enough available resources.
- The target HMC and the PowerVM® NovaLink must be defined in the PowerHA® SystemMirror® cluster.

- In PowerHA® SystemMirror® Version 7.2.1 SP1, or later, ROHA supports resources that can be released asynchronously if the source node and the target node are located on different systems.

Note: If you were using LPAR nodes in your cluster before using the ROHA function, you might need to revise and rewrite your existing pre-event and post-event scripts.

Prerequisites for using ROHA

Before using the ROHA function in PowerHA® SystemMirror®, you must review the following information:

Verify Software and hardware levels

You must verify that the system is configured to use the required software and hardware for the DLPAR, On/Off CoD, and EPCoD functions. HMC version 8.40 uses different architecture than HCM version 8.50, or later, to manage Enterprise pools. Therefore, you cannot mix HMC version 8.40 with HMC version 8.50, or later.

Verify LPAR node name

The hostname of the AIX® operating system and the HMC LPAR do not have to match if your environment is configured as follows:

- IBM® AIX® 7.1 with Technology Level 4, or later, or AIX® Version 7.2, or later
- Power® Firmware SC840 for Enterprise POWER8® (E870 and E880)

For any other environment configurations, the host name of the AIX® operating system and the HMC LPAR name must match.

Verify what DLPAR resources are available and what CoD licenses you can access

PowerHA® SystemMirror® does not identify what resources are available. PowerHA® SystemMirror® has no control over whether the resources are physically available on Power Systems™ or whether they are unallocated and available. PowerHA® SystemMirror® provides dynamic allocations only for CPU and Memory resources. PowerHA® SystemMirror® does not support dynamic changes of the I/O slots.

Identify the type of CoD function

Create the EPCoD on the HMC. Enter the license key (also called activation code) for either the EPCoD or On/Off CoD on the HMC.

Establish secure connections to the HMC

PowerHA® SystemMirror® must communicate securely with the LPAR nodes through HMC. You must install SSH for PowerHA® SystemMirror® to access the HMC without entering a user name and password. If you want to use SSH for a secure connection, from the HMC, select **Users and Security › Systems and Console Security**, and then click **Enable Remote Command Execution**. In the Remote Command Execution window, select **Enable remote command execution using the ssh facility**. The AIX® operating system must have SSH installed to generate the public and private keys.

Note: PowerHA® SystemMirror® uses the root user on the cluster node to issue the SSH commands. HMC system runs command based on login credentials that are specified by the user. However, if no user is specified, PowerHA® SystemMirror® uses the default hscroot user.

Verify HMC access

You must verify that all nodes and LPARs can access the HMC that manages all nodes and LPARs. If the node or LPAR cannot access the HMC, the resource allocation fails.

Note: PowerHA® SystemMirror® ROHA does not releases more resources than the resources received from the resource release operation while using EPCoD or On/Off CoD.

Configuring HMC or PowerVM® NovaLink to work with Resource Optimized High Availability

If you are running PowerHA® SystemMirror® Version 7.2.2 for AIX® or later, and if you want to use the Resource Optimized High Availability (ROHA) function, you must configure each HMC or PowerVM® NovaLink LPAR link to use Secure Shell (SSH). However, if you are running PowerHA® SystemMirror® Version 7.2.2 for AIX® or later, you can also use Representational State Transfer (REST) application programming interface (API) to connect to the HMC instead of SSH. You must also configure a backup HMC. The ROHA function can be used only with HMC version 8.40, or later.

SSH communication with HMC or PowerVM® NovaLink

For systems that are running PowerHA® SystemMirror® Version 7.2.2 for AIX® or later, LPARs must use SSH to communicate with the Hardware Management Console (HMC) or PowerVM® NovaLink.

You must configure SSH to not require a password when PowerHA® SystemMirror® is communicating with the HMC or PowerVM® NovaLink. To configure SSH communication, you can run the **ssh-keygen** command on each LPAR node to generate a public and private key pair. The public key must be copied to the HMCs or PowerVM® NovaLink public key file that is authorized. The following example displays how to set up SSH from the LPAR:

```
# /usr/bin/ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save key (//.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in //.ssh/id_rsa.
Your public key has been saved in //.ssh/id_rsa.pub.
The key fingerprint is:
9c:00:9f:61:d9:40:60:0c:1d:6b:89:ac:f9:8e:fc:f5 root@4ndc1
# mykey=`cat ~/.ssh/id_rsa.pub`
# ssh hmcuser@cuodhmc mkauthkeys -a \"$mykey\"
```

Verify that the SSH communication is configured correctly by running the **ssh <HMCuser>@hmcname ls /tmp** command on LPAR, where **HMCuser** is the login ID configured in the HMC and **hmcname** is the name of the HMC or PowerVM® NovaLink.

Note: PowerHA® SystemMirror® Version 7.2.5 for AIX® or later supports nonroot user to connect with the HMC by using SSH communications. PowerHA® SystemMirror® Version 7.2.4 for AIX®, or earlier, only supports hscroot as HMC username.

To add nonroot user support, run the **clmgr** command.
To add a username, while adding the HMC definition to PowerHA® SystemMirror® configuration, enter the following command:

```
clmgr add hmc <HMC> USER_NAME=<hmcuser>
```

To add a username, if HMC is already added to the PowerHA® SystemMirror® configuration, enter the following command:

```
clmgr modify hmc <HMC> USER_NAME=<hmcuser>
```

REST API communication with HMC

In PowerHA® SystemMirror® Version 7.2.2 for AIX®, or later, you can use SSH or a REST API to communicate securely with an HMC. The REST APIs are provided in HMC Version 8.40, or later, and are based on HTTP

protocols. An HMC user should have privileges to run the query and to acquire, and release resource operations when a nonroot user is used.

You can configure the REST API by using one of the following methods:

SMIT

1. From the command line, enter `smit sysmirror`.
2. Select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › HMC Configuration › Change/Show Default HMC Tunables**, and press Enter.
3. Change Connection Type as REST API.

clmgr command

To specify the connection type, run the `clmgr manage cluster hmc` command.
Example: `clmgr manage cluster hmc CONNECTION_TYPE=rest`

To add the user name and password, enter the following command:

```
clmgr add hmc
```

To modify the user name and password, enter the following command:

```
clmgr modify hmc
```

To view help information about modifying the HMC parameters, enter the following command:

```
clmgr manage cluster hmc –h
```

HMC REST API support

1. You must use the HMC Version 8.4.3, HMC Version 8.5.3, HMC Version 8.6.2, or later.
2. You can download the **curl** package from the IBM® **AIX® Toolbox for Linux® Applications** website. You also need to install ca-certificates during **curl** package installation. The **ssl** and **crypto** library versions that are shipped with the AIX®® operation system must be compatible with the **curl** package. The following versions of AIX® operating system are tested:
   - AIX® 7.2 with Technology Level 0 with SP 2
   - AIX® 7.2 Technology Level 1 with SP 1
   - AIX® 7.1 with Technology Level 4 with SP 4

   > **Note:** The port number that is used for HMC REST API is 12443

   The latest versions of HMC supports both 12443 and 443 ports. Starting with PowerHA® SystemMirror® Version 7.2.8, or later, you can modify the HMC port by making following entry in the `/etc/environment` file.

   ```
   HMC_PORT=<port number>
   ```

   By default, the port 12443 is used if the port entry is not made in the `/etc/environment` file.

REST API communication with PowerVM® NovaLink

In PowerHA® SystemMirror® Version 7.2.2, you cannot use REST API to communicate with the PowerVM® NovaLink. Therefore, the PowerHA® SystemMirror® cannot verify the REST API communication.

Timeout and retry mechanism

If the HMC cannot get the CEC lock, the LPAR requests that to the HMC might fail. If a failure occurs between the LPAR and the HMC, PowerHA® SystemMirror® tries to re-establish a connection. You can configure the

frequency of times PowerHA® SystemMirror® attempts to establish a connection by using one of the following methods:

SMIT

1. From the command line, enter `smit sysmirror`.
2. Select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › HMC Configuration › Change/Show Default HMC Tunables**, and press Enter.

clmgr command

To display the HMC parameters, run the `clmgr query cluster roha` command.

To view help information about modifying the HMC parameters, run the `clmgr manage cluster hmc -h` command.

> **Note:** HMC 8.2, or later, queues all HMC commands if the HMC cannot get the CEC lock. This function suppresses the requirement to retry and establish a connection. Therefore, the retry mechanism parameters are only used if the HMC is 8.2 or earlier.

If the PowerVM® NovaLink is not reachable, the LPAR requests that to the PowerVM® NovaLink mightPowerVM® fail. If a failure occurs between the LPAR and the PowerVM® NovaLink, PowerHA® SystemMirror® retries to establish a connection. You can configure the frequency of times PowerHA® SystemMirror® attempts to establish a connection by using one of the following methods:

SMIT

1. From the command line, enter `smit sysmirror`.
2. Select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › NovaLink Configuration › Change/Show Default NovaLink Tunables**, and press Enter.
3. Change **Connection Type** as **REST API**.x

clmgr command

To display the PowerVM® NovaLink parameters, run the `clmgr query cluster nova` command.

To view help information about modifying the PowerVM® NovaLink parameters, run the `clmgr manage cluster nova -h` command.

Configuring a backup HMC

You can configure more than one HMC. If one HMC fails to respond, the Resource Optimized High Availability (ROHA) function can switch to another HMC. The HMCs are used in the order they are listed in the HMC list. For example, in the HMC list, if you have three systems that are listed in the following order: *HMC1*, *HMC2*, and *HMC3*. If *HMC1* fails, ROHA stops trying to communicate with *HMC1* and starts to communicate with the next HMC in the list (*HMC2*). The currently used HMC is saved in the cluster configuration so that the ROHA function skips any failing HMCs and communicates with a working HMC. At the end of the session, the falling HMC information is cleared and *HMC1* is listed as the first HMC system in the list again.

To change the order of HMCs in a list, use one of the following methods:

SMIT

1. From the command line, enter `smit sysmirror`.
2. Select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › HMC Configuration › Change/Show Default HMC Tunables**, and press Enter.

clmgr command

To display the HMC parameters, run the `clmgr query cluster hmc` command.

To view help information about modifying the HMC parameters run, the `clmgr manage cluster hmc -h` command.

Primary HMC and backup HMC

In HMC Version 8.40, or earlier, PowerHA® SystemMirror® can perform an Enterprise Pool CoD (EPCoD) operation only on a primary HMC. An EPCoD can work with a pair of HMCs. One HMC is the primary and performs changes on the EPCoD, and the other HMC is the backup that can only send query requests to the EPCoD. If the primary HMC fails (does not respond or is not functioning), PowerHA® SystemMirror® uses another HMC. The new HMC can be considered as the new primary HMC only if this new HMC was previously designated to the EPCoD as a backup HMC. If the HMC was designated to the EPCoD as the backup HMC, PowerHA® SystemMirror® can transform this backup HMC to the primary HMC. If the HMC was not designated to the EPCoD as the backup HMC, PowerHA® SystemMirror® cannot transform this backup HMC to the primary HMC unless you provide an XML file that contains the EPCoD definition that was used to create the EPCoD pool. Therefore, it is important to save this XML file in a directory that can be accessed easily.

> **Note:** The concept of a primary and backup HMC is applicable only for EPCoD pools and operations.

The HMC that you use to create the EPCoD pool is the primary HMC. When you create the primary HMC, you can create a second HMC and configure it as the backup HMC for the primary HMC. The second HMC can be a backup HMC only if it manages all of the servers that belong to the EPCoD pool.

If the primary HMC is online, you must use the primary HMC to set the backup HMC. You can set a backup HMC by running the **`chcodpool -p epcodpoolname -o update -a "backup_master_mc_name=backup_hmc"`** command, where *backup_hmc* is the name of the backup HMC.

In HMC Version 8.50, or later, all servers in a EPCoD pool are not required to be managed by the same HMC pair. When the HMC primary is defined, all EPCoD pool requests are routed through the primary HMC. All managing HMCs must have a network connection with the primary HMC. Requests that are initiated on the managing HMC are first sent to the primary HMC. If the primary HMC is not managing the EPCoD pool or if it is not connected to the target server, all requests are sent to an HMC that is managing that server. Servers that are managed by HMCs send events to the primary HMC. The primary HMC sends the EPCoD pool data to all other HMCs every time the EPCoD pool resources are updated.

> **Note:** The backup primary HMCs are deprecated in HMC Version 8.50, or later.

In HMC Version 8.50, or later, if partitions exist at the HMC group level and if two primary HMC are active at the same time, the resources might move to an `uncompliance` state when resource acquisition occurs. After communication is restored between first primary HMC and the new primary HMC, the new primary HMC takes over as the HMC primary for the entire EPCoD pool. In this scenario, the first primary HMC is removed and all HMCs that are managing the EPCoD pool are notified. The resources that are assigned to the servers in the EPCoD pool might not be same as the resources identified by the new primary HMC. Therefore, you must resync all resources with the new primary HMC.

> **Note:** Verify and synchronize the cluster after all changes have been made, use any of the following methods.
>
> **SMIT:**
>
> 1. From the command line enter `smit sysmirror`.
> 2. Select **Cluster Applications and Resources- › Verify and Synchronize Cluster configuration** and press **Enter**.
>
> **clmgr command:**
>
> To verify and synchronize the cluster, enter the following command:
>
> ```
> clmgr sync cluster
> ```

Configuring Resource Optimized High Availability

You can use the SMIT interface to configure the Resource Optimized High Availability function.

To configure Resource Optimized High Availability, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability**, and press Enter.
3. Select from the following options:
   HMC Configuration

   Configure the HMC that is used by your cluster. If you do not have an HMC associated with a node in the cluster, PowerHA® SystemMirror® uses the default HMC settings for the configuration.

   Hardware Resource Provisioning for Application Controller

   Configure the CPU and memory resources that are used for an application controller.

   Change/Show Default Cluster Tunables

   Configure the parameters for DLPAR, CoD On/Off, and Enterprise Pool CoD.

Adding an HMC definition for Resource Optimized High Availability

You can use the SMIT interface to add an Hardware Management Console (HMC) definition that is used by the PowerHA® SystemMirror® cluster.

To add an HMC definition, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › HMC Configuration › Add HMC Definition** , then select the **Change/Show Default HMC Tunables** and enter the new list of HMCs in the `HMC list` field and press Enter.
3. Complete the following fields, and press Enter.
   HMC name

   Enter the name or IP address for the HMC.

   DLPAR operations timeout

   Enter a timeout value in minutes for the DLPAR commands that are run on an HMC. If you do not specify a value, the default values, which are specified in the **Change/Show Default HMC Tunables** SMIT panel, are used. The default values are:

   ```
   DLPAR operations timeout (in minutes)                    10
   Number of retries                                        5
   Delay between retries (in seconds)                       10
   Connection Type                                          SSH
   ```

   Number of retries

   Enter the number of times you want one HMC command to be retried before the HMC is considered as not responding. The next HMC in the list is used after the number of retries that you entered have failed. If you do not specify a value, the default values, which are specified in the **Change/Show Default HMC Tunables** SMIT panel, are used.

   Delay between retries

   Enter the number of seconds you want to delay before attempting to retry to send an HMC command. If you do not specify a value, the default values, which are specified in the **Change/Show Default HMC Tunables** SMIT panel, are used.

Nodes

Enter the nodes that use the specified HMC. You only need to specify nodes in this field if the HMC is specific to these nodes. If the HMC is used by all nodes in the cluster, you do not need to specify any nodes in this field.

Sites

Enter the sites that use the specified HMC. You only need to specify sites in this field if the HMC is specific to these sites. If the HMC is used by both sites, you do not need to specify any sites in this field.

> All nodes that belong to the site use the HMC defined for the site.

User name

If you are using Security Shell (SSH) to establish a connection with the HMC, specify a user name. However, if you are using REST API to establish a connection with the HMC, specify a user name along with the password.

> If you create an HMC without specifying user name and password, the default `ssh` user ID is used for HMC communication.

> If you create an HMC without specifying any nodes or sites, the HMC is used by all nodes in the cluster and both sites.

Changing hardware provisioning for an application controller

You can use the SMIT interface to change the hardware provisioning for an application controller.

When an application requires additional resources to be allocated on a node, PowerHA® SystemMirror® determines whether only DLPAR resource from the free pool on the system are required, or Capacity on Demand (CoD) resources are also required. CoD resources can belong to EPCoD pool or the On/Off CoD pool.

During verification, PowerHA® SystemMirror® verifies that the resource values specified are less than LPAR maximum values for CPU and memory resources. PowerHA® SystemMirror® also verifies that the total of required resources for all application controllers that can run concurrently on the LPAR is less than the LPAR maximum. For example, if the LPAR node is already hosting application controllers that require additional DLPAR and CoD resources, it is possible that the LPAR cannot contain the additional resources because it has reached its LPAR maximum. In this case, PowerHA® SystemMirror® displays a warning message.

To change the hardware provisioning for an application controller, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optmized High Availability › Hardware Resource Provisioning for Application Controller › Change/Show Hardware Resource Provisioning of an Application Controller** , and press Enter.
3. Select the application controller you want to change from the list.
4. Change the following fields:
   Used desired level from the LPAR profile

   Select **Yes** if you want the LPAR hosting your node to reach the desired level specified in the LPAR profile. Select **No** if you want to enter a specific value for CPU and memory resources.
   If you have a mixed configuration with this field set to both **Yes** and **No**, the allocation performed is the sum of the LPAR profile value and the various optimal values that you specified.

   Optimal amount of gigabytes of memory

Enter the amount of memory that PowerHA® SystemMirror® attempts to allocate to the node before starting the specified application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. You can specify values in increments of 0.25 GB, 0.5 GB, 0.75 GB, or 1 GB. For example, a value of 1.5 represents 1.5 GB or 1536 MB. If the amount of memory is not satisfied, PowerHA® SystemMirror® performs recovery actions to move the resource group with its applications to another node, or PowerHA® SystemMirror® might allocate less memory depending on the setting for the **Start RG even if resources are insufficient** tunable.

Optimal number of dedicated processors

Enter the amount of processors that PowerHA® SystemMirror® attempts to allocate to the node before starting the application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. If the amount of CPUs is not satisfied, PowerHA® SystemMirror® performs recovery actions to move the resource group with its applications to another node, or PowerHA® SystemMirror® might allocate fewer CPUs depending on the setting for the **Start RG even if resources are insufficient** tunable.

Optimal number of processing units

Enter the amount of processing units that PowerHA® SystemMirror® attempts to allocate to the node before starting the application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. You can specify a value up to two decimal places in the range 0.01 - 255.99. This value is used only on nodes that support allocation of processing units. If the amount of processing units is not satisfied, PowerHA® SystemMirror® performs recovery actions to move the resource group with its applications to another node, or PowerHA® SystemMirror® might allocate fewer processing units depending on the setting for the **Start RG even if resources are insufficient** tunable.

Optimal number of virtual processors

Enter the amount of virtual processors that PowerHA® SystemMirror® attempts to allocate to the node before starting the application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. This value is used only on nodes that support allocation of processing units. If the amount of virtual CPUs is not satisfied, PowerHA® SystemMirror® performs recovery actions to move the resource group with its applications to another node, or PowerHA® SystemMirror® might allocate fewer virtual CPUs depending on the settings for the **Start RG even if resources are insufficient** tunable.

Changing the default cluster tunables

You can use the SMIT interface to change the settings for dynamic LPARs and On/Off CoD.

To change the default cluster tunables, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optmized High Availability › Change/Show Default Cluster Tunables**, and press Enter.
3. Change the following fields:
   Always start resource groups

   If you select **Yes**, the resource might be started when the total requested resource exceeds the maximum value for the LPAR profile or the combined available resources value. Therefore, selecting **Yes**, ensures the best possible allocation of resources. The default value for this field is **Yes**.
   If you select **No**, the resource groups are not started.

   Resource allocation order

   Specifies the order in which resources are allocated to the node. You can select the following options:

   Free pool before the Enterprise pool

   Select this option to first allocate resources from the free pool on the system. If you choose this option, the resources are allocated in the following order:

   a. PowerHA® SystemMirror® allocates resources from the free pool on the system.

b. If the free pool on the system has insufficient resources, PowerHA® SystemMirror® requests more resources from the Enterprise pool and the required resources are allocated.

c. If the Enterprise pool has insufficient resources, PowerHA® SystemMirror® requests more resources from other systems in the same Enterprise pool first and then requests the Capacity on Demand (CoD) pool to allocate the remaining resources.

Enterprise pool before free pool

Select this option to allocate resources from the Enterprise pool on the system. If you choose this option, the resources are allocated in the following order:

a. PowerHA® SystemMirror® allocates resources from the Enterprise pool on the system.

b. If the Enterprise pool on the system has insufficient resources, PowerHA® SystemMirror® considers resources from the free pool and allocates the required resources.

c. If the Enterprise pool has insufficient resources, PowerHA® SystemMirror® requests more resources from other systems in the same Enterprise pool first and then requests the Capacity on Demand (CoD) pool to allocate the remaining resources.

All Enterprise pool before free pool

Select this option to allocate resources from the Enterprise pool on the system. If you choose this option, the resources are allocated in the following order:

a. PowerHA® SystemMirror® allocates resources from the Enterprise pool on the system.

b. If the Enterprise pool has insufficient resources, PowerHA® SystemMirror® requests more resources from other systems in the same Enterprise pool first and then requests the Capacity on Demand (CoD) pool to allocate the remaining resources.

c. If the Enterprise pool on the system has insufficient resources, PowerHA® SystemMirror® considers resources from the free pool and allocates the required resources.

> When more resources are requested from an Enterprise pool, the state of the Enterprise pool might change to **Out of compliance**. For more information about Enterprise pool states, see Power® Enterprise Pool compliance.

The default option of this field is **Free pool before the Enterprise pool**. The resources are released in the reverse order in which they are allocated.

Adjust Shared Processor Pool size if required

Select **Yes** to authorize to dynamically change the maximum shared processor pool value. If required, the allocation process increases the maximum limit of the shared processor pool during the allocation process.
Select **no** to not have adjust the shared processor pool size.

Force synchronous release of DLPAR resources

Select **yes** to have release the CPU and memory resource synchronously. By default, automatically detects if the resource is released by checking whether the active node and backup node are on the same or different CECs.
Select **no** to not force synchronous release of DLPAR resources. Using asynchronous releases instead of synchronous releases does not cause a delay during the takeover process.

I agree to use On/Off CoD and be billed for extra costs

Select **yes** to have use the On/Off CoD function to obtain enough resources to fulfill the optimal amount of resources that your application requires. You must enter an activation code in the Hardware Management Console (HMC) to use the On/Off CoD function.
Select **no** if you do not want to use the On/Off CoD function.

Number of activating days for On/Off CoD requests

Specify the number of days that you want for activating On/Off CoD requests. If your environment has a multi-cluster deployment, you must use the same value for this field on all the different clusters. A

multi-cluster deployment is an environment with multiple clusters deployed across two different systems.

4. Verify and synchronize the cluster by using one of the following methods.
   SMIT:
   a. From the command line enter `smit sysmirror`.
   b. Select **Cluster Applications and Resources- › Verify and Synchronize Cluster configuration** and press **Enter**.

   clmgr command:

   To verify and synchronize the cluster, enter the following command:

   ```
   clmgr sync cluster
   ```

Troubleshooting Resource Optimized High Availability

You can use commands to troubleshoot Resource Optimized High Availability (ROHA) operations in your cluster.

To view of all data related to ROHA, run the **`clmgr view report roha`** command.

To run consistency checks on all nodes in your cluster, run the **`clmgr verify cluster`** command.

PowerHA® SystemMirror® logs the following information in the `/var/hacmp/log/hacmp.out` file:

- The results of the query, compute, identify, and apply phases of the allocation and release processes.
- All allocations and releases that are performed
- All data persisted into the ODM entry HACMPdynresop. This entry contains the results of the last ROHA operation that was performed. You can display the entry by running the **`clodmget HACMPdynresop`** command.
- The event summaries that you can use to track the processing.
- All ROHA operations.

> **Note:** The ROHA operations contain the *ROHALOG* string. To find ROHA results in the `hacmp.out` file, you can run the **`grep ROHALOG /var/hacmp/log/hacmp.out`** command.

You can view general information from the HMC by running the **`clmgr view report roha`** command on a PowerHA® SystemMirror® node. You can use the following commands on the HMC for more detailed information:

chcod

Run the CoD operations on the HMC outside of PowerHA® SystemMirror® and to manually change the Trial CoD, On/Off CoD, and other types of CoD. You might use this command if PowerHA® SystemMirror® issues an error or a warning during the verification process, or if you requested to use DLPAR and On/Off CoD resources in PowerHA® SystemMirror®.

chcodpool

Run the EPCoD operations on the HMC outside of PowerHA® SystemMirror® and to manually change the Enterprise pool capacity resources. You might use this command if PowerHA® SystemMirror® issues an error or a warning during the verification process, or if you requested to use DLPAR, On/Off CoD, or EPCoD resources in PowerHA® SystemMirror®.

chhwres

Run the DLPAR operations on the HMC outside of PowerHA® SystemMirror® and to manually change the LPAR minimum, LPAR minimum, and LPAR required values for the LPAR. You might use this command if PowerHA® SystemMirror® issues an error or a warning during the verification process, or if you requested to use to use DLPAR and CoD resources in PowerHA® SystemMirror®.

lscod

View the systems CoD configuration.

lscodpool

View the systems Enterprise Pool CoD (EPCoD) configuration.

lshwres

View the LPAR minimum, LPAR maximum, and the total amount of memory and the number of CPUs that are currently allocated to the LPAR.

lssyscfg

Verify that the LPAR node is DLPAR capable.

Application provisioning in PowerHA® SystemMirror®

This section describes the flow of actions in the PowerHA® SystemMirror® cluster, if the application provisioning function through DLPAR and CoD is configured. It also includes several examples that illustrate how resources are allocated, depending on different resource requirements.

In addition, the section provides some recommendations on using pre-and post-scripts.

Overview of application provisioning

When you configure an LPAR on the HMC (outside of PowerHA® SystemMirror®), you provide LPAR minimum and LPAR maximum values for the number of CPUs and amount of memory. You can obtain these values by running the commands on the HMC. The stated minimum values of the resources must be available when an LPAR node starts. If more resources are available in the free pool on the frame, an LPAR can allocate up to the stated desired values. During dynamic allocation operations, the system does *not* allow the values for CPU and memory to go less than the minimum or more than the maximum amounts specified for the LPAR.

PowerHA® SystemMirror® obtains the LPAR minimums and LPAR maximums amounts and uses them to allocate and release CPU and memory when application controllers are started and stopped on the LPAR node.

PowerHA® SystemMirror® requests the DLPAR resource allocation on the HMC before the application controllers are started, and releases the resources after the application controllers are stopped. The Cluster Services waits for the completion of these events before continuing the event processing in the cluster.

These considerations are important:

- Once PowerHA® SystemMirror® has acquired additional resources for the application controller, when the application controller moves again to another node, PowerHA® SystemMirror® releases only those resources that are no longer necessary to support this application on the node.
- PowerHA® SystemMirror® does *not* start and stop LPAR nodes.

Stopping LPAR nodes

When the Cluster Manager is forced down on an LPAR node, and that LPAR is then shutdown (outside of PowerHA® SystemMirror®), the CPU and memory DLPAR resources are released (not by PowerHA® SystemMirror®) and becomes available for other resource groups that are running on other LPARs. However, On/Off CoD resources are not released.

PowerHA® SystemMirror® tracks CPU and memory resources through various operations (DLPAR, On/Off CoD, and EPCoD) that were allocated to the LPAR before the LPAR stops. This tracking is performed so that when the LPAR node restarts, the resources can be dynamically and automatically released (if required). The automatic release of resources after an LPAR failure occurs so that unnecessary resources are not retained.

If a shutdown or a failure occurs while hardware resources are allocated, the resources are automatically released only when the LPAR restarts. This mechanism is called *automatic release after failure*. Without the automatic release after failure mechanism, the On/Off resources remain turned off and the EPCoD resources are not allocated and are not released.

Releasing DLPAR and CoD resources

When the application controller is stopped on the LPAR node (the resource group moves to another node), PowerHA® SystemMirror® releases only those resources that are no longer necessary to support this application controller on the node.

Releasing CPU and memory resources through a DLPAR operation can take time to complete. The Hardware Management Console (HMC) communication tunable sets the timeout value that PowerHA® SystemMirror® waits for the DLPAR operation to complete. The **TIMEOUT** parameter for the HMC adds one minute per gigabyte that is released from the resource. For example, if the communication tunable is set to 10 minutes and to release 100 GB of memory, the actual timeout value is set to 110 minutes.

PowerHA® SystemMirror® can release DLPAR and CoD resources by using synchronous mode or asynchronous mode. Asynchronous mode releases the resources in the background and allows the takeover to start sooner. The default setting for releasing resources is asynchronous mode.

The release mode (synchronous or asynchronous) that is used to release the DLPAR resources is automatically computed by PowerHA® SystemMirror®. For example, if two LPARs are on the same CEC, PowerHA® SystemMirror® performs synchronous mode for releasing the DLPAR resources. In this example, PowerHA® SystemMirror® considers that standby node might need DLPAR resources that are still being used by an active node. Therefore, wait for these DLPAR resources to be released first (synchronous mode).

The order in which DLPAR and CoD resources are released is computed to optimize the release process. In the asynchronous mode, EPCoD resources are released first so that they can be used by another frame. Delaying the release of the EPCoD resources after the DLPAR releases would have delayed the takeover process. On/Off CoD resources are released after the DLPAR resources are released.

Automatic release of DLPAR and CoD resources after failure

There are two different methods for automatically releasing resources.

The first method (normal method) for automatically releasing resources is done during a planned takeover on the active node. Depending on your cluster topology, the release process is either synchronous or asynchronous. The releases process for the normal method consists of releasing DLPAR, On/Off CoD, and EPCoD resources. On/Off CoD resources are released after the DLPAR resources are released. EPCoD resources are released before the DLPAR resources are released for both synchronous or asynchronous release processes. Thus, the EPCoD resources are quickly made available to another frame.

The release process for the normal method is started from the active node before the standby node starts its acquisition process.

If the release process is synchronous, the DLPAR resources that belonged to the active node are released before the allocation process on the standby node starts. The synchronous release is required when both LPARs (active and standby) are hosted on the same frame. If the LPARs are hosted on different frames, the release of the DLPAR resources can be asynchronous because the standby node does not need the DLPAR resources. The asynchronous release process speeds up the takeover process because it is not required for the release process to be finished to start the acquisition of resources by the standby node.

The second method for automatically releasing resources is performed during an unplanned takeover. For this method, PowerHA® SystemMirror® provide another way for releasing resources called *automatic release after failure*. In this scenario, the failing node cannot perform the release process because it failed, and the release process is not done when the standby node starts its takeover process. This process might prevent the standby node from acquiring resources that are held by the failing node. If any EPCoD resources were held by the failing frame, those resources can be applied to the standby node (under certain condition) because the standby node can release these resources for the failing frame. When the failing frame is restarted, any DLPAR and On/Off CoD resources that are previously owned by that frame are automatically released.

During PowerHA® SystemMirror® startup, when a node that previously failed joins the cluster (either manually or automatically) PowerHA® SystemMirror® analyzes all resource group policy options and generates a plan for resource groups to be acquired and for the resource groups to be released. The resources that are released during the automatic release process might be required by the node.

Acquiring DLPAR and CoD resources

If you configure an application controller that requires a minimum and a required amount of resources (CPU or memory), PowerHA® SystemMirror® determines if additional resources need to be allocated for the node and allocates them if possible.

In general, PowerHA® SystemMirror® tries to allocate as many resources as possible to meet the required amount for the application, and uses CoD, if allowed, to do this.

The LPAR node has the LPAR minimum

If the node owns only the minimum amount of resources, PowerHA® SystemMirror® requests additional resources through DLPAR and CoD.

In general, PowerHA® SystemMirror® starts counting the extra resources that are required for the application from the minimum amount. That is, the minimum resources are retained for the node's overhead operations, and are not used to host an application.

The LPAR node has enough resources to host an application

The LPAR node that is about to host an application may already contain enough resources (in addition to the LPAR minimum) to meet the required amount of resources for this application.

In this case, PowerHA® SystemMirror® does not allocate any additional resources and the application can be successfully started on the LPAR node. PowerHA® SystemMirror® also calculates that the node has enough resources for this application in addition to hosting all other application controllers that could be currently running on the node.

Resources requested from the free pool and from the CoD pool

If the amount of resources in the free pool is insufficient to satisfy the total amount requested for allocation (minimum requirements for one or more applications), PowerHA® SystemMirror® requests resources from CoD.

If PowerHA® SystemMirror® meets the requirement for a minimum amount of resources for the application controllers, application controller processing continues. Application controller processing continues even if the total required resources (for one or more applications) have not been met or are only partially met. In general, PowerHA® SystemMirror® attempts to acquire up to the required amount of resources requested for an application.

If the amount of resources is insufficient to host an application, PowerHA® SystemMirror® starts resource group recovery actions to move the resource group to another node.

Resources are requested from free pool and from the CoD pool. You can use the **Resource allocation order** field to specify the order in which the resources are allocated. The resources are released in the reverse order in which they are allocated. The default value for this field is **Free pool before Enterprise pool**.

The minimum amount requested for an application cannot be satisfied

In some cases, even after PowerHA® SystemMirror® requests to use resources from the CoD pool, the amount of resources it can allocate is less than the minimum amount specified for an application.

If the amount of resources is still insufficient to host an application, PowerHA® SystemMirror® starts resource group recovery actions to move the resource group to another node.

The LPAR node is hosting application controllers

In all cases, PowerHA® SystemMirror® checks whether the node is already hosting application controllers that required application provisioning, and that the LPAR maximum for the node is not exceeded:

- Upon subsequent fallovers, PowerHA® SystemMirror® checks if the minimum amount of requested resources for yet another application controller plus the amount of resources already allocated to applications residing on the node exceeds the LPAR maximum.
- In this case, PowerHA® SystemMirror® attempts resource group recovery actions to move the resource group to another LPAR. Note that when you configure the DLPAR and CoD requirements for this application controller, then during cluster verification, PowerHA® SystemMirror® warns you if the total number of resources requested for all applications exceeds the LPAR maximum.

Allocation of resources in a cluster with multiple applications

If you have multiple applications in different resource groups in the cluster with LPAR nodes, and more than one application is configured to potentially request additional resources through the DLPAR and CoD function, the resource allocation in the cluster becomes more complex.

Based on the resource group processing order, some resource groups (hence the applications) might not be started.

Changing the max size of the shared processor pool

For an LPAR that is configured with the **Shared processing mode** setting, the maximum (max) size of the shared processor pool (SSP) can be a physical limit for PowerHA® SystemMirror®.

To authorize PowerHA® SystemMirror® to dynamically adjust the max size of the SSP, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availablity › Change/Show Default Cluster Tunables**, and press Enter.
3. In the **Adjust Shared Processor Pool size if requested** field, specify **yes**.
4. Verify and synchronize the cluster.

AFTER COMPLETING THE TASKPowerHA® SystemMirror® can dynamically change the max size of a SSP. The SSP are typically used with third-party software that requires a license that charges on a per-processor basis. For example, if you pay for 7 CPU Oracle licenses and if you have 6 CPUs on the active CEC with one CPU on the backup CEC, when a failover occurs and the takeover process starts, you expect the shared processor pool size to be identical on the new CEC that is used due to the takeover. PowerHA® SystemMirror®, as part of the resource release process, decreases the size of the SSP to its original size on the active node and increases the size of the SSP to the expected size on the standby node. This process verifies that your Oracle applications are still contained on the same number of CPUs than before the takeover (6 CPUs on the new active node and 1 CPU on the previously active node).

Examples of using DLPAR and CoD resources

These examples show CPU allocation and release.

Memory allocation process is similar.

It is important to remember that once PowerHA® SystemMirror® acquires additional resources for an application controllers, when the server moves again to another node, it takes the resources with it, that is, the LPAR node releases all the additional resources it acquired, and remains with just the minimum.

The configuration is an 8 CPU frame, with a two-node (each an LPAR) cluster. There are 2 CPUs available in the CoD pool, that is through the CoD activations. The nodes have the following characteristics:

Node Name	LPAR Minimum	LPAR Maximum
Node1	1	9
Node2	1	5

The following application controllers are defined in separate resource groups:

Application controller name	CPU Desired	CPU Minimum	Allow to Use CoD?
AS1	1	1	Yes
AS2	2	2	No
AS3	4	4	No

Example 1: No CPUs are allocated at application controller start, some CPUs are released at server stop

Current® configuration settings:

- Node1 has 3 CPUs allocated.
- Node2 has 1 CPU allocated.
- The free pool has 4 CPUs.

PowerHA® SystemMirror® starts application controllers as follows:

- Node1 starts AS2, no CPUs are allocated to meet the requirement of 3 CPUs. (3 CPUs is equal to the sum on Node1's LPAR minimum of 1 plus AS2 desired amount of 2).

- Node1 stops AS2. 2 CPUs are released, leaving 1 CPU, the minimum requirement. (Since no other application controllers are running, the only requirement is Node1 LPAR minimum of 1).

Example 2: Failure to allocate CPUs due to resource group processing order

Current® configuration settings:

- Node1 has 3 CPUs allocated.
- Node2 has 1 CPU allocated.
- The free pool has 4 CPUs.

PowerHA® SystemMirror® starts application controllers as follows:

- Node1 starts AS1, no CPUs are allocated since the requirement of 2 is met.
  Node1 starts AS3, 3 CPUs are allocated to meet the requirement of 6. There is now 1 CPU in the free pool.
- Node1 attempts to start AS2. After Node 1 has acquired AS1 and AS3, the total amount of CPUs Node1 must now own to satisfy these requirements is 6, which is the sum of Node1 LPAR minimum of 1 plus AS1 desired amount of 1 plus AS3 desired amount of 4.
  Since AS2 minimum amount is 2, in order to acquire AS2, Node1 needs to allocate 2 more CPUs, but there is only 1 CPU remaining in the free pool and it does *not* meet the minimum requirement of 2 CPUs for AS2. The resource group with AS2 goes into error state since there is only 1 CPU in the free pool and CoD use is *not* allowed.

Example 3: Successful CoD resources allocation and release

Current® configuration settings:

- Node1 has 3 CPUs allocated.
- Node2 has 1 CPU allocated.
- The free pool has 4 CPUs.

PowerHA® SystemMirror® starts application controllers as follows:

- Node1 starts AS3, 2 CPUs are allocated to meet the requirement of 5.
- Node1 starts AS2, 2 CPUs are allocated to meet the requirement of 7. There are now no CPUs in the free pool.
- Node1 starts AS1, 1 CPU is taken from CoD and allocated to meet the requirement of 8.
- Node1 stops AS3, 4 CPUs are released and 1 of those CPUs is put back into the CoD pool.

Example 4: Resource group failure (the minimum for the server Is not met, but the LPAR maximum for the node Is reached)

Current® configuration settings:

- Node1 has 1 CPU allocated.
- Node2 has 1 CPU allocated.
- The free pool has 6 CPUs.

PowerHA® SystemMirror® starts application controllers as follows:

- Node2 starts AS3, 4 CPUs are allocated to meet the requirement of 5. There are now 2 CPUs in the free pool.
- Node2 attempts to start AS2, but AS2 goes into error state since the LPAR maximum for Node2 is 5 and Node2 cannot acquire more CPUs.

Example 5: Resource group fallover

Current® configuration settings:

- Node1 has 3 CPUs allocated.
- Node2 has 1 CPU allocated.
- The Free pool has 4 CPUs.

PowerHA® SystemMirror® starts application controllers as follows:

- Node1 starts AS2, no CPUs are allocated to meet the requirement of 3.
- The resource group with AS2 falls over from Node1 to Node2.
- Node1 stops AS2. 2 CPUs are released, leaving 1 CPU on the LPAR, the minimum requirement for the node.
- Node2 start AS2, 2 CPUs are allocated to meet the requirement of 3.

Using pre-event and post-event scripts

The existing pre-event and post-event scripts that you were using in a cluster with LPARs (before using the CoD integration with PowerHA® SystemMirror®) may need to be modified or rewritten, if you plan to configure CoD and DLPAR requirements in PowerHA® SystemMirror®.

Keep in mind the following:

- PowerHA® SystemMirror® performs all the DLPAR operations before the application controllers are started, and after they are stopped. You may need to rewrite the scripts to account for this.
- Since PowerHA® SystemMirror® takes care of the resource calculations, requests additional resources from the DLPAR operations and, if allowed, from CoD, you may get rid of the portions of your scripts that do that.
- PowerHA® SystemMirror® only takes into consideration the free pool on a single frame. If your cluster is configured within one frame, then modifying the scripts is sufficient.

However, if a cluster is configured with LPAR nodes residing on two frames, you may still require the portions of the existing pre-event and post-event scripts that deal with dynamically allocating resources from the free pool on one frame to the node on another frame, should the application require these resources.

Resource Optimized High Availability in Cloud

Resource Optimized High Availability (ROHA) in Cloud is a function in PowerHA® SystemMirror® that automatically and dynamically manages Dynamic Logical Partitions (DLPARs) resources. You can configure ROHA in Cloud with virtual server instances in Cloud, hardware resource provisioning, and cluster tunable configurations.

PowerHA® SystemMirror® 7.2.6 SP1 supports ROHA feature with a standard cluster. Starting with PowerHA® SystemMirror® 7.2.7, it also supported site-based cluster. The Shared processor pool (SPP) feature is supported with ROHA in Cloud.

Starting with the PowerHA® SystemMirror® 7.2.8 and later, the cloud administrator must create a shared processor pool and configure the PowerVS instances by selecting the shared processor pool.

PowerHA® SystemMirror®7.2.8 does not require special configuration. A cloud administrator must select SPP when configuring a virtual machine. The IBM infrastructure team controls SPP when allocating the resources to the virtual machine.

For more information about shared processor pool, see IBM Cloud documentation.

For a reference example of shared processor pool, see IBM Cloud documentation.

Before you use ROHA in Cloud, you must add all virtual server instances that are in Cloud onto all associated clusters in PowerHA® SystemMirror®. You must also plan for the necessary resources for your applications, identify your workloads and your requirements for physical resources (CPU cores, virtual CPUs, and memory). After you identify all these requirements, you must configure ROHA in Cloud.

Before you use the ROHA in Cloud function for the first time, you must complete the following steps:

1. Install the following applications on your cluster.
    - Python Interpreter version 2.7.x
    - Representational State Transfer (REST) application programming interface (API) for AIX® module (libcurl.a) on Cloud logical partitions (LPARs)

2. Add Cloud instance details for all the associated cluster nodes by using the **clmgr manage cluster cloudroha** command. PowerHA® SystemMirror® automatically collects and adds Cloud instance details for all the associated cluster nodes.
3. Create hardware resource provisioning for each application controller that you identified for your workloads by running the **clmgr add roha** command for cluster.

> **Note:** The ROHA in Cloud feature is supported only on IBM®Cloud.

Depending on the configuration of your environment, you might have to complete the following optional tasks:

- Define virtual server instances in Cloud at the site level or cluster level if your topology requires them.
- Change the default value of the virtual server tunable such as the timeout value for DLPAR operations.
- Verify virtual server instances in Cloud that are used by the cluster and by each node.
- Change the clusters ROHA tunables for DLPAR acquisition and release operations.

Planning for ROHA in Cloud

If you plan to use the Resource Optimized High Availability (ROHA) in Cloud function in a PowerHA® SystemMirror® cluster, you must plan and allocate resources to the Logical Partitions (LPARs) by using a Cloud account.

Review the following planning information about ROHA in Cloud in a PowerHA® SystemMirror® cluster:

- Obtain the LPAR resources information and resource group policies information:
  - How much memory and resources the applications that are supported by your cluster require when they run on their regular hosting nodes. Under normal running conditions, check how much memory and what number of CPUs each application uses to run with optimum performance on the LPAR node on which its resource group resides normally (home node for the resource group).
  - Determine the startup, fallover, and fallback policies of the resource group that contains the application controller by using the **clRGinfo** command. This identifies the LPAR node to which the resource group will fall over, in case of a failure.
  - How much memory and what number of CPUs are allocated to the LPAR node on which the resource group will fall over, in case of a failure. This LPAR node is referred to as a standby node. With these numbers in mind, consider whether the application's performance will be impaired on the standby node, if the application is running with fewer resources.
  - Check the existing values for the LPAR minimums, LPAR maximums, and LPAR desired amounts (resources and memory) specified by using the **lparstat** command on the standby node.
- Estimate the resources that are required for the application:
  - For each standby node that can host a resource group, you must estimate the optimal amount of resources (CPU and memory) that this node requires for the application to run successfully. The optimal amount of resources that you identify are specified in PowerHA® SystemMirror®. PowerHA® SystemMirror® verifies that the optimal amount is contained within the boundaries of the LPAR maximums that are configured outside of PowerHA® SystemMirror® for each LPAR.
  - In PowerHA® SystemMirror® Version 7.2.6 SP1, or later, ROHA in Cloud supports resources that are acquired from DLPAR. Currently, ROHA in Cloud feature does not support Enterprise Pool Capacity on Demand (EPCoD) license.
- Revise existing pre-event and post-event scripts that were used to allocated DLPAR resources.
- The following limitations are applicable for ROHA in Cloud multi-cluster deployments. A multi-cluster deployment is an environment with multiple clusters deployed across two different systems. In a multi-cluster deployment environment, each cluster has one node on an active system, and another node on the standby system.
  - In Cloud, we will not be able to identify how many resources are free before we request for acquire operation, hence you should check with Cloud support team before you plan your resources.

- ROHA in Cloud feature does not support resource groups that have a startup policy of `Online on All Available Nodes`.
- The resources are released asynchronously in ROHA in Cloud environment. If one release operation is in progress, and another release operation is triggered simultaneously, the second release operation might fail. You must specify proper delay after first release operation (timeout) so that the second release operation succeeds. To avoid such failure, you must perform the acquire or release operations for the resources groups during cluster start, cluster stop, or failover operations.

Note: If you were using LPAR nodes in your cluster before using the ROHA function, you might need to revise and rewrite your existing pre-event and post-event scripts.

Configuring Cloud instances for cluster nodes to work with ROHA

You must configure Cloud instances on all PowerHA® cluster nodes to use with Resource Optimized High Availability (ROHA) in Cloud. The information about ROHA in Cloud configuration is used during acquire or release operations of the resources (memory or processing units) in the Cloud environment.

You can use the SMIT interface and the **`clmgr`** utility to configure the Cloud instance for IBM® Power® Systems Virtual Servers that is used by the PowerHA® SystemMirror® cluster. This feature is supported in PowerHA® SystemMirror® 7.2.6 SP1, and later.

Note: In PowerHA® SystemMirror® 7.2.6 SP1, the ROHA in Cloud configuration is supported only for the standard type of cluster configuration.

The following figure describes the high-level details of ROHA in Cloud configuration by using SMIT:

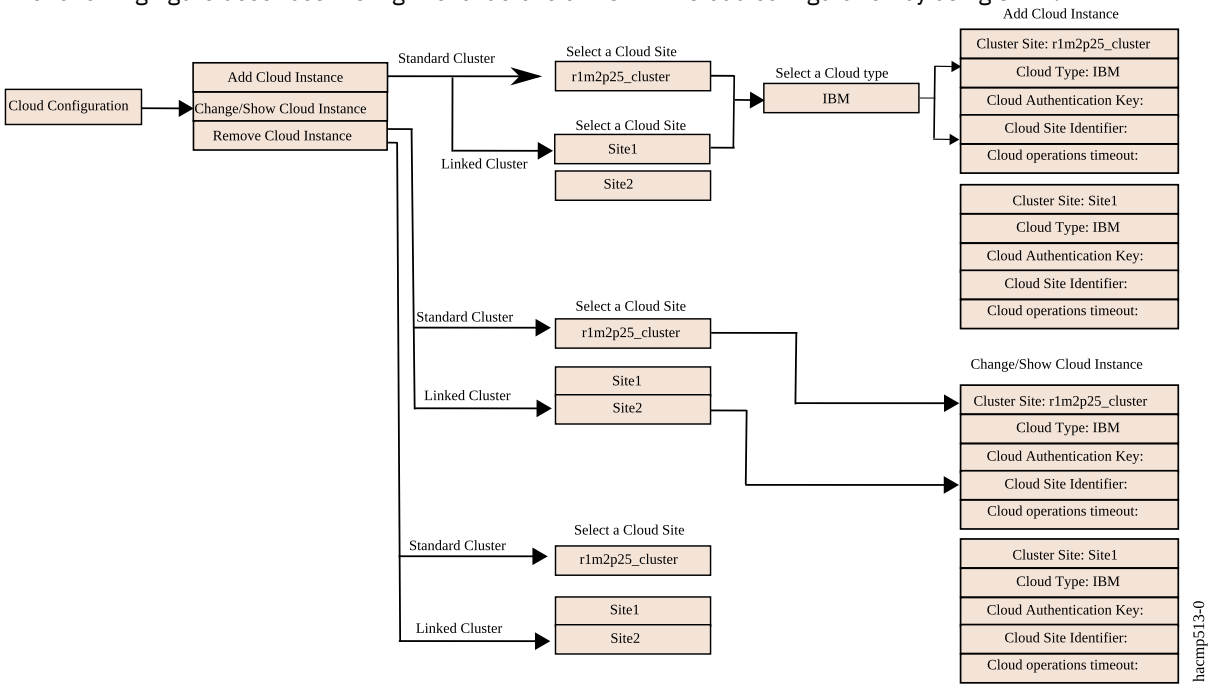


Figure 13: ROHA in Cloud configuration process

To configure ROHA in Cloud by using SMIT, complete the following steps.

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Cloud Configuration**, and press Enter.
3. Select from the following options:
   Add Cloud Instance

Add the Cloud instance for Power® Systems Virtual Servers that is used by the PowerHA® SystemMirror® cluster.

Change/Show Cloud Instance

Change the Cloud instance parameters for a PowerHA® SystemMirror® site or cluster.

Remove Cloud Instance

Delete the ROHA in Cloud configuration.

Verify and Sync cluster configuration

Verify and sync the cluster configuration.

Adding a Cloud instance to use IBM® Power® Systems Virtual Servers for ROHA

You can use SMIT interface to add a Cloud instance for Power® Systems Virtual Servers for Resource Optimized High Availability (ROHA).

To add a Cloud instance, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Cloud Configuration › Add Cloud Instance**, and press Enter.
3. Complete the following fields, and press Enter.
   Cloud Site

   This field is auto-selected based on the cluster-type in PowerHA® SystemMirror®. For the standard cluster type, you must specify value in the`Cluster Name`field. For the linked cluster-type, select the specific site name from the SMIT selection screen.

   Cloud Type

   From the SMIT selection screen, choose the Cloud type that you want to associate with the Power® Systems Virtual Servers nodes that are located under a PowerHA® cluster or a PowerHA® site. The default Cloud type is IBM® Cloud.

   Cloud Authentication Key

   Enter the authentication key for the Cloud Identity and Access Management (IAM) user. For IBM® Cloud, enter the application programming interface (API) key of the cloud account where the Cloud instance is defined. The API key is used to extract the IAM token for starting the Representational State Transfer (REST) API calls provided by IBM® cloud. The API key is encrypted and stored in the PowerHA® database.

   Cloud Site Identifier

   Enter the cloud service identifier of a region or a zone in which Power® Systems Virtual Servers instance is defined. For IBM® Cloud, specify the Cloud Resource Name (CRN) that you received from the Cloud service. For more information about CRN, see IBM® documentation.

   Cloud operations timeout (in minutes)

   Enter a timeout period in minutes for the Cloud ROHA commands that ran on the Cloud node of Power® Systems Virtual Servers during cloud operations such as resource acquire or resource release. If you do no set the timeout value, a default timeout value of 5 minutes is used.

> After completion, Cloud configuration is set on to all Cloud nodes of Power® Systems Virtual Servers in a PowerHA® site or a PowerHA® cluster. The Cloud instance IDs of Power® Systems Virtual Servers are collected dynamically based on the specified fields and are stored in the PowerHA® database.

Changing the Cloud instance for IBM® Power® Systems Virtual Servers for ROHA

You can use SMIT interface to modify or update the Cloud instance for Power® Systems Virtual Servers for Resource Optimized High Availability (ROHA).

To change a Cloud instance, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Cloud Configuration › Change/Show Cloud Instance**, and press Enter.
3. Update the required fields, and press Enter.

Removing the Cloud instance for IBM® Power® Systems Virtual Servers for ROHA

You can use SMIT interface to remove the Cloud instance for Power® Systems Virtual Servers for Resource Optimized High Availability (ROHA).

To remove a Cloud instance, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Cloud Configuration › Remove Cloud Instance**, and press Enter.
3. Select a PowerHA® site or a PowerHA® cluster, and press Enter.

Verifying the Cloud instance for IBM® Power® Systems Virtual Servers for ROHA

You can use SMIT interface to verify the Cloud instance for Power® Systems Virtual Servers for Resource Optimized High Availability (ROHA).

To verify a Cloud instance, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Cloud Configuration › Verify and Sync cluster configuration**, and press Enter.
3. Select a PowerHA® site or a PowerHA® cluster, and press Enter.
   The verification process completes the following check:
   - Checks whether the configured cloud instances of Power® Systems Virtual Servers are reachable by using the service broker. If the configured cloud instances are not reachable, you must check the Cloud configuration.
   - Checks whether the Cloud logical partition (LPAR) status of Power® Systems Virtual Servers is ACTIVE. The verification process also checks whether the health of Power® Systems Virtual Servers is OK. If the verification process fails, the Dynamic Logical Partition (DLPAR) operations might fail.

Changing hardware provisioning for an application controller in ROHA in Cloud configuration

You can use the SMIT interface to change the hardware provisioning for an application controller in the Cloud Resource Optimized High Availability (ROHA) in Cloud configuration.

When an application requires more resources to be allocated on a node, PowerHA® SystemMirror® determines whether only Dynamic Logical Partition (DLPAR) resources from the free pool on the system are required.

During verification, PowerHA® SystemMirror® verifies that the resource values specified are below LPAR maximum values for CPU and memory resources. PowerHA® SystemMirror® also verifies that the total of required resources for all application controllers that can run concurrently on the LPAR is less than the LPAR maximum. For example, if the LPAR node is already hosting application controllers that require an additional

DLPAR, it is possible that the LPAR cannot contain the additional resources because it has reached its LPAR maximum. In this case, PowerHA® SystemMirror® displays a warning message.

To change the hardware provisioning for an application controller, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Resources › Configure User Applications (Scripts and Monitors) › Resource Optimized High Availability › Hardware Resource Provisioning for Application Controller › Change/Show Hardware Resource Provisioning of an Application Controller** , and press Enter.
3. Select the application controller that you want to change from the list.
4. Change the following fields:
   Used desired level from the LPAR profile

   Select **Yes** if you want the LPAR hosting your node to reach the desired level specified in the LPAR profile. Select **No** if you want to enter a specific value for CPU and memory resources.
   If you have a mixed configuration with this field set to both **Yes** and **No**, the allocation that is performed is the sum of the LPAR profile value and the various optimal values that you specified.

   Optimal amount of gigabytes of memory

   Enter the amount of memory that PowerHA® SystemMirror® attempts to allocate to the node before you start the specified application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. You can specify values in increments of 0.25 GB, 0.5 GB, 0.75 GB, or 1 GB. For example, a value of 1.5 represents 1.5 GB or 1536 MB.

   Optimal number of dedicated processors

   Do not change this field when ROHA is configured in the IBM® Power® Systems Virtual Servers Cloud environment.

   Optimal number of processing units

   Enter the number of processing units that PowerHA® SystemMirror® attempts to allocate to the node before you start the application controller. You can change this field only if the **Used desired level from the LPAR profile** field is set to **No**. You can specify a value up to two decimal places in the range 0.01 - 255.99. This value is used only on nodes that support allocation of processing units.

   Optimal number of virtual processors

   Do not change this field when ROHA is configured in the Power® Systems Virtual Servers Cloud environment.

> If the provisioned resource (amount of memory or processing_units) is not sufficient, PowerHA® SystemMirror® performs recovery actions to move the resource group along with the applications to another node, or PowerHA® SystemMirror® might allocate fewer resources depending on the setting for the Start resource group even if the resources have insufficient tunable.

Troubleshooting ROHA in Cloud

You can use commands to troubleshoot Resource Optimized High Availability (ROHA) in Cloud operations in your cluster.

To view all data that is related to ROHA in Cloud, run the **`clmgr view report roha`** command. The ROHA in Cloud configuration might delay the runtime of the **`clmgr view report roha`** command. If this command fails with a timeout error, you can manually prefix the timeout duration by specifying the **`CLMGR_TIMEOUT=##seconds`** attribute while rerunning the **`clmgr view report roha`** command.

You can view general information in the ROHA in Cloud configuration by running the **`clodmget HACMPnode`** command on a PowerHA® SystemMirror® node.

To run consistency checks on all nodes in your cluster, run the **`clmgr verify cluster`** command.

PowerHA® SystemMirror® logs the following information in the `/var/hacmp/log/hacmp.out` file:

- The results of the query, compute, identify, and apply phases of the allocation and release processes.

- All software allocations and releases that are installed.
- All data that persisted into the HACMPdynresop Object Data Manager (ODM) entry. The ODM entry contains the results of the last ROHA operation that was completed. You can display the ODM entry by running the **clodmget HACMPdynresop** command.
- The event summaries that you can use to track the processing.
- All ROHA operations that are completed.

> **Note:** The ROHA operations contain the *ROHALOG* string. You can find the output of ROHA operations in the `hacmp.out` file by running the **grep ROHALOG /var/hacmp/log/hacmp.out** command.

PowerHA® SystemMirror® logs the following information in the `/var/hacmp/log/cloudroha.log` file:

- Authentication issues that are related to Cloud (if any).
- Data that is related to the Cloud operations such as query, acquires, or release performed on logical partitions (LPARs) of IBM Power® Systems Virtual Servers.

SAP high availability management with PowerHA® SystemMirror®

Using PowerHA® SystemMirror®, you can mange the high availability of SAP environments in the cluster. You can use PowerHA® SystemMirror® management interfaces to configure high availability policies and use methods to start, stop, and monitor the managed instances of your environment.

PowerHA® SystemMirror® supports SAP environments with various software components, such as DB2®, Oracle, and NFS 4, and the General Parallel File System (GPFS) that are running on various nodes in the cluster.

> **Note:** All SAP functions are available in the SMIT interface.

SAP high availability infrastructure

You can configure a shared file system to monitor for high availability by using the SMIT interface.

PowerHA® SystemMirror® provides high availability agents that you can use to manage SAP environments. A high availability agent is called a Smart Assist agent. Using a Smart Assists agent, you can discover and configure high availability policies and monitor the health of your entire network.

SAP liveCache Hot Standby with PowerHA® SystemMirror®

With the SAP liveCache Hot Standby function of SAP, you can maintain a standby instance of an SAP liveCache environment so that when a failure occurs, the standby instance can take over the SAP primary service immediately. The take over occurs without rebuilding the memory structure or losing the data that is written to the database log files.

With PowerHA® SystemMirror®, you can use the Smart Assist for SAP liveCache Hot Standby to set up SAP liveCache Hot Standby instances for a high availability environment.

PowerHA® SystemMirror® SAP liveCache Hot Standby wizard

You can use the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard in the SMIT interface to complete the initial configuration of SAP liveCache Hot Standby.

Prerequisites for using the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard

Before you can use the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard, you must install the PowerHA® SystemMirror® file sets and Smart Assist file sets on the PowerHA® SystemMirror® cluster and all nodes in the cluster.

> **Important:** Review the information in the Planning for Smart Assist for SAP liveCache Hot Standby topic before you use the wizard.

The wizard has the following limitations:

- The wizard is available when cluster services are running and working correctly on all nodes in the cluster.
- Do not upgrade SAP liveCache Hot Standby before using the wizard.
- The wizard cannot run on a virtual SCSI disk.
- Only one Hardware Management Console (HMC) be added using the wizard. The HMC can access the storage system that is used for automation. After you run the wizard, you can manually added a second HMC for redundancy.
- The wizard only functions in a two-node configuration.
- The SAP Advanced Planner and Optimizer (SAP APO) only accepts letters and numbers for the official host name.
- Host names must be the same as the node names in the PowerHA® SystemMirror® cluster.
- The SAP liveCache instance name is limited to a maximum of 7 characters.

Before you use the wizard, you must know the following information about your environment:

- SAP liveCache instance name
- SAP liveCache administrative user (typically control)
- SAP liveCache XUSER for control user is enabled for root and sdb user on both nodes
- For IBM® SAN Volume Controller (SVC) access, the Secure Shell (SSH) keys for both nodes
- Storage IP address
- Storage types
- Primary node
- Service IP label that is used to configure SAP liveCache high availability
- Shared file system that is used as the LOCK directory for the SAP liveCache instance
- Disks that are used for the log volume group and data volume group
- Primary SAP liveCache log volume group.

  > **Note:** The SAP liveCache log volume group must be created and imported on all nodes in the cluster and must be concurrently active. The raw logical volumes that belong to the SAP liveCache log volume group must be controlled by the **SdbOwner** user and the **SdbGroup** group.

- Primary SAP liveCache data volume group

  > **Note:** The SAP liveCache data volume group must be created and imported on the primary node in the cluster and must be active. The raw logical volumes that belong to the SAP liveCache data

volume group must be controlled by the **SdbOwner** user and the **SdbGroup** group on the primary node.

Creating an XUSER

An XUSER entry contains the logon data and stores the data as user keys. When logging on to a database, you specify the user key. An XUSER entry is stored separately for each operating system user.

To create an XUSER variable with control user name credentials, use the following syntax once as a sdb user and once as a root user on both nodes in the PowerHA® SystemMirror® SAP liveCache Hot Standby configuration:

```
MAXDB_INDEP_PROGRAM_PATH/bin/xuser -U <XUSER_NAME> -u <DBM_USERNAME>,<DBM_PASSWORD> -d
<LIVECACHE_NAME> -n <NODE_NAME>
```

Note: The `MAXDB_INDEP_PROGRAM_PATH` can be found in the **/etc/opt/sdb** file.

Configuring the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard

Use the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard to create a highly available SAP liveCache environment.

To configure the PowerHA® SystemMirror® SAP liveCache Hot Standby wizard, complete the following steps:

1. From the command line, enter `smit sysmirror`
2. From the SMIT interface, select **Cluster Applications and Resources › Make Applications Highly Available (Use Smart Assists) › SAP liveCache Hot Standby Configuration Wizard** , and press Enter.
3. From the list of node names, select the two nodes you want to use for the SAP liveCache Hot Standby configuration and press Enter.
4. Select the storage subsystem, and press Enter.

   The supported storage subsystems are DS8000® and SAN Volume Controller (SVC).

5. Complete the following fields, which are based on the storage subsystem you selected in step 4.

Fields for storage subsystem	
Field	**Value**
liveCache Instance Name	Enter the name of the primary SAP liveCache instance that you want to configure for SAP liveCache Hot Standby.
SAP liveCache Hot Standby DBM user XUSER	Enter the name of the XUSER, created with DBM user crendentials.
Storage (HMC) type	Display the type of storage subsystem that you selected in step 4. This field cannot be cahnged.
Storage Server (HMC) IP	Enter the IP address of the Hardware Management Console (HMC).
Storage Server (HMC) User	Enter the storage user name of the HMC. This filed is only available if you selected the DS8000® storages subsystem in step 4.

Field	Value
Storage (HMC) Password	Enter the password for the storage user name of the HMC. This filed is only available if you selected the DS8000® storages subsystem in step 4.
liveCache Global Filesystem Mount point	Enter a global file system mount point that is configured for PowerHA® SystemMirror®. The file system is used as a lock directory for SAP liveCache Hot Standby. The mount point must always be available on all nodes in the cluster.
Primary Node	Select the node where the primary SAP liveCache database instance is installed.
Service Interface	Enter the logical host name that is used to configure the SAP liveCache database instance.
Node Names	Display the node names that you selected in step 3. This field cannot be changed.
Log Volume Group	Select the volume group that is associated with the SAP liveCache Hot Standby log volume groups.
Data Volume Group	Select the volume group that is associated with the SAP liveCache Hot Standby data volume groups.

HDISK of Data Volume Group	Enter the hdisk information for the disks that you selected in the **Data Volume Group** field. Enter the hdisk pair information in the following format:
	`primary node disk-->secondary node disk,pirmary node disk-->secondary node disk`
	In the following example, `hdisk1` on the primary node is mapped to `hdisk1` on the secondary node, and `hdisk3` on the primary node is mapped to `hdisk4` on the secondary node.
	`hdisk1-->hdisk1,hdisk3-->hdisk4`
	You must input the information for data volume groups in a specific format. For example, data volume group 1 (DATAVG1) has `hdisk1`, `hdisk3`, and `hdisk4` on the primary node and `hdisk10`, `hdisk11`, and `hdisk13` are the corresponding hdisks on the secondary node. Data volume group 2 (DATAVG2) has `hdisk2` on the primary node and `hdisk20` is the corresponding hdisk on the secondary node. Data volume group 3 (DATAVG3) has `hdisk5` and `hdisk7` on the primary node and `hdisk21` and `hdisk22` are the corresponding hdisks on the secondary node. In this example, the datavg order is DATAVG1, DATAVG2, and DATAVG3. Therefore, your hdisk pairs would be in the following order:
	`|---------- DATAVG1 ----------|---DATAVG2-----|----------DATAVG3----------|`
	`hdisk1-->hdisk10,hdisk3-->hdisk11,hdisk4-->hdisk13,hdisk2-->hdisk20,hdisk5-->hdisk21,hdisk7-->hdisk22`

6. Verify and synchronize the cluster.

Customizing the PowerHA® SystemMirror®SAP liveCache Hot Standby wizard

You can customize your environment by changing the `/usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS` file in the PowerHA® SystemMirror® cluster. Customizing your environment changes how the wizard configures the SAP liveCache Hot Standby instance.

Changing the /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS file

You must change the `/usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS` file before you use the PowerHA® SystemMirror®SAP liveCache Hot Standby wizard.

The /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS file exports a set of environmental variables that are used by the wizard. These variables carry default values that you can change to customize your environment when you use the wizard.

> **Attention:** Changing the variables in the /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS file can cause failure in your cluster. You must test your changes to the /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS file before you apply the changes to your cluster.

The following table displays information about the /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS file.

Settings for the /usr/es/sbin/cluster/sa/hswizard/sbin/GLOBALS *file*		
Variable name	**Default value**	**Description**
MAXDB_REF_FILE	/etc/opt/sdb	SAP MaxDB populates this file with a variable that you can change, such as SAP MaxDB user, group, independent program path, and independent data path. The path of this file cannot be changed.
MAXDB_PROGRAM_PATH		To find this path, run the **grep IndepPrograms MAXDB_REF_FILE** command.
MAXDB_INDEP_DATA_PATH		To find this path, run the **grep IndepData MAXDB_REF_FILE** command.
MAXDB_DEP_PATH		This path is found in the MAXDB_INDEP_DATA_PATH/config/Databases.ini file.
MAXDB_DBMCLI_COMMAND	MAXDB_PROGRAM_PATH/bin/dbmcli	The database manager used a command-line-oriented client called the Database Manager CLI (DBMCLI).
MAXDB_X_USER	MAXDB_PROGRAM_PATH/bin/xuser	Using the XUSER database tool, you can store user login data and access database instances by using simplified logon. When logging on to a database instance, you must specify the user key.
LC_CONFIG_FILE	/usr/es/sbin/cluster/sa/hswizard/sbin/lc_param_config	You can manually customize this file to change parameters for the SAP liveCache that is created by the wizard.
HSS_LIB_PATH	/opt/ibm/ibmsap	Path where the HSS library is installed.
HSS_CONNECTORS_SVC	/opt/ibm/ibmsap/connectors/HSS2145	Location of the connector scripts for SVC.
HSS_CONNECTORS_DS	/opt/ibm/ibmsap/connectors/HSS2107	Location of the connector scripts for DS.
LIB_DS	libHSSibm2107.so	Name of the DS library.
LIB_SVC	libHSSibm2145.so	Name of the SVC library.
PRI_MAPNAME_SVC		Name of the flash copy consistency group that goes from primary to secondary nodes.

Variable name	Default value	Description
SEC_MAPNAME_SVC		Name of the flash copy consistency group that goes from secondary to primary nodes.
PRI_MAPNAME_DS	3333	A 4-digit number is used while creating a flash copy consistency group between primary and secondary nodes.
SEC_MAPNAME_DS	5555	A 4-digit number is used while creating a flash copy consistency group between secondary and primary nodes.
DSCLI_DIR	/opt/ibm/dscli	Location of DSCLI.
DSCLI	/opt/ibm/dscli/dscli	DSCLI command to run on the Hardware Management Console (HMC).
LSHOSTVOL	/opt/ibm/dscli/bin/lshostvol.sh	Path to the lshostvol.sh script.

Changing the RTEHSS_config.txt file

The RTEHSS_config.txt file is located in the `/opt/ibm/ibmsap/instance_name/` directory, where `instance_name` is the name of the SAP liveCache instance.

The `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` file is created when you start the PowerHA® SystemMirror®SAP liveCache Hot Standby wizard.

To help you understand how to configure the parameters in the RTEHSS_config.txt file, you can view the `/opt/ibm/ibmsap/RTEHSS_config_sample.txt` sample file.

The `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` file contains the storage parameters that are used by the HSS library to manage DS8000® and SAN Volume Controller (SVC) storage. The wizard creates this file to identify the location of the HSS library, Hardware Management Console (HMC) user credentials, HMC IP address, source volume IDs, and target volume IDs.

> **Attention:** Changing the variables in the `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` file can cause failure in your cluster. You must test your changes to the `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` file before you apply the changes to your cluster.

The following table displays information about the `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` file.

Settings for the `/opt/ibm/ibmsap/instance_name/RTEHSS_config.txt` *file*		
Variable name	**Default value**	**Description**
CSmode	FC	This variable copies server services.
IBMclidir	**DS8000®** /opt/ibm/DScli **SAN Volume Controller (SVC)** /opt/ibm/ibmsap/connectors/HSS2145	This variable displays the directory for the storage command-line interface. For the DS8000®, this variable displays the path to the DSCLI For the SVC, this variable displays the path to the storage interface.
IBMsapapodir	Discovered by wizard	This variable displays the install directory for the storage-dependent runtime library.

Variable name	Default value	Description
MlCLogVdiskID	Discovered by wizard	This variable displays the ID of the log volume on the primary node. For DS8000®, this variable displays the ID for the volume. The ID is a 4-digit hexadecimal number (0000 - FFFF). For SVC, this variable displays the vdisk_id or vdisk_name.
SlCLogVdiskID	Discovered by wizard	This variable displays the ID of the log volume on the secondary node. For DS8000®, this variable displays the ID for the volume. The ID is a 4-digit hexadecimal number (0000 - FFFF). For SVC, this variable displays the vdisk_id or vdisk_name.
MlCDataVdiskID	Discovered by wizard	This variable displays the ID of the data volume on the primary liveCache server. For multiple values, use a comma to separate the IDs. For DS8000®, this variable displays the ID for the volume. The ID is a 4-digit hexadecimal number (0000 - FFFF). For SVC, this variable displays the vdisk_id or vdisk_name.
SlCDataVdiskID	Discovered by wizard	This variable displays the ID of the data volume on the first standby liveCache server. For multiple values, use a comma to separate the IDs. For DS8000®, this variable displays the ID for the volume. The ID is a four-digit hexadecimal number (0000 - FFFF). For SVC, this variable displays the vdisk_id or vdisk_name.
CSaIP	Value is entered when using the wizard	This variable displays the copy server IP address. For DS8000®, this variable displays the IP address of the Hardware Management Console (HMC). For SVC, this variable displays the IP address of the SVC.
CSaUID	Value is entered when using the wizard	This variable displays the copy server user ID for the administrator. For DS8000®, this variable displays the user ID to perform copy service task.

Variable name	Default value	Description
		For SVC, this variable displays the ID name for the SSH connection to SVC.
CSapwd	Value is entered when using the wizard	This variable copies the DS8000® password for the user ID. If you have an SVC disk, the value must be blank.
DSdevID	Discovered by wizard	This variable displays the ID for DS8000® disk. If you have an SVC disk, the value must be blank.
HSS_NODE_001	Primary node entered in wizard	This variable displays the primary node name.
HSS_NODE_002	Extracted from node list	This variable displays the standby node.
EstDataCST_001_002	3333	Defines the copy server tasks when a storage system uses flash copy to copy the data volumes from HSS_NODE_001 to HSS_NODE_002. The flash copy relationship is built dynamically. You cannot create the flash copy relationship. DS8000®: The sequence number that is used to copy the data volume from the current primary node (HS_NODE_001) to the requesting standby node (HS_NODE_002). The sequence number is a 4-digit hexadecimal number (0000 - FFFF). The task is built dynamically. SVC: The unique name that is used to name the dynamically created FC relationship to copy the data volume from the current primary node (HS_NODE_001) to the requesting standby node (HS_NODE_002). The name cannot begin with a digit. It must begin with a letter.
EstDataCST_002_001	5555	
TermDataCST_001_002	*instance_name*'_01', where *instance_name* is the name of the SAP liveCache instance.	Define the task names to end the flash copy relationship between the HS_NODE_001 volume and of HS_NODE_002 volume.
TermDataCST_002_001	*instance_name*'_02', where *instance_name* is the name of the SAP liveCache instance.	

Live Partition Mobility

You can use Live Partition Mobility (LPM) to migrate partitions that are running the AIX® operating system. You can also migrate applications from one physical server to another without disrupting the infrastructure services.

The migration operation maintains complete system transactional integrity. The migration transfers the entire system environment, including the processor state, memory, attached virtual devices, and connected users.

LPM provides the facility for no required downtime for planned hardware maintenance. However, LPM does not offer the same for software maintenance or unplanned downtime. You can use PowerHA® SystemMirror® within a partition that is capable of being moved with LPM. This does not mean that PowerHA® SystemMirror® uses LPM in anyway, and it is treated as another application within the partition.

For PowerHA® SystemMirror® clusters configured to use heartbeating and application monitoring with short checking intervals, you must complete testing to validate that the period of suspension during LPM does not cause unwanted cluster events. You can greatly minimize the chance of unwanted cluster events by stopping cluster services with the **Unmanage Resource Group** option in SMIT on the node in the cluster that LPM is going to be performed on. You do not want to interfere with any applications during the LPM process. While the cluster is in an unmanaged state, PowerHA® SystemMirror® does not monitor any applications. Therefore, you must monitor the applications during the LPM process. If an LPAR failure occurs during the LPM process, you can start the workload on a standby node.

PowerHA® SystemMirror® automates some of the LPM steps by registering a script with the LPM framework.

PowerHA® SystemMirror® listens to LPM events and automates steps in PowerHA® SystemMirror® to handle the LPAR freeze that might occur during the LPM process. As part of the automation, PowerHA® SystemMirror® provides a few variables that can be changed based on the requirements for your environment.

You can also set LPM_POLICY to UNMANAGED state which will stop the cluster services in UNMANAGED state before LPM operation, and restart the cluster services after LPM operation. While restarting the cluster services, the application monitors are run to check whether any applications need to be restarted. A failure or disruption might occur if the application monitor is not configured for application monitoring or if the application monitor script disrupts the cluster services. To avoid failure or disruption after LPM operation, you can use the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable.

> **Note:** You do not have to take any actions or change any application monitor or application controller scripts after LPM operation.

If you use the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable, any cluster resource processing is ignored after cluster services are restarted. After LPM operation, the entire partition and all cluster resources are restarted and are set automatically to the previous state. However, if you use the SKIP_EVENT_PROCESSING_MANAGE_MODE tunable and if you manually change the configuration of the cluster resources when the cluster services are in the UNMANAGED state, you must manually restore the cluster resources to its previous state. PowerHA® SystemMirror® does not determine automatically the state of the cluster service and set the state of the cluster service to the previous state before you use the **Unmanaged resource groups** option. For more information, see "Starting cluster services on a node with a resource group in the UNMANAGED state" on page 165.

Configuring SAN communication with LPM

If a PowerHA® SystemMirror® cluster is configured and deployed with storage area network (SAN) communication, you might have to complete a few extra steps before and after you use Live Partition Mobility (LPM). SAN communication is not required for various cluster operations.

PowerHA® SystemMirror® uses network-based communication for typical cluster management. If your cluster environment is configured for SAN communications and a critical network failure occurs, the cluster functions switch from network-based communication to SAN communication. If SAN communications fail or are not available, the cluster functions use repository disk-based health management.
You can perform LPM on a PowerHA® SystemMirror® LPAR that is configured with SAN communication. You can use LPM when network communication is configured for SAN communication. However, when you use LPM the SAN communication is not automatically migrated to the destination system. You must configure SAN communication on the destination system before you use LPM.

Perform LPM operations on an LPAR only when the network communication is healthy from the LPAR across the cluster.

To use LPM on a node in the cluster that is configured with SAN communication, complete the following steps:

1. Verify that the TME flag for the FC adapters inside VIOS is set to **yes**.

If you change the TME flag, yout might need to reboot the system because the change requires the adapter to be reinitialized. Therefore, access to the storage disk through the adapter is disrupted. You must proactively plan for this disruption and enable the TME flag before you start the LPM process.

2. Start the LPM process. You can ignore warning messages on the target side for SAN communication. Specifically, you can ignore any warning messages about the missing VLAN port 3358.

   On the destination VIOS system, enter the **`lsdev -Ct storfwork`** command to verify whether a sfwcommX device is already identified as a VLAN storage framework communication.

3. On the destination system, reestablish SAN communication between VIOS and the client LPAR. To configure SAN communication on the destination system, you must configure the virtual LAN adapters that are related to SAN communication for the VLAN 3358 adapter on VIOS. To reestablish VLAN communication between the client LPAR and SAN communication module, use the **`cfgmgr`** command from VIOS.

   The time taken to reestablish communication for the SAN communication routing tables depends on the number of hops or elements in the SAN fabric between the host systems. The time taken to reestablish communication does not affect cluster operations.

Performing LPM operation on a cluster node and modifying LPM variables

PowerHA® SystemMirror® automates some of the Live Partition Mobility (LPM) steps by registering a script with the LPM framework.

PowerHA® SystemMirror® listens to LPM events and automates steps in PowerHA® SystemMirror® to handle the LPAR freeze that might occur during the LPM process. As part of the automation, PowerHA® SystemMirror® provides a few variables that can be changed based on the requirements for your environment.

You can change the following LPM variables in PowerHA® SystemMirror® that provide LPM automation:

Variable	Description	Usage options
• Command line: *HEARTBEAT_FREQUENCY_DURING_LPM* • SMIT interface: **Node Failure Detection Timeout during LPM**	You can use this variable to increase in seconds the node failure detection time (in seconds) across the cluster, during the LPM process. Ensure that the LPM process that is related to LPAR freeze times does not exceed the node failure detection time. The LPAR freeze time is different and it is a smaller value as compared to the entire duration needed for the LPM process. If you do not specify a value for this variable, then the maximum value of 600 seconds is used during the LPM process.	• Command line: `clmgr modify cluster HEARTBEAT_FREQUENCY_DURING_LPM=<node_timeout>` • SMIT interface: **Custom Cluster Configuration › Cluster Nodes and Networks › Manage the Cluster › Cluster heartbeat settings** and select the **Node Failure Detection Timeout during LPM** field.
• Command line: *LPM_POLICY* • SMIT interface: **LPM Node Policy**	You can use this variable to set the resource groups on the LPAR to be in an **unmanage** state or in a **manage** state during the LPM process. The default value for this variable is **manage**. **Note:** Do not modify any applications during the LPM process. While the cluster is in an **unmanage** state, PowerHA® SystemMirror® does not monitor any applications. If an LPAR failure occurs during the	• Command line: `clmgr modify cluster LPM_POLICY=unmanage` • SMIT interface: **Custom Cluster Configuration › Cluster Nodes and Networks › Manage the Cluster › Cluster heartbeat settings** and select the **LPM Node Policy** field.

Variable	Description	Usage options
	LPM process, you can start the application on a standby node. You can reduce the occurrence of an unwanted cluster event by disabling cluster services during the LPM process. To disable cluster services, specify **unmanage** for this variable.	

Note: You must verify and synchronize the cluster for any changes that you make to the LPM variables.

If you have not set the *LPM_POLICY* variable to **unmanage**, you can manually stop cluster services and perform the LPM operation on a cluster node by completing the following steps:

1. From the command line, enter **smitty cl_stop**.
2. In the SMIT interface, select **PowerHA® SystemMirror® Services › Stop Cluster Services**, and press Enter.
3. Select the **Select an Action on Resource Groups** field and press F4.
4. From the list, select **Unmanage Resource Groups** and press Enter to stop cluster services.
5. Perform the LPM process.
6. After the LPM process is complete, from the command line, enter **smitty clstart**.
7. Select the value for the **Manage Resource Groups** field as `Automatically` and press Enter.

Backing up and restoring an application data in PowerHA® SystemMirror® by using cloud storage

In a high availability environment, you can use redundant resources to recover critical applications by switching the workload to a different set of resources. In addition to redundant resources, such as nodes and networks, you can back up application data that you can access after a failure occurs. You can use different options to back up application data, such as a local solution (shared disks) or a remote backup solution that uses storage hardware-based backup mechanisms.

In PowerHA® SystemMirror® Version 7.2.3, or later, you can add another layer of data backup by using IBM Cloud® and Amazon Web Services (AWS) cloud backup solutions.

You must configure a primary backup profile for your IBM® SAN Volume Controller (SVC) storage, then you can configure the cloud solution to function as a backup for the remote copy of the data on the local SVC storage device.

Starting with PowerHA® SystemMirror® Version 7.2.7, or later, you can take a backup of volume group data that is stored on IBM® System Storage® DS8000® (DS8K) disks, upload it to cloud services, and restore on any type of storage disks.

Planning for backing up data by using cloud storage in PowerHA® SystemMirror®

PowerHA® SystemMirror® Version 7.2.3, or later, supports IBM Cloud® and Amazon Web Services (AWS) cloud storage solutions. Before you configure PowerHA® SystemMirror® to use a cloud storage solution, you must understand the capabilities and requirements for the available cloud storage options.

Both IBM Cloud® and AWS cloud storage solutions use an external internet connection from your cluster environment to access cloud services. Therefore, you must plan how your cluster environment network can

access the external could services. For example, you can enabling your cluster environment to access the internet through a firewall.

You must configure a primary backup profile for your IBM® SAN Volume Controller (SVC) storage, then you can configure the cloud solution to function as a backup for the remote copy of the data on the local SVC storage device.

Starting with PowerHA® SystemMirror® Version 7.2.7, or later, you can take a backup of volume group data that is stored on IBM® System Storage® DS8000® (DS8K) disks, upload it to cloud services, and restore on any type of storage disks.

Limitations

The following are the limitations apply while using PowerHA® SystemMirror® Version 7.2.3, or later, with cloud storage solutions:

- Only resource groups are available to backup data. A volume group that you want to backup, must be part of a resource group.
- Only IBM® SAN Volume Controller (SVC) storage devices are support for the local data backup.
- You must define the FlashCopy® mapping in the SVC storage device. The FlashCopy® must be associated with the SVC consistency group.
- Only server-side encryption is supported.
- Concurrent resource groups are not supported.
- The backup profile names must use the following names:

  IBM Cloud®

  ```
  ibm_profile
  ```

  AWS

  ```
  aws_profile
  ```

- You can restore data from a cloud backup file by using cloud backup restoration feature. For more information, see Restoring backup profiles.

Prerequisites

The following prerequisites apply while using PowerHA® SystemMirror® Version 7.2.3, or later with cloud storage solutions:

- The Backup management function requires Python 2.7.x or Python 3.x.x with the Boto software development kit. You can install the required Python modules from the IBM® COS SDK for Python website. For more information about configuring the Boto software development kit, see the Boto 3 Documentation website.
- By default, Python 3 is installed on IBM AIX® 7.3. You must use the Rational Portfolio Manager (RPM) filesets to install any Python 3.x.x with related Boto software development kit. It is necessary to use a soft link between the old and the new Python versions after installation. Here is an example of a soft link.

  ```
  ln -sf /opt/freeware/bin/python3 /usr/bin/python3
  ```

- You must install and use Secure Socket Shell (SSH) to create a secure connection to the SVC/DS8K storage devices. To run SSH commands on the PowerHA® SystemMirror® cluster, you must have root access. For more information about running SSH commands, see Setting up an SSH client the topic.
- For DS8K, install the DS8000® data storage command line interface (DSCLI).
- Before you use the cloud backup function, complete the following steps:
  a. Configure a primary backup profile for your storage. The data that is backed up on the local storage is uploaded to the cloud storage environment.
  b. Use SSH to verify that the nodes in the cluster can access the storage devices. For more information about configuring SSH, see the Setting up an SSH client topic.

c. Create a storage system for the backup data by running the **clmgr add storage_system** command or by using the **Cluster Applications and Resources › Cloud Backup Configuration › Stroage Configuration** SMITH path.

d. Only for SVC storage, create a flash copy mapping to SVC storage devices that correspond to the volume groups.

e. Create a backup profile by running the **clmgr add backup_profile** command or by using the SMIT options that are described in the "Configuring backup profiles" on page 391 topic.

f. Verify and synchronize the cluster.

Configuring PowerHA® SystemMirror® for cloud storage

You can use the SMIT interface and the **clmgr** command to configure PowerHA® SystemMirror® Version 7.2.3, or later, to work with cloud storage options.

If you are not familiar with the processes for using IBM Cloud™ or Amazon Web Services (AWS), see the following websites:

- AWS documentation
- IBM Cloud™ documentation

Configuring Secured Shell (SSH) for SVC/DS8K storage

Configure the SSH protocol for each node that communicates with storage system. Public or private key pairs must be created on the cluster nodes and distributed to storage systems. PowerHA® SystemMirror® uses SSH protocol to manage the backup process on storage system. For more information about how to configure SSH for storage system, see the storage documentation.

Configuring the Boto interface

1. To use IBM Cloud™ or AWS storage solutions, you must configure the Boto interface. For more information about the Boto interface, see the Boto 3 Documentation website.
2. Configure authentication credentials in the Boto interface.
   a. Log in to the Identity and Access Management (IAM) Console
   b. If you have an installation of AWS CLI, run the **aws configure** command.
   c. Add the following information in the `/.aws/credentials` file for the corresponding cloud solution:

   > **Note:** You must use the exact profile names that are specified in the following examples.

   Examples:
   IBM Cloud™

   ```
   [ibm_profile]
   aws_access_key_id = YOUR_IBM_ACCESS_KEY
   aws_secret_access_key = YOUR_IBM_SECRET_KEY
   ```

   AWS

   ```
   [aws_profile]
   aws_access_key_id = YOUR_ACCESS_KEY
   aws_secret_access_key = YOUR_SECRET_KEY
   ```

3. Configure the storage regions. You must update the storage region in the `<user_home>/.aws/config` file for the corresponding cloud solution to identify the cloud bucket is used.

Note: You must use the exact profile names that are specified in the following examples.

Examples:

IBM Cloud™

If your bucket exists under the AP cross region use the following.

```
[profile ibm_profile]
ibm_profile_endpoint=https://s3.ap-geo.objectstorage.softlayer.net
```

AWS

If your bucket exist under the Asia Pacific (Mumbai) region use the following:

```
[profile aws_profile]
region=ap-south-1
```

Note: Ensure that your node time is synchronized with the cloud time to avoid out of synchronization errors. Use Network Time Protocol (NTP) to synchronize your node time with cloud time.

Restoring data from a backup profile

When you restore data from a backup profile, the target disk must be of the same size as the disk that was used when you created the backup profile. To know the size of your disk, run the following command:

```
bootinfo -s <hdisk>
```

The naming convention for the backup image file is: *<cluster name>_<resource group name>_<volume group name>_<disk name>_<timestamp>*.img.gz

For example: `cluster11_RG1_VG1cbm_1G_2_hdisk5.2018-07-07_05:45:16.449614.mg.gz`

To restore data from the backup profile, complete the following steps:

1. Extract the compressed backup image file by running the following command. Do not perform this step if the backup image is not in the compressed format.

   ```
   tar -xvf <image.gz>
   ```

2. Transfer all data to the target disk by running the following command:

   ```
   dd if=/backup_file of=/dev/<target disk> bs=1024
   ```

   The backup data is now accessible on the target disk. However, the **dd** command might change the PVID of the volume group.

3. Clear the PVID from the Volume Group Descriptor Area (VGDA) of the volume group and then reassign the PVID to the volume group on all cluster nodes by running the following commands:

   ```
   chdev -l hdisk$i -a pv=clear
   chdev -l hdisk$i -a pv=yes
   ```

4. Import the disk into the existing volume group by running the following command:

   ```
   importvg -V <majornumber> -y <vgname> <hdisk>
   ```

5. If the disks are duplicated from another set of disks that belong to a specific volume group, you might need to re-create the volume group by running the **recreatevg** command.

6. If you used your backup profile to back up the `rootvg` volume group, you can verify whether the backup restoration was successful by changing the operating system **`bootlist`** to point to the restored disks and by rebooting the node.

Popular tasks

The following list contains popular task that can help you to configure and troubleshoot your PowerHA® SystemMirror® environment for cloud storage:

- To run consistency checks in your cluster, run the **`clmgr verify cluster`** command.
- To view the storage connectivity status for the backup profile, run the **`smitty hacmp`** command and select **Problem Determination Tools › Storage Connectivity**.
- For information about the consistency group operations, image file creation, and cloud storage upload details, view the `/var/hacmp/log/clutils.log` file.
- To view the status of the backup process and the schedule for when the next backup occurs, run the **`clmgr query backup_profile <backup_profile>`** command.
- To view all files that were uploaded to the cloud storage service, run the **`clmgr query backup_files <backup_profile>`** command.
- To verify that the connection to the cloud storage device is available, run the **`curl <address>`** command.

Configuring backup profiles

You can use the SMIT interface or the **`clmgr add backup_profile`** command to configure backup profiles.

To configure backup profiles in PowerHA® SystemMirror® Version 7.2.3, or later, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. From the SMIT interface, select **Cluster Applications and Resources › Cloud Backup Configuration › Backup Profiles** , and press Enter.
3. Complete the following fields, and then press Enter.

   Backup Profile

   Specify the name of the backup profile. You can select any of the configured resource groups or the rootvg profile as the backup profile name. If you want to back up all resource groups in a cluster, select **ALL**.

   Enable Backup

   Select this option to enable or disable the backup profile. You can also use this option to temporarily postpone the backup operation. The default value is **Yes**.

   Backup Method

   Select this option to specify the backup solution. Select **Cloud** to use IBM Cloud™™ or AWS. Select **Remote Storage** to use SVC storage based replication. The default value is **Cloud**. The following fields are displayed if you select the backup method as **Cloud**:

   Bucket Name

   Specify the name of an existing cloud storage bucket to upload the files into the cloud storage device.

   Target Location

   Specify a directory or file system that must be used to temporarily store backup files. A backup file is created in the specified location with data from the selected volume groups, and then the backup file is uploaded to the bucket in the cloud storage that you want to use. The last version of the backup file that was uploaded to the cloud storage is available in the target location. You must have adequate free space to store backup files of all volume groups that are specified in the selected backup profile.

   Cloud Service

   Specify **IBM** or **AWS** as the cloud service provider. The default value is **IBM**.

   Compression

Specify the compression option for the backup data. Select **Enable** to automatically compress all backup data. Select **Disable** if you do not want to compress data. The default value is **Disable**. The compression option is not supported for file sizes that are larger than 10 GB.

> **Note:** The **gzip** program is used for data compression.

Backup Frequency

Specify the number of days for which you want PowerHA® SystemMirror® to collect the data for backup. Acceptable values for this field are in the range 0 - 999. For example, if you specify the value as 7, the data is backed up once a week. If you specify the value as 0, the data is backed up only once.

Backup Schedule

Specify the time when you want the backup process to start. Acceptable values for this field are in the range 00:00 - 23:59. The default value is 00:42.

Incremental Backup Frequency

Specify the frequency, in hours, when you want PowerHA® SystemMirror® to collect the incremental backup data. The value for this field cannot be higher than the value specified in the **Backup frequency** field. Acceptable values for this field are in the range 0 - 999. You can use the **Incremental Backup Frequency** option along with the **Backup Frequency** option.

Encryption Algorithm

Specify **Yes** and select the encryption type as **KMS** or **AES** to encrypt the backup data. Specify **Disable** if you do not want to encrypt the backup data. The default value is **Disabled**. If the selected cloud service is`IBM`, only **AES** can be used as the encryption algorithm.

Volume Group(s)

Select the volume groups that apply to the backup profile. If you have selected a resource group as the back up profile name, the volume groups that you specify must be a part of the selected resource group. If you want to back up all volume groups for a selected resource group, select **ALL**.

Replicated Resources

This field is available only if you are using IBM® SAN Volume Controller (SVC) or IBM System Storage DS8000 (DS8K) storage devices.

For SVC storage

- If the value of the **Backup Method** field is **Cloud**, you must specify a FlashCopy® consistency group. The specified FlashCopy® consistency group must contain FlashCopy® mappings between the source disks and the target disks that are created in the storage system. The source disks and the target disks must be of the same size. The source disks and the target disks must be mapped to all nodes in a specific resource group before starting the FlashCopy® operation. The FlashCopy® consistency group must not be part of any resource group in the cluster.
- If the value of the **Backup Method** field is **Remote Storage**, you must specify a SVC Peer-to-Peer Remote Copy (PPRC) consistency group. The specified consistency group must contain relationships between the source disks and the target disks that are created in the storage system. The target disk is not required to be mapped to any cluster nodes. The SVC PPRC consistency group must not be part of any resource group in the cluster.

For DS8K storage

Create a mapping file anywhere in the node with all the target and source disks listed down there. There are two columns in the mapping file, first column contains source disks Volume ID's list, and second column consists of target disks Volume ID's list. You must specify this mapping file name in the **Replicated Resources** option.

Storage Name

Select the storage system names that contain the replicated resources that are defined in the backup profile. If you are using IBM® SVC storage device for a remote copy operation, you can use only the

SVC PPRC primary console. If you are using IBM® PowerHA® SystemMirror® Enterprise Edition configuration, you must define the storage devices separately for the backup operation.

Notify Method

Specify a custom method (script) to be called if the backup operation fails. You must specify the full path name to the notify method. The script that you specify must exist on all cluster nodes and you must be able to run the script. When the notify method is called, the script receives two parameters as input strings: backup profile name and the reason for failure. The reason for failure is similar to the following example:

```
Flash copy operation failed for Backup profile demo_backup. Check clutils.log for
more details.
```

.

4. Verify and synchronize the changes across the cluster.

Restoring backup profiles

In PowerHA® SystemMirror® Version 7.2.7, or later, you can restore the backup profiles on Cloud to any cluster in PowerHA® SystemMirror® by using the backup restoration feature.

The backup restoration feature restores the data to a volume group by using Cloud backup files that are uploaded to the Cloud during Cloud backup operation. You must specify fields such as volume group name, disks, Cloud bucket name, Cloud service, and temporary directory path where the cloud backup files are temporarily copied. The backup restoration feature is not restricted to any specific storage, and it can be used for all types of storage disks.

> **Note:** Remote backup method cannot be used to restore the backups profile.

You can use the SMIT interface or the **`clmgr add backup_profile`** command to restore the backup profiles.

For more information about restoring backup profiles on Cloud by using the **`clmgr add backup_profile`** command, see clmgr command.

To restore the backup profiles in PowerHA® SystemMirror® Version 7.2.7, or later, complete the following steps:

1. From the command line, enter `smit sysmirror`.
2. From the SMIT interface, select **Cluster Applications and Resources › Cloud Backup Configuration › Restore Cloud Backup** and press Enter.
3. Complete the following fields and then press Enter:
   Backup file

   Name of the backup file that is used to identify the specific backup.

   Cloud Service

   Name of the Cloud service where the specified cloud bucket exists. The supported values are IBM® and AWS. The default value is IBM®.

   Download path

   Directory path where the downloaded files are temporarily stored, and automatically deleted later. The directory must have sufficient space to store all temporary files. Restoration of backup files might fail due to insufficient storage in the directory.

   Node names

   List of node names where the backup files are restored.

   Source disks

   List of disk names where the backup files are stored. To select the target disk for backup restoration, the disk size is also displayed along with the disk name.

   Target disks

Select disks into which a volume group must to be restored. The volume group is created on the selected disks. The target disk count must match with number of source disks.

Volume group

Name of a volume group that is created from the backup files. You must specify a unique new name that is different from names of volume groups that exists on the cluster.

Note: After you restore the volume group on the PowerHA® SystemMirror® cluster, you can check the names of the associated file system and the logical volume. You can change these names by using the **clmgr** command or the **smitty cspoc** option.

Understanding network instability

PowerHA SystemMirror relies on **Cluster Aware AIX** (CAA) and **Reliable Scalable Cluster Technology** (RSCT) to monitor the health of the network interfaces on the cluster nodes.

When a network interface fails, or begins operating normally, PowerHA SystemMirror runs a cluster event for that interface in return. If the change of state affects an entire network, PowerHA SystemMirror might also run a network event.

The cluster events might take additional actions if cluster resources are affected by the change of state.

For example, a change of state for a network interface might cause a persistent IP address to be moved to a different network interface.

Under certain circumstances, CAA and RSCT might generate multiple change of state events in a short period of time. If these change of state events affect cluster resources, multiple, unnecessary and disruptive recovery actions might be required. Therefore, the **network stability** feature of a PowerHA SystemMirror allows you to define the threshold number of events that can occur within a certain period of time.

If PowerHA SystemMirror receives more than the expected threshold number of events within the specified time, it will run a **network_unstable** event instead of continuing to run individual network interfaces and network events. Running a single network event instead of continuing to run multiple network events avoids unnecessary and potentially disruptive recovery actions. The **network_unstable** event continues to run as long as PowerHA SystemMirror continues to receive additional change of state notifications within the specified time period. Messages are stored in the `hacmp.out` log file. The message indicates that the network that is unstable and indicate suitable recovery actions. You can also configure pre events, post events and notification events for the `network_unstable` event

When PowerHA SystemMirror is no longer receiving multiple, continuous events, the **network_unstable** event is completed and a **network_stable** event is run. If the affected network is up at this time, a **network_up** event is run.

Notices

This information was developed for products and services offered in the US.

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

IBM® may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM® Intellectual Property Department in your country or send inquiries, in writing, to:
Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

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

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:
IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM® under terms of the IBM® Customer Agreement, IBM® International Program License Agreement or any equivalent agreement between us.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM® has not tested those products and cannot confirm the

accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All IBM® prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

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

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM®, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM®, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM® shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work must include a copyright notice as follows:

© (your company name) (year).

Portions of this code are derived from IBM® Corp. Sample Programs.

© Copyright IBM® Corp. _enter the year or years_.

Privacy policy considerations

IBM® Software products, including software as a service solutions, (“Software Offerings”) may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering’s use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as the customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, see IBM®’s Privacy Policy at http://www.ibm.com/privacy and IBM®’s Online Privacy Statement at http://www.ibm.com/privacy/details the section entitled “Cookies, Web Beacons and Other Technologies” and the “IBM® Software Products and Software-as-a-Service Privacy Statement” at http://www.ibm.com/software/info/product-privacy.

Trademarks

IBM®, the IBM® logo, and ibm.com® are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM® or other companies. A current list of IBM® trademarks is available on the web at Copyright and trademark information at www.ibm.com/legal/copytrade.shtml.

The registered trademark Linux® is used pursuant to a sublicense from the Linux® Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

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

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

Index

/

/etc/hosts 15
/etc/inittab 15
/etc/services 15
/etc/snmpd.conf 15, 15
/etc/snmpdv3.conf 15
/etc/syslog.conf 19, 15
/tmp/clconvert.log 306
/usr/es/sbin/cluster/snapshots 305
/var/hacmp/clverify/clverify.log 119
/var/spool/cron/crontabs/root 15

7

7.1.2 or earlier 243
7x24 maintenance 310
7x24 maintenance
  hardware maintenance 319
  planning 310
  preventative maintenance 321
  runtime maintenance 316

A

Adding
  Node 43
adding
  attribute
    extended configuration path 89
  custom configuration
    tape resources 66
  custom verification method 127
  resource
    extended configuration path 89
  user-defined resource 68
AIX
  /etc/hosts 15
  /etc/inittab 15
  /etc/services 15
  /etc/snmpd.conf 15, 15
  /etc/snmpdv3.conf 15
  /etc/syslog.conf 19, 15
  /var/spool/cron/crontabs/root 15
AIX Live Update 41
AIX Logical Volume Manager
  C-SPOC 194
  managing 193
AIX Workload Manager
  configuring 84
  reconfiguration 86
  shutdown 86
  startup 86
  verifying configuration 85
AIX Workload Partitions
  running
    resource group 97
application
  monitoring 180
application controller
  extended configuration path
    configuring 54
  reconfiguring 254
  standard configuration path
    configuring 25
application monitor
  changing 255
  custom 55
  custom configuration
    configuring custom 63
    configuring multiple 55
    configuring process 59
  process 55
  removing 255
application service interface
  extended configuration path
    configuring 46

C

C-SPOC 194, 232
capacity on demand
  examples 369
  overview 347
changing
  application monitors 255
  cluster
    name 246
  communication interfaces configuration 249
  custom remote notification 109
  custom verification method 127
  file collection using SMIT 125, 126
  network configuration 247
  node configuration 246
  passwords 285
  resource group 262
  snapshot of cluster configuration 309
Changing host name 244, 243
cldisk command 185
clinfo 170
Cloud instance for Power Systems Virtual Servers
  Adding definition 374
  Changing instance 375
  Removing definition 375
  Removing instance 375
clstat 173
CLSTRMGR_KILL command 150
cluster 307
cluster
  automatic verification 110
  changing name 246
  configuration

changing a snapshot 309
removing snapshot 310
restoring from snapshot 307
configuring
standard configuration path 21, 22
configuring events 13
configuring options 11
configuring security 289
custom testing 137
events 100
extended configuration path
configuring 40
resetting tunables 43
format of snapshot 305
maintaining 13
maintaining cluster information services 170
managing 279
managing group 286
monitoring 14
monitoring
with clstat 173
monitoring tools 172
reconfiguring dynamically 235, 253
restoring 304
running verification 110
saving 304
starting cluster services 162
stopping cluster services 166
synchronizing configuration 251
synchronizing resources 261
testing 13, 131
testing automatically 133
testing overview 131
testing tool 131
verifying
manually 113
using SMIT 113
verifying configuration 13
clverify.log 119, 127
CoD 347
collection
Configuration_Files 121
HACMP_Files 121
command
post-event 101
pre-event 101
communication interface
changing configuration 249
managing 237
Configuration_Files 121
configure
snapshot 210
storage 209
Configure
Resource Optimized High Availability 361, 373
configuring
AIX Workload Manager 84
AIX Workload Manager
verifying 85
cluster events 13
cluster security 289
custom remote notification method 105
delayed fallback timer 88
extended configuration path
application controllers 54
application monitors 55
application service interface 46
cluster 40
custom application monitor 63
file system 55
IP-based network 45
logical volume 55
persistent node IP labels/addresses 47
process application monitor 59
resource group 77
resource group runtime policies 75
resource groups 71
resources 50
service IP labels 51
tape drives as resources 66
topology 40
volume group 55
heartbeats 41
LVM split-site mirroring 229
LVM split-site mirroring for a new volume group 230
LVM split-site mirroring for an existing volume group 230
message authentication 292
message encryption 292
mirror pools 231
options 11
settling time
resource group 86
standard configuration path 24
standard configuration path
application controllers 25
file system 27
logical volume 27
overview 21
resource groups 30
resources 25, 31
service IP addresses 26
service IP labels 26
steps 22
viewing configuration 34
volume group 27
user-defined resource 29
user-defined resource types 27
Configuring
AIX Live Update 41
critical volume groups 207
creating 307
creating

custom remote notification 106
file collection
using SMIT 123
shared volume groups 200
test plan 138
custom configuration
adding
tape resources 66
configuring
application controllers 54
application service interface 46
cluster 40
file system 55
logical volume 55
persistent node IP labels/addresses 47
resource group 77
resource groups 71
resources 50
service IP labels 51
topology 40
volume group 55
options 40
resetting
cluster tunables 43
synchronizing
tape drive configuration 67
verifying
tape drive configuration 67
custom remote notification
changing 109
configuring 105
creating 106
deleting 109
custom verification method
adding 127
changing 127
removing 127
showing 127

D

defining
delayed fallback timer 87
standard configuration path
topology 25
delayed fallback timer
configuring 88
defining 87
deleting
custom remote notification 109
DLPAR 347
dynamic logical partitioning
examples 369
overview 347
dynamic reconfiguration 251

E

enable
safeguard 211
enabling
safeguard 209
event
duration
tuning 103
post 100
pre 100
example
CoD 369
DLPAR 369
location dependency 333
resource group behavior 333
test plan 152
extended configuration path
configuring
custom application monitor 63
IP-based network 45
multiple application monitors 55
process application monitor 59
resource group runtime policies 75
tape drives as resources 66
extending
LVM split-site mirroring 231

F

file collection
changing using SMIT 125, 126
creating using SMIT 123
managing 121
propagating 122
removing using SMIT 126
setting automatic timer using SMIT 125
synchronizing using SMIT 126
verifying using SMIT 126
file system
custom configuration
configuring 55
standard configuration path
configuring 27
forcing
varyon
volume group 96
volume group
forcing varyon 96

G

group
managing 286

H

HACMP_Files 121
hardware maintenance 319
heartbeats
configuring 41
HMC
Adding definition 361
Configure with Resource Optimized High Availability 357
host name 245

I

inactive component 120
IP address 245
IP security filter rules 289

L

Live Partition Mobility (LPM) 385
location dependency
  example 333
log
  error logs 154
logical volume
  custom configuration
    configuring 55
  maintaining 212
  maintaining concurrent 232
  standard configuration path
    configuring 27
LVM 193
LVM split-site mirroring
  configuring 229
  extending 231

M

maintaining
  cluster 13
  cluster information services 170
  concurrent access volume groups 234
  concurrent logical volumes 232
  logical volumes 212
  physical volumes 223
  raw volumes 222
  shared volume groups 196
managing
  communication interfaces 237
  file collections 121
  groups 286
  keys 293
  password changes 282
  persistent node IP labels 250
  shared LVM components 193
  user accounts 279
message
  configuring
    authentication 292
    encryption 292
Migrating
  Oracle RAC cluster 208
monitoring
  applications 180
  cluster 14
  cluster
    with clstat 173
  tools 172
moving
  resource group 273

N

network
  changing configuration 247
  custom configuration
    configuring IP-based 45
  testing 136
  testing for IP network 142
  testing interface 145
NFS 94, 260
Node
  Adding 43
node
  changing configuration 246
  testing 140

O

overview
  CoD 347
  configuring
    standard configuration path 21
  DLPAR 347
  testing 131

P

password 282
password
  changing 285
persistent node IP address
  extended configuration path
    configuring 47
persistent node IP label
  extended configuration path
    configuring 47
  managing 250
physical volume
  maintaining 223
planning
  7x24 maintenance 310
  custom test procedure 137
  test procedure 137
PowerVM NovaLink
  Adding definition 44
preventative maintenance 321

R

raw volume
  maintaining 222
reconfiguring
  application controllers 254
  cluster 235
  cluster
    dynamically 253
  dynamically 235
  resources 260
  service IP label 257
  tape drive resource 259
recovering
  resource groups 332
removing
  application monitor 255

custom verification method 127
file collection using SMIT 126
snapshot 310
report
inactive components 120
repository disk
replacing 50, 193
resetting
custom configuration
cluster tunables 43
resource
custom configuration
configuring tape drives 66
dynamic reconfiguration 251
extended configuration path
adding tape 66
configuring 50
reconfiguring 260
standard configuration path
configuring 25
resource group
acquisition failures 330
changing 262
configuring
settling time 86
custom configuration
configuring 71, 77
configuring runtime policies 75
event handling 324
example of behavior 333
moving 273
processing order 82
recovering 332
selective fallover 327
standard configuration path
configuring 30
configuring resources 31
testing 136, 146
resources
user-defined 27, 29, 68
restoring
cluster configuration 304
cluster configuration from snapshot 307
running
cluster verification 110
custom test procedures 152
resource
in AIX Workload Partitions 97
runtime maintenance 316

S
safeguard
enable 211
enabling 209
safeguarded
volume group 211
SAN
Live Partition Mobility (LPM) 385
saving
cluster configuration 304
script
post-event 100
pre-event 100
security
configuring 289
managing keys 293
standard mode 291
selective fallover
resource groups 327
service IP address
standard configuration path
configuring 26
service IP label
custom configuration
configuring 51
reconfiguring 257
standard configuration path
configuring 26
settling time
resource group
configuring 86
shard processor pool 369
site
adding 242
SMIT
automatic timer for file collections 125
changing file collection 125, 126
creating file collection 123
removing file collection 126
synchronizing file collection 126
verifying file collection 126
snapshot
changing 306, 309
configure 210
creating 307
defining custom 306
format 305
removing 306, 310
restoring from 307
snapshot of configuration 307
standard configuration path
configuring
application controller 25
cluster 22
file system 27
logical volume 27
resource 25
resource groups 30
resources in resource group 31
service IP addresses 26
service IP labels 26
volume group 27
defining
topology 25
synchronizing 33
verifying 33
viewing
configuration 34
starting

cluster services 162
stopping
cluster services 166
storage
configure 209
synchronizing
cluster configuration 251
cluster resources 261
configuration 13
extended configuration path
tape drive configuration 67
file collections using SMIT 126
standard configuration path 33

T
tape drive
custom configuration
synchronizing configuration 67
verifying configuration 67
reconfiguring resource 259
temporary host name 244
testing
cluster 13, 131
cluster
automatically 133
custom 137
creating test plan 138
custom configuration
customized resource 70
error logging 154
evaluating results 153
example test plan 152
fixing problems 159
general 149
IP network 142
network interface 145
networks 136
node 140
overview 131
planning 137
planning custom 137
resource group 136
resource groups 146
running custom 152
syntax 140
topology 135
volume group 137
volume groups 148
topology
dynamic reconfiguration 251
extended configuration path
configuring 40
standard configuration path
defining 25
testing 135
viewing 236
tuning
event duration time 103

U
user
changing password 285
managing passwords 282
managing user accounts 279

V
verifying
AIX Workload Manager
configuration 85
cluster
automatically 110
manually 113
using SMIT 113
configuration 13
extended configuration path
tape drive configuration 67
file collection using SMIT 126
standard configuration path 33
viewing
standard configuration path
configuration 34
topology 236
volume group
creating shared 200
extended configuration path
configuring 55
maintaining concurrent access 234
maintaining shared 196
safeguarded 211
standard configuration path
configuring 27
testing 137, 148

W
WLM 84
Workload Manager 84
Workload Partitions 97
WPAR 97

© Copyright International Business Machines Corporation 2017, 2025
US Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp