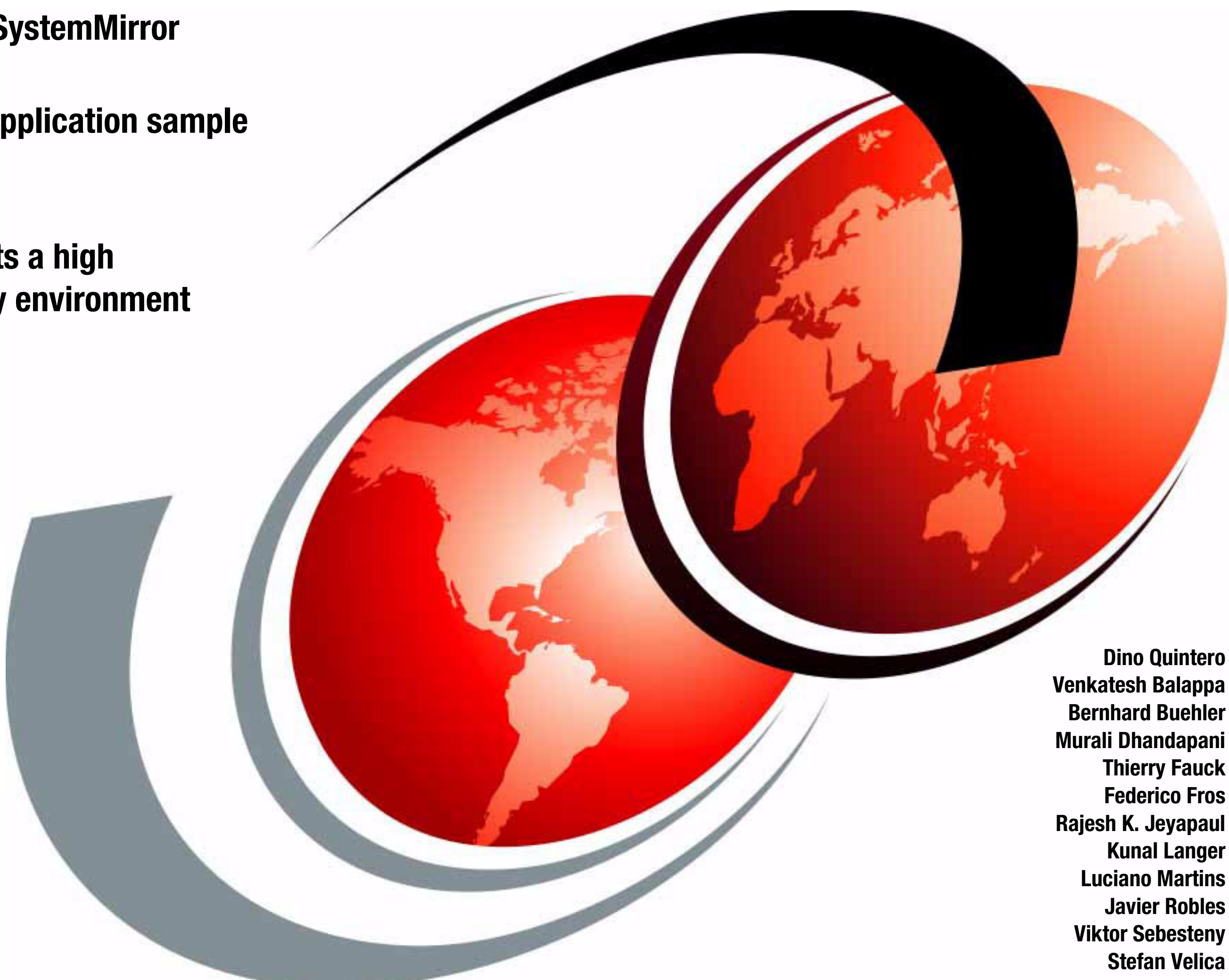
# IBM PowerHA SystemMirror Standard Edition 7.1.1 for AIX Update

**Introduces the latest features of IBM PowerHA SystemMirror**

**Provides application sample scenarios**

**Implements a high availability environment**




Dino Quintero
Venkatesh Balappa
Bernhard Buehler
Murali Dhandapani
Thierry Fauck
Federico Fros
Rajesh K. Jeyapaul
Kunal Langer
Luciano Martins
Javier Robles
Viktor Sebesteny
Stefan Velica
Klaus Micha Zuehlke



Redbooks

ibm.com/redbooks

International Technical Support Organization

# IBM PowerHA SystemMirror Standard Edition 7.1.1 for AIX Update

October 2012

SG24-8030-00

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

**First Edition (October 2012)**

This edition applies to these products:

- IBM AIX 7100-01-03-1207
- IBM AIX 7.1 TL01 SP02
- IBM AIX 6100-07-00-0000
- IBM AIX 6100-06-00-0000
- IBM PowerHA 7.1.1.1
- IBM PowerHA 7.1.0 SP5
- IBM PowerHA SystemMirror plug-in V7.1.1
- IBM DB2 V9.7.0.4, s110330, IP23236, Fix Pack 4
- IBM Tivoli Directory Server V6.2.1
- IBM Tivoli Directory Server V6.3
- IBM Tivoli Storage Manager Server V6.2.3
- EWAS Version 7.0
- WebSphere MQ V7.1.0
- UNIX
- SAP kernel release 720, kernel make variant, 720_REL, compiled on AIX 2 5 00092901D600 for rs6000_64, compiled for 64 BIT, compilation mode UNICODE, compile time May 23 2011 21:24:01, update level 0, patch number 90, source id 0.090, liveCache: Version 7.7.07

# Contents



# Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM 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 supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which 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.

# Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

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

AIX®
DB2 Universal Database™
DB2®
DS4000®
DS6000™
DS8000®
eServer™
FileNet®
FlashCopy®
Global Technology Services®
GPFS™
HACMP™
IBM®
Jazz™
Lotus®
MQSeries®
Power Systems™
POWER6®
POWER7®
PowerHA®
PowerVM®
POWER®
pSeries®
Rational®
Redbooks®
Redbooks (logo) ®
solidDB®
Storwize®
System i®
System p®
System Storage®
SystemMirror®
Tivoli®
UC™
WebSphere®
XIV®

The following terms are trademarks of other companies:

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

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

Snapshot, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

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

# Preface

This IBM® Redbooks® publication helps you install, tailor, and configure the new IBM PowerHA® SystemMirror® for AIX® 7.1.1 Standard Edition. This book gives an understanding of the Cluster Aware AIX (CAA). This book helps you design a solution to migrate from the previous version of the IBM PowerHA.

This IBM Redbooks publication is targeted toward technical professionals (consultants, technical support staff, IT architects, and IT specialists) responsible for providing continuous availability solutions and support.

## The team who wrote this book

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

**Dino Quintero** is a Senior Project Leader and an IT generalist with the International Technical Support Organization (ITSO) in Poughkeepsie, NY. His areas of knowledge include enterprise continuous availability planning and implementation, enterprise systems management, virtualization, and clustering solutions. He is currently an Open Group Master Certified IT Specialist. Dino holds a Master of Computing Information Systems degree and a Bachelor of Science degree in Computer Science from Marist College.

**Venkatesh Balappa** is Systems Software Engineer in IBM PowerHA SystemMirror Service Pack Test team Bangalore, India. I have more than 4.5 years of experience in AIX and PowerHA SystemMirror testing. I hold a Bachelor of Engineering degree in Computer Science from B V Bhoomaraddi college of Engineering and Technology in Hubli, India. My areas of expertise include the Hardware Management Console (HMC), Lightweight Directory Access Protocol (LDAP), PowerHA SystemMirror, and AIX.

**Bernhard Buehler** is an IT Specialist in Germany. He currently works for IBM STG Lab Services in La Gaude, France. He has worked at IBM for 31 years and has 22 years of experience in AIX and the availability field. His areas of expertise include AIX, PowerHA, HA architecture, script programming, and AIX security. He is a co-author of several IBM Redbooks publications. He is also a co-author of several courses in the IBM AIX curriculum.

**Murali Dhandapani** is Certified IT Specialist in IBM India. He is currently working for IBM India Software Lab Operations, where he is a technical lead for IBM Rational® Jazz™ products infrastructure, high availability, and disaster recovery deployment. His areas of expertise include Linux, AIX, IBM POWER® Virtualization, PowerHA SystemMirror, System Management, and Rational tools. Murali has a Master of Computer Science degree. He is an IBM Certified Specialist in System p® administration and an IBM eServer™ Certified Systems Expert - pSeries® High Availability Cluster Multi-Processing (IBM HACMP™).

**Thierry Fauck** is a Certified IT Specialist in France. He has 25 years of experience in support of major high-performance computing (HPC) providers as a system administrator of the French development lab. His expertise includes AIX, VIOS, SAN and IBM PowerVM®. He currently leads a FVT development team for WPAR and WPAR mobility features. He is a co-author of several IBM Redbooks publications including AIX, WPAR, and PowerHA.

**Federico Fros** is an IT Specialist who currently works for IBM Global Technologies Services in Uruguay. He has worked in IBM for eight years, including six years of experience in IBM Power Systems™. He is an IBM Certified Systems Expert for UNIX and High Availability. His areas of expertise include AIX, Linux, PowerHA SystemMirror, PowerVM, and IBM System Storage®.

**Rajesh K. Jeyapaul** is the technical lead for IBM Systems Director POWER Server management. His focus is on the PowerHA SystemMirror plug-in and the PowerRF interface for Virtualization Management Control (VMC) Plug-in on System Director. He is part of the Technical advocate team that works closely work with clients to tackle their POWER Server-related issues. Rajesh holds a Master in Software Systems degree from the University of BITS, India, and a Master of Business Administration (MBA) degree from the University of MKU, India.

**Kunal Langer** is a Technical Consultant - Power Systems and AIX for IBM India. He currently works for IBM STG Lab Services in Bangalore, India. He has worked in IBM for more than five years. He worked with HACMP for over five years in testing, support, and development. Kunal is a Certified Specialist for IBM System p Administration and HACMP for AIX. His areas of expertise include IBM System p Servers, AIX, and PowerHA. Kunal holds a Bachelor degree in Computer Science Engineering from VTU, India.

**Luciano Martins** is a Senior IT Specialist working in IBM Global Technology Services® in IBM Brazil. He has 13 years of experience in the IT industry, mainly focused on IBM Power and IBM Storage System solutions. He is a Certified Technical Expert for AIX/PowerHA and POWER Systems, and he is also an IBM Certified Cloud Computing Infrastructure Architect. His areas of expertise include AIX, PowerVM, PowerHA, IBM General Parallel File System (GPFS™), Cloud Computing, and Storage Systems. Luciano holds a degree in Computer Science from Brazil, with academic work in Virtualization and High Availability.

**Javier Robles** is an IT Specialist, who currently works for GBM, an IBM Business Partner in Costa Rica. He has more than 10 years experience in Power Systems. He is a Certified Systems Expert for AIX and HACMP. His areas of expertise include AIX, PowerHA, PowerVM, and IBM DB2®. Javier holds a degree in Computer Science from Costa Rica.

**Viktor Sebesteny** is an IT Specialist in Hungary. He has worked for IBM for 16 years. He supports AIX and PowerHA for Hungarian and international clients from the CEEMEA region. He holds a Bachelor degree in Computer Science from KKMF University, Budapest, Hungary. His areas of expertise include Power Systems, AIX, PowerHA, PowerVM, and Linux. He has co-authored two previous PowerHA/HACMP IBM Redbooks publications.

**Stefan Velica** is an IT Specialist who currently works for IBM Global Technologies Services in Romania. He has five years of experience in Power Systems. He is a Certified Specialist for IBM System p Administration, HACMP for AIX, High-end and Entry/Midrange DS Series, and Storage Networking Solutions. His areas of expertise include IBM System Storage, PowerVM, AIX, and PowerHA. Stefan holds a Bachelor degree in Electronics and Telecommunications Engineering from Politechnical Institute of Bucharest.

**Klaus Micha Zuehlke** is an IT Specialist, who currently works for IBM Global Technology Services in Germany. He has more than 15 years experience in Power Systems. He is a Certified Technical Expert for AIX and HACMP/PowerHA. His areas of expertise include AIX, PowerHA, PowerVM, IBM System Storage, and Tivoli® Systems Management. He is also a Certified Technology Consultant for SAP NetWeaver 6.40 and 7.00 and for OS/DB Migration. Klaus holds a degree in Mathematics from Germany.

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

Richard Conway, David Bennin
International Technical Support Organization, Poughkeepsie Center

Ravikiran Moningi, Dimpu Nath
IBM India

Mike Coffey, Paul Moyer, Christopher Tulloch, Duane Witherspoon
IBM Poughkeepsie

Katharina Probst
IBM Germany

Ravi A. Shankar, Tom Weaver, Matthew Ochs, Rajeev Nimmagadda, Rohit Krishna Prasad
IBM Austin

Shawn Bodily
IBM Dallas

Claudio Marcantoni
IBM Italy

# Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

**ibm.com**/redbooks/residencies.html

# Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  **ibm.com**/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

## Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

# Introducing IBM PowerHA SystemMirror 7.1.1

This chapter describes the mechanism on which PowerHA SystemMirror works to allow enterprise-level business services to become highly available.

This chapter covers PowerHA SystemMirror concepts, including these new features:

- PowerHA architecture concepts
- Hardware environment
- Cluster scenarios
- New features on PowerHA 7.1.1
- PowerHA 7.1.1 installation
- Migrating to PowerHA 7.1.1

**Resources:** For more information about PowerHA 7.1.1, see the IBM Information Center for PowerHA SystemMirror:

http://publib.boulder.ibm.com/infocenter/aix/v7r1/topic/com.ibm.aix.powerha.navigation/powerha_main.htm

## 1.1 PowerHA architecture concepts

Before you start with the PowerHA features, we suggest a good understanding of the main goals and special PowerHA concepts.

One of the PowerHA solution main goals is to help continuous business services operations even after one (or more) components experience failures. Unexpected failures can happen and they can be related to human errors or other errors. Either way, the PowerHA design phase is intended to remove any *Single Point of Failure (SPOF)* from the environment by using redundant components and automated PowerHA procedures.

It is important to remember that any hardware component can experience failures and cause application disruptions. So, when you plan a highly available environment, you must check all components, from disk access to power circuits, for redundancy.

See Figure 1-1 and Figure 1-2 on page 3 for examples of how an environment can be configured with or without redundant components, and what behavior can be expected in each scenario.

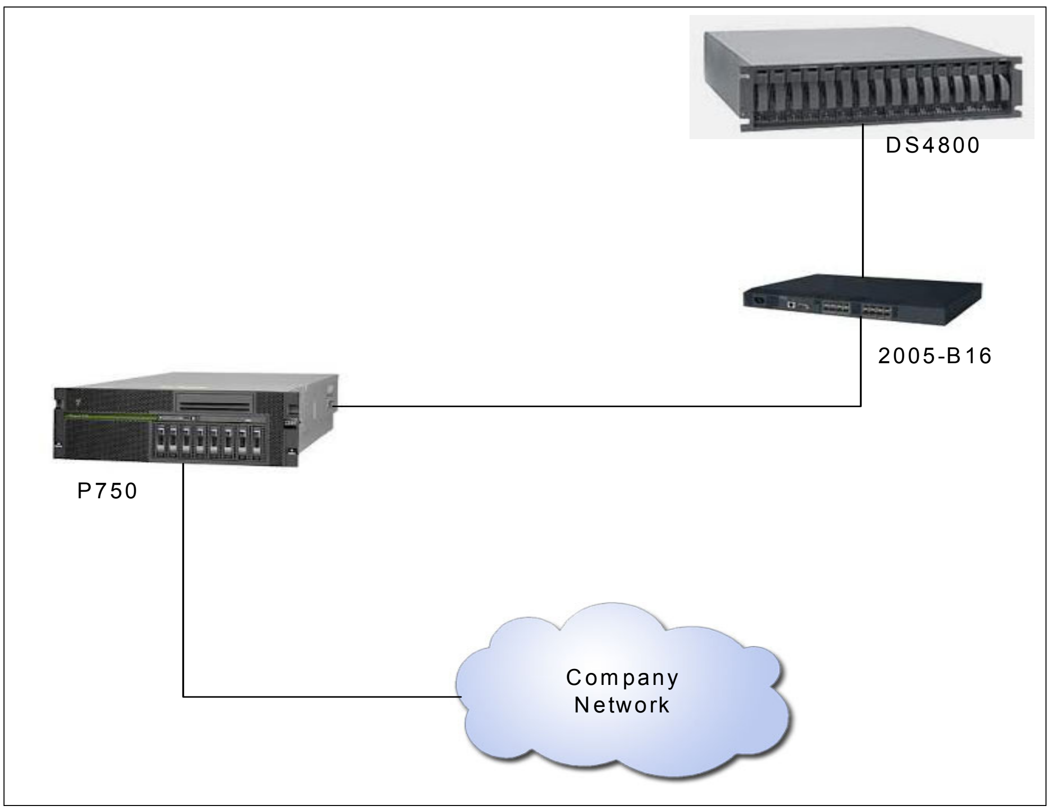



*Figure 1-1 A sample client environment with no fault tolerance*

Figure 1-1 shows a client environment with no fault tolerance mechanisms. So, if any component fails, for example, the company network switch, or the SAN switch, the application that runs on the IBM Power 750 server becomes unavailable due to lack of redundancy.

If failures occur with this configuration, the IBM Power 750 server experiences a disruption until all failing components are replaced or fixed. Depending on which component fails, it can take from hours to weeks to fix it, which affects service availability and in the worst case, data availability.

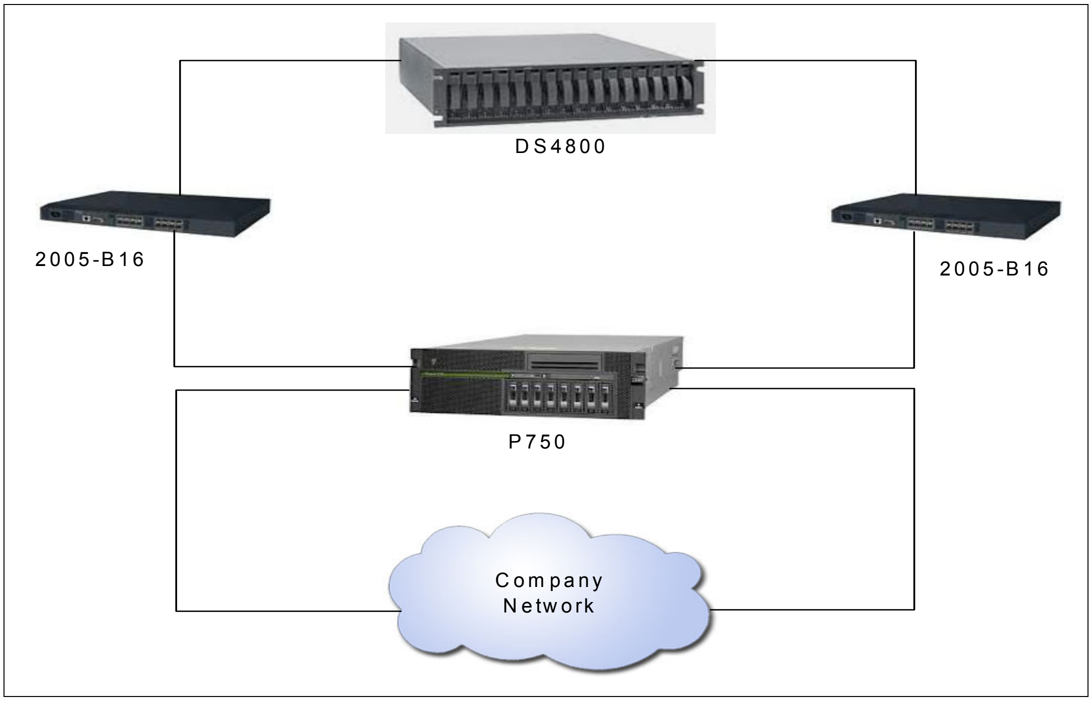


*Figure 1-2 A sample client environment with redundant components*

Figure 1-2 shows a sample client environment with redundant network connections and dual SAN switches for disk access. The configuration in Figure 1-2 enables the IBM Power 750 server to be more resilient to environmental issues. This resiliency keeps business services available even with failures in parts of the company infrastructure.

Even without using PowerHA features, this configuration is resilient to certain possible failures. If an IP network switch becomes unavailable, the IBM Power 750 server has a secondary network connection on a secondary network switch. If a SAN switch is not available, the IBM Power 750 server can reach its data smoothly through the secondary SAN switch. Therefore, the client environment becomes more resilient to unexpected issues. The client environment allows business services and all their data to be active, accurate, and available to any user all the time.

PowerHA, as requirement, needs redundancy on many components, for example:

- Network access
- SAN disk access
- SAN disk formatting as Redundant Array of Independent Disks (RAID)

When a production environment is to be migrated to a clustered infrastructure, all possible components must be assessed to address all necessary redundancies before the cluster setup. It is important to avoid issues that are caused by no redundant components or SPOF.

For more details about infrastructure requirements for the PowerHA cluster buildup, see Chapter 3, “PowerHA 7.1.1 basic installation and configuration” on page 73.

**High availability solution:** It is important to mention that a high availability solution such as IBM PowerHA SystemMirror for AIX helps with the failure of any component, such as hardware, software, or system management. A high availability solution prevents the application and its data from being inaccessible to the user community through the elimination or masking of both planned and unplanned downtime. High availability solutions help eliminate SPOFs through appropriate design, planning, selection of hardware, configuration of software, and carefully controlled change management discipline. High availability does not mean that there is no interruption to the application; therefore, it is called fault resilient instead of tolerant.

Before you start with PowerHA features and configurations, it is important to pay attention to the fundamental PowerHA concepts. The concepts allow a better understanding of all scenarios, configurations, and features that are explained in this publication.

This book is focused on special features in PowerHA 7.1.1 implementation. The book briefly explains the important concepts in PowerHA.

**Resources:** If you want more details about PowerHA architecture and concepts, see the *PowerHA SystemMirror Concepts Guide*:

http://pic.dhe.ibm.com/infocenter/aix/v6r1/index.jsp?topic=%2Fcom.ibm.aix.powerha.concepts%2Fha_concepts.htm

### 1.1.1 Reliable Scalable Cluster Technology (RSCT)

RSCT is a set of low-level operating system components that allow clustering technologies implementation such as PowerHA and General Parallel File System (GPFS). The implementation of RSCT is part of the IBM AIX operating system structure. On the current AIX release, AIX 7.1, RSCT is on Version 3.1.2.0.

All RSCT functionalities are based on the following RSCT components:

- Resource Monitoring and Control (RMC) subsystem: It is considered the backbone of RSCT. It runs on each server and provides a common abstraction about server resources (a hardware or software component that provides services to some other component).
- RSCT Core Resource Manager: It is a software layer between a resource and RMC. Resource Manager maps the abstractions that are defined by RMC to real calls and commands for each resource.
- RSCT Security Services: It provides the security infrastructure that is required so that RSCT components can authenticate the identity of other parties.
- Topology Services subsystem: It provides the node and network monitoring and failure detection.

  **Important:** After installing PowerHA and Cluster Aware AIX (CAA) filesets, RSCT Topology Services subsystem is deactivated and all its functionality is performed by CAA.

- Group Services subsystem: It coordinates cross-node operations on cluster environments. This subsystem is responsible to span changes across all cluster nodes and to ensure that all of the changes finished completely with all modifications performed.

Figure 1-3 shows the RSCT filesets that are installed during the AIX 7.1 OS installation with descriptions of their functions.

```
Fileset                      Level  State  Type  Description (Uninstaller)
 ----------------------------------------------------------------------------
 rsct.basic.hacmp           3.1.2.0    C     F    RSCT Basic Function (HACMP/ES
                                                  Support)
 rsct.basic.rte             3.1.2.0    C     F    RSCT Basic Function
 rsct.basic.sp              3.1.2.0    C     F    RSCT Basic Function (PSSP
                                                  Support)
 rsct.compat.basic.hacmp    3.1.2.0    C     F    RSCT Event Management Basic
                                                  Function (HACMP/ES Support)
 rsct.compat.basic.rte      3.1.2.0    C     F    RSCT Event Management Basic
                                                  Function
 rsct.compat.basic.sp       3.1.2.0    C     F    RSCT Event Management Basic
                                                  Function (PSSP Support)
 rsct.compat.clients.hacmp  3.1.2.0    C     F    RSCT Event Management Client
                                                  Function (HACMP/ES Support)
 rsct.compat.clients.rte    3.1.2.0    C     F    RSCT Event Management Client
                                                  Function
 rsct.compat.clients.sp     3.1.2.0    C     F    RSCT Event Management Client
                                                  Function (PSSP Support)
 rsct.core.auditrm          3.1.2.0    C     F    RSCT Audit Log Resource
                                                  Manager
 rsct.core.errm             3.1.2.0    C     F    RSCT Event Response Resource
                                                  Manager
 rsct.core.fsrm             3.1.2.0    C     F    RSCT File System Resource
                                                  Manager
 rsct.core.gui              3.1.2.0    C     F    RSCT Graphical User Interface
 rsct.core.hostrm           3.1.2.0    C     F    RSCT Host Resource Manager
 rsct.core.lprm             3.1.2.0    C     F    RSCT Least Privilege Resource
                                                  Manager
 rsct.core.microsensor      3.1.2.0    C     F    RSCT MicroSensor Resource
                                                  Manager
 rsct.core.rmc              3.1.2.0    C     F    RSCT Resource Monitoring and
                                                  Control
 rsct.core.sec              3.1.2.0    C     F    RSCT Security
 rsct.core.sensorrm         3.1.2.0    C     F    RSCT Sensor Resource Manager
 rsct.core.sr               3.1.2.0    C     F    RSCT Registry
 rsct.core.utils            3.1.2.0    C     F    RSCT Utilities
```

*Figure 1-3 RSCT filesets that are installed during the AIX 7.1 installation*

Figure 1-4 on page 6 shows the placement of the RSCT layer in relationship to the AIX operating system environment.

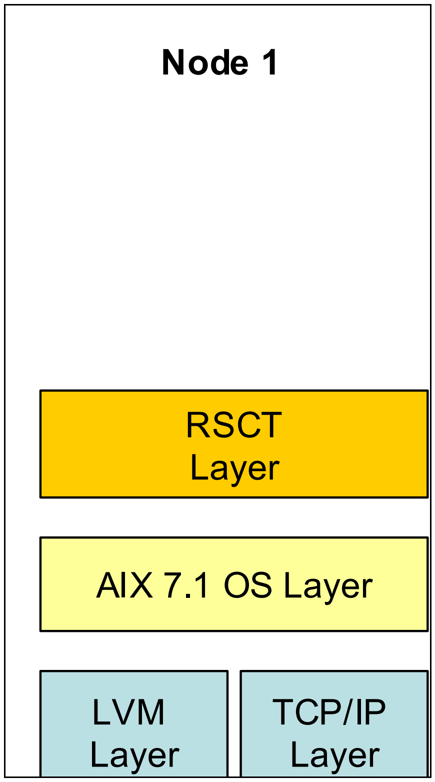


*Figure 1-4 RSCT placement on an AIX server*

**Resources:** For more details about the RSCT structure, see the *RSCT Version 3.1.2.0 Administration Guide,* SA22-7889:

http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=%2Fcom.ibm.cluster.related_libraries.doc%2Frelated.htm&path=3_6

## 1.1.2 Cluster Aware AIX (CAA)

AIX 7.1 (and 6.1 TL6) introduced a new built-in clustering capability called *Cluster Aware AIX (CAA)*. This feature enables system administrators to create clusters from a group of AIX instances by using commands and programming APIs. It includes kernel-based heartbeating, monitoring, and an event infrastructure.

CAA is primarily intended to provide a more reliable layer for clustering products such as PowerHA SystemMirror. Also, clients can directly use CAA layer functionalities to enhance their management tasks in their own computer environments.

CAA also introduces a new component, which is required for PowerHA cluster environments, called *Cluster Repository Disk*. It is a central repository for all cluster topology-related information, and it must be shared by all servers that form the cluster.

In PowerHA 7.1.0, if the repository disk failed, the nodes shut down automatically. A new feature in PowerHA 7.1.1 called *Repository Disk Resilience* allows administrators to perform cluster maintenance tasks, for example, cluster fallover and fallback, even after the repository disk experiences a failure. Now, CAA supports online repository disk replacement with no cluster effects. For more details about Repository Disk Resilience and the procedure to replace a failing repository disk, see 2.2.7, “Repository disk resiliency” on page 44.

> **Important:** At the time of writing this publication, CAA does not support IPv6-based network environments. Consider this aspect of CAA when you design your cluster or migrate to later PowerHA releases.

Figure 1-5 shows how AIX, RSCT, CAA, and PowerHA services interact from a software stack perspective.

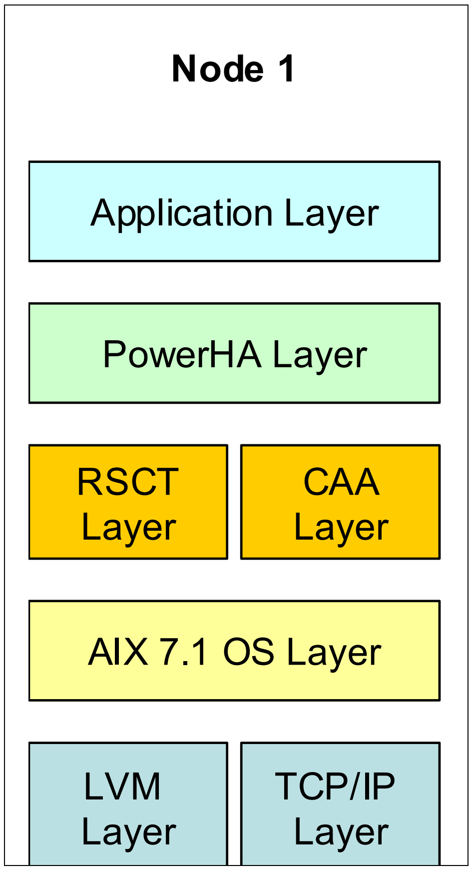



*Figure 1-5 Interaction between AIX, RSCT, CAA, and PowerHA components*

> **For more information:** For more detailed information about CAA features, functionalities, and operations, see 1.1.2, "Cluster Aware AIX (CAA)" on page 6. Consider the deprecated CAA command parameters on PowerHA 7.1.1, specifically with the **`chcluster`**, **`clusterconf`**, and **`rmcluster`** commands.

## 1.1.3 PowerHA cluster components

The following sections provide information about the PowerHA cluster components.

### PowerHA cluster

A *cluster* is a set of connected computer systems that access commonly attached storage. They can be in the same geographical place or data center. Or, they can be separate, such as one data center in the US and another data center in Brazil.

By adopting cluster technologies, companies can increase service reliability to their customers or make disasters not apparent to their customers. So, from the delivered service quality perspective, a clustered environment enables companies to be better service providers. A clustered environment can continue to operate automatically without being affected by a disaster.

When an initial PowerHA cluster is configured, you assign a cluster name (Figure 1-6). The name is used by PowerHA methods and procedures to work with a specific group of machines, services, and information.

```
COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]
Cluster Name: sapnfs
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.30
There are 2 node(s) and 2 network(s) defined

NODE sapnfs1:
        Network net_ether_01
                sapnfssvc1      172.16.21.65
[MORE...21]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 1-6 Cluster name of a running PowerHA cluster*

Figure 1-6 shows the cluster name of a running PowerHA cluster. It can be checked by using **`smitty sysmirror`** and choosing **Cluster Nodes and Networks** → **Manage the Cluster** → **Display PowerHA SystemMirror Configuration** options.

Also, the same output can be obtained by using the **`/usr/es/sbin/cluster/utilities/cltopinfo`** command.

### PowerHA cluster nodes

A *PowerHA cluster node* is any AIX based system that runs PowerHA services and is part of a PowerHA cluster (Figure 1-7 on page 9).

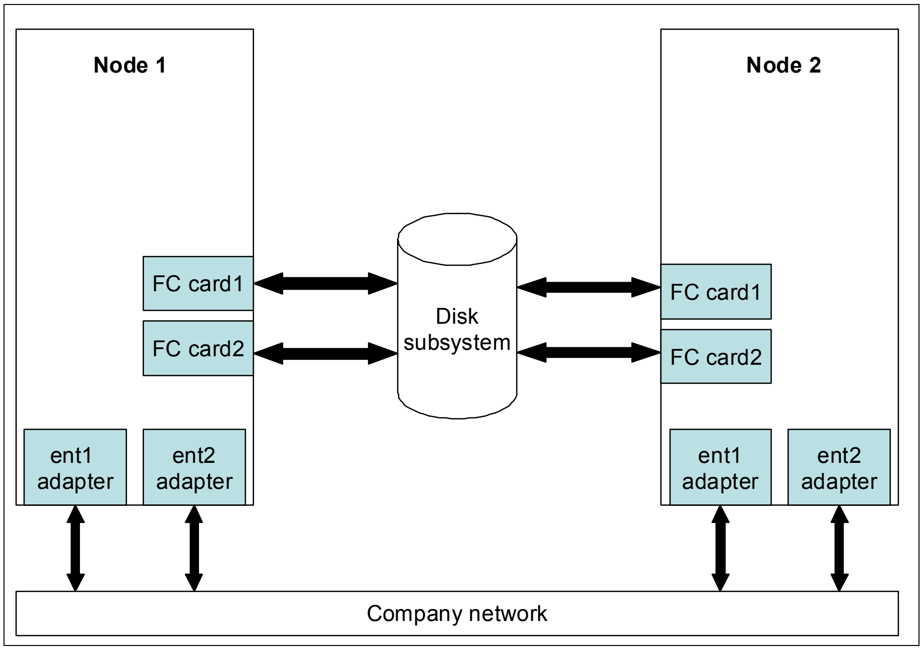



*Figure 1-7 Standard 2-node cluster hardware configuration*

Figure 1-7 shows a standard cluster configuration with two nodes. The network and the SAN accesses are redundant and data is shared due to the use of shared disk subsystems.

In PowerHA 7.1.1, up to 16 nodes can be included in a single cluster. PowerHA 7.1.1 supports cluster nodes, POWER servers that are divided into logical partitions (LPARs), IBM System i®, Blades, stand-alone POWER servers, or a combination.

A cluster requires infrastructure components that are shared among all cluster nodes. For further details about hardware requirements and configuration, see Chapter 3, "PowerHA 7.1.1 basic installation and configuration" on page 73.

**Important:** You cannot dynamically change the host names of any cluster node. To change the name, you must remove the cluster, change the host names, and reconfigure the cluster.

## PowerHA networks

For PowerHA, networks are the paths through which cluster nodes communicate with each other. In addition, all CAA heartbeat messages are sent between the nodes. Also, a network is used by customers or application services to reach services that run on top of the PowerHA cluster layer.

When you define a network, you can choose any name for the network. Make it easy to be identified when you view the cluster topology. Otherwise, PowerHA automatically assigns a network name with the `net_ether_XX` pattern, as shown in Figure 1-8 on page 10.

Starting with PowerHA 7.1.1, the networks can be *Public* or *Private*. The main difference between Public and Private networks is that CAA does not perform heartbeat operations across Private networks.

```
                        Change/Show a Network

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

                                                        [Entry Fields]
* Network Name                                        net_ether_01
  New Network Name                                    []
* Network Type                                        [ether] +
* Netmask(IPv4)/Prefix Length(IPv6)                   [255.255.254.0]
* Network attribute                                   public +

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

*Figure 1-8 Cluster network configuration*

### PowerHA IP addresses and IP labels

In a clustered environment, different IP addresses labels can be used:

- Boot (or base) IP label: It is related to the IP that is physically configured to the Ethernet adapters. It is the first IP that is configured on nodes when they finish a boot process.
- Service IP label: It is related to the IP address that is used by application services users to get into application functionalities and data. Normally, it moves between cluster nodes, according to which node is the current holder of the cluster services.
- Persistent IP label: In many cluster configurations, as in the cluster configuration that is used to develop this book, the boot IP addresses are part of a non-routed network. For specific operating system maintenance tasks where the IT support team needs to reach specific nodes, do not use the service IP addresses. To ensure that the IT support team can reach the cluster nodes even when the cluster is not up, persistent IP addresses are used.

### PowerHA service IP addresses distribution policies

Because cluster nodes can have multiple Ethernet adapters, there are seven policies that can be used to service IP distribution:

- Anti-collocation: The standard cluster policy for allocation of the service IP addresses. PowerHA services allocate the service IP addresses in the maximum number of Ethernet adapters that share the network.
- Collocation: PowerHA allocates all service IP addresses that share the same network on the same Ethernet adapter.
- Anti-collocation with persistent label: PowerHA distributes the service IP addresses to all Ethernet adapters that are not hosting the persistent IP address. Service IP and persistent IP addresses share the same Ethernet adapter only if there is a shortage of interfaces.
- Collocation with persistent label: PowerHA allocates all service IP addresses on the Ethernet adapters that are hosting the node persistent IP address. It is useful for environments where certain firewall restrictions are applied and only specific interfaces are allowed to communicate with external networks.
- Anti-collocation with source: Introduced in PowerHA 7.1, service IP addresses are allocated by using the anti-collocation policy but, if there are not enough adapters, more than one service IP address can be put on one adapter. Then, one service IP address is chosen as the *source address* for communication.

- Collocation with source: Introduced in PowerHA 7.1, service IP addresses are allocated by using the collocation policy. Then, one service IP address is chosen as the *source address* for communication on the outgoing shared adapter.
- Anti-collocation with persistent label and source: Introduced in PowerHA 7.1, service IP addresses are allocated by using the anti-collocation with persistent label policy. One service IP address can be chosen as the source address if there are multiple service IP addresses on the same boot adapter. A common 2-node network cluster is shown in Figure 1-9.

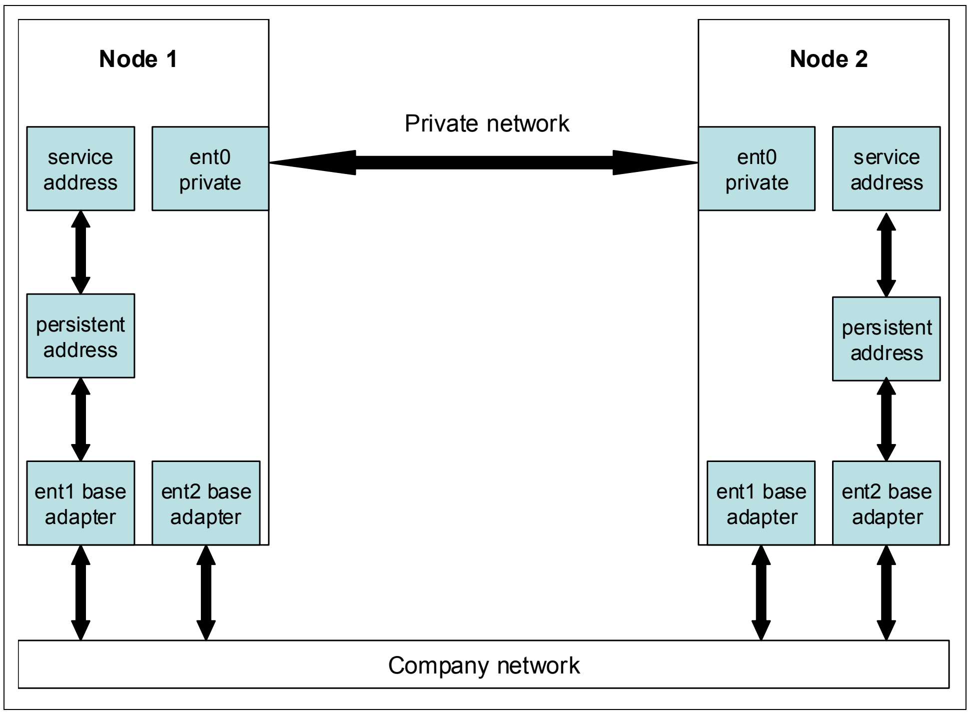



*Figure 1-9 Common network configuration in a 2-node cluster environment*

## IP address (IPAT)

In most environments, when a service is released to user access, it must be a routable IP address. So, when you design a clustered environment, you must consider how the application services are reached by users on any cluster node.

PowerHA implements one mechanism called *IP Address Takeover* (IPAT) in which PowerHA services manage all IP addresses that move between cluster nodes by using IP aliases.

When a resource group is moved to a secondary node for any reason, the service IP address that is assigned to the resource group is activated on the secondary node as an IP address alias. The users do not need to care where the services are running. The IP address that is used on the application layer is the same IP address on any cluster node that runs the resource group services.

## PowerHA resources and resource groups

When a client considers purchasing a cluster solution, the goal is to keep specific business applications highly available. Any application, database, or middleware product can be considered a business application.

A *Resource* is any component that is required to bring up one service application and that, by using PowerHA mechanisms, can be moved from one cluster node to another. A resource can consist of many things:

- File systems
- Raw devices
- Volume groups (VGs)
- IP addresses
- Network File System (NFS) mounts
- Workload partitions (WPARs)
- Applications

To bring up one application, a set of resources is usually required. This set is called the *Resource Group*.

Figure 1-10 shows a sample cluster scenario with basic shared components.

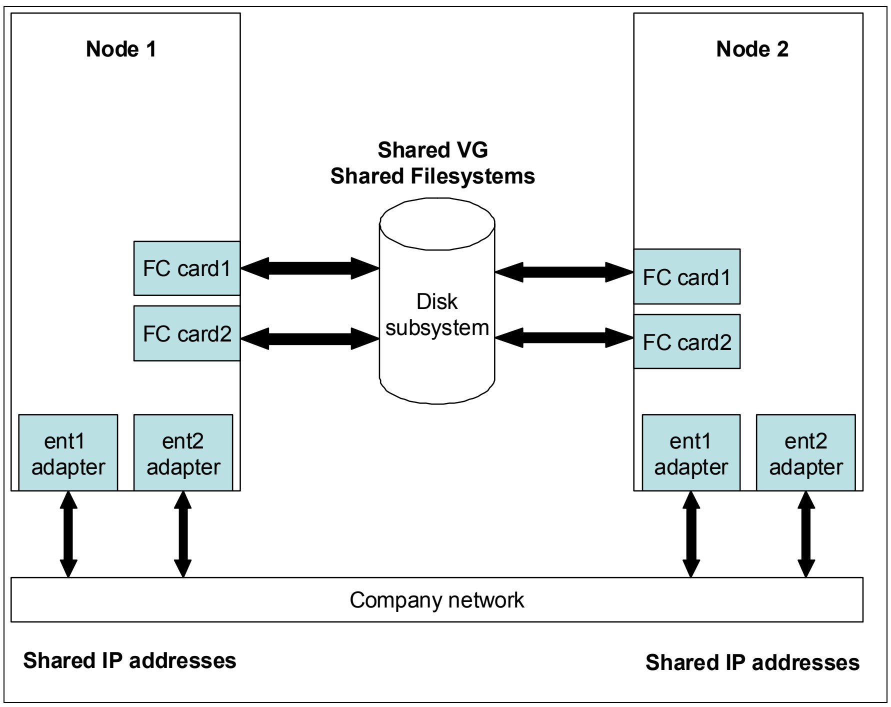



*Figure 1-10 A sample cluster scenario with basic shared components*

Consider Figure 1-10 to create a sample scenario inside a PowerHA. The shared components (IP addresses, VGs, and file systems) can be put together in a single resource group. The shared components are shared and moved between the two cluster nodes.

> **Maximum:** By using PowerHA 7.1.1, you can configure environments up to a total of 64 resource groups.

## PowerHA resource group policies

When you design a PowerHA cluster, you need to plan how you want the cluster to behave in the event of failure. To make it easier, PowerHA uses methodologies to automatically manage this requirement.

When you define a resource group by using **`smitty`**, normally, you see a window similar to Figure 1-11.

```
 Add a Resource Group

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

                                                        [Entry Fields]
* Resource Group Name                                [rg1]
* Participating Nodes (Default Node Priority)        [node1 node2]          +

  Startup Policy                                      Online On Home Node O> +
  Fallover Policy                                     Fallover To Next Prio> +
  Fallback Policy                                     Fallback To Higher Pr> +

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

*Figure 1-11 Resource Group policy definitions on smitty sysmirror menu*

As shown in Figure 1-11, there are three types of resource group policies that must be configured during cluster implementation.

The first policy to be configured in the resource groups is called the *Startup Policy*. It defines when and on which node the resource group is brought online when the cluster is started. The following options are available for the Startup Policy:

- Online on home node only: When this policy is chosen, the resource group is brought online only on the node that is called the *home node*, which is the first one from the left on the Participant Nodes field in Figure 1-11. By using Figure 1-11 as an example, if this policy is chosen, the resource group rg1 is online only when node1 is online.
- Online on first available node: When this policy is chosen, the resource group is brought online on the first participant node that comes online. By using Figure 1-11 as an example, if this policy is chosen, the resource group rg1 is online on node1 if it is the first available node. Or, it is online on node2 if node2 becomes available first.
- Online using distribution policy: When this policy is chosen, the resource group is brought online by following one of these two methods. The resource groups are distributed, trying to keep only one resource group online on each participant node online (*node-based resource group distribution*). Or, they are distributed trying to keep only one resource group per node and per network (*network-based resource group distribution*).
- Online on all available nodes: When this policy is chosen, the resource group is brought online on all available nodes in the cluster. To avoid data corruption or any application issue, ensure that the components included on the resource group can be used concurrently.

For resource group startup, there is a parameter that can be customized on the PowerHA cluster side. It is called Settling Time. By using Settling Time, any cluster node waits for the configured time to ensure that any other higher priority node is not about to join the cluster. It is an important parameter to use when you have a multiple node cluster, and all of the nodes start simultaneously.

> **Setting Time parameter:** To change or define the Setting Time on a PowerHA cluster, type `smitty sysmirror`. Then, choose **Cluster Applications and Resources → Resource Groups → Configure Resource Group Run-Time Policies → Configure Setting Time for Resource Groups**.

The second policy is called the *Fallover Policy*. It defines the behavior of resource groups when the node resource group owning node fails. The following options are available for the Fallover Policy:

- Fallover to next priority node in the list: When the owning node online resource group fails, if the resource group is not online on all available nodes, it is brought online on the next node according to the resource group's participant nodes list (as shown in Figure 1-11 on page 13).
- Fallover using dynamic node priority: When the node owning node online resource group fails, the resource group is moved to another node according to the *Dynamic Node Priority* policy that is defined. These policies are based on RSCT variables, for example, the node with most free memory. Remember that if you choose this option without a defined Dynamic Node Priority policy, you get an error when you perform a cluster configuration synchronization between the cluster nodes.
- Bring offline (on error node only): When the owning node online resource fails, no fallover action is taken. If the resource group is online at one node at a time, the services are unavailable until any administrator action. When the resource group is online on all available nodes, the resource is offline on the failing node only. The resource group continues to work correctly on all another nodes.

The third policy to be configured is called the *Fallback Policy*. This policy defines what happens with a resource group when a higher priority node, which experienced a failure, rejoins on the cluster. The following options are available for the Fallback Policy:

- Fallback to higher priority node in the list: When you use this policy, if a higher priority node returns to the cluster from a previous failure, the resource group is brought offline anywhere it is resident and brought online on the higher priority node. It is important to remember when you use this automatic fallback method that if an intermittent issue occurs on the higher priority node, the cluster applications start an infinite loop of movement between nodes.
- Never fallback: When you use this policy, even if a higher priority node returns from a previous failure, the resource group remains on the lower priority node until a manual resource group move is performed by the cluster administrator. It is an important configuration to consider while you design the cluster because there is a short disruption while the resource groups are moved to a higher priority node.

Another policy that you might configure when you create a resource group, which is not mandatory, is the *Fallback Timer Policy*. By using the Fallback Timer Policy, you can configure on which specific frequency a fallback operation can be performed. The following options are available for this policy:

- Daily: Fallback operations are performed on a daily basis, on the hour and date that are set up by the system administrator.
- Weekly: Fallback operations are performed on a weekly basis, on the day, hour, and time that are specified by the system administrator. You can choose only one week day.
- Monthly: Fallback operations are performed on a monthly basis, on the day of the month, hour, and time that are specified by the system administrator. You can choose only one day of the month.

- Yearly: Fallback operations are performed on a yearly basis, on the month, day of the month, hour, and time that are specified by the system administrator. You can choose only a single year, date, and time.

> **Fallback Timer Policy:** To configure the Fallback Timer Policy, type `smitty sysmirror`. Then, choose **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Delayed Fallback Timer Policies**. Or, use the `smitty cm_timer_menu` fast path menu.

### PowerHA resource group dependencies

In some cases, multiple applications are distributed together between cluster nodes. This design becomes an issue if you do not check whether any type of relationship exists between applications. This design also becomes an issue if one application must be started first to allow another to start correctly.

To address this issue, PowerHA allows administrators to define the order to start the resource group and any restrictions about how to bring up the nodes. These features are *Parent/Child Dependencies*, *Location Dependencies,* and *StartAfter/StopAfter* policies.

With Parent/Child Dependencies, PowerHA allows administrators to define a specific order to start the resource groups. If one web service depends on the availability of a database, you can define which resource groups must be started first by using PowerHA features. Figure 1-12 on page 16 shows multiple levels of Parent/Child Dependency where the web services run in a resource group that is called Web RG. Web RG depends on the application services inside a resource group called Application RG to start first. Also, the application services require that the database data is available.

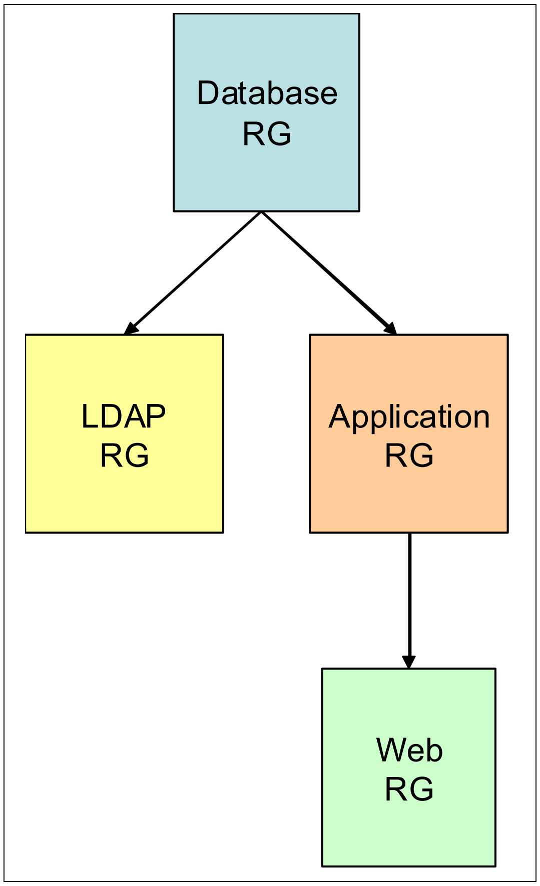


*Figure 1-12 Parent/Child dependency in PowerHA resource groups*

**Resource Group Parent/Child Dependencies configuration:** Enter `smitty sysmirror`. Select **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Parent/Child Dependency**. Or, use the `smitty cm_rg_dependencies` fast path.

Also, there are cases when different application tiers cannot run on the same server due to internal application restrictions, such as shared memory or specific network port usage. To address this requirement, PowerHA allows Resource Group Location Dependencies. This function enables restrictions on the configuration of resource groups where you can define which resource groups cannot be brought online at the same time on the same cluster node (Figure 1-13 on page 17).

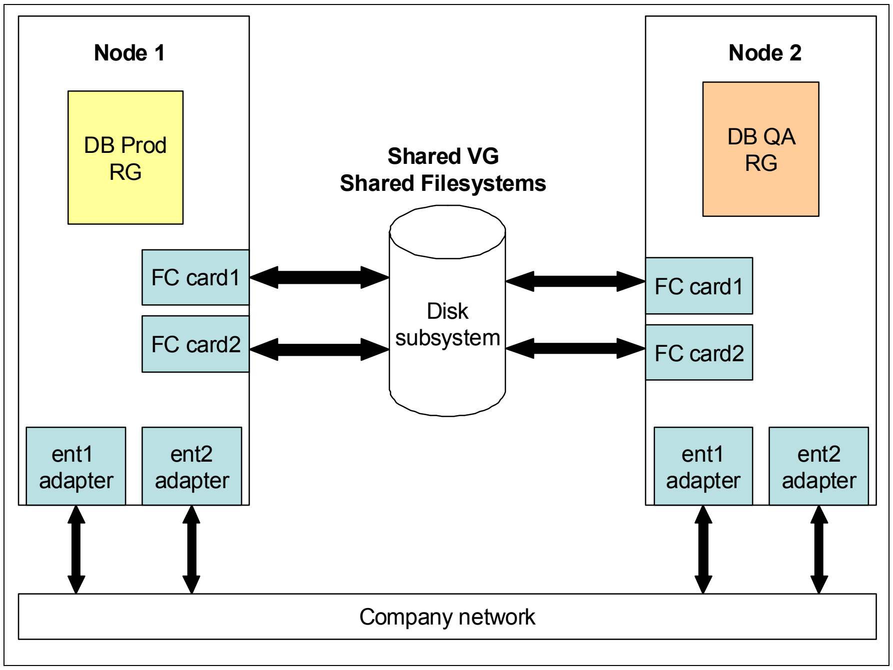



*Figure 1-13 A sample cluster environment with location dependency policies*

Figure 1-13 shows a sample environment where a company has two services that run inside a cluster: a production database and a quality assurance database. Due to the client requirements, the database services that run on DB Prod RG use the same network ports as the database services that run on DB QA RG. An attempt to put one RG online when another RG is already running on a node results in a failure. To prevent this failure, PowerHA allows the configuration to be restricted by using the Resource Group Location Dependencies Policy.

**Resource Group Parent/Child Dependencies:** Access Resource Group Parent/Child Dependencies by selecting **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Online on the Same Node Dependency** and by using `smitty sysmirror`.

In some cases, application dependencies are not covered by Parent/Child or Location policies. PowerHA V7.1 offers two more policies: *StartAfter* and *StopAfter*. The StartAfter policy allows a special dependency, where one resource group is brought online when another resource group is already started. The StopAfter policy allows a dependency where one resource group can be brought offline only when another one is already shut down.

## PowerHA applications

To PowerHA, any process or service that provides information to users is called an *application*. Because each application can have specific procedures for startup and shutdown, PowerHA requires specific shell scripts to perform application start and stop operations. The PowerHA functions that control structure are the *Application Controller Scripts*.

**Terminology:** Before PowerHA 7.1, the terminology and SMIT menus used *Application Server*. From PowerHA 7.1 forward, we use the new terminology, which is *Application Controller*.

As shown in Figure 1-14, specify which scripts are used to start and stop the application services when they are brought up or down by PowerHA.

```
 Add Application Controller Scripts

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

                                                           [Entry Fields]
* Application Controller Name                           [application01]
* Start Script                                          [/fs1/app01_start.ksh]
* Stop Script                                           [/fs1/app01_stop.ksh]
  Application Monitor Name(s)                                                +
  Application startup mode                              [background]         +

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

*Figure 1-14 Defining Application Controller scripts by using smitty menus*

For each application, you can also create *Application Monitoring Methods*. Application monitoring methods are scripts that perform automatic checks on the application to verify that the application functionalities work correctly. If an application monitoring method fails, by default, the cluster moves the resource group to another node. However, it can also be changed to only notify someone.

PowerHA applications can be created by using **`smitty sysmirror`**. Select **Cluster Nodes and Networks** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Controller Scripts** → **Add Application Controller Scripts**. Or, use the **`smitty cm_add_app_scripts`** fast path.

**Important:** Be careful with application monitoring methods because the default is for a resource group fallover to occur when a monitoring script ends with an error. Any inconsistency that is in these scripts might result in an unnecessary fallover.

## PowerHA cluster events

If you consider all involved components, PowerHA provides ways to monitor any part of the cluster structure. Also, according to the output of these monitoring methods, the PowerHA cluster takes any automatic action. These actions can be a notification or even a resource group fallover.

PowerHA allows the customization of predefined cluster events and also allows the creation of events. When you create events, it is important to check whether there is a predefined event that covers the desired action to avoid creating unnecessary events ending up in duplications.

All cluster events have their own meaning and behavior. Some examples of cluster events can be seen in Table 1-1 on page 19.

*Table 1-1 Examples of standard cluster events*

| Event name | Event type | Quick description |
|---|---|---|
| **node_up** | Nodes joining cluster | A **node_up** event starts when a node joins or rejoins the cluster. |
| **node_down** | Nodes leaving cluster | A **node_down** event starts when the cluster is not receiving heartbeats from a node. It considers the node gone and starts a **node_down** event. |
| **network_up** | Network-related events | A **network_up** event starts when the cluster detects that a network is available and ready for cluster usage (for a service IP address activation, for example). |
| **network_down** | Network-related events | A **network_down** event starts when a specific network is unreachable. It can be a **network_down_local**, when only a specific node lost its connectivity for a network. Or, it can be a **network_down_global**, when all nodes lost connectivity. |
| **swap_adapter** | Network-related events | A **swap_adapter** event starts when the interface that hosts one service IP address fails. If there are other boot networks available on the same node, the **swap_adapter** event moves the service IP address to another boot interface and refreshes the network routing table. |
| **fail_interface** | Interface-related issues | A **fail_interface** event starts when any node interface fails. If the interface has no service IP defined, only the **fail_interface** event runs. If the failing interface hosts a service IP address and there is no other boot interface available to host it, an **rg_move** event is triggered. |
| **join_interface** | Interface-related issues | A **join_interface** event starts when a boot interface becomes available or when it recovers from a failure. |
| **fail_standby** | Interface-related issues | A **fail_standby** event starts when a boot interface that hosts no service IP address fails. |
| **join_standby** | Interface-related issues | A **join_standby** event starts when a boot interface becomes available or when it recovers from a failure. |
| **rg_move** | Resource group changes | An **rg_move** event starts when a resource group operation from one node to another starts. |
| **rg_up** | Resource group changes | An **rg_up** event starts when a resource group is successfully brought online at a node. |
| **rg_down** | Resource group changes | An **rg_down** event starts when a resource group is brought offline. |

**Events:** All events have detailed usage descriptions inside their script files. All events are in the `/usr/es/sbin/cluster/events` directory. In a case of status change, For an example Node_down. PowerHA will log the event in the cluster log files. For example, Sep 12 08:01:16 EVENT START: node_up <NODENAME>.

## PowerHA cluster configurations

PowerHA is flexible and allows many cluster configurations to provide high availability. Several potential configurations are listed with examples to help you understand how they work.

### *Standby configuration*

In this simple cluster configuration, one node runs all services for a resource group while the other nodes are idle. The other nodes are ready to host the resource group services if the main node fails (Figure 1-15).

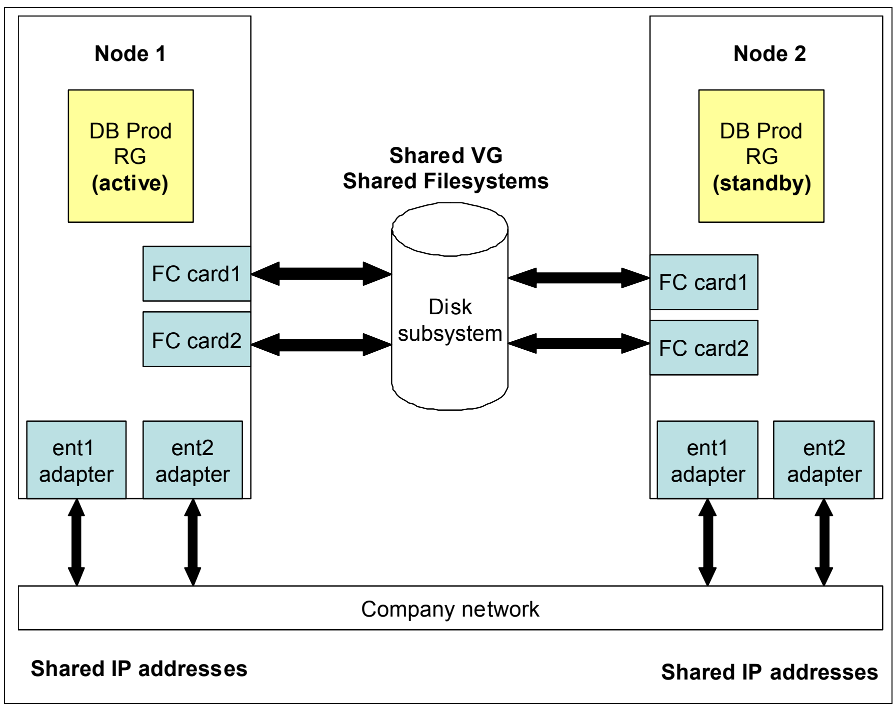



*Figure 1-15 Sample standby cluster configuration*

As shown on Figure 1-15, all cluster services are designed to be up on Node 1, since it is the active cluster node. Node 2 remains idle with no cluster services running on it. Only in case of a Node 1 failure, the DB Prod RG resource group automatically moves to Node 2

### *Takeover configuration*

This configuration allows more efficient hardware usage. All cluster nodes run parts of the production workload. A takeover configuration can be split in two possible subconfigurations: *One-Sided Takeover* or *Mutual Takeover*. Further details about the configurations are shown in Figure 1-16 on page 21 and Figure 1-17 on page 21.

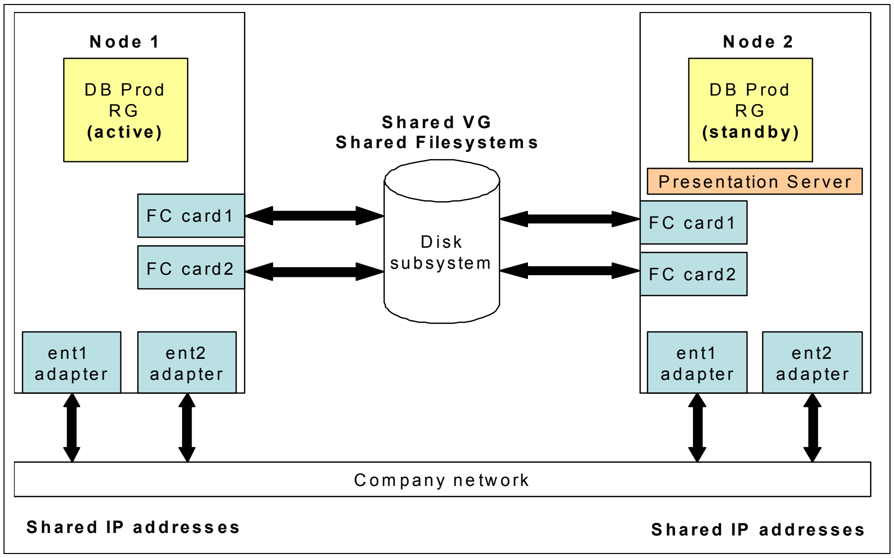



*Figure 1-16 Sample one-sided takeover cluster configuration*

As shown in Figure 1-16, on a one-sided takeover cluster configuration, some application parts are brought up as highly available parts and managed by a resource group. In this example, the DB Prod RG, and some application parts run stand-alone, with no high availability behavior and run outside the cluster structure. If Node 1 fails, its services are automatically brought online on Node 2. But if Node 2 fails, its services remain unavailable until it comes back up and runs in production again.

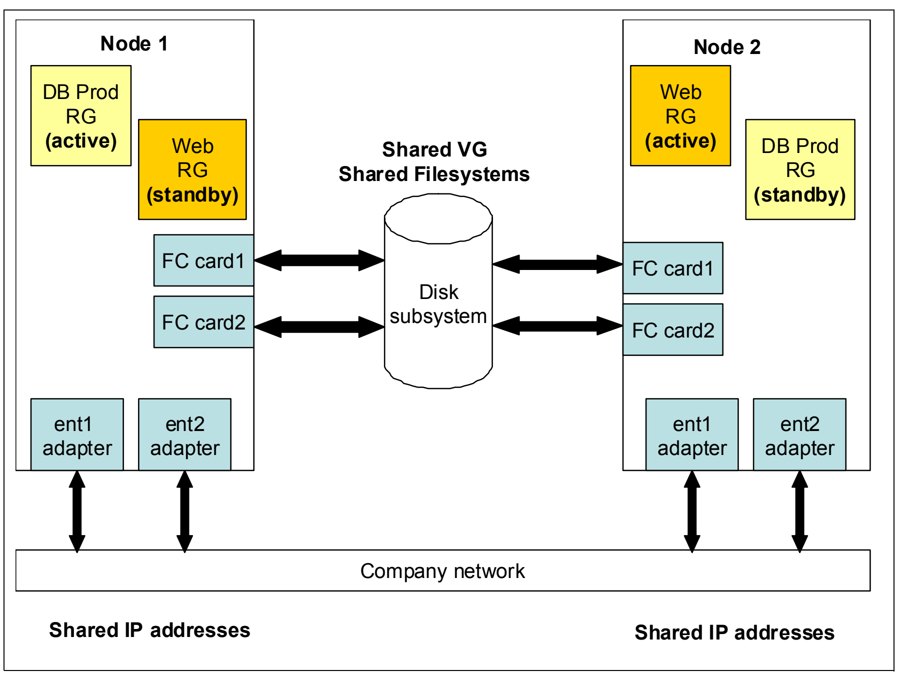



*Figure 1-17 Sample mutual takeover cluster configuration*

As shown in Figure 1-17, on a mutual takeover cluster configuration, all application parts are highly available and managed by a resource group, for example, the DB Prod RG and Web RG. If Node 1 fails, its services are automatically brought online on Node 2. And if Node 2 fails, its services are automatically brought online on Node 1. Either node failure can be covered by the PowerHA cluster structure with minimal effect to the users.

## PowerHA cluster single point of control (C-SPOC)

When you manage a cluster environment, some administrative tasks become more difficult due to the number of managed clusters and managed nodes. Inconsistencies can appear, especially in relationship to the Logical Volume Manager (LVM) structure or user and group management. To avoid these issues, PowerHA provides a way to facilitate administrative tasks on all nodes inside a PowerHA cluster that is called *Cluster Single Point of Control* (C-SPOC).

By using C-SPOC, you can perform the following tasks on all cluster nodes:

- PowerHA services control: startup and shutdown
- Cluster resource group and applications management
- Cluster nodes communication interface management
- File collection management
- Log viewing and management
- AIX operating system user, group, and password management across all cluster nodes
- LVM management
- General Parallel File System (GPFS) tasks
- Smitty session opened on any specific node
- Encrypted File System (EFS) management
- Lightweight Directory Access Protocol (LDAP) integration
- Cluster security

> **Resource:** Throughout this book, many tasks are performed by using C-SPOC functionalities to show specific PowerHA features and behavior. For more detailed information about C-SPOC features, see *PowerHA SystemMirror system management with C-SPOC*:
>
> http://pic.dhe.ibm.com/infocenter/aix/v6r1/index.jsp?topic=%2Fcom.ibm.aix.powerha.concepts%2Fha_concepts_hacmp_cspoc.htm

## PowerHA Smart Assists

Smart Assists are PowerHA tools that help system administrators include applications in a cluster infrastructure. By using Smart Assists, you can create the cluster application environment. Smart Assists take care of the start and stop scripts for the application. Smart Assists work specifically with each application. Individual Smart Assist packages must be installed in addition to the PowerHA base software to allow support for specific applications.

If an application that needs to be included on a cluster environment has no specific Smart Assist product, PowerHA provides a general application Smart Assist (GASA). GASA helps to include those applications in a clustered environment.

Requirements must be addressed before you use Smart Assists:

1. The Smart Assist fileset must be installed on all cluster nodes.
2. Before you use Smart Assists, a basic cluster must be created by using **`smitty`** or the IBM Director interface.
3. Before you configure an application inside the cluster by using Smart Assists, you must ensure that the application already can run manually with no issues on all cluster nodes.
4. We strongly suggest that you configure the application with Smart Assist on the cluster node where the application runs currently.
5. Read the PowerHA documentation to satisfy any documented limitation or consideration.

```
Make Applications Highly Available (Use Smart Assists)

  +--------------------------------------------------------------------------+
  |      Select a Smart Assist From the List of Available Smart Assists      |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   DB2 UDB non-DPF Smart Assist     # smass1 smass2                       |
  |   DHCP Smart Assist                # smass1 smass2                       |
  |   DNS Smart Assist                 # smass1 smass2                       |
  |   Lotus Domino smart assist        # smass1 smass2                       |
  |   FileNet P8 Smart Assist          # smass1 smass2                       |
  |   IBM HTTP Server Smart Assist     # smass1 smass2                       |
  |   SAP MaxDB Smart Assist           # smass1 smass2                       |
  |   Oracle Database Smart Assist     # smass1 smass2                       |
  |   Oracle Application Server Smart Assist # smass1 smass2                 |
  |   Print Subsystem Smart Assist     # smass1 smass2                       |
  |   SAP Smart Assist                 # smass1 smass2                       |
  |   Tivoli Directory Server Smart Assist # smass1 smass2                   |
  |   TSM admin smart assist           # smass1 smass2                       |
  |   TSM client smart assist          # smass1 smass2                       |
  |   TSM server smart assist          # smass1 smass2                       |
  |   WebSphere Smart Assist           # smass1 smass2                       |
  |   WebSphere MQ Smart Assist        # smass1 smass2                       |
  +--------------------------------------------------------------------------+
```

*Figure 1-18 Smart Assists available in IBM PowerHA 7.1.1*

Figure 1-18 shows many Smart Assists that can be used other than the GASA.

Chapter 5, "Smart Assists for PowerHA SystemMirror" on page 227 contains detailed explanations about Smart Assist usage with practical examples.

## 1.2 Hardware environment

In the following chapters, we adopted a common environment (Figure 1-19 on page 24) to perform the scenarios that are presented in this book. In this section, the environment description is introduced from the hardware structure to the virtual environment. We include the naming convention that we used with all partitions and SAN disks.

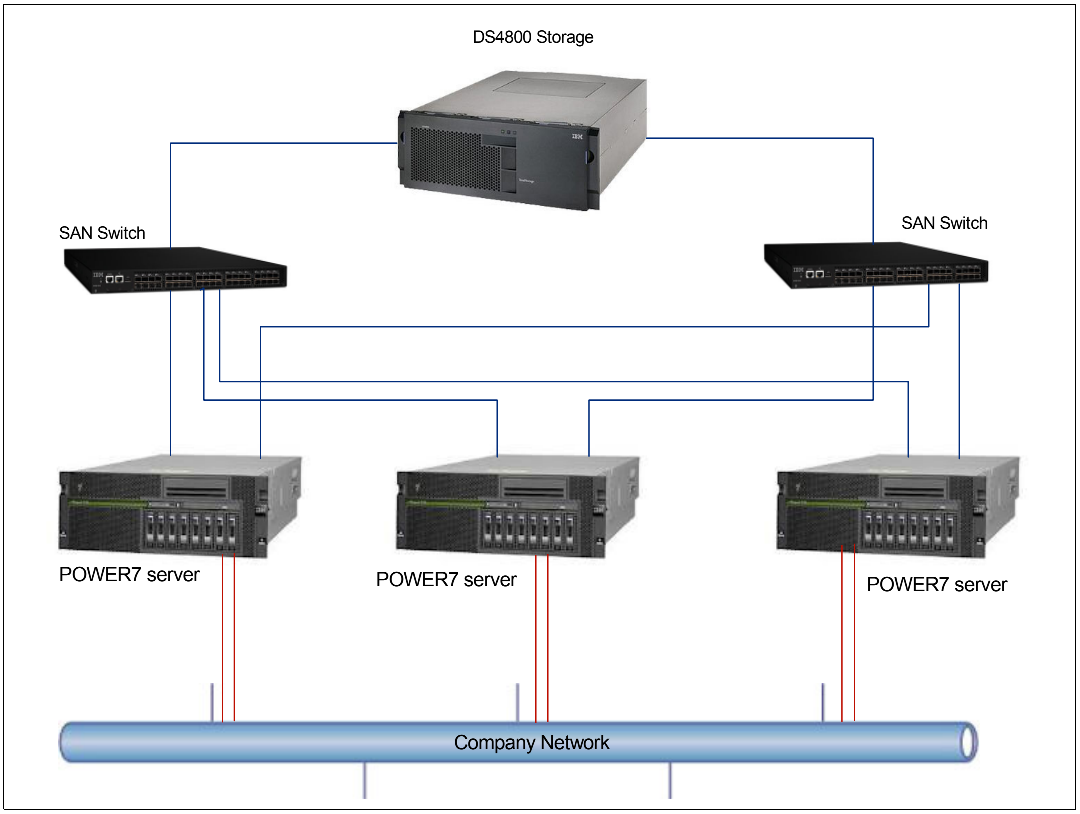


*Figure 1-19 ITSO residency hardware environment*

Figure 1-19 shows the hardware environment that is used in this book. It consists of an IBM TotalStorage DS4800 connected to a dual SAN environment, which is built with two IBM 2005-B16 switches. The servers are IBM POWER7® 750 servers. Figure 1-20 shows the Power 750 servers in the Hardware Management Console (HMC).

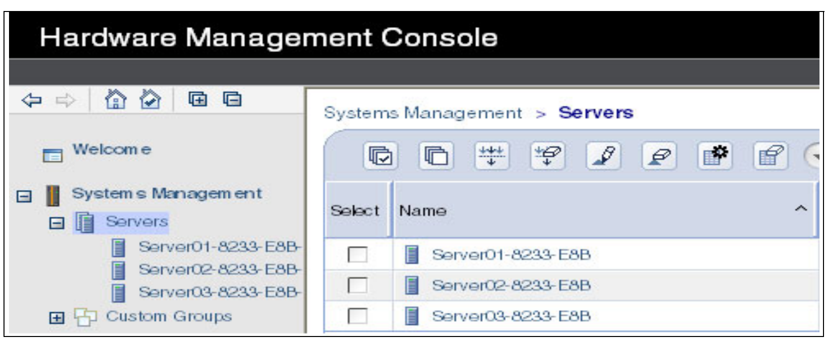


*Figure 1-20 POWER7 servers in the HMC*

## Virtualized environment

To mimic a real-world scenario, we built redundant Virtual I/O (VIO) servers on each frame with Shared Ethernet Adapter (SEA) failover and N-Port ID Virtualization (NPIV) for all client logical partitions (LPARs).

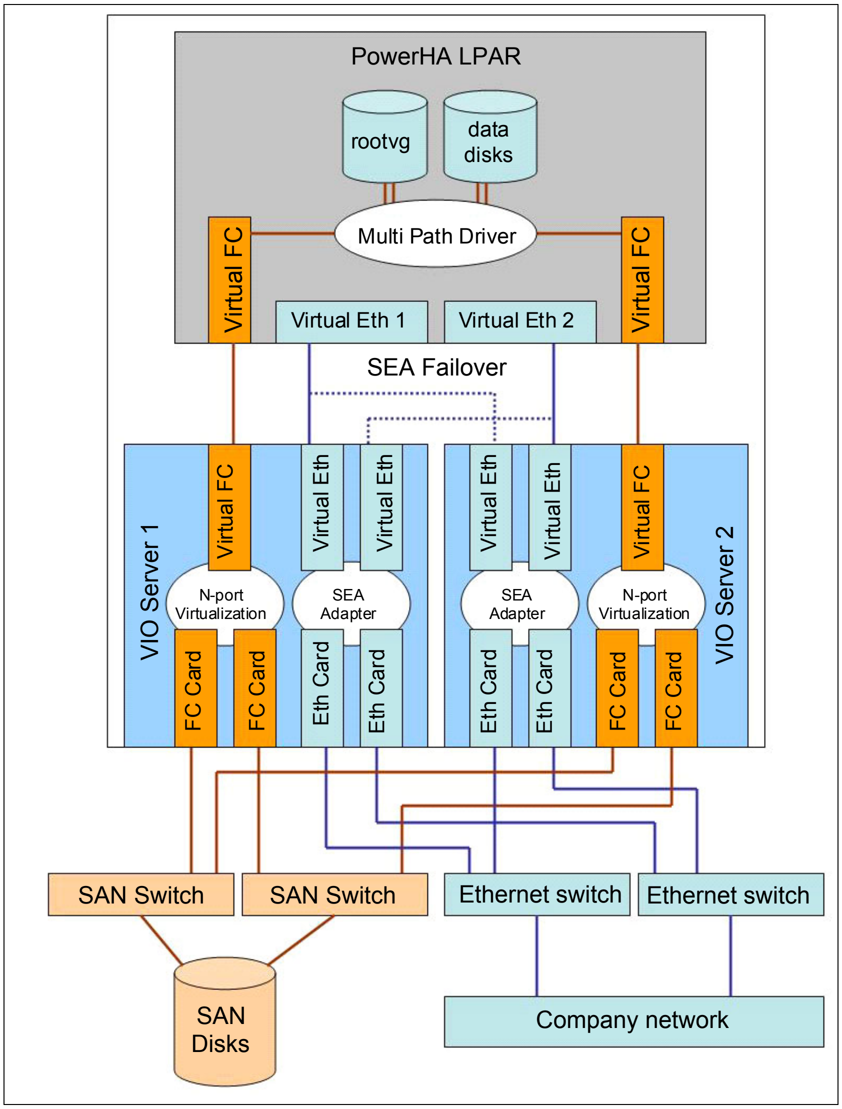



*Figure 1-21 Hardware lab virtual environment*

The virtualization structure in Figure 1-21 shows that all client LPARs have dual Ethernet connections through client SEAs and dual SAN connection through dual NPIV Fibre Channel (FC) adapters. This structure helps provide the correct hardware redundancy for all scenarios that are presented in this book.

## Client LPAR structure

Because we use three physical servers to build the clustered environments for this book, we partitioned the Power 750 frames to create 10 client LPARs per frame that follow the naming standard LPAR<*num*>.<*frame num*>, where <*num*> is the LPAR sequence number and <*frame num*> is the frame sequence number (Figure 1-22 on page 26).

| LPAR Name | Hostname | OS Level | OS Disk Type | OS Disk Amount | FC Type |
|---|---|---|---|---|---|
| VIOS01.01 | vios0101 | 2.2.1.3 | Internal SAS | 3x146GB | Physical HBA |
| VIOS02.01 | vios0201 | 2.2.1.3 | Internal SAS | 3x146GB | Physical HBA |
| LPAR01.01 | lpar0101 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR02.01 | lpar0201 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR03.01 | lpar0301 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR04.01 | lpar0401 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR05.01 | lpar0501 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR06.01 | lpar0601 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR07.01 | lpar0701 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR08.01 | lpar0801 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR09.01 | lpar0901 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |
| LPAR10.01 | lpar1001 | 7.1 TL01SP03 | Virtual SCSI | 2x30GB (one per VIO) | NPIV Adapter |

*Figure 1-22 LPAR partitioning on the hardware lab*

As shown in Figure 1-22, on frame Server01 (further details in Figure 1-20 on page 24), we created two redundant VIO servers and 10 client LPARs. The same design is adopted for Server02 and Server03 frames.

> **Resource**: For further information about IBM Power Systems PowerVM virtualization features and configuration, see *IBM PowerVM Virtualization Introduction and Configuration,* SG24-7940:
>
> `http://www.redbooks.ibm.com/abstracts/sg247940.html?Open`

## SAN storage environment

All scenarios that are shown in this book are implemented with SAN shared disks for data that is used by PowerHA resource groups.

All disks are created on an IBM TotalStorage DS4800 midrange disk system with 1.7 TB of raw data. All data on the system is formatted in RAID-5 and originates from a single large array called *PW2201_Array*. All logical drives that are used on the cluster scenarios are created inside this RAID-5 array with <*Cluster Scenario*>-<*Disk Sequence*> as the naming convention. For performance purposes, all logical volumes are created to keep the load balancing between DS4800 Controller A and Controller B.

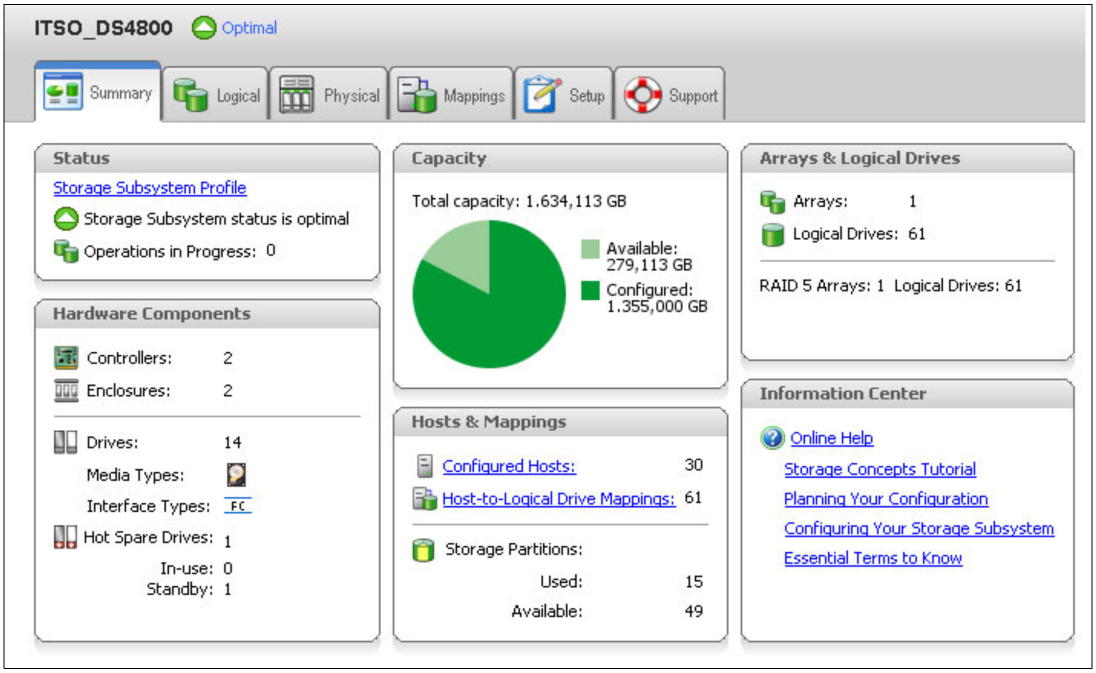



*Figure 1-23 Summary DS4800 disk system view from the IBM Storage Manager tool*

Figure 1-23 shows the general configuration of the DS4800 disk system that is used during the book scenarios. For each scenario, specific host groups are created to ensure that only the correct nodes can reach logical unit number (LUN) disk sets, as shown in Figure 1-24 on page 28.

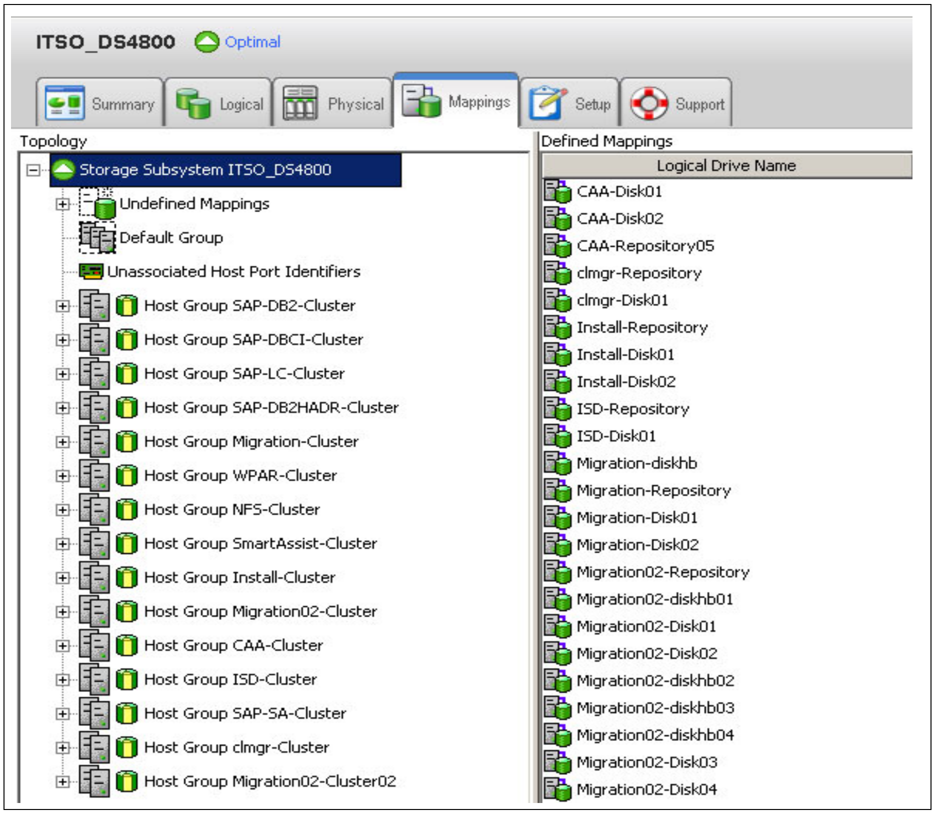


*Figure 1-24 Host groups that are defined for cluster scenarios*

## LPARs access to SAN disks

All hardware connections are configured for redundancy. All clustered LPARs have two NPIV adapters and two paths for disk access, as shown in Figure 1-25 and Figure 1-26 on page 29.

```
root@sapnfs1 / # mpio_get_config -Av
Frame id 0:
    Storage Subsystem worldwide name: 60ab800114632000048ed17e
    Controller count: 2
    Partition count: 1
    Partition 0:
    Storage Subsystem Name = 'ITSO_DS4800'
        hdisk#           LUN #   Ownership          User Label
        hdisk2               0   A (preferred)      NFS-Repository
        hdisk3               1   B (preferred)      NFS-Disk01
        hdisk4               2   A (preferred)      NFS-Disk02
        hdisk5               3   B (preferred)      NFS-Disk03
        hdisk6               4   A (preferred)      NFS-Disk04
```

*Figure 1-25 LUN disks that are assigned to NFS cluster nodes*

```
root@sapnfs1 / # lspath
Enabled hdisk0 vscsi0
Enabled hdisk1 vscsi1
Enabled hdisk2 fscsi0
Enabled hdisk3 fscsi0
Enabled hdisk4 fscsi0
Enabled hdisk2 fscsi1
Enabled hdisk3 fscsi1
Enabled hdisk4 fscsi1
Enabled hdisk5 fscsi0
Enabled hdisk6 fscsi0
Enabled hdisk5 fscsi1
Enabled hdisk6 fscsi1
```

*Figure 1-26 LPAR NPIV adapters that are configured on LPARs*

## 1.3 Cluster scenarios

We created multiple cluster scenarios to implement the many features of the IBM PowerHA 7.1.1. Table 1-2 shows the list of all cluster scenarios that are covered by this book.

*Table 1-2 Cluster scenarios that are covered in this book*

| Cluster name | Cluster nodes | LPAR names | Description |
|---|---|---|---|
| SAP DB2 | sapdb1<br>sapdb2 | LPAR01.01<br>LPAR01.03 | SAP DB2 database cluster details are at 7.4, “Scenario for a complete SAP and liveCache environment” on page 388. |
| SAP CI | sapci1<br>sapci2 | LPAR01.02<br>LPAR02.01 | SAP Central Instance cluster details are at 7.6, “Cluster 2: Database instance” on page 431. |
| SAP liveCache | saplc1<br>saplc2 | LPAR02.02<br>LPAR02.03 | SAP liveCache Hotstandby cluster details are at 7.7, “Cluster 3: liveCache” on page 440. |
| SAP DB2 HADR | sapdbhr3<br>sapdbhr4 | LPAR03.01<br>LPAR03.03 | SAP DB2 database with HADR cluster details are at 7.8, “DB2 HADR cluster solution” on page 463. |
| SAP SmartAssist cluster | sapsma1<br>sapsma2 | LPAR09.01<br>LPAR09.02 | Cluster that uses Smart Assist for SAP details are at 7.3.3, “Starting Smart Assist for SAP: Global file system (NFS)” on page 373. |
| Migration clusters | migr1<br>migr2<br>migr3<br>migr4<br>migr5<br>migr6 | LPAR03.02<br>LPAR04.01<br>LPAR07.01<br>LPAR07.02<br>LPAR10.02<br>LPAR10.03 | Multiple clusters are used to perform all possible PowerHA migrations from early versions to the 7.1.1 version. The details are at Chapter 6, “Migration scenarios” on page 255. |
| WPAR cluster | wpar1<br>wpar2 | LPAR04.02<br>LPAR04.03 | WPAR clusters use PowerHA 7.1.1. Details are at Chapter 8, “Workload partition and PowerHA scenario” on page 487. |

| Cluster name | Cluster nodes | LPAR names | Description |
|---|---|---|---|
| Installation cluster | inst1<br>inst2 | LPAR06.02<br>LPAR06.03 | PowerHA installation drills. The details are at Chapter 3, “PowerHA 7.1.1 basic installation and configuration” on page 73. |
| SmartAssist cluster | smas1<br>smas2 | LPAR05.03<br>LPAR06.01 | PowerHA scenarios that use the Smart Assist tool. The details are at Chapter 5, “Smart Assists for PowerHA SystemMirror” on page 227. |
| CAA cluster | caa1<br>caa2 | LPAR07.03<br>LPAR08.01 | PowerHA cluster that is dedicated to CAA-related tasks and tests. The details are at Chapter 2, “Cluster Aware AIX” on page 35. |
| ISD cluster | isd1<br>isd2 | LPAR08.02<br>LPAR08.03 | PowerHA cluster for IBM System Directory Server. The details are at 5.5, “Smart Assist for IBM Tivoli Directory Server” on page 246. |
| clmgr cluster | clmgr1<br>clmgr2 | LPAR09.03<br>LPAR10.01 | PowerHA cluster that is mounted by using only the **`clmgr`** command. The details are at Appendix B, “Configuring the PowerHA cluster by using clmgr” on page 551. |

## 1.4 New features on PowerHA 7.1.1

PowerHA 7.1.1 introduced the following features:

- Federated Security: Many security features are included on PowerHA 7.1.1 that allow a cluster-wide single point of control:
  - Encrypted File System (EFS) support
  - Role Based Access Control (RBAC) support
  - Authentication by using LDAP methods
- LVM and C-SPOC: PowerHA 7.1.1 enhanced the administrative activities for LVM and for C-SPOC with these new features:
  - EFS management by C-SPOC.
  - Support for *mirror pools*.
  - Disk renaming support inside the cluster.
  - PowerHA 7.1.1 now supports EMC, Hitachi, and HP disk subsystems multipathing LUN as a clustered repository disk.
  - PowerHA 7.1.1 capability to display disk Universally Unique Identifier (UUID).
  - PowerHA 7.1.1 filesystem mounting feature called *Mount Guard*. After one filesystem is mounted by using this feature, AIX ensures that no other node can improperly mount the same filesystem, at the same time, which can cause data corruption.
- Application management: A new application startup option is available on PowerHA 7.1.1. When you add an application controller, you can choose the Application Startup Mode. Now, you can choose background startup mode, which is the default and where the cluster activation moves forward with an application start script that runs in the background. Or,

you can choose foreground startup mode. When you choose the application controller option, the cluster activation is sequential, which means that cluster events hold application startup script execution. If the application script ends with a failure (non-zero return code), the cluster activation is considered to failed, as well.

### Network features

PowerHA 7.1.1 offers more network management features. It is now possible to define a network as *private* and more network tunables are provided.

> **Network features:** For more details about the new administrative and network management features, see Chapter 4, "Cluster administration" on page 149.

### Smart Assist features

Within PowerHA 7.1.1, new Smart Assists are included for the following applications:

- SAP liveCache Hot Standby for fast failover with IBM DS8000® and SAN Volume Controller
- MQ Series Smart Assist

> **Smart Assist:** For more details about Smart Assists usage and new features, see Chapter 5, "Smart Assists for PowerHA SystemMirror" on page 227 and 7.3.2, "Installation and configuration steps before you use Smart Assist for SAP" on page 366.

## 1.5 PowerHA 7.1.1 installation

When you install PowerHA 7.1.1, it is important to have a supported environment. This chapter provides a brief introduction and description for all topics that relate to PowerHA 7.1.1 product installation. For detailed information about this topic, see Chapter 3, "PowerHA 7.1.1 basic installation and configuration" on page 73.

### Supported hardware

PowerHA 7.1.1 supports POWER5, IBM POWER6®, POWER7, and Power Blades servers.

From disk subsystems, PowerHA allows the usage of multiple disk solutions as shared disks between cluster nodes. The following IBM TotalStorage systems are supported:

- DS3000 family
- IBM DS4000® family
- DS5000 family
- IBM DS6000™ family
- DS8000 family
- IBM XIV®
- SAN Volume Controller
- IBM Storwize® V7000

For all information that relates to version compatibility between AIX, PowerHA, and IBM disk subsystems, see this website:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105638

For the latest firmware upgrades for server and disk subsystems, see IBM Fix Central:

http://support.ibm.com/fixcentral

**Matrix:** For a complete list of supported hardware for PowerHA 7.1.1 and as the minimum operating system level that is required, see the PowerHA Hardware Support Matrix:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105638

### Supported AIX OS versions

The minimum AIX level that is required to install PowerHA 7.1.1 is AIX 6.1 TL7 SP2 or AIX 7.1 TL1 SP2. We suggest that you keep your environment on the newest Technology Level that is available for your AIX release before you start with PowerHA 7.1.1.

**Important:** There is no support for AIX 5.3. Any environment that still runs AIX 5.3 must be upgraded to AIX 6.1 TL7 SP02 at a minimum before you install or migrate to PowerHA 7.1.1.

The latest AIX updates are at IBM Fix Central:

http://support.ibm.com/fixcentral

### Disk space requirements

PowerHA 7.1.1 requires a specific amount of free filesystem space for installation. Check that the following requirements are satisfied before you start the PowerHA 7.1.1 filesets installation:

- The / filesystem must have at least 1 MB of free disk space.
- The `/usr` filesystem must have at least 82 MB of free disk space.

The cluster logs are stored inside the `/var/log` and `/var/hacmp` directories. It is important to keep adequate free space available in the `/var` filesystem, or it can lead to serious problems within the PowerHA cluster.

### PowerHA 7.1.1 filesets

The PowerHA 7.1.1 installation media contains several filesets. Figure 1-27 shows the names and descriptions of these filesets.

| Package Name | Package Description |
|---|---|
| cluster.adt.es | PowerHA SystemMirror Client Samples |
| cluster.doc.en_US.assist.db2.pdf | PowerHA SystemMirror Smart Assist documentation |
| cluster.doc.en_US.assist.oracle.pdf | PowerHA SystemMirror Smart Assist documentation |
| cluster.doc.en_US.assist.websphere.pdf | PowerHA SystemMirror Smart Assist documentation |
| cluster.doc.LOCALE.es | PowerHA SystemMirror general documentation |
| cluster.es.assist | PowerHA SystemMirror Smart Assist binaries |
| cluster.es.cfs | Cluster File System Support |
| cluster.es.client | PowerHA SystemMirror Client |
| cluster.es.cspoc | CSPOC |
| cluster.es.nfs | NFS Support |
| cluster.es.server | PowerHA SystemMirror Server Utilities |
| cluster.license | License |
| cluster.es.migcheck | Migration verification enablement |
| cluster.es.director.agent | PowerHA SystemMirror Director Agent binaries |
| cluster.man.LOCALE.es | Man pages |
| cluster.man.LOCALE.assist | Man pages for Smart Assists |
| cluster.msg.LOCALE.es | General messages |

*Figure 1-27 PowerHA 7.1.1 filesets*

### Obtaining PowerHA 7.1.1 updates

After you install PowerHA 7.1.1, a common practice is to apply the latest service packs (SPs) on the system. All updates and SPs that relate to any IBM product, including PowerHA 7.1.1, are in IBM Support Fix Central:

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

> **For more information:** For further details about planning and the design steps both before and during a PowerHA 7.1.1 installation, see Chapter 3, "PowerHA 7.1.1 basic installation and configuration" on page 73.

## 1.6 Migrating to PowerHA 7.1.1

If your current PowerHA environment runs at 5.4.1 or later, you can migrate your cluster to the current PowerHA Version 7.1.1. PowerHA supports migrations for earlier versions to current releases:

- Offline migration: PowerHA services are brought offline on all nodes before any upgrade task is performed. During the upgrade period, no resources are available to users.
- Snapshot migration: The following actions occur with this upgrade approach:
  - A cluster snapshot is taken.
  - PowerHA cluster services are brought offline on all nodes.
  - The current PowerHA release is uninstalled.
  - PowerHA 7.1.1 is installed.
  - The snapshot is converted by using the **`clconvert_snapshot`** utility.
  - The cluster configuration from the snapshot is restored.
- Rolling migration: During a rolling migration only one node is stopped and updated at a time. The next node is only upgraded when the previous one is successfully upgraded and rejoins the cluster.

### Migration requirements

It is important to know the supported PowerHA migration methods. The following guidelines apply to a PowerHA 7.1.1 migration:

- PowerHA 7.1.1 supports automatic migration from the following prior releases:
  - PowerHA 5.4.1
  - PowerHA 5.5
  - PowerHA 6.1
  - PowerHA 7.1.0
- To migrate from PowerHA 7.1.0 to PowerHA 7.1.1, you can use only offline migration or snapshot migration.
- Due to CAA characteristics, PowerHA 7.1.1 does not support a migration from an old PowerHA release that uses IPv6-based networks.
- After all migration steps are done, apply all latest updates for the PowerHA cluster product.

> **For more information:** All detailed steps for PowerHA cluster migration from earlier (supported) versions to PowerHA 7.1.1 are fully described in Chapter 6, "Migration scenarios" on page 255.

2

# Cluster Aware AIX

This chapter provides information about Cluster Aware AIX (CAA). This chapter includes new features for Release 2 (R2), and the differences from Release 1 (R1).

This chapter discusses the following topics:

- Introduction
- Changes and new features
- Security
- Deadman switch (DMS)
- The central repository disk
- Repository disk for third-party multipathing
- Repository disk resiliency
- CAA start and stop commands
- CAA tunables
- Troubleshooting CAA
- Statistics
- Gathering more cluster information
- CAA cluster commands and force option
- Restricting interfaces from clustering use
- VLAN tagging
- Round-trip time and heartbeats
- End-to-end monitoring
- CAA daemons
- CAA multicast

## 2.1 Introduction

Cluster Aware AIX (CAA) introduces fundamental clustering capabilities into the base AIX operating system. These capabilities include the creation and definition of the set of nodes that comprise the cluster. CAA provides the tools and monitoring capabilities for the detection of node and interface health.

> **Filesets:** CAA is provided by the AIX filesets `bos.cluster.rte`, `bos.ahafs`, and `bos.cluster.solid`. These filesets are included in AIX 6.1 TL6 and later, and in the AIX 7.1 installation media.
>
> For more information about CAA, see the information center:
>
> http://pic.dhe.ibm.com/infocenter/aix/v7r1/topic/com.ibm.aix.clusteraware/claware_main.htm

CAA provides a set of tools and APIs to enable clustering on the AIX operating system. CAA does not provide the application monitoring and resource failover capabilities that PowerHA provides. Other applications and software programs can use the APIs and command-line interfaces (CLIs) that CAA provides to make their applications and services “Cluster Aware”.

Figure 2-1 shows a high-level architectural view of how IBM high availability (HA) applications PowerHA and Virtual I/O Server (VIOS) clustered storage use Resource Monitoring and Control (RSCT) and the CAA architecture.

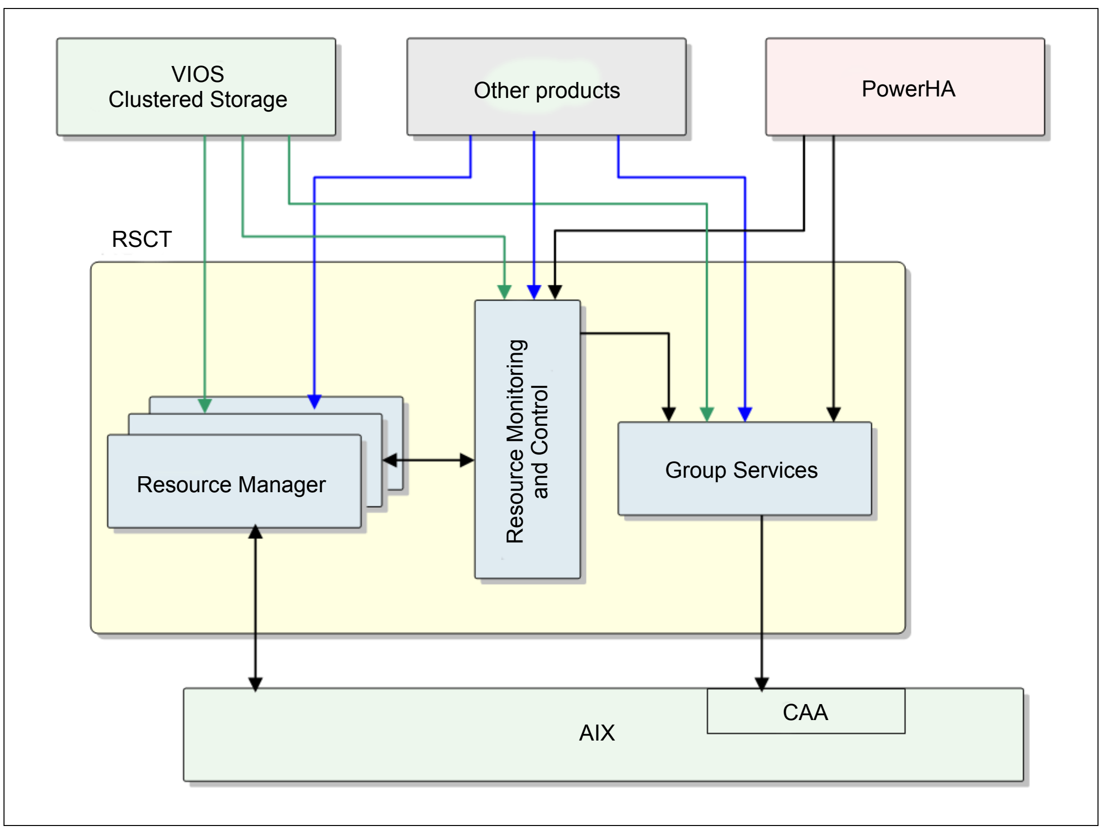



*Figure 2-1 HA applications that use RSCT and CAA*

The following products use the CAA technology:

- RSCT (3.1 and later)
- PowerHA (7.1 and later)
- VIOS 2.2.0.11-FP-24 SP-01 and later

CAA provides the following features among others:

- Central repository:
  - Configuration
  - Security
- Quorumless (CAA does not require a quorum to be up and operational)
- Monitoring capabilities for custom actions
- Fencing aids:
  - Network
  - Storage
  - Applications

The following sections explain some of the concepts of the CAA central repository, changes from previous versions, and how PowerHA 7.1.1 uses CAA.

> **Important:** If you use PowerHA, you do not need to manually configure the CAA cluster. Any configuration changes in the CAA cluster are managed by the PowerHA configuration tools.
>
> CAA does not support IPv6-based network environments.

## 2.2 Changes and new features

We describe the changes and new features in the 2011 release of CAA, and its impact on PowerHA 7.1.1 functionality:

- Changes
- Security
- Deadman switch
- The central repository disk
- CAA and IBM solidDB®
- Repository disk for third-party multipath disks
- Repository disk resilience
- Synchronous configuration changes
- Start and stop commands
- CAA tunables
- Troubleshooting CAA - snap & logs
- Statistics
- How to gather more cluster information
- CAA cluster commands and force option
- The use of the ifrestrict file
- Added support for VLAN tagging
- Round-trip time and heartbeats

## 2.2.1 Changes

The following list summarizes the changes and updates in the 2011 release of CAA:

- Upgrading AIX in a CAA cluster to AIX 6.1 TL07 or to AIX 7.1 TL01 is not supported.

  To upgrade from AIX 6.1 TL6 to TL7 or from AIX 7 to TL1, first remove the CAA cluster, and then install AIX 6 with 6100-07 or install AIX 7 with 7100-01 on all nodes that are part of the new cluster. See Appendix A, “AIX upgrades and PowerHA SystemMirror 7.1.0” on page 531.
- CAA no longer uses embedded IBM solidDB database. The solid and solidhac daemons are no longer used by CAA.
- The CAA infrastructure now provides limited support for third-party multipathing software. No disk events are available for these disks, but they can be configured into a cluster as a repository or as shared disks.
- CAA commands no longer support force cleanup options. The following list shows the options, by command, that are not supported in the 2011 release:
  - **`chcluster -f`**
  - **`clusterconf -f`**, **`-s`**, **`-u`**
  - **`rmcluster -f`**

  The `clctrl` command can be used for tuning the cluster subsystem.

> **Important:** Only tune the cluster subsystem under the direction of IBM customer support.

## 2.2.2 Security

Secure the core communication between nodes in the cluster by using one of the following encryption mechanisms.

The following encryption keys are supported:

- Message Digest 5 (MD5) with Data Encryption Standard (DES)
- MD5 with Triple DES
- MD5 with Advanced Encryption Standard (AES)

The following certificates are supported:

- Open Secure Sockets Layer (SSL)
- Self-signed certificates
- Secure Shell (SSH) certificates

You can configure the security options and options to distribute encryption keys by using the SMIT interface or the `clctrl` command:

- **`smit clustsec`**
- **`clctrl -sec`**

## 2.2.3 Deadman switch (DMS)

A *deadman switch* is an action that occurs when CAA detects that a node is isolated in a multinode environment. No network or disk communication occurs between nodes.

While it might not seem obvious, the reason for implementing the DMS is to protect the data on the external disks.

The AIX operating system reacts differently depending on the DMS (deadman_mode) tunable. The deadman switch mode can be set to either force a system crash or generate an Autonomic Health Advisor File System (AHAFS) event. See Example 2-1 for the DMS tunable options help and Example 2-2 for a listing of the current DMS tunable status.

The CAA DMS tunable (deadman_mode) allows two actions:

- Assert (crash) the system
- Generate AHAFS event

*Example 2-1 DMS tunable options*

```
root@caa2 / # clctrl -tune -h deadman_mode
Help for tunable deadman_mode:
Purpose:
Controls the behavior of the deadman timer.  The interval associated with this
timer is the node_timeout tunable.
Scope: clusterwide, per node
Values:
        Default: a
        Range: a, e
        Unit:
Tuning:
When the value is set to "a" (assert), the node will crash upon the deadman
timer popping.

When the value is set to "e" (event), an AHAFS event is generated.

A node-specific setting trumps the cluster-wide setting and can be used to
override the behavior on a per-node basis.
```

*Example 2-2 Listing DMS tunable status*

```
root@caa1 / # clctrl -tune -L deadman_mode
NAME                      DEF    MIN    MAX    UNIT           SCOPE
     ENTITY_NAME(UUID)                                                CUR
--------------------------------------------------------------------------------
deadman_mode              a                                   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     a
--------------------------------------------------------------------------------
```

By default, the CAA deadman_mode option is “a”. If the deadman timeout is reached, the node crashes immediately to prevent a partitioned cluster and data corruption.

Next, we present a scenario of a node isolation in a running mutual takeover PowerHA cluster, a consequent deadman switch is triggered, and a node crashes. Look for the behavior of PowerHA and CAA clusters on the surviving node.

Figure 2-2 on page 40 shows the initial status of the PowerHA cluster.

```
              clstat - PowerHA SystemMirror Cluster Status Monitor
                -------------------------------------

Cluster: caa_cl (1089725824)
Thu Apr 26 14:43:00 EDT 2012
                State: UP               Nodes: 2
                SubState: STABLE

        Node: caa1              State: UP
           Interface: caa1 (0)                  Address: 10.1.1.48
                                                State:   UP
           Interface: caa1b2 (1)                Address: 10.1.2.48
                                                State:   UP
           Interface: caasvc1 (0)               Address: 172.16.21.69
                                                State:   UP
           Resource Group: db_RG                        State:  On line

        Node: caa2              State: UP
           Interface: caa2 (0)                  Address: 10.1.1.33
                                                State:   UP
           Interface: caa2b2 (1)                Address: 10.1.2.33
                                                State:   UP
           Interface: caasvc2 (1)               Address: 10.10.20.69
                                                State:   UP
           Resource Group: internet_RG                  State:  On line
```

*Figure 2-2 Status of the cluster before the deadman switch*

We unmap the repository disk from the storage for both nodes to simulate the loss of the disk. Then, we detach the Ethernet interfaces on the caa1 node, type the next commands (Example 2-3) to simulate the network failure, and stop any communication between the nodes.

*Example 2-3 Ifconfig commands to disconnect the network*

```
root@caa1 / #ifconfig en2 down detach
root@caa1 / #ifconfig en1 down detach
root@caa1 / #ifconfig en0 down detach
root@caa1 / #
```

After some seconds, the caa1 node crashes (Example 2-5 on page 41) with the `KERNEL_PANIC` and the `UNEXPECTED SYSTEM HALT` messages in the AIX error log.

Example 2-4 shows the system configuration attribute that causes the reboot after the DMS crashes the system.

*Example 2-4 Autorestart system configuration attribute*

```
root@caa1 / # lsattr -El sys0|grep restart
autorestart     true Automatically REBOOT OS after a crash           True
root@caa1 / #
```

Example 2-5 on page 41 shows the message in the AIX error log when the DMS kernel extension is triggered.

*Example 2-5 AIX error log: DMS notification*

```
---------------------------------------------------------------------------
LABEL:          KERNEL_PANIC
IDENTIFIER:     225E3B63

Date/Time:       Tue Apr 10 23:34:36 EDT 2012
Sequence Number: 1017243
Machine Id:      00F61AB24C00
Node Id:         caa2
Class:           S
Type:            TEMP
WPAR:            Global
Resource Name:   PANIC

Description
SOFTWARE PROGRAM ABNORMALLY TERMINATED

        Recommended Actions
        PERFORM PROBLEM DETERMINATION PROCEDURES

Detail Data
ASSERT STRING

PANIC STRING
Deadman timer triggered.

---------------------------------------------------------------------------
```

Figure 2-3 on page 42 shows the status of the PowerHA cluster after the DMS is triggered. You can see the caa1 node in the `DOWN` state, and the caa2 node runs both resource groups.

```
                  clstat - PowerHA SystemMirror Cluster Status Monitor
                -------------------------------------

Cluster: caa_cl (1089725824)
Thu Apr 26 15:08:38 EDT 2012
                State: UP              Nodes: 2
                SubState: STABLE

        Node: caa1             State: DOWN
           Interface: caa1 (0)                Address: 10.1.1.48
                                              State:   DOWN
           Interface: caa1b2 (1)              Address: 10.1.2.48
                                              State:   DOWN

        Node: caa2             State: UP
           Interface: caa2 (0)                Address: 10.1.1.33
                                              State:   UP
           Interface: caa2b2 (1)              Address: 10.1.2.33
                                              State:   UP
           Interface: caasvc2 (1)             Address: 10.10.20.69
                                              State:   UP
           Interface: caasvc1 (0)             Address: 172.16.21.69
                                              State:   UP
           Resource Group: db_RG                      State:  On line
           Resource Group: internet_RG                State:  On line
```

*Figure 2-3 Status of the PowerHA cluster after the DMS is triggered*

### 2.2.4 The central repository disk

PowerHA SystemMirror uses a shared disk across the cluster as a central repository for managing the configuration of the cluster. This disk must be accessible by all of the nodes in the cluster.

With CAA, you can use the cluster repository disk for the following purposes:

- Cluster-wide configuration management
- Cluster messaging and heartbeating

The amount of configuration information that is stored on this repository disk is directly dependent on the number of cluster entities, such as shared disks, number of nodes, and number of adapters in the environment. You must ensure that you have enough space for the following components when you determine the size of a repository disk:

- Node-to-node communication
- Cluster topology management
- All migration processes

You must have at least 512 MB and no more than 460 GB of disk space that is allocated for the repository disk. Normally, 1 GB is enough for any configuration.

The repository disk is partitioned by the cluster to maintain redundant copies of the information, and to provide high availability for the configuration data and cluster communication. Use RAID technology to protect this disk from any disk failure. Use the PowerHA SystemMirror snapshot function to back up or restore any cluster configurations.

**Important:** The loss of the cluster repository disk does not affect the normal PowerHA cluster operation. This information is true only for PowerHA 7.1.1; it is not true for PowerHA 7.1.0). However, changes to the PowerHA cluster configuration are not possible if the repository disk fails. See 2.3, “End-to-end monitoring” on page 66.

The AIX operating system manages repository disks in a special way during the configuration of a cluster. AIX cluster services recognize the repository disk during the AIX device configuration process and can perform special actions on the disk. Cluster services also generate detailed events in relationship to the health of the repository disk. For example, the repository disk is used by the cluster services to enable disk-based communication across hosts. Disk communication involves one host that writes to the disk and the other host or hosts retrieving the message through a polling mechanism.

**Important:** Users must not work with the repository disk and the Volume Group (VG) that is created on it by using **`varyonvg`**, **`varyoffvg`**, or **`mklv`**. This disk and VG are for the exclusive use of CAA.

### 2.2.5 CAA and the IBM solidDB

CAA support is either removed or deprecated for these previously available features:

- IBM solidDB is no longer used by CAA.
- The logical volumes fslv00 and fslv01 with file system mount points of `/clrepos_private1` and `/clrepos_private2` no longer exist.
- CAA no longer attempts to create a consistent repository disk device view across the cluster.

With R2 of CAA, solidDB is no longer used. The solid and solidhac subsystems are present but dormant on the system. The fileset `bos.cluster.solid` is installed but not used.

Example 2-6 shows the Logical Volume Manager (LVM) and file system structure of a CAA R1 repository disk.

*Example 2-6 Content of previous caavg_private volume group*

```
# lsvg -l caavg_private
caavg_private:
LV NAME TYPE LPs PPs PVs LV STATE MOUNT POINT
caalv_private1 boot 1 1 1 closed/syncd N/A
caalv_private2 boot 1 1 1 closed/syncd N/A
caalv_private3 boot 4 4 1 open/syncd N/A
fslv00 jfs2 4 4 1 closed/syncd /clrepos_private1
fslv01 jfs2 4 4 1 open/syncd /clrepos_private2
powerha_crlv boot 1 1 1 closed/syncd N/A
```

As shown in Example 2-7, the CAA R2 does not create file systems `/clrepos_private1`, `/clrepos_private2`, and their associated logical volumes, fslv00 and fslv01. With this change, CAA makes raw I/O requests to the repository disk.

*Example 2-7 Content of the caavg_private volume group*

```
# lsvg -l caavg_private
caavg_private:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
```

```
caalv_private1      boot       1       1       1    closed/syncd  N/A
caalv_private2      boot       1       1       1    closed/syncd  N/A
caalv_private3      boot       4       4       1    open/syncd    N/A
powerha_crlv        boot       1       1       1    closed/syncd  N/A
```

## 2.2.6 Repository disk for third-party multipathing

The AIX operating system can configure the following types of disks as repository disks:

- AIX multipath disks (AIX MPIO)

  These disks can be automatically created and used as repository disks.

- Third-party multipath disks

  These disks follow the guidelines of AIX multipathing concepts, but provide their own multipathing device drivers and software. These disks can be configured as repository disks when the relevant Object Data Manager (ODM) information is available. Disks that are managed by EMC PowerPath and Hitachi Dynamic Link Manager (HDLM) can be configured as repository disks by using this method.

## 2.2.7 Repository disk resiliency

Repository resiliency is a new feature that allows the replacement of a failed repository disk on a running PowerHA SystemMirror cluster with no downtime. CAA repopulates the new disk with cluster-related information and starts to use that disk as the repository disk.

The repository disk resiliency function requires these software levels at a minimum:

- PowerHA 7.1.1 SP1
- AIX 6.1 TL7 SP3
- AIX 7.1 TL1 SP3

CAA can handle these failures without affecting critical cluster functions. The PowerHA cluster operates in a restricted mode until you replace the failed repository disk. In this mode of operation, most topology-related operations are not allowed. For example, a node cannot be added or a node cannot join the PowerHA cluster. However, critical PowerHA cluster functions, such as moving a resource group from an active node to a standby node, are handled by the cluster.

**Important:** You must configure a repository disk through multipathing and RAID configurations if you want a highly available environment.

### Viewing the status of the repository disk

CAA runs verification checks to detect when a repository disk failure occurs and generates a notification message in the cluster event log file (`hacmp.out`). The notification messages continue until you replace the failed repository disk with a new repository disk. These verification checks are performed when the CAA configuration daemon runs, every 10 minutes, and also every time that the disk is required for a particular function (cluster creation, cluster deletion, or cluster modification). These checks include disk existence and repository signature validation.

Parallel to CAA, the storage framework also checks for repository disk existence and notifies CAA when it loses connectivity to the repository disk. When this situation occurs, CAA looks to the repos_mode tunable for the user-desired assertion: currently either logging the event through the AIX error log facilities or asserting (deliberately crashing) the system. See 2.2.10, “CAA tunables” on page 54.

### Status of the cluster before the repository disk fails

Figure 2-4 shows the hdisk2 that is configured as the repository disk with the caavg_private VG in the active state.

```
root@caa1 / # lspv
hdisk0          00f74d47f963be5f                    rootvg          active
hdisk1          00f74d47fa93539e                    rootvg          active
hdisk2          00f61ab23678ae74                    caavg_private   active
hdisk3          00f74d472b9f9c73                    dbvg            concurrent
```

*Figure 2-4 Status of the repository disk*

Figure 2-5 shows the output for the storage interface. You can see that there is one disk that is configured (hdisk2) with the UP state and the REPDISK type.

```
root@caa1 / # lscluster -d
Storage Interface Query

Cluster Name:  caa_cl
Cluster uuid:  0f945780-857b-11e1-8cf4-b6fcc11f846f
Number of nodes reporting = 2
Number of nodes expected = 2
Node caa1
Node uuid = 0f5bfa5c-857b-11e1-8cf4-b6fcc11f846f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 92c9ec23-176e-7097-59ab-0b2323ec640e
          type  : REPDISK
Node caa2
Node uuid = 0f9001b2-857b-11e1-8cf4-b6fcc11f846f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 92c9ec23-176e-7097-59ab-0b2323ec640e
          type  : REPDISK
```

*Figure 2-5 Output for the lscluster -d command*

In Figure 2-6 on page 46, you can see the Disk Ping interface, dpcom, which is also often referred to as “*pingcomm*” with the state UP for both cluster nodes.

```
root@caa1 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  caa_cl
Cluster uuid:  247a062c-7383-11e1-802c-b6fcc11f846f
Number of nodes reporting = 2
Number of nodes expected = 2
Node caa1
Node uuid = 242fabd6-7383-11e1-802c-b6fcc11f846f
Number of interfaces discovered = 3
...
...
Interface number 3 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN

Node caa2
Node uuid = 2475cdf0-7383-11e1-802c-b6fcc11f846f
Number of interfaces discovered = 3
...
...
Interface number 3 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN
```

*Figure 2-6 Status of the dpcom cluster interfaces*

This entity is used by CAA to communicate over the repository disk. This entity is treated as a secondary means of communication if the LAN and SAN technologies go offline (Example 2-8).

*Example 2-8 Status of the configuration of the cluster nodes*

```
root@caa1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2
```

```
        Node name: caa1
        Cluster shorthand id for node: 1
        uuid for node: 0f5bfa5c-857b-11e1-8cf4-b6fcc11f846f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        caa_cl             local        0f945780-857b-11e1-8cf4-b6fcc11f846f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: caa2
        Cluster shorthand id for node: 2
        uuid for node: 0f9001b2-857b-11e1-8cf4-b6fcc11f846f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        caa_cl             local        0f945780-857b-11e1-8cf4-b6fcc11f846f

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
         dpcom  UP  RESTRICTED
         en0  UP
         en1  UP
```

### Failure of the repository disk

For this test, to simulate the repository disk failure, we unmap the logical unit number (LUN) from the storage for both nodes.

Almost immediately after we simulate the failure, we start to see error messages in the PowerHA event log file and the AIX error log, as usual, as shown in Figure 2-7 on page 48.

```
ERROR: rep_disk_notify : Wed Mar 21 17:37:51 EDT 2012 : Node
"caa2"(0x2475CDF0738311E1802CB6FCC11F846F) on Cluster caa_cl has lost access to
repository disk hdisk2.
 Please recover from this error or replace the repository disk using smitty.
child process for 1.
ERROR: rep_disk_notify : Wed Mar 21 17:37:52 EDT 2012 : Node
"caa1"(0x242FABD6738311E1802CB6FCC11F846F) on Cluster caa_cl has lost access to
repository disk hdisk2.
 Please recover from this error or replace the repository disk using smitty.
ERROR: rep_disk_notify : Wed Mar 21 17:38:21 EDT 2012 : Node
"caa2"(0x2475CDF0738311E1802CB6FCC11F846F) on Cluster caa_cl has lost access to
repository disk hdisk2.
 Please recover from this error or replace the repository disk using smitty.
ERROR: rep_disk_notify : Wed Mar 21 17:38:22 EDT 2012 : Node
"caa1"(0x242FABD6738311E1802CB6FCC11F846F) on Cluster caa_cl has lost access to
repository disk hdisk2.
 Please recover from this error or replace the repository disk using smitty.
```

*Figure 2-7 Repository disk error messages in hacmp.out log file*

These AIX error log messages are logged after the failure of the repository disk as shown in Figure 2-8.

```
B6267342   0321174512 P H hdisk2         DISK OPERATION ERROR
E86653C3   0321174512 P H LVDD           I/O ERROR DETECTED BY LVM
B6267342   0321174512 P H hdisk2         DISK OPERATION ERROR
E86653C3   0321174512 P H LVDD           I/O ERROR DETECTED BY LVM
B6267342   0321174512 P H hdisk2         DISK OPERATION ERROR
E86653C3   0321174512 P H LVDD           I/O ERROR DETECTED BY LVM
B6267342   0321174512 P H hdisk2         DISK OPERATION ERROR
E86653C3   0321174512 P H LVDD           I/O ERROR DETECTED BY LVM
B6267342   0321174512 P H hdisk2         DISK OPERATION ERROR
E86653C3   0321174512 P H LVDD           I/O ERROR DETECTED BY LVM
```

*Figure 2-8 Repository disk error messages in AIX error log*

The CAA cluster status after the failure of the repository disk is shown in Figure 2-9 on page 49.

```
root@caa2 / # lscluster -d
Storage Interface Query

Cluster Name:  caa_cl
Cluster uuid:  4eb6abec-9094-11e1-b181-b6fcc11f846f
Number of nodes reporting = 1
Number of nodes expected = 1
Node caa2
Node uuid = 4eb2255e-9094-11e1-b181-b6fcc11f846f
Number of disk discovered = 1
        hdisk2
          state : DOWN
          uDid  :
          uUid  : 0976f93c-64b1-7ed7-e006-e6976d893940
          type  : REPDISK
```

*Figure 2-9 Status of the CAA repository disk after the disk fail*

To confirm that access to the repository disk is lost, use the **lquerypv** command as shown in Example 2-9. If there is no output, the disk is unavailable. Check both nodes to be sure.

*Example 2-9 lquerypv command*

```
root@caa1 / # lquerypv -h /dev/hdisk2
root@caa1 / #
```

## Testing critical cluster functions

For this test, we start with the cluster active and running in a mutual takeover configuration as shown in Figure 2-10. We simulate the failure of the repository disk by unmapping the disk from the Storage Manager.

Next, we show the output for some commands, such as **clRGinfo** and **lscluster**, so that we can check the status of the cluster and to show that the cluster still provides critical functions (Figure 2-10):

- Move a resource group (RG) to the other node
- Bring an RG offline
- Bring an RG online on the home node again

```
root@caa1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                          Node
-----------------------------------------------------------------------------
db_RG          ONLINE                         caa1
               OFFLINE                        caa2

internet_RG    ONLINE                         caa2
               OFFLINE                        caa1
```

*Figure 2-10 Status of the resource groups*

### *Moving a resource group to the other node*

In this first test, we move the `db_RG` from the primary node to the standby node as shown in Figure 2-11 on page 50.

```
                                COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Attempting to move resource group db_RG to node caa2.

Waiting for the cluster to process the resource group movement request....

Waiting for the cluster to stabilize.........

Resource group movement successful.
Resource group db_RG is online on node caa2.
Cluster Name: caa_cl

Resource Group Name: db_RG
Node                         State
---------------------------- ---------------
caa1                         OFFLINE
caa2                         ONLINE

Resource Group Name: internet_RG
Node                         State
---------------------------- ---------------
caa2                         ONLINE
caa1                         OFFLINE
```

*Figure 2-11 SMIT window shows the RG movement*

### *Bringing a resource group offline*

We bring the `db_RG` resource group offline in the standby node as shown in Figure 2-12 on page 51.

```
                              COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Attempting to bring group db_RG offline on node caa2.

Waiting for the cluster to process the resource group movement request....

Waiting for the cluster to stabilize.......

Resource group movement successful.
Resource group db_RG is offline on node caa2.



Cluster Name: caa_cl

Resource Group Name: db_RG
Primary instance(s):
The instance is temporarily offline upon user requested rg_move performed on
Wed Mar 21 17:42:13 2012

Node                         State
---------------------------- ---------------
caa1                         OFFLINE
caa2                         OFFLINE

Resource Group Name: internet_RG
Node                         State
---------------------------- ---------------
caa2                         ONLINE
caa1                         OFFLINE
```

*Figure 2-12 Bring a resource group offline*

### *Bringing a resource group online on the home node*

We bring the `db_RG` resource group online again on the home node as shown in Figure 2-13 on page 52.

```
                              COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Attempting to bring group db_RG online on node caa1.

Waiting for the cluster to process the resource group movement request....

Waiting for the cluster to stabilize......

Resource group movement successful.
Resource group db_RG is online on node caa1.



Cluster Name: caa_cl

Resource Group Name: db_RG
Node                         State
---------------------------- ---------------
caa1                         ONLINE
caa2                         OFFLINE

Resource Group Name: internet_RG
Node                         State
---------------------------- ---------------
caa2                         ONLINE
caa1                         OFFLINE
```

*Figure 2-13 Bring a resource group online*

## Replacing the failed repository disk in a running cluster

In this section, details are provided about how to replace a failed repository disk in a running cluster.

### *Replacing a repository disk with SMIT*

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** → **Select a new Cluster repository disk**, and press Enter.
3. In the repository disk field, press F4 (Lists of available new repository disks). SMIT shows only shared disks whose size is 512 MB - 460 GB. Select an available disk from all nodes in the cluster, or enter the name of the disk. Press Enter to set up the selected disk as the new repository disk for the cluster.
4. Verify and synchronize the cluster configuration.
5. After all nodes are updated, verify that the new repository disk works by running the **`/usr/sbin/lscluster -d`** command.

### *Replacing a repository disk from the command line*

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

1. From the command line, use the **clmgr** command to replace the disk:

   ```
   clmgr modify cluster caa_cl REPOSITORY=hdiskX
   ```

2. Verify and synchronize the PowerHA configuration:

   ```
   clmgr verify cluster CHANGES_ONLY=yes SYNC=yes
   ```

3. After all nodes are updated, verify that the new repository disk works by running the **/usr/sbin/lscluster -d** command.

## 2.2.8 CAA configuration changes are synchronous

All cluster configuration events occur through a new process that is called "*coordinated configuration*". This process is a synchronous sequence of steps in which all nodes of the cluster must agree before committing the change. The change is committed to all nodes simultaneously. Only one coordinated configuration operation is allowed at a time. The coordinated configuration operation stops on the first failure. This new feature is introduced to prevent situations where the cluster state and status differ between the nodes.

The following configuration tasks are coordinated across the cluster:

- Cluster creation
- Add or remove nodes or disks
- Cluster removal
- Stop or start the node
- Node reboot

## 2.2.9 CAA start and stop commands

The **clctrl** command provides a set of subcommands for managing a CAA cluster. The start and stop subcommands are new in R2 of CAA:

- The **-stop** subcommand is used to take one or more nodes offline for maintenance. Stopping a node causes the other nodes to consider it as down.
- A stopped node does not send or receive heartbeat messages, and it remains in the stopped state, even across a reboot operation, until a **-start** subcommand causes it to rejoin the cluster.
- The **-stop** subcommand can also be issued while a node is powered off to prevent it from rejoining the cluster when it is rebooted.
- The **-start** subcommand is used to bring one or more nodes online after they are offline for maintenance. Starting a node allows it to rejoin the cluster and have the other nodes consider it as up.
- The **-start** subcommand can also be issued while a node is powered off to allow it to rejoin the cluster when it is rebooted.

To stop a node, you can issue the command as shown in Example 2-10 on page 54 from any active node in the cluster. The stop is persistent across node reboots until you issue the start command.

*Example 2-10 Stop a cluster node*

```
clctrl -stop -m <node>
```

The **stop** command is useful, for example, if you need to install device drivers on the node or maintain the node.

To start a node, issue the command as shown in Example 2-11.

*Example 2-11 Start a cluster node*

```
clctrl - start -m <node>
```

## 2.2.10 CAA tunables

With the **clctrl** command and by using the **tune** subcommand, you can list or change some configuration options across cluster nodes. The **-tune** subcommand is used to display or set cluster tunable values. Table 2-1 shows the flags that control the **-tune** subcommand.

*Table 2-1 Flags that control the -tune subcommand*

| Option | Description |
|---|---|
| -a | Displays values for all tunables, one per line. |
| -D | Resets all tunables to their default values. |
| -d tunable | Resets tunable to its default value. |
| -F | Forces the display of restricted tunables when option (-a, -L, or -x) is specified alone on the command line. |
| -h | Displays help about the command and its arguments. |
| -h tunable | Displays help about a tunable. |
| -L tunable | Lists information about one or all tunables in a table format. |
| -o tunable | Displays the current value of a tunable. |
| -o tunable=value | Sets tunable to the value. |
| -p | Makes the changes (-D, -d, or -o) or displays (-a or -o) apply to a permanent (current and nextboot) value. |
| -r | Makes the changes (-D, -d, or -o) or displays (-a or -o) apply to the next boot value. |
| -x tunable | Lists information about one or all tunables in a comma-separated format. |
| -y | Suppresses a confirmation prompt before executing the **bosboot** command. |

### Changing CAA tunables

To set a tunable value, you can use the following command syntax:

```
clctrl -tune -o <tunable>=<value>
```

This command modifies the tunable across the cluster in the repository disk. The new value becomes active at the next start of the node.

Example 2-12 shows the help for the repository disk repos_mode tunable. This tunable defines the action of a node that is part of a CAA cluster if the access to the repository disk is lost for the node.

*Example 2-12 clctrl repos_mode tunable help*

```
root@caa2 / # clctrl -tune -h repos_mode
Help for tunable repos_mode:
Purpose:
Controls node behavior when cluster repository access is lost.
Scope: clusterwide, per node
Values:
        Default: e
        Range: a, e
        Unit:
Tuning:
When the value is set to "a" (assert), the node will crash upon losing access
to the cluster repository.

When the value is set to "e" (event), an AHAFS event is generated.

A node-specific setting trumps the cluster-wide setting and can be used to
override the behavior on a per-node basis.
```

To see the actual configuration of the repos_mode tunable, use the command in Example 2-13. In this case, the tunable is configured with the default value of “e”. Therefore, when a node loses access to the repository disk only, an AHAFS event is generated.

*Example 2-13 Listing the status of the repos_mode tunable*

```
root@caa2 / # clctrl -tune -L repos_mode
NAME                      DEF    MIN    MAX    UNIT           SCOPE
     ENTITY_NAME(UUID)                                                 CUR
--------------------------------------------------------------------------------
repos_mode                e                                   c n
     caa_cl(a7db7a9e-6fed-11e1-a9b9-6e8dd007696f)                      e
--------------------------------------------------------------------------------
```

We can use the command in Example 2-14 to change this behavior and change the value from “e” (event) to “a” (assert). After changing this tunable, reboot the node to apply the changes. If the repository disk is lost, the node crashes instead of generating an AHAFS event.

*Example 2-14 repos_mode tunable change*

```
root@caa2 / # clctrl -tune -o repos_mode=a
1 tunable updated on cluster caa_cl.
```

## Listing all CAA tunables

Figure 2-14 on page 56 displays a detailed listing of tunables.

```
root@caa1 / # clctrl -tune -L
NAME                      DEF    MIN    MAX    UNIT           SCOPE
     ENTITY_NAME(UUID)                                                CUR
--------------------------------------------------------------------------------
config_timeout            240    0      2G-1   seconds        c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     240
--------------------------------------------------------------------------------
deadman_mode              a                                   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     a
--------------------------------------------------------------------------------
hb_src_disk               1      -1     3                     c
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     1
--------------------------------------------------------------------------------
hb_src_lan                1      -1     3                     c
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     1
--------------------------------------------------------------------------------
hb_src_san                2      -1     3                     c
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     2
--------------------------------------------------------------------------------
link_timeout              0      0      0      milliseconds   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     0
--------------------------------------------------------------------------------
node_down_delay           10000  5000   600000 milliseconds   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     10000
--------------------------------------------------------------------------------
node_timeout              20000  10000  600000 milliseconds   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     20000
--------------------------------------------------------------------------------
repos_mode                e                                   c n
     caa_cl(25ebea90-784a-11e1-a79d-b6fcc11f846f)                     e
--------------------------------------------------------------------------------

n/a means parameter not supported by the current platform or kernel

Scope codes:
    c = clusterwide: applies to the entire cluster
    s = per site: may be applied to one or more sites
    n = per node: may be applied to one or more nodes
    i = per interface: may be applied to one or more communication interfaces

Value conventions:
    K = Kilo: 2^10       G = Giga: 2^30       P = Peta: 2^50
    M = Mega: 2^20       T = Tera: 2^40       E = Exa: 2^60
```

*Figure 2-14 clctrl -tune output*

## 2.2.11 Troubleshooting CAA

We provide details about troubleshooting CAA by using the **snap** command. We review the syslog facility.

### SNAP command

The clustering subsystem provides a **snap** script to help you collect logs and provides data configuration that you can use to help troubleshoot problems.

The **snap caa** command collects data and log files in the directory structure as shown in Example 2-15 on page 57.

*Example 2-15 snap caa command*

```
root@caa2 / # snap caa
********Checking and initializing directory structure
Creating /tmp/ibmsupt/caa directory tree... done.
Creating /tmp/ibmsupt/testcase directory tree... done.
Creating /tmp/ibmsupt/other directory tree... done.
********Finished setting up directory /tmp/ibmsupt

Checking Space requirement for caa
Checking for enough free space in filesystem... done.

Gathering caa data
```

### Log file

CAA uses the syslog facility to log debug information and errors. Example 2-16 shows the default syslog configuration for CAA.

*Example 2-16 CAA syslog configuration /etc/syslog.conf file*

```
caa.info /var/adm/ras/syslog.caa rotate size 1m files 10
```

You can edit the `/etc/syslog.conf` file if you want to troubleshoot any particular problem. For example, you can change from *.info to the *.debug facility, which provides more information.

> **Refresh:** You need to refresh the syslog daemon if you modify the configuration file.

The following information is in the CAA log file:

- All configuration commands (**mkcluster**, **rmcluster**, and **chcluster**)
- All storage events and their delivery to the cluster network
- Error events

## 2.2.12 Statistics

To display the cluster statistics, use the **lscluster -s** command as shown in Example 2-17.

*Example 2-17 Displaying cluster statistics*

```
root@caa1 / # lscluster -s
Cluster Statistics:

Cluster Network Statistics:

pkts seen:14003431                      pkts passed:1811887
IP pkts:12826417                        UDP pkts:12240064
gossip pkts sent:2906895                gossip pkts recv:5479896
cluster address pkts:0          CP pkts:12191555
bad transmits:82                        bad posts:0
short pkts:0                    multicast pkts:12150150
cluster wide errors:0                   bad pkts:0
dup pkts:405                    pkt fragments:0
fragments queued:0              fragments freed:0
pkts pulled:0                   no memory:0
rxmit requests recv:5           requests found:5
requests missed:4               ooo pkts:34
requests reset sent:4            reset  recv:2
```

```
requests lnk reset send :0       reset  lnk recv:0
rxmit requests sent:1850
alive pkts sent:0               alive pkts recv:0
ahafs pkts sent:11551           ahafs pkts recv:11280
nodedown pkts sent:0            nodedown pkts recv:0
socket pkts sent:32875          socket pkts recv:31804
cwide pkts sent:797             cwide pkts recv:792
socket pkts no space:0          pkts recv notforhere:0
Pseudo socket pkts sent:0       Pseudo socket pkts recv:0
Pseudo socket pkts dropped:0
arp pkts sent:2                 arp pkts recv:1
stale pkts recv:41              other cluster pkts:2
storage pkts sent:1             storage pkts recv:1
disk pkts sent:7357             disk pkts recv:7305
unicast pkts sent:52712         unicast pkts recv:41418
out-of-range pkts recv:0
```

## 2.2.13 Gathering more cluster information

We describe how to gather more cluster information.

### *Obtaining the node and repository disk Universally Unique Identifier*

Use the command that is shown in Example 2-18 to list the node and the repository disk Universally Unique Identifier (UUID).

*Example 2-18 Node and repository disk ID*

```
root@caa1 / # lsattr -El cluster0
clvdisk   4d621527-d912-3479-5e16-5c9a03980323 Cluster repository disk identifier True
node_uuid 25bc9128-784a-11e1-a79d-b6fcc11f846f OS image identifier                True
```

### *Listing valid cluster repository disks*

Use the command in Example 2-19 to list the valid cluster repository disks.

*Example 2-19 List valid cluster repository disks*

```
root@caa1 / # /usr/lib/cluster/clras lsrepos
hdisk2 has a cluster repository signature.
Found 1 cluster repository disk.
```

### *Displaying storage framework UUID for disks*

Use the command in Example 2-20 to display the repository disk identifier.

*Example 2-20 Display UUIDs for disk*

```
root@caa1 / # /usr/lib/cluster/clras sfwinfo -d hdisk2
hdisk2  4d621527-d912-3479-5e16-5c9a03980323
```

### *Displaying content of cluster repository disk*

To display the content of the repository disk, including the CAA tunable values and cluster configuration, use the **clras dumprepos** command as shown in Example 2-21.

*Example 2-21 Content of the repository disk*

```
root@caa1 / # /usr/lib/cluster/clras dumprepos
HEADER
        CLUSRECID:      0xa9c2d4c2
```

```
        Name:          caa_cl
        UUID:          25ebea90-784a-11e1-a79d-b6fcc11f846f
        SHID:          0x0
        Data size:     3584
        Checksum:      0x64e2
        Num zones:     0
        Dbpass:
        Multicast:     228.1.1.48
        config_timeout : 240
        node_down_delay : 10000
        node_timeout : 20000
        link_timeout : 0
        deadman_mode : a
          repos_mode : e
          hb_src_lan : 1
          hb_src_san : 2
          hb_src_disk : 1
             site_up : e
             site_down : e

DISKS none

NODES numcl   numz    uuid                                  shid    flags          name
        1       0     25bc9128-784a-11e1-a79d-b6fcc11f846f    1       00000000       caa1
        1       0     25e7a7dc-784a-11e1-a79d-b6fcc11f846f    2       00000000       caa2

ZONES none
```

### 2.2.14 CAA cluster commands and force option

In the R2 release of CAA, the **-f** flag is not used to force the commands. See 2.2.1, "Changes" on page 38 for a list of commands that are not supported. If you want to force some CAA commands, the environment variable *CAA_FORCE_ENABLE* is set to 1 before executing the command instead of using the **-f** flag.

Example 2-22 shows the **rmcluster** with force to scrub the repository disk and ignore all errors.

*Example 2-22 rmcluster command with force*

```
# export CAA_FORCE_ENABLED=1
# rmcluster -r hdisk2
```

See Table 2-2 on page 60 for a list of CAA commands and actions with and without the force option.

*Table 2-2 CAA commands*

| Command | No force | With force |
|---|---|---|
| `chcluster -m|-d` | Adds or removes the node or disk. It fails on an error. | Adds or removes the node or disk. It ignores and continues on errors. |
| `rmcluster -n` | Removes the cluster. It fails on an error. | Removes the cluster. It ignores and continues on errors. |
| `rmcluster -r` | Only use with force. | Scrubs the repository disk and ignores all failures. |
| `clusterconf -u` | Only use with force. | Performs leave_sinc(), and removes the node from the cluster without going through a coordinated configuration sequence. |

## 2.2.15 Restricting interfaces from clustering use

When you create a cluster, all network interfaces that are configured with an IP address are used by CAA to exchange multicast packets between nodes. To restrict specific interfaces, you can configure them as type private. Every interface of a network that is defined as private in the PowerHA configuration is listed in the `/etc/cluster/ifrestrict` file, and it is not used by the CAA multicast communication.

> **Important:** Interfaces that are *not defined* in the HA topology can still be restricted. A typical example is a virtual LAN (VLAN) that is dedicated to backups and that is not included in the HA topology. If a client wants to reduce or avoid unnecessary traffic, the client adds the VLAN to the `ifrestrict` (even if it is not in the HA topology). After the client manually edits the `ifrestrict` file, the client must also run **`clusterconf`** to instruct CAA to include the change (even with HA active).

The following AIX levels are required to support this functionality:

- AIX 6.1 TL6 SP4 (bos.cluster.rte 6.1.6.4) and up
- AIX 7.1 TL0 SP3 (bos.cluster.rte 7.1.0.15) and up

### Restricting interfaces

We describe how to restrict interfaces in a 2-node cluster with two networks that are configured as public with two interfaces each (Figure 2-15 on page 61).

> **Private networks:** The private network is reintroduced in PowerHA SystemMirror 7.1.1. It is used for Oracle interconnect purposes only. No other external or user traffic is allowed to use it. There is no heartbeating or any PowerHA related communication on it. The private adapters are listed in the `/etc/cluster/ifrestrict` file. You cannot mark the interface that has the host name as private.

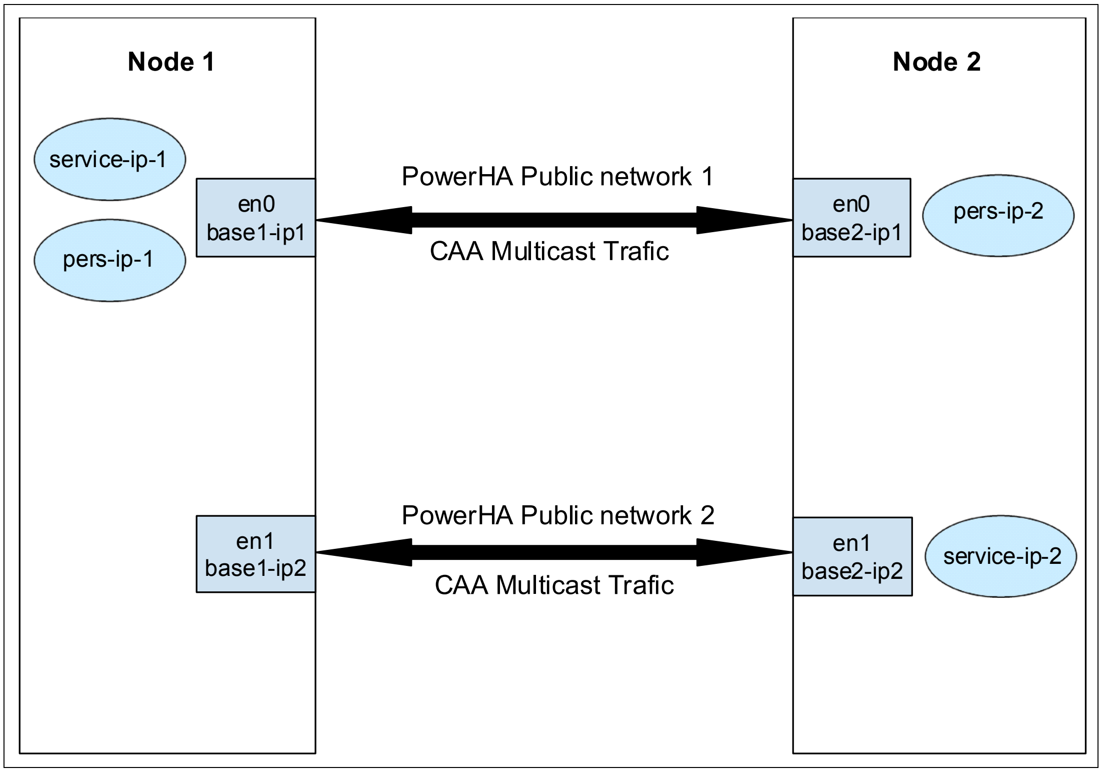



*Figure 2-15 Two-node cluster with two defined networks*

Example 2-23 shows the AIX network configuration for both nodes.

*Example 2-23 AIX network configuration for both nodes*

```
root@caa1 / # ifconfig -a
en0:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFL
OAD(ACTIVE),CHAIN>
        inet 10.1.1.48 netmask 0xfffffe00 broadcast 10.1.1.255
        inet 172.16.21.48 netmask 0xfffffe00 broadcast 172.16.21.255
        inet 172.16.21.69 netmask 0xfffffe00 broadcast 172.16.21.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
en1:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFL
OAD(ACTIVE),CHAIN>
        inet 10.1.2.48 netmask 0xffffff00 broadcast 10.1.2.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1

root@caa2 / # ifconfig -a
en0:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFL
OAD(ACTIVE),CHAIN>
        inet 10.1.1.33 netmask 0xfffffe00 broadcast 10.1.1.255
        inet 172.16.21.33 netmask 0xfffffe00 broadcast 172.16.21.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
en1:
flags=1e080863,10480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OF
FLOAD(ACTIVE),CHAIN>
        inet 10.10.20.69 netmask 0xffffff00 broadcast 10.10.20.255
        inet 10.1.2.33 netmask 0xffffff00 broadcast 10.1.2.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
```

Example 2-24 shows a 2-node PowerHA cluster topology.

*Example 2-24 Two-node PowerHA cluster topology*

```
root@caa2 /etc/cluster # cltopinfo
Cluster Name: caa_cl
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk5
Cluster IP Address: 228.1.1.48
There are 2 node(s) and 2 network(s) defined

NODE caa1:
        Network net_ether_01
                caasvc1 172.16.21.69
                caa1    10.1.1.48
        Network net_ether_02
                caa1b2  10.1.2.48

NODE caa2:
        Network net_ether_01
                caasvc1 172.16.21.69
                caa2    10.1.1.33
        Network net_ether_02
                caa2b2  10.1.2.33
```

Figure 2-16 shows the initial state of the en1 interface on the caa1 node at the CAA cluster level.

```
Interface number 2 en1
               ifnet type = 6 ndd type = 7
               Mac address length = 6
               Mac address =  b6.fc.c1.1f.84.70
               Smoothed rrt across interface = 7
               Mean Deviation in network rrt across interface = 3
               Probe interval for interface = 100 ms
               ifnet flags for interface = 0x1e080863
               ndd flags for interface = 0x21081b
               Interface state  UP
               Number of regular addresses configured on interface = 1
               IPV4 ADDRESS: 10.1.2.48  broadcast 10.1.2.255  netmask 255.255.255.0
               Number of cluster multicast addresses configured on interface = 1
               IPV4 MULTICAST ADDRESS: 228.1.1.48  broadcast 0.0.0.0  netmask 0.0.0.0
```

*Figure 2-16 Output of lscluster -i command for the en1 interface*

To restrict specific interfaces for cluster usage, follow these steps:

1. Define the network as a private network in the PowerHA configuration (Figure 2-17 on page 63). Use **smitty sysmirror**. Select **Cluster Nodes and Networks** → **Manage Networks and Network Interfaces** → **Networks** → **Change/Show a Network**.

```
 +--------------------------------------------------------------------------+
 |                      Select a Network to Change/Show                     |
 |                                                                          |
 | Move cursor to desired item and press Enter.                             |
 |                                                                          |
 |   net_ether_01 (172.16.20.0/23 10.1.0.0/23)                              |
 |   net_ether_02 (10.1.2.0/24)                                             |
 |                                                                          |
 | F1=Help                F2=Refresh             F3=Cancel                  |
 | F8=Image               F10=Exit               Enter=Do                   |
 | /=Find                 n=Find Next                                       |
 +--------------------------------------------------------------------------+

                            Change/Show a Network

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

                                                        [Entry Fields]
* Network Name                                        net_ether_02
  New Network Name                                   []
* Network Type                                       [ether] +
* Netmask(IPv4)/Prefix Length(IPv6)                  [255.255.255.0]
* Network attribute                                   private +
```

*Figure 2-17 Change the network attribute to private*

2. Stop the cluster services in all nodes.
3. Verify and synchronize the cluster configuration.

> **Important:** If you try to synchronize the cluster configuration with the cluster services running, the synchronization fails with this error:
>
> `cldare: Changes made to the networking configuration will result in IP aliasing support being enabled or disabled. Please note that changing these network parameters via a DARE is not supported.`

4. Start the cluster services.

> **Restricting an interface:** The `/etc/cluster/ifrestrict` file is created every time that PowerHA starts. This file can be edited manually, which is the only way that a client can restrict an interface that is not in the PowerHA topology.

## Testing the restricted interfaces

After you perform these steps, verify that the `/etc/cluster/ifrestrict` file is created in every cluster node (Figure 2-18 on page 64). This file contains the interfaces that belong to the private network in the PowerHA configuration and that are restricted to CAA.

```
root@caa1 # cat /etc/cluster/ifrestrict
en1
----
root@caa2 # cat /etc/cluster/ifrestrict
en1
```

*Figure 2-18 Content of /etc/cluster/ifrestrict file*

We see the en1 interface with the status DOWN RESTRICTED in the **lscluster** command output. This status means that CAA does not use this interface for multicasting any longer as shown in Figure 2-19.

```
 Interface number 2 en1
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.c1.1f.84.70
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  DOWN  RESTRICTED  SOURCE HARDWARE RECEIVE  SOURCE
HARDWARE TRANSMIT
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 10.1.2.48  broadcast 10.1.2.255  netmask 255.255.255.0
                Number of cluster multicast addresses configured on interface = 1
                IPV4 MULTICAST ADDRESS: 228.1.1.48  broadcast 0.0.0.0  netmask 0.0.0.0
```

*Figure 2-19 Output of the lscluster -i command for the restricted interface*

You can also use the **tcpdump** command to verify that there is no multicast traffic in the restricted interface as shown in Figure 2-20.

```
root@caa2 /etc/cluster # tcpdump -t -i2 -v ip and host 228.1.1.48
tcpdump: listening on en1, link-type 1, capture size 96 bytes

820 packets received by filter
0 packets dropped by kernel
```

*Figure 2-20 Tcpdump command output for the restricted interface*

You can also use the **tcpdump** command for the non-restricted interface. Figure 2-21 on page 65 shows how the multicast traffic goes through the interface.

```
root@caa2 /etc/cluster # tcpdump -t -i1 -v ip and host 228.1.1.48
tcpdump: listening on en0, link-type 1, capture size 96 bytes
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450
...
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450
IP (tos 0x0, ttl  32, id 0, offset 0, flags [none], proto: UDP (17), length: 1478)
caa2.drmsfsd > 228.1.1.48.drmsfsd: UDP, length 1450

460 packets received by filter
0 packets dropped by kernel
```

*Figure 2-21 Tcpdump command output in a non-restricted interface*

**Important:** You cannot configure the network as private in PowerHA if you have a mix of Boot and Service or Persistent IPs that are configured in the network. It fails with the following error:

```
ERROR: Network net_ether_02 has a mix of interface types.
Only networks with all boot or all service labels can be converted to private.
```

After the changes are completed, the new scenario for the 2-node cluster with one Public and one Private Network is shown in Figure 2-22.

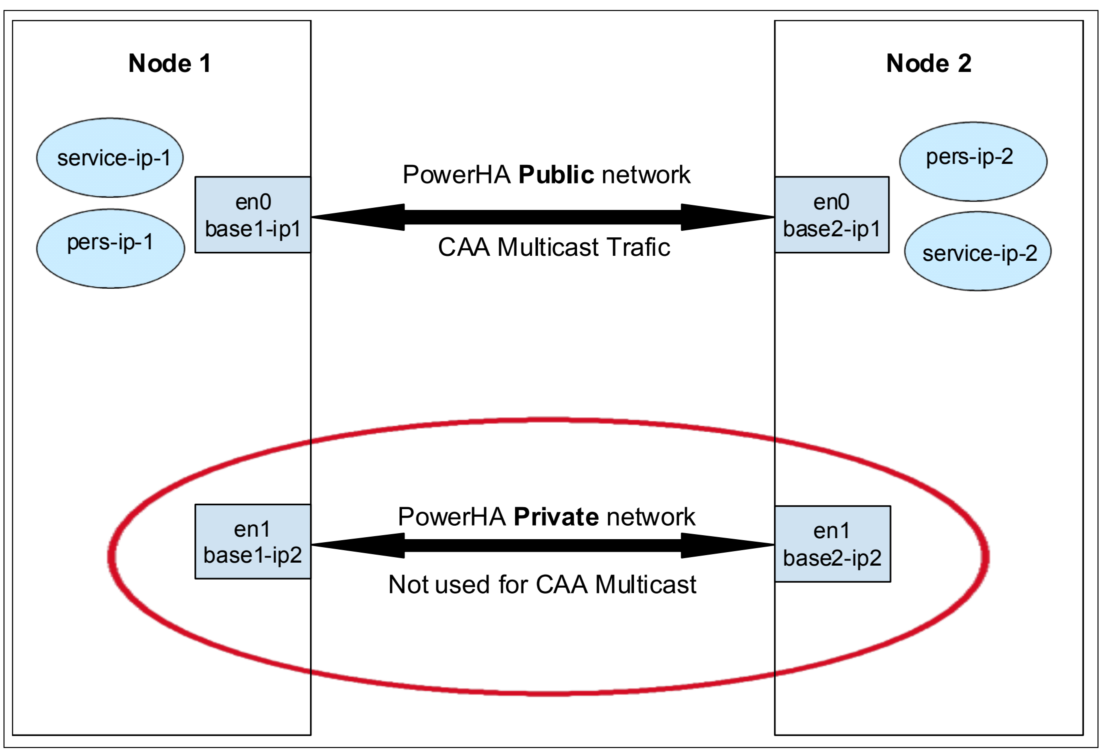



*Figure 2-22 Two-node cluster with a public network and a private network*

## 2.2.16 VLAN tagging

In previous versions of AIX 7.1 and AIX 6.1, CAA did not accept packets with VLAN tags in the Ethernet header. Any network interfaces that are defined on top of an AIX VLAN adapter must be “restricted” through the CAA `ifrestrict` mechanism.

As of AIX 7.1 TL1 SP03 and AIX 6.1 TL7 SP03, CAA fully supports VLAN tagging, including the use of AIX pseudo VLAN devices.

## 2.2.17 Round-trip time and heartbeats

Unlike previous versions of PowerHA, normally heartbeat tuning is unnecessary. CAA monitors the interfaces of each node by using the multicast protocol and gossip packets. Gossip packets are periodically sent from each node in the cluster for timing purposes. These gossip packets are automatically replied to by the other nodes in the cluster. The packet exchanges are used to calculate the round-trip time.

The round-trip time (`rtt`) value is shown in the output of the **`lscluster -i`** and **`lscluster -m`** commands. The mean deviation in network `rtt` is the average round-trip time, which is automatically managed by CAA.

To change the cluster heartbeat settings, modify the failure cycle and the grace period for the PowerHA cluster from the custom cluster configuration in the SMIT sysmirror panel (Example 2-25). Use **`smit sysmirror`**. Select **Custom Cluster Configuration** → **Cluster Nodes and Networks** → **Manage the Cluster** → **Cluster heartbeat settings**.

*Example 2-25 Cluster heartbeat settings SMIT panel*

```
                            Cluster heartbeat settings

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

                                                        [Entry Fields]
* Grace Period                                       [0]
* Failure Cycle                                      [0]
```

In Example 2-25, the following definitions apply:

- *Grace period*: The grace period is the amount of time in seconds that the node waits before the node marks the remote node as DOWN. This value can be 5 - 30 seconds. The default value is zero seconds.
- *Failure cycle*: The failure cycle determines the frequency of the heartbeat. This value can be 1 - 20 seconds. The default value is zero seconds.

# 2.3 End-to-end monitoring

We provide the monitoring details.

### Point-of-contact

A *point-of-contact* is a receive-side entity in software that is used to indicate connectivity from the current node to another node in the cluster. It correlates with an actual communications interface from which traffic from the remote node is received. A point-of-contact is most important for when CAA performs unicast communications between the current node and a specific other node.

A point-of-contact indicates that a node actually received communication packets across this interface from another node. This communication process allows the application that is monitoring the health of a node to act discretely based on near real-time event notification. You can also monitor the storage devices to provide UP and DOWN events for any recovery actions that are identified as necessary by the monitoring application.

## Point-of-contact status

The point-of-contact is marked UP when the first packet of any type is received in the interface. The point-of-contact is marked down when no communication packets are received for the node down period of time.

CAA monitors both the communication interface states and the points-of-contact between nodes on a node-by-node basis (Example 2-26).

> **Monitoring:** The ability to monitor this particular condition is important. An interface in the UP state and a point-of-contact in a down state can occur because of the hardware or other network issues between these particular nodes.

*Example 2-26 lscluster - m command output*

```
root@caa1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: caa1
        Cluster shorthand id for node: 1
        uuid for node: 0f5bfa5c-857b-11e1-8cf4-b6fcc11f846f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        caa_cl             local        0f945780-857b-11e1-8cf4-b6fcc11f846f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: caa2
        Cluster shorthand id for node: 2
        uuid for node: 0f9001b2-857b-11e1-8cf4-b6fcc11f846f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        caa_cl             local        0f945780-857b-11e1-8cf4-b6fcc11f846f

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
```

```
dpcom  UP  RESTRICTED
en0  UP
en1  UP
```

## 2.4 CAA daemons

When CAA is active, the following daemon services run as shown in Figure 2-23.

```
root@caa2 / # lssrc -g caa
Subsystem         Group            PID          Status
 clcomd           caa              6553806      active
 clconfd          caa              6619352      active
```

*Figure 2-23 CAA services*

CAA includes the following services:

- clcomd: This daemon is the cluster communications daemon, which changed in PowerHA 7.1. In previous versions of PowerHA, it was called clcomdES. The location of the `rhosts` file that PowerHA uses also changed. The `rhosts` file that is used by the clcomd service is in the `/etc/cluster/rhosts` file. The old clcomdES `rhosts` file in the `/usr/es/sbin/cluster/etc` directory is not used.
- clconfd: The clconfd subsystem runs on each node of the cluster. The clconfd daemon wakes up every 10 minutes to synchronize any necessary cluster changes.

> **Daemons:** The solid and solidhac daemons are no longer used by CAA.

## 2.5 CAA multicast

*Multicast* consists of sending messages or information from the source to a group of hosts simultaneously in a single transmission. This communication uses network infrastructure efficiently. The source sends a packet only one time, even if it needs to be delivered to many receivers. The other nodes in the network replicate the packet to reach multiple receivers only when necessary.

CAA uses a multicast IP-based network communication mechanism to communicate between nodes in a cluster. A special "*gossip*" protocol is used to determine node information and implement scalable reliable multicast. For cluster communication, you can manually configure a multicast address or have CAA automatically select the multicast address.

### 2.5.1 Testing multicast in a network

Do not attempt to create a PowerHA cluster until you verify that multicast packets can be sent successfully across all nodes that are part of the PowerHA cluster.

To test end-to-end multicast communication across all nodes in your network, run the **mping** command to send and receive packets between nodes.

If you run PowerHA SystemMirror 7.1.1, or later, you cannot create a cluster if the **mping** command fails. If the **mping** command fails, your network is not set up correctly for multicast communication. If your network is not set up correctly for multicast communication, review the documentation for your switches and routers to enable multicast communication.

You can run the **mping** command with a specific multicast address; otherwise, the command uses a default multicast address. You must use the multicast address that is used to create the cluster as input for the **mping** command (Figure 2-24).

> **Tip:** To see this IP, use the **lscluster -c** command to display it.

```
root@caa2 / # mping
Usage: mping -r|-s [-v] [-a address] [-p port] [-t ttl]

-r|-s        Receiver or sender. Required argument,
             mutually exclusive
-a address   Multicast address to listen/send on,
             overrides the default. (must be < 16 characters long)
-p port      Multicast port to listen/send on,
             overrides the default of 4098.
-t ttl       Multicast Time-To-Live to send,
             overrides the default of 1.
-v           Verbose mode
-c count     Count of packets to send/receive before quitting,
             default is to quit on receiving interrupt signal
```

*Figure 2-24 mping command options*

Example 2-27 shows the **mping** command for the multicast address 228.1.1.34, where node A is the receiver and node B is the sender.

*Example 2-27 mping command*

```
root@caa1 / # mping -v -r -c 5 -a 228.1.1.34
mping version 1.1
Localhost is caa1, 10.1.1.48
Listening on 228.1.1.34/4098:

Replying to mping from 10.1.1.33 bytes=32 seqno=1 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.33 bytes=32 seqno=2 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.33 bytes=32 seqno=3 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.33 bytes=32 seqno=4 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.33 bytes=32 seqno=5 ttl=32
Discarding receiver packet

root@caa2 / # mping -v -s -c 5 -a 228.1.1.34
mping version 1.1
Localhost is caa2, 10.1.1.33
mpinging 228.1.1.34/4098 with ttl=32:

Discarding sender packet
32 bytes from 10.1.1.48: seqno=1 ttl=32 time=0.471 ms
```

```
Discarding sender packet
32 bytes from 10.1.1.48: seqno=2 ttl=32 time=0.321 ms
Discarding sender packet
32 bytes from 10.1.1.48: seqno=3 ttl=32 time=0.394 ms
Discarding sender packet
32 bytes from 10.1.1.48: seqno=4 ttl=32 time=0.360 ms
Discarding sender packet
32 bytes from 10.1.1.48: seqno=5 ttl=32 time=0.458 ms

--- mping statistics ---
5 packets transmitted, 5 packets received
round-trip min/avg/max = 0.321/0.401/0.471 ms
```

CAA selects a default multicast address if you do not specify a multicast address when you create the cluster. The default multicast address is created by combining the logical OR of the value (228.0.0.0) with the lower 24 bits of the IP address of the node as shown in Example 2-28.

*Example 2-28 Default multicast address*

```
if the node's IP address is 10.1.1.33, then the default multicast address would be
228.1.1.33

root@caa2 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  caa_cl
Cluster uuid:  cfd568b8-7207-11e1-9211-b6fcc11f846f
Number of nodes reporting = 2
Number of nodes expected = 2
Node caa2
Node uuid = cf98b95e-7207-11e1-9211-b6fcc11f846f
Number of interfaces discovered = 3
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.d5.65.89.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 172.16.21.33  broadcast 172.16.21.255  netmask 255.255.254.0
                IPV4 ADDRESS: 10.1.1.33  broadcast 10.1.1.255  netmask 255.255.254.0
                Number of cluster multicast addresses configured on interface = 1
                IPV4 MULTICAST ADDRESS: 228.1.1.33 broadcast 0.0.0.0  netmask 0.0.0.0
```

## 2.5.2 Troubleshooting multicast

Use the **mping** command to test whether your nodes can send and receive multicast packets. If the **mping** command fails, you need to identify the problem in your network environment.

To troubleshoot multicast problems, review the following guidelines:

- Review the documentation for the switches that are used for multicast communication.

- Disable the Internet Group Management Protocol (IGMP) snooping on the switches that are used for multicast communication.
- Eliminate any cascaded switches between the nodes in the cluster. Have only a single switch between the nodes in the cluster.
- Verify that an entry exists in the routing table for the multicast IP address.

**Troubleshooting:** If your network infrastructure does not allow IGMP snooping to be disabled permanently, you might be able to troubleshoot problems by temporarily disabling snooping on the switches and then adding more network components one at a time.

**Error:** The **mping** command fails with the following message if there is no routing entry for the multicast address:

```
root@caa1 / # mping -r -a 228.1.1.34
mping version 1.1
setsockopt() failed: Can't assign requested address
```

3

# PowerHA 7.1.1 basic installation and configuration

This chapter describes step-by-step how to design and implement a PowerHA cluster.

We include the following topics:

- Introduction
- Designing your cluster
- Hardware requirements
- Important considerations for VIOS
- Networking
- Disk considerations
- PowerHA SystemMirror installation and prerequisites
- Configuring PowerHA SystemMirror topology
- Configuring PowerHA resources
- Configuring the application for PowerHA
- Dynamic LPAR and capacity on demand resources
- Test scenarios
- Troubleshooting

## 3.1 Introduction

We show you how to design, install, and configure a 2-node, mutual takeover PowerHA SystemMirror cluster. We focus on a typical, widely used cluster configuration, which covers the usual client requirements.

We produced a detailed, step-by-step installation guide, which is suitable for system administrators without prior clustering experience. We assume that you have a basic understanding of AIX and that you read Chapter 1, "Introducing IBM PowerHA SystemMirror 7.1.1" on page 1.

Cluster Aware AIX (CAA) is the new underlying layer of PowerHA SystemMirror 7.1.1. You do not need to be afraid of CAA. You do not have to deal directly with it. PowerHA commands take care of it. We use only one command to check the initial setup of the CAA cluster.

Figure 3-1 shows a typical, 2-node, mutual takeover PowerHA SystemMirror 7.1.1 cluster.

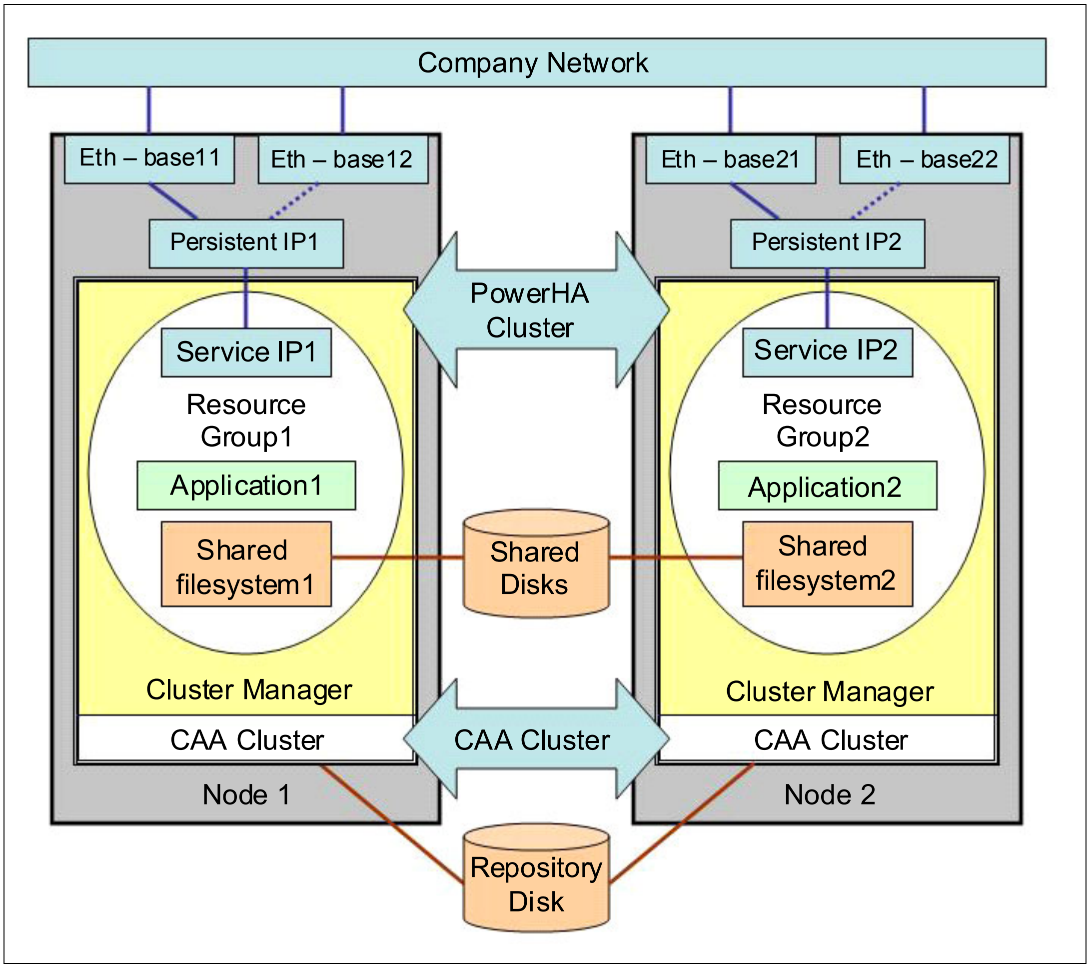



*Figure 3-1 Typical PowerHA SystemMirror 7.1.1 cluster*

### 3.1.1 Our sample cluster

We install a typical, 2-node, mutual takeover cluster.

We summarize what we know about our cluster configuration at the beginning of the design phase:

- We have a 2-node cluster.
- Application1 runs on the first node.
- Application2 runs on the second node.
- In a node failure, the application is recovered in the other node.
- We use a fully virtualized environment.

Next, we describe how to collect all required configuration data for the PowerHA implementation.

## 3.2 Designing your cluster

In the beginning of the design phase, usually limited information about our cluster configuration is available. You might know only the application and its basic requirements: required disk size, mount points, and the designated service IP address. Starting from this information, you can gradually collect all required data for the cluster configuration. Follow the steps to design and implement your own cluster.

Now, we show you the two most important design concepts of the PowerHA cluster:

- Avoid a single point of failure (SPOF).
- Keep it simple.

> **Important:** Even with simple cluster designs, anyone who operates the cluster without adequate AIX and PowerHA skills often introduces downtime.

### 3.2.1 Avoid a single point of failure

When you implement a PowerHA cluster, ensure that there are no hardware or software components that are a single point of failure. Check all components of the installation for redundancy:

- The server has a redundant power supply that is connected to two independent electrical circuits.
- There are redundant fan modules or cooling in your server.
- The storage area network (SAN) consists of two fully independent networks.
- Use two network switches per server for redundant Ethernet connectivity.
- Mirror your system disks (rootvg).
- Have two Ethernet cards for each network.
- Have two SAN adapters for each node.
- Use a redundant power supply for the Ethernet and SAN switches.
- Use a redundant storage system.
- Use disk multipath I/O (MPIO).
- For the multiport adapters, ensure that you do not use different ports of the same adapter to access the same network.
- For further reducing the possible downtime, consider buying a server with hot-plug extension cards and disks.

You need the following components if you plan to use a Virtual I/O Server (VIOS):

- Two VIOS per frame
- Two Ethernet cards for each network on each VIOS
- Two SAN adapter cards on each VIOS

### 3.2.2 Keep it simple

Try to keep the PowerHA configuration as simple as possible. As you introduce more components, for example, networks and resource groups, your configuration becomes more tangled and requires more effort to keep it working. In most cases, the simple 2-node cluster with two resource groups is sufficient. It is rare that you need to configure more than two resource groups or more than two networks.

## 3.3 Hardware requirements

PowerHA SystemMirror 7.1.1 can be installed on any hardware that is supported by AIX 6.1 TL7 SP2 or AIX 7.1 TL1 SP2. For supported IBM servers, extension cards, and storage subsystems, see the PowerHA hardware support matrix:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105638

The "Setting up cluster storage communication" page provides a list of supported Fibre Channel (FC) cards for SAN heartbeating:

http://publib.boulder.ibm.com/infocenter/aix/v7r1/index.jsp?topic=/com.ibm.aix.clusteraware/claware_comm_setup.htm

**Important:** It is highly suggested but not technically mandatory that all cluster nodes are on different hardware servers.

PowerHA supports full system partitions as well as fully virtualized logical partitions (LPARs). You can use a mixture of physical and virtual interfaces, too. If you plan to use physical adapters, ensure that you have the required redundancy.

## 3.4 Important considerations for VIOS

PowerHA 7.1.1 supports a virtualized environment. For example, you can use virtual Ethernet, virtual SCSI disks, and FC N-Port ID Virtualization (NPIV).

The PowerHA is indifferent to the device that you use: a real hardware device or virtual adapter. However, there are important considerations about using VIOS:

- Have two VIOS on each frame for redundancy.
- Configure Shared Ethernet Adapter (SEA) failover or network interface backup in the VIOS.
- You have to configure the `netmon.cf` file to check the status of the network behind the virtual switch.
- Have two independent Ethernet switches for fully redundant network access.
- Configure MPIO for disk access for the VIOS.
- Configure MPIO for disk access for the PowerHA LPAR.

This configuration is common for all virtualized environments so you use these features on your server anyway. Before you start installing and configuring the PowerHA cluster, check and test that the VIOSs, SEA failover, and MPIO work correctly.

For further information about VIOS configuration, see *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

### 3.4.1 Virtual Ethernet: SEA failover, network interface backup, and the number of virtual Ethernet adapters

During the developing of this book, we repeatedly asked “What is the suggested virtual Ethernet configuration?” The authors all had their own opinions, and there were many long debates on this topic. Finally, we agreed that there is no specific or suggested virtual Ethernet configuration because all redundant configurations work well in a PowerHA environment.

We suggest the following virtual Ethernet configuration:

- Two VIOSs per physical server.
- Use the servers that are already configured virtual Ethernet settings because no special modification is required. For a VLAN-tagged network, the preferred solution is to use SEA failover; otherwise, use the network interface backup.
- One client-side virtual Ethernet interface simplifies the configuration; however, PowerHA misses network events.
- Two virtual Ethernet interfaces need to be on the cluster LPAR to enable PowerHA to receive the network events. This configuration results in a more stable cluster.

> **Important:** The VIOS provides Ethernet network redundancy by using the SEA failover or Network Interface Backup (NIB) feature. However, the SEA failover and the NIB do not check the reachability of the specified ping address through the backup interface if the primary interface is up and running. So, you do not know that your backup path is working until the primary interface fails.

You might want two Ethernet interfaces in the cluster LPARs so that PowerHA can track the network changes, similar to physical network cards. Ensure that the two virtual Ethernet cards use different SEAs, VIOSs, and virtual Ethernet switches as their primary path.

In the following section, we describe some of the possible virtual Ethernet solutions and their properties. The only common component in the following configurations is the dual VIOSs.

#### Two Ethernet adapters in the PowerHA network (no SEA failover or NIB)

Figure 3-2 on page 78 shows the simplest configuration: two virtual Ethernet adapters without SEA failover or NIB:

- No SEA failover or NIB, which is easy to configure on the VIOS side.
- Two virtual Ethernet adapters per PowerHA network.
- Redundancy that is provided by PowerHA.
- Similar to the traditional, dual physical adapter configuration.
- If the VIOS is down, the PowerHA network interface swap happens in the cluster.

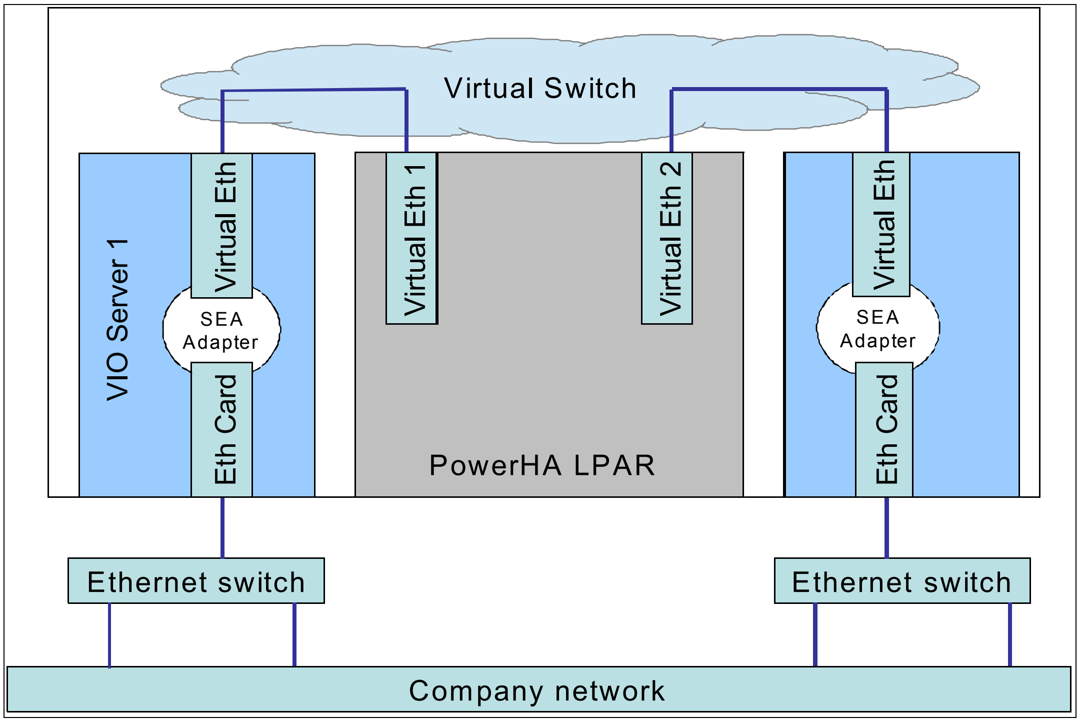



*Figure 3-2 Configuration with two virtual Ethernet adapters without SEA failover or NIB*

### NIB and a single Ethernet adapter in PowerHA network

The following characteristics apply to this scenario:

- NIB provides load balancing across the VIOS.
- This scenario can be used only when there is no VLAN tagging on the network.
- Redundancy is provided by the NIB EtherChannel active-backup setting.
- Only one Ethernet adapter is allowed per PowerHA network.
- PowerHA cannot see the network events.

See Figure 3-3 on page 79 for the design of the simple NIB configuration.

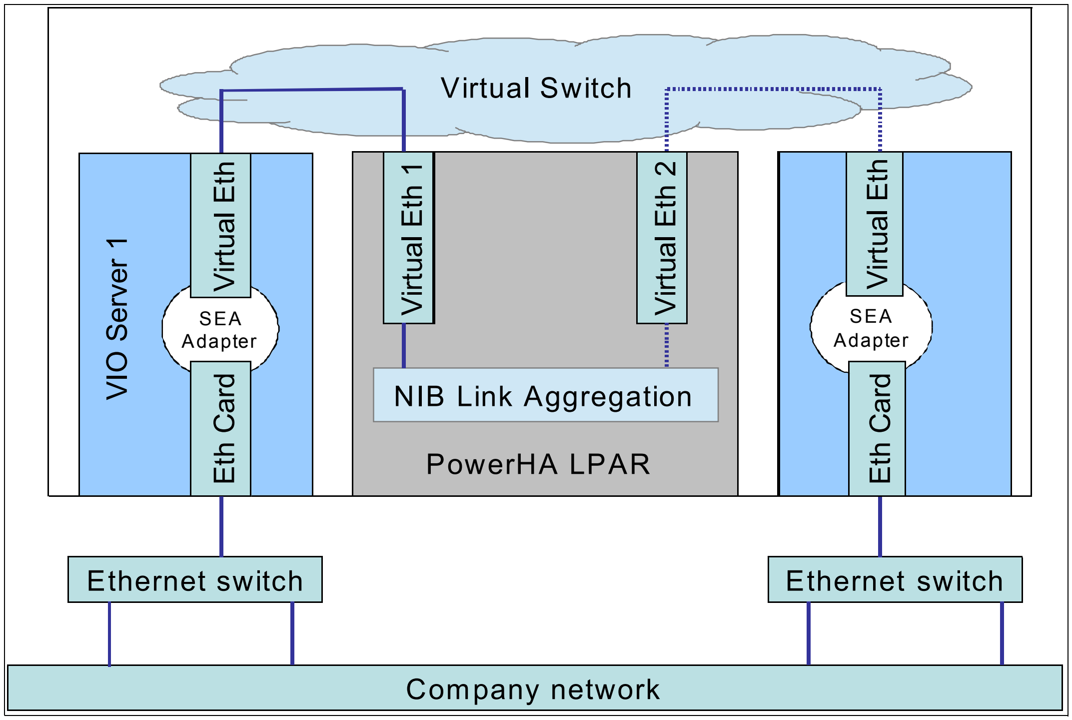



*Figure 3-3 Simple NIB configuration*

### NIB and two Ethernet adapters per PowerHA network

See Figure 3-4 on page 80 for the dual virtual Ethernet NIB configuration with the following characteristics:

- NIB provides load balancing across the VIOSs.
- This scenario can be used only when there is no VLAN tagging on the network.
- Two Ethernet adapters are required per PowerHA network.
- Two virtual switches are required.
- Double redundancy: PowerHA and NIB EtherChannel exist.
- PowerHA can track the network events.

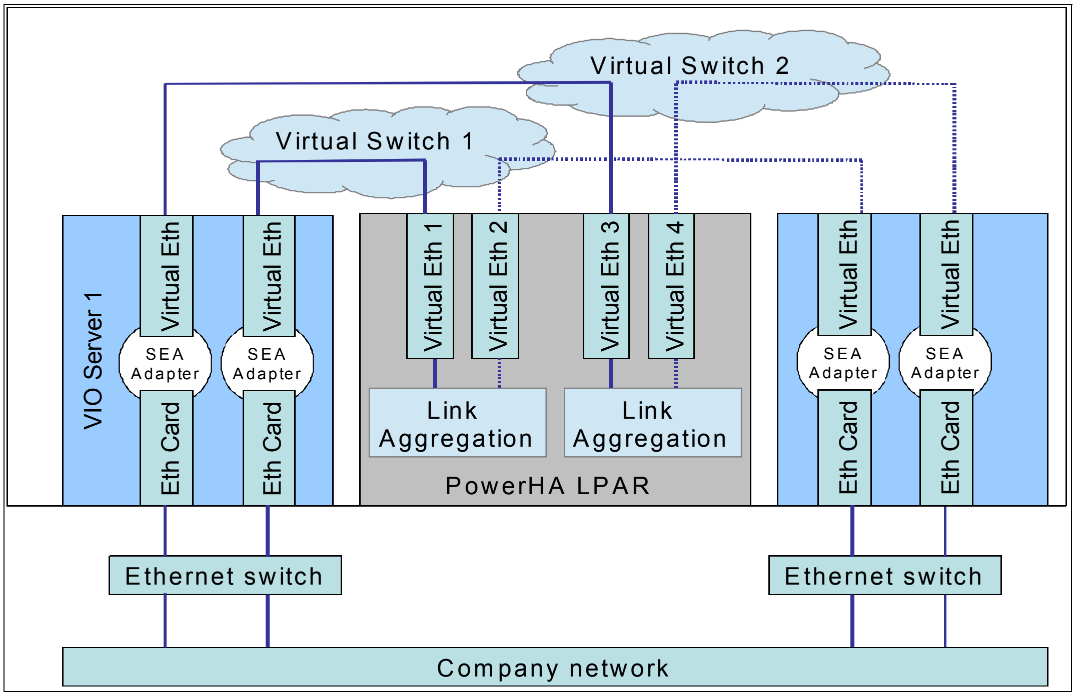



*Figure 3-4 NIB with dual Ethernet adapters on the client LPAR*

## SEA failover and one virtual Ethernet adapter on the client side

Figure 3-5 on page 81 shows the simple SEA failover configuration, which has these characteristics:

- It works with VLAN tagging.
- SEA shared mode provides load sharing between VLANs.
- When you use SEA shared mode, the default path might change after every reboot.
- Only one virtual Ethernet adapter exists per network.
- Redundancy is provided by SEA failover.
- PowerHA cannot track the network events.

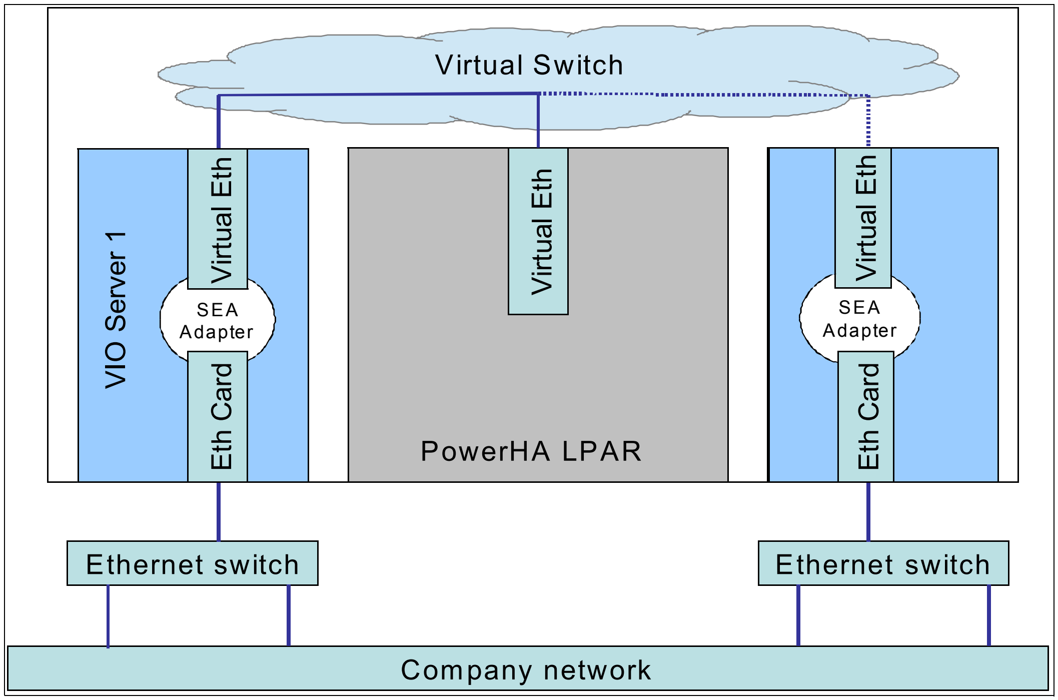



*Figure 3-5 Simple SEA failover configuration*

### SEA failover with two virtual Ethernet interfaces in the cluster LPAR

This configuration has the following characteristics:

- There are two virtual switches.
- Each virtual adapter connects to its own virtual switch and SEAs - no common devices.
- It works with VLAN tagging.
- SEA shared mode provides load sharing between VLANs.
- When you use SEA shared mode, the default path can change after every reboot.
- Two virtual Ethernet adapters exist per network.
- Dual redundancy is available with SEA failover and PowerHA.
- PowerHA tracks the network events.

Figure 3-6 on page 82 shows a drawing of the SEA failover with the virtual Ethernet interfaces in the cluster LPAR.

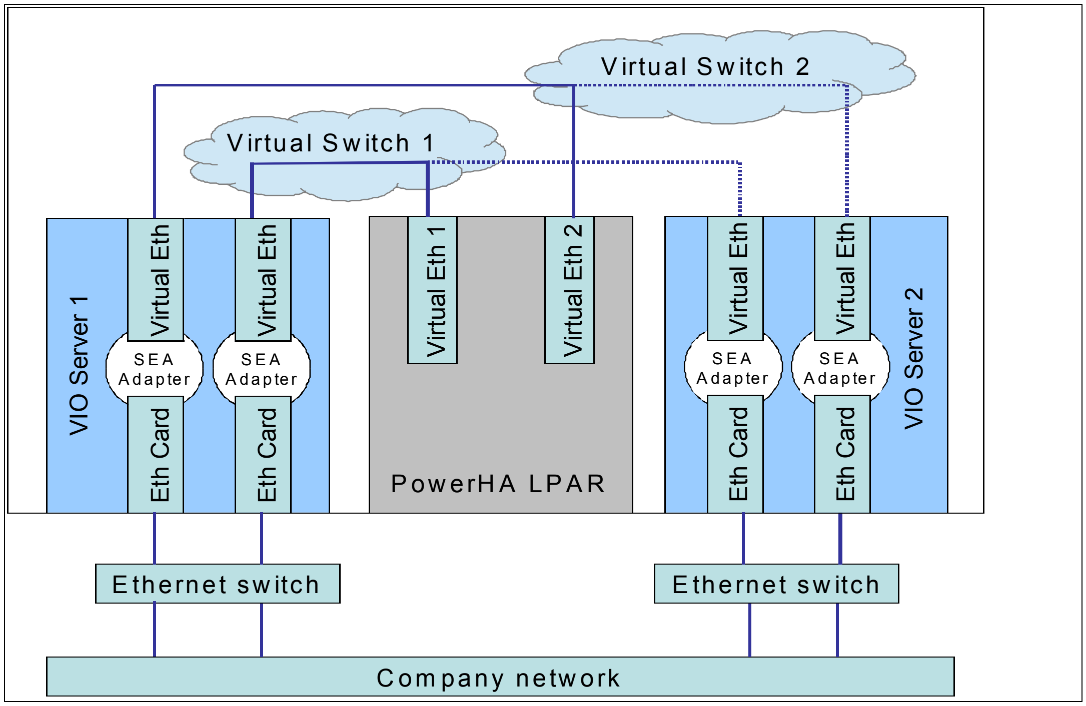



*Figure 3-6 SEA failover with two virtual Ethernet adapters on the client side*

## 3.4.2 Our sample cluster virtual environment

Figure 3-7 on page 83 shows our sample configuration, a typical redundant VIOS environment. All components are duplicated on the VIOS and client LPAR level, too. We have two, fully redundant Ethernet networks that are connected to the same subnet or VLAN. One network consists of one physical adapter on each VIOS, one dedicated virtual switch, and the corresponding SEA and virtual adapters. Thus, there are no single points of failure.

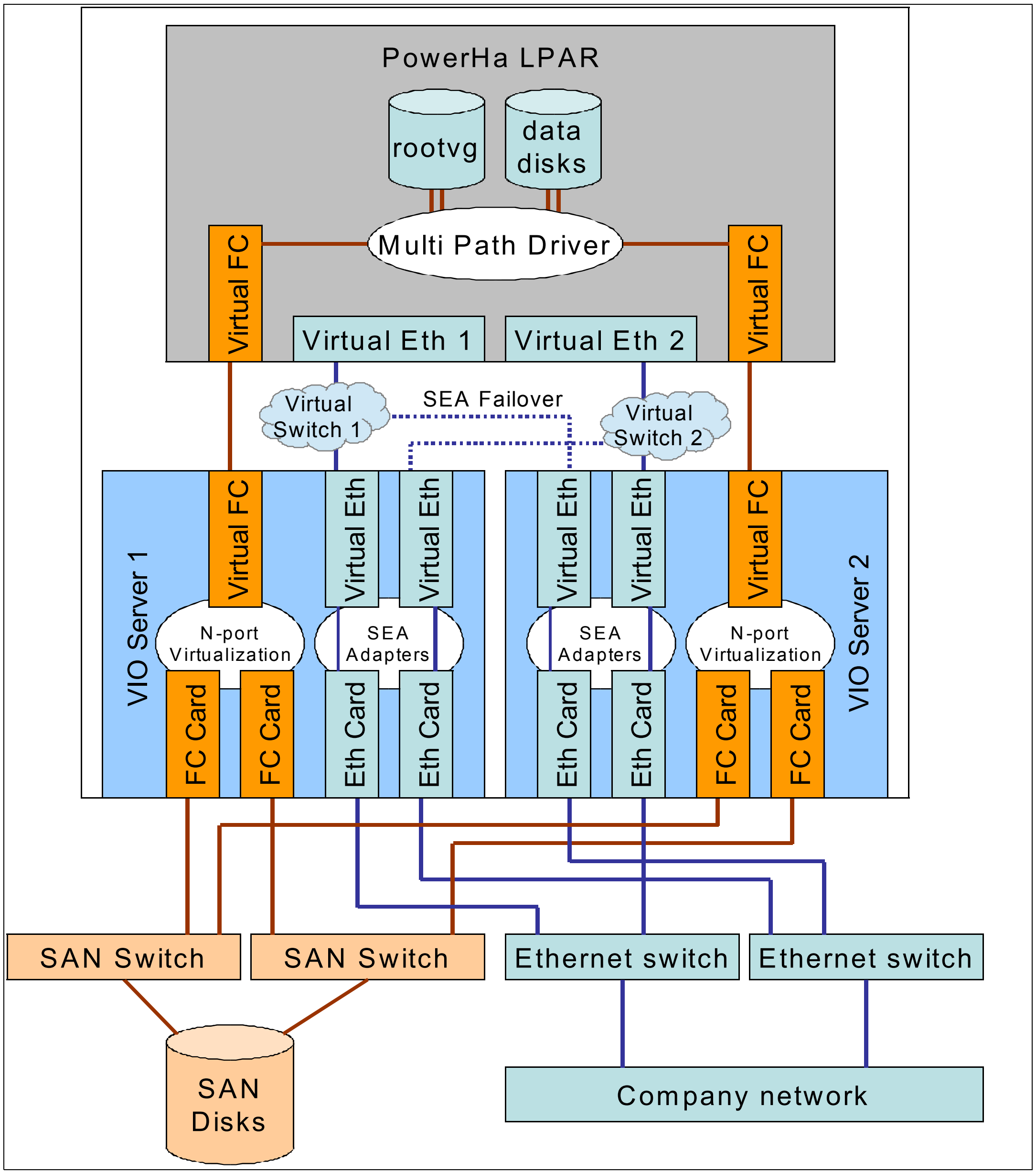



*Figure 3-7 Redundant VIOS environment of our sample cluster*

Under normal circumstances, the first network connects through VIOS 1, and the second network uses VIOS 2. If an Ethernet switch fails or a VIOS fails, the Shared Ethernet failover ensures that both connections remain online through the second switch or VIOS.

For rootvg, we use Virtual SCSI disks. For data disk access, we choose NPIV. Both solutions provide multipath capability through two VIOSs for the client LPARs. Therefore, a controller, VIOS, or SAN switch failure does not affect our system.

## 3.5 Networking

Networking is the most challenging part of the PowerHA configuration. First, we clarify the following PowerHA networking terms.

PowerHA uses two network types:

- IP network: Ethernet network for user external access
- Non-IP network: SAN-based heartbeat communication and disk heartbeat through the repository disk

PowerHA has two types of network functions:

- Service network: For normal user communication
- Private network: For Oracle Real Application Cluster (RAC) internal communication

The PowerHA network configuration (Figure 3-8) has the following network interfaces and functions:

- One or more service addresses
- One persistent address for each node (optional)
- One or two base addresses for each network that is protected by PowerHA

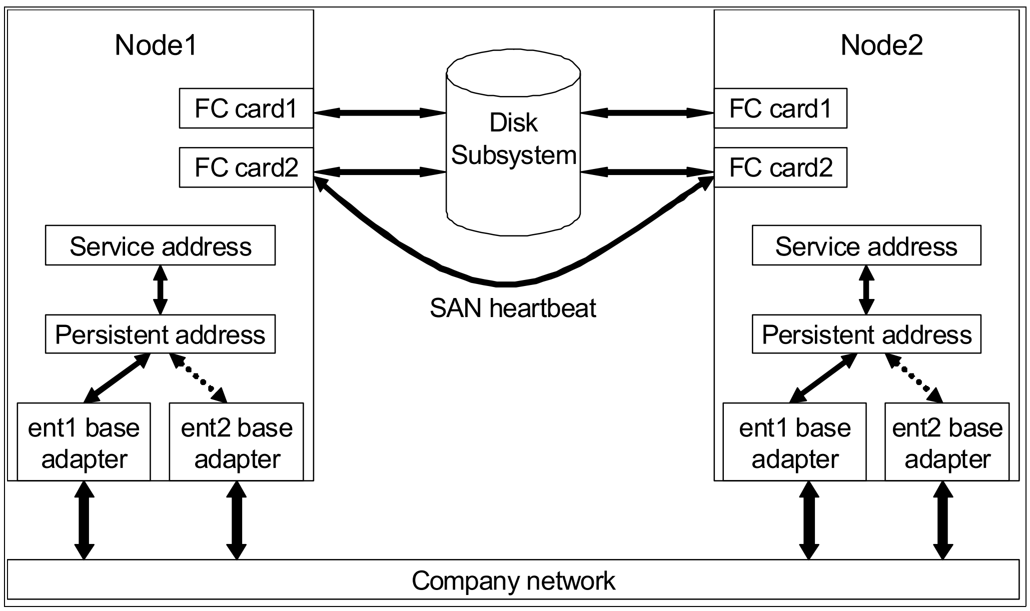


*Figure 3-8 Typical cluster network topology*

Figure 3-8 shows the typical PowerHA cluster network topology. For better viewing and understanding, we omitted the VIOS and its components.

> **Important:** The network must able to handle multiple IP addresses on the adapters. The same IP address can appear on different Ethernet cards and servers with a Media Access Control (MAC) address, which, on a typical network environment, is not a problem.

In the following section, we show the function of all interfaces and how to choose the correct IP address for them.

## 3.5.1 PowerHA network

A *PowerHA network* is a collection of interfaces and IP addresses that connect to the same "physical" network or VLAN.

On each node, a traditional PowerHA network configuration consists of the following components:

- Two Ethernet adapter cards with their base IP addresses
- One optional persistent address: An IP alias that is in one of the base adapters
- A service address: An IP alias that is in one of the base adapters

See Figure 3-8 on page 84 for a graphical representation of the PowerHA network components.

## 3.5.2 Number of Ethernet adapters per node per network

Although you can create a cluster with only one Ethernet adapter per service network per node, we highly suggest that you have two adapter cards for each network on each node. We recognized that the cluster topology and network error detection and correction works better with the dual-adapter configuration.

### Physical network adapters

If you use physical network adapters, always use two adapter cards per node for each PowerHA network. Otherwise, you lose the swap adapter functionality, so a single network error starts a resource group takeover and application outage.

### Virtual Ethernet

Even if you use a fully virtualized environment with SEA failover, we advise that you have two virtual Ethernet adapters for each node. For our suggested virtual Ethernet configuration, see 3.4.1, “Virtual Ethernet: SEA failover, network interface backup, and the number of virtual Ethernet adapters” on page 77.

### EtherChannel interfaces

PowerHA supports EtherChannel interfaces as the base adapter. However, try to avoid the use of EtherChannel in Active-Backup mode because PowerHA misses all network connectivity-related events. The EtherChannel protocol does not check the status of the backup adapter if the primary link is active. Another disadvantage is that you might have to use the same switch for both adapters. So, if you have two network cards, we advise that you configure them as base adapters for PowerHA and not use EtherChannel in Active-Backup mode.

We advise that you use 802.3AD link aggregation for higher network bandwidth requirements and for load balancing between the network cards. If possible, have a third interface as a secondary base adapter in PowerHA.

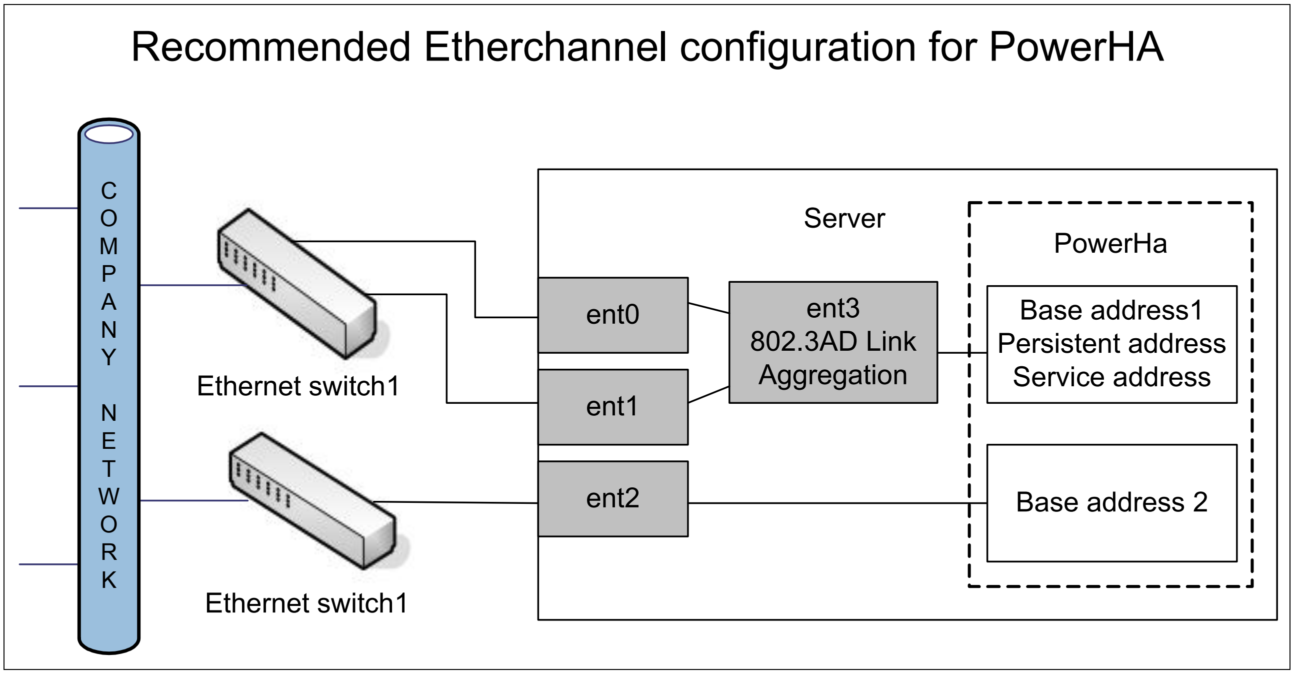



*Figure 3-9 Suggested EtherChannel configuration for PowerHA cluster*

Figure 3-9 shows the suggested EtherChannel configuration for the PowerHA cluster. We have three network cards (ent0, ent1, and ent2). Ent0 and ent1 are used for 802.3AD link aggregation. The EtherChannel interface is ent3, and we use ent3 as a base adapter in the cluster configuration. We can configure the persistent IP and the service IP address to be on this adapter, by default. The ent2 interface is used as a base interface too in a total EtherChannel failure or network switch failure. Then, the service address is relocated here.

### 3.5.3 Service network

The service network is used for external user communication. All of its interfaces are used for heartbeating; PowerHA constantly monitors their status. The Ethernet network is the service network, by default.

### 3.5.4 Private network

The private network is reintroduced in PowerHA SystemMirror 7.1.1. It is used for Oracle interconnect purposes only. No other external or user traffic is allowed to use it. There is no heartbeating or any PowerHA related communication on it. The private adapters are listed in the `/etc/cluster/ifrestrict` file. You cannot mark the interface private that has the host name.

### 3.5.5 Service address

The application users use the service IP address to connect to the service by the name: service address. This IP address is configured in a resource group and PowerHA brings it up and down with the corresponding application. In a takeover, this address is relocated to the secondary or backup node. If PowerHA is down, this address is down, too.

The service address is an IP alias that is in one of the base network interfaces. This address must be in the normal company data room network range and must be accessible from the outside network.

### 3.5.6 Persistent address

When the PowerHA services are down, or the application resource group is on the other node, we cannot use the service address for login. The persistent address is an IP alias, which is permanently connected to one of the base addresses of a node. It is always on the node. PowerHA does not move or change this address in a takeover or node that is down. This address can be used to distinguish a node. The persistent address is used to log on to the server during installation, when the PowerHA service address is down or when you want to connect to a certain node. This interface is used by the system administrators only; users never log on through it. We advise that the persistent IP address is in the same IP subnet with the service address.

> **Tip:** If you have the boot IP in the same subnet as the service IP, you can configure a default gateway or static routes without a persistent address.

### 3.5.7 Base addresses

Because PowerHA uses IP aliases for the service and persistent addresses, there must be an IP address that is associated with the network interface cards (NICs). This IP address is the base address. Each network that is defined in the PowerHA topology has one or (preferably) two NICs (physical or virtual). Each card has its own base IP address in a different IP subnet. The base IP address is used for CAA heartbeating only; there is no incoming or outgoing external connection for the base IP address. We advise you to use non-routable IP subnets for the base address.

### 3.5.8 Why use different subnets for the two base IP addresses

Because of TCP/IP routing rules, if an AIX server is connected to the same IP subnet by more than one network interface only, the first interface is used for outgoing packets. Both interfaces can receive packets. Figure 3-10 shows an AIX server that is connected to same subnet by using two interfaces.

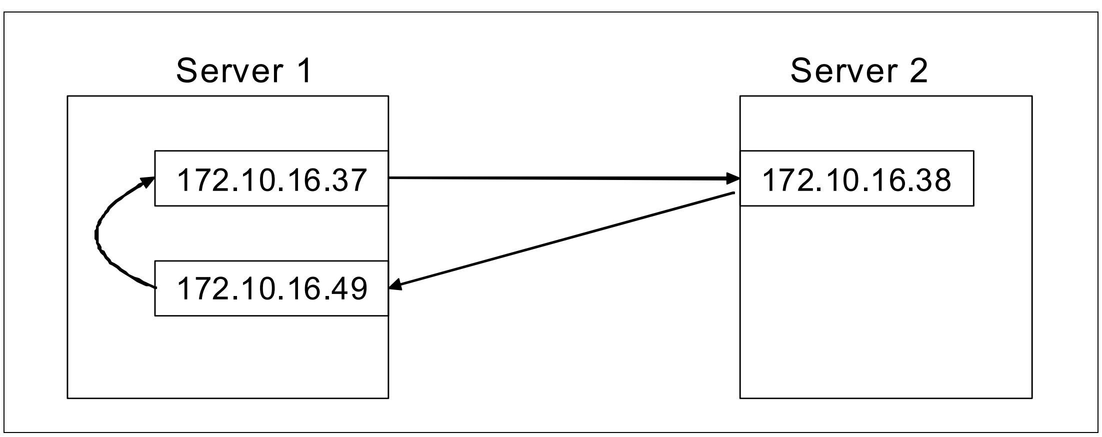



*Figure 3-10 AIX server that is connected to same subnet by using two interfaces*

Figure 3-10 on page 87 shows that when we address the second interface from an outside address, the TCP/IP packets are received on the second interface. But, the reply packet always uses the first interface as a source address. So, PowerHA is not able to verify the full functionality of the network cards.

In the preferred dual-network card PowerHA configuration, the service and persistent address must be in a separate IP subnet because they are on top of the base addresses.

### 3.5.9 Hostname and node name

Unlike earlier versions, now PowerHA SystemMirror 7.1.1 has strict rules for which interface can be the hostname due to the new CAA layer requirements.

**Important:**

- The hostname cannot be an alias in the `/etc/hosts` file.
- The name resolution for the hostname must work for both ways, therefore a limited set of characters can be used.
- The IP address that belongs to the hostname must be reachable on the server, even when PowerHA is down.
- The hostname cannot be a service address.
- The hostname cannot be an address located on a network which is defined as private in PowerHA.
- The hostname, the CAA node name, and the “communication path to a node” must be the same.
- By default, the PowerHA, nodename, the CAA nodename, and the “communication path to a node” are set to the same name.
- The hostname and the PowerHA nodename can be different.
- The hostname cannot be changed after the cluster configuration.

The rules leave the base addresses and the persistent address as candidates for the hostname. You can use the persistent address as the hostname only if you set up the persistent alias manually before you configure the cluster topology.

On the Domain Name System (DNS) server, you can specify any “external” name for the service addresses so that the clients can use this name when they connect to the application that runs in the cluster.

See our sample cluster hostname configuration in 3.7.5, “Setting up the hostname” on page 96.

**Note:** We could not configure a PowerHA 7.1.1 cluster because of a mismatch in the way the hostname was used. The hostname was defined using lowercase characters but in the /etc/hosts file it was written using uppercase. All the default TCP/IP commands worked fine. But the cluster setup failed. We updated the hosts file and then cluster setup worked fine.

### 3.5.10 IP subnet rules

We summarize the IP subnet rules for PowerHA SystemMirror.

**Important:**

- The two base IP address pairs for a specific physical network must be on separate subnets.
- All service and persistent addresses must be on a separate subnet from any of the base subnets.
- The service IP addresses can all be in the same or different subnets.
- The base addresses and the corresponding persistent and service addresses must have the same subnet mask.
- The persistent IP address can be in the same or a different subnet from the service IP addresses.
- If you use separate subnets for the persistent and the service IP addresses, be aware of the TCP/IP routing issues.
- If you have a single adapter network configuration, both the base and service addresses are allowed on the same subnet.

To clarify this concept, here is our sample cluster network configuration:

- We use one IP network with two virtual Ethernet network interfaces per node. We use two service IP addresses (one for each application) from the company data room network, which is 172.16.20.0/23. We know that our service IP address for application1 is 172.16.21.67, and for application2 is 172.16.21.72. The netmask is 255.255.254.0 and the default gateway is 172.16.20.1.
- Now, we need two persistent addresses. They must be in the same subnet with the service IP. Our persistent addresses are 172.16.21.39 and 172.16.21.47. The corresponding netmask is 255.255.254.0.
- We have two Ethernet interfaces on each node as a base adapter. We selected two subnets that are not used elsewhere in the company network and the routers do not forward them. These networks are 10.1.0.0/23 and 10.10.0.0/23. The first pair of interfaces (one from each node) uses the 10.1.1.39 and 10.1.1.47 addresses. The other two interfaces use the 10.10.1.39 and 10.10.1.47 addresses. Because our company data room network has a 255.255.254.0 subnet mask, we use the same subnet mask on the base adapters, too.

### 3.5.11 Multicast heartbeat

PowerHA uses multicast addressing for heartbeating. Therefore, the network infrastructure must correctly handle the multicast traffic (check the multicast configuration with your network administrator):

- Enable multicast traffic on all switches that are used by the cluster nodes.
- Check the available multicast traffic IP address allocation.
- Ensure that the multicast traffic is correctly forwarded by the network infrastructure (firewalls and routers) between the cluster nodes.

The multicast address is generated automatically the first time that the cluster is synchronized. The CAA changes the first octet of the base address to 228 to create the multicast address:

```
base address: x.y.z.t
multicast address: 228.y.z.t
```

Alternatively, you can specify your multicast address during the cluster installation. For more information about the multicast heartbeat, see 2.5, “CAA multicast” on page 68.

## 3.5.12 Netmon.cf file

There are changes in PowerHA SystemMirror 7.1.1 on the use of the `/usr/es/sbin/cluster/netmon.cf` file:

- If you plan to use a network with one physical adapter (not virtual) per node, you do not need to use the `netmon.cf` file any more.
- If you use virtual Ethernet in your cluster, you must configure `netmon.cf`.

## 3.5.13 Netmon.cf file configuration for virtual Ethernet environment

> **Important:** This configuration requires APAR IV14422.

When the cluster nodes have virtual Ethernet adapters, the correct configuration of the `/usr/es/sbin/cluster/netmon.cf` file is required for PowerHA to monitor the status of the network outside of the virtual switch.

In the virtual Ethernet configuration, the LPARs are connected to a virtual switch that hides the underlying physical adapters and VIOS components from the client LPARs. When the VIOS network cables are unplugged, the virtual switch remains intact, up, and running inside the frame. The LPAR clients cannot recognize the external network failure without knowing the entire network topology.

If something happens to the VIOS physical network interfaces or SEAs, the Virtual IO clients are not notified; therefore, PowerHA does not know the status of the external network. For example, when all network cables are removed from the VIOS, PowerHA is not notified of the network failure and does not run any event.

## 3.5.14 Firewalls

Do not put firewalls between the cluster nodes. If you use a firewall, you must open the firewall for the persistent and service addresses. The firewall must handle the IP address changes and relocations in the cluster.

Optionally, you can define a permanent source address of the outgoing traffic. For more details, see 3.9.2, “Configuring the service IP address distribution preference” on page 123. In this case, when the cluster is up and running, all outbound TCP/IP packets have the predefined source address.

### Changing PowerHA networks

To change network behavior, you can use **`smitty sysmirror`** command and go through **Cluster Nodes and Networks** → **Manage Networks and Network Interfaces** → **Networks** → **Change/Show a Network**. Then, select the network for which you want to change the Network attribute configuration, as shown in Figure 3-11.

```
                            Change/Show a Network

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

                                                          [Entry Fields]
* Network Name                                          net_ether_01
  New Network Name                                      []
* Network Type                                          [ether] +
* Netmask(IPv4)/Prefix Length(IPv6)                     [255.255.254.0]
* Network attribute                                     public +

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

*Figure 3-11 Cluster network configuration*

# 3.6 Disk considerations

For data disks, there is no practical limitation in PowerHA SystemMirror.

## 3.6.1 Repository disk

PowerHA stores the CAA cluster configuration on a designated disk, which is called the *repository disk*. This disk is shared between the cluster nodes; therefore, it must support concurrent access. The repository disk size must be 512 MB - 460 GB. For a 2-node cluster, about 1 GB of disk space is sufficient.

> **Important:** The repository disk cannot be mirrored by using AIX Logical Volume Manager (LVM) mirror. Ensure that the repository disk is a hardware RAID-1 or RAID-5 protected logical unit number (LUN).

PowerHA SystemMirror 7.1.1 now supports EMC, Hitachi, and HP storage for repository disks. For the supported storage subsystems, see the PowerHA hardware support matrix:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105638

# 3.7 PowerHA SystemMirror installation and prerequisites

In this section, we perform the actual installation steps for our sample PowerHA cluster. Our sample cluster has fully virtualized nodes; however, the same procedure applies to a cluster with physical Ethernet and FC adapters.

We provide a detailed step-by-step configuration that is easy to follow. We advise you to double-check every step to avoid any configuration errors.

## 3.7.1 Our sample cluster

Figure 3-12 on page 92 displays the detailed configuration for our sample cluster.

### Application and resource group configuration

We have two applications, app1 and app2, in a mutual takeover configuration. Each application has a service IP address (instsvc1 and instsvc2), a volume group (app1vg, app2vg), and a file system (`/data1` and `/data2`). The cluster consists of two resource groups, one for each application. The resource groups are named inst1rg and inst2rg. The resource groups contain the application, related volume groups, file systems, and service IP addresses.

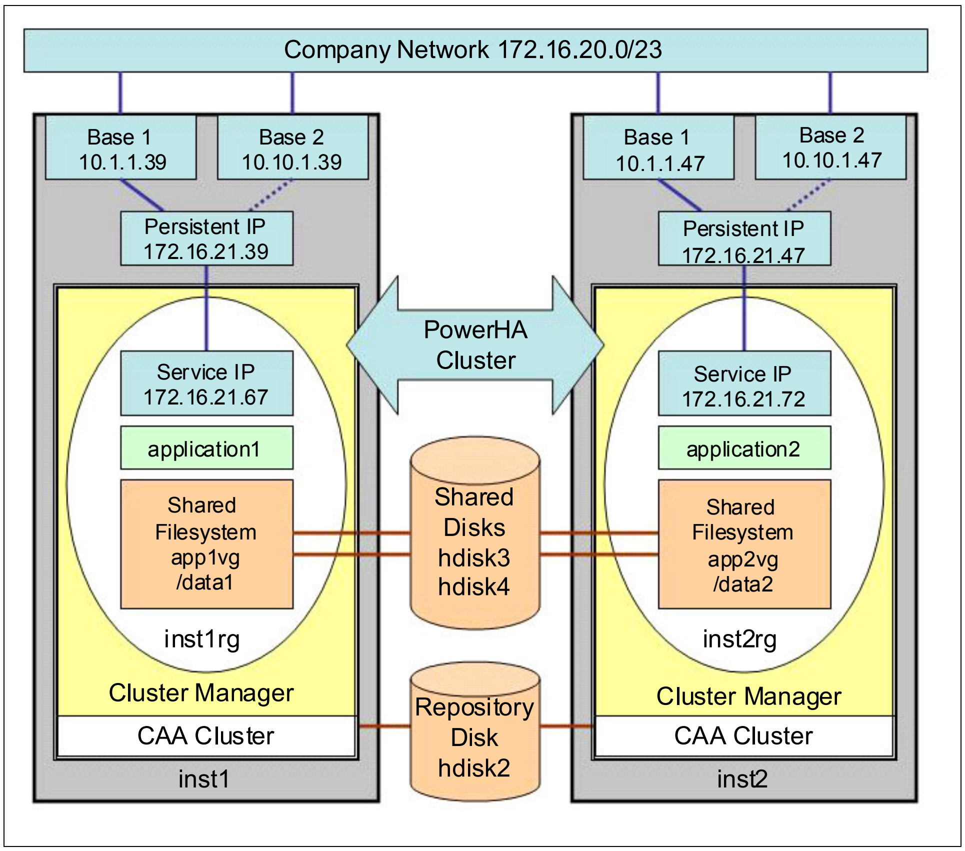



*Figure 3-12 Our sample cluster configuration*

### Network configuration

Our main service network is 172.16.20.0/23 and the default gateway is 172.16.20.1. This network is used for application client communication. We configure a persistent and service IP for each node from this subnet. The persistent IP label is the hostname.

The base IP ranges are 10.1.0.0/23 and 10.10.0.0/23. These subnets are not used anywhere else in our company network; the routers do not forward them.

Table 3-1 summarizes the service network IP configuration for our sample cluster. We use the default `net_ether_01` name for this service network in the cluster configuration.

*Table 3-1 IP addresses that are used by our sample cluster*

| Function | Node1 IP address | Node1 IP label | Node2 IP address | Node2 IP label |
|---|---|---|---|---|
| service | 172.16.21.67 | instsvc1 | 172.16.21.72 | instsvc2 |
| persistent | 172.16.21.39 | **inst1** | 172.16.21.47 | **inst2** |
| base 1 | 10.1.1.39 | inst1b1 | 10.1.1.47 | inst2b1 |
| base 2 | 10.10.1.39 | inst1b2 | 10.10.1.47 | inst2b2 |

We use the persistent IP labels for the hostname and node name, which are bold in the table.

We have an additional network in our servers, which is used for accessing the Tivoli Storage Manager backup servers only. The network is configured in the PowerHA topology as `net_ether_02`. However, there is no service address for that network; therefore, we do not configure this network as part of a resource group. PowerHA does not protect this network. Table 3-2 shows the backup network IP settings.

*Table 3-2 Backup network IP configuration*

| Function | Node1 IP address | Node 1 IP label | Node2 IP address | Node2 IP label |
|---|---|---|---|---|
| base | 10.1.2.39 | inst1backup | 10.1.2.47 | inst2backup |

We advise you to create a similar network configuration table for your cluster.

### Hardware configuration

We use two IBM Power 750 servers. We have two VIOSs on each system. The VIOSs are connected to our Ethernet network by two adapter cards. For SAN connectivity, we use two Fibre Channel (FC) adapters per VIOS. All LPARs are fully virtualized with virtual Ethernet and FC adapters.

### Virtual I/O Server configuration

Our servers have the following VIOS configuration:

- Two VIOSs on each server
- SEA failover for all Ethernet interfaces
- Virtual SCSI disks for rootvg
- NPIV port virtualization for SAN connectivity

For the detailed VIOS configuration, see 3.4.2, “Our sample cluster virtual environment” on page 82.

## 3.7.2 Starting point for the installation

For the start of the PowerHA SystemMirror installation, we assume that you have the following environment:

- Two servers
- Fully working and tested redundant VIOS configuration
- Ethernet connectivity and outside network access that works well
- Storage access, including SAN zones and LUN mapping
- Two installed AIX 7.1 LPARs
- No IP that is configured on the LPARs

- No default gateway that is set
- Mirrored rootvg

We advise you to hardcode the host name into the PS1 environment variable (Example 3-1) for the root and other administrative users so the system administrators always see from the prompt on which cluster node they are signed up. We know of several cases when the system administrator stopped the wrong cluster node.

*Example 3-1 Setting up PS1 environment variable for the root user*

```
inst1:/# cat .profile
set -o vi
export PS1='inst1:$PWD#'
```

## 3.7.3 Base adapter network setup

First, we must configure the network settings for the base adapters. You have to log on through the Hardware Management Console (HMC) virtual terminal session because we are going to change the IP addresses. Identify your network adapters, and then use **smit mktcpip** to configure the two base adapters on both cluster nodes. Select the interface that you want and enter the TCP/IP parameters. *Do not set your default gateway yet.* See Figure 3-13 to set up the base address.

```
 Minimum Configuration & Startup

 To Delete existing configuration data, please use Further Configuration menus

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

                                                        [Entry Fields]
* HOSTNAME                                           [inst1]
* Internet ADDRESS (dotted decimal)                  [10.1.1.39]
  Network MASK (dotted decimal)                      [255.255.254.0]
* Network INTERFACE                                   en0
  NAMESERVER
           Internet ADDRESS (dotted decimal)         []
           DOMAIN Name                               []
  Default Gateway
       Address (dotted decimal or symbolic name)     []
       Cost                                          []                       #
       Do Active Dead Gateway Detection?              no                      +
  Your CABLE Type                                     N/A                     +
  START Now                                           no                      +

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

*Figure 3-13 Setting up the base interface*

## 3.7.4 Persistent IP and default gateway setup

To log on to the cluster nodes through the company network, we must configure the persistent addresses on the LPARs. Enter **smit inetalias**. Click **Add an IPV4 Network Alias**. Select one of the base interfaces, and enter the persistent IP address and the corresponding netmask. Figure 3-14 on page 95 shows how to configure an IPV4 alias for a persistent address.

```
 Add an IPV4 Network Alias

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

                                                        [Entry Fields]
  Network INTERFACE                                   en0

* IPV4 ADDRESS (dotted decimal)                      [172.16.21.39]
  Network MASK (hexadecimal or dotted decimal)       [255.255.255.0]

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

*Figure 3-14 Configuring the IP alias for the persistent address*

Now, configure your default gateway. Start **smitty mktcpip**, select any of the base interfaces, and fill out the default gateway entry field (see Figure 3-15). Set the persistent address IP label to your hostname.

```
 Minimum Configuration & Startup

 To Delete existing configuration data, please use Further Configuration menus

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

                                                         [Entry Fields]
* HOSTNAME                                            [inst1]
* Internet ADDRESS (dotted decimal)                   [10.1.1.39]
  Network MASK (dotted decimal)                       [255.255.254.0]
* Network INTERFACE                                    en0
  NAMESERVER
           Internet ADDRESS (dotted decimal)          []
           DOMAIN Name                                []
  Default Gateway
       Address (dotted decimal or symbolic name)      [172.16.20.1]
       Cost                                           []                      #
       Do Active Dead Gateway Detection?               no                     +
  Your CABLE Type                                      N/A                    +
  START Now                                            no                     +

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

*Figure 3-15 Configuring the default gateway*

Log on to your cluster nodes from the company network.

### 3.7.5 Setting up the hostname

CAA has strict rules for hostname conventions.

**Important:** Unlike in earlier versions of PowerHA, the hostname must not be an alias in the `/etc/hosts` file. The name resolution for the hostname must work for both ways, therefore a limited set of characters can be used. The IP address that belongs to the hostname must be reachable on the server:

- The hostname cannot be a service address.
- The hostname cannot be an alias in the `/etc/hosts` file.
- The hostname cannot be an address located on a network which is defined as private in PowerHA.
- The hostname, the CAA node name, and the "communication path to a node" must be the same.
- By default, the PowerHA, nodename, the CAA nodename, and the "communication path to a node" are set to the same name (see 3.8.3, "Initial cluster setup" on page 114).
- The hostname of the cluster nodes cannot change after the cluster is configured.

We advise you to use one of the base addresses or the persistent address as the hostname. You can use the persistent address as the hostname only if you set up the persistent alias manually before you configure the cluster topology.

We use the persistent IP labels for hostname and node name, inst1 and inst2.

## 3.7.6 Preparing the /etc/hosts file and changing the name resolution order

This area is the most problematic part of the cluster configuration. If your `/etc/hosts` file is incorrect, you cannot define your PowerHA cluster. Now, you know the cluster topology, networks, and all IP labels. During the node discovery, PowerHA reads the `/etc/hosts` file and compiles a list of IP labels. You must prepare the `/etc/hosts` file so that you can select the correct IP labels with the SMIT dialogs during the cluster configuration.

**Important:**

- All IP labels that are used by the PowerHA cluster must be resolvable.
- Do not use aliases in the `/etc/hosts` file because of CAA considerations.
- The cluster IP addresses and IP labels must be the same on all cluster nodes.

Example 3-2 displays the `/etc/hosts` file on our PowerHA cluster. Our hostnames are *inst1* and *inst2*, and they are configured as the persistent addresses.

*Example 3-2 /etc/hosts file for the cluster*

```
#Base IP addresses (10.1.0.0/23)
10.1.1.39       inst1b1
10.1.1.47       inst2b1

#Base IP addresses (10.10.0.0/23)
10.10.1.39      inst1b2
10.10.1.47      inst2b2

#Persistent IP Addresses (172.16.20.0/23)
172.16.21.39    inst1
172.16.21.47    inst2
```

```
#Service IP Addresses (172.16.20.0/23)
172.16.21.67    instsvc1
172.16.21.72    instsvc2

#Backup network IP addresses (10.1.2.0/24)
10.1.2.39       inst1backup
10.1.2.47       inst2backup
```

If you use a nameserver, we strongly advise that you change the name resolution order in the `/etc/netsvc.conf` file:

```
hosts=local,bind
```

Therefore, AIX looks for an IP label/IP address in the local `/etc/hosts` file first, and then in the nameserver that is defined in `/etc/resolv.conf` file.

## 3.7.7 Check your network settings

Before you continue, double-check the network settings. From our experience, the wrong network configuration is the most common source of PowerHA configuration problems. Example 3-3 shows how to check your network settings.

*Example 3-3 Checking the network settings*

```
root@inst2 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.98.6e.92.6f   374658     0    53425     0     0
en0   1500  10.1        inst1b1            374658     0    53425     0     0
en0   1500  172.16.20   inst1              374658     0    53425     0     0
en1   1500  link#3      2e.47.98.6e.92.70    10846     0      129     0     0
en1   1500  10.1.2      inst1backup         10846     0      129     0     0
en2   1500  link#4      2e.47.98.6e.92.71   310118     0       47     0     0
en2   1500  10.10       inst1b2            310118     0       47     0     0
lo0   16896 link#1                          76325     0    76325     0     0
lo0   16896 127         loopback            76325     0    76325     0     0
lo0   16896 loopback                        76325     0    76325     0     0

root@inst2 / # ifconfig -a
en0:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT
,CHECKSUM_OFFLOAD(ACTIVE),CHAIN>
        inet 10.1.1.47 netmask 0xfffffe00 broadcast 10.1.1.255
        inet 172.16.21.47 netmask 0xfffffe00 broadcast 172.16.21.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
en1:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT
,CHECKSUM_OFFLOAD(ACTIVE),CHAIN>
        inet 10.1.2.47 netmask 0xffffff00 broadcast 10.1.2.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
en2:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT
,CHECKSUM_OFFLOAD(ACTIVE),CHAIN>
        inet 10.10.1.47 netmask 0xfffffe00 broadcast 10.10.1.255
         tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1
```

```
lo0:
flags=e08084b,c0<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,LAR
GESEND,CHAIN>
        inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
        inet6 ::1%1/0
         tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1

root@inst2 / # netstat -rn|grep default
default            172.16.20.1       UG        1       562 en0      -      -
```

Ensure that the network connectivity works:

- Log on to the nodes from the company network by using the persistent address.
- Ping the base adapter pairs.
- Ping or access the default gateway.
- Ensure that name resolution works for the hostname, all IP labels, and addresses that are used in the cluster configuration.

## 3.7.8 Time synchronization

We suggest that you use time synchronization on the cluster nodes. Some applications require time synchronization, but by running **xntpd**, the cluster administration and debugging are easier.

### Configuring NTP clients

If your network already has an established time server, you can set up the cluster nodes to get the accurate time information from it:

1. Modify the /etc/ntp.conf file to contain these lines:

   ```
   server 192.169.1.254 # your ntp time server's IP address goes here
   driftfile /etc/ntp.drift
   tracefile /etc/ntp.trace
   ```

2. Start the xntpd daemon: **startsrc -s xntpd**
3. Set up xntpd to start automatically at boot time. Remove the comment mark from the beginning of the line in the /etc/rc.tcpip file:

   ```
   start /usr/sbin/xntpd "$src_running"
   ```

### Setting up a time server

If you do not have a time server, you can easily set up a time server in the cluster. Without external time information, the time of the servers is not accurate, but identical. Follow these steps:

1. Choose one of the nodes to act as a Network Time Protocol (NTP) server, preferably the node with the highest priority in the cluster.
2. Modify the /etc/ntp.conf file on the NTP server:

   ```
   disable auth
   server 127.127.1.1 prefer # use the local clock as preferred
   fudge 127.127.1.1 stratum 4
   driftfile /etc/ntp.drift
   ```

3. Edit the /etc/ntp.conf file on the other nodes:

   ```
   server 10.10.1.1 # your cluster's ntp time server's base IP address goes here
   driftfile /etc/ntp.drift
   ```

```
tracefile /etc/ntp.trace
```

We suggest that you use the NTP server node base address or a persistent interface, if any.

4. Edit the `/etc/rc.tcpip` file on all nodes so that xntpd starts automatically:

   ```
   start /usr/sbin/xntpd "$src_running"
   ```

5. Start xntpd on all nodes:

   ```
   startsrc -s xntpd
   ```

## 3.7.9 Installing the AIX prerequisites

Install the following filesets from the AIX media:

```
bos.adt
bos.ahafs
bos.cluster
bos.clvm
bos.net.tcp.client
bos.net.tcp.server
devices.common.IBM.storfwork.rte
rsct.basic.rte
rsct.compat.basic.hacmp
rsct.compat.client.hacmp
```

If you plan to use Network File System (NFS), you must install the NFS client and server packages, too:

```
bos.net.nfs.client
bos.net.nfs.server
```

For the optional message authentication and encryption communication between the cluster nodes, install one of the following packages from the AIX Expansion Pack media:

- `rsct.crypt.des` - For Data Encryption Standard (DES)
- `rsct.crypt.3des` - For Triple DES
- `rsct.crypt.aes256` - For Advanced Encryption Standard (AES)

## 3.7.10 Installing the latest AIX fixes

This point is the best time to install the latest AIX fixes. You can download them from IBM FixCentral:

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

Also, remember to download the latest PowerHA fixes. We install them later. After you apply the AIX fixes, reboot the servers.

> **Important:** If you plan to use virtual Ethernet adapters in the cluster, install APAR IV14422.

## 3.7.11 Setting up the required SAN zones for SAN-based heartbeat

We must set up an additional SAN zoning that is required for the SAN-based heartbeat. We need to add a zone that contains all SAN adapters of the cluster nodes in the fabric so that they can access each other.

**Important:**

- If you use physical FC adapters, you must zone together that cards' worldwide name (WWN).
- If you use virtual FC adapters, you must to zone together the physical adapters' WWN from the VIOSs.

Example 3-4 shows how to query the WWNs of the FC cards. In this case, the cards are physical FC adapters on a cluster node.

*Example 3-4 Listing the WWN addresses of the FC cards*

```
root@inst1 / # lscfg -vpl fcs0|grep Network
        Network Address.............1050760506190014
root@inst1 / # lscfg -vpl fcs1|grep Network
        Network Address.............1050760506190015
-------------------------------------------------------------
root@inst2 / # lscfg -vpl fcs0|grep Network
        Network Address.............10507605061A0078
root@inst2 / # lscfg -vpl fcs1|grep Network
        Network Address.............10507605061A0079
```

If you use virtual FC adapters, you have to use the physical adapters on the VIOS. Example 3-5 shows how to display the WWN addresses of the FC cards on the VIOSs.

*Example 3-5 Displaying the WWN addresses of the FC cards on the VIOSs*

```
$ lsdev -dev fcs0 -vpd | grep Network
        Network Address.............10000000C9E331E8
$ lsdev -vpd -dev fcs1 |grep Network
        Network Address.............10000000C9E331E9
```

Example 3-6 illustrates the required zones on the SAN switches for a 2-physical card configuration.

*Example 3-6 The required SAN zones*

```
Fabric1:
zone: inst1fcs0_inst2fcs0
         C0:50:76:05:06:19:00:14
         C0:50:76:05:06:1A:00:78
-------------------------------------------------------------
Fabric2:
zone: inst1fcs1_inst2fcs1
         C0:50:76:05:06:19:00:15
         C0:50:76:05:06:1A:00:79
```

## 3.7.12 Configuring the physical FC adapters for SAN-based heartbeat

We describe how to configure SAN-based heartbeat for physical FC adapters. If you use virtual FC adapters (NPIV), see 3.7.13, "Configuring SAN heartbeating in a virtual environment" on page 103.

You can configure SAN-based heartbeat only if you have one of the supported FC adapters. For more information, see the "Setting up cluster storage communication" page:

http://publib.boulder.ibm.com/infocenter/aix/v7r1/index.jsp?topic=/com.ibm.aix.clusteraware/claware_comm_setup.htm

To correctly configure the FC adapters for the cluster SAN-based communication, follow these steps:

> **X in fcsX:** In the following steps, the X in `fcsX` represents the number of the FC adapters. You must complete this procedure for each FC adapter that is involved in the cluster SAN-based communication.

1. Change the `tme` attribute value to `yes` in the `fcsX` definition:

   ```
   chdev -P -l fcsX -a tme=yes
   ```

2. Enable the dynamic tracking, and the fast-fail error recovery policy on the corresponding `fscsiX` device:

   ```
   chdev -P -l fscsiX -a dyntrk=yes -a fc_err_recov=fast_fail
   ```

3. Restart the server.
4. Verify the configuration changes by running the following commands:

   ```
   lsdev -C | grep -e fcsX -e sfwcommX
   lsattr -El fcsX | grep tme
   lsattr -El fscsiX | grep -e dyntrk -e fc_err_recov
   ```

Example 3-7 illustrates the procedure for port `fcs0` on node `inst1`.

*Example 3-7 SAN-based communication channel setup*

```
inst1:/ # lsdev -l fcs0
fcs0 Available 00-00 8Gb PCI Express Dual Port FC Adapter (df1000f114108a03)

sinst1:/ # lsattr -El fcs0|grep tme
tme           no         Target Mode Enabled                              True

inst1:/ # rmdev -Rl fcs0
fcnet1 Defined
sfwcomm0 Defined
fscsi0 Defined
fcs0 Defined

inst1:/ # chdev -l fcs0 -a tme=yes
fcs0 changed

inst1:/ # chdev -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail
fscsi0 changed

inst1:/ # cfgmgr -l fcs0;cfgmgr -l sfwcomm0

inst1:/ # lsdev -C|grep -e fcs0 -e sfwcomm0
fcs0       Available 01-00       8Gb PCI Express Dual Port FC Adapter
(df1000f114108a03)
sfwcomm0   Available 01-00-02-FF Fibre Channel Storage Framework Comm
```

```
inst1:/ # lsattr -El fcs0|grep tme
tme           yes        Target Mode Enabled True

inst1:/ # lsattr -El fscsi0|grep -e dyntrk -e fc_err_recov
dyntrk       yes       Dynamic Tracking of FC Devices        True
fc_err_recov fast_fail FC Fabric Event Error RECOVERY Policy True
```

No other configuration is required in PowerHA. When your cluster is up and running, you can check the status of the SAN heartbeat with the **lscluster -i sfwcom** command. See Example 3-10 on page 105.

### 3.7.13 Configuring SAN heartbeating in a virtual environment

If you do not have a physical FC adapter that is assigned to the cluster LPARs, you still can use SAN heartbeating through the VIOSs. This solution does not require virtual FC adapters on the client LPARS, virtual SCSI clients can use it, too. Figure 3-16 on page 104 shows the SAN heartbeat configuration and data flow. For clarity, we omitted the second VIOS and the redundant VIOS components from both frames.

The client LPARs cannot use the FC adapters of the VIOSs directly for heartbeating. A number of forward devices (sfwcommX) and a special virtual Ethernet connection are needed between the LPARs and the VIOS. This connection requires a dedicated virtual Ethernet adapter with VLAN ID 3358 on the VIOSs and all cluster LPARs. The PowerHA underlying CAA cluster uses the sfwcomm VLAN Storage Framework Communication device to send the heartbeats. The heartbeats go through the virtual Ethernet switch. Then, the VIOS’s sfwcomm device forwards it to the physical FC adapter and to the SAN network. To bypass AIX and the VIOS TCP/IP subsystem, sfwcomm devices use the Ethernet data link layer communication.

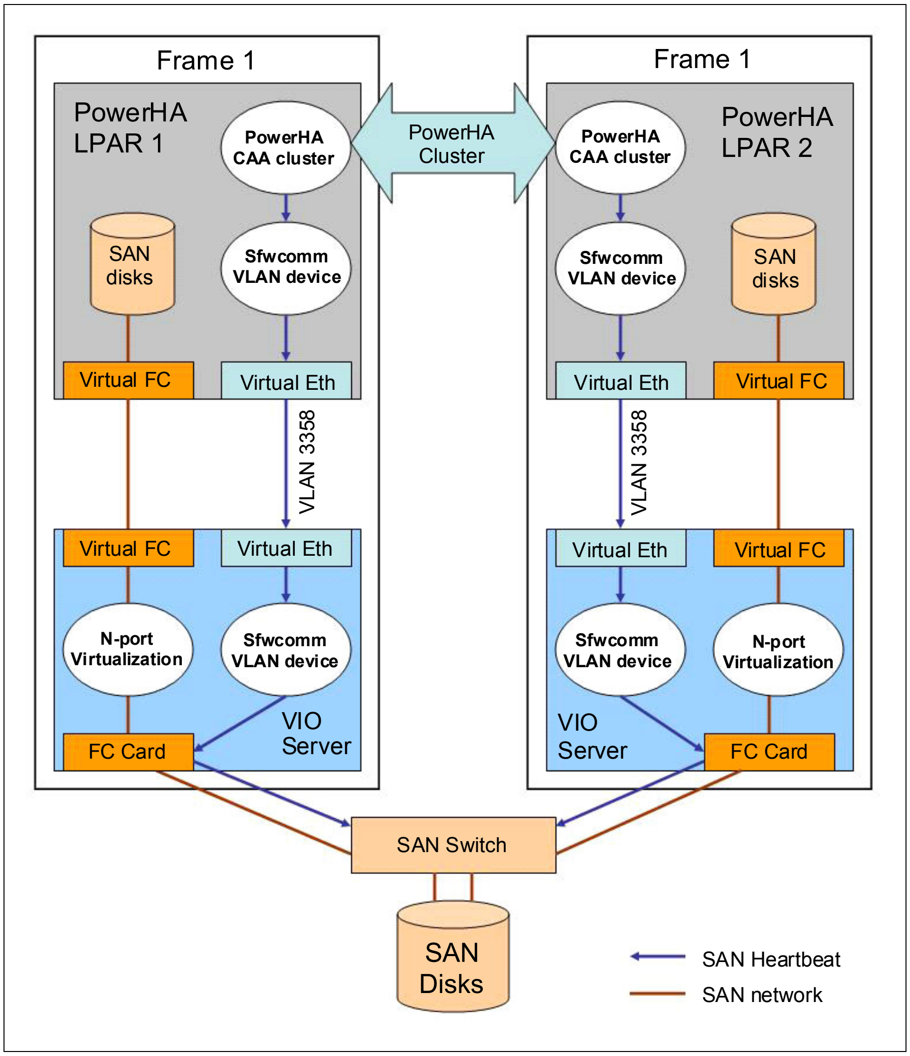



*Figure 3-16 SAN heartbeating in a virtual environment*

**Important:** To configure SAN-based heartbeating in a virtual environment, VIOS level 2.2.0.11-FP24SP01 or newer is required. You can download the latest version of VIOS from IBM Fix Central:

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

Follow these configuration steps:

1. Check that the FC adapter in the VIOS supports target mode. See the “Setting up cluster storage communication” page for the supported FC adapters:

   http://publib.boulder.ibm.com/infocenter/aix/v7r1/index.jsp?topic=/com.ibm.aix.clusteraware/claware_comm_setup.htm

2. Configure the FC adapters for SAN heartbeating on the VIOSs:

   ```
   chdev -dev fcs0 -attr tme=yes -perm
   ```

   Repeat this step for all FC cards.

3. Reboot the VIO server.

**Important:** Repeat step 2 and 3 on all VIOSs serving the cluster LPARs.

4. On the HMC, create a new virtual Ethernet adapter for each cluster LPAR and VIOS. Set the VLAN ID to 3358. Do not put another VLAN ID or any other traffic to this network interface. Save the LPAR profile. For further information about how to create a virtual Ethernet interface for an LPAR on the HMC, see *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.
5. You do not have to set up any IP address on the new virtual Ethernet interface.
6. On the VIOS, run **cfgdev**, and then check for the new Ethernet and sfwcomm devices with the **lsdev** command. See Example 3-8.

*Example 3-8 VLAN Storage Framework Communication devices in the VIOS*

```
$ lsdev|grep sfw
sfw0            Available   Storage Framework Module
sfwcomm0        Available   Fibre Channel Storage Framework Comm
sfwcomm1        Available   Fibre Channel Storage Framework Comm
sfwcomm2        Available   SAS Storage Framework Comm
sfwcomm3        Available   vLAN Storage Framework Comm
```

7. On the cluster nodes, run **cfgmgr**, and check for the virtual Ethernet and sfwcomm devices with the **lsdev** command (Example 3-9).

*Example 3-9 VLAN storage framework communication devices in a cluster LPAR*

```
root@inst1 / # lsdev -C|grep sfw
sfw0           Available             Storage Framework Module
sfwcomm0       Available 81-T1-01-FF Fibre Channel Storage Framework Comm
sfwcomm1       Available 82-T1-01-FF Fibre Channel Storage Framework Comm
sfwcomm2       Available             vLAN Storage Framework Comm
```

8. No other configuration is required in PowerHA. When your cluster is up and running, you can check the status of the SAN heartbeat with the **lscluster -i sfwcom** command (Example 3-10).

*Example 3-10 Checking the status of the SAN-based heartbeat*

```
root@inst1 / # lscluster -i sfwcom
Network/Storage Interface Query

Cluster Name:  inst1_cluster
Cluster uuid:  0a21480a-7502-11e1-ad51-2e47986e9270
Number of nodes reporting = 1
Number of nodes expected = 1
Node inst1
Node uuid = 09d7df1c-7502-11e1-ad51-2e47986e9270
Number of interfaces discovered = 1
        Interface number 1 sfwcom
                ifnet type = 0 ndd type = 304
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 0
                Mean Deviation in network rrt across interface = 0
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x0
```

```
ndd flags for interface = 0x9
Interface state  UP
```

## 3.7.14 Shared storage

Now, we explain the process of creating an enhanced concurrent volume group for the application data.

### Configuring the FC adapters

If dynamic tracking and the fast-fail error recovery policy are not enabled, enable the dynamic tracking and the fast-fail error recovery policy on the fscsi devices, and then, restart the server. This task applies to virtual FC adapters, too.

```
chdev -P -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail
```

### Disabling the SCSI reservation

**Important:** You must disable the SCSI reservation locks for all shared cluster disks.

### Disabling the SCSI locks on the data disks on the storage subsystem

You must disable the SCSI locks for the data disks. If you use physical or virtual FC adapters to connect to the storage subsystem, you have to check this feature on the storage. However, on most storage systems, if you set the host type correctly, you do not have to do anything. For more information, see the storage documentation.

### Disabling the SCSI reservation locks on the VIOS

If your cluster uses data disks through a virtual SCSI adapter from a VIOS, disable the SCSI locks on the VIOS:

1. Log on to the VIOS and check the reserve policy for the hdisk that is used for the virtual SCSI backing device for the cluster nodes. Example 3-11 shows how to check the reserve policy for an hdisk.

*Example 3-11 Checking the reserve policy for virtual SCSI backing disks*

```
$ lsdev -dev hdisk1 -attr
attribute       value                                      description
user_settable
PCM             PCM/friend/scsiscsd                        Path Control Module          False
algorithm       fail_over                                  Algorithm                    True
dist_err_pcnt   0                                          Distributed Error Percentage True
dist_tw_width   50                                         Distributed Error Sample Time True
hcheck_interval 0                                          Health Check Interval        True
hcheck_mode     nonactive                                  Health Check Mode            True
max_transfer    0x100000                                   Maximum TRANSFER Size        True
pvid            00f61ab28892d2a50000000000000000           Physical volume identifier   False
queue_depth     16                                         Queue DEPTH                  True
reserve_policy  single_path                                Reserve Policy               True
size_in_mb      146800                                     Size in Megabytes            False
unique_id       2A1135000C5002399FC0B0BST9146852SS03IBMsas Unique device identifier     False
ww_id           5000c5002399fc0b                           World Wide Identifier        False
```

2. Change the reserve policy for the virtual SCSI backing device:

   ```
   chdev -dev hdisk1 -attr reserve_policy=no_reserve
   ```

3. Repeat this step for all virtual SCSI backing disk devices on all VIOSs.

### Disabling SCSI locks for the data disks on the cluster nodes

You must check the *reserve_policy* for all your cluster data and repository disks and change it to *no_reserve,* if necessary. Example 3-12 shows how to change the reserve policy to *no_reserve*.

*Example 3-12 Checking and changing the reserve policy for a cluster data disk*

```
root@inst1 / # lsattr -El hdisk4|grep reserve
reserve_policy  single_path      Reserve Policy             True

root@inst1 / # chdev -a reserve_policy=no_reserve -l hdisk4
```

### Configuring the hdisk Physical Volume ID on all cluster nodes

We advise you to set up the Physical Volume ID (PVID) on the cluster data and repository disks on all cluster nodes by using the commands that are shown in Example 3-13. This step speeds up the shared disk configuration process later because PowerHA identifies the disks by the PVID.

*Example 3-13 Setting up the PVID for the cluster data and repository disks (first node)*

```
root@inst1 / # lspv
hdisk0          00f74d45f95e5beb                    rootvg          active
hdisk1          00f74d45fa932b47                    rootvg          active
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None

root@inst1 / # chdev -l hdisk2 -a pv=yes
hdisk2 changed
root@inst1 / # chdev -l hdisk3 -a pv=yes
hdisk3 changed
root@inst1 / # chdev -l hdisk4 -a pv=yes
hdisk4 changed

root@inst1 / # lspv
hdisk0          00f74d45f95e5beb                    rootvg          active
hdisk1          00f74d45fa932b47                    rootvg          active
hdisk2          00f74d45367ec638                    None
hdisk3          00f74d45367ed24c                    None
hdisk4          00f74d45367edf3b                    None
```

Repeat the previous steps on the other cluster node so that we have the same PVIDs for the data disks (Example 3-14).

*Example 3-14 Setting up the PVID for the cluster data and repository disks (second node)*

```
root@inst2 / # lspv
hdisk0          00f74d47f96263a2                    rootvg          active
hdisk1          00f74d47fa934fef                    rootvg          active
hdisk2          00f74d45367ec638                    None
hdisk3          00f74d45367ed24c                    None
hdisk4          00f74d45367edf3b                    None
```

## Creating data volume groups and file systems

Perform these steps on the primary node only. To create a scalable, enhanced concurrent access volume group from the data disks, use `smit mkvg`. Then, select **Add a Scalable Volume Group**. Figure 3-17 displays how to create a volume group.

```
 Add a Scalable Volume Group

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

                                                        [Entry Fields]
  VOLUME GROUP name                                   [app1vg]
  Physical partition SIZE in megabytes                                     +
* PHYSICAL VOLUME names                               [hdisk3]             +
  Force the creation of a volume group?                no                  +
  Activate volume group AUTOMATICALLY                  no                  +
    at system restart?
  Volume Group MAJOR NUMBER                           [100]                +#
  Create VG Concurrent Capable?                        enhanced concurrent +
  Max PPs per VG in units of 1024                      32                  +
  Max Logical Volumes                                  256                 +
  Enable Strict Mirror Pools                           No                  +
  Infinite Retry Option                                no                  +

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

*Figure 3-17 Creating a scalable, enhanced concurrent volume group*

Ensure that you supply a unique major number for each volume group. The available major numbers can differ on the cluster nodes. To get the list of the available major numbers, move the cursor to the Volume Group MAJOR NUMBER line and press F4 or issue the **`lvlstmajor`** command. Select a number that is available on all cluster nodes. A good starting number can be 100.

The major number represents the pointer to the device in the kernel. Normally, you are not concerned with this number, but PowerHA forces you to use the same major number for all volume groups in the cluster configuration.

Example 3-15 shows how to check a volume group major number.

*Example 3-15 Checking a volume group major number*

```
root@inst1 / # ls -al /dev/app1vg
crw-rw----    1 root     system      100,  0 Mar 21 14:49 /dev/app1vg
```

## Creating logical volumes and file systems

You have to issue the **`varyonvg app1vg`** command manually before you can start to create the logical volumes and file systems. Do not vary on the same volume group on the backup node. During logical volume and file system creation, ensure that you use unique names for the LVM components. If you plan to use separate journaled file system (JFS)/JFS2 log volumes, all volume groups must have their own log volume. We advise you to use an inline log for the JFS2 file systems. When you create a file system set “Mount AUTOMATICALLY at system restart?” to `no`.

Alternatively, you can add your file systems later when the cluster is up and running by using **smit cspoc**. For information about how to manage the shared LVM components from PowerHA Cluster Single Point of Control (C-SPOC), see 4.15, "C-SPOC cluster user and group management" on page 202.

### Importing the volume groups to the backup node

After you create the necessary logical volumes and file systems, you must import the volume group to the backup node:

1. Unmount the file systems and vary off the data volume groups on the primary node:

   ```
   varyoffvg app1vg
   ```

2. Identify the physical volume IDs of the data volume on both cluster nodes. See Example 3-16 and Example 3-17.

   *Example 3-16 Identify the PVIDs of the app1vg volume group on the primary node*

   ```
   inst1:/# lspv
   hdisk0          00f74d45f95e5beb                    rootvg          active
   hdisk1          00f74d45fa932b47                    rootvg          active
   hdisk2          00f74d45367ec638                    None
   hdisk3          00f74d45367ed24c                    app1vg
   hdisk4          00f74d45367edf3b                    app2vg
   ```

   *Example 3-17 Identify the PVID of the data volume group on the backup node*

   ```
   root@inst2 / # lspv
   hdisk0          00f74d47f96263a2                    rootvg          active
   hdisk1          00f74d47fa934fef                    rootvg          active
   hdisk2          00f74d45367ec638                    None
   hdisk3          00f74d45367ed24c                    None
   hdisk4          00f74d45367edf3b                    None
   ```

3. Import the volume group to the backup site: **importvg -V major_number -y volume_group_name -c physical_volume**. Example 3-18 shows how to use the **importvg** command.

   *Example 3-18 Importing a data volume group to the secondary node*

   ```
   root@inst2 / # importvg -V 100 -y app1vg -c hdisk3
   app1vg
   0516-783 importvg: This imported volume group is concurrent capable.
           Therefore, the volume group must be varied on manually.
   ```

   Ensure that you use the same major number that you use on the primary node.

#### *Converting an existing volume group to an enhanced concurrent volume group*

If data volume groups are already created, you can change them to enhanced concurrent capable: run **smit chvg**, then press F4, and select the volume group. Set the volume group type to *enhanced concurrent* and disable the automatic activation. For information to change a volume group to enhanced concurrent, see Figure 3-18 on page 110.

> **Important:** PowerHA 7.1 supports shared VGs of type enhanced concurrent only (unlike previous PowerHA versions).

```
 Change a Volume Group

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

                                                        [Entry Fields]
* VOLUME GROUP name                                   app2vg
* Activate volume group AUTOMATICALLY                 no                     +
    at system restart?
* A QUORUM of disks required to keep the volume       yes                    +
   group on-line ?
  Concurrent Capable                                  enhanced concurrent    +
  Change to big VG format?                            no                     +
  Change to scalable VG format?                       no                     +
  LTG Size in kbytes                                  128                    +
  Set hotspare characteristics                        n                      +
  Set synchronization characteristics of stale        n                      +
   partitions
  Max PPs per VG in units of 1024                     32                     +
  Max Logical Volumes                                 256                    +
  Mirror Pool Strictness                                                     +
  Infinite Retry Option                               no                     +

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

*Figure 3-18 Changing an existing volume group to the enhanced concurrent type*

Now, you can import the volume group on the other node. Ensure that you use the same major number on both nodes. If the volume group has a major number that is already used on the other node, export and import the volume group on the primary node, too. Example 3-19 shows how you can synchronize the volume group major number between the cluster nodes.

*Example 3-19 How to synchronize the volume group major numbers*

```
Primary node:
root@inst1 / # ls -al /dev/app2vg
crw-rw----    1 root     system 33,  0 Mar 21 15:35 /dev/app2vg
root@inst1 / # lvlstmajor
39..99,101...

Backup node:
root@inst2 / # lvlstmajor
35..99,101...

Primary node:
inst1:/# exportvg app2vg
inst1:/# importvg -V 101 -y app2vg -c hdisk4
app2vg
0516-783 importvg: This imported volume group is concurrent capable.
        Therefore, the volume group must be varied on manually.

Backup node:
root@inst2 / # importvg -V 101 -y app2vg -c hdisk4
```

```
app2vg
0516-783 importvg: This imported volume group is concurrent capable.
        Therefore, the volume group must be varied on manually.
```

### 3.7.15 Installing PowerHA filesets

The following PowerHA filesets are on the installation media:

- `cluster.adt.es`: Clinfo and Clstat samples, include files, and a web-based monitor demo
- `cluster.doc.en_US.assist`: Smart Assist PDF documentation
- `cluster.doc.en_US.es`: PowerHA SystemMirror PDF documentation
- `cluster.es.assist`: Smart Assist filesets
- `cluster.es.cfs`: General Parallel File System (GPFS) support
- `cluster.es.client`: Cluster client binaries, libraries, and web-based Smit for PowerHA
- `cluster.es.cspoc`: Cluster Single Point of Control (C-SPOC) and the **`dsh`** command
- `cluster.es.director.agent`: PowerHA SystemMirror Director common agent services (CAS) agent
- `cluster.es.migcheck`: Migration support
- `cluster.es.nfs`: NFS server support
- `cluster.es.server`: Base cluster filesets
- `cluster.es.worksheets`: Online planning worksheets
- `cluster.hativoli`: PowerHA SystemMirror Tivoli Server and Client
- `cluster.license`: Electronic license file
- `cluster.man.en_US.es`: Man pages - US English
- `cluster.msg.Ja_JP.assist`: Smart Assist messages - Japanese
- `cluster.msg.Ja_JP.es`: Japanese message catalog
- `cluster.msg.en_US.assist`: US English Smart Assist messages
- `cluster.msg.en_US.es`: US English message catalog

Use **`smit installp`** to install the PowerHA filesets. The following filesets are the minimum required filesets for the PowerHA SystemMirror installation:

- `cluster.es.client.clcomd`
- `cluster.es.client.lib`
- `cluster.es.client.rte`
- `cluster.es.client.utils`
- `cluster.es.cspoc`
- `cluster.es.migcheck`
- `cluster.es.server`
- `cluster.license`

Remember to set “ACCEPT new license agreements?” to `yes`.

### 3.7.16 Installing PowerHA fixes

Install the latest PowerHA fixes. You can download them from IBM FixCentral:

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

## 3.8 Configuring PowerHA SystemMirror topology

The PowerHA SystemMirror topology is the skeleton of the cluster. It describes the cluster networks and network interfaces. CAA always monitors all interfaces, no matter which ones are in the HA topology.

After you complete 3.7, "PowerHA SystemMirror installation and prerequisites" on page 91, then you can proceed with the PowerHA SystemMirror configuration.

There are several ways to create and configure your cluster:

- SMIT: SMIT is the easiest and most convenient way to manage your cluster. We use this method in this section. Figure 3-19 displays the main SMIT menu for PowerHA SystemMirror (**`smit sysmirror`**).
- The clmgr command-line interface: New in PowerHA SystemMirror 7.1. You can perform all cluster management tasks with **`clmgr`**. This interface is an excellent tool to write scripts for large-scale deployments and daily cluster management tasks. See Appendix B, "Configuring the PowerHA cluster by using clmgr" on page 551 for more information about **clmgr** usage.
- PowerHA SystemMirror plug-in for IBM System Director: This System Director plug-in provides a graphical user interface for cluster management.
- Web-based SMIT: A graphical, web-based SMIT for PowerHA menus only. You have to install the `cluster.es.client.wsm` fileset; however, we do not advise you to use this method.

```
 PowerHA SystemMirror

Move cursor to desired item and press Enter.

  Cluster Nodes and Networks
  Cluster Applications and Resources

  System Management (C-SPOC)
  Problem Determination Tools
  Custom Cluster Configuration

  Can't find what you are looking for ?
  Not sure where to start ?


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 3-19 SMIT PowerHA SystemMirror main menu*

> **Important:** The PowerHA SMIT windows have a new design since the earlier version. However, some help text and man pages still refer to the old version.

Familiarize yourself with the SMIT PowerHA menus. Some of the functions are not available until you set up the corresponding cluster objects. Press F1 anytime for context-sensitive help.

## 3.8.1 Propagating the /etc/cluster/rhosts file

The first step is to create and propagate the `/etc/cluster/rhosts` file on the nodes. Unlike in earlier version of PowerHA SystemMirror, now you have to manually edit this file. Follow these steps:

1. Create the `/etc/cluster/rhosts` file and write all IP addresses or IP labels from the cluster nodes into it. See Example 3-20 for our sample cluster `rhosts` file. We insert all base, persistent, and service addresses from both nodes.

   *Example 3-20 /etc/cluster/rhosts file*

   ```
   root@inst1 / # cat /etc/cluster/rhosts
   10.1.1.39
   10.1.1.47
   10.10.1.39
   10.10.1.47
   172.16.21.39
   172.16.21.47
   172.16.21.67
   172.16.21.72
   ```

2. Copy the file to all nodes.
3. Restart the clcomd daemon on all nodes: **`refresh -s clcomd`**.

## 3.8.2 Configuring the netmon.cf file

If you use virtual Ethernet adapters in the cluster, you have to create the `/usr/es/sbin/cluster/netmon.cf` file.

> **Important:** This configuration requires APAR IV14422.

Select a good target IP address for each of your networks. We advise you to use the IP address of the default gateway, firewall, or any other fundamental device of your network. The address must answers for the ping requests.

Edit the `/usr/es/sbin/cluster/netmon.cf` file. The file format is shown:

```
!REQD base_ip1 target_IP1
!REQD base_ip2 target_IP2
!REQD persistent_ip target_IP3
```

Consider the following information as you edit this file:

- The first IP is always the monitored adapter; the second IP is the target that we want to ping.
- You can add as many lines as you want. If there are more than one entry for an adapter, PowerHA tries to ping all of them. If at least one target replies, the interface is marked “good”.
- Ensure that you add at least one line for each base adapter. It is suggested to add a ping target for your persistent addresses.
- The file can be different on the nodes.

You can see our `nemton.cf` files from our sample cluster nodes in Example 3-21. We use the default gateway as our target IP address for the persistent IP. Because the IP addresses on the base interfaces cannot access the default gateway, we use another PowerHA host IP for their target.

*Example 3-21 /usr/es/sbin/cluster/netmon.cf files from our cluster*

```
root@inst1 / # cat /usr/es/sbin/cluster/netmon.cf
!REQD 10.1.1.39 10.1.1.55
!REQD 10.10.1.39 10.10.1.73
!REQD 172.16.21.39 172.16.20.1

root@inst2 / # cat /usr/es/sbin/cluster/netmon.cf
!REQD 10.1.1.47 10.1.1.55
!REQD 10.10.1.47 10.10.1.73
!REQD 172.16.21.47 172.16.20.1
```

## 3.8.3 Initial cluster setup

Start **`smit sysmirror`**. Select **Cluster Nodes and Networks** → **Initial Cluster Setup (Typical)** → **Setup a Cluster, Nodes and Networks**. Provide a cluster name (avoid extra or international characters), add new nodes by moving the cursor down to the `New nodes` line, and press F4. SMIT shows you the host names from your `/etc/hosts` file. Select the hostname of the new node with F7. See Figure 3-20 on page 115.

PowerHA performs the following actions:

- Connects to the nodes through the specific communication path.
- Discovers the current network settings.
- Parses the `/etc/hosts` file.
- Discovers the shared disk settings.
- Defines the boot addresses in the cluster topology.
- Creates the network definitions for the cluster topology. By default, the new network name is `net_ether_01`, and it contains all base adapters.

This information is used in the SMIT menus to help users accurately select the existing components.

**Important:** Ensure that the following information applies when you add a new node to the cluster:

- The hostname can be resolved.
- There is no alias for the hostname in the `/etc/hosts` file.
- The hostname is not an alias.
- The hostname and the PowerHA node name are the same.
- The hostname and the “communication path to the node” are the same.

```
 Setup Cluster, Nodes and Networks (Typical)

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

                                                        [Entry Fields]
* Cluster Name                                       [installation_cluster]
  New Nodes (via selected communication paths)       []                      +
  Currently Configured Node(s)                        inst1

+--------------------------------------------------------------------------+
  |              New Nodes (via selected communication paths)                 |
  |                                                                           |
  | Move cursor to desired item and press F7.                                 |
  |     ONE OR MORE items can be selected.                                    |
  | Press Enter AFTER making all selections.                                  |
  |                                                                           |
  |   nimres1 (10.1.1.251)                                                    |
  |   sapnfs1 (172.16.21.30)                                                  |
  |   inst2b1 (10.1.1.47)                                                     |
  |   inst1b1 (10.1.1.39)                                                     |
  |   inst1b2 (10.10.1.39)                                                    |
  |   inst2b2 (10.10.1.47)                                                    |
  |   inst1p  (172.16.21.39)                                                  |
  |   inst2p  (172.16.21.47)                                                  |
  |   inst1   (172.16.21.67)                                                  |
  |   inst2   (172.16.21.72)                                                  |
  |   inst1backup     (10.1.2.39)                                             |
  |   inst2backup     (10.1.2.47)                                             |
  |                                                                           |
  | F1=Help               F2=Refresh             F3=Cancel                    |
F1| F7=Select             F8=Image               F10=Exit                     |
F5| Enter=Do              /=Find                 n=Find Next                  |
F9+---------------------------------------------------------------------------+
```

*Figure 3-20 Set up the cluster*

We created our `installation_cluster` with two nodes: inst1 and inst2. We selected the persistent adapter (inst2) of the second node as the communication path to the new node. When the setup is finished, the SMIT output shows the network and disk information of the discovered nodes.

## 3.8.4 Defining the repository disk

Define the repository disk and the cluster IP address:

1. Ensure that the repository disk has the same PVID on all cluster nodes. See “Configuring the hdisk Physical Volume ID on all cluster nodes” on page 107.
2. Start **`smit sysmirror`**. Select **Cluster Nodes and Networks** → **Initial Cluster Setup (Typical)** → **Define Repository Disk and Cluster IP Address**. Figure 3-21 on page 116 shows the SMIT window for defining the repository disk.
3. Press F4 and select the repository disk from the pickup list.

4. You can specify the cluster IP address that is the multicast address that is used for the heartbeat. If you leave it blank, PowerHA creates one for you (See “Multicast heartbeat” on page 89).

```
 Define Repository and Cluster IP Address

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

                                                        [Entry Fields]
* Cluster Name                                        inst1_cluster
* Repository Disk                                    [hdisk2]                +
  Cluster IP Address                                 []


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

*Figure 3-21 Defining the repository disk and cluster IP address*

## 3.8.5 Adding persistent IP labels

We describe how to add persistent IP labels:

1. Start **`smit sysmirror`**. Select **Cluster Nodes and Networks** → **Manage Nodes** → **Configure Persistent Node IP Label/Addresses** → **Add a Persistent Node IP Label/Address** → **Select a Node**. See Figure 3-22 on page 117 for the corresponding SMIT window.
2. Press F4 to list the available networks. At least one network is defined in the PowerHA topology: `net_ether_01`. This network is created during the initial cluster setup. By default, this network contains the IP range of the base addresses.
3. Select the network that contains the base adapters where the persistent address is located.
4. Go to the `Node IP Label/Address` line and press F4. The list shows the discovered and available IP labels from the `/etc/hosts` file. Choose the persistent address.
5. You can specify the netmask, or PowerHA automatically selects the netmask that is used on the base adapters.
6. Repeat this process for all persistent IP labels.

```
 Add a Persistent Node IP Label/Address

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

                                                        [Entry Fields]
* Node Name                                            inst1
* Network Name                                       [net_ether_01]         +
* Node IP Label/Address                              [inst1]                +
  Netmask(IPv4)/Prefix Length(IPv6)                  []

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

*Figure 3-22 Add a persistent address*

## 3.8.6 Checking the cluster topology information

Before you proceed, double-check your cluster topology. Start **`smit sysmirror`**. Select **Cluster Nodes and Networks** → **Manage the Cluster** → **Display PowerHA SystemMirror Configuration** to verify the current PowerHA configuration. See Figure 3-23 on page 118.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Cluster Name: inst1_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.32
There are 2 node(s) and 2 network(s) defined
NODE inst1:
        Network net_ether_01
                inst1b1 10.1.1.39
                inst1b2 10.10.1.39
        Network net_ether_02
                inst1backup     10.1.2.39
NODE inst2:
        Network net_ether_01
                inst2b2 10.10.1.47
                inst2b1 10.1.1.47
        Network net_ether_02
                inst2backup     10.1.2.47

No resource groups defined

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 3-23   Checking the PowerHA topology configuration*

There are other useful PowerHA commands to verify the cluster configuration. For example, **/usr/es/sbin/cluster/utilities/cllsif** is the best way to check the cluster network settings. See Example 3-22 for the output of the **cllsif** command on our sample cluster.

*Example 3-22   Verifying the cluster network settings with cllsif command*

```
root@inst1 /usr/es/sbin/cluster/utilities # ./cllsif
Adapter              Type       Network    Net Type   Attribute  Node       IP Address
Hardware Address Interface Name   Global Name      Netmask         Alias for HB Prefix Length

inst1b1          boot      net_ether_01 ether     public    inst1     10.1.1.39
en0                         255.255.254.0              23
inst1b2          boot      net_ether_01 ether     public    inst1     10.10.1.39
en2                         255.255.254.0              23
inst1backup     boot    net_ether_02 ether   public   inst1   10.1.2.39
255.255.255.0               24
inst2b1          boot      net_ether_01 ether     public    inst2     10.1.1.47
en0                         255.255.254.0              23
inst2b2          boot      net_ether_01 ether     public    inst2     10.10.1.47
en2                         255.255.254.0              23
```

```
inst2backup     boot    net_ether_02 ether   public   inst2   10.1.2.47
255.255.255.0               24
```

The persistent addresses are not shown in the **cllsif** output.

### 3.8.7 Verifying and synchronizing the cluster topology

We verify and deploy the cluster topology configuration. The PowerHA verification and synchronization performs the following tasks:

- Checks the new configuration for errors.
- Contacts the cluster nodes and downloads the current PowerHA configuration, if any.
- Compares the new and the current configuration.
- Synchronizes the new cluster configuration to all nodes.
- Creates and starts the CAA cluster.
- If the cluster is already running, the PowerHA applies the new modification dynamically.

Start **smit sysmirror** and select **Cluster Nodes and Networks** → **Verify and Synchronize Cluster Configuration** to synchronize the cluster. See Figure 3-24 for the output of the cluster verification function.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]
Timer object autoclverify already exists

Verification to be performed on the following:
        Cluster Topology
        Cluster Resources

Verification will interactively correct verification errors.

Retrieving data from available cluster nodes.  This could take a few minutes.

        Start data collection on node inst1
        Start data collection on node inst2
        Collector on node inst2 completed
        Collector on node inst1 completed
[MORE...44]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 3-24 PowerHA verification and synchronization*

**Important:** If you encounter any error or warning messages now, rerun the cluster verification and synchronization before you continue the installation process. For more information about how to debug your cluster problems, see 3.8.9, “If a problem occurs during the initial cluster configuration” on page 121.

### 3.8.8 Checking the CAA cluster

The underlying CAA cluster must be up and running. We can check it with the **lscluster** command. See Example 3-23 on how to check the CAA cluster status.

*Example 3-23 Checking the CAA cluster status*

```
root@inst1 / # lscluster -c
Cluster query for cluster inst1_cluster returns:
Cluster uuid: 0a21480a-7502-11e1-ad51-2e47986e9270
Number of nodes in cluster = 2
        Cluster id for node inst1 is 1
        Primary IP address for node inst1 is 172.16.21.39
        Cluster id for node inst2 is 2
        Primary IP address for node inst2 is 172.16.21.47
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.32

root@inst1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2

        Node name: inst1
        Cluster shorthand id for node: 1
        uuid for node: 09d7df1c-7502-11e1-ad51-2e47986e9270
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        inst1_cluster      local        0a21480a-7502-11e1-ad51-2e47986e9270

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: inst2
        Cluster shorthand id for node: 2
        uuid for node: 0a1b8e06-7502-11e1-ad51-2e47986e9270
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        inst1_cluster      local        0a21480a-7502-11e1-ad51-2e47986e9270
```

```
Number of points_of_contact for node: 4
Point-of-contact interface & contact state
 dpcom  UP  RESTRICTED
 en1  UP
 en0  UP
 en2  UP
```

Also, check the status of the communication interfaces with the **lscluster -i** command.

### 3.8.9 If a problem occurs during the initial cluster configuration

If a problem occurs during the cluster topology definition and synchronization, check the following information:

- The network interfaces are set correctly, and they follow the networking rules that are described in "Networking" on page 83.
- A common reason for failures is that multicasting does not work on the client network.
- The `/etc/hosts` file is correct.
- You follow the rules for the hostname (see "Initial cluster setup" on page 114).
- The name resolution works.
- All communication interface pairs can ping each other.
- The `inetd` is enabled.
- The clcomd and clstrmgrES daemons are active on all nodes.
- For node connectivity, `clrsh` works between the nodes, see Example 3-24.

*Example 3-24 Checking node connectivity with the clrsh command*

```
root@inst1 /usr/es/sbin/cluster/utilities # clrsh inst2 date
   Fri Mar 23 13:52:52 EDT 2012
```

During the early stages of the cluster configuration, you can check the following log files:

- `/var/log/clcomd/clcomd.log`: Cluster communication daemon log file, which is useful to check node connectivity and Global Object Data Manager (ODM) problems
- `/var/hacmp/log/cspoc.log`: C-SPOC commands log file, which is useful to debug why a PowerHA command failed on a (remote) node
- `/var/hacmp/adm/cluster.log`: PowerHA cluster log file, mostly Reliable Scalable Cluster Technology (RSCT) messages
- `/var/adm/ras/syslog.caa`: CAA cluster log file, which is useful to check why the **mkcluster** command fails

On 3.13, "Troubleshooting" on page 144, there are more cluster debugging tips and how to contact IBM for support.

## 3.9 Configuring PowerHA resources

We configure the PowerHA SystemMirror resources for high availability. We assume that you successfully configured and synchronized your PowerHA cluster topology.

### 3.9.1 Configuring the service IP labels

We follow these steps to configure the service IP labels:

1. Start `smit sysmirror`. Select **Cluster Applications and Resources → Resources → Configure Service IP Labels/Addresses → Add a Service IP Label/Address → Select the network name**. See Figure 3-25 for the SMIT window.

```
 Configure Service IP Labels/Addresses

Move cursor to desired item and press Enter.

  Add a Service IP Label/Address
  Change/ Show a Service IP Label/Address
  Remove Service IP Label(s)/Address(es)
  Configure Service IP Label/Address Distribution Preferences

   +--------------------------------------------------------------------------+
   |                               Network Name                               |
   |                                                                          |
   | Move cursor to desired item and press Enter.                             |
   |                                                                          |
   |   net_ether_01 (10.1.0.0/23 10.10.0.0/23)                                |
   |   net_ether_02 (10.1.2.0/24)                                             |
   |                                                                          |
   | F1=Help                F2=Refresh             F3=Cancel                  |
   | F8=Image               F10=Exit               Enter=Do                   |
F1=| /=Find                 n=Find Next                                       |
F9=+--------------------------------------------------------------------------+
```

*Figure 3-25 Select the network for the service IP label*

2. Select the network that has the IP range of the corresponding base addresses.
3. Press F4 for the list of discovered and available IP labels.
4. Choose your service IP label. See Figure 3-26 on page 123.
5. Optional: Specify the netmask. By default, PowerHA uses the subnet mask of the underlying base adapters.

```
 Add a Service IP Label/Address

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

                                                        [Entry Fields]
* IP Label/Address                                    instsvc1
+
  Netmask(IPv4)/Prefix Length(IPv6)                  []
* Network Name                                        net_ether_01


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

*Figure 3-26 Add a service IP label*

6. Repeat this process for all service IP labels.

## 3.9.2 Configuring the service IP address distribution preference

This distribution preference defines which base adapter is used for the service, the persistent IP address, and where the external traffic leaves the server. For the available options, see "PowerHA service IP addresses distribution policies" on page 10. We advise you to use `Collocation with Persistent Label`. Normally, you access the default gateway through your persistent address (as in our sample cluster). In this case, the persistent and the service addresses connect to same base adapter; therefore, you can avoid routing problems.

The optional `Source IP Label for outgoing packets` explicitly defines the source IP address that is used in the TCP/IP pockets for the outgoing traffic. If you have strict firewall rules, use this option. See also 3.5.14, "Firewalls" on page 90.

Follow these steps to configure the service IP address distribution preference:

1. Start **`smit sysmirror`**. Select **Cluster Applications and Resources** → **Resources** → **Configure Service IP Labels/Addresses** → **Configure Service IP Label/Address Distribution Preferences**. Then, select the network.
2. Press F4 and select the distribution preference. See Figure 3-27 on page 124.
3. Optional: Define the source IP label for the outgoing traffic. Press F4 for the available service IP labels.
4. Repeat these steps for all networks with service IP labels.

```
 Configure Service IP Labels/Address Distribution Preference

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

                                                         [Entry Fields]
* Network Name                                         net_ether_01
* Distribution Preference                              Anti-Collocation +
  Source IP Label for outgoing packets +

   +--------------------------------------------------------------------------+
   |                          Distribution Preference                         |
   |                                                                          |
   | Move cursor to desired item and press Enter.                             |
   |                                                                          |
   |   Anti-Collocation                                                       |
   |   Anti-Collocation with Source                                           |
   |   Collocation                                                            |
   |   Collocation with Source                                                |
   |   Collocation with Persistent Label                                      |
   |   Anti-Collocation with Persistent Label                                 |
   |   Anti-Collocation with Persistent Label and Source                      |
   |                                                                          |
   | F1=Help                F2=Refresh             F3=Cancel                  |
F1=| F8=Image               F10=Exit               Enter=Do                   |
F5=| /=Find                 n=Find Next                                       |
F9=+--------------------------------------------------------------------------+
```

*Figure 3-27   Configure service IP label distribution preference*

### 3.9.3 Verifying the service IP settings

Verify the service IP address settings in PowerHA SystemMirror topology network settings with the **cllsif** command (Example 3-25). Because a service address can be on both nodes, the **cllsif** lists them twice.

*Example 3-25   Checking the service adapter settings in the PowerHA topology*

```
root@inst1 / # /usr/es/sbin/cluster/utilities/cllsif
Adapter              Type       Network    Net Type   Attribute  Node       IP Address
Hardware Address Interface Name   Global Name      Netmask          Alias for HB Prefix Length

inst1b2          boot     net_ether_01 ether     public    inst1     10.10.1.39
en2                        255.255.254.0              23
inst1b1          boot     net_ether_01 ether     public    inst1     10.1.1.39
en0                        255.255.254.0              23
instsvc2      service  net_ether_01 ether   public   inst1   172.16.21.72
255.255.254.0              23
instsvc1      service  net_ether_01 ether   public   inst1   172.16.21.67
255.255.254.0              23
inst1backup       boot     net_ether_02 ether     public    inst1     10.1.2.39
en1                        255.255.255.0              24
inst2b2          boot     net_ether_01 ether     public    inst2     10.10.1.47
en2                        255.255.254.0              23
```

```
inst2b1          boot     net_ether_01 ether     public    inst2     10.1.1.47
en0                        255.255.254.0              23
instsvc2      service  net_ether_01 ether   public   inst2   172.16.21.72
255.255.254.0              23
instsvc1      service  net_ether_01 ether   public   inst2   172.16.21.67
255.255.254.0              23
inst2backup       boot     net_ether_02 ether     public    inst2     10.1.2.47
en1                        255.255.255.0              24
```

### 3.9.4 Creating the resource group

We describe how to create a resource group:

1. Start **smit sysmirror**. Then, select **Cluster Applications and Resources** → **Resource Groups** → **Add a Resource Group**. See Figure 3-28.
2. Supply the resource group name. We advise that you provide a meaningful name for the resource group. For example, include the primary node name or the application name in the resource group name. We use inst1rg and inst2rg for the names in our sample cluster.
3. Select the participating node names. The node name order is important because it is the node priority order for the resource group. You can select the nodes from the pickup list by pressing F4.

```
 Add a Resource Group

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

                                              [Entry Fields]
* Resource Group Name                         [inst1rg]
* Participating Nodes (Default Node Priority) []                             +

  Startup Policy               Online On Home Node Only                      +
  Fallover Policy              Fallover To Next Priority Node In The List    +
  Fallback Policy              Fallback To Higher Priority Node In The List  +

         +----------------------------------------------------------------+
         |              Participating Nodes (Default Node Priority)        |
         |                                                                 |
         | Move cursor to desired item and press F7.                       |
         |     ONE OR MORE items can be selected.                          |
         | Press Enter AFTER making all selections.                        |
         |                                                                 |
         | > inst1                                                         |
         | > inst2                                                         |
         |                                                                 |
         | F1=Help              F2=Refresh              F3=Cancel          |
F1=Help  | F7=Select            F8=Image                F10=Exit           |
F5=Reset | Enter=Do             /=Find                  n=Find Next        |
F9=Shell +----------------------------------------------------------------+
```

*Figure 3-28 Add a resource group in smit sysmirror*

4. Select the startup policies:
   - Online On Home Node Only
   - Online On First Available Node
   - Online Using Distribution Policy
   - Online On All Available Nodes
5. Choose the fallover policies:
   - Fallover To Next Priority Node In The List
   - Fallover Using Dynamic Node Priority
   - Bring Offline (On Error Node Only)
6. Select the fallback policies:
   - Fallback To Higher Priority Node In The List
   - Never Fallback

For a detailed description of these resource group policies, see "PowerHA resource group policies" on page 12.

We created two resource groups with the following characteristics:

- Resource group name: `inst1rg`:
  - Participating nodes (default node priority): `inst1, inst2`
  - Startup policy: `Online On First Available Node`
  - Fallover policy: `Fallover To Next Priority Node In The List`
  - Fallback policy: `Fallback To Higher Priority Node In The List`
- Resource group name: `inst2rg`:
  - Participating nodes (default node priority): `inst2, inst1`
  - Startup policy: `Online On Home Node Only`
  - Fallover policy: `Fallover To Next Priority Node In The List`
  - Fallback policy: `Never Fallback`

## 3.9.5 Adding resources and attributes for the resource groups

In this step, we define the cluster resources, such as the service address, shared volume group, and attributes for the resource groups.

Start **`smit sysmirror`**. Select **Cluster Applications and Resources** → **Resource Groups** → **Change/Show Resources and Attributes for a Resource Group** → **Select the resource group**. See Figure 3-29 on page 127 for the SMIT panel.

```
 Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
  Resource Group Name                                  inst1rg
  Participating Nodes (Default Node Priority)          inst1 inst2

  Startup Policy                                       Online On First Avail>
  Fallover Policy                                      Fallover To Next Prio>
  Fallback Policy                                      Fallback To Higher Pr>
  Fallback Timer Policy (empty is immediate)          []                      +

  Service IP Labels/Addresses                         [instsvc1]              +
  Application Controllers                             []                      +

  Volume Groups                                       [app1vg]                +
  Use forced varyon of volume groups, if necessary     false                  +
  Automatically Import Volume Groups                   false                  +

  Filesystems (empty is ALL for VGs specified)        []                      +
  Filesystems Consistency Check                        fsck                   +
  Filesystems Recovery Method                          sequential             +
  Filesystems mounted before IP configured             false                  +

  Filesystems/Directories to Export (NFSv2/3)         []                      +
  Filesystems/Directories to NFS Mount                []
  Network For NFS Mount                               []                      +

  Tape Resources                                      []                      +
  Raw Disk PVIDs                                      []                      +

  Primary Workload Manager Class                      []                      +
  Secondary Workload Manager Class                    []                      +

  Miscellaneous Data                                  []
  WPAR Name                                           []                      +
  User Defined Resources                              [ ]                     +

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

*Figure 3-29 Add resources for a resource group*

You can specify the following parameters:

- `Fallback Timer Policy`: See "PowerHA resource group policies" on page 12 for more information.
- `Service IP Labels/Addresses`: You can add one or more service addresses. Press F4 for the list of the available service labels.
- `Application Controllers`: Press F4 for the defined application servers. We do not use this option during this setup.

- `Volume Groups`: F4 gives the list of the discovered, shared, and enhanced concurrent volume groups.
- `Use forced varyon of volume groups, if necessary`: Select `true` only if you use AIX LVM mirroring in the shared volume groups.
- `Automatically Import Volume Groups`: Leave it `false`. This setting is only important when you add a volume group to the cluster dynamically.
- `File systems`: If you want PowerHA to mount all file systems from the shared volume groups, leave this field empty.
- `File systems Consistency Check`: You can select between the `fsck` and `logredo` commands to perform the consistency check. We advise that you use `fsck`.
- `File systems Recovery Method`: It can be `sequential` or `parallel`. If you have many shared file systems, set it to `parallel`.
- `File systems mounted before IP configured`: When you use NFS in the cluster, always set it to `true`. Otherwise, leave it `false`.
- `File systems/Directories to Export (NFSv2/3)`: Enter the directories to NFS export inside the cluster.
- `File systems/Directories to NFS Mount`: All nodes in the resource group mount this file system. The format is `global_nfs_mount;local_filesystem`. For the correct syntax, see “NFS cross mount” on page 129.
- `Network For NFS Mount`: The preferred network to mount over the NFS directories.
- `Tape Resources`: You can add your tape resources here.
- `Raw Disk PVIDs`: If your application uses raw disks without volume groups, you must specify the physical volume IDs for the disks.
- `Primary Workload Manager Class`: Optional Workload Manager Class that is used by the primary node for this resource group.
- `Secondary Workload Manager Class`: Optional Workload Manager Class that is used by the backup nodes for this resource group.
- `Miscellaneous Data`: You can specify a string here that is placed in the environment with the resource group information during script execution. Leave it empty.
- `WPAR Name`: You can specify a WPAR name where this resource group application runs. For more information about WPARs and PowerHA, see Chapter 8, “Workload partition and PowerHA scenario” on page 487.
- `User Defined Resources`: You can have your own type of resources. Virtually, it is a set of start, stop, recovery, and notification scripts for special purposes.

We create the following resources:

- Resources for `inst1rg`:
  - Service IP label: `instsvc1`
  - Volume groups: `app1vg`
- Resources for `inst2rg`:
  - Service IP label: `instsvc2`
  - Volume groups: `app2vg`

All other options are empty or use defaults.

### 3.9.6 NFS cross mount

If you want to use NFS cross mount in the cluster, you need extra preparation steps.

**Important:** The NFS cross mount requires that you install the following filesets:

- `cluster.es.nfs`
- `bos.net.nfs.client`
- `bos.net.nfs.server`

For example, you want to import a file system that is called `/nfsdata` to all cluster nodes from the primary node (inst1). You must create the file system in a shared volume group. But, do not name it `/nfsdata`. Use a different name for it: `/nfsdata_local`. PowerHA mounts this file system locally on the primary node, then it exports and mounts it on all nodes under the `/nfsdata` name. On the local node, `/nfsdata_local` is mounted to `/nfsdata` through NFS loopback (Example 3-26). Your application must use `/nfsdata` on all nodes.

*Example 3-26 NFS cross mount on the primary node*

```
root@inst1 / # mount
  node       mounted        mounted over    vfs       date        options
-------- ---------------  ---------------  ------ ------------ ---------------
         /dev/hd4         /                jfs2   Mar 23 14:55 rw,log=/dev/hd8
         /dev/hd2         /usr             jfs2   Mar 23 14:55 rw,log=/dev/hd8
         /dev/hd9var      /var             jfs2   Mar 23 14:56 rw,log=/dev/hd8
         /dev/hd3         /tmp             jfs2   Mar 23 14:56 rw,log=/dev/hd8
         /dev/hd1         /home            jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /dev/hd11admin   /admin           jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /proc            /proc            procfs Mar 23 14:57 rw
         /dev/hd10opt     /opt             jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /aha             /aha             ahafs  Mar 23 14:57 rw
         /dev/nfslv       /nfsdata_local   jfs2   Mar 23 14:57 rw,log=/dev/loglv10
inst1    /nfsdata_local   /nfsdata         nfs3   Mar 28 15:23
```

The backup nodes mount the `/nfsdata` like a normal NFS share (Example 3-27).

*Example 3-27 NFS cross mount on the backup node*

```
root@inst2 / # mount
  node       mounted        mounted over    vfs       date        options
-------- ---------------  ---------------  ------ ------------ ---------------
         /dev/hd4         /                jfs2   Mar 23 14:55 rw,log=/dev/hd8
         /dev/hd2         /usr             jfs2   Mar 23 14:55 rw,log=/dev/hd8
         /dev/hd9var      /var             jfs2   Mar 23 14:56 rw,log=/dev/hd8
         /dev/hd3         /tmp             jfs2   Mar 23 14:56 rw,log=/dev/hd8
         /dev/hd1         /home            jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /dev/hd11admin   /admin           jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /proc            /proc            procfs Mar 23 14:57 rw
         /dev/hd10opt     /opt             jfs2   Mar 23 14:57 rw,log=/dev/hd8
         /aha             /aha             ahafs  Mar 23 14:57 rw
inst1    /nfsdata_local   /nfsdata         nfs3   Mar 28 15:23
```

Create a resource group with the following resources for the NFS cross mount. The following resources are for `inst1rg`:

- Participating nodes: `inst1` and `inst2`
- Service IP label: `instsvc1`
- Volume group: `app1vg`
- File systems are mounted before the IP is configured: `true`
- File systems/directories to export (NFSv2/3): `/nfsdata_local`
- File systems/directories to NFS mount: `/nfsdata;/nfsdata_local`
- Network for NFS mount: `net_ether_01`

## 3.9.7 Checking the cluster resources

Review all modifications that you made before you performed the cluster synchronization. You can check the cluster resources by starting **`smit sysmirror`**. Select **Cluster Applications and Resources** → **Resource Groups** → **Show All Resources by Node or Resource Group** → **Show Resource Information by Resource Group** → **Select resource group**. See Figure 3-30.

```
COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]
Resource Group Name            inst1rg
Participating Node Name(s)     inst1 inst2
Startup Policy                 Online On First Available Node
Fallover Policy                Fallover To Next Priority Node In The List
Fallback Policy                Fallback To Higher Priority
Node In The List
Site Relationship              ignore
Node Priority
Service IP Label               instsvc1
Filesystems                    ALL
Filesystems Consistency Check  fsck
Filesystems Recovery Method    sequential
[MORE...27]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 3-30 Displaying resource group configuration*

## 3.9.8 Verifying and synchronizing the cluster resource configuration

Verify and synchronize the PowerHA cluster resources by starting **`smit sysmirror`**. Select **Cluster Applications and Resources** → **Verify and Synchronize Cluster Configuration**.

The PowerHA verification and synchronization performs the following tasks:

- Check the new configuration for errors.
- Contacts the cluster nodes and downloads the current PowerHA configuration.
- Compares the new and the current configuration.
- Synchronizes the new cluster configuration to all nodes.
- If the cluster is already running, it applies the new modification dynamically.

**Important:** If you encounter any error or warning messages here, correct them and rerun cluster verification and synchronization before you proceed with the installation process. For more information about how to debug your cluster problems, go to 3.13, “Troubleshooting” on page 144.

## 3.9.9 Starting the cluster

It is time to start the cluster. Start **`smit sysmirror`**. Select **System Management (C-SPOC)** → **PowerHA SystemMirror Services** → **Start Cluster Services**, or use **`smit clstart`** to start the cluster services. By pressing F4, select both nodes to start. See Figure 3-31 for the window.

```
 Start Cluster Services

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

                                                        [Entry Fields]
* Start now, on system restart or both                now                     +
  Start Cluster Services on these nodes              [inst1,inst2]            +
* Manage Resource Groups                              Automatically           +
  BROADCAST message at startup?                       false                   +
  Startup Cluster Information Daemon?                 false                   +
  Ignore verification errors?                         false                   +
  Automatically correct errors found during           Interactively           +
  cluster start?

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

*Figure 3-31 smit clstart window*

The following startup options are available:

- `Start now, on system restart or both`: Normally, you select `now`. However, you can specify that the cluster automatically starts at system reboot, which is not advised.
- `Start Cluster Services on these nodes`: Select the cluster nodes where you want to start the cluster. Select all nodes.
- `Manage Resource Groups`: If you specify `Automatically`, PowerHA starts all resource groups. `Manually` means that you have to bring up the resource group by using **`smit cspoc`**. Then, select **Resource Groups and Applications** → **Bring a Resource Group Online**.
- `BROADCAST message at startup`: The cluster manager sends a “wall” broadcast notification to all terminal sessions.

- `Startup Cluster Information Daemon`: If you plan to use the **clstat** utility or any other Simple Network Management Protocol (SNMP)-based monitoring tool, you must start the **clinfo** daemon. For more information about **clstat**, see "Clstat" on page 154.
- `Ignore verification errors`: During the cluster startup, PowerHA runs a full cluster verification and synchronization. If this process encounters any error, you can force the cluster manager to ignore it. This option is useful if you lost one of your nodes or the cluster configuration is broken.
- `Automatically correct errors found during cluster start`: You can select how the cluster manager corrects the verification errors: `automatically`, `interactively` (prompts for yes or no), or not at all.

**Tip:** We advise that you monitor the output of the `/var/hacmp/log/hacmp.out` log file during the cluster startup, shutdown, and resource group movement operation.

## 3.9.10 Verifying PowerHA services

After you successfully start PowerHA on all nodes, check the status of the cluster:

- Cluster resource group status (**/usr/es/sbin/cluster/utilities/clRGinfo -m**).
- Service addresses are configured (**netstat -in**).
- Shared volume groups are activated in concurrent mode (**lspv**).
- Shared file systems are mounted (**mount**).
- Repository disk is activated and `/aha` file system is mounted (**lspv**, **mount**).
- Log on to the nodes from an external server by using the service address of the nodes.

See Example 3-28 to check the status of the cluster by using various AIX commands.

*Example 3-28 Checking the status of the PowerHA cluster*

```
root@inst1 / #  /usr/es/sbin/cluster/utilities/clRGinfo -m
----------------------------------------------------------------------
Group Name     Group State          Application state         Node
----------------------------------------------------------------------
inst1rg        ONLINE                                         inst1
inst2rg        ONLINE                                         inst2

root@inst1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.98.6e.92.6f   641615     0   105907     0     0
en0   1500  10.1        inst1b1            641615     0   105907     0     0
en0   1500  172.16.20   inst1              641615     0   105907     0     0
en0   1500  172.16.20   instsvc1           641615     0   105907     0     0
en1   1500  link#3      2e.47.98.6e.92.70    31450     0     1006     0     0
en1   1500  10.1.2      inst1backup         31450     0     1006     0     0
en2   1500  link#4      2e.47.98.6e.92.71   539828     0     3443     0     0
en2   1500  10.10       inst1b2            539828     0     3443     0     0
lo0   16896 link#1                         240536     0   240537     0     0
lo0   16896 127         loopback           240536     0   240537     0     0
lo0   16896 loopback                       240536     0   240537     0     0

root@inst1 / # lspv
hdisk0         00f74d45f95e5beb                    rootvg          active
```

```
hdisk1          00f74d45fa932b47                    rootvg          active
hdisk2          00f74d45367ec638                    caavg_private   active
hdisk3          00f74d45367ed24c                    app1vg          concurrent
hdisk4          00f74d45367edf3b                    app2vg          concurrent

root@inst1 / # df
Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4         2097152   1593992   24%    10786     6% /
/dev/hd2         6291456   2098312   67%    49451    18% /usr
/dev/hd9var      4194304   3313960   21%     8789     3% /var
/dev/hd3         6291456   6281720    1%      143     1% /tmp
/dev/hd1         4194304   4192976    1%        5     1% /home
/dev/hd11admin    2097152   2096112    1%        7     1% /admin
/proc                  -         -    -         -     -  /proc
/dev/hd10opt     4194304   3810920   10%     7065     2% /opt
/dev/livedump     524288    523552    1%        4     1% /var/adm/ras/livedump
/aha                   -         -    -        45     1% /aha
/dev/data1lv      163840    163152    1%        4     1% /data1
```

### 3.9.11 Testing the cluster functionality without an application

Before we install the application, we must test the PowerHA functionality. Without a configured application, we can ensure that the cluster works and that the takeover and resource group movements work. To start a resource group relocation, start **`smit sysmirror`**. Select **System Management (C-SPOC)** → **Resource Groups and Applications** → **Move Resource Groups to Another Node**. Select the resource group to move from the list, and then, select the destination node (see Figure 3-32).

```
 Move Resource Group(s) to Another Node

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

                                                        [Entry Fields]
  Resource Group(s) to be Moved                       inst1rg
  Destination Node                                    inst2


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

*Figure 3-32 Move a resource group to another node*

For more details about testing PowerHA SystemMirror, see 3.12, “Test scenarios” on page 143. If you encountered any problem during the cluster tests, check the `/var/hacmp/log/hacmp.out` log file for any obvious errors or `config_too_long_events`. For more information about how to debug PowerHA problems, see 3.13, “Troubleshooting” on page 144.

### 3.9.12 Stopping the cluster

Follow these steps if you want to stop the PowerHA SystemMirror:

1. Start **smit clstop**. See Figure 3-33.
2. Select the nodes where you want to stop the cluster.
3. Choose the shutdown mode:
   - Bring Resource Groups Offline: Normal shutdown of the cluster. The application stops.
   - Move Resource Groups: Takeover. The resource groups from the node are moved to the other nodes.
   - Unmanage Resource Groups: The PowerHA stops, but all resources, the application, shared volume groups, and service IP remain available on the cluster node. This method is a good way to install the PowerHA fixes without stopping the application.

```
 Stop Cluster Services

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

                                                        [Entry Fields]
* Stop now, on system restart or both                 now                    +
  Stop Cluster Services on these nodes               [inst1,inst2]           +
  BROADCAST cluster shutdown?                         true                   +
* Select an Action on Resource Groups                 Bring Resource Groups> +


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

*Figure 3-33 Stop PowerHA SystemMirror*

## 3.10 Configuring the application for PowerHA

Now, our cluster is ready for the application installation. Follow these steps:

1. Start the cluster and bring up all resource groups on its primary node.
2. Hand over the cluster to the application administrators.
3. Install the application on the primary node.
4. Create the application startup and shutdown scripts. For more information, see “Creating the application startup and shutdown scripts” on page 135.
5. Stop the application on the primary node.
6. Move the resource groups to the backup server.
7. Install the application on the backup server.
8. Stop the application and revert the cluster.

Ensure that the following information is correct:

- You have the same users and settings cluster wide. User names and the corresponding user ID (UID) must be the same on all nodes. See “Showing the disk UUID” on page 182.
- You install all application components on all nodes.
- The application configuration is the same everywhere.
- The application survives an abrupt end and restart. You might need to enable application redo logging.
- You have to maintain the application on all cluster nodes during the full lifecycle of the system. If you install application upgrades or new components, you must apply the changes on all cluster nodes.

## 3.10.1 Creating the application startup and shutdown scripts

You must write your own application startup and shutdown scripts. Follow these guidelines:

- Put your scripts in a separate directory for easy management.
- Name them meaningfully.
- Do not use interactive scripts or startup and shutdown programs.
- The startup script must start all of the required application components.
- The shutdown script must stop all application components. *No process can be running.*
- The scripts can differ on the nodes, but they must have the same name and path.
- The scripts must finish in a timely manner and return exit 0 on success. For example, with an Oracle Database, use shutdown abort to stop the database.
- During execution, PowerHA redirects the output of the script to the `/var/hacmp/log/hacmp.out` log. We advise you to log all pertinent information to a separate application startup and shutdown log file.
- Test run your scripts before you add them to the cluster configuration.

## 3.10.2 Creating the application monitoring scripts

PowerHA SystemMirror supports application monitoring. There are two methods:

- Process monitoring: You have to specify a process name, the number of instances, and the process owner name. PowerHA checks for this specific process.
- Custom monitoring: You can write your own application monitor script.

Requirements for the custom application monitor script:

- Check the application availability by running some internal application process. For example, in Oracle, you can run an SQL select on the system tables.
- The script must finish in a few seconds.
- The monitor must return exit code 0 on success.
- The monitor must return a non-zero exit code on failure.
- If the application fails, print a meaningful description of the error.
- The scripts can differ on the nodes, but they must have the same name and path.
- PowerHA redirects the output of the monitoring script to the `/var/hacmp/log` directory. The file name is `clappmond.`<*application_name*>`.`<*resource_group_name*>`.monitor.log`.

## 3.10.3 Configuring the application control scripts in PowerHA

We describe how to configure the application control scripts in PowerHA:

1. Start `smit sysmirror`. Select **Cluster Applications and Resources → Resources → Configure User Applications (Scripts and Monitors) → Application Controller Scripts → Add Application Controller Scripts**. See Figure 3-34.
2. Enter the name of the application server (use a meaningful name).
3. Provide the full path name of the startup and stop scripts.
4. You cannot add the application monitor name yet, so leave it empty. We configure that name in the next step (“Configuring the application monitor in PowerHA” on page 136).
5. Select the application startup mode. The default and suggested mode is `background`. If you specify `foreground`, the PowerHA event execution waits for the termination of the application startup script.

```
 Add Application Controller Scripts

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

                                                        [Entry Fields]
* Application Controller Name                        [app1]
* Start Script                                       [/usr/ha/start_app1]
* Stop Script                                        [/usr/ha/stop_app1]
  Application Monitor Name(s)                                             +
  Application startup mode                           [background]         +

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

*Figure 3-34 Configuring application controller scripts*

We created two application servers:

- Application controller name: `app1`
  - Start script: `/usr/ha/start_app1`
  - Stop script: `/usr/ha/stop_app1`
- Application controller name: `app2`
  - Start script: `/usr/ha/start_app2`
  - Stop script: `/usr/ha/stop_app2`

## 3.10.4 Configuring the application monitor in PowerHA

**Important:** Ensure that the application administrators are aware of the application monitoring. They must know how to suspend and resume the monitoring.

During application changes and upgrades, you must suspend the monitoring. For instructions to suspend and resume the application monitoring, see 4.7, “Monitoring the cluster” on page 154.

Configure the application monitor in PowerHA SystemMirror:

1. Start `smit sysmirror`. Select **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Monitors** → **Configure Custom Application Monitors** → **Add a Custom Application Monitor**. See Figure 3-35 on page 138.
2. Enter a name for the monitor.
3. Select the corresponding application.
4. Choose between `Long-running monitoring`, `Startup Monitoring`, or `Both`:
   - `Long-running Monitoring`: The default mode. PowerHA periodically checks the availability of the application. The checks start after the specified stabilization interval passed.
   - `Startup Monitoring`: In this case, the cluster manager checks that the application successfully started during the stabilization interval. This option is useful if you have a child resource group and an application that depends on this application.
   - `Both`: The combination of the prior two methods.
5. Enter the full path name of the monitor script.
6. Supply the following parameters (all interval values are in seconds):
   - `Monitor Interval`: The monitor runs periodically at this interval. Usually, 60 seconds is fine.
   - `Hung Monitor Signal`: The signal is sent to stop the Monitor Method if it does not return within the monitor interval. By default, it is SIGKILL (9).
   - `Stabilization Interval`: The estimated start time of the application. There is no monitoring during this period for `Long-running monitoring`. In `Startup Monitoring` mode, the application must be up and running within this period. If not, PowerHA starts a takeover. To be cautious, specify a generous time, for example 180 seconds.
   - `Restart Count`: The number of times to restart the application before you start a takeover.
   - `Restart Interval`: The number of seconds that the application must remain stable after the stabilization interval resets the restart counter. The formula is:

     `restart interval >= 1.1 * restart count* (monitor Interval + stabilization Interval)`

     Eventually, PowerHA calculates this interval for you if you leave it blank.
   - `Action on Application Failure`: Select `Takeover` or `Notify`. If you choose notify, you have to supply a script that runs if an application has an error.
   - `Notify Method`: The full path name of the notification script. This method is required only if you choose Notify earlier.
   - `Cleanup Method`: You can leave it blank; PowerHA uses the application server stop script. Alternatively, you can specify another application stop script.
   - `Restart Method`: If blank, PowerHA uses the application startup script, but you can have a different startup script.

```
 Add Custom Application Monitor

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

                                                        [Entry Fields]
* Monitor Name                                       [app1monitor]
* Application Controller(s) to Monitor                app1                   +
* Monitor Mode                                       [Long-running monitori> +
* Monitor Method                                     [/usr/ha/monitor_app1]
  Monitor Interval                                   [60]                     #
  Hung Monitor Signal                                []                       #
* Stabilization Interval                             [180] #
* Restart Count                                      [3]                      #
  Restart Interval                                   [600]                    #
* Action on Application Failure                      [fallover]              +
  Notify Method                                      []
  Cleanup Method                                     []                       /
  Restart Method                                     []                       /

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

*Figure 3-35   Add a custom application monitor*

We created two application monitors (Figure 3-35):

- Monitor name: `app1monitor`
  - Monitor mode: `Long-running monitoring`
  - Monitor method: `/usr/ha/monitor_app1`
  - Monitor interval: `60`
  - Stabilization interval: `180`
  - Restart count: `3`
  - Action on application failure: `fallover`
- Monitor name: `app2monitor`
  - Monitor mode: `Long-running monitoring`
  - Monitor method: `/usr/ha/monitor_app1`
  - Monitor interval: `120`
  - Stabilization interval: `360`
  - Restart count: `3`
  - Action on application failure: `fallover`

## 3.10.5 Add the application to the resource groups

We add the application to the resource groups:

1. You can add your application to the cluster by opening **`smit sysmirror`**. Select **Cluster Applications and Resources → Resource Groups → Change/Show All Resources and Attributes for a Resource Group**.
2. Select the resource group that you want.
3. Go to the `Application Controllers` line and press F4 for the list of configured application servers. See Figure 3-36.

```
 Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
  Resource Group Name                                 inst1rg
  Participating Nodes (Default Node Priority)         inst1 inst2

  Startup Policy                                      Online On First Avail>
  Fallover Policy                                     Fallover To Next Prio>
  Fallback Policy                                     Fallback To Higher Pr>
  Fallback Timer Policy (empty is immediate)         []                     +

  Service IP Labels/Addresses                        [instsvc1]             +
  Application Controllers                            [app1]                 +

  Volume Groups                                      [app1vg ]              +
  Use forced varyon of volume groups, if necessary    false                 +
  Automatically Import Volume Groups                  false                 +

  Filesystems (empty is ALL for VGs specified)       [ ]                    +
  Filesystems Consistency Check                       fsck                  +
  Filesystems Recovery Method                         sequential            +
  Filesystems mounted before IP configured            false                 +

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

*Figure 3-36 Adding an application server to a resource group*

## 3.10.6 Cluster verification and synchronization

To apply the application configuration changes, we must run cluster verification and synchronization again. This process checks for configuration errors and deploys the new cluster configuration. If everything is correct, PowerHA starts the application.

Start **`smit sysmirror`**. Select **Cluster Applications and Resources → Verify and Synchronize Cluster Configuration**.

### 3.10.7 Checking the application integration

As the final step in the cluster application integration, you have to ensure that the application starts and stops correctly in the PowerHA cluster. We advise you to perform a full cluster test, as described in 3.12, "Test scenarios" on page 143.

You can check the resource group and application status with the **clRGinfo** command (Example 3-29).

*Example 3-29 Checking the cluster applications*

```
root@inst1 / # /usr/es/sbin/cluster/utilities/clRGinfo -m
-----------------------------------------------------------------------------
Group Name     Group State              Application state          Node
-----------------------------------------------------------------------------
inst1rg        ONLINE                                              inst1
 app1                                   ONLINE MONITORED

inst2rg        ONLINE                                              inst2
 app2                                   ONLINE MONITORED
```

## 3.11 Dynamic LPAR and capacity on demand resources

PowerHA supports Dynamic LPAR, capacity on demand (CoD) CPU, and memory resources. You can configure the minimum and desired number of CPUs, virtual CPUs, and memory for each application. When an application is activated on a node, PowerHA contacts the HMC to acquire this resource in addition to the current resources of the LPAR.

In our final example, we show how to set up Dynamic LPAR and CoD resources. If you plan to use CoD resources, ensure that they are activated on the HMC.

1. Example 3-30 shows how to set up passwordless Secure Shell (SSH) communication to the HMC.

   *Example 3-30 Set up passwordless SSH communication to the HMC*

   ```
   root@inst1 /.ssh # /usr/bin/ssh-keygen -t rsa
   Generating public/private rsa key pair.
   Enter file in which to save the 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@inst1

   root@inst1 / # mykey=`cat ~/.ssh/id_rsa.pub`

   root@inst1 / # ssh hscroot@172.16.20.113 mkauthkeys -a \"$mykey\"
   ```

2. Add the HMC communication path to a node.

   Start **smit sysmirror**. Select **Cluster Applications and Resources → Resources → Configure User Applications (Scripts and Monitors) → Configure Application for Dynamic LPAR and CoD Resources → Configure Communication Path to HMC → Add HMC IP addresses for a node**. See Figure 3-37 on page 141.

You have to enter the node name, the corresponding HMC IP address, and the managed system name (server name).

```
 Add HMC IP addresses for a node

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

                                                        [Entry Fields]
* Node Name                                          [inst1]                 +
* HMC IP Address(es)                                 [172.16.20.113]         +
Managed System Name                              [Server02-8233-E8B-SN104D45R]

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

*Figure 3-37 Adding an HMC IP address for a node*

3. Configure the Dynamic LPAR or CoD resources:
   a. Start **`smit sysmirror`**. Select **Cluster Applications and Resources → Resources → Configure User Applications (Scripts and Monitors) → Configure Application for Dynamic LPAR and CoD Resources → Configure Dynamic LPAR and CoD Resources for Applications → Add Dynamic LPAR and CoD Resources for Applications**. See Figure 3-38 on page 142.
   b. Select the application.
   c. Supply the following parameters:
      - `Minimum number of CPUs`: Minimum number of virtual CPUs to acquire when the resource group is activated. If the system cannot fulfill this request, PowerHA starts the resource group recovery actions and takes over.
      - `Desired number of CPUs`
      - `Minimum number of processing units`: Minimum number of processing units to acquire for the resource group. If the system does not have enough CPU capacity, PowerHA performs a takeover.
      - `Desired number of processing units`
      - `Minimum amount of memory`: The minimum memory that is required for the application. If the server does not have enough memory, PowerHA performs a takeover.
      - `Desired amount of memory`
      - `Use CUoD if resources are insufficient`: Select yes, if you want to use CoD resources.
      - `I agree to use CUoD resources`: Select yes one more time if you plan to use CoD.

```
 Add Dynamic LPAR and CoD Resources for Applications

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

                                                      [Entry Fields]
* Application Controller Name                        app1

* Minimum number of CPUs                             [4]                      #
* Desired number of CPUs                             [4]                      #

  Parameters for Shared Processor partitions
*     Minimum number of processing units             [  1.00]
*     Desired number of processing units             [  2.00]

* Minimum amount of memory (in megabytes)            [0]                      #
* Desired amount of memory (in megabytes)            [0]                      #

* Use CUoD if resources are insufficient?            [no]                    +

* I agree to use CUoD resources                      [no]                    +
  (Using CUoD may result in extra costs)

  You must ensure that
  * CoD enablement keys are activated
  * CoD resources are not used for any other purpose

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

*Figure 3-38 Configuring Dynamic LPAR and CoD resources*

In Figure 3-38, we configured four virtual CPUs and two processing units for `app1`. We can see the actual CPU capacity numbers on the backup node (`inst2`) when `app1rg` is not activated (Example 3-31).

*Example 3-31 Entitled capacity and the number of online CPUs when app1 is not activated*

```
root@inst2 / # lparstat -i|grep  "Entitled Capacity"
Entitled Capacity                         : 1.00

root@inst2 / # lparstat -i|grep  "Online Virtual CPU"
Online Virtual CPUs                       : 4
```

After the takeover application `app1` is activated on node `inst2`, and the required CPU and memory settings are acquired for the application, we can see the virtual CPU and entitled capacity settings on Example 3-32 on page 143.

*Example 3-32 Entitled capacity and virtual CPU settings when app1 is activated*

```
root@inst2 / # lparstat -i|grep  "Entitled Capacity"
Entitled Capacity                          : 3.00

root@inst2 / # lparstat -i|grep  "Online Virtual CPU"
Online Virtual CPUs                        : 8
```

## 3.12 Test scenarios

Table 3-3 provides a comprehensive list of PowerHA SystemMirror failover tests. We advise you to run all tests at the end of the cluster configuration.

*Table 3-3 PowerHA SystemMirror test scenarios*

| Tests | Expected results | Results |
|---|---|---|
| PowerHA SystemMirror start:<br>Start the cluster. | PowerHA SystemMirror starts on all nodes. The resource groups are available on their primary nodes. | |
| Resource group move:<br>Move a resource group to the backup node. Repeat this test for all resource groups. | The resource group moves to the backup node while the other resource groups stay online. | |
| Resource group move:<br>Move a resource group back to the primary node. Repeat this test for all resource groups. | The resource group moves back to the primary node while the other resource groups stay online. | |
| PowerHA takeover:<br>Halt the primary node. | The backup node takes over the resource groups of the primary node. The resource groups of the backup server stay online. | |
| Primary node restarts. | According to the resource group definition, either the resource groups of the primary server automatically go back to the original location or you have to manually move them back. | |
| PowerHA takeover:<br>Halt the backup node. | The resource groups of the backup node move to the primary node. The resource groups of the primary node are not affected by the halt of the secondary server. | |
| Backup node restarts. | The resource groups of the primary node are not affected by the restart of the secondary server. | |
| Ethernet card error:<br>Remove one of the Ethernet cables from the server. | The service IP must relocate to the remaining online network interface. | |
| Ethernet network error:<br>Remove all network cables from the primary server. | The resource groups of the primary node are taken over by the backup hosts. | |

| Tests | Expected results | Results |
|---|---|---|
| Application error:<br>Shut down the application to check the application monitoring. | PowerHA must restart the application. | |
| VIOS failover:<br>Turn off one of the VIOS. | All virtual IO must work through the second VIOS. | |
| SAN redundancy:<br>Remove one of the SAN cables. | All disk IO must work through the backup adapter. | |

After you perform each test, ensure that the following capabilities work:

- You can connect to the application through its service IP.
- The shared volume groups and file systems are mounted.
- The resource groups are active on the correct node.
- Applications are running.

# 3.13 Troubleshooting

We introduce the basic PowerHA problem determination techniques.

## 3.13.1 Cluster start-up problems

If your cluster cannot start, check the following areas:

- Network interfaces are set correctly; therefore, they follow the networking rules that are described in “Networking” on page 83.
- Check that multicasting works correctly (most common problem with PowerHA 7.1 at client sites).
- The `/etc/hosts` file is correct.
- You follow the rules for the hostname (see “Initial cluster setup” on page 114).
- Name resolution works.
- All communication interface pairs can ping each other.
- `inetd` is enabled.
- The clcomd and clstrmgrES daemons are active on all nodes.
- Node connectivity: `clrsh` works between the nodes.
- Check the AIX error log (`errpt`).
- Check the cluster configuration on all nodes.
- Run cluster verification and synchronization.

## 3.13.2 Resource group errors

If you encounter resource group errors, check the cluster status first:

- Cluster resource group status (**`/usr/es/sbin/cluster/utilities/clRGinfo -m`**).
- Application status (**`ps -ef`**).
- Persistent and service addresses are configured (**`netstat -in`**).

- Shared volume groups are activated in concurrent mode (**lspv**).
- Shared file systems are mounted (**mount**).
- Repository disk is activated and /aha file system is mounted (**lspv**, **mount**).
- You can log on to the nodes from an external server by using the service address of the nodes.
- Check the /var/hacmp/log/hacmp.out log for errors. For more information, see "The hacmp.out file" on page 146.
- Check the AIX error log (**errpt -a**).

## 3.13.3 Recovering from a script error

If you can locate the problem and resolve it, you can enable PowerHA to continue to process the resource group. For example, a typical problem is that the application cannot stop because the stop script does not handle all application components. If you stop the remaining application processes, you can recover the cluster from the script error:

1. Run **smit sysmirror**. Select **Problem Determination Tools** → **Recover From PowerHA SystemMirror Script Failure**.
2. Select the node where the error occurred.

   Alternatively, you can run **/usr/es/sbin/cluster/utilities/clruncmd <*nodename*>** command.
3. Check the /var/hacmp/log/hacmp.out log file for the result that you want.
4. You might have to run the script recovery on all cluster nodes.

## 3.13.4 Releasing locks that are set by dynamic reconfiguration

If you encounter an error during dynamic reconfiguration and the cluster is locked, you can release the DARE locks. Start **smit sysmirror**. Select **Problem Determination Tools** → **Release Locks Set By Dynamic Reconfiguration**.

## 3.13.5 Comparing active and default configurations

You can compare the active, running configuration of the cluster with the ODM-based configuration. With this option, you can identify the changes that are made since the last cluster verification. Start **smit sysmirror**. Select **Problem Determination Tools** → **Compare Active and Default Configurations**.

Example 3-33 shows the output of the **cl_dare_compare** command after we added a service label (instsvc2) to the inst1rg resource group.

*Example 3-33 Comparing the active and default cluster configurations*

```
root@inst2 / # /usr/es/sbin/cluster/utilities/cl_dare_compare
No unsynchronized changes exist - nothing to compare against.
Cluster services are active
Comparing configuration in
/etc/es/objrepos
against
/usr/es/sbin/cluster/etc/objrepos/active
Contents of HACMPadapter do not match
Current:                                  | New:
```

```
HACMPresource:                          <
        group = "inst1rg"               <
        type = ""                       <
        name = "SERVICE_LABEL"          <
        value = "instsvc2"              <
        id = 20                         <

Found 1 difference in the configuration
```

## 3.13.6 Restoring the PowerHA SystemMirror configuration database from the active configuration

If you unintentionally changed the PowerHA SystemMirror configuration and did not synchronize the cluster, you can restore the original settings from the active, running configuration.

Start **`smit sysmirror`** and select **Problem Determination Tools** → **Restore PowerHA SystemMirror Configuration Database from Active Configuration**.

## 3.13.7 The hacmp.out file

The `/var/hacmp/log/hacmp.out` log file is our best source of information about the cluster events. If you encounter any cluster startup, shutdown, or takeover (resource group movement) problem, you have to look in this file for details. Use the **`grep "EVENT START" /var/hacmp/log/hacmp.out`** command to see what happened on the cluster. Example 3-34 shows the PowerHA events during a node startup.

*Example 3-34 Looking for events in the /var/hacmp/log.hacmp.out file*

```
root@inst1 /var/hacmp/log # grep "EVENT START" hacmp.out
Mar 26 13:56:23 EVENT START: node_up inst1
Mar 26 13:56:26 EVENT START: rg_move_fence inst1 1
Mar 26 13:56:26 EVENT START: rg_move_acquire inst1 1
Mar 26 13:56:26 EVENT START: rg_move inst1 1 ACQUIRE
Mar 26 13:56:27 EVENT START: acquire_service_addr
Mar 26 13:56:28 EVENT START: acquire_aconn_service en2 net_ether_01
Mar 26 13:56:31 EVENT START: rg_move_complete inst1 1
Mar 26 13:56:31 EVENT START: start_server app1
Mar 26 13:56:35 EVENT START: node_up_complete inst1
Mar 26 13:56:46 EVENT START: node_up inst2
Mar 26 13:56:50 EVENT START: rg_move_fence inst1 2
Mar 26 13:56:50 EVENT START: rg_move_acquire inst1 2
Mar 26 13:56:50 EVENT START: rg_move inst1 2 ACQUIRE
Mar 26 13:56:59 EVENT START: rg_move_complete inst1 2
Mar 26 13:57:03 EVENT START: node_up_complete inst2
```

Because the hacmp.out file can be large, the best way to work with it is to search for the “`!!!!!!!!!! ERROR !!!!!!!!!!`” string. The real error messages are usually a few lines above this string. Example 3-35 shows a part of the `hacmp.out` file with **`varyonvg`** errors.

*Example 3-35 hacmp.out errors*

```
+inst1rg:cl_pvo:app1vg[17] rc=20
+inst1rg:cl_pvo:app1vg[18] : exit status of varyonvg -n -c -P app1vg is: 20
```

```
+inst1rg:cl_pvo:app1vg[20] (( 20 == 20 ))
+inst1rg:cl_pvo:app1vg[25] cl_mirrorset app1vg
+inst1rg:cl_mirrorset[+52] [[ high = high ]]
+inst1rg:cl_mirrorset[+52] version=1.5
+inst1rg:cl_mirrorset[+55] vgname=app1vg
+inst1rg:cl_mirrorset[+56] typeset -i mirrorset=1
+inst1rg:cl_mirrorset[+59] [[ -z  ]]
+inst1rg:cl_mirrorset[+62] grep -w HACMP_MIRROR_VARYON /etc/environment
+inst1rg:cl_mirrorset[+62] grep -iw TRUE
+inst1rg:cl_mirrorset[+62] eval
+inst1rg:cl_mirrorset[+65] echo
+inst1rg:cl_mirrorset[+65] grep -iqw TRUE
+inst1rg:cl_mirrorset[+65] [[  = false ]]
+inst1rg:cl_mirrorset[+65] [[ -z inst1rg ]]
+inst1rg:cl_mirrorset[+87] +inst1rg:cl_mirrorset[+87] odmget -q group = inst1rg
and name = FORCED_VARYON and    value = true HACMPresource
FORCED_VARYON=
+inst1rg:cl_mirrorset[+87] [[ -z  ]]
+inst1rg:cl_mirrorset[+89] return 1
+inst1rg:cl_pvo:app1vg[34] (( 20 != 0 ))
+inst1rg:cl_pvo:app1vg[35] cl_log 296 'cl_pvo: Failed to vary on volume group
app1vg in passive mode' cl_pvo app1vg
+inst1rg:cl_log[+50] version=1.10
+inst1rg:cl_log[+94] SYSLOG_FILE=/var/hacmp/adm/cluster.log
***************************
Mar 8 2012 14:06:04 !!!!!!!!!! ERROR !!!!!!!!!!
***************************
Mar 8 2012 14:06:04 cl_pvo: Failed to vary on volume group app1vg in passive mode
+inst1rg:cl_pvo:app1vg[44] return 0
+inst1rg:cl_pvo:app1vg[504] return 0
+inst1rg:clvaryonvg[827] : Let us assume that the old style synclvodm would sync
all the PV/FS changes.
+inst1rg:clvaryonvg[829] expimpvg_notrequired=1
```

### 3.13.8 Application monitoring log files

The application monitor files are in the `/var/hacmp/log` directory. The file name is `clappmond.application_name.resource_group_name.monitor.log`.

### 3.13.9 Other log files

We show the locations of more log files:

- `/var/log/clcomd/clcomd.log`: Cluster communication daemon log file, which is useful to check node connectivity and global ODM problems
- `/var/hacmp/log/cspoc.log`: C-SPOC commands log file helps you debug why a PowerHA command failed on a node
- `/var/hacmp/adm/cluster.log`: PowerHA cluster log file, which is mostly RSCT messages
- `/var/adm/ras/syslog.caa`: CAA cluster log file, which is useful to check why the **`mkcluster`** command failed
- `/var/hacmp/log/clstrmgr.debug`: Cluster manager debug file
- `/var/hacmp/log/clavan.log`: Cluster events

- /var/hacmp/log/cl_event_summaries.txt: Cluster events summary
- /var/hacmp/log/autoverify.log: Daily automatic cluster verification log file

### 3.13.10 Contacting IBM support

If you encounter a serious issue with your cluster, and you want to contact IBM support, collect the following information:

- IBM customer number
- Server type and serial number
- PowerHA snap
- Description of the events as clearly as possibly
- Indication of the time of the events, which is important because sometimes we cannot decide which cluster test or event caused the problem

### 3.13.11 Collecting PowerHA snap

We show how to collect the PowerHA snap:

1. Remove the earlier snap information from the nodes: Run **snap -r** on all cluster nodes to clean up the /tmp/ibmsupt directory structure.
2. Collect AIX snap: Execute **snap -a** on all nodes. The snap files are collected under the /tmp/ibmsupt directory.
3. Run PowerHA snap: Issue the **snap -e** command only on one node to collect the PowerHA configuration from all nodes. The output files are in the /tmp/ibmsupt/hacmp directory.
4. Gather the snap files from the /tmp/ibmsupt directory structure into a tar archive. The file name must contain the node name, for example, inst1.snap.tar.Z.
5. Rename the files so that they start with the IBM PMR number. The PMR number is provided by the IBM Technical Support personnel.

> **Important:** The PowerHA **snap -e** must be collected as soon as possible after the problem to avoid log file wraps and to prevent the loss of useful information to debug the problem.

# 4

# Cluster administration

In this chapter, we discuss the administration tasks that relate to the PowerHA SystemMirror daily operations:

- AIX and PowerHA SystemMirror service pack upgrade
- Reconfiguring the cluster online
- Cross-site mirroring and AIX LVM Mirror pools
- Critical volume groups (voting disks) for Oracle RAC
- Federated security for cluster-wide security management

## 4.1 Introduction

Most of the daily administration tasks can be performed from the **smit cspoc** menu as seen in Figure 4-1. Expert administrators can use the Cluster Single Point of Control (C-SPOC) command line. For easy understanding, we use **smit cspoc** in all of our examples.

```
 System Management (C-SPOC)

Move cursor to desired item and press Enter.

  Storage
  PowerHA SystemMirror Services
  Communication Interfaces
  Resource Groups and Applications
  PowerHA SystemMirror Logs
  File Collections
  Security and Users
  LDAP

  Open a SMIT Session on a Node

F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-1 smit cspoc*

## 4.2 AIX and PowerHA SystemMirror service pack upgrade

**Important:** Install the latest AIX and PowerHA fixes regularly (twice a year).

You can download the AIX and PowerHA SystemMirror service packs from the IBM FixCentral website:

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

Install the latest AIX and PowerHA SystemMirror fixes regularly. Because the AIX service packs contain a new kernel, you cannot avoid a system restart and a short downtime. Also, this scheduled downtime is an opportunity to run takeover tests and ensure that the PowerHA configuration is correct.

You can install PowerHA SystemMirror fixes without any downtime by following these steps:

1. Stop the cluster services on a node that hosts the application.

   After all upgraded PowerHA SystemMirror cluster nodes are running and the cluster is stable, bring node A offline without bringing its associated resource groups offline by stopping cluster services by using the Unmanage Resource Groups option. (Before PowerHA SystemMirror 5.4, we called this approach "*forced down*".)

2. Install the PowerHA SystemMirror Service Pack.

   On node A, apply the PowerHA SystemMirror Service Pack update, which updates the PowerHA SystemMirror configuration database (Object Data Managers (ODMs)).

3. Start cluster services on the upgraded node.

   Start cluster services on node A. Node A is running the updated PowerHA SystemMirror version while nodes B, C, and others are running the previous version of PowerHA SystemMirror. The upgraded version of PowerHA SystemMirror starts monitoring the running applications and resources on node A. When you run the **`clstart`** command, PowerHA SystemMirror executes the application controller start script, unless there is a configured application monitor.

4. Verify the upgraded cluster definition.

   After all nodes are running and the cluster is stable, run cluster verification on the upgraded PowerHA SystemMirror cluster.

## 4.3 Reconfiguring the cluster online

You can perform any cluster resource reconfiguration tasks online, while the cluster is up and running and the application is online. In our installation 3.7, "PowerHA SystemMirror installation and prerequisites" on page 91, we reconfigured the resources groups while the cluster was up and running.

We encourage you to do all PowerHA SystemMirror reconfiguration while the cluster is running. For example, you can add, change, or remove the following objects:

- Nodes
- Networks
- Resources
- Application servers
- Application monitors
- Resource groups

There is only one limitation: you cannot change a resource while it serves an online resource group. For example, you cannot change the service address of the online resource group. The online reconfiguration is safe. PowerHA warns you or prevents the modification that risks the online resource groups and application.

After the configuration changes, you have to run PowerHA cluster synchronization and verification. PowerHA cluster synchronization and verification starts the Dynamic Reconfiguration (DARE) process, which activates the changes.

## 4.4 Verifying the cluster automatically

> **Important:** Enable the daily automatic cluster verification. Remember to check the output file regularly.

The PowerHA SystemMirror can run automatic cluster verification every day. You can set it up by starting **`smit sysmirror`**. Select **Problem Determination Tools → PowerHA SystemMirror Verification → Automatic Cluster Configuration Monitoring**. See Figure 4-2 on page 152. You can modify the daily schedule and the node that runs the test and hosts the log files. The `Default` node is the first node in alphabetical order in the cluster configuration.

```
 Automatic Cluster Configuration Monitoring

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

                                                        [Entry Fields]
* Automatic cluster configuration verification        Enabled              +
* Node name                                           Default              +
* HOUR (00 - 23)                                     [00]                  +#
  Debug                                               no                   +

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

*Figure 4-2 Automatic cluster configuration monitoring setup*

In a verification error, PowerHA sends a broadcast message to all users. The automatic cluster verification log files are at these locations:

- `/var/hacmp/log/autoclstrcfgmonitor.out`
- `/var/hacmp/log/autoverify.log`
- `/var/hacmp/clverify.log`

## 4.5 Regularly testing your cluster

Regularly check your cluster and run takeover tests. For a list of test cases, see 3.12, “Test scenarios” on page 143. From our experience, we know that most cluster configurations and applications evolve over time. Not all application and system administrators are aware of the related required cluster configuration tasks. For example, we show some of the typical user errors on long-running clusters that can be avoided by performing regular takeover tests:

- User and user ID mismatch between nodes
- Missing users on the backup node
- Forgotten passwords on the backup nodes
- Missing application components on the backup nodes
- Different versions of the application across the cluster nodes
- Application startup and shutdown scripts no longer work correctly
- Incorrectly mounted non-clustered file systems, Network File System (NFS), or Samba-shares and exports (the CD-ROM is mounted under the shared file system)
- New firewalls, or firewall rules between the cluster nodes
- Erroneous `/etc/hosts` file
- Ping address that is used in `netmon.cf` or in the EtherChannel configuration no longer works

Check the AIX error log (**`errpt`**) daily to avoid trouble, such as realizing that the disk failure errors started a year ago.

## 4.6 Backing up the cluster configuration

PowerHA SystemMirror snapshot facility provides a convenient method to back up and restore the cluster configuration. If you plan to modify your cluster, you can create a snapshot first. If you are not happy with the new configuration, you can easily restore the configuration from the snapshot.

The snapshot creates two files: `info` and `odm`:

- The `info` file contains readable output of the cluster configuration and useful information, including the network setup and the shared Logical Volume Manager (LVM) information. You can use this file to document the cluster.
- The `odm` file is a full copy of the PowerHA ODM files.

> **Important:** Back up your cluster configuration data regularly.

The following steps create a PowerHA snapshot:

1. Start **`smit sysmirror`** and select **Cluster Nodes and Networks → Manage the Cluster → Snapshot Configuration → Create a Snapshot of the Cluster Configuration**.
2. On the SMIT panel (Figure 4-3), enter a meaningful name for the snapshot. This name is the name of the snapshot files (`name.info` and `name.odm`).
3. You can specify a custom snapshot method, and a short description of the backup.
4. Do not include the cluster log files in the snapshot.
5. The snapshot files are created in the `/usr/es/sbin/cluster/snapshots` directory.

```
 Create a Snapshot of the Cluster Configuration

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

                                                         [Entry Fields]
* Cluster Snapshot Name                               [inst-20120404]         /
  Custom-Defined Snapshot Methods                     []                     +
  Save Cluster Log Files in snapshot                   No                    +
* Cluster Snapshot Description                        [Installation demo clu>

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

*Figure 4-3 PowerHA snapshot*

For data safety, we perform this backup regularly:

1. Back up the system by using the **`mksysb`** command.
2. Create an AIX snap with the **`snap -ac`** command and save the `/tmp/ibmsupt/snap.pax.Z` file in a safe place. The snap data contains all AIX configuration settings.
3. Back up the PowerHA cluster configuration by using the snapshot facility. Save the files in a safe place.

### 4.6.1 Restoring the cluster configuration from a snapshot

We describe how to restore the cluster configuration from a PowerHA snapshot:

1. Start **`smit sysmirror`** and select **Cluster Nodes and Networks → Manage the Cluster → Snapshot Configuration → Restore the Cluster Configuration From a Snapshot**.
2. Select the snapshot to restore. The snapshot files are in the `/usr/es/sbin/cluster/snapshots` directory.
3. Select **Yes** to “Un/Configure Cluster Resources”. See Figure 4-4 for the screen capture of the SMIT panel.

```
 Restore the Cluster Snapshot

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

                                                        [Entry Fields]
  Cluster Snapshot Name                              inst-20120404
  Cluster Snapshot Description                       Installation demo clu>
  Un/Configure Cluster Resources?                   [Yes]                   +
  Force apply if verify fails?                      [No]                    +

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

*Figure 4-4 Restore the cluster snapshot*

## 4.7 Monitoring the cluster

There are several commands that show the cluster and its resource group status. We show the most useful commands.

### 4.7.1 Clstat

**`Clstat`** is an interactive utility that shows the cluster status online. It requires the clinfo daemon and Simple Network Management Protocol (SNMP) Version 3.

Figure 4-5 on page 155 shows the output of the **`clstat`** command.

```
              clstat - HACMP Cluster Status Monitor
              -------------------------------------

Cluster: inst1_cluster  (1089556542)
Fri Apr  6 15:43:02 EDT 2012
              State: UP               Nodes: 2
              SubState: STABLE


       Node: inst1           State: UP
          Interface: inst1b1 (0)            Address: 10.1.1.39
                                            State:   UP
          Interface: inst1backup (1)        Address: 10.1.2.39
                                            State:   UP
          Interface: inst1b2 (0)            Address: 10.10.1.39
                                            State:   UP
          Interface: instsvc1 (0)           Address: 172.16.21.67
                                            State:   UP
          Resource Group: inst1rg                   State:  On line

       Node: inst2           State: UP
          Interface: inst2b1 (0)            Address: 10.1.1.47
                                            State:   UP
          Interface: inst2backup (1)        Address: 10.1.2.47
                                            State:   UP
          Interface: inst2b2 (0)            Address: 10.10.1.47
                                            State:   UP
************************ f/forward, b/back, r/refresh, q/quit *****************
```

*Figure 4-5 Clstat output*

### 4.7.2 Cldump

**Cldump** commands provide similar detailed information about the cluster status. It is not an interactive program; therefore, you can use it in shell scripts. It requires **clinfo** and SNMP Version 1. See Example 4-1 for the **cldump** output in our sample cluster.

*Example 4-1 Cldump output*

```
root@inst2 /usr/es/sbin/cluster/utilities # ./cldump

Obtaining information via SNMP from Node: inst2...

_____________________________________________________________________________
Cluster Name: inst1_cluster
Cluster State: UP
Cluster Substate: STABLE
_____________________________________________________________________________

Node Name: inst1               State: UP

  Network Name: net_ether_01      State: UP
```

```
    Address: 10.1.1.39       Label: inst1b1           State: UP
    Address: 10.10.1.39      Label: inst1b2           State: UP
    Address: 172.16.21.67    Label: instsvc1          State: UP

  Network Name: net_ether_02       State: UP

    Address: 10.1.2.39       Label: inst1backup       State: UP

Node Name: inst2                State: UP

  Network Name: net_ether_01       State: UP

    Address: 10.1.1.47       Label: inst2b1           State: UP
    Address: 10.10.1.47      Label: inst2b2           State: UP

  Network Name: net_ether_02       State: UP

    Address: 10.1.2.47       Label: inst2backup       State: UP


Cluster Name: inst1_cluster

Resource Group Name: inst1rg
Startup Policy: Online On First Available Node
Fallover Policy: Fallover To Next Priority Node In The List
Fallback Policy: Fallback To Higher Priority Node In The List
Site Policy: ignore
Primary instance(s):
The following node temporarily has the highest priority for this instance:
inst1, user-requested rg_move performed on Tue Apr  3 15:57:48 2012

Node                         Group State
---------------------------- ---------------
inst1                        ONLINE
inst2                        OFFLINE
```

### 4.7.3 ClRGinfo

The **/usr/es/sbin/cluster/utilities/clRGinfo** command shows the resource group and the application status. The command does not require SNMP or **clinfo**. **clRGinfo -s** creates colon-separated output that is easy to process in a shell script. Example 4-2 shows the **clRGinfo** command.

*Example 4-2 Checking the status of the resource groups*

```
root@inst2 / # /usr/es/sbin/cluster/utilities/clRGinfo  -m
---------------------------------------------------------------------------
Group Name     Group State             Application state          Node
---------------------------------------------------------------------------
inst1rg        OFFLINE
 app1                                  OFFLINE

inst2rg        ONLINE                                             inst2
 app2                                  ONLINE MONITORED
```

```
root@inst2 / # /usr/es/sbin/cluster/utilities/clRGinfo  -s
inst1rg:OFFLINE:inst1:non-concurrent:OFAN:FNPN:FBHPN:ignore:OFFLINE:: : :::
inst1rg:OFFLINE:inst2:non-concurrent:OFAN:FNPN:FBHPN:ignore:OFFLINE:: : :::
inst2rg:ONLINE:inst2:non-concurrent:OHN:FNPN:NFB:ignore::: : :::
inst2rg:OFFLINE:inst1:non-concurrent:OHN:FNPN:NFB:ignore::: : :::
```

## 4.8 Resource group and application management

In this chapter, we describe the resource group and application management tools in PowerHA SystemMirror. We present the following topics:

- Bringing a resource group online
- Bringing a resource group offline
- Moving a resource group
- Suspending or resuming application monitoring
- Application analysis

### 4.8.1 Bringing a resource group online

To start a resource group and its application, execute the following commands:

1. Start **smitty cspoc** and click **Resource Groups and Applications** → **Bring a Resource Group Online**. See Figure 4-6.
2. Select the resource group. Only the offline resource groups are shown in the pickup list.
3. Select the destination node. An asterisk shows the primary node for that resource group.
4. Wait for the cluster to stabilize.

```
 Bring a Resource Group Online

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

                                                         [Entry Fields]
  Resource Group to Bring Online                       inst1rg
  Node on Which to Bring Resource Group Online         inst1

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

*Figure 4-6 Bring a resource group online*

### 4.8.2 Bringing a resource group offline

You can take a resource group and its corresponding application and resources offline:

1. Start **smitty cspoc** and select **Resource Groups and Applications** → **Bring a Resource Group Offline**. See Figure 4-7 on page 158.
2. Select the resource group.
3. Select the online node, which is only important when you have concurrent resource groups.

4. Wait for the cluster to stabilize.

```
 Bring a Resource Group Offline

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

                                                          [Entry Fields]
  Resource Group to Bring Offline                       inst1rg
  Node On Which to Bring Resource Group Offline         inst1

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

*Figure 4-7 Bring a resource group offline*

## 4.8.3 Moving a resource group

To move a resource group and an application between the nodes, follow these steps:

1. Start **`smitty cspoc`** and click **Resource Groups and Applications** → **Move Resource Groups to Another Node**. See Figure 4-8.
2. Select the resource group. The pickup list shows the current location and status of the resource group.
3. Select the destination node. An asterisk shows the primary node for that resource group.
4. Wait for the cluster to stabilize.

```
 Move Resource Group(s) to Another Node

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

                                                          [Entry Fields]
  Resource Group(s) to be Moved                         inst1rg
  Destination Node                                      inst2

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

*Figure 4-8 Move a resource group to another node*

## 4.8.4 Suspending or resuming application monitoring

**Important:** Your application administrator must be aware of the PowerHA application monitoring feature.

Before you perform any changes to the application environment that require application downtime, suspend the application monitoring.

You have to suspend the application monitoring if you plan to perform any administration task that requires application downtime while the PowerHA resources are active, such as applying fixes for the application.

### Suspending application monitoring

The following steps show how to suspend the application monitoring:

1. Start `smitty cspoc` and select **Resource Groups and Applications** → **Suspend/Resume Application Monitoring** → **Suspend Application Monitoring**. See Figure 4-9.
2. Select the application controller.
3. Select the resource group for the application. One application controller can be in more than one resource group.

```
 Suspend Application Monitoring

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

                                                        [Entry Fields]
* Application Controller Name                         app1
* Resource Group                                     [inst1rg]            +

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

*Figure 4-9 Suspend application monitoring*

### Resuming application monitoring

The following steps show how to resume application monitoring:

1. Start `smitty cspoc` and select **Resource Groups and Applications** → **Suspend/Resume Application Monitoring** → **Resume Application Monitoring**.
2. Select the application controller.
3. Select the resource group for the application.

## 4.8.5 Application analysis

This tool provides regular reports about the application availability. Follow these steps to configure this tool:

1. Start `smitty cspoc`. Select **Resource Groups and Applications** → **Application Availability Analysis**.
2. Select the application.
3. Provide the start date for the analysis.
4. Enter the end date.

See the application availability analysis report for our sample cluster as shown in Figure 4-10 on page 160.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.


Analysis begins:           Friday, 23-March-2012, 00:00
Analysis ends:             Friday, 30-March-2012, 23:59
Application analyzed:      app2

Total time:                7 days, 23 hours, 59 minutes, 59 seconds

Uptime:
        Amount:            4 days, 12 hours, 30 minutes, 42 seconds
        Percentage:        56.52%
        Longest period:    4 days,  7 hours, 35 minutes, 17 seconds

Downtime:
        Amount:            3 days, 11 hours, 29 minutes, 17 seconds
        Percentage:        43.48%
        Longest period:    3 days, 10 hours, 29 minutes, 14 seconds

Application monitoring was suspended for 45.20% of the time period analyzed.

Application monitor failed during the time period analyzed.

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 4-10 Application availability analysis*

## 4.9 New application controller foreground startup

In versions before PowerHA 7.1.1, the start scripts of the application controllers run in the background and their exit code is ignored. It is still the default of PowerHA 7.1.1, but you can choose now for foreground execution instead. Foreground execution causes the cluster event processing to wait for the completion of the application controller start script. With foreground execution, your application controller start script is called synchronously during the node up event, and execution of the event stops until the controller start script completes. If the start script exits with a non-zero exit code, the node up event fails.

The exit code of the script is not checked in the PowerHA 7.1.1 base version. SP1 changed this situation and a non-zero exit causes now an event error. Also, SP1 added timestamps and tracing of the startup sequence in `hacmp.out`.

Foreground startup simplifies the design of start scripts and allows the sequencing of resource groups with dependencies. In multiple resource groups with dependencies or sequential processing, the user can ensure that one is started before the next one.

Poorly designed scripts can cause the script to stop (`config_too_long`). Examine carefully the start script before you configure foreground startup. If there is any possibility of the script stopping, use a combination of the background startup option with a startup monitor instead of foreground startup.

### 4.9.1 Configuring application controller foreground startup

The application controller startup mode parameter can be selected in the application controller configuration panel as shown in Figure 4-11. The parameter is available with the path **smit sysmirror** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Controller Scripts** → **Add Application Controller Scripts**.

```
 Add Application Controller Scripts

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

                                                     [Entry Fields]
* Application Controller Name                     [db_app]
* Start Script                                    [/HA/db_start.sh]
* Stop Script                                     [/HA/db_stop.sh]
  Application Monitor Name(s)                                            +
  Application startup mode                        [foreground]           +





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

*Figure 4-11 Application controller configuration panel*

We configure two application controllers, one for each of the resource groups of our test cluster as detailed in Example 4-3. The resource groups are in a parent/child dependency.

*Example 4-3 Details for application controllers and resource groups*

```
root@migr3 / # cllsserv -h
Name     Start Script      Stop Script     Startup mode
db_app   /HA/db_start.sh   /HA/db_stop.sh  foreground
ihs_app  /HA/ihs_start.sh  /HA/ihs_stop.sh background
root@migr3 / #
root@migr3 / # clrgdependency -t'PARENT_CHILD' -sl
#Parent                          Child
db_rg                            ihs_rg
root@migr3 / #
```

The test cluster configuration is illustrated in Figure 4-12 on page 162.

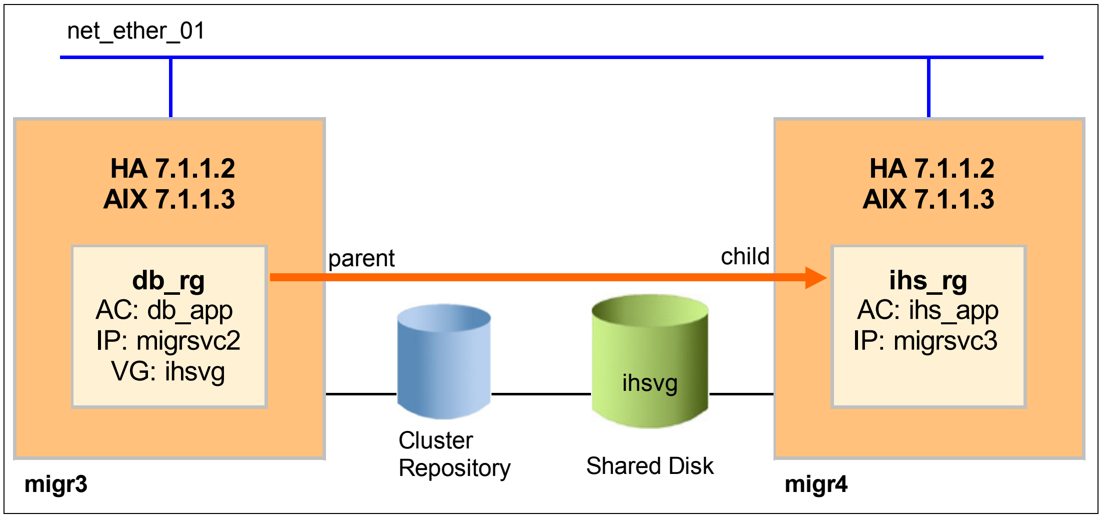



*Figure 4-12 Test cluster configuration*

Example 4-4 shows the cluster topology and the cluster resource details.

*Example 4-4 Cluster topology and resource details*

```
root@migr3 / # cldisp
Cluster: migr34
   Cluster services: inactive

#############
APPLICATIONS
#############
   Cluster migr34 provides the following applications: db_app ihs_app
      Application: db_app
      db_app is started by /HA/db_start.sh
      db_app is stopped by /HA/db_stop.sh
         No application monitors are configured for db_app.
         This application is part of resource group 'db_rg'.
            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 db_app: migr3 migr4
            Resources associated with db_app:
               Service Labels
                  migrsvc2(172.16.21.68) {}
                     Interfaces configured to provide migrsvc2:
                        migr3 {}
                           with IP address: 172.16.21.32
                           on interface: en0
                           on node: migr3 {}
                           on network: net_ether_01 {}
                        migr4 {}
                           with IP address: 172.16.21.40
                           on interface: en0
                           on node: migr4 {}
                           on network: net_ether_01 {}
               Shared Volume Groups:
```

```
                 ihsvg

      Application: ihs_app
      ihs_app is started by /HA/ihs_start.sh
      ihs_app is stopped by /HA/ihs_stop.sh
         No application monitors are configured for ihs_app.
         This application is part of resource group 'ihs_rg'.
            Resource group policies:
               Startup: on first available node
               Fallover: to next priority node in the list
               Fallback: never
            Nodes configured to provide ihs_app: migr4 migr3
            Resources associated with ihs_app:
               Service Labels
                  migrsvc3(172.16.21.69) {}
                     Interfaces configured to provide migrsvc3:
                        migr4 {}
                           with IP address: 172.16.21.40
                           on interface: en0
                           on node: migr4 {}
                           on network: net_ether_01 {}
                        migr3 {}
                           with IP address: 172.16.21.32
                           on interface: en0
                           on node: migr3 {}
                           on network: net_ether_01 {}

#############
TOPOLOGY
#############
   migr34 consists of the following nodes: migr3 migr4
      migr3
         Network interfaces:
            migr3 {}
               with IP address: 172.16.21.32
               on interface: en0
               on network: net_ether_01 {}
      migr4
         Network interfaces:
            migr4 {}
               with IP address: 172.16.21.40
               on interface: en0
               on network: net_ether_01 {}
root@migr3 / #
```

### 4.9.2 Testing application controller foreground startup

For this test, both resource groups are temporarily configured with the same home node. This way, their application scripts log messages in the same file so that you can see the detailed sequence of their start and finish.

Example 4-5 on page 164 shows the start and stop scripts. The syslog configuration is made to log the messages through the `local7` facility in the `/var/hacmp/log/foreground_startup.log` log file.

*Example 4-5 Dummy start and stop scripts*

```
root@migr3 /HA # ls
clstrmgr      db_start.sh   db_stop.sh    ihs_start.sh  ihs_stop.sh
root@migr3 /HA # cat db_start.sh
#!/bin/ksh
fp="local7.info"
file="`expr "//$0" : '.*/\([^/]*\)'`"

# cleanup
if [ -f /ihsfs/db.lck ]; then rm /ihsfs/db.lck; fi
logger -t"$file" -p$fp "Starting up DB... "
sleep 50
echo "DB started at:\n\t`date`">/ihsfs/db.lck
logger -t"$file" -p$fp "DB is running!"
exit 0

root@migr3 /HA # cat db_stop.sh
#!/bin/ksh
fp="local7.info"
file="`expr "//$0" : '.*/\([^/]*\)'`"

logger -t"$file" -p$fp "Shutting down APP... "
sleep 2
# cleanup
if [ -f /ihsfs/app.lck ]; then rm /ihsfs/app.lck; fi
logger -t"$file" -p$fp "APP stopped!"
exit 0

root@migr3 /HA # cat ihs_start.sh
#!/bin/ksh
fp="local7.info"
file="`expr "//$0" : '.*/\([^/]*\)'`"

# cleanup
if [ -f /ihsfs/app.lck ]; then rm /ihsfs/app.lck; fi
logger -t"$file" -p$fp "Starting up APP..."
sleep 10
echo "APP started at:\n\t`date`">/ihsfs/app.lck
logger -t"$file" -p$fp "APP is running!"
exit 0

root@migr3 /HA # cat ihs_stop.sh
#!/bin/ksh
fp="local7.info"
file="`expr "//$0" : '.*/\([^/]*\)'`"

logger -t"$file" -p$fp "Shutting down DB... "
sleep 20
# cleanup
if [ -f /ihsfs/db.lck ]; then rm /ihsfs/db.lck; fi
logger -t"$file" -p$fp "DB stopped!"
exit 0

root@migr3 /HA # grep local7 /etc/syslog.conf
local7.info /var/hacmp/log/foreground_startup.log rotate size 256k files 4
root@migr3 /HA #
```

With the db_app application controller in the background, the APP startup script is launched before the DB startup script returns as shown in Example 4-6.

*Example 4-6 Startup script markers in background mode*

```
Apr 25 13:12:21 migr3 local7:info db_start.sh: Starting up DB...
Apr 25 13:12:26 migr3 local7:info ihs_start.sh: Starting up APP...
Apr 25 13:12:36 migr3 local7:info ihs_start.sh: APP is running!
Apr 25 13:13:11 migr3 local7:info db_start.sh: DB is running!
```

You can see in Example 4-7 the state of the resource groups at various moments interlaced with the messages of the startup scripts.

*Example 4-7 Startup sequence in background mode*

```
Wed Apr 25 13:12:17 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         OFFLINE                      migr3
               OFFLINE                      migr4
db_rg          OFFLINE                      migr3
               OFFLINE                      migr4

Wed Apr 25 13:12:18 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         OFFLINE                      migr3
               OFFLINE                      migr4
db_rg          ACQUIRING                    migr3
               OFFLINE                      migr4

Apr 25 13:12:21 migr3 local7:info db_start.sh: Starting up DB...

Wed Apr 25 13:12:22 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         OFFLINE                      migr3
               OFFLINE                      migr4
db_rg          ONLINE                       migr3
               OFFLINE                      migr4

Wed Apr 25 13:12:24 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ACQUIRING                    migr3
               OFFLINE                      migr4
db_rg          ONLINE                       migr3
               OFFLINE                      migr4

Apr 25 13:12:26 migr3 local7:info ihs_start.sh: Starting up APP...

Wed Apr 25 13:12:27 EDT 2012
```

```
-------------------------------------------------------------------------------
Group Name     Group State                 Node
-------------------------------------------------------------------------------
ihs_rg         ONLINE                      migr3
               OFFLINE                     migr4
db_rg          ONLINE                      migr3
               OFFLINE                     migr4

Apr 25 13:12:36 migr3 local7:info ihs_start.sh: APP is running!
Apr 25 13:13:11 migr3 local7:info db_start.sh: DB is running!
```

In application controller foreground mode, the APP startup script is launched after the DB startup script returns as shown in Example 4-8.

*Example 4-8   Startup script markers in foreground mode*

```
Apr 25 12:50:46 migr3 local7:info db_start.sh: Starting up DB...
Apr 25 12:51:36 migr3 local7:info db_start.sh: DB is running!
Apr 25 12:51:41 migr3 local7:info ihs_start.sh: Starting up APP...
Apr 25 12:51:51 migr3 local7:info ihs_start.sh: APP is running!
```

Example 4-9 shows the state of the resource groups at various moments interlaced with the messages of the startup scripts.

*Example 4-9   Startup sequence in foreground mode*

```
Wed Apr 25 12:50:41 EDT 2012
-------------------------------------------------------------------------------
Group Name     Group State                 Node
-------------------------------------------------------------------------------
ihs_rg         OFFLINE                     migr3
               OFFLINE                     migr4
db_rg          OFFLINE                     migr3
               OFFLINE                     migr4

Wed Apr 25 12:50:43 EDT 2012
-------------------------------------------------------------------------------
Group Name     Group State                 Node
-------------------------------------------------------------------------------
ihs_rg         OFFLINE                     migr3
               OFFLINE                     migr4
db_rg          ACQUIRING                   migr3
               OFFLINE                     migr4

Apr 25 12:50:46 migr3 local7:info db_start.sh: Starting up DB...
Apr 25 12:51:36 migr3 local7:info db_start.sh: DB is running!

Wed Apr 25 12:51:37 EDT 2012
-------------------------------------------------------------------------------
Group Name     Group State                 Node
-------------------------------------------------------------------------------
ihs_rg         OFFLINE                     migr3
               OFFLINE                     migr4
db_rg          ONLINE                      migr3
               OFFLINE                     migr4
```

```
Wed Apr 25 12:51:39 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ACQUIRING                    migr3
               OFFLINE                      migr4
db_rg          ONLINE                       migr3
               OFFLINE                      migr4

Apr 25 12:51:41 migr3 local7:info ihs_start.sh: Starting up APP...

Wed Apr 25 12:51:42 EDT 2012
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ONLINE                       migr3
               OFFLINE                      migr4
db_rg          ONLINE                       migr3
               OFFLINE                      migr4

Apr 25 12:51:51 migr3 local7:info ihs_start.sh: APP is running!
```

# 4.10 Disk management in PowerHA SystemMirror

You cannot use standard AIX LVM commands for disks that are configured in the cluster. PowerHA SystemMirror C-SPOC provides the same LVM functionality as standard AIX LVM. We introduce the Cluster Logical Volume Manager (CLVM) and describe how to perform the typical daily LVM tasks.

The basic parameters for the CLVM commands are the same as the corresponding AIX LVM commands. If you are in doubt, check the man pages for the normal LVM command.

## 4.10.1 Adding a disk to the cluster nodes

We describe how to add a disk to the cluster nodes:

1. Create the LUN and map it to the cluster nodes.
2. Run **`cfgmgr`** on all nodes, one-by-one.
3. Check for the SCSI Reserve policy by using **`lsattr -El hdiskX`**. If the `reserve_policy` is not set to `no_reserve`, change it by using **`chdev -a reserve_policy=no_reserve -l hdiskX`**. You have to perform this step on all cluster nodes.
4. Create the Physical Volume Identifier by running the **`chdev -l hdiskX pv=yes`** command on all nodes.

## 4.10.2 Creating a shared volume group

We show how to create a shared volume group:

1. Start **`smit cspoc`**, and select **Storage** → **Volume Groups** → **Create a Volume Group**. See Figure 4-13 on page 168.

2. Select the node names where you want to create the volume group. By default, you have to select all nodes.
3. Select the disk. PowerHA C-SPOC shows the physical volume ID (PVID) and the disk name to help you.
4. Choose the type of volume group. We suggest scalable volume groups.
5. Optional: You can provide a resource group name. You can select one of the existing resource groups or if you enter a new resource group name, C-SPOC creates a resource group. However, this method is not the best way to create a resource group.

```
 Create a Scalable Volume Group

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

                                                         [Entry Fields]
  Node Names                                           inst1,inst2
  Resource Group Name                                 [inst1rg]             +
  PVID                                                 00f74d474f69ea30
  VOLUME GROUP name                                   [databasevg]
  Physical partition SIZE in megabytes                 64                   +
  Volume group MAJOR NUMBER                           [101]                  #
  Enable Fast Disk Takeover or Concurrent Access       Fast Disk Takeover   +
  Volume Group Type                                    Scalable
  CRITICAL volume group?                               no                   +

  Max PPs per VG in units of 1024                      32                   +
  Max Logical Volumes                                  256                  +

  Enable Strict Mirror Pools                           No                   +
  Mirror Pool name                                    []

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

*Figure 4-13 Creating a shared volume group*

6. Type the volume group name.
7. Enter the typical volume group parameters (you can leave the default value):
   - Physical partition (PP) size
   - Max PPs per VG
   - Max Logical Volumes
   - Enable Strict Mirror Pools
   - Mirror Pool name. Check “Cross-site mirroring and AIX LVM Mirror pools” on page 183.
8. Supply the major number. The major number must be available on all nodes. For more information about the major number, see “Creating data volume groups and file systems” on page 108.
9. Select Enable Fast Disk Takeover or Concurrent Access:
   - Enable Fast Disk Takeover: This option is the normal shared enhanced concurrent volume group.

– Concurrent Access: Parallel disk access for Oracle RAC. This option is used only with the concurrent resource group.

10. If you use Concurrent Access, you can set this volume group as a critical one. For more information about critical volume groups, see “Critical volume groups (voting disks) for Oracle RAC” on page 192.
11. If you selected a resource group or created a resource group, you must verify and synchronize the cluster. Start **`smit sysmirror`**. Select **Cluster Applications and Resources** → **Verify and Synchronize Cluster Configuration**.

**Important:** If you make any changes in the resource group configuration, you must verify and synchronize the cluster.

## 4.10.3 Adding a disk to a shared volume group

We show how to add a disk to a shared volume group:

1. Start **`smit cspoc`**. Select **Storage** → **Volume Groups** → **Set Characteristics of a Volume Group** → **Add a Volume to a Volume Group**. See Figure 4-14.
2. Select the volume group to extend.
3. Choose the disks to add.
4. Optional: Specify the mirror pool name for the new disks.

```
 Add a Volume to a Volume Group

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

                                                        [Entry Fields]
  VOLUME GROUP name                                   app1vg
  Resource Group Name                                 inst1rg
  Node List                                           inst1,inst2
  VOLUME names                                        hdisk4 hdisk5 hdisk6
  Physical Volume IDs                                 00f74d45367edf3b 00f7>
  Mirror Pool for New Volumes                        []                      +

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

*Figure 4-14   Extending a volume group with new disks*

## 4.10.4 Importing a shared volume group

We show how to import a shared volume group:

1. Start **`smit cspoc`**. Select **Storage** → **Volume Groups** → **Import a Volume Group**. See Figure 4-15 on page 170.
2. Choose the volume group to import. It must be a shared volume group.
3. Select the referencing disk.

4. Supply the major number. The major number must be available on all nodes. For more information about the major number, see “Creating data volume groups and file systems” on page 108
5. Select **yes** to make this VG Concurrent Capable.

```
 Import a Volume Group

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

                                                        [Entry Fields]
  VOLUME GROUP name                                   app1vg
  Resource Group Name                                 inst1rg
  Node List                                           inst1,inst2
  Reference node                                      inst1
  PHYSICAL VOLUME name                                hdisk3
  Volume group MAJOR NUMBER                          [102]                  +#
  Make this VG Concurrent Capable?                    yes                   +
  Make default varyon of VG Concurrent?               no                    +

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

*Figure 4-15 Import a volume group*

## 4.10.5 Adding a logical volume

We show how to add a logical volume:

1. Start **`smit cspoc`**. Select **Storage** → **Logical Volumes** → **Add a Logical Volume**.
2. Choose the volume group.
3. You can specify the hard disks for the logical volume or leave it Auto-select.
4. Fill out the dialog with the usual logical volume parameters (Figure 4-16 on page 171).

```
 Add a Logical Volume

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

[TOP]                                                    [Entry Fields]
  Resource Group Name                                   inst1rg
  VOLUME GROUP name                                     app1vg
  Node List                                             inst1,inst2
  Reference node                                        inst1
* Number of LOGICAL PARTITIONS                         [10]                     #
  PHYSICAL VOLUME names                                 hdisk3
  Logical volume NAME                                  [datalv01]
  Logical volume TYPE                                  [jfs2]                  +
  POSITION on physical volume                           outer_middle           +
  RANGE of physical volumes                             minimum                +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                   []                       #
    to use for allocation
  Number of COPIES of each logical                      1                      +
    partition

[MORE...19]

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

*Figure 4-16 Creating a logical volume*

### 4.10.6 Increasing the size of a logical volume

We show how to increase the size of a logical volume:

1. Start **smit cspoc**. Select **Storage** → **Logical Volumes** → **Set Characteristics of a Logical Volume** → **Increase the Size of a Logical Volume**.
2. Choose the volume group.
3. Select the logical volume.
4. You can specify a disk where the extension is placed, or leave it Auto-select.
5. Enter the number of additional physical partitions, and the other usual LVM parameters. See Figure 4-17 on page 172.

```
 Increase the Size of a Logical Volume

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

                                                         [Entry Fields]
  Volume Group Name                                    app1vg
  Resource Group Name                                  inst1rg
  LOGICAL VOLUME name                                  data2lv

  Reference node
* Number of ADDITIONAL logical partitions             [10]                    #
  PHYSICAL VOLUME names
  POSITION on physical volume                          outer_middle          +
  RANGE of physical volumes                            minimum               +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                  [1024]                  #
    to use for allocation
  Allocate each logical partition copy                 yes                   +
    on a SEPARATE physical volume?

  File containing ALLOCATION MAP                      []                      /

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

*Figure 4-17 Increase the size of a logical volume*

### 4.10.7 Changing a logical volume

We show how to change a logical volume:

1. Start **`smit cspoc`** and select **Storage** → **Logical Volumes** → **Change a Logical Volume**.
2. Choose the volume group.
3. Select the logical volume.
4. Enter the usual LVM parameters. See Figure 4-18 on page 173.

```
 Change a Logical Volume on the Cluster

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

                                                        [Entry Fields]
  Volume Group Name                                   app1vg
  Resource Group Name                                 inst1rg
* Logical volume NAME                                 data2lv

  Logical volume TYPE                                [jfs2]                +
  POSITION on physical volume                         outer_middle         +
  RANGE of physical volumes                           minimum              +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                 [1024]                 #
    to use for allocation
  Allocate each logical partition copy                yes                  +
    on a SEPARATE physical volume?
  RELOCATE the logical volume during reorganization?  yes                  +
  Logical volume LABEL                               [None]
  MAXIMUM NUMBER of LOGICAL PARTITIONS               [512]

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

*Figure 4-18 Change a logical volume*

## 4.10.8 Adding a file system

We show how to add a file system:

1. Run **smit cspoc**, and select **Storage** → **File Systems** → **Add a File System**.
2. Select the volume group.
3. Select the file system type: **Enhanced Journaled File System**.
4. Enter the mount point, the size of the file system, and other usual parameters. See Figure 4-19 on page 174. The new file system that underlies logical volume name is lv*XX*, for example, lv01.

```
 Add an Enhanced Journaled File System

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

                                                        [Entry Fields]
  Resource Group                                      inst1rg
* Node Names                                          inst1,inst2
  Volume group name                                   app1vg

  SIZE of file system
          Unit Size                                   Gigabytes              +
*         Number of units                            [10]                     #
* MOUNT POINT                                        [/data2]                 /
  PERMISSIONS                                         read/write             +
  Mount OPTIONS                                      []                      +
  Block Size (bytes)                                  4096                   +
  Inline Log?                                         yes                    +
  Inline Log size (MBytes)                           []                       #
  Logical Volume for Log                                                     +
  Extended Attribute Format                           Version 1              +
  ENABLE Quota Management?                            no                     +
  Enable EFS?                                         no                     +

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

*Figure 4-19 Create a cluster file system*

## 4.10.9 Increasing the size of a file system

We show how to increase the size of a file system:

1. Run **smit cspoc**, and select **Storage** → **File Systems** → **Change / Show Characteristics of a File System**.
2. Select the file system.
3. Change the LVM values that you want to change. See Figure 4-20 on page 175.

```
 Change/Show Characteristics of an Enhanced Journaled File System

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

                                                        [Entry Fields]
   Volume group name                                  app1vg
  Resource Group Name                                 inst1rg
* Node Names                                          inst2,inst1

* File system name                                    /data2
  NEW mount point                                    [/data2]                 /
  SIZE of file system
          Unit Size                                   512bytes               +
          Number of units                            [412992]                 #
  Mount GROUP                                        []
  Mount AUTOMATICALLY at system restart?              no                     +
  PERMISSIONS                                         read/write             +
  Mount OPTIONS                                      []                      +
  Start Disk Accounting?                              no                     +
  Block Size (bytes)                                  4096
  Inline Log?                                         yes
  Inline Log size (MBytes)                           [1]                      #
  Extended Attribute Format                          [v1]
  ENABLE Quota Management?                            no                     +
  Allow Small Inode Extents?                         [yes]                   +
  Logical Volume for Log                              yes                    +
  Encrypted File System                               no

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

*Figure 4-20 Change a file system*

## 4.10.10 Mirroring a logical volume

We show how to mirror a logical volume:

1. Execute **`smit cspoc`** and select **Storage** → **Logical Volumes** → **Set Characteristics of a Logical Volume** → **Add a Copy to a Logical Volume**.
2. Choose the volume group.
3. Select the logical volume.
4. You can select disks for the mirror copy.
5. Enter the number of copies and other usual LVM parameters, including the mirror pool names. See Figure 4-21 on page 176.

```
 Add a Copy to a Logical Volume

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

                                                         [Entry Fields]
  Volume Group Name                                    app1vg
  Resource Group Name                                  inst1rg
* LOGICAL VOLUME name                                  data1lv

  Reference node
* NEW TOTAL number of logical partition                2                       +
    copies
  PHYSICAL VOLUME names
  POSITION on physical volume                          outer_middle            +
  RANGE of physical volumes                            minimum                 +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                  []                        #
    to use for allocation
  Allocate each logical partition copy                 yes                     +
    on a SEPARATE physical volume?
  SYNCHRONIZE the data in the new                      no                      +
    logical partition copies?

  File containing ALLOCATION MAP                      []                        /

  Mirror Pool for First Copy                          []                       +
  Mirror Pool for Second Copy                         []                       +
  Mirror Pool for Third Copy                          []                       +

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

*Figure 4-21 Mirror a logical volume*

### 4.10.11 Mirroring a volume group

We show how to mirror a volume group:

1. Start **`smit cspoc`** → **Storage** → **Volume Groups** → **Mirror a Volume Group**.
2. Choose the volume group.
3. Select the physical volumes. In most cases, Auto-select is fine.
4. You can modify the usual **`mirrorvg`** parameters such as the number of copies and the mirror pool settings. See Figure 4-22 on page 177.

```
 Mirror a Volume Group

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

                                                          [Entry Fields]
* VOLUME GROUP name                                     app1vg
  Resource Group Name                                   inst1rg
  Node List                                             inst1,inst2
  Reference node
  PHYSICAL VOLUME names

  Mirror sync mode                                      Foreground             +
  Number of COPIES of each logical                      2                      +
    partition
  Keep Quorum Checking On?                              no                     +
  Create Exact LV Mapping?                              no                     +

  Mirror Pool for First Copy                           []                      +
  Mirror Pool for Second Copy                          []                      +
  Mirror Pool for Third Copy                           []                      +

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

*Figure 4-22 Mirror a volume group*

### 4.10.12 Synchronizing the LVM mirror

We show how to synchronize the LVM mirror (Figure 4-23 on page 178):

1. Run **smit cspoc**, and select **Storage** → **Volume Groups** → **Synchronize LVM Mirrors** → **Synchronize by Volume Group**.
2. Choose the volume group.
3. You can change the following parameters:
   – Number of Partitions to Sync in Parallel: You can specify a number between 1 - 32. Do not select a number higher than the number of disks in the mirror.
   – Synchronize All Partitions: Select yes for mirroring all partitions regardless of their current synchronization status.
   – Delay Writes to VG from other cluster nodes during this Sync: If the volume group belongs to a concurrent resource group, you can defer the writes on the other nodes until the sync is finished.

```
 Synchronize LVM Mirrors by Volume Group

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

                                                        [Entry Fields]
  VOLUME GROUP name                                   app1vg
  Resource Group Name                                 inst1rg
* Node List                                           inst1,inst2

  Number of Partitions to Sync in Parallel           [1]                    +#
  Synchronize All Partitions                          no                    +
  Delay Writes to VG from other cluster nodes during  no                    +
   this Sync

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

*Figure 4-23 Synchronize LVM mirrors*

### 4.10.13 Unmirroring a logical volume

We show how to unmirror a logical volume:

1. Execute **smit cspoc** and select **Storage** → **Logical Volumes** → **Set Characteristics of a Logical Volume** → **Remove a Copy from a Logical Volume**.
2. Choose the volume group.
3. Select the logical volume.
4. Select the disk that contains the mirror to remove (Figure 4-24).
5. Enter the new number of logical volume copies.

```
 Remove a Copy from a Logical Volume

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

                                                        [Entry Fields]
  Volume Group Name                                   app1vg
  Resource Group Name                                 inst1rg
* LOGICAL VOLUME name                                 data1lv
  Reference node

* NEW maximum number of logical partition             1                     +
    copies
  PHYSICAL VOLUME name(s) to remove copies from

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

*Figure 4-24 Remove a copy from a logical volume*

## 4.10.14 Unmirroring a volume group

We show how to unmirror a volume group:

1. Start **smit cspoc**, and select **Storage** → **Volume Groups** → **Unmirror a Volume Group**.
2. Choose the volume group.
3. Select the disk that contains the mirror to remove.
4. Set Number of COPIES of each logical partition to the value that you want, which is usually 1. See Figure 4-25.

```
 Unmirror a Volume Group

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

                                                        [Entry Fields]
  VOLUME GROUP name                                   app1vg
  Resource Group Name                                 inst1rg
  Node List                                           inst1,inst2
  Reference node                                      inst1
  PHYSICAL VOLUME names                               hdisk3

  Number of COPIES of each logical                    1                       +
    partition

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

*Figure 4-25 Unmirror a volume group*

## 4.10.15 Removing a file system

We show how to remove a file system:

**Important:** Before you start the following SMIT operation, you must manually unmount the file system.

1. Start **smit cspoc** and select **Storage** → **File Systems** → **Remove a File System**.
2. Select the file system to remove. See Figure 4-26 on page 180.

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

                                                        [Entry Fields]
   Volume group name                                  app1vg
  Resource Group Name                                 inst1rg
* File system name                                    /lv02
* Node Names                                          inst2,inst1

  Remove Mount Point                                  yes                    +

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

*Figure 4-26 Remove a file system*

## 4.10.16 Removing a logical volume

We show how to remove a logical volume:

1. Start **smit cspoc** and select **Storage** → **Logical Volumes** → **Remove a Logical Volume**.
2. Select the volume group.
3. Select the logical volume to remove. See Figure 4-27.

```
 Remove a Logical Volume

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

                                                        [Entry Fields]
  Volume Group Name                                   app1vg
  Resource Group Name                                 inst1rg
* LOGICAL VOLUME name                                 data2lv


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

*Figure 4-27 Remove a logical volume*

## 4.10.17 Removing a physical disk from a shared volume group

We remove a physical disk from a shared volume group:

1. Start **smit cspoc** and select **Storage** → **Volume Groups** → **Set Characteristics of a Volume Group** → **Remove a Volume from a Volume Group**.
2. Select the volume group.
3. Select the physical disk to remove.

## 4.10.18 Replacing a cluster disk

PowerHA SystemMirror C-SPOC provides an easy to use method to exchange a broken disk for a new disk in a shared volume group. In a disk failure, execute the following steps:

1. Configure a new disk or LUN to the cluster nodes. See "Adding a disk to the cluster nodes" on page 167.
   - The disk must be the same size or larger than the source disk regardless of the used space in the source.
   - It cannot be a volume group member.
2. Start **`smit cspoc`** and select **Storage** → **Physical Volumes** → **Cluster Disk Replacement**. See Figure 4-28.
3. Select the hard disk to replace.
4. Select the destination disk.

The Cluster Disk Replacement performs the following steps:

- Adds the destination disk to the volume group
- Moves all data from the source disk to the new disk by using LVM mirroring
- Removes the failed disk from the resource group

```
 Cluster Disk Replacement

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

                                                         [Entry Fields]
  SOURCE DISK
* Volume Group                                         applvg
* hdisk                                                hdisk5
* PVID                                                 00f74d47415d3ac4
* Nodename                                             inst1
  Resource Group                                       inst1rg
* Node List                                            inst1,inst2

  DESTINATION DISK
* hdisk                                                hdisk6
* PVID                                                 00f74d474f69ea30

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

*Figure 4-28 Cluster disk replacement*

## 4.10.19 Removing a volume group

We show how to remove a volume group:

Start **`smit cspoc`** and select **Storage** → **Volume Groups** → **Remove a Volume Group**. Select the volume group to remove. You cannot remove a volume group that is used in an active resource group.

### 4.10.20 Showing the disk UUID

The Universal Unique ID (UUID) is a low-level disk identifier that is used in the Cluster Aware AIX (CAA) Storage Framework. UUID is based on an Open Software Foundation Distributed Computing Environment standard. This ID can be useful for analyzing low-level traces. You can display the disk UUID by running **`smit cspoc`** and selecting **Storage** → **Physical Volumes** → **Show UUID for a Physical Volume**. See Figure 4-29 for more details.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

 Disk name:                    hdisk4
 Disk UUID:                    5d39d2dd4f51d721 e4ffa0fe5c804aba
 Fence Group UUID:             0000000000000000 0000000000000000 - Not in a
Fen
ce Group
 Disk device major/minor number:       18, 2
 Fence height:                 0 (Read/Write)
 Reserve mode:                 0 (No Reserve)
 Disk Type:                    0x01 (Local access only)
 Disk State:                   32791

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 4-29 Displaying a disk UUID*

### 4.10.21 Renaming a physical volume

Managing a high number of disks can be a daunting task. Before PowerHA SystemMirror 7.1, the only way to match the hdisk numbers between the nodes is by checking the PVIDs. If you have more than 10 hdisks, this task is challenging.

The new physical volume rename function provides a convenient method to manage cluster disks. You can rename any of the shared disks from hdisk*X* to any meaningful name that you want to use. Therefore, you can have a common hard disk naming scheme in the cluster. Follow these steps to rename a physical volume:

1. Start **`smit cspoc`** and select **Storage** → **Physical Volumes** → **Rename a Physical Volume**. See Figure 4-30 on page 183.
2. Select the hdisk to rename from the pickup list. You can rename only shared disks that are not in a volume group.
3. Enter the new disk name.
4. For Change all Physical Volumes with this PVID, select **yes**. With this option, C-SPOC changes the name of the hdisk on all nodes regardless of the current disk name.

```
 Rename a Physical Volume

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

                                                        [Entry Fields]
  Physical Volume Name                                  hdisk5
  Physical Volume Identifier                            00f74d47415d3ac4
  Physical Volume is Known on these Nodes               inst1,inst2

* New Physical Volume Name                             [dorka1]
  Change all Physical Volumes with this PVID?           yes +

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

*Figure 4-30 Rename a physical volume*

Example 4-10 shows the result of the physical volume rename.

*Example 4-10 Checking the new name of an hdisk*

```
root@inst2 /usr/es/sbin/cluster/utilities # lspv
hdisk0          00f74d47f96263a2                    rootvg          active
hdisk1          00f74d47fa934fef                    rootvg          active
hdisk2          00f74d45367ec638                    caavg_private   active
hdisk3          00f74d45367ed24c                    app1vg
hdisk4          00f74d45367edf3b                    None
dorka1          00f74d47415d3ac4                    None
hdisk6          00f74d474f69ea30                    None

root@inst2 /usr/es/sbin/cluster/utilities # lsdev -Ccdisk
dorka1 Available 81-T1-01 MPIO DS4800 Disk
hdisk0 Available          Virtual SCSI Disk Drive
hdisk1 Available          Virtual SCSI Disk Drive
hdisk2 Available 82-T1-01 MPIO DS4800 Disk
hdisk3 Available 82-T1-01 MPIO DS4800 Disk
hdisk4 Available 81-T1-01 MPIO DS4800 Disk
hdisk6 Available 81-T1-01 MPIO DS4800 Disk
```

# 4.11 Cross-site mirroring and AIX LVM Mirror pools

The AIX LVM Mirror Pools feature helps to maintain the physical layout of the LVM mirrors. It is an AIX feature since Version 6.1. We show how to use LVM Mirror Pools in the PowerHA environment.

The Mirror pools feature supports two types of operation modes: synchronous and asynchronous. We focus on synchronous mirror pools because async and remote mirror pools require PowerHA SystemMirror Extended Edition.

## 4.11.1 Introducing PowerHA with Cross-site mirroring

With LVM mirror pools, you can create a multiple site disaster-proof cluster by using PowerHA SystemMirror Standard Edition. The basic configuration is the same as a local cluster. The only difference is that you have two storage subsystems (one on each location) and you have to set up AIX mirroring for all shared volume groups. In a site failure, the surviving nodes still have access to the local copy of the shared data; therefore, they can perform a successful takeover.

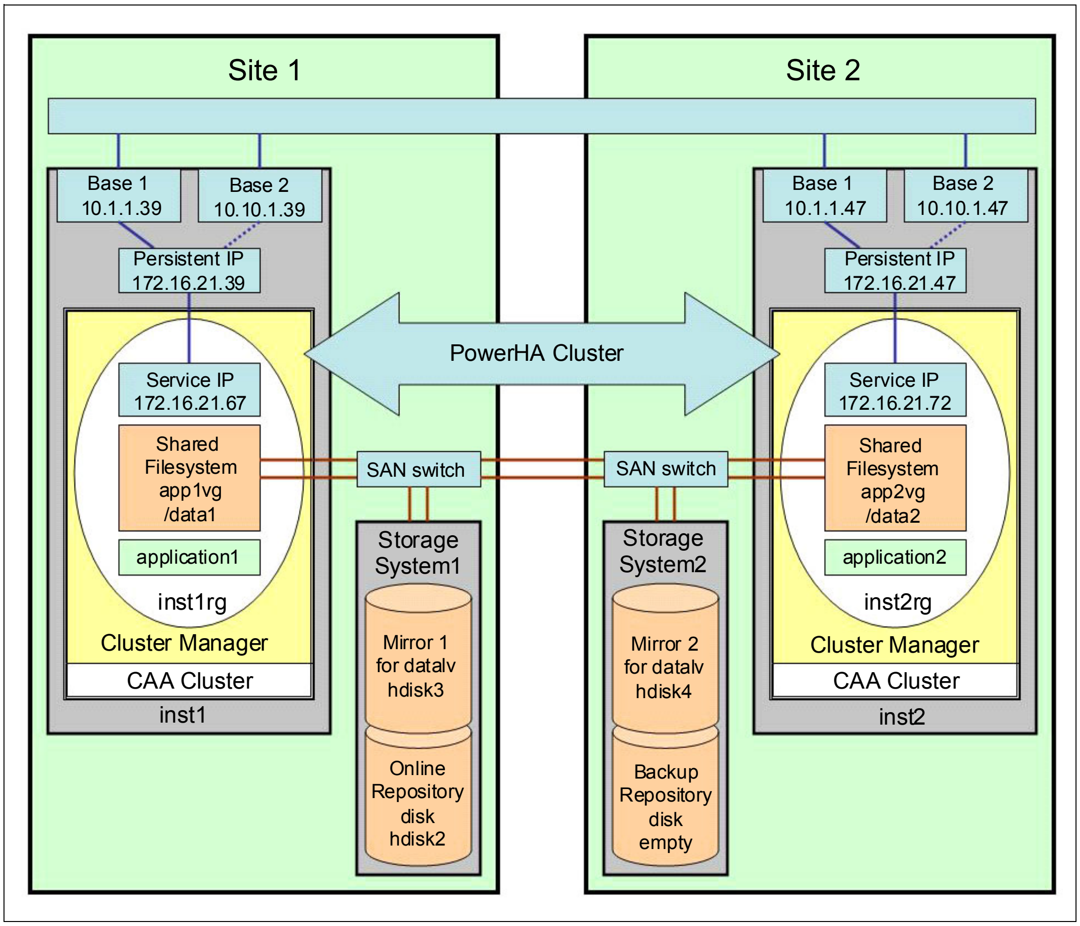



*Figure 4-31 Cross-site mirroring solution with PowerHA SystemMirror*

However, there are some important limitations:

- Requires a storage subsystem on each site.
- Requires a reliable, fast SAN and network connection between the sites.
- The servers must connect to the same logical network, and they must have the same subnet and default gateway.
- The maximum distance between the computers is limited. If the distance is longer than 30 - 50 kilometers, the storage-based mirroring performs better.
- You cannot use storage-based mirroring such as the IBM DS8000 Geomirror feature.
- Geographic Logical Volume Manager (GLVM) and asynchronous remote mirror pools are not supported in PowerHA SystemMirror Standard Edition.

> **Important:** You cannot mirror the repository disk, and in a site or storage subsystem failure, you have to rely on the repository disk resilience.

IBM PowerHA SystemMirror Enterprise Edition provides a long-distance disaster recovery solution without the preceding limitation.

When you set up a cross-site mirroring configuration, ensure that the following characteristics are true:

- All shared volume groups are fully mirrored.
- The mirror copies are on the corresponding storage subsystem.
- Quorum checking is turned off for the shared volume groups (**`chvg -Q n vgname`**).
- After a full site down event, you have to check the volume group reintegration and you might need to manually synchronize the volume group mirrors.

## 4.11.2 The multiple storage mirroring problem

By default, AIX does not track the physical locations of the hdisk devices during LVM mirroring, logical volume, file system creation, and an extension operation. LVM ensures that each copy of a logical volume is on a separate hdisk but it does not matter where a certain hdisk is located. Figure 4-32 shows an example without using AIX Mirror Pools. LVM mirrors the logical volumes without taking care of a disk physical location.

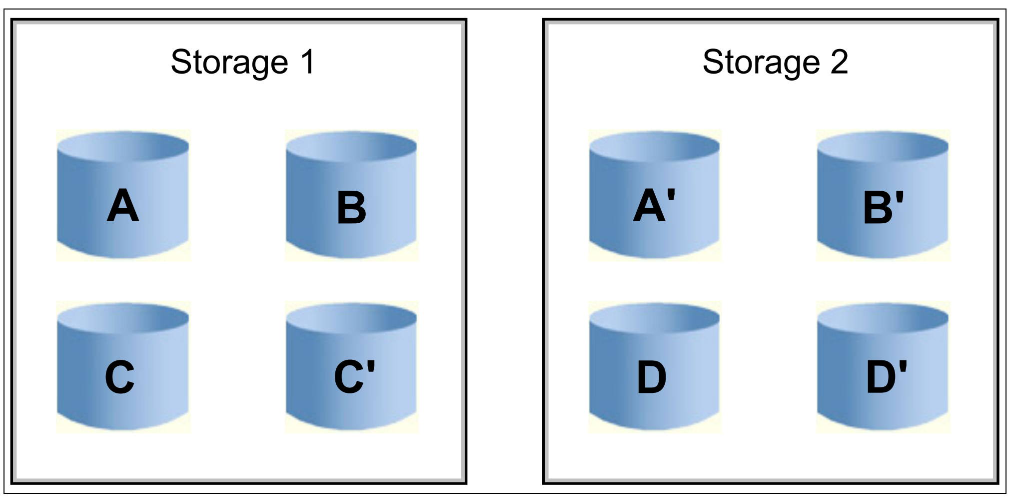



*Figure 4-32 LVM mirroring without AIX Mirror Pools*

The mirror pools are user-defined tags for physical volumes. If you specify them for all disks in a volume group, AIX lays out the mirror copies according to this label. In Example 4-11, hdisk3 and hdisk5 are in the same storage in our computer data room. We label them as *mirror1*. Hdisk4 and hdisk5 are in another building; we use the *mirror2* label to differentiate them. When we create our shared, mirrored logical volumes (for example, `bubbalv`), we specify LVM to use the *mirror1* and *mirror2* mirror pools.

*Example 4-11 Using mirror pools*

```
root@bluemoon # lspv -P
Physical Volume   Volume Group      Mirror Pool
hdisk0            rootvg
hdisk1            rootvg
```

```
hdisk2          caavg_private
hdisk3          applvg          mirror1
hdisk4          applvg          mirror2
hdisk5          applvg          mirror1
hdisk6          applvg          mirror2

root@bluemoon / # lslv bubbalv
LOGICAL VOLUME:     bubbalv                VOLUME GROUP:   applvg
LV IDENTIFIER:      00f74d4500004c00000001363697f4c7.6 PERMISSION:     read/write
VG STATE:           active/complete        LV STATE:       closed/syncd
TYPE:               jfs2                   WRITE VERIFY:   off
MAX LPs:            512                    PP SIZE:        8 megabyte(s)
COPIES:             2                      SCHED POLICY:   parallel
LPs:                5                      PPs:            10
STALE PPs:          0                      BB POLICY:      relocatable
INTER-POLICY:       minimum                RELOCATABLE:    yes
INTRA-POLICY:       middle                 UPPER BOUND:    1024
MOUNT POINT:        N/A                    LABEL:          None
DEVICE UID:         0                      DEVICE GID:     0
DEVICE PERMISSIONS: 432
MIRROR WRITE CONSISTENCY: on/ACTIVE
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?:     NO
INFINITE RETRY:     no
DEVICESUBTYPE:      DS_LVZ
COPY 1 MIRROR POOL: mirror1
COPY 2 MIRROR POOL: mirror2
COPY 3 MIRROR POOL: None
```

### 4.11.3 Assigning the mirror pool names

You can assign the mirror pool names in several ways, for example, when you create or extend a volume group. All standard AIX LVM and PowerHA C-SPOC LVM commands support the mirror pool names:

1. Start **smit cspoc** and select **Storage** → **Volume Groups** → **Manage Mirror Pools for Volume Groups** → **Add Disks to a Mirror Pool**.
2. Select the mirror pool. If it is a new mirror pool, choose **<New>**.
3. In the smit panel (Figure 4-33 on page 187), enter the name of the new mirror pool.
4. Select the disks to add. Press F4 for the list of available, untagged disks.

```
 Add Disks to a Mirror Pool

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

                                                        [Entry Fields]
* Mirror Pool                                        [mirror1]              +
  VOLUME GROUP name                                   app1vg
  Resource Group Name                                 inst1rg
  Node Names                                          inst1,inst2

* Disks to Add to the Mirror Pool                    []                     +

  +--------------------------------------------------------------------------+
  |                      Disks to Add to the Mirror Pool                     |
  |                                                                          |
  | Move cursor to desired item and press F7.                                |
  |     ONE OR MORE items can be selected.                                   |
  | Press Enter AFTER making all selections.                                 |
  |                                                                          |
  | > 00f74d45367ed24c ( hdisk3 on all cluster nodes )                       |
  |   00f74d45367edf3b ( hdisk4 on all cluster nodes )                       |
  | > 00f74d47415d3ac4 ( hdisk5 on all cluster nodes )                       |
  |                                                                          |
  | F1=Help                F2=Refresh              F3=Cancel                 |
F1| F7=Select              F8=Image                F10=Exit                  |
F5| Enter=Do               /=Find                  n=Find Next               |
F9+--------------------------------------------------------------------------+
```

*Figure 4-33 Add disks to a mirror pool*

## 4.11.4 Extending a volume group

When you extend a volume group, you can specify the mirror pool names for the new disks. Be careful to extend the volume group with disks from only one mirror pool at a time:

1. Start `smit cspoc`, and select **Storage** → **Volume Groups** → **Set Characteristics of a Volume Group** → **Add a Volume to a Volume Group**.
2. Select the volume group to extend.
3. Select the physical volumes to add. Select disks from only one storage pool at the same time.
4. Press F4 and select the existing mirror pool name for the new disks. See Figure 4-34 on page 188.

```
 Add a Volume to a Volume Group

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

                                                         [Entry Fields]
  VOLUME GROUP name                                    app1vg
  Resource Group Name                                  inst1rg
  Node List                                            inst1,inst2
  VOLUME names                                         hdisk6
  Physical Volume IDs                                  00f74d474f69ea30
  Mirror Pool for New Volumes                         []                      +

  +--------------------------------------------------------------------------+
  |                      Mirror Pool for New Volumes                         |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   mirror1                                                                |
  |   mirror2                                                                |
  |                                                                          |
  | F1=Help                 F2=Refresh              F3=Cancel                |
F1| F8=Image                F10=Exit                Enter=Do                 |
F5| /=Find                  n=Find Next                                      |
F9+--------------------------------------------------------------------------+
```

*Figure 4-34 Extending a volume group*

## 4.11.5 Checking the mirror pool assignments

You can check the mirror pool assignments with the **lspv -P**, **lsvg -P vgname**, and **lslv lvname** commands. See Example 4-12.

*Example 4-12 Checking the mirror pools*

```
root@bluemoon # lspv -P
Physical Volume   Volume Group      Mirror Pool
hdisk0            rootvg
hdisk1            rootvg
hdisk2            caavg_private
hdisk3            app1vg            mirror1
hdisk4            app1vg            mirror2
hdisk5            app1vg            mirror1
hdisk6            app1vg            mirror2

root@bluemoon / # lsvg -P app1vg
Physical Volume   Mirror Pool
hdisk3            mirror1
hdisk4            mirror2
hdisk5            mirror1
hdisk6            mirror2

root@bluemoon / # lslv  bubbalv
LOGICAL VOLUME:     bubbalv                VOLUME GROUP:   app1vg
```

```
LV IDENTIFIER:      00f74d4500004c00000001363697f4c7.6 PERMISSION:     read/write
VG STATE:           active/complete       LV STATE:       closed/stale
TYPE:               jfs2                  WRITE VERIFY:   off
MAX LPs:            512                   PP SIZE:        8 megabyte(s)
COPIES:             2                     SCHED POLICY:   parallel
LPs:                5                     PPs:            10
STALE PPs:          5                     BB POLICY:      relocatable
INTER-POLICY:       minimum               RELOCATABLE:    yes
INTRA-POLICY:       middle                UPPER BOUND:    1024
MOUNT POINT:        N/A                   LABEL:          None
DEVICE UID:         0                     DEVICE GID:     0
DEVICE PERMISSIONS: 432
MIRROR WRITE CONSISTENCY: on/ACTIVE
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?:     NO
INFINITE RETRY:     no
DEVICESUBTYPE:      DS_LVZ
COPY 1 MIRROR POOL: mirror1
COPY 2 MIRROR POOL: mirror2
COPY 3 MIRROR POOL: None
```

### 4.11.6 Creating a mirrored logical volume with mirror pools

After your disks are assigned to mirror pools, you can specify the mirror pools instead of the physical volume names in all LVM mirroring-related commands:

1. Start **smit cspoc** and select **Storage** → **Logical Volumes** → **Add a Logical Volume**.
2. Choose the volume group to contain the new logical volume.
3. Specify the hard disks for the logical volume or leave it Auto-select.
4. Fill out the dialog with the usual logical volume parameters, including the number of mirror copies.
5. In the bottom of the SMIT panel, select the mirror pool names for the first, second, and third copies. See Figure 4-35 on page 190.

```
 Add a Logical Volume

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

[TOP]                                                    [Entry Fields]
  Resource Group Name                                   inst1rg
  VOLUME GROUP name                                     app1vg
  Node List                                             inst1,inst2
  Reference node
* Number of LOGICAL PARTITIONS                         [10]                    #
  PHYSICAL VOLUME names
  Logical volume NAME                                  [m1lv]
  Logical volume TYPE                                  [jfs2]                 +
  POSITION on physical volume                           outer_middle          +
  RANGE of physical volumes                             minimum               +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                   []                      #
    to use for allocation
  Number of COPIES of each logical                      2                     +
    partition
  Mirror Write Consistency?                             active                +
  Allocate each logical partition copy                  yes                   +
    on a SEPARATE physical volume?
  RELOCATE the logical volume during reorganization?  yes                   +
  Logical volume LABEL                                 []
  MAXIMUM NUMBER of LOGICAL PARTITIONS                 [512]                   #
  Enable BAD BLOCK relocation?                          yes                   +
  SCHEDULING POLICY for reading/writing                 parallel              +
    logical partition copies
  Enable WRITE VERIFY?                                  no                    +
  File containing ALLOCATION MAP                       []                      /
  Stripe Size?                                         [Not Striped]          +
  Serialize I/O?                                        no                    +
  Make first block available for applications?          no                    +
  Mirror Pool for First Copy                            mirror1               +
  Mirror Pool for Second Copy                           mirror2               +
  Mirror Pool for Third Copy                                                  +
[MORE...2]

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

*Figure 4-35 Creating a logical volume with mirror pools*

## 4.11.7 Mirroring a volume group

We show how to mirror a volume group:

1. Start `smit cspoc` and select **Storage** → **Volume Groups** → **Mirror a Volume Group**.
2. Choose the volume group to mirror.
3. Select the physical volumes. When you use mirror pools, Auto-select is fine.
4. Set the number of copies and the other LVM parameters to the values that you want.

5. Select the mirror pools for the first, second, and third mirror copies. See Figure 4-36 on page 191.

```
 Mirror a Volume Group

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

                                                        [Entry Fields]
* VOLUME GROUP name                                   applvg
  Resource Group Name                                 inst1rg
  Node List                                           inst1,inst2
  Reference node
  PHYSICAL VOLUME names

  Mirror sync mode                                    Foreground             +
  Number of COPIES of each logical                    2                      +
    partition
  Keep Quorum Checking On?                            no                     +
  Create Exact LV Mapping?                            no                     +

  Mirror Pool for First Copy                         [mirror1]               +
  Mirror Pool for Second Copy                        [mirror2]               +
  Mirror Pool for Third Copy                         []                      +

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

*Figure 4-36 Mirror a volume group*

### 4.11.8 Mirroring a logical volume by using mirror pools

We show how to mirror a logical volume by using mirror pools:

1. Execute **smit cspoc** and select **Storage** → **Logical Volumes** → **Set Characteristics of a Logical Volume** → **Add a Copy to a Logical Volume**.
2. Choose the volume group.
3. Select the logical volume to mirror.
4. You can select disks for the mirror copy or leave it on Auto-select.
5. Enter the number of copies and other usual LVM parameters.
6. Select the mirror pool names for each mirror copy. See Figure 4-37 on page 192.

```
 Add a Copy to a Logical Volume

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

                                                         [Entry Fields]
  Volume Group Name                                    app1vg
  Resource Group Name                                  inst1rg
* LOGICAL VOLUME name                                  bubbalv

  Reference node
* NEW TOTAL number of logical partition                2                       +
    copies
  PHYSICAL VOLUME names
  POSITION on physical volume                          outer_middle            +
  RANGE of physical volumes                            minimum                 +
  MAXIMUM NUMBER of PHYSICAL VOLUMES                  []                        #
    to use for allocation
  Allocate each logical partition copy                 yes                     +
    on a SEPARATE physical volume?
  SYNCHRONIZE the data in the new                      no                      +
    logical partition copies?

  File containing ALLOCATION MAP                      []                        /

  Mirror Pool for First Copy                          [mirror1]                +
  Mirror Pool for Second Copy                         [mirror2]                +
  Mirror Pool for Third Copy                          []                       +

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

*Figure 4-37 Mirror a logical volume by using mirror pools*

## 4.12 Critical volume groups (voting disks) for Oracle RAC

The PowerHA SystemMirror critical volume group function provides multiple node disk heartbeat functionality for the Oracle Real Application Cluster (RAC) voting disk. This feature is new in PowerHA 7.1; it is a replacement for the old Multi-Node Disk Heart Beat technology.

Critical volume groups safeguard the Oracle RAC voting disks. PowerHA continuously monitors the read-write accessibility of the voting disks. You can set up one of the following recovery actions if you lose access to a volume group:

- Notify only.
- Halt the node.
- Fence the node so that the node remains up but it cannot access the Oracle database.
- Shut down cluster services and bring all resource groups offline.

**Important:** The critical volume groups and the Multi-Node Disk Heart Beat do not replace the SAN-based disk heartbeat. These different technologies are used for separate purposes.

Do not use critical volume groups instead of the SAN-based heartbeat.

If you have Oracle RAC, have at least one designated volume group for voting. Follow these steps to configure a critical volume group:

1. You have to set up a concurrent resource group for two or more nodes:
   - Startup policy: Online on all available nodes
   - Fallover policy: Bring offline
   - Fallback policy: Never fallback
2. Create an enhanced concurrent volume group, which is accessible for all nodes in the resource group. This volume group stores the Oracle RAC voting files.
3. Add the volume group to the concurrent resource group.
4. Synchronize your cluster.
5. Start `smit cspoc` and select **Storage** → **Volume Groups** → **Manage Critical Volume Groups** → **Mark a Volume Group as Critical**.
6. Select the volume group from the pickup list.
7. Configure the failure action: Start `smit cspoc` and select **Storage** → **Volume Groups** → **Manage Critical Volume Groups** → **Configure failure actions for Critical Volume Groups**.
8. Select the volume group.
9. Select the recovery action on the loss of disk access:
   - Notify Only
   - Halt the node
   - Fence the node
   - Shutdown Cluster Services and bring all Resource Groups Offline
10. Synchronize the cluster.

**Important:** If you change the critical volume groups, verify and synchronize the cluster.

## 4.13 File collections

PowerHA SystemMirror provides cluster-wide file synchronization capabilities through C-SPOC file collection functions. A *file collection* is a user-defined set of files and synchronization rules.

PowerHA provides three ways to propagate your files:

- Manually: You can synchronize your files manually at any time. The files are copied from the local node to the remote one.
- Automatically during cluster verification and synchronization: The files are propagated from the node where you start the PowerHA verification.

- Automatically when changes are detected: PowerHA checks the file collection on all nodes periodically. If a file changed, PowerHA synchronizes this file across the cluster. You can set up a timer for how frequently PowerHA checks the file collections.

PowerHA retains the permissions, ownership, and time stamp of the file and propagates them to the remote nodes. You can specify filenames, wildcard filenames, and directories. You cannot add the following information:

- Symbolic links
- Wildcard directory names
- Pipes
- Sockets
- Device files (`/dev/*`)
- Files from the `/proc` directory
- ODM files from `/etc/objrepos/*` and `/etc/es/objrepos/*`

Always use the full path names. Each file can be added to one file collection only, except those files that automatically are added to the HACMP_Files collection. The files must not exist on the remote nodes; PowerHA creates them during the first synchronization.

PowerHA creates a backup copy of the modified files during synchronization on all nodes. These backups are stored in the `/var/hacmp/filebackup` directory. Only one previous version is retained and you can only restore them manually.

The file collection logs are stored in the `/var/hacmp/log/clutils.log` file.

> **Important:** It is your responsibility to ensure that files on the local node (where you start the propagation) are the most recent and are not corrupted.

## 4.13.1 Predefined file collections

PowerHA provides two file collections by default: Configuration_Files and HACMP_Files. None of them is set up for automatic synchronization by default. You can enable them by setting either the “Propagate files during cluster synchronization” or “Propagate files automatically when changes are detected” option to **Yes** in the SMIT Change/Show a file collection menu. See “Changing a file collection” on page 197.

### Configuration_Files

This collection contains the essential AIX configuration files:

- `/etc/hosts`
- `/etc/services`
- `/etc/snmpd.conf`
- `/etc/snmpdv3.conf`
- `/etc/rc.net`
- `/etc/inetd.conf`
- `/usr/es/sbin/cluster/netmon.cf`
- `/usr/es/sbin/cluster/etc/clhosts`
- `/etc/cluster/rhosts`
- `/usr/es/sbin/cluster/etc/clinfo.rc`

You can easily add files to or remove files from this collection. For more information, see “Adding files to a file collection” on page 198.

### HACMP_Files

This file collection automatically collects all user-defined scripts from the PowerHA configuration. If you define any of the following files in your cluster configuration, these files are automatically included in the HACMP_Files file collection:

- Application server start script
- Application server stop script
- Event notify script
- Pre-event script
- Post-event script
- Event error recovery script
- Application monitor notify script
- Application monitor cleanup script
- Application monitor restart script
- Pager text message file
- SNA Link start and stop scripts
- X.25 Link start and stop scripts
- HA tape support start script
- HA tape support stop script
- User-defined event recovery program
- Custom snapshot method script

We show an example of a file collection. Our cluster has an application server, DB2. Its start script is `/usr/ha/db2.start` and its stop script is `/usr/ha/db2.stop`. Also, we have a custom post-event script for the node_up event that is called `/usr/ha/post.node_up`. These three files are automatically added to the HACMP_Files file collection when we defined them during HACMP configuration. You can check that the files are added:

1. Start SMIT PowerHA file collection management by entering **`smit cspoc`** and selecting **File Collections** → **Change/Show a File Collection**.
2. Select **HACMP_Files** from the pop-up list and press Enter.
3. Scroll down to the Collection files field and press F4. As you can see in Figure 4-38 on page 196, the application start and stop scripts and the post-event command are automatically added to this file collection.

```
 Change/Show a File Collection

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

                                                         [Entry Fields]
  File Collection Name                                 HACMP_Files
  New File Collection Name                            []
  File Collection Description                         [User-defined scripts >
  +-----------------------------------------------------------------------+
  ¦                              Collection files                          ¦
  ¦                                                                        ¦
  ¦ The value for this entry field must be in the                          ¦
  ¦ range shown below.                                                     ¦
  ¦ Press Enter or Cancel to return to the entry field,                    ¦
  ¦ and enter the desired value.                                           ¦
  ¦                                                                        ¦
  ¦   /usr/ha/db2.start                                                    ¦
  ¦   /usr/ha/db2.stop                                                     ¦
  ¦   /usr/ha/post.node_up                                                 ¦
  ¦                                                                        ¦
  ¦ F1=Help                F2=Refresh              F3=Cancel               ¦
F1¦ F8=Image               F10=Exit                Enter=Do                ¦
F5¦ /=Find                 n=Find Next                                     ¦
F9+-----------------------------------------------------------------------+
```

*Figure 4-38 HACMP_Files file collection*

**Important:** You cannot add or remove files directly to this file collection. If you start to use the HACMP_Files collection, ensure that your scripts can work correctly on all nodes.

If you do not want to synchronize all user-defined scripts or they are not the same on all nodes, disable this file collection and create another one that includes only the required files.

## 4.13.2 Managing file collections

We describe how you can create, modify, and remove a file collection.

### Adding a file collection

We show how to add a file collection:

1. Start SMIT by entering **`smit cspoc`**. Select **File Collection**.

   Or, you can start PowerHA File Collection Management by entering **`smit cm_filecollection_menu`**.

2. Select **Manage File Collections** → **Add a File Collection**.
3. Supply the requested information (see Figure 4-39 on page 197):
   – File Collection Name: Unique name for file collection.
   – File Collection Description: A short description of this file collection.
   – Propagate files during cluster synchronization: If you set this value to yes, PowerHA propagates this file collection during cluster synchronization. This option is a

convenient solution for cluster-related files. For example, your application startup scripts automatically synchronize after you change the cluster configuration.

– Propagate files automatically when changes are detected: If you select yes, PowerHA checks regularly the files in this collection. If any of them changed, it repropagates them.

If both of the options are left to **No,** no automatic synchronization takes place.

```
 Add a File Collection

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

                                                        [Entry Fields]
* File Collection Name                               [app_files]
  File Collection Description                        [Application config fi>
  Propagate files during cluster synchronization?     yes +
  Propagate files automatically when changes are det  no +
  ected?

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

*Figure 4-39 Add a file collection*

## Changing a file collection

We show how to change a file collection:

1. Start PowerHA File Collection Management by entering the **`smit cspoc`** fast path. Select **File Collections** → **Manage File Collections** → **Change/Show a File Collections**.
2. Select a file collection from the pop-up list.
3. Now, you can change the following information (see SMIT window on Figure 4-40 on page 198):
   – File collection name.
   – Description.
   – Propagate files during cluster synchronization (yes/no).
   – Propagate files automatically when changes are detected (yes/no).
   – Collection files: Press F4 here to see the list of files in this collection.

   See “Adding a file collection” on page 196 for an explanation of the fields.

```
 Change/Show a File Collection

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

                                                          [Entry Fields]
  File Collection Name                                  Configuration_Files
  New File Collection Name                             []
  File Collection Description                          [AIX and HACMP configu>
  Propagate files during cluster synchronization?       no +
  Propagate files automatically when changes are det  no +
  ected?
  Collection files +

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

*Figure 4-40 Change a file collection*

## Removing a file collection

We show how to remove a file collection:

1. Enter **smit cspoc** and select **File Collections** → **Manage File Collections** → **Remove a File Collection**.
2. Select a file collection from the pop-up list.
3. Press Enter again to confirm the deletion of the file collection.

## Changing the automatic update timer

You can set the timer for how frequently PowerHA checks the files in the collections for changes. Only one timer can be set for all file collection:

1. Enter **smit cspoc** and select **File Collections** → **Manage File Collections** → **Change/Show Automatic Update Time**.
2. Select a file collection from the pop-up list.
3. Supply the Automatic File Update Time in minutes. The value must be between 10 minutes - 1440 minutes (one day).

## Adding files to a file collection

We show how to add files to a file collection:

1. Start PowerHA File Collection Management by entering **smit cm_filecollection_menu** fast path.
2. Select **Manage File in File Collections** → **Add Files to a File Collection**.
3. Select a file collection from the pop-up list end press Enter.
4. On the SMIT panel, you can check the current file list or you can add new files (See Figure 4-41 on page 199):
   - To get the list of current files in this collection, scroll down to the Collection Files field and press F4.
   - To add new files, go to the New files field and type the file name here that you want to add to the file collection. You can add only one file at a time. The file name must start

with “/”. You can specify ordinary files only here, but you cannot add symbolic links, a directory, a pipe, a socket, a device file (`/dev/*`), files from `/proc` directory, and ODM files from `/etc/objrepos/*` and `/etc/es/objrepos/*`.

```
 Add Files to a File Collection

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

                                                        [Entry Fields]
  File Collection Name                                 app_files
  File Collection Description                          Application Config Fi>
  Propagate files during cluster synchronization?     no
  Propagate files automatically when changes are det  no
  ected?
  Collection files +
* New files                                           [/db2/app.config] /

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

*Figure 4-41 Adding files to a file collection*

**Important:** You cannot add files here to the HACMP_Files collection.

### Removing files from a file collection

We show how to remove files from a file collection:

1. Start **`smit cspoc`** and select **Manage File in File Collections** → **Remove Files from a File Collection**.
2. Select a file collection from the pop-up list end press Enter.
3. Select one or more files from the list and press Enter. See Figure 4-42 on page 200.

```
 Manage File in File Collections

Move cursor to desired item and press Enter.

  Add Files to a File Collection
  Remove Files from a File Collection

  +--------------------------------------------------------------------------+
  ¦       Select one or more files to remove from this File Collection       ¦
  ¦                                                                          ¦
  ¦ Move cursor to desired item and press F7.                                ¦
  ¦     ONE OR MORE items can be selected.                                   ¦
  ¦ Press Enter AFTER making all selections.                                 ¦
  ¦                                                                          ¦
  ¦   /db2/appconfig1.txt                                                    ¦
  ¦   /db2/databases.conf                                                    ¦
  ¦   /db2/appdata.conf                                                      ¦
  ¦                                                                          ¦
  ¦ F1=Help                F2=Refresh              F3=Cancel                 ¦
  ¦ F7=Select              F8=Image                F10=Exit                  ¦
F1¦ Enter=Do               /=Find                  n=Find Next               ¦
F9+--------------------------------------------------------------------------+
```

*Figure 4-42 Remove files from a file collection*

**Important:** You cannot remove files from the HACMP_Files collection this way.

## Manually propagating files in a file collection

You can manually synchronize any file collection (see Figure 4-43 on page 201):

1. Start PowerHA SystemMirror File Collection Management by entering the **`smit cm_filecollection_menu`** fast path.
2. Select **Propagate Files in File Collections**.
3. Select a file collection from the pop-up list and press Enter.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Manual file collection propagation called.
The following file collections will be processed:
app_files
Starting file propagation to remote node p650n01.
Successfully propagated file /db2/appconfig1.txt to node p650n01.
Successfully propagated file /db2/databases.conf to node p650n01.
Successfully propagated file /db2/appdata.conf to node p650n01.
Total number of files propagated to node p650n01: 3

F1=Help            F2=Refresh         F3=Cancel          F6=Command
F8=Image           F9=Shell           F10=Exit           /=Find
n=Find Next
```

*Figure 4-43 Manual propagation of a file collection*

## 4.14 Replacing the repository disk

If you encounter a hardware error on the repository disk or you have to move it to a new storage system, follow these steps:

> **Minimum software levels:** The “repository disk replacement” function is not available in the HA 7.1.1 base code. The following minimum software levels are required for repository disk replacement:
>
> - AIX 6.1 TL7 SP3
> - AIX 7.1 TL1 SP3
> - PowerHA 7.1.1 SP1

1. Ensure that you have a new shared disk that is accessible by all cluster nodes. See “Adding a disk to the cluster nodes” on page 167.
2. Go to **smit sysmirror** and select **Problem Determination Tools** → **Select a new Cluster repository disk**.
3. Select the new repository disk by pressing F4. See Figure 4-44 on page 202.

```
 Select a new Cluster repository disk

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

                                                        [Entry Fields]
* Cluster Name                                        inst1_cluster
* Repository Disk                                    [hdisk2]              +

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

*Figure 4-44 Replacing the cluster repository disk*

# 4.15 C-SPOC cluster user and group management

PowerHA SystemMirror C-SPOC provides an easy-to use interface for cluster-wide user management. You do not have to manually synchronize the user settings between the cluster hosts. It supports both local and Lightweight Directory Access Protocol (LDAP) authentication. We describe the local user management only. For LDAP, see 4.16, “Federated security for cluster-wide security management” on page 208.

The C-SPOC user and group management use the clustered, multiple node version of the standard AIX user and group management commands. All C-SPOC commands use the same parameters as their stand-alone counterparts. The following functions are implemented:

- Add a user
- Change a user
- Remove a user
- List users
- Add a group
- Change a group
- Remove a group
- List groups
- Change your password
- Change a cluster user password
- Manage the users who can change their cluster-wide password

When you use **`smit cspoc`** and select **Security and Users**, you have to select the nodes by the resource group. With this method, you can create users and groups for a specific resource group or application. For example, if you have a 3-node cluster, and one resource group of your application is configured on only two nodes, you can create users who can log on to only these two nodes.

## 4.15.1 Setting up cluster-wide password management with C-SPOC

If you plan to use C-SPOC user and password management, you have to change the default AIX **`passwd`** command to the **`/usr/es/sbin/cluster/utilities/clpasswd`** command:

1. Start **`smit cspoc`** and select **Security and Users** → **Passwords in a PowerHA SystemMirror cluster** → **Modify System Password Utility**.
2. Press F4 and select **Link to Cluster Password Utility**. See Figure 4-45 on page 203.

```
 Modify System Password Utility

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

                                                        [Entry Fields]
* /bin/passwd utility is                             [Link to Cluster Passw> +

  Select nodes by Resource Group                                             +
     *** No selection means all nodes! ***
+--------------------------------------------------------------------------+
  |                        /bin/passwd utility is                            |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   Original AIX System Command                                            |
  |   Link to Cluster Password Utility                                       |
  |                                                                          |
  | F1=Help                 F2=Refresh              F3=Cancel                |
F1| F8=Image                F10=Exit                Enter=Do                 |
F5| /=Find                  n=Find Next                                      |
F9+--------------------------------------------------------------------------+
```

*Figure 4-45 Modify the system password utility*

## 4.15.2 Enabling cluster-wide password management for a user

You have to enable the cluster-wide password management for the users; otherwise, the users can have different passwords on the cluster nodes. Follow these steps:

1. Start `smit cspoc` and select **Security and Users** → **Passwords in a PowerHA SystemMirror cluster** → **Manage List of Users Allowed to Change Password**.
2. Press F4 for the list of available users.
3. Select individual users by pressing F7 to enable the user for cluster-wide password management. You have to reselect the previously enabled users because only the actual selected users are allowed to use the cluster password management feature.
4. Alternatively, you can select ALL_USERS to enable this feature for all users. See Figure 4-46 on page 204.

```
 Manage List of Users Allowed to Change Password

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

                                                      [Entry Fields]
  Users allowed to change password cluster-wide      [bela bubba]           +

F4 lists users defined to all cluster nodes
  +--------------------------------------------------------------------------+
  |              Users allowed to change password cluster-wide               |
  |                                                                          |
  | Move cursor to desired item and press F7.                                |
  |     ONE OR MORE items can be selected.                                   |
  | Press Enter AFTER making all selections.                                 |
  |                                                                          |
  |   ALL_USERS                                                              |
  |   root                                                                   |
  |   adm                                                                    |
  |   esaadmin                                                               |
  |   pconsole                                                               |
  |   sshd                                                                   |
  | > bela                                                                   |
  | > bubba                                                                  |
  |                                                                          |
  | F1=Help                 F2=Refresh              F3=Cancel                |
F1| F7=Select               F8=Image                F10=Exit                 |
F5| Enter=Do                /=Find                  n=Find Next              |
F9+--------------------------------------------------------------------------+
```

*Figure 4-46 Enable users to have cluster-wide passwords*

### 4.15.3 Listing users to use cluster-wide password management

To list the users whom are enabled for cluster-wide password management execute `smit cspoc`, then select **Security and Users** → **Passwords in a PowerHA SystemMirror cluster** → **List Users Allowed to Change Password**.

### 4.15.4 Adding a user to the cluster

We show how to add a user to the cluster:

1. Start `smit cspoc` and select **Security and Users** → **Users in a PowerHA SystemMirror cluster** → **Add a User to the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA creates the user on all cluster nodes. Otherwise, select a resource group, and the C-SPOC creates the user only on the nodes where the resource group is configured.
4. Enter the user name and all other pertinent information. See Figure 4-47 on page 205. Leave the User ID field empty so that PowerHA can select an ID that is free on all nodes.

5. If you want to enable cluster-wide password management for the user, follow the steps in "Enabling cluster-wide password management for a user" on page 203.

```
 Add a User to the Cluster

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

[TOP]                                                    [Entry Fields]
  Select nodes by resource group                        inst1rg
   *** No selection means all nodes! ***

* User NAME                                            [bubba]
  User ID                                              []                   #
  ADMINISTRATIVE USER?                                  true               +
  Primary GROUP                                        [staff]             +
  Group SET                                            [security]          +
  ADMINISTRATIVE GROUPS                                []                  +
  Another user can SU TO USER?                          true               +
  SU GROUPS                                            [ALL]               +
  HOME directory                                       [/data1/bubba]
  Initial PROGRAM                                      []
  User INFORMATION                                     [Bubba Jitsu]
  EXPIRATION date (MMDDhhmmyy)                         [0]
  Is this user ACCOUNT LOCKED?                          false              +
  User can LOGIN?                                       true               +
  User can LOGIN REMOTELY(rsh,tn,rlogin)?               true               +
  Allowed LOGIN TIMES                                  []
  Number of FAILED LOGINS before                       [0]                  #
       user account is locked
[MORE...32]

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

*Figure 4-47 Add a user to the cluster*

## 4.15.5 Changing a cluster user

We show how to change a cluster user:

1. Start **`smit cspoc`** and select **Security and Users** → **Users in a PowerHA SystemMirror cluster** → **Change / Show Characteristics of a User in the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA list users from all cluster nodes.
4. Enter the user name or press F4 for a list of available users.
5. Change to the values that you want and press Enter. See Figure 4-48 on page 206.

```
 Change User Attributes on the Cluster

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

[TOP]                                                    [Entry Fields]
  Resource group
* User NAME                                            bubba
  User ID                                             [204]              #
  ADMINISTRATIVE USER?                                 true             +
  Primary GROUP                                       [staff]           +
  Group SET                                           [staff,security]  +
  ADMINISTRATIVE GROUPS                               []                +
  Another user can SU TO USER?                         true             +
  SU GROUPS                                           [ALL]             +
  HOME directory                                      [/data1/bubba]
  Initial PROGRAM                                     [/usr/bin/ksh]
[MORE...45]

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

*Figure 4-48 Change a cluster user*

## 4.15.6 Removing a user from the cluster

We show how to remove a user from the cluster:

1. Execute the **smit cspoc** command, then go to **Security and Users** → **Users in an PowerHA SystemMirror cluster** → **Remove a User from the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA selects users from all cluster nodes.
4. Enter the user name or press F4 for a list of available users.
5. Select **yes** to “Remove AUTHENTICATION information”. This way, C-SPOC removes all user information from the /etc/security/passwd file.

## 4.15.7 Listing users in the cluster

We show how to list the users in the cluster:

1. Start **smit cspoc** and select **Security and Users** → **Users in an PowerHA SystemMirror cluster** → **List Users in the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA lists users from all cluster nodes.

## 4.15.8 Adding a group in the cluster

We show how to add a group in the cluster:

1. Start **smit cspoc** and select **Security and Users** → **Groups in an PowerHA SystemMirror cluster** → **Add a Group to the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA creates the group on all cluster nodes. Otherwise, select a resource group, and the C-SPOC creates the group only on the nodes where the resource group is configured.
4. Enter the group name and all other required information. See Figure 4-49. Leave the Group ID field empty so that PowerHA can select an ID that is free on all nodes.

```
 Add a Group to the Cluster

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

                                                        [Entry Fields]
  Select nodes by resource group
   *** No selection means all nodes! ***

* Group NAME                                         [appusers]
  ADMINISTRATIVE group?                               false                  +
  Group ID                                           []                       #
  USER list                                          [bela,bubba]            +
  ADMINISTRATOR list                                 [bubba]                 +
  Initial Keystore Mode                              []                      +
  Keystore Encryption Algorithm                      []                      +
  Keystore Access                                    []                      +

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

*Figure 4-49 Add a group to the cluster*

## 4.15.9 Changing a group in the cluster

We show how to change a group in the cluster:

1. Start **smit cspoc** and select **Security and Users** → **Groups in an PowerHA SystemMirror cluster** → **Change / Show Characteristics of a Group in the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES).**
3. Select nodes by resource group: If you leave it empty, PowerHA lists groups from all cluster nodes.
4. Enter the group name or press F4 for the list of available groups.
5. Change to the values that you want.

## 4.15.10 Removing a group from the cluster

We show how to remove a group from the cluster:

1. Execute **smit cspoc**, and select **Security and Users** → **Groups in an PowerHA SystemMirror cluster** → **Remove a Group from the Cluster**.

2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA selects groups from all cluster nodes.
4. Enter the group name or press F4 for the list of available groups.
5. Press Enter to remove the selected group from the cluster.

### 4.15.11 Listing groups from the cluster

We show the step to list groups from the cluster:

1. Start `smit cspoc` and select **Security and Users** → **Groups in an PowerHA SystemMirror cluster** → **List All Groups in the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA selects groups from all cluster nodes.

### 4.15.12 Changing a user password in the cluster

We show how to change a user password in the cluster:

1. Start `smit cspoc`, then select **Security and Users** → **Passwords in an PowerHA SystemMirror cluster** → **Change a User's Password in the Cluster**.
2. Select the authentication and registry mode: **LOCAL(FILES)**.
3. Select nodes by resource group: If you leave it empty, PowerHA selects users from all cluster nodes.
4. Enter the user name or press F4 for the list of available users.
5. Change the password for the user.

### 4.15.13 Changing your password in the cluster

You can use `/usr/es/sbin/cluster/utilities/clpasswd` to change your password. If you are enabled to have cluster-wide password management, your password is changed on all nodes.

## 4.16 Federated security for cluster-wide security management

The AIX operating system provides a rich set of security capabilities. The goal of federated security is to enable the security administration of AIX security features across the cluster.

Federated security addresses Lightweight Access Protocol (LDAP), role-based access control (RBAC), and Encrypted file system (EFS) integration into cluster management.

Through the federated security cluster, users are able to manage roles and the encryption of data across the cluster.

## 4.16.1 Federated security components

Federated security integrates components and features such as LDAP and RBAC into the cluster management. We look into the functional value of each component in the cluster management as shown in Figure 4-50.

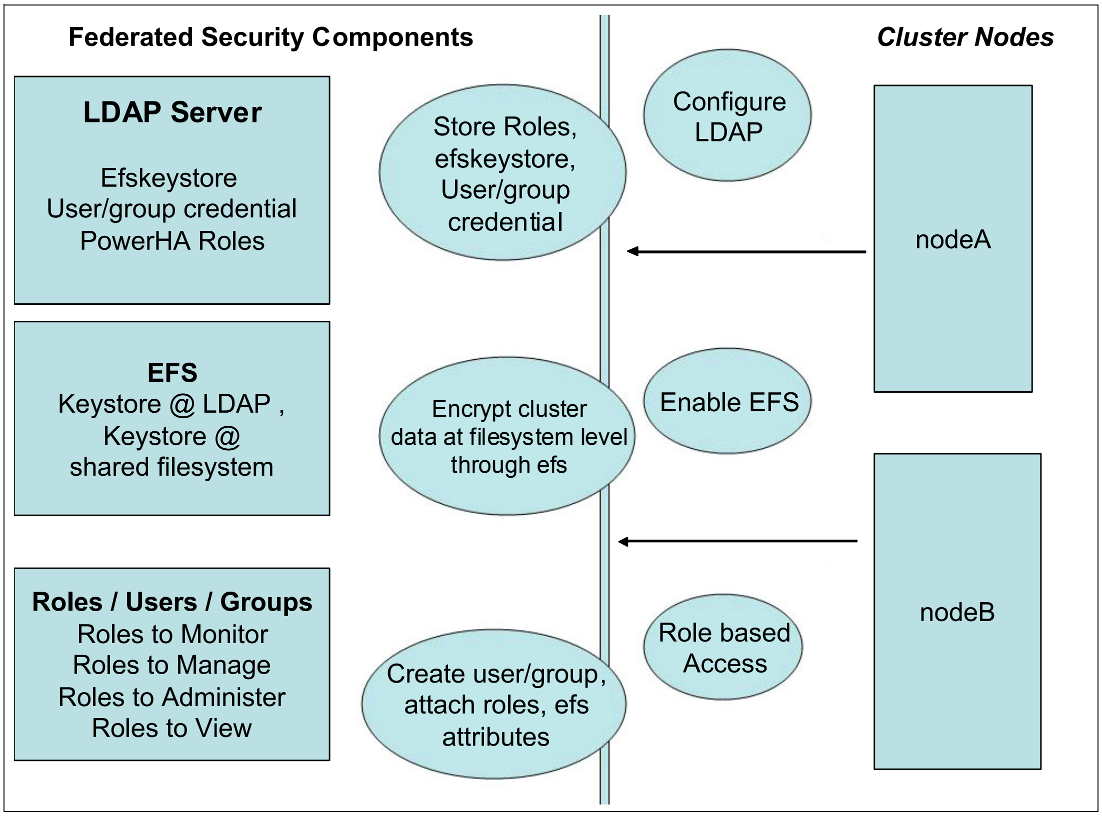



*Figure 4-50 Federated security components*

### LDAP

The LDAP method is used by cluster nodes to allow centralized security authentication and access to user and group information.

The following information is stored by the federated security at LDAP:

- PowerHA roles: As part of LDAP client configuration through PowerHA (details of PowerHA roles and LDAP client configuration are explained in the following section).
- EFS keystore: Exports EFS keystore data to LDAP. Stores the local user and group keystores to LDAP when EFS is enabled through PowerHA. Details are in “EFS” on page 210.
- Stores user and group account information for authorization.

Details of the LDAP server and client configuration are explained in “Configuring LDAP” on page 211.

The following supported LDAP servers can be configured for federated security:

- IBM Tivoli Director server
- Windows Active Directory server

All cluster nodes must be configured with the LDAP server and the client filesets. PowerHA provides options to configure the LDAP server and client across all cluster nodes.

> **SSL:** Secure Sockets Layer (SSL) connection is mandatory for binding LDAP clients to servers. Remember to configure SSL in the cluster nodes.

## LDAP

The LDAP server and client configuration is provided through the PowerHA smitty option and the System Director PowerHA plug-in.

For the LDAP server and client setup, SSL must be configured. The SSL connection is mandatory for binding LDAP clients to servers.

> **LDAP server:** The LDAP server needs to be configured on all cluster nodes. If there is an existing LDAP server, it can be incorporated into PowerHA for federated security usage.

## RBAC

Cluster administration is an important aspect of high availability operations, and security in the cluster is an inherent part of most system administration functions. Federated security integrates the AIX RBAC features to enhance the operational security.

During LDAP client configuration, four PowerHA defined roles are created in LDAP:

- `ha_admin`: Provides ADMIN authorization for the relevant cluster functionality. For example, taking a cluster snapshot is under administrator authorization.
- `ha_op`: Provides OPERATOR authorization for the relevant cluster functionality. For example, “move cluster resource group” is under operator authorization.
- `ha_mon`: Provides MONITOR authorization for the relevant cluster functionality. For example, the command **`clRGinfo`** is under monitor authorization.
- `ha_view`: Provides VIEW authorization. It has all read permissions for the cluster functionality.

These roles can be assigned to the user to provide restricted access to the cluster functionality based on the role.

> **Roles:** PowerHA roles are created while you configure the LDAP client in the cluster nodes.

## EFS

The Encrypted File System (EFS) enables users on the system to encrypt their data in the journaled file system 2 (JFS2) through their individual keystores. The keys are stored in a cryptographically protected keystore. On a successful login, the user keys are loaded into the kernel and associated with the process credentials.

From the federated security perspective, the EFS keystores are stored in LDAP. There is an option to store the keystores through a shared file system in the cluster environment if LDAP is not configured in the cluster.

**Tip**: Store the EFS keystore in LDAP. As an option, if the LDAP environment is not configured, the keystore can be stored in a Network File System (NFS) mounted file system. In order to avoid the risk of grace period expired during fallover event of the NFSv4 resource group, the resource group run-time policies from the default of parallel into serial release and acquisition order is suggested. This will enforce the NFS un-mounting first on local node before the resource group fallover. This helps avoid NFS cross-mount.

## 4.16.2 Federated security configuration requirement

The following prerequisites are necessary for a complete federated security environment:

- LDAP Configuration:
  - DB2 V9.7
  - GSKit filesets, preferably version 8
  - LDAP Server filesets (Tivoli Director server 6.3 version)
  - LDAP Client filesets (Tivoli Director server 6.3 version)
- RBAC configuration
- EFS environment

The filesets for RBAC and EFS are available by default in AIX 6.1 and later versions, and nothing specific is required. The challenge is to configure LDAP.

### Configuring LDAP

Use the following steps to install the LDAP configuration:

1. Install and Configure DB2.
2. Install `GSkit` filesets.
3. Install Tivoli Director server (LDAP server and client) filesets.

### Installing DB2

We provide the DB2 installation steps as shown in Example 4-13.

*Example 4-13 DB2 installation steps*

```
# ./db2_install

Default directory for installation of products – /opt/IBM/db2/V9.7
Do you want to choose a different directory to install [yes/no] ?
no
Specify one of the following keywords to install DB2 products.
ESE <<<<<< Select ESE >>>>>>
CLIENT
RTCL
Enter “help” to redisplay product names.
Enter “quit” to exit.
***********************************************************

ESE <<<< selected option >>>>>
DB2 installation is being initialized.
Total number of tasks to be performed: 46
Total estimated time for all tasks to be performed: 2369
Task #1 start
Description: Checking license agreement acceptance
```

```
Estimated time 1 second(s)
Task #1 end
Task #47 end
The execution completed successfully.
```

### GSkit filesets

Ensure that the `GSkit` filesets are installed in both server and client nodes, basically in all cluster nodes. See Example 4-14 on page 212.

*Example 4-14 GSkit fileset installation*

```
Installing GSkit (64-bit)
installp -acgXd . GSKit8.gskcrypt64.ppc.rte
installp -acgXd . GSKit8.gskssl64.ppc.rte
Installing GSkit (32-bit)
installp -acgXd . GSKit8.gskcrypt32.ppc.rte
installp -acgXd . GSKit8.gskssl32.ppc.rte
Install AIX Certificate and SSL base
installp -acgXd . gsksa.rte
installp -acgXd . gskta.rte
```

Ensure that the SSL filesets are configured as shown in Example 4-15.

*Example 4-15 SSL filesets*

```
# lslpp -l | grep ssl
  GSKit8.gskssl32.ppc.rte   8.0.14.7  COMMITTED  IBM GSKit SSL Runtime With
  GSKit8.gskssl64.ppc.rte   8.0.14.7  COMMITTED  IBM GSKit SSL Runtime With
  openssl.base            0.9.8.1100  COMMITTED  Open Secure Socket Layer
  openssl.license          0.9.8.801  COMMITTED  Open Secure Socket License
  openssl.man.en_US       0.9.8.1100  COMMITTED  Open Secure Socket Layer
  openssl.base            0.9.8.1100  COMMITTED  Open Secure Socket Layer
```

### Tivoli Directory Server (LDAP) filesets

We show the Tivoli Directory Server (LDAP) filesets as shown in Example 4-16.

*Example 4-16 LDAP client and server filesets*

```
# lslpp -l | grep idsldap
idsldap.clt32bit63.rte 6.3.0.3 COMMITTED Directory Server - 32 bit
idsldap.clt64bit63.rte 6.3.0.3 COMMITTED Directory Server - 64 bit
idsldap.clt_max_crypto32bit63.rte
idsldap.clt_max_crypto64bit63.rte
idsldap.cltbase63.adt 6.3.0.3 COMMITTED Directory Server - Base Client
idsldap.cltbase63.rte 6.3.0.3 COMMITTED Directory Server - Base Client
```

**More information:** For complete DB2 and LDAP configuration details, see this website:

`https://www.ibm.com/developerworks/mydeveloperworks/wikis/home/wiki/PowerHA%20SystemMirror/page/PowerHA%20Cluster%20with%20Federated%20Security?lang=en`

## 4.16.3 Federated security configuration details

After the required filesets are installed, the federated security can be configured by using the following two options:

- PowerHA smitty panel
- PowerHA SystemMirror PowerHA plug-in for Systems Director

### LDAP configuration

The LDAP configuration by using the smitty panel can be reached via **System Management (C-SPOC)** → **LDAP** as shown in Figure 4-51 on page 213.

```
          System Management (C-SPOC)

Move cursor to desired item and press Enter.

  Storage
  PowerHA SystemMirror Services
  Communication Interfaces
  Resource Group and Applications
  PowerHA SystemMirror Logs
  PowerHA SystemMirror File Collection Management
  Security and Users
  LDAP
  Configure GPFS

  Open a SMIT Session on a Node


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-51 LDAP configuration panel through C-SPOC*

### LDAP server configuration

Under the LDAP server configuration, two options are provided (Figure 4-52 on page 214):

- Configure a new LDAP server
- Add an existing LDAP server

If an LDAP server is already configured, the cluster nodes can use the existing LDAP server or configure a new LDAP server.

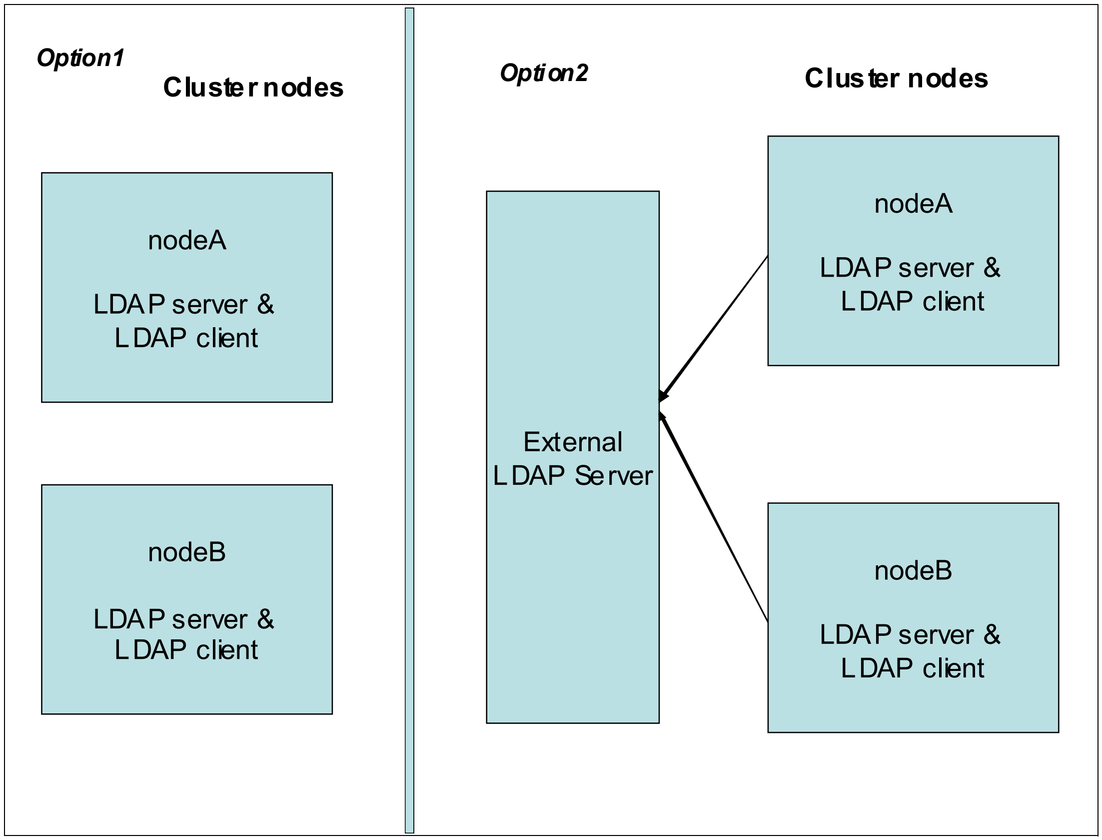



*Figure 4-52 LDAP server configuration option*

## Configuring a new LDAP server

To configure a new LDAP server (Figure 4-53 on page 215), enter **`smitty hacmp`** and select **System Management (C-SPOC)** → **LDAP** → **LDAP Server** → **Configure a new LDAP server**.

```
         Configure a new LDAP server

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

                                                     [Entry Fields]
* Hostname(s)                                                              +
* LDAP Administrator DN                             [cn=admin]
* LDAP Administrator password                       []
  Schema type                                        rfc2307aix
* Suffix / Base DN                                  [cn=aixdata,o=ibm]
* Server port number                                [636]                   #
* SSL Key path                                      []                      /
* SSL Key password                                  []
* Version                                                                  +
* DB2 instance password                             []
* Encryption seed for Key stash files               []

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

*Figure 4-53 New LDAP server configuration*

Consider these key points:

- The existence of any LDAP instance is verified. The configuration continues only if the instance name is *not* ldapdb2. Have only one instance for federated security purposes.
- Internally, the configuration creates a peer-to-peer configuration to avoid LDAP instance failure. Therefore, a minimum of two nodes are expected as input.
- The configuration loads the local user and group information into LDAP.
- The configuration loads the RBAC AIX tables into LDAP.
- The configuration loads the EFS keystore that is defined for users and groups into LDAP.
- Various data trees that are created in LDAP are in the `/etc/security/ldap/sectoldif.cfg` file.

**Encryption:** The encryption seed must be a minimum of 12 characters.

The success of the LDAP configuration can be verified by using the ODM command in Example 4-17.

*Example 4-17 ODM command to verify LDAP configuration for federated security*

```
# odmget -q "group=LDAPServer and name=ServerList" HACMPLDAP

HACMPLDAP:
        group = "LDAPServer"
        type = "IBMExisting"
        name = "ServerList"
        value = "selma06,selma07"
```

## Adding an existing LDAP server configuration

To add an existing LDAP server (Figure 4-54), enter **`smitty hacmp`** and select **System Management (C-SPOC)** → **LDAP** → **LDAP Server** → **Add an existing LDAP server**.

```
          Add an existing LDAP server

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

                                                         [Entry Fields]
* LDAP server(s)                                      []
* Bind DN                                             [cn=admin]
* Bind password                                       []
* Suffix / Base DN                                    [cn=aixdata,o=ibm]
* Server port number                                  [636]                #
* SSL Key path                                        []                   /
* SSL Key password                                    []

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

*Figure 4-54 Adding an existing LDAP Server*

Consider these key points:

- Temporarily, the LDAP client is configured to verify the LDAP server input parameters. The compatible LDAP client fileset must be at least at version 6.2.
- It adds the user/group RBAC tables and EFS keystore into the existing LDAP server.

The success of adding an existing LDAP server is verified with the ODM command as shown in Example 4-18.

*Example 4-18 ODM command to verify the existing LDAP server configuration for federated security*

```
# odmget -q "group=LDAPServer and name=ServerList" HACMPLDAP

HACMPLDAP:
        group = "LDAPServer"
        type = "IBMExisting"
        name = "ServerList"
        value = "selma06,selma07"
```

## LDAP client configuration

To configure the LDAP client (Figure 4-55 on page 217 and Figure 4-56 on page 217), enter **`smitty hacmp`** and select **System Management (C-SPOC)** → **LDAP** → **LDAP Client** → **Configure LDAP client**.

```
 LDAP Client

Move cursor to desired item and press Enter.

  Configure LDAP client
  Show LDAP client configuration
  Delete the LDAP client



F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-55 LDAP client configuration*

```
         Configure LDAP client

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

                                                       [Entry Fields]
* LDAP server(s)                                    []                      +
* Bind DN                                           [cn=admin]
* Bind password                                     []
  Authentication type                                ldap_auth
* Suffix / Base DN                                  [cn=aixdata,o=ibm]
* +--------------------------------------------------------------------------+#
* |                              LDAP server(s)                              |/
* |                                                                          |
  | Move cursor to desired item and press F7.                                |
  |     ONE OR MORE items can be selected.                                   |
  | Press Enter AFTER making all selections.                                 |
  |                                                                          |
  |   quimby06                                                               |
  |                                                                          |
  | F1=Help                 F2=Refresh              F3=Cancel                |
F1| F7=Select               F8=Image                F10=Exit                 |
F5| Enter=Do                /=Find                  n=Find Next              |
F9+--------------------------------------------------------------------------+
```

*Figure 4-56 LDAP client configuration parameters*

Consider these key points:

- Ensure that the LDAP client filesets and `GSkit` filesets are installed as described in “Federated security configuration requirement” on page 211. The minimum compatible LDAP client filesets must be version 6.2.
- The RBAC is configured during client configuration. The PowerHA defined roles are created in the LDAP server.
- It generates SSL keys, extracts the server certificate, and binds with SSL as part of the LDAP client configuration.
- It enables the home directory automatically at user login, which is required by LDAP users.

Verify the client configuration by using the ODM command as shown in Example 4-19.

*Example 4-19 ODM command to verify LDAP client configuration*

```
# odmget -q "group=LDAPClient and name=ServerList" HACMPLDAP

HACMPLDAP:
        group = "LDAPClient"
        type = "ITDSClinet"
        name = "ServerList"
        value = "selma06,selma07"
```

Also, the client configuration can be verified by checking the LDAP client daemon status by using the command as shown in Example 4-20.

*Example 4-20 Verify the client daemon status after LDAP client configuration*

```
# ps -eaf | grep secldapclntd
    root  4194478        1   2 04:30:09      -  0:10 /usr/sbin/secldapclntd
```

## RBAC

During the LDAP client configuration, the PowerHA defined roles are created in the LDAP server.

Verify the configuration of the RBAC roles in the LDAP server by using the ODM command as shown in Example 4-21.

*Example 4-21 ODM command to verify RBAC configuration into LDAP server*

```
# odmget -q "group=LDAPClient and name=RBACConfig" HACMPLDAP

HACMPLDAP:
        group = "LDAPClient"
        type = "RBAC"
        name = "RBACConfig"
        value = "YES"
```

Verify the four PowerHA defined roles that are created in LDAP as shown in Example 4-22.

*Example 4-22 Roles that are defined by PowerHA*

```
# lsrole -a ALL | grep ha*
ha_admin
ha_op
ha_mon
ha_view
```

Example 4-22 shows that the RBAC is configured and can be used by the cluster users and groups. The usage scenario of roles by cluster users and groups are defined in the following section.

## EFS

The EFS management scenario is shown in the flow chart in Figure 4-57 on page 219.

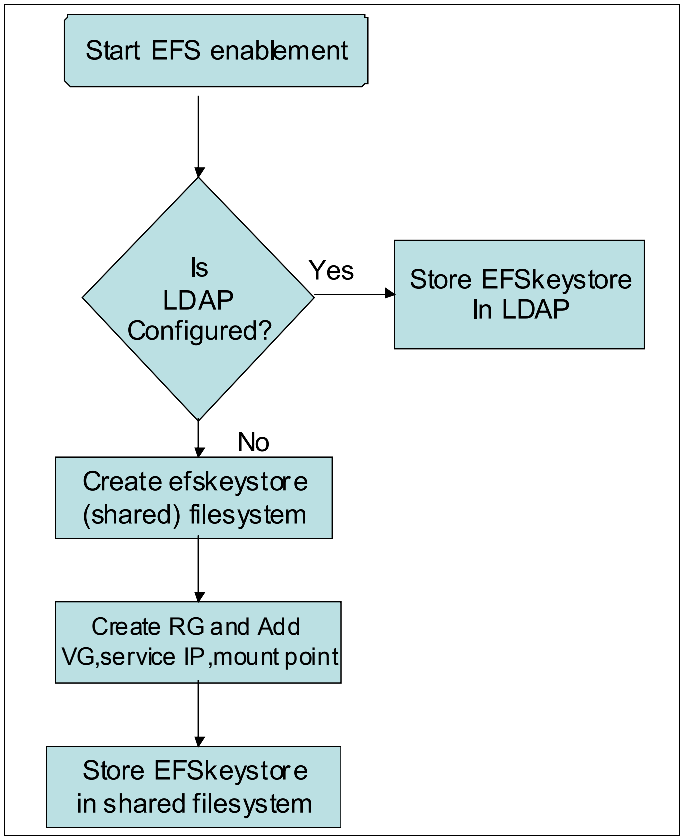



*Figure 4-57 EFS management*

To configure the EFS management configuration (Figure 4-58), enter **`smitty hacmp`** and select **System Management (C-SPOC)** → **Security and Users** → **EFS Management**.

```
          Security and Users

Move cursor to desired item and press Enter.

  PowerHA SystemMirror Cluster Security
  Users in an PowerHA SystemMirror cluster
  Groups in an PowerHA SystemMirror cluster
  Passwords in an PowerHA SystemMirror cluster
  EFS Management


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-58 EFS management*

Under EFS management, the options are provided to enable EFS and to store keystores either in LDAP or a shared file system.

**Important:** Federated security mandates that the LDAP configuration creates roles and stores EFS keystores. You can store EFS keystores under the shared file system only if LDAP is not configured.

### EFS keystore in LDAP

If LDAP is configured, only the LDAP option is available to store the EFS keystore (Figure 4-59).

```
          Enable EFS Keystore

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

                                                         [Entry Fields]
* EFS keystore mode                                    LDAP              +
  EFS admin password                                  []
  Volume group for EFS Keystore                       []                 +
  Service IP                                          []                 +


  +--------------------------------------------------------------------------+
  |                            EFS keystore mode                             |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   LDAP                                                                   |
  |   Shared Filesystem                                                      |
  |                                                                          |
  | F1=Help                F2=Refresh             F3=Cancel                  |
F1| F8=Image               F10=Exit               Enter=Do                   |
F5| /=Find                 n=Find Next                                       |
F9+--------------------------------------------------------------------------+
```

*Figure 4-59 EFS keystore mode*

**Important:** The volume group and service IP are invalid and ignored in LDAP mode.

Consider these key points:

- Enables EFS by using the **/usr/sbin/efsenable -a -d cn=aixdata,o=ibm** command.
- Prompts the user to enter the password to protect the initial keystore.

Verify the EFS enablement as understood by the cluster as shown in Example 4-23.

*Example 4-23 ODM command to verify the EFS enablement status*

```
# odmget  -q "group=EFSKeyStore AND name=mode"  HACMPLDAP

HACMPLDAP:
        group = "EFSKeyStore"
        type = "EFS"
        name = "mode"
        value = "1"
```

### EFS keystore in a shared file system

If LDAP is not configured but you want to use EFS to encrypt the cluster data, federated security provides an option to store the EFS keystore in a shared file system.

As shown in Figure 4-13 on page 168, to enable EFS and to store the EFS keystore in the shared file system, provide the volume group and service IP details:

- The volume group to store the EFS keystore in a file system
- The service IP to mount the file system where the keystore is stored so that it is highly available to cluster nodes

Consider these key points during this configuration:

- Creates the EFS keystore file system in the specified volume group
- Creates the EFS mount point on all cluster nodes
- Creates the resource group to include the NFS exports with fallback as an NFS option
- Adds a specified volume group in the resource group
- Adds a service IP and a mount point in the resource group

**Important:** The file system creation, mount point, and NFS export are performed internally under the EFS keystore in a shared file system option.

Verify the configuration by using the ODM command as shown in Example 4-24.

*Example 4-24 ODM command to verify EFS configuration in shared file system mode*

```
# odmget  -q "group=EFSKeyStore AND name=mode"  HACMPLDAP

HACMPLDAP:
        group = "EFSKeyStore"
        type = "EFS"
        name = "mode"
        value = "2"
```

## 4.16.4 User and group management under federated security

From the federated security perspective, the users and groups that are created can be authenticated through LDAP.

### User management

We show user management in PowerHA (Figure 4-60 on page 222).

```
        Users in an PowerHA SystemMirror cluster

Move cursor to desired item and press Enter.

  Add a User to the Cluster
  Change / Show Characteristics of a User in the Cluster
  Remove a User from the Cluster
  List Users in the Cluster



F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-60 User management*

To reach user management, enter **smitty hacmp** and select **System Management (C-SPOC) → Security and Users → Users in an PowerHA SystemMirror cluster**.

To create a user, you can set the authentication and registry mode to either LDAP or local as shown in Figure 4-61.

```
 Users in an PowerHA SystemMirror cluster

Move cursor to desired item and press Enter.

  Add a User to the Cluster
  Change / Show Characteristics of a User in the Cluster
  Remove a User from the Cluster
  List Users in the Cluster


  +--------------------------------------------------------------------------+
  |                Select an Authentication and registry mode                |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   LOCAL(FILES)                                                           |
  |   LDAP                                                                   |
  |                                                                          |
  | F1=Help                F2=Refresh             F3=Cancel                  |
  | F8=Image               F10=Exit               Enter=Do                   |
F1| /=Find                 n=Find Next                                       |
F9+--------------------------------------------------------------------------+
```

*Figure 4-61 User creation with registry mode as LDAP or local*

The new user can be assigned with the PowerHA roles that are created as part of the federated security LDAP configuration as shown in Figure 4-62 on page 223.

```
         Add a User to the LDAP

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

[TOP]                                                  [Entry Fields]
* User NAME                                         []
  User ID                                           []                       #
* Roles                                             [ha_admin]              +
* Registry                                           LDAP
* Login Authentication Grammar                       LDAP
  Keystore Access                                    LDAP
  ADMINISTRATIVE USER?                               false                  +
  Primary GROUP                                     []                      +
  Group SET                                         []                      +
  ADMINISTRATIVE GROUPS                             []                      +
  Another user can SU TO USER?                       true                   +
  SU GROUPS                                         [ALL]                   +
  HOME directory                                    []
[MORE...38]

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

*Figure 4-62 User with PowerHA defined role*

## Group management

The group management panel can be reached (Figure 4-63) by starting **smitty hacmp**. Then, select **System Management (C-SPOC)** → **Security and Users** → **Groups in an PowerHA SystemMirror cluster**.

```
        Groups in an PowerHA SystemMirror cluster

Move cursor to desired item and press Enter.

  Add a Group to the Cluster
  Change / Show Characteristics of a Group in the Cluster
  Remove a Group from the Cluster
  List All Groups in the Cluster

F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

*Figure 4-63 Group management*

Similar to user management, group management can set the registry and authentication mode as LDAP or local as shown in Figure 4-61 on page 222.

# 4.17 IBM System Director plug-in update

The PowerHA Standard Edition configuration can be done via the System Director SysMirror Plug-in.

The architecture of the IBM System Director SysMirror Plug-in is shown for quick reference in Figure 4-64.

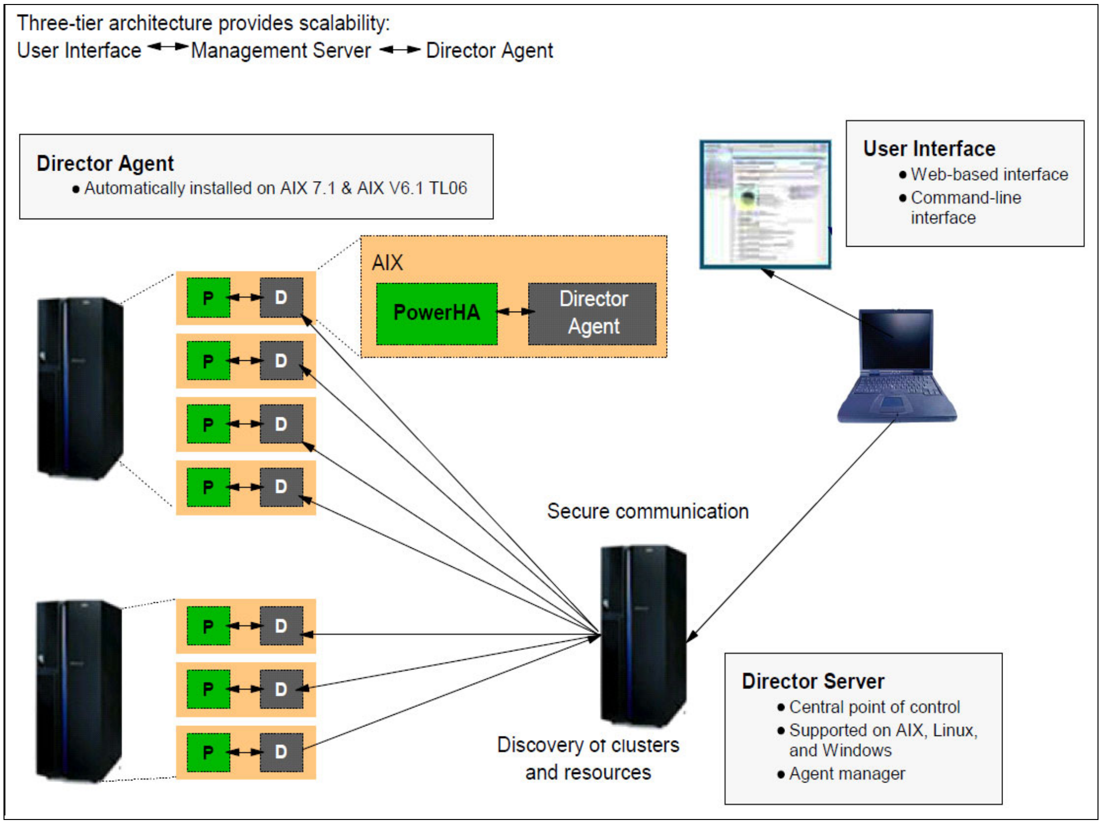



*Figure 4-64 System Director architecture that involves the PowerHA plug-in*

Figure 4-65 on page 225 shows the plug-in compatibility matrix with System Director versions.

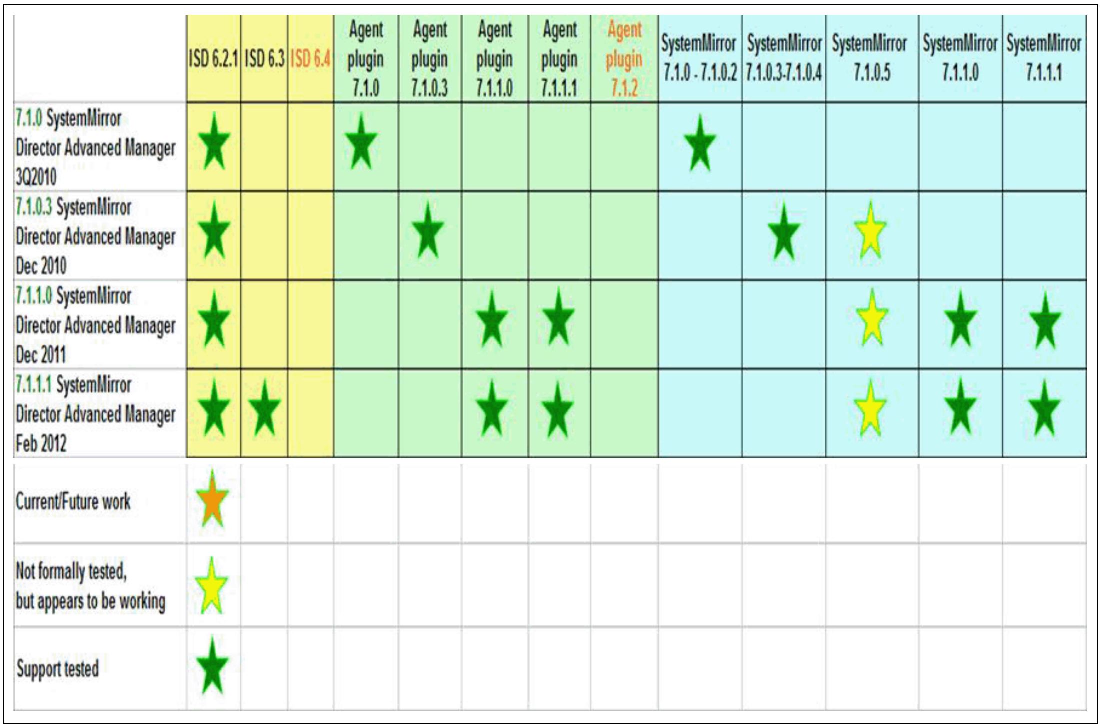



*Figure 4-65 System Director support matrix*

The SystemMirror Director Advanced Manager is the plug-in that needs to be deployed at the System Director. The Agent Plug-in needs to be deployed at the cluster nodes.

The System Director plug-in offers these advantages:

- A single, centralized view into all PowerHA SystemMirror clusters:
  - Centralized and secure access point
  - Single sign-on (SSO) capability
- Scheduling

  For example, functionality, such as Create Cluster and Create Resource Group, can be scheduled.
- Event tracking

  For example, cluster availability can be monitored by using event monitors and event automation plans.
- Multiple cluster management
- Easy to configure through wizards

5

# Smart Assists for PowerHA SystemMirror

Smart Assists for PowerHA SystemMirror 7.1.1 manages a collection of PowerHA SystemMirror components that you identify to support a particular application. You can view these collections of PowerHA SystemMirror components as a single entity. In PowerHA SystemMirror, the entity is represented by an application name.

This chapter explains how to install and configure a hot standby 2-node IBM PowerHA SystemMirror 7.1.1 cluster by using two of the new Smart Assists that are introduced in this version. Also, there is an example of using the Smart Assist for IBM Tivoli Storage Manager servers. The lab cluster smass is used for the examples with the participating nodes smass1 and smass2.

This chapter describes how to install and configure the following Smart Assist applications:

- Smart Assist for IBM Tivoli Storage Manager
- Smart Assist for IBM MQ Series
- Smart Assist for IBM Tivoli Directory Server

## 5.1 Installing PowerHA SystemMirror Smart Assist

You can install PowerHA SystemMirror Smart Assist from media, a hard drive, or an installation server. The media for installing PowerHA SystemMirror Smart Assist contains the filesets that are shown in Table 5-1.

*Table 5-1 Smart Assists filesets on installation media*

| Fileset | Description |
|---|---|
| cluster.es.assist.common | PowerHA SystemMirror Smart Assist Common Files |
| cluster.es.assist.db2 | PowerHA SystemMirror Smart Assist for DB2 |
| cluster.es.assist.dhcp | PowerHA SystemMirror Smart Assist for Dynamic Host Configuration Protocol (DHCP) |
| cluster.es.assist.dns | PowerHA SystemMirror Smart Assist for Domain Name System (DNS) |
| cluster.es.assist.filenet | PowerHA SystemMirror Smart Assist for IBM FileNet® P8 |
| cluster.es.assist.ihs | PowerHA SystemMirror Smart Assist for IBM HTTP Server |
| cluster.es.assist.tds | PowerHA SystemMirror Smart Assist for IBM Tivoli Directory Server |
| cluster.es.assist.wmq | PowerHA SystemMirror Smart Assist for MQ Series |
| cluster.es.assist.oracle | PowerHA SystemMirror Smart Assist for Oracle |
| cluster.es.assist.oraappsrv | PowerHA SystemMirror Smart Assist for Oracle Application Server |
| cluster.es.assist.printServer | PowerHA SystemMirror Smart Assist for Print Subsystem |
| cluster.es.assist.sap | PowerHA SystemMirror Smart Assist for SAP |
| cluster.es.assist.maxdb | PowerHA SystemMirror Smart Assist for SAP MaxDB |
| cluster.es.assist.websphere | PowerHA SystemMirror Smart Assist for IBM WebSphere® |
| cluster.es.assist.domino | PowerHA SystemMirror SmartAssist for IBM Lotus® domino server |
| cluster.es.assist.tsmadmin | PowerHA SystemMirror SmartAssist for IBM Tivoli Storage Manager Admin center |
| cluster.es.assist.tsmclient | PowerHA SystemMirror SmartAssist for IBM Tivoli Storage Manager Client |
| cluster.es.assist.tsmserver | PowerHA SystemMirror SmartAssist for IBM Tivoli Storage Manager Server |

Verify that your environment meets the specific requirements for the solution deployment before you install and use the various PowerHA SystemMirror Smart Assist applications, for example:

- The node has at least 5 MB of space in the `/usr` directory.
- The systems run PowerHA SystemMirror 7.1.1 SP 02, or later.
- The environment uses AIX Version 7.1 with the 7100-01 SP3 Technology Level, or later.
- You have a PowerHA SystemMirror cluster, shared storage, and connectivity.

**Prerequisites:** Read the Smart Assist manuals carefully for any application-specific prerequisite.

## 5.2 Scenario

In our scenario, we use two AIX logical partitions (LPARs) that are called smass1 and smass2. They both share the storage disk access that is used for the repository disk and data disk.

In each case that we tested, we clear the cluster configuration before installation and configure the next required PowerHA Smart Assist. The base diagram of the configuration is shown in Figure 5-1.

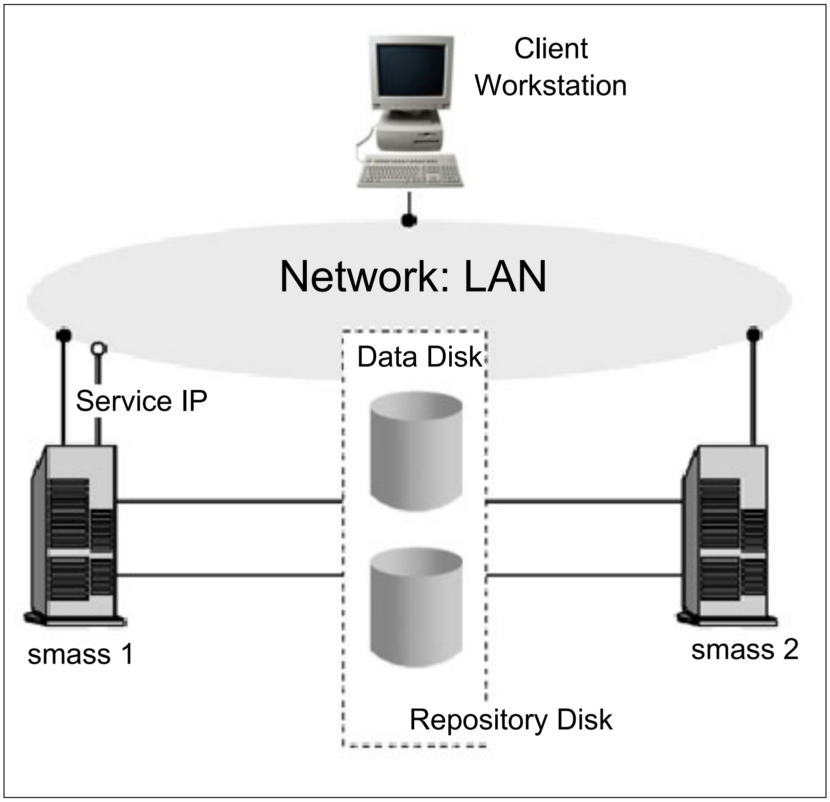



*Figure 5-1 Configuration that is used to test PowerHA Smart Assist*

Figure 5-2 on page 230 shows the cluster configuration after we create the base cluster and before we install and configure any PowerHA Smart Assist products.

```
Cluster Name: smass
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.46
There are 2 node(s) and 1 network(s) defined
NODE smass1:
        Network net_ether_01
                smass1  10.1.1.46
NODE smass2:
        Network net_ether_01
                smass2  10.1.1.31

No resource groups defined
```

*Figure 5-2 Base cluster configuration*

## 5.3 Smart Assist for IBM Tivoli Storage Manager

Before you start planning, you need to have a basic understanding of the different components of the IBM Tivoli Storage Manager.

### Tivoli Storage Manager server

The Tivoli Storage Manager server provides backup, archive, and space management services to the clients. You can set up multiple servers in your environment to balance storage, processor, and network resources. By using Smart Assist for Tivoli Storage Manager, you can configure the IBM Tivoli Storage Manager server instance for high availability.

### Tivoli Storage Manager client

A client node can be a workstation, a personal computer, a file server, or another IBM Tivoli Storage Manager server. The client node uses installed Tivoli Storage Manager client software and is registered with the server.

### Tivoli Storage Manager backup-archive client

With the backup-archive client, you can maintain backup versions of files, which you can use to restore the original files if they are lost or damaged. You can also archive files for long-term storage and retrieve the archived files when necessary. You can register workstations and file servers as client nodes with the Tivoli Storage Manager server.

### Tivoli Storage Manager for Space Management

Tivoli Storage Manager for Space Management provides you with space management services for workstations. The space management function is essentially a more automated version of archive. Tivoli Storage Manager for Space Management automatically migrates files that are less frequently used to server storage, freeing space on the workstation.

### Tivoli Storage Manager admin center

The Tivoli Storage Manager admin center and integrated solutions console are included in the IBM Tivoli Storage Manager product distribution. They are installed with the IBM Tivoli Storage Manager server (as an optional component).

## 5.3.1 Planning for Smart Assist for Tivoli Storage Manager Server

Before you can use the Smart Assist for Tivoli Storage Manager server, review the following information:

- The Tivoli Storage Manager server instance must have the service IP that can be pinged by the Smart Assist for Tivoli Storage Manager client.
- The Tivoli Storage Manager server instance name and Tivoli Storage Manager server user name must be identical and configured on all nodes in the cluster.
- The following information is shared between the nodes in the cluster for an IBM Tivoli Storage Manager server instance:
  - Database directories
  - Instance log volumes
  - Storage pool volumes
  - Instance directory
  - Instance user directory
- Any number of Tivoli Storage Manager server instances can be configured by using the Smart Assist for Tivoli Storage Manager. However, the volume groups that are used for each Tivoli Storage Manager server instance must be different from other Tivoli Storage Manager server instances.
- When you add a Tivoli Storage Manager server instance, it must be defined in the `dsm.sys` and `dsm.opt` files in the `/opt/tivoli/tsm/server/bin/tsmdiag` directory on all nodes of the cluster that are used for starting, stopping, and monitoring the Tivoli Storage Manager server instance. For example, the `dsm.sys` file must contain the following information:
  - servername `tsmsmass`
  - commmethod `tcpip`
  - tcpport `1500`
  - tcpserveraddress `abc.us.com`

  The `dsm.opt` file must contain the servername `tsmsmass`.

Figure 5-3 on page 232 displays a typical 2-node configuration for a Tivoli Storage Manager Server.

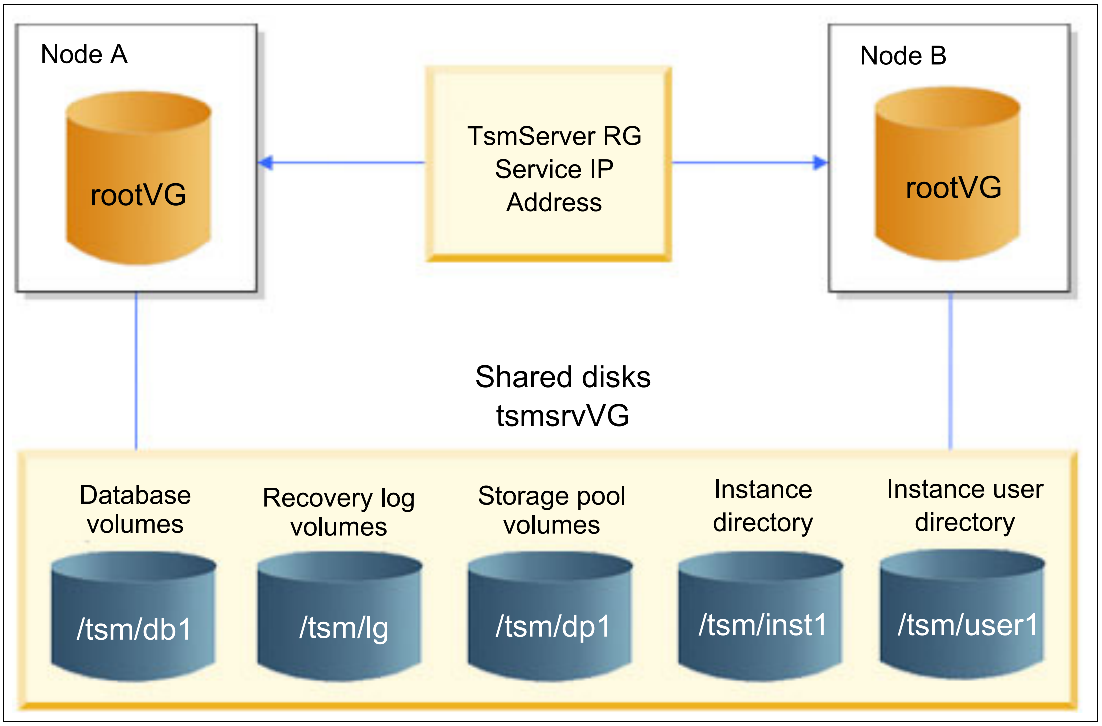



*Figure 5-3 Tivoli Storage Manager server typical 2-node configuration*

### 5.3.2 Prerequisites

We show how to install the prerequisites:

1. Install the required filesets on all nodes as shown in Example 5-1.

*Example 5-1 Tivoli Storage Manager Smart Assist fileset installation*

```
root@smass1 /mnt/nfs/HA711 # installp -acNgXd . cluster.es.assist.tsmserver

. . . (a few lines for this output have been omitted)

Finished processing all filesets.  (Total time:  12 secs).

+-----------------------------------------------------------------------------+
                                Summaries:
+-----------------------------------------------------------------------------+

Installation Summary
--------------------
Name                        Level           Part        Event       Result
-------------------------------------------------------------------------------
cluster.es.assist.common    7.1.1.0         USR         APPLY       SUCCESS
cluster.es.assist.tsmserver 7.1.1.0         USR         APPLY       SUCCESS
cluster.es.assist.tsmserver 7.1.1.0         ROOT        APPLY       SUCCESS
```

2. Check that a shared volume group is available between the nodes. Create a file system by using Cluster Single Point of Control (C-SPOC) tools to hold the installation of the instance in the shared volume group (Example 5-2 on page 233).

*Example 5-2 Shared volume group for Tivoli Storage Manager instance*

```
root@smass1 / # clcmd lspv
-------------------------------
NODE smass1
-------------------------------
hdisk0          00f61ab2f915452b                    rootvg          active
hdisk3          00f74d472c0e5965                    tsmvg

-------------------------------
NODE smass2
-------------------------------
hdisk0          00f74d47f961b168                    rootvg          active
hdisk3          00f74d472c0e5965                    tsmvg
```

3. Remember to install Tivoli Storage Manager server 6.2.3 in both nodes and with the same level as shown in Example 5-3.

*Example 5-3 Tivoli Storage Manager server installation verification*

```
root@smass1 / # clcmd lslpp -L tivoli.tsm.server

-------------------------------
NODE smass1
-------------------------------
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  tivoli.tsm.server          6.2.3.0    C     F    TSM Server
-------------------------------
NODE smass2
-------------------------------
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  tivoli.tsm.server          6.2.3.0    C     F    TSM Server
```

4. Create the same user on both nodes as the owner of the instance and make it the owner of the file system on the shared volume group. The binaries remain local and the instance is created in the shared file system as shown in Example 5-4.

*Example 5-4 User that owns the file system that is used by Tivoli Storage Manager*

```
root@smass1 / # clcmd lsuser -a id home tsminst
-------------------------------
NODE smass2
-------------------------------
tsminst id=208 home=/home/tsminst

-------------------------------
NODE smass1
-------------------------------
tsminst id=208 home=/home/tsminst

root@smass1 / #
root@smass1 / # chown tsminst /tsmdata
```

5. Create the Tivoli Storage Manager server instance as a normal installation in the *first node* with the configuration tool **dsmicfgx** by using a file system in the shared disk.

Complete the information with the required directories, and check **Do not automatically start the server** when asked.

6. Manually start the server and check that everything works as shown in Example 5-5. Stop the Tivoli Storage Manager server, unmount the share file system, and vary off the shared volume group.

*Example 5-5 Verify that the Tivoli Server Manager server is running*

```
root@smass1 / # ps -ef | grep tsm
    root 7077888 4915216   0 12:02:04  pts/1  0:00 grep tsm
  tsminst 7209026 7602308   0 12:01:52      -  0:00 db2vend (PD Vendor Process
- 258)
  tsminst 7602308 7733304  12 12:01:51      -  0:00 db2sysc 0
  tsminst 8323262 8781884   1 12:01:43  pts/0  0:01
/opt/tivoli/tsm/server/bin/dsmserv -u tsminst -i /tsmdata/ -q
  tsminst 9240688 7733304   0 12:01:53      -  0:00 db2acd 0
root@smass1 / # su - tsminst -c 'db2 list applications' | wc -l
      15
```

7. Create the Tivoli Storage Manager server instance as a normal installation in the *second node*. Indicate that it is a "backup node on a cluster configuration" so that the instance already created on the shared volume group is imported and ready to use. Ensure that the Tivoli Storage Manager server starts and works. Stop the Tivoli Storage Manager server as shown in Example 5-5. After verification, unmount the share file system and vary off the shared volume group.

**Restart:** Remember to restart the AIX machine after you install and configure the Tivoli Storage Manager 6.2.3 server.

### 5.3.3 Configuring the Smart Assist for Tivoli Storage Manager

After you plan for and install the Smart Assist for Tivoli Storage Manager, you can configure it.

#### Automatic discovery and configuration for the Smart Assist for Tivoli Storage Manager server

With minimal input, you can use the Smart Assist for Tivoli Storage Manager server to automatically discover and configure nodes. To automatically discover and configure Tivoli Storage Manager server instances, 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)** → **Add an Application to the PowerHA SystemMirror Configuration** and press Enter.
3. From the list of applications, select **TSM server smart assist** and press Enter.
4. Select **Automatic Discovery and Configuration** and press Enter.
5. From the list, select **TSM server** and press Enter.
6. Select the server instance that you want to discover and configure and press Enter.
7. Enter the information for the fields that are listed in Table 5-2 on page 235.

Table 5-2 Tivoli Storage Manager Smart Assist required information

| Fields | Values |
|---|---|
| Application name | Specify the name for the collection of PowerHA SystemMirror components that represent the IBM Tivoli Storage Manager server. The name is limited to 64 characters and cannot contain spaces. |
| Tivoli Storage Manager instance owning node | Specify the highest priority node that can own the resource group that is created by the Smart Assist for Tivoli Storage Manager server. |
| Take over nodes | Specify the lowest priority node in order that can own the resource group that the Tivoli Storage Manager server configuration assistant creates. You must leave a space between the node names. |
| Tivoli Storage Manager server instance name | The selected Tivoli Storage Manager server instance to be made highly available by the Tivoli Storage Manager server configuration assistant is displayed. |
| Tivoli Storage Manager server instance directory | Specify the Tivoli Storage Manager directory where the Tivoli Storage Manager instance logs are stored. This directory is in the shared disk. |
| Shared volume groups for Tivoli Storage Manager server instance | Specify the shared volume groups where the Tivoli Storage Manager servers log files and the configuration information is stored. |
| Service IP label | Specify the service IP label to use by the Tivoli Storage Manager server instance. |
| Netmask (IPv4)/Prefix Length (IPv6) | If you use the IPv4 service interface, enter the network mask for the address. If you use the IPv6 service interface, enter the prefix length for the address. If you do not enter a value, the value of the underlying network is used. |
| Tivoli Storage Manager server Administrator user ID | Specify the Tivoli Storage Manager server instance administrator user ID that is used for monitoring server instances. |
| Tivoli Storage Manager server Administrator password | Specify the Tivoli Storage Manager server password that is used for monitoring server instances. |

In our case, the completed form is shown in Figure 5-4.

```
 Add a TSM server Highly Available Instance Resource Group

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

* Application Name                                      [tsminst_smass]
* TSM instance Owning Node                               smass1
* Take over Node(s)                                     [smass2]
* TSM server Instance Name                               tsminst1
* TSM server Instance Directory                         [/tsmdata/tsminst]
* Shared VGs for TSM server instance                    [tsmvg]
* Service IP Label                                      [smasssvc1]
  Netmask(IPv4)/Prefix Length(IPv6)                     []
* TSM server Administrator user id                      [tsmadmin]
* TSM server Administrator password                     [tsmadmin00]
```

Figure 5-4 Tivoli Storage Manager Smart Assist values that are used on the Smart Assist window

If the configuration is correct, we see the output that is shown in Example 5-6.

*Example 5-6 Tivoli Storage Manager Smart Assist successful installation*

```
Command: OK            stdout: yes           stderr: no

Adding TSM server instance=tsminst
Adding TSM server instance Resource Group TSM_SERV_RG_tsminst to PowerHA
SystemMirror configuration
Generating verify script for TSM server instance tsminst
```

After the configuration, the created resource group contains the information that is shown in Figure 5-5.

```
 Change/Show All Resources and Attributes for a Resource Group

Resource Group Name                                   TSM_SERV_RG_tsminst
  Participating Nodes (Default Node Priority)           smass1 smass2

  Startup Policy                                        Online On Home Node Only
  Fallover Policy                                       Fallover To Next Priority
Node In The List
  Fallback Policy                                       Fallback To Higher
Priority Node In The List
  Fallback Timer Policy (empty is immediate)           [] +

  Service IP Labels/Addresses                          [smasssvc1] +
  Application Controllers                              [TSM_SERV_APP_tsminst] +

  Volume Groups                                        [tsmvg ] +
```

*Figure 5-5 Resource group that is created automatically by Smart Assist*

**For more information:** If you need more information about what happens during configuration, see the log file in /var/hacmp/log.

After the creation of the resources by the Smart Assist, run the Verify and Synchronize process in the cluster to make the new configuration and changes available on all nodes. It is also a good idea to make a snapshot of the new configuration.

## 5.3.4 Smart Assist for Tivoli Storage Manager resources

After the Tivoli Storage Manager is configured through the Smart Assist for Tivoli Storage Manager, PowerHA SystemMirror creates some resources.

### Tivoli Storage Manager server resources

The resources that are created by PowerHA SystemMirror Smart Assist are shown in Table 5-3 on page 237.

*Table 5-3 Tivoli Storage Manager server resources*

| **PowerHA SystemMirror resource** | **Name** |
|---|---|
| Resource group | `TSM_SERV_RG_`*`instanceName`*, where *`instanceName`* is the name of the Tivoli Storage Manager server instance name. |
| Backup archive client application server | `TSM_SERV_APP_`*`instanceName`*, where *`instanceName`* is the name of the Tivoli Storage Manager server instance name |
| Backup archive client custom monitor | `TSM_SERV_APP_MON_`*`instanceName`*, where *`instanceName`* is the name of the Tivoli Storage Manager server instance name. |
| Backup archive client start script | `cl_tsmserverstart`. This file is in the directory `/usr/es/sbin/cluster/sa/tsmserver/sbin`. |
| Backup archive client stop script | `cl_tsmserverstop`. This file is in the directory `/usr/es/sbin/cluster/sa/tsmserver/sbin`. |
| Backup archive client monitor script | `cl_tsmservermonitor`. This file is in the directory `/usr/es/sbin/cluster/sa/tsmserver/sbin`. |

Table 5-4 describes the default settings that are associated with the custom monitor (`TSM_SERV_APP_MON_`*`instanceName`*) that is displayed in Table 5-3.

*Table 5-4 Default settings for Tivoli Storage Manager application monitoring*

| **Field** | **Value** |
|---|---|
| Name | `TSM_SERV_APP_MON_`*`instanceName`* |
| Application Servers to Monitor | `TSM_SERV_APP_`*`instanceName`* |
| Monitor method | `/usr/es/sbin/cluster/sa/tsmserver/sbin/ cl_tsmservermonitor –i` <*Instance Name*> |
| Mode | Long-running monitoring |
| Interval | 180 sec |
| Hung Monitor Signal | 9 |
| Stabilization interval | 180 Sec |
| Restart count | 3 |
| Restart interval | 1200 Sec |
| Action on Application Failure | Fallover |
| Cleanup method | `/usr/es/sbin/cluster/sa/tsmserver/sbin/ cl_tsmserverstop -i` <*instance Name*> |
| Restart Method | `/usr/es/sbin/cluster/sa/tsmserver/sbin/ cl_tsmserverstart -i` <*instance Name*> `-d` <*instance Directory*> |

**Log data:** The Smart Assist for Tivoli Storage Manager writes log data to the `/var/hacmp/log/tsm_server.log` file.

## 5.4 Smart Assist for IBM MQ Series

IBM WebSphere MQ is middleware for messaging and queuing network communication. You can use WebSphere MQ to increase the speed of distributed applications by simplifying application development and testing.

WebSphere MQ runs in various platforms. WebSphere MQ applications use a consistent application programming interface (API) across all platforms. With WebSphere MQ, your applications can communicate with each other across a network of different components such as processors, subsystems, operating systems, and communication protocols.

### 5.4.1 Planning for WebSphere MQ

Install the operating system, PowerHA SystemMirror, and WebSphere MQ by using normal procedures on all systems in the cluster.

Install WebSphere MQs into internal disks (non-shared) on each of the nodes and do not share a single installation on shared disks. It is important that you run identical versions of software on all cluster nodes, except during a rolling upgrade.

When you install WebSphere MQ in a cluster, the *mqm username* and the *mqm groupname* must be created and have the same numeric value on all cluster nodes.

Figure 5-6 displays a typical 2-node configuration for a WebSphere MQ installation.

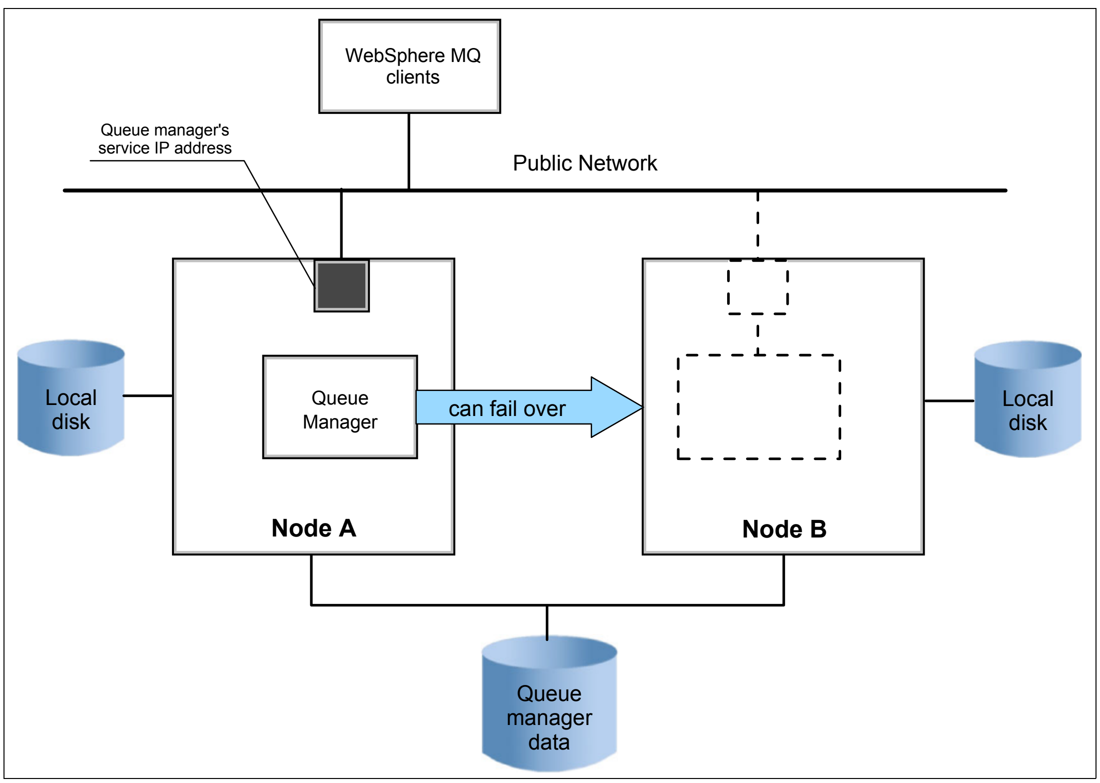



*Figure 5-6 Typical 2-node WebSphere MQ cluster configuration*

The smallest unit of failover for WebSphere MQ is a queue manager because you cannot move part of a queue manager without moving all of it. Place each queue manager in a separate resource group with the resources on which it depends. The resource group must contain the following resources:

- The shared volume group that is used by the queue manager
- The IP address that is used to connect to the queue manager (the service address)
- The object that represents the queue manager

**Multiple queue managers:** You can put multiple queue managers into the same resource groups, but they might fallover to another node together, even if the problem that causes the fallover is confined to one queue manager.

A queue manager that is used in a PowerHA SystemMirror cluster must have its recovery logs and data on shared disks so that they can be accessed by a surviving node in a node failure. A node that runs a queue manager must also maintain a number of files on internal (non-shared) disks. These files include files that relate to all queue managers on the node such as the `/var/mqm/mqs.ini` file, and queue manager-specific files that are used to generate internal control information.

## 5.4.2 Software requirements

Your environment must run the following software to successfully implement the Smart Assist for WebSphere MQ:

- WebSphere MQ Version 7.0.1, or later
- PowerHA SystemMirror Version 7.1.1, or later

## 5.4.3 Installing WebSphere MQ

Before you install WebSphere MQ, verify that the `mqm` *`username`* and the `mqm` *`groupname`* are created and have the same numeric value on all of the cluster nodes. Set the primary group of user `mqm` to `mqm`.

To make the WebSphere MQ queue manager highly available on a node, install the following filesets from the WebSphere MQ media installer:

- `mqm.base.runtime`
- `mqm.base.sdk`
- `mqm.client.rte`
- `mqm.java.rte`
- `mqm.jre.rte`
- `mqm.keyman.rte`
- `mqm.server.rte`
- `mqm.txclient.rte`
- `mqm.base.runtime`

Install the required Smart Assist filesets as shown in Example 5-7.

*Example 5-7 IBM MQSeries® Smart Assist installation*

```
root@smass1 /mnt/nfs/HA711 # installp -acNgXd . cluster.es.assist.wmq
+-----------------------------------------------------------------------------+
. . . (more)
Installation Summary
--------------------
```

```
Name                       Level          Part        Event       Result
--------------------------------------------------------------------------------
cluster.es.assist.wmq      7.1.1.0        USR         APPLY       SUCCESS
cluster.es.assist.wmq      7.1.1.0        ROOT        APPLY       SUCCESS
root@smass1 /mnt/nfs/HA711 #
```

**Important:** Ensure that you add `/usr/mqm/bin` to the global PATH environment.

## 5.4.4 Configuring WebSphere MQ

Before you can configure WebSphere MQ, verify that PowerHA SystemMirror is installed and configured on all nodes in the cluster.

### Configuring a shared disk for WebSphere MQ

A WebSphere MQ queue manager in a PowerHA SystemMirror cluster requires that data files and log files are in a common named remote file system on a shared disk (Example 5-8).

To configure a shared disk to work with WebSphere MQ, complete the following steps:

1. Create a fast disk takeover-enabled volume group by using the C-SPOC utilities. This volume group is used for the data and log files of the queue manager. This volume group is managed by the PowerHA SystemMirror cluster in the same resource group as the queue manager.

*Example 5-8 Share volume group for MQ installation*

```
root@smass1 / # clcmd lspv
-------------------------------
NODE smass2
-------------------------------
hdisk0         00f61ab2f915452b                   rootvg          active
hdisk4         00f74d472c0e5a93                   mqvg

-------------------------------
NODE smass1
-------------------------------
hdisk0         00f74d47f961b168                   rootvg          active
hdisk4         00f74d472c0e5a93                   mqvg
```

2. Create a file system for data and another file system for logs by using the volume group that is created in step 1. You can use the C-SPOC utilities so that both nodes have the correct file system configuration.

### Creating a queue manager for a PowerHA SystemMirror cluster

Before you can use a queue manager in a PowerHA SystemMirror cluster, you must create a queue manager on one of the nodes in the cluster.

To create a queue manager for a PowerHA SystemMirror cluster, complete the following steps:

1. Select one of the nodes in the cluster on which you want to create the queue manager and log in as a root user.
2. Vary on the shared volume group and mount the data file system and log file system that you created in the shared disks, for example, `/MQHA`.

3. Change the owner permissions of the data file system and log file system by running the following commands:

```
chown -R mqm:mqm file_system_name
chmod -R 2775 file_system_name
```

The *file_system_name* is the name of the file system.

4. As `mqm user`, create the queue manager as shown in Example 5-9.

*Example 5-9 How to create queue manager*

```
/usr/mqm/bin/crtmqm -md /MQHA/qmgrjcrm/data -ld /MQHA/qmgrjcrm/log qmgrjcrm
```

where qmgrjcrm is the name of the queue manager. And check the output:

```
WebSphere MQ queue manager created.
Directory '/MQHA/qmgrjcrm/data/qmgrjcrm' created.
The queue manager is associated with installation 'Installation1'.
Creating or replacing default objects for queue manager 'qmgrjcrm'.
Default objects statistics : 71 created. 0 replaced. 0 failed.
Completing setup.
Setup completed.
```

5. Display the **addmqinf** command output as the `mqm user` by running the following command:

```
dspmqinf -o command qmgrjcrm
```

`qmgrjcrm` is the name of the queue manager.

6. Verify that you populated the `/var/mqm/mqs.ini` file with the correct information as shown in Example 5-10.

*Example 5-10 Information in mqs.ini file*

```
$ cat /var/mqm/mqs.ini
#*******************************************************************#
. . . (a few lines for this output have been omitted)
#*******************************************************************#
   DefaultPrefix=/var/mqm

LogDefaults:
   LogDefaultPath=/var/mqm/log

QueueManager:
   Name=qmgrjcrm
   Prefix=/var/mqm
   Directory=qmgrjcrm
   DataPath=/MQHA/qmgrjcrm/data/qmgrjcrm
   InstallationName=Installation1
```

If you did not, run the **addmqinf** command that is displayed as a result of running the **dspmqinf** command in step 5. The **addmqinf** command that you run is similar to the following example:

```
addmqinf -sQueueManager -vName=qmgrjcrm -vDirectory=qmgrjcrm
-vPrefix=/var/mqm -vDataPath=/MQHA/qmgrjcrm/data/qmgrjcrm
```

7. Unmount the queue manager file systems and vary off the shared volume group.

**Important:** If you plan to create more than one qmanager, it is a good idea to create a separate directory in the shared file system for each one.

### Adding a queue manager configuration to other nodes

After you create a queue manager, you can add the configuration from the queue manager that you created to other nodes in the PowerHA SystemMirror cluster.

To add the configuration information for the queue manager to each of other nodes in the PowerHA SystemMirror cluster, complete the following steps on each of the other nodes:

1. Vary on the shared volume group and mount the queue manager data file system and log file system.
2. Add the queue manager configuration information, as `mqm user`, to the node by editing the `/var/mqs.ini` file, or by running the **`addmqinf`** command that is displayed when you run the **`dspmqinf -o`** command.
3. As the `mqm user`, start and stop the queue manager to verify that the configuration works correctly; see Example 5-11.

*Example 5-11 How to test that the queue manager works correctly*

```
$ /usr/mqm/bin/dspmq
QMNAME(qmgrjcrm)                                           STATUS(Ended
normally)
$ /usr/mqm/bin/strmqm qmgrjcrm
WebSphere MQ queue manager 'qmgrjcrm' starting.
The queue manager is associated with installation 'Installation1'.
5 log records accessed on queue manager 'qmgrjcrm' during the log replay phase.
Log replay for queue manager 'qmgrjcrm' complete.
Transaction manager state recovered for queue manager 'qmgrjcrm'.
WebSphere MQ queue manager 'qmgrjcrm' started using V7.1.0.0.
$ /usr/mqm/bin/dspmq
QMNAME(qmgrjcrm)                                           STATUS(Running)
$ /usr/mqm/bin/endmqm qmgrjcrm
Quiesce request accepted. The queue manager will stop when all outstanding work
is complete.
$ /usr/mqm/bin/dspmq
QMNAME(qmgrjcrm)                                           STATUS(Quiescing)
$ /usr/mqm/bin/dspmq
QMNAME(qmgrjcrm)                                           STATUS(Ended
normally)
```

4. Unmount the queue manager data file system and log file system, and vary off the shared volume group.

## 5.4.5 Configuring Smart Assist for WebSphere MQ

After you install and configure WebSphere MQ for your environment, you can configure Smart Assist for WebSphere MQ.

### Automatic discovery and configuration for Smart Assist for WebSphere MQ

By using SMIT, you can set up the Smart Assist for WebSphere MQ to automatically discover a WebSphere MQ instance that runs in the cluster with its resources, such as shared volume groups and service IP addresses.

To set up automatic discovery and configuration, 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)** → **Add an Application to the PowerHA SystemMirror Configuration**, and press Enter.
3. From the list of applications, select **WebSphere MQ Smart Assist**, and press Enter.
4. Select **Automatic Discovery and Configuration**, and press Enter.
5. Select **WebSphere MQ Server**, and press Enter.
6. Select the queue manager that you want to automatically discover and configure, and press Enter.
7. Enter the information for the fields that are shown in Table 5-5.

*Table 5-5 Smart Assist for MQSeries required information*

| Fields | Values |
|---|---|
| Application Name | Enter the name for the collection of PowerHA SystemMirror resources that represent the WebSphere MQ queue manager component. |
| WebSphere MQ Manager Name | Displays the name of the queue manager that you selected in step 5. This field cannot be modified. |
| Primary Node | Displays the name of the primary node where the queue manager is active. |
| Takeover Node(s) | Enter (or select from the pick list) the name of one or more cluster nodes to which the application might fallover. |
| Service Interface | Enter or select from the list the service IP label or service IP address that the clients use to communicate with the queue manager. |
| Netmask(IPv4)/Prefix Length(IPv6) | For the IPv4 service interface, enter the network mask for the address. For the IPv6 service interface, enter the prefix length for the address. |

Verify that all fields are correct and press Enter. In our case, the completed form is shown in Figure 5-7 on page 244.

```
 Add a WebSphere MQ Manager Highly Available Resource Group

Application Name                                    [qmgrjcrm_app]
  WebSphere MQ Manager Name                            qmgrjcrm
* Primary Node                                        [smass1]
* Takeover Node(s)                                    [smass2]
* Service Interface                                   [smasssvc1]
  Netmask(IPv4)/Prefix Length(IPv6)                   []
```

*Figure 5-7 MQ Series Smart Assist values used*

After the configuration is created, the resource group contains the information that is shown in Figure 5-8.

```
 Change/Show All Resources and Attributes for a Resource Group

Resource Group Name                                  qmgrjcrm_RG
Participating Nodes (Default Node Priority)          smass1 smass2

Startup Policy                                       Online On Home Node Only
Fallover Policy                                      Fallover To Next Priority
Node In The List
Fallback Policy                                      Fallback To Higher Priority
Node In The List
Fallback Timer Policy (empty is immediate)          []

Service IP Labels/Addresses                         [smasssvc1]
Application Controllers                             [qmgrjcrm_app]

Volume Groups                                       [mqvg ]
```

*Figure 5-8 Resource group that is created automatically by Smart Assist for MQSeries*

8. Verify and synchronize your cluster to activate the configuration changes on all nodes.

## 5.4.6 Smart Assist for WebSphere MQSeries resources

The Smart Assist for WebSphere MQSeries uses a standard naming convention so that it is easy to identify various PowerHA SystemMirror resources that are created for Smart Assist for WebSphere MQSeries.

Table 5-6 on page 245 lists the PowerHA SystemMirror resources that are created for Smart Assist for WebSphere MQSeries.

*Table 5-6 Resources that are created for Smart Assist for WebSphere MQSeries*

| PowerHA SystemMirror resources | Name |
|---|---|
| Resource group | *Queue_mgr*_RG, where *Queue_mgr* is the queue manager name. |
| Application controller | *Application_Name*, where *Application_Name* is the application controller name. |
| Custom monitor | *Application_name*_mon, where *Application_name* is the application name. The relevant script is the `cl_wmq_monitor` file that is in the `/usr/es/sbin/cluster/sa/wmq/sbin` directory. |
| Start script | The relevant script is the `cl_wmq_start` file that is in the `/usr/es/sbin/cluster/sa/wmq/sbin` directory. |
| Stop script | The relevant script is the `cl_wmq_stop` file that is in the `/usr/es/sbin/cluster/sa/wmq/sbin` directory. |

## Custom application monitors

Smart Assist for WebSphere MQSeries configures a custom monitor for the WebSphere MQ applications that you configure in your PowerHA SystemMirror environment. Custom application monitors check the health of an application at user-specified polling intervals.

Table 5-7 displays the default settings for the Smart Assist for WebSphere MQSeries custom monitor.

*Table 5-7 Smart Assist for WebSphere MQSeries custom monitor settings*

| Field | Value |
|---|---|
| Name | *Application_name*_mon, where *Application_name* is the application name |
| Application controllers to monitor | *Application_controller*, where *Application_controller* is the name of the application controller |
| Monitor method | `/usr/es/sbin/cluster/sa/wmq/sbin/ cl_wmq_monitor -u` <*mqm_user*> `-m` <*MQM_NAME*>, where *mqm_user* is the user for the queue manager and *MQM_NAME* is the queue manager name |
| Mode | Long running monitoring |
| Interval | 120 sec |
| Hung monitor signal | 9 |
| Stabilization interval | 180 sec |
| Restart interval | 900 sec |
| Restart count | 3 |
| Action on application failure | Fallover |
| Cleanup method | `/usr/es/sbin/cluster/sa/wmq/sbin/cl_wmq_stop -u` <*mqm_user*> `-m` <*MQM_NAME*>, where *mqm_user* is the user for the queue manager and *MQM_NAME* is the queue manager name |
| Restart method | `/usr/es/sbin/cluster/sa/wmq/sbin/cl_wmq_start -u`<*mqm_user*> `-m` <*MQM_NAME*>, where *mqm_user* is the user for the queue manager and *MQM_NAME* is the queue manager name |

**Log:** The Smart Assist for MQSeries writes log data to the `/var/hacmp/log/wmqsa.log` file.

## 5.5 Smart Assist for IBM Tivoli Directory Server

IBM Tivoli Directory Server is a powerful, security-rich, and standards-compliant enterprise directory for corporate intranets and the Internet. Tivoli Directory Server is the identity data foundation for the rapid development and deployment of web applications and security by including strong management, replication, and security features. Tivoli Directory Server is a powerful Lightweight Directory Access Protocol (LDAP) infrastructure. It provides a foundation for deploying comprehensive identity management applications and advanced software architectures.

With the Smart Assist for Tivoli Directory Server, you can automatically configure PowerHA SystemMirror where Tivoli Directory Server is already installed.

### Smart Assist for Tivoli Directory Server concepts

Smart Assist for Tivoli Directory Server supports stand-alone directory configurations, distributed directory configurations, and peer-to-peer directory configurations.

#### *Stand-alone directory configuration*

In a stand-alone directory configuration, there are multiple Tivoli Directory Server instances that run on a single node (Node 1). When a failure occurs in Node 1, another node (Node 2) obtains the resources of Node 1 and takes over running the Tivoli Directory Server instance. Node 2 uses the same disk set and the same IP address for the Tivoli Directory Server instance that Node 1 used. When Node 1 fails, a database instance that is created on a shared volume group is started and used by Node 2. Also, when Node 1 fails, its service IP label is taken over by Node 2.

#### *Distributed directory configuration*

In a distributed directory configuration, there are multiple Tivoli Directory Server instances that run on each node in a PowerHA SystemMirror cluster. Each Tivoli Directory Server instance has its peer servers and replica servers that run on a node in the cluster. In a distributed directory configuration, client requests are routed to the correct Tivoli Directory Server by a proxy server that runs in a node in the cluster. The proxy server is not controlled by PowerHA SystemMirror. When a failure occurs on any of the Tivoli Directory Server instances, the instance is restarted by PowerHA SystemMirror. If a failure occurs on a node in the cluster, PowerHA SystemMirror is not supported because fallover and fallback are handled by the Tivoli Directory Server. When a failure occurs for a node and data is replicated into a different Tivoli Directory Server, there is no need for a shared volume group or service IP label in the cluster.

#### *Peer-to-peer directory configuration*

In a peer-to-peer directory configuration, the proxy server is not available, and there is no data partition and data distribution between the Tivoli Directory Server instances. However, the Tivoli Directory Server instances have access to peer server instances and replica server instances on a node in the cluster. When a failure occurs on any of the Tivoli Directory Server instances, the instance is restarted by PowerHA SystemMirror. If failure occurs on a node in the cluster, PowerHA SystemMirror is not supported because fallover and fallback are handled by the Tivoli Directory Server.

## 5.5.1 Planning Smart Assist for Tivoli Directory Server

You must install all Tivoli Directory Server software on all nodes in the cluster before you install Smart Assist for Tivoli Directory Server. You must install the Tivoli Directory Server files in the `/opt/IBM/ldap/V6.3` default directory because the Smart Assist looks for the Tivoli Directory Server installation in the default directory during the discovery process.

If you use a stand-alone directory configuration, you must create every database instance on a shared volume group. If you use a peer-to-peer directory configuration or a distributed directory configuration, you do not need to create any database instances on a shared volume group, you can create the database instance on a local node.

## 5.5.2 Software requirements

In our example, we configure a stand-alone directory and then use the Smart Assist to create the needed resource. You must have the following software installed before you install Smart Assist for Tivoli Directory Server:

- Tivoli Directory Server Version 6.3, or later
- PowerHA SystemMirror Version 7.1.1, or later
- Embedded WebSphere Application Server Version 7.0.0.7, or later

You must follow the next steps before you configure Smart Assist for Tivoli Directory Server:

1. Check that a shared volume group is available between the nodes. Create a file system by using C-SPOC tools in the volume group to hold the installation of the LDAP instance as shown in Example 5-12.

*Example 5-12 Shared volume group for Tivoli Storage Manager instance*

```
root@smass1 / # clcmd lspv
-------------------------------
NODE smass1
-------------------------------
hdisk0 00f61ab2f915452b rootvg          active
hdisk5 00f74d472c0e5bc0 ldapvg

-------------------------------
NODE smass2
-------------------------------
hdisk0 00f74d47f961b168 rootvg          active
hdisk5 00f74d472c0e5bc0 ldapvg
```

2. Install the Tivoli Directory Server server filesets on a share file system in the *first node*. Remember that the same user is created on both nodes, and the home directory for the user is on the share file system on the share volume group.
3. Install embedded WebSphere Application Server for Tivoli Directory Server. It is used to access the server and perform administrative operations later. After you download and unpack embedded WebSphere Application Server, run the installer as shown in Example 5-13.

*Example 5-13 Embedded WebSphere Application Server for Tivoli Directory Server installation*

```
root@smass1 /mnt/nfs/ITDS_eWAS/tdsV6.3/appsrv # ./install.sh -installRoot
/opt/IBM/ldap/V6.3/appsrv

+--------------------------------------+
```

```
+       EWAS Version 7.0 Install       +
+--------------------------------------+

 Validating target directory ...
 Copying files ...
 Setting permissions ...
 Installation complete.
root@smass1 /mnt/nfs/ITDS_eWAS/tdsV6.3/appsrv #
```

copy the application tools to the required directory with the command

```
root@smass1 /opt/IBM/ldap/V6.3/idstools # cp IDSWebApp.war
/opt/IBM/ldap/V6.3/appsrv/installableApps/.
```

deploy el tool issuing the command

```
/opt/IBM/ldap/V6.3/idstools/deploy_IDSWebApp
```

4. Run the LDAP configuration tool and fill out the required fields. Create the instance by using the user that was previously created as shown in Figure 5-9.

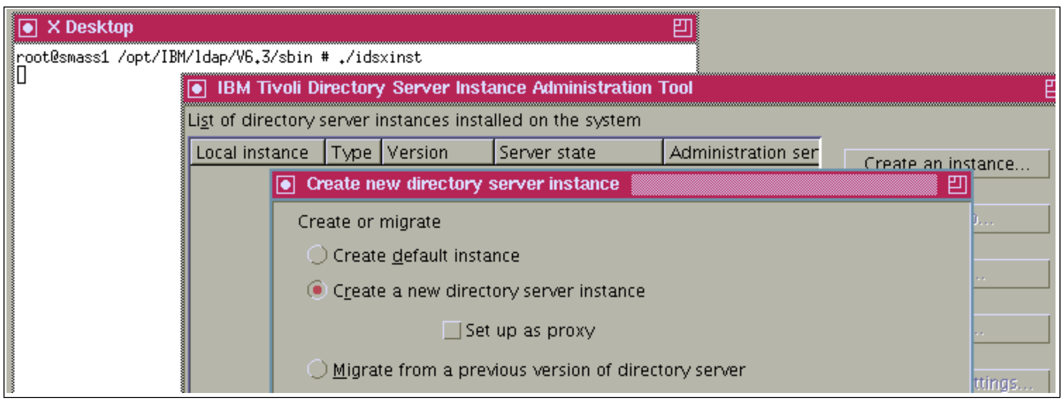



*Figure 5-9 LDAP instance configuration*

5. Start the server with the admin tool and verify that it works and is functional (Figure 5-10 on page 249).

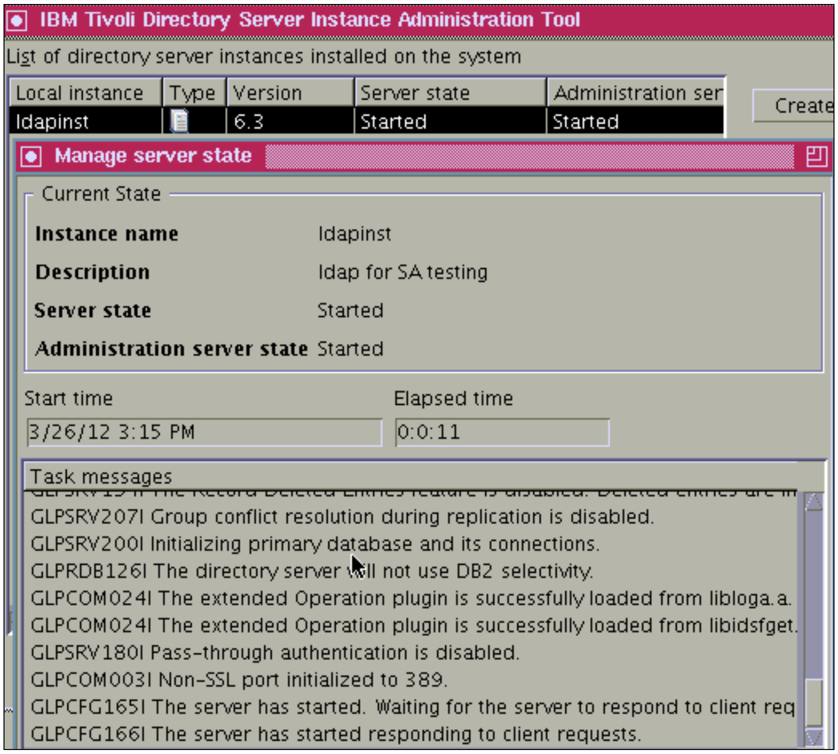



*Figure 5-10 LDAP instance starting process*

6. Connect to AIX as the owner user of the instance and issue a search to validate that the data is accessible as shown in Example 5-14.

*Example 5-14 Validate that the ldap is accessible*

```
$ idsldapsearch -D cn=root -w admin -s base objectclass=* | grep config
ibm-configurationnamingcontext=CN=CONFIGURATION
ibm-slapdisconfigurationmode=FALSE
```

7. Ensure that the service is down, unmount the shared file system, and vary off the shared volume group in the first node.
8. In the second node, use **varyonvg** to vary on the shared volume group and mount the shared file system. Edit the `/etc/services` file and include the information from the file in the first node. In our case, Example 5-15 shows the necessary lines that are based on *username*.

*Example 5-15 Lines added after you create the LDAP instance*

```
root@smass1 /tdsdata # cat /etc/services | grep ldapinst
DB2_ldapinst    60004/tcp
DB2_ldapinst_1  60005/tcp
DB2_ldapinst_2  60006/tcp
DB2_ldapinst_END        60007/tcp
ldapinstsvcids  3708/tcp
ldapinstsvcidsi 3737/tcp
root@smass1 /tdsdata #
```

9. Add an entry on the `inittab` file with the previously created entry in the first node as shown in Example 5-16 on page 250.

*Example 5-16 inittab entry for LDAP instance*

```
root@smass1 / # lsitab -a | grep ids
ids0:2345:once:/opt/IBM/ldap/V6.3/sbin/ibmdiradm -I ldapinst > /dev/null 2>&1
#Autostart IBM LDAP Admin Daemon Instance
```

10. Copy the `ldif` file with definitions from the first node to the second node as shown in Example 5-17.

*Example 5-17 ldif file with instance definitions*

```
root@smass1 /opt/IBM/ldap/idsinstinfo # cat ids*
charset: ISO-8859-1
version: 1

dn: CN=IDS INSTANCES
cn: IDS INSTANCES
objectClass: TOP
objectClass: CONTAINER
dn: cn=ldapinst, CN=IDS INSTANCES

cn: ldapinst
ids-instanceDesc: ldap for Tivoli Directory Server Smart Assist
ids-instanceLocation: /tdsdata/ldapinst
ids-instanceVersion: 6.3
objectClass: TOP
objectClass: ids-instance
```

11. Start the server as shown in step 5 on page 248. Issue a search to validate that the server works in the second node as shown in Example 5-18.

*Example 5-18 Validate LDAP access on the second node*

```
$ idsldapsearch -D cn=root -w admin -s base objectclass=* | grep config
ibm-configurationnamingcontext=CN=CONFIGURATION
ibm-slapdisconfigurationmode=FALSE
```

12. You can also start the webAdmin tool and check that everything runs and is functional as shown in Figure 5-11.

```
root@smass1 /opt/IBM/ldap/V6.3/idstools/bin # ./startWebadminApp
/opt/IBM/ldap/V6.3/appsrv/profiles/TDSWebAdminProfile/bin/startServer.sh
server1
ADMU0116I: Tool information is being logged in file

/opt/IBM/ldap/V6.3/appsrv/profiles/TDSWebAdminProfile/logs/server1/startServ
er.log
ADMU0128I: Starting tool with the TDSWebAdminProfile profile
ADMU3100I: Reading configuration for server: server1
ADMU3200I: Server launched. Waiting for initialization status.
ADMU3000I: Server server1 open for e-business; process id is 13697238
```

*Figure 5-11 WebAdmin tool process startup*

13. Connect by using a browser and validate that Tivoli Directory Server is up as shown in Figure 5-12 on page 251.

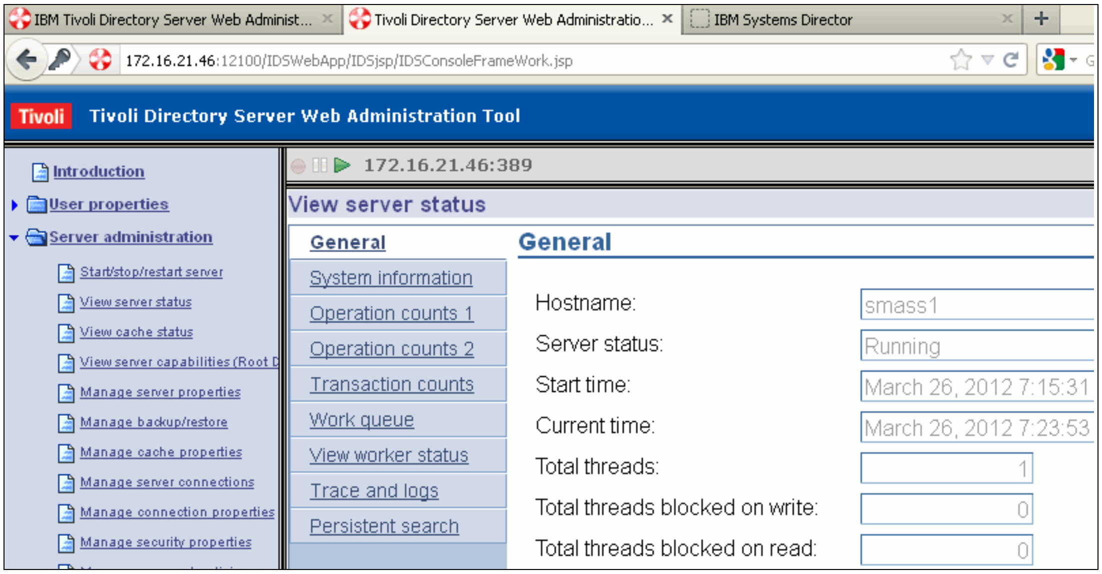



*Figure 5-12 Tivoli Directory Server status by using the webadmin tool*

14. Remember to stop all processes in the second node, unmount the file system, and vary off the shared volume group.

## 5.5.3 Configuring the Smart Assist for Tivoli Directory Server

After you plan for and install the Smart Assist for Tivoli Directory Server, you can configure the Smart Assist for Tivoli Directory Server. The process to configure the Smart Assist for Tivoli Directory Server differs for distributed directory and stand-alone Tivoli Directory Server configurations.

### Automatic discovery and configuration of Smart Assist for Tivoli Directory Server

You can use automatic discovery to configure Smart Assist for Tivoli Directory Server with minimal input. To automatically discover and configure Smart Assist for Tivoli Directory Server, 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)** → **Add an Application to the PowerHA SystemMirror Configuration** and press Enter.
3. From the list of applications, select **Tivoli Directory Server Smart Assist** and press Enter.
4. Select **Automatic Discovery and Configuration** and press Enter.
5. Select **Tivoli Directory Server** and press Enter.
6. Enter the information for the fields that are shown in Table 5-8 on page 252.

*Table 5-8 Smart Assist for Tivoli Directory Server required information*

| Field | Value |
|---|---|
| Application name | Specify an application name. A unique application name is required for every Tivoli Directory Server instance because a different resource group is created for every Tivoli Directory Server instance. |
| Primary node | Specify the primary node for the Smart Assist. This field is available for stand-alone configurations only. |
| Takeover node(s) | Enter the name of one or more cluster nodes to which the application might fallover. This field is available for stand-alone configurations only. |
| Service interface | Specify the service IP label that is used with the Tivoli Directory Server client. This field is available for stand-alone configurations only. |
| Tivoli Directory Server password | Specify the password for the Tivoli Directory Server. |
| Tivoli Directory Server port | Specify the port that is configured with Tivoli Directory Server. |
| DB2 instance name | Specify the DB2 instance name that is used with the Tivoli Directory Server. This field is available for stand-alone configurations only. |
| Number of Tivoli Directory Server instances | Specify the number of instances on the node. This field is available for distributed directory configurations only. |
| Tivoli Directory Server instance name | Specify the name of the Tivoli Directory Server instance. This field is available for stand-alone configurations only. |

In our case, the completed form is shown in Figure 5-13.

```
 Add a Tivoli Directory Server to the Cluster

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

                                                        [Entry Fields]
* Application Name                                   [TDS_smass1]
  Primary Node                                        smass1
* Takeover Node(s)                                   [smass2]
* Service Interface                                  [smasssvc1]
* Tivoli Directory Server Password                   [secret]
* Tivoli Directory Server Port                       [389]
* DB2 Instance Name                                  [ldapinst]
* Tivoli Directory Server Instance Name              [ldapinst]
```

*Figure 5-13 Information that is needed by Smart Assist*

If everything is configured correctly, the automatic discovery results are similar to Figure 5-14 on page 253.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Mon Mar 26 13:54:27 EDT 2012 - Creating Service label smasssvc1.
----------------------------------------------------------------------
Configuring service IP label: smasssvc1
Service IP label configuration complete.
----------------------------------------------------------------------
Mon Mar 26 13:54:28 EDT 2012 - Application Server tdsas_smass1_ldapinst
created.
Mon Mar 26 13:54:28 EDT 2012 - Resource Group tdsrg_smass1_ldapinst created.
Mon Mar 26 13:54:28 EDT 2012 - Application Monitor for application server
tdsas_smass1_ldapinst created.
Auto Discover/Import of Volume Groups was set to true.
Gathering cluster information, which may take a few minutes.
Mon Mar 26 13:54:30 EDT 2012 - Resources added to resource group
tdsrg_smass1_ldapinst.
Mon Mar 26 13:54:30 EDT 2012 - TDS import create complete.
```

*Figure 5-14 Automatic discovery that uses Smart Assist for Tivoli Directory Server*

7. Verify and synchronize your cluster to make the configuration changes available in all nodes.

### 5.5.4 Smart Assist for Tivoli Directory Server resources

After Smart Assist for Tivoli Directory Server runs, the standard naming convention is used to make it easy to identify various PowerHA SystemMirror resources that are created for Smart Assist for Tivoli Directory Server.

Figure 5-15 on page 254 shows PowerHA SystemMirror resources that are created for Smart Assist for Tivoli Directory Server.

```
 Change/Show All Resources and Attributes for a Custom Resource Group

Resource Group Name                                   tdsrg_smass1_ldapinst
Participating Nodes (Default Node Priority)           smass1 smass2

Startup Policy                                        Online On Home Node Only
Fallover Policy                                       Fallover To Next Priority
Node In The List
Fallback Policy                                       Never Fallback

Service IP Labels/Addresses                          [smasssvc1] +
Application Controllers                              [tdsas_smass1_ldapinst] +

Volume Groups                                        [tdsvg] +
Use forced varyon of volume groups, if necessary      false +
Automatically Import Volume Groups                    false +
```

*Figure 5-15 Resource group that is created by Smart Assist*

To control and monitor the Tivoli Directory Server system, two scripts are used and defined as shown in Figure 5-16.

```
Change/Show Application Controller Scripts

Application Controller Name                           tdsas_smass1_ldapinst
New Name                                             [tdsas_smass1_ldapinst]
Start Script [/usr/es/sbin/cluster/sa/tds/sbin/startTDS>
Stop Script [/usr/es/sbin/cluster/sa/tds/sbin/stopTDS>
Application Monitor Name(s)                           tdsas_smass1_ldapinst
Application startup mode                             [background] +
```

*Figure 5-16 Controller scripts for Tivoli Directory Server*

More scripts and logs are in `/usr/es/sbin/cluster/sa/tds`, but do not modify any of this information.

**Log:** Smart Assist for Tivoli Directory Server writes log data to the `/var/hacmp/log/tdssa.log` file.

# 6

# Migration scenarios

This chapter describes the following topics about migration to IBM PowerHA SystemMirror 7.1.1:

- Considerations before you migrate to PowerHA 7.1.1
- Migration from pre-7.1 versions:
  - Requirements
  - Stages of migration and the clmigcheck command
  - Offline migration from PowerHA 5.5
  - Snapshot migration from PowerHA 5.5
  - Rolling migration from PowerHA 6.1
- Migration from PowerHA 7.1.0:
  - Offline migration from PowerHA 7.1.0 version
  - Snapshot migration from 7.1.0 version

## 6.1 Considerations before you migrate to PowerHA 7.1.1

IBM introduced a major release of PowerHA SystemMirror in 2010 as version 7.1 of the product. A minor release update, also called a Technology Level (TL), is released as PowerHA SystemMirror 7.1 TL01 or 7.1.1. Before you migrate your existing cluster to PowerHA SystemMirror 7.1.1, you must be aware of the following considerations:

- The minimum required software versions:
  - AIX 6.1 TL7 Service Pack 3 (SP3) with Reliable Scalable Cluster Technology (RSCT) 3.1.2.0 when you use AIX Version 6.1
  - AIX 7.1 TL1 SP3 with RSCT 3.1.2.0 when you use AIX Version 7.1
  - Virtual I/O Server (VIOS) 2.2.0.1 - FP24 SP01 when the cluster nodes are VIOS client logical partitions (LPARs)

  > **SP3:** SP3 of AIX 7.1 TL1 and AIX 6.1 TL7 is the only prerequisite for SP1 of PowerHA 7.1.1. You can install the PowerHA 7.1.1 base at SP2 of AIX 6.1 TL7 or AIX 7.1 TL1.

  We strongly advise you to use the latest versions and service packs (SPs) that are available for AIX, VIOS, and PowerHA.

- A shared disk must be dedicated as the cluster repository

  Choose an appropriate size. Ensure that the storage subsystem that hosts the repository disk is supported. Also, ensure that the adapters and the multipath driver that are used for the connection to the repository disk are supported. See 3.6.1, “Repository disk” on page 91 for more details.

  If you choose a previously used disk, you must ensure that it is clean.

  > **Previously used disk:** The disk must be a clean logical unit number (LUN). It must not contain previous Cluster Aware AIX (CAA) or Logical Volume Manager (LVM) structures that are known to any of our cluster nodes or to a third party.

- Multicast must be enabled in your cluster network infrastructure

  Confirm with the network administrator that the multicast IP address that you intend to use is available or ask for an unused one. Ensure that the multicast traffic generated by any of the cluster nodes is correctly forwarded by the network infrastructure toward the other cluster nodes. See 2.5, “CAA multicast” on page 68 for more details.

- Non-IP heartbeating is accomplished differently in PowerHA 7.1 compared with previous versions.

  A non-IP heartbeat channel is automatically configured through the repository disk. More SAN-based heartbeating can be configured through the physical Fibre Channel (FC) or serial-attached SCSI (SAS) adapters. The SAS adapters do not require a special setup. In direct-attached FC adapters, more setup is required. For more details, see 3.7.12, “Configuring the physical FC adapters for SAN-based heartbeat” on page 101. A similar non-IP SAN-based communication channel can be configured for PowerHA 7.1 LPARs/nodes in N-Port ID Virtualization (NPIV) or virtual Small Computer System Interface (vSCSI) virtualized environments. The VIOS release must be at least 2.2.0.11-FP24SP01. Setup instructions are available in DOC APAR IV03643:

  `http://www-01.ibm.com/support/docview.wss?uid=isg1IV03643`

  See 3.7.13, “Configuring SAN heartbeating in a virtual environment” on page 103 for a configuration example.

You can migrate directly to PowerHA 7.1.1 from PowerHA versions 5.4.1, 5.5, and 6.1 by using one of the following migration methods:

- Offline migration (as explained in “Offline migration from PowerHA 5.5”)
- Snapshot migration (as explained in “Snapshot migration from PowerHA 5.5”)
- Rolling migration (as explained in “Rolling migration from PowerHA 6.1”)

If you run a version that is earlier than PowerHA 5.4.1, you must upgrade to one of the three versions that we mentioned first.

In case you migrate from PowerHA 7.1.0, only the offline or snapshot migration method is available (see 6.3, “Migration from PowerHA 7.1.0” on page 321).

**Important:** Rolling migration from PowerHA 7.1.0 to PowerHA 7.1.1 is not supported.

# 6.2 Migration from pre-7.1 versions

Before you begin migration, you must understand the process and all possible scenarios. The process of migration from PowerHA 6.1or prior versions to PowerHA 7.1.1 is similar to the migration process to PowerHA 7.1.0. Both processes differ from the earlier migration processes that are used to upgrade clusters to PowerHA 6.1 or before. The main reason is the introduction of the new Cluster Aware AIX (CAA) feature, initially in AIX 6.1 TL06 and AIX 7.1 TL00, which is used by PowerHA 7.1.

## 6.2.1 Requirements

As in migration to PowerHA 7.1.0, not all configurations can be migrated to PowerHA 7.1.1 from version 6.1 and earlier.

The support for the following outdated network technologies is removed:

- Asynchronous transfer mode (ATM)
- Fiber Distributed Data Interface (FDDI)
- Token Ring
- X.25

Cluster networks that are based on these network types must be removed or replaced with Ethernet networks before the migration.

Configurations that use the following features cannot be migrated:

- IP address takeover (IPAT) via Replacement
- Hardware Media Access Control (MAC) Address Takeover
- Heartbeat over IP aliases

These features must be removed from the cluster configuration before you start the migration. You can exchange IPAT via Replacement with IPAT via Aliases. Site concept and IPv6 functionality are not supported in this version.

Non-IP heartbeating is accomplished differently in PowerHA 7.1.1 so the previously used non-IP networks are not supported anymore:

- RS232
- HACMP Target Mode SCSI heartbeat (TMSCSI)
- Target Mode Serial Storage Architecture (TMSSA)

- Disk Heartbeat (DISKHB)
- Multinode disk heartbeat (MNDHB)

Although possible, you do not need to remove the non-IP networks from the configuration before the migration. The migration process warns you and removes any identified non-IP network.

The “Communication Path to a Node” parameter for each PowerHA cluster node must match the host name of that node. If they do not match, you have to reconfigure the cluster so that they match. The definition for the host name of a node is the output of the command **hostname** when run on that node.

The cluster node host name that resolves to a persistent IP address is not supported. The host name of the cluster nodes must resolve to a base IP. While you plan for migration from PowerHA 6.1or earlier to PowerHA 7.1.1, if a cluster node host name resolves to a persistent IP address, reconfigure the cluster so that the host name resolves to a base IP. For more details, see 3.7.5, “Setting up the hostname” on page 96.

Any standard shared volume group must be converted to Enhanced Concurrent Mode before you start the migration.

## 6.2.2 Stages of migration and the clmigcheck command

Migration to PowerHA 7.1.1 from PowerHA 5.4.1, 5.5, or 6.1involves the following stages:

1. Upgrade to AIX 6.1 TL07 or AIX 7.1 TL01.

   Before you can migrate PowerHA to version 7.1.1, you must upgrade the AIX operating system. You can upgrade all nodes to the appropriate AIX level first before you upgrade the PowerHA filesets on any node. You can either use a rolling migration approach, while you keep the cluster up on at least one node, or a total cluster outage. Only for the rolling migration, you can upgrade both the AIX operating system and the PowerHA filesets in the same step on a node before you go to the next node.

2. Perform the premigration check (**clmigcheck**).

   During this stage, you use the **clmigcheck** command to check whether the current PowerHA cluster configuration can be migrated to PowerHA 7.1.1:

   – Run the **clmigcheck** command on the first node to validate the current cluster configuration for migration. If the configuration is invalid, the **clmigcheck** program notifies you of any unsupported elements, such as disk heartbeating or IPAT via Replacement. If the configuration is supported, continue by choosing option 3 to specify the repository disk and the IP multicast address.

   – `The` **clmigcheck** command is executed on each remaining node in the cluster during the migration process. However, the menu options do not appear. You get a message that states “`OK to install the new version of PowerHA SystemMirror`.”

   > **The `clmigcheck` command:** The **clmigcheck** command is not a PowerHA command. It is part of `bos.cluster` and is in the `/usr/sbin` directory. When it is run on the last node of the PowerHA cluster, the **clmigcheck** automatically creates a CAA cluster.

3. Upgrade to PowerHA 7.1.1.

   After stage 2 is complete, you can upgrade the PowerHA filesets to version 7.1.1. In offline and snapshot migration procedures, you must upgrade them on all nodes before you start PowerHA cluster services on any node. With the rolling migration, you can perform all three stages on a node before you go to the next node.

## 6.2.3 Offline migration from PowerHA 5.5

This migration process involves bringing the PowerHA cluster down, installing PowerHA 7.1.1, and restarting PowerHA cluster services one node at a time.

Before you migrate PowerHA, AIX needs to be upgraded or migrated on all cluster nodes. All cluster nodes must have one shared disk that is used for the cluster repository. Ensure that the current network setup is multicast enabled.

PowerHA SystemMirror 7.1 and higher supports *only* Enhanced Concurrent Mode volume groups. Therefore, bos.clvm.enh is a prerequisite to convert non-concurrent volume groups that are used in previous PowerHA versions.

We describe the steps of an offline migration example from PowerHA 5.5 to PowerHA 7.1.1 for a 2-node cluster.

### Test environment

The nodes are VIOS client LPARs on different frames that share the SAN LUNs through NPIV. The backing devices are LUNs on a DS4800 storage subsystem.

The environment begins with AIX 6.1.5.7 and PowerHA 5.5.0.10. The migration leads to AIX 6.1.7.3 and PowerHA 7.1.1.0.

### Starting configuration

Figure 6-1 on page 260 shows a simplified layout of a cluster that is migrated in this scenario. Both the systems run AIX 6.1 TL5 SP7. The installed PowerHA is version 5.5 SP10.

The cluster layout is a hot-standby configuration. The `sydney` system is the primary server for the DB2 application. The `perth` system is the takeover or standby node.

Because of resource limitations, the disk heartbeat uses one of the existing shared disks. Three networks are defined:

- The `net_ether_01` network is used by the application and its users.
- The `net_ether_02` network is an IPAT via Replacement network.
- The `net_diskhb_01` network is the disk heartbeating network.

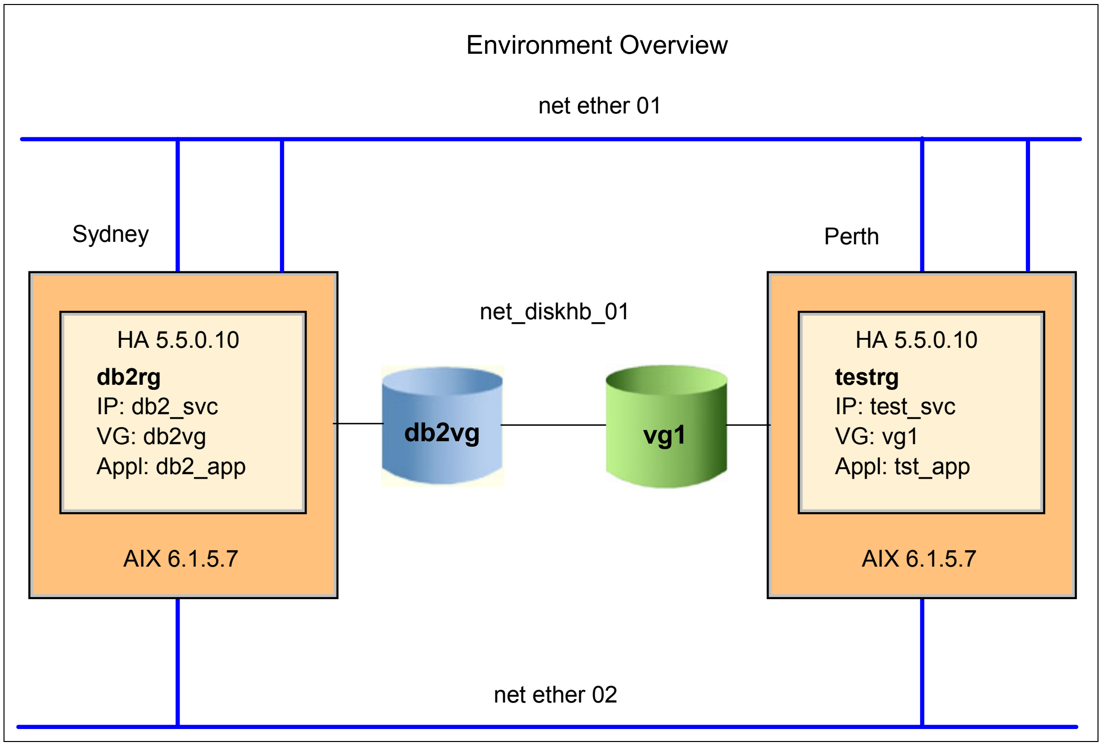



*Figure 6-1 Initial cluster configuration*

Figure 6.1 shows the cluster configuration before the migration begins. The cluster contains two resource groups: `db2rg` and `testrg`. The `db2rg` resource group hosts the DB2 database application and uses IP Aliasing for IP Address takeover for its service address. The `testrg` hosts a dummy test script application and uses IPAT via Replacement for the service IP/label. The home node for `db2rg` is `sydney`. The `perth` node is home for `testrg`. Figure 6-2 on page 261 shows the AIX and cluster level and the resource group states that are migrated in this scenario.

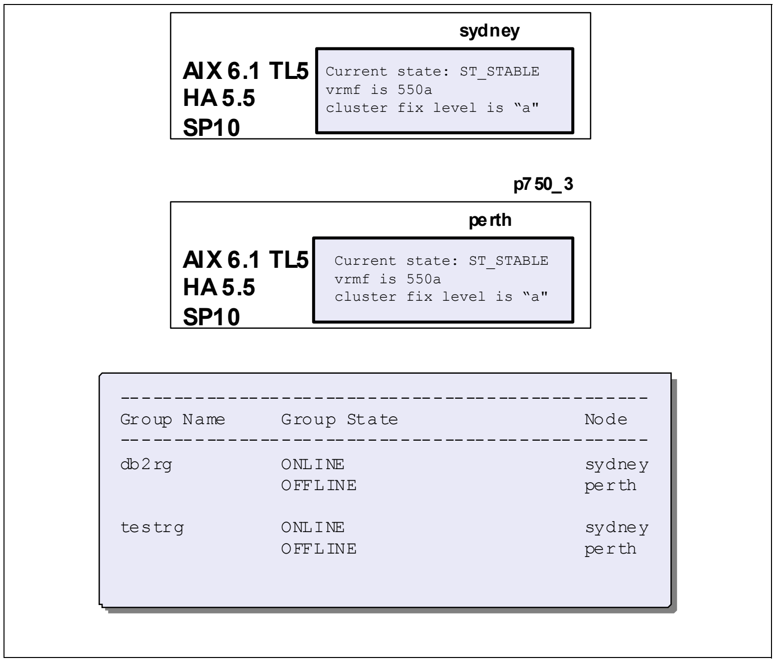



*Figure 6-2 Two-node cluster configuration before migration*

Figure 6-3 shows the **cllsif** output before cluster migration.

```
root@sydney /# /usr/es/sbin/cluster/utilities/cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address Interface Name
Global Name Netmask Alias for HB Prefix Length

perth_hdisk3_01   service net_diskhb_01  diskhb serial perth /dev/hdisk3 hdisk3
perth_b1 boot net_ether_01 ether public perth 10.1.1.55 en0 255.255.255.0 24

db2_svc service  net_ether_01  ether public perth 172.16.21.71 255.255.255.0 24

test_svc service  net_ether_01 ether public perth 10.10.20.71 255.255.255.0 24

perth_b2 boot    net_ether_02 ether public perth 10.1.2.55 en1 255.255.255.0 24

sydney_hdisk3_01 service net_diskhb_01 diskhb serial sydney /dev/hdisk3 hdisk3

sydney_b1 boot net_ether_01 ether public sydney 10.1.1.54 en0 255.255.255.0 24

db2_svc service net_ether_01 ether public sydney 172.16.21.71 255.255.255.0 24

test_svc service net_ether_01 ether public sydney 10.10.20.71 255.255.255.0 24

sydney_b2 boot net_ether_02 ether public sydney 22.22.22.47 en1 255.255.255.0 24
root@sydney /#
```

*Figure 6-3 cllsif output before cluster migration*

## Planning the offline migration

Offline migration is best when you have a maintenance window.

Your migration planning must begin with a careful analysis of the cluster to avoid starting a migration from an unsupported configuration. After the unsupported elements are identified, downtime must be scheduled to convert the configuration to a supported one.

The **`clmigcheck`** command can assist you with this analysis, but this command needs the AIX operating system to be upgraded, at least on one node. A possible approach is to first upgrade the AIX operating system on all nodes, one node at a time. Keep the cluster running on the other nodes, or upgrade AIX on all nodes by having a cluster downtime.

An additional shared disk is required for the new CAA repository disk. Figure 6-4 shows the results of the completed migration. To migrate, see “Migrating the nodes” on page 266.

After migration, the cluster configuration looks like Figure 6-4.

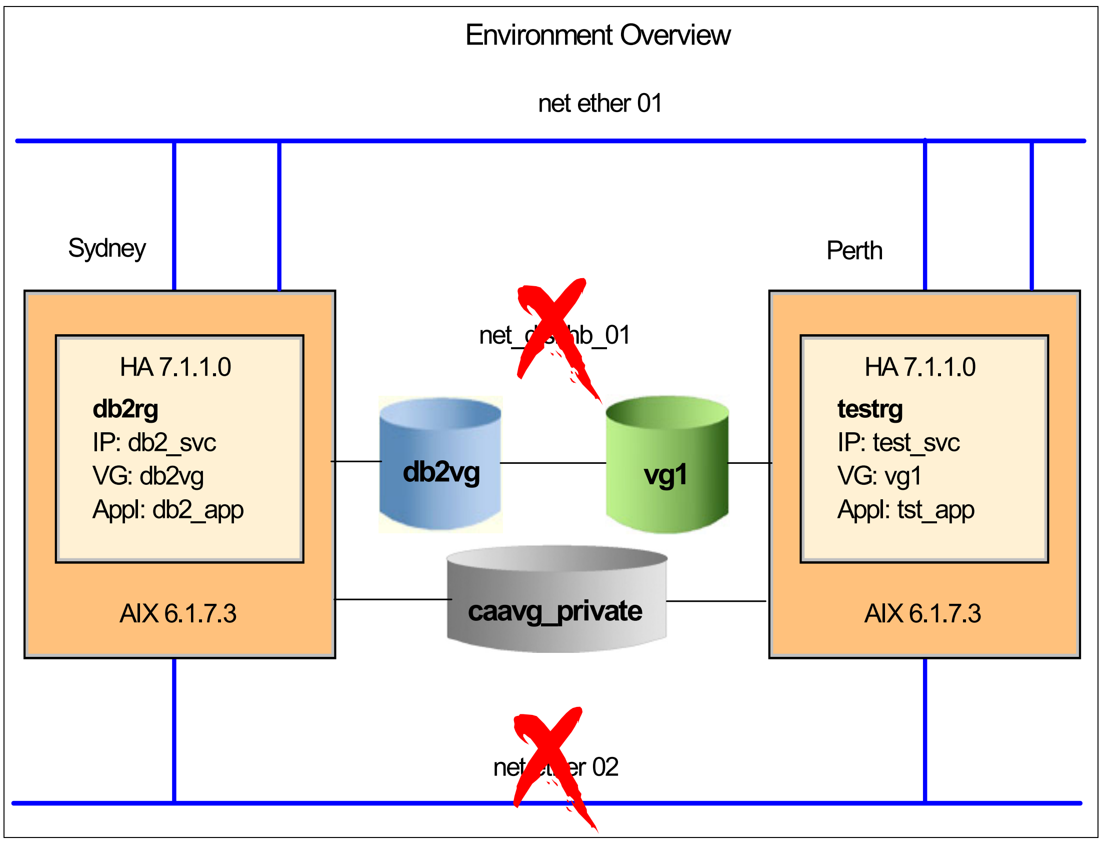



*Figure 6-4 Cluster configuration after migration*

Figure 6-5 on page 263 shows the AIX and PowerHA levels after migration is completed. It also shows resource group states.

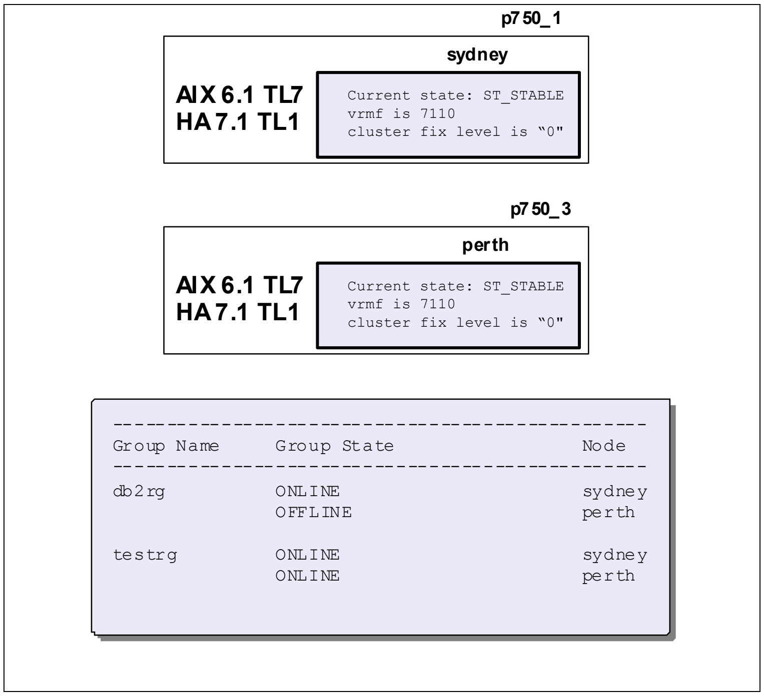



*Figure 6-5 Two-node cluster configuration after migration*

The cluster configuration, after migration, does not have a disk-based heartbeat network and it does not have an IPAT via Replacement type of network (`net_ether_02`) either. The **`cllsif`** output that is shown in Figure 6-6 shows the migrated configuration.

```
root@sydney /# /usr/es/sbin/cluster/utilities/cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address Interface Name
Global Name Netmask Alias for HB Prefix Length

perth_b1 boot net_ether_01 ether public perth 10.1.1.55 en0 255.255.255.0 24

db2_svc service net_ether_01 ether public perth 172.16.21.71 255.255.255.0 24

test_svc service net_ether_01 ether public perth 10.10.20.71 255.255.255.0 24

perth_b2 boot net_ether_02 ether public perth 10.1.2.55 en1 255.255.255.0 24

sydney_b1 boot net_ether_01 ether public sydney 10.1.1.54 en0 255.255.255.0 24

db2_svc service net_ether_01 ether public sydney 172.16.21.71 255.255.255.0 24

test_svc service net_ether_01 ether public sydney 10.10.20.71 255.255.255.0 24

sydney_b2 boot net_ether_02 ether public sydney 22.22.22.47 en1 255.255.255.0 24
root@sydney /#
```

*Figure 6-6 cllsif output after cluster migration completes*

A new volume group, `caavg_private`, is created when the CAA cluster is configured.

Figure 6-7 on page 264 and Figure 6-8 on page 264 show an example of using the **`clmigcheck`** command to identify unsupported elements in the configuration, in this case, IPAT via Replacement and disk heartbeat network.

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_diskhb_01.  This will be
removed from the configuration during the migration
 to PowerHA System Mirror 7.1.


Hit <Enter> to continue
```

*Figure 6-7 clmigcheck warning message for disk-based heartbeat network*

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-ERROR: The configuration contains unsupported options:  IP Address
Takeover via Replacement.  The PowerHA network name is net_ether_02.

This will have to be removed from the configuration before
 migration to PowerHA System Mirror
```

*Figure 6-8 clmigcheck error for IPAT via Replacement*

Any set of actions that involves modifications at the operating system level or cluster level must be preceded by a back-out/reversion plan that includes the following operations:

- Back up all data and binaries of all applications.
- Take a cluster snapshot and save it locally and to another machine.
- Save a copy of any custom script files locally and to another machine.
- Perform a **mksysb** backup of each involved node.

A back-out plan allows easy restoration of the application data and binaries, cluster, and AIX configuration in case you run into problems.

## Preparing the offline migration

After you finish the analysis, take the following steps:

- Ensure that all nodes are at the same level of the operating system and cluster software (including PTFs).
- Check that the cluster software is committed (and not merely applied) by using **lslpp** command.
- Ensure that the software is consistent on all nodes by using the **lppchk** command.

Before you start the actual migration, check that the cluster has no pending changes on any of its nodes since the last cluster startup and also check that the cluster is in a stable state:

1. Run Verification and Synchronization (Figure 6-9 on page 265) on each cluster node by following the path **smitty hacmp** → **Extended Configuration Extended Verification and Synchronization**. Select **Yes** for the “Verify changes only?” field and press Enter.

```
 HACMP Verification and Synchronization (Active Cluster Nodes Exist)

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

                                                      [Entry Fields]
* Verify changes only?                              [Yes]              +
* Logging                                           [Standard]         +


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

*Figure 6-9 Verifying and synchronizing changes, if any*

2. Confirm that the “Verify changes only?” operations end with an “OK” message and that no pending cluster configuration changes are detected by cldare on each of the cluster nodes. Figure 6-10 shows this result for one of the cluster nodes.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...85]
Adding any necessary HACMP for AIX entries to /etc/inittab and /etc/rc.net for
IP Address Takeover on node sydney.
Adding any necessary HACMP for AIX entries to /etc/inittab and /etc/rc.net for
IP Address Takeover on node perth.


Verification has completed normally.

cldare: No changes detected in Cluster Topology or Resources.
...completed.

[MORE...18]
F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-10 No change detected during verification*

Any other message from cldare means that changes are pending for the cluster or that errors are detected. You must perform more examination and finally ensure that the cluster is in the required configuration with no errors or pending changes.

3. Ensure that you start from a stable state of the cluster (Example 6-1 on page 266).

*Example 6-1 Checking that the cluster is stable*

```
root@sydney / # lssrc -ls clstrmgrES | grep state
Current state: ST_STABLE
root@sydney / #
```

### Migrating the nodes

To migrate the cluster, follow the steps:

1. Stop the cluster services on all nodes as shown in Figure 6-11.

   Use the **Bring Resource Groups Offline** option in the **`smitty clstop`** panel.

```
 Stop Cluster Services

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

                                                     [Entry Fields]
* Stop now, on system restart or both                 now                    +
  Stop Cluster Services on these nodes           [sydney, perth]
+
  BROADCAST cluster shutdown?                         true                   +
* Select an Action on Resource Groups                  Bring Resource Groups
Offline+




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

*Figure 6-11 Stop cluster services on all nodes with Bring Resource Groups Offline*

   Check that the cluster services are stopped on all nodes and that resource groups are offline as shown in Example 6-2.

*Example 6-2 Checking the cluster state*

```
root@sydney/ # lssrc -ls clstrmgrES | grep state
Current state: ST_INIT
root@sydney / # clRGinfo
Cluster IPC error: The cluster manager on node clmgr1 is in ST_INIT or
NOT_CONFIGURED state and cannot process the IPC request.
root@sydney / #

root@perth / # lssrc -ls clstrmgrES | grep state
Current state: ST_INIT
root@perth / #
```

2. Upgrade AIX on all nodes to version 6.1 TL7 SP2 (or later).

Use any supported procedure to upgrade the operating system. See the AIX installation guide for the procedure. Then, install the following filesets if they are not already installed:

- `bos.cluster` and `bos.ahafs` are mandatory CAA required filesets.
- `bos.clvm.enh` is mandatory for PowerHA SystemMirror 7.1.x.
- `devices.common.IBM.storfwork.rte`

> **Mandatory filesets**: You must install the CAA specific `bos.cluster`, `bos.ahafs`, and all other mandatory filesets.

Ensure that you restart the node and check the operating system level as shown in Example 6-3.

*Example 6-3 Checking operating system level and AIX cluster fileset level*

```
root@sydney / # oslevel -s
6100-07-03-1207
root@sydney / #
root@sydney / # lslpp -L bos.cluster\* bos.ahafs bos.clvm.enh
devices.common.IBM.storfwork.rte
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  bos.ahafs                  6.1.7.2    C     F    Aha File System
  bos.cluster.rte            6.1.7.3    C     F    Cluster Aware AIX
  bos.cluster.solid          6.1.7.0    C     F    Cluster Aware AIX SolidDB
  bos.clvm.enh               6.1.7.0    C     F    Enhanced Concurrent Logical
                                                   Volume Manager
  devices.common.IBM.storfwork.rte
                             6.1.7.2    C     F    Storage Framework Module

root@perth / # oslevel -s
6100-07-03-1207
root@perth / #
root@perth / # lslpp -L bos.cluster\* bos.ahafs bos.clvm.enh
devices.common.IBM.storfwork.rte
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  bos.ahafs                  6.1.7.2    C     F    Aha File System
  bos.cluster.rte            6.1.7.3    C     F    Cluster Aware AIX
  bos.cluster.solid          6.1.7.0    C     F    Cluster Aware AIX SolidDB
  bos.clvm.enh               6.1.7.0    C     F    Enhanced Concurrent Logical
                                                   Volume Manager
  devices.common.IBM.storfwork.rte
                             6.1.7.2    C     F    Storage Framework Module
```

3. Install `clic.rte 4.7.0.0` after the AIX upgrade.
4. Decide on the shared disk to use as the cluster repository disk. If a previously defined volume group is on any disk, you must remove it. For more information, see Chapter 2, “Cluster Aware AIX” on page 35.

> **Previous volume disk group**: The disk must be a clean LUN that does not contain a previous volume group that is known to any of our cluster nodes or to a third party. The volume group information must be removed from any of the cluster nodes, too.

5. Enable CAA logging. Edit the `/etc/syslog.conf` file on all migrated nodes to add an entry such as “`*.info /tmp/syslog.out rotate size 10m files 10`”. Verify that the specified file exists; otherwise, create it. Then, restart or refresh the syslogd daemon (Example 6-4).

*Example 6-4 Updating the syslog.conf file to capture CAA events*

```
root@sydney / # grep *.info /etc/syslog.conf
*.info /tmp/syslog.out rotate size 10m files 10
root@sydney / # ls -l /tmp/syslog.out
ls: 0653-341 The file /tmp/syslog.out does not exist.
root@sydney / # > /tmp/syslog.out
root@sydney / # ls -l /tmp/syslog.out
-rw-r--r--    1 root     system           0 Mar 27 05:06 /tmp/syslog.out
root@sydney / # refresh -s syslogd
0513-095 The request for subsystem refresh was completed successfully.
root@sydney / # cat /tmp/syslog.out
Mar 27 05:06:36 clmgr1 syslog:info syslogd: restart
root@sydney / #

root@perth / # grep *.info /etc/syslog.conf
*.info /tmp/syslog.out rotate size 10m files 10
root@perth / # ls -l /tmp/syslog.out
ls: 0653-341 The file /tmp/syslog.out does not exist.
root@perth / # > /tmp/syslog.out
root@perth / # ls -l /tmp/syslog.out
-rw-r--r--    1 root     system           0 Mar 27 05:06 /tmp/syslog.out
root@perth / # refresh -s syslogd
0513-095 The request for subsystem refresh was completed successfully.
root@perth / # cat /tmp/syslog.out
Mar 27 05:06:36 clmgr1 syslog:info syslogd: restart
root@perth / #
```

Run the **`/usr/sbin/clmigcheck`** command on the first node to check for any unsupported element. You see the initial panel as shown in Figure 6-12.

```
------------[ PowerHA System Mirror Migration Check ]-------------

Please select one of the following options:

1       = Check ODM configuration.

2       = Check snapshot configuration.

3       = Enter repository disk and multicast IP addresses.


Select one of the above,"x"to exit or "h" for help:
```

*Figure 6-12 clmigcheck panel*

Follow these steps:

a. Select option **1** (Check ODM configuration).

   When you choose this option, the **`clmigcheck`** command checks your cluster configuration and reports any unsupported element that might affect the migration.

This migration scenario uses a disk-based heartbeat path between two nodes and the IPAT via Replacement type of network. If you have more than one disk-based heartbeat network, it is reported, as well.

The **clmigcheck** command detects these situations and warns you that this configuration cannot be migrated as shown in Figure 6-13 and Figure 6-14.

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_diskhb_01.  This will
be
removed from the configuration during the migration
 to PowerHA System Mirror 7.1.



Hit <Enter> to continue
```

*Figure 6-13 clmigcheck warning for disk heartbeat network*

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-ERROR: The configuration contains unsupported options:  IP Address
Takeover via Replacement.  The PowerHA network name is net_ether_02.

This will have to be removed from the configuration before
 migration to PowerHA System Mirror
```

*Figure 6-14 clmigcheck error for IPAT via Replacement network*

b. Remove the IPAT via Replacement type of network from the configuration. Use **smitty hacmp**. Select **Extended Configuration** → **Extended Topology Configuration** → **Configure HACMP Networks** → **Remove a Network from HACMP Cluster** to remove the unsupported networks. Select the IPAT via Replacement network, **net_ether_02**, in this example as shown in Figure 6-15 on page 270.

```
+--------------------------------------------------------------------------+
|                        Select a Network to Remove
|
|
| Move cursor to desired item and press Enter.
|
|
|
|
|   net_diskhb_01
|   net_ether_01 (172.16.21.0/24 10.10.20.0/24 10.1.1.0/24)            |
|   net_ether_02
|
|
|
| F1=Help                F2=Refresh             F3=Cancel
|
| F8=Image               F10=Exit               Enter=Do
|
| /=Find                 n=Find Next
|
+--------------------------------------------------------------------------+
```

*Figure 6-15 Pop-up menu that displays HACMP networks*

After you press Enter, net_ether_02 is removed from the configuration as shown in the SMIT panel in Figure 6-16.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Network removed: net_ether_02

F1=Help                              F2=Refresh
F3=Cancel                                 F6=Command
F8=Image                             F9=Shell
F10=Exit                                  /=Find
```

*Figure 6-16 SMIT menu that shows success status for network removal*

c. Verify and synchronize the cluster. Start **smitty hacmp**. Select **Extended Configuration** → **Extended Verification and Synchronization**.

d. Run **clmigcheck** again and select option **3** to enter the repository disk as shown in Figure 6-17 on page 271.

```
------------[ PowerHA System Mirror Migration Check ]-------------

Select the disk to use for the repository

        1       = 00f74d4733c9370e(hdisk3)

Select one of the above or "x" to exit:1
```

*Figure 6-17 clmigcheck panel to select repository disk*

e. Enter the multicast IP address to be used for network monitoring as shown in Figure 6-18.

```
------------[ PowerHA System Mirror Migration Check ]-------------

PowerHA System Mirror uses multicast address for internal cluster communi-
cation and monitoring. These must be in the multicast range, 224.0.0.0 -
239.255.255.255.

If you make a NULL entry, AIX will generate an appropriate address for you.
You should only specify an address if you have an explicit reason to do
so, but are cautioned that this address cannot be changed once the
configuration is activated (i.e. migration is complete).


h = help

Enter the multicast IP address to use for network monitoring: 228.1.1.36
```

*Figure 6-18 clmigcheck panel to enter multicast IP address*

The repository disk and the multicast IP address are saved in the `/var/clmigcheck/clmigcheck.txt` file on each node of the cluster.

f. Exit the **clmigcheck** tool.

6. Upgrade PowerHA on the `sydney` node to the PowerHA 7.1.1 base release.

Perform an update_all installation by using the base PowerHA 7.1.1 filesets. Do not apply any full service pack or individual fix on top of the base filesets until all nodes in the cluster are upgraded to the new base release. The only exception is the application of interim fixes that are supplied specifically for the base filesets, which is allowed. Example 6-5 shows the PowerHA filesets that are updated to the base 7.1.1 release.

*Example 6-5 Check PowerHA fileset levels*

```
root@sydney / # lslpp -L cluster*
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.adt.es.client.include
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0    C     F    PowerHA SystemMirror Client
```

```
                                                CLINFO Samples
  cluster.adt.es.client.samples.clstat
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Clstat Samples
  cluster.adt.es.client.samples.libcl
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   LIBCL Samples
  cluster.adt.es.java.demo.monitor
                             7.1.1.0    C     F    Web Based Monitor Demo
  cluster.es.client.clcomd   7.1.1.0    C     F    Cluster Communication
                                                   Infrastructure
  cluster.es.client.lib      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Libraries
  cluster.es.client.rte      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Runtime
  cluster.es.client.utils    7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Utilities
  cluster.es.client.wsm      7.1.1.0    C     F    Web based Smit
  cluster.es.cspoc.cmds      7.1.1.0    C     F    CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0    C     F    CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.0    C     F    CSPOC Runtime Commands
  cluster.es.director.agent.rte
                             7.1.1.0    C     F    PowerHA SystemMirror
Director
                                                   CAS agent
  cluster.es.migcheck        7.1.1.0    C     F    PowerHA SystemMirror
Migration
                                                   support
  cluster.es.nfs.rte         7.1.1.0    C     F    NFS Support
  cluster.es.server.cfgast   7.1.1.0    C     F    Two-Node Configuration
                                                   Assistant
  cluster.es.server.diag     7.1.1.0    C     F    Server Diags
  cluster.es.server.events   7.1.1.0    C     F    Server Events
  cluster.es.server.rte      7.1.1.0    C     F    Base Server Runtime
  cluster.es.server.testtool
                             7.1.1.0    C     F    Cluster Test Tool
  cluster.es.server.utils    7.1.1.0    C     F    Server Utilities
  cluster.license            7.1.1.0    C     F    PowerHA SystemMirror
                                                   Electronic License
  cluster.man.en_US.es.data  7.1.1.0    C     F    Man Pages - U.S. English
```

We upgraded the operating system to AIX 6.1 TL7 SP3 earlier and restarted the system. A restart is not required. The update process converts the existing cluster configuration to be compatible with the new software. Conversion details are logged in the `/tmp/clconvert.log` file.

7. Populate the `/etc/cluster/rhosts` file on `sydney` with `hostname` IP addresses as shown in Example 6-6.

   You must include the addresses that resolve to the host name of the cluster nodes.

*Example 6-6 Updating /etc/cluster/rhosts file*

```
root@sydney / # cat /etc/cluster/rhosts
10.1.1.54
10.1.1.55
root@sydney / # host  `hostname`
```

```
sydney is 10.1.1.54
root@sydney / #
```

As you populate the `/etc/cluster/rhosts` file, you must refresh the clcomd subsystem as shown in Example 6-7.

*Example 6-7 Refreshing clcomd*

```
root@sydney / # lssrc -s clcomd
Subsystem         Group            PID          Status
 clcomd           caa              14024890     active
root@sydney / # refresh -s clcomd
root@sydney / #
```

8. Populate the `/etc/cluster/rhosts` file in the `perth` node with the `hostname` IP addresses.

   You must include the addresses that resolve to the host name of the cluster nodes as shown in Example 6-8.

*Example 6-8 Updating /etc/cluster/rhosts on the last node*

```
root@perth / # cat /etc/cluster/rhosts
10.1.1.54
10.1.1.55
root@perth / # host  `hostname`
perth is 10.1.1.55
root@perth / #
```

After you populate the `/etc/cluster/rhosts` file, you must refresh the clcomd subsystem as shown in Example 6-9.

*Example 6-9 Refreshing clcomd on the last node*

```
root@perth / # lssrc -s clcomd
Subsystem         Group            PID          Status
 clcomd           caa              15401162     active
root@perth / # refresh -s clcomd
root@perth / #
```

9. Run the **clmigcheck** tool on the last node of the cluster as shown in Example 6-10.

*Example 6-10 clmigcheck message that indicates the last node of the cluster*

```
root@perth /# clmigcheck
Verifying clcomd communication, please be patient.

Verifying multicast IP communication, please be patient.

mping version 1.1
Localhost is sydney, 10.1.1.54
Listening on 228.168.101.43/4098:

Replying to mping from 10.1.1.55 bytes=32 seqno=2 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.55 bytes=32 seqno=3 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.55 bytes=32 seqno=4 ttl=32
Discarding receiver packet
```

```
Replying to mping from 10.1.1.55 bytes=32 seqno=5 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.55 bytes=32 seqno=1 ttl=32
Discarding receiver packet
clmigcheck: Running
/usr/sbin/rsct/install/bin/ct_caa_set_disabled_for_migration on each node in
the cluster

Creating CAA cluster, please be patient.

------------[ PowerHA System Mirror Migration Check ]-------------

You can install the new version of PowerHA System Mirror.


Hit <Enter> to continue

root@perth /#
```

10. The **clmigcheck** tool detects the last node of the cluster. Then, it creates the CAA infrastructure on all nodes.

    You can run the **/usr/sbin/lscluster -m** command to check the CAA cluster configuration as shown in Example 6-11.

*Example 6-11 Checking the CAA cluster*

```
root@sydney / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: sydney
        Cluster shorthand id for node: 1
        uuid for node: e064d038-77d9-11e1-8da0-b6fcc07d1d70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        cls1               local        e0901e5a-77d9-11e1-8da0-b6fcc07d1d70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: perth
        Cluster shorthand id for node: 2
        uuid for node: e08d371c-77d9-11e1-8da0-b6fcc07d1d70
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
```

```
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        cls1               local        e0901e5a-77d9-11e1-8da0-b6fcc07d1d70

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
         dpcom  UP  RESTRICTED
         en1  UP
         en0  UP
root@sydney / #
```

11.Upgrade PowerHA on the `perth` node to the PowerHA 7.1.1 base release.

Perform an **`update_all`** installation by using the base PowerHA 7.1.1 filesets. Do not apply any full service pack or individual fix on top of the base filesets until all nodes in the cluster are upgraded to the new base release. The only exception is the application of interim fixes that are supplied specifically for the base filesets, which is allowed. Example 6-12 shows the PowerHA filesets that are updated to the base 7.1.1 release.

*Example 6-12 Checking PowerHA fileset level on the last node*

```
root@perth / # lslpp -L cluster*
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.adt.es.client.include
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   CLINFO Samples
  cluster.adt.es.client.samples.clstat
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Clstat Samples
  cluster.adt.es.client.samples.libcl
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   LIBCL Samples
  cluster.adt.es.java.demo.monitor
                             7.1.1.0    C     F    Web Based Monitor Demo
  cluster.es.client.clcomd   7.1.1.0    C     F    Cluster Communication
                                                   Infrastructure
  cluster.es.client.lib      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Libraries
  cluster.es.client.rte      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Runtime
  cluster.es.client.utils    7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Utilities
  cluster.es.client.wsm      7.1.1.0    C     F    Web based Smit
  cluster.es.cspoc.cmds      7.1.1.0    C     F    CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0    C     F    CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.0    C     F    CSPOC Runtime Commands
  cluster.es.migcheck        7.1.1.0    C     F    PowerHA SystemMirror
Migration
                                                   support
  cluster.es.nfs.rte         7.1.1.0    C     F    NFS Support
  cluster.es.server.cfgast   7.1.1.0    C     F    Two-Node Configuration
```

```
                                                    Assistant
cluster.es.server.diag     7.1.1.0    C     F    Server Diags
cluster.es.server.events   7.1.1.0    C     F    Server Events
cluster.es.server.rte      7.1.1.0    C     F    Base Server Runtime
cluster.es.server.testtool
                           7.1.1.0    C     F    Cluster Test Tool
cluster.es.server.utils    7.1.1.0    C     F    Server Utilities
cluster.license            7.1.1.0    C     F    PowerHA SystemMirror
                                                 Electronic License
cluster.man.en_US.es.data  7.1.1.0    C     F    Man Pages - U.S. English
cluster.msg.en_US.es.client
                           7.1.1.0    C     F    PowerHA SystemMirror Client
                                                 Messages - U.S. English
cluster.msg.en_US.es.server
                           7.1.1.0    C     F    Recovery Driver Messages -
                                                 U.S. English
```

12..Start the cluster services, one node at a time, and check that each node successfully joins the cluster.

You see warning messages about a mixed level cluster when the node reintegrates to the cluster (Figure 6-19 on page 277).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Cluster services are running at different levels across
the cluster.  Verification will not be invoked in this environment.

Starting Cluster Services on node: sydney
This may take a few minutes.  Please wait...
sydney: start_cluster: Starting PowerHA SystemMirror
sydney:  4456498      -  0:00 syslogd
sydney: Setting routerevalidate to 1
sydney: 0513-059 The topsvcs Subsystem has been started. Subsystem PID is
8585320.
sydney: 0513-059 The grpsvcs Subsystem has been started. Subsystem PID is
8388776.
sydney: 0513-059 The emsvcs Subsystem has been started. Subsystem PID is
7667948.
sydney: 0513-059 The clevmgrdES Subsystem has been started. Subsystem PID is
7078138.
sydney: 0513-059 The gsclvmd Subsystem has been started. Subsystem PID is
7471112.
sydney: Mar 26 2012 15:05:44 Starting execution of
/usr/es/sbin/cluster/etc/rc.cluster
sydney: with parameters: -boot -N -A -C interactive -P cl_rc_cluster
sydney:
sydney: Mar 26 2012 15:05:48 Checking for srcmstr active...
[MORE...13]

F1=Help                F2=Refresh              F3=Cancel
Esc+6=Command
Esc+8=Image            Esc+9=Shell             Esc+0=Exit
/=Find
n=Find Next
```

*Figure 6-19 Warning message about mixed level cluster*

Run **lssrc -ls clstrmgrES** or **odmget HACMPcluster** on the node to check the cluster version as shown in Example 6-13.

*Example 6-13 Checking cluster version on the first node*

```
root@sydney / # lssrc -ls clstrmgrES | grep version
CLversion: 13
root@sydney / # odmget HACMPcluster | grep -e "cluster_version"
        cluster_version = 13
root@sydney / #
```

13. After the node sydney successfully joins the cluster, start cluster services on the second node of the cluster (perth).

You can use **smit clstart** to start the cluster services. When the cluster services start on this node, the warning message about a mixed level cluster disappears.

You can run **lssrc -ls clstrmgrES** or **odmget HACMPcluster** on the node to check the cluster version as shown in Example 6-14.

*Example 6-14 Checking cluster version on the second node*

```
root@perth / # lssrc -ls clstrmgrES | grep version
CLversion: 13
root@perth / # odmget HACMPcluster | grep -e "cluster_version"
        cluster_version = 13
root@perth / #
```

Check that the cluster is stable and that the resource groups are online as expected (per the participating node list and startup policy). See Example 6-15.

*Example 6-15 clRGinfo command output*

```
root@clmgr1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
db2rg          ONLINE                       sydney
               OFFLINE                      perth

testrg         ONLINE                       perth
               OFFLINE                      syndey
```

## 6.2.4 Snapshot migration from PowerHA 5.5

We use the cluster snapshot to migrate from PowerHA 5.5 to PowerHA SystemMirror 7.1.1. In this migration, we need to stop the cluster services on all nodes gracefully, uninstall PowerHA 5.5, and install PowerHA SystemMirror 7.1.1. There is a minimum downtime window during the migration. Plan this downtime before you migrate the cluster. We use the **clmigcheck** utility to check whether the cluster snapshot complies with PowerHA SystemMirror 7.1.1. We need to remove all unsupported elements in the configuration before we migrate to PowerHA SystemMirror 7.1.1.

### Cluster configuration

The test environment is a 2-node cluster with two Ethernet networks, one disk heartbeat network, and one resource group as shown in Figure 6-20 on page 279.

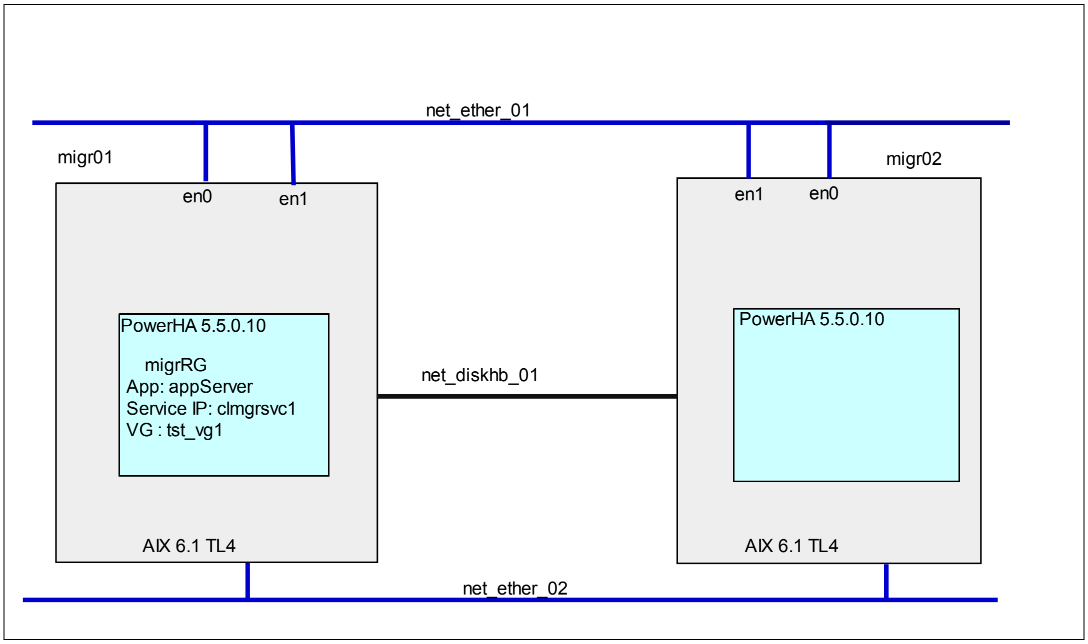



*Figure 6-20 Cluster diagram*

The Ethernet network `net_ether_01` is used by the application server and IP address takeover via aliasing. The resource group contains the user application server, volume group, and service IP. The user application server script writes the date entry into the file that is hosted by the volume group `tst_vg1` every 3 seconds. It has start and stop scripts. We show the various scripts that are used in this test environment. The application server script is shown in Example 6-16.

*Example 6-16 Application server script that writes date entry into the file every 3 seconds*

```
#!/bin/ksh
#  this is just an application we use to see if the vg, lv, fs and application
#  continue running after a takeover
#
set -x

outfile=/ha_testfs/test.out

rm $outfile

count=0

# check if the filesystem is mounted
while [[ -z $flag ]]
do
        mounted=`/usr/sbin/mount | grep ha_testfs`
        if [[ -n $mounted ]]
        then
                flag=true
        fi

   let count=count+3
   if [ $count -gt 360 ]
```

```
   then
      flag=false
   fi
done

if [[ $flag = true ]]
then

#loop to write to the new output file
        while (true); do
                echo `date` >> $outfile
                sleep 3
        done
else
        exit 1
fi
exit 0
```

The script that is used to start the user application server is shown in Example 6-17.

*Example 6-17 Startup script to start application server*

```
#!/bin/ksh
count=0
while [[ -z $flag ]]
do
   tmp=`/usr/sbin/lsvg -o | /bin/grep tst_vg1`
   if [[ -n $tmp ]]
   then
      flag=true
   fi
   let count=count+3
   if [ $count -gt 360 ]
   then
      flag=false
   fi
done
#
if [[ $flag = true ]]
then
   /scripts/appscript
else
   exit 1
fi
#
exit 0
```

The script that is used to stop the user application server is shown in Example 6-18.

*Example 6-18 Stop script to stop application server*

```
#!/bin/ksh
#
set -x
PIDS=`ps -ef | grep /scripts/appscript | grep -v grep | awk '{print $2}'`
```

```
#
echo $PIDS is the pid for appscript
if [ -n $PIDS ]
then
   for pid in $PIDS
   do
      /usr/bin/kill -TERM $pid
   done
fi
```

The script that is used to monitor the user application is shown in Example 6-19.

*Example 6-19 Script that is used to monitor the user application server*

```
#!/bin/ksh
proc=`ps -ef | grep /scripts/appscript | grep -v grep`
rc=$?
if [[ $rc != 0 ]]; then
        return 1
fi
return 0
```

### Taking the system image backup

Take the system backup by using the **mksysb** command before you start the migration so that you can restore the previous system image if the migration fails. For information about taking system backups by using the **mksysb** command, see “AIX 6.1 Installation and migration” at this website:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.install/doc/insgdrf/insgdrf_pdf.pdf

Migration consists of various steps. We go through these steps in detail.

### Updating the AIX level

PowerHA SystemMirror 7.1.1 requires AIX 6.1 with TL 7 (6100-07) plus SP02 or higher version of the `bos.cluster.rte` (6.1.7.2) or AIX 7.1 with TL 1 (7100-01) plus the SP02 or higher version of `bos.cluster.rte` (7.1.1.2) and RSCT 3.1.2.0. Update AIX on all nodes in the cluster to AIX 6.1 TL 7 or AIX 71. TL 1. Then, install the following filesets if they are not already installed:

- `bos.cluster`
- `bos.ahafs`
- `bos.clvm.enh`
- `bos.data`

> **Required:** `bos.cluster`, `bos.ahafs`, and `bos.clvm.enh` are mandatory filesets for PowerHA SystemMirror 7.1.1.

Restart the nodes after you update the AIX level. Check that all required filesets are installed. Ensure that the filesets installed cleanly as shown in Example 6-20.

*Example 6-20 Mandatory filesets for SystemMirror 7.1.1*

```
[c665f1sq07][/]> lppchk -v
[c665f1sq07][/]> lslpp -L bos.cluster\* bos.clvm.enh bos.data bos.ahafs
  Fileset                      Level  State  Type  Description (Uninstaller)
```

```
---------------------------------------------------------------------------
bos.ahafs                  6.1.7.1    C     F    Aha File System
bos.cluster.rte            6.1.7.3    C     F    Cluster Aware AIX
bos.cluster.solid          6.1.7.1    C     F    POWER HA Business Resiliency
                                                 SolidDB
bos.clvm.enh               6.1.7.0    C     F    Enhanced Concurrent Logical
                                                 Volume Manager
bos.data                  6.1.6.15    C     F    Base Operating System Data
```

## Creating a snapshot

Create a cluster snapshot by using the **smit** command. Save the snapshot in the /tmp/migration folder. Start **smit hacmp**. Select **Extended Configuration** → **Snapshot Configuration** → **Create a Snapshot of the Cluster Configuration**. Enter the cluster snapshot name and description as shown in Figure 6-21, and press Enter.

```
 Create a Snapshot of the Cluster Configuration

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

                                                        [Entry Fields]
* Cluster Snapshot Name                             [cluster_migration]
/
  Custom Defined Snapshot Methods                      []
+
  Save Cluster Log Files in snapshot                   No
+
* Cluster Snapshot Description                              [used for migration]



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

*Figure 6-21 Smit panel to create the cluster snapshot*

The command creates a snapshot under the /usr/es/sbin/cluster/snapshots/ directory as shown in Figure 6-22 on page 283. It creates two files: migration_cluster.odm and migration_cluster.info. If the cluster services are running, stop the cluster services gracefully.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.


clsnapshot: Creating file
/usr/es/sbin/cluster/snapshots/migration_cluster.odm...

clsnapshot: Creating file
/usr/es/sbin/cluster/snapshots/migration_cluster.info...

clsnapshot: Executing clsnapshotinfo command on node: migr01...

clsnapshot: Executing clsnapshotinfo command on node: migr02...

clsnapshot: Succeeded creating Cluster Snapshot: migration_cluster.



F1=Help               F2=Refresh             F3=Cancel
F6=Command
F8=Image              F9=Shell               F10=Exit              /=Find
n=Find Next
```

*Figure 6-22 Shows snapshot that is created successfully*

## Checking snapshot compliance

There are a few limitations in the snapshot migration. Not all cluster configurations are compliant for migration. First, we need to remove the unsupported elements from the cluster configuration before we migrate. The following configuration items are not supported:

- Asynchronous transfer mode (ATM), Fiber Distributed Data Interface (FDDI), and token ring interfaces
- Two-node or multinode disk heartbeat network
- Two sites
- Extended distance (XD) configuration
- IPv6
- IPAT via IP Replacement
- Enhanced security mode
- Heartbeat via aliasing

### *The clmigcheck utility*

We use the **clmigcheck** utility to check whether the cluster snapshot is compliant with PowerHA SystemMirror 7.1.1. This utility is at `/usr/sbin`. You can check the path by using the **which** command as shown in Figure 6-23 on page 284.

```
[c665f1sq09][/]> which clmigcheck
/usr/sbin/clmigcheck
```

*Figure 6-23 clmigcheck path*

Run the **clmigcheck** command, which prompts for user input as shown in Figure 6-24.

```
------------[ PowerHA System Mirror Migration Check ]-------------

Please select one of the following options:

1       = Check ODM configuration.

2       = Check snapshot configuration.

3       = Enter repository disk and multicast IP addresses.


Select one of the above,"x"to exit or "h" for help:
```

*Figure 6-24 clmigcheck command prompt*

With the **clmigcheck** command, you can perform these functions:

1. Check the ODM configuration.

   It checks whether the Object Data Manager (ODM) entries are compliant with PowerHA SystemMirror 7.1.1. This option is used in other migration methods. We are not using this option in the snapshot migration.

2. Check the snapshot configuration.

   It checks whether the snapshot is compliant with PowerHA SystemMirror 7.1.1.

3. Enter repository disk and multicast IP addresses.

   PowerHA SystemMirror supports CAA, which needs a shared repository disk to store cluster configuration. We need to specify the shared repository disk details in this step. For more information about the repository disk, see 3.6.1, “Repository disk” on page 91.

Run the snapshot configuration check. The **clmigcheck** command asks for the snapshot name. Enter the snapshot name that you created earlier. It processes the snapshot. If your configuration has unsupported elements, it displays configuration errors or warnings as shown in Figure 6-25 on page 285.

```
------------[ PowerHA System Mirror Migration Check ]-------------

h = help

Enter snapshot name (in /usr/es/sbin/cluster/snapshots): migration_cluster

clsnapshot: Removing any existing temporary HACMP ODM entries...
clsnapshot: Creating temporary HACMP ODM object classes...
clsnapshot: Adding HACMP ODM entries to a temporary directory...
clsnapshot: Succeeded generating temporary ODM containing Cluster Snapshot:
migration_cluster

------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_diskhb_01.  This will be
removed from the configuration during the migration
 to PowerHA System Mirror 7.1.


Hit <Enter> to continue
```

*Figure 6-25 Shows that ODM has unsupported elements*

Check the `/tmp/clmigcheck/clmigcheck.log` file for errors or warnings. If the cluster contains any unsupported elements, remove all unsupported elements from the cluster configuration. Take a fresh snapshot and run the **clmigcheck** command again until the ODM has no unsupported elements as shown in Figure 6-26.

```
------------[ PowerHA System Mirror Migration Check ]-------------

The ODM has no unsupported elements.

Hit <Enter> to continue
```

*Figure 6-26 Shows that ODM has no unsupported elements*

After you check the snapshot compliance, provide the repository disk and the multicast address details. Select option **3** from the PowerHA SystemMirror Migration Check window. Choose the appropriate shared repository disk as shown in Figure 6-27 on page 286.

```
------------[ PowerHA System Mirror Migration Check ]-------------

Please select one of the following options:

1       = Check ODM configuration.

2       = Check snapshot configuration.

3       = Enter reopsitory disk and multicast IP addresses.


Select one of the above,"x"to exit or "h" for help: 3
Verifying clcomd communication, please be patient.

------------[ PowerHA System Mirror Migration Check ]-------------

Select the disk to use for the repository

        1       = 00f74d4733c9370e(hdisk3)

Select one of the above or "x" to exit:  1
```

*Figure 6-27 Repository disk details*

After you select the repository disk, the **clmigcheck** utility asks the user to provide the multicast address. Enter the multicast address, if you reserved one for your cluster communication in your company network. Otherwise, enter a NULL value as shown in Example 6-21. The CAA creates the multicast address for you based on the base IP address. For more information about the multicast address, see 3.5.11, "Multicast heartbeat" on page 89.

*Example 6-21 Multicast address panel*

```
------------[ PowerHA System Mirror Migration Check ]-------------

PowerHA System Mirror uses multicast address for internal
cluster communication and monitoring. These must be in the
multicast range, 224.0.0.0 - 239.255.255.255.

If you make a NULL entry, AIX will generate an appropriate address for you.
You should only specify an address if you have an explicit reason to do
so, but are cautioned that this address cannot be changed once the
configuration is activated (i.e. migration is complete).


h = help

Enter the multicast IP address to use for network monitoring: NULL
```

The **clmigcheck** utility collects data from the user and creates the `/var/clmigcheck/clmigcheck.txt` on all nodes in the cluster.

### Uninstalling PowerHA 5.5

Uninstall existing PowerHA 5.5 filesets from all cluster nodes by using the **installp** command as shown in Example 6-22.

*Example 6-22 Uninstalling PowerHA 5.5 filesets*

```
root@clmgr1 / # installp -ug cluster.adt.* cluster.es.* cluster.license
+-----------------------------------------------------------------------------+
                    Pre-deinstall Verification...
+-----------------------------------------------------------------------------+
Verifying selections...done
Verifying requisites...done
Results...

SUCCESSES
---------
  Filesets listed in this section passed pre-deinstall verification
  and will be removed.

  Selected Filesets
  -----------------
  cluster.adt.es.client.include 5.5.0.0       # ES Client Include Files
  cluster.adt.es.client.samples.clinfo 5.5.0.0 # ES Client CLINFO Samples
  cluster.adt.es.client.samples.clstat 5.5.0.0 # ES Client Clstat Samples
  cluster.adt.es.client.samples.libcl 5.5.0.0 # ES Client LIBCL Samples
  cluster.adt.es.java.demo.monitor 5.5.0.0    # ES Web Based Monitor Demo
  cluster.es.cfs.rte 5.5.0.0                  # ES Cluster File System Support
  cluster.es.client.clcomd 5.5.0.0            # ES Cluster Communication Inf...
  cluster.es.client.lib 5.5.0.0               # ES Client Libraries
  cluster.es.client.rte 5.5.0.0               # ES Client Runtime
  cluster.es.client.utils 5.5.0.0             # ES Client Utilities
  cluster.es.client.wsm 5.5.0.0               # Web based Smit
  cluster.es.cspoc.cmds 5.5.0.0               # ES CSPOC Commands
  cluster.es.cspoc.dsh 5.5.0.0                # ES CSPOC dsh
  cluster.es.cspoc.rte 5.5.0.0                # ES CSPOC Runtime Commands
  cluster.es.server.cfgast 5.5.0.0            # ES Two-Node Configuration As...
  cluster.es.server.diag 5.5.0.0              # ES Server Diags
  cluster.es.server.events 5.5.0.0            # ES Server Events
  cluster.es.server.rte 5.5.0.0               # ES Base Server Runtime
  cluster.es.server.simulator 5.5.0.0         # ES Cluster Simulator
  cluster.es.server.testtool 5.5.0.0          # ES Cluster Test Tool
  cluster.es.server.utils 5.5.0.0             # ES Server Utilities
  cluster.license 5.5.0.0                     # HACMP Electronic License

  << End of Success Section >>
```

Check that you uninstalled all PowerHA 5.5 filesets cleanly with the commands that are shown in Figure 6-28 on page 288.

```
root@clmgr1 / # lslpp -l | grep cluster.*
root@clmgr1 / #
root@clmgr1 / # installp -C

0503-439 installp:  No filesets were found in the Software Vital
        Product Database that could be cleaned up.
```

*Figure 6-28 Shows a clean uninstall*

## /etc/syslog.conf and /etc/cluster/rhosts files

We need to edit the `/etc/syslog.conf` and the `/etc/cluster/rhosts` files before we install PowerHA SystemMirror 7.1.1.

PowerHA SystemMirror 7.1.0 and later use the `/etc/cluster/rhosts` file for communication. This file is not populated automatically. We must edit it manually and refresh the clcomd daemon. In this file, each line contains only cluster nodes IPs. For more information, see 3.8.1, "Propagating the /etc/cluster/rhosts file" on page 113.

We need to edit the `/etc/syslog.conf` file to enable the capture of the CAA log information. After we add the entry in the `syslog.conf` file, we refresh the syslogd daemon as shown in Example 6-23.

*Example 6-23 Editing /etc/syslog.conf file*

```
caa.info /var/adm/ras/syslog.caa rotate size 1m files 10
*.info /tmp/syslog.out rotate size 10m files 10
```

## PowerHA SystemMirror 7.1.1 installation

Now, install the PowerHA SystemMirror 7.1.1 filesets on all cluster nodes. The PowerHA SystemMirror 7.1.1 filesets are on the installation media. See 3.7.15, "Installing PowerHA filesets" on page 111.

## Converting a verified snapshot to PowerHA SystemMirror 7.1.1

We created the snapshot by using PowerHA 5.5. We need to convert it to PowerHA SystemMirror 7.1.1. We use the **`/usr/es/sbin/cluster/conversion/clconvert_snapshot`** command to convert the snapshot to the PowerHA SystemMirror 7.1.1 version. We need to pass the version number from which we are migrating and the snapshot name as shown in Example 6-24.

*Example 6-24 Converting cluster snapshot from 5.5 to 7.1.1*

```
root@clmgr1 / # /usr/es/sbin/cluster/conversion/clconvert_snapshot -v 5.5 -s
migration_cluster
Extracting ODM's from snapshot file... done.
Converting extracted ODM's... done.
Rebuilding snapshot file... done.
root@clmgr1 / #
```

**Command usage**: The command uses this syntax:

`/usr/es/sbin/cluster/conversion/clconvert_snapshot -v [release] [-s [snap_file]]`

Include these options:

`-v [`*`version`*`]` version from which we migrated

`-s [`*`snap_file`*`]` snapshot file

**Warning**: If you do not know your previous version, *do not* run this command.

### Restoring the cluster

We restore the cluster configuration from the converted snapshot:

1. We use the **`smit`** path to restore the cluster configuration. Start **`smit hacmp`**. Select **Cluster Nodes and Networks** → **Manage Cluster** → **Snapshot Configuration** → **Restore the Cluster Configuration From a Snapshot**.

   Select the snapshot name that we converted earlier as shown in Example 6-25 and press Enter.

   *Example 6-25 Restoring the cluster snapshot panel*

```
Restore the Cluster Snapshot

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

                                                          [Entry Fields]
  Cluster Snapshot Name                                 migration_cluster
  Cluster Snapshot Description                          Used -- for migration
  Un/Configure Cluster Resources?                     [Yes]
+
  Force apply if verify fails?                        [No]
+




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

2. If the restore is successful, the smit panel shows the `OK` message as shown in Figure 6-29 on page 290.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

clsnapshot: Removing any existing temporary PowerHA SystemMirror ODM entries...
clsnapshot: Creating temporary PowerHA SystemMirror ODM object classes...
clsnapshot: Adding PowerHA SystemMirror ODM entries to a temporary
directory..ODMDIR set to /tmp/snapshot


Verification has completed normally.

clsnapshot: Removing current PowerHA SystemMirror ODM entries...
clsnapshot: Adding new PowerHA SystemMirror ODM entries...
clsnapshot: Synchronizing cluster configuration to all cluster nodes...
/etc/es/objrepos
Timer object autoclverify already exists

Committing any changes, as required, to all available nodes...
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab and
/etc/rc.net for IP Address Takeover on node migr01.
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab and
/etc/rc.net for IP Address Takeover on node migr02.

Verification has completed normally.

clsnapshot: Succeeded applying Cluster Snapshot: Migration_cluster

[BOTTOM]

F1=Help                                F2=Refresh
F3=Cancel                                       F6=Command
F8=Image                               F9=Shell
F10=Exit                                        /=Find
n=Find Next
```

*Figure 6-29   Snapshot that restored successfully*

3. Check that the CAA cluster is created successfully by using the **lscluster** command as shown in Example 6-26.

*Example 6-26   Ensuring that the CAA cluster is created successfully*

```
root@clmgr1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: clmgr1
        Cluster shorthand id for node: 2
```

```
        uuid for node: 49c0b4a2-79b5-11e1-af6f-b6fcc07d1d70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        migcluster         local        4a508e6a-79b5-11e1-af6f-b6fcc07d1d70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: clmgr2
        Cluster shorthand id for node: 3
        uuid for node: 4a4add08-79b5-11e1-af6f-b6fcc07d1d70
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        migcluster         local        4a508e6a-79b5-11e1-af6f-b6fcc07d1d70

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en0  UP
         en1  UP
```

4. Also, check the cluster configuration with the **cltopinfo** command as shown in Example 6-27.

*Example 6-27 Ensuring the cluster is created in PowerHA*

```
root@clmgr1 / # cltopinfo
Cluster Name: Migration_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk3
Cluster IP Address: 228.1.1.54
There are 2 node(s) and 1 network(s) defined
NODE migr01:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr1b2        10.1.2.54
                clmgr1  10.1.1.54
NODE migr02:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr2b2        10.1.2.55
                clmgr2  10.1.1.55
```

```
Resource Group migrRG
        Startup Policy   Online On Home Node Only
        Fallover Policy  Fallover To Next Priority Node In The List
        Fallback Policy  Never Fallback
        Participating Nodes      migr01 migr02
        Service IP Label                 clmgrsvc1
```

5. Start the cluster services on all nodes by using the **smit clstart** command. If your cluster application server uses the clinfoES daemon, or if you want to check the cluster status by using the **clstat** command, set “Startup Cluster Information Daemon” to **true** as shown in Figure 6-30.

```
 Start Cluster Services

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

                                                         [Entry Fields]
* Start now, on system restart or both                 now
  Start Cluster Services on these nodes                [migr01,migr02]
* Manage Resource Groups                               Automatically
  BROADCAST message at startup?                        false
  Startup Cluster Information Daemon?                  true
  Ignore verification errors?                          false
  Automatically correct errors found during            Interactively
  cluster start?

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

*Figure 6-30 Starting the cluster services*

6. The command output shows OK as the status message. Check that the cluster services are up and running and that all nodes joined the cluster and are stable as shown in Example 6-28.

*Example 6-28 Checking cluster status*

```
root@clmgr1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
migrRG         ONLINE                       migr01
               OFFLINE                      migr02

root@clmgr1 / # lssrc -ls clstrmgrES
Current state: ST_STABLE
sccsid = "@(#)36    1.135.1.107 src/43haes/usr/sbin/cluster/hacmprd/main.C,
hacmp.pe, 61haes_r711, 1150E_hacmp711 1/31/12 15:48:17"
build = "Feb  6 2012 16:41:07 1150F_hacmp711"
i_local_nodeid 0, i_local_siteid -1, my_handle 2
.........
```

```
.........
local node vrmf is 7112
cluster fix level is "2"

root@clmgr1 / # lspv | grep caa
hdisk3          00f74d4733c9370e                    caavg_private   active

root@clmgr1 / # lsvg -o
tst_vg1
caavg_private
```

We successfully restored the cluster configuration.

### 6.2.5 Rolling migration from PowerHA 6.1

During a rolling migration, you perform a series of operations on each node, one node at a time. Therefore, for a node, you first stop cluster services with the takeover or with the move resource groups option. Then, you upgrade the operating system and the cluster software. Then, you reintegrate the node into the cluster, which brings its resource groups back in the initial status.

Rolling migration helps you when the main goal is minimum application downtime. You experience a short application interruption when the resource group is moved to a peer node. Also, a second short interruption is experienced when the resource group moves back to the home node after this node is migrated.

Rolling migration is allowed only to the base version of PowerHA 7.1.1. After the whole cluster is upgraded to the base version of PowerHA 7.1.1, the latest Service Pack can be applied on each node without closing the applications. During the update to the latest Service Pack, the applications and resources continue to run on the node, although they are not highly available during the update. For more details, see 4.2, "AIX and PowerHA SystemMirror service pack upgrade" on page 150. You can also apply the Service Pack in a rolling approach, one node at a time. We describe the steps of a rolling migration example from PowerHA 6.1 to PowerHA 7.1.1 for a 3-node cluster.

#### Test environment

The nodes are VIOS client LPARs on different frames that share SAN LUNs through NPIV. The backing devices are LUNs that are carved in a DS4800 storage subsystem. The `net_ether_01` is a single adapter logical network and the application is available on the `migrsvc2` service IP label. Each node has two virtual Ethernet adapters in network `net_ether_02`, all are in the same VLAN, which is distinct from the VLAN of `net_ether_01`. Example 6-29 shows the cluster topology.

*Example 6-29 Three-node initial cluster topology*

```
root@migr3 / # cllsif
Adapter              Type       Network    Net Type   Attribute  Node       IP
Address       Hardware Address Interface Name   Global Name      Netmask
Alias for HB Prefix Length

migr3_hdisk3         service    net_dskhb_34 diskhb     serial     migr3
/dev/hdisk3                    hdisk3
migr3_hdisk7         service    net_dskhb_35 diskhb     serial     migr3
/dev/hdisk7                    hdisk7
```

```
migr3              boot      net_ether_01 ether     public     migr3
10.1.1.32                   en0                    255.255.254.0
23
migrsvc2           service   net_ether_01 ether     public     migr3
172.16.21.68                                       255.255.254.0
23
migr3b3            boot      net_ether_02 ether     public     migr3
10.1.3.32                   en2                    255.255.255.0
24
migr3b2            boot      net_ether_02 ether     public     migr3
10.1.2.32                   en1                    255.255.255.0
24
migrsvc4           service   net_ether_02 ether     public     migr3
10.10.20.68                                        255.255.255.0
24
migr4_hdisk3       service   net_dskhb_34 diskhb    serial     migr4
/dev/hdisk3                  hdisk3
migr4_hdisk6       service   net_dskhb_45 diskhb    serial     migr4
/dev/hdisk6                  hdisk6
migr4              boot      net_ether_01 ether     public     migr4
10.1.1.40                   en0                    255.255.254.0
23
migrsvc2           service   net_ether_01 ether     public     migr4
172.16.21.68                                       255.255.254.0
23
migr4b2            boot      net_ether_02 ether     public     migr4
10.1.2.40                   en1                    255.255.255.0
24
migr4b3            boot      net_ether_02 ether     public     migr4
10.1.3.40                   en2                    255.255.255.0
24
migrsvc4           service   net_ether_02 ether     public     migr4
10.10.20.68                                        255.255.255.0
24
migr5_hdisk7       service   net_dskhb_35 diskhb    serial     migr5
/dev/hdisk7                  hdisk7
migr5_hdisk6       service   net_dskhb_45 diskhb    serial     migr5
/dev/hdisk6                  hdisk6
migr5              boot      net_ether_01 ether     public     migr5
10.1.1.72                   en0                    255.255.254.0
23
migrsvc2           service   net_ether_01 ether     public     migr5
172.16.21.68                                       255.255.254.0
23
migr5b2            boot      net_ether_02 ether     public     migr5
10.1.2.72                   en1                    255.255.255.0
24
migr5b3            boot      net_ether_02 ether     public     migr5
10.1.3.72                   en2                    255.255.255.0
24
migrsvc4           service   net_ether_02 ether     public     migr5
10.10.20.68                                        255.255.255.0
24
root@migr3 / #
```

Example 6-30 shows the entries in the `/etc/hosts` file, which is the same on all nodes.

*Example 6-30 Entries in /etc/hosts*

```
root@migr3 / # cat /etc/hosts|grep -v ^#

127.0.0.1              loopback localhost      # loopback (lo0) name/address
172.16.20.40    nimres1
10.1.1.32       migr3 migr3b1
10.1.1.40       migr4 migr4b1
10.1.1.72       migr5 migr5b1
10.1.2.32       migr3b2
10.1.2.40       migr4b2
10.1.2.72       migr5b2
10.1.3.32       migr3b3
10.1.3.40       migr4b3
10.1.3.72       migr5b3
172.16.21.32    migr3p1 lpar0701
172.16.21.40    migr4p1 lpar0702
172.16.21.75    migr5p1 lpar1003
172.16.21.68    migrsvc2
10.10.20.68     migrsvc4
root@migr3 / #
```

The cluster contains two resource groups: `ihs_rg` and `tstipat_rg`. The `ihs_rg` resource group hosts an IBM HTTP Server application. The `tstipat_rg` resource group hosts a dummy test script application. Both resource groups use IPAT via Aliases for the service IP. The home node for `ihs_rg` is node `migr3`. The `migr4` node is the home node for `tstipat_rg`. The `migr5` node is in a standby role. As shown in Figure 6-31, the Startup Policy and Fallback Policy attributes of the `ihs_rg` resource group are intentionally changed from their defaults to “Online On First Available Node” and “Never Fallback”.

```
Resource Group Name                                   ihs_rg
Participating Node Name(s)                            migr3 migr5 migr4
Startup Policy                                        Online On First Available Node
Fallover Policy                                       Fallover To Next Priority Node
Fallback Policy                                       Never Fallback
Service IP Label                                      migrsvc2
Volume Groups                                         ihsvg
Application Servers                                   ihsas

Resource Group Name                                   tstipat_rg
Participating Node Name(s)                            migr4 migr5 migr3
Startup Policy                                        Online On Home Node Only
Fallover Policy                                       Fallover To Next Priority Node
Fallback Policy                                       Fallback To Higher Priority Node
Service IP Label                                      migrsvc4
```

*Figure 6-31 Resource group attributes*

Figure 6-32 on page 296 shows the AIX and cluster initial software versions and the status of the running cluster and its resource groups.

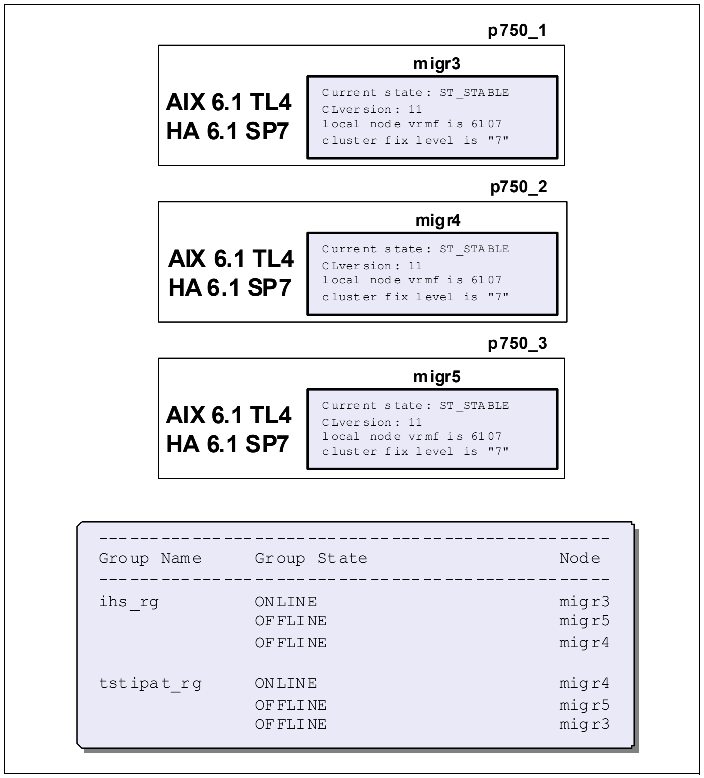



*Figure 6-32 Three-node cluster status and software versions before migration*

## Planning the rolling migration

Begin the migration planning by reviewing 6.1, "Considerations before you migrate to PowerHA 7.1.1" on page 256. Then, continue with a careful analysis of the cluster to avoid starting a migration from an unsupported configuration. For more details, see 6.2.1, "Requirements" on page 257. Downtime must be scheduled to convert this configuration to a supported one.

Any set of actions that involves modifications at the operating system or cluster level must be preceded by back-out (reversion) planning that includes the following operations:

- Back up all data and binaries of the applications.
- Take a cluster snapshot and save it locally and also to another machine.
- Save a copy of any custom script files locally and also to another machine.
- Perform a **mksysb** backup of each involved node.

A back-out plan allows easy restoration of the application data, binaries, cluster, and AIX configuration if you encounter problems. We assume that you start a rolling migration and you get `CONFIG-ERROR` messages while you run the **clmigcheck** command on the first node. See For an example of this message, see Figure 6-33 on page 297.

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-ERROR: The configuration contains unsupported options:  IP Address
Takeover via Replacement.  The PowerHA network name is net_ether_02.

This will have to be removed from the configuration before
 migration to PowerHA System Mirror
```

*Figure 6-33 clmigcheck error for IPAT via Replacement*

Then, you have two choices. If the downtime is acceptable, try to reconfigure the cluster to eliminate the errors and then continue the rolling migration. Otherwise, execute the back-out plan and then reevaluate the entire migration planning and analysis by considering the errors that you encountered.

## Preparing the rolling migration

After you finish the analysis, take the following steps:

- Ensure that all nodes are at the same level of the operating system and cluster software. Check that the cluster software is committed (and not merely applied). Use the **oslevel -s** and **lslpp -L cluster*** commands. Verify that the software is consistent on all nodes by using the **lppchk -v** command.
- Decide on the shared disk to be used as a cluster repository and coordinate with the network administrator on the multicast communication as described in “Considerations before you migrate to PowerHA 7.1.1” on page 256.

Before you start the actual migration, ensure that the cluster has no pending changes on any of its nodes or configuration errors since the last cluster startup. Also, check that all nodes joined the cluster and are in a stable state. Follow these steps:

1. Ensure that the cluster has no pending changes on any of its nodes.

   Run **odmget HACMPcluster | grep handle** on all nodes (Example 6-31).

   *Example 6-31 Checking pending changes on cluster nodes*

   ```
   root@migr3 / # odmget HACMPcluster | grep handle
           handle = 1
   root@migr3 / #

   root@migr4 / #odmget HACMPcluster | grep handle
           handle = 2
   root@migr4 / #

   root@migr5 / #odmget HACMPcluster | grep handle
           handle = 3
   root@migr5 / #
   ```

   A non-zero value for the handle parameter on each node means that there are no pending changes so you can go to the next step. A zero value for the handle parameter on a node means that changes are pending on that node.

   If changes are pending on one node and you choose to apply them on top of the current configuration, check them, decide on a final configuration, and run a Verification and Synchronization operation. In an active cluster, certain changes might not be allowed. If you really need these changes, you must stop the cluster services.

If you decide to cancel any pending changes on any node and to keep the currently active configuration, run on either node **smit hacmp** and select **Problem Determination Tools → Restore HACMP Configuration Database from Active Configuration**. Select the **Verification and Synchronization** operation on the same node.

This way, you get a clean and homogeneous configuration on all nodes. Any pending changes on other nodes are canceled by the final successful synchronization.

2. Check that the cluster has no configuration errors.

   Run the Verify Cluster Configuration operation on either cluster node by following the path **smitty hacmp → Problem Determination Tools → HACMP Verification → Verify HACMP Configuration**. Use **No** for the "Verify changes only?" field and press Enter. The expected results are shown in Figure 6-34.

```
COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...164]
   ihs_app                           ihs_rg
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.

[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure 6-34 Verification ends with an OK message on a cluster node*

3. Check that all nodes joined the cluster and are in a stable state (Example 6-32).

*Example 6-32 Checking that the cluster is stable*

```
root@migr3 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 6107
cluster fix level is "7"
root@migr3 / #

root@migr4 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 6107
```

```
cluster fix level is "7"
root@migr4 / #

root@migr5 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 6107
cluster fix level is "7"
root@migr5 / #
```

### Migrating the first node

To migrate the first node, follow these steps:

1. Stop the cluster services on the first node (`migr3` in our scenario).

   Use the Move Resource Groups option in the `Stop Cluster Services` (**`smitty clstop`**) panel (Figure 6-35).

```
 Stop Cluster Services

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

                                                        [Entry Fields]
* Stop now, on system restart or both                 now +
  Stop Cluster Services on these nodes               [migr3] +
  BROADCAST cluster shutdown?                         false +
* Select an Action on Resource Groups                 Move Resource Groups +

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

*Figure 6-35 Stop cluster services with the Move Resource Groups option*

Check that the resource group moved on one of the remaining active nodes and that the cluster is stable (Example 6-33).

*Example 6-33 Cluster state after the resource group movement*

```
root@migr5 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         OFFLINE                      migr3
               ONLINE                       migr5
               OFFLINE                      migr4

tstipat_rg     ONLINE                       migr4
               OFFLINE                      migr3
               OFFLINE                      migr5

root@migr5 / # lssrc -ls clstrmgrES|grep state
Current state: ST_STABLE
root@migr5 / #
```

Check that the cluster services are successfully stopped on the first node (Example 6-34). Look for the ST_INIT status.

*Example 6-34 Cluster services are stopped on the first node*

```
root@migr3 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr3 / #
```

2. Upgrade AIX on the first node to version 7.1 TL1 SP2 or later.

   Use a supported procedure to upgrade the operating system. Then, install the following filesets if they are not already installed:

   – `bos.cluster` and `bos.ahafs`
   – `bos.clvm.enh`

> **Mandatory filesets:** You must install the CAA-specific `bos.cluster` and `bos.ahafs`. Another mandatory fileset is `bos.clvm.enh`. Even though it is not a prerequisite for the PowerHA 7.1. installation, this fileset is needed for Enhanced Concurrent Mode volume groups, which is the only supported mode in PowerHA 7.1.

   Ensure that you reboot the node and make the final checks (Example 6-35).

*Example 6-35 Checking the AIX upgrade*

```
root@migr3 / # uptime
  09:05PM   up 3 mins,  2 users,  load average: 0.49, 0.54, 0.25
root@migr3 / # oslevel -s
7100-01-03-1207
root@migr3 / # lppchk -v
root@migr3 / # lslpp -L bos.cluster\* bos.ahafs bos.clvm.enh|grep -p Level
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  bos.ahafs                  7.1.1.2    C     F    Aha File System
  bos.cluster.rte            7.1.1.3    C     F    Cluster Aware AIX
  bos.cluster.solid          7.1.1.0    C     F    Cluster Aware AIX SolidDB
  bos.clvm.enh               7.1.1.0    C     F    Enhanced Concurrent Logical
                                                   Volume Manager
root@migr3 / #
```

3. Populate the `/etc/cluster/rhosts` file with all cluster IP addresses. You must include the addresses that resolve to the host name of the cluster nodes (Example 6-36).

*Example 6-36 Adding all cluster IPs in /etc/cluster/rhosts*

```
root@migr3 / # cat /etc/cluster/rhosts
10.1.1.32
10.1.1.40
10.1.1.72
10.1.2.32
10.1.2.40
10.1.2.72
10.1.3.32
10.1.3.40
10.1.3.72
172.16.21.32
172.16.21.40
```

```
172.16.21.75
172.16.21.68
10.10.20.68
root@migr3 / #

root@migr3 / # host `hostname`
migr3 is 10.1.1.32
root@migr3 / #

root@migr4 / # host `hostname`
migr4 is 10.1.1.40
root@migr4 / #

root@migr5 / # host `hostname`
migr5 is 10.1.1.72
root@migr5 / #
```

After you change `/etc/cluster/rhosts`, you must refresh **clcomd** (Example 6-37).

*Example 6-37 Refreshing clcomd*

```
root@migr3 / # refresh -s clcomd
0513-095 The request for subsystem refresh was completed successfully.
root@migr3 / #
```

4. Run the **clmigcheck** command on the first node.

   You see the initial PowerHA System Mirror Migration Check (**clmigcheck**) panel as shown in Figure 6-36.

```
------------[ PowerHA System Mirror Migration Check ]-------------

Please select one of the following options:

1       = Check ODM configuration.

2       = Check snapshot configuration.

3       = Enter reopsitory disk and multicast IP addresses.


Select one of the above,"x"to exit or "h" for help:
```

*Figure 6-36 clmigcheck initial panel*

   Follow these steps:

   a. On Figure 6-36, select option **1** (Check ODM configuration).

      When you choose this option, the **clmigcheck** command checks your configuration and reports any problems that might affect the migration. Our migration scenario uses three different point-to-point disk-based heartbeat paths, one for each pair of nodes.

      The **clmigcheck** command detects a situation and warns you that this configuration is removed during migration (Figure 6-37 on page 302).

```
------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_dskhb_34.  This will
be removed from the configuration during the migration to PowerHA System
Mirror 7.1.

Hit <Enter> to continue


------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_dskhb_35.  This will
be removed from the configuration during the migration to PowerHA System
Mirror 7.1.

Hit <Enter> to continue


------------[ PowerHA System Mirror Migration Check ]-------------

CONFIG-WARNING: The configuration contains unsupported hardware:  Disk
Heartbeat network.  The PowerHA network name is net_dskhb_45.  This will
be removed from the configuration during the migration to PowerHA System
Mirror 7.1.

Hit <Enter> to continue
```

*Figure 6-37 clmigcheck heartbeat disk warning*

You do not need to act because the disk-based heartbeating is automatically removed during migration.

If `CONFIG-ERROR` messages are displayed and if the downtime is acceptable, try to reconfigure the cluster to eliminate the error. Then, repeat the **`clmigcheck`** verification. Otherwise, execute the back-out plan and redo the entire analysis and consider the encountered error.

If no errors are detected, you see the output in Figure 6-38.

```
------------[ PowerHA System Mirror Migration Check ]-------------

The ODM has no unsupported elements.

Hit <Enter> to continue
```

*Figure 6-38 Migration-compliant configuration message*

Press Enter after this last panel, and return to the main menu.

b. Choose option **3** to enter the repository disk and then select the repository disk that you decided to use (Figure 6-39 on page 303).

```
------------[ PowerHA System Mirror Migration Check ]-------------

Select the disk to use for the repository

        1       = 00f61ab22b9e1f5c(hdisk2)

Select one of the above or "x" to exit:
```

*Figure 6-39 Selecting the repository disk*

c. If you agreed on a multicast IP address with your network manager, use it. Otherwise, let AIX generate an appropriate one (Figure 6-40). Whatever you choose, press Enter to get to the next window.

```
------------[ PowerHA System Mirror Migration Check ]-------------

PowerHA System Mirror uses multicast address for internal cluster
communication and monitoring. These must be in the multicast range,
224.0.0.0 - 239.255.255.255.

If you make a NULL entry, AIX will generate an appropriate address for
you. You should only specify an address if you have an explicit reason to
do so, but are cautioned that this address cannot be changed once the
configuration is activated (i.e. migration is complete).


h = help

Enter the multicast IP address to use for network monitoring:
```

*Figure 6-40 Specifying the multicast IP address*

d. The next window is the initial menu where you choose **x** to exit the **clmigcheck** tool (Example 6-38).

*Example 6-38 PowerHA SystemMirror migration check*

```
------------[ PowerHA System Mirror Migration Check ]-------------

Please select one of the following options:

1       = Check ODM configuration.

2       = Check snapshot configuration.

3       = Enter reopsitory disk and multicast IP addresses.


Select one of the above,"x"to exit or "h" for help:
```

e. Verify whether you are ready for the PowerHA upgrade on the first node by running the **clmigcheck** tool again. If you are ready, you see the panel that is shown in Figure 6-41 on page 304.

```
------------[ PowerHA System Mirror Migration Check ]-------------

clmigcheck: This is not the first node or last node clmigcheck was run
on. No further checking is required on this node.  You can install the
new version of PowerHA System Mirror.


Hit <Enter> to continue
```

*Figure 6-41 Confirmation that PowerHA filesets can be upgraded*

You can see more details about the actions that are performed by the **clmigcheck** command that is repeatedly invoked throughout this step in the log file that it creates on the first node, /tmp/clmigcheck/clmigcheck.log.

5. Upgrade PowerHA on the first node to PowerHA 7.1.1 base release.

   Perform an update_all installation by using the base PowerHA 7.1.1 filesets. Do not apply any full service pack or individual fix on top of the base filesets until all nodes in the cluster are upgraded to the new base release. The only exception is the application of interim fixes that are supplied specifically for the base filesets. This application is allowed.

   Example 6-39 shows the PowerHA filesets that are updated to the base 7.1.1 release. Run the commands **lppchk -v**, **lppchk -c "cluster.*"**, and **lslpp -L cluster*** to verify that the upgrade is successful, as shown in Example 6-39.

*Example 6-39 Checking the PowerHA software consistency after the upgrade*

```
root@migr3 / # lppchk -v
root@migr3 / # lppchk -c "cluster.*"
root@migr3 / # lslpp -L cluster\*|grep -p Level
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.adt.es.client.include
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   CLINFO Samples
  cluster.adt.es.client.samples.clstat
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Clstat Samples
  cluster.adt.es.client.samples.libcl
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   LIBCL Samples
  cluster.adt.es.java.demo.monitor
                             7.1.1.0    C     F    Web Based Monitor Demo
  cluster.es.client.clcomd   7.1.1.0    C     F    Cluster Communication
                                                   Infrastructure
  cluster.es.client.lib      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Libraries
  cluster.es.client.rte      7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Runtime
  cluster.es.client.utils    7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Utilities
  cluster.es.client.wsm      7.1.1.0    C     F    Web based Smit
```

```
  cluster.es.cspoc.cmds      7.1.1.0    C     F    CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0    C     F    CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.0    C     F    CSPOC Runtime Commands
  cluster.es.migcheck        7.1.1.0    C     F    PowerHA SystemMirror
Migration
                                                   support
  cluster.es.server.cfgast   7.1.1.0    C     F    Two-Node Configuration
                                                   Assistant
  cluster.es.server.diag     7.1.1.0    C     F    Server Diags
  cluster.es.server.events   7.1.1.0    C     F    Server Events
  cluster.es.server.rte      7.1.1.0    C     F    Base Server Runtime
  cluster.es.server.testtool
                             7.1.1.0    C     F    Cluster Test Tool
  cluster.es.server.utils    7.1.1.0    C     F    Server Utilities
  cluster.license            7.1.1.0    C     F    PowerHA SystemMirror
                                                   Electronic License
  cluster.man.en_US.es.data  7.1.1.0    C     F    Man Pages - U.S. English
  cluster.msg.en_US.cspoc    6.1.0.0    C     F    HACMP CSPOC Messages - U.S.
                                                   English
  cluster.msg.en_US.es.client
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Messages - U.S. English
  cluster.msg.en_US.es.server
                             7.1.1.0    C     F    Recovery Driver Messages -
                                                   U.S. English
```

The update process converts the existing local cluster configuration to be compatible with the new software. The conversion details are logged in the `/tmp/clconvert.log` file. Check this log file to confirm that the conversion succeeded.

6. Start cluster services on this first node by issuing the **`smitty clstart`** command.

   You receive a warning message about mixed versions when the node reintegrates into the cluster (Figure 6-42 on page 306).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Cluster services are running at different levels across the cluster.
Verification will not be invoked in this environment.

Starting Cluster Services on node: migr3
This may take a few minutes.  Please wait...
migr3: start_cluster: Starting PowerHA SystemMirror
migr3:  4325534      -  0:00 syslogd
migr3: Setting routerevalidate to 1
migr3: 0513-059 The topsvcs Subsystem has been started. Subsystem PID is 733410
migr3: 0513-059 The grpsvcs Subsystem has been started. Subsystem PID is 812694
[MORE...26]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-42 Starting cluster services on the first node*

7. Check that the cluster is stable and runs the new software (Example 6-40).

*Example 6-40 Checking status and new software version on the first node*

```
root@migr3 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 7110
cluster fix level is "0"
root@migr3 / #
```

8. Move the resource groups back to the host node (`migr3`) before you start the migration as shown in Figure 6-43 on page 307.

```
Move Resource Group(s) to Another Node

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

                                                          [Entry Fields]
  Resource Group(s) to be Moved                            ihs_rg
  Destination Node                                         migr3


F1=Help                F2=Refresh          F3=Cancel            F4=List
Esc+5=Reset            Esc+6=Command       Esc+7=Edit           Esc+8=Image
Esc+9=Shell            Esc+0=Exit          Enter=Do
```

*Figure 6-43 Moving the resource group*

9. Check that the cluster is stable and the resource groups are online as expected (Example 6-41).

*Example 6-41 Checking the state after the resource group movement*

```
root@migr3 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ONLINE                       migr3
               OFFLINE                      migr5
               OFFLINE                      migr4

tstipat_rg     ONLINE                       migr4
               OFFLINE                      migr5
               OFFLINE                      migr3

root@migr3 / # lssrc -ls clstrmgrES|grep state
Current state: ST_STABLE
root@migr3 / #
```

## Migrating an intermediate node

We present the migration procedure for the second node of our 3-node cluster. If your cluster has more than three nodes, this procedure is repeated for each remaining node except the last one. We call them *intermediate* nodes. If you have a 2-node cluster skip this process, and go to “Migrating the last node” on page 310. Here are the steps for an intermediate node:

1. Stop the cluster services (our only intermediate node is `migr4`).

   Use the move resource group option in the **`smitty clstop`** panel. You can also move the resource group on the second node to another node and then stop the cluster services. Check that the resource group moved onto the next node that is available in the priority list and that the cluster is stable (Example 6-42).

*Example 6-42 Checking the cluster state after resource group movement*

```
root@migr5 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
```

```
-----------------------------------------------------------------------------
ihs_rg         ONLINE                     migr3
               OFFLINE                    migr5
               OFFLINE                    migr4

tstipat_rg     OFFLINE                    migr4
               ONLINE                     migr5
               OFFLINE                    migr3

root@migr5 / #lssrc -ls clstrmgrES|grep state
Current state: ST_STABLE
root@migr5 / #
```

Check that the cluster services are successfully stopped on the current migrated node (Example 6-43). Look for the ST_INIT status.

*Example 6-43 Cluster services are stopped on the intermediate node*

```
root@migr4 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr4 / #
```

2. Upgrade AIX on this intermediate node similarly to the process that you used for the first node. Ensure that you installed the required filesets and that AIX is restarted (Example 6-44).

*Example 6-44 Checking the AIX upgrade*

```
root@migr4 / # uptime
  12:05AM   up 1 min,  1 user,  load average: 2.04, 0.51, 0.18
root@migr4 / # oslevel -s
7100-01-03-1207
root@migr4 / # lppchk -v
root@migr4 / # lslpp -L bos.cluster\* bos.ahafs bos.clvm.enh|grep -p Level
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  bos.ahafs                  7.1.1.2    C     F    Aha File System
  bos.cluster.rte            7.1.1.3    C     F    Cluster Aware AIX
  bos.cluster.solid          7.1.1.0    C     F    Cluster Aware AIX SolidDB
  bos.clvm.enh               7.1.1.0    C     F    Enhanced Concurrent Logical
                                                   Volume Manager
```

3. Populate the `/etc/cluster/rhosts` file so that it is the same as on the first node that you upgraded. You can copy it from the first node. Refresh the clcomd daemon. See Example 6-45.

*Example 6-45 Checking the rhost file and refreshing the clcomd daemon*

```
root@migr4 / # cat /etc/cluster/rhosts
10.1.1.32
10.1.1.40
10.1.1.72
10.1.2.32
10.1.2.40
10.1.2.72
10.1.3.32
10.1.3.40
```

```
10.1.3.72
172.16.21.32
172.16.21.40
172.16.21.75
172.16.21.68
10.10.20.68
root@migr4 / # refresh -s clcomd
0513-095 The request for subsystem refresh was completed successfully.
root@migr4 / #
```

4. Run the **clmigcheck** command on the intermediate node (Example 6-46). This step is important to ensure that you can proceed with the PowerHA upgrade.

*Example 6-46 Running clmigcheck on an intermediate node*

```
root@migr4 / # clmigcheck

------------[ PowerHA System Mirror Migration Check ]-------------

clmigcheck: This is not the first node or last node clmigcheck was run on.
No further checking is required on this node.  You can install the new
version of PowerHA System Mirror.


Hit <Enter> to continue
root@migr4 / #
```

5. Upgrade PowerHA on the intermediate node to PowerHA 7.1.1 base release.

   Follow the same steps as on the first node to upgrade PowerHA filesets to version 7.1.1.

6. Start the cluster services on the intermediate node.

   Issue the **smitty clstart** command and then check that the cluster is stable and runs the new software (Example 6-47).

*Example 6-47 Checking status and new software version on the intermediate node*

```
root@migr4 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 7110
cluster fix level is "0"
root@migr4 / #
```

Move the resource groups back to the intermediate node before you start the migration (Example 6-48). Depending on the fallback policy, the resource group might move automatically on the host node.

*Example 6-48 Checking the resource group that moved back on the intermediate node*

```
root@migr4 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ONLINE                       migr3
               OFFLINE                      migr5
               OFFLINE                      migr4
```

```
tstipat_rg     ONLINE                       migr4
               OFFLINE                      migr5
               OFFLINE                      migr3

root@migr4 / #
```

By performing this task, you finished the migration for an intermediate node. Use the same steps for any other remaining node except the last one if your cluster has more than three nodes.

You are ready to proceed with the migration of the final node in the cluster.

### Migrating the last node

To migrate the last node (`migr5` in our scenario), follow these steps:

1. Stop cluster services on the last node.

   If it hosts an online resource group, use the Move resource group option in the **`smitty clstop`** panel. You can also move the resource group to another node and then stop the cluster services. Check that the resource group moved onto one of the remaining active nodes and that the cluster is stable.

   In our case, the last node does not have a resource group. We check that the cluster services are successfully stopped on the last node only (Example 6-49).

   *Example 6-49 Checking that the last node cluster is stopped*

   ```
   root@migr5 / # lssrc -ls clstrmgrES|grep state
   Current state: ST_INIT
   root@migr5 / #
   ```

2. Upgrade AIX on the last node similar to the process that you used for the previous nodes. Ensure that you installed the required filesets and that AIX is restarted (Example 6-50).

   *Example 6-50 Checking the AIX upgrade on the last node*

   ```
   root@migr5 / # uptime
     02:12AM   up 1 min,  2 users,  load average: 1.30, 0.36, 0.13
   root@migr5 / # oslevel -s
   7100-01-03-1207
   root@migr5 / # lppchk -v
   root@migr5 / # lslpp -L bos.cluster\* bos.ahafs bos.clvm.enh|grep -p Level
     Fileset                      Level  State  Type  Description (Uninstaller)
     ----------------------------------------------------------------------------
     bos.ahafs                  7.1.1.2    C     F    Aha File System
     bos.cluster.rte            7.1.1.3    C     F    Cluster Aware AIX
     bos.cluster.solid          7.1.1.0    C     F    Cluster Aware AIX SolidDB
     bos.clvm.enh               7.1.1.0    C     F    Enhanced Concurrent Logical
                                                      Volume Manager
   ```

3. Populate the `/etc/cluster/rhosts` file so that it is the same as on the previous nodes that you upgraded. You can copy the file from one of the other nodes. Refresh clcomd (each time that `/etc/cluster/rhosts` is changed, clcomd must be refreshed).

4. Then, run the **`clmigcheck`** command for the last time.

   When the **`clmigcheck`** command runs, it recognizes that this node is the last node of the cluster to migrate and it starts to create the CAA cluster. You see the message that is shown in Example 6-51 on page 311.

*Example 6-51 Running clmigcheck on the last node*

```
root@migr5 / # clmigcheck
Verifying clcomd communication, please be patient.

Verifying multicast IP communication, please be patient.

mping version 1.1
Localhost is migr3, 10.1.1.32
Listening on 228.168.101.43/4098:

Replying to mping from 10.1.1.72 bytes=32 seqno=2 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=3 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=4 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=5 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=1 ttl=32
Discarding receiver packet
mping version 1.1
Localhost is migr4, 10.1.1.40
Listening on 228.168.101.43/4098:

Replying to mping from 10.1.1.72 bytes=32 seqno=2 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=3 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=4 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=5 ttl=32
Discarding receiver packet
Replying to mping from 10.1.1.72 bytes=32 seqno=1 ttl=32
Discarding receiver packet
clmigcheck: Running
/usr/sbin/rsct/install/bin/ct_caa_set_disabled_for_migration on each node in
the cluster

Creating CAA cluster, please be patient.

------------[ PowerHA System Mirror Migration Check ]-------------

You can install the new version of PowerHA System Mirror.


Hit <Enter> to continue
root@migr5 / #
```

Verify that the CAA infrastructure is started by running the **lscluster** command with various options as shown in Example 6-52.

*Example 6-52 Checking the CAA cluster state*

```
root@migr3 / # lscluster -c
Cluster query for cluster clmigr345 returns:
Cluster uuid: 6f16435a-9cfd-11e1-8322-b6fcca1bda6f
```

```
Number of nodes in cluster = 3
        Cluster id for node migr3 is 1
        Primary IP address for node migr3 is 10.1.1.32
        Cluster id for node migr4 is 2
        Primary IP address for node migr4 is 10.1.1.40
        Cluster id for node migr5 is 3
        Primary IP address for node migr5 is 10.1.1.72
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.72
root@migr3 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 3


        Node name: migr3
        Cluster shorthand id for node: 1
        uuid for node: 6edddb0a-9cfd-11e1-8322-b6fcca1bda6f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr345          local        6f16435a-9cfd-11e1-8322-b6fcca1bda6f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: migr4
        Cluster shorthand id for node: 2
        uuid for node: 6edde21c-9cfd-11e1-8322-b6fcca1bda6f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr345          local        6f16435a-9cfd-11e1-8322-b6fcca1bda6f

        Number of points_of_contact for node: 4
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en2  UP
         en1  UP
         en0  UP

------------------------------

        Node name: migr5
        Cluster shorthand id for node: 3
        uuid for node: 6edde91a-9cfd-11e1-8322-b6fcca1bda6f
        State of node:  UP
```

```
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr345          local        6f16435a-9cfd-11e1-8322-b6fcca1bda6f

        Number of points_of_contact for node: 4
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en0  UP
         en2  UP
         en1  UP
root@migr3 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  clmigr345
Cluster uuid:  6f16435a-9cfd-11e1-8322-b6fcca1bda6f
Number of nodes reporting = 3
Number of nodes expected = 3
Node migr3
Node uuid = 6edddb0a-9cfd-11e1-8322-b6fcca1bda6f
Number of interfaces discovered = 4
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.d6.15.e0.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 3
                IPV4 ADDRESS: 10.1.1.32  broadcast 10.1.1.255  netmask
255.255.254.0
                IPV4 ADDRESS: 172.16.21.32  broadcast 172.16.21.255  netmask
255.255.254.0
                IPV4 ADDRESS: 172.16.21.68  broadcast 172.16.21.255  netmask
255.255.254.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 2 en1
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.d6.15.e0.70
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
```

```
                IPV4 ADDRESS: 10.1.2.32  broadcast 10.1.2.255  netmask
255.255.255.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 3 en2
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.d6.15.e0.71
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 10.1.3.32  broadcast 10.1.3.255  netmask
255.255.255.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 4 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN
Node migr5
Node uuid = 6edde91a-9cfd-11e1-8322-b6fcca1bda6f
Number of interfaces discovered = 5
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.ca.1b.da.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.1.1.72  broadcast 10.1.1.255  netmask
255.255.254.0
                IPV4 ADDRESS: 172.16.21.75  broadcast 172.16.21.255  netmask
255.255.254.0
               Number of cluster multicast addresses configured on interface =
1
```

```
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 2 en1
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.ca.1b.da.70
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 10.1.2.72  broadcast 10.1.2.255  netmask
255.255.255.0
                Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 3 en2
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.ca.1b.da.71
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 10.1.3.72  broadcast 10.1.3.255  netmask
255.255.255.0
                Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 4 sfwcom
                ifnet type = 0 ndd type = 304
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 0
                Mean Deviation in network rrt across interface = 0
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP
        Interface number 5 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
```

```
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN
Node migr4
Node uuid = 6edde21c-9cfd-11e1-8322-b6fcca1bda6f
Number of interfaces discovered = 5
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  2e.47.92.d5.43.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.1.1.40  broadcast 10.1.1.255  netmask 
255.255.254.0
                IPV4 ADDRESS: 172.16.21.40  broadcast 172.16.21.255  netmask 
255.255.254.0
               Number of cluster multicast addresses configured on interface = 
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask 
0.0.0.0
        Interface number 2 en1
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  2e.47.92.d5.43.70
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.10.20.68  broadcast 10.10.20.255  netmask 
255.255.255.0
                IPV4 ADDRESS: 10.1.2.40  broadcast 10.1.2.255  netmask 
255.255.255.0
               Number of cluster multicast addresses configured on interface = 
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask 
0.0.0.0
        Interface number 3 en2
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  2e.47.92.d5.43.71
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
```

```
                IPV4 ADDRESS: 10.1.3.40  broadcast 10.1.3.255  netmask
255.255.255.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.1.1.72  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 4 sfwcom
                ifnet type = 0 ndd type = 304
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 0
                Mean Deviation in network rrt across interface = 0
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP
        Interface number 5 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN
root@migr3 / # lscluster -d
Storage Interface Query

Cluster Name:  clmigr345
Cluster uuid:  6f16435a-9cfd-11e1-8322-b6fcca1bda6f
Number of nodes reporting = 3
Number of nodes expected = 3
Node migr3
Node uuid = 6edddb0a-9cfd-11e1-8322-b6fcca1bda6f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 6037ba18-552b-67dd-284c-74ca9bcf55f6
          type  : REPDISK
Node migr5
Node uuid = 6edde91a-9cfd-11e1-8322-b6fcca1bda6f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 6037ba18-552b-67dd-284c-74ca9bcf55f6
          type  : REPDISK
Node migr4
Node uuid = 6edde21c-9cfd-11e1-8322-b6fcca1bda6f
Number of disk discovered = 1
        hdisk2
```

```
         state : UP
         uDid  :
         uUid  : 6037ba18-552b-67dd-284c-74ca9bcf55f6
         type  : REPDISK
root@migr3 / #
```

5. Upgrade PowerHA on the last node to PowerHA 7.1.1 base release.

   Follow the same steps that are used on the previous nodes.

6. Start the cluster services on the last node by issuing the **smitty clstart** command. Figure 6-44 shows the output of the command.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Cluster services are running at different levels across
the cluster.  Verification will not be invoked in this environment.

Starting Cluster Services on node: migr5
This may take a few minutes.  Please wait...
migr5: start_cluster: Starting PowerHA SystemMirror
migr5:  4980894      -  0:00 syslogd
migr5: Setting routerevalidate to 1
migr5: 0513-059 The topsvcs Subsystem has been started. Subsystem PID is
8192032.
migr5: 0513-059 The grpsvcs Subsystem has been started. Subsystem PID is
8650788.
migr5: 0513-059 The emsvcs Subsystem has been started. Subsystem PID is
9306112.
[MORE...17]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-44   Starting PowerHA cluster services on the last node*

Check that the cluster is stable and runs the new PowerHA version (Example 6-53).

*Example 6-53   Checking the state and software version on the last node*

```
root@migr5 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_BARRIER
CLversion: 11
local node vrmf is 7110
cluster fix level is "0"
root@migr5 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 11
local node vrmf is 7110
cluster fix level is "0"
```

```
root@migr5 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7110
cluster fix level is "0"
root@migr5 / #

root@migr3 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7110
cluster fix level is "0"
root@migr3 / #

root@migr4 / # lssrc -ls clstrmgrES|grep -e state -e version -e vrmf -e fix
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7110
cluster fix level is "0"
root@migr4 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_rg         ONLINE                       migr3
               OFFLINE                      migr5
               OFFLINE                      migr4

tstipat_rg     ONLINE                       migr4
               OFFLINE                      migr5
               OFFLINE                      migr3

root@migr4 / #
```

Move back any resource group hosted by node `migr5` before starting the migration. In our case, there was no resource group hosted by the `migr5` node.

Run a verification of your migrated cluster to ensure that it operates with no problems. Use the command **smit cl_sync** with default options. The output for a successful run of this command is shown in Figure 6-45 on page 320.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...140]
Cluster Manager Current state: ST_BARRIER
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_CBARRIER
Cluster Manager Current state: ST_UNSTABLE
Cluster Manager Current state: ST_UNSTABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
...completed.

[BOTTOM]
F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-45 Verifying the migrated cluster*

By performing this task, you finished the rolling migration for the entire cluster. The cluster migration is now complete. Figure 6-46 shows how the migrated cluster looks now.

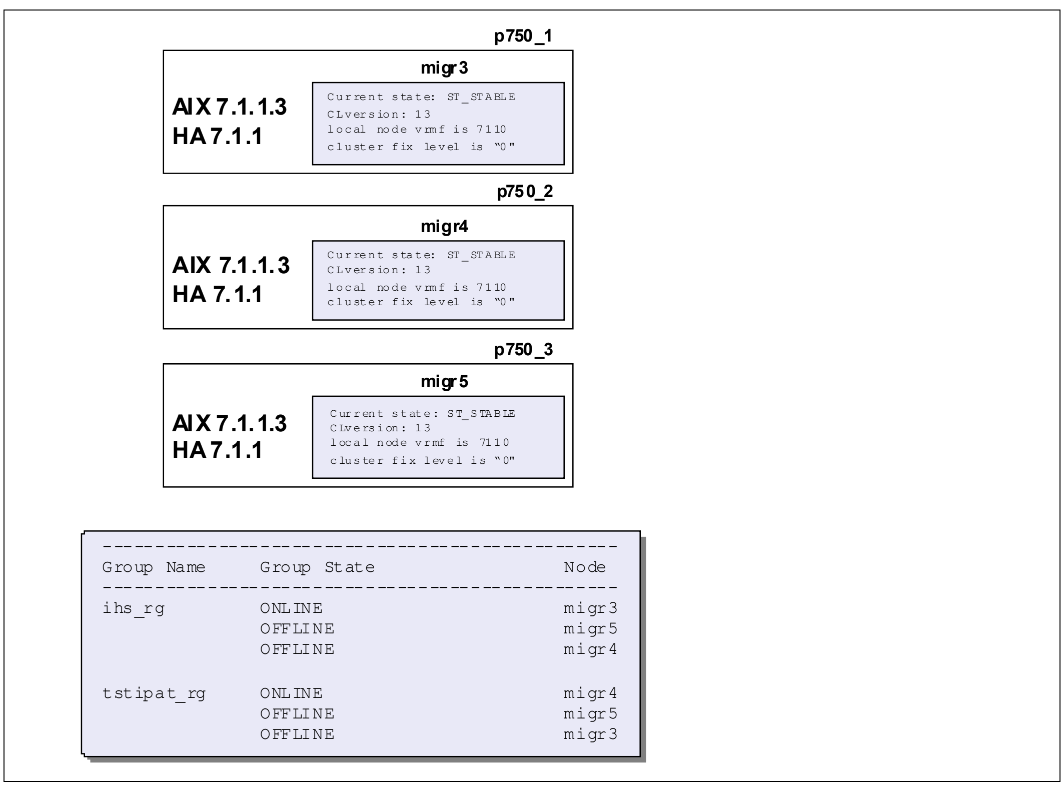



*Figure 6-46 Cluster status and software versions after migration*

## 6.3 Migration from PowerHA 7.1.0

When you migrate a PowerHA 7.1.0 cluster that runs on top of AIX 7.1 TL00 or AIX 6.1 TL06 to PowerHA 7.1.1, you can use either the offline migration method or the snapshot migration method. Rolling migration is not supported in this case.

A PowerHA cluster outage is required to recreate the underlying CAA cluster. A predefined AIX 7.1 TL00 or AIX 6.1 TL06 CAA cluster cannot be migrated to an AIX 7.1 TL01 or AIX 6.1 TL07 CAA cluster. You must first remove the old CAA cluster, and then upgrade the AIX operating system on all involved nodes. Then, you recreate the CAA cluster by using a PowerHA command. This way, you obtain a new cluster instance, by using AIX 7.1 TL01 or AIX 6.1 TL07 code, while you keep the configuration parameters of the previous CAA cluster.

As a preliminary step, apply the latest updates (service packs) that are available for the running versions of AIX (AIX 7.1 TL00 or AIX 6.1 TL06) and PowerHA SystemMirror 7.1.0.

To remove the old CAA cluster instance, you must run the **rmcluster -n <*clustername*>** command as root after you stop all PowerHA cluster services in all nodes.

In the offline migration case, you create the CAA cluster instance after you upgrade both the AIX and the PowerHA filesets. Use the **clmkcaa** PowerHA utility script for this purpose. This script takes the appropriate parameters from the PowerHA configuration files and passes them to the CAA **mkcluster** command. You do not use the **clmigcheck** command.

In the snapshot migration case, after you save a cluster snapshot as a first step, you completely uninstall the PowerHA 7.1.0 filesets and then remove the CAA cluster. The next step is to install the new filesets of the PowerHA 7.1.1. With the new code installed, you convert the saved snapshot to PowerHA version 7.1.1 and restore the cluster configuration from the converted snapshot. A new CAA cluster is automatically created during the restoration process.

> **Demonstration:** A demonstration of a 7.1.0 to 7.1.1 snapshot migration is available at this website:
>
> http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS4941

A related scenario is possible when the nodes of a PowerHA SystemMirror 7.1.0 cluster must be upgraded from AIX 7.1 TL00 to AIX 7.1 TL01 (or AIX 6.1 TL06 to AIX 6.1 TL07) while you keep the cluster at version 7.1.0. The steps to perform this upgrade are documented in APAR IV08163:

http://www-01.ibm.com/support/docview.wss?uid=isg1IV08163

Appendix A, “AIX upgrades and PowerHA SystemMirror 7.1.0” on page 531 presents the results of running the steps of this scenario in our test environment.

> **Demonstration:** A demonstration of a CAA update is at this website:
>
> http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS4895

### 6.3.1 Offline migration from PowerHA 7.1.0 version

With the offline migration method, you can migrate the cluster definitions, individually on each node, by updating the PowerHA filesets. The ODM files on the node that runs the update process are converted to the new version.

## Test environment

In our scenario, we begin with PowerHA 7.1.0.1 on top of AIX 7.1.0.1. The target versions are PowerHA 7.1.1.2 and AIX 7.1.1.3, which are the latest available versions at the time of writing this book. The nodes are VIOS client LPARs on different frames that share the SAN LUNs through NPIV. The backing devices are LUNs that are carved in a DS4800 storage subsystem. The network topology is simple, with only one interface per node, which is typical for a virtualized environment. The initial cluster configuration is illustrated in Figure 6-47.

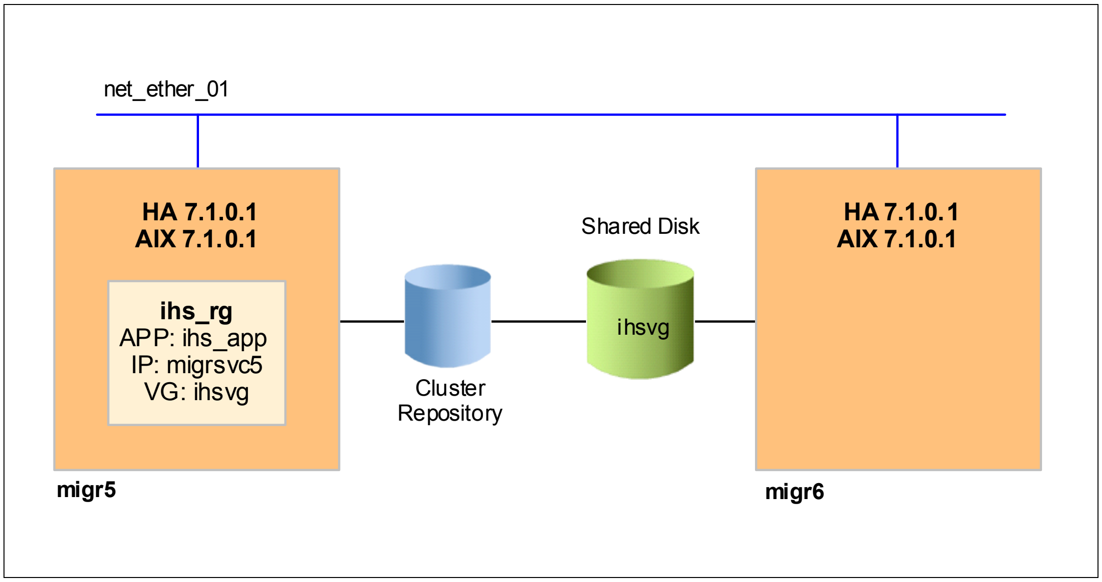



*Figure 6-47 Initial cluster configuration*

You can see the topology details of our cluster in Example 6-54.

*Example 6-54 Topology details*

```
root@migr5 / # cllsif
Adapter              Type       Network     Net Type   Attribute  Node       IP
Address       Hardware Address Interface Name   Global Name      Netmask
Alias for HB Prefix Length

migr5                boot       net_ether_01 ether      public     migr5
172.16.21.75                   en0                      255.255.254.0
23
migrsvc5             service    net_ether_01 ether      public     migr5
172.16.21.77                                            255.255.254.0
23
migr6                boot       net_ether_01 ether      public     migr6
172.16.21.76                   en0                      255.255.254.0
23
migrsvc5             service    net_ether_01 ether      public     migr6
172.16.21.77                                            255.255.254.0
23
root@migr5 / #
```

The cluster contains one resource group, `ihs_rg`, which hosts an IBM HTTP Server application. The resource group uses IPAT via Alias for the service IP. The home node for `ihs_rg` is node `migr5`. Node `migr6` is in a standby role. Figure 6-48 on page 323 shows more details about the configured resource group.

```
Resource Group Name                  ihs_rg
Participating Node Name(s)           migr5 migr6
Startup Policy                       Online On First Available Node
Fallover Policy                      Fallover To Next Priority Node In The List
Fallback Policy                      Never Fallback
Service IP Label                     migrsvc5
Volume Groups                        ihsvg
Application Servers                  ihs_app
```

*Figure 6-48 Resource group details*

### Preparing the offline migration from PowerHA 7.1.0

Any set of actions that involves modifications at the operating system or cluster level must be preceded by a back-out (reversion) planning, including mandatory backup operations:

- Back up all data and binaries of the applications.
- Take a cluster snapshot and save it locally and to another machine.
- Save a copy of any custom script files locally and to another machine.
- Perform a **mksysb** backup of each involved node.

This back-out plan allows easy restoration of the application data, binaries, cluster configuration, and AIX configuration if you encounter into problems.

Perform the usual software consistency checks:

- Ensure that all nodes are at the same level of operating system and cluster software. Check that the cluster software is committed (and not merely applied). Use the **oslevel -s** and the **lslpp -L cluster*** commands.
- Verify that the software is consistent on all nodes by using the **lppchk -v** command.

Before you start the actual migration, check the overall status of your cluster:

1. Ensure that the cluster is started on all nodes, is stable, and in a normal operating state. Run the **clmgr view report status** command on either node and check the state of the cluster, resource groups, interfaces, and System Resource Controller (SRC) subsystems (Example 6-55).

   *Example 6-55 Cluster state before migration*

```
root@migr5 / # clmgr view report status

Obtaining information via SNMP from Node: migr5...

_____________________________________________________________________________
Cluster Name: clmigr56
Cluster State: UP
Cluster Substate: STABLE
_____________________________________________________________________________

Node Name: migr5                State: UP

  Network Name: net_ether_01       State: UP

    Address: 172.16.21.75    Label: migr5             State: UP
    Address: 172.16.21.77    Label: migrsvc5          State: UP
```

```
Node Name: migr6                State: UP

  Network Name: net_ether_01       State: UP

    Address: 172.16.21.76    Label: migr6              State: UP


Cluster Name: clmigr56

Resource Group Name: ihs_rg
Startup Policy: Online On First Available Node
Fallover Policy: Fallover To Next Priority Node In The List
Fallback Policy: Never Fallback
Site Policy: ignore
Node                         Group State
---------------------------- ---------------
migr5                        ONLINE
migr6                        OFFLINE


Status of the RSCT subsystems used by HACMP:
Subsystem         Group            PID          Status
 cthags           cthags           16384048     active
 ctrmc            rsct             5439686      active

Status of the HACMP subsystems:
Subsystem         Group            PID          Status
 clstrmgrES       cluster          15663252     active
 clcomd           caa              7471316      active

Status of the optional HACMP subsystems:
Subsystem         Group            PID          Status
 clinfoES         cluster          11534382     active
root@migr5 / #
```

You might get the equivalent status information by using other Simple Network Management Protocol (SNMP)-based commands such as **cldump** or **clstat -o** that are combined with **clshowsrv -v**. Also, instead of SNMP-based commands, you can use **lssrc -ls clstrmgrES**, **clRGinfo**, and **netstat -in**.

2. Ensure that the cluster has no pending configuration changes on any of its nodes.

   Use the command **clcmd odmget HACMPcluster | egrep "NODE|handle"** (Example 6-56).

   *Example 6-56 Checking pending configuration changes*

```
root@migr5 / # clcmd odmget HACMPcluster | egrep "NODE|handle"
NODE migr6
        handle = 2
NODE migr5
        handle = 1
root@migr5 / #
```

   A non-zero value for the handle parameter on each node means that there are no pending changes so you can go to the next step.

   A zero value for the handle parameter on a node means that changes are pending on that node. If changes are pending on one node and you choose to apply them on top of the current configuration, check them, decide on a final configuration, and run a “Verify and Synchronize Cluster Configuration” operation. In an active cluster, there are some changes that might not be allowed. If you really need these changes, you have to stop the cluster services.

If you decide to cancel any pending changes on any node and to keep the currently active configuration, on either node, run **smit sysmirror**. Select **Problem Determination Tools → Restore PowerHA SystemMirror Configuration Database from Active Configuration**. Select **Verify and Synchronize Cluster Configuration** on the same node. You might avoid the last synchronization by restoring the default cluster configuration from the active configuration on each of the cluster nodes.

3. Check that the cluster has no configuration errors.

   Select **Verify Cluster Configuration** on either cluster node by starting **smitty sysmirror** and selecting **Problem Determination Tools → PowerHA SystemMirror Verification → Verify Cluster Configuration**. Use `No` for the “Verify changes only?” field and press Enter. The result is shown in Figure 6-49.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...164]
   ihs_app                           ihs_rg
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.

[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure 6-49 Verifying cluster configuration*

4. Confirm that the CAA cluster state is error-free (Example 6-57).

*Example 6-57 Checking CAA cluster state*

```
root@migr5 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  clmigr56
Cluster uuid:  435ee888-7913-11e1-a13d-b6fcca1bda70
Number of nodes reporting = 2
Number of nodes expected = 2
```

```
Node migr5
Node uuid = 436337ee-7913-11e1-a13d-b6fcca1bda70
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.ca.1b.da.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 172.16.21.75  broadcast 172.16.23.255  netmask
255.255.252.0
                IPV4 ADDRESS: 172.16.21.77  broadcast 172.16.21.255  netmask
255.255.254.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.16.21.75  broadcast 0.0.0.0
netmask 0.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
Node migr6
Node uuid = 656d55e0-7913-11e1-8d55-2e479566c670
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  2e.47.95.66.c6.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 172.16.21.76  broadcast 172.16.23.255  netmask
255.255.252.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.16.21.75  broadcast 0.0.0.0
netmask 0.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
```

```
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
root@migr5 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 436337ee-7913-11e1-a13d-b6fcca1bda70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 656d55e0-7913-11e1-8d55-2e479566c670
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 1
        Point-of-contact interface & contact state
         en0  UP
root@migr5 / #

root@migr6 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 436337ee-7913-11e1-a13d-b6fcca1bda70
        State of node:  UP
        Smoothed rtt to node: 7
```

```
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 1
        Point-of-contact interface & contact state
         en0  UP

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 656d55e0-7913-11e1-8d55-2e479566c670
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a
root@migr6 / #
```

### Performing the offline migration from PowerHA 7.1.0

To perform the offline migration, follow these steps:

1. Ensure that all nodes in the cluster run the same and most recent version of the AIX and PowerHA SystemMirror software.

   Identify the latest updates (service packs) that are available for the current versions of AIX and PowerHA SystemMirror on your cluster nodes. Apply these service packs (SPs) to eliminate any known problem that might affect the migration.

   > **PowerHA level and SP:** Consider the following procedure to discover the PowerHA level and SP that are installed (the **halevel** command has a man page):
   >
   > ```
   > # /usr/es/sbin/cluster/utilities/halevel
   > 7.1.0
   > # /usr/es/sbin/cluster/utilities/halevel -s
   > 7.1.0 SP5
   > ```

   In our test environment, we start from AIX 7.1 TL00 SP01 and PowerHA SystemMirror 7.1.0 SP01 (Example 6-57).

   *Example 6-58 Initial cluster state and software versions*

   ```
   root@migr5 / # oslevel -s
   7100-00-01-1037
   root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
   Current state: ST_STABLE
   CLversion: 12
   ```

```
local node vrmf is 7101
cluster fix level is "1"
root@migr5 / #

root@migr6 / # oslevel -s
7100-00-01-1037
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7101
cluster fix level is "1"
root@migr6 / #
```

We update to the latest levels that are available at the time of writing this book: AIX 7.1 TL00 SP04 and PowerHA SystemMirror 7.1.0 SP05. Then, we ensure that everything is OK after the update by running again a "Verify and Synchronize" operation, followed by a cluster startup and status check (Example 6-59).

*Example 6-59   Cluster state and software versions after we apply the latest service packs*

```
root@migr5 / # oslevel -s
7100-00-04-1140
root@migr5 / # lppchk -v
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr5 / #

root@migr5 / # oslevel -s
7100-00-04-1140
root@migr5 / # lppchk -v
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr5 / #
```

2. Stop the cluster services on all nodes.

   Use the **smitty clstop** command. Select all nodes and choose **Bring Resource Groups Offline** as an action on resource groups (Figure 6-50 on page 330).

```
 Stop Cluster Services

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

                                                        [Entry Fields]
* Stop now, on system restart or both                 now                    +
  Stop Cluster Services on these nodes               [migr5,migr6]           +
  BROADCAST cluster shutdown?                         true                   +
* Select an Action on Resource Groups                 Bring Resource Groups> +

F1=Help             F2=Refresh          F3=Cancel           F4=List
Esc+5=Reset         Esc+6=Command       Esc+7=Edit          Esc+8=Image
Esc+9=Shell         Esc+0=Exit          Enter=Do
```

*Figure 6-50 Stopping cluster services on both nodes*

Ensure that the PowerHA cluster services are stopped on all nodes (Example 6-60).

*Example 6-60 Checking that the cluster is stopped*

```
root@migr5 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr5 / #

root@migr6 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr6 / #
```

3. Remove CAA.

   Use **lscluster** and **lspv** to obtain information about the cluster name and repository disk PVID (Example 6-61).

*Example 6-61 CAA cluster name and repository disk information*

```
root@migr5 / # lscluster -d
Storage Interface Query

Cluster Name:  clmigr56
Cluster uuid:  b450137e-83f8-11e1-9636-b6fcca1bda70
Number of nodes reporting = 2
Number of nodes expected = 2
Node migr5
Node uuid = 8e041058-797b-11e1-8b14-b6fcca1bda70
Number of disk discovered = 1
        caa_private0
          state : UP
          uDid  :
          uUid  : acffc231-fb7e-bb0d-b3f2-08eebe201a81
          type  : REPDISK
Node migr6
Node uuid = ae4b0e20-797b-11e1-9063-2e479566c670
Number of disk discovered = 1
        caa_private0
          state : UP
```

```
          uDid  :
          uUid  : acffc231-fb7e-bb0d-b3f2-08eebe201a81
          type  : REPDISK
root@migr5 / # lspv
caa_private0    00f74d47797a84d6                    caavg_private   active
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
hdisk0          00f74d473088fca1                    rootvg          active
hdisk1          00f74d4731ff9753                    rootvg          active
root@migr5 / #
```

Use the command **rmcluster -n *<clustername>*** on either node to remove the CAA cluster. Then, check that it is successfully removed with the **lscluster -m** command on both nodes (Example 6-62).

*Example 6-62 Removing the CAA cluster*

```
root@migr6 / # rmcluster -n clmigr56
rmcluster: Removed cluster shared disks are automatically renamed to names such
        as hdisk10, [hdisk11, ...] on all cluster nodes.  However, this cannot
        take place while a disk is busy or on a node which is down or not
        reachable.  If any disks cannot be renamed now, you must manually
        rename them.
root@migr6 / # lscluster -m
Cluster services are not active.
root@migr6 / #

root@migr5 / # lscluster -m
Cluster services are not active.
root@migr5 / #
```

The previous repository disk, with `00f74d47797a84d6` as the PVID, is now clean of any CAA LVM structure on any node. In our case, it is renamed to `hdisk2` (Example 6-63).

*Example 6-63 Checking that the previous repository disk is clean*

```
root@migr5 / # lspv
hdisk2          00f74d47797a84d6                    None
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
hdisk0          00f74d473088fca1                    rootvg          active
hdisk1          00f74d4731ff9753                    rootvg          active
root@migr5 / #

root@migr6 / # lspv
hdisk2          00f74d47797a84d6                    None
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
hdisk0          00f74d45501b2428                    rootvg          active
hdisk1          00f74d455084775a                    rootvg          active
root@migr6 / #
```

**Repository disk PVID:** If problems appear, call IBM support. You must not try any action that might change the PVID of the repository disk. The PVID is required later to recreate the CAA cluster with the same repository disk.

4. Upgrade the AIX and PowerHA SystemMirror software on all nodes in the cluster.

   Upgrade the AIX to version 7.1 TL01 SP2 or later by using a supported procedure. In our scenario, we upgrade to the latest level that is available at the time of this writing, AIX 7.1 TL01 SP03. Ensure that you rebooted the systems after the upgrade (Example 6-64).

*Example 6-64 Checking AIX upgrade*

```
root@migr5 / # uptime
  11:16AM   up   1:38,  1 user,  load average: 1.27, 1.24, 1.17
root@migr5 / # oslevel -s
7100-01-03-1207
root@migr5 / # lppchk -v
root@migr5 / #

root@migr6 / # uptime
  11:16AM   up   1:38,  1 user,  load average: 1.19, 1.26, 1.16
root@migr6 / # oslevel -s
7100-01-03-1207
root@migr6 / # lppchk -v
root@migr6 / #
```

> **AIX Upgrade:** Use a supported upgrade procedure and appropriate software sources (download or media); otherwise, you might overwrite configuration files. For example, if you use AIX 7.1 TL01 base media to update from AIX7.1 TL00, you might overwrite the `/etc/cluster/rhosts` file.

Proceed with the upgrade of the PowerHA filesets to version 7.1.1 or later. In our scenario, we also apply the SP02 Service Pack for PowerHA 7.1.1, which is the latest available SP at the time of developing this publication (Example 6-65).

*Example 6-65 Checking PowerHA filesets upgrade*

```
root@migr5 / # lslpp -L cluster\*
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.adt.es.client.include
                             7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   CLINFO Samples
  cluster.adt.es.client.samples.clstat
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Clstat Samples
  cluster.adt.es.client.samples.libcl
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   LIBCL Samples
  cluster.adt.es.java.demo.monitor
                             7.1.1.0    C     F    Web Based Monitor Demo
  cluster.es.client.clcomd   7.1.1.0    C     F    Cluster Communication
                                                   Infrastructure
  cluster.es.client.lib      7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Libraries
  cluster.es.client.rte      7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Runtime
```

```
  cluster.es.client.utils    7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Utilities
  cluster.es.client.wsm      7.1.1.0    C     F    Web based Smit
  cluster.es.cspoc.cmds      7.1.1.2    C     F    CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0    C     F    CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.2    C     F    CSPOC Runtime Commands
  cluster.es.migcheck        7.1.1.0    C     F    PowerHA SystemMirror
Migration
                                                   support
  cluster.es.server.cfgast   7.1.1.0    C     F    Two-Node Configuration
                                                   Assistant
  cluster.es.server.diag     7.1.1.1    C     F    Server Diags
  cluster.es.server.events   7.1.1.1    C     F    Server Events
  cluster.es.server.rte      7.1.1.1    C     F    Base Server Runtime
  cluster.es.server.testtool
                             7.1.1.0    C     F    Cluster Test Tool
  cluster.es.server.utils    7.1.1.1    C     F    Server Utilities
  cluster.license            7.1.1.0    C     F    PowerHA SystemMirror
                                                   Electronic License
  cluster.man.en_US.es.data  7.1.1.0    C     F    Man Pages - U.S. English
  cluster.msg.en_US.es.client
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Messages - U.S. English
  cluster.msg.en_US.es.server
                             7.1.1.1    C     F    Recovery Driver Messages -
                                                   U.S. English
root@migr5 / # lppchk -v
```

Review the `/tmp/clconvert.log` file to ensure that the conversion of the PowerHA ODMs finished successfully.

5. Create the CAA cluster.

   Check that the `/etc/cluster/rhosts` and clcomd are in a good state (Example 6-66).

*Example 6-66 Checking the clcomd communication*

```
root@migr5 / # clrsh migr5 hostname
migr5
root@migr5 / # clrsh migr6 hostname
migr6
root@migr5 / #

root@migr6 / # clrsh migr5 hostname
migr5
root@migr6 / # clrsh migr6 hostname
migr6
root@migr6 / #
```

Run the **/usr/es/sbin/cluster/utilities/clmkcaa** command (Example 6-67).

*Example 6-67 Recreating CAA cluster*

```
root@migr5 / # /usr/es/sbin/cluster/utilities/clmkcaa
Verifying clcomd communication, please be patient.

Creating CAA cluster, please wait.
```

```
CLUSTER_OVERRIDE=yes /usr/sbin/mkcluster -n clmigr56 -r hdisk2 -m
migr5{cle_globid=1},migr6{cle_globid=2} -s 228.16.21.75

clmkcaa: lscluster output:

Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 18b1ebe2-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 18e38bc0-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP
        Smoothed rtt to node: 134
        Mean Deviation in network rtt to node: 189
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 2
        Point-of-contact interface & contact state
         dpcom  UP  RESTRICTED
         en0  UP
root@migr5 / #
```

**Important:** Now, it is not possible to recreate the CAA cluster by using a "Verify and Synchronize" operation. We are in the middle of the migration because the ODM files are converted during the upgrade of the PowerHA filesets.

Verify that the CAA services are active on each node by running the **/usr/sbin/lscluster -m** command (Example 6-68).

*Example 6-68 Verifying the new CAA cluster state*

```
root@migr5 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2
```

```
        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 18b1ebe2-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 18e38bc0-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 2
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en0  UP
root@migr5 / #

root@migr6 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 18b1ebe2-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 2
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
```

```
         en0  UP

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 18e38bc0-850e-11e1-93e3-b6fcca1bda70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        18e87f40-850e-11e1-93e3-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a
root@migr6 / #
```

Verify that the RSCT services are active on each node by running the **lssrc -s cthags** command (Example 6-69).

*Example 6-69 Verifying the cthags state*

```
root@migr5 / # lssrc -s cthags
Subsystem         Group            PID          Status
 cthags           cthags           7209186      active
root@migr5 / #

rroot@migr6 / # lssrc -s cthags
Subsystem         Group            PID          Status
 cthags           cthags           6029330      active
root@migr6 / #
```

6. Start the PowerHA cluster services, one node at a time.

   Use **smitty clstart** on the first node (Figure 6-51 on page 337).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Cluster services are running at different levels across
the cluster.  Verification will not be invoked in this environment.

Starting Cluster Services on node: migr5
This may take a few minutes.  Please wait...
migr5: start_cluster: Starting PowerHA SystemMirror
migr5:  4849664      -  0:00 syslogd
migr5: Setting routerevalidate to 1
migr5: 0513-059 The topsvcs Subsystem has been started. Subsystem PID is
8454278
.
migr5: 0513-059 The grpsvcs Subsystem has been started. Subsystem PID is
8519904
[MORE...24]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure 6-51 Cluster starts OK on the first node*

Then, ensure that the node successfully joins the cluster (Example 6-70). Although the software is updated at the 7.1.1 level, the node still runs at version 12.

*Example 6-70 First node joined the cluster*

```
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7112
cluster fix level is "2"
root@migr5 / #
```

Repeat the procedure for each node of the cluster, one node at a time.

After you start the cluster services on the latest node and it joins the cluster, you can check the cluster version update (Example 6-71).

*Example 6-71 Last node joins the cluster*

```
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_JOINING
CLversion: 12
local node vrmf is 0
cluster fix level is "ffffffff"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_BARRIER
```

```
CLversion: 12
local node vrmf is 7112
cluster fix level is "2"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7112
cluster fix level is "2"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7112
cluster fix level is "2"
root@migr6 / #

root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7112
cluster fix level is "2"
root@migr5 / #
```

7. Select a final **Verify and Synchronize Cluster Configuration** to ensure that the cluster runs error-free (Figure 6-52).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...193]
Cluster Manager Current state: ST_BARRIER
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_CBARRIER
Cluster Manager Current state: ST_UNSTABLE
Cluster Manager Current state: ST_UNSTABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
...completed.


[BOTTOM]

F1=Help              F2=Refresh           F3=Cancel            Esc+6=Command
Esc+8=Image          Esc+9=Shell          Esc+0=Exit           /=Find
n=Find Next
```

*Figure 6-52 Final verification of the migrated cluster*

## 6.3.2 Snapshot migration from 7.1.0 version

With the snapshot migration, we preserve the entire PowerHA SystemMirror 7.1.0 cluster configuration by creating a snapshot of the configuration. Here, we uninstall the PowerHA 7.1.0 after we take the snapshot. The CAA cluster must be removed before we upgrade AIX. The snapshot is later used after the PowerHA SystemMirror 7.1.1 is installed in the cluster.

### Test environment

In our scenario, we begin with PowerHA 7.1.0.5 on top of AIX 7.1.0.1. The target versions are PowerHA 7.1.1.2 and AIX 7.1.1.3, which are the latest version available at the time of this writing. The nodes are VIOS client LPARs on different frames that share the SAN LUNs from the storage device DS4800 through NPIV. The initial cluster configuration is illustrated in Figure 6-53.

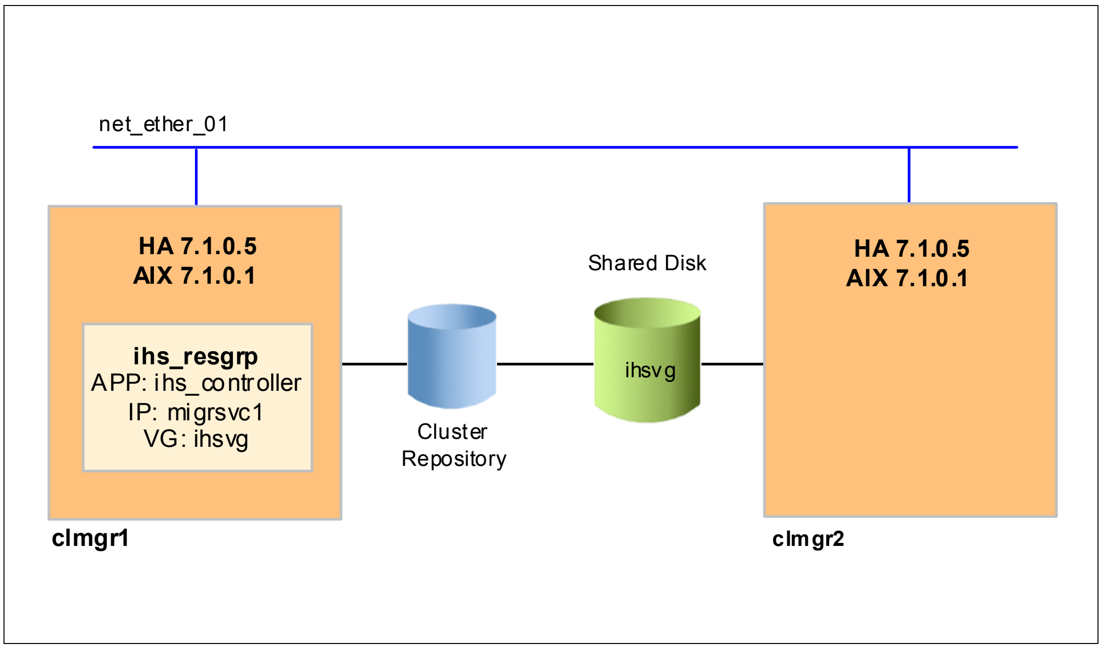



*Figure 6-53 Initial cluster configuration*

You can see the topology of our cluster in Example 6-72.

*Example 6-72 Topology details*

```
root@clmgr1 / # cllsif
Adapter                Type       Network    Net Type   Attribute  Node       IP
Address        Hardware Address Interface Name   Global Name       Netmask
Alias for HB Prefix Length

clmgr1                 boot       net_ether_01 ether      public     clmgr1
10.1.1.54                       en0                       255.255.254.0
23
clmgrsvc1              service    net_ether_01 ether      public     clmgr1
172.16.21.71                                              255.255.254.0
23
clmgr2                 boot       net_ether_01 ether      public     clmgr2
10.1.1.55                       en0                       255.255.254.0
23
```

```
clmgrsvc1           service    net_ether_01 ether      public     clmgr2
172.16.21.71                                           255.255.254.0
23
root@clmgr1 / #
```

### PowerHA SystemMirror 7.1.0 cluster configuration

The cluster contains one resource group `ihs_resgrp`, which hosts the IBM HTTP Server application. The resource uses IPAT via IP Alias for the service IP. The home node for `ihs_resgrp` is `clmgr1`. The `clmgr2` node has a standby role. Figure 6-54 displays the existing configuration of PowerHA SystemMirror 7.1.0, which is planned for snapshot migration.

```
Cluster Name: ihs_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: caa_private0
Cluster IP Address: 228.1.1.36
There are 2 node(s) and 1 network(s) defined
NODE clmgr1:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr1  10.1.1.54
NODE clmgr2:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr2  10.1.1.55

Resource Group ihs_resgrp
        Startup Policy   Online On Home Node Only
        Fallover Policy  Fallover To Next Priority Node In The List
        Fallback Policy  Never Fallback
        Participating Nodes      clmgr1 clmgr2
        Service IP Label                 clmgrsvc1
```

*Figure 6-54 Existing cluster configuration*

### Preparing the snapshot migration from PowerHA 7.1.0

Take a snapshot of the cluster configuration from the existing cluster environment. The snapshot still needs to be restored on a new version of PowerHA. To avoid any issues during the restoration process and also to have a revert back solution, ensure that you take a backup of all application data and binaries and any custom scripts that are used in the cluster configuration and a **mksysb** backup of all nodes.

Before you create the snapshot of the existing cluster configuration, verify the following information:

1. Check whether the cluster status is stable as shown in Example 6-73.

   *Example 6-73 Cluster state before migration*

```
root@clmgr1 / # lssrc -ls clstrmgrES | grep state ; clRGinfo
Current state: ST_STABLE
-----------------------------------------------------------------------------
```

```
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_resgrp     ONLINE                       clmgr1
               OFFLINE                      clmgr2
```

2. Check whether the cluster has any pending changes on any node as shown in Example 6-74.

*Example 6-74 Checking pending configuration changes*

```
root@clmgr1 / # clcmd odmget HACMPcluster | egrep "NODE|handle"
NODE clmgr2
        handle = 2
NODE clmgr1
        handle = 1
```

If both nodes display a non-zero value for the handle parameter as shown in Example 6-74, the cluster does not have pending changes on any node.

3. Check whether the cluster has any configuration errors.

Run **Verify Cluster Configuration** on either cluster node by starting **smitty sysmirror** and selecting **Problem Determination Tools** → **PowerHA SystemMirror Verification** → **Verify Cluster Configuration**. Use No for the "Verify changes only?" field and press Enter. The result is similar to Figure 6-55.

```
 COMMAND STATUS

Command: OK            stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

[MORE...150]
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.

[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-55 Verifying cluster configuration*

4. Check that the CAA cluster is in an error-free state as shown in Example 6-75 on page 342.

*Example 6-75 Checking CAA cluster state*

```
root@clmgr1 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  ihs_cluster
Cluster uuid:  0b24cf30-8ffe-11e1-ad2a-b6fcc07d1d70
Number of nodes reporting = 2
Number of nodes expected = 2
Node clmgr1
Node uuid = c0269ec0-8f74-11e1-9fff-b6fcc07d1d70
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.c0.7d.1d.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 3
                IPV4 ADDRESS: 10.1.1.54  broadcast 10.1.1.255  netmask
255.255.2
54.0
                IPV4 ADDRESS: 172.16.21.54  broadcast 172.16.21.255  netmask
255
.255.254.0
                IPV4 ADDRESS: 172.16.21.71  broadcast 172.16.21.255  netmask
255
.255.254.0
   Numberofclustermulticastaddressesconfiguredoninterface=
1
                IPV4 MULTICAST ADDRESS: 228.1.1.36  broadcast 0.0.0.0  netmask
0
.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
Node clmgr2
Node uuid = 91c788b8-8f75-11e1-8992-6e8dde31d770
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.de.31.d7.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
```

```
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.1.1.55  broadcast 10.1.1.255  netmask
255.255.2
54.0
                IPV4 ADDRESS: 172.16.21.55  broadcast 172.16.21.255  netmask
255
.255.254.0
   Numberofclustermulticastaddressesconfiguredoninterface=
1
                IPV4 MULTICAST ADDRESS: 228.1.1.36  broadcast 0.0.0.0  netmask
0
.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
root@clmgr1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: clmgr1
        Cluster shorthand id for node: 1
        uuid for node: c0269ec0-8f74-11e1-9fff-b6fcc07d1d70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        ihs_cluster        local        0b24cf30-8ffe-11e1-ad2a-b6fcc07d1d70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: clmgr2
        Cluster shorthand id for node: 2
        uuid for node: 91c788b8-8f75-11e1-8992-6e8dde31d770
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
```

```
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        ihs_cluster        local        0b24cf30-8ffe-11e1-ad2a-b6fcc07d1d70

        Number of points_of_contact for node: 1
        Point-of-contact interface & contact state
         en0  UP
root@clmgr1 / #
```

5. Ensure that all nodes in the cluster have the same version of AIX and PowerHA SystemMirror with the latest service packs to eliminate any known problem that might affect the migration.

   The AIX software and PowerHA SystemMirror software that are used in the test environment are shown in Example 6-76. Both products have the latest levels of the software at the time of writing this publication.

*Example 6-76   Cluster state and software version with the latest service package*

```
root@clmgr1 / # oslevel -s
7100-00-04-1140
root@clmgr1 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@clmgr1 / #

root@clmgr2 / # oslevel -s
7100-00-04-1140
root@clmgr2 / #  lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@clmgr2 / #
```

### Creating a Snapshot on PowerHA SystemMirror 7.1.0

After you verify the cluster status and existing cluster configuration, you can proceed to create a snapshot of the cluster configuration by using **smit sysmirror**. Select **Cluster Nodes and Networks** → **Manage the cluster** → **Snapshot Configuration** → **Create a Snapshot of the Cluster Configuration**.

Enter the snapshot name and description as shown in Figure 6-56 on page 345.

```
 Create a Snapshot of the Cluster Configuration

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

                                                        [Entry Fields]
* Cluster Snapshot Name                              [710Snapshot]          /
  Custom-Defined Snapshot Methods                    []                     +
  Save Cluster Log Files in snapshot                  No                    +
* Cluster Snapshot Description                       [710 Snapshot Migration]



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

*Figure 6-56 Creating a snapshot*

The snapshot creates two files, `710Snapshot.odm` and `710Snapshot.info`, as shown in Figure 6-57. Keep a backup copy of these snapshot files and preserve it until the migration is completed successfully. We use these files after PowerHA SystemMirror 7.1.1 is installed on the cluster. Stop the cluster services on all nodes.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.


clsnapshot: Creating file /usr/es/sbin/cluster/snapshots/710Snapshot.odm.

clsnapshot: Creating file /usr/es/sbin/cluster/snapshots/710Snapshot.info.

clsnapshot: Executing clsnapshotinfo command on node: clmgr2...

clsnapshot: Executing clsnapshotinfo command on node: clmgr1...

clsnapshot: Succeeded creating Cluster Snapshot: 710Snapshot



F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-57 Creation of snapshot files*

## Removal of the CAA cluster

Use the **lscluster** and the **lspv** commands to get information about the cluster name and repository disk PVID as shown in Example 6-77.

*Example 6-77 CAA cluster information details*

```
root@clmgr1 /opt/IBM# lscluster -d
Storage Interface Query

Cluster Name:  ihs_cluster
Cluster uuid:  2355f8de-904b-11e1-a908-b6fcc07d1d70
Number of nodes reporting = 2
Number of nodes expected = 2
Node clmgr1
Node uuid = c0269ec0-8f74-11e1-9fff-b6fcc07d1d70
Number of disk discovered = 1
        caa_private0
          state : UP
          uDid  :
          uUid  : 39597db5-d0b8-5741-3e06-095cd9da92e7
          type  : REPDISK
Node clmgr2
Node uuid = 91c788b8-8f75-11e1-8992-6e8dde31d770
Number of disk discovered = 1
        caa_private0
          state : UP
          uDid  :
          uUid  : 39597db5-d0b8-5741-3e06-095cd9da92e7
          type  : REPDISK
root@clmgr1 /opt/IBM#

root@clmgr1 /opt/IBM# lspv
hdisk0          00f74d471c3abc13                    rootvg          active
hdisk1          00f74d4733c60118                    rootvg          active
caa_private0    00f74d4733c91f39                    caavg_private   active
hdisk3          00f74d4733c9370e                    ihsvg
root@clmgr1 /opt/IBM#

root@clmgr1 # lspv | grep caa
caa_private0    00f74d4733c91f39                    caavg_private   active
bash-3.00#
```

Use the command **rmcluster -n *<cluster name>*** on either node to remove the CAA cluster, then check whether it is removed with the **cluster -m** command on both nodes (Example 6-78).

*Example 6-78 Removing the CAA cluster*

```
root@clmgr1#  lscluster -c
Cluster query for cluster ihs_cluster returns:
Cluster uuid: d5de798c-82ee-11e1-90df-b6fcc07d1d70
Number of nodes in cluster = 2
        Cluster id for node clmgr1 is 1
        Primary IP address for node clmgr1 is 10.1.1.54
        Cluster id for node clmgr2 is 2
```

```
        Primary IP address for node clmgr2 is 10.1.1.55
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.36

root@clmgr1# rmcluster -n ihs_cluster
rmcluster: Removed cluster shared disks are automatically renamed to names such
        as hdisk10, [hdisk11, ...] on all cluster nodes.  However, this cannot
        take place while a disk is busy or on a node which is down or not
        reachable.  If any disks cannot be renamed now, you must manually
        rename them by removing them from the ODM database and then running
        the cfgmgr command to recreate them with default names.  For example:
        rmdev -l cldisk1 -d
        rmdev -l cldisk2 -d
        cfgmgr

root@clmgr1# lscluster -m
Cluster services are not active
root@clmgr1#

root@clmgr2# lscluster -m
Cluster services are not active
root@clmgr2#
```

The successful execution of the **rmcluster** command removes the volume group `caavg_private` on all nodes of the cluster as shown in Figure 6-58.

```
root@clmgr1# lspv | grep hdisk2
hdisk2          00f74d4733c91f39                    None

root@clmgr2# lspv | grep hdisk2
hdisk2          00f74d4733c91f39                    None
```

*Figure 6-58 Repository disk status on cluster nodes*

## Uninstall PowerHA SystemMirror 7.1.0

Uninstall the PowerHA 7.1.0 cluster filesets from all nodes by using **installp** as shown in Figure 6-59.

```
root@clmgr1# installp -ug cluster.adt.* cluster.es.* cluster.doc.*
cluster.license
```

*Figure 6-59 Uninstalling cluster filesets*

Ensure that you uninstalled the cluster filesets on all nodes by using the commands that are shown in Figure 6-60 on page 348.

```
root@clmgr1# lslpp -l | grep cluster.*
  bos.cluster.rte            7.1.0.1  COMMITTED  Cluster Aware AIX
  bos.cluster.solid          7.1.0.0  COMMITTED  Cluster Aware AIX SolidDB
root@clmgr1# installp -C

installp: No filesets were found in the Software Vital
        Product Database that could be cleaned up.
```

*Figure 6-60 Verifying the uninstalled cluster filesets*

## Upgrading AIX from AIX 7.1 TL 00 to AIX 7.1 TL 01

Upgrade AIX to AIX 7.1 TL 01 SP 3 (which is the latest version available during the writing of this publication) on all nodes and reboot all cluster nodes as shown in Example 6-79.

*Example 6-79 AIX level*

```
Oslevel after removing the caa on all nodes of the cluster.

root@clmgr1# oslevel -s
7100-00-04-1140

Oslevel after upgrading the TL on all nodes of the cluster.

root@clmgr1# oslevel -s
7100-01-03-1207
```

## Installing PowerHA SystemMirror 7.1.1

Install the base filesets of PowerHA SystemMirror 7.1.1 on all nodes and install the latest service packs. For more details, see 3.7, “PowerHA SystemMirror installation and prerequisites” on page 91.

## Converting the snapshot to support PowerHA SystemMirror 7.1.1

By using the previously taken snapshot configuration files, execute the command that is shown in Example 6-80. The successful execution of this command renames the existing `710Snapshot.odm` file to `710Snapshot.odm.old` and creates a `710Snapshot.odm` file.

*Example 6-80 Converting the snapshot file*

```
root@clmgr1# /usr/es/sbin/cluster/conversion/clconvert_snapshot -v 7.1.0 -s
710Snapshot
Extracting ODM's from snapshot file... done.
Converting extracted ODM's... done.
Rebuilding snapshot file... done.
root@clmgr1#

root@clmgr1# cd /usr/es/sbin/cluster/snapshots/
root@clmgr1# ls -l
total 272
-rw-r--r--    1 root     system        77863 Apr 10 07:16 710Snapshot.info
-rw-r--r--    1 root     system        51502 Apr 10 07:15 710Snapshot.odm
-rw-------    1 root     system           51 Apr 10 07:15 clsnapshot.log
root@clmgr1# ls -l
total 376
```

```
-rw-r--r--    1 root     system        77863 Apr 10 07:16 710Snapshot.info
-rw-r--r--    1 root     system        51505 Apr 10 10:37 710Snapshot.odm
-rw-r--r--    1 root     system        51502 Apr 10 07:15 710Snapshot.odm.old
-rw-------    1 root     system           51 Apr 10 07:15 clsnapshot.log
root@clmgr1#
```

### Syslog and CAA

Ensure that you have an entry in the `/etc/syslog.conf` file to capture CAA logs. Ensure that the host names of all cluster nodes that resolve to their respective IP addresses are added in the `/etc/cluster/rhosts` file. Restart **clcomd** as shown in the Example 6-81.

*Example 6-81 Syslog and CAA rhosts file entries*

```
root@clmgr1 / # tail -1 /etc/syslog.conf
*.info /var/adm/ras/syslog.caa rotate size 1m files 10

root@clmgr1 / # tail -2 /etc/cluster/rhosts
clmgr1
clmgr2

root@clmgr1 / # stopsrc -s clcomd; sleep 2; startsrc -s clcomd;
```

### Restoring the cluster by using the converted snapshot file

We can restore the converted snapshot cluster configuration file by using **smit sysmirror**. Start **smit sysmirror** and select **Cluster Nodes and Networks** → **Manage the cluster** → **Snapshot Configuration** → **Restore the Cluster Configuration From a Snapshot**.

Select the snapshot file name and press Enter as shown in Figure 6-61.

```
 Restore the Cluster Snapshot

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

                                                         [Entry Fields]
  Cluster Snapshot Name                                 710Snapshot
  Cluster Snapshot Description                          Migration
  Un/Configure Cluster Resources?                      [Yes]                +
  Force apply if verify fails?                         [No]                 +



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

*Figure 6-61 Restoring the snapshot configuration*

The successful restoration of the snapshot configuration files restores the cluster configuration of PowerHA 7.1.1 and creates the volume group `caavg_private` in the repository disk. You can verify the `caavg_private` volume group on either cluster node as shown in Example 6-82 on page 350.

*Example 6-82 Verifying the caavg_private vg*

```
root@clmgr1# lspv | grep caa
hdisk2          00f74d4733c91f39                    caavg_private   active
```

Check the cluster configuration on all nodes and confirm that the CAA cluster state is error free as shown in Example 6-83.

*Example 6-83 Verifying the CAA cluster state*

```
# lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: clmgr1
        Cluster shorthand id for node: 2
        uuid for node: f8b7ec58-8970-11e1-b362-b6fcc07d1d6f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        ihs_cluster        local        f8e343d0-8970-11e1-b362-b6fcc07d1d6f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: clmgr2
        Cluster shorthand id for node: 3
        uuid for node: f8e03d98-8970-11e1-b362-b6fcc07d1d6f
        State of node:  UP
        Smoothed rtt to node: 17
        Mean Deviation in network rtt to node: 16
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        ihs_cluster        local        f8e343d0-8970-11e1-b362-b6fcc07d1d6f

        Number of points_of_contact for node: 2
        Point-of-contact interface & contact state
         dpcom  UP  RESTRICTED
         en0  UP


root@clmgr1# lscluster -i
Network/Storage Interface Query

Cluster Name:  ihs_cluster
Cluster uuid:  f8e343d0-8970-11e1-b362-b6fcc07d1d6f
Number of nodes reporting = 2
```

```
Number of nodes expected = 2
Node clmgr1
Node uuid = f8b7ec58-8970-11e1-b362-b6fcc07d1d6f
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.c0.7d.1d.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.1.1.54  broadcast 10.1.1.255  netmask
255.255.254.0
                IPV4 ADDRESS: 172.16.21.54  broadcast 172.16.21.255  netmask
255.255.254.0
                Number of cluster multicast addresses configured on interface = 1
                IPV4 MULTICAST ADDRESS: 228.1.1.36  broadcast 0.0.0.0  netmask
0.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
        Pseudo Interface
                 Interface State DOWN
Node clmgr2
Node uuid = f8e03d98-8970-11e1-b362-b6fcc07d1d6f
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  6e.8d.de.31.d7.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 10.1.1.55  broadcast 10.1.1.255  netmask
255.255.254.0
                IPV4 ADDRESS: 172.16.21.55  broadcast 172.16.21.255  netmask
255.255.254.0
                Number of cluster multicast addresses configured on interface = 1
                IPV4 MULTICAST ADDRESS: 228.1.1.36  broadcast 0.0.0.0  netmask
0.0.0.0
```

```
Interface number 2 dpcom
        ifnet type = 0 ndd type = 305
        Mac address length = 0
        Mac address =  0.0.0.0.0.0
        Smoothed rrt across interface = 667
        Mean Deviation in network rrt across interface = 1208
        Probe interval for interface = 18750 ms
        ifnet flags for interface = 0x0
        ndd flags for interface = 0x9
        Interface state  UP  RESTRICTED  AIX_CONTROLLED
Pseudo Interface
         Interface State DOWN
```

### Checking the PowerHA SystemMirror configuration

The following smit menu helps you check the PowerHA configuration. Start **`smit sysmirror`** and select **Problem Determination Tools** → **PowerHA SystemMirror Verification** → **Verify PowerHA SystemMirror Configuration**. Select **No** for the "Verify changes only?" field and press Enter. A successful execution displays the status as shown in Figure 6-62.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]
HACMPnode ODM on node clmgr2 verified.

HACMPnetwork ODM on node clmgr2 verified.

HACMPcluster ODM on node clmgr2 verified.

HACMPnim ODM on node clmgr2 verified.

HACMPadapter ODM on node clmgr2 verified.

HACMPtopsvcs ODM on node clmgr2 verified.

[MORE...173]

F1=Help             F2=Refresh          F3=Cancel           F6=Command
F8=Image            F9=Shell            F10=Exit            /=Find
n=Find Next
```

*Figure 6-62 Verifying the PowerHA SystemMirror configuration*

### Starting the Cluster Services

After verifying the CAA cluster and PowerHA SystemMirror configuration state, start the Cluster Services node by node by using **`smitty clstart`** as shown in Figure 6-63 on page 353.

```
 Start Cluster Services

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

                                                        [Entry Fields]
* Start now, on system restart or both               now                    +
  Start Cluster Services on these nodes             [clmgr1]                +
* Manage Resource Groups                             Automatically          +
  BROADCAST message at startup?                      false                  +
  Startup Cluster Information Daemon?                true                   +
  Ignore verification errors?                        false                  +
  Automatically correct errors found during          Interactively          +
  cluster start?



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

*Figure 6-63 Starting Cluster Services*

Verify the state of the cluster on all nodes and ensure that the resource group is available and that the application is accessible (Example 6-84).

*Example 6-84 Verifying the cluster status after we restore the snapshot configuration*

```
root@clmgr1# lssrc -ls clstrmgrES|grep -e state -e vrmf -e fix
Current state: ST_STABLE
local node vrmf is 7112
cluster fix level is "2"

root@clmgr2# lssrc -ls clstrmgrES|grep -e state -e vrmf -e fix
Current state: ST_STABLE
local node vrmf is 7112
cluster fix level is "2"

root@clmgr1# clRGinfo
-----------------------------------------------------------------------------
Group Name     Group State                  Node
-----------------------------------------------------------------------------
ihs_resgrp     ONLINE                       clmgr1
               OFFLINE                      clmgr2

bash-3.00# cltopinfo
Cluster Name: ihs_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.36
There are 2 node(s) and 1 network(s) defined
```

```
NODE clmgr1:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr1  10.1.1.54
NODE clmgr2:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr2  10.1.1.55

Resource Group ihs_resgrp
        Startup Policy   Online On Home Node Only
        Fallover Policy  Fallover To Next Priority Node In The List
        Fallback Policy  Never Fallback
        Participating Nodes      clmgr1 clmgr2
        Service IP Label                 clmgrsvc1
```

We successfully restored the cluster configuration by using the snapshot migration.

7

# IBM PowerHA SystemMirror Smart Assist for SAP

This chapter describes the functionality of the PowerHA SystemMirror Smart Assist for SAP, which provides high availability for SAP services.

This chapter includes the following topics:

- PowerHA SystemMirror Smart Assist for SAP
- Standard steps for all clusters in this chapter
- Simple cluster solution with IBM PowerHA SystemMirror Smart Assist for monolithic installed SAP systems
- Scenario for a complete SAP and liveCache environment:
  - Cluster 1: SAP Supply Chain Management
  - Cluster 2: Database instance
  - Cluster 3: liveCache
- DB2 HADR cluster solution

## 7.1 PowerHA SystemMirror Smart Assist for SAP

PowerHA SystemMirror Smart Assist for SAP is included in the Standard Edition software at no additional charge. It simplifies and minimizes the time and effort to make an SAP system highly available. The Smart Assist automatically discovers SAP instances and databases and creates start and stop scripts for the instances. The Smart Assist also creates process and custom PowerHA application monitors that help to keep the SAP instances highly available.

The first Smart Assist for SAP is offered with PowerHA SystemMirror 7.1.

This chapter is presented in two parts. The first part shows a one-cluster solution that combines the SAP services and the database in one cluster. The second part explains how to configure a three-cluster SAP Supply Chain Management (SCM)/Advanced Planning and Optimization (APO) solution with a hot-standby liveCache installation. All three clusters are two-node IBM PowerHA SystemMirror 7.1.1 clusters that use the Smart Assist for SAP.

The simple cluster solution is to provide a solution for older monolithic SAP installations if no separate instance or installation for ABAP SAP Central Services (ASCS) or the database exists. For new installations, you can also set up a one-cluster solution. However, all SAP instances (database, ASCS, Enqueue Replication Server (ERS), and primary application server) must be installed separately and be able to operate independently. If not, you interfere with the SAP strategy for high availability solutions.

The three cluster solution is the preferred setup. It complies with the SAP strategy to use a stand-alone enqueue server with an optional replicated enqueue. Also, from the PowerHA perspective, this design is the focal solution.

7.8, "DB2 HADR cluster solution" on page 463 provides a sample configuration for a DB2 high availability disaster recovery (HADR) cluster. You can exchange the Smart Assist cluster 2 with this solution if you need faster takeover times. There is no Smart Assist for this solution.

**Terminology:** For better differentiation, in this book, the SAP central instance without a message and enqueue server (in the past also called DVEBMGS) is called the primary application server. Further dialog instances are called additional application servers.

### 7.1.1 Different Smart Assists

Different Smart Assists in the PowerHA SystemMirror 7.1.1 for AIX Standard Edition are used for this chapter. The central component is the Smart Assist for SAP, which you can install with the `cluster.es.assist.sap` fileset. But this Smart Assist for SAP is built on top of another Smart Assist. The Standard Smart Assist is the basic part, and it is installed with the `cluster.es.assist.common` fileset.

The Smart Assist for Network File System (NFS) is in the `cluster.es.nfs.rte` fileset.

If you want to make the database highly available with the Smart Assist for SAP, you also must install the following filesets, depending on the database:

- `cluster.es.assist.db2` for DB2
- `cluster.es.assist.oracle` for Oracle
- `cluster.es.assist.maxdb` for MaxDB

For the Smart Assist for SAP liveCache Hot Standby, one more Smart Assist is necessary. It is installed with the `cluster.es.assist.maxdb` fileset.

But everything is configured inside the Smart Assist for SAP except the liveCache part. There are different subtypes inside the Smart Assist for SAP:

- Global file system (GFS):

  This subtype discovers and configures the NFS shared file systems in an SAP environment (`/usr/sap/trans` and `/sapmnt`) and uses as the base layer the Smart Assist for NFS.

  > **Important:** This GFS subtype is mandatory as the basic subtype to discover all other Smart Assist for SAP subtypes. It uses only NFS version 4.

- Central services (SCS):

  This subtype discovers and configures the central services instances, which are the SCS and also ASCS instances.

- Enqueue Replication Server (ERS):

  This subtype discovers and configures the enqueue replication instance for a central service instance, which is the ERS named instance.

- Application server (AS):

  This subtype discovers and configures all application server instances, including both primary application server instances and additional application server instances.

- Database (DB)

  This subtype discovers and configures a highly available environment for a DB2 or Oracle database and uses as a base layer the Smart Assist for DB2 or the Smart Assist for Oracle.

## 7.1.2 Prerequisites of the solution

The solution requires these prerequisites:

- The use of the NFS service of the Smart Assist for SAP is required with this version of PowerHA. The SAP globals must be configured as the base of NFS 4 inside the cluster for the SAP central services. This requirement applies for `/sapmnt/<SID>` and `/usr/sap/trans`. It prevents you from setting up a second cluster with Smart Assist for SAP in a three-system SAP landscape with development, quality assurance, and a productive system, except that you use a remote transport system within SAP. The Smart Assist for SAP subtype NFS only supports NFS version 4.
- For the Smart Assists to work correctly, you must not have a second file system whose mount point begins with `/usr/sap/trans` or `/sapmnt`. The Smart Assist for SAP is not able to detect the NFS service for these file systems.
- There is no Smart Assist implementation for DB2 HADR. However, the use of the solution without Smart Assist that is shown in “DB2 HADR cluster solution” is possible.
- The Smart Assist for SAP supports ABAP and dual-stack SAP installations only. It does not support the pure JAVA stack.
- The service addresses are always configured on the first known PowerHA network in the cluster. If you use more networks and want to configure your service addresses on these networks, you must set up the service address on the correct network after the Smart Assist runs manually.

## 7.1.3 Supported versions by Smart Assist

In this section, you see a listing of the supported versions of SAP, DB2, and MaxDB for Smart Assist for SAP, and its dependent Smart Assist:

The supported version by the PowerHA Smart Assist for SAP is SAP NetWeaver 2004s.

The supported version by the PowerHA Smart Assist for DB2 is IBM DB2 Universal Database™ Enterprise Server Edition versions 8.1, 8.2, 9.1, and 9.5.

The supported version by the PowerHA Smart Assist for MaxDB is SAP MaxDB 7.6 and 7.7.

The supported stack by the PowerHA Smart Assist for SAP liveCache Hot Standby at the time of writing this publication includes the following products:

- The suggested version for AIX is SAP MaxDB Version 7.7.07.39 or later
- SAP SCM 7.0 EHP1
- SAP Database DB2 V9.7 or Oracle 10g2
- Storage:
  - SAN Volume Controller
  - V7000
  - DS8000 with flashcopy

Obtain a current list of supported SCMs in the Product Availability Matrix (PAM) on the SAP marketplace, if SCM is the clustered base on the Smart Assist.

> **Important:** Use the latest PowerHA service pack because of improved logic to control the applications.

## 7.1.4 Versions in our demonstration environment

In our environment, we installed the following versions:

- AIX: It is mandatory to install at a minimum all listed efixes as shown in Example 7-1.

*Example 7-1 Installed AIX version*

```
root@saplc2 / # oslevel -s
7100-01-03-1207

root@saplc2 / # emgr -l

ID  STATE LABEL      INSTALL TIME      UPDATED BY ABSTRACT
=== ===== ========== ================= ==========
======================================
1    S    IV16769    04/10/12 05:27:53           Avoided duplicated mountguard
define
2    S    IV12624s03 04/10/12 05:29:23           PCM stays on low priority
path
3    S    IV13083s03 04/10/12 05:30:02           SFWOBJ_T RETURNS INCORRECT
HEIGHT
4   S   IV17272s03 04/10/12 05:31:40        IV17272 for AIX 7.1 TL01 SP03
5    S    CAA_LOCK_F 05/07/12 10:13:09           Reboot of 1 node, rest are
down.
6    S    ABCDAPR    05/07/12 10:13:42          ABCD fix
```

```
7   S   IV14422c  05/07/12 10:16:00        Fix for IV14422 at RSCT 3.1.2
8    S    IV18343    05/07/12 10:16:42          fix for clRGinfo output
```

- Example 7-2 shows the installed PowerHA version.

*Example 7-2 Installed PowerHA version*

```
root@saplc2 / # halevel -s
7.1.1 SP1

root@saplc2 / # lslpp -L cluster*
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.adt.es.client.include
                             7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   CLINFO Samples
  cluster.doc.en_US.es.pdf   7.1.1.0    C     F    PowerHA SystemMirror PDF
                                                   Documentation - U.S. English
  cluster.es.assist.common   7.1.1.1    CE    F    PowerHA SystemMirror Smart
                                                   Assist Common Files
  cluster.es.assist.maxdb    7.1.1.1    CE    F    PowerHA SystemMirror Smart
                                                   Assist for SAP MaxDB
  cluster.es.assist.sap      7.1.1.0    C     F    PowerHA SystemMirror Smart
                                                   Assist for SAP
  cluster.es.client.clcomd   7.1.1.0    C     F    Cluster Communication
                                                   Infrastructure
  cluster.es.client.lib      7.1.1.1    C     F    PowerHA SystemMirror Client
                                                   Libraries
  cluster.es.client.rte      7.1.1.1    C     F    PowerHA SystemMirror
Client
                                                   Runtime
  cluster.es.client.utils    7.1.1.1    C     F    PowerHA SystemMirror
Client
                                                   Utilities
  cluster.es.client.wsm      7.1.1.0    C     F    Web based Smit
  cluster.es.cspoc.cmds      7.1.1.2    C     F    CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0    C     F    CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.2    C     F    CSPOC Runtime Commands
  cluster.es.migcheck        7.1.1.0    C     F    PowerHA SystemMirror
Migration
                                                   support
  cluster.es.server.cfgast   7.1.1.0    C     F    Two-Node Configuration
                                                   Assistant
  cluster.es.server.diag     7.1.1.1    CE    F    Server Diags
  cluster.es.server.events   7.1.1.1    CE    F    Server Events
  cluster.es.server.rte      7.1.1.1    CE    F    Base Server Runtime
  cluster.es.server.testtool
                             7.1.1.0    C     F    Cluster Test Tool
  cluster.es.server.utils    7.1.1.1    CE    F    Server Utilities
  cluster.license            7.1.1.0    C     F    PowerHA SystemMirror
                                                   Electronic License
  cluster.man.en_US.es.data  7.1.1.0    C     F    Man Pages - U.S. English
```

```
  cluster.msg.en_US.assist   7.1.1.1    AE    F    PowerHA SystemMirror Smart
                                                   Assist Messages - U.S.
English
  cluster.msg.en_US.es.client
                             7.1.1.0    C     F    PowerHA SystemMirror Client
                                                   Messages - U.S. English
  cluster.msg.en_US.es.server
                             7.1.1.1    C     F    Recovery Driver Messages -
                                                   U.S. English
```

- SAP NetWeaver 7.0 EHP2
- SAP SCM 7.0 EHP1
- DB2 V9.7 SP4
- MaxDB 7.7.07.39

Example 7-3 shows the installed DB2 version.

*Example 7-3 Installed DB2 version after SAPinst*

```
sapsma1:db2te1 1> db2pd -v
Instance db2te1 uses 64 bits and DB2 code release SQL09074
with level identifier 08050107
Informational tokens are DB2 v9.7.0.4, s110330, IP23236, Fix Pack 4.
```

Example 7-4 shows the installed SAP kernel version.

*Example 7-4 Installed SAP kernel version after SAPinst*

```
sapsma1:te1adm 1> disp+work -version
--------------------
disp+work information
--------------------
kernel release                720
kernel make variant           720_REL
compiled on                   AIX 2 5 00092901D600 for rs6000_64
compiled for                  64 BIT
compilation mode              UNICODE
compile time                  May 23 2011 21:24:01
update level                  0
patch number                  90
source id                     0.090
---------------------
supported environment
---------------------
database (SAP, table SVERS)   700
                              710
                              701
                              702
                              703
                              711
                              720
                              730
                              731
operating system
```

```
AIX 2 5
AIX 3 5
AIX 1 6
AIX 1 7
```

### 7.1.5 SAP system landscape in the demonstration environment

Table 7-1 shows the one-cluster solution with the SAP IDs.

*Table 7-1 SAP IDs in the one-cluster demonstration environment*

| Cluster name | SAP or DB ID | Type | Instance number |
|---|---|---|---|
| sapsma_cluster | TE1 | DVEBMGS<br>DB2 database | 02 |

Table 7-2 shows the users and groups in the one-cluster environment.

*Table 7-2 Users and groups in the one-cluster demonstration environment*

| SID | AIX user | AIX groups |
|---|---|---|
| TE1 | db2te1 | dbte1ctl |
| TE1 | te1adm | sapsys and dbte1ctl |
| TE1 | sapsr3 | sapsys and dbte1ctl |
| TE1 | sapadm | sapsys |

Table 7-3 shows the three-cluster solution with the SAP IDs.

*Table 7-3 SAP IDs in the three-cluster solution in the demonstration environment*

| Cluster name | SAP or DB ID | Type | Instance number |
|---|---|---|---|
| sapci_cluster | TC1 | ASCS<br>ERS<br>Primary application server | 10<br>11<br>12 |
| sapdb_cluster | TC1 | DB2 database | |
| saplc_cluster | TL1 | liveCache | |

Table 7-4 shows the users and groups in the three-cluster environment.

*Table 7-4 Users and groups in the three-cluster solution in the demonstration environment*

| SID | AIX user | AIX groups |
|---|---|---|
| TC1 | db2tc1 | dbtc1ctl |
| TC1 | tc1adm | sapsys, dbtc1ctl, and sdba |
| TC1 | sapsr3 | sapsys, dbtc1ctl, and sdba |
| TC1 | sapadm | sapsys |
| TC1 | sdb | sdba |

| SID | AIX user | AIX groups |
|---|---|---|
| TC1 | tl1adm | sapsys |

# 7.2 Standard steps for all clusters in this chapter

In this section, we describe the standard steps for the installation of all clusters in this chapter.

## 7.2.1 Installing the required PowerHA SystemMirror Smart Assist filesets

The following filesets are installed in the demonstration environment as shown in Example 7-5 before the Smart Assist for SAP 7.1.1. can be used. All Smart Assist filesets are not necessary on all clusters. See 7.1.1, “Different Smart Assists” on page 356.

*Example 7-5 Additional filesets that are required for installing Smart Assist for SAP*

```
root@sapsma1 / # clcmd lslpp -L cluster.es.assist.common cluster.es.assist.sap
cluster.es.assist.db2 cluster.es.nfs.rte

-------------------------------
NODE sapsma2b1
-------------------------------
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.es.assist.common   7.1.1.1    C     F    PowerHA SystemMirror Smart
                                                   Assist Common Files
  cluster.es.assist.db2      7.1.1.0    C     F    PowerHA SystemMirror Smart
                                                   Assist for DB2
  cluster.es.assist.sap      7.1.1.0    C     F    PowerHA SystemMirror Smart
                                                   Assist for SAP
  cluster.es.nfs.rte         7.1.1.0    C     F    NFS Support

-------------------------------
NODE sapsma1b1
-------------------------------
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  cluster.es.assist.common   7.1.1.1    C     F    PowerHA SystemMirror Smart
                                                   Assist Common Files
  cluster.es.assist.db2      7.1.1.0    C     F    PowerHA SystemMirror Smart
                                                   Assist for DB2
  cluster.es.assist.sap      7.1.1.0    C     F    PowerHA SystemMirror Smart
                                                   Assist for SAP
  cluster.es.nfs.rte         7.1.1.0    C     F    NFS Support
```

## 7.2.2 Creating the users

The SAPinst creates all necessary users for DB2, SAP, or MaxDB if they cannot be found on the local system. SAPinst creates the user only on the first node and sets the SAP default home directories and the master password for all these users. If you want special directories and different passwords, you create the users before you start the installation. Or, you can use Cluster Single Point of Control (C-SPOC) to create the users with PowerHA SystemMirror functionality. If you do not use C-SPOC to create the users and groups, you have to ensure that the same users (with the same passwords) and groups are created on the second node.

**Important:** The users and groups have to be created with the same ID on both nodes.

Complete sample scripts to create the users and groups (before SAPinst is started and without using C-SPOC) can be found in "Scripts to create the SAP users" on page 604.

## 7.2.3 Creating volume groups, logical volumes, and file systems

You must create the volume groups, logical volumes, and the file systems before you start the SAPinst. Use the C-SPOC functionality with the PowerHA instruments or write your own script. You have to import the shared volume groups on the second node. If you use this method, use the same major number on the second node to prevent NFS problems. C-SPOC performs all required tasks on the second node automatically.

Example scripts for each environment in the demonstration environment can be found in "Scripts to create the volume groups, logical volumes, and file systems" on page 605.

**Important:** For the Smart Assists to work correctly, do not have a second file system whose mount point begins with `/usr/sap/trans` or `/sapmnt`; otherwise, the Smart Assist cannot detect the NFS service.

**Important:** Smart Assist for SAP with subtype DB2 and liveCache automatically imports the volume groups on the backup node if the volume group is not manually imported earlier. GFS and all other subtypes do not work with this automation. In this case, you must manually import the volume group in advance.

## 7.2.4 Updating the /etc/services file on the secondary node

When the instance is created on the primary node, the `/etc/services` file is updated with information for SAP usage. You must also add the lines that the sapinst created on the installing node to the `/etc/services` file on the secondary node. There are two ways to make these entries on the second side available. The first and the easiest way is to copy the complete file simply to the second node, ideally with the same timestamp. This method can be used only if both files are the same before the start of the SAPinst after the SAPinst finished and there are only additional lines on the first node:

```
root@sapsma1 / # clrcp /etc/services sapsma2:/etc/services
```

A better method is to use this command:

```
root@sapsma1 / # scp -p /etc/services sapsma2:/etc/services
```

The second method is to add every necessary line on the second node with an AIX command, for example:

```
root@sapsma2 / # chservices -a -v sapgw99s -p tcp -n 4899 -u "SAP System Gateway
Security Port"
```

### 7.2.5 Copying more directories to the second node

SAP and DB2 can be started on the second cluster, and you have to copy more files and directories manually to the other node. In our environment, we copy the following files and directories that are local on the first node to the second node:

- `root@sapsma1 / # scp -pr /var/db2 sapsma2:/var/db2`
- `root@sapsma1 / # scp -pr /usr/sap/* sapsma2:/usr/sap/`

> **Important:** Use the second command only if all shared file systems are unmounted.

> **Important:** You cannot use the **`clrcp`** command for this step because it works with files only, not with directories.

## 7.3 Simple cluster solution with IBM PowerHA SystemMirror Smart Assist for monolithic installed SAP systems

We provide a simple cluster solution description with IBM PowerHA Smart Assist for monolithic installed SAP systems.

### 7.3.1 Overview

We show how you can make an existing SAP system highly available. In this environment, we use the SAP NetWeaver 7.0 EHP2 installation with DB2 as the underlying database. We present two alternatives to make the SAP system highly available.

Alternative A for existing SAP installations is shown in Figure 7-1 on page 365.

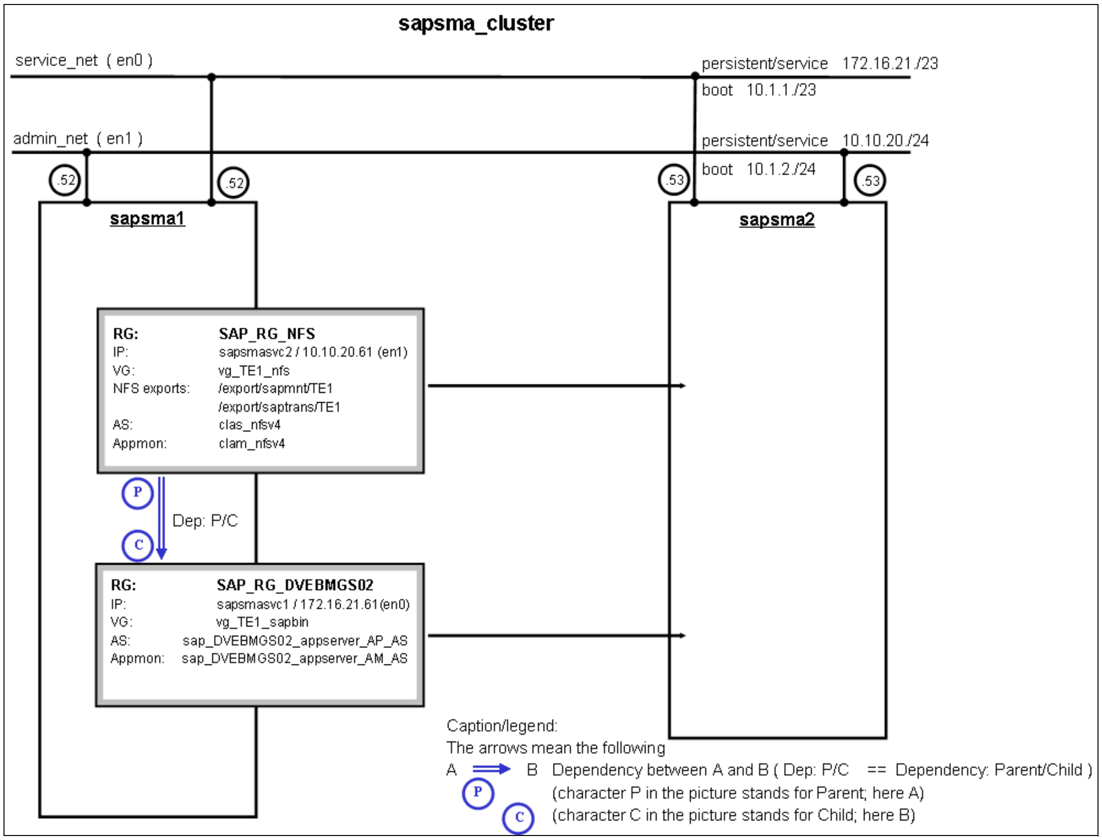



*Figure 7-1 Schema overview alternative A for sapsma_cluster*

In Figure 7-1, the SAP system is installed as a monolithic SAP central system with an included database with only one volume group (VG) for the database and the SAP system and only one service address for both. The Smart Assist for SAP can only be used to set up an application server instance. Because of this consideration, there is no monitoring that is configured by the Smart Assist for SAP for the database instance.

Figure 7-1 shows these main steps:

- Installation and configuration steps before you use Smart Assist for SAP
- Starting Smart Assist for SAP: Global file system (NFS)
- Starting Smart Assist for SAP: Application server instance (AS)
- Completing the configuration

Alternative B is shown in Figure 7-2 on page 366.

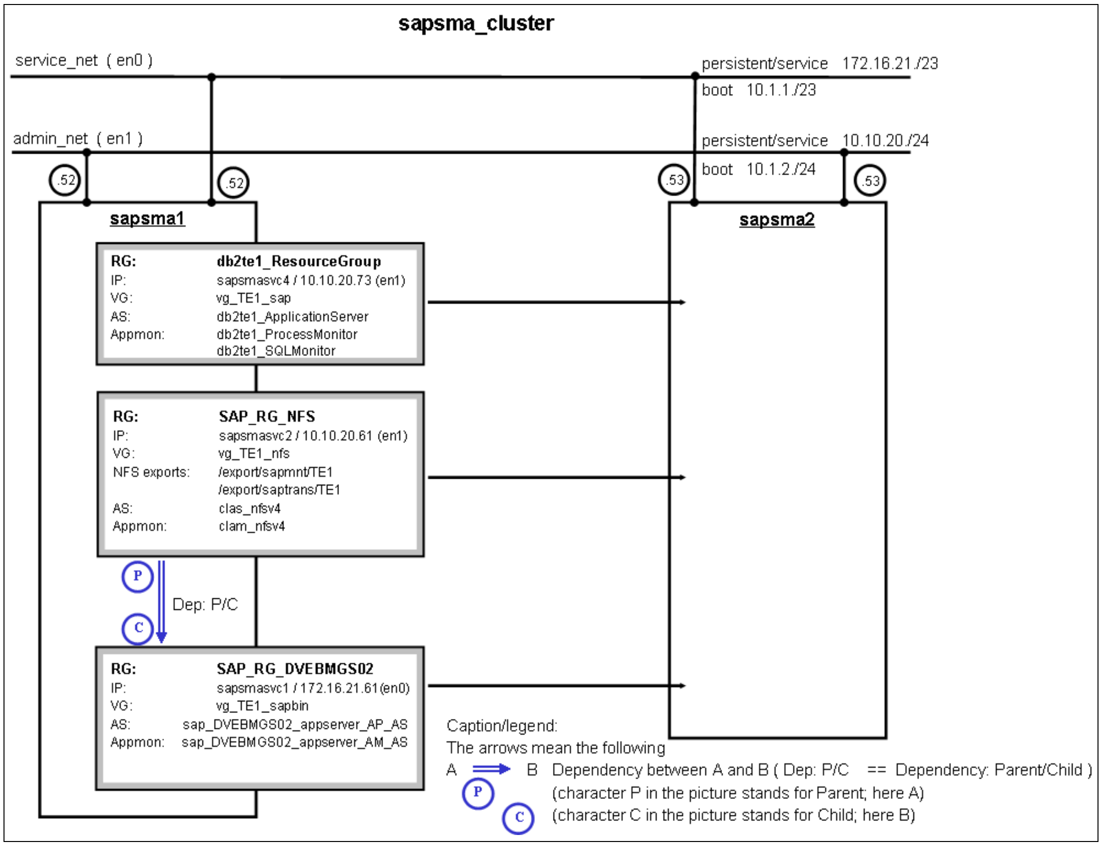



*Figure 7-2 Schema overview alternative B for sapsma_cluster*

Figure 7-2 uses a separate installation for the SAP central instance (DVEBMGS). The message and enqueue server (with no separated ASCS like alternative A) and the database each has its own volume group and service address. In this example, the Smart Assist for SAP configures monitoring for the database instance and the SAP system.

Figure 7-2 has these main steps:

- Installation and configuration steps before you use Smart Assist for SAP
- Starting Smart Assist for SAP: Global file system (NFS)
- Alternative B only: Starting Smart Assist for SAP with the database instance
- Starting Smart Assist for SAP: Application server instance (AS)
- Completing the configuration

Unless otherwise explained, the step is for both alternatives.

## 7.3.2 Installation and configuration steps before you use Smart Assist for SAP

The steps for the installation and configuration preparation are listed with some steps that are shown only as a link to the detailed described steps in 7.2, “Standard steps for all clusters in this chapter” on page 362. We explain the preliminary steps that are required before you can start Smart Assist for SAP:

1. Install the required PowerHA Smart Assist filesets.

   See 7.2.1, “Installing the required PowerHA SystemMirror Smart Assist filesets” on page 362.

2. Configure the base IBM PowerHA SystemMirror.

   You must configure the topology of the PowerHA cluster before you use the Smart Assist for SAP as shown in the previous chapters. Example 7-6 shows the cluster `sapsma_cluster` that is configured with two Ethernet interfaces in each node.

   *Example 7-6 Cluster network configuration*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # lscluster -c
Cluster query for cluster sapsma_cluster returns:
Cluster uuid: 46837bf4-71f3-11e1-a1c9-6e8dd5fb9f6f
Number of nodes in cluster = 2
        Cluster id for node sapsma1b1 is 3
        Primary IP address for node sapsma1b1 is 10.1.1.52
        Cluster id for node sapsma2b1 is 4
        Primary IP address for node sapsma2b1 is 10.1.1.53
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.27

root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address
Interface Name Global Name Netmask Alias for HB Prefix Length

sapsma1b2 boot admin_net ether public sapsma1 10.1.2.52 en1 255.255.255.0 24
sapsma1b1 boot service_net ether public sapsma1 10.1.1.52 en0 255.255.254.0 23
sapsma2b2 boot admin_net ether public sapsma2 10.1.2.53 en1 255.255.255.0 24
sapsma2b1 boot service_net ether public sapsma2 10.1.1.53 en0 255.255.254.0 23
```

3. Configure the network.

   The IP addresses for the NFS server (`sapsmasvc2`) and the SAP system (`sapsmasvc1`) are configured on the node where SAP is active.

   Alternative A: Example 7-7 shows the network settings for alternative A.

   *Example 7-7 Network settings for alternative A*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.d5.fb.9f.6f  3569880     0 14219178     0     0
en0   1500  10.1        sapsma1b1          3569880     0 14219178     0     0
en0   1500  172.16.20   sapsma1p1          3569880     0 14219178     0     0
en0   1500  172.16.20   sapsmasvc1         3569880     0 14219178     0     0
en1   1500  link#3      6e.8d.d5.fb.9f.70    17243     0     1083     0     0
en1   1500  10.1.2      sapsma1b2            17243     0     1083     0     0
en1   1500  10.10.20    sapsmasvc2           17243     0     1083     0     0
lo0   16896 link#1                          200994     0   200994     0     0
lo0   16896 127         loopback            200994     0   200994     0     0
lo0   16896 loopback                        200994     0   200994     0     0
```

   Alternative B: An additional service address (`sapsmasvc4`) for the database is configured as shown in Example 7-8.

   *Example 7-8 Network settings for alternative B*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.d5.fb.9f.6f  3569880     0 14219178     0     0
en0   1500  10.1        sapsma1b1          3569880     0 14219178     0     0
```

```
en0   1500  172.16.20   sapsma1p1         3569880     0 14219178     0      0
en0   1500  172.16.20   sapsmasvc1        3569880     0 14219178     0      0
en1   1500  link#3      6e.8d.d5.fb.9f.70   17243     0     1083     0      0
en1   1500  10.1.2      sapsma1b2           17243     0     1083     0      0
en1   1500  10.10.20    sapsmasvc2          17243     0     1083     0      0
en1   1500  10.10.20    sapsmasvc4          27705     0     1747     0      0
lo0   16896 link#1                         200994     0   200994     0      0
lo0   16896 127         loopback           200994     0   200994     0      0
lo0   16896 loopback                       200994     0   200994     0      0
```

4. Create the users and groups.

   See 7.2.2, "Creating the users" on page 362 and the script to create the users in the "Script to create users and groups for sapsma_cluster" on page 604.

5. Create volume groups, logical volumes, and file systems.

   See 7.2.3, "Creating volume groups, logical volumes, and file systems" on page 363, and the script to create the volume groups and file systems in the "Script to create VGs and LVs for sapsma_cluster" on page 605.

   Example 7-9 shows the volume groups that are known to both nodes.

> **Important:** For the Smart Assists for SAP subtype NFS to work correctly, do not have a second file system whose mount point begins with `/usr/sap/trans` or `/sapmnt`. Smart Assist cannot detect the NFS service correctly.

*Example 7-9 Volume groups that are known on both nodes*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # clcmd lspv

-------------------------------
NODE sapsma2b1
-------------------------------
hdisk0       00f74d451c3a5596              rootvg         active
hdisk1       00f74d452b615d3e              rootvg         active
hdisk2       00f61ab22c211a1e              caavg_private  active
hdisk3       00f61ab22c0f4e26              vg_TE1_sap
hdisk4       00f61ab22c0f4ee5              vg_TE1_sap
hdisk5       00f61ab22c0f4f8c              vg_TE1_sap
hdisk6         00f61ab23bd8e232                vg_TE1_nfs

-------------------------------
NODE sapsma1b1
-------------------------------
hdisk0       00f61ab21c38c7c7              rootvg         active
hdisk1       00f61ab22b620078              rootvg         active
hdisk2       00f61ab22c211a1e              caavg_private  active
hdisk3       00f61ab22c0f4e26              vg_TE1_sap     active
hdisk4       00f61ab22c0f4ee5              vg_TE1_sap     active
hdisk5       00f61ab22c0f4f8c              vg_TE1_sap     active
hdisk6         00f61ab23bd8e232                vg_TE1_nfs      active
```

The shared volume groups have to be *active* with file systems that are mounted on one node. It does not matter on which node they are mounted. Mount the file systems as shown in Example 7-10 so that Smart Assist for SAP can discover the available NFS services, SAP instances, and databases:

– Alternative A

*Example 7-10 Alternative A: Checking for mounted file systems in node sapsma1*

```
root@sapsma1 / # lsvg -l vg_TE1_nfs
vg_TE1_nfs:
LV NAME            TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TE1.lvsapmnt       jfs2 64 64 1    open/syncd    /export/sapmnt/TE1
TE1.lvsaptrans     jfs2 32 32 1    open/syncd    /export/saptrans/TE1

root@sapsma1 / # lsvg -l vg_TE1_sap
vg_TE1_sap:
LV NAME            TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TE1.lvusrsap       jfs2 160 160 1    open/syncd    /usr/sap/TE1
TE1.lvdb2binary    jfs2 128 128 1    open/syncd    /db2/TE1
TE1.lvdb2logdir    jfs2 64 64 1      open/syncd    /db2/TE1/log_dir
TE1.lvdb2logarc    jfs2 64 64 1      open/syncd    /db2/TE1/log_archive
TE1.lvdb2dat.01    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata1
TE1.lvdb2dat.02    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata2
TE1.lvdb2dat.03    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata3
TE1.lvdb2dat.04    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata4
```

– Alternative B

An additional volume group is configured. Example 7-11 shows the mounted file systems in node `sapsma1`.

*Example 7-11 Alternative B: Checking for mounted file systems in node sapsma1*

```
root@sapsma1 / # lsvg -l vg_TE1_nfs
vg_TE1_nfs:
LV NAME            TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TE1.lvsapmnt       jfs2 64 64 1      open/syncd    /export/sapmnt/TE1
TE1.lvsaptrans     jfs2 32 32 1      open/syncd    /export/saptrans/TE1

root@sapsma1 /usr/sap/TE1 # lsvg -l vg_TE1_sapbin
vg_TE1_sapbin:
LV NAME            TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TE1.lvusrsap       jfs2 160 160 1    open/syncd    /usr/sap/TE1

root@sapsma1 / # lsvg -l vg_TE1_sap
vg_TE1_sap:
LV NAME            TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TE1.lvdb2binary    jfs2 128 128 1    open/syncd    /db2/TE1
TE1.lvdb2logdir    jfs2 64 64 1      open/syncd    /db2/TE1/log_dir
TE1.lvdb2logarc    jfs2 64 64 1      open/syncd    /db2/TE1/log_archive
TE1.lvdb2dat.01    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata1
TE1.lvdb2dat.02    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata2
TE1.lvdb2dat.03    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata3
TE1.lvdb2dat.04    jfs2 160 160 1    open/syncd    /db2/TE1/sapdata4
```

6. Set up NFS.

   Configure an NFS domain on both nodes because the Smart Assist for SAP uses NFS version 4:

   ```
   root@sapsma1 / # clcmd chnfsdom sapsma_cluster
   ```

   The NFS server exports all needed file systems as shown in Example 7-12.

   *Example 7-12 NFS server on sapsma1*

   ```
   root@sapsma1 / # exportfs
   /export/sapmnt/TE1
   -sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapsma1:sapsma2:sapsma1b2:sapsma2b2:sapsma
   svc2:sapsmasvc4
   /export/saptrans/TE1
   -sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapsma1:sapsma2:sapsma1b2:sapsma2b2:sapsma
   svc2:sapsmasvc4
   ```

   The NFS client mounts are available as shown in Example 7-13.

   *Example 7-13 NFS mounts on sapsma1*

   ```
   root@sapsma1 / # mount|grep nfs
   sapsmasvc2 /export/sapmnt/TE1 /sapmnt/TE1        nfs3   Mar 23 10:33
   sapsmasvc2 /export/saptrans/TE1 /usr/sap/trans   nfs3   Mar 23 10:34
   ```

7. Install SAP.

   We performed a normal SAP installation with the virtual SAP host name. Only the usual SAP parameters are set as shown in Example 7-14.

   *Example 7-14 SAPinst environment parameters*

   ```
   export TMPDIR=/tmp/TE1
   mkdir $TMPDIR
   export JAVA_HOME=/usr/java14_64
   ./sapinst SAPINST_USE_HOSTNAME=sapsmasvc1
   ```

   Complete a standard SAPinst installation.

   > **Important:** Smart Assist SAP with subtype DB2 and Smart Assist for DB2 support only non-Distributed Print Function (DPF) and non-HADR DB2 databases.

   After the SAPinst installation, start the SAP instance if it is not running as shown in Example 7-15.

   *Example 7-15 Start SAP on sapsma1*

   ```
   sapsma1:te1adm 1> startsap sapsmasvc1

   Checking db6 db Database
   ------------------------------
    Database is not available via R3trans
    Running /usr/sap/TE1/SYS/exe/run/startdb
   03/23/2012 19:04:29     0   0   SQL1063N  DB2START processing was successful.
   SQL1063N  DB2START processing was successful.
   Database activated
   *** WARNING Logretain mode is disabled
   ```

```
    You will be unable to fully recover your database in case
    of a media failure
 /usr/sap/TE1/SYS/exe/run/startdb completed successfully

Starting Startup Agent sapstartsrv
-----------------------------
OK
 Instance Service on host sapsma1 started

starting SAP Instance DVEBMGS02
------------------------------
 Startup-Log is written to /usr/sap/TE1/home/te1adm/startsap_DVEBMGS02.log
/usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02 -function Start
 Instance on host sapsma1 started
```

Check whether DB2 is running as shown in Example 7-16.

*Example 7-16 DB2 is running on sapsma1*

```
sapsma1:db2te1 1> db2pd -

Database Partition 0 -- Active -- Up 0 days 00:02:01 -- Date 03/23/2012
19:06:23
```

Check whether the SAP instance is active as shown in Example 7-17.

*Example 7-17 SAP is running on sapsma1*

```
sapsma1:te1adm 5> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 02
-function GetProcessList

23.03.2012 19:19:40
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
msg_server, MessageServer, GREEN, Running, 2012 03 23 19:05:02, 0:14:38,
10420250
disp+work, Dispatcher, GREEN, Running, Message Server connection ok, Dialog
Queue time: 0.00 sec, 2012 03 23 19:05:02, 0:14:38, 10616962
igswd_mt, IGS Watchdog, GREEN, Running, 2012 03 23 19:05:02, 0:14:38, 9764996
```

**Important:** The Smart Assist for DB2 uses the **db2gcf** command to start the database in a PowerHA environment. It does *not* change the db2nodes.cfg directly.

8. Update the /etc/services file on the secondary node.

   See 7.2.4, “Updating the /etc/services file on the secondary node” on page 363.

9. Copy the additional directories.

   See 7.2.5, “Copying more directories to the second node” on page 364.

10. Set up the DB2 environment variable.

    Find the path for the binary files and then export the variable as shown in Example 7-18 on page 372. The DSE_INSTALL_DIR environment variable is exported as a root user with the actual path for the DB2 binary files. If more than one DB2 version is installed, choose the version that you use for this highly available instance.

*Example 7-18   Finding the DB2 binary files and exporting them*

```
root@sapsma1 / # /db2/TE1/db2te1/db2_software/bin/db2level
DB21085I  Instance "db2te1" uses "64" bits and DB2 code release "SQL09074" with
level identifier "08050107".
Informational tokens are "DB2 v9.7.0.4", "s110330", "IP23236", and Fix Pack
"4".
Product is installed at "/db2/db2te1/db2_software".

root@sapsma1 / # export DSE_INSTALL_DIR=/db2/db2te1/db2_software
```

**Important:** The setting of this variable is necessary. Without the seeded variable DSE_INSTALL_DIR, the Smart Assist for SAP cannot find the appropriate database.

11. Check what the Smart Assist for SAP can discover as shown in Example 7-19.

    You can check the possible types on the command line. All located types have a "1" at the end. The following parameters are available for the **`cl_sapdiscover -t`** command:

    - GFS
    - AS
    - SCS
    - ERS
    - DB

*Example 7-19   Checking the Smart Assist for SAP discover function*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t GFS
SAP Smart Assist:SAPNW_7.0:1.SAP NW 7.0 Global Filesystem:SAPNW_7.0_SAPGFS:1

root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t AS
SAP Smart Assist:SAPNW_7.0:4.SAP NW 7.0 AS Instance:SAPNW_7.0_ASINSTANCE:1

root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t SCS
SAP Smart Assist:SAPNW_7.0:2.SAP NW 7.0 SCS Instance:SAPNW_7.0_SCSINSTANCE:0

root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t ERS
SAP Smart Assist:SAPNW_7.0:3.SAP NW 7.0 ERS Instance:SAPNW_7.0_ERSINSTANCE:0

root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t DB
SAP Smart Assist:SAPNW_7.0:5.SAP Database Instance:SAPNW_7.0_DBINSTANCE:1
```

12. Get trace information.

    More information can be collected by setting an additional environment variable as shown in Example 7-20.

*Example 7-20   Setting enlarged output for cl_sapdiscover*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # export VERBOSE_LOGGING="high"
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t GFS
...
```

13. Determine the PowerHA state.

    You can start the PowerHA Smart Assist on a running cluster or on a cluster in the INIT state, but both nodes must be in the same state.

### 7.3.3 Starting Smart Assist for SAP: Global file system (NFS)

After you complete the steps in the previous sections, you are ready to start Smart Assist for SAP as explained in the following steps:

1. Launch Smart Assist for SAP by using the path for sapsma1: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-3), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  DB2 UDB non-DPF Smart Assist     # sapsma1 sapsma2
  SAP MaxDB Smart Assist           # sapsma1 sapsma2
  SAP Smart Assist                 # sapsma1 sapsma2
```

*Figure 7-3 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-4), select **Automatic Discovery and Configuration**.

```
                           Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-4 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-5), select **1.SAP NW 7.0 Global Filesystem**.

```
              Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  4.SAP NW 7.0 AS Instance         # sapsma1
  5.SAP Database Instance          # sapsma1 sapsma2
  1.SAP NW 7.0 Global Filesystem   # sapsma1
```

*Figure 7-5 Selecting the configuration to create*

5. The Add SAP Global Filesystem Details panel appears (Figure 7-6 on page 374).

```
                        Add SAP Global Filesystem Details

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

                                                          [Entry Fields]
* SAP Global File System Owning Node                    sapsma1
* Takeover Nodes                                       []
* Service IP Label                                     []
* Shared Volume Groups                                 [vg_TE1_nfs]
* Filesystems/Directories to Export (NFSv2/3)          [/export/sapmnt/TE1
/export/saptrans/TE1]
* Filesystems/Directories to NFS Mount
[/sapmnt/TE1;/export/sapmnt/TE1 /usr/sap/trans;/export/saptrans/TE1]
```

*Figure 7-6 SAP global filesystem details*

6. Using the available pick lists (F4), edit the Takeover Nodes field to show **sapsma2** and the Service IP Label field to show **sapsmasvc2**, as shown in Figure 7-7. Press Enter.

```
                        Add SAP Global Filesystem Details

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

                                                          [Entry Fields]
* SAP Global File System Owning Node                    sapsma1
* Takeover Nodes                                       [sapsma2]
* Service IP Label                                     [sapsmasvc2]
* Shared Volume Groups                                 [vg_TE1_nfs]
* Filesystems/Directories to Export (NFSv2/3)          [/export/sapmnt/TE1
/export/saptrans/TE1]
* Filesystems/Directories to NFS Mount
[/sapmnt/TE1;/export/sapmnt/TE1 /usr/sap/trans;/export/saptrans/TE1]
```

*Figure 7-7 Additional entries for the Smart Assist*

**Important:** There is an automatic synchronization after you press Enter; therefore, press Enter only if the cluster is not running.

**Important:** If you configure the SAP global file system details on a running cluster, errors occur. The cluster synchronizes and tries to start the new resource group, then it has problems when it tries to unmount the manually mounted file systems. To prevent this error, unmount everything manually before you press Enter.

7. The following warning might appear as shown in Figure 7-8 on page 375.

```
WARNING: Node sapsma2 has cluster.es.nfs.rte installed however grace periods
are not fully enabled on this node. Grace periods must be enabled before
NFSv4 stable storage
 can be used.

PowerHA SystemMirror will attempt to fix this opportunistically when
acquiring NFS resources on this node however the change won't take effect
until the next time that nf
sd is started.

If this warning persists, the administrator should perform the following
steps to enable grace periods on sapsma2 at the next planned downtime:
  1. stopsrc -s nfsd
  2. smitty nfsgrcperiod
  3. startsrc -s nfsd
```

*Figure 7-8 Cluster synchronization and verification message*

8. The NFS server on all cluster nodes must be restarted.
9. A new PowerHA resource group, `SAP_RG_NFS`, is created. The volume group `vg_TE1_vg` and the service IP label `sapsmasvc2` are automatically added to the resource group as shown in Example 7-21.

*Example 7-21 Configured resource group for the SAP NFS instance*

```
root@sapsma2 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_NFS

root@sapsma2 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TE1 /export/saptrans/TE1"
FILESYSTEM=""
FORCED_VARYON="false"
FSCHECK_TOOL="fsck"
FS_BEFORE_IPADDR="true"
MOUNT_FILESYSTEM="/sapmnt/TE1;/export/sapmnt/TE1
/usr/sap/trans;/export/saptrans/TE1"
RECOVERY_METHOD="sequential"
SERVICE_LABEL="sapsmasvc2"
SSA_DISK_FENCING="false"
VG_AUTO_IMPORT="false"
VOLUME_GROUP="vg_TE1_nfs"
USERDEFINED_RESOURCES=""
```

10. If you have a second interface for the service IP label, you have to configure the correct network after the use of the Smart Assist for SAP on type NFS:

    Follow the path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resources** → **Configure Service IP Labels/Addresses** → **Change/Show a Service IP Label/Address**.
11. Select the service IP address and press Enter as shown in Figure 7-9 on page 376.

```
              Select a Service IP Label/Address to Change/Show

Move cursor to desired item and press Enter.

  sapsmasvc2
```

Figure 7-9 Select a Service IP Label/Address to Change/Show

12. Change the service address to the correct network as shown in Figure 7-10.

```
               Change/Show a Service IP Label/Address(standard)

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

                                                       [Entry Fields]
  IP Label/Address                                    sapsmasvc2
* New IP Label/Address                               [sapsmasvc2] +
  Netmask(IPv4)/Prefix Length(IPv6)                  [24]
* Network Name                                       [admin_net] +
  Resource Group Name                                 SAP_RG_NFS
```

Figure 7-10 Change/Show a Service IP Label/Address (standard)

13. Synchronize the cluster (Figure 7-11).

    Follow the path: **smitty sysmirror** → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.

```
             PowerHA SystemMirror Verification and Synchronization

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

                                                       [Entry Fields]
* Verify, Synchronize or Both                        [Both] +
* Include custom verification library checks        [Yes] +
* Automatically correct errors found during          [Yes] +
  verification?

* Force synchronization if verification fails?       [No] +
* Verify changes only?                               [No] +
* Logging                                            [Standard]
```

Figure 7-11 Change/Show a Service IP Label/Address (standard)

14. Administrator task: Verify that the start and stop scripts are created for the resource group:

    a. To verify the scripts, use the **cllsserv** commands or the SMIT tool as shown in Example 7-22 on page 377.

*Example 7-22 Verifying the start and stop scripts*

```
root@sapsma2 / # cllsserv
clas_nfsv4  /usr/es/sbin/cluster/apps/clas_nfsv4/start
/usr/es/sbin/cluster/apps/clas_nfsv4/stop background
```

b. Follow the path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Controller Scripts** → **Change/Show Application Controller Scripts**.

c. Select the application controller and press Enter. The characteristics of the application controller are displayed as shown in Figure 7-12.

```
                 Change/Show Application Controller Scripts

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

                                                          [Entry Fields]
  Application Controller Name                           clas_nfsv4
  Start Script /usr/es/sbin/cluster/apps/clas_nfsv4/start
  Stop Script /usr/es/sbin/cluster/apps/clas_nfsv4/stop
  Application Monitor Name(s)                           clam_nfsv4
  Application startup mode                             [background]
```

*Figure 7-12 Change/Show Application Controller Scripts panel*

d. Administrator task: Verify which custom and process application monitors are created by the Smart Assist for SAP. In our example, the application monitor is `clas_nfsv4`.

e. Run the following path for `seoul`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Monitors** → **Configure Custom Application Monitors** → **Configure Custom Application Monitors** → **Change/Show Custom Application Monitor**.

f. In the Application Monitor to Change panel (Figure 7-13), select **clam_nfsv4** and press Enter.

```
                        Application Monitor to Change

Move cursor to desired item and press Enter.

  clam_nfsv4

F1=Help                 F2=Refresh              F3=Cancel
Esc+8=Image             Esc+0=Exit              Enter=Do
/=Find                  n=Find Next
```

*Figure 7-13 Selecting the application monitor to change*

g. In the Change/Show Custom Application Monitor panel (Figure 7-14), you see the attributes of the application monitor.

```
 Change/Show Custom Application Monitor

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

                                                        [Entry Fields]
* Monitor Name                                        clam_nfsv4
  Application Controller(s) to Monitor                clas_nfsv4
* Monitor Mode                                        longrunning
* Monitor Method
/usr/es/sbin/cluster/apps/clam_nfsv4/monitor
  Monitor Interval                                   [60] #
  Hung Monitor Signal                                [9] #
* Stabilization Interval                             [15] #
  Restart Count                                      [0] #
  Restart Interval                                   [0] #
* Action on Application Failure                      [fallover] +
  Notify Method                                      []
  Cleanup Method                                     [] /
  Restart Method                                     [] /
```

*Figure 7-14 Change/Show Custom Application Monitor panel*

## 7.3.4 Alternative B only: Starting Smart Assist for SAP with the database instance

We describe how to start the Smart Assist for SAP with the database instance:

1. Launch Smart Assist for SAP by using the path for `sapsma1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-15), select **SAP Smart Assist**.

```
 Select an Application from the List of Discovered Applications Below

  Move cursor to desired item and press Enter.

   DB2 UDB non-DPF Smart Assist     # sapsma1 sapsma2
   SAP MaxDB Smart Assist           # sapsma1 sapsma2
   SAP Smart Assist                 # sapsma1 sapsma2
```

*Figure 7-15 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-16), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

 Move cursor to desired item and press Enter.

   Automatic Discovery And Configuration
   Manual Configuration
```

*Figure 7-16 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-17), select **5. SAP Database Instance**.

```
             Select The Specific Configuration You Wish to Create

 Move cursor to desired item and press Enter.

   4.SAP NW 7.0 AS Instance         # sapsma1
   5.SAP Database Instance          # sapsma1 sapsma2
   1.SAP NW 7.0 Global Filesystem   # sapsma1
```

*Figure 7-17 Selecting the configuration to create*

**SAP database instance:** If there is no entry for the SAP Database Instance, you have not set the necessary DB2 variable, for example:

```
export DSE_INSTALL_DIR=/db2/db2te1/db2_software
```

5. Select the line with your db2 user, `db2tel` (Figure 7-18).

```
               Select a Configured DB2 Instance Resource Group

Move cursor to desired item and press Enter.

  db2te1
```

*Figure 7-18 Select a Configured DB2 Instance Resource Group*

6. By using the available pick lists (F4), edit the DB2 Instance Owning Node, Takeover Node, DB2 Instance Name, and DB2 Instance Database to Monitor as shown in Figure 7-19 even if incorrect values are listed and press Enter.

```
                            DB2 Instance Owning Node

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

                                                         [Entry Fields]
* Application Name                                    [db2te1]

* DB2 Instance Owning Node                            [Not a terminal] +
* Takeover Node(s)                                    [db2te1] +
* DB2 Instance Name
cl_db2sa_add_single_instance +
* DB2 Instance Database to Monitor +
* Service IP Label                                    []
```

*Figure 7-19 DB2 Instance Owning Node before selection*

Figure 7-20 shows the DB2 Instance Owning Node after your selections.

```
                            DB2 Instance Owning Node

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

                                                         [Entry Fields]
* Application Name                                    [db2te1]

* DB2 Instance Owning Node                            [sapsma1] +
* Takeover Node(s)                                    [sapsma2] +
* DB2 Instance Name                                    db2te1 +
* DB2 Instance Database to Monitor                     TE1 +
* Service IP Label                                    [sapsmasvc4]
```

*Figure 7-20 DB2 Instance Owning Node after your selections*

7. There might be a warning error as shown in Figure 7-21.

```
WARNING: Volume group: vg_TE1_sap on node: sapsma2 does not exist, but is
set to auto import in resource group: db2te1_ResourceGroup.

WARNING: Filesystem /db2/TE1 on node sapsma2 does not exist for resource
group: db2te1_ResourceGroup.
Resource group db2te1_ResourceGroup is set to automatically import.
....
```

*Figure 7-21 Cluster verification message*

This message means that the volume group is automatically imported by the synchronization.

8. A new PowerHA resource group, `db2te1_ResourceGroup`, is created. The volume group `vg_TE1_sap` and the service IP label `sapsmasvc4` are automatically added to the resource group as shown in Example 7-23.

*Example 7-23 Configured resource group for the SAP DB2 instance*

```
root@sapsma1 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_NFS
db2te1_ResourceGroup

root@sapsma1 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4 db2te1_ApplicationServer"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TE1 /export/saptrans/TE1"
FILESYSTEM=" "
FORCED_VARYON="false false"
FSCHECK_TOOL="fsck fsck"
FS_BEFORE_IPADDR="true false"
MOUNT_FILESYSTEM="/sapmnt/TE1;/export/sapmnt/TE1
/usr/sap/trans;/export/saptrans/TE1"
RECOVERY_METHOD="sequential parallel"
SERVICE_LABEL="sapsmasvc2 sapsmasvc4"
SSA_DISK_FENCING="false false"
VG_AUTO_IMPORT="false true"
VOLUME_GROUP="vg_TE1_nfs vg_TE1_sap"
USERDEFINED_RESOURCES=""
```

9. In this configuration, we have to change the file system recovery method.

> **Important:** We use nested file systems for this DB2 installation. In this case, the PowerHA documentation suggests that we must set the recovery method from parallel to sequential to guarantee the correct mount order. If we do not change this setting, a fatal error occurs about the mounting and the DB2 database cannot start.

Launch resource group settings by using the path for `sapsma1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Change/Show Resources and Attributes for a Resource Group**.

Select the line with the DB2 resource group as shown in Figure 7-22.

```
          Change/Show Resources and Attributes for a Resource Group

Move cursor to desired item and press Enter.

  SAP_RG_DVEBMGS02
  SAP_RG_NFS
  db2te1_ResourceGroup
```

*Figure 7-22 Change/Show Resources and Attributes for a Resource Group*

Change the file system recovery method from parallel to **sequential** as shown in Figure 7-23 on page 382.

```
 Change/Show All Resources and Attributes for a Resource Group

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

[TOP]                                                        [Entry Fields]
  Resource Group Name                                      db2te1_ResourceGroup
  Participating Nodes (Default Node Priority)              sapsma1 sapsma2

  Startup Policy                                         Online On Home Node Only
  Fallover Policy                    Fallover To Next Priority Node In The List
  Fallback Policy                                          Never Fallback

  Service IP Labels/Addresses                             [sapsmasvc4] +
  Application Controllers                            [db2te1_ApplicationServer] +

  Volume Groups                                           [vg_TE1_sap ] +
  Use forced varyon of volume groups, if necessary         false +
  Automatically Import Volume Groups                       false +

  Filesystems (empty is ALL for VGs specified)            [ ] +
  Filesystems Consistency Check                            fsck +
  Filesystems Recovery Method                              sequential +
  Filesystems mounted before IP configured                 false +
```

*Figure 7-23 Change/Show All Resources and Attributes for a Resource Group*

Then, press Enter.

10. Manually verify and synchronize at the end.

> **Import the volume groups:** If you did not import the volume groups for this resource group, you must import the volume groups now if you want to synchronize with the running cluster. Or, you have to stop the entire cluster and use the autocorrect functionality of PowerHA.

### 7.3.5 Starting Smart Assist for SAP: Application server instance (AS)

We describe how to start the Smart Assist for SAP with the application server instance (AS):

1. Launch Smart Assist for SAP by using the path for `sapsma1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.

2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-24), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  DB2 UDB non-DPF Smart Assist     # sapsma1 sapsma2
  SAP MaxDB Smart Assist           # sapsma1 sapsma2
  SAP Smart Assist                 # sapsma1 sapsma2
```

*Figure 7-24 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-25), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-25 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-26), select **4. SAP NW 7.0 AS Instance**.

```
             Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  4.SAP NW 7.0 AS Instance         # sapsma1
  5.SAP Database Instance          # sapsma1 sapsma2
  1.SAP NW 7.0 Global Filesystem   # sapsma1
```

*Figure 7-26 Selecting the configuration to create*

5. Select the line with your application instance (Figure 7-27).

```
                         Select an Application instance

Move cursor to desired item and press Enter.

DVEBMGS02
```

*Figure 7-27 Select an application instance*

6. By using the available pick lists (F4), edit the Takeover Nodes as shown in Figure 7-28, even if incorrect values are listed and then press Enter.

```
                    Add SAP Application Instance Details

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

                                                        [Entry Fields]
* SAP SYSTEM ID                                       TE1
* SAP Application Server Instance Name                DVEBMGS02
* SAP Application Server Instance No.                 02
* Application Name                                   [sap_DVEBMGS02_appserver]
* Primary Node                                        sapsma1
* Takeover Nodes                                     [sapsma2] +
* Service IP Label                                    sapsmasvc1
* Shared Volume Groups                               [vg_TE1_sapbin]
```

*Figure 7-28 Add SAP Application Instance Details*

7. A new PowerHA resource group, `SAP_RG_DVEBMGS02`, is created. The volume group `vg_TE1_sapbin` and the service IP label `sapsmasvc1` are automatically added to the resource group as shown in Example 7-24.

*Example 7-24 Configured resource group for the SAP DVEBMGS02 instance*

```
root@sapsma1 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_DVEBMGS02
SAP_RG_NFS
db2te1_ResourceGroup

root@sapsma1 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4 db2te1_ApplicationServer
sap_DVEBMGS02_appserver_AP_AS"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TE1 /export/saptrans/TE1"
FILESYSTEM="  "
FORCED_VARYON="false false false"
FSCHECK_TOOL="fsck fsck fsck"
FS_BEFORE_IPADDR="true false false"
MOUNT_FILESYSTEM="/sapmnt/TE1;/export/sapmnt/TE1
/usr/sap/trans;/export/saptrans/TE1"
RECOVERY_METHOD="sequential sequential sequential"
SERVICE_LABEL="sapsmasvc2 sapsmasvc4 sapsmasvc1"
SSA_DISK_FENCING="false false false"
VG_AUTO_IMPORT="false false false"
VOLUME_GROUP="vg_TE1_nfs vg_TE1_sap vg_TE1_sapbin"
USERDEFINED_RESOURCES=""
```

**Import the volume groups:** If you did not import the volume groups for this resource group, you have to import them now if you want to synchronize with the running cluster. Or, you have to stop the entire cluster and use the autocorrect functionality of PowerHA.

8. Change the resource group dependency from parent/child to Start After.

With the default setting, every move of the resource group for NFS service from one node to the other node manually or by an automatic takeover also restarts the central instance. To prevent this restart, remove the Parent/Child dependency (Figure 7-29) and create instead a Startafter dependency (Figure 7-30).

Follow the path: `smitty sysmirror` → **Cluster Applications and Resources → Resource Groups → Configure Resource Group Run-Time Policies → Configure Dependencies between Resource Groups → Configure Parent/Child Dependency → Remove Parent/Child Dependency between Resource Groups**.

```
          Select a Parent/Child Resource Group Dependency to Delete

 Move cursor to desired item and press Enter.

   # Parent                          Child
   SAP_RG_NFS                        SAP_RG_DVEBMGS02
```

*Figure 7-29 Remove a parent/child dependency*

Select the only configured line and press Enter. Or, you can use the command line as shown in Example 7-25.

*Example 7-25 Command to remove a parent/child resource group dependency*

```
/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'PARENT_CHILD' \
   -d \
   -p'SAP_RG_NFS' \
   -c'SAP_RG_DVEBMGS02'
```

Follow this path: `smitty sysmirror` → **Cluster Applications and Resources → Resource Groups → Configure Resource Group Run-Time Policies → Configure Dependencies between Resource Groups → Configure Start After Resource Group Dependency → Add Start After Resource Group Dependency**.

```
                       Select the Source Resource Group

 Move cursor to desired item and press Enter.

SAP_RG_DVEBMGS02
SAP_RG_NFS
db2te1_ResourceGroup
```

*Figure 7-30 Select the Source Resource Group*

In Figure 7-30, select the **SAP_RG_DVEBMGS02** as the source resource group (Figure 7-31 on page 386) and press Enter.

```
                        Select the Target Resource Group

 Move cursor to desired item and press F7.
     ONE OR MORE items can be selected.
 Press Enter AFTER making all selections.

   SAP_RG_NFS
   db2te1_ResourceGroup
```

*Figure 7-31 Select the Target Resource Group*

Select the **SAP_RG_NFS** as the target resource group (Figure 7-32) and press Enter.

```
                   Add Start After Resource Group Dependency

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

                                                        [Entry Fields]
* Source Resource Group                  [SAP_RG_DVEBMGS02]
* Target Resource Group                   [SAP_RG_NFS]
```

*Figure 7-32 Add Start After Resource Group Dependency*

Create the dependency by pressing Enter again. Or, you can use the command line in Example 7-26.

*Example 7-26 Command to add a Start_After resource group dependency*

```
/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'START_AFTER' \
   -a \
   -c'SAP_RG_DVEBMGS02' \
   -p'SAP_RG_NFS'
```

## 7.3.6 Completing the configuration

After the Smart Assist for SAP is finished, complete the configuration:

1. Stop the DB2 instance on the primary node as shown in Example 7-27. Remember that it is active only during the Smart Assist for DB2 discovery process.

*Example 7-27 Stopping the DB2 instance*

```
sapsma1:/ # su - db2te1

sapsma1:db2te1 1> db2stop
03/24/2012 12:02:56     0   0   SQL1064N  DB2STOP processing was successful.
SQL1064N  DB2STOP processing was successful.
```

2. Unmount the shared file systems as shown in Example 7-28.

*Example 7-28 Unmounting the shared file systems*

```
sapsma1:/ # lsvg -l vg_TE1_sap
vg_TE1_sap:
LV NAME             TYPE LPs     PPs     PVs  LV STATE      MOUNT POINT
TE1.lvdb2binary     jfs2 48      48      1 closed/syncd    /db2/TE1
TE1.lvdb2logdir     jfs2 64      64      1 closed/syncd    /db2/TE1/log_dir
TE1.lvdb2logarc     jfs2 32      32      1 closed/syncd    /db2/TE1/log_archive
TE1.lvdb2dat.01     jfs2 160     160     1 closed/syncd    /db2/TE1/sapdata1
TE1.lvdb2dat.02     jfs2 160     160     1 closed/syncd    /db2/TE1/sapdata2
TE1.lvdb2dat.03     jfs2 160     160     1 closed/syncd    /db2/TE1/sapdata3
TE1.lvdb2dat.04     jfs2 160     160     1 closed/syncd    /db2/TE1/sapdata4
TE1.lvsapmnt_o      jfs2 19      19      1 closed/syncd    /sapmn__old/TE1
TE1.lvsaptran_o     jfs2 2       2       1 closed/syncd    /usr/sap/tran__old
TE1.lvdb2bin        jfs2 32      32      1 closed/syncd    /db2/db2te1
```

3. Deactivate the shared volume group as shown in Example 7-29.

*Example 7-29 Deactivating the shared volume group of vg_TE1_sap*

```
sapsma1:/ # varyoffvg vg_TE1_sap

sapsma1:/ # lsvg -o
caavg_private
rootvg
```

4. Synchronize the PowerHA cluster by using SMIT:
   a. Follow the path: **`smitty sysmirror`** → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.
   b. In the PowerHA SystemMirror Verification and Synchronization panel (Figure 7-33), press Enter to accept the default option.

```
           PowerHA SystemMirror Verification and Synchronization

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Verify, Synchronize or Both                       [Both]           +
* Include custom verification library checks        [Yes]            +
* Automatically correct errors found during         [Yes]            +
  verification?
* Force synchronization if verification fails?      [No]             +
* Verify changes only?                              [No]             +
* Logging                                           [Standard]       +
```

*Figure 7-33 Accepting the default actions on the Verification and Synchronization panel*

5. Start the cluster on both nodes `sapsma1` and `sapsma21` by running **`smitty clstart`**.
6. In the Start Cluster Services panel (Figure 7-34 on page 388), complete these steps:
   a. For Start now, on system restart or both, select **now**.
   b. For Start Cluster Services on these nodes, enter [`sapsma1 sapsma2`].
   c. For Manage Resource Groups, select **Automatically**.
   d. For BROADCAST message at startup, select **false**.

e. For Startup Cluster Information Daemon, select **true**.
f. For Ignore verification errors, select **false**.
g. For Automatically correct errors found during cluster start, select **yes**.

Press Enter.

```
                              Start Cluster Services

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Start now, on system restart or both                now                      +
  Start Cluster Services on these nodes               [sapsma1 sapsma2]        +
* Manage Resource Groups                              Automatically            +
  BROADCAST message at startup?                       false                    +
  Startup Cluster Information Daemon?                 true                     +
  Ignore verification errors?                         false                    +
  Automatically correct errors found during           yes                      +
  cluster start?
```

*Figure 7-34 Specifying the options for starting cluster services*

> **Tip:** The log file for the Smart Assist is in the `/var/hacmp/log/sa.log` file. You can use the **`clmgr`** utility to easily view the log, as in the following example:
>
> `clmgr view log sa.log`

When the PowerHA cluster starts, the DB2 instance is automatically started. The application monitors start after the defined stabilization interval as shown in Example 7-30.

*Example 7-30 Checking the status of the highly available cluster and SAP and DB2 instances*

```
root@sapsma1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                        Node
-----------------------------------------------------------------------------
SAP_RG_NFS     ONLINE                       sapsma1
               OFFLINE                      sapsma2

db2te1_Resourc ONLINE                       sapsma1
               OFFLINE                      sapsma2

SAP_RG_DVEBMGS ONLINE                       sapsma1
               OFFLINE                      sapsma2
```

Your SAP instance and DB2 database are now configured for high availability in a PowerHA SystemMirror configuration.

## 7.4 Scenario for a complete SAP and liveCache environment

We provide the scenario for a complete setup of the SAP and liveCache environment.

## 7.4.1 Overview

In this scenario, we want to illustrate the preferred environment setup for an SAP SCM/APO system that includes liveCache. This outlined solution exists within three separate clusters that accommodate each part of the entire solution. Through these separate individual components, the solution can be implemented, configured, and operated independently. A faster takeover is also a result of this implementation in the event of a node failure.

**Tip:** Consider the use of a virtual IP address for all SAP application server and instances to easily relocate them for maintenance and later SAP automation through SAP landscape virtualization manager (LVM).

The concept also provides two separate networks to isolate the communication between the server itself, and the normal communication between the clients and the servers. This concept also hides some parts of the cluster resources from the clients that are not necessary for the clients.

All needed users and groups have to be created with the same ID. To simplify the environment, all needed users and groups for the three-cluster solution are created on all six nodes. But for higher security, the users and groups must be created only on the systems where they are necessary.

Figure 7-35 is an overview of the components in this three-cluster solution.

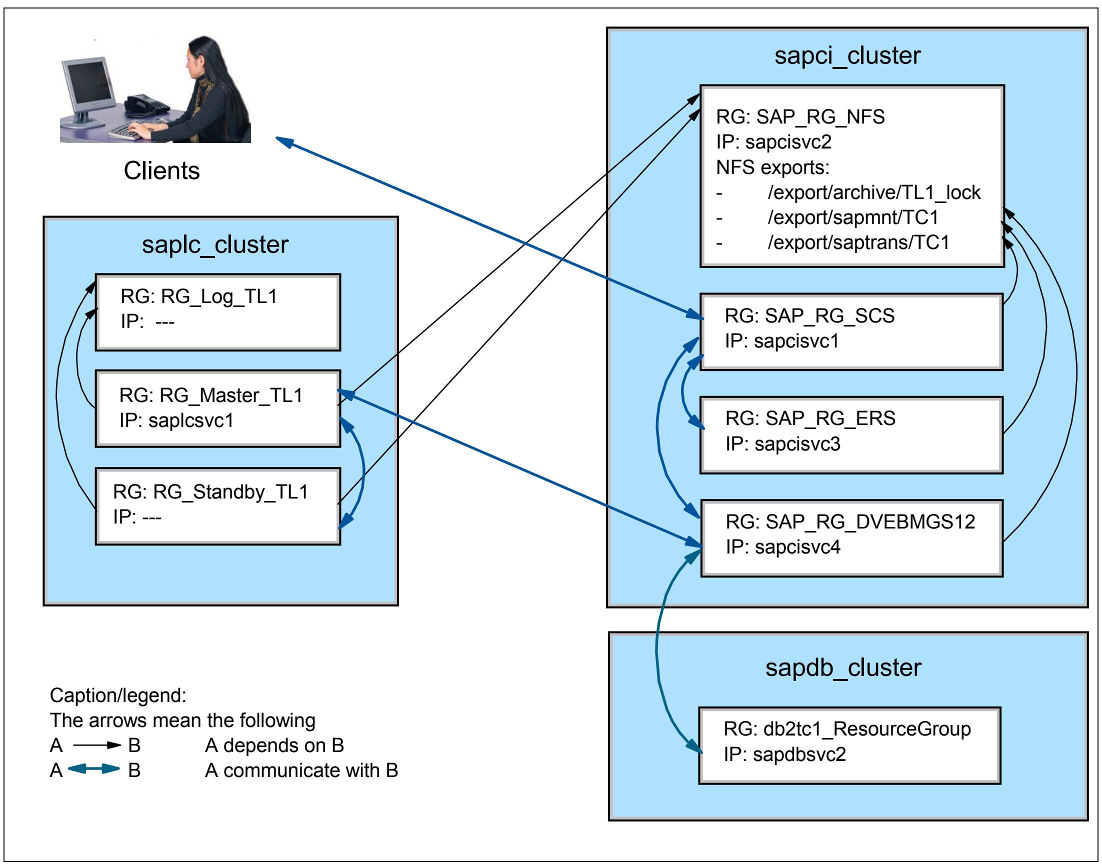



*Figure 7-35 Overview of the scenario for a complete three-cluster solution for SAP and liveCache*

A second variant to this drafted solution is the local installation of an application server on every PowerHA node and to not illustrate only one primary application server inside the cluster that is taken over in a failure situation. The advantage of the solution is easier administration of only one instance inside the cluster. The disadvantage is the extended takeover time through the start of the primary application server on the backup node. This extended time influences the solution because the ABAP system, which is a double-stack SAP system, needs to start on the second node. Due to the JAVA stack, it needs a long time to be fully activated.

In both variants, we do not consider the additional application servers, which are necessary in a high availability concept. Because these application servers are installed outside the cluster on separate systems or LPARs, we do not describe them in this book. These servers can be installed as usual. In the cluster solution with Smart Assist for SAP, we do not describe the start or stop procedure of these additional application servers. This process must be manually performed outside the cluster.

### 7.4.2 Assignment of the three clusters

The clusters have the following assignments:

- Cluster 1: `sapci_cluster`

  In this cluster environment, the entire SAP instances are installed and controlled to enumerate an ASCS, an ERS, and a primary application server. This cluster houses an NFS service, which is a mandatory condition for the use of the Smart Assist for SAP. It uses only NFS version 4.

  So that we do not have to set up an additional NFS server for the demonstration environment, the NFS server from the Smart Assist for SAP is also applied for the lock management of the hot-standby liveCache system. This approach is not usual and not required for a normal environment. It can differ from solution to solution. This part can also be handled by an external NFS service.

- Cluster 2: `sapdb_cluster`

  In this cluster, the database is deployed and the SAP system connects to `sapcisvc4`. In the demonstration environment, one DB2 database is installed that uses shared disks between the cluster nodes for failover.

  This solution is easy to implement and administer, but it is not optimized for fast takeover times. In a takeover, the remaining node needs time to recover.

  If this solution does not meet your time constraints, a DB2 HADR solution is implemented in 7.8, “DB2 HADR cluster solution” on page 463 that permits shorter takeover times. This alternative can be substituted for the cluster 2 solution, but it is not a Smart Assist solution. It uses the scripts that are implemented by the International SAP IBM Competence Center (ISICC) team. The scripts are described in the “Cluster scripts for DB2 HADR” on page 645.

- Cluster 3: `saplc_cluster`

  This cluster implements a hot-standby liveCache solution. Therefore, two MaxDB databases with local not shared disks for the data area and concurrent shared disks for the log area are deployed. Both the data area and the log area have to stay on raw devices. This Smart Assist solution also implements an accelerated reestablishment of the standby functionality of the backup node. The data area of the backup node is cloned through IBM FlashCopy® from the active system if this reactivated system is not in synchronization with the running system.

In the demonstration environment, a DS8000 is selected. If you are interested in an implementation with the SAN Volume Controller and want more information, see `PowerHA7.1.1_HotStandbyImplementationPath_v1.pub.pdf` (version at press date), which is published by the ISICC:

http://w3-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100677

## 7.4.3 Considerations and preconditions of this solution

The following considerations and prerequisites relate to implementing this solution:

- The use of the NFS service for the Smart Assist for SAP is *required* with this version of PowerHA. It prevents you from setting up a second cluster with Smart Assist in a three-system SAP landscape with development, quality assurance, and productive system, except that you use a remote transport system within SAP. The Smart Assist for SAP subtype NFS only supports *NFS version 4*.
- There is no Smart Assist implementation for DB2 HADR. However, the use of the solution without Smart Assist that is shown in “DB2 HADR cluster solution” on page 463 is possible.
- A cluster with the liveCache hot-standby solution is only possible with the SAN Volume Controller, V7000, or DS8000 storage systems. The liveCache MaxDB can only use raw devices.

## 7.4.4 Installation path for the entire three-cluster environment

We describe the installation path for the entire three-cluster environment.

### Path for the installation and configuration steps before you use Smart Assist

Table 7-5 shows the path for installation and configuration steps for the Smart Assist. In this table, we show the best sequence/path to set up the whole three-cluster environment. The number behind the cross reference lists the step number inside the referenced chapter. We show with this table that some steps have to/must be done at the same time on both nodes in the three clusters and which steps have to follow each other.

*Table 7-5 Path for installation and configuration steps before you use Smart Assist*

| sapci_cluster | | sapdb_cluster | | saplc_cluster | |
|---|---|---|---|---|---|
| **sapci1** | **sapci2** | **sapdb1** | **sapdb2** | **saplc1** | **saplc2** |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-1.) | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-1.) | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431-1.) | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431-1.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-1.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-1.) |
| | | | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-2.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-2.) |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-2.) | | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431 - 2.) | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442 - 3.) | |

| sapci_cluster | | sapdb_cluster | | saplc_cluster | |
|---|---|---|---|---|---|
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 3.) + 4.) | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 3.) + 4.) | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431 3.) + 4.) | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431 3.) + 4.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442 4.) + 5.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442 4.) + 5.) |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-5.) | | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431-5.) | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-6.) | |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-6.) | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-6.) | | | | |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-7.) | | | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-7.) | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-7.) |
| | | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431-6.) | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-8.) | |
| | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-8.) | | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431 7.) + 8.) | | |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395-9.) | | 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431 9.) - 11.) | | | |
| | | | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442-9.) | |
| | | | | | 7.7.4, “Installation and configuration steps before you use Smart Assist for MaxDB” on page 442 10.) + 11.) |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 10.) | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 10.) | | | | |

| sapci_cluster | | sapdb_cluster | | saplc_cluster | |
|---|---|---|---|---|---|
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 11.) | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 12.) + 13.) | | | | |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 14.) + 15.) | | | | | |
| | 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 16.) + 17.) | | | | |
| 7.5.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 395 18.) + 19.) | | | | 7.7.5, “Preliminary steps” on page 451 | |

## Installation path for starting Smart Assists

Table 7-6 shows the installation path for starting Smart Assist.

*Table 7-6 Installation path for starting Smart Assist*

| sapci_cluster | | sapdb_cluster | | saplc_cluster | |
|---|---|---|---|---|---|
| **sapci1** | **sapci2** | **sapdb1** | **sapdb2** | **saplc1** | **saplc2** |
| 7.5.3, “Starting the Smart Assist for SAP: Global file system (GFS)” on page 406 | | | | | |
| 7.5.4, “Starting Smart Assist for SAP: Central services (SCS)” on page 413 | | | | | |
| 7.5.5, “Starting Smart Assist for SAP: Enqueue replication server instance (ERS)” on page 416 | | | | | |
| 7.5.6, “Starting Smart Assist for SAP: Application server instance (AS)” on page 421 | | | | | |

| sapci_cluster | | sapdb_cluster | | saplc_cluster | |
|---|---|---|---|---|---|
| 7.5.7, “Completing the configuration” on page 425 | | | | | |
| | | 7.6.3, “Starting Smart Assist for SAP: Database instance (DB)” on page 435 | | | |
| | | 7.6.4, “Completing the configuration” on page 437 | | | |
| | | | | 7.7.6, “Starting SAP liveCache Hot Standby Wizard” on page 453 | |
| | | | | 7.7.7, “Configuring the cluster by using the Smart Assist” on page 458 | |
| | | | | 7.7.8, “Verifying the Smart Assist settings” on page 460 | |
| | | | | 7.7.9, “Completing the configuration” on page 462 | |

# 7.5 Cluster 1: SAP Supply Chain Management

We provide an overview of the SAP Supply Chain Management (SCM) cluster 1.

## 7.5.1 Overview

This cluster uses the Smart Assist for SAP: GFS, SCS, ERS, and AS to make an SAP SCM system highly available. Figure 7-36 on page 395 presents the cluster solution.

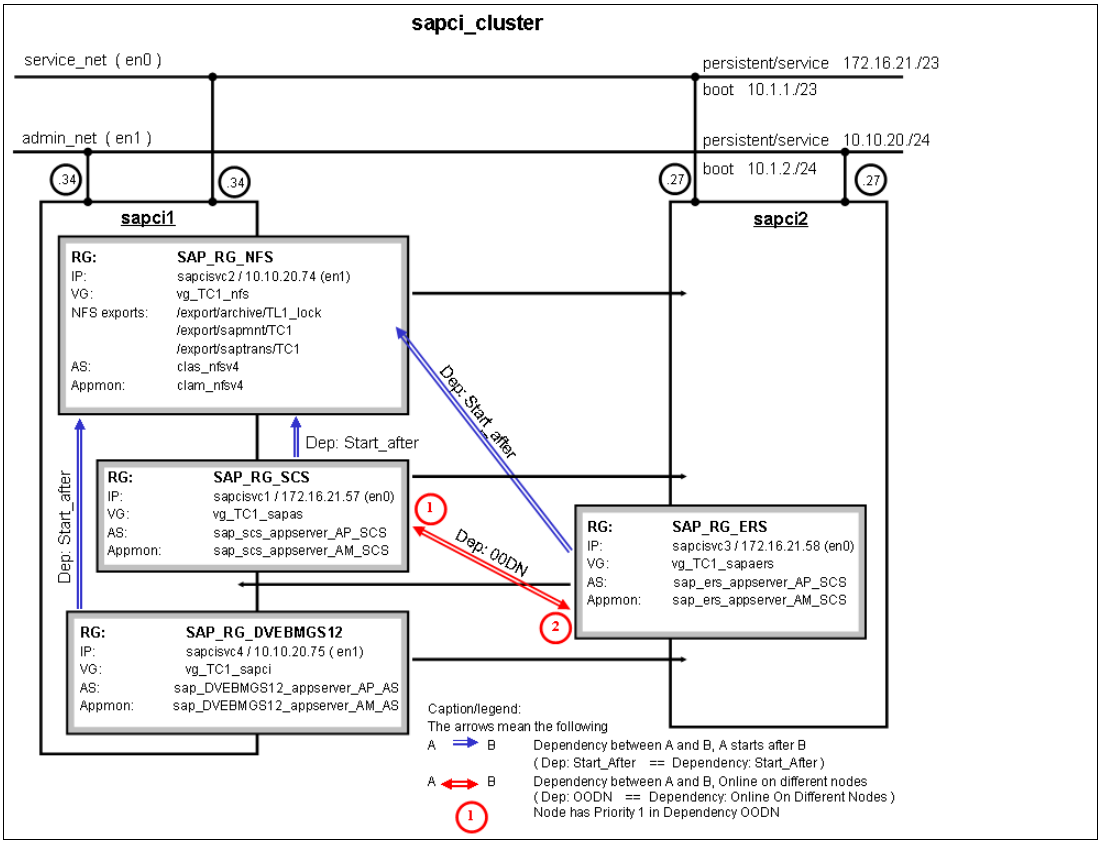



*Figure 7-36 Schema overview for sapci_cluster*

## 7.5.2 Installation and configuration steps before you use Smart Assist for SAP

Cluster 1 in this complex environment uses the Smart Assist for SAP with subtypes GFS, SCS (in particular for ASCS), ERS, and AS (for primary application server instance). We explain the preliminary required steps before you start the Smart Assist for SAP. Then, we explain how to start the Smart Assist for SAP. Follow these steps:

1. Install the required filesets.

   See 7.2.1, “Installing the required PowerHA SystemMirror Smart Assist filesets” on page 362.

2. Configure the base IBM PowerHA SystemMirror.

   You must configure the topology of the PowerHA cluster before you use the Smart Assist for SAP. Example 7-31 shows the cluster `sapci_cluster` that is configured with two Ethernet interfaces in each node.

   *Example 7-31 Cluster network configuration*

```
root@sapci1 / # lscluster -c
Cluster query for cluster sapci_cluster returns:
Cluster uuid: 62856a40-7209-11e1-81c5-2e4799b2bf6f
Number of nodes in cluster = 2
        Cluster id for node sapci1b1 is 1
        Primary IP address for node sapci1b1 is 10.1.1.34
        Cluster id for node sapci2b1 is 2
        Primary IP address for node sapci2b1 is 10.1.1.27
Number of disks in cluster = 0
```

```
Multicast address for cluster is 228.1.1.24

root@sapci1 / # cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address
Interface Name Global Name Netmask Alias for HB Prefix Length
sapci1b2 boot admin_net  ether public sapci1 10.1.2.34 en1 255.255.255.0 24
sapci1b1 boot service_net ether public sapci1 10.1.1.34 en0 255.255.254.0 23
sapci2b2 boot admin_net ether public sapci2 10.1.2.27 en1 255.255.255.0 24
sapci2b1 boot service_net ether public sapci2 10.1.1.27 en0 255.255.254.0 23
```

3. Configure the network.

   The IP addresses for the NFS server (`sapcisvc2`) and the SAP system (`sapcisvc1`) are configured as an alias on the first network of the first node `sapci1` where SAP AS and that primary application server are installed. The NFS service address, as shown in Example 7-32, is configured on a separated network also as an alias.

   *Example 7-32 Network settings for sapci1*

```
root@sapci1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.99.b2.bf.6f  7753726     0  3251371     0     0
en0   1500  10.1        sapci1b1           7753726     0  3251371     0     0
en0   1500  172.16.20   sapci1p1           7753726     0  3251371     0     0
en0   1500  172.16.20   sapcisvc1          7753726     0  3251371     0     0
en1   1500  link#3      2e.47.99.b2.bf.70   286960     0     9661     0     0
en1   1500  10.1.2      sapci1b2            286960     0     9661     0     0
en1   1500  10.10.20    sapci1p2            286960     0     9661     0     0
en1   1500  10.10.20    sapcisvc2           286960     0     9661     0     0
lo0   16896 link#1                         2462408     0  2462407     0     0
lo0   16896 127         loopback           2462408     0  2462407     0     0
lo0   16896 loopback                       2462408     0  2462407     0     0
```

   The IP addresses for the SAP ERS system (`sapcisvc3`) are configured on the second node (`sapci2`) where the SAP ERS is installed. See Example 7-33.

   *Example 7-33 Network settings for sapci2*

```
root@sapci2 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c8.81.6f  5070991     0  1212018     0     0
en0   1500  10.1        sapci2b1           5070991     0  1212018     0     0
en0   1500  172.16.20   sapci2p1           5070991     0  1212018     0     0
en0   1500  172.16.20   sapcisvc3          5070991     0  1212018     0     0
en1   1500  link#3      6e.8d.dd.c8.81.70   302567     0     7542     0     0
en1   1500  10.1.2      sapci2b2            302567     0     7542     0     0
en1   1500  10.10.20    sapci2p2            302567     0     7542     0     0
lo0   16896 link#1                         1690441     0  1690441     0     0
lo0   16896 127         loopback           1690441     0  1690441     0     0
lo0   16896 loopback                       1690441     0  1690441     0     0
```

4. Create the users.

   See 7.2.2, “Creating the users” on page 362. The script to create the users is in the “Script to create users and groups for sapci_cluster, sapdb_cluster, and saplc_cluster” on page 604.

5. Create the volume groups, logical volumes, and file systems.

See 7.2.3, “Creating volume groups, logical volumes, and file systems” on page 363 and validate them on both nodes. The scripts to create the volume groups and file systems are in “Script to create VGs and LVs for sapci_cluster” on page 606. See Example 7-34.

> **Prerequisite:** `/export/sapmnt/<SID>` and `/export/saptrans/<SID>` are fixed naming conventions. Use these naming conventions because the Smart Assist can only find them by using these naming conventions.

*Example 7-34 Volume groups at sapci_cluster*

```
root@sapci1 / # lsvg -l vg_TC1_nfs
vg_TC1_nfs:
LV NAME             TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TC1.lvsapmnt        jfs2 64  64  1    open/syncd    /export/sapmnt/TC1
TC1.lvsaptrans      jfs2 32  32  1    open/syncd    /export/saptrans/TC1
TL1.lvlock          jfs2 10  10  1    open/syncd    /export/archive/TL1_lock

root@sapci1 / # lsvg -l vg_TC1_sapas
vg_TC1_sapas:
LV NAME             TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TC1.lvusrsap        jfs2 160 160 1    open/syncd    /usr/sap/TC1/ASCS10

root@sapci1 / # lsvg -l vg_TC1_sapers
vg_TC1_sapers:
LV NAME             TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TC1.lvusrsapers     jfs2 10  10  1    closed/syncd  /usr/sap/TC1/ERS11
root@sapci1 / # lsvg -l vg_TC1_sapci
vg_TC1_sapci:
LV NAME             TYPE LPs PPs PVs  LV STATE      MOUNT POINT
TC1.lvusrsapci      jfs2 160 160 1    open/syncd    /usr/sap/TC1/DVEBMGS12
TC1.lvdb2binary     jfs2 128 128 1    open/syncd    /db2/TC1
```

> **Directories:** The directories `/usr/sap` and `/db2` must be created locally as discrete file systems.

6. Configure NFS.

   An NFS domain has to be configured on both nodes because the Smart Assist for SAP uses the NFS4 version.

   `root@sapci1 / # clcmd chnfsdom sapci_cluster`

   Set these necessary settings for NFS Smart Assist:

   - `stopsrc -s nfsd`
   - `smitty nfsgrcperiod`
   - `startsrc -s nfsd`

   Create a cluster exports file with the necessary mounts as shown in Figure 7-37 on page 398.

```
/export/sapmnt/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:
sapcisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/saptrans/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:
sapcisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/archive/TL1_lock
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=saplc1:saplc2:saplc1b2:saplc2b2:
saplcsvc1:saplcsvc2:saplc1p2:saplc2p2
```

*Figure 7-37 cat /usr/es/sbin/cluster/etc/exports*

Copy this file to the second node:

```
root@sapci1 / # clrcp /usr/es/sbin/cluster/etc/exports
sapci2:/usr/es/sbin/cluster/etc/exports
```

The NFS server exports all needed file systems as shown in Example 7-35.

*Example 7-35 NFS server on sapci1*

```
root@sapci1 / # exportfs -a

root@sapci1 / # exportfs
/export/sapmnt/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/saptrans/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/archive/TL1_lock
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=saplc1:saplc2:saplc1b2:saplc2b2:sap
lcsvc1:saplcsvc2:saplc1p2:saplc2p2
```

Mount as the NFS client:

```
root@sapci1 / # mount -o vers=4,hard,intr sapcisvc2:/export/sapmnt/TC1
/sapmnt/TC1
root@sapci1 / # mount -o vers=4,hard,intr sapcisvc2:/export/saptrans/TC1
/usr/sap/trans
```

The following NFS client mounts are available as shown in Example 7-36.

*Example 7-36 NFS mounts on sapci1*

```
root@sapci1 / # mount|grep nfs
sapcisvc2 /export/sapmnt/TC1 /sapmnt/TC1      nfs4   Apr 05 06:44
vers=4,hard,intr
sapcisvc2 /export/saptrans/TC1 /usr/sap/trans nfs4   Apr 05 06:45
vers=4,hard,intr
```

7. Install the SAP ASCS instance on the first node.

   We installed SAP normally with a virtual SAP host name (`sapcisvc1`). Only the normal SAP parameters are set as shown in Example 7-37 on page 399.

*Example 7-37 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc1
```

Start the ASCS installation as shown in Figure 7-38. Select **ASCS Instance**.

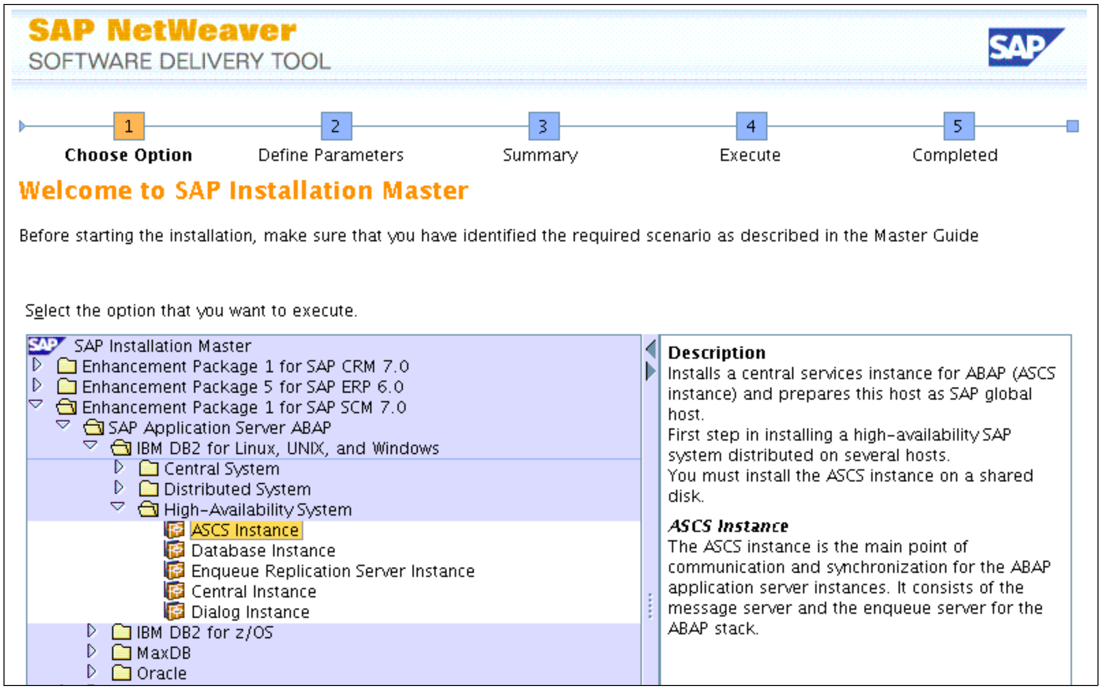



*Figure 7-38 SAPinst ASCS installation screen capture 1: Start installation*

For a detailed view of the complete installation process, see “sapci_cluster: Installing the SAP ASCS instance on the first node” on page 609.

8. Install the SAP enqueue replication service instance on the second node.

   For the ERS installation with SAPinst, a virtual SAP host name (`sapcisvc3`) is also used. It is mandatory from SAP in case it is cluster controlled, but this situation is not the default for the Smart Assist for SAP. Therefore, we set up the ERS without a service IP.

   The following SAP parameters are set as shown in Example 7-38.

*Example 7-38 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc3
```

Start the ERS installation as shown in Figure 7-39 on page 400. Select **Enqueue Replication Server Instance**.

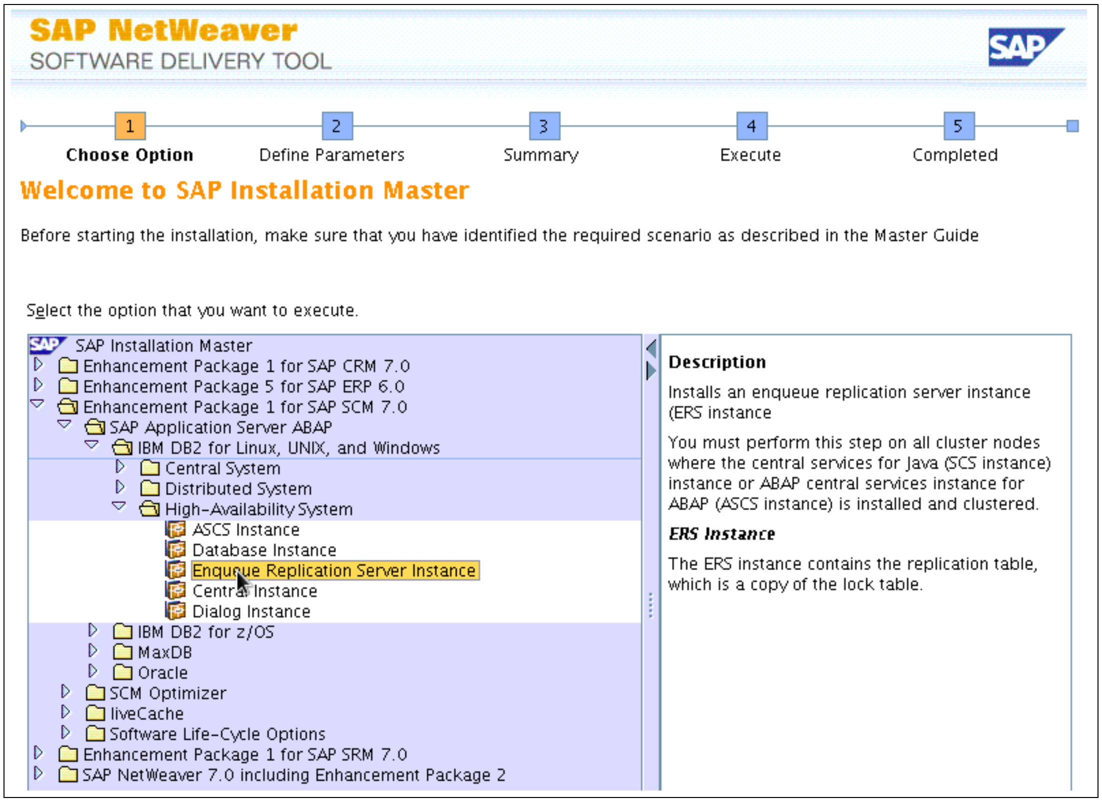



*Figure 7-39 SAPinst ERS installation screen capture 1: Start installation*

For a detailed view of the complete installation process, see Appendix E: “sapci_cluster: Installing the SAP Enqueue Replication Service (ERS) instance on the second node” on page 623.

So that the Smart Assist for SAP later can identify this virtualized installation, we have to create links in the profile directory. This task is not necessary for the SAP functionality because the replicated enqueue is installed with a virtual IP address.

> **Important:** The Smart Assist waits for profiles for the ERS with the host name at the end, not with a service address.

We create these links:

```
root@sapci1 / # cd /sapmnt/TC1/profile
root@sapci1 /sapmnt/TC1/profile # ln -s START_ERS11_sapcisvc3
START_ERS11_sapci1
root@sapci1 /sapmnt/TC1/profile # ln -s START_ERS11_sapcisvc3
START_ERS11_sapci2
root@sapci1 /sapmnt/TC1/profile # ln -s TC1_ERS11_sapcisvc3 TC1_ERS11_sapci1
root@sapci1 /sapmnt/TC1/profile # ln -s TC1_ERS11_sapcisvc3 TC1_ERS11_sapci2
```

9. Install the SAP primary application server instance on the first node.

   For the primary application server installation with `sapinst`, a virtual SAP host name (`sapcisvc4`) is used. The following SAP parameters are set as shown in Example 7-39 on page 401.

*Example 7-39 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc4
```

Start the installation as shown in Figure 7-40. Select **Central Instance**.

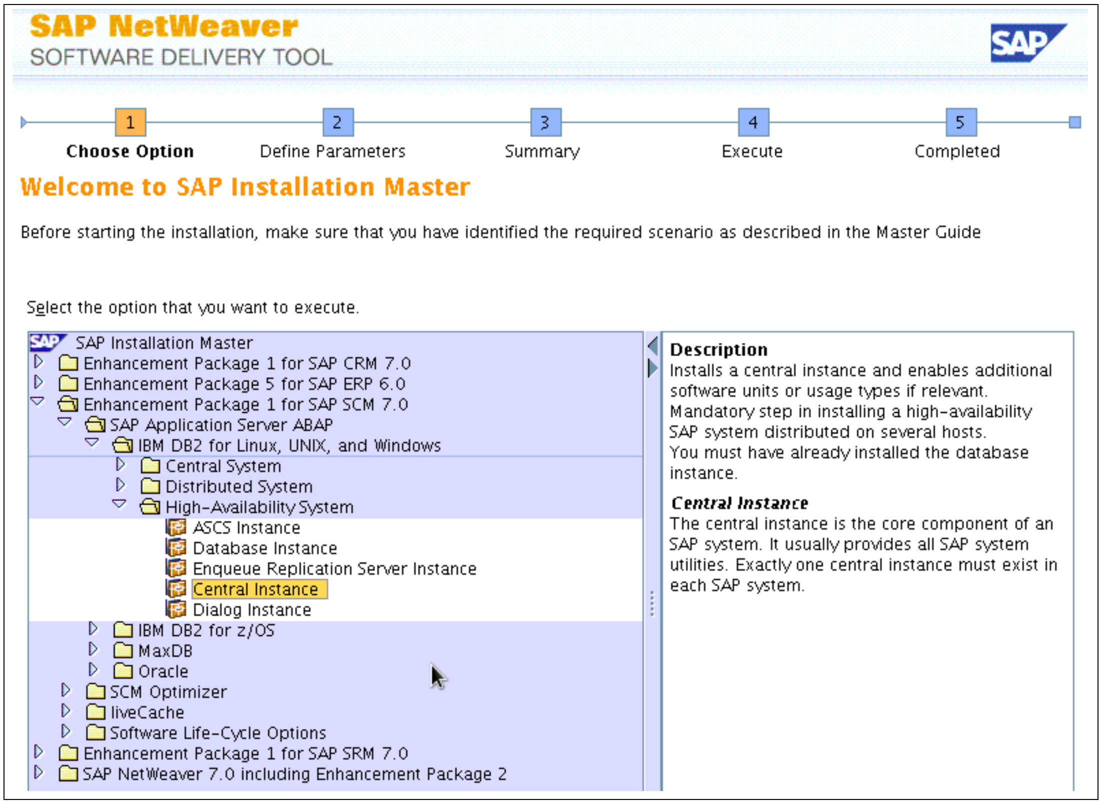



*Figure 7-40 SAPinst CI installation screen capture 1: Start installation*

For a detailed view of the complete installation process, see "sapci_cluster: Installing the SAP central instance on the first node" on page 626.

10. Stop SAP on both nodes.

    The SAP system has to be stopped before you import the volume groups on the node:

    – Stop SAP on node sapci1:
      - `root@sapci1 / # su - tc1adm`
      - `sapci1:tc1adm 1> stopsap sapcisvc4`
      - `sapci1:tc1adm 1> stopsap sapcisvc1`

    – Stop SAP on node sapci2:
      - `root@sapci1 / # su - tc1adm`
      - `sapci1:tc1adm 1> stopsap sapcisvc3`

11. Change the SAP profile parameter to meet the high availability needs.

    Change the Start Option for the SAP enqueue server from "Restart" to "Start" as shown in Figure 7-41 on page 402.

```
[...]
#-----------------------------------------------------------------------
# Start SAP message server
#-----------------------------------------------------------------------
_MS = ms.sap$(SAPSYSTEMNAME)_$(INSTANCE_NAME)
Execute_01 = local rm -f $(_MS)
Execute_02 = local ln -s -f $(DIR_EXECUTABLE)/msg_server$(FT_EXE) $(_MS)
Restart_Program_00 = local $(_MS) pf=$(_PF)
#-----------------------------------------------------------------------
# Start SAP enqueue server
#-----------------------------------------------------------------------
_EN = en.sap$(SAPSYSTEMNAME)_$(INSTANCE_NAME)
Execute_03 = local rm -f $(_EN)
Execute_04 = local ln -s -f $(DIR_EXECUTABLE)/enserver$(FT_EXE) $(_EN)
Start_Program_01 = local $(_EN) pf=$(_PF)
[...]
```

*Figure 7-41 Edit /sapmnt/TC1/profile/START_ASCS10_sapcisvc1*

Change the following parameters in the DEFAULT profile as shown in Figure 7-42.

```
[...]
enque/deque_wait_answer = TRUE
enque/con_retries = 120
ms/conn_timeout = 1800
[...]
```

*Figure 7-42 Edit /sapmnt/TC1/profile/DEFAULT.PFL*

Also, disable polling, enable all application servers for automatic reconnect, and enable ERS to be started for the correct ASCS instance. In some SAPinst versions, there are errors with these settings.

For more information, see these websites:

- http://www.ibm.com/developerworks/wikis/display/WikiPtype/Advanced+HA+Implementation
- http://help.sap.com/saphelp_nw70ehp1/helpdata/en/47/e0208d86983c85e10000000a42189c/frameset.htm
- http://help.sap.com/saphelp_nw70ehp1/helpdata/en/47/e929cd3d7001cee10000000a421937/content.htm

12. Update the `/etc/services` file on the secondary node.

    See 7.2.4, “Updating the /etc/services file on the secondary node” on page 363.

13. Copy and merge the directories.

    For SAP to be started on the second cluster, you have to copy more files and directories manually to the other node. In the test environment, we copied several files and directories that are local on the first node to the second node:

    ```
    root@sapci1 / # scp -pr /usr/sap/ccms/* sapci2:/usr/sap/ccms/
    ```

**Important:** The command **`clrcp`** works with files only not with directories.

The /usr/sap/sapservices file has to be merged between the two nodes because on the first node only entries for ASCS and DVEBMGS exist, and only entries for ERS exist on the second node.

Figure 7-43 shows the file /usr/sap/sapservices on node sapci1 before it is merged.

```
#!/bin/sh
LIBPATH=/usr/sap/TC1/ASCS10/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/ASCS10/exe/sapstartsrv
pf=/usr/sap/TC1/SYS/profile/START_ASCS10_sapcisvc1 -D -u tc1adm
LIBPATH=/usr/sap/TC1/DVEBMGS12/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/DVEBMGS12/exe/sapstartsrv
pf=/usr/sap/TC1/SYS/profile/START_DVEBMGS12_sapcisvc1 -D -u tc1adm
```

*Figure 7-43 /usr/sap/sapservices on node sapci1*

Figure 7-44 shows the file /usr/sap/sapservices on node sapci2 before it is merged.

```
#!/bin/sh
LIBPATH=/usr/sap/TC1/ERS11/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/ERS11/exe/sapstartsrv
pf=/usr/sap/TC1/ERS11/profile/START_ERS11_sapcisvc3 -D -u tc1adm
```

*Figure 7-44 /usr/sap/sapservices on node sapci2*

Figure 7-45 shows the merged file /usr/sap/sapservices.

```
#!/bin/sh
LIBPATH=/usr/sap/TC1/ASCS10/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/ASCS10/exe/sapstartsrv
pf=/usr/sap/TC1/SYS/profile/START_ASCS10_sapcisvc1 -D -u tc1adm
LIBPATH=/usr/sap/TC1/DVEBMGS12/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/DVEBMGS12/exe/sapstartsrv
pf=/usr/sap/TC1/SYS/profile/START_DVEBMGS12_sapcisvc1 -D -u tc1adm
LIBPATH=/usr/sap/TC1/ERS11/exe:$LIBPATH; export LIBPATH;
/usr/sap/TC1/ERS11/exe/sapstartsrv
pf=/usr/sap/TC1/ERS11/profile/START_ERS11_sapcisvc3 -D -u tc1adm
```

*Figure 7-45 Merged file for both nodes*

14. Start the SAP ASCS and DVEBMGS instance on sapci1 if it is not already running as shown in Example 7-40.

*Example 7-40 Start SAP on sapci1*

```
sapci1:tc1adm 2> startsap sapcisvc1

Starting Startup Agent sapstartsrv
-----------------------------
OK
 Instance Service on host sapci1 started

starting SAP Instance ASCS10
------------------------------
```

```
 Startup-Log is written to /usr/sap/TC1/home/tc1adm/startsap_ASCS10.log
/usr/sap/TC1/ASCS10/exe/sapcontrol -prot NI_HTTP -nr 10 -function Start
 Instance on host sapci1 started

sapci1:tc1adm 4> startsap sapcisvc4

Checking db Database
------------------------------
 Database is running

Starting Startup Agent sapstartsrv
-----------------------------
OK
 Instance Service on host sapci1 started

starting SAP Instance DVEBMGS12
------------------------------
 Startup-Log is written to /usr/sap/TC1/home/tc1adm/startsap_DVEBMGS12.log
/usr/sap/TC1/DVEBMGS12/exe/sapcontrol -prot NI_HTTP -nr 12 -function Start
 Instance on host sapci1 started
```

15. Check whether the SAP instance is active (Example 7-41).

*Example 7-41 SAP is running on sapsma1*

```
sapci1:tc1adm 7> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 10
-function GetProcessList

05.04.2012 10:57:32
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
msg_server, MessageServer, GREEN, Running, 2012 04 05 10:54:30, 0:03:02,
14549130
enserver, EnqueueServer, GREEN, Running, 2012 04 05 10:54:30, 0:03:02, 4653292

sapci1:tc1adm 5> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 12
-function GetProcessList

05.04.2012 10:56:41
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
disp+work, Dispatcher, GREEN, Running, Message Server connection ok, Dialog
Queue time: 0.00 sec, 2012 04 05 10:55:07, 0:01:34, 13369350
igswd_mt, IGS Watchdog, GREEN, Running, 2012 04 05 10:55:07, 0:01:34, 8388776
```

16. Start the SAP ERS instance on sapci2 if it is not already running (Example 7-42).

*Example 7-42 Start SAP on sapci2*

```
sapci2:tc1adm 1> startsap sapcisvc3

Starting Startup Agent sapstartsrv
-----------------------------
```

```
OK
 Instance Service on host sapci2 started

starting SAP Instance ERS11
------------------------------
 Startup-Log is written to /usr/sap/TC1/home/tc1adm/startsap_ERS11.log
/usr/sap/TC1/ERS11/exe/sapcontrol -prot NI_HTTP -nr 11 -function Start
 Instance on host sapci2 started
```

17. Check whether the SAP ERS instance is active. See Example 7-43.

*Example 7-43 SAP is running on sapsma1*

```
sapci2:tc1adm 4> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 11
-function GetProcessList

05.04.2012 11:09:08
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
enrepserver, EnqueueReplicator, GREEN, Running, 2012 04 05 11:04:11, 0:04:57,
15401158
```

18. Check to see what instances the Smart Assist for SAP can discover (Example 7-44 and Example 7-45). You can check the possible types on the command line. All types have a **1** at the end. The possible parameters for the **cl_sapdiscover** command are **-t GFS/AS/SCS/ERS/DB**.

*Example 7-44 Checking the Smart Assist for SAP discover function on sapci1*

```
root@sapci1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t GFS
SAP Smart Assist:SAPNW_7.0:1.SAP NW 7.0 Global Filesystem:SAPNW_7.0_SAPGFS:1

root@sapci1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t AS
SAP Smart Assist:SAPNW_7.0:4.SAP NW 7.0 AS Instance:SAPNW_7.0_ASINSTANCE:1

root@sapci1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t SCS
SAP Smart Assist:SAPNW_7.0:2.SAP NW 7.0 SCS Instance:SAPNW_7.0_SCSINSTANCE:1
```

*Example 7-45 Checking the Smart Assist for SAP discover function on sapci2*

```
root@sapci2 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t ERS
SAP Smart Assist:SAPNW_7.0:3.SAP NW 7.0 ERS Instance:SAPNW_7.0_ERSINSTANCE:1
```

19. If an expected type appears, more information can be collected by setting an additional environment variable as shown in Example 7-46.

*Example 7-46 Setting enlarged output for cl_sapdiscover*

```
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # export VERBOSE_LOGGING="high"
root@sapsma1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t GFS
...
```

Resolve all errors until all instances can be discovered.

20. You can start the PowerHA Smart Assist on a running cluster or on a cluster in the INIT state. Both nodes must be in the same state. We suggest that you configure the cluster in the INIT state.

## 7.5.3 Starting the Smart Assist for SAP: Global file system (GFS)

After you complete the steps in the previous sections, you are ready to start the Smart Assist for SAP as explained in the following steps:

1. Launch Smart Assist for SAP by using the path for `sapci1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-46), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  SAP Smart Assist                 # sapci1 sapci2
```

*Figure 7-46 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-47), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-47 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-48), select **1. SAP NW 7.0 Global Filesystem**.

```
             Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  4.SAP NW 7.0 AS Instance         # sapci1
  3.SAP NW 7.0 ERS Instance        # sapci1 sapci2
  1.SAP NW 7.0 Global Filesystem   # sapci1 sapci2
  2.SAP NW 7.0 SCS Instance        # sapci1
```

*Figure 7-48 Selecting the configuration to create*

The Add SAP Global Filesystem Details window opens as shown in Figure 7-49 on page 407.

```
                      Add SAP Global Filesystem Details

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

                                                          [Entry Fields]
* SAP Global File System Owning Node                    sapci1
* Takeover Nodes                                       []
* Service IP Label                                     []
* Shared Volume Groups                                 [vg_TC1_nfs]
* Filesystems/Directories to Export (NFSv2/3)          [/export/sapmnt/TC1
/export/saptrans/TC1 /export/archive/TL1_lock]
* Filesystems/Directories to NFS Mount
[/sapmnt/TC1;/export/sapmnt/TC1 /usr/sap/trans;/export/saptrans/TC1]
```

*Figure 7-49 SAP global file system details*

5. By using the available pick lists (F4), edit the Takeover Node by entering `sapci2` and the Service IP Label by entering `sapcisvc2` as shown in Figure 7-50. Press Enter.

```
                      Add SAP Global Filesystem Details

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

                                                          [Entry Fields]
* SAP Global File System Owning Node                    sapci1
* Takeover Nodes                                       [sapci2]
* Service IP Label                                     [sapcisvc2]
* Shared Volume Groups                                 [vg_TC1_nfs]
* Filesystems/Directories to Export (NFSv2/3)          [/export/sapmnt/TC1
/export/saptrans/TC1 /export/archive/TL1_lock]
* Filesystems/Directories to NFS Mount
[/sapmnt/TC1;/export/sapmnt/TC1 /usr/sap/trans;/export/saptrans/TC1]
```

*Figure 7-50 Additional entries for the Smart Assist*

> **Important:** If you try to configure this solution on a running cluster, configuration errors appear. The cluster synchronizes automatically and tries to start the new resource group, which results in errors when it is unmounted. To prevent these errors, you have to unmount everything manually before you press the Enter key in a running cluster or synchronize in a cluster in the INIT state.

A new PowerHA resource group, `SAP_RG_NFS`, is created. The volume group `vg_TE1_vg` and the service IP label `sapcisvc2` are automatically added to the resource group as shown in Example 7-47.

*Example 7-47 Configured resource group for the SAP NFS instance*

```
root@sapci2 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_NFS

root@sapci2 / # /usr/es/sbin/cluster/utilities/cllsres
```

```
APPLICATIONS="clas_nfsv4"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TC1 /export/saptrans/TC1
/export/archive/TL1_lock"
FILESYSTEM=""
FORCED_VARYON="false"
FSCHECK_TOOL="fsck"
FS_BEFORE_IPADDR="true"
MOUNT_FILESYSTEM="/sapmnt/TC1;/export/sapmnt/TC1
/usr/sap/trans;/export/saptrans/TC1"
RECOVERY_METHOD="sequential"
SERVICE_LABEL="sapcisvc2"
SSA_DISK_FENCING="false"
VG_AUTO_IMPORT="false"
VOLUME_GROUP="vg_TC1_nfs"
USERDEFINED_RESOURCES=""

root@sapci1 / # clshowres -g SAP_RG_NFS
Resource Group Name                                 SAP_RG_NFS
Participating Node Name(s)                          sapci1 sapci2
Startup Policy                                      Online On Home Node Only
Fallover Policy                                     Fallover To Next Priority
Node In The List
Fallback Policy                                     Never Fallback
Site Relationship                                   ignore
Dynamic Node Priority
Service IP Label                                    sapcisvc2
Filesystems                                         ALL
Filesystems Consistency Check                       fsck
Filesystems Recovery Method                         sequential
Filesystems/Directories to be exported (NFSv2/NFSv3)
Filesystems/Directories to be exported (NFSv4)      /export/sapmnt/TC1
/export/saptrans/TC1 /export/archive/TL1_lock
Filesystems to be NFS mounted
/sapmnt/TC1;/export/sapmnt/TC1 /usr/sap/trans;/export/saptrans/TC1
Network For NFS Mount
Filesystem/Directory for NFSv4 Stable Storage
Volume Groups                                       vg_TC1_nfs
Concurrent Volume Groups
Use forced varyon for volume groups, if necessary   false
Disks
GMVG Replicated Resources
GMD Replicated Resources
PPRC Replicated Resources
SVC PPRC Replicated Resources
EMC SRDF? Replicated Resources
TRUECOPY Replicated Resources
GENERIC XD Replicated Resources
Connections Services
Fast Connect Services
Shared Tape Resources
Application Servers                                 clas_nfsv4
Highly Available Communication Links
Primary Workload Manager Class
Secondary Workload Manager Class
Delayed Fallback Timer
```

```
Miscellaneous Data
Automatically Import Volume Groups                   false
Inactive Takeover
SSA Disk Fencing                                     false
Filesystems mounted before IP configured             true
WPAR Name
Run Time Parameters:

Node Name                                            sapci1
Debug Level                                          high
Format for hacmp.out                                 Standard

Node Name                                            sapci2
Debug Level                                          high
Format for hacmp.out                                 Standard

root@sapci1 / # cllsserv
clas_nfsv4          /usr/es/sbin/cluster/apps/clas_nfsv4/start
/usr/es/sbin/cluster/apps/clas_nfsv4/stop background

root@sapci1 / # cllsappmon
clam_nfsv4      user
```

6. If you have a second interface/network for the service IP label, you have to configure the correct network after the use of the Smart Assist for SAP on type NFS:

   Follow the path: `smitty sysmirror` → **Cluster Applications and Resources** → **Resources** → **Configure Service IP Labels/Addresses** → **Change/Show a Service IP Label/Address**.

   Select the IP address and press Enter as shown in Figure 7-51.

```
               Select a Service IP Label/Address to Change/Show

Move cursor to desired item and press Enter.

  sapcisvc2
```

*Figure 7-51 Select a Service IP Label/Address to Change/Show*

Change the network to the correct network as shown in Figure 7-52.

```
               Change/Show a Service IP Label/Address(standard)

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

                                                         [Entry Fields]
  IP Label/Address                                     sapcisvc2
* New IP Label/Address                                [sapcisvc2] +
  Netmask(IPv4)/Prefix Length(IPv6)                   [24]
* Network Name                                        [admin_net] +
  Resource Group Name                                  SAP_RG_NFS
```

*Figure 7-52 Change/Show a Service IP Label/Address (standard)*

7. Change the resource group dependency from Parent/Child to Startafter.

   With the default setting, every manual move or automatic takeover of the resource group for NFS service from one node to the other also restarts the central services and the central instance. To prevent this situation, remove the Parent/Child dependency and create instead a Startafter dependency as shown in Figure 7-53.

   Follow the path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Parent/Child Dependency** → **Remove Parent/Child Dependency between Resource Groups**.

```
         Select a Parent/Child Resource Group Dependency to Delete

  Move cursor to desired item and press Enter.

    # Parent                            Child
    SAP_RG_NFS                          SAP_RG_SCS
    SAP_RG_NFS                          SAP_RG_DVEBMGS12
```

   *Figure 7-53 Remove a parent/child dependency*

   In Figure 7-53, select both configured lines one at a time and press Enter. Or, you can use the command line as shown in Example 7-48.

   *Example 7-48 Command to remove a parent/child resource group dependency*

```
/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'PARENT_CHILD' \
   -d \
   -p'SAP_RG_NFS' \
   -c'SAP_RG_SCS'

/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'PARENT_CHILD' \
   -d \
   -p'SAP_RG_NFS' \
   -c'SAP_RG_DVEBMGS12'
```

8. Follow this path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Start After Resource Group Dependency** → **Add Start After Resource Group Dependency**.

   We have to go through these steps twice. The first procedure is for the central service (Figure 7-54 on page 411).

```
                       Select the Source Resource Group

 Move cursor to desired item and press Enter.

SAP_RG_DVEBMGS12
SAP_RG_ERS
SAP_RG_NFS
SAP_RG_SCS
```

*Figure 7-54 Select the Source Resource Group*

In Figure 7-54, select **SAP_RG_SCS** as the source resource group and press Enter.

9. Select the target resource group (Figure 7-55).

```
                       Select the Target Resource Group

 Move cursor to desired item and press F7.
     ONE OR MORE items can be selected.
 Press Enter AFTER making all selections.

  SAP_RG_DVEBMGS12
  SAP_RG_ERS
  SAP_RG_NFS
```

*Figure 7-55 Select the Target Resource Group*

In Figure 7-55, select **SAP_RG_NFS** as the target resource group and press Enter.

10. Next, you add a start after the resource group dependency (Figure 7-56).

```
                  Add Start After Resource Group Dependency

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

                                                        [Entry Fields]
* Source Resource Group                              [SAP_RG_SCS]
* Target Resource Group                              [SAP_RG_NFS]
```

*Figure 7-56 Add Start After Resource Group Dependency*

Create the dependency by pressing Enter again as shown in Figure 7-56.

11. Go through these steps again for the central instance (Figure 7-57 on page 412).

```
                         Select the Source Resource Group

 Move cursor to desired item and press Enter.

SAP_RG_DVEBMGS12
SAP_RG_ERS
SAP_RG_NFS
SAP_RG_SCS
```

*Figure 7-57 Select the Source Resource Group*

Select **SAP_RG_DVEBMGS12** as source resource group and press Enter (Figure 7-57). The Select the Target Resource Group panel that is shown in Figure 7-58 opens.

```
                         Select the Target Resource Group

 Move cursor to desired item and press F7.
     ONE OR MORE items can be selected.
 Press Enter AFTER making all selections.

  SAP_RG_SCS
  SAP_RG_ERS
  SAP_RG_NFS
```

*Figure 7-58 Select the Target Resource Group*

12. In Figure 7-58, select **SAP_RG_NFS** as the target resource group and press Enter. The Add Start After Resource Group Dependency panel opens (Figure 7-59).

```
                    Add Start After Resource Group Dependency

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

                                                         [Entry Fields]
* Source Resource Group                               [SAP_RG_DVEBMGS12]
* Target Resource Group                               [SAP_RG_NFS]
```

*Figure 7-59 Add Start After Resource Group Dependency*

13. Create the dependency by pressing Enter again. Or, you can use the command line as shown in Example 7-49.

*Example 7-49 Command to add a Start_After resource group dependency*

```
/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'START_AFTER' \
   -a \
   -c'SAP_RG_SCS' \
   -p'SAP_RG_NFS'

/usr/es/sbin/cluster/utilities/clrgdependency \
```

```
-t'START_AFTER' \
-a \
-c'SAP_RG_DVEBMGS12' \
-p'SAP_RG_NFS'
```

### 7.5.4 Starting Smart Assist for SAP: Central services (SCS)

We describe starting the Smart Assist for SAP central services (SCS).

> **Smart Assist for SAP:** The Smart Assist for SAP supports only pure ABAP and dual-stack systems. The Smart Assist for SAP with subtype SCS implies ASCS and SCS.

Follow these steps:

1. Launch the Smart Assist for SAP by using the path for `sapci1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-60), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  SAP Smart Assist                 # sapci1 sapci2
```

*Figure 7-60 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-61), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-61 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-62), select **2.SAP NW 7.0 SCS Instance**.

```
              Select The Specific Configuration You Wish to Create

 Move cursor to desired item and press Enter.

   4.SAP NW 7.0 AS Instance        # sapci1
   3.SAP NW 7.0 ERS Instance       # sapci1 sapci2
   1.SAP NW 7.0 Global Filesystem  # sapci1 sapci2
   2.SAP NW 7.0 SCS Instance       # sapci1
```

*Figure 7-62 Selecting the configuration to create*

5. Select the line with your ASCS instance as shown in Figure 7-63.

```
                          Select a SCS/ASCS instance

Move cursor to desired item and press Enter.

ASCS10
```

*Figure 7-63 Select an ASCS instance*

6. By using the available pick lists (F4), edit the Takeover Nodes to **sapci1** as shown in Figure 7-64. Check all other values for correctness even if there are values that are listed, and then press Enter.

```
                      Add SAP SCS/ASCS Instance(s) Details

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

                                                        [Entry Fields]
* SAP SYSTEM ID                                       TC1
* SAP SCS/ASCS Instance(s) Name(s)                    ASCS10
* SAP SCS/ASCS Instance No(s).                        10
* Application Name                                   [sap_scs_appserver]
* Primary Node                                        sapci1
* Takeover Nodes                                     [sapci2] +
* Service IP Label                                    sapcisvc1
* Shared Volume Groups                               [vg_TC1_sapas]
```

*Figure 7-64 Add SAP SCS/ASCS Instance Details*

A new PowerHA resource group, `SAP_RG_SCS`, is created. The volume group `vg_TC1_sapas` and the service IP label `sapcisvc1` are automatically added to the resource group as shown in Example 7-50.

*Example 7-50 Configured resource group for the SAP ASCS instance*

```
root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_NFS
SAP_RG_SCS
```

```
root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4 sap_scs_appserver_AP_SCS"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TC1 /export/saptrans/TC1"
FILESYSTEM=" "
FORCED_VARYON="false false"
FSCHECK_TOOL="fsck fsck"
FS_BEFORE_IPADDR="true false"
MOUNT_FILESYSTEM="/sapmnt/TC1;/export/sapmnt/TC1
/usr/sap/trans;/export/saptrans/TC1"
RECOVERY_METHOD="sequential sequential"
SERVICE_LABEL="sapcisvc2 sapcisvc1"
SSA_DISK_FENCING="false false"
VG_AUTO_IMPORT="false false"
VOLUME_GROUP="vg_TC1_nfs vg_TC1_sapas"
USERDEFINED_RESOURCES=""

root@sapci1 /usr/sap/TC1/DVEBMGS12/work # clshowres -g SAP_RG_SCS
Resource Group Name                                 SAP_RG_SCS
Participating Node Name(s)                          sapci1 sapci2
Startup Policy                                      Online On Home Node Only
Fallover Policy                                     Fallover To Next Priority
Node In The List
Fallback Policy                                     Never Fallback
Site Relationship                                   ignore
Dynamic Node Priority
Service IP Label                                    sapcisvc1
Filesystems                                         ALL
Filesystems Consistency Check                       fsck
Filesystems Recovery Method                         sequential
Filesystems/Directories to be exported (NFSv2/NFSv3)
Filesystems/Directories to be exported (NFSv4)
Filesystems to be NFS mounted
Network For NFS Mount
Filesystem/Directory for NFSv4 Stable Storage
Volume Groups                                       vg_TC1_sapas
Concurrent Volume Groups
Use forced varyon for volume groups, if necessary   false
Disks
GMVG Replicated Resources
GMD Replicated Resources
PPRC Replicated Resources
SVC PPRC Replicated Resources
EMC SRDF? Replicated Resources
TRUECOPY Replicated Resources
GENERIC XD Replicated Resources
Connections Services
Fast Connect Services
Shared Tape Resources
Application Servers                                 sap_scs_appserver_AP_SCS
Highly Available Communication Links
Primary Workload Manager Class
Secondary Workload Manager Class
Delayed Fallback Timer
Miscellaneous Data
```

```
Automatically Import Volume Groups                  false
Inactive Takeover
SSA Disk Fencing                                    false
Filesystems mounted before IP configured            false
WPAR Name
Run Time Parameters:

Node Name                                           sapci1
Debug Level                                         high
Format for hacmp.out                                Standard

Node Name                                           sapci2
Debug Level                                         high
Format for hacmp.out                                Standard

root@sapci1 / # cllsserv
clas_nfsv4          /usr/es/sbin/cluster/apps/clas_nfsv4/start
/usr/es/sbin/cluster/apps/clas_nfsv4/stop background
sap_scs_appserver_AP_SCS       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartSCS
-a sap_scs_appserver       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopSCS -a
sap_scs_appserver background


root@sapci1 / # cllsappmon
clam_nfsv4      user
sap_scs_appserver_AM_SCS        user
```

### 7.5.5 Starting Smart Assist for SAP: Enqueue replication server instance (ERS)

We describe how to start the Smart Assist for SAP enqueue replication server instance (ERS). Follow these steps:

1. Launch Smart Assist for SAP from the `sapci1` node: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-65), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  SAP Smart Assist                 # sapci1 sapci2
```

*Figure 7-65 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-66), select **Automatic Discovery and Configuration**.

```
                    Select Configuration Mode

 Move cursor to desired item and press Enter.

   Automatic Discovery And Configuration
   Manual Configuration
```

*Figure 7-66 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-67), select **3.SAP NW 7.0 ERS Instance**.

```
          Select The Specific Configuration You Wish to Create

 Move cursor to desired item and press Enter.

   4.SAP NW 7.0 AS Instance         # sapci1
   3.SAP NW 7.0 ERS Instance        # sapci2
   1.SAP NW 7.0 Global Filesystem   # sapci1 sapci2
   2.SAP NW 7.0 SCS Instance        # sapci1
```

*Figure 7-67 Selecting the configuration to create*

5. Select the line with your ERS instance as shown in Figure 7-68.

```
                       Select ERS instance

Move cursor to desired item and press Enter.

ERS11
```

*Figure 7-68 Select an ERS instance*

6. By using the available pick lists (F4), edit the Takeover Nodes to select **sapci1** as shown in Figure 7-69 and then press Enter.

```
                        Add SAP ERS Instance(s) Details

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

                                                        [Entry Fields]
* SAP SYSTEM ID                                       TC1
* SAP ERS Instance(s) Name(s)                         ERS11
* SAP ERS Instance No(s).                             11
* Application Name                                   [sap_ers_appserver]
* Primary Node                                       [sapci2] +
* Takeover Nodes                                     [sapci1] +
* Shared Volume Groups                               [vg_TC1_sapers]
```

*Figure 7-69 Add SAP ERS Instance Details*

7. A new PowerHA resource group, `SAP_RG_ERS`, is created. The volume group `vg_TC1_sapers` is automatically added to the resource group as shown in Example 7-51. However, *no service address* is configured by the Smart Assist for SAP.

*Example 7-51 Configured resource group for the SAP ASCS instance*

```
root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_ERS
SAP_RG_NFS
SAP_RG_SCS

root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4 sap_scs_appserver_AP_SCS sap_ers_appserver_AP_SCS"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TC1 /export/saptrans/TC1"
FILESYSTEM="  "
FORCED_VARYON="false false false"
FSCHECK_TOOL="fsck fsck fsck"
FS_BEFORE_IPADDR="true false false"
MOUNT_FILESYSTEM="/sapmnt/TC1;/export/sapmnt/TC1
/usr/sap/trans;/export/saptrans/TC1"
NODE_PRIORITY_POLICY="cl_highest_udscript_rc"
RECOVERY_METHOD="sequential sequential sequential"
SDNP_SCRIPT_PATH="/usr/es/sbin/cluster/sa/sap/sbin/cl_SCSFailoverNodeCheck"
SDNP_SCRIPT_TIMEOUT="360"
SERVICE_LABEL="sapcisvc2 sapcisvc1"
SSA_DISK_FENCING="false false false"
VG_AUTO_IMPORT="false false false"
VOLUME_GROUP="vg_TC1_nfs vg_TC1_sapers vg_TC1_sapas"
USERDEFINED_RESOURCES=""

root@sapci1 / # clshowres -g SAP_RG_ERS

Resource Group Name                                    SAP_RG_ERS
Participating Node Name(s)                             sapci2 sapci1
Startup Policy                                         Online On Home Node Only
```

```
Fallover Policy                                     Fallover To Next Priority
Node In The List
Fallback Policy                                     Never Fallback
Site Relationship                                   ignore
Dynamic Node Priority
Service IP Label
Filesystems                                         ALL
Filesystems Consistency Check                       fsck
Filesystems Recovery Method                         sequential
Filesystems/Directories to be exported (NFSv2/NFSv3)
Filesystems/Directories to be exported (NFSv4)
Filesystems to be NFS mounted
Network For NFS Mount
Filesystem/Directory for NFSv4 Stable Storage
Volume Groups                                       vg_TC1_sapers
Concurrent Volume Groups
Use forced varyon for volume groups, if necessary   false
Disks
GMVG Replicated Resources
GMD Replicated Resources
PPRC Replicated Resources
SVC PPRC Replicated Resources
EMC SRDF? Replicated Resources
TRUECOPY Replicated Resources
GENERIC XD Replicated Resources
Connections Services
Fast Connect Services
Shared Tape Resources
Application Servers                                 sap_ers_appserver_AP_SCS
Highly Available Communication Links
Primary Workload Manager Class
Secondary Workload Manager Class
Delayed Fallback Timer
Miscellaneous Data
Automatically Import Volume Groups                  false
Inactive Takeover
SSA Disk Fencing                                    false
Filesystems mounted before IP configured            false
WPAR Name
Run Time Parameters:

Node Name                                           sapci2
Debug Level                                         high
Format for hacmp.out                                Standard

Node Name                                           sapci1
Debug Level                                         high
Format for hacmp.out                                Standard

root@sapci1 / # cllsserv
clas_nfsv4          /usr/es/sbin/cluster/apps/clas_nfsv4/start
/usr/es/sbin/cluster/apps/clas_nfsv4/stop background
sap_ers_appserver_AP_SCS       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartERS
-a sap_ers_appserver       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopERS -a
sap_ers_appserver background
```

```
sap_scs_appserver_AP_SCS       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartSCS
-a sap_scs_appserver      /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopSCS -a
sap_scs_appserver background

root@sapci1 / # cllsappmon
clam_nfsv4      user
sap_ers_appserver_AM_SCS        user
sap_scs_appserver_AM_SCS        user
```

8. To configure a service address, we have to launch the following smit path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resources** → **Configure Service IP Labels/Addresses** → **Add a Service IP Label/Address**.

   Select the desired network and press Enter (Figure 7-70).

```
                                Network Name

Move cursor to desired item and press Enter.

   admin_net (10.10.20.0/24 10.1.2.0/24)
   service_net (172.16.20.0/23 10.1.0.0/23)
```

*Figure 7-70 Select a Service IP Label/Address to Add*

   Add the Service IP as shown in Figure 7-71 and press Enter.

```
                        Add a Service IP Label/Address

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

                                                       [Entry Fields]
* IP Label/Address                                    sapcisvc3 +
  Netmask(IPv4)/Prefix Length(IPv6)                  []
* Network Name                                        service_net
```

*Figure 7-71 Add a Service IP Label/Address*

9. Add the service address to the resource group by following the smit path: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Change/Show Resources and Attributes for a Resource Group**.

   Select the desired resource group **SAP_RG_ERS** and press Enter (Figure 7-72).

```
           Change/Show Resources and Attributes for a Resource Group

Move cursor to desired item and press Enter.

   SAP_RG_ERS
   SAP_RG_NFS
   SAP_RG_SCS
```

*Figure 7-72 Select the resource group to change*

Add the Service IP Label/Addresses as shown in Figure 7-73 and press Enter.

```
        Change/Show All Resources and Attributes for a Resource Group

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

[TOP]                                                   [Entry Fields]
  Resource Group Name                                 SAP_RG_ERS
  Participating Nodes (Default Node Priority)         sapci2 sapci1

  Startup Policy                Online On Home Node Only
  Fallover Policy               Fallover To Next Priority Node In The List
  Fallback Policy               Never Fallback

  Service IP Labels/Addresses   [sapcisvc3] +
  Application Controllers       [sap_ers_appserver_AP_SCS] +

  Volume Groups                                      [vg_TC1_sapers ] +
  Use forced varyon of volume groups, if necessary    false +
  Automatically Import Volume Groups                  false +

  Filesystems (empty is ALL for VGs specified)       [ ] +
  Filesystems Consistency Check                       fsck +
  Filesystems Recovery Method                         sequential +
  Filesystems mounted before IP configured            false +

  Filesystems/Directories to Export (NFSv2/3)        [] +
  Filesystems/Directories to Export (NFSv4)          [] +
....
```

*Figure 7-73 Add a service IP label to the resource group*

## 7.5.6 Starting Smart Assist for SAP: Application server instance (AS)

We describe how to configure the Smart Assist for SAP application server instance (AS).

**Subtype:** The Smart Assist for SAP does not differentiate between a central instance, a primary application server instance, and an additional application server. So, you have to choose the subtype AS.

Follow these steps:

1. Launch the Smart Assist for SAP by using the path for `sapci1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.

2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-74), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  SAP Smart Assist                 # sapci1 sapci2
```

*Figure 7-74 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-75), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-75 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-76), select **4.SAP NW 7.0 AS Instance**.

```
             Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  4.SAP NW 7.0 AS Instance         # sapci1
  3.SAP NW 7.0 ERS Instance        # sapci1
  1.SAP NW 7.0 Global Filesystem   # sapci1 sapci2
  2.SAP NW 7.0 SCS Instance        # sapci1
```

*Figure 7-76 Selecting the configuration to create*

5. Select the line with your application instance (Figure 7-77).

```
                        Select an Application instance

Move cursor to desired item and press Enter.

DVEBMGS12
```

*Figure 7-77 Select an application instance*

6. By using the available pick lists (F4), edit the Takeover Nodes by selecting **sapci2** as shown in Figure 7-78 and then press Enter.

```
                    Add SAP Application Instance Details

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

                                                          [Entry Fields]
* SAP SYSTEM ID                                         TC1
* SAP Application Server Instance Name                  DVEBMGS12
* SAP Application Server Instance No.                   12
* Application Name                                     [sap_DVEBMGS12_appserver]
* Primary Node                                          sapci1
* Takeover Nodes                                       [sapci2] +
* Service IP Label                                      sapcisvc4
* Shared Volume Groups                                 [vg_TC1_sapci]
```

*Figure 7-78 Add SAP Application Instance Details*

A new PowerHA resource group, `SAP_RG_DVEBMGS12`, is created. The volume group `vg_TC1_sapci` and the service IP label `sapcisvc4` are automatically added to the resource group as shown in Example 7-52.

*Example 7-52 Configured resource group for the SAP application instance*

```
root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsgrp
SAP_RG_DVEBMGS12
SAP_RG_ERS
SAP_RG_NFS
SAP_RG_SCS

root@sapci1 / # /usr/es/sbin/cluster/utilities/cllsres
APPLICATIONS="clas_nfsv4 sap_scs_appserver_AP_SCS sap_ers_appserver_AP_SCS
sap_DVEBMGS12_appserver_AP_AS"
EXPORT_FILESYSTEM_V4="/export/sapmnt/TC1 /export/saptrans/TC1"
FILESYSTEM="   "
FORCED_VARYON="false false false false"
FSCHECK_TOOL="fsck fsck fsck fsck"
FS_BEFORE_IPADDR="true false false false"
MOUNT_FILESYSTEM="/sapmnt/TC1;/export/sapmnt/TC1
/usr/sap/trans;/export/saptrans/TC1"
NODE_PRIORITY_POLICY="cl_highest_udscript_rc"
RECOVERY_METHOD="sequential sequential sequential sequential"
SDNP_SCRIPT_PATH="/usr/es/sbin/cluster/sa/sap/sbin/cl_SCSFailoverNodeCheck"
SDNP_SCRIPT_TIMEOUT="360"
SERVICE_LABEL="sapcisvc2 sapcisvc1 sapcisvc3 sapcisvc4"
SSA_DISK_FENCING="false false false false"
VG_AUTO_IMPORT="false false false false"
VOLUME_GROUP="vg_TC1_nfs vg_TC1_sapers vg_TC1_sapas vg_TC1_sapci"
USERDEFINED_RESOURCES=""

root@sapci1 / # clshowres -g SAP_RG_DVEBMGS12
Resource Group Name                                    SAP_RG_DVEBMGS12
Participating Node Name(s)                             sapci1 sapci2
```

```
Startup Policy                                      Online On Home Node Only
Fallover Policy                                     Fallover To Next Priority
Node In The List
Fallback Policy                                     Never Fallback
Site Relationship                                   ignore
Dynamic Node Priority
Service IP Label                                    sapcisvc4
Filesystems                                         ALL
Filesystems Consistency Check                       fsck
Filesystems Recovery Method                         sequential
Filesystems/Directories to be exported (NFSv2/NFSv3)
Filesystems/Directories to be exported (NFSv4)
Filesystems to be NFS mounted
Network For NFS Mount
Filesystem/Directory for NFSv4 Stable Storage
Volume Groups                                       vg_TC1_sapci
Concurrent Volume Groups
Use forced varyon for volume groups, if necessary   false
Disks
GMVG Replicated Resources
GMD Replicated Resources
PPRC Replicated Resources
SVC PPRC Replicated Resources
EMC SRDF? Replicated Resources
TRUECOPY Replicated Resources
GENERIC XD Replicated Resources
Connections Services
Fast Connect Services
Shared Tape Resources
Application Servers
sap_DVEBMGS12_appserver_AP_AS
Highly Available Communication Links
Primary Workload Manager Class
Secondary Workload Manager Class
Delayed Fallback Timer
Miscellaneous Data
Automatically Import Volume Groups                  false
Inactive Takeover
SSA Disk Fencing                                    false
Filesystems mounted before IP configured            false
WPAR Name
Run Time Parameters:

Node Name                                           sapci1
Debug Level                                         high
Format for hacmp.out                                Standard

Node Name                                           sapci2
Debug Level                                         high
Format for hacmp.out                                Standard

root@sapci1 / # cllsserv
clas_nfsv4          /usr/es/sbin/cluster/apps/clas_nfsv4/start
/usr/es/sbin/cluster/apps/clas_nfsv4/stop background
```

```
sap_DVEBMGS12_appserver_AP_AS  /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartAS
-a sap_DVEBMGS12_appserver  /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopAS -a
sap_DVEBMGS12_appserver background
sap_ers_appserver_AP_SCS       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartERS
-a sap_ers_appserver       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopERS -a
sap_ers_appserver background
sap_scs_appserver_AP_SCS       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStartSCS
-a sap_scs_appserver       /usr/es/sbin/cluster/sa/sap/sbin/cl_sapStopSCS -a
sap_scs_appserver background

root@sapci1 / # cllsappmon
clam_nfsv4      user
sap_DVEBMGS12_appserver_AM_AS   user
sap_ers_appserver_AM_SCS        user
sap_scs_appserver_AM_SCS        user
```

### 7.5.7 Completing the configuration

After the Smart Assist for SAP is started, complete the configuration:

1. Stop the complete SAP instance on the primary and secondary node as shown in Example 7-53 and Example 7-54 on page 426. Remember that it is active only for the Smart Assist for SAP discovery process.

*Example 7-53 Stopping the SAP instance on sapci1*

```
root@sapci1 / # su - tc1adm
sapci1:tc1adm 1> stopsap sapcisvc4

Checking db Database
------------------------------
 Database is running

stopping the SAP instance DVEBMGS12
----------------------------------
 Shutdown-Log is written to /usr/sap/TC1/home/tc1adm/stopsap_DVEBMGS12.log
/usr/sap/TC1/DVEBMGS12/exe/sapcontrol -prot NI_HTTP -nr 12 -function Stop
 Instance on host sapci1 stopped
 Waiting for cleanup of resources................

sapci1:tc1adm 2> stopsap sapcisvc1

stopping the SAP instance ASCS10
----------------------------------
 Shutdown-Log is written to /usr/sap/TC1/home/tc1adm/stopsap_ASCS10.log
/usr/sap/TC1/ASCS10/exe/sapcontrol -prot NI_HTTP -nr 10 -function Stop
 Instance on host sapci1 stopped
 Waiting for cleanup of resources.......

root@sapci1 / # /etc/rc.d/rc2.d/Ksapinit stop
saphostexec is already running (pid=7077960). Stopping...Stopped
```

*Example 7-54 Stopping the SAP instance on sapci2*

```
root@sapci2 / # su - tc1adm
sapci2:tc1adm 1> stopsap sapcisvc3

stopping the SAP instance ERS11
----------------------------------
 Shutdown-Log is written to /usr/sap/TC1/home/tc1adm/stopsap_ERS11.log
/usr/sap/TC1/ERS11/exe/sapcontrol -prot NI_HTTP -nr 11 -function Stop
 Instance on host sapci2 stopped
 Waiting for cleanup of resources......

root@sapci2 / # /etc/rc.d/rc2.d/Ksapinit stop
saphostexec is already running (pid=6947054). Stopping...Stopped
```

2. Example 7-55 and Example 7-56 show the unmount of the NFS client file systems on both nodes, if mounted.

*Example 7-55 Unmount NFS client file systems on sapci1*

```
root@sapci1 / # umount /sapmnt/TC1
root@sapci1 / # umount /usr/sap/trans
```

*Example 7-56 Unmount NFS client file systems on sapci2*

```
root@sapci2 / # umount /sapmnt/TC1
root@sapci2 / # umount /usr/sap/trans
```

3. Unexport the NFS server directories. Check which directories are exported as shown in Example 7-57.

*Example 7-57 Exported directories*

```
root@sapci1 / # exportfs
/export/sapmnt/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/saptrans/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
```

Unexport all directories for the SAP NFS service as shown in Example 7-58.

*Example 7-58 Unexport directories on the NFS server*

```
root@sapci1 / # exportfs -u /export/sapmnt/TC1
root@sapci1 / # exportfs -u /export/saptrans/TC1
```

Verify that no directory is exported as shown in Example 7-59.

*Example 7-59 NFS exports*

```
root@sapci1 / # exportfs
exportfs: 1831-182 nothing exported
```

4. Unmount the shared file systems as shown in Example 7-60 on page 427 and Example 7-61 on page 427.

*Example 7-60 Unmounting the shared file systems on sapci1*

```
root@sapci1 / # umount -t TC1

root@sapci1 / # lsvg -o|grep vg_TC1|lsvg -l -i
vg_TC1_sapas:
LV NAME             TYPE LPs PPs     PVs  LV STATE      MOUNT POINT
TC1.lvusrsap        jfs2 160 160     1    closed/syncd  /usr/sap/TC1/ASCS10
vg_TC1_sapci:
LV NAME             TYPE LPs PPs     PVs  LV STATE      MOUNT POINT
TC1.lvusrsapci      jfs2 160 160     1    closed/syncd  /usr/sap/TC1/DVEBMGS12
vg_TC1_nfs:
LV NAME             TYPE LPs PPs     PVs  LV STATE      MOUNT POINT
TC1.lvsapmnt        jfs2 64 64       1    closed/syncd  /export/sapmnt/TC1
TC1.lvsaptrans      jfs2 32 32       1    closed/syncd  /export/saptrans/TC1
```

*Example 7-61 Unmounting the shared file systems on sapci2*

```
root@sapci2 / # umount -t TC1

root@sapci2 / # lsvg -o|grep vg_TC1|lsvg -l -i
vg_TC1_sapers:
LV NAME             TYPE LPs     PPs     PVs  LV STATE      MOUNT POINT
TC1.lvusrsapers     jfs2 10      10      1    closed/syncd  /usr/sap/TC1/ERS11
```

5. Deactivate the shared volume group as shown in Example 7-62 and Example 7-63.

*Example 7-62 Deactivating the shared volume group on sapci1*

```
root@sapci1 / # lsvg -o|grep vg_TC1|xargs -I{} varyoffvg {}
root@sapci1 / # lsvg -o
caavg_private
rootvg
```

*Example 7-63 Deactivating the shared volume group on sapci2*

```
root@sapci2 / # lsvg -o|grep vg_TC1|xargs -I{} varyoffvg {}

root@sapci2 / # lsvg -o
caavg_private
rootvg
```

6. Unconfigure the network alias on both nodes as shown in Example 7-64 and Example 7-65 on page 428.

*Example 7-64 Network settings on sapci1*

```
root@sapci1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.99.b2.bf.6f  329104     0   155135     0     0
en0   1500  10.1        sapci1b1           329104     0   155135     0     0
en0   1500  172.16.20   sapci1p1           329104     0   155135     0     0
en0   1500  172.16.20   sapcisvc1          329104     0   155135     0     0
en1   1500  link#3      2e.47.99.b2.bf.70  244142     0   289877     0     0
en1   1500  10.1.2      sapci1b2           244142     0   289877     0     0
en1   1500  10.10.20    sapci1p2           244142     0   289877     0     0
en1   1500  10.10.20    sapcisvc2          244142     0   289877     0     0
en1   1500  10.10.20    sapcisvc4          244142     0   289877     0     0
```

```
lo0   16896 link#1                         1091286     0  1091285     0     0
lo0   16896 127         loopback           1091286     0  1091285     0     0
lo0   16896 loopback                       1091286     0  1091285     0     0
root@sapci1 / # ifconfig en0 delete sapcisvc1
root@sapci1 / # ifconfig en1 delete sapcisvc2
root@sapci1 / # ifconfig en1 delete sapcisvc4
root@sapci1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.99.b2.bf.6f  330445     0   156016     0     0
en0   1500  10.1        sapci1b1           330445     0   156016     0     0
en0   1500  172.16.20   sapci1p1           330445     0   156016     0     0
en1   1500  link#3      2e.47.99.b2.bf.70  244157     0   289877     0     0
en1   1500  10.1.2      sapci1b2           244157     0   289877     0     0
en1   1500  10.10.20    sapci1p2           244157     0   289877     0     0
lo0   16896 link#1                         1091582     0  1091581     0     0
lo0   16896 127         loopback           1091582     0  1091581     0     0
lo0   16896 loopback                       1091582     0  1091581     0     0
```

*Example 7-65 Network settings on sapci2*

```
root@sapci2 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c8.81.6f  226211     0    55603     0     0
en0   1500  10.1        sapci2b1           226211     0    55603     0     0
en0   1500  172.16.20   sapci2p1           226211     0    55603     0     0
en0   1500  172.16.20   sapcisvc3          226211     0    55603     0     0
en1   1500  link#3      6e.8d.dd.c8.81.70  295227     0   238833     0     0
en1   1500  10.1.2      sapci2b2           295227     0   238833     0     0
en1   1500  10.10.20    sapci2p2           295227     0   238833     0     0
lo0   16896 link#1                           77709     0    77709     0     0
lo0   16896 127         loopback             77709     0    77709     0     0
lo0   16896 loopback                         77709     0    77709     0     0
root@sapci2 / # ifconfig en0 delete sapcisvc3
root@sapci2 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c8.81.6f  226317     0    55643     0     0
en0   1500  10.1        sapci2b1           226317     0    55643     0     0
en0   1500  172.16.20   sapci2p1           226317     0    55643     0     0
en1   1500  link#3      6e.8d.dd.c8.81.70  295227     0   238833     0     0
en1   1500  10.1.2      sapci2b2           295227     0   238833     0     0
en1   1500  10.10.20    sapci2p2           295227     0   238833     0     0
lo0   16896 link#1                           77727     0    77727     0     0
lo0   16896 127         loopback             77727     0    77727     0     0
lo0   16896 loopback                         77727     0    77727     0     0
```

7. Synchronize the PowerHA cluster by using SMIT:

   a. Follow the path: `smitty sysmirror` → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.

   b. In the PowerHA SystemMirror Verification and Synchronization panel (Figure 7-79 on page 429), change Automatically correct errors found during verification `to` **Yes** `and` press Enter to accept the other default options.

```
PowerHA SystemMirror Verification and Synchronization

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Verify, Synchronize or Both                       [Both]          +
* Include custom verification library checks        [Yes]           +
* Automatically correct errors found during         [Yes]           +
  verification?
* Force synchronization if verification fails?      [No]            +
* Verify changes only?                              [No]            +
* Logging                                           [Standard]      +
```

*Figure 7-79 Accepting the actions on the Verification and Synchronization panel*

8. Start the cluster on both nodes, sapci1 and sapci2, by running **smitty clstart**.
9. In the Start Cluster Services panel (Figure 7-80), complete these steps:
   a. For Start now, on system restart or both, select **now**.
   b. For Start Cluster Services on these nodes, enter [sapci1,sapci2].
   c. For Manage Resource Groups, select **Automatically**.
   d. For BROADCAST message at startup, select **false**.
   e. For Startup Cluster Information Daemon, select **true**.
   f. For Ignore verification errors, select **false**.
   g. For Automatically correct errors found during cluster start, select **yes**.

   Press Enter.

```
                          Start Cluster Services

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Start now, on system restart or both              now                 +
  Start Cluster Services on these nodes             [sapci1,sapci2]     +
* Manage Resource Groups                            Automatically       +
  BROADCAST message at startup?                     false               +
  Startup Cluster Information Daemon?               true                +
  Ignore verification errors?                       false               +
  Automatically correct errors found during         yes                 +
  cluster start?
```

*Figure 7-80 Specifying the options for starting cluster services*

**Tip:** The log file for the Smart Assist is in the /var/hacmp/log/sa.log or /var/hacmp/logsapsa.log file. You can use the **clmgr** utility to view the log as shown in the following example (this utility does not work for sapsa.log):

```
clmgr view log sa.log
```

When the PowerHA cluster starts, the entire SAP instance is automatically started. The application monitors the start after the defined stabilization interval as shown in Example 7-66 on page 430 and Example 7-67 on page 430.

*Example 7-66 Checking the status of the highly available cluster and the SAP instance on sapci1*

```
root@sapci1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                        Node
-----------------------------------------------------------------------------
SAP_RG_NFS     ONLINE                       sapci1
               OFFLINE                      sapci2

SAP_RG_SCS     ONLINE                       sapci1
               OFFLINE                      sapci2

SAP_RG_ERS     ONLINE                       sapci2
               OFFLINE                      sapci1

SAP_RG_DVEBMGS ONLINE                       sapci1
               OFFLINE                      sapci2

root@sapci1 / # ps -ef | grep /usr/es/sbin/cluster/clappmond | grep -v grep
    root  7078066 11141218   0 09:59:13      -  0:00
/usr/es/sbin/cluster/clappmond clam_nfsv4
    root 12189878 15532260   0 10:04:42      -  0:00
/usr/es/sbin/cluster/clappmond sap_scs_appserver_AM_SCS
    root 12976290  8585320   0 10:10:42      -  0:00
/usr/es/sbin/cluster/clappmond sap_DVEBMGS12_appserver_AM_AS

root@sapci1 / # su - tc1adm
sapci1:tc1adm 1> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 10 -function
GetProcessList

06.04.2012 10:18:27
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
msg_server, MessageServer, GREEN, Running, 2012 04 06 09:59:27, 0:19:00, 9568304
enserver, EnqueueServer, GREEN, Running, 2012 04 06 09:59:27, 0:19:00, 9240828

sapci1:tc1adm 2> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 12 -function
GetProcessList

06.04.2012 10:18:37
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
disp+work, Dispatcher, GREEN, Running, Message Server connection ok, Dialog Queue
time: 0.00 sec, 2012 04 06 10:00:01, 0:18:36, 12058850
igswd_mt, IGS Watchdog, GREEN, Running, 2012 04 06 10:00:01, 0:18:36, 15073450
```

*Example 7-67 Checking the status of the highly available cluster and the SAP instance on sapci2*

```
root@sapci2 / # ps -ef | grep /usr/es/sbin/cluster/clappmond | grep -v grep
    root 16187472 10813538   0 10:05:30      -  0:00
/usr/es/sbin/cluster/clappmond sap_ers_appserver_AM_SCS

root@sapci2 / # su - tc1adm
```

```
sapci2:tc1adm 1> /usr/sap/hostctrl/exe/sapcontrol -prot NI_HTTP -nr 11 -function
GetProcessList

06.04.2012 10:17:49
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
enrepserver, EnqueueReplicator, GREEN, Running, 2012 04 06 10:01:16, 0:16:33,
12910770
```

Your SAP instances are now configured for high availability in a hot-standby PowerHA SystemMirror configuration.

# 7.6 Cluster 2: Database instance

We describe the database instance in cluster 2.

## 7.6.1 Overview

This cluster uses the Smart Assist for SAP database instance (DB) to make a DB2 database highly available (Figure 7-81).

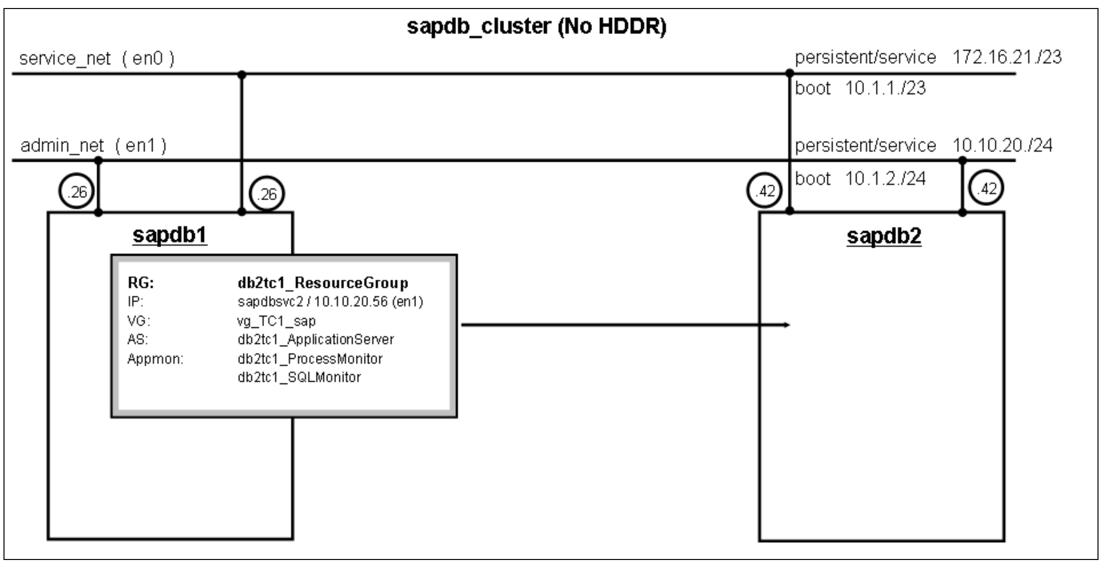



*Figure 7-81 Schema overview for sapdb_cluster*

Figure 7-81 outlines the resource groups and network settings for this cluster.

## 7.6.2 Installation and configuration steps before you use Smart Assist for SAP

Cluster 2 uses the Smart Assist for SAP database instance (DB). Follow these steps:

1. Install the required filesets.

   See 7.2.1, “Installing the required PowerHA SystemMirror Smart Assist filesets” on page 362.

2. Configure the base IBM PowerHA SystemMirror.

You must configure the topology of the PowerHA cluster before you use the Smart Assist for SAP as shown in previous sections. Example 7-68 shows the cluster `sapdb_cluster` that is configured with two Ethernet interfaces in each node.

*Example 7-68 Cluster network configuration*

```
root@sapdb1 / # lscluster -c
Cluster query for cluster sapdb_cluster returns:
Cluster uuid: 4bb355a6-720a-11e1-8613-6e8dddc6b66f
Number of nodes in cluster = 2
        Cluster id for node sapdb1b1 is 1
        Primary IP address for node sapdb1b1 is 10.1.1.26
        Cluster id for node sapdb2b1 is 2
        Primary IP address for node sapdb2b1 is 10.1.1.42
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.23

root@sapdb1 / # cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address
Interface Name Global Name Netmask Alias for HB Prefix Length

sapdb1b2 boot admin_net  ether public sapdb1 10.1.2.26 en1 255.255.255.0 24
sapdb1b1 boot service_net ether public sapdb1 10.1.1.26 en0 255.255.254.0 23
sapdb2b2 boot admin_net  ether public sapdb2 10.1.2.42 en1 255.255.255.0 24
sapdb2b1 boot service_net ether public sapdb2 10.1.1.42 en0 255.255.254.0 23
```

3. Configure network.

   The IP addresses for the DB2 server (`sapdbsvc2`) is configured on the first node `sapdb1` where DB2 is installed (Example 7-69).

   *Example 7-69 Network settings for sapdb1*

```
root@sapdb1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c6.b6.6f  8955245     0  3359862     0     0
en0   1500  10.1        sapdb1b1           8955245     0  3359862     0     0
en0   1500  172.16.20   sapdb1p1           8955245     0  3359862     0     0
en1   1500  link#3      6e.8d.dd.c6.b6.70   331797     0    37242     0     0
en1   1500  10.1.2      sapdb1b2            331797     0    37242     0     0
en1   1500  10.10.20    sapdb1p2            331797     0    37242     0     0
en1   1500  10.10.20    sapdbsvc2           331797     0    37242     0     0
lo0   16896 link#1                         2618976     0  2618971     0     0
lo0   16896 127         loopback           2618976     0  2618971     0     0
lo0   16896 loopback                       2618976     0  2618971     0     0
```

   No additional IP address is configured on the second node `sapdb2`.

4. Create the users.

   See 7.2.2, “Creating the users” on page 362 and the script to create the users in “Script to create users and groups for sapci_cluster, sapdb_cluster, and saplc_cluster” on page 604.

5. Create volume groups, logical volumes, and file systems as shown in Example 7-70 on page 433.

   See 7.2.3, “Creating volume groups, logical volumes, and file systems” on page 363, and the script to create the volume groups and file systems in “Script to create VGs and LVs for sapdb_cluster” on page 608.

*Example 7-70 Volume group at sapdb_cluster*

```
root@sapdb1 / # lsvg -l vg_TC1_sap
vg_TC1_sap:
LV NAME             TYPE       LPs  PPs  PVs  LV STATE      MOUNT POINT
TC1.lvdb2binary     jfs2       128  128  1    open/syncd    /db2/TC1
TC1.lvdb2logdir     jfs2        64   64  1    open/syncd    /db2/TC1/log_dir
TC1.lvdb2logarc     jfs2        64   64  1    open/syncd    /db2/TC1/log_archive
TC1.lvdb2dat.01     jfs2       320  320  1    open/syncd    /db2/TC1/sapdata1
TC1.lvdb2dat.02     jfs2       320  320  1    open/syncd    /db2/TC1/sapdata2
TC1.lvdb2dat.03     jfs2       320  320  1    open/syncd    /db2/TC1/sapdata3
TC1.lvdb2dat.04     jfs2       320  320  1    open/syncd    /db2/TC1/sapdata4
```

6. Install the SAP database instance on the first node of this cluster.

   We installed SAP normally with a virtual SAP host name. We set the usual SAP parameters only, as shown in Example 7-71.

*Example 7-71 SAPinst environment parameters*

```
export TMPDIR=/tmp/TE1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapdbsvc2
```

Start the database installation as shown in Figure 7-82. Select **Database Instance**.

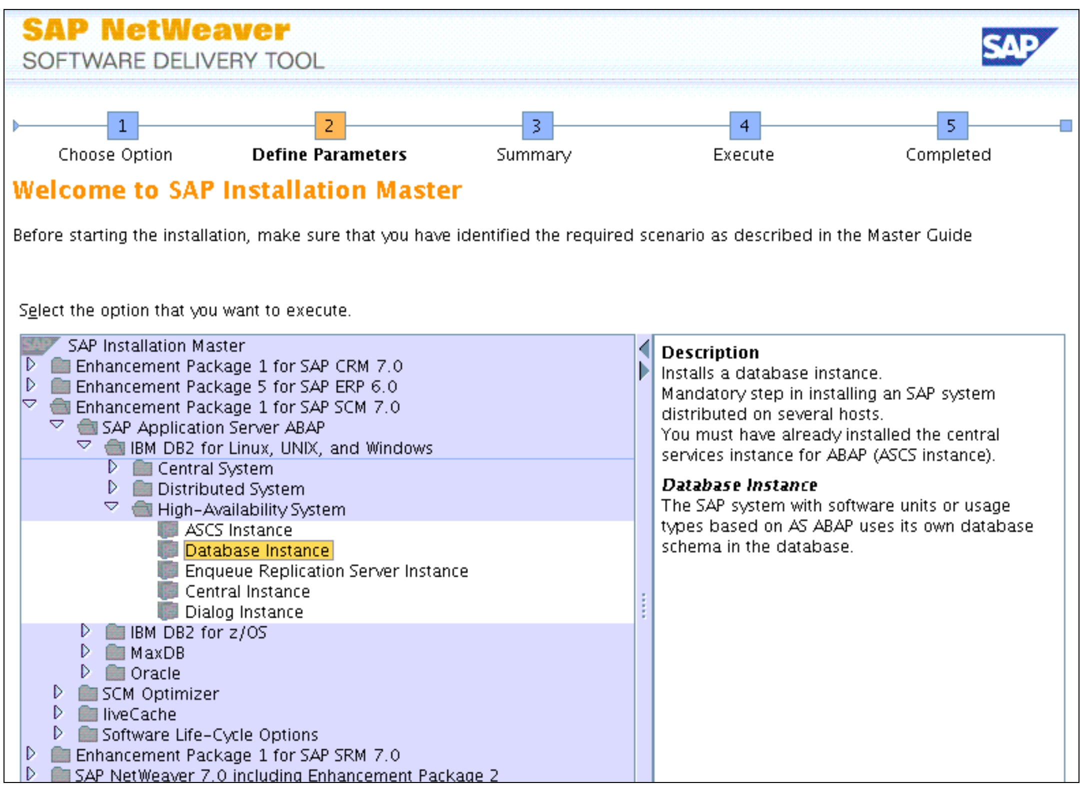



*Figure 7-82 SAPinst database installation screen capture 1: Start installation*

For a detailed view, we documented the complete installation process in "sapdb_cluster: Installing the SAP database instance on the first node" on page 614.

> **Important:** Smart Assist for DB2 supports only non-DPF and non-HADR DB2 databases.

7. Update the `/etc/services` file on the secondary node.

   See 7.2.4, “Updating the /etc/services file on the secondary node” on page 363.

8. Copy additional directories.

   See 7.2.5, “Copying more directories to the second node” on page 364.

9. Find the path for the binary files and then export the variable as shown in Example 7-72. The **DSE_INSTALL_DIR** environment variable is exported as the `root` user with the actual path for the DB2 binary files. If more than one DB2 version is installed, choose the version that you want to use for this highly available instance.

*Example 7-72 Finding the DB2 binary files and exporting them*

```
root@sapdb1 / # /db2/TC1/db2tc1/db2_software/bin/db2level
DB21085I  Instance "db2tc1" uses "64" bits and DB2 code release "SQL09074" with
level identifier "08050107".
Informational tokens are "DB2 v9.7.0.4", "s110330", "IP23236", and Fix Pack
"4".
Product is installed at "/db2/TC1/db2tc1/db2_software".

root@sapdb1 / # export DSE_INSTALL_DIR=/db2/TC1/db2tc1/db2_software
```

> **Important:** The setting of this variable is mandatory. Without this seeded variable, the Smart Assist for SAP cannot find the appropriate database.

10. Check for the types that the Smart Assist for SAP can discover as shown in Example 7-73.

    You can check the possible types on the command line. All types have a `1` at the end. The possible parameters for the **cl_sapdiscover** command are **-t GFS/AS/SCS/ERS/DB**.

*Example 7-73 Checking the Smart Assist for SAP discover function*

```
root@sapdb1 / # /usr/es/sbin/cluster/sa/sap/sbin/cl_sapdiscover -t DB
SAP Smart Assist:SAPNW_7.0:5.SAP Database Instance:SAPNW_7.0_DBINSTANCE:1
```

11. If an expected type is not found, more information can be collected by setting an additional environment variable as shown in Example 7-74.

*Example 7-74 Setting enlarged output for cl_sapdiscover*

```
root@sapdb1 /usr/es/sbin/cluster/sa/sap/sbin # export VERBOSE_LOGGING="high"
root@sapdb1 /usr/es/sbin/cluster/sa/sap/sbin # ./cl_sapdiscover -t DB
...
```

12. PowerHA state

    You can start the PowerHA Smart Assist in a running cluster or a cluster in the INIT state. Both nodes must be in the same state.

## 7.6.3 Starting Smart Assist for SAP: Database instance (DB)

After you complete the steps that are outlined in the previous sections, you are ready to start the Smart Assist for SAP as explained in the following steps:

1. Launch the Smart Assist for SAP by using the path for `sapsma1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-83), select **SAP Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

  DB2 UDB non-DPF Smart Assist     # sapsma1 sapsma2
  SAP MaxDB Smart Assist           # sapsma1 sapsma2
  SAP Smart Assist                 # sapsma1 sapsma2
```

*Figure 7-83 Selecting SAP Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-84), select **Automatic Discovery and Configuration**.

```
                           Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-84 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-85), select **5.SAP Database Instance**.

```
              Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  5.SAP Database Instance          # sapdb1
```

*Figure 7-85 Selecting the configuration to create*

**DB2 variable:** If there is no entry for the SAP Database Instance, you have not set the necessary DB2 variable, for example:

```
export DSE_INSTALL_DIR=/db2/TC1/db2tc1/db2_software
```

5. Select the line with your db2 user as shown in Figure 7-86 on page 436.

```
                Select a Configured DB2 Instance Resource Group

Move cursor to desired item and press Enter.

  db2tc1
```

*Figure 7-86 Select a Configured DB2 Instance Resource Group*

6. By using the available pick lists (F4), edit the DB2 Instance Owning Node by entering `[Not a terminal]`, Takeover Node by entering `[db2tc1]`, DB2 Instance Name by selecting **cl_db2sa_add_single_instance** and DB2 Instance Database to Monitor to blank as shown in Figure 7-87. Press Enter.

```
                           DB2 Instance Owning Node

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

                                                        [Entry Fields]
* Application Name                                    [db2tc1]

* DB2 Instance Owning Node                            [Not a terminal] +
* Takeover Node(s)                                    [db2tc1] +
* DB2 Instance Name                           cl_db2sa_add_single_instance +
* DB2 Instance Database to Monitor +
* Service IP Label                                    []
```

*Figure 7-87 DB2 Instance Owning Node before selection*

> **Important:** Check all values for correctness, even if there are values listed. Use the pickup list to select the correct entries.

Figure 7-88 shows the DB2 Instance Owning Node window after the selection.

```
                           DB2 Instance Owning Node

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

                                                        [Entry Fields]
* Application Name                                    [db2tc1]

* DB2 Instance Owning Node                            [sapdb1] +
* Takeover Node(s)                                    [sapdb2] +
* DB2 Instance Name                                    db2tc1 +
* DB2 Instance Database to Monitor                     TC1 +
* Service IP Label                                    [sapdbsvc2]
```

*Figure 7-88 DB2 Instance Owning Node after selection*

You might also see the cluster verification message that is shown in Figure 7-89 on page 437.

```
...
Importing volume group vg_TC1_sap to the following node(s):  sapdb2
sapdb2: 0516-783 importvg: This imported volume group is concurrent capable.
sapdb2:         Therefore, the volume group must be varied on manually.
sapdb2: 0516-1804 chvg: The quorum change takes effect immediately.
sapdb2: Volume group vg_TC1_sap has been imported.
Verification to be performed on the following:
        Cluster Topology
        Cluster Resources
...
```

*Figure 7-89 Cluster verification message*

The messages in Figure 7-89 explain that the volume group is automatically imported by the verification.

A new PowerHA resource group, `db2tc1_ResourceGroup`, is created. The volume group `vg_TC1_sap` and the service IP label `sapdbsvc2` are automatically added to the resource group as shown in Example 7-75.

*Example 7-75 Configured resource group for the SAP DB2 instance*

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # cllsgrp
db2tc1_ResourceGroup

root@sapdb1 /usr/es/sbin/cluster/sa/db2 # cllsres
APPLICATIONS="db2tc1_ApplicationServer"
FILESYSTEM=""
FORCED_VARYON="false"
FSCHECK_TOOL="fsck"
FS_BEFORE_IPADDR="false"
RECOVERY_METHOD="parallel"
SERVICE_LABEL="sapdbsvc2"
SSA_DISK_FENCING="false"
VG_AUTO_IMPORT="true"
VOLUME_GROUP="vg_TC1_sap"
USERDEFINED_RESOURCES=""
```

## 7.6.4 Completing the configuration

After the DB2 database is configured with the Smart Assist for SAP subtype DB, complete the configuration:

1. Stop the DB2 instance on the primary node if it is running, as shown in Example 7-76.

*Example 7-76 Stopping the DB2 instance*

```
sapdb1:/ # su - db2tc1

sapdb1:/db2/db2tc1 # db2stop
```

2. Unmount the shared file systems as shown in Example 7-77.

*Example 7-77 Unmounting the shared file systems*

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # umount -t TC1
```

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # lsvg -l vg_TC1_sap
vg_TC1_sap:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
TC1.lvdb2binary     jfs2       128     128     1    closed/syncd  /db2/TC1
TC1.lvdb2logdir     jfs2       64      64      1    closed/syncd
/db2/TC1/log_dir
TC1.lvdb2logarc     jfs2       64      64      1    closed/syncd
/db2/TC1/log_archive
TC1.lvdb2dat.01     jfs2       320     320     1    closed/syncd
/db2/TC1/sapdata1
TC1.lvdb2dat.02     jfs2       320     320     1    closed/syncd
/db2/TC1/sapdata2
TC1.lvdb2dat.03     jfs2       320     320     1    closed/syncd
/db2/TC1/sapdata3
TC1.lvdb2dat.04     jfs2       320     320     1    closed/syncd
/db2/TC1/sapdata4
```

3. Deactivate the shared volume group as shown in Example 7-78.

*Example 7-78 Deactivating the shared volume group*

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # varyoffvg vg_TC1_sap
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # lsvg -o
caavg_private
rootvg
```

4. Unconfigure the network alias on both nodes as shown in Example 7-79.

*Example 7-79 Network settings on sapdb1*

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c6.b6.6f  9237187     0  3543150     0     0
en0   1500  10.1        sapdb1b1           9237187     0  3543150     0     0
en0   1500  172.16.20   sapdb1p1           9237187     0  3543150     0     0
en1   1500  link#3      6e.8d.dd.c6.b6.70   344897     0    42429     0     0
en1   1500  10.1.2      sapdb1b2            344897     0    42429     0     0
en1   1500  10.10.20    sapdb1p2            344897     0    42429     0     0
en1   1500  10.10.20    sapdbsvc2           344897     0    42429     0     0
lo0   16896 link#1                         2784035     0  2784030     0     0
lo0   16896 127         loopback           2784035     0  2784030     0     0
lo0   16896 loopback                       2784035     0  2784030     0     0
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # ifconfig en1 delete sapdbsvc2
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.dd.c6.b6.6f  9237681     0  3543223     0     0
en0   1500  10.1        sapdb1b1           9237681     0  3543223     0     0
en0   1500  172.16.20   sapdb1p1           9237681     0  3543223     0     0
en1   1500  link#3      6e.8d.dd.c6.b6.70   344902     0    42429     0     0
en1   1500  10.1.2      sapdb1b2            344902     0    42429     0     0
en1   1500  10.10.20    sapdb1p2            344902     0    42429     0     0
lo0   16896 link#1                         2784360     0  2784355     0     0
lo0   16896 127         loopback           2784360     0  2784355     0     0
lo0   16896 loopback                       2784360     0  2784355     0     0
```

5. Synchronize the PowerHA cluster by using SMIT:

   a. Follow the path: **`smitty sysmirror`** → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.

   b. In the PowerHA SystemMirror Verification and Synchronization panel (Figure 7-90), press Enter to accept the default options.

```
           PowerHA SystemMirror Verification and Synchronization

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Verify, Synchronize or Both                         [Both]          +
* Include custom verification library checks          [Yes]           +
* Automatically correct errors found during           [Yes]           +
  verification?
* Force synchronization if verification fails?        [No]            +
* Verify changes only?                                [No]            +
* Logging                                             [Standard]      +
```

*Figure 7-90 Accepting the default actions on the Verification and Synchronization panel*

6. Verify that the IP alias is on the correct network.
7. Start the cluster on both nodes, `sapdb1` and `sapdb2`, by running **`smitty clstart`**.
8. In the Start Cluster Services panel (Figure 7-91 on page 440), complete these steps:

   a. For Start now, on system restart or both, select **now**.
   b. For Start Cluster Services on these nodes, enter `[sapdb1 sapdb2]`.
   c. For Manage Resource Groups, select **Automatically**.
   d. For BROADCAST message at startup, select **false**.
   e. For Startup Cluster Information Daemon, select **true**.
   f. For Ignore verification errors, select **false**.
   g. For Automatically correct errors found during cluster start, select **yes**.

   Press Enter.

```
                              Start Cluster Services

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Start now, on system restart or both                now                    +
  Start Cluster Services on these nodes               [sapdb1 sapdb2]        +
* Manage Resource Groups                              Automatically          +
  BROADCAST message at startup?                       false                  +
  Startup Cluster Information Daemon?                 true                   +
  Ignore verification errors?                         false                  +
  Automatically correct errors found during           yes                    +
  cluster start?
```

*Figure 7-91 Specifying the options for starting cluster services*

**Tip:** The log file for the Smart Assist is in the `/var/hacmp/log/sa.log` file. You can use the **`clmgr`** utility to easily view the log:

```
clmgr view log sa.log
```

When the PowerHA cluster starts, the DB2 instance starts automatically. The application monitors the start after the defined stabilization interval as shown in Example 7-80.

*Example 7-80 Checking the status of the highly available cluster and the DB2 instance*

```
root@sapdb1 /usr/es/sbin/cluster/sa/db2 # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                        Node
-----------------------------------------------------------------------------
db2tc1_Resourc ONLINE                       sapdb1
               OFFLINE                      sapdb2

root@sapdb1 /usr/es/sbin/cluster/sa/db2 # ps -ef | grep
/usr/es/sbin/cluster/clappmond | grep -v grep
    root 18874568 29884530   0 14:30:53      -  0:00
/usr/es/sbin/cluster/clappmond db2tc1_ProcessMonitor
```

Your DB2 instance and database are now configured for high availability in a hot-standby PowerHA SystemMirror configuration.

## 7.7 Cluster 3: liveCache

This cluster makes an SAP MaxDB liveCache available as a hot standby. This book shows a solution with a DS8000 as the storage system. Inside the DS8000, a snapshot mechanism is available to create a fast backup of the HotStandby liveCache, if necessary.

More information about the SAP liveCache installation is in “Invincible Supply Chain - Implementation Guide for SAP HotStandby liveCache with PowerHA 7.1.1”. The file, `PowerHA7.1.1_HotStandbyImplementationPath_v1.pub.pdf` (version at press date), is on techdocs and is published by the International SAP IBM Competence Center (ISICC):

http://w3-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100677

In this document, you can obtain the following information:

- Planning your liveCache installation
- Implementing a SAN Volume Controller-based solution instead of a DS8000-based solution
- Sizing, tuning, and administering your liveCache

## 7.7.1 Overview

This cluster uses the Smart Assist for MaxDB: liveCache to make a liveCache/MaxDB database highly available (Figure 7-92).

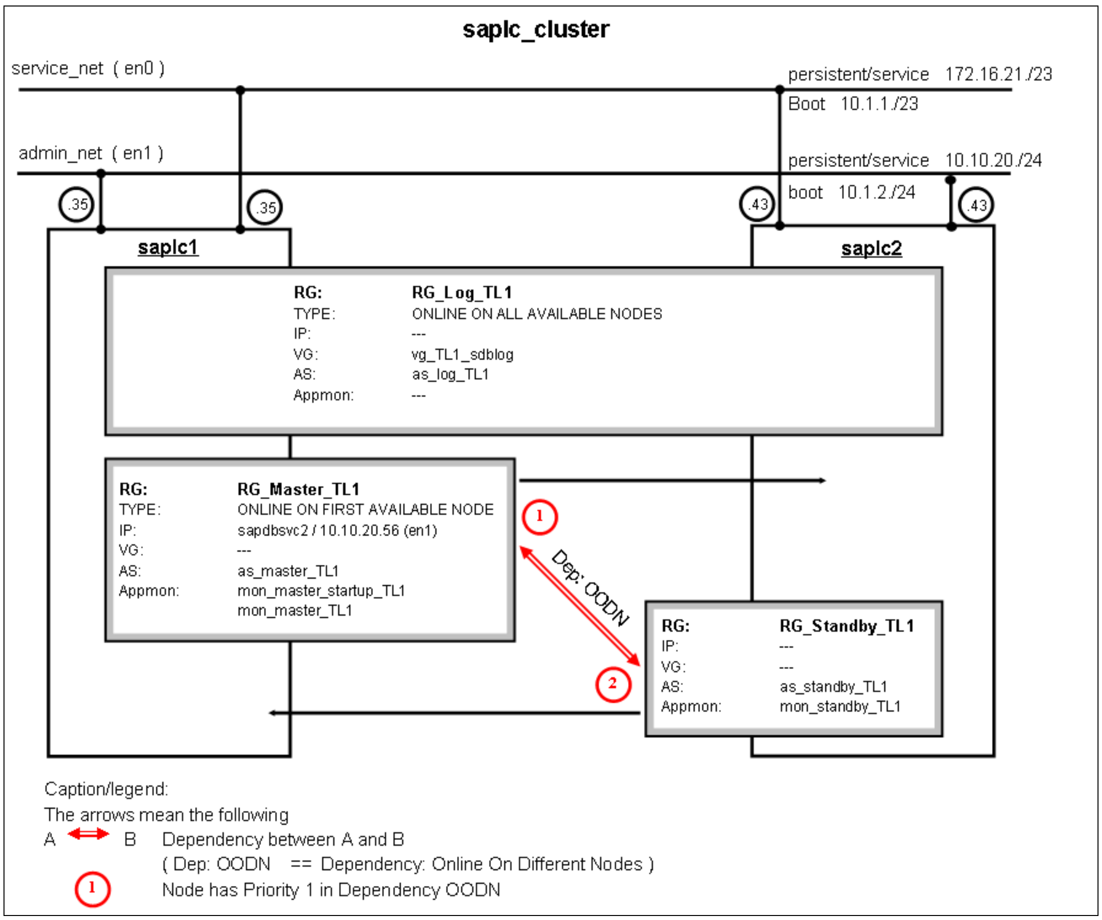



*Figure 7-92 Schema overview of saplc_cluster*

Figure 7-92 shows the entire environment for the `saplc_cluster`.

## 7.7.2 Prerequisites for the hot-standby liveCache Smart Assist

We show the prerequisites for the hot-standby liveCache Smart Assist:

- Executables directory `/sapdb` is a local directory or even better a local file system on both nodes.
- Data volumes for liveCache are local raw devices.

- Log volumes for liveCache are shared raw devices because both nodes need access to that data at the same time.
- The Smart Assist for MaxDB only discovers NPIV disks automatically. If vscsi is used, the manual steps are documented on techdocs.
- The naming conventions that are provided by the framework must not be changed.
- The amount and size of hdisks, logical volumes, and volume groups are limited (see also the techdocs document).
- The wizard only allows one Hardware Management Console (HMC) to access the storage system to be used during the automation. For redundancy, a second HMC can be manually added after the wizard is run.
- The liveCache version 7.7.07.39 or later is required.
- The wizard only functions in a two-node configuration.
- The automation works for SAN Volume Controller-based storage attachment through Secure Shell (ssh) and dscli for DS8000.
- A notification method must be included in the PowerHA application monitors.

## 7.7.3 PowerHA SystemMirror supported disk configurations

The PowerHA SystemMirror SAP liveCache HotStandby solution is tested for many disks, volume groups, and volumes. The configuration options that are shown in Table 7-7 are supported.

*Table 7-7 PowerHA SystemMirror supported disk configuration*

| Component | Size |
|---|---|
| Logical volume size | Up to 128 GB |
| Number of Disks/Volume Group | Up to six |
| Number of Volumes/Volume Group | Up to 12 |
| Number of Volume Groups in Log or Data Volume Groups that are activated between the Master and the Standby | Up to three |

For any relevant limits on the disk and volume sizes that SAP advises, see the SAP documentation.

The SAP liveCache supports individual raw device sizes up to 128 GB. In the typical layout, a raw device is represented by a raw logical volume. Therefore, the physical disks can be greater than 128 GB and broken down into multiple logical volumes, not to exceed 12 logical volumes per volume group.

On the storage level only, one volume is supported by this version because no consistency group is created automatically on the DS8000 storage level.

## 7.7.4 Installation and configuration steps before you use Smart Assist for MaxDB

We show the installation and configuration steps before you use the Smart Assist for MaxDB:

1. Install the required filesets.

See 7.2.1, “Installing the required PowerHA SystemMirror Smart Assist filesets” on page 362 for the additional file for Smart Assist for MaxDB.

2. Set the global required operating system (OS) and TCP/IP parameter.

   On both nodes, set the following OS and TCP/IP parameters:

   – Operating system tuning

     Increase the maximum number of processes to at least 2,000. Use the following operating system command:

     ```
     chdev -l sys0 -a maxuproc='2000'
     ```

   – TCPIP tuning

     To avoid batch jobs that stop as a result of the failover, the socket handling needs to be adjusted. The time until the connection is closed needs to be reduced on the liveCache LPARs and on all application servers that connect to the liveCache. The following suggestions must be validated against other prerequisites that might be set by applications that are not part of this book to ensure that there are no unexpected side effects. It is essential to make the changes permanent to span a reboot.

     The following values are set:

     ```
     no -p -o tcp_keepintvl=10
     no -p -o tcp_keepidle=300
     no -p -o tcp_keepinit=20
     ```

     To make them permanent after the reboot, the **-p** (or **-r**) flag is required.

   – Change root user limits

     The same installation and configuration steps run as user `root`. Set the following soft and hard limits for CPU time, file size, data segment size, Receive Side Scaling (RSS) size, and stack size to unlimited for user `root`. Use the following command:

     ```
     chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu_hard='-1'
     fsize_hard='-1' data_hard='-1' stack_hard='-1' rss_hard='-1' root
     ```

3. Configure the base IBM PowerHA SystemMirror.

   You must configure the topology of the PowerHA cluster before you use the Smart Assist for SAP. Example 7-81 shows the cluster `saplc_cluster` that is configured with two Ethernet interfaces in each node.

*Example 7-81 Cluster network configuration*

```
root@saplc1 / # lscluster -c
Cluster query for cluster saplc_cluster returns:
Cluster uuid: c7843bf4-720b-11e1-b3ad-2e479785d36f
Number of nodes in cluster = 2
        Cluster id for node saplc1b1 is 1
        Primary IP address for node saplc1b1 is 10.1.1.35
        Cluster id for node saplc2b1 is 2
        Primary IP address for node saplc2b1 is 10.1.1.43
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.25

root@saplc1 / # cllsif
Adapter              Type       Network    Net Type   Attribute  Node       IP
Address       Hardware Address Interface Name   Global Name     Netmask
Alias for HB Prefix Length

saplc1b2 boot admin_net  ether public saplc1 10.1.2.35 en1 255.255.255.0 24
```

```
saplc1b1 boot service_net ether public saplc1 10.1.1.35 en0 255.255.254.0 23
saplc2b2 boot admin_net  ether public saplc2 10.1.2.43 en1 255.255.255.0 24
saplc2b1 boot service_net ether public saplc2 10.1.1.43 en0 255.255.254.0 23
```

4. Configure the network (Example 7-82).

   The IP addresses for the SAP liveCache system (`saplcsvc1`) are configured on the node where the SAP liveCache is installed.

*Example 7-82 Network settings*

```
oot@saplc1 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      2e.47.97.85.d3.6f 11049701     0   502349     0     0
en0   1500  10.1        saplc1b1          11049701     0   502349     0     0
en0   1500  172.16.20   saplc1p1          11049701     0   502349     0     0
en0   1500  172.16.20   saplcsvc1         11049701     0   502349     0     0
en1   1500  link#3      2e.47.97.85.d3.70  4346133     0    67538     0     0
en1   1500  10.1.2      saplc1b2           4346133     0    67538     0     0
en1   1500  10.10.20    saplc1p2           4346133     0    67538     0     0
en2   1500  link#4      2e.47.97.85.d3.71  2064876     0     6958     0     0
en2   1500  9.12        saplc1             2064876     0     6958     0     0
lo0   16896 link#1                         1572475     0  1572476     0     0
lo0   16896 127         loopback           1572475     0  1572476     0     0
lo0   16896 loopback                       1572475     0  1572476     0     0
```

5. Create the users and groups.

   See 7.2.2, “Creating the users” on page 362 and the script to create the users at “Script to create users and groups for sapci_cluster, sapdb_cluster, and saplc_cluster” on page 604.

   The user account `sdb` must be locked. If not, the liveCache installation fails because this user account is checked:

   `chuser account_locked=true sdb`

6. Create the volume groups, logical volumes, and file systems.

   See 7.2.3, “Creating volume groups, logical volumes, and file systems” on page 363. Also, see the script to create the volume groups and file systems in the “Script to create users and groups for sapsma_cluster” on page 604.

   The volume group `vg_TL1_sdbdat` is local on both nodes, and each node has the logical volumes for the liveCache data. The volume group `vg_TL1_sdblog` is shared between the two nodes, and each node has the logical volumes for the log volumes for the MaxDB logging. See Example 7-83.

*Example 7-83 Local volume group and /sapdb file system at saplc1*

```
root@saplc1 / # lsvg -l vg_TL1_sdbdat
vg_TL1_sdbdat:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_data1        jfs2       1596    1596    2    open/syncd    N/A
sapdblv             jfs2       64      128     2    open/syncd    /sapdb
```

Example 7-84 on page 445 shows the local volume group and the `/sapdb` file system for `saplc2`.

*Example 7-84 Local volume group and /sapdb file system at saplc2*

```
root@saplc1 / # lsvg -l vg_TL1_sdbdat
vg_TL1_sdbdat:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_data1        jfs2       1596    1596    2    open/syncd    N/A
sapdblv             jfs2       64      128     2    open/syncd    /sapdb
```

Example 7-85 shows the shared volume group for the `saplc_cluster`.

*Example 7-85 Shared volume groups at saplc_cluster*

```
root@saplc1 / # clcmd lsvg -l vg_TL1_sdblog

-------------------------------
NODE saplc2b1
-------------------------------
vg_TL1_sdblog:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_log1         jfs2       399     399     1    open/syncd    N/A
lv_TL1_log2         jfs2       399     399     1    open/syncd    N/A


-------------------------------
NODE saplc1b1
-------------------------------
vg_TL1_sdblog:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_log1         jfs2       399     399     1    open/syncd    N/A
lv_TL1_log2         jfs2       399     399     1    open/syncd    N/A
```

7. Configure the NFS for the application state information.

> **Important:** You can use every highly available NFS server for the directory service for the liveCache application state information. But in the demonstration environment, we used the NFS server from the `sapci_cluster` as an NFS version 4 service because there is no external NFS server that is available in the demonstration environment.

An NFS domain must be configured on both nodes because we use the NFS server from the `sapci_cluster` in the demonstration environment. Because the Smart Assist for SAP uses NFS version 4, we have to configure NFS version 4 and use the same NFS domain as the NFS cluster:

```
root@saplc1 / # clcmd chnfsdom sapci_cluster
```

Set the necessary settings for NFS version 4:

```
stopsrc -g nfs
smitty nfsgrcperiod
startsrc -g nfs
```

Check whether the NFS server that we use from the `sapci_cluster` exports all needed file systems as shown in Example 7-86 on page 446.

*Example 7-86 NFS server on sapci_cluster*

```
root@sapci1 / # exportfs
/export/sapmnt/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/saptrans/TC1
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=sapci1:sapci2:sapci1b2:sapci2b2:sap
cisvc2:sapcisvc4:sapci1p2:sapci2p2
/export/archive/TL1_lock
-vers=4,sec=sys:krb5p:krb5i:krb5:dh,rw,root=saplc1:saplc2:saplc1b2:saplc2b2:sap
lcsvc1:saplcsvc2:saplc1p2:saplc2p2
```

Mount as the NFS client on both nodes:

```
root@saplc1 / # mount -o vers=4,hard,intr sapcisvc2:/export/archive/TL1_lock
/archive/sapdb/TL1_LOCK
root@saplc2 / # mount -o vers=4,hard,intr sapcisvc2:/export/archive/TL1_lock
/archive/sapdb/TL1_LOCK
```

The following NFS client mounts must be available as shown in Example 7-87.

*Example 7-87 NFS mounts on saplc1*

```
root@saplc1 / # mount|grep nfs
sapci1   /export/archive/TL1_lock /archive/sapdb/TL1_LOCK nfs4   Apr 12 03:49
vers=4,hard,intr
```

The easiest way to make this mount available after each reboot is to configure the automountd as shown in Example 7-88.

*Example 7-88 Automounter config on saplc_cluster*

```
root@saplc1 / # cat /etc/auto*
/-             /etc/autofs_root.SAP   -nobrowse
#/sapmnt       /etc/autofs_import.SAP -nobrowse
/-             /etc/autofs_root.SAP   -nobrowse
#/usr/sap/trans -fstype=nfs,timeo=1,retry=3,rw,hard,intr
sapnfssvc1:/export/saptrans
/archive/sapdb/TL1_LOCK -fstype=nfs,timeo=1,retry=3,rw,hard,intr,vers=4
sapcisvc2:/export/archive/TL1_lock
```

8. Set up the cluster state lock directory.

   The configured directory `/archive/sapdb/TL1_LOCK` in the previous section needs to be dedicated to this particular liveCache instance. The following users need to have read and write access:

   – root
   – sdb

   These users can have read and write access by setting `chmod 777 /archive/sapdb/TL1_LOCK` or by setting `chown sdb:sdba /archive/sapdb/TL1_LOCK`.

   The configured directory `/archive/sapdb/TL1_LOCK` must be attached to both nodes as shown in the previous section and highly available.

   The mount point of this directory (in our example, `/archive/sapdb/TL1_LOCK`) is one of the configuration parameters of the Smart Assist automation.

9. Install the SAP liveCache on the first node.

We performed a standard SAP installation. We set only the usual SAP parameters as shown in Example 7-89.

*Example 7-89 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst
```

The SAPinst in the installation CD-ROM provides options when you create a liveCache instance. Important entries are highlighted along with the screen captures. It is possible to first install the MAXDB database, but the easier way is to let the SAPinst install it for you.

Start the liveCache installation as shown in Figure 7-93. Select **liveCache Server Installation**.

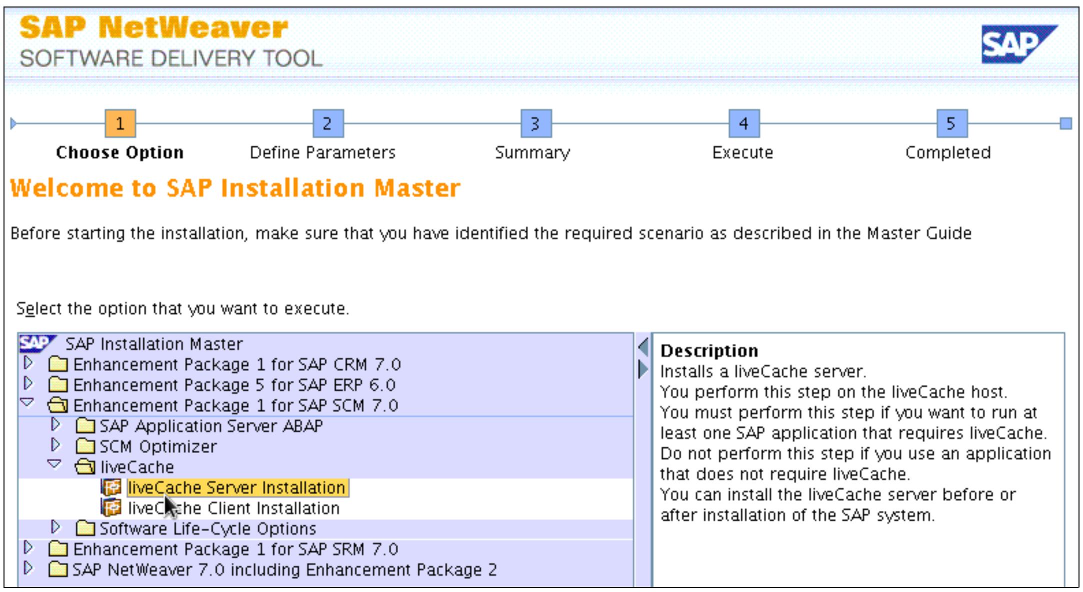



*Figure 7-93 SAPinst liveCache installation screen capture 1: Start the database installation*

The complete screen captures from the installation are in “sapci_cluster: Installing the SAP ASCS instance on the first node” on page 609.

10. Install the liveCache database executables on the second node.

In Example 7-90, the liveCache executables are installed from an image that is downloaded from the SAP marketplace by using the command-line installation **SDBINST**. The installation images in Example 7-90 run from a shared file system.

Use SAP **SDBINST** to install the liveCache software and select the option for **Server + Client**. In this example, we selected the dependent data path according to the SID that is selected for the liveCache instance that is being implemented: `/sapdb/TL1`. The result is that all dependent data that is generated for the implementation is placed under the relevant SID in the `/sapdb` file system.

*Example 7-90 liveCache executables installation on saplc2*

```
root@saplc2
/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_7
_07_39 # ./SDBINST
```

```
        Installation of SAP liveCache Software
        ***************************************


starting installation Th, Apr 12, 2012 at 05:34:43
operating system: AIX PowerPC 7.1.0.0
callers working directory:
/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_7
_07_39
installer directory:
/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_7
_07_39
archive directory:
/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_7
_07_39


existing component groups:
0:  Server + Client
1:  Client
2:  Custom
3:  none
please enter component group id: 0


starting preparing phase of package Base 7.7.07.39 64 bit
---------------------------------------------------------

no updatable installation of package "Base" found
please enter group name for database programs [sdba]:
please enter owner name for database programs [sdb]:
please enter independent data path [/var/opt/sdb/data]: /sapdb/data
directory "/sapdb/data" does not exist, create? (y/n) y
please enter independent program path [/opt/sdb/programs]: /sapdb/programs
directory "/sapdb/programs" does not exist, create? (y/n) y
checking interferences to other packages... ok

collecting data finished:

independent data path: /sapdb/data
independent program path: /sapdb/programs
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/SDBBAS.TGZ"
checking mmap ...
mmap check OK

package Base successfully checked

starting preparing phase of package SAP Utilities 7.7.07.39 64 bit
------------------------------------------------------------------
```

```
checking interferences to other packages... ok

collecting data finished:

: /sapdb/data
independent program path: /sapdb/programs
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/SAPUTL.TGZ"

package SAP Utilities successfully checked

starting preparing phase of package Server Utilities 7.7.07.39 64 bit
---------------------------------------------------------------------

checking interferences to other packages... ok

collecting data finished:

independent program path: /sapdb/programs
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/SDBUTL.TGZ"

package Server Utilities successfully checked

starting preparing phase of package Database Kernel 7.7.07.39 64 bit
--------------------------------------------------------------------

no updatable installation of package "Database Kernel" found
please enter dependent path [/opt/sdb/7707]: /sapdb/TL1/db
directory "/sapdb/TL1/db" does not exist, create? (y/n) y
checking interferences to other packages... ok

collecting data finished:

dependent path: /sapdb/TL1/db
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/SDBKRN.TGZ"

package Database Kernel successfully checked

starting preparing phase of package JDBC 7.6.06.05
--------------------------------------------------

checking interferences to other packages... ok

collecting data finished:
```

```
java driver path: /sapdb/programs
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/SDBJDBC.TGZ"

package JDBC successfully checked

starting preparing phase of package SQLDBC 7.7.07.39 64 bit
-----------------------------------------------------------

[...]

starting preparing phase of package APO LC APPS 7.00.019 64 bit
---------------------------------------------------------------

checking interferences to other packages... ok

collecting data finished:

apo com path: /sapdb/TL1/db/sap
owner: sdb
group: sdba
start extraction test run of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/APOLCA.TGZ"

package APO LC APPS successfully checked


checking filesystem "/sapdb"... free disk space ok

starting installation phase of package Base 7.7.07.39 64 bit
------------------------------------------------------------

[...]

starting installation phase of package APO LC APPS 7.00.019 64 bit
------------------------------------------------------------------

start real extraction of
"/install/SAP/Database/MAXDB/MAXDB_7.7.07.39/lca700019_livecache-aix5-64bit-ppc-7_
7_07_39/APOLCA.TGZ"
extracting: -r--r--r--         315 2012-01-09 05:10:23 lcapps.pkg
extracting: -r-xr-xr-x    41614799 2012-01-09 08:43:29 liball_api.so
extracting: -r-xr-xr-x    37342503 2012-01-09 09:48:11 liball_api_apo.so
extracting: -r-xr-xr-x    27150990 2012-01-09 07:40:09 liball_aps.so
extracting: -r-xr-xr-x     2024155 2012-01-09 05:21:23 libaps_base.so
extracting: -r-xr-xr-x    13337639 2012-01-09 07:19:47 libaps_peg.so
extracting: -r-xr-xr-x    12125004 2012-01-09 07:50:35 libaps_scheduler.so
extracting: -r-xr-xr-x    25568812 2012-01-09 07:03:39 libaps_dam.so
extracting: -r-xr-xr-x     6290026 2012-01-09 07:09:17 librpm_exports.so
extracting: -r-xr-xr-x    15664682 2012-01-09 05:35:23 libtsdp_api.so
```

```
extracting: -r-xr-xr-x      1940641 2012-01-09 05:36:03 libtsdp_test.so
extracting: -r-xr-xr-x      9810932 2012-01-09 05:18:19 libCOMBase.so
extracting: -r--r--r--          790 2012-01-09 05:10:23 libSAPBAS.lst
extracting: -r-xr-xr-x      2310896 2012-01-09 05:19:13 libSAPBAS.so
extracting: -r--r--r--         4219 2012-01-09 05:10:23 libSAPAPO.lst
extracting: -r-xr-xr-x     19618300 2012-01-09 10:01:33 libSAPAPO.so
extracting: -r--r--r--         1267 2012-01-09 05:10:23 libSAPATP.lst
extracting: -r-xr-xr-x     58080724 2012-01-09 06:44:39 libSAPATP.so
extracting: -r--r--r--          353 2012-01-09 05:10:23 libSAPLCK.lst
extracting: -r-xr-xr-x      4423894 2012-01-09 05:41:09 libSAPLCK.so
extracting: -r--r--r--          538 2012-01-09 05:10:23 libSAPREP.lst
extracting: -r-xr-xr-x      4506394 2012-01-09 08:43:19 libSAPREP.so
extracting: -r--r--r--         1429 2012-01-09 05:10:23 libSAPRPM.lst
extracting: -r-xr-xr-x      5678809 2012-01-09 10:07:57 libSAPRPM.so
extracting: -r--r--r--          305 2012-01-09 05:10:23 libSAPSIM.lst
extracting: -r-xr-xr-x      3251344 2012-01-09 05:22:47 libSAPSIM.so
extracting: -r--r--r--         1973 2012-01-09 05:10:23 libSAPTS.lst
extracting: -r-xr-xr-x      7061723 2012-01-09 05:39:25 libSAPTS.so
extracting: -r-xr-xr-x      1478520 2011-10-11 11:51:37 niping
checking unpacked archive... ok
installation of SAP liveCache Software finished successfully Th, Apr 12, 2012 at
05:37:56
```

11. Copy the additional directories.

    So that the SAP liveCache can be started with the correct user environment on the second cluster node with user `tl1adm`, you have to copy additional files and directories manually to the other node:

    ```
    root@saplc1 / # scp -pr /home/sdb/* saplc2:/home/sdb/
    root@saplc1 / # scp -pr /home/sdb/.* saplc2:/home/sdb/
    ```

    > **Important:** The command **`clrcp`** works with files only, not directories.

12. Verify that the liveCache kernel versions on both nodes are the same.

### 7.7.5 Preliminary steps

We explain the required preliminary steps before you start the wizard and how to start the Smart Assist for SAP.

Before you start the Smart Assist for MaxDB, complete the following steps:

1. You can start the PowerHA SAP liveCache Hot Standby Wizard on a running cluster only.
2. The local file system `/sapdb` has to be online on both nodes. The shared volume groups for the liveCache logging have to be *concurrent* on both nodes `saplc1` and `saplc2` (Example 7-91). The local volume groups for the database data exist only on the primary node and have to be active there (Example 7-92 on page 452). On the backup node, no volume group for liveCache has to be configured.

   *Example 7-91 Activate concurrent volume group on both node saplc1 and node saplc2*

   ```
   varyonvg -c vg_TL1_sdblog
   ```

*Example 7-92 Checking for active volume groups on node saplc1*

```
root@saplc1 / # lspv
hdisk0          00f74d45f95ba523                    rootvg          active
hdisk1          00f74d45fa931a97                    rootvg          active
hdisk2          00f74d451796915b                   caavg_private   active
hdisk4          00f74d45179693fb                   vg_TL1_sdbdat   active
hdisk5          00f74d45179695a2                   vg_TL1_sdbdat   active
hdisk6          00f74d4531d87200                    pagingvg        active
hdisk7          00f74d4531d87400                    pagingvg        active
hdisk8          00f74d45a6f36283                   vg_TL1_sdblog   concurrent
hdisk3          00f74d451796925f                    None
root@saplc1 / # lsvg -o|grep TL1|lsvg -l -i
vg_TL1_sdblog:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_log1         jfs2       399     399     1    open/syncd    N/A
lv_TL1_log2         jfs2       399     399     1    open/syncd    N/A
vg_TL1_sdbdat:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_data1        jfs2       1596    1596    2    open/syncd    N/A
```

Example 7-93 checks for the active volume groups on node `saplc2`.

*Example 7-93 Checking for active volume groups on node saplc2*

```
root@saplc2 /opt/ibm/dscli # lspv
hdisk0          00f74d47f960f2b6                    rootvg          active
hdisk1          00f74d47fa933eba                    rootvg          active
hdisk2          00f74d451796915b                   caavg_private   active
hdisk4          00f74d47ac05cbf1                    None
hdisk5          00f74d47ac05cfa5                    None
hdisk6          00f74d4731d8c9d0                    pagingvg        active
hdisk7          00f74d4731d8cbbb                    pagingvg        active
hdisk8          00f74d45a6f36283                 vg_TL1_sdblog   concurrent

root@saplc2 /opt/ibm/dscli # lsvg -o|grep TL1|lsvg -l -i
vg_TL1_sdbdat:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
lv_TL1_data1        jfs2       1596    1596    2    open/syncd    N/A
```

3. The log and data volumes of a HotStandby liveCache are accessed as raw devices. Therefore, the liveCache user (`sdb`) must have the authority to access the devices at the raw level. Set the permissions for the liveCache raw devices to user `sdb` and group `sdba`. This step remains the same regardless of the liveCache SID and the administration user `<SID>adm`. Normally, the SAPinst set the permissions, but you have to check the permissions on the second node and maybe also if the wizard fails.

   Example 7-94 on page 453 shows three logical volumes: two volumes for the log and one volume for the data. They are named `lv_TL1_log1`, `lv_TL1_log2`, and `lv_TL1_data1`.

```
chown sdb.sdba /dev/rlv_TL1_log1
chown sdb.sdba /dev/rlv_TL1_log2
chown sdb.sdba /dev/rlv_TL1_data1
```

*Example 7-94 Raw logical volumes on saplc1*

```
root@saplc1 / # ls -al /dev/rlv_*
crw-rw----    1 sdb      sdba        101,  1 Apr 13 10:26 /dev/rlv_TL1_data1
crw-rw----    1 sdb      sdba        102,  1 Apr 13 10:26 /dev/rlv_TL1_log1
crw-rw----    1 sdb      sdba        102,  2 Apr 12 08:51 /dev/rlv_TL1_log2
```

4. The MAXDB/liveCache database on the primary node has to be started as shown in Example 7-95.

*Example 7-95 Start liveCache database on the first node*

```
root@saplc1 / # su - sdb
$ /sapdb/programs/bin/dbmcli -d TL1 -u control,pw2201tc1
/sapdb/programs/bin/dbmcli on TL1>db_state
OK
State
ONLINE

---
```

5. Run the following command as user `root` and `sdb` on both nodes to prepare the database to connect to the SAPDB database:

   `/sapdb/programs/bin/xuser -U <SID>_XUSER -u <control>,password -d <SID>`

   For our example, we used the following command:

   `/sapdb/programs/bin/xuser -U TL1_XUSER -u control,pw2201tc1 -d TL1`

6. Download and install the DS8000 dscli lppsource on both AIX nodes as described in the corresponding `readme` file. No further customization is necessary.

### 7.7.6 Starting SAP liveCache Hot Standby Wizard

After you complete the steps in the preceding sections, you are ready to start the SAP liveCache Hot Standby Configuration Wizard as explained in the following steps:

1. Launch the wizard by using the path for `saplc1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **SAP liveCache Hot Standby Configuration Wizard**.
2. In the Select Node Names panel (Figure 7-94), select both nodes, **saplc1** and **saplc2**.

```
                              Select Node Names

  Move cursor to desired item and press F7.
        ONE OR MORE items can be selected.
  Press Enter AFTER making all selections.

  > saplc1
  > saplc2
```

*Figure 7-94 Selecting the nodes*

3. In the Storage subsystems found panel (Figure 7-95), select **DS8000**.

```
                         Storage subsystem(s) found

    Move cursor to desired item and press Enter.

       DS8000
```

*Figure 7-95 Storage subsystems found*

4. Enter all necessary information as shown in Figure 7-96:
   - For liveCache Instance Name (SID), type [TL1].
   - For MaxDB DBM User XUSER, type [TL1_XUSER].
   - For Storage (HMC) type, enter [DS8000].
   - For Storage server (HMC) IP, type [9.12.6.17].
   - For Storage (HMC) User, type [pw2201copy].
   - For Storage (HMC) Password, type [lunc0py].
   - For liveCache Global Filesystem Mount point, type [/archive/sapdb/TL1_LOCK].

```
                    SAP liveCache Hot Standby Configuration

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

                                                         [Entry Fields]
* liveCache Instance Name (SID)                       [TL1]
* MaxDB DBM User XUSER                                [TL1_XUSER]
* Storage (HMC) type                                  [DS8000]
* Storage server (HMC) IP                             [9.12.6.17]
* Storage (HMC) User                                  [pw2201copy]
* Storage (HMC) Password                              [lunc0py]
* liveCache Global Filesystem Mount point           [/archive/sapdb/TL1_LOCK]
* Primary Node                                         saplc1 +
* Service IP Label                                     saplcsvc1 +
* Node Names                                           saplc1,saplc2
* Log Volume Group(s)                                  vg_TL1_sdblog +
* Data Volume Group(s)                                 vg_TL1_sdbdat +
* HDISK(s) pairs of Data Volume Group(s)   [hdisk4-->hdisk4,hdisk5-->hdisk5]
```

*Figure 7-96 SAP liveCache Hot Standby Configuration*

Alternatively, you can use the command line as shown in Example 7-96.

*Example 7-96 liveCache wizard command line*

```
/usr/es/sbin/cluster/sa/hswizard/sbin/cl_hotstandby_createInstance \
   -I'TL1' \
   -D'TL1_XUSER' \
   -t'DS8000' \
   -H'9.12.6.17' \
   -S'pw2201copy' \
   -T'lunc0py' \
   -F'/archive/sapdb/TL1_LOCK' \
   -p'saplc1' \
   -s'saplcsvc1' \
```

```
  -n'saplc1,saplc2' \
  -l'vg_TL1_sdblog' \
  -d'vg_TL1_sdbdat' \
  -o'hdisk4-->hdisk4,hdisk5-->hdisk5'
```

**Important:** The SHMC password is inserted in a readable format, which you must remember if you provide information externally, such as smitty logs.

Example 7-97 shows the command output from the smitty panel.

*Example 7-97 Smitty command output liveCache wizard*

```
Fri Apr 13 10:12:00 EDT 2012 - INFO_MSG : Initializing SAP liveCache Hot
Standby Wizard....
Fri Apr 13 10:12:04 EDT 2012 - INFO_MSG : Validating User Input..
Fri Apr 13 10:12:04 EDT 2012 - INFO_MSG : Validating Nodelist ....
Fri Apr 13 10:12:04 EDT 2012 - INFO_MSG : Verifying if all the nodes belong to
same site
Fri Apr 13 10:12:04 EDT 2012 - INFO_MSG : Sites are not Configured.
Proceeding..
Fri Apr 13 10:12:04 EDT 2012 - INFO_MSG : Validating Disks....
Fri Apr 13 10:12:06 EDT 2012 - INFO_MSG : Verifying if all the disks are of
same size
DS8000:hdisk4@saplc1
DS8000:hdisk4@saplc2
DS8000:hdisk5@saplc1
DS8000:hdisk5@saplc2
Fri Apr 13 10:12:06 EDT 2012 - INFO_MSG : Verifying if same number of disks are
selected from each node
Fri Apr 13 10:12:08 EDT 2012 - INFO_MSG : Validating MaxDB Configuration....
Fri Apr 13 10:12:08 EDT 2012 - INFO_MSG : Verifying MaxDB version on chosen
nodes
Fri Apr 13 10:12:08 EDT 2012 - INFO_MSG : Verifying UID and GID on chosen nodes
Fri Apr 13 10:12:09 EDT 2012 - TRACE_MSG : MaxDB Version on "saplc1" is
"7.7.07.39".
Fri Apr 13 10:12:09 EDT 2012 - TRACE_MSG : MaxDB Version on "saplc2" is
"7.7.07.39".
Fri Apr 13 10:12:09 EDT 2012 - INFO_MSG : Validating Instance name....
Fri Apr 13 10:12:09 EDT 2012 - INFO_MSG : Verifying if a liveCache instance
already exists with given name
Fri Apr 13 10:12:10 EDT 2012 - INFO_MSG : Validating HMC....
Fri Apr 13 10:12:10 EDT 2012 - INFO_MSG : Verifying if HMC can be reached
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : Verifying if DSCLI is installed on
all nodes
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : Searching for DSCLI on node "saplc1"
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : DSCLI found on node "saplc1"
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : Searching for DSCLI on node "saplc2"
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : DSCLI found on node "saplc2"
Fri Apr 13 10:12:11 EDT 2012 - INFO_MSG : Retrieving Storage ID
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Validating Library libHSS....
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Verifying if library HSS in installed
on chosen nodes
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Starting SAP liveCache Hot Standby
Setup
```

```
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Discovering DS8000 source and target
volume IDs
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Successfully discovered source and
target volume IDs
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Configuring Library libHSS
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Customizing RTEHSS_config.txt
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Successfully Configured Library
libHSS on chosen nodes
Fri Apr 13 10:12:25 EDT 2012 - INFO_MSG : Initiating Flash Copy on DS8000
Source and target Volumes
rsExecuteTask: /opt/ibm/dscli/dscli -user pw2201copy -passwd lunc0py -hmc1
9.12.6.17 mkflash -dev IBM.2107-75BALB1 -seqnum 3333 1000:1001
Date/Time: April 13, 2012 10:12:31 AM EDT IBM DSCLI Version: 6.6.20.230 DS:
IBM.2107-75BALB1
CMUC00137I mkflash: FlashCopy pair 1000:1001 successfully created.
rsExecuteTask: /opt/ibm/dscli/dscli -user pw2201copy -passwd lunc0py -hmc1
9.12.6.17 mkflash -dev IBM.2107-75BALB1 -seqnum 3333 1100:1101
Date/Time: April 13, 2012 10:12:39 AM EDT IBM DSCLI Version: 6.6.20.230 DS:
IBM.2107-75BALB1
CMUC00137I mkflash: FlashCopy pair 1100:1101 successfully created.
Fri Apr 13 10:12:42 EDT 2012 - INFO_MSG : Flash Copy consistency group
established Successfully
Fri Apr 13 10:12:51 EDT 2012 - INFO_MSG : Flash Copy in progress : "817112"
tracks to be synchronized
Fri Apr 13 10:15:00 EDT 2012 - INFO_MSG : Flash Copy in progress : "681312"
tracks to be synchronized
Fri Apr 13 10:17:09 EDT 2012 - INFO_MSG : Flash Copy in progress : "531432"
tracks to be synchronized
Fri Apr 13 10:19:17 EDT 2012 - INFO_MSG : Flash Copy in progress : "381312"
tracks to be synchronized
Fri Apr 13 10:21:25 EDT 2012 - INFO_MSG : Flash Copy in progress : "225587"
tracks to be synchronized
Fri Apr 13 10:23:34 EDT 2012 - INFO_MSG : Flash Copy in progress : "67408"
tracks to be synchronized
Fri Apr 13 10:25:42 EDT 2012 - INFO_MSG : Flash Copy finished successfully
Fri Apr 13 10:25:42 EDT 2012 - INFO_MSG : Assigning PVID to target hdisks
Fri Apr 13 10:25:42 EDT 2012 - INFO_MSG : Importing VG "vg_TL1_sdbdat" onto
other nodes
saplc2: vg_TL1_sdbdat
saplc2: 0516-306 getlvodm: Unable to find volume group vg_TL1_sdbdat in the
Device
saplc2:         Configuration Database.
saplc2: 0516-942 varyoffvg: Unable to vary off volume group vg_TL1_sdbdat.
saplc2: 0516-306 getlvodm: Unable to find volume group vg_TL1_sdbdat in the
Device
saplc2:         Configuration Database.
saplc2: 0516-772 exportvg: Unable to export volume group vg_TL1_sdbdat
Fri Apr 13 10:25:46 EDT 2012 - INFO_MSG : Assigning PVID to target hdisks
Fri Apr 13 10:25:46 EDT 2012 - INFO_MSG : Importing VG "vg_TL1_sdbdat" onto
other nodes
Fri Apr 13 10:25:47 EDT 2012 - INFO_MSG : Removing Flash Copy mappings
Date/Time: April 13, 2012 10:25:52 AM EDT IBM DSCLI Version: 6.6.20.230 DS:
IBM.2107-75BALB1
CMUC00140I rmflash: FlashCopy pair 1000:1001 successfully removed.
CMUC00140I rmflash: FlashCopy pair 1100:1101 successfully removed.
```

```
Fri Apr 13 10:25:55 EDT 2012 - INFO_MSG : Setting up Hot Standby parameters
Fri Apr 13 10:25:55 EDT 2012 - INFO_MSG : Starting X-server on chosen nodes
saplc2:     12916  XSERVER  Found other running x_server with version 'U64/AIX
7.7.07   Build 039-123-243-619'
saplc2:     12902  XSERVER  started, 'already...'
OK
OK
OK
OK
OK
OK
HotStandbyStorageDLLPath        libHSSibm2107.so
OfficialNodeName        SAPLCSVC1
CURRENT_NODE    SAPLC1
HotStandbyNodeName001   SAPLC1                                          0
HotStandbyNodeName002   SAPLC2                                          0
Fri Apr 13 10:26:18 EDT 2012 - INFO_MSG : Preparing Standby Instance
OK
OK
OK
OK
Fri Apr 13 10:27:11 EDT 2012 - INFO_MSG : Flash Copy in progress : "816536"
tracks to be synchronized
Fri Apr 13 10:28:15 EDT 2012 - INFO_MSG : Flash Copy in progress : "745903"
tracks to be synchronized
Fri Apr 13 10:29:20 EDT 2012 - INFO_MSG : Flash Copy in progress : "681408"
tracks to be synchronized
Fri Apr 13 10:30:24 EDT 2012 - INFO_MSG : Flash Copy in progress : "608651"
tracks to be synchronized
Fri Apr 13 10:31:29 EDT 2012 - INFO_MSG : Flash Copy in progress : "532656"
tracks to be synchronized
Fri Apr 13 10:32:33 EDT 2012 - INFO_MSG : Flash Copy in progress : "457448"
tracks to be synchronized
Fri Apr 13 10:33:38 EDT 2012 - INFO_MSG : Flash Copy in progress : "379592"
tracks to be synchronized
Fri Apr 13 10:34:42 EDT 2012 - INFO_MSG : Flash Copy in progress : "306652"
tracks to be synchronized
Fri Apr 13 10:35:46 EDT 2012 - INFO_MSG : Flash Copy in progress : "228076"
tracks to be synchronized
Fri Apr 13 10:36:51 EDT 2012 - INFO_MSG : Flash Copy in progress : "148776"
tracks to be synchronized
Fri Apr 13 10:37:55 EDT 2012 - INFO_MSG : Flash Copy in progress : "69395"
tracks to be synchronized
Fri Apr 13 10:38:59 EDT 2012 - INFO_MSG : Flash Copy finished successfully
OK
OK
Fri Apr 13 10:39:13 EDT 2012 - INFO_MSG : Finished Setting up Standby Instance
Fri Apr 13 10:39:13 EDT 2012 - INFO_MSG : Configuring lccluster for integration
with APO
Fri Apr 13 10:39:13 EDT 2012 - INFO_MSG : Finished Configuring lccluster
Fri Apr 13 10:39:13 EDT 2012 - INFO_MSG : Successfully Configured SAP liveCache
Hot Standby Instance "TL1"
```

5. Problem determination.

   You might see an error message such as the following example:

```
...
CMUN03042E mkflash: 1100:1101: Copy Services operation failure: already a
FlashCopy target
...
```

If you see this error, you must remove the FlashCopy pair manually:

```
root@saplc2 / # /opt/ibm/dscli/dscli -user pw2201copy -passwd lunc0py -hmc1
9.12.6.17 rmflash 1100:1101
```

### 7.7.7 Configuring the cluster by using the Smart Assist

After you complete the prerequisites, you can start Smart Assist for MaxDB as explained in the following steps:

1. Launch the wizard by using the path for `saplc1`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Make Applications Highly Available (Use Smart Assists)** → **Add an Application to the PowerHA SystemMirror Configuration**.
2. In the Select an Application from the List of Discovered Applications Below panel (Figure 7-97), select **SAP MaxDB Smart Assist**.

```
Select an Application from the List of Discovered Applications Below

 Move cursor to desired item and press Enter.

   SAP MaxDB Smart Assist           # saplc1 saplc2
   SAP Smart Assist                 # saplc1 saplc2
```

*Figure 7-97 Selecting MaxDB Smart Assist*

3. In the Select Configuration Mode panel (Figure 7-98), select **Automatic Discovery and Configuration**.

```
                          Select Configuration Mode

Move cursor to desired item and press Enter.

  Automatic Discovery And Configuration
  Manual Configuration
```

*Figure 7-98 Selecting the configuration mode*

4. In the Select the Specific Configuration You Wish to Create panel (Figure 7-99), select **SAP liveCache Hot Standby Database Instance(s) # saplc1**.

```
             Select The Specific Configuration You Wish to Create

Move cursor to desired item and press Enter.

  SAP liveCache Hot Standby Database Instance(s) # saplc1
  SAP MaxDB Database Instance(s)   # saplc1 saplc2
```

*Figure 7-99 Selecting the configuration to create*

5. In the Select The Specific Configuration You Wish to Create instance panel (Figure 7-100), select **TL1**.

```
              Select The Specific Configuration You Wish to Create

 Move cursor to desired item and press Enter.

   TL1
```

*Figure 7-100 Selecting the specific configuration to create*

6. Use the available pick lists (F4) or insert the necessary information as shown in Figure 7-101:
   - For SAP liveCache Hot Standby Instance DBM User XUSER, enter [TL1_XUSER].
   - For liveCache Global Filesystem Mount point, enter [/archive/sapdb/TL1_LOCK].
   - For Takeover Node(s), select **saplc2**.
   - For Data Volume Group(s), enter [vg_TL1_sdbdat].
   - For Log Volume Group(s), enter [vg_TL1_sdblog].

   Then, press Enter.

```
                 Add a SAP MaxDB Hot Standby Database Instance

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

                                                        [Entry Fields]
* SAP liveCache Hot Standby Instance Name             TL1
* SAP liveCache Hot Standby Instance DBM User XUSER  [TL1_XUSER]
* liveCache Global Filesystem Mount point        [/archive/sapdb/TL1_LOCK]
* Primary Node                                        saplc1 +
* Takeover Node(s)                                    saplc2 +
* Service IP Label                                    saplcsvc1 +
  Netmask(IPv4)/Prefix Length(IPv6)                  []
* Data Volume Group(s)                               [vg_TL1_sdbdat] +
* Log Volume Group(s)                                [vg_TL1_sdblog] +
```

*Figure 7-101 Additional entries for the Smart Assist*

Alternatively, you can use the command line as shown in Example 7-98.

*Example 7-98 liveCache Smart Assist command line*

```
/usr/es/sbin/cluster/sa/maxdb/sbin/cl_maxdb_addInstance \
   -H \
   -i'TL1' \
   -U'TL1_XUSER' \
   -F'/archive/sapdb/TL1_LOCK' \
   -o'saplc1' \
   -t'saplc2' \
   -s'saplcsvc1' \
   -d'vg_TL1_sdbdat' \
   -l'vg_TL1_sdblog' \
```

Example 7-99 on page 460 shows the command output from the smitty panel.

*Example 7-99 Smitty command output liveCache Smart Assist*

```
-----------------------------------------------------------------------
Configuring service IP label: saplcsvc1
Service IP label configuration complete.
-----------------------------------------------------------------------
NAME="TL1_instFiles"
DESCRIPTION="Instance Related files for MaxDB Instance name TL1"
SYNC_WITH_CLUSTER="true"
SYNC_WHEN_CHANGED="true"
Removing files from PowerHA SystemMirror File Collection TL1_instFiles...
PowerHA SystemMirror File Collection TL1_instFiles has been removed.
NAME="TL1_instFiles"
DESCRIPTION=""
SYNC_WITH_CLUSTER="true"
SYNC_WHEN_CHANGED="false"
FILES="/usr/es/sbin/cluster/etc/config/verify/liveCache_Master_TL1_MaxDB_7.6_Ma
xDB_Hot_Standby.ver"
NAME="TL1_instFiles"
DESCRIPTION=""
SYNC_WITH_CLUSTER="true"
SYNC_WHEN_CHANGED="false"
FILES="/usr/es/sbin/cluster/etc/config/verify/liveCache_Master_TL1_MaxDB_7.6_Ma
xDB_Hot_Standby.ver"
Removing files from PowerHA SystemMirror File Collection TL1_instFiles...
File
/usr/es/sbin/cluster/etc/config/verify/liveCache_Master_TL1_MaxDB_7.6_MaxDB_Hot
_Standby.ver has been removed from PowerHA SystemMirror File Collection.
PowerHA SystemMirror File Collection TL1_instFiles has been removed.
```

## 7.7.8 Verifying the Smart Assist settings

We verify the Smart Assist settings. Follow these steps:

1. Verify whether the service address is configured on the correct network.
2. In some older versions, the Smart Assist configures a start after the resource group dependencies. Remove these dependencies because these dependencies disappeared in the actual versions.

   Launch PowerHA by using the path for `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Start After Resource Group Dependency** → **Remove Start After Resource Group Dependency**.

   Select one of the dependencies. It does not matter which dependency because the second dependency must also be removed as shown in Figure 7-102 on page 461.

```
          Select a Start After Resource Group Dependency to Remove

Move cursor to desired item and press Enter.

# Source                           Target
  RG_Master_TL1                     RG_Log_TL1
  RG_Standby_TL1                    RG_Master_TL1
```

*Figure 7-102 Select a Start After Resource Group Dependency to Remove*

Or, you can use the command line as shown in Example 7-100.

*Example 7-100 Command to remove a start after resource group dependency*

```
/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'START_AFTER' \
   -d \
   -c'RG_Standby_TL1' \
   -p'RG_Master_TL1'

/usr/es/sbin/cluster/utilities/clrgdependency \
   -t'START_AFTER' \
   -d \
   -c'RG_Master_TL1' \
   -p'RG_Log_TL1'
```

3. Verify or configure the resource group processing order (Figure 7-103).

   Launch PowerHA by using the path for sapdbhr3: **smitty sysmirror** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Resource Group Processing Ordering**.

```
                 Change/Show Resource Group Processing Order

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

                                                         [Entry Fields]
Resource Groups Acquired in Parallel
  Serial Acquisition Order          RG_Log_TL1 RG_Master_TL1 RG_Standby_TL1
  New Serial Acquisition Order  [RG_Log_TL1 RG_Master_TL1 RG_Standby_TL1 ] +

  Resource Groups Released in Parallel
  Serial Release Order              RG_Standby_TL1 RG_Master_TL1 RG_Log_TL1
  New Serial Release Order     [ RG_Standby_TL1 RG_Master_TL1 RG_Log_TL1 ] +
```

*Figure 7-103 Change/Show Resource Group Processing Order*

Or, you can use the command line as shown in Example 7-101 on page 462.

*Example 7-101 Command to change the resource group processing order*

```
/usr/es/sbin/cluster/utilities/clrgorder \
   -c \
   -a 'RG_Log_TL1 RG_Master_TL1 RG_Standby_TL1' \
   -r 'RG_Standby_TL1 RG_Master_TL1 RG_Log_TL1'
```

### 7.7.9 Completing the configuration

> **Tip:** It is not necessary to stop the MaxDB liveCache environment. You can synchronize with a running database.

After the Smart Assist for MaxDB is configured, complete the configuration:

1. Synchronize the PowerHA cluster by using SMIT:
   a. Follow the path: **`smitty sysmirror`** → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.
   b. In the PowerHA SystemMirror Verification and Synchronization panel (Figure 7-104), press Enter to accept the default options.

```
   PowerHA SystemMirror Verification and Synchronization (Active Cluster
                            Nodes Exist)

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

                                                        [Entry Fields]
* Verify changes only?                              [No] +
* Logging                                           [Standard] +
```

*Figure 7-104 Accepting the default actions on the Verification and Synchronization panel*

> **Tip:** The log file for the liveCache Smart Assist for SAP is in the `/var/hacmp/log/maxdbsa.log` file. You can use the **`clmgr`** utility to easily view the log as shown in the following example:
>
> `clmgr view log maxdbsa.log`

2. After the synchronization finishes, the resource groups are online. The application monitors start after the defined stabilization interval as shown in Example 7-102.

*Example 7-102 Cluster state and active application monitor*

```
root@saplc1 / # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                          Node
-----------------------------------------------------------------------------
RG_Log_TL1     ONLINE                         saplc1
               ONLINE                         saplc2

RG_Master_TL1  OFFLINE                        saplc1
               ONLINE                         saplc2
```

```
RG_Standby_TL1 OFFLINE                     saplc2
               ONLINE                      saplc1

root@saplc1 / # ps -ef | grep /usr/es/sbin/cluster/clappmond | grep -v grep
    root 18612394 16646194   0 14:42:28      -  0:00
/usr/es/sbin/cluster/clappmond mon_master_TL1

root@saplc2 / # ps -ef | grep /usr/es/sbin/cluster/clappmond | grep -v grep
    root 15138950 11337888   0 14:41:39      -  0:00
/usr/es/sbin/cluster/clappmond mon_standby_TL1
```

Your MaxDB instance and database are now configured for high availability in a hot-standby PowerHA SystemMirror configuration.

# 7.8 DB2 HADR cluster solution

This solution does not use the IBM PowerHA SystemMirror Smart Assist. Therefore, we need to use the script set from the ISICC team. This solution is an alternative solution for cluster 2 in the SAP liveCache scenario that is shown in 7.6, “Cluster 2: Database instance” on page 431.

As described in the referenced “Invincible SCM” paper, the value of “back to production in 5 minutes” can be created on the base of End-to-End top to bottom HotStandby or active/passive solutions only.

Therefore, a solution for DB2 HADR is added, as well, to complete the stack.

## 7.8.1 Overview

We provide an overview of the solution (Figure 7-105 on page 464).

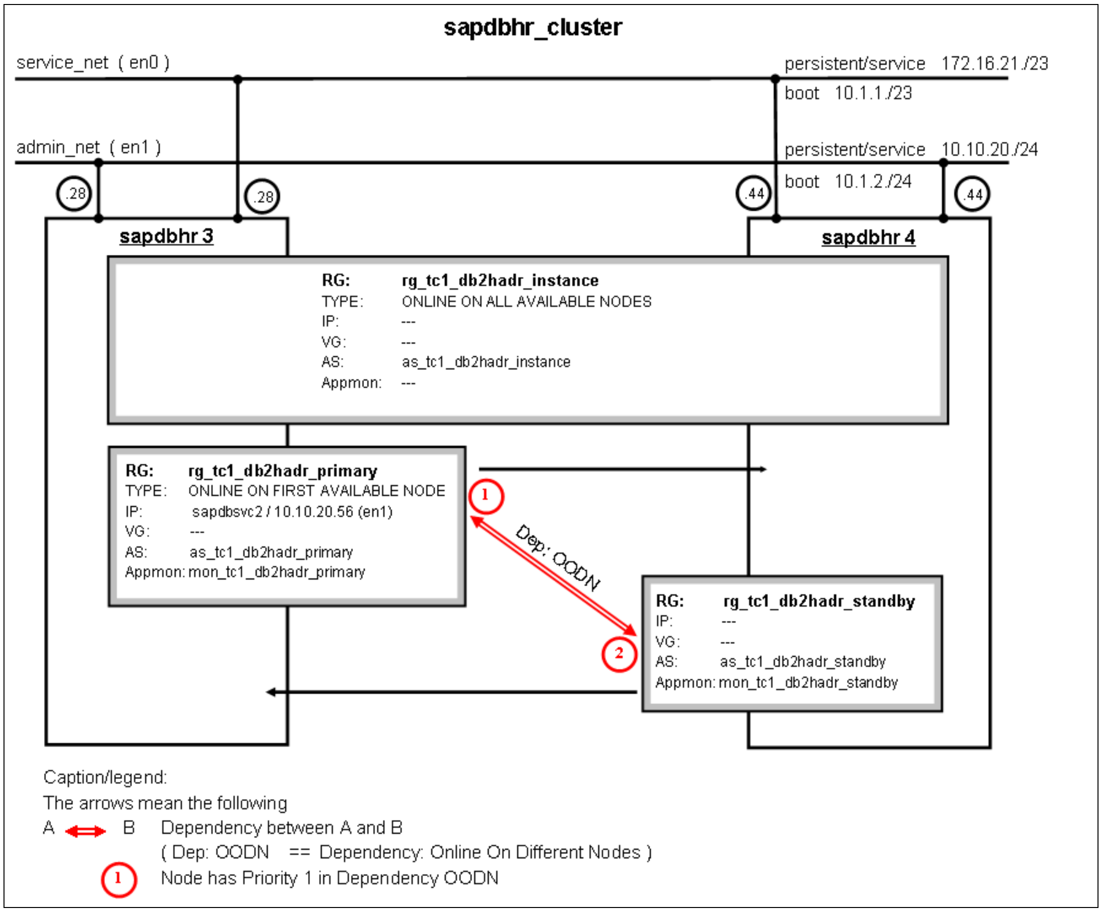



*Figure 7-105 Schema overview for sapdbhr_cluster*

Figure 7-105 shows the DB2 HADR solution for this cluster `sapdbhr`.

## Cluster design

This DB2 HADR solution uses three resource groups:

- `rg_tc1_db2hadr_instance`

  This resource group starts the local DB2 instance. It is started on all available nodes. No IP address and no volume group are configured in this resource group. The necessary volume group(s) and file systems are activated automatically after the boot. It is monitored by the **`db2fm`** tool.

- `rg_tc1_db2hadr_primary`

  This resource group starts the database as the primary HADR instance. It is started on the first available node and has a service address for the connection of the external SAP always with the same IP address. It is monitored from the PowerHA cluster.

- `rg_tc1_db2hadr_standby`

  This resource group starts the database as the standby HADR instance. No additional resources are associated with this resource group. It is monitored from the PowerHA cluster. It is configured “online on a different node” as second priority to the resource group `rg_tc1_db2hadr_primary`.

All the scripts that are used are listed in “Cluster scripts for DB2 HADR” on page 645.

### Startup

During startup, the cluster needs to be handled differently. If you bring up the pair, you have to ensure that the standby is started first and is in the peer state one time.

### Maintenance and application awareness

To not interfere with database administrative tasks, the logic is aware of the status (start/stop) to which the application is set by the database or SAP administrator for the following cases:

1. The administrator must use only one of the following commands:
   - `sapdbctrl start`
   - `sapdbctrl stop`
   - `startdb`
   - `stopdb`
   - `db2start`
   - `db2stop`
   - `db2gcf`
   - `db2fm`
   - `db2 activate`
   - `db2 deactivate`
2. The tool **`db2fm`** is enabled.
3. The flag “APPLICATION_INTEGRATION” in `sapha_SID.cfg` is set to “1”.

> **Important:** Never use the **`hadr`** commands in this PowerHA clustered environment to manually move, start, or stop. You must use the cluster tools to manually move, start, or stop.

## 7.8.2 Installation and configuration steps before you set up PowerHA

We describe the installation and configuration steps before we set up PowerHA. Follow these steps:

1. Request the latest set of scripts as described in “Cluster scripts for DB2 HADR” on page 645. The only change is in the file `sapha_<SID>.cfg`. Follow these steps:
   a. Rename the file `sapha_<SID>.cfg` to include the SID that is used for the DB2 installation, in this case, `TC1: sapha_TC1.cfg`.
   b. Open the file and adapt the names to your naming conventions. The details are in “sapha_TL1_cfg” on page 669.
2. Create the users.

   See 7.2.2, “Creating the users” on page 362.
3. Create volume groups, logical volumes, and file systems (Example 7-103).

   You must create the volume groups, logical volumes, and the file systems before you start the sapinst.

   You can use the same command that is in 7.2.3, “Creating volume groups, logical volumes, and file systems” on page 363, but you have to create the volume group on both nodes.

   *Example 7-103 Volume group at sapdbhr_cluster*

```
root@sapdbhr3 / # lsvg -l vg_TC1_sap
vg_TC1_sap:
LV NAME             TYPE       LPs PPs PVs  LV STATE      MOUNT POINT
```

```
TC1.lvdb2binary     jfs2       64 64 1    open/syncd    /db2/TC1
TC1.lvdb2logdir     jfs2       64 64 1    open/syncd    /db2/TC1/log_dir
TC1.lvdb2logarc     jfs2       64 64 1    open/syncd    /db2/TC1/log_archive
TC1.lvdb2dat.01     jfs2     140 140 1    open/syncd    /db2/TC1/sapdata1
TC1.lvdb2dat.02     jfs2     140 140 1    open/syncd    /db2/TC1/sapdata2
TC1.lvdb2dat.03     jfs2     140 140 1    open/syncd    /db2/TC1/sapdata3
TC1.lvdb2dat.04     jfs2     140 140 1    open/syncd    /db2/TC1/sapdata4
```

4. Configure the network (Example 7-104).

   This configuration is an alternate installation of cluster 2, but we use a separate IP configuration. The IP addresses for the DB2 server (`sapdbhrsvc2`) are configured on the first node `sapdbhr3`, where the DB2 primary is installed. But, we need this IP address for the communication between primary application server instance and the database only, not for the database.

*Example 7-104 Network settings for sapdbhr3*

```
root@sapdbhr3 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.da.d9.10.6f 19831419     0  3412628     0     0
en0   1500  10.1        sapdbhr3b1        19831419     0  3412628     0     0
en0   1500  172.16.20   sapdbhr3p1        19831419     0  3412628     0     0
en1   1500  link#3      6e.8d.da.d9.10.70  1639564     0     5226     0     0
en1   1500  10.1.2      sapdbhr3b2         1639564     0     5226     0     0
en1   1500  10.10.20    sapdbhr3p2         1639564     0     5226     0     0
en1   1500  10.10.20    sapdbsvc2          1639564     0     5226     0     0
lo0   16896 link#1                          563851     0   563851     0     0
lo0   16896 127         loopback            563851     0   563851     0     0
lo0   16896 loopback                        563851     0   563851     0     0
```

   No additional IP address is configured on the second node `sapdbhr4`.

## 7.8.3 Installing the DB2 HADR database instance

There are two ways to install DB2 HADR:

- Use the SAPinst for the installation of the DB2 HADR database on the first node, as well as on the second node.
- Use the installation approach that is shown in “sapdbhr_cluster: Installing DB2 high availability disaster recovery” on page 640.

> **Node names must be the same:** For this solution with the ISICC script set, the instances must have the same name on both nodes.

## 7.8.4 Configuring the DB2 fault monitor

The db2fm monitor needs to be started for our SAP DB2 instance on both nodes. The instance must not be started at boot time. The cluster starts the instance. Use these commands to start the db2fm monitor:

```
su - db2tc1
db2fm -i db2tc1 -f yes
```

You can check whether everything is correctly configured in the config file as shown in Example 7-105.

> **Path:** The path `/db2/TC1/db2tc1/sqllib` depends on the db2 installation.

*Example 7-105 cat /db2/TC1/db2tc1/sqllib/fm.sapdbhr3.reg*

```
sapdbhr3:db2tc1 14> cat /db2/TC1/db2tc1/sqllib/fm.sapdbhr3.reg
FM_ON = yes # updated by db2fm
FM_ACTIVE = yes # default
START_TIMEOUT = 600 # default
STOP_TIMEOUT = 600 # default
STATUS_TIMEOUT = 20 # default
STATUS_INTERVAL = 20 # default
RESTART_RETRIES = 3 # default
ACTION_RETRIES = 3 # default
NOTIFY_ADDRESS = db2tc1@sapdbhr3 # default
```

Also, the db2fmc has to be configured to be started at every system boot as shown in Example 7-106.

*Example 7-106 Configure and start db2fmc*

```
root@sapdbhr3 / # /db2/TC1/db2tc1/sqllib/bin/db2iauto -on db2tc1

root@sapdbhr3 / # /db2/TC1/db2tc1/sqllib/bin/db2fmcu -u -p
/db2/TC1/db2tc1/sqllib/bin/db2fmcd

root@sapdbhr3 / # /db2/TC1/db2tc1/sqllib/bin/db2greg -v -updinstrec
instancename=db2tc1!startatboot=1
Update record search criteria:
   Service      = |N/A|
   Version      = |N/A|
   InstanceName = |db2tc1|
   InstancePath = |N/A|
   Usage        = |N/A|
   StartAtBoot  = N/A
   Maintenance  = N/A
   InstallPath  = |N/A|
   RemoteProf   = |N/A|
   Comment      = |N/A|
Record to be updated to:
   Service      = |DB2|
   Version      = |9.7.0.4|
   InstanceName = |db2tc1|
   InstancePath = |/db2/TC1/db2tc1/sqllib|
   Usage        = |N/A|
   StartAtBoot  = 1
   Maintenance  = 0
   InstallPath  = |/db2/TC1/db2tc1/db2_software|
   RemoteProf   = |N/A|
   Comment      = |N/A|
Review the db2greg.log file for more information.
```

Now, reboot the node and validate the correct functionality as shown in Example 7-107.

*Example 7-107 Verify db2fmc*

```
root@sapdbhr3 /db2 # ps -ef|grep fm
  db2tc1  8454248 10223760   0 15:00:27      -  0:00 db2fmp (C) 0
  db2tc1  9568324        1   0 05:06:07      -  0:00
/db2/TC1/db2tc1/db2_software/bin/db2fmd -i db2tc1 -m
/db2/TC1/db2tc1/db2_software/lib64/libdb2gcf.a
    root 15204466        1   0 05:06:07      -  0:00
/db2/TC1/db2tc1/sqllib/bin/db2fmcd
```

You have a new entry in `/etc/inittab` that the DB2 fault monitor coordinator is started after every system boot as shown in Example 7-108.

*Example 7-108 lsitab fmc*

```
root@sapdbhr3 /db2 # lsitab fmc
fmc:2:respawn:/db2/TC1/db2tc1/sqllib/bin/db2fmcd #DB2 Fault Monitor Coordinator
```

Validate the function of the DB2 fault monitor as shown in Example 7-109.

*Example 7-109 Validate DB2 fault monitor*

```
root@sapdbhr3 / # su - db2tc1
sapdbhr3:db2tc1 1> db2fm -i db2tc1 -s
Gcf module '/db2/TC1/db2tc1/db2_software/lib64/libdb2gcf.a' state is AVAILABLE
sapdbhr3:db2tc1 3> db2fm -i db2tc1 -S
Gcf module 'fault monitor' state is AVAILABLE
```

**Message:** If the following message appears, it is also correct because the '`cl_db2_start_local`' script fixes this bug:

```
"Gcf module ... is INSTALLED PROPERLY, BUT NOT ALIVE"
```

You must perform the same steps on the second node `sapdbhr4`.

### 7.8.5 Configuring the base IBM PowerHA SystemMirror

At first, we configure the topology of the PowerHA cluster. In Example 7-110, the cluster `sapdbhr_cluster` is configured with two Ethernet interfaces in each node.

*Example 7-110 Cluster network configuration*

```
root@sapdbhr3 /db2/TC1 # lscluster -c
Cluster query for cluster sapdbhr_cluster returns:
Cluster uuid: 49e85828-720c-11e1-98f2-6e8ddad91070
Number of nodes in cluster = 2
        Cluster id for node sapdbhr3b1 is 1
        Primary IP address for node sapdbhr3b1 is 10.1.1.28
        Cluster id for node sapdbhr4b1 is 2
        Primary IP address for node sapdbhr4b1 is 10.1.1.44
Number of disks in cluster = 0
Multicast address for cluster is 228.1.1.26
root@sapdbhr3 /db2/TC1 # cllsif
Adapter Type Network Net Type Attribute Node IP Address Hardware Address Interface
Name   Global Name      Netmask          Alias for HB Prefix Length
```

```
sapdbhr3b2 boot admin_net  ether public sapdbhr3   10.1.2.28 en1 255.255.255.0 24
sapdbhr3b1 boot service_net ether public sapdbhr3   10.1.1.28 en0 255.255.254.0 23
sapdbhr4b2 boot admin_net  ether public sapdbhr4   10.1.2.44 en1 255.255.255.0 24
sapdbhr4b1 boot service_net ether public sapdbhr4   10.1.1.44 en0 255.255.254.0 23
```

## 7.8.6 Cluster configuration

We describe the cluster configuration.

### Configuring the application controller scripts

We describe how to configure the application controller scripts:

1. Launch PowerHA by using the path for sapdbhr3: **smitty sysmirror** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Controller Scripts** → **Add Application Controller Scripts**.
2. Configure the application controller scripts for the resource group rg_tc1_db2hadr_instance as shown in Figure 7-106.

```
                       Add Application Controller Scripts

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

                                                        [Entry Fields]
* Application Controller Name                   [as_tc1_db2hadr_instance]
* Start Script       [/usr/es/sbin/cluster/sap/cl_db2_start_local DB2 TC1]
* Stop Script         [/usr/es/sbin/cluster/sap/cl_db2_stop_local DB2 TC1]
  Application Monitor Name(s)                                            +
  Application startup mode                          [foreground]       +
```

*Figure 7-106 Add application controller scripts for rg_tc1_db2hadr_instance*

Example 7-111 shows the command-line option.

*Example 7-111 Command to add application controller scripts for rg_tc1_db2hadr_instance*

```
/usr/es/sbin/cluster/utilities/claddserv \
   -s'as_tc1_db2hadr_instance' \
   -b'/usr/es/sbin/cluster/sap/cl_db2_start_local DB2 TC1' \
   -e'/usr/es/sbin/cluster/sap/cl_db2_stop_local DB2 TC1'
   -O ‘foreground’
```

3. Configure the application controller scripts for the resource group rg_tc1_db2hadr_primary as shown in Figure 7-107 on page 470.

```
                        Add Application Controller Scripts

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

                                                        [Entry Fields]
* Application Controller Name                      [as_tc1_db2hadr_primary]
* Start Script     [/usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1]
* Stop Script       [/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1]
  Application Monitor Name(s)                                             +
  Application startup mode                            [foreground]        +
```

*Figure 7-107 Add application controller scripts for rg_tc1_db2hadr_primary*

Example 7-112 shows the command-line option.

*Example 7-112 Command to add application controller scripts for rg_tc1_db2hadr_primary*

```
/usr/es/sbin/cluster/utilities/claddserv \
   -s'as_tc1_db2hadr_primary' \
   -b'/usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1' \
   -e'/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1'
   -O ‘foreground’
```

4. Configure the application controller scripts for the resource group `rg_tc1_db2hadr_standby` as shown in Figure 7-108.

```
                        Add Application Controller Scripts

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

                                                        [Entry Fields]
* Application Controller Name                      [as_tc1_db2hadr_standby]
* Start Script     [/usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1]
* Stop Script       [/usr/es/sbin/cluster/sap/cl_db2_stop_hadr STANDBY TC1]
  Application Monitor Name(s)                                             +
  Application startup mode                            [foreground]        +
```

*Figure 7-108 Add application controller scripts for rg_tc1_db2hadr_standby*

Example 7-113 shows the command-line option.

*Example 7-113 Command to add application controller scripts for rg_tc1_db2hadr_standby*

```
/usr/es/sbin/cluster/utilities/claddserv \
   -s'as_tc1_db2hadr_standby' \
   -b'/usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1' \
   -e'/usr/es/sbin/cluster/sap/cl_db2_stop_hadr STANDBY TC1'
   -O ‘foreground’
```

5. Check the settings as shown in Example 7-114.

*Example 7-114 Check the settings with cllsserv*

```
root@sapdbhr3 /usr/es/sbin/cluster # cllsserv
```

```
as_tc1_db2hadr_instance  /usr/es/sbin/cluster/sap/cl_db2_start_local DB2 TC1
/usr/es/sbin/cluster/sap/cl_db2_stop_local DB2 TC1 foreground
as_tc1_db2hadr_primary  /usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1
/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1 foreground
as_tc1_db2hadr_standby  /usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1
/usr/es/sbin/cluster/sap/cl_db2_stop_hadr STANDBY TC1 foreground
```

### Configuring application monitors

We describe how to configure the application monitors:

1. Launch PowerHA by using the path for `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resources** → **Configure User Applications (Scripts and Monitors)** → **Application Monitors** → **Configure Custom Application Monitors** → **Add a Custom Application Monitor**.
2. Configure the application monitors for the resource group `rg_tc1_db2hadr_primary` as shown in Figure 7-109.

```
 Add a Custom Application Monitor

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

                                                        [Entry Fields]
* Monitor Name                                    [mon_tc1_db2hadr_primary]
* Application Controller(s) to Monitor              as_tc1_db2hadr_primary +
* Monitor Mode                                    [Long-running monitoring] +
* Monitor Method [/usr/es/sbin/cluster/sap/cl_db2_monitor_hadr PRIMARY TC1]
  Monitor Interval                                     [30] #
  Hung Monitor Signal                                  [9] #
* Stabilization Interval                               [60] #
* Restart Count                                        [0] #
  Restart Interval                                     [0] #
* Action on Application Failure                        [fallover] +
  Notify Method                                        []
  Cleanup Method  [/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1] /
  Restart Method  [/usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1] /
```

*Figure 7-109 Add application monitors for rg_tc1_db2hadr_primary*

Or, you can use the command line as shown in Example 7-115.

*Example 7-115 Command to add application monitors for rg_tc1_db2hadr_primary*

```
/usr/es/sbin/cluster/utilities/claddappmon \
   MONITOR_TYPE=user \
   name='mon_tc1_db2hadr_primary' \
   RESOURCE_TO_MONITOR='as_tc1_db2hadr_primary' \
   INVOCATION='longrunning' \
   MONITOR_METHOD='/usr/es/sbin/cluster/sap/cl_db2_monitor_hadr PRIMARY TC1' \
   MONITOR_INTERVAL='30' \
   HUNG_MONITOR_SIGNAL='9' \
   STABILIZATION_INTERVAL='60' \
   RESTART_COUNT='0' \
   RESTART_INTERVAL='0' \
   FAILURE_ACTION='fallover' \
```

```
  CLEANUP_METHOD='/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1' \
  RESTART_METHOD='/usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1'
```

**Important:** The provided intervals must be adjusted to fit the individual installation.

**Important:** The command line is the only way to configure a cleanup and restart method that uses parameters. The smitty menu did not accept it.

3. Configure the application monitors for the resource `group rg_tc1_db2hadr_standby` as shown in Figure 7-110.

```
 Add a Custom Application Monitor

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

                                                          [Entry Fields]
* Monitor Name                                    [mon_tc1_db2hadr_standby]
* Application Controller(s) to Monitor              as_tc1_db2hadr_standby +
* Monitor Mode                                    [Long-running monitoring] +
* Monitor Method [/usr/es/sbin/cluster/sap/cl_db2_monitor_hadr STANDBY TC1]
  Monitor Interval                                     [30] #
  Hung Monitor Signal                                  [9] #
* Stabilization Interval                               [60] #
* Restart Count                                        [3] #
  Restart Interval                                     [264] #
* Action on Application Failure                        [fallover] +
  Notify Method                                        []
  Cleanup Method                                       [] /
  Restart Method  [/usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1] /
```

*Figure 7-110 Add application monitors for rg_tc1_db2hadr_standby*

Or, you can use the command line as shown in Example 7-116.

*Example 7-116 Command to add application monitors for rg_tc1_db2hadr_standby*

```
/usr/es/sbin/cluster/utilities/claddappmon \
  MONITOR_TYPE=user \
  name='mon_tc1_db2hadr_standby' \
  RESOURCE_TO_MONITOR='as_tc1_db2hadr_standby' \
  INVOCATION='longrunning' \
  MONITOR_METHOD='/usr/es/sbin/cluster/sap/cl_db2_monitor_hadr STANDBY TC1' \
  MONITOR_INTERVAL='30' \
  HUNG_MONITOR_SIGNAL='9' \
  STABILIZATION_INTERVAL='60' \
  RESTART_COUNT='3' \
  RESTART_INTERVAL='264' \
  FAILURE_ACTION='fallover' \
  RESTART_METHOD='/usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1'
```

**Important:** The command line is the only way to configure a cleanup and restart method that uses parameters. The smitty menu did not accept it.

4. Check the settings as shown in Example 7-117.

*Example 7-117 Check the settings with cllsappmon*

```
root@sapdbhr3 /usr/es/sbin/cluster # cllsappmon
mon_tc1_db2hadr_primary user
mon_tc1_db2hadr_standby user

root@sapdbhr3 /usr/es/sbin/cluster # cllsappmon mon_tc1_db2hadr_primary
mon_tl1_db2hadr_primary user    /usr/es/sbin/cluster/sap/cl_db2_monitor_hadr
PRIMARY TC1       10     longrunning    9      60     fallover       0
0       /usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1
/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1
as_tc1_db2hadr_primary

root@sapdbhr3 /usr/es/sbin/cluster # cllsappmon mon_tc1_db2hadr_standby
mon_tl1_db2hadr_standby user    /usr/es/sbin/cluster/sap/cl_db2_monitor_hadr
STANDBY TC1       10     longrunning    9      60     fallover       3
264   /usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1
as_tc1_db2hadr_standby
```

## Configuring the service IP labels/address

We guide you to configure the service IP labels/address:

1. Launch PowerHA by using the path for sapdbhr3: **smitty sysmirror** → **Cluster Applications and Resources** → **Resources** → **Configure Service IP Labels/Addresses** → **Add a Service IP Label/Address**.
2. Select the network name of the network for which you want to configure the service label as shown in Figure 7-111.

```
                              Network Name

Move cursor to desired item and press Enter.

   admin_net (10.10.20.0/24 10.1.2.0/24)
   service_net (10.1.0.0/23)
```

*Figure 7-111 Network name for service IP label*

3. Configure a service IP label for the resource group `rg_tc1_db2hadr_primary` as shown in Figure 7-112 on page 474.

```
                              Add a Service IP Label/Address

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

                                                        [Entry Fields]
IP Label/Address                                    sapdbsvc2
Netmask(IPv4)/Prefix Length(IPv6)                  []
* Network Name                                        admin_net
```

*Figure 7-112 Add a service ip label/address for resource group rg_tc1_db2hadr_primary*

Or, you can use the command line as shown in Example 7-118.

*Example 7-118 Command to add a service IP label/address*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/clmgr add service_ip
'sapdbsvc2' NETWORK='admin_net'
```

## Adding resource groups

We show how to add resource groups:

1. Launch PowerHA by using the path for `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Add a Resource Group**.
2. Add the resource group `rg_tc1_db2hadr_instance` as shown in Figure 7-113.

```
                                 Add a Resouce Group

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

                                                        [Entry Fields]
* Resource Group Name                              [rg_tc1_db2hadr_instance]
* Participating Nodes (Default Node Priority)         [sapdbhr3 sapdbhr4] +

  Startup Policy                                Online On All Available Nodes +
  Fallover Policy                           Bring Offline (On Error Node Only) +
  Fallback Policy                                       Never Fallback +
```

*Figure 7-113 Add a resource group rg_tc1_db2hadr_instance*

Or, you can use the command line as shown in Example 7-119.

*Example 7-119 Command to add a resource group rg_tc1_db2hadr_instance*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddgrp  -g
'rg_tc1_db2hadr_instance' -n 'sapdbhr3 sapdbhr4' -S 'OAAN' -O 'BO' -B 'NFB'
```

3. Add the resource group `rg_tc1_db2hadr_primary` as shown in Figure 7-114 on page 475.

```
                              Add a Resouce Group

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

                                                        [Entry Fields]
* Resource Group Name                              [rg_tc1_db2hadr_primary]
* Participating Nodes (Default Node Priority)         [sapdbhr3 sapdbhr4] +

  Startup Policy                            Online On First Available Node +
  Fallover Policy                     Fallover To Next Priority Node In The List +
  Fallback Policy                                        Never Fallback +
```

*Figure 7-114 Add a resource group rg_tc1_db2hadr_primary*

Or, you can use the command line as shown in Example 7-120.

*Example 7-120 Command to add a resource group rg_tc1_db2hadr_primary*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddgrp  -g
'rg_tc1_db2hadr_primary' -n 'sapdbhr3 sapdbhr4' -S 'OFAN' -O 'FNPN' -B 'NFB'
```

4. Add the resource group `rg_tc1_db2hadr_standby` as shown in Figure 7-115.

```
                              Add a Resouce Group

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

                                                        [Entry Fields]
* Resource Group Name                              [rg_tc1_db2hadr_standby]
* Participating Nodes (Default Node Priority)         [sapdbhr4 sapdbhr3] +

  Startup Policy                            Online On First Available Node +
  Fallover Policy                     Fallover To Next Priority Node In The List +
  Fallback Policy                                        Never Fallback +
```

*Figure 7-115 Add a resource group rg_tc1_db2hadr_standby*

Or, you can use the command line as shown in Example 7-121.

*Example 7-121 Command to add a resource group rg_tc1_db2hadr_standby*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddgrp  -g
'rg_tc1_db2hadr_standby' -n 'sapdbhr4 sapdbhr3' -S 'OFAN' -O 'FNPN' -B 'NFB'
```

## Changing and showing resources and attributes for a resource group

We describe how to change and show resource and attributes for a resource group:

1. Launch PowerHA by using the path for `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Change/Show Resources and Attributes for a Resource Group**
2. Select the resource group that you want to change as shown in Figure 7-116 on page 476.

```
Change/Show Resources and Attributes for a Resource Group

Move cursor to desired item and press Enter.

   rg_tc1_db2hadr_instance
   rg_tc1_db2hadr_primary
   rg_tc1_db2hadr_standby
```

*Figure 7-116 Change/Show Resources and Attributes for a Resource Group*

3. Change and show resources and attributes for the resource group `rg_tc1_db2hadr_instance` and select the appropriate application controller (Figure 7-117).

```
        Change/Show All Resources and Attributes for a Resource Group

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

                                                         [Entry Fields]
Resource Group Name                                  rg_tc1_db2hadr_instance
  Participating Nodes (Default Node Priority)          sapdbhr3 sapdbhr4

  Startup Policy                             Online On All Available Nodes
  Fallover Policy                       Bring Offline (On Error Node Only)
  Fallback Policy                                      Never Fallback

  Concurrent Volume Groups                            [] +
  Use forced varyon of volume groups, if necessary     false +
  Automatically Import Volume Groups                   false +

  Application Controllers                 [as_tc1_db2hadr_instance] +
...
```

*Figure 7-117 Change resources and attributes for resource group rg_tc1_db2hadr_instance*

Or, you can use the command line as shown in Example 7-122.

*Example 7-122 Command to change resources for resource group rg_tc1_db2hadr_instance*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddres -g
'rg_tc1_db2hadr_instance' APPLICATIONS='as_tc1_db2hadr_instance'
```

4. Press F3.
5. Select the resource group that you want to change as shown in Figure 7-118.

```
Change/Show Resources and Attributes for a Resource Group

Move cursor to desired item and press Enter.

   rg_tc1_db2hadr_instance
   rg_tc1_db2hadr_primary
   rg_tc1_db2hadr_standby
```

*Figure 7-118 Change/Show Resources and Attributes for a Resource Group*

6. Change and show resources and attributes for the resource group `rg_tc1_db2hadr_primary` and select the appropriate application controller (Figure 7-119).

```
        Change/Show All Resources and Attributes for a Resource Group

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

                                                         [Entry Fields]
Resource Group Name                                  rg_tc1_db2hadr_primary
  Participating Nodes (Default Node Priority)          sapdbhr3 sapdbhr4

  Startup Policy                        Online On First Available Node
  Fallover Policy           Fallover To Next Priority Node In The List
  Fallback Policy                                       Never Fallback

  Service IP Labels/Addresses                             [sapdbsvc2] +
  Application Controllers                  [as_tc1_db2hadr_primary] +

  Volume Groups                                        [] +
  Use forced varyon of volume groups, if necessary      false +
  Automatically Import Volume Groups                    false +

...
```

*Figure 7-119 Change resources and attributes for resource group rg_tc1_db2hadr_primary*

Or, you can use the command line as shown in Example 7-123.

*Example 7-123 Command to change resources for resource group rg_tc1_db2hadr_primary*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddres -g
'rg_tc1_db2hadr_primary' SERVICE_LABEL='sapdbsvc2'
APPLICATIONS='as_tc1_db2hadr_primary'
```

7. Press F3.
8. Select the resource group that you want to change (Figure 7-120).

```
Change/Show Resources and Attributes for a Resource Group

Move cursor to desired item and press Enter.

   rg_tc1_db2hadr_instance
   rg_tc1_db2hadr_primary
   rg_tc1_db2hadr_standby
```

*Figure 7-120 Change/Show Resources and Attributes for a Resource Group*

9. Change and show resources and attributes for the resource group `rg_tc1_db2hadr_standby` and select the correct application controller (Figure 7-121 on page 478).

```
        Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
Resource Group Name                                 rg_tc1_db2hadr_standby
  Participating Nodes (Default Node Priority)         sapdbhr4 sapdbhr3

  Startup Policy                        Online On First Available Node
  Fallover Policy            Fallover To Next Priority Node In The List
  Fallback Policy                                       Never Fallback

  Service IP Labels/Addresses                                       [] +
  Application Controllers                   [as_tc1_db2hadr_standby] +

  Volume Groups                                       [] +
  Use forced varyon of volume groups, if necessary      false +
  Automatically Import Volume Groups                    false +

...
```

*Figure 7-121 Change resources and attributes for resource group rg_tc1_db2hadr_standby*

Or, you can use the command line as shown in Example 7-124.

*Example 7-124 Command to change resources for resource group rg_tc1_db2hadr_standby*

```
root@sapdbhr3 / # /usr/es/sbin/cluster/utilities/claddres -g
'rg_tc1_db2hadr_standby' APPLICATIONS='as_tc1_db2hadr_standby'
```

## Configuring the settling time

We describe how to configure the settling time:

1. Launch PowerHA by using the path for node `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Settling Time for Resource Groups** (Figure 7-122).

```
                 Configure Settling Time for Resource Groups

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

                                                        [Entry Fields]
S Settling Time (in seconds)                            [60]
```

*Figure 7-122 Configure Settling Time for Resource Groups*

Or, you can use the command line as shown in Example 7-125.

*Example 7-125 Command to change the settling time*

```
/usr/es/sbin/cluster/utilities/clsettlingtime change ‘60’
```

## Configuring a resource group processing order

We help configure the resource group processing order:

1. Launch PowerHA by using the path for node `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Resource Group Processing Ordering**.
2. The Change/Show Resource Group Processing Order panel opens (Figure 7-123).

```
                    Change/Show Resource Group Processing Order

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

                                                        [Entry Fields]
Resource Groups Acquired in Parallel
rg_tc1_db2hadr_instance rg_tc1_db2hadr_primary rg_tc1_db2hadr_standby
  Serial Acquisition Order
  New Serial Acquisition Order
[rg_tc1_db2hadr_instance rg_tc1_db2hadr_standby rg_tc1_db2hadr_primary] +

  Resource Groups Released in Parallel
rg_tc1_db2hadr_instance rg_tc1_db2hadr_primary rg_tc1_db2hadr_standby
  Serial Release Order
  New Serial Release Order
[rg_tc1_db2hadr_primary rg_tc1_db2hadr_standby rg_tc1_db2hadr_instance]
```

*Figure 7-123 Change/Show Resource Group Processing Order*

Or, you can use the command line as shown in Example 7-126.

*Example 7-126 Command to change resource group processing order*

```
/usr/es/sbin/cluster/utilities/clrgorder \
-c \
-a 'rg_tc1_db2hadr_instance rg_tc1_db2hadr_standby rg_tc1_db2hadr_primary' \
-r 'rg_tc1_db2hadr_primary rg_tc1_db2hadr_standby rg_tc1_db2hadr_instance'
```

## Configuring online on dependencies of different nodes

We configure the online on the dependencies of different nodes (Figure 7-124 on page 480):

1. Launch PowerHA by using the path for `sapdbhr3`: **`smitty sysmirror`** → **Cluster Applications and Resources** → **Resource Groups** → **Configure Resource Group Run-Time Policies** → **Configure Dependencies between Resource Groups** → **Configure Online on Different Nodes Dependency**.
2. Enter `rg_tc1_db2hadr_primary` for the High Priority Resource Group. Enter `rg_tc1_db2hadr_standby` for the Low Priority Resource Group.

```
                Configure Online on Different Nodes Dependency

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

                                                              [Entry Fields]
High Priority Resource Group(s)                       [rg_tc1_db2hadr_primary] +
  Intermediate Priority Resource Group(s)             []                        +
  Low Priority Resource Group(s)                      [rg_tc1_db2hadr_standby] +
```

*Figure 7-124 Configure Online on Different Nodes Dependency*

Or, you can use the command line as shown in Example 7-127.

*Example 7-127 Command to add parent/child dependency between resource groups*

```
/usr/es/sbin/cluster/utilities/clrgdependency  \
   -t'ANTICOLLOCATION' \
   -u \
   -hp'rg_tc1_db2hadr_primary' \
   -ip'' \
   -lp'rg_tc1_db2hadr_standby'
```

## Checking the complete resource group configuration

We check the complete resource group configuration (Example 7-128).

*Example 7-128 Check the resource group configuration*

```
root@sapdbhr3 /usr/es/sbin/cluster # cllsgrp
rg_tc1_db2hadr_instance
rg_tc1_db2hadr_primary
rg_tc1_db2hadr_standby

root@sapdbhr3 / # clrgdependency -t ANTICOLLOCATION -sl
#HIGH:INTERMEDIATE:LOW
rg_tc1_db2hadr_primary::rg_tc1_db2hadr_standbyroot@sapdbhr3 / #

root@sapdbhr3 /usr/es/sbin/cluster # cldisp
Cluster: sapdbhr_cluster
   Cluster services: inactive

#############
APPLICATIONS
#############
   Cluster sapdbhr_cluster provides the following applications:
as_tc1_db2hadr_instance as_tc1_db2hadr_primary as_tc1_db2hadr_standby
      Application: as_tc1_db2hadr_instance
      as_tc1_db2hadr_instance is started by
/usr/es/sbin/cluster/sap/cl_db2_start_local DB2 TC1
      as_tc1_db2hadr_instance is stopped by
/usr/es/sbin/cluster/sap/cl_db2_stop_local DB2 TC1
         No application monitors are configured for as_tc1_db2hadr_instance.
         This application is part of resource group 'rg_tc1_db2hadr_instance'.
            Resource group policies:
               Startup: on all available nodes
```

```
            Fallover: bring offline on error node
            Fallback: never
          Nodes configured to provide as_tc1_db2hadr_instance: sapdbhr3 sapdbhr4
          No resources are associated with as_tc1_db2hadr_instance.

      Application: as_tc1_db2hadr_primary
      as_tc1_db2hadr_primary is started by
/usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY TC1
      as_tc1_db2hadr_primary is stopped by
/usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY TC1
         Application monitor of as_tc1_db2hadr_primary: mon_tc1_db2hadr_primary
            Monitor name: mon_tc1_db2hadr_primary
               Type: custom
               Monitor method: user
               Monitor interval: 10 seconds
               Hung monitor signal: 9
               Stabilization interval: 60 seconds
               Retry count: 0 tries
               Restart interval: 0 seconds
               Failure action: fallover
               Cleanup method: /usr/es/sbin/cluster/sap/cl_db2_stop_hadr PRIMARY
TC1
               Restart method: /usr/es/sbin/cluster/sap/cl_db2_start_hadr PRIMARY
TC1
         This application is part of resource group 'rg_tc1_db2hadr_primary'.
            Resource group policies:
               Startup: on first available node
               Fallover: to next priority node in the list
               Fallback: never
            Nodes configured to provide as_tc1_db2hadr_primary: sapdbhr3 sapdbhr4
            Resources associated with as_tc1_db2hadr_primary:
               Service Labels
                  sapdbsvc2(10.10.20.56) {}
                     Interfaces configured to provide sapdbsvc2:
                        sapdbhr3b2 {}
                           with IP address: 10.1.2.28
                           on interface: en1
                           on node: sapdbhr3 {}
                           on network: admin_net {}
                        sapdbhr4b2 {}
                           with IP address: 10.1.2.44
                           on interface: en1
                           on node: sapdbhr4 {}
                           on network: admin_net {}

      Application: as_tc1_db2hadr_standby
      as_tc1_db2hadr_standby is started by
/usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY TC1
      as_tc1_db2hadr_standby is stopped by
/usr/es/sbin/cluster/sap/cl_db2_stop_hadr STANDBY TC1
         Application monitor of as_tc1_db2hadr_standby: mon_tc1_db2hadr_standby
            Monitor name: mon_tc1_db2hadr_standby
               Type: custom
               Monitor method: user
               Monitor interval: 10 seconds
```

```
          Hung monitor signal: 9
          Stabilization interval: 60 seconds
          Retry count: 3 tries
          Restart interval: 264 seconds
          Failure action: fallover
          Cleanup method:
          Restart method: /usr/es/sbin/cluster/sap/cl_db2_start_hadr STANDBY
TC1
      This application is part of resource group 'rg_tc1_db2hadr_standby'.
        Resource group policies:
          Startup: on first available node
          Fallover: to next priority node in the list
          Fallback: never
        Nodes configured to provide as_tc1_db2hadr_standby: sapdbhr3 sapdbhr4
        No resources are associated with as_tc1_db2hadr_standby.

#############
TOPOLOGY
#############
  sapdbhr_cluster consists of the following nodes: sapdbhr3 sapdbhr4
    sapdbhr3
      Network interfaces:
        sapdbhr3b2 {}
          with IP address: 10.1.2.28
          on interface: en1
          on network: admin_net {}
        sapdbhr3b1 {}
          with IP address: 10.1.1.28
          on interface: en0
          on network: service_net {}
    sapdbhr4
      Network interfaces:
        sapdbhr4b2 {}
          with IP address: 10.1.2.44
          on interface: en1
          on network: admin_net {}
        sapdbhr4b1 {}
          with IP address: 10.1.1.44
          on interface: en0
          on network: service_net {}
```

### 7.8.7 Completing the configuration

After the resource groups are configured, complete the configuration:

1. Stop the DB2 instance on the primary node as shown in Example 7-129, if it is running.

*Example 7-129 Stopping the DB2 instance on the primary node*

```
sapdbhr3:/ # su - db2tc1 -c db2gcf -d -p 0 -i db2tc1

Instance  : db2tc1
DB2 Stop : Success
Partition 0 : Success
```

2. Stop the DB2 instance on the standby node if it is running, as shown in Example 7-130.

*Example 7-130 Stopping the DB2 instance on the standby node*

```
sapdbhr4:/ # su - db2tc1 -c db2gcf -d -p 0 -i db2tc1

Instance  : db2tc1
DB2 Stop : Success
Partition 0 : Success
```

3. Unconfigure the network alias on the primary node as shown in Example 7-131.

*Example 7-131 Network settings on sapdbhr3*

```
root@sapdbhr3 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.da.d9.10.6f 22407414     0 11005526     0     0
en0   1500  10.1        sapdbhr3b1        22407414     0 11005526     0     0
en0   1500  172.16.20   sapdbhr3p1        22407414     0 11005526     0     0
en1   1500  link#3      6e.8d.da.d9.10.70  1306148     0     2525     0     0
en1   1500  10.1.2      sapdbhr3b2         1306148     0     2525     0     0
en1   1500  10.10.20    sapdbhr3p2         1306148     0     2525     0     0
en1   1500  10.10.20    sapdbsvc2           344897     0    42429     0     0
lo0   16896 link#1                          352882     0   352882     0     0
lo0   16896 127         loopback            352882     0   352882     0     0
lo0   16896 loopback                        352882     0   352882     0     0

root@sapdbhr3 / # ifconfig en1 delete sapdbsvc2
root@sapdbhr3 / # netstat -i
Name  Mtu   Network     Address            Ipkts Ierrs    Opkts Oerrs  Coll
en0   1500  link#2      6e.8d.da.d9.10.6f 22407414     0 11005526     0     0
en0   1500  10.1        sapdbhr3b1        22407414     0 11005526     0     0
en0   1500  172.16.20   sapdbhr3p1        22407414     0 11005526     0     0
en1   1500  link#3      6e.8d.da.d9.10.70  1306148     0     2525     0     0
en1   1500  10.1.2      sapdbhr3b2         1306148     0     2525     0     0
en1   1500  10.10.20    sapdbhr3p2         1306148     0     2525     0     0
lo0   16896 link#1                          352882     0   352882     0     0
lo0   16896 127         loopback            352882     0   352882     0     0
lo0   16896 loopback                        352882     0   352882     0     0
```

4. Synchronize the PowerHA cluster by using SMIT:

   a. Follow the path: `smitty sysmirror` → **Custom Cluster Configuration** → **Verify and Synchronize Cluster Configuration (Advanced)**.

   b. In the PowerHA SystemMirror Verification and Synchronization panel (Figure 7-125 on page 484), press Enter to accept the default options.

```
              PowerHA SystemMirror Verification and Synchronization

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Verify, Synchronize or Both                          [Both]          +
* Include custom verification library checks           [Yes]           +
* Automatically correct errors found during            [Yes]           +
  verification?
* Force synchronization if verification fails?         [No]            +
* Verify changes only?                                 [No]            +
* Logging                                              [Standard]      +
```

*Figure 7-125 Accepting the default actions on the Verification and Synchronization panel*

5. Start the cluster on both nodes, `sapdbhr3` and `sapdbhr4`, by running **`smitty clstart`**.
6. In the Start Cluster Services panel (Figure 7-126), complete these steps:
   a. For Start now, on system restart or both, select **now**.
   b. For Start Cluster Services on these nodes, enter [`sapdbhr3,sapdbhr4`].
   c. For Manage Resource Groups, select **Automatically**.
   d. For BROADCAST message at startup, select **false**.
   e. For Startup Cluster Information Daemon, select **true**.
   f. For Ignore verification errors, select **false**.
   g. For Automatically correct errors found during cluster start, select **yes**.

   Press Enter.

```
                              Start Cluster Services

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[Entry Fields]
* Start now, on system restart or both                now                    +
  Start Cluster Services on these nodes               [sapdbhr3 sapdbhr4]    +
* Manage Resource Groups                              Automatically          +
  BROADCAST message at startup?                       false                  +
  Startup Cluster Information Daemon?                 true                   +
  Ignore verification errors?                         false                  +
  Automatically correct errors found during           yes                    +
  cluster start?
```

*Figure 7-126 Specifying the options for starting cluster services*

When the PowerHA cluster starts, the DB2 instance is automatically started. The application monitors start after the defined stabilization interval as shown in Example 7-132.

*Example 7-132 Checking the status of the highly available cluster and the DB2 instance*

```
root@sapdbhr3 /var/hacmp/log # clRGinfo
-----------------------------------------------------------------------------
Group Name     State                        Node
-----------------------------------------------------------------------------
rg_tc1_db2hadr ONLINE                       sapdbhr3
               ONLINE                       sapdbhr4
```

```
rg_tc1_db2hadr ONLINE                       sapdbhr3
               OFFLINE                      sapdbhr4

rg_tc1_db2hadr ONLINE                       sapdbhr4
               OFFLINE                      sapdbhr3

root@sapdbhr3 /var/hacmp/log # ps -ef|grep clappmon|grep -v grep
    root 16318592  8454352   0 03:37:51      -  0:00 run_clappmond -sport 1000
-result_node sapdbhr3 -script_id 0 -command_id 9 -command mon_tc1_db2hadr_primary
-environment
?VERBOSE_LOGGING=high??FAIL_COUNT=0??CLUSTER_VERSION=13??GS_NODEID=1??APPLICATION_
SERVER=mon_tc1_db2hadr_primary??MISC_DATA=??GROUPNAME=rg_tc1_db2hadr_primary??RESO
URCE_GROUP=rg_tc1_db2hadr_primary??RESTART_METHOD=/usr/es/sbin/cluster/sap/cl_db2_
start_hadr PRIMARY TC1??CLEANUP_METHOD=/usr/es/sbin/cluster/sap/cl_db2_stop_hadr
PRIMARY
TC1??NOTIFY_METHOD=??MONITOR_METHOD=/usr/es/sbin/cluster/sap/cl_db2_monitor_hadr
PRIMARY
TC1??FAILURE_ACTION=fallover??RESTART_INTERVAL=0??HUNG_MONITOR_SIGNAL=9??RESTART_C
OUNT=0??STABILIZATION_INTERVAL=60??MONITOR_INTERVAL=10??INSTANCE_COUNT=0??PROCESS_
OWNER=??PROCESSES=??MONITOR_TYPE=user??HACMP_VERSION=__PE__??PATH=/usr/bin:/etc:/u
sr/sbin:/usr/ucb:/usr/bin/X11:/sbin??ODMDIR=/etc/es/objrepos??LC_FASTMSG=true??PIN
G_IP_ADDRESS=
??LOCALNODEID=sapdbhr3??LOCALNODENAME=sapdbhr3??CM_CLUSTER_NAME=sapdbhr_cluster??C
M_CLUSTER_ID=1089684365?
    root 16973828 16318592   0 03:37:51      -  0:00
/usr/es/sbin/cluster/clappmond mon_tc1_db2hadr_primary
```

Your DB2 instance and database are now configured for high availability in a hot-standby DB2 HADR PowerHA SystemMirror configuration.

# 8

# Workload partition and PowerHA scenario

This chapter describes scenarios that relate to workload partitions (WPARs) in an IBM PowerHA SystemMirror configuration for Standard Edition 7.1.1 for AIX.

This chapter presents the following sections:

- Introduction to WPARs
- Planning for high availability
- Support for a WPAR in PowerHA
- Scenario with a local WPAR
- SAP scenario on AIX 7.1 NFS WPAR
- NFS versioned 5.2 WPAR

## 8.1 Introduction to WPARs

Workload partitions (WPARs) are software-created virtualized operating system environments within a single instance of the AIX operating system. WPARs secure and isolate the environment for the processes and signals that are used by enterprise applications.

There are multiple WPARs types: Application WPARs or System WPARs. System WPARs are autonomous virtual system environments with their own private file systems, users and groups, login, network space, and administrative domain.

By default, a system WPAR shares the two file systems named `/usr` and `/opt` from the global environment by using read-only namefs mounts. You can configure WPARs to have a non-shared, writable `/usr` file system and `/opt` file system. The WPARs are also called private.

For more information about IBM AIX WPARs, see *Exploiting IBM AIX Workload Partitions*, SG24-7955.

In AIX Version 7, administrators now can create WPARs that can run AIX 5.2 inside an AIX 7 operating system instance. It is supported on the POWER7 server platform. PowerHA support is announced with the January 21, 2011 PowerHA Support Flashes for VIOS 2.2 and new Versioned 5.2 WPAR Support Flash.

For the announcement details, see the website:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10737

> **Important:** Versioned WPARs can be non-shared system WPARs only.

We describe three user scenarios: a scenario a local (namefs) WPAR, a scenario with an NFS shared WPAR, and a scenario with a versioned WPAR 5.2.

## 8.2 Planning for high availability

The WPAR offering is supported by IBM PowerHA SystemMirror since version 5.4.1. However, particularly in the planning phase, be careful because the combination of WPARs and PowerHA in an environment can potentially introduce new single points of failure (SPOFs).

> **PowerHA:** PowerHA does not manage or monitor the WPAR. It manages and monitors only the applications that run within the WPAR.

### 8.2.1 General considerations

In PowerHA, you can have a mixture of normal resource groups and resource groups that run in a WPAR. Figure 8-1 on page 489 shows an example. In this example, we have two resource groups. One resource group runs in the Global AIX or Global WPAR environment. The second resource group runs inside a WPAR. Both resource groups have two defined application servers and an application monitor for each resource group.

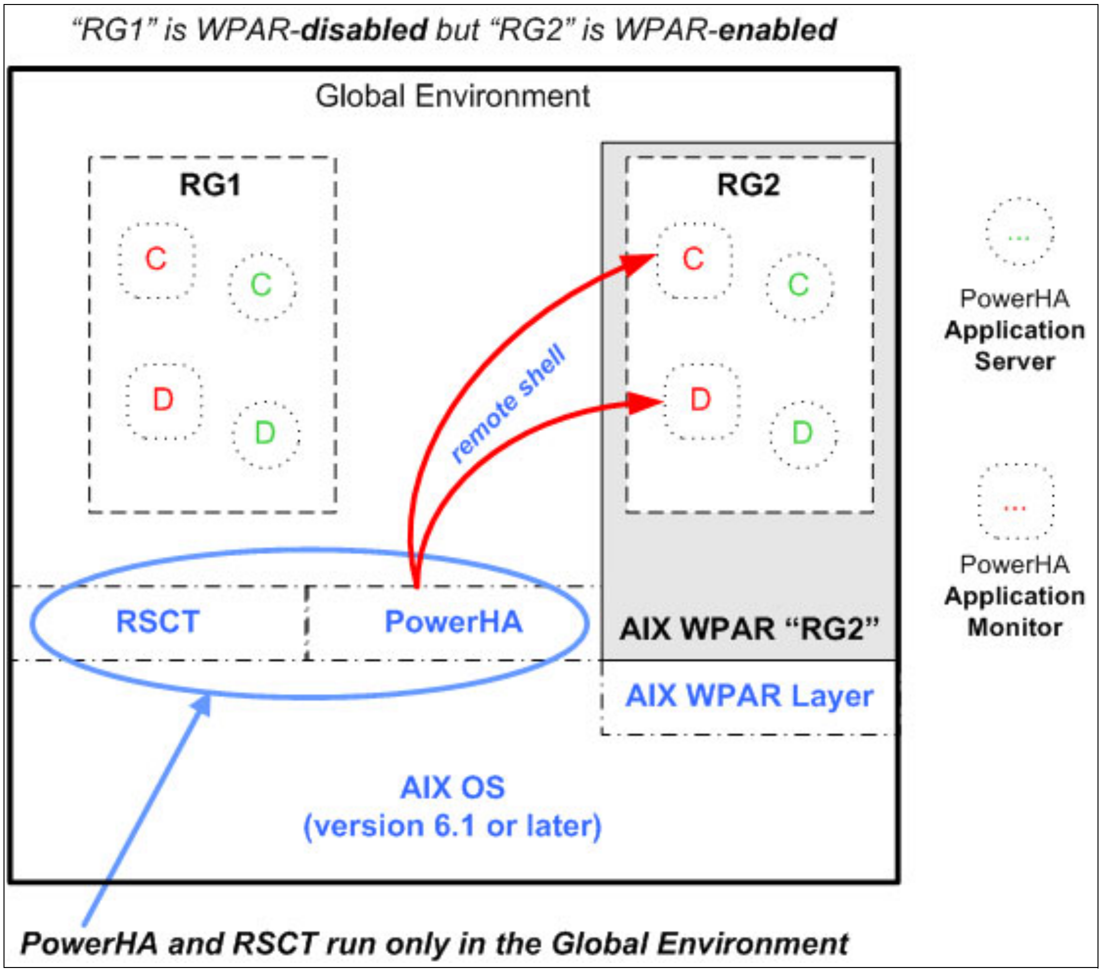



*Figure 8-1 PowerHA and WPAR basics*

## 8.2.2 PowerHA and rootvg WPARs

A system WPAR, which is configured with its own dedicated root volume group, is called a rootvg WPAR. There are several restrictions for managing rootvg WPARs. Rootvg WPARs are *not supported* by PowerHA.

> **Important:** PowerHA does *not* have integration support to manage rootvg WPARs, although it has integration support for system WPARs.

Configuring a rootvg WPAR by using one or more storage devices gives the WPAR administrator complete control of the storage devices that are exported to the WPAR, the volume groups on these devices, and the logical volumes and file systems within these volume groups. A system WPAR, which is not a rootvg WPAR, does not have its own root volume group. It has the same file system layout that is created in logical volumes, which are created externally in another place such as a Network File System (NFS) server (NFS WPARs) or on a volume group of the global system (local WPAR).

## 8.2.3 WPAR on local disk

This solution has a limited use. All data of the WPAR is on a local disk. This solution might be appropriate for any application that can use NFS or General Parallel File System (GPFS) shared file systems, for example, an application server.

If PowerHA creates the WPAR, this type of installation and configuration results. For more details, see 8.3.2, “Creating a WPAR with the Resource Group menu” on page 494.

## 8.2.4 Planning for NFS-based file systems

We describe the necessary considerations when you plan to use NFS for your WPAR in your PowerHA cluster.

### Planning steps

In this section, we summarize the setup sequence and necessary planning:

1. Set up the NFS server.

   We include the main items:

   – If you use a dedicated NFS server or network-attached storage (NAS) system, ensure that it has the same or better availability capabilities as your systems.

   – Remember to check whether you have enough disk space available on your NFS or NAS server.

   – Ensure that the root equivalency is defined. See Example 8-1 on page 491 for details.

   – Create the directory for your WPAR.

2. Configure the WPAR.

   The file system for a WPAR is structured from a starting directory, for instance, `/wpars/<wpar_name>`. This directory contains subdirectories for each private file system in the WPAR.

   The starting directory of the WPAR must be created in the NFS server before the execution of the **mkwpar** command.

   > **Important:** The wpar_name must equal the PowerHA resource group name that you plan to use.

   For an NFS-based WPAR, each file system must be specified at creation time. These file systems include `/`, `/home`, `/var/hacmp/adm`, `/var`, and, optionally, `/usr` and `/opt` for a private system WPAR.

   For an example, see “Defining WPAR” on page 491.

3. Configure PowerHA.

### NFS setup

For an NFS-based WPAR in an PowerHA environment, each node in the cluster must have root access to the NFS shared file systems that contain the WPAR data. Example 8-1 on page 491 shows how the entry in the `/etc/exports` might look. In this example, the PowerHA cluster nodes are `sys5lpar3` and `sys5lpar4`. The NFS server is a third system (not part of the cluster).

*Example 8-1 Content of /etc/exports on a dedicated NFS server*

```
cat /etc/exports
/wpars -sec=sys:krb5p:krb5i:krb5:dh,rw,root=sys5lpar3:sys5lpar4
#
```

Before you can create your WPAR, you have to check that you created the main WPAR directory in your NFS server. In our example, it is named `testwpar`. So we performed a **`mkdir testwpar`**. Then, we used the command that is shown in Example 8-2.

## Defining WPAR

For an NFS-based WPAR, each file system must be specified at creation time. These file systems include `/`, `/home`, `/var/hacmp/adm`, `/var`, and, optionally, `/usr` and `/opt` for a private system WPAR.

*Example 8-2 Create a WPAR on the first node*

```
# mkwpar -r -a -N address=10.12.154.175 -n testwpar -h testwpar \
> -M directory=/ vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/ \
> -M directory=/var vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/var \
> -M directory=/var/hacmp/adm vfs=nfs host=hg5lpar1
dev=/wpars/testwpar/var/hacmp/adm \
> -M directory=/home vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/home
#
```

When the WPAR is created on the first node, you can define it on the next node or nodes by adding the **`-p`** option to the command that is used in Example 8-2. If you forget the **`-p`** option, you get an error message. Example 8-3 shows the command that we used.

*Example 8-3 Create a WPAR on the second node*

```
# mkwpar -r -a -p -N address=10.12.154.175 -n testwpar -h testwpar \
> -M directory=/ vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/ \
> -M directory=/var vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/var \
> -M directory=/var/hacmp/adm vfs=nfs host=hg5lpar1
dev=/wpars/testwpar/var/hacmp/adm \
> -M directory=/home vfs=nfs host=hg5lpar1 dev=/wpars/testwpar/home
#
```

## Configuring the resource group in PowerHA

The important part here is that you check the WPAR name is the same as the resource name, and these two are equal to the name you decided to use in the "NFS setup" on page 490.

Example 8-4 shows the important fields in the Change/Show window for resource groups (Example 8-4). The full listing is in Example F-1 on page 700.

*Example 8-4 Resource group settings for WPAR*

```
Change/Show All Resources and Attributes for a Resource Group

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

                                                       [Entry Fields]
  Resource Group Name                                 testwpar
  Participating Nodes (Default Node Priority)         sys5lpar3 sys5lpar4
```

```
  Startup Policy                                     Online On Home Node Only
  Fallover Policy                                    Fallover To Next Priority
Node In The List
  Fallback Policy                                    Fallback To Higher Priority
Node In The List
  Fallback Timer Policy (empty is immediate)        []                          +

  Service IP Labels/Addresses                       [localwpar]                 +
  Application Controllers                           [ApplicationB]              +

  Volume Groups                                     []                          +

...

Miscellaneous Data                                  []
  WPAR Name                                         [testwpar]                  +
  User Defined Resources                            [ ]
```

## 8.2.5 Planning for a versioned WPAR

The goal is to run an existing AIX 5.2 environment with a small application within a versioned WPAR 5.2 and to describe how the PowerHA configuration allows a failover transaction. Running an existing AIX 5.2 environment inside of an AIX 7 operating system instance requires the use of a versioned WPAR 5.2. It is supported on the POWER7 server platform.

A versioned WPAR provides a different version of the runtime environment than the global system. Support for AIX 5.2 or AIX 5.3 versioned WPARs requires the installation of more licensed program products:

- AIX 5.2 WPARs for AIX 7
- AIX 5.3 WPARs for AIX 7

A **mksysb** backup of a system that runs an earlier version of AIX is used to create the versioned WPAR. Applications that run in a versioned WPAR use the commands and libraries from the operating system files where the backup is made to create the versioned WPAR, for example, AIX 5.2 or AIX 5.3. These versioned WPARs own writable `/opt` and `/usr` file systems. Applications that run in the versioned WPAR do not need to know that the global system has a different version of the AIX operating system.

### Requirements for versioned WPARs

The following requirements are necessary to create a versioned WPAR:

- POWER7 hardware.
- The minimum level of AIX 5.2 that can be used within a versioned AIX 5.2 WPAR is AIX 5.2 with Technology Level (TL) 10 and Service Pack (SP) 8. Therefore, any backup image that is used to create an AIX 5.2 WPAR must be from an AIX 5.2 system that runs the latest version.
- The minimum level of AIX 5.3 that can be used within a versioned AIX 5.3 WPAR is AIX 5.3 with TL 12. Therefore, any backup image that is used to create an AIX 5.3 WPAR must be from an AIX 5.3 system that runs TL 12 or later.

### Installing support for a versioned WPAR

The versioned WPAR product that is associated with the level of AIX WPAR to be created must be installed on the system.

The product media contains the required installation images that are called `vwpar.images` to support the creation of versioned WPARs. The product media contains optional software that provides System Management Interface Tool (SMIT) support to create and manage versioned WPARs.

### Creating a versioned WPAR

You can create a new versioned WPAR with the **mkwpar** command, the smitty interface, or the System Director plug-in for WPAR.

Each WPAR has an isolated network environment with unique IP addresses and a unique host name. You can access WPARs through standard networking programs, such as telnet, FTP, and rlogin.

The following example shows the command-line command to create the WPAR:

```
mkwpar -n WPARname -C -B /mksysb_images/backupname
```

The command creates the WPAR according to your backup. The initial output of the **mkwpar** command looks similar to the following example:

```
mkwpar: Extracting file system information from backup...
    mkwpar: Creating file systems...
Creating file system '/' specified in image.data
     /bff
    Creating file system '/bff' specified in image.data
     /home
    Creating file system '/home' specified in image.data
```

## 8.3 Support for a WPAR in PowerHA

The current support of WPAR in PowerHA is oriented toward the basic WPARs:

- Currently, support is available for local (`namefs` file systems) and NFS WPARs only. WPARs can be shared or private. Versioned WPARs are also supported.
- When a WPAR-enabled resource group (RG) is brought online, all its associated resources are activated within the corresponding WPAR. The WPAR-enabled RG is associated with a WPAR based on their common name. If a resource group called `wpar_rg` is WPAR-enabled, it is associated with a WPAR with the name `wpar_rg`.
- When an RG is WPAR-enabled, all user scripts, such as application start and stop scripts must be accessible within the WPAR, at the paths that are specified in the PowerHA configuration. It is the responsibility of the user to verify that these scripts are executable and return `0`.
- A WPAR-enabled RG can consist of some nodes that are not WPAR capable so you do not need to upgrade all nodes of the RG to the latest AIX operating system version. And when a WPAR-enabled RG comes online on a WPAR-incapable node, it behaves as if the WPAR property for the RG is not set. However, you must ensure that all user-defined scripts are accessible at the same path as previously specified in the PowerHA configuration.

- A WPAR-enabled RG supports the following resources: service label, application servers, and file systems. The service address is mandatory. The service address is allocated to the WPAR when PowerHA starts the RG.
- When a WPAR-enabled RG is deleted, the corresponding WPAR on the nodes of the RG are unaffected (that is, the corresponding WPAR is not deleted).
- All the supported resource types that are supported for a WPAR-enabled RG can be DARE added and removed from a WPAR-enabled RG. If the WPAR property of an RG is changed through DARE (when the RG is online), the effect takes place when the RG is brought online the next time.
- PowerHA configuration verification checks that all WPAR-capable nodes of a WPAR-enabled RG have a WPAR that is configured for the RG (that is, a WPAR with the same name as the RG). If the PowerHA configuration verification is run with corrective action enabled, you are prompted to fix the WPAR-related verification errors through PowerHA corrective action. It might mean the creation of a local WPAR on all nodes that are specified in the RG modification menu.
- When a WPAR-enabled RG is brought online on a WPAR-capable node, PowerHA (which runs in the global WPAR) automatically sets up **rsh** access to the corresponding WPAR to manage various resources that are associated with the RG.

**Important:** PowerHA automatically assigns and unassigns resources to and from a WPAR as the corresponding WPAR-enabled resources come online (or go offline). You must not assign any PowerHA resources to a WPAR.

### Considerations

Consider the following important information:

- PowerHA Smart Assist scripts are not supported for a WPAR-enabled RG. Therefore, any application server or application monitoring script that uses the PowerHA Smart Assist scripts cannot be configured as a part of a WPAR-enabled RG.
- Process application monitoring is not supported for WPAR-enabled RGs.
- For every WPAR-capable node that is a part of a WPAR-enabled RG and contains a WPAR for a WPAR-enabled RG, at least one of the service labels (of the WPAR-enabled RG) must be accessible from the corresponding global WPAR.

**Important:** Only the Global instance can run PowerHA. A WPAR can be considered an RG of the type WPAR-enabled RG only.

## 8.3.1 Creating a WPAR before you define a Resource Group

We highly advise that you create your WPAR before you add it to an RG in PowerHA.

## 8.3.2 Creating a WPAR with the Resource Group menu

To create an RG, you enter this command: **smit hacmp**. Select **Cluster Applications and Resources** → **Resource Groups** → **Add a Resource Group**. Or, use the fast path: **smitty cm_add_resource_group**.

The Add a Resource Group panel opens as shown in Figure 8-2 on page 495. Use this menu to specify the RG name and the startup and stop script full path that is available in the Global instance.

```
                              Add a Resource Group

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

                                                        [Entry Fields]
* Resource Group Name                                [ApplA]
* Participating Nodes (Default Node Priority)        [wpar1 wpar2]          +

  Startup Policy                                      Online On Home Node O> +
  Fallover Policy                                     Fallover To Next Prio> +
  Fallback Policy                                     Fallback To Higher Pr> +



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

*Figure 8-2 Adding an RG*

The WPAR-enable specification is added through the Change/Show resources and Attributes for a Resource Group menu. After you specify the application name that is entered in the Resource Group menu, you are shown a complete menu to specify the nodes, service address, and WPAR name specification. In our example, we specified a two-node list `wpar1` and `wpar2`, a service IP address as `wpar1sap`, a set of scripts that is part of the application controller group `ApplA`, and the WPAR named `ApplA`.

The path to access this SMIT panel (Figure 8-3 on page 496) is **`smit hacmp`**. Select **Cluster Applications and Resources → Resource Groups → Change/Show Resources and Attributes for a Resource Group**. Or, use fast path: **`smitty cm_resource_groups`** and select **Change/Show All Resources and Attributes for a Resource Group**.

```
       Change/Show All Resources and Attributes for a Resource Group

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

[TOP]                                                    [Entry Fields]
  Resource Group Name                                  ApplA
  Participating Nodes (Default Node Priority)          wpar1 wpar2

  Startup Policy                                       Online On First Avail>
  Fallover Policy                                      Fallover To Next Prio>
  Fallback Policy                                      Never Fallback

  Service IP Labels/Addresses                         [wpar1sap]           +
  Application Controllers                             [ApplA]              +

  Volume Groups                                       []             +
  Use forced varyon of volume groups, if necessary     false               +
  Automatically Import Volume Groups                   false               +

  Filesystems (empty is ALL for VGs specified)        [ ]                  +
  Filesystems Consistency Check                        fsck                +
  Filesystems Recovery Method                          sequential          +
  Filesystems mounted before IP configured             false               +

  Filesystems/Directories to Export (NFSv2/3)         []                   +
  Filesystems/Directories to Export (NFSv4)           []                   +
  Stable Storage Path (NFSv4)                         []                   +
  Filesystems/Directories to NFS Mount                []
  Network For NFS Mount                               []                   +

  Tape Resources                                      []                   +
  Raw Disk PVIDs                                      []                   +

  Primary Workload Manager Class                      []                   +

  Miscellaneous Data                                  []
  WPAR Name                                           [ApplA]              +
  User Defined Resources                              [ ]                  +
```

*Figure 8-3 Adding a WPAR-enabled RG*

> **Important:** If the WPAR name does not exist when you synchronize the configuration, you are asked to correct the error and the system creates a simple WPAR by using the command **`mkwpar -n WPAR-name`** on the specified node. The service address is attached to the WPAR when the RG is brought online.

If the WPAR name did not exist or you have a typo in the WPAR Name field, the WPAR is defined on the `rootvg` on all nodes that are part of this RG. See Example 8-5 on page 497.

*Example 8-5   WPAR fs in rootvg*

```
# svg -l rootvg
rootvg:
LV NAME             TYPE     LPs  PPs  PVs  LV STATE      MOUNT POINT
hd5                 boot     2    2    1    closed/syncd  N/A
...
fslv00              jfs2     6    6    1    closed/syncd /wpars/ApplA
fslv01              jfs2     2    2    1    closed/syncd /wpars/ApplA/home
fslv03              jfs2     6    6    1    closed/syncd
/wpars/ApplA/var/hacmp/adm
fslv04              jfs2     8    8    1    closed/syncd /wpars/ApplA/var
#
```

When you execute a **lswpar -M ApplA** command on your nodes, you get output as shown in Example 8-6.

*Example 8-6   lswpar on local disk*

```
# lswpar -M ApplA
Name     MountPoint           Device         Vfs     Nodename  Options
-----------------------------------------------------------------------
ApplA    /wpars/ApplA       /dev/fslv00     jfs2
ApplA    /wpars/ApplA/home  /dev/fslv01     jfs2
ApplA    /wpars/ApplA/opt   /opt            namefs            ro
ApplA    /wpars/ApplA/proc  /proc           namefs            rw
ApplA    /wpars/ApplA/var/hacmp/adm    /dev/fslv03     jfs2
ApplA    /wpars/ApplA/usr   /usr            namefs            ro
ApplA    /wpars/ApplA/var   /dev/fslv04     jfs2
#
```

Next, we describe the WPAR scenario example.

# 8.4 Scenario with a local WPAR

Creating a local WPAR does not allow migration nor duplication. However, you can create the same WPAR on two nodes that use a shared disk. We describe this scenario, which is represented by Figure 8-4 on page 498.

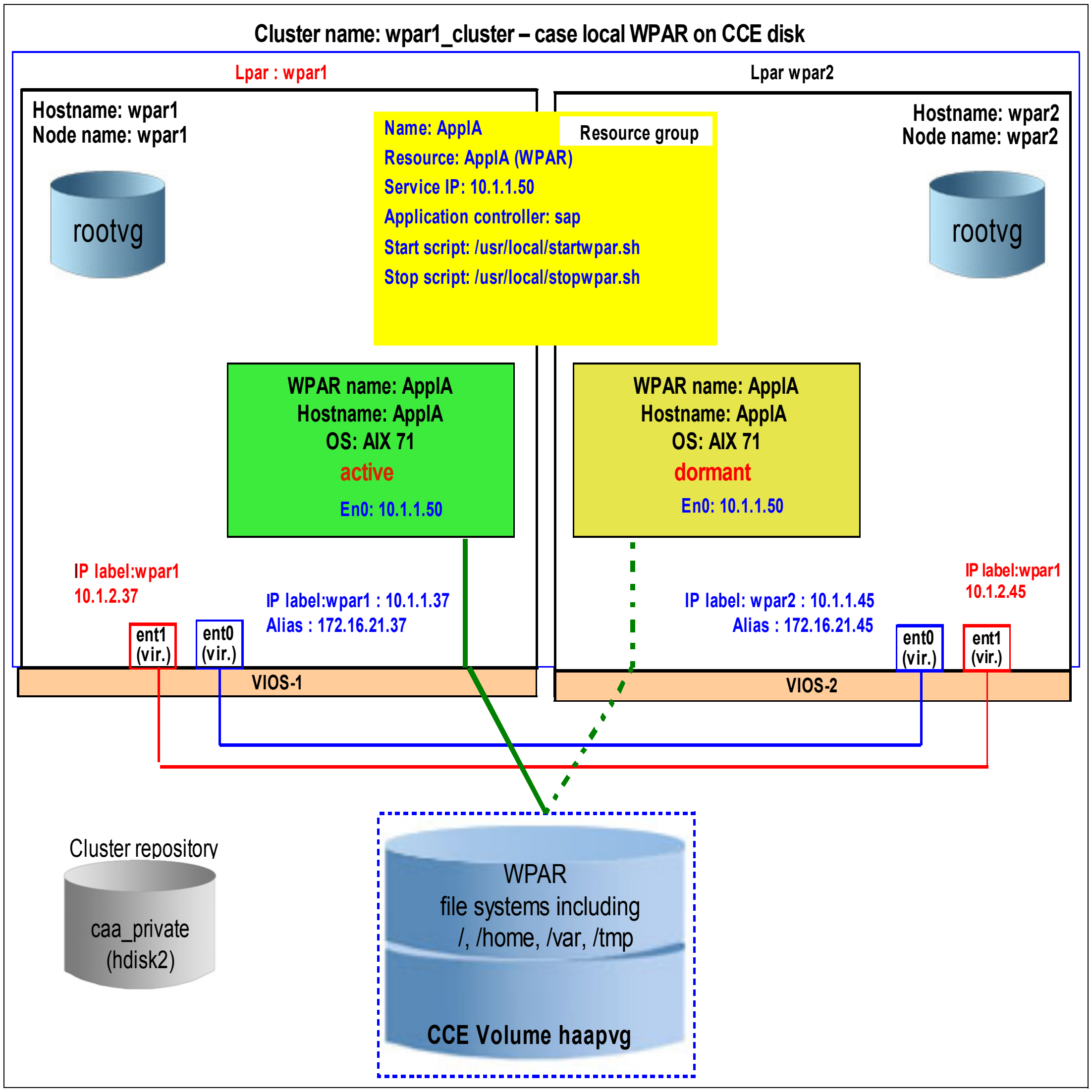



*Figure 8-4 Overview of the local (namefs) WPAR environment*

**Requirement:** It requires some AIX admin knowledge as well as some WPAR knowledge.

## 8.4.1 Creating a local WPAR on two nodes

The first creation of a local WPAR that uses a local shared disk is described as a reference but it is supposed to be known. It requires a volume group on the shared disk and an address for the WPAR (`10.1.1.50` as shown in Example 8-7).

*Example 8-7 Creation of a local WPAR on a shared volume group*

```
#Create the concurrent capable volume group
exportvg haapvg
mkvg -f -S -y haapvg -V 111 -C -n hdisk7
varyonvg haapvg

# Create the standalone wpar on volume haapvg with address 10.1.1.50
# and name ApplA
```

```
mkwpar -n ApplA -g haapvg -a -F -N address=10.1.1.50
   mkwpar: Creating file systems...
    /
    /home
    /opt
    /proc
    /var/hacmp/adm
    /usr
    /var
   Mounting all workload partition file systems.
   x ./usr
   x ./lib
   x ...........
   Workload partition ApplA created successfully.
   mkwpar: 0960-390 To start the workload partition, execute the following as
   root: startwpar [-v] ApplA

# Change lv names for further processing
lsvg -l haapvg
/usr/sbin/chlv -n'wrg1' fslv00
/usr/sbin/chlv -n'wrg1_home' fslv01
/usr/sbin/chlv -n'wrg1_var/hacmp/adm' fslv02
/usr/sbin/chlv -n'wrg1_var' fslv03

# Ensure wpar is able to start
startwpar ApplA
```

The WPAR is started and functional. We can create a specfile configuration file to use on the other node (Example 8-8).

*Example 8-8 Creation of a specfile*

```
mkwpar -w -e ApplA -o ApplA.spec
```

We need to vary off the volume group for the other node to use it and create the WPAR again as though it is new (Example 8-9).

*Example 8-9 Moving the volume to the other node*

```
varyoffvg haapvg

# GOTO WPAR2
lspv
importvg -y haapvg -V 111 hdisk1
varyonvg haapvg

root@wpar2 / # lsvg -l haapvg
haapvg:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
wrg1                jfs2       12      12      1    closed/syncd  /wpars/ApplA
wrg1_home           jfs2       4       4       1    closed/syncd  /wpars/ApplA/home
wrg1_var/hacmp/adm  jfs2       12      12      1    closed/syncd
/wpars/ApplA/var/hacmp/adm
wrg1_var            jfs2       16      16      1    closed/syncd  /wpars/ApplA/var
```

The volume is imported and the /etc/filesystems (Example 8-10) is populated with entries for the WPAR, but the WPAR does not exist. You need to remove the entries as shown in Example 8-10.

*Example 8-10 Removing the file systems that are imported*

```
rmfs /dev/wrg1
rmfs /dev/wrg1_home
rmfs /dev/wrg1_var/hacmp/adm
rmfs /dev/wrg1_var
```

Remove the /wpar/ApplA directory. You can create the WPAR again from the beginning by using the **mkwpar -f ApplA.spec** command (Example 8-11).

*Example 8-11 Creating the WPAR again by using a specfile*

```
# mkwpar -f ApplA.cfg
mkwpar: Creating file systems...
 /
 /home
 /opt
 /proc
 /var/hacmp/adm
 /usr
 /var
......
Workload partition ApplA created successfully.
mkwpar: 0960-390 To start the workload partition, execute the following as root:
startwpar [-v] ApplA
```

Create the file systems again as seen in the initial node by using **chlv** as shown in Example 8-12.

*Example 8-12 lsvg of the volume haapvg*

```
# Change the lv name to match initial node specification
lsvg -l haapvg
/usr/sbin/chlv -n'wrg1' fslv00
/usr/sbin/chlv -n'wrg1_home' fslv02
/usr/sbin/chlv -n'wrg1_var/hacmp/adm' fslv03
/usr/sbin/chlv -n'wrg1_var' fslv04
```

The WPAR is created again on node wpar2, and it can be started. To start the WPAR on node wpar1, vary the volume offline and vary the volume online.

The WPAR is defined on both nodes and can be started on the node where the volume haagvg is active.

> **Administrator:** Any modification to the configuration of the WPAR must be propagated to the other node. Any modification that uses the **chwpar** command must be issued on the two nodes.

We can configure the cluster.

## 8.4.2 Configuring PowerHA

For details to create a simple cluster, see Chapter 3, “PowerHA 7.1.1 basic installation and configuration” on page 73.

All commands can be issued by using tools, such as **smit**, **clmgr** on the command line, or the System Director plug-in. In some cases, where you can use the command line, it is listed.

We create a simple cluster with two nodes, one repository disk, and we create the appropriate RG for PowerHA functionality with the WPAR.

The following steps are listed for reference:

1. Set the routing table and persistent addresses on your systems.
2. Update `/etc/cluster/rhosts` with the two host names, `wpar1` and `wpar2`, and all service, persistent, and base addresses. Ensure that `/usr/es/sbin/cluster/etc/hosts` matches `/etc/cluster/rhosts`.
3. Create the cluster by using the smit **cm_setup_cluster_nodes_networks** fastpath or the **clmgr add cluster** command. For example, create the cluster `wpar1_cluster` with the two nodes, `wpar1` and `wpar2`.
4. Add the repository disk and the multicast address by using **clmgr modify cluster** or the **smitty cm_define_repos_ip_addr** fast path. In our example, we add `hdisk2` and `228.1.1.29`.
5. Check the addresses by using **/usr/es/sbin/cluster/utilities/cllsif**.
6. Add the persistent addresses for the nodes and the applications by using the smit panel **cm_add_interfaces**.
7. Check the topology by using **cltopinfo** as shown in Example 8-13.

*Example 8-13 Simple cluster topology output*

```
root@wpar1 / # cltopinfo
Cluster Name: wpar1_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.29
There are 2 node(s) and 2 network(s) defined

NODE wpar1:
        Network net_ether_01
                wparsvc1        172.16.21.63
                wpar1p1 172.16.21.37
                wpar1   10.1.1.37
        Network net_ether_02
                wpar1b2 10.1.2.37

NODE wpar2:
        Network net_ether_01
                wparsvc1        172.16.21.63
                wpar2p1 172.16.21.45
                wpar2   10.1.1.45
        Network net_ether_02
```

8. Verify and synchronize the cluster configuration.
9. Check the CAA cluster by using the **lscluster** commands as shown in Example 8-14.

*Example 8-14 lscluster output*

```
root@wpar1 / # lscluster -d
Storage Interface Query

Cluster Name:  wpar1_cluster
Cluster uuid:  41320306-7aa8-11e1-96d5-2e4791550c6f
Number of nodes reporting = 2
Number of nodes expected = 2
Node wpar1
Node uuid = 40f2ab20-7aa8-11e1-96d5-2e4791550c6f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : a6a85ae8-1e89-8ab5-bafc-5a27dc82aa5a
          type  : REPDISK
Node wpar2
Node uuid = 412cbb4e-7aa8-11e1-96d5-2e4791550c6f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : a6a85ae8-1e89-8ab5-bafc-5a27dc82aa5a
          type  : REPDISK
root@wpar1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: wpar1
        Cluster shorthand id for node: 1
        uuid for node: 40f2ab20-7aa8-11e1-96d5-2e4791550c6f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        wpar1_cluster      local        41320306-7aa8-11e1-96d5-2e4791550c6f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: wpar2
        Cluster shorthand id for node: 2
        uuid for node: 412cbb4e-7aa8-11e1-96d5-2e4791550c6f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
```

```
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        wpar1_cluster      local        41320306-7aa8-11e1-96d5-2e4791550c6f

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en1  UP
         en0  UP
```

10.Create the `ApplA` application controller scripts by using **`cm_add_app_scripts`** and ensure that they are executable on both nodes. Examples of these scripts are shown in Example 8-15.

*Example 8-15 Sample scripts to start and stop the RG ApplA*

```
# cat /usr/local/ha/StartA
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#
Name=$(basename $0 )

if [ "$Name" = "StartA" ]
then
  echo "$(date) \"Application A started\" " >> /var/hacmp/adm/applA.log
  touch /var/hacmp/adm/AppAup
  nohup /usr/local/bin/ApplA &
  exit 0
elif [ "$Name" = "StopA" ]
then
  rm -f /var/hacmp/adm/AppAup
  echo "$(date) \"Application A stopped\" " >> /var/hacmp/adm/applA.log
  exit 0
else
  echo "$(date) \"ERROR - Application A start/stop script called with wrong
name\" " >> /var/hacmp/adm/applA.log
  exit 999
fi

#---------------------------------------------------#
# cat /usr/local/ha/StopA
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#
Name=$(basename $0 )

if [ "$Name" = "StartA" ]
then
  echo "$(date) \"Application A started\" " >> /var/hacmp/adm/applA.log
  touch /var/hacmp/adm/AppAup
  nohup /usr/local/bin/ApplA &
  exit 0
elif [ "$Name" = "StopA" ]
```

```
then
  rm -f /var/hacmp/adm/AppAup
  echo "$(date) \"Application A stopped\" " >> /var/hacmp/adm/applA.log
  exit 0
else
  echo "$(date) \"ERROR - Application A start/stop script called with wrong
name\" " >> /var/hacmp/adm/applA.log
  exit 1
fi
```

11. Create the application monitor by using the Add Custom Application Monitor menu (Figure 8-5). The **cm_cfg_custom_appmon** command is the command-line command that brings up the menu that is shown in Figure 8-5. In our example, we add a script that is called `MonA` as shown in Example 8-16.

*Example 8-16 Local custom application monitor*

```
> cat /usr/local/ha/MonA
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#

if ( $(ps -ef | grep -w ApplA | grep -vq grep) )
then
  echo "MON:ApplA is running on $(uname -n)\n" >>/var/hacmp/adm/mon.log
  exit 0
else
  echo "MON:ApplA is NOT running on $(uname -n)\n" >>/var/hacmp/adm/mon.log
  exit 1
fi
```

```
 Add Custom Application Monitor

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

                                                        [Entry Fields]
* Monitor Name                                       [ApplA]
* Application Controller(s) to Monitor                ApplA                +
* Monitor Mode                                       [Long-running monitori> +
* Monitor Method                                     [/usr/local/ha/MonA]
  Monitor Interval                                   [2]                    #
  Hung Monitor Signal                                [9]                    #
* Stabilization Interval                             [1]                    #
* Restart Count                                      [3]                    #
  Restart Interval                                   []                     #
* Action on Application Failure                      [fallover]            +
  Notify Method                                      []
  Cleanup Method                                     [/usr/local/ha/StopA]
  Restart Method                                     [/usr/local/ha/StartA]


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

*Figure 8-5 Add Custom Application Monitor menu*

12. Add the application controller scripts (scripts for starting and stopping the WPAR) by using the **cm_add_app_scripts** smit command, and the menu is shown in Figure 8-6. Perform a quick check by using the **cllsserv** command.

```
Add Application Controller Scripts

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

                                                        [Entry Fields]
* Application Controller Name                        [ApplA]
* Start Script                                       [/usr/local/ha/StartA]
* Stop Script                                        [/usr/local/ha/StopA]
  Application Monitor Name(s)
+
  Application startup mode                           [background]
+

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

*Figure 8-6 cm_add_app_scripts menu*

13. Add a service label for the WPAR address `10.1.1.50` by using the command line or the smit **cm_add_a_service_ip_label_address.select_net** (Example 8-17).

*Example 8-17 Add a service label*

```
/usr/es/sbin/cluster/utilities/clmgr add service_ip 'wpar1sap' NETMASK |
  |   ='255.255.254.0' NETWORK='net_ether_01'
```

14. Add the RG `ApplA` by using smit **cm_add_resource_group**. The output can be checked with the **cltopinfo** command as shown in Example 8-18.

*Example 8-18 cltopinfo output*

```
Cluster Name: wpar1_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.29
There are 2 node(s) and 2 network(s) defined

NODE wpar1:
        Network net_ether_01
                wpar1sap        10.1.1.50
                wparsvc1        172.16.21.63
                wpar1   10.1.1.37
        Network net_ether_02
                wpar1b2 10.1.2.37salut

NODE wpar2:
        Network net_ether_01
```

```
            wpar1sap        10.1.1.50
            wparsvc1        172.16.21.63
            wpar2   10.1.1.45
        Network net_ether_02

Resource Group ApplA
        Startup Policy   Online On First Available Node
        Fallover Policy  Fallover To Next Priority Node In The List
        Fallback Policy  Never Fallback
        Participating Nodes      wpar1 wpar2
        Service IP Label                 wpar1sap
```

15. Modify the RG to be a WPAR RG by using the smit panel **cm_change_show_rg_resources**. Specify the service IP address (wpar1sap), the application controller name (ApplA), and the WPAR name (ApplA). We also specified the vary online of the volume group (Figure 8-7 on page 507).

```
Change/Show All Resources and Attributes for a Resource Group

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

                                                      [Entry Fields]
  Resource Group Name                                 ApplA
  Participating Nodes (Default Node Priority)         wpar1 wpar2

  Startup Policy                                      Online On First Avail>
  Fallover Policy                                     Fallover To Next Prio>
  Fallback Policy                                     Never Fallback

  Service IP Labels/Addresses                        [wpar1sap]             +
  Application Controllers                            [ApplA]                +

  Volume Groups                                      []             +
  Use forced varyon of volume groups, if necessary    true +
  Automatically Import Volume Groups                  true +

  Filesystems (empty is ALL for VGs specified)       [ ]                    +
  Filesystems Consistency Check                       fsck                  +
  Filesystems Recovery Method                         sequential            +
  Filesystems mounted before IP configured            false                 +

  Filesystems/Directories to Export (NFSv2/3)        []                     +
  Filesystems/Directories to Export (NFSv4)          []                     +
  Stable Storage Path (NFSv4)                        []                     +
  Filesystems/Directories to NFS Mount               []
  Network For NFS Mount                              []                     +

  Tape Resources                                     []                     +
  Raw Disk PVIDs                                     []                     +

  Primary Workload Manager Class                     []                     +
  Secondary Workload Manager Class                   []                     +

  Miscellaneous Data                                 []
  WPAR Name                                          [ApplA]                +
  User Defined Resources                             [ ]                    +

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

*Figure 8-7 cm_change_show_rg_resources smit panel configuration*

16. Verify and synchronize the configuration.
17. Start the cluster. By default, it starts your RG (Example 8-19).
18. Check that the WPAR is created. Check that the application is active by using the log file `/var/hacmp/adm/applA.log`.

*Example 8-19 clRGinfo when RG online*

```
                                COMMAND STATUS

Command: OK            stdout: yes           stderr: no
```

```
Before command completion, additional instructions may appear below.

[MORE...17]

Node                         State
---------------------------- ---------------
wpar1                        OFFLINE
wpar2                        OFFLINE

Resource Group Name: ApplA
Node                         State
---------------------------- ---------------
wpar1                        ONLINE
wpar2                        OFFLINE
```

19.Move the RG to the other node. The volume group must be varied online on the other node, the WPAR must be started, and the application must be running. Select the RG to move as shown in Figure 8-8.

```
 Resource Group and Applications

Move cursor to desired item and press Enter.

  Show the Current State of Applications and Resource Groups
  +--------------------------------------------------------------------------+
  |                         Select a Resource Group                          |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   #                                                                      |
  |   # Resource Group                 State               Node(s) / Site    |
  |   #                                                                      |
  |   ApplA                            ONLINE              wpar1 /           |
  |                                                                          |
  |   #                                                                      |
  |   # Resource groups in node or site collocation configuration:           |
  |   # Resource Group(s)                          State    Node / Site      |
  |   #                                                                      |
  |                                                                          |
  | F1=Help                F2=Refresh              F3=Cancel                 |
  | F8=Image               F10=Exit                Enter=Do                  |
F1| /=Find                 n=Find Next                                       |
F9+--------------------------------------------------------------------------+
```

*Figure 8-8 Select the RG to move*

20.Select the target note for the move as shown in Figure 8-9 on page 509.

```
 Resource Group and Applications

Move cursor to desired item and press Enter.

  Show the Current State of Applications and Resource Groups
  Bring a Resource Group Online
  Bring a Resource Group Offline
  Move Resource Groups to Another Node

  Suspend/Resume Application Monitoring
  Application Availability Analysis

  +--------------------------------------------------------------------------+
  |                        Select a Destination Node                         |
  |                                                                          |
  | Move cursor to desired item and press Enter.                             |
  |                                                                          |
  |   # *Denotes Originally Configured Highest Priority Node                 |
  |   wpar1                                                                  |
  |                                                                          |
  | F1=Help                 F2=Refresh              F3=Cancel                |
  | F8=Image                F10=Exit                Enter=Do                 |
F1| /=Find                  n=Find Next                                      |
F9+--------------------------------------------------------------------------+

                    Move Resource Group(s) to Another Node

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

                                                       [Entry Fields]
  Resource Group(s) to be Moved                      ApplA
  Destination Node                                   wpar1

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

*Figure 8-9 Selecting the node for the RG to be moved*

The result is a summary of the RG status as shown in Figure 8-10 on page 510.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

--------------------------------------------------------------------------------
-------------------------------------
Group Name     State                         Application state           Node

--------------------------------------------------------------------------------
-------------------------------------
ApplA          ONLINE                                                    wpar2

 ApplA                                        ONLINE NOT MONITORED
```

*Figure 8-10   Status of the WPAR-enabled RG after the move*

**WPAR address:** When the WPAR is handled by PowerHA and is *online,* it gets the address up and running that is associated to the WPAR:

```
# lswpar ApplA ; lswpar -N ApplA
Name   State  Type  Hostname  Directory     RootVG WPAR
--------------------------------------------------------
ApplA  A      S     ApplA     /wpars/ApplA  no

Name   Interface  Address(6)  Mask/Prefix    Broadcast
--------------------------------------------------------
ApplA  en0        10.1.1.50   255.255.255.0  10.1.1.255
```

When the Resource group is *offline,* the address is no longer associated to the WPAR. PowerHA internally remembers the association:

```
# clRGinfo
-----------------------------------------------------------------------------
Group Name     State                        Node
-----------------------------------------------------------------------------
ApplA          OFFLINE                      wpar1
               OFFLINE                      wpar2

root@wpar1 /var/hacmp/log # lswpar -N ApplA
lswpar: 0960-538 ApplA has no network configuration.
```

## 8.5 SAP scenario on AIX 7.1 NFS WPAR

We describe the use of the SAP environment as installed in Chapter 7, “IBM PowerHA SystemMirror Smart Assist for SAP” on page 355 within a WPAR.

That scenario illustrates the following information:

- NFS WPARs overview
- Specific commands to fit the SAP environment
- SAP installation

## 8.5.1 NFS WPARs overview

An NFS system WPAR, which is configured with its own file systems, is on one or more dedicated shared storage devices. A set of file systems that represents a rootvg volume must be created on the NFS server. More file systems can be added later.

Figure 8-11 shows the network configuration that includes the WPAR.

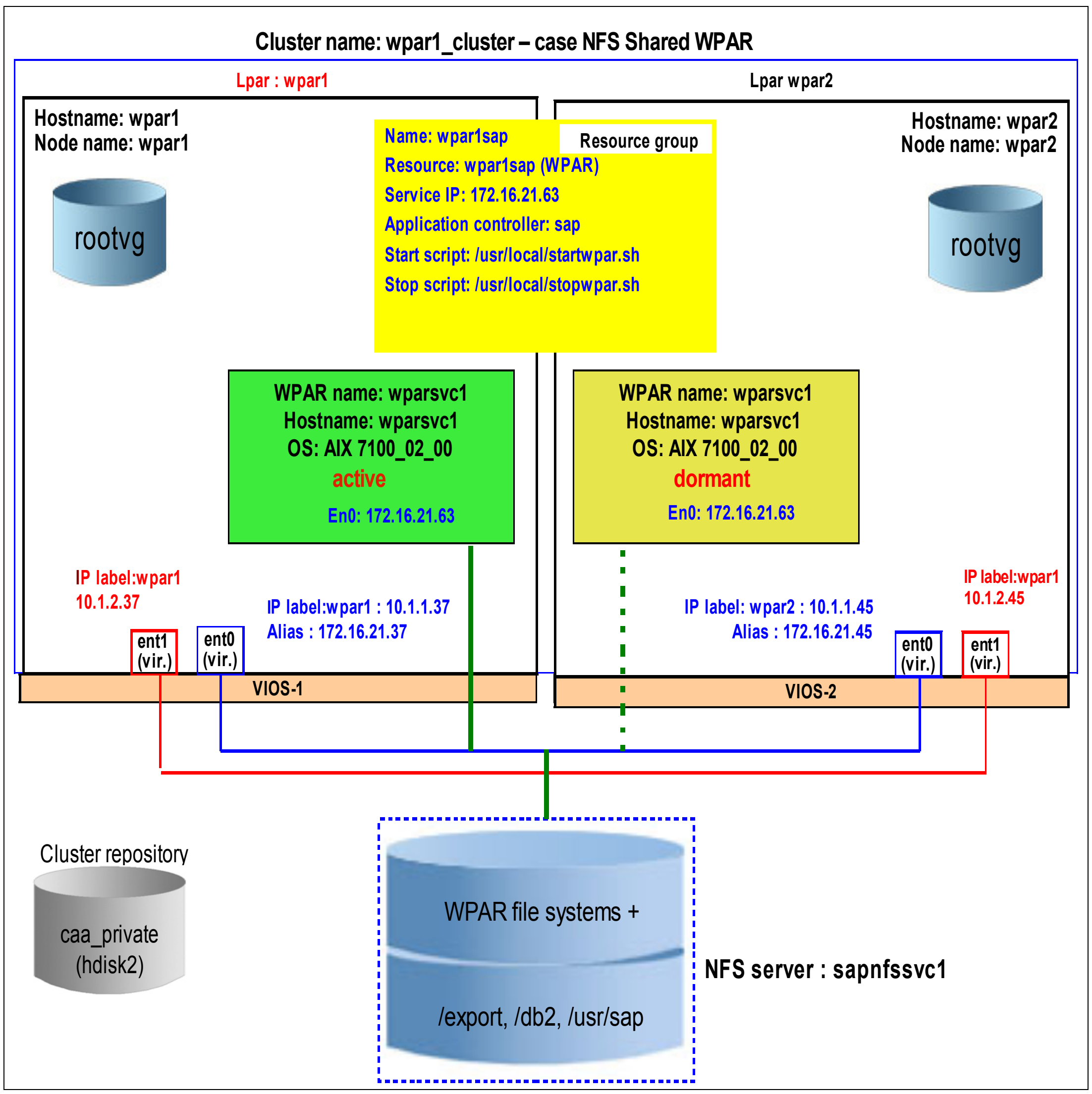



*Figure 8-11 Configuration with an NFS WPAR*

## 8.5.2 Specific commands to fit the SAP environment

Because we created the SAP environment on another cluster and we are cloning the environment, we have to consider a few areas to get our WPAR ready:

- Consider a shared or non-shared (private) WPAR
- Create a writable directory under a shared directory
- Make links
- Allocate addresses

### Consider a shared or non-shared (private) WPAR

By default, a system WPAR shares the `/usr` file system and the `/opt` file system from the global environment by using read-only namefs mounts. You can configure WPARs to have a non-shared, writable `/usr` file system and `/opt` file system.

In our SAP scenario (typical installation), we run multiple specific file systems that are named and mounted as shown in Example 8-20.

*Example 8-20 Set of file systems to be mounted within WPAR*

```
# df
Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
172.16.21.65:/install/wpar/wpars/wparsvc1/  309329920  93195536   70%    27138     1% /
172.16.21.65:/install/wpar/wpars/wparsvc1/db2  309329920  93195536   70%    27138     1%
/db2
172.16.21.65:/install/wpar/wpars/wparsvc1/export  309329920  93195536   70%    27138     1%
/export
172.16.21.65:/install/wpar/wpars/wparsvc1/home  309329920  93195536   70%    27138     1%
/home
Global           4194304   3810112   10%     7076     2% /opt
Global                 -         -    -         -     -  /proc
172.16.21.65:/install/wpar/wpars/wparsvc1/var/hacmp/adm  309329920  93195536  70%   27138
1% /var/hacmp/adm
Global           6291456   1798928   72%    51159    20% /usr
172.16.21.65:/install/wpar/wpars/wparsvc1/usrsap  309329920  93195536   70%    27138     1%
/usr/sap
Global            262144    235296   11%      358     2% /var
```

The `/usr` and `/opt` file systems can be shared with the Global environment.

### Make links

For our scenario, we need to create the `/usr/sap` directory under the Global environment. This `/usr/snap` directory is seen within the WPAR and can be overmounted within the WPAR.

> **Administrator:** File systems must be created correctly for the `/etc/filesystem` file to mount them in the correct order because multiple file system overmount themselves.

### Allocate addresses

The WPAR service address that is allocated to our network address, for example, `172.16.21.63`, is named `wparsvc1` in `/etc/hosts`. There is no need to create another specific address for the WPAR. The PowerHA environment uses its own set of addresses as explained in 3.5, “Networking” on page 83.

## 8.5.3 SAP installation

The SAP installation process is described at a high level in Chapter 7, “IBM PowerHA SystemMirror Smart Assist for SAP” on page 355.

In our scenario, we use the file systems that we created during that process and included them in a WPAR. So, we describe only the mandatory steps to make SAP work within the WPAR.

### Configuring I/O completion ports (IOCP) within the Global instance

Use the **smit iocp** fastpath from the Global instance to enable IOCP.

### Creating the WPAR

The WPAR creation is issued with the NFS disk by using the following specifications:

- Own service address of `172.16.21.63`
- Named `wpar1svc1` to match `/etc/host` entry for address `172.16.21.63`

The command line to create the WPAR is shown in Example 8-21. It also can be created by using the smit **wpar** fast path.

*Example 8-21 Simple SAP NFS WPAR creation*

```
#!/usr/bin/ksh
WPARNAME=wparsvc1
addr=172.16.21.63
NFSHOST=172.16.21.65
NFSROOT=/install/wpar/wpars

         # Mount the shared file system  located on NFS server as local
          mount $NFSHOST:/install  /install
          LNFSROOT=$NFSROOT

echo Creating wpar $WPARNAME with address $addr on server $NFSHOST in $NFSROOT

mkdir -p $LNFSROOT/$WPARNAME || error
chmod +x $LNFSROOT/$WPARNAME
mkdir $LNFSROOT/$WPARNAME/var
mkdir $LNFSROOT/$WPARNAME/var/hacmp/adm
mkdir $LNFSROOT/$WPARNAME/home
mkdir $LNFSROOT/$WPARNAME/opt
mkdir $LNFSROOT/$WPARNAME/usr
mkdir $LNFSROOT/$WPARNAME/usr/sap
mkdir $LNFSROOT/$WPARNAME/db2
mkdir $LNFSROOT/$WPARNAME/export

#!/usr/bin/ksh
WPARNAME=wparsvc1
addr=172.16.21.63
NFSHOST=172.16.21.65
NFSROOT=/install/wpar/wpars

# Mount the shared file system  located on NFS server as local
mount $NFSHOST:/install  /install
LNFSROOT=$NFSROOT
```

```
echo Creating wpar $WPARNAME with address $addr on server $NFSHOST in $NFSROOT

mkdir -p $LNFSROOT/$WPARNAME || error
chmod +x $LNFSROOT/$WPARNAME
mkdir $LNFSROOT/$WPARNAME/var
mkdir $LNFSROOT/$WPARNAME/var/hacmp/adm
mkdir $LNFSROOT/$WPARNAME/home
mkdir $LNFSROOT/$WPARNAME/opt
mkdir $LNFSROOT/$WPARNAME/usr
mkdir $LNFSROOT/$WPARNAME/usr/sap
mkdir $LNFSROOT/$WPARNAME/db2
mkdir $LNFSROOT/$WPARNAME/export

mkwpar -F -l  -r -N address=$addr interface=en0 interface=en0 -n
$WPARNAME \
-M directory=/ vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/  \
-M directory=/var vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/var  \
-M directory=/var/hacmp/adm vfs=nfs host=$NFSHOST
dev=$NFSROOT/$WPARNAME/var/hacmp/adm  \
-M directory=/home vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/home  \
-M directory=/usr vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/usr  \
-M directory=/opt vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/opt \
-M directory=/usr/sap vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/usr/sap  \
-M directory=/db2 vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/db2 \
-M directory=/export vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/export
```

When the WPAR is created, you can use the **lswpar** command to check for the disk and the main parameters as shown in Example 8-22.

*Example 8-22 Main lswpar command that is used to check creation*

```
root@wpar1 /install/wpar # lswpar -M
Name      MountPoint              Device                                Vfs
Nodename      Options
--------------------------------------------------------------------------------
-----------------------
wparsvc1  /wpars/wparsvc1         /install/wpar/wpars/wparsvc1/        nfs     172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/db2     /install/wpar/wpars/wparsvc1/db2     nfs     172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/export  /install/wpar/wpars/wparsvc1/export  nfs     172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/home    /install/wpar/wpars/wparsvc1/home    nfs     172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/opt     /opt                                 namefs                ro
wparsvc1  /wpars/wparsvc1/proc    /proc                                namefs                rw
wparsvc1  /wpars/wparsvc1/var/hacmp/adm     /install/wpar/wpars/wparsvc1/var/hacmp/adm    nfs    172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/usr     /usr                                 namefs                ro
wparsvc1  /wpars/wparsvc1/usr/sap  /install/wpar/wpars/wparsvc1/usrsap  nfs     172.16.21.65  rw
wparsvc1  /wpars/wparsvc1/var     /dev/fslv00                          jfs2
root@wpar1 /install/wpar # lswpar -G
=================================================================
wparsvc1 - Active
=================================================================
Type:                    S
RootVG WPAR:             no
Owner:                   root
Hostname:                wparsvc1
WPAR-Specific Routing:   no
Virtual IP WPAR:
Directory:               /wpars/wparsvc1
Start/Stop Script:
```

```
Auto:                   no
Private /usr:           no
Checkpointable:         no
Application:

OStype:                 0
UUID:                   68f2fcdc-ca7e-4835-b9e1-9edaa2a36dcd
```

### Steps to configure the SAP within the WPAR

Follow these steps to configure SAP within a WPAR:

1. Create the proper user as seen in the SAP reference system. Example 8-23 on page 515 shows the full script to create users for SAP:
   - `sapsr3` - in `/usr/sap/TE1/home/sapsr3`
   - `db2te1` - in `/db2/db2te1`
   - `sapadm` - in `/usr/sap/TE1/homesapadm`
   - `da1adm` - in `/usr/sap/TE1/home/da1adm`
   - `te1adm` is created by the installation process in `/usr/sap/TE1/home/te1adm`.

*Example 8-23 Full script to create users for SAP*

```
mkgroup -A id=300 sapsys
mkgroup -A id=301 sapinst
mkgroup -A id=303 dbte1ctl
mkgroup -A id=304 dbte1mnt
mkgroup -A id=305 dbte1adm
mkgroup -A id=306 dbte1mon
mkuser id=300 pgrp=dbte1ctl shell='/bin/csh' groups='dbte1ctl,staff'
home='/home/temp' gecos='SAP Database Administrator' db2te1
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' db2te1
chuser home='/db2/db2te1' db2te1
echo db2te1:A300B0 | chpasswd -c
mkuser id=310 pgrp=sapsys shell='/bin/csh' groups='sapsys,staff,sapinst,dbte1ctl'
home='/home/temp' gecos='SAP System Administrator' te1adm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' te1adm
chuser home='/usr/sap/TE1/home/te1adm' te1adm
echo te1adm:A310B0 | chpasswd -c
mkuser id=305 pgrp=sapsys shell='/bin/csh' groups='sapsys,dbte1ctl'
home='/home/temp' gecos='ABAP Database Connect User' sapsr3
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapsr3
chuser home='/usr/sap/TE1/home/sapsr3' sapsr3
echo sapsr3:A305B0 | chpasswd -c
mkuser id=320 pgrp=sapsys shell='/bin/csh' groups='sapsys,sapinst'
home='/home/temp' gecos='SAP System Administrator' sapadm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapadm
chuser home='/usr/sap/TE1/home/sapadm' sapadm
echo sapadm:A320B0 | chpasswd -c
mkuser id=323 pgrp=sapsys shell='/bin/csh' groups='sapsys,sapinst'
home='/home/temp' gecos='SAP System Administrator' da1adm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' da1adm
chuser home='/usr/sap/TE1/home/da1adm' da1adm
echo da1adm:A323B0 | chpasswd -c
chgroup users=sapsr3 dbte1ctl
```

**Important:** To change the user, you need to set the `CLUSTER_OVERRIDE=yes` variable.

The resulting `/etc/password` is as shown in Example 8-24.

*Example 8-24 /etc/password entry for SAP users*

```
db2te1:!:300:305:SAP Database Administrator:/db2/db2te1:/bin/csh
te1adm:!:310:300:SAP System Administrator:/usr/sap/TE1/home/te1adm:/bin/csh
sapsr3:!:305:300:ABAP Database Connect User:/usr/sap/TE1/home/sapsr3:/bin/csh
sapadm:!:320:300:SAP System Administrator:/home/sapadm:/bin/csh
da1adm:!:323:300:SAP System Administrator:/home/da1adm:/bin/csh
```

2. Create the correct file directories by using the script that is shown in Example 8-25.

*Example 8-25 Script to create home user directories*

```
mkdir -p /usr/sap/TE1/home/sapadm
mkdir -p /usr/sap/TE1/home/da1adm
chown -R te1adm /usr/sap/TE1/sapadm
chown -R te1adm /usr/sap/TE1/da1adm
mkdir -p /db2/TE1/home/db2te1
mkdir -p /usr/sap/TE1/home/te1adm
chown -R db2te1 /db2/TE1
chown -R te1adm /usr/sap/TE1
```

3. Create the following links:
   - `cd /usr/sap`; ln `-s /export/saptrans trans`
   - `mkdir /sapmnt`; `cd /sapmnt`; ln `-s /export/sapmnt/TE1 TE1`
   - `cd /db2`; ln `-s TE1/db2te1 db2te1`
   - Copy the following file systems from the SAP installation to the WPAR:
     - `/usr/sap/TE1`
     - `/usr/sap/DA1`
     - `/usr/sap/ccms`
     - `/usr/sap/hostctrl`
     - `/usr/sap/sapservices`
     - `/usr/sap/trans`
     - `/export/saptrans`
     - `/export/sapmnt`
     - `/db2/TE1`
     - `/var/db2`
     - `/etc/rc.d/rc2.d/*sap*`
     - `/etc/services`
     - `/home/sapadm`
     - `/home/da1adm`
4. Ensure that the profiles are up-to-date with the WPAR name `wparsvc1`. The content and names must reflect the current host name in the following directories (Example 8-26):
   - `/sapmnt/TE1/profile`

*Example 8-26 Host name change in /sapmnt/TE1/profile*

```
root@wparsvc1 /sapmnt/TE1/profile # ls
DEFAULT.1.PFL             START_DVEBMGS02_wpar1svc
DEFAULT.PFL               TE1_DVEBMGS02_wpar1svc
root@wparsvc1 /sapmnt/TE1/profile # grep wparsvc1 *
DEFAULT.1.PFL:SAPDBHOST = wparsvc1
DEFAULT.1.PFL:j2ee/dbhost = wparsvc1
DEFAULT.1.PFL:SAPGLOBALHOST = wparsvc1
DEFAULT.1.PFL:rdisp/mshost = wparsvc1
DEFAULT.PFL:SAPDBHOST = wparsvc1
DEFAULT.PFL:j2ee/dbhost = wparsvc1
DEFAULT.PFL:SAPGLOBALHOST = wparsvc1
DEFAULT.PFL:rdisp/mshost = wparsvc1
START_DVEBMGS02_wpar1svc:SAPLOCALHOST = wparsvc1
START_DVEBMGS02_wpar1svc:_PF = $(DIR_PROFILE)/TE1_DVEBMGS02_wpar1svc
TE1_DVEBMGS02_wpar1svc:SAPLOCALHOST = wparsvc1
```

- //usr/sap/TE1/home/te1adm

5. Modify the DB2 configuration files to match the WPAR name:
   - /db2/TE1/db2te1/sqllib/db2nodes.cfg
   - /usr/sap/TE1/SYS/global/db6/db2cli.ini

## Starting SAP

To start the SAP environment, you need to be user `te1adm` that you created earlier. It executes its own profile and you can start, monitor, and stop SAP. Example 8-27 shows starting SAP by using the **`startsap wparsvc1`** command.

*Example 8-27 Output of SAP start*

```
# su - te1adm
wparsvc1:te1adm 1> startsap wparsvc1

Checking db6 db Database
------------------------------
 Database is not available via R3trans
 Running /usr/sap/TE1/SYS/exe/run/startdb
03/22/2012 11:18:57     0   0   SQL1063N  DB2START processing was successful.
SQL1063N  DB2START processing was successful.
Database activated
*** WARNING Logretain mode is disabled
    You will be unable to fully recover your database in case
    of a media failure
 /usr/sap/TE1/SYS/exe/run/startdb completed successfully

Starting Startup Agent sapstartsrv
-----------------------------
OK
 Instance Service on host wparsvc1 started

starting SAP Instance DVEBMGS02
------------------------------
 Startup-Log is written to /usr/sap/TE1/home/te1adm/startsap_DVEBMGS02.log
/usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02 -function Start
 Instance on host wparsvc1 started
```

## Check that SAP works correctly

To check that SAP works, we can use actual applications. For our purposes, we use SAP commands to check the status of SAP and to verify that it is running. Example 8-28 shows the output of the SAP command.

*Example 8-28 Checking the status of the SAP instance*

```
wparsvc1:te1adm 2> /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02
-function GetProcessList

22.03.2012 11:27:23
GetProcessList
OK
name, description, dispstatus, textstatus, starttime, elapsedtime, pid
msg_server, MessageServer, GREEN, Running, 2012 03 22 11:21:59, 0:05:24, 1572936
disp+work, Dispatcher, GREEN, Running, Message Server connection ok, Dialog Queue
time: 0.00 sec, 2012 03 22 11:21:59, 0:05:24, 2097170
igswd_mt, IGS Watchdog, GREEN, Running, 2012 03 22 11:21:59, 0:05:24, 2621502
```

The **GREEN** indicator that is shown in Example 8-28 on page 518 indicates that our SAP instance runs within the WPAR.

## 8.5.4 Setting the cluster

Because we created the cluster earlier, we only need to create an RG that is called wparsvc1 (name of the WPAR and the name of the RG must be identical).

To see how to create a simple cluster, see Chapter 3, “PowerHA 7.1.1 basic installation and configuration” on page 73.

All commands can be issued by using the usual tools, such as **smit**, **clmgr** on the command line, or the System Director plug-in. In some cases, where the command line is usable, it is listed as an option.

We create a cluster with two nodes and one repository disk, and we create the appropriate RG for the PowerHA function with the WPAR.

We highlight specific tasks:

- Create the sap application controller scripts by using the **cm_add_app_scripts** and check that they can execute in both nodes. Examples of these scripts are shown in Example 8-29.

*Example 8-29 Example of start and stop scripts*

```
root@wpar1 /usr/local # cat startwpar.sh
#!/usr/bin/ksh93
if df |grep -w /usr/sap
then
        mount /usr/sap
fi
if [[ ! -d /usr/sap/TE1/home/te1adm ]]
then
        print "te1adm home directory doesn't exist"
        exit 1
fi
```

```
su - te1adm -c startsap wpar1sap
su - te1adm -c /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02
-function GetProcessList
su - te1adm -c /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02
-function GetProcessList |grep -w GREEN
rc=$?
if [[ $rc == 0 ]]
then
        print "SAP started and is GREEN"
fi
return $rc
#-----------------------------------------------#
root@wpar1 /usr/local # cat stopwpar.sh
#!/usr/bin/ksh93
savebase
export CLUSTER_OVERRIDE=yes
su - te1adm -c stopsap wpar1sap
su - te1adm -c /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot NI_HTTP -nr 02
-function GetProcessList
umount -f /export
umount -f /db2
umount -f /usr/sap
return 0
```

- Add the scripts for starting and stopping the WPAR by using **`cm_cludrestype_add`** (Example 8-30). Check by using the **`cllsserv`** command.

*Example 8-30 cm_cludrestype_add menu*

```
Add a User Defined Resource Type

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

                                                    [Entry Fields]
* Resource Type Name                              [wparsvc1]
* Processing order                                [WPAR]                 +
  Verification Method                             []
    Verification Type                             [Script]               +
* Start Method                                    [/usr/local/starwpar.sh]
* Stop Method                                     [/usr/local/stowpar.sh]
  Monitor Method
[/usr/local/monitorwpar.sh]
  Cleanup Method                                  []
  Restart Method
[/usr/local/restartwapr.sh]
  Failure Notification Method                     []
  Required Attributes                             []
  Optional Attributes                             []
  Description                                     [wpar]


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

- Add a service label for the WPAR address `10.1.1.50` by using the command line or smit **`cm_add_a_service_ip_label_address.select_net`**(Example 8-31).

*Example 8-31 Add a service label*

```
/usr/es/sbin/cluster/utilities/clmgr add service_ip 'wparsvc1' NETMASK |
  |   ='255.255.254.0' NETWORK='net_ether_01'
```

- Add the RG `wparsvc1` by using smit **`cm_add_resource_group`**. Check the output by using the **`cllsclstr`** command that is shown in Example 8-32.

*Example 8-32 Cllsclstr output*

```
root@wpar1 /usr/local # /usr/es/sbin/cluster/utilities/cllsclstr
ID      Name    Security        Persist Repository      Cluster IP Address

1088917986      wpar1_cluster   Standard                hdisk2  228.1.1.29
```

- Modify the RG to be a WPAR RG by using the smit panel **`cm_change_show_rg_resources`** (Example 8-33 on page 520).

*Example 8-33 cm_change_show_rg_resources smit panel configuration*

```
Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
  Resource Group Name                                 wparsvc1
  Participating Nodes (Default Node Priority)         wpar1 wpar2

  Startup Policy                                      Online On Home Node O>
  Fallover Policy                                     Fallover To Next Prio>
  Fallback Policy                                     Never Fallback

  Service IP Labels/Addresses                        [10.1.1.50]          +
  Application Controllers                            [sap]        +

  Volume Groups                                      []         +
  Use forced varyon of volume groups, if necessary    false +
  Automatically Import Volume Groups                  false +

  Filesystems (empty is ALL for VGs specified)       []                   +
  Filesystems Consistency Check                       fsck                +
  Filesystems Recovery Method                         sequential          +
  Filesystems mounted before IP configured            false               +

  Filesystems/Directories to Export (NFSv2/3)        []                   +
  Filesystems/Directories to Export (NFSv4)          []                   +
  Stable Storage Path (NFSv4)                        []                   +
  Filesystems/Directories to NFS Mount               []
  Network For NFS Mount                              []                   +

  Tape Resources                                     []                   +
  Raw Disk PVIDs                                     []                   +

  Primary Workload Manager Class                     []                   +
```

```
Secondary Workload Manager Class                    []                 +

Miscellaneous Data                                  []
WPAR Name                                           [wparsvc1]         +
User Defined Resources                              [ ]                +

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

- Start the cluster by using the **/usr/es/sbin/cluster/utilities/clstart** command or smitty.
- Bring the resource group online.
- Check failover to the other node. The WPAR restarts.

### 8.5.5 Using the command line to create the cluster

To reproduce the scenario, we list the commands that are required to install, configure, and test the cluster as shown in Example 8-34.

*Example 8-34 Short commands to create the cluster*

```
# /usr/es/sbin/cluster/utilities
ssh wpar1 ps
ssh wpar2 ps
/usr/es/sbin/cluster/utilities/cl_rsh wpar1 p
/usr/es/sbin/cluster/utilities/cl_rsh wpar2 ps
clmgr add cluster wpar1_cluster NODES=’’
clmgr modify cluster wpar1_cluster REPOSITORY=hdisk2 CLUSTER_IP='228.1.1.29'

clmgr -f add node 'wpar2' COMMPATH='10.1.1.45'

clmgr add interface 'wpar1p1' TYPE='ether' NETWORK='net_ether_01' NODE='wpar1'
INTERFACE='en0'

clmgr add network net_ether_01 TYPE=ether '255.255.254.0'
clmgr add interface 'wpar1' TYPE='ether' NETWORK='net_ether_01' NODE='10.1.1.37'
clmgr add interface 'wpar2' TYPE='ether' NETWORK='net_ether_01' NODE='10.1.1.45'

cltopinfo -w
claddserv -s'sap' -b'/usr/local/startwpar.sh' -e'/usr/local/stopwpar.sh' -O 'foreground'
clmgr add service_ip 'wparsvc1' NETWORK='net_ether_01'
claddgrp -g 'wparsvc1' -n 'wpar1' -S 'OHN' -O 'FNPN' -B 'NFB'
/usr/es/sbin/cluster/utilities/claddres -g 'wparsvc1' SERVICE_LABEL='wparsvc1' A
PPLICATIONS='sap' VOLUME_GROUP= FORCED_VARYON='false' VG_AUTO_IMPORT='false' FIL
ESYSTEM= FSCHECK_TOOL='fsck' RECOVERY_METHOD='sequential' FS_BEFORE_IPADDR='fals
e' EXPORT_FILESYSTEM= EXPORT_FILESYSTEM_V4= STABLE_STORAGE_PATH= MOUNT_FILESYSTE
M= NFS_NETWORK= SHARED_TAPE_RESOURCES= DISK= MISC_DATA= WPAR_NAME='wparsvc1' USE
RDEFINED_RESOURCES=
clmgr sync cluster wpar1_cluster
cldare -t
clmgr start cluster wpar1_cluster
/usr/es/sbin/cluster/utilities/cldare -rt -V 'normal' -C'interactive'
clRGinfo
clshowsrv -av
lswpar
# In case setting disabled auto start
```

```
# clmgr start resource_group wparsvc1
# clRGinfo
# lswpar

# Check application has started ok
clogin wparsvc1 " su - te1adm -c /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot
NI_HTTP -nr 02 -function GetProcessList"
#Move resource_group to second node
clmgr move resource_group wparsvc1

# Check application has started ok
clogin wparsvc1 " su - te1adm -c /usr/sap/TE1/DVEBMGS02/exe/sapcontrol -prot
NI_HTTP -nr 02 -function GetProcessList"
```

## 8.6 NFS versioned 5.2 WPAR

This scenario uses an NFS versioned WPAR with PowerHA. Check the planning information in 8.2.5, "Planning for a versioned WPAR" on page 492. We use a **mksysb** image of an AIX 5.2 release.

The configuration of the network is similar to the previous NFS scenario. Only the WPAR changed because it moved from a standard shared WPAR to a private NFS versioned WPAR as seen in figure Figure 8-12 on page 524.

### Application to run in the WPAR

The application can be the same as in 8.4, "Scenario with a local WPAR" on page 497. The start and stop scripts are shown in Example 8-35.

*Example 8-35 Sample start and stop controller scripts for application ApplA*

```
# cat /usr/local/ha/StartA
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#
Name=$(basename $0 )

if [ "$Name" = "StartA" ]
then
  echo "$(date) \"Application A started\" " >> /var/hacmp/adm/applA.log
  touch /var/hacmp/adm/AppAup
  nohup /usr/local/bin/ApplA &
  exit 0
elif [ "$Name" = "StopA" ]
then
  rm -f /var/hacmp/adm/AppAup
  echo "$(date) \"Application A stopped\" " >> /var/hacmp/adm/applA.log
  exit 0
else
  echo "$(date) \"ERROR - Application A start/stop script called with wrong
name\" " >> /var/hacmp/adm/applA.log
  exit 999
fi
```

```
#--------------------------------------------------#
# cat /usr/local/ha/StopA
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#
Name=$(basename $0 )

if [ "$Name" = "StartA" ]
then
  echo "$(date) \"Application A started\" " >> /var/hacmp/adm/applA.log
  touch /var/hacmp/adm/AppAup
  nohup /usr/local/bin/ApplA &
  exit 0
elif [ "$Name" = "StopA" ]
then
  rm -f /var/hacmp/adm/AppAup
  echo "$(date) \"Application A stopped\" " >> /var/hacmp/adm/applA.log
  exit 0
else
  echo "$(date) \"ERROR - Application A start/stop script called with wrong
name\" " >> /var/hacmp/adm/applA.log
  exit 1
fi
```

The user application is shown in Example 8-36.

*Example 8-36 Sample user application ApplA*

```
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x

while [ -a /var/hacmp/adm/AppAup ]
do
  echo "$(date) \"Application A is running\" " >> /var/hacmp/adm/applA.log
  sleep 10
done
```

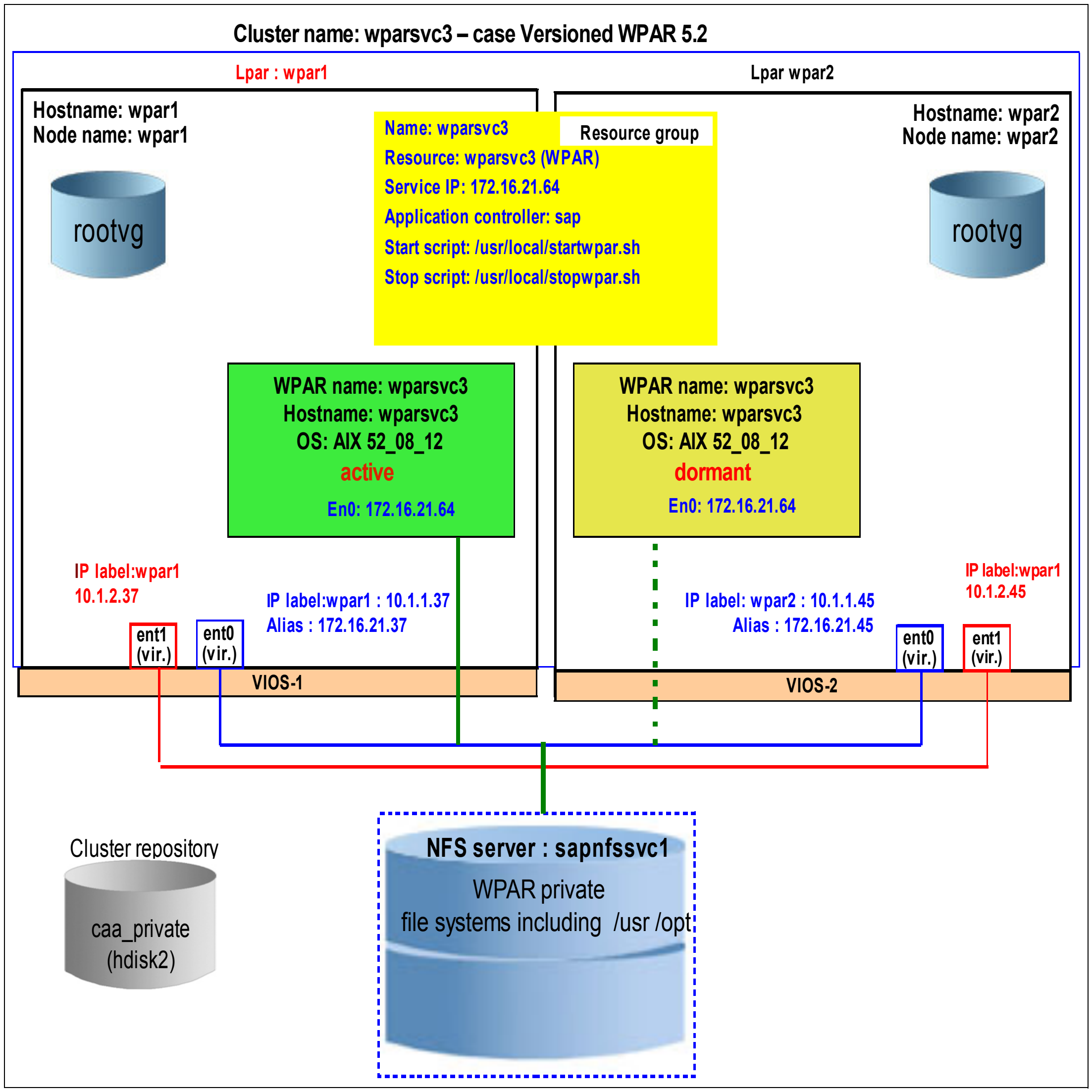



*Figure 8-12 Configuration with NFS versioned 5.2 WPAR*

## Creating the WPAR

The WPAR is created on an NFS server that uses the following specs:

- It has its own service address of `172.16.21.64`.
- The server is named `wpar1svc3` to match `/etc/host` entry for address `172.16.21.64`.
- The **`mksysb`** image is called `AIX52New.mksysb`.

The command-line script to create the WPAR is shown in Example 8-37. Or, you can use the smit wpar fast path.

*Example 8-37 Simple SAP NFS WPAR creation*

```
#!/usr/bin/ksh

WPARNAME=wparsvc3
addr=172.16.21.64
NFSHOST=172.16.21.65
NFSROOT=/install/wpar/wpars
```

```
MKSYSB=/install/wpar/AIX52New.mksysb

#Local mount the shared file system located on NFS server
mount $NFSHOST/install /install
LNFSROOT=$NFSROOT

echo Creating wpar $WPARNAME with address $addr on server $NFSHOST in $NFSROOT
from mksysb $MKSYSB

# location of NFS wpar

NFSVERS=3

# definition for private /usr and /opt
PRIVATE_FLAG=-l
NFS_EXTRA_PRIVATE="\
-M directory=/usr vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/usr
mountopts=vers=$NFSVERS \
-M directory=/opt vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/opt
mountopts=vers=$NFSVERS"

mkdir -p $LNFSROOT/$WPARNAME || error
chmod +x $LNFSROOT/$WPARNAME
mkdir $LNFSROOT/$WPARNAME/var
mkdir $LNFSROOT/$WPARNAME/var/hacmp/adm
mkdir $LNFSROOT/$WPARNAME/home
mkdir $LNFSROOT/$WPARNAME/opt
mkdir $LNFSROOT/$WPARNAME/usr

mkwpar -F $PRIVATE_FLAG -r -N address=$addr interface=en0 interface=en0 -n
$WPARNAME \
 -M directory=/ vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/
mountopts=vers=$NFSVERS \
-M directory=/var vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/var
mountopts=vers=$NFSVERS  \
-M directory=/var/hacmp/adm vfs=nfs host=$NFSHOST
dev=$NFSROOT/$WPARNAME/var/hacmp/adm mountopts=vers=$NFSVERS  \
-M directory=/home vfs=nfs host=$NFSHOST dev=$NFSROOT/$WPARNAME/home
mountopts=vers=$NFSVERS  \
$NFS_EXTRA_PRIVATE  \
-C -B $MKSYSB
```

The major difference is the use of parameters "**-C -B -l**". For more explanation, see *Exploiting IBM AIX Workload Partitions*, SG24-7955.

When the WPAR is created, you can use the **lswpar** command to check for the disk and the main parameters as presented in Example 8-38.

*Example 8-38 Main lswpar command that is used to check the WPAR creation*

```
# lswpar -M wparsvc3
Name      MountPoint              Device                              Vfs     Nodename
Options
-----------------------------------------------------------------------------------------
-------------
```

```
wparsvc3  /wpars/wparsvc3          /install/wpar/wpars/wparsvc3/      nfs
172.16.21.65  vers=3
wparsvc3  /wpars/wparsvc3/home      /install/wpar/wpars/wparsvc3/home  nfs
172.16.21.65  vers=3
wparsvc3  /wpars/wparsvc3/nre/opt   /opt  namefs               ro
wparsvc3  /wpars/wparsvc3/nre/sbin  /sbin namefs               ro
wparsvc3  /wpars/wparsvc3/nre/usr   /usr  namefs               ro
wparsvc3  /wpars/wparsvc3/opt       /install/wpar/wpars/wparsvc3/opt   nfs
172.16.21.65  vers=3
wparsvc3  /wpars/wparsvc3/proc      /proc namefs               rw
wparsvc3  /wpars/wparsvc3/var/hacmp/adm     /install/wpar/wpars/wparsvc3/var/hacmp/adm
nfs     172.16.21.65  vers=3
wparsvc3  /wpars/wparsvc3/usr       /install/wpar/wpars/wparsvc3/usr   nfs
172.16.21.65  vers=3
wparsvc3  /wpars/wparsvc3/var       /install/wpar/wpars/wparsvc3/var   nfs
172.16.21.65  vers=3


# # lswpar -G wparsvc3
=================================================================
wparsvc3 - Defined
=================================================================
Type:                      S
RootVG WPAR:               no
Owner:                     root
Hostname:                  wparsvc3
WPAR-Specific Routing:     no
Virtual IP WPAR:
Directory:                 /wpars/wparsvc3
Start/Stop Script:
Auto:                      no
Private /usr:              yes
Checkpointable:            no
Application:

OStype:                    1
UUID:                      2767381f-5de7-4cb7-a43c-5475ecde54f6

# df |grep wparsvc3
172.16.21.65:/install/wpar/wpars/wparsvc3/  309329920  39976376   88%   118848     3%
/wpars/wparsvc3
172.16.21.65:/install/wpar/wpars/wparsvc3/home  309329920  39976376   88%   118848
3% /wpars/wparsvc3/home
172.16.21.65:/install/wpar/wpars/wparsvc3/opt  309329920  39976376   88%   118848     3%
/wpars/wparsvc3/opt
/proc                  -          -     -          -     -  /wpars/wparsvc3/proc
172.16.21.65:/install/wpar/wpars/wparsvc3/var/hacmp/adm  309329920  39976376   88%
118848     3% /wpars/wparsvc3/var/hacmp/adm
172.16.21.65:/install/wpar/wpars/wparsvc3/usr  309329920  39976376   88%   118848     3%
/wpars/wparsvc3/usr
172.16.21.65:/install/wpar/wpars/wparsvc3/var  309329920  39976376   88%   118848     3%
/wpars/wparsvc3/var
/usr            6291456   1787184   72%    50979    20% /wpars/wparsvc3/nre/usr
/opt            4194304   3809992   10%     7078     2% /wpars/wparsvc3/nre/opt
/sbin           2097152   1566912   26%    11155     6%
/wpars/wparsvc3/nre/sbin
```

Because we created the WPAR, it is possible to create it on the other system by using the **mkwpar -pf** command.

### Creating the WPAR on the second node

Create the specfile on the first node and create the WPAR on the second node by using the specfile that you created. The commands are listed in Example 8-39.

*Example 8-39 Create the WPAR by using the specfile*

```
mkwpar -w -e wparsvc3 -o /install/wpar/CFG/wparsvc3.spec

# ON THE OTHER NODE
mkwpar -pf /install/wpar/CFG/wparsvc3.spec
   **********************************************************************
   Warning
   mkwpar: 0960-125 network.address: 172.16.21.64/255.255.254.0 is not in the same
   network as any of the global interfaces.

   **********************************************************************
   mkwpar: Creating file systems...
    /
    /home
    /nre/opt
    /nre/sbin
    /nre/usr
    /opt
    /proc
    /var/hacmp/adm
    /usr
    /var
   Workload partition wparsvc3 created successfully.
   mkwpar: 0960-390 To start the workload partition, execute the following as
   root: startwpar [-v] wparsvc3
```

WPAR `wparsvc3` is now defined and running in both nodes.

## 8.6.1 Creating the resource group

For more details about the RG creation, see 8.5.4, “Setting the cluster” on page 518.

Follow these steps:

1. Create the RG `wparsvc3` (Figure 8-13 on page 528).

```
Add a Resource Group

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

                                                        [Entry Fields]
* Resource Group Name                                  [wparsvc3]
* Participating Nodes (Default Node Priority)          [wpar1 wpar2]          +

  Startup Policy                                        Online On Home Node O> +
  Fallover Policy                                       Fallover To Next Prio> +
  Fallback Policy                                       Never Fallback         +

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

*Figure 8-13 Create RG wparsvc3*

2. Create the associated service IP address (Figure 8-14).

```
Add a Service IP Label/Address

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

                                                        [Entry Fields]
* IP Label/Address                                      wparsvc3               +
  Netmask(IPv4)/Prefix Length(IPv6)                    []
* Network Name                                          net_ether_01

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

*Figure 8-14 Create the service IP address wparsvc3*

3. Modify the RG to make it WPAR-enabled by using the same application controller as shown in Example 8-40 with the service address `wparsvc3`. Both RGs are independently started. You might want to create a separate controller group and have specific scripts for each WPAR.

*Example 8-40 Enable the RG as WPAR-enabled*

```
Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
  Resource Group Name                                   wparsvc3
  Participating Nodes (Default Node Priority)           wpar1 wpar2

  Startup Policy                                        Online On Home Node O>
  Fallover Policy                                       Fallover To Next Prio>
  Fallback Policy                                       Never Fallback
```

```
  Service IP Labels/Addresses                          [wparsvc3]          +
  Application Controllers                              [ApplA]             +

  Volume Groups                                        []                  +
  Use forced varyon of volume groups, if necessary      false              +
  Automatically Import Volume Groups                    false              +

  Filesystems (empty is ALL for VGs specified)         []                  +
  Filesystems Consistency Check                         fsck               +
  Filesystems Recovery Method                           sequential         +
  Filesystems mounted before IP configured              false              +

  Filesystems/Directories to Export (NFSv2/3)          []                  +
  Filesystems/Directories to Export (NFSv4)            []                  +
  Stable Storage Path (NFSv4)                          []                  +
  Filesystems/Directories to NFS Mount                 []
  Network For NFS Mount                                []                  +

  Tape Resources                                       []                  +
  Raw Disk PVIDs                                       []                  +

  Primary Workload Manager Class                       []                  +
  Secondary Workload Manager Class                     []                  +

  Miscellaneous Data                                   []
  WPAR Name                                            [wparsvc3]          +
  User Defined Resources                               [ ]                 +


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

4. Verify and synchronize the cluster configuration.
5. Bring the RG online.
6. Check the application output.
7. Move the RG to the other node.

A

# AIX upgrades and PowerHA SystemMirror 7.1.0

This appendix provides a supported upgrade procedure of the AIX operating system from AIX 7.1 Technology Level (TL) 00 to AIX 7.1 TL01 (or from AIX 6.1 TL06 to AIX 6.1 TL07). The procedure must be used when you have a PowerHA SystemMirror 7.1.0 cluster that is deployed in your AIX environment.

Because AIX 7.1 TL01 and AIX 6.1 TL07 technology levels modified Cluster Aware AIX (CAA) significantly, the CAA cluster must be removed before the AIX upgrade and recreated after the upgrade. Finally, PowerHA remains at version 7.1.0, but after the operating system upgrade is running on top of the new AIX 7.1 TL01 or AIX 6.1 TL07. See the detailed procedure steps in “Upgrading AIX” on page 532.

The last section of this appendix,“Upgrading to PowerHA 7.1.1” on page 544, presents our results of using a simple procedure by which we are able to successfully upgrade the PowerHA 7.1.0 already running on top of AIX 7.1 TL1 to PowerHA 7.1.1.

# Upgrading AIX

Use the following steps to upgrade AIX while you keep PowerHA at 7.1.0 level:

1. Prepare the AIX upgrade:
   a. Save a cluster snapshot, make the usual backups (data, binaries, and **mksysb**), and prepare a back-out plan.
   b. Check software consistency and ensure that the cluster state and configuration are error-free on all nodes.
2. Perform the AIX upgrade:
   a. Apply the latest AIX Service Pack for the running version.
   b. Stop PowerHA cluster services on all nodes.
   c. Remove the CAA cluster.
   d. Upgrade the AIX on all nodes in the cluster.
   e. Recreate the CAA cluster from PowerHA SystemMirror.
   f. Restart the PowerHA cluster services.

## Preparing the AIX upgrade

Plan to stop the cluster and perform the following steps to update the AIX operating system technology levels (AIX 6.1 TL07, AIX 7.1 TL01, or later AIX releases). The cluster configuration information is retained and reused. You do not need to reenter it.

Any set of actions that involves modifications at the operating system or cluster level must be preceded by back-out (reversion) planning, including the following operations:

- Back up the data and binaries of all applications.
- Take a cluster snapshot, save it locally, and save it to another machine.
- Save a copy of any custom script files locally and to another machine.
- Perform a **mksysb** backup of each involved node.

With the back-out plan, you can restore the application data and binaries easily and the cluster and AIX configuration in case you run into problems.

Then, perform the usual software consistency checks:

- Ensure that all nodes are at the same level of operating system and cluster software. Check that the cluster software is committed (and not merely applied). Use the **oslevel -s** and **lslpp -L cluster*** commands.
- Verify that the software is consistent on all nodes by using the **lppchk -v** command.

Before you start the actual upgrade, check the overall status of your cluster:

1. Ensure that the cluster is started on all nodes, stable, and in a normal operating state. Run the **clmgr view report status** command on either node and check the state of the cluster, resource groups, interfaces, and System Resource Controller (SRC) subsystems (Example A-1).

*Example A-1 Initial cluster state*

```
root@migr5 / # clmgr view report status

Obtaining information via SNMP from Node: migr5...
```

```
Cluster Name: clmigr56
Cluster State: UP
Cluster Substate: STABLE
```

```
Node Name: migr5              State: UP

  Network Name: net_ether_01      State: UP

    Address: 172.16.21.75    Label: migr5            State: UP
    Address: 172.16.21.77    Label: migrsvc5         State: UP

Node Name: migr6              State: UP

  Network Name: net_ether_01      State: UP

    Address: 172.16.21.76    Label: migr6            State: UP



Cluster Name: clmigr56

Resource Group Name: ihs_rg
Startup Policy: Online On First Available Node
Fallover Policy: Fallover To Next Priority Node In The List
Fallback Policy: Never Fallback
Site Policy: ignore
Node                         Group State
---------------------------- ---------------
migr5                        OFFLINE
migr6                        ONLINE


Status of the RSCT subsystems used by HACMP:
Subsystem         Group            PID          Status
 cthags           cthags           16384048     active
 ctrmc            rsct             5439686      active

Status of the HACMP subsystems:
Subsystem         Group            PID          Status
 clstrmgrES       cluster          15663252     active
 clcomd           caa              7471316      active

Status of the optional HACMP subsystems:
Subsystem         Group            PID          Status
 clinfoES         cluster          11534382     active
root@migr5 / #
```

You might get equivalent status information by using other Simple Network Management Protocol (SNMP)-based commands, such as **cldump** or **clstat -o**, that are combined with **clshowsrv -v**. Also, instead of SNMP-based commands, you can use **lssrc -ls clstrmgrES**, **clRGinfo**, and **netstat -in**.

2. Ensure that the cluster has no pending changes on any of its nodes. Run a verification from SMIT, which is explained in step 3.

   If changes are pending on one node and you choose to apply them on top of the current configuration, check them, decide on a final configuration, and run a Verify and Synchronize Cluster Configuration operation. In an active cluster, some changes might not be allowed. If you really need these changes, you must stop the cluster services.

   If you decide to cancel any pending changes on any node and to keep the currently active configuration, run on either node `smit sysmirror` → **Problem Determination Tools** → **Restore PowerHA SystemMirror Configuration Database from Active Configuration**. Then, select **Verify and Synchronize Cluster Configuration** on the same node. You might avoid the last synchronization by restoring the default cluster configuration from the active configuration on each of the cluster nodes.

3. Check that the cluster has no configuration errors.

   Run the Verify Cluster Configuration operation on either cluster node by following the path `smitty sysmirror` → **Problem Determination Tools** → **PowerHA SystemMirror Verification** → **Verify Cluster Configuration**. Select **No** for the “Verify changes only?” field and press Enter. The result is shown in Figure A-1.

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...164]
   ihs_app                           ihs_rg
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.

[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure A-1 Verifying the cluster configuration*

4. Confirm that the CAA cluster state is error-free (Example A-2).

*Example A-2 Checking the CAA cluster state*

```
root@migr5 / # lscluster -i
Network/Storage Interface Query

Cluster Name:  clmigr56
Cluster uuid:  435ee888-7913-11e1-a13d-b6fcca1bda70
```

```
Number of nodes reporting = 2
Number of nodes expected = 2
Node migr5
Node uuid = 436337ee-7913-11e1-a13d-b6fcca1bda70
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  b6.fc.ca.1b.da.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 2
                IPV4 ADDRESS: 172.16.21.75  broadcast 172.16.23.255  netmask
255.255.252.0
                IPV4 ADDRESS: 172.16.21.77  broadcast 172.16.21.255  netmask
255.255.254.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.16.21.75  broadcast 0.0.0.0
netmask 0.0.0.0
        Interface number 2 dpcom
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
Node migr6
Node uuid = 656d55e0-7913-11e1-8d55-2e479566c670
Number of interfaces discovered = 2
        Interface number 1 en0
                ifnet type = 6 ndd type = 7
                Mac address length = 6
                Mac address =  2e.47.95.66.c6.6f
                Smoothed rrt across interface = 7
                Mean Deviation in network rrt across interface = 3
                Probe interval for interface = 100 ms
                ifnet flags for interface = 0x1e080863
                ndd flags for interface = 0x21081b
                Interface state  UP
                Number of regular addresses configured on interface = 1
                IPV4 ADDRESS: 172.16.21.76  broadcast 172.16.23.255  netmask
255.255.252.0
               Number of cluster multicast addresses configured on interface =
1
                IPV4 MULTICAST ADDRESS: 228.16.21.75  broadcast 0.0.0.0
netmask 0.0.0.0
        Interface number 2 dpcom
```

```
                ifnet type = 0 ndd type = 305
                Mac address length = 0
                Mac address =  0.0.0.0.0.0
                Smoothed rrt across interface = 750
                Mean Deviation in network rrt across interface = 1500
                Probe interval for interface = 22500 ms
                ifnet flags for interface = 0x0
                ndd flags for interface = 0x9
                Interface state  UP  RESTRICTED  AIX_CONTROLLED
root@migr5 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 436337ee-7913-11e1-a13d-b6fcca1bda70
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 656d55e0-7913-11e1-8d55-2e479566c670
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 1
        Point-of-contact interface & contact state
         en0  UP
root@migr5 / #

root@migr6 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 436337ee-7913-11e1-a13d-b6fcca1bda70
```

```
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 1
        Point-of-contact interface & contact state
         en0  UP

------------------------------

        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 656d55e0-7913-11e1-8d55-2e479566c670
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        435ee888-7913-11e1-a13d-b6fcca1bda70

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a
root@migr6 / #
```

## Performing the AIX update

To perform the update, follow these steps:

1. Ensure that all nodes in the cluster are running the same and most recent version of the AIX and PowerHA SystemMirror software.

   Obtain out the latest updates (service packs) that are available for the current versions of AIX and PowerHA SystemMirror in your cluster nodes. Apply these service packs in order to eliminate any known bug that might affect the migration.

   In our test environment, we start from AIX 7.1 TL00 SP1 and PowerHA SystemMirror 7.1.0 SP1 (Example A-3).

   *Example A-3 Initial cluster state and software version*

```
root@migr5 / # oslevel -s
7100-00-01-1037

Or gather the information using the following command:
# /usr/es/sbin/cluster/utilities/halevel -s

root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7101
```

```
cluster fix level is "1"
root@migr5 / #

root@migr6 / # oslevel -s
7100-00-01-1037
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7101
cluster fix level is "1"
root@migr6 / #
```

We update to the latest levels that are available at the time of writing this book, AIX 7.1 TL00 SP04 and PowerHA SystemMirror 7.1.0 SP05. Then, we ensure that everything is fine after the update by running again a Verify and Synchronize operation, followed by a cluster startup and status check (Example A-4).

*Example A-4 Cluster state and software versions after we apply the latest service packs*

```
root@migr5 / # oslevel -s
7100-00-04-1140
root@migr5 / # lppchk -v
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr5 / #

root@migr5 / # oslevel -s
7100-00-04-1140
root@migr5 / # lppchk -v
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr5 / #
```

2. Stop PowerHA cluster services on all nodes.

   Use the **smitty clstop** command. Select all nodes and choose “Bring Resource Groups Offline” as an action on the resource groups (Figure A-2 on page 539).

```
Stop Cluster Services

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

                                                     [Entry Fields]
* Stop now, on system restart or both               now                    +
  Stop Cluster Services on these nodes             [migr5,migr6]           +
  BROADCAST cluster shutdown?                       true                   +
* Select an Action on Resource Groups               Bring Resource Groups> +

F1=Help             F2=Refresh          F3=Cancel           F4=List
Esc+5=Reset         Esc+6=Command       Esc+7=Edit          Esc+8=Image
Esc+9=Shell         Esc+0=Exit          Enter=Do
```

*Figure A-2 Stopping cluster services on both nodes*

Ensure that the PowerHA cluster services are stopped in all nodes (Example A-5).

*Example A-5 Checking that the cluster is stopped*

```
root@migr5 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr5 / #

root@migr6 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr6 / #
```

3. Remove the CAA cluster.

   Use **lscluster** and **lspv** to get information about the cluster name and repository disk port VLAN (virtual local area network) identifier (PVID) (Example A-6).

*Example A-6 Finding the CAA cluster name and repository disk PVID*

```
root@migr5 / # lscluster -c
Cluster query for cluster clmigr56 returns:
Cluster uuid: 7d2b1d10-7b21-11e1-a317-b6fcca1bda70
Number of nodes in cluster = 2
        Cluster id for node migr5 is 1
        Primary IP address for node migr5 is 172.16.21.75
        Cluster id for node migr6 is 2
        Primary IP address for node migr6 is 172.16.21.76
Number of disks in cluster = 0
Multicast address for cluster is 228.16.21.75
root@migr5 / # lspv
caa_private0    00f74d4750f55449                    caavg_private   active
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
hdisk0          00f74d473088fca1                    rootvg          active
hdisk1          00f74d4731ff9753                    rootvg          active
root@migr5 / #
```

Use the **rmcluster –n *<clusterName>*** command in one node to remove the cluster. Then, check that the cluster is successfully removed with the **lscluster -m** command (Example A-7).

*Example A-7 Removing the CAA cluster*

```
root@migr5 / # rmcluster -n clmigr56
rmcluster: Removed cluster shared disks are automatically renamed to names such
        as hdisk10, [hdisk11, ...] on all cluster nodes.  However, this cannot
        take place while a disk is busy or on a node which is down or not
        reachable.  If any disks cannot be renamed now, you must manually
        rename them.
root@migr5 / # lscluster -m
Cluster services are not active.
root@migr5 / #

root@migr6 / # lscluster -m
Cluster services are not active.
root@migr6 / #
```

The repository disk must now be cleaned of any CAA Logical Volume Manager (LVM) structure on any node. See `hdisk2` in Example A-8.

*Example A-8 Checking that the previous repository disk is clean*

```
root@migr5 / # lspv
hdisk0          00f74d473088fca1                    rootvg          active
hdisk1          00f74d4731ff9753                    rootvg          active
hdisk2          00f74d4750f55449                    None
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
root@migr5 / #

root@migr6 / # lspv
hdisk2          00f74d4750f55449                    None
hdisk3          00f74d4750f55592                    ihsvg
hdisk4          00f74d4750f556a5                    ihsvg
hdisk0          00f74d45501b2428                    rootvg          active
hdisk1          00f74d455084775a                    rootvg          active
root@migr6 / #
```

**Repository disk PVID:** If problems appear here, call IBM support. Do not try any action that might change the PVID of the repository disk. The PVID is needed later to recreate the CAA cluster with the same repository disk.

4. Upgrade AIX on all nodes in the cluster.

   Upgrade the AIX to version 7.1 TL01 or later by using a supported procedure. In our scenario, we upgrade to the latest level available at the time of writing this book, AIX 7.1 TL01 SP3. Ensure that you restart the systems after the upgrade (Example A-9 on page 541).

**Configuration files:** Use a supported upgrade procedure and appropriate software sources (download or media); otherwise, you might overwrite the configuration files. For example, if you use AIX 7.1 TL01 base media to update from AIX 7.1 TL00, you might overwrite the `/etc/cluster/rhosts` file.

*Example A-9 Checking the AIX upgrade*

```
root@migr5 / # uptime
  11:16AM   up   1:38,  1 user,  load average: 1.27, 1.24, 1.17
root@migr5 / # oslevel -s
7100-01-03-1207
root@migr5 / # lppchk -v
root@migr5 / #

root@migr6 / # uptime
  11:16AM   up   1:38,  1 user,  load average: 1.19, 1.26, 1.16
root@migr6 / # oslevel -s
7100-01-03-1207
root@migr6 / # lppchk -v
root@migr6 / #
```

5. Recreate the CAA cluster from PowerHA SystemMirror.

   Ensure that `/etc/cluster/rhosts` and **`clcomd`** are in a good state (Example A-10).

*Example A-10 Checking the clcomd communication*

```
root@migr5 / # clrsh migr5 hostname
migr5
root@migr5 / # clrsh migr6 hostname
migr6
root@migr5 / #

root@migr6 / # clrsh migr5 hostname
migr5
root@migr6 / # clrsh migr6 hostname
migr6
root@migr6 / #
```

Run **`smitty sysmirror`** → **Cluster Nodes and Networks** → **Verify and Synchronize Cluster Configuration** (Figure A-3 on page 542).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...75]
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.
Committing any changes, as required, to all available nodes...
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab
and /e
tc/rc.net for IP Address Takeover on node migr5.
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab
and /e
tc/rc.net for IP Address Takeover on node migr6.

Verification has completed normally.

[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure A-3 Synchronizing the PowerHA cluster*

Check that the CAA cluster is recreated (Example A-11).

*Example A-11 Verifying the new CAA cluster state*

```
root@migr5 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: migr5
        Cluster shorthand id for node: 1
        uuid for node: 98d20982-7b5a-11e1-9a04-b6fcca1bda6f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        9908a3fc-7b5a-11e1-9a04-b6fcca1bda6f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------
```

```
        Node name: migr6
        Cluster shorthand id for node: 2
        uuid for node: 990400fe-7b5a-11e1-9a04-b6fcca1bda6f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmigr56           local        9908a3fc-7b5a-11e1-9a04-b6fcca1bda6f

        Number of points_of_contact for node: 2
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en0  UP
root@migr5 / #
```

6. Restart the PowerHA cluster services.

   Start the cluster by using **smitty clstart** and check its status (Example A-12).

*Example A-12 Final cluster status*

```
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr5 / #

root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7105
cluster fix level is "5"
root@migr6 / # clmgr view report status

Obtaining information via SNMP from Node: migr6...
_____________________________________________________________________________
Cluster Name: clmigr56
Cluster State: UP
Cluster Substate: STABLE
_____________________________________________________________________________

Node Name: migr5                State: UP

  Network Name: net_ether_01       State: UP

    Address: 172.16.21.75    Label: migr5            State: UP
    Address: 172.16.21.77    Label: migrsvc5         State: UP

Node Name: migr6                State: UP

  Network Name: net_ether_01       State: UP

    Address: 172.16.21.76    Label: migr6            State: UP
```

```
Cluster Name: clmigr56

Resource Group Name: ihs_rg
Startup Policy: Online On First Available Node
Fallover Policy: Fallover To Next Priority Node In The List
Fallback Policy: Never Fallback
Site Policy: ignore
Node                         Group State
---------------------------- ---------------
migr5                        ONLINE
migr6                        OFFLINE


Status of the RSCT subsystems used by HACMP:
Subsystem         Group            PID          Status
 cthags           cthags           10289194     active
 ctrmc            rsct             6881416      active

Status of the HACMP subsystems:
Subsystem         Group            PID          Status
 clstrmgrES       cluster          5963996      active
 clcomd           caa              5767344      active

Status of the optional HACMP subsystems:
Subsystem         Group            PID          Status
 clinfoES         cluster          11665610     active
root@migr6 / #
```

# Upgrading to PowerHA 7.1.1

We present a simple procedure by which we can upgrade the PowerHA 7.1.0 that already runs on top of AIX 7.1 TL1 SP3 to PowerHA 7.1.1. This procedure is adapted from the official procedure for offline migration of a PowerHA 7.1.0 cluster that runs on top of AIX 7.1 TL00, which is presented in 6.3.1, “Offline migration from PowerHA 7.1.0 version” on page 321. The AIX and the CAA cluster are already at the required versions for PowerHA 7.1.1. So, we skip the step of removing the CAA cluster and the step of upgrading the AIX. After we upgrade the PowerHA filesets, we also skip the step of recreating the CAA cluster by using the **clmkcaa** utility script. We used the following high-level steps of the simplified procedure:

1. Prepare the PowerHA upgrade:
   a. Save a cluster snapshot, make the usual backups (data, binaries, and **mksysb**), and prepare a back-out plan.
   b. Check the software consistency and ensure that the cluster state and configuration are error-free on all nodes.

   These steps are similar to the steps in “Preparing the AIX upgrade” on page 532, which are already documented in this appendix.

2. Perform the PowerHA upgrade to 7.1.1:
   a. Stop PowerHA cluster services on all nodes.
   b. Upgrade PowerHA filesets on all nodes to 7.1.1, which is the latest service pack.
   c. Restart the PowerHA cluster services, node-by-node.

   We present these three steps in depth.

## Perform the PowerHA upgrade to 7.1.1

With all preparations already performed, we follow these steps:

1. Stop PowerHA cluster services on all nodes.

   Use the **smitty clstop** command. Select all nodes and choose **Bring Resource Groups Offline** as an action on the resource groups (Figure A-4).

```
Stop Cluster Services

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

                                                         [Entry Fields]
* Stop now, on system restart or both                now                     +
  Stop Cluster Services on these nodes              [migr5,migr6]            +
  BROADCAST cluster shutdown?                        true                    +
* Select an Action on Resource Groups                Bring Resource Groups>  +

F1=Help             F2=Refresh          F3=Cancel           F4=List
Esc+5=Reset         Esc+6=Command       Esc+7=Edit          Esc+8=Image
Esc+9=Shell         Esc+0=Exit          Enter=Do
```

*Figure A-4 Stopping cluster services on both nodes*

   Ensure that the PowerHA cluster services are stopped on all nodes (Example A-13).

*Example A-13 Checking that the cluster is stopped*

```
root@migr5 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr5 / #

root@migr6 / # lssrc -ls clstrmgrES|grep state
Current state: ST_INIT
root@migr6 / #
```

2. Upgrade the PowerHA filesets on all nodes to version 7.1.1 and apply the latest service pack.

   In our scenario, after we apply SP02 for PowerHA 7.1.1, which is the latest available SP at the time of writing this book. Example A-14 shows the PowerHA 7.1.1 fileset versions after the upgrade.

*Example A-14 Checking the upgrade of the PowerHA filesets*

```
root@migr6 / # lppchk -v cluster\*
root@migr6 / # lslpp -l cluster\*|grep -p Level
  Fileset                      Level  State      Description
```

```
  ---------------------------------------------------------------------------
Path: /usr/lib/objrepos
  cluster.adt.es.client.include
                             7.1.1.1  COMMITTED  PowerHA SystemMirror Client
                                                 Include Files
  cluster.adt.es.client.samples.clinfo
                             7.1.1.0  COMMITTED  PowerHA SystemMirror Client
                                                 CLINFO Samples
  cluster.adt.es.client.samples.clstat
                             7.1.1.0  COMMITTED  PowerHA SystemMirror Client
                                                 Clstat Samples
  cluster.adt.es.client.samples.libcl
                             7.1.1.0  COMMITTED  PowerHA SystemMirror Client
                                                 LIBCL Samples
  cluster.adt.es.java.demo.monitor
                             7.1.1.0  COMMITTED  Web Based Monitor Demo
  cluster.es.client.clcomd   7.1.1.0  COMMITTED  Cluster Communication
                                                 Infrastructure
  cluster.es.client.lib      7.1.1.1  COMMITTED  PowerHA SystemMirror Client
                                                 Libraries
  cluster.es.client.rte      7.1.1.1  COMMITTED  PowerHA SystemMirror Client
                                                 Runtime
  cluster.es.client.utils    7.1.1.1  COMMITTED  PowerHA SystemMirror Client
                                                 Utilities
  cluster.es.client.wsm      7.1.1.0  COMMITTED  Web based Smit
  cluster.es.cspoc.cmds      7.1.1.2  COMMITTED  CSPOC Commands
  cluster.es.cspoc.dsh       7.1.1.0  COMMITTED  CSPOC dsh
  cluster.es.cspoc.rte       7.1.1.2  COMMITTED  CSPOC Runtime Commands
  cluster.es.migcheck        7.1.1.0  COMMITTED  PowerHA SystemMirror Migration
                                                 support
  cluster.es.server.cfgast   7.1.1.0  COMMITTED  Two-Node Configuration
                                                 Assistant
  cluster.es.server.diag     7.1.1.1  COMMITTED  Server Diags
  cluster.es.server.events   7.1.1.1  COMMITTED  Server Events
  cluster.es.server.rte      7.1.1.1  COMMITTED  Base Server Runtime
  cluster.es.server.testtool
                             7.1.1.0  COMMITTED  Cluster Test Tool
  cluster.es.server.utils    7.1.1.1  COMMITTED  Server Utilities
  cluster.license            7.1.1.0  COMMITTED  PowerHA SystemMirror
                                                 Electronic License
  cluster.msg.en_US.es.client
                             7.1.1.0  COMMITTED  PowerHA SystemMirror Client
                                                 Messages - U.S. English
  cluster.msg.en_US.es.server
                             7.1.1.1  COMMITTED  Recovery Driver Messages -
                                                 U.S. English

root@migr6 / # lppchk -v
root@migr6 / #
```

3. Restart the PowerHA cluster services, node-by-node.

   Use **smitty clstart** on the first node (Figure A-5 on page 547).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Cluster services are running at different levels across
the cluster.  Verification will not be invoked in this environment.

Starting Cluster Services on node: migr5
This may take a few minutes.  Please wait...
migr5: start_cluster: Starting PowerHA SystemMirror
migr5:   4980904      -  0:00 syslogd
migr5: Setting routerevalidate to 1
migr5: INFORMATION: must wait at least 2 minutes before cluster restart
migr5: Sleeping 2 minutes.
migr5: 0513-059 The topsvcs Subsystem has been started. Subsystem PID is
7995586
[MORE...24]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure A-5 Cluster starts up correctly on the first node*

Then, ensure that the node successfully joins the cluster (Example A-15). Although the software is updated at the 7.1.1 level, the node still runs at version 12.

*Example A-15 First node joined the cluster*

```
root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7112
cluster fix level is "2"
root@migr5 / #
```

Repeat the procedure for each node of the cluster, one node at a time.

After you start the cluster services on the latest node and it joins the cluster, you can check the cluster version update (Example A-16).

*Example A-16 Last node joins the cluster*

```
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_JOINING
CLversion: 12
local node vrmf is 0
cluster fix level is "ffffffff"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_BARRIER
CLversion: 12
local node vrmf is 7112
```

```
cluster fix level is "2"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 12
local node vrmf is 7112
cluster fix level is "2"
root@migr6 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7112
cluster fix level is "2"
root@migr6 / #

root@migr5 / # lssrc -ls clstrmgrES|egrep "state|version|vrmf|fix"
Current state: ST_STABLE
CLversion: 13
local node vrmf is 7112
cluster fix level is "2"
root@migr5 / #
```

Select a final **Verify and Synchronize Cluster Configuration** to ensure that the cluster runs error-free (Figure A-6).

```
 COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[MORE...220]
Cluster Manager Current state: ST_BARRIER
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_RP_RUNNING
Cluster Manager Current state: ST_CBARRIER
Cluster Manager Current state: ST_UNSTABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
Cluster Manager Current state: ST_STABLE
...completed.


[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

*Figure A-6   Final verification of the migrated cluster*

**Alternate path:** By combining the steps in “Upgrading AIX” on page 532 and “Upgrading to PowerHA 7.1.1” on page 544, we can perform a complete PowerHA migration from version 7.1.0 that is running on top of AIX 7.1.0. This alternate path resembles the standard offline migration. The **clmkcaa** utility script is not used.

B

# Configuring the PowerHA cluster by using clmgr

This appendix provides information about the **clmgr** command-line utility available in PowerHA 7.1.1. First introduced in PowerHA 7.1.0, this utility provides a consistent, reliable interface for performing PowerHA cluster operations, such as cluster configuration and management.

We describe these topics in this appendix:

- Introduction to clmgr
- Full syntax for the clmgr command and basic format
- Cluster configuration topology
- Cluster configuration resources
- Cluster verification and synchronization
- Cluster start and stop
- Cluster management

# Introduction to clmgr

The **clmgr** is a command-line utility that can manage all PowerHA capabilities (for example, all tasks that can be managed through SMIT). The purpose of **clmgr** is to provide a consistent, reliable interface for performing PowerHA cluster operations by using a terminal or script. In release 7.1.1, the **clmgr** is enhanced to support all Cluster Single Point of Control (C-SPOC) functionalities that are missing in the previous release. This utility is improved to be *case insensitive*; therefore, it allows users to type `CLUSTER`, `Cluster`, or `cluster` to get the same results.

One earlier limitation of the **clmgr** command is that it requires all actions, classes, and attribute labels to be fully spelled out by the user. This limitation is addressed and removed with the new **clmgr** utility. Now, users can type aliases for actions, classes, and attributes to perform the tasks. For more details, see the **clmgr** man page.

The output from the **clmgr** command is displayed in the format “`ATTRIBUTE-NAME=VALUE`”.

All **clmgr** command operations are logged in the `clutils.log` file. This file saves the operation in verbose-enabled format. This format includes the name of the command that is executed, the start and stop times, and the name that initiated the command. Example B-1 is sample data from the `clutils.log` file.

*Example B-1 Output from the clutils.log file*

```
CLMGR STARTED  (2562:6029388:6750276):  2012-03-16T19:51:22.583539
CLMGR USER     (2562:6029388:6750276):  ::root:system
    isEnterprise()[10](.000): /usr/bin/lslpp -lc "cluster.xd.*"
    lslpp RC: 1
    get_local_node_label()[21](.140):
/usr/es/sbin/cluster/utilities/get_local_nodename
    Warning: There is no cluster found.
    cllsclstr: No cluster defined.
    cllsclstr: Error reading configuration.
    Warning: There is no cluster found.
    cllsnode: Error reading configuration.
    get_local_nodename RC: 0; localnode == ""
    KLIB_HACMP_get_cluster_name()[11](.220):
/usr/es/sbin/cluster/utilities/cllsclstr -cS
    Warning: There is no cluster found.
    cllsclstr: No cluster defined.
    cllsclstr: Error reading configuration.
    cllsclstr RC: 255; name == ""
    isEnterprise()[10](.010): /usr/bin/lslpp -lc "cluster.xd.*"
    lslpp RC: 1
```

**Sites:** At the time of writing this book, site support is not available in PowerHA so **clmgr** does not allow users to configure and manage sites.

# Full syntax for the clmgr command and basic format

The following syntax is the full syntax of the **clmgr** command:

```
clmgr [-c|-x] [-S] [-v] [-f] [-D] [-l {low|med|high|max}] [-T <ID>]
        [-a {<ATTR#1>,<ATTR#2>,...}] <ACTION> <CLASS> [<NAME>]
```

```
        [-h | <ATTR#1>=<VALUE#1> <ATTR#2>=<VALUE#2> <ATTR#n>=<VALUE#n>]

       ACTION={add|modify|delete|query|online|offline|...}
        CLASS={cluster|node|network|resource_group|...}

clmgr {-h|-?} [-v]
clmgr [-v] help
```

Use the following basic format the **clmgr** command:

**clmgr** *<ACTION>* *<CLASS>* [*<NAME>*] [*<ATTRIBUTES...>*]

The command uses the following flags:

- ACTION

  This flag describes the operation to be performed. ACTION is not case-sensitive. Aliases are provided for command-line use and *must not be used in script*. Four ACTIONs are available:

  - Add (Alias: a)
  - Query (Alias: q)
  - Modify (Aliases: mod, ch, and set)
  - Delete (Aliases: de, rm, and er)

  For a detailed list of the available ACTIONs, see the **clmgr** manual page.

- CLASS

  This flag contains the type of object on which the ACTION is performed. Aliases are provided for command-line use and *must not be used in scripts*. Some of the available CLASSes are listed:

  - Cluster (Alias: cl)
  - Node (Alias: no)
  - Interface (Aliases: in and if)
  - Network (Aliases: ne and nw)

  A detailed list is available in the **clmgr** man page.

- Name

  This flag contains the specific object of type CLASS, on which ACTION is to be performed.

- ATTR=value

  This flag is optional and has attribute pairs and value pairs that are specific to the ACTION+CLASS combination.

## Cluster configuration topology

We describe how to configure the topology components of the PowerHA cluster by using **clmgr**. The topology-related components are listed:

- Cluster
- Nodes
- Network
- Interfaces
- Persistent IP/label

## Cluster

The `cluster` action is specified to configure a PowerHA cluster with a user-specified name. When this command is executed without specifying any node name, it takes the local node name or the host from where **clmgr** runs as the first node of the cluster (Example B-2).

Use this syntax:

**clmgr** add cluster *<CLS-name>* [REPOSITORY=*<hdisk#>* CLUSTER_IP=*<multicast IP address>*]

*Example B-2 Command to create a cluster*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add cluster clmgr_cluster
REPOSITORY=hdisk2 CLUSTER_IP=228.1.1.36
Warning: since no nodes were specified for this cluster, a one-node cluster will
be created with this system: "clmgr1"
Cluster Name: clmgr_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: None
Cluster IP Address:
There are 1 node(s) and 1 network(s) defined
NODE clmgr1:
        Network net_ether_02
                clmgr1b2        10.1.2.54

No resource groups defined
clharvest_vg: Initializing....
Gathering cluster information, which may take a few minutes...
clharvest_vg: Processing...
Storing the following information in file
/usr/es/sbin/cluster/etc/config/clvg_config

clmgr1:

Hdisk:          hdisk0
PVID:           00f74d471c3abc13
VGname:         rootvg
VGmajor:        10
Conc-capable:   No
VGactive:       Yes
Quorum-required:Yes
Hdisk:          hdisk1
PVID:           00f74d4733c60118
VGname:         None
VGmajor:        0
Conc-capable:   No
VGactive:       No
Quorum-required:No
Hdisk:          hdisk2
PVID:           00f74d4733c91f39
VGname:         None
VGmajor:        0
Conc-capable:   No
```

```
VGactive:       No
Quorum-required:No
Hdisk:          hdisk3
PVID:           00f74d4733c9370e
VGname:         None
VGmajor:        0
Conc-capable:   No
VGactive:       No
Quorum-required:No
FREEMAJORS:     35...
Cluster Name: clmgr_cluster
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.36
There are 1 node(s) and 1 network(s) defined
NODE clmgr1:
        Network net_ether_02
                clmgr1b2        10.1.2.54

No resource groups defined

Warning: There is no cluster found.
cllsclstr: No cluster defined
cllsclstr: Error reading configuration
Communication path clmgr1 discovered a new node. Hostname is clmgr1. Adding it to
the configuration with Nodename clmgr1.
Discovering IP Network Connectivity

Retrieving data from available cluster nodes.  This could take a few minutes.

        Start data collection on node clmgr1
        Collector on node clmgr1 completed
        Data collection complete
        Completed 10 percent of the verification checks
        Completed 20 percent of the verification checks
        Completed 30 percent of the verification checks
        Completed 40 percent of the verification checks
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
Discovered [3] interfaces
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks
IP Network Discovery completed normally


Current cluster configuration:


Discovering Volume Group Configuration
```

```
Current cluster configuration:

root@clmgr1 / #
```

The warning message states that a one-node cluster is configured. The system, from which **clmgr** is executed, is added as the first node of the cluster. The example configures a one-node cluster named `clmgr_cluster` with node `clmgr1` as the first node of the cluster.

## Node

The `node` action is used to add nodes or servers to the cluster (Example B-3).

Use this syntax:

**clmgr** add node <*node-name*> [COMMPATH=<*communication path to the node*>]

*Example B-3 Command to add a node*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add node clmgr2
Attempting to discover available resources on "clmgr2"...

root@clmgr1 / #
```

Example B-3 adds node `clmgr2` to the cluster. Execute this command for all cluster nodes in your configuration.

## Network

The `network` or `nw` action is used to create or add a network in the PowerHA cluster (Example B-4).

Use this syntax:

**clmgr** add network <*network-name*> [TYPE={ether|infiniband} NETMASK=<*255.255.x.x*> | PREFIX=1..128]

*Example B-4 Command to add a network*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add network net_ether_01
TYPE=ether

root@clmgr1 / #
```

This example creates an `ether`-type network that is labeled `net_ether_01` and adds it to the cluster configuration. If your cluster configuration needs more networks, you can execute the same command.

## Interface

The `interface` action is used to add communication interfaces to the network in the PowerHA cluster (Example B-5 on page 557).

Use this syntax:

**clmgr** add interface <*interface-label*> NETWORK=<*network-name*> [NODE=<*node*> TYPE={ether|infiniband} INTERFACE=<*network-interface*>]

*Example B-5 Command to add an interface to the network*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add interface clmgr2b2
NETWORK=net_ether_02 NODE=clmgr2 INTERFACE=en1

root@clmgr1 / #
```

To eliminate a communication interface from becoming a single point of failure (SPOF), add more interfaces.

### Persistent IP/label

The `persistent_ip` action is used to add a persistent IP address for the network in the PowerHA cluster. One persistent IP address or label is allowed per node per network. In a two-node cluster configuration with the `ether`-type network, at least two persistent IPs are required, one each for node (Example B-6).

Use this syntax:

**clmgr** add persistent_ip <*persistent_IP*> NETWORK=<*network-name*> [NODE=<*node*>]

*Example B-6 Command to add a persistent IP label*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add persistent clmgr1p1
NETWORK=net_ether_01 NODE=clmgr1

root@clmgr1 / #
```

You can run the **cllsif -p** command to check the persistent IPs that are added to the configuration as shown in Example B-7.

*Example B-7 Output of cllsif command*

```
root@clmgr1 / # cllsif -p
Adapter              Type       Network    Net Type   Attribute  Node       IP
Address       Hardware Address Interface Name   Global Name      Netmask
Alias for HB Prefix Length

clmgr1               boot       net_ether_01 ether      public     clmgr1
10.1.1.54                     en0                      255.255.255.0
24
clmgr1p1             persistent net_ether_01 ether      public     clmgr1
172.16.21.54                                           255.255.255.0
24
clmgr1b2             boot       net_ether_02 ether      public     clmgr1
10.1.2.54                     en1                      255.255.254.0
23
clmgr2               boot       net_ether_01 ether      public     clmgr2
10.1.1.55                     en0                      255.255.255.0
24
clmgr2p1             persistent net_ether_01 ether      public     clmgr2
172.16.21.55                                           255.255.255.0
24
clmgr2b2             boot       net_ether_02 ether      public     clmgr2
10.1.2.55                     en1                      255.255.254.0
23
```

```
root@clmgr1 / #
```

## Cluster configuration resources

In this section, we show how to configure resources and resource groups by using the **clmgr** utility. Resource components are the entities that are made highly available when the cluster is in the production state. We configure the following resource components:

- Service IP/label
- Application controller
- Application monitor
- Volume group
- Resource group

### Service IP/label

To configure the service IP/label, we use `service_ip`. When this command is executed without specifying any node name, it uses the local node name or the host, from where **clmgr** is run, as the first node of the cluster (Example B-8).

Use this syntax:

**clmgr** add service_ip <*Service-label*> NETWORK=<*network-name*> [NETMASK=<*255.x.x.x*>]

*Example B-8 Command to create the service IP/Label*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add service_ip clmgrsvc1
NETWORK=net_ether_01

root@clmgr1 / #
```

Example B-8 adds the service IP/label `clmgrsvc1` to the network `net_ether_01`. This service IP can be added as a resource to the resource group.

### Application controller

The term *application controller* is used in PowerHA to refer to the application server start and stop scripts that are used to start and stop the application when the resource group comes online. To configure the application server, we specify the `application_controller` action or aliases `ac` or `app` (Example B-9).

Use this syntax:

**clmgr** add application_controller <*app-server-name*> STARTSCRIPT="<*absolutepath*>" STOPSCRIPT="<*absolutepath*>" [STARTUP_MODE={background|foreground}]

*Example B-9 Command to add the application controller*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add application_controller
test_app1 STARTSCRIPT="/home/apps/start1.sh" STOPSCRIPT="/home/apps/stop1.sh"
STARTUP_MODE=background

root@clmgr1 / #
```

Example B-9 configures an application controller `test_app1`. Repeat the command if you need more application servers in your configuration.

## Application monitor

Application monitors are optional but highly suggested resource components in PowerHA. Application monitors help ensure that applications function correctly and they are also useful during a forced down or UNMANAGE stop of cluster. You can configure more than one application monitor for the same application controller. The ACTION to be used to configure the application monitor is `application_monitor` or `am` (Example B-10).

Use this syntax:

```
clmgr add application_monitor <monitor-name> TYPE={Process|Custom}
APPLICATIONS=<appctrl#1>[,<appctrl#2...>] MODE={longrunning|startup|both}
[STABILIZATION="1 .. 3600" ] [RESTARTCOUNT="0 .. 100"]
[FAILUREACTION={notify|fallover} ]

Process Arguments:
PROCESSES="pmon1,dbmon,..."  \
OWNER="<processes_owner_name>" \
[ INSTANCECOUNT="1 .. 1024" ] \
[ RESTARTINTERVAL="1 .. 3600" ] \
[ NOTIFYMETHOD="</script/to/notify>" ] \
[ CLEANUPMETHOD="</script/to/cleanup>" ] \
[ RESTARTMETHOD="</script/to/restart>" ]

Custom Arguments:
MONITORMETHOD="/script/to/monitor" \
[ MONITORINTERVAL="1 .. 1024" ] \
[ HUNGSIGNAL="1 .. 63" ] \
[ RESTARTINTERVAL="1 .. 3600"  ] \
[ NOTIFYMETHOD="</script/to/notify>" ] \
[ CLEANUPMETHOD="</script/to/cleanup>" ] \
[ RESTARTMETHOD="</script/to/restart>" ]
```

*Example B-10 Command to create the application monitor*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add application_monitor
monitor1 TYPE=Custom APPLICATIONS=test_app1 MODE=both STABILIZATION=360
RESTARTCOUNT=2 FAILUREACTION=fallover MONITORMETHOD="/home/apps/appmon1.sh"
HUNGSIGNAL=3 RESTARTINTERVAL=30 CLEANUPMETHOD="/home/apps/stop1.sh"
RESTARTMETHOD="/home/apps/start1.sh"

root@clmgr1 / #
```

Example B-10 creates an application monitor `monitor1` for application controller `test_app1`. This application monitor is a custom-based monitor. You can configure more than one application monitor for the same application controller.

## Volume group

The `volume_group` action can be used to create a shared volume group in your configuration as shown in Example B-11 on page 560.

Use this syntax:

```
clmgr add volume_group <vg-name> NODES="<node#1>,<node#2>[,...>]"
PHYSICAL_VOLUMES="<disk#1>[,<disk#2>,...]" [TYPE={original|big|scalable|legacy} ]
[MAJOR_NUMBER=## ] [CONCURRENT_ACCESS={false|true}]
[ACTIVATE_ON_RESTART={false|true}] [CRITICAL={false|true} ] \
```

```
[ FAILURE_ACTION={halt|notify|fence|stoprg|moverg} ] \
[ NOTIFYMETHOD=</file/to/invoke> ]
```

*Example B-11 Command to create a shared volume group*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add volume_group test_vg1
NODES="clmgr1,clmgr2" PHYSICAL_VOLUMES=hdisk3 TYPE=original MAJOR_NUMBER=35
ACTIVATE_ON_RESTART=false
clmgr1: test_vg1

clmgr1: mkvg: This concurrent capable volume group must be varied on manually.
clmgr1: synclvodm: No logical volumes in volume group test_vg1.
clmgr1: Volume group test_vg has been updated.
clmgr2: synclvodm: No logical volumes in volume group test_vg1.
clmgr2: 0516-783 importvg: This imported volume group is concurrent capable.
clmgr2:         Therefore, the volume group must be varied on manually.
clmgr2: 0516-1804 chvg: The quorum change takes effect immediately.
clmgr2: Volume group test_vg has been imported.
cl_mkvg: Discovering Volume Group Configuration...
clmgr1: 0516-1804 chvg: The quorum change takes effect immediately.
root@clmgr1 / #
```

Example B-11 creates a shared volume group, `test_vg1`, on nodes `clmgr1` and `clmgr2` of type `concurrent`.

## Resource group

The *resource group* is the holding entity that contains resources, such as service IP, volume group, application controller, and file system. The action that is used to configure or add a resource group is `resource_group` (Example B-12).

Use this syntax:

```
clmgr add resource_group <RG-name> NODES="<node#1>,<node#2>[,...>]"
[STARTUP={OHN|OFAN|OAAN|OUDP}] [FALLOVER={FNPN|FUDNP|BO}]
[FALLBACK={NFB|FBHPN}] [SERVICE_LABEL=service_ip#1[,service_ip#2,...]]
[APPLICATIONS=appctlr#1[,appctlr#2,...]] [VOLUME_GROUP=<VG>[,<VG#2>,...]]
[FILESYSTEM=/file_system#1[,/file_system#2,...]]
[EXPORT_FILESYSTEM=/expfs#1[,/expfs#2,...]]
[EXPORT_FILESYSTEM_V4=/expfs#1[,/expfs#2,...]] [ STABLE_STORAGE_PATH="/fs3"]
[NFS_NETWORK="nfs_network"] [MOUNT_FILESYSTEM=/nfs_fs1;/expfs1,/nfs_fs2;,..]
```

*Example B-12 Command to add a resource group to the cluster*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add resource_group clmgr_RG1
NODES="clmgr1,clmgr2" STARTUP=OHN FALLOVER=FNPN FALLBACK=NFB VOLUME_GROUP=test_vg
SERVICE_LABLE=clmgrsvc1 APPLICATIONS=test_app1

Auto Discover/Import of Volume Groups was set to true.
Gathering cluster information, which may take a few minutes.
root@clmgr1 / #
```

Example B-12 creates a resource group that is named `clmgr_RG1` with participating nodes, `clmgr1` and `clmgr2`, with startup policy `Online on Home Node`, fallover policy `Fallover to Next Priority Node`, and fallback set to `Never Fallback` with a service label, `clmgrsvc1`. When the resource group starts, the volume group is activated, the file systems, if any, are mounted, and the service IP `serv_ip1` is aliased on one of the boot interfaces and made accessible. Run the **`clmgr add resource_group`** command to add more resource groups to your configuration, if needed.

## Cluster verification and synchronization

In this section, we look at the commands and actions for cluster verification and synchronization.

### Cluster verification

*Verification* is the process to check to ensure that there are no unsupported elements in the configuration. Verification checks to ensure that all resources that are used by PowerHA are configured correctly. Verification checks to ensure that the rules about resource ownership and resource takeover agree across all nodes. Verification must be executed on configuration and reconfiguration (Example B-13).

Use this syntax:

```
clmgr verify cluster [CHANGES_ONLY={no|yes}] [DEFAULT_TESTS={yes|no}]
[METHODS=<method#1>[,<method#2>,...]] [FIX={no|yes}] [LOGGING={standard|verbose}]
[LOGFILE=<PATH_TO_LOG_FILE>] [SYNC={no|yes}] [FORCE={no|yes} ]
```

*Example B-13   Command to run the cluster verification*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr verify cluster
CHANGES_ONLY=no FIX=yes LOGGING=standard
Verifying clcomd communication, please be patient.

Verifying multicast communication with mping.


Verification to be performed on the following:
        Cluster Topology
        Cluster Resources

Verification will automatically correct verification errors.


Retrieving data from available cluster nodes.  This could take a few minutes.

        Start data collection on node clmgr1
        Start data collection on node clmgr2
        Collector on node clmgr2 completed
        Collector on node clmgr1 completed
        Data collection complete

Verifying Cluster Topology...

        Completed 10 percent of the verification checks

WARNING: Multiple communication interfaces are recommended for networks that
use IP aliasing in order to prevent the communication interface from
becoming a single point of failure. There are fewer than the recommended
number of communication interfaces defined on the following node(s) for
the given network(s):
```

```
    Node:                              Network:
    ---------------------------------- ----------------------------------
    clmgr1                             net_ether_01
    clmgr2                             net_ether_01

        Completed 20 percent of the verification checks
WARNING: There are IP labels known to HACMP and not listed in file
/usr/es/sbin/cluster/etc/clhosts.client on node: clmgr1. Clverify can
automatically populate this file to be used on a client node, if executed in
auto-corrective mode.
WARNING: There are IP labels known to HACMP and not listed in file
/usr/es/sbin/cluster/etc/clhosts.client on node: clmgr2. Clverify can
automatically populate this file to be used on a client node, if executed in
auto-corrective mode.
Starting Corrective Action: cl_topology_modify_clhosts_client_entry.
<01> WARNING: File /usr/es/sbin/cluster/etc/clhosts.client does not exist on node:
clmgr1
<02> WARNING: Backing up /usr/es/sbin/cluster/etc/clhosts.client on node clmgr1 to
file /usr/es/sbin/cluster/etc/clhosts.client.03_22_2012: FAIL
<03> Adding entry(s) '172.16.21.71      #clmgrsvc1
10.1.1.54       #clmgr1
172.16.21.54    #clmgr1p1
10.1.2.54       #clmgr1b2
10.1.1.55       #clmgr2
172.16.21.55    #clmgr2p1
10.1.2.55       #clmgr2b2
' to /usr/es/sbin/cluster/etc/clhosts.client on node clmgr1: PASS
<04> WARNING: File /usr/es/sbin/cluster/etc/clhosts.client does not exist on node:
clmgr2
<05> WARNING: Backing up /usr/es/sbin/cluster/etc/clhosts.client on node clmgr2 to
file /usr/es/sbin/cluster/etc/clhosts.client.03_22_2012: FAIL
<06> Adding entry(s) '172.16.21.71      #clmgrsvc1
10.1.1.54       #clmgr1
172.16.21.54    #clmgr1p1
10.1.2.54       #clmgr1b2
10.1.1.55       #clmgr2
172.16.21.55    #clmgr2p1
10.1.2.55       #clmgr2b2
' to /usr/es/sbin/cluster/etc/clhosts.client on node clmgr2: PASS
WARNING: Network option "nonlocsrcroute" is set to 0 and will be set to 1 on
during HACMP startup on the following nodes:

        clmgr1
        clmgr2

WARNING: Network option "ipsrcrouterecv" is set to 0 and will be set to 1 on
during HACMP startup on the following nodes:

        clmgr1
        clmgr2

Starting Corrective Action: cl_resource_set_net_option.
<01> Successfully set network option nonlocsrcroute="1" on node clmgr1.
<02> Successfully set network option ipsrcrouterecv="1" on node clmgr1.
```

```
<03> Successfully set network option nonlocsrcroute="1" on node clmgr2.
<04> Successfully set network option ipsrcrouterecv="1" on node clmgr2.
        Completed 30 percent of the verification checks

Verifying Cluster Resources...

        Completed 40 percent of the verification checks
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.
root@clmgr1 / #
```

The command that is shown in Example B-13 on page 561 runs verification against the cluster configuration.

### Cluster synchronization

*Cluster synchronization* is the process to make all cluster nodes aware of the cluster configuration. This process updates PowerHA Object Data Managers (ODMs) on cluster nodes with the latest or updated configuration information as shown in Example B-14.

Use this syntax:

**clmgr** sync cluster [VERIFY={yes|no}] [CHANGES_ONLY={no|yes}] [ DEFAULT_TESTS={yes|no}] [METHODS=<*method#1*>[,<*method#2*>,...]] [FIX={no|yes} ] [LOGGING={standard|verbose}] [LOGFILE=<*PATH_TO_LOG_FILE*>] [MAX_ERRORS=##] [FORCE={no|yes}]

*Example B-14 Command that shows how to synchronize the cluster*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr sync cluster CHANGES_ONLY=no
FIX=yes LOGGING=standard
Saving existing /var/hacmp/clverify/ver_mping/ver_mping.log to
/var/hacmp/clverify/ver_mping/ver_mping.log.bak
Verifying clcomd communication, please be patient.

Verifying multicast communication with mping.


Committing any changes, as required, to all available nodes...
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab and
/etc/rc.net for IP Address Takeover on node clmgr1.
Adding any necessary PowerHA SystemMirror for AIX entries to /etc/inittab and
/etc/rc.net for IP Address Takeover on node clmgr2.

Verification has completed normally.

Verification to be performed on the following:
        Cluster Topology
```

```
        Cluster Resources

Verification will automatically correct verification errors.



Retrieving data from available cluster nodes.  This could take a few minutes.

        Start data collection on node clmgr1
        Start data collection on node clmgr2
        Collector on node clmgr2 completed
        Collector on node clmgr1 completed
        Data collection complete

Verifying Cluster Topology...

        Completed 10 percent of the verification checks

WARNING: Multiple communication interfaces are recommended for networks that
use IP aliasing in order to prevent the communication interface from
becoming a single point of failure. There are fewer than the recommended
number of communication interfaces defined on the following node(s) for
the given network(s):

    Node:                                  Network:
    ----------------------------------     ----------------------------------
    clmgr1                                 net_ether_01
    clmgr2                                 net_ether_01

        Completed 20 percent of the verification checks
        Completed 30 percent of the verification checks

Verifying Cluster Resources...

        Completed 40 percent of the verification checks
        Completed 50 percent of the verification checks
        Completed 60 percent of the verification checks
        Completed 70 percent of the verification checks
        Completed 80 percent of the verification checks
        Completed 90 percent of the verification checks
        Completed 100 percent of the verification checks

Remember to redo automatic error notification if configuration has changed.

Verification has completed normally.
root@clmgr1 / #
```

The command that is shown in Example B-14 on page 563 runs the synchronization and populates other cluster nodes with configuration information. You can run **`lscluster -m`** to check the CAA cluster and **`cltopinfo`** commands to check the PowerHA cluster configuration as shown in Example B-15.

*Example B-15   Output of cltopinfo and lscluster -m commands*

```
root@clmgr1 / # cltopinfo
Cluster Name: clmgr_cluster
```

```
Cluster Connection Authentication Mode: Standard
Cluster Message Authentication Mode: None
Cluster Message Encryption: None
Use Persistent Labels for Communication: No
Repository Disk: hdisk2
Cluster IP Address: 228.1.1.36
There are 2 node(s) and 2 network(s) defined
NODE clmgr1:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr1  10.1.1.54
        Network net_ether_02
                clmgr1b2        10.1.2.54
NODE clmgr2:
        Network net_ether_01
                clmgrsvc1       172.16.21.71
                clmgr2  10.1.1.55
        Network net_ether_02
                clmgr2b2        10.1.2.55

Resource Group clmgr_RG1
        Startup Policy   Online On Home Node Only
        Fallover Policy  Fallover To Next Priority Node In The List
        Fallback Policy  Never Fallback
        Participating Nodes      clmgr1 clmgr2
        Service IP Label                 clmgrsvc1
root@clmgr1 / # lscluster -m
Calling node query for all nodes
Node query number of nodes examined: 2


        Node name: clmgr1
        Cluster shorthand id for node: 1
        uuid for node: 8f03938a-73fd-11e1-951a-b6fcc07d1d6f
        State of node:  UP  NODE_LOCAL
        Smoothed rtt to node: 0
        Mean Deviation in network rtt to node: 0
        Number of zones this node is a member in: 0
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmgr_cluster      local        8f42c3ac-73fd-11e1-951a-b6fcc07d1d6f

        Number of points_of_contact for node: 0
        Point-of-contact interface & contact state
         n/a

------------------------------

        Node name: clmgr2
        Cluster shorthand id for node: 2
        uuid for node: 8f3cb804-73fd-11e1-951a-b6fcc07d1d6f
        State of node:  UP
        Smoothed rtt to node: 7
        Mean Deviation in network rtt to node: 3
        Number of zones this node is a member in: 0
```

```
        Number of clusters node is a member in: 1
        CLUSTER NAME       TYPE  SHID   UUID
        clmgr_cluster      local        8f42c3ac-73fd-11e1-951a-b6fcc07d1d6f

        Number of points_of_contact for node: 3
        Point-of-contact interface & contact state
         dpcom  DOWN  RESTRICTED
         en0  UP
         en1  UP
root@clmgr1 / #
```

## Cluster start and stop

We describe two actions that can be used to start and stop the cluster services on all or some of the cluster nodes:

- Online
- Offline

### Online

The `online` action can be used to start the cluster services on some or all cluster nodes. If you want to start the entire cluster, you can use `online cluster`. If the cluster services are to be started one node at a time or some of the nodes, you can use `online node <node#1>[,<node#2>,...]` (Example B-16).

Use this syntax1 to start the cluster services on all nodes:

**clmgr** online cluster [ WHEN={now|restart|both} MANAGE={auto|manual} BROADCAST={false|true} CLINFO={false|true|consistent} FORCE={false|true} FIX={no|yes|interactively} TIMEOUT=<*seconds_to_wait_for_completion*>]

*Example B-16 Command to start cluster services on all nodes*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr online cluster WHEN=now
MANAGE=auto BROADCAST=true CLINFO=true


/usr/es/sbin/cluster/diag/cl_ver_alias_topology[41] [[ high = high ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[41] version=%I% $Source:
61haes_r711 43haes/usr/sbin/cluster/clverify/clver/cl_ver_alias_topology.sh 2$
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[43]
UTIL_DIR=/usr/es/sbin/cluster/utilities
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[45] typeset -i status=0
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[48] typeset -i aliasing=0
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50] cut -f3 -d :
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50]
/usr/es/sbin/cluster/utilities/cllsnw -cSw
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50] [[ true = true ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] (( aliasing=aliasing+1 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] [[ true = true ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] (( aliasing=aliasing+1 ))
. . .
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[216] (( boots=boots+1 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[224] (( 1 > 1 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[334] return 0
```

```
WARNING: Multiple communication interfaces are recommended for networks that
use IP aliasing in order to prevent the communication interface from
becoming a single point of failure. There are fewer than the recommended
number of communication interfaces defined on the following node(s) for
the given network(s):

    Node:                                  Network:
    ----------------------------------    ----------------------------------
    clmgr1                                 net_ether_01
    clmgr2                                 net_ether_01


WARNING: Network option "routerevalidate" is set to 0 and will be set to 1 on
during HACMP startup on the following nodes:

        clmgr1

/usr/es/sbin/cluster/diag/clwpardata[24] [[ high == high ]]
/usr/es/sbin/cluster/diag/clwpardata[24] version='%I% $Source: 61haes_r711
43haes/usr/sbin/cluster/wpar/wpar_utils.sh 3$'
/usr/es/sbin/cluster/diag/clwpardata[26] .
/usr/es/sbin/cluster/wpar/wpar_common_funcs
/usr/es/sbin/cluster/diag/clwpardata[23] [[ high == high ]]
/usr/es/sbin/cluster/diag/clwpardata[23] set -x
. . .
/usr/es/sbin/cluster/diag/clwpardata[325] exit 0


clmgr1: start_cluster: Starting PowerHA SystemMirror
clmgr1:   3997834      -  0:00 syslogd
clmgr1: Setting routerevalidate to 1
clmgr1: 0513-059 The clevmgrdES Subsystem has been started. Subsystem PID is
9240704.
clmgr2: start_cluster: Starting PowerHA SystemMirror
clmgr2:   4390976      -  0:00 syslogd
clmgr2: Setting routerevalidate to 1
clmgr2: 0513-059 The clevmgrdES Subsystem has been started. Subsystem PID is
12845228.
clmgr2: 0513-059 The gsclvmd Subsystem has been started. Subsystem PID is
10420476.
clmgr2: 0513-059 The clinfoES Subsystem has been started. Subsystem PID is
10092648.

The cluster is now online.


:get_local_nodename[+45] [[ high = high ]]
:get_local_nodename[+45] version=%I% $Source: 61haes_r711
43haes/usr/sbin/cluster/utilities/get_local_nodename.sh 1$
:get_local_nodename[+46] :get_local_nodename[+46] cl_get_path
HA_DIR=es
:get_local_nodename[+47] :get_local_nodename[+47] cl_get_path -S
OP_SEP=~
:get_local_nodename[+49] AIXODMDIR=/etc/objrepos
```

```
:get_local_nodename[+50] HAODMDIR=/etc/es/objrepos
:get_local_nodename[+52] :get_local_nodename[+52] uname -m
UNAME=00F74D474C00
:get_local_nodename[+58] export PLATFORM=__AIX__
:get_local_nodename[+64] export ODMDIR=/etc/es/objrepos
. . .
Starting Cluster Services on node: clmgr1
This may take a few minutes.  Please wait...
clmgr1: Mar 22 2012 05:11:59 Starting execution of
/usr/es/sbin/cluster/etc/rc.cluster
clmgr1: with parameters: -boot -N -A -b -i -P cl_rc_cluster
clmgr1:
clmgr1: Mar 22 2012 05:12:01 Checking for srcmstr active...
clmgr1: Mar 22 2012 05:12:01 complete.
clmgr1: Mar 22 2012 05:12:01
clmgr1: /usr/es/sbin/cluster/utilities/clstart: called with flags -m -G -i -b -P
cl_rc_cluster -B -A
clmgr1:
clmgr1:         2012-03-22T05:12:06.251194 hats_adapter_notify
clmgr1: 2012-03-22T05:12:06.302602 hats_adapter_notify
clmgr1: Mar 22 2012 05:12:06
clmgr1: Completed execution of /usr/es/sbin/cluster/etc/rc.cluster
clmgr1: with parameters: -boot -N -A -b -i -P cl_rc_cluster.
clmgr1: Exit status = 0
clmgr1:

Starting Cluster Services on node: clmgr2
This may take a few minutes.  Please wait...
clmgr2: Mar 22 2012 05:12:06 Starting execution of
/usr/es/sbin/cluster/etc/rc.cluster
clmgr2: with parameters: -boot -N -A -b -i -P cl_rc_cluster
clmgr2:
clmgr2: Mar 22 2012 05:12:08 Checking for srcmstr active...
clmgr2: Mar 22 2012 05:12:08 complete.
clmgr2: Mar 22 2012 05:12:08
clmgr2: /usr/es/sbin/cluster/utilities/clstart: called with flags -m -G -i -b -P
cl_rc_cluster -B -A
clmgr2:
clmgr2:         2012-03-22T05:12:12.657263 hats_adapter_notify
clmgr2: 2012-03-22T05:12:12.672820 hats_adapter_notify
clmgr2: Mar 22 2012 05:12:12
clmgr2: Completed execution of /usr/es/sbin/cluster/etc/rc.cluster
clmgr2: with parameters: -boot -N -A -b -i -P cl_rc_cluster.
clmgr2: Exit status = 0
clmgr2:
root@clmgr1 / #
```

Example B-16 on page 566 shows the start of cluster services on all nodes.

Or, you can use this syntax to start services on one or some of the nodes:

```
clmgr online node <node#1>[,<node#2>,. .] [ WHEN={now|restart|both}
MANAGE={auto|manual} BROADCAST={false|true} CLINFO={false|true|consistent}
FORCE={false|true} FIX={no|yes|interactively}
TIMEOUT=<seconds_to_wait_for_completion>]
```

*Example B-17 Command to start services on one or some of the nodes*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr online node clmgr1 WHEN=now
MANAGE=auto CLINFO=true


/usr/es/sbin/cluster/diag/cl_ver_alias_topology[41] [[ high = high ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[41] version=%I% $Source:
61haes_r711 43haes/usr/sbin/cluster/clverify/clver/cl_ver_alias_topology.sh 2$
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[43]
UTIL_DIR=/usr/es/sbin/cluster/utilities
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[45] typeset -i status=0
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[48] typeset -i aliasing=0
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50] cut -f3 -d :
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50]
/usr/es/sbin/cluster/utilities/cllsnw -cSw
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[50] [[ true = true ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] (( aliasing=aliasing+1 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] [[ true = true ]]
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[54] (( aliasing=aliasing+1 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[58] (( aliasing == 0 ))
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[63] echo LOG 7322 "Verifying
cluster topology for IP aliasing.\\n"
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[67] cut -f1-3 -d :
. . .
/usr/es/sbin/cluster/diag/cl_ver_alias_topology[334] return 0

WARNING: Multiple communication interfaces are recommended for networks that
use IP aliasing in order to prevent the communication interface from
becoming a single point of failure. There are fewer than the recommended
number of communication interfaces defined on the following node(s) for
the given network(s):

    Node:                               Network:
    ----------------------------------  ----------------------------------
    clmgr1                              net_ether_01
    clmgr2                              net_ether_01


/usr/es/sbin/cluster/diag/clwpardata[24] [[ high == high ]]
/usr/es/sbin/cluster/diag/clwpardata[24] version='%I% $Source: 61haes_r711
43haes/usr/sbin/cluster/wpar/wpar_utils.sh 3$'
/usr/es/sbin/cluster/diag/clwpardata[26] .
/usr/es/sbin/cluster/wpar/wpar_common_funcs
/usr/es/sbin/cluster/diag/clwpardata[23] [[ high == high ]]
/usr/es/sbin/cluster/diag/clwpardata[23] set -x
/usr/es/sbin/cluster/diag/clwpardata[24] [[ high == high ]]
/usr/es/sbin/cluster/diag/clwpardata[24] version='%I% $Source: 61haes_r711
43haes/usr/sbin/cluster/wpar/wpar_common_funcs.sh 1$'
/usr/es/sbin/cluster/diag/clwpardata[26]
PATH=/usr/es/sbin/cluster:/usr/es/sbin/cluster/utilities:/usr/es/sbin/cluster/even
ts:/usr/es/sbin/cluster/events/utils:/usr/es/sbin/cluster/events/cmd:/usr/es/sbin/
cluster/diag:/usr/es/sbin/cluster/etc:/usr/es/sbin/cluster/sbin:/usr/es/sbin/clust
er/cspoc:/usr/es/sbin/cluster/conversion:/usr/es/sbin/cluster/events/emulate:/usr/
es/sbin/cluster/events/emulate/driver:/usr/es/sbin/cluster/events/emulate/utils:/u
```

```
sr/es/sbin/cluster/tguides/bin:/usr/es/sbin/cluster/tguides/classes:/usr/es/sbin/c
luster/tguides/images:/usr/es/sbin/cluster/tguides/scripts:/usr/es/sbin/cluster/gl
vm/utils:/usr/es/sbin/cluster/wpar:/bin:/usr/bin:/usr/es/sbin/cluster/utilities:/u
sr/es/sbin/cluster/diag:/bin:/usr/bin:/usr/sbin
/usr/es/sbin/cluster/diag/clwpardata[27] export PATH
/usr/es/sbin/cluster/diag/clwpardata[29] typeset usageErr invalArgErr internalErr
. . .
/usr/es/sbin/cluster/diag/clwpardata[325] exit 0


clmgr1: start_cluster: Starting PowerHA SystemMirror
clmgr1:   3997834      -  0:00 syslogd
clmgr1: Setting routerevalidate to 1
clmgr1: 0513-059 The clevmgrdES Subsystem has been started. Subsystem PID is
10354912.


Broadcast message from root@clmgr1 (tty) at 05:35:43 ...

Starting Event Manager (clevmgrdES) subsystem on clmgr1


clmgr1: 0513-059 The clinfoES Subsystem has been started. Subsystem PID is
14090466.


Broadcast message from root@clmgr1 (tty) at 05:35:44 ...

Starting Cluster Information Services (clinfoES) subsystem on clmgr1


"clmgr1" is now online.

:get_local_nodename[+45] [[ high = high ]]
:get_local_nodename[+45] version=%I% $Source: 61haes_r711
43haes/usr/sbin/cluster/utilities/get_local_nodename.sh 1$
:get_local_nodename[+46] :get_local_nodename[+46] cl_get_path
HA_DIR=es
:get_local_nodename[+47] :get_local_nodename[+47] cl_get_path -S
OP_SEP=~
:get_local_nodename[+49] AIXODMDIR=/etc/objrepos
:get_local_nodename[+50] HAODMDIR=/etc/es/objrepos
:get_local_nodename[+52] :get_local_nodename[+52] uname -m
UNAME=00F74D474C00
:get_local_nodename[+58] export PLATFORM=__AIX__
:get_local_nodename[+64] export ODMDIR=/etc/es/objrepos
:get_local_nodename[+66] :get_local_nodename[+66]
/usr/es/sbin/cluster/utilities/cllsclstr -N
nodename=clmgr1
:get_local_nodename[+68] :get_local_nodename[+68] cut -d: -f1
:get_local_nodename[+68] cllsnode -cS
NODENAME=clmgr1
clmgr2
:get_local_nodename[+72] [[ clmgr1 = clmgr1 ]]
:get_local_nodename[+75] print clmgr1
```

```
:get_local_nodename[+76] exit 0

Starting Cluster Services on node: clmgr1
This may take a few minutes.  Please wait...
clmgr1: Mar 22 2012 05:35:38 Starting execution of
/usr/es/sbin/cluster/etc/rc.cluster
clmgr1: with parameters: -boot -N -A -b -i -P cl_rc_cluster
clmgr1:
clmgr1: Mar 22 2012 05:35:41 Checking for srcmstr active...
clmgr1: Mar 22 2012 05:35:41 complete.
clmgr1: Mar 22 2012 05:35:41
clmgr1: /usr/es/sbin/cluster/utilities/clstart: called with flags -m -G -i -b -P
cl_rc_cluster -B -A
clmgr1:
clmgr1:         2012-03-22T05:35:44.810829 hats_adapter_notify
clmgr1: 2012-03-22T05:35:44.834661 hats_adapter_notify
clmgr1: Mar 22 2012 05:35:44
clmgr1: Completed execution of /usr/es/sbin/cluster/etc/rc.cluster
clmgr1: with parameters: -boot -N -A -b -i -P cl_rc_cluster.
clmgr1: Exit status = 0
clmgr1:
root@clmgr1 / #
```

Example B-17 on page 569 shows the start of cluster services on node `clmgr1`. You can run a similar command to start the cluster services on other nodes, as required.

## Offline

The `offline` action is used to stop the services on all or some of the cluster nodes (Example B-18).

Use this syntax to stop cluster services on all nodes:

**clmgr** offline cluster [WHEN={now|restart|both} MANAGE={offline|move|unmanage} BROADCAST={true|false} TIMEOUT=<*seconds_to_wait_for_completion*>]

*Example B-18 Command to stop cluster services on all nodes*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr offline cluster
MANAGE=offline

Warning: "WHEN" must be specified. Since it was not,
       a default of "now" will be used.


Broadcast message from root@clmgr1 (tty) at 05:30:09 ...

HACMP/ES for AIX on clmgr1 shutting down.
Please exit any cluster applications...

clmgr1: 0513-044 The clinfoES Subsystem was requested to stop.
clmgr1: 0513-044 The clevmgrdES Subsystem was requested to stop.
clmgr2: 0513-044 The clinfoES Subsystem was requested to stop.
clmgr2: 0513-044 The clevmgrdES Subsystem was requested to stop.
```

```
The cluster is now offline.


clmgr1: Mar 22 2012 05:30:09 /usr/es/sbin/cluster/utilities/clstop: called with
flags -N -g
clmgr2: Mar 22 2012 05:30:12 /usr/es/sbin/cluster/utilities/clstop: called with
flags -N -g
root@clmgr1 / #
```

Example B-18 on page 571 stops the cluster services on all nodes.

Or, use this syntax to stop cluster services on some nodes:

**clmgr** offline node <*node#1*>[,<*node#2*>,...] [ WHEN={now|restart|both} MANAGE={offline|move|unmanage} BROADCAST={false|true} FORCE={false|true} FIX={no|yes|interactively} TIMEOUT=<*seconds_to_wait_for_completion*>]

*Example B-19 Command to stop cluster services on some nodes*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr offline node clmgr1 when=NOW
BROADCAST=false MANAGE=offline
clmgr1: 0513-044 The clinfoES Subsystem was requested to stop.
clmgr1: 0513-044 The clevmgrdES Subsystem was requested to stop.


"clmgr1" is now offline.

clmgr1: Mar 22 2012 05:39:33 /usr/es/sbin/cluster/utilities/clstop: called with
flags -N -s -g
root@clmgr1 / #
```

Example B-19 shows that cluster services are stopped on node `clmgr1`. You can run a similar command to stop cluster services on other nodes, as needed.

## Cluster management

We describe several aspects of cluster management. Cluster management includes various tasks, such as snapshot creation, resource group movement, and the move service IP. We describe the following management tasks:

- Modify cluster label
- Manage cluster
- Manage node
- Move resource group
- Move service IP/label
- Manage application controller
- Add RG dependency
- Cluster snapshot

### Modify the cluster label

The `modify` action can be used to change the cluster label as shown in Example B-20 on page 573.

Use this syntax:

```
clmgr modify cluster [NAME=<new_cluster_label> NODES=<host>[,<host#2>,..]
CLUSTER_IP=<IP_address> DAILY_VERIFICATION={Enabled|Disabled}
VERIFICATION_NODE={Default|<node>} VERIFICATION_HOUR=<00..23>
VERIFICATION_DEBUGGING={Enabled|Disabled}]
```

*Example B-20 Command to modify the cluster label*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr modify cluster
NAME=my_new_cls_label

root@clmgr1 / #
```

Example B-20 changes the cluster name to the new label that is specified. For the change to occur on other cluster nodes, run the cluster verification tool.

**Important:** Do not change the cluster label or cluster name while services are running.

## Manage cluster

The `manage` action can be used to discover the cluster, reset cluster tunables, and unlock the cluster on a DARE failure as shown in Example B-21, Example B-22, and Example B-23.

Use this syntax:

```
clmgr manage cluster {discover|reset|unlock}
```

*Example B-21 Command to run discovery*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage cluster discover

root@clmgr1 / #
```

Example B-21 runs discovery to fetch PowerHA related information.

*Example B-22 Command to remove DARE locks*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage cluster unlock
cldare: Succeeded removing all DARE locks.

root@clmgr1 / #
```

*Example B-23 Command to reset cluster tunables*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage cluster reset
clsnapshot: ERROR: The Cluster Services must be stopped with either
 'Bring Resource Groups Offline' or 'Move Resource Groups' option
 for 'Select an Action on Resource Groups' field in the 'Stop Cluster
 Services' SMIT screen before resetting the cluster tunable.

root@clmgr1 / #
```

The message in Example B-23 displays when cluster tunables are reset when the cluster services are running.

**Important:** The cluster must be stopped before you can reset the cluster tunables.

### Manage node

The `manage` action for the node class can be used to undo the changes for that node as shown in Example B-24.

Use this syntax:

**clmgr** manage node undo_changes

*Example B-24 Command to undo changes to a node*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage node undo_changes

clsnapshot: Creating file
/usr/es/sbin/cluster/snapshots/Restored_From_ACD.Mar.22.06.26.24.odm.

clsnapshot: Succeeded creating Cluster Snapshot: Restored_From_ACD.Mar.22.06.26.24

Successfully restored Default Configuration from Active Configuration.
root@clmgr1 / #
```

Example B-24 restores the default configuration from the active configuration.

### Move resource group

The `move` action can be used to bring the resource group offline or online on a specific node as shown in Example B-25.

Use this syntax:

**clmgr** move resource_group <*resource_group*>,[<*rg#2*>,..] NODE=<*node_label*> [STATE={offline|online}]

*Example B-25 Command to move the resource group*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr move resource_group clmgr_RG1
NODE=clmgr2
Attempting to move resource group clmgr_RG to node clmgr2.

Waiting for the cluster to process the resource group movement request....

Waiting for the cluster to stabilize......

Resource group movement successful.
Resource group clmgr_RG is online on node clmgr2.



Cluster Name: clmgr_cluster

Resource Group Name: clmgr_RG1
Node                         Group State
---------------------------- ---------------
```

```
clmgr1                          OFFLINE
clmgr2                          ONLINE

root@clmgr1 / #
```

Example B-25 on page 574 moves the resource group `clmgr_RG` from node `clmgr1` to node `clmgr2`.

## Move service IP label

The `move` action for the service IP label moves the service address from one interface to another. This option can be used if you need to replace the adapter with a new adapter or perform a maintenance task (Example B-26).

Use this syntax:

**clmgr** move service_ip <*service_ip*> INTERFACE=<*new_interface*>

*Example B-26 Command to move the service IP*

```
root@clmgr2 / # /usr/es/sbin/cluster/utilities/clmgr move service_ip clmgrsvc1
INTERFACE=en1

swap_adapter clmgr2 net_ether_01 10.1.2.55 clmgrsvc1
swap_adapter_complete clmgr2 net_ether_01 10.1.2.55 clmgrsvc1

root@clmgr2 / #
```

Example B-26 moves the service IP/label to the new interface `en1`.

## Manage application controller

The `manage` action for the application controller can be used to suspend or resume application server monitoring for a specific application. Or, it can be used to suspend or resume application server monitoring for all application servers in the resource group as shown in Example B-27.

Use this syntax:

**clmgr** manage application_controller {suspend|resume}
Resource_Group=<*resource_group*> <*application_controller*> | ALL

*Example B-27 Command to suspend application monitoring*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage application_controller
suspend test_app1 RESOURCE_GROUP="clmgr_RG1"
2012-03-23T20:41:08.570866
2012-03-23T20:41:08.579757
monitor1

Mar 23 2012 20:41:08 cl_RMupdate: Completed request to suspend monitor(s) for
application test_app1.
Mar 23 2012 20:41:08 cl_RMupdate: The following monitor(s) are in use for
application test_app1:
root@clmgr1 / #
```

Example B-27 on page 575 suspends application monitoring.

*Example B-28 Command to resume application monitoring*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr manage application_controller
resume test_app1 RESOURCE_GROUP="clmgr_RG2"
2012-03-23T20:41:08.570866
2012-03-23T20:41:08.579757
monitor1

Mar 23 2012 20:41:08 cl_RMupdate: Completed request to suspend monitor(s) for
application test_app1.
Mar 23 2012 20:41:08 cl_RMupdate: The following monitor(s) are in use for
application test_app1:

root@clmgr1 / #
```

Example B-28 resumes application monitoring.

## Add a resource group dependency

The `add` action for the dependency class can be used to add resource group dependencies, such as parent-child, node collocation, Anti collocation, and start/stop after (Example B-29).

Use this syntax:

- `# Temporal dependency (parent ==> child)`

  **clmgr** add dependency PARENT=<*rg#1*> CHILD="<*rg#2*>[,<*rg#3*>,...]"

- `# Temporal dependency (start/stop after)`

  **clmgr** add dependency {STOP|START}="<*rg#2*>[,<*rg#3*>,...]" AFTER=<*rg#1*>

- `# Temporal dependency (collocation)`

  **clmgr** add dependency SAME={NODE} GROUPS="<*rg1*>,<*rg2*>[,<*rg#n*>]"

- `# Temporal dependency (anti-collocation)`

  **clmgr** add dependency SAME={NODE} GROUPS="<*rg1*>,<*rg2*>[,<*rg#n*>]"

*Example B-29 Command to add parent-child dependency*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add dependency
PARENT="clmgr_RG1" CHILD="clmgr_RG2"

root@clmgr1 / #
```

Example B-29 adds a parent-child dependency for the resource groups.

*Example B-30 Command to add node-collocation dependency*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add dependency SAME=NODE
GROUPS="clmgr_RG1 clmgr_RG2"
```

```
root@clmgr1 / #
```

Example B-30 on page 576 adds a same-node (collocation) dependency between resource groups.

## Cluster snapshot

To create a cluster snapshot, we can use the `add` action with the CLASS `snapshot` as shown in Example B-31.

Use this syntax:

**clmgr** add snapshot <*snapshot-name*> DESCRIPTION=”<*snapshot-description*>” [METHODS=”method1, method2, .. “] [SAVE_LOGS={false|true}]

*Example B-31 Command to create the cluster snapshot*

```
root@clmgr1 / # /usr/es/sbin/cluster/utilities/clmgr add snapshot mySNAP
DESCRIPTION="my first snapshot"

clsnapshot: Creating file /usr/es/sbin/cluster/snapshots/mySNAP.odm.

clsnapshot: Creating file /usr/es/sbin/cluster/snapshots/mySNAP.info.

clsnapshot: Executing clsnapshotinfo command on node: clmgr1...

clsnapshot: Executing clsnapshotinfo command on node: clmgr2...

clsnapshot: Succeeded creating Cluster Snapshot: mySNAP
root@clmgr1 / #
```

Example B-31 creates the cluster snapshot.

C

# Creating the hardware environment by using command-line tools

In this appendix, all scripts that are used to create the hardware environment that is used in this book are shown and explained. In some client environments, there are security or bandwidth network restrictions that do not allow system administrators to use graphical tools to perform tasks. Also, sometimes the environment is too large to be created by using graphical tools.

So, we focus on complex environment creation by using the command-line tools only. No graphical tool is used; only scripts are generated to cover all required tasks.

These scripts cover the following tasks:

- Virtual I/O Server (VIOS) creation that includes Shared Ethernet Adapters (SEA) and N-Port ID Virtualization (NPIV) configuration
- Client logical partition (LPARs) creation that includes Shared Ethernet Adapters (SEA) and N-Port ID Virtualization (NPIV) configuration
- Network Installation Management (NIM) client definition and NIM BOS installation procedures
- Storage area network (SAN) zoning
- DS4800 array, logical volume, and host group creation

## Hardware details

We describe the hardware configuration details (Figure C-1).

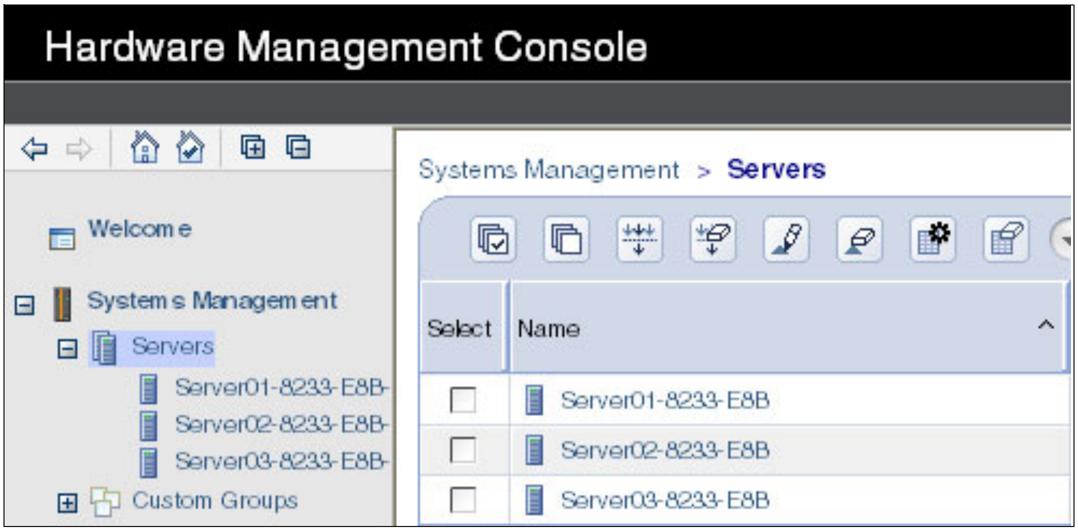



*Figure C-1 Power 750 servers are used in the hardware environment*

As shown in Figure C-1 and explained in Chapter 1, "Introducing IBM PowerHA SystemMirror 7.1.1" on page 1, we use three Power 750 servers as part of the hardware environment. For more details about the hardware design, see 1.2, "Hardware environment" on page 23.

> **Important:** Because all server configurations are similar, this appendix only covers the configuration scripts for *Server01-8233-E8B*. The same scripts are used on the other two Power 750 servers, only we change the server and LPAR names.

## Virtual I/O Server creation

Because the hardware environment used in this publication is virtualized, the first partitions that are created are the VIOS partitions.

> **Important:** All VIOS creation commands must be used in the Hardware Management Console (HMC) shell.

Example C-1 shows the scripts that are perform all tasks for the VIOS creation.

*Example C-1 VIOS profile creation*

```
mksyscfg -r lpar -m Server01-8233-E8B -i
'name=VIOS01.01,profile_name=VIOS01.01,lpar_env=vioserver,min_mem=2048,desired_mem
=2048,max_mem=8192,mem_mode=ded,proc_mode=shared,min_proc_units=1.0,desired_proc_u
nits=1.0,max_proc_units=16.0,min_procs=1,desired_procs=1,max_procs=16,sharing_mode
=uncap,uncap_weight=128,lpar_io_pool_ids=none,max_virtual_slots=150,"virtual_seria
l_adapters=0/server/1/any//any/1,1/server/1/any//any/1",virtual_scsi_adapters=none
,virtual_eth_adapters=none,vtpm_adapters=none,hca_adapters=none,"virtual_fc_adapte
rs=3/server/3//31//1,4/server/4//41//1,5/server/5//51//1,6/server/6//61//1,7/serve
r/7//71//1,8/server/8//81//1,9/server/9//91//1,10/server/10//101//1",boot_mode=nor
m,conn_monitoring=0,auto_start=0,lpar_proc_compat_mode=default'
```

```
mksyscfg -r lpar -m Server01-8233-E8B -i
'name=VIOS02.01,profile_name=VIOS02.01,lpar_env=vioserver,min_mem=2048,desired_mem
=2048,max_mem=8192,mem_mode=ded,proc_mode=shared,min_proc_units=1.0,desired_proc_u
nits=1.0,max_proc_units=16.0,min_procs=1,desired_procs=1,max_procs=16,sharing_mode
=uncap,uncap_weight=128,lpar_io_pool_ids=none,max_virtual_slots=150,"virtual_seria
l_adapters=0/server/1/any//any/1,1/server/1/any//any/1",virtual_scsi_adapters=none
,virtual_eth_adapters=none,vtpm_adapters=none,hca_adapters=none,"virtual_fc_adapte
rs=3/server/3//32//1,4/server/4//42//1,5/server/5//52//1,6/server/6//62//1,7/serve
r/7//72//1,8/server/8//82//1,9/server/9//92//1,10/server/10//102//1",boot_mode=nor
m,conn_monitoring=0,auto_start=0,lpar_proc_compat_mode=default'
```

Example C-2 shows the steps to create virtual components in the VIOS: Virtual SCSI, virtual Ethernet, and Virtual Fibre Channel Adapters for client LPARs.

*Example C-2 Virtual component creation for VIOS*

```
#Virtual SCSI Adapters

chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS01.01,lpar_id=1,"virtual_scsi_adapters=30/server/3/LPAR01.01/34/1,40/ser
ver/4/LPAR02.01/44/1,50/server/5/LPAR03.01/54/1,60/server/6/LPAR04.01/64/1,70/serv
er/7/LPAR05.01/74/1,80/server/8/LPAR06.01/84/1,90/server/9/LPAR07.01/94/1,100/serv
er/10/LPAR08.01/104/1,110/server/11/LPAR09.01/114/1,120/server/12/LPAR10.01/124/1"
'
chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS02.01,lpar_id=2,"virtual_scsi_adapters=30/server/3/LPAR01.01/35/1,40/ser
ver/4/LPAR02.01/45/1,50/server/5/LPAR03.01/55/1,60/server/6/LPAR04.01/65/1,70/serv
er/7/LPAR05.01/75/1,80/server/8/LPAR06.01/85/1,90/server/9/LPAR07.01/95/1,100/serv
er/10/LPAR08.01/105/1,110/server/11/LPAR09.01/115/1,120/server/12/LPAR10.01/125/1"
'

# Virtual Ethernet Adapters
chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS01.01,lpar_id=1,"virtual_eth_adapters=111/0/1//1/1/ETHERNET0//all/none,1
13/0/10//1/1/ETHERNET0//all/none,112/0/99//0/1/ETHERNET0//all/none,114/0/98//0/1/E
THERNET0//all/none"'

chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS02.01,lpar_id=2,"virtual_eth_adapters=111/0/1//1/1/ETHERNET0//all/none,1
13/0/10//1/1/ETHERNET0//all/none,112/0/99//0/1/ETHERNET0//all/none,114/0/98//0/1/E
THERNET0//all/none"'

#Fibre Channel NPIV Adapters
chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS01.01,lpar_id=1,"virtual_fc_adapters=3/server/3/LPAR01.01/31//1,4/server
/4/LPAR02.01/41//1,5/server/5/LPAR03.01/51//1,6/server/6/LPAR04.01/61//1,7/server/
7/LPAR05.01/71//1,8/server/8/LPAR06.01/81//1,9/server/9/LPAR07.01/91//1,10/server/
10/LPAR08.01/101//1,11/server/11/LPAR09.01/121//1,12/server/12/LPAR10.01/131//1"'

chsyscfg -r prof -m Server01-8233-E8B -i
'name=VIOS02.01,lpar_id=2,"virtual_fc_adapters=3/server/3/LPAR01.01/32//1,4/server
/4/LPAR02.01/42//1,5/server/5/LPAR03.01/52//1,6/server/6/LPAR04.01/62//1,7/server/
7/LPAR05.01/72//1,8/server/8/LPAR06.01/82//1,9/server/9/LPAR07.01/92//1,10/server/
10/LPAR08.01/102//1,11/server/11/LPAR09.01/122//1,12/server/12/LPAR10.01/132//1"'
```

## Client LPAR creation

We show all commands to create the client LPARs that are used across this book.

**HMC:** All client LPAR creation commands must be used in the HMC shell.

Example C-3 shows the commands that are used to create the client LPAR profiles.

*Example C-3 Client LPAR profile creation*

```
mksyscfg -r lpar -m Server01-8233-E8B -i
'name=LPAR01.01,profile_name=LPAR01.01,lpar_env=aixlinux,min_mem=10240,desired_mem
=10240,max_mem=20480,mem_mode=ded,proc_mode=shared,min_proc_units=1.0,desired_proc
_units=1.0,max_proc_units=16.0,min_procs=4,desired_procs=4,max_procs=64,sharing_mo
de=uncap,uncap_weight=128,lpar_io_pool_ids=none,max_virtual_slots=150,boot_mode=no
rm,conn_monitoring=0,auto_start=0,lpar_proc_compat_mode=default'

mksyscfg -r lpar -m Server01-8233-E8B -i
'name=LPAR02.01,profile_name=LPAR02.01,lpar_env=aixlinux,min_mem=10240,desired_mem
=10240,max_mem=20480,mem_mode=ded,proc_mode=shared,min_proc_units=1.0,desired_proc
_units=1.0,max_proc_units=16.0,min_procs=4,desired_procs=4,max_procs=64,sharing_mo
de=uncap,uncap_weight=128,lpar_io_pool_ids=none,max_virtual_slots=150,boot_mode=no
rm,conn_monitoring=0,auto_start=0,lpar_proc_compat_mode=default'
```

Example C-4 shows the commands that are used to create virtual adapters for client LPARs in our environment.

*Example C-4 Creating virtual adapters in the client LPARs*

```
# Virtual SCSI Adapters
chsyscfg -r prof -m Server01-8233-E8B-SN061AB2P -i
'name=LPAR01.01,lpar_id=3,"virtual_scsi_adapters=34/client/1/VIOS01.01/30/1,35/cli
ent/2/VIOS02.01/30/1"'

chsyscfg -r prof -m Server01-8233-E8B-SN061AB2P -i
'name=LPAR02.01,lpar_id=4,"virtual_scsi_adapters=44/client/1/VIOS01.01/40/1,45/cli
ent/2/VIOS02.01/40/1"'

# Virtual Ethernet Adapters

chsyscfg -r prof -m Server01-8233-E8B-SN061AB2P -i
'name=LPAR01.01,lpar_id=3,"virtual_eth_adapters=111/0/1//0/1/ETHERNET0//all/none,1
12/0/10//0/1/ETHERNET0//all/none"'
chsyscfg -r prof -m Server01-8233-E8B-SN061AB2P -i
'name=LPAR02.01,lpar_id=4,"virtual_eth_adapters=111/0/1//0/1/ETHERNET0//all/none,1
12/0/10//0/1/ETHERNET0//all/none"'

# Fibre Channel NPIV Adapters

chsyscfg -r prof -m Server01-8233-E8B -i
'name=LPAR01.01,lpar_id=3,virtual_fc_adapters="31/client/1/VIOS01.01/11//1,32/clie
nt/2/VIOS02.01/11//1"'
```

```
chsyscfg -r prof -m Server01-8233-E8B -i
'name=LPAR02.01,lpar_id=4,virtual_fc_adapters="41/client/1/VIOS01.01/12//1,42/clie
nt/2/VIOS02.01/12//1"'
```

## NIM client definition

After all profiles are correctly set up in the HMC, the next step is to prepare NIM to create new LPARs.

> **NIM:** For more information about the NIM environment configuration and concepts, see the NIM documentation at the IBM information center:
>
> http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.install/doc/insgdrf/nim_intro.htm

Example C-5 shows the steps in the HMC and in the NIM master server to allow all LPAR configuration and BOS installation.

*Example C-5 HMC and NIM master server steps*

```
# Getting LPARs ethernet adapter physical addresses using HMC

lpar_netboot -M -n -t ent "LPAR01.01" "LPAR01.01" "Server01-8233-E8B"

# That command shows an output similar to that below:

 # Connecting to LPAR01.01
 # Connected
 # Checking for power off.
 # Power off complete.
 # Power on LPAR01.01 to Open Firmware.
 # Power on complete.
 # Getting adapter location codes.
 # Type  Location Code   MAC Address
 ent U9124.720.100486A-V5-C2-T1 bad3f0005002

# Using the MAC address above, register server on NIM Server:

nim -o define -t'standalone' -a platform=chrp -a netboot_kernel=64 -a if1="nim172
lpar0101 bad3f0005002 ent" -a cable_type1=bnc lpar0101
```

After the NIM client definition, the BOS installation procedure needs to be performed as shown in Example C-6 on page 584. After you perform all required configuration steps (Network File System (NFS) share permissions, bootp, and tftp configuration files), the BOS installation on the NIM master is correctly set up.

*Example C-6 Performing BOS Installation operation on the NIM master*

```
# The following command performs all steps required for a BOS Install

nim -o bos_inst -a source=spot -a spot=aix71 -a lpp_source=aix71_full -a
accept_licenses=yes -a bosinst_data=bosinst-jfs2-64 -a preserve_res=yes -a
no_client_boot=yes -a set_bootlist=no -a force_push=no lpar0101
```

With the NIM client correctly defined and the NIM BOS installed, perform the client LPAR network boot as shown in Example C-7.

*Example C-7 Network boot on client LPARs*

```
# Run the following command from the HMC command-line prompt

lpar_netboot -t ent -m bad3f0005002 -s auto -d auto -S 172.16.20.40 -G 172.16.20.1
-C 172.16.21.26 "LPAR01.01" "LPAR01.01" "Server01-8233-E8B"
```

> **Important:** Because many questions are asked during a standard AIX operating system installation, the NIM environment allows the use of `bosinst_data scripts`. These scripts are used to answer the mandatory AIX operating system installation questions. The `bosinst_data script` that is used to build the environment is in Example C-8.

*Example C-8 bosinst_data script*

```
root /nimrepo/scripts # cat bosinst.data.jfs264
control_flow:
   CONSOLE = Default
   INSTALL_METHOD = overwrite
   PROMPT = no
   EXISTING_SYSTEM_OVERWRITE = yes
   RUN_STARTUP = no
   RM_INST_ROOTS = no
   ERROR_EXIT =
   CUSTOMIZATION_FILE =
   TCB = no
   BUNDLES =
   RECOVER_DEVICES = Default
   BOSINST_DEBUG = no
   ACCEPT_LICENSES = yes
   INSTALL_CONFIGURATION =
   DESKTOP = CDE
        INSTALL_DEVICES_AND_UPDATES = yes
        IMPORT_USER_VGS = yes
        ENABLE_64BIT_KERNEL = yes
        CREATE_JFS2_FS = yes
        ALL_DEVICES_KERNELS = yes
        GRAPHICS_BUNDLE = no
        DOC_SERVICES_BUNDLE = no
        NETSCAPE_BUNDLE = yes
        HTTP_SERVER_BUNDLE = yes
        KERBEROS_5_BUNDLE = yes
        SERVER_BUNDLE = yes
        ALT_DISK_INSTALL_BUNDLE = yes
        REMOVE_JAVA_118 = no
target_disk_data:
```

```
   PVID =
   CONNECTION =
   LOCATION =
   SIZE_MB =
   HDISKNAME = hdisk0

locale:
   BOSINST_LANG = en_US
   CULTURAL_CONVENTION = en_US
   MESSAGES = en_US
   KEYBOARD = en_US
large_dumplv:
   DUMPDEVICE = lg_dumplv
   SIZE_GB = 1
```

# SAN zoning

With the client LPARs correctly installed, configure the SAN environment and allocate the SAN logical unit number (LUN) disks that are required for the cluster scenarios.

On the hardware environment as described in 1.2, “Hardware environment” on page 23, we use two IBM 2005-B16 switches.

**Switch models:** All switch syntax commands that are shown work for the same switch model and all compatible models.

## 8.6.2 Alias creation

The first step is the creation of aliases for all hosts and storage ports on the SAN fabric. *All aliases must relate to only the hosts that attach to the switch that is configured*. To keep things simple, we grouped all aliases together as shown in Example C-9.

*Example C-9 Alias creation on SAN switches*

```
# The commands below create the alias definitions on IBM 2005-B16 switches

## Switch 01
# Aliases Creation

alicreate DS4800_CA_CH1,20:12:00:a0:b8:11:a6:62
alicreate lpar0101_fcs0,c0:50:76:03:03:9e:00:5e
alicreate lpar0201_fcs0,c0:50:76:03:03:9e:00:60
alicreate lpar0301_fcs0,c0:50:76:03:03:9e:00:64
alicreate lpar0401_fcs0,c0:50:76:03:03:9e:00:68
alicreate lpar0501_fcs1,c0:50:76:03:03:9e:00:6e
alicreate lpar0601_fcs1,c0:50:76:03:03:9e:00:72
alicreate lpar0701_fcs1,c0:50:76:03:03:9e:00:76
alicreate lpar0801_fcs1,c0:50:76:03:03:9e:00:7a
alicreate lpar0901_fcs0,c0:50:76:03:03:9e:00:7c
alicreate lpar1001_fcs0,c0:50:76:03:03:9e:00:80
alicreate lpar0102_fcs0,c0:50:76:05:06:19:00:02
alicreate lpar0202_fcs0,c0:50:76:05:06:19:00:04
```

```
alicreate lpar0302_fcs0,c0:50:76:05:06:19:00:08
alicreate lpar0402_fcs0,c0:50:76:05:06:19:00:0c
alicreate lpar0502_fcs1,c0:50:76:05:06:19:00:12
alicreate lpar0602_fcs1,c0:50:76:05:06:19:00:16
alicreate lpar0702_fcs1,c0:50:76:05:06:19:00:1a
alicreate lpar0802_fcs1,c0:50:76:05:06:19:00:1e
alicreate lpar0902_fcs0,c0:50:76:05:06:19:00:20
alicreate lpar1002_fcs0,c0:50:76:05:06:19:00:24
alicreate lpar0103_fcs0,c0:50:76:05:06:1a:00:02
alicreate lpar0203_fcs0,c0:50:76:05:06:1a:00:04
alicreate lpar0303_fcs0,c0:50:76:05:06:1a:00:08
alicreate lpar0403_fcs0,c0:50:76:05:06:1a:00:0c
alicreate lpar0503_fcs1,c0:50:76:05:06:1a:00:12
alicreate lpar0603_fcs1,c0:50:76:05:06:1a:00:16
alicreate lpar0703_fcs1,c0:50:76:05:06:1a:00:1a
alicreate lpar0803_fcs1,c0:50:76:05:06:1a:00:1e
alicreate lpar0903_fcs0,c0:50:76:05:06:1a:00:20
alicreate lpar1003_fcs0,c0:50:76:05:06:1a:00:24

# Zones Creation

zonecreate lpar0101_fcs0_DS4800_CA_CH1,"lpar0101_fcs0;DS4800_CA_CH1"
zonecreate lpar0201_fcs0_DS4800_CA_CH1,"lpar0201_fcs0;DS4800_CA_CH1"
zonecreate lpar0301_fcs0_DS4800_CA_CH1,"lpar0301_fcs0;DS4800_CA_CH1"
zonecreate lpar0401_fcs0_DS4800_CA_CH1,"lpar0401_fcs0;DS4800_CA_CH1"
zonecreate lpar0501_fcs1_DS4800_CA_CH1,"lpar0501_fcs1;DS4800_CA_CH1"
zonecreate lpar0601_fcs1_DS4800_CA_CH1,"lpar0601_fcs1;DS4800_CA_CH1"
zonecreate lpar0701_fcs1_DS4800_CA_CH1,"lpar0701_fcs1;DS4800_CA_CH1"
zonecreate lpar0801_fcs1_DS4800_CA_CH1,"lpar0801_fcs1;DS4800_CA_CH1"
zonecreate lpar0901_fcs0_DS4800_CA_CH1,"lpar0901_fcs0;DS4800_CA_CH1"
zonecreate lpar1001_fcs0_DS4800_CA_CH1,"lpar1001_fcs0;DS4800_CA_CH1"
zonecreate lpar0102_fcs0_DS4800_CA_CH1,"lpar0102_fcs0;DS4800_CA_CH1"
zonecreate lpar0202_fcs0_DS4800_CA_CH1,"lpar0202_fcs0;DS4800_CA_CH1"
zonecreate lpar0302_fcs0_DS4800_CA_CH1,"lpar0302_fcs0;DS4800_CA_CH1"
zonecreate lpar0402_fcs0_DS4800_CA_CH1,"lpar0402_fcs0;DS4800_CA_CH1"
zonecreate lpar0502_fcs1_DS4800_CA_CH1,"lpar0502_fcs1;DS4800_CA_CH1"
zonecreate lpar0602_fcs1_DS4800_CA_CH1,"lpar0602_fcs1;DS4800_CA_CH1"
zonecreate lpar0602_fcs1_lpar0603_fcs1,"lpar0602_fcs1;lpar0603_fcs1"
zonecreate lpar0702_fcs1_DS4800_CA_CH1,"lpar0702_fcs1;DS4800_CA_CH1"
zonecreate lpar0802_fcs1_DS4800_CA_CH1,"lpar0802_fcs1;DS4800_CA_CH1"
zonecreate lpar0902_fcs0_DS4800_CA_CH1,"lpar0902_fcs0;DS4800_CA_CH1"
zonecreate lpar1002_fcs0_DS4800_CA_CH1,"lpar1002_fcs0;DS4800_CA_CH1"
zonecreate lpar0103_fcs0_DS4800_CA_CH1,"lpar0103_fcs0;DS4800_CA_CH1"
zonecreate lpar0203_fcs0_DS4800_CA_CH1,"lpar0203_fcs0;DS4800_CA_CH1"
zonecreate lpar0303_fcs0_DS4800_CA_CH1,"lpar0303_fcs0;DS4800_CA_CH1"
zonecreate lpar0403_fcs0_DS4800_CA_CH1,"lpar0403_fcs0;DS4800_CA_CH1"
zonecreate lpar0503_fcs1_DS4800_CA_CH1,"lpar0503_fcs1;DS4800_CA_CH1"
zonecreate lpar0603_fcs1_DS4800_CA_CH1,"lpar0603_fcs1;DS4800_CA_CH1"
zonecreate lpar0703_fcs1_DS4800_CA_CH1,"lpar0703_fcs1;DS4800_CA_CH1"
zonecreate lpar0803_fcs1_DS4800_CA_CH1,"lpar0803_fcs1;DS4800_CA_CH1"
zonecreate lpar0903_fcs0_DS4800_CA_CH1,"lpar0903_fcs0;DS4800_CA_CH1"
zonecreate lpar1003_fcs0_DS4800_CA_CH1,"lpar1003_fcs0;DS4800_CA_CH1"

# Switch Configuration
```

```
cfgcreate ITSO_2005,lpar0101_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0201_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0301_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0401_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0501_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0601_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0701_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0801_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0901_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar1001_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0102_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0202_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0302_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0402_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0502_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0602_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0702_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0802_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0902_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar1002_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0103_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0203_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0303_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0403_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0503_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0603_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0703_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0803_fcs1_DS4800_CA_CH1
cfgadd ITSO_2005,lpar0903_fcs0_DS4800_CA_CH1
cfgadd ITSO_2005,lpar1003_fcs0_DS4800_CA_CH1

cfgvave // run this command from switch console

You are about to save the Defined zoning configuration. This
action will only save the changes on Defined configuration.
Any changes made on the Effective configuration will not
take effect until it is re-enabled.
Do you want to save Defined zoning configuration only?  (yes, y, no, n): [no] y

cfgenable ITSO_2005 // run this command from switch console

You are about to enable a new zoning configuration.
This action will replace the old zoning configuration with the
current configuration selected. If the update includes changes
to one or more traffic isolation zones, the update may result in
localized disruption to traffic on ports associated with
the traffic isolation zone changes
Do you want to enable 'ITSO_2005' configuration  (yes, y, no, n): [no] y

## Switch 02

# Aliases Creation

alicreate DS4800_CB_CH1,20:13:00:a0:b8:11:a6:62
alicreate DS8000_02,50:05:07:63:04:14:c1:2c
```

```
alicreate lpar0101_fcs1,c0:50:76:03:03:9e:00:5c
alicreate lpar0201_fcs1,c0:50:76:03:03:9e:00:62
alicreate lpar0301_fcs1,c0:50:76:03:03:9e:00:66
alicreate lpar0401_fcs1,c0:50:76:03:03:9e:00:6a
alicreate lpar0901_fcs1,c0:50:76:03:03:9e:00:7e
alicreate lpar0501_fcs0,c0:50:76:03:03:9e:00:6c
alicreate lpar0601_fcs0,c0:50:76:03:03:9e:00:70
alicreate lpar0701_fcs0,c0:50:76:03:03:9e:00:74
alicreate lpar0801_fcs0,c0:50:76:03:03:9e:00:78
alicreate lpar1001_fcs1,c0:50:76:03:03:9e:00:82
alicreate lpar0102_fcs1,c0:50:76:05:06:19:00:00
alicreate lpar0202_fcs1,c0:50:76:05:06:19:00:06
alicreate lpar0302_fcs1,c0:50:76:05:06:19:00:0a
alicreate lpar0402_fcs1,c0:50:76:05:06:19:00:0e
alicreate lpar0502_fcs0,c0:50:76:05:06:19:00:10
alicreate lpar0602_fcs0,c0:50:76:05:06:19:00:14
alicreate lpar0702_fcs0,c0:50:76:05:06:19:00:18
alicreate lpar0802_fcs0,c0:50:76:05:06:19:00:1c
alicreate lpar0902_fcs1,c0:50:76:05:06:19:00:22
alicreate lpar1002_fcs1,c0:50:76:05:06:19:00:26
alicreate lpar0103_fcs1,c0:50:76:05:06:1a:00:00
alicreate lpar0203_fcs1,c0:50:76:05:06:1a:00:06
alicreate lpar0303_fcs1,c0:50:76:05:06:1a:00:0a
alicreate lpar0403_fcs1,c0:50:76:05:06:1a:00:0e
alicreate lpar0503_fcs0,c0:50:76:05:06:1a:00:10
alicreate lpar0603_fcs0,c0:50:76:05:06:1a:00:14
alicreate lpar0703_fcs0,c0:50:76:05:06:1a:00:18
alicreate lpar0803_fcs0,c0:50:76:05:06:1a:00:1c
alicreate lpar0903_fcs1,c0:50:76:05:06:1a:00:22
alicreate lpar1003_fcs1,c0:50:76:05:06:1a:00:26

# Zones Creation

zonecreate lpar0101_fcs1_DS4800_CB_CH1,"lpar0101_fcs1;DS4800_CB_CH1"
zonecreate lpar0201_fcs1_DS4800_CB_CH1,"lpar0201_fcs1;DS4800_CB_CH1"
zonecreate lpar0301_fcs1_DS4800_CB_CH1,"lpar0301_fcs1;DS4800_CB_CH1"
zonecreate lpar0401_fcs1_DS4800_CB_CH1,"lpar0401_fcs1;DS4800_CB_CH1"
zonecreate lpar0501_fcs0_DS4800_CB_CH1,"lpar0501_fcs0;DS4800_CB_CH1"
zonecreate lpar0601_fcs0_DS4800_CB_CH1,"lpar0601_fcs0;DS4800_CB_CH1"
zonecreate lpar0701_fcs0_DS4800_CB_CH1,"lpar0701_fcs0;DS4800_CB_CH1"
zonecreate lpar0801_fcs0_DS4800_CB_CH1,"lpar0801_fcs0;DS4800_CB_CH1"
zonecreate lpar0901_fcs1_DS4800_CB_CH1,"lpar0901_fcs1;DS4800_CB_CH1"
zonecreate lpar1001_fcs1_DS4800_CB_CH1,"lpar1001_fcs1;DS4800_CB_CH1"
zonecreate lpar0102_fcs1_DS4800_CB_CH1,"lpar0102_fcs1;DS4800_CB_CH1"
zonecreate lpar0202_fcs1_DS4800_CB_CH1,"lpar0202_fcs1;DS4800_CB_CH1"
zonecreate lpar0302_fcs1_DS4800_CB_CH1,"lpar0302_fcs1;DS4800_CB_CH1"
zonecreate lpar0402_fcs1_DS4800_CB_CH1,"lpar0402_fcs1;DS4800_CB_CH1"
zonecreate lpar0502_fcs0_DS4800_CB_CH1,"lpar0502_fcs0;DS4800_CB_CH1"
zonecreate lpar0602_fcs0_DS4800_CB_CH1,"lpar0602_fcs0;DS4800_CB_CH1"
zonecreate lpar0702_fcs0_DS4800_CB_CH1,"lpar0702_fcs0;DS4800_CB_CH1"
zonecreate lpar0802_fcs0_DS4800_CB_CH1,"lpar0802_fcs0;DS4800_CB_CH1"
zonecreate lpar0902_fcs1_DS4800_CB_CH1,"lpar0902_fcs1;DS4800_CB_CH1"
zonecreate lpar1002_fcs1_DS4800_CB_CH1,"lpar1002_fcs1;DS4800_CB_CH1"
zonecreate lpar0103_fcs1_DS4800_CB_CH1,"lpar0103_fcs1;DS4800_CB_CH1"
zonecreate lpar0203_fcs1_DS4800_CB_CH1,"lpar0203_fcs1;DS4800_CB_CH1"
```

```
zonecreate lpar0303_fcs1_DS4800_CB_CH1,"lpar0303_fcs1;DS4800_CB_CH1"
zonecreate lpar0403_fcs1_DS4800_CB_CH1,"lpar0403_fcs1;DS4800_CB_CH1"
zonecreate lpar0503_fcs0_DS4800_CB_CH1,"lpar0503_fcs0;DS4800_CB_CH1"
zonecreate lpar0603_fcs0_DS4800_CB_CH1,"lpar0603_fcs0;DS4800_CB_CH1"
zonecreate lpar0703_fcs0_DS4800_CB_CH1,"lpar0703_fcs0;DS4800_CB_CH1"
zonecreate lpar0803_fcs0_DS4800_CB_CH1,"lpar0803_fcs0;DS4800_CB_CH1"
zonecreate lpar0903_fcs1_DS4800_CB_CH1,"lpar0903_fcs1;DS4800_CB_CH1"
zonecreate lpar1003_fcs1_DS4800_CB_CH1,"lpar1003_fcs1;DS4800_CB_CH1"

# Switch Configuration

cfgcreate ITSO_2005,lpar0101_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0201_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0301_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0401_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0501_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0601_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0701_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0801_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0901_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar1001_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0102_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0202_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0302_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0402_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0502_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0602_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0702_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0802_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0902_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0103_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0203_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0303_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0403_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0503_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0603_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0703_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0803_fcs0_DS4800_CB_CH1
cfgadd ITSO_2005,lpar0903_fcs1_DS4800_CB_CH1
cfgadd ITSO_2005,lpar1003_fcs1_DS4800_CB_CH1

cfgvave // run this command from switch console

You are about to save the Defined zoning configuration. This
action will only save the changes on Defined configuration.
Any changes made on the Effective configuration will not
take effect until it is re-enabled.
Do you want to save Defined zoning configuration only?  (yes, y, no, n): [no] y

cfgenable ITSO_2005 // run this command from switch console

You are about to enable a new zoning configuration.
This action will replace the old zoning configuration with the
current configuration selected. If the update includes changes
to one or more traffic isolation zones, the update may result in
```

```
localized disruption to traffic on ports associated with
the traffic isolation zone changes
Do you want to enable 'ITSO_2005' configuration  (yes, y, no, n): [no] y
```

## IBM DS4800 arrays, logical volumes, and host group creation

The last task that is executed to build the environment that we used in this book relates to the DS4800 storage system. Manually, as a first step, we created one large array, `PW2201_Array`, with all DS4800 disks. All subsequent tasks are performed by using this array to create the logical volumes.

> **Important:** In a production environment, you need to assess the correct way to create arrays and logical volumes. Ensure that I/O-bound applications do not access the same physical disks and affect performance.

You must execute all of these commands from the *Storage Manager for DS4000 script editor*. For more information about the Storage Manager tool, see *IBM System Storage Solutions Handbook*, SG24-5250:

http://www.redbooks.ibm.com/abstracts/sg245250.html?Open

We configure the following DS4800 objects (Example C-10):

- DS4800 host groups
- DS4800 hosts
- DS4800 logical volumes/LUN disks

*Example C-10 Host group and host creation*

```
# Host Groups and Hosts Creation

create hostGroup userLabel="SAP-DB2-Cluster";
create host userLabel="lpar0101" hostGroup="SAP-DB2-Cluster";
create hostPort identifier="c0507603039e005e" userLabel="lpar0101-fcs0"
host="lpar0101" interfaceType=FC;
create hostPort identifier="c0507603039e005c" userLabel="lpar0101-fcs1"
host="lpar0101" interfaceType=FC;
create host userLabel="lpar0103" hostGroup="SAP-DB2-Cluster";
create hostPort identifier="c0507605061a0002" userLabel="lpar0103-fcs0"
host="lpar0103" interfaceType=FC;
create hostPort identifier="c0507605061a0000" userLabel="lpar0103-fcs1"
host="lpar0103" interfaceType=FC;

create hostGroup userLabel="SAP-DBCI-Cluster";
create host userLabel="lpar0102" hostGroup="SAP-DBCI-Cluster";
create hostPort identifier="c050760506190002" userLabel="lpar0102-fcs0"
host="lpar0102" interfaceType=FC;
create hostPort identifier="c050760506190000" userLabel="lpar0102-fcs1"
host="lpar0102" interfaceType=FC;
create host userLabel="lpar0201" hostGroup="SAP-DBCI-Cluster";
create hostPort identifier="c0507603039e0060" userLabel="lpar0201-fcs0"
host="lpar0201" interfaceType=FC;
create hostPort identifier="c0507603039e0062" userLabel="lpar0201-fcs1"
host="lpar0201" interfaceType=FC;
```

```
create hostGroup userLabel="SAP-LC-Cluster";
create host userLabel="lpar0202" hostGroup="SAP-LC-Cluster";
create hostPort identifier="c050760506190004" userLabel="lpar0202-fcs0"
host="lpar0202" interfaceType=FC;
create hostPort identifier="c050760506190006" userLabel="lpar0202-fcs1"
host="lpar0202" interfaceType=FC;
create host userLabel="lpar0203" hostGroup="SAP-LC-Cluster";
create hostPort identifier="c0507605061a0004" userLabel="lpar0203-fcs0"
host="lpar0203" interfaceType=FC;
create hostPort identifier="c0507605061a0006" userLabel="lpar0203-fcs1"
host="lpar0203" interfaceType=FC;

create hostGroup userLabel="SAP-DB2HADR-Cluster";
create host userLabel="lpar0301" hostGroup="SAP-DB2HADR-Cluster";
create hostPort identifier="c0507603039e0064" userLabel="lpar0301-fcs0"
host="lpar0301" interfaceType=FC;
create hostPort identifier="c0507603039e0066" userLabel="lpar0301-fcs1"
host="lpar0301" interfaceType=FC;
create host userLabel="lpar0303" hostGroup="SAP-DB2HADR-Cluster";
create hostPort identifier="c0507605061a0008" userLabel="lpar0303-fcs0"
host="lpar0303" interfaceType=FC;
create hostPort identifier="c0507605061a000a" userLabel="lpar0303-fcs1"
host="lpar0303" interfaceType=FC;

create hostGroup userLabel="Migration-Cluster";
create host userLabel="lpar0302" hostGroup="Migration-Cluster";
create hostPort identifier="c050760506190008" userLabel="lpar0302-fcs0"
host="lpar0302" interfaceType=FC;
create hostPort identifier="c05076050619000a" userLabel="lpar0302-fcs1"
host="lpar0302" interfaceType=FC;
create host userLabel="lpar0401" hostGroup="Migration-Cluster";
create hostPort identifier="c0507603039e0068" userLabel="lpar0401-fcs0"
host="lpar0401" interfaceType=FC;
create hostPort identifier="c0507603039e006a" userLabel="lpar0401-fcs1"
host="lpar0401" interfaceType=FC;

create hostGroup userLabel="WPAR-Cluster";
create host userLabel="lpar0402" hostGroup="WPAR-Cluster";
create hostPort identifier="c05076050619000c" userLabel="lpar0402-fcs0"
host="lpar0402" interfaceType=FC;
create hostPort identifier="c05076050619000e" userLabel="lpar0402-fcs1"
host="lpar0402" interfaceType=FC;
create host userLabel="lpar0403" hostGroup="WPAR-Cluster";
create hostPort identifier="c0507605061a000c" userLabel="lpar0403-fcs0"
host="lpar0403" interfaceType=FC;
create hostPort identifier="c0507605061a000e" userLabel="lpar0403-fcs1"
host="lpar0403" interfaceType=FC;

create hostGroup userLabel="SmartAssist-Cluster";
create host userLabel="lpar0503" hostGroup="SmartAssist-Cluster";
create hostPort identifier="c0507605061a0012" userLabel="lpar0503-fcs0"
host="lpar0503" interfaceType=FC;
create hostPort identifier="c0507605061a0010" userLabel="lpar0503-fcs1"
host="lpar0503" interfaceType=FC;
```

```
create host userLabel="lpar0601" hostGroup="SmartAssist-Cluster";
create hostPort identifier="c0507603039e0072" userLabel="lpar0601-fcs0"
host="lpar0601" interfaceType=FC;
create hostPort identifier="c0507603039e0070" userLabel="lpar0601-fcs1"
host="lpar0601" interfaceType=FC;

create hostGroup userLabel="Install-Cluster";
create host userLabel="lpar0602" hostGroup="Install-Cluster";
create hostPort identifier="c050760506190014" userLabel="lpar0602-fcs0"
host="lpar0602" interfaceType=FC;
create hostPort identifier="c050760506190016" userLabel="lpar0602-fcs1"
host="lpar0602" interfaceType=FC;
create host userLabel="lpar0603" hostGroup="Install-Cluster";
create hostPort identifier="c0507605061a0014" userLabel="lpar0603-fcs0"
host="lpar0603" interfaceType=FC;
create hostPort identifier="c0507605061a0016" userLabel="lpar0603-fcs1"
host="lpar0603" interfaceType=FC;

create hostGroup userLabel="Migration02-Cluster";
create host userLabel="lpar0701" hostGroup="Migration02-Cluster";
create hostPort identifier="c0507603039e0074" userLabel="lpar0701-fcs0"
host="lpar0701" interfaceType=FC;
create hostPort identifier="c0507603039e0076" userLabel="lpar0701-fcs1"
host="lpar0701" interfaceType=FC;
create host userLabel="lpar0702" hostGroup="Migration02-Cluster";
create hostPort identifier="c050760506190018" userLabel="lpar0702-fcs0"
host="lpar0702" interfaceType=FC;
create hostPort identifier="c05076050619001a" userLabel="lpar0702-fcs1"
host="lpar0702" interfaceType=FC;

create hostGroup userLabel="CAA-Cluster";
create host userLabel="lpar0703" hostGroup="CAA-Cluster";
create hostPort identifier="c0507605061a0018" userLabel="lpar0703-fcs0"
host="lpar0703" interfaceType=FC;
create hostPort identifier="c0507605061a001a" userLabel="lpar0703-fcs1"
host="lpar0703" interfaceType=FC;
create host userLabel="lpar0801" hostGroup="CAA-Cluster";
create hostPort identifier="c0507603039e0078" userLabel="lpar0801-fcs0"
host="lpar0801" interfaceType=FC;
create hostPort identifier="c0507603039e007a" userLabel="lpar0801-fcs1"
host="lpar0801" interfaceType=FC;

create hostGroup userLabel="ISD-Cluster";
create host userLabel="lpar0802" hostGroup="ISD-Cluster";
create hostPort identifier="c05076050619001c" userLabel="lpar0802-fcs0"
host="lpar0802" interfaceType=FC;
create hostPort identifier="c05076050619001e" userLabel="lpar0802-fcs1"
host="lpar0802" interfaceType=FC;
create host userLabel="lpar0803" hostGroup="ISD-Cluster";
create hostPort identifier="c0507605061a001c" userLabel="lpar0803-fcs0"
host="lpar0803" interfaceType=FC;
create hostPort identifier="c0507605061a001e" userLabel="lpar0803-fcs1"
host="lpar0803" interfaceType=FC;

create hostGroup userLabel="SAP-SA-Cluster";
```

```
create host userLabel="lpar0901" hostGroup="SAP-SA-Cluster";
create hostPort identifier="c0507603039e007c" userLabel="lpar0901-fcs0"
host="lpar0901" interfaceType=FC;
create hostPort identifier="c0507603039e007e" userLabel="lpar0901-fcs1"
host="lpar0901" interfaceType=FC;
create host userLabel="lpar0902" hostGroup="SAP-SA-Cluster";
create hostPort identifier="c050760506190020" userLabel="lpar0902-fcs0"
host="lpar0902" interfaceType=FC;
create hostPort identifier="c050760506190022" userLabel="lpar0902-fcs1"
host="lpar0902" interfaceType=FC;

create hostGroup userLabel="clmgr-Cluster";
create host userLabel="lpar0903" hostGroup="clmgr-Cluster";
create hostPort identifier="c0507605061a0020" userLabel="lpar0903-fcs0"
host="lpar0903" interfaceType=FC;
create hostPort identifier="c0507605061a0022" userLabel="lpar0903-fcs1"
host="lpar0903" interfaceType=FC;
create host userLabel="lpar1001" hostGroup="clmgr-Cluster";
create hostPort identifier="c0507603039e0080" userLabel="lpar1001-fcs0"
host="lpar1001" interfaceType=FC;
create hostPort identifier="c0507603039e0082" userLabel="lpar1001-fcs1"
host="lpar1001" interfaceType=FC;
```

After all hosts and host groups are correctly created, the logical volumes are created and allocated to host groups (Example C-11).

**Important:** All logical volumes are created by using the DS4800 controller A or DS4800 controller B to provide better performance after both storage system controllers are used concurrently for I/O access.

*Example C-11 Logical volume creation and assignment*

```
# Logical volumes creation and assignment

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SAP-DB2-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DB2-Disk01" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DB2-Disk02" owner=B;
set logicaldrive ["SAP-DB2-Repository"] logicalUnitNumber=0
hostGroup="SAP-DB2-Cluster";
set logicaldrive ["SAP-DB2-Disk01"] logicalUnitNumber=1
hostGroup="SAP-DB2-Cluster";
set logicaldrive ["SAP-DB2-Disk02"] logicalUnitNumber=2
hostGroup="SAP-DB2-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SAP-DBCI-Repository" owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DBCI-Disk01" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DBCI-Disk02" owner=B;
```

```
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DBCI-Disk03" owner=A;
set logicaldrive ["SAP-DBCI-Repository"] logicalUnitNumber=0
hostGroup="SAP-DBCI-Cluster";
set logicaldrive ["SAP-DBCI-Disk01"] logicalUnitNumber=1
hostGroup="SAP-DBCI-Cluster";
set logicaldrive ["SAP-DBCI-Disk02"] logicalUnitNumber=2
hostGroup="SAP-DBCI-Cluster";
set logicaldrive ["SAP-DBCI-Disk03"] logicalUnitNumber=3
hostGroup="SAP-DBCI-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SAP-LC-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-LC-Disk01" owner=B;
set logicaldrive ["SAP-LC-Repository"] logicalUnitNumber=0
hostGroup="SAP-LC-Cluster";
set logicaldrive ["SAP-LC-Disk01"] logicalUnitNumber=1 hostGroup="SAP-LC-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SAP-DB2HADR-Repository" owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DB2HADR-Disk01" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-DB2HADR-Disk02" owner=B;
set logicaldrive ["SAP-DB2HADR-Repository"] logicalUnitNumber=0
hostGroup="SAP-DB2HADR-Cluster";
set logicaldrive ["SAP-DB2HADR-Disk01"] logicalUnitNumber=1
hostGroup="SAP-DB2HADR-Cluster";
set logicaldrive ["SAP-DB2HADR-Disk02"] logicalUnitNumber=2
hostGroup="SAP-DB2HADR-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="Migration-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=1GB userLabel="Migration-diskhb"
owner=B;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Migration-Disk01" owner=A;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Migration-Disk02" owner=B;
set logicaldrive ["Migration-diskhb"] logicalUnitNumber=0
hostGroup="Migration-Cluster";
set logicaldrive ["Migration-Repository"] logicalUnitNumber=1
hostGroup="Migration-Cluster";
set logicaldrive ["Migration-Disk01"] logicalUnitNumber=2
hostGroup="Migration-Cluster";
set logicaldrive ["Migration-Disk02"] logicalUnitNumber=3
hostGroup="Migration-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB userLabel="WPAR-Repository"
owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB userLabel="WPAR-Disk01"
owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB userLabel="WPAR-Disk02"
owner=A;
```

```
create logicalDrive array="PW2201_Array" capacity=50.00GB userLabel="WPAR-Disk03"
owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB userLabel="WPAR-Disk04"
owner=A;
set logicaldrive ["WPAR-Repository"] logicalUnitNumber=0 hostGroup="WPAR-Cluster";
set logicaldrive ["WPAR-Disk01"] logicalUnitNumber=1 hostGroup="WPAR-Cluster";
set logicaldrive ["WPAR-Disk02"] logicalUnitNumber=2 hostGroup="WPAR-Cluster";
set logicaldrive ["WPAR-Disk03"] logicalUnitNumber=3 hostGroup="WPAR-Cluster";
set logicaldrive ["WPAR-Disk04"] logicalUnitNumber=4 hostGroup="WPAR-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SmartAssist-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SmartAssist-Disk01" owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SmartAssist-Disk02" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SmartAssist-Disk03" owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SmartAssist-Disk04" owner=A;
set logicaldrive ["SmartAssist-Repository"] logicalUnitNumber=0
hostGroup="SmartAssist-Cluster";
set logicaldrive ["SmartAssist-Disk01"] logicalUnitNumber=1
hostGroup="SmartAssist-Cluster";
set logicaldrive ["SmartAssist-Disk02"] logicalUnitNumber=2
hostGroup="SmartAssist-Cluster";
set logicaldrive ["SmartAssist-Disk03"] logicalUnitNumber=3
hostGroup="SmartAssist-Cluster";
set logicaldrive ["SmartAssist-Disk04"] logicalUnitNumber=4
hostGroup="SmartAssist-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="Install-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Install-Disk01" owner=B;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Install-Disk02" owner=A;
set logicaldrive ["Install-Repository"] logicalUnitNumber=0
hostGroup="Install-Cluster";
set logicaldrive ["Install-Disk01"] logicalUnitNumber=1
hostGroup="Install-Cluster";
set logicaldrive ["Install-Disk02"] logicalUnitNumber=2
hostGroup="Install-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="Migration02-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="Migration02-diskhb" owner=B;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Migration02-Disk01" owner=A;
create logicalDrive array="PW2201_Array" capacity=10.00GB
userLabel="Migration02-Disk02" owner=B;
set logicaldrive ["Migration02-Repository"] logicalUnitNumber=0
hostGroup="Migration02-Cluster";
```

```
set logicaldrive ["Migration02-diskhb"] logicalUnitNumber=1
hostGroup="Migration02-Cluster";
set logicaldrive ["Migration02-Disk01"] logicalUnitNumber=2
hostGroup="Migration02-Cluster";
set logicaldrive ["Migration02-Disk02"] logicalUnitNumber=3
hostGroup="Migration02-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB userLabel="CAA-Repository"
owner=A;
create logicalDrive array="PW2201_Array" capacity=10.00GB userLabel="CAA-Disk01"
owner=B;
create logicalDrive array="PW2201_Array" capacity=10.00GB userLabel="CAA-Disk02"
owner=A;
set logicaldrive ["CAA-Repository"] logicalUnitNumber=0 hostGroup="CAA-Cluster";
set logicaldrive ["CAA-Disk01"] logicalUnitNumber=1 hostGroup="CAA-Cluster";
set logicaldrive ["CAA-Disk02"] logicalUnitNumber=2 hostGroup="CAA-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB userLabel="ISD-Repository"
owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB userLabel="ISD-Disk01"
owner=B;
set logicaldrive ["ISD-Repository"] logicalUnitNumber=0 hostGroup="ISD-Cluster";
set logicaldrive ["ISD-Disk01"] logicalUnitNumber=1 hostGroup="ISD-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB
userLabel="SAP-SA-Repository" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-SA-Disk01" owner=B;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-SA-Disk02" owner=A;
create logicalDrive array="PW2201_Array" capacity=50.00GB
userLabel="SAP-SA-Disk03" owner=B;
set logicaldrive ["SAP-SA-Repository"] logicalUnitNumber=0
hostGroup="SAP-SA-Cluster";
set logicaldrive ["SAP-SA-Disk01"] logicalUnitNumber=1 hostGroup="SAP-SA-Cluster";
set logicaldrive ["SAP-SA-Disk02"] logicalUnitNumber=2 hostGroup="SAP-SA-Cluster";
set logicaldrive ["SAP-SA-Disk03"] logicalUnitNumber=3 hostGroup="SAP-SA-Cluster";

create logicalDrive array="PW2201_Array" capacity=1GB userLabel="clmgr-Repository"
owner=A;
create logicalDrive array="PW2201_Array" capacity=10.00GB userLabel="clmgr-Disk01"
owner=A;
set logicaldrive ["clmgr-Repository"] logicalUnitNumber=0
hostGroup="clmgr-Cluster";
set logicaldrive ["clmgr-Disk01"] logicalUnitNumber=2 hostGroup="clmgr-Cluster";
```

D

# Replacing the failed repository disk if any nodes are not joined to the cluster

This appendix describes the steps that you can follow to replace the Cluster Aware AIX (CAA) cluster repository disk in a PowerHA environment if a node is not joined to the cluster. This situation might happen if any node in the cluster is restarted or halted without a correctly configured repository disk.

# Introduction

Without the repository disk, a node cannot join the cluster, so if you reboot any of the nodes without the configured repository disk, you cannot start the cluster services in this node. Next, we show steps to address this situation and bring up the cluster services.

Follow the next steps to recover your PowerHA cluster:

1. Stop the cluster services in every node in the cluster.

   Follow this path: **`smit sysmirror`** → **System Management (C-SPOC)** → **PowerHA SystemMirror Services** → **Stop Cluster Services**. Select all nodes in the PowerHA cluster as shown in Figure D-1.

```
     Stop Cluster Services

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

                                                          [Entry Fields]
    * Stop now, on system restart or both               now +
      Stop Cluster Services on these nodes             [caa1,caa2] +
      BROADCAST cluster shutdown?                       false +
    * Select an Action on Resource Groups               Bring Resource Groups Offline +
```

*Figure D-1 Stop cluster services*

2. Save a snapshot of the cluster.

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks** → **Manage the Cluster** → **Snapshot Configuration** → **Create a Snapshot of the Cluster Configuration** (Figure D-2).

```
                 Create a Snapshot of the Cluster Configuration

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

                                 [Entry Fields]
* Cluster Snapshot Name                                  [no-disk-test01]
  Custom Defined Snapshot Methods                        []
  Save Cluster Log Files in snapshot                      No
* Cluster Snapshot Description                           [Test01]
```

*Figure D-2 Snapshot smit menu*

3. Remove the cluster for every node.

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks -> Manage the Cluster** → **Remove the Cluster Definition (**Figure D-3 on page 599).

```
                          Remove the Cluster Definition

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

                                                        [Entry Fields]
  Cluster Name                                        caa_cl
* Remove definition from all nodes                   [Yes]

  NOTE: All user-configured cluster information
  WILL BE DELETED by this operation.
```

*Figure D-3 Remove the cluster menu*

Verify that no cluster definition exists for all nodes as shown in Figure D-4.

```
root@caa1 / # cltopinfo
Warning: There is no cluster found.
cltopinfo: Error reading configuration.
-----
root@caa2 / # cltopinfo
Warning: There is no cluster found.
cltopinfo: Error reading configuration.
```

*Figure D-4 cltopinfo output*

4. Remove and replace the failed disk in the operating system for all nodes.

   Remove the failed disk (Figure D-5).

```
root@caa1 / # rmdev -dl hdisk6
hdisk6 deleted
-----
root@caa2 / # rmdev -dl hdisk6
hdisk6 deleted
```

*Figure D-5 Remove the failed disk*

5. Use the **cfgmgr** command to configure the new disk in both nodes. See Figure D-6 on page 600.

```
root@caa1 / # lspv
hdisk0          00f74d47f963be5f                    rootvg          active
hdisk1          00f74d47fa93539e                    rootvg          active
hdisk3          00f74d472b9f9c73                    dbvg
hdisk4          00f74d472b9f9d6e                    internetvg
root@caa1 / # cfgmgr
root@caa1 / # lspv
hdisk0          00f74d47f963be5f                    rootvg          active
hdisk1          00f74d47fa93539e                    rootvg          active
hdisk3          00f74d472b9f9c73                    dbvg
hdisk4          00f74d472b9f9d6e                    internetvg
hdisk2          none                                None
```

*Figure D-6   Configure the new repository disk in AIX*

6. Configure a basic cluster (only cluster and nodes).

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks** → **Initial Cluster Setup (Typical)** → **Setup a Cluster, Nodes and Networks** (Figure D-7).

```
 Set up a Cluster, Nodes and Networks

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

                                                        [Entry Fields]
* Cluster Name                                       [caa_cl]
  New Nodes (via selected communication paths)       [caa2] +
  Currently Configured Node(s)                        caa1
```

*Figure D-7   Configure a basic cluster*

7. Add the repository disk.

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks** → **Initial Cluster Setup (Typical)** → **Define Repository Disk and Cluster IP Address** (Figure D-8).

```
                      Define Repository Disk and Cluster IP Address

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

                                                        [Entry Fields]
* Cluster Name                                        caa_cl
* Repository Disk                                    [hdisk2] +
  Cluster IP Address                                 [228.1.1.34]
```

*Figure D-8   Define cluster repository disk*

8. Verify and synchronize the cluster configuration.

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks** → **Verify and Synchronize Cluster Configuration**.

9. Recover the PowerHA cluster from the previously saved snapshot.

   Select your previously saved snapshot from step 2 and select the default option (**yes**) for Un/Configure Cluster Resources? (Figure D-9).

   Follow this path: **`smit sysmirror`** → **Cluster Nodes and Networks** → **Manage the Cluster** → **Snapshot Configuration** → **Restore the Cluster Configuration from a Snapshot**.

```
 Restore the Cluster Configuration from a Snapshot

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

                                                        [Entry Fields]
  Cluster Snapshot Name                               no-disk-test01
  Cluster Snapshot Description                        Test01
  Un/Configure Cluster Resources?                    [Yes] +
  Force apply if verify fails?                       [No] +
```

*Figure D-9 Apply previously saved snapshot*

10. Start Cluster Services on both nodes.

    Follow this path: **`smit sysmirror`** → **System Management (C-SPOC)** → **PowerHA SystemMirror Services** → **Start Cluster Services** (Figure D-10).

```
                                   Start Cluster Services

Type or select values in entry fields.
Press Enter AFTER making all desired changes.
                                                        [Entry Fields]
* Start now, on system restart or both                now +
  Start Cluster Services on these nodes              [caa1,caa2] +
* Manage Resource Groups                              Automatically +
  BROADCAST message at startup?                       false +
  Startup Cluster Information Daemon?                 true +
  Ignore verification errors?                         false +
  Automatically correct errors found during           Interactively +
  cluster start?
```

*Figure D-10 Starting the cluster services*

11. Verify the new repository disk.

    After all nodes are updated, verify that the new repository disk works by running the **`/usr/sbin/lscluster -d`** command as shown in Figure D-11 on page 602.

```
root@caa2 / # lscluster -d
Storage Interface Query

Cluster Name:  caa_cl
Cluster uuid:  964943a0-7e9e-11e1-83ac-b6fcc11f846f
Number of nodes reporting = 2
Number of nodes expected = 2
Node caa2
Node uuid = 96432b00-7e9e-11e1-83ac-b6fcc11f846f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 4d621527-d912-3479-5e16-5c9a03980323
          type  : REPDISK
Node caa1
Node uuid = 9600aa0a-7e9e-11e1-83ac-b6fcc11f846f
Number of disk discovered = 1
        hdisk2
          state : UP
          uDid  :
          uUid  : 4d621527-d912-3479-5e16-5c9a03980323
          type  : REPDISK
```

*Figure D-11 lscluster -d output*

E

# IBM PowerHA SystemMirror Smart Assist for SAP additional materials for Chapter 7

In this appendix, we describe additional materials that are used in Chapter 7, "IBM PowerHA SystemMirror Smart Assist for SAP" on page 355.

# Scripts to create the SAP users

The following section describes the scripts that are used to create SAP users.

## Script to create users and groups for sapsma_cluster

Example E-1 shows the script to create all users and groups for the `sapsma_cluster`.

*Example E-1 Create all needed users and groups for sapsma_cluster*

```
mkgroup -A id=300 sapsys
mkgroup -A id=301 sapinst
mkgroup -A id=303 dbte1ctl
mkgroup -A id=304 dbte1mnt
mkgroup -A id=305 dbte1adm
mkgroup -A id=306 dbte1mon

mkuser id=300 pgrp=dbte1ctl shell='/bin/csh' groups='dbte1ctl,staff'
home='/db2/db2te1' gecos='SAP Database Administrator' db2te1
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' db2te1
echo db2te1:A300B0 | chpasswd -c

mkuser id=310 pgrp=sapsys shell='/bin/csh' groups='sapsys,staff,sapinst,dbte1ctl'
home='/usr/sap/TE1/home/te1adm' gecos='SAP System Administrator' te1adm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' te1adm
echo te1adm:A310B0 | chpasswd -c

mkuser id=305 pgrp=sapsys shell='/bin/csh' groups='sapsys,dbte1ctl'
home='/usr/sap/TE1/home/sapsr3' gecos='ABAP Database Connect User' sapsr3
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapsr3
echo sapsr3:A305B0 | chpasswd -c

mkuser id=320 pgrp=sapsys shell='/bin/csh' groups='sapsys,sapinst'
home='/usr/sap/TE1/home/sapadm' gecos='SAP System Administrator' sapadm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapadm
echo sapadm:A320B0 | chpasswd -c
```

## Script to create users and groups for sapci_cluster, sapdb_cluster, and saplc_cluster

Example E-2 shows the script to create all needed users and groups for the `sapci_cluster`, `sapdb_cluster`, and `saplc_cluster`.

*Example E-2 Create all needed user and groups for sapci_cluster, sapdb_cluster, and saplc_cluster*

```
mkgroup -A id=300 sapsys
mkgroup -A id=301 sapinst
mkgroup -A id=303 dbtc1ctl
mkgroup -A id=304 dbtc1mnt
mkgroup -A id=305 dbtc1adm
mkgroup -A id=306 dbtc1mon
```

```
mkuser id=301 pgrp=dbtc1ctl shell='/bin/csh' groups='dbtc1ctl,staff'
home='/db2/db2tc1' gecos='SAP Database Administrator' db2tc1
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' db2tc1
echo db2tc1:A301B0 | chpasswd -c

mkuser id=311 pgrp=sapsys shell='/bin/csh' groups='sapsys,staff,sapinst,dbtc1ctl'
home='/usr/sap/TC1/home/tc1adm' gecos='SAP System Administrator' tc1adm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' tc1adm
echo tc1adm:A311B0 | chpasswd -c

mkuser id=305 pgrp=sapsys shell='/bin/csh' groups='sapsys,dbtc1ctl'
home='/usr/sap/TC1/home/sapsr3' gecos='ABAP Database Connect User' sapsr3
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapsr3
echo sapsr3:A305B0 | chpasswd -c

mkuser id=320 pgrp=sapsys shell='/bin/csh' groups='sdba,staff' home='/home/sdb'
gecos='SAP Database Administrator' sdb
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sdb
echo sapadm:A320B0 | chpasswd -c

mkuser id=321 pgrp=sapsys shell='/bin/csh' groups='sapsys,sapinst'
home='/usr/sap/TC1/home/sapadm' gecos='SAP System Administrator' sapadm
chuser fsize='-1' data='-1' stack='-1' rss='-1' cpu='-1' nofiles='8000' sapadm
echo sapadm:A321B0 | chpasswd -c
```

# Scripts to create the volume groups, logical volumes, and file systems

The following section shows the script to create volume groups (VGs) and logical volumes (LVs) for the `sapsma_cluster`.

## Script to create VGs and LVs for sapsma_cluster

Example E-3 shows the script to create the VGs, LVs, and file systems.

*Example E-3 Create VGs, LVs, and file systems*

```
mkvg -f -S -y vg_TE1_sap -s 64 -V 102 hdisk3 hdisk4 hdisk5
varyonvg vg_TE1_sap
chvg -ay -Qy -P 262144 vg_TE1_sap

mklv -y TE1.lvusrsap    -t'jfs2' vg_TE1_sap 160
crfs -v jfs2 -d TE1.lvusrsap -m /usr/sap/TE1 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TE1.lvdb2binary -t'jfs2' vg_TE1_sap 128
crfs -v jfs2 -d TE1.lvdb2binary -m /db2/TE1   -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mount /db2/TE1

mklv -y TE1.lvdb2logdir -t'jfs2' vg_TE1_sap 64
```

```
crfs -v jfs2 -d TE1.lvdb2logdir -m /db2/TE1/log_dir       -u TE1  -A'no' -p'rw' -a
agblksize='512'  -a logname='INLINE'

mklv -y TE1.lvdb2logarc -t'jfs2' vg_TE1_sap 64
crfs -v jfs2 -d TE1.lvdb2logarc -m /db2/TE1/log_archive   -u TE1  -A'no' -p'rw' -a
agblksize='512'  -a logname='INLINE'

mklv -y TE1.lvdb2dat.01 -t'jfs2' vg_TE1_sap 400
crfs -v jfs2 -d TE1.lvdb2dat.01 -m /db2/TE1/sapdata1 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TE1.lvdb2dat.02 -t'jfs2' vg_TE1_sap 400
crfs -v jfs2 -d TE1.lvdb2dat.02 -m /db2/TE1/sapdata2 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TE1.lvdb2dat.03 -t'jfs2' vg_TE1_sap 400
crfs -v jfs2 -d TE1.lvdb2dat.03 -m /db2/TE1/sapdata3 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TE1.lvdb2dat.04 -t'jfs2' vg_TE1_sap 400
crfs -v jfs2 -d TE1.lvdb2dat.04 -m /db2/TE1/sapdata4 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mkvg -f -S -y vg_TE1_nfs -s 64 -V 103 hdisk6
varyonvg vg_TE1_nfs
chvg -ay -Qy -P 262144 vg_TE1_nfs

mklv -y TE1.lvsapmnt    -t'jfs2' vg_TE1_nfs 64
crfs -v jfs2 -d TE1.lvsapmnt -m /export/sapmnt/TE1 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TE1.lvsaptrans  -t'jfs2' vg_TE1_nfs 32
crfs -v jfs2 -d TE1.lvsaptrans -m /export/saptrans/TE1 -u TE1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mount -t TE1
mkdir -p /db2/TE1/home/db2te1
mkdir -p /usr/sap/TE1/home/te1adm
mkdir -p /usr/sap/TE1/home/sapadm

chown -R db2te1 /db2/TE1
chown -R te1adm:sapsys /usr/sap/TE1
```

## Script to create VGs and LVs for sapci_cluster

We provide a script to create volume groups and logical volumes for the `sapci_cluster`.

### Volume groups and file systems for Network File System (NFS), Application server (AS), and Central Instance (CI)

Example E-4 on page 607 shows a script to create VGs, LVs, and the file systems for the `sapci_cluster`.

*Example E-4 Create VGs, LVs, and file systems on sapci_cluster*

```
mkvg -C -f -S -y vg_TC1_sapas -s 64 -V 102 hdisk3
varyonvg vg_TC1_sapas
chvg -ay -Qy -P 262144 vg_TC1_sapas

mklv -y TC1.lvusrsapas  -t'jfs2' vg_TC1_sapas 160
crfs -v jfs2 -d TC1.lvusrsapas -m /usr/sap/TC1/ASCS10 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mkvg -C -f -S -y vg_TC1_sapci -s 64 -V 101 hdisk4
varyonvg vg_TC1_sapci
chvg -ay -Qy -P 262144 vg_TC1_sapci

mklv -y TC1.lvusrsapci  -t'jfs2' vg_TC1_sapci 160
crfs -v jfs2 -d TC1.lvusrsapci -m /usr/sap/TC1/DVEBMGS12 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mkvg -C -f -S -y vg_TC1_nfs -s 64 -V 103 hdisk5
varyonvg vg_TC1_nfs
chvg -ay -Qy -P 262144 vg_TC1_nfs
mklv -y TC1.lvsapmnt    -t'jfs2' vg_TC1_nfs 64
crfs -v jfs2 -d TC1.lvsapmnt -m /export/sapmnt/TC1 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mklv -y TC1.lvsaptrans  -t'jfs2' vg_TC1_nfs 32
crfs -v jfs2 -d TC1.lvsaptrans -m /export/saptrans/TC1 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mklv -y TL1.lvlock -t'jfs2' vg_TC1_nfs 32
crfs -v jfs2 -d TL1.lvlock -m /export/lock/TL1 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'

mount -t TC1
mkdir -p /usr/sap/TC1/home/tc1adm
mkdir -p /usr/sap/TC1/home/sapadm
mkdir -p /home/sdb

chown -R tc1adm:sapsys /usr/sap/TC1
chown -R sapadm:sapsys /usr/sap/TC1/home/sapadm
chown -R sdb:sapsys /home/sdb
```

## Volume groups and file systems for Enqueue Replication Server (ERS)

Example E-5 shows the script to create VGs, LVs, and file systems for the `sapci_cluster`.

*Example E-5 Create VGs, LVs, and file systems on sapci_cluster*

```
mkvg -C -f -S -y vg_TC1_sapers -s 64 -V 104 hdisk6
varyonvg vg_TC1_sapers
chvg -ay -Qy -P 262144 vg_TC1_sapers

mklv -y TC1.lvusrsapers    -t'jfs2' vg_TC1_sapers 10
crfs -v jfs2 -d TC1.lvusrsapers -m /usr/sap/TC1/ERS11 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mount -t TC1
chown -R tc1adm /usr/sap/TC1
```

## Script to create VGs and LVs for sapdb_cluster

Example E-6 shows the script to create VGs, LVs, and file systems for the `sapdb_cluster`.

*Example E-6 Create VGs, LVs, and file systems on sapdb_cluster*

```
mkvg -f -S -y vg_TC1_sap -s 64 -V 102 hdisk3 hdisk4
varyonvg vg_TC1_sap
chvg -ay -Qy -P 262144 vg_TC1_sap
mklv -y TC1.lvdb2binary -t'jfs2' vg_TC1_sap 128
crfs -v jfs2 -d TC1.lvdb2binary -m /db2/TC1   -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mount /db2/TC1
mklv -y TC1.lvdb2logdir -t'jfs2' vg_TC1_sap 64
crfs -v jfs2 -d TC1.lvdb2logdir -m /db2/TC1/log_dir      -u TC1  -A'no' -p'rw' -a
agblksize='512'  -a logname='INLINE'
mklv -y TC1.lvdb2logarc -t'jfs2' vg_TC1_sap 64
crfs -v jfs2 -d TC1.lvdb2logarc -m /db2/TC1/log_archive   -u TC1  -A'no' -p'rw' -a
agblksize='512'  -a logname='INLINE'
mklv -y TC1.lvdb2dat.01 -t'jfs2' vg_TC1_sap 320
crfs -v jfs2 -d TC1.lvdb2dat.01 -m /db2/TC1/sapdata1 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mklv -y TC1.lvdb2dat.02 -t'jfs2' vg_TC1_sap 320
crfs -v jfs2 -d TC1.lvdb2dat.02 -m /db2/TC1/sapdata2 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mklv -y TC1.lvdb2dat.03 -t'jfs2' vg_TC1_sap 320
crfs -v jfs2 -d TC1.lvdb2dat.03 -m /db2/TC1/sapdata3 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mklv -y TC1.lvdb2dat.04 -t'jfs2' vg_TC1_sap 320
crfs -v jfs2 -d TC1.lvdb2dat.04 -m /db2/TC1/sapdata4 -u TC1 -A'no' -p'rw' -a
agblksize='4096' -a logname='INLINE'
mount -t TC1
mkdir -p /db2/TC1/home/db2tc1
chown -R db2tc1 /db2/TC1
```

## Script to create VGs and LVs for the saplc_cluster

We create the `/sapdb` file system in the demonstration environment inside the rootvg because we do not have enough disks to create a dedicated volume group. See Example E-7.

*Example E-7 Create local VGs, LVs, and file systems on saplc1 and saplc2*

```
mkvg -f -S -y vg_TL1_sdbdat -s 64 -V 101 hdisk4 hdisk5
varyonvg vg_TL1_sdbdat
chvg -ay -Qy -P 262144 vg_TL1_sdbdat
mklv -y 'lv_TL1_data1' -t'jfs2' vg_TL1_sdbdat 1596
mklv -y 'sapdblv' -t'jfs2' -c'2' vg_TL1_sdbdat 32
crfs -v jfs2 -d'sapdblv' -m'/sapdb' -A'y' -p'rw' -a agblksize='4096' -a
logname='INLINE'
chown -R sdb:sdba /sapdb
```

Example E-8 shows the creation of the shared VGs and LVs for the log volumes.

*Example E-8 Create shared VGs, LVs, and file systems on saplc_cluster*

```
mkvg -f -S -y vg_TL1_sdblog -s 64 -V 102 hdisk3
```

```
varyonvg vg_TL1_sdblog
chvg -ay -Qy -P 262144 vg_TL1_sdblog
mklv -y lv_TL1_log1 -t'jfs2' vg_TL1_sdblog 399
mklv -y lv_TL1_log2 -t'jfs2' vg_TL1_sdblog 399
```

# SAPinst installation

This section contains the details for the SAPinst installation.

## sapci_cluster: Installing the SAP ASCS instance on the first node

We performed a normal SAP installation with a virtual SAP host name (`sapcisvc1`). Only the normal SAP parameters are set as you see in Example E-9.

*Example E-9 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc1
```

Follow these steps:

1. Start the ABAP SAP Central Services (ASCS) installation as shown in Figure E-1. Select **liveCache Server Installation**. The panels might appear in a different order. It depends on the SAPinst version that you use.

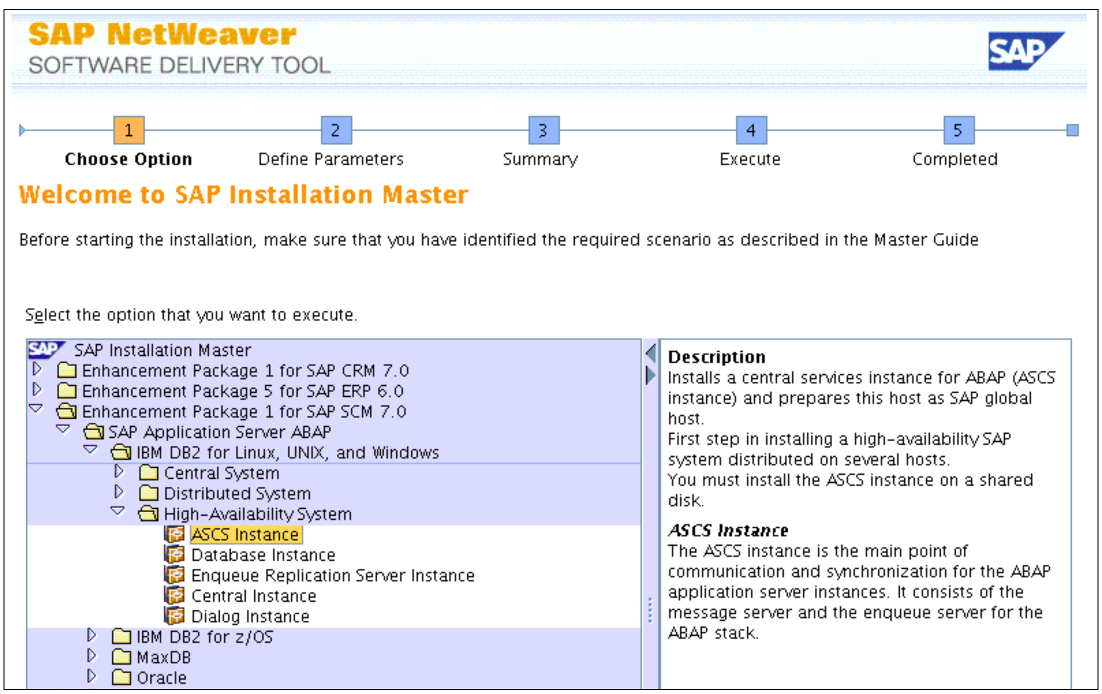



*Figure E-1 SAPinst ASCS installation screen capture 1: Start installation*

2. Always select **Custom** installation as shown in Figure E-2 on page 610 so that you can change some default values.

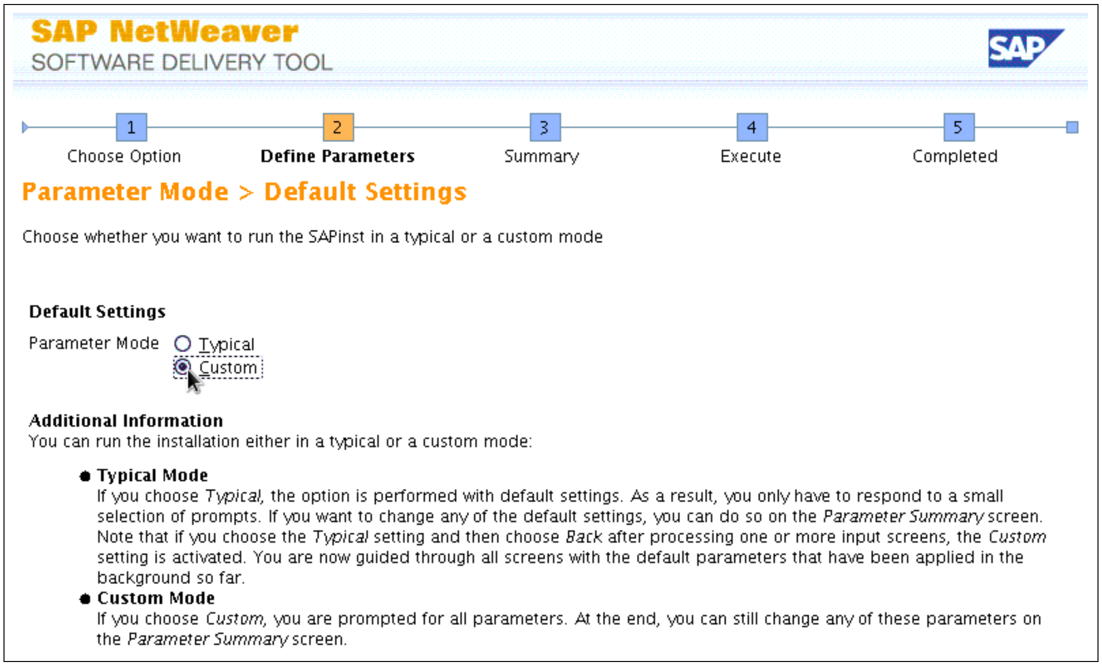



*Figure E-2 SAPinst ASCS installation screen capture 2: Custom installation*

3. In the next window, enter the SAP system ID that you want to use for this installation. The SAP system mount directory on a UNIX system must always be the default. Select **/sapmnt** as shown in Figure E-3.


*Figure E-3 SAPinst ASCS installation screen capture 3: SAP system ID and /sapmnt directory*

4. In the demonstration environment, we do not use a Domain Name System (DNS) server, so we clear this selection in the window that is shown in Figure E-4 on page 611.

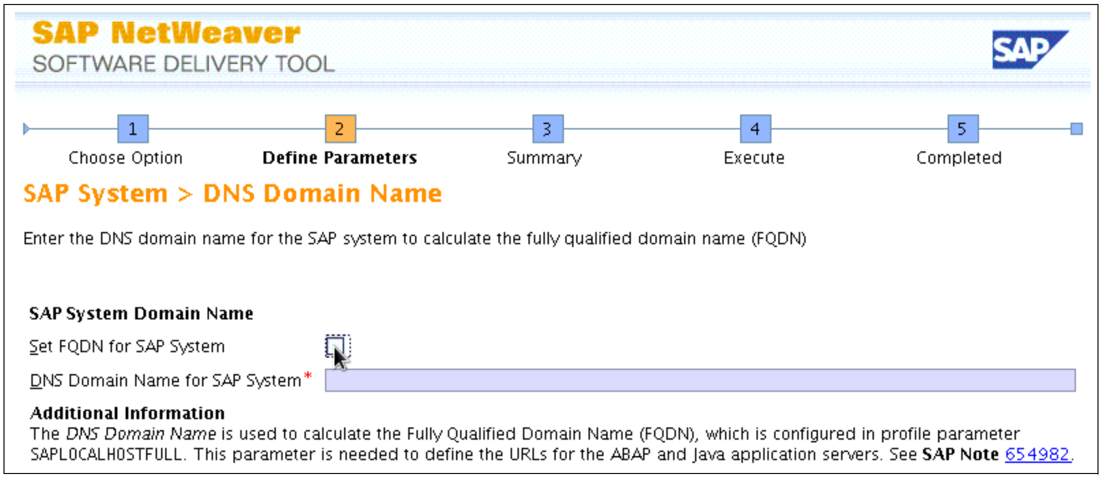



*Figure E-4 SAPinst ASCS installation screen capture 4: FQDN and DNS domain*

5. You have to set up a master password. This password is used by all users that the SAPinst creates as shown in Figure E-5.

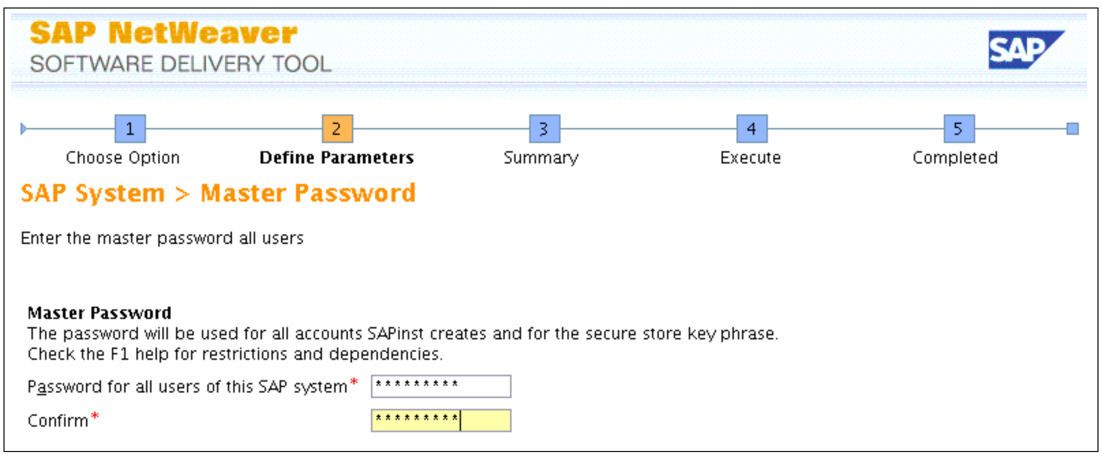



*Figure E-5 SAPinst ASCS installation screen capture 5: Master password*

6. SAPinst now asks for the location of the IBM UC™ Kernel NetWeaver 7.20 as shown in Figure E-6.

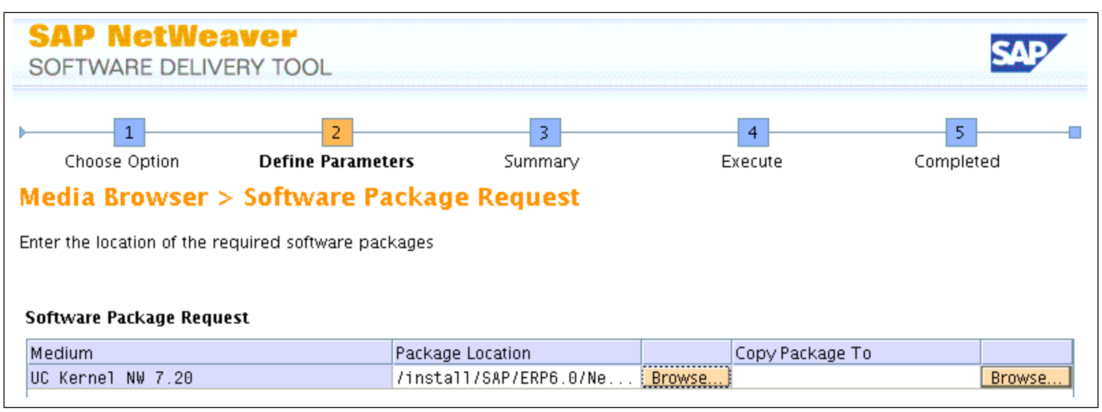



*Figure E-6 SAPinst ASCS installation screen capture 5: UC Kernel NW 7.20*

7. Enter the elected instance number as shown in Figure E-7 on page 612.

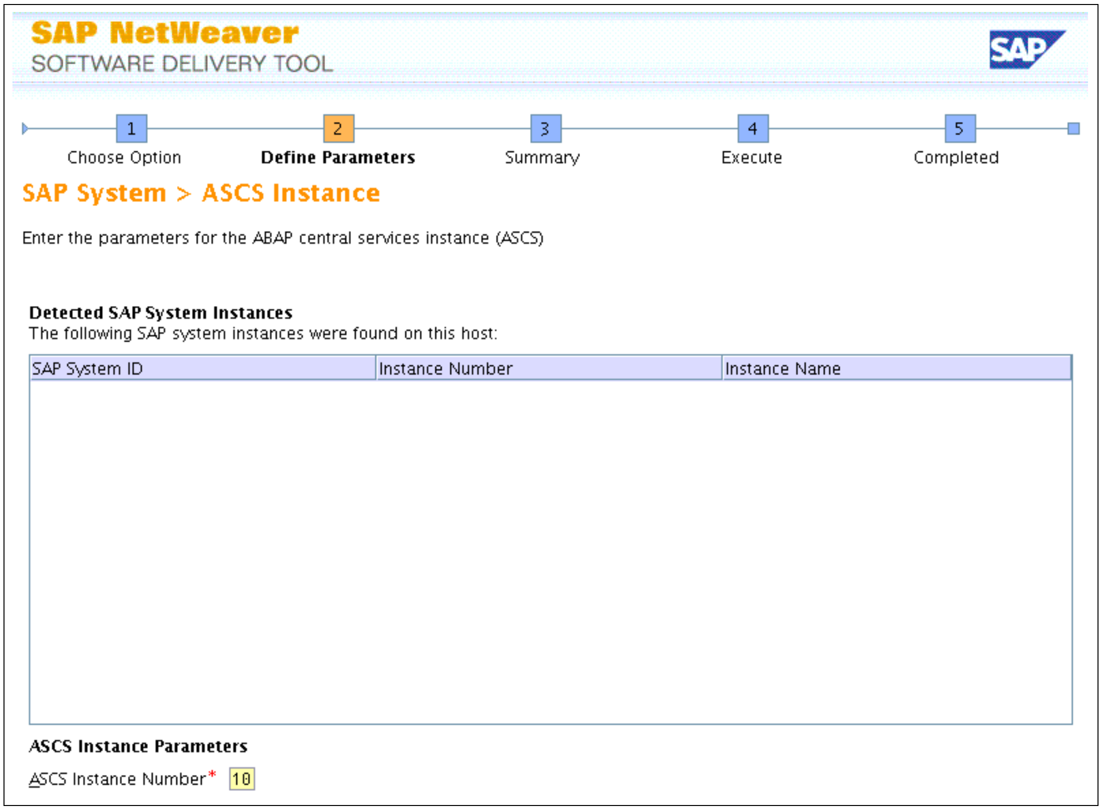



*Figure E-7 SAPinst ASCS installation screen capture 6: ASCS instance number*

8. Accept the default message server ports as shown in Figure E-8.

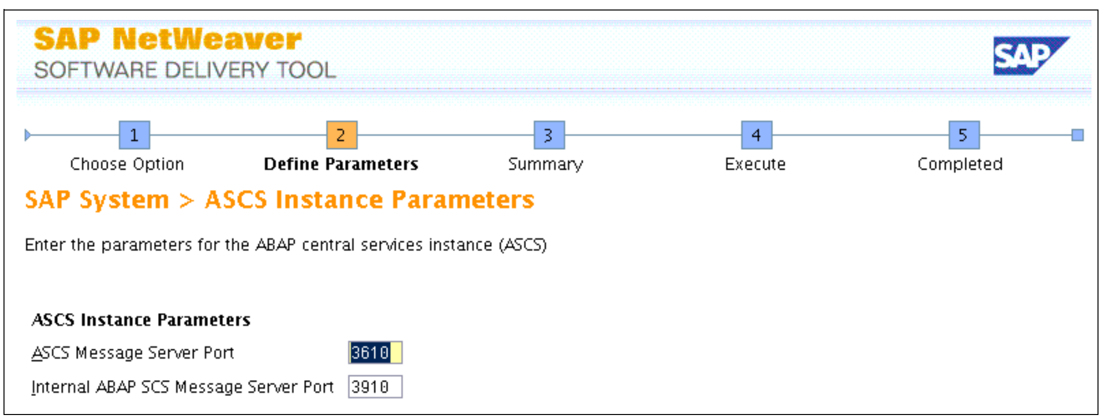



*Figure E-8 SAPinst ASCS installation screen capture 7: ASCS message ports*

9. The SAP Cryptographic Software is not installed in the environment as shown in Figure E-9 on page 613.

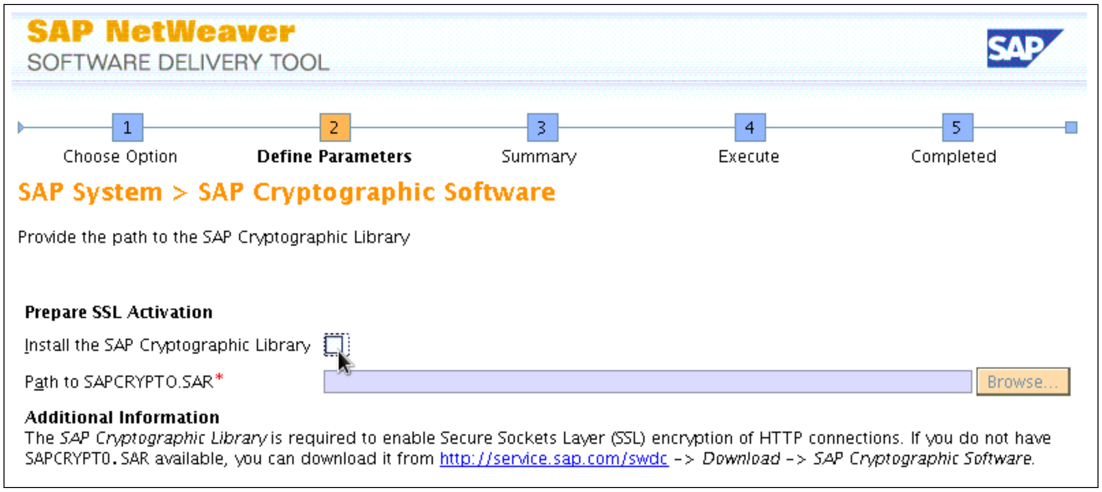



*Figure E-9 SAPinst ASCS installation screen capture 8: SAP Cryptographic Software*

10. Click **Next** for SAPinst to unpack the necessary archive (.SAR) files as shown in Figure E-10.


*Figure E-10 SAPinst ASCS installation screen capture 9: SAP archives to unpack*

11. After several more panels, SAPinst starts the actual installation. We installed the following SAP version in our demonstration system as shown in Example E-10.

*Example E-10 SAP version installed*

```
sapci1:tc1adm 1> disp+work -version

--------------------
disp+work information
--------------------

kernel release                720

kernel make variant           720_REL

compiled on                   AIX 2 5 00092901D600 for rs6000_64

compiled for                  64 BIT
```

```
compilation mode              UNICODE

compile time                  May 23 2011 21:24:01

update level                  0

patch number                  90

source id                     0.090


---------------------
supported environment
---------------------

database (SAP, table SVERS)   700
                              710
                              701
                              702
                              703
                              711
                              720
                              730
                              731

operating system
AIX 2 5
AIX 3 5
AIX 1 6
AIX 1 7
```

## sapdb_cluster: Installing the SAP database instance on the first node

We performed a normal SAP installation with a virtual SAP host name. Only the usual SAP parameters are set as you can see in Example E-11.

*Example E-11 SAPinst environment parameters*

```
export TMPDIR=/tmp/TE1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapdbsvc2
```

Follow these steps:

1. Start the database installation as shown in Figure E-11 on page 615. Select **Database Instance**.

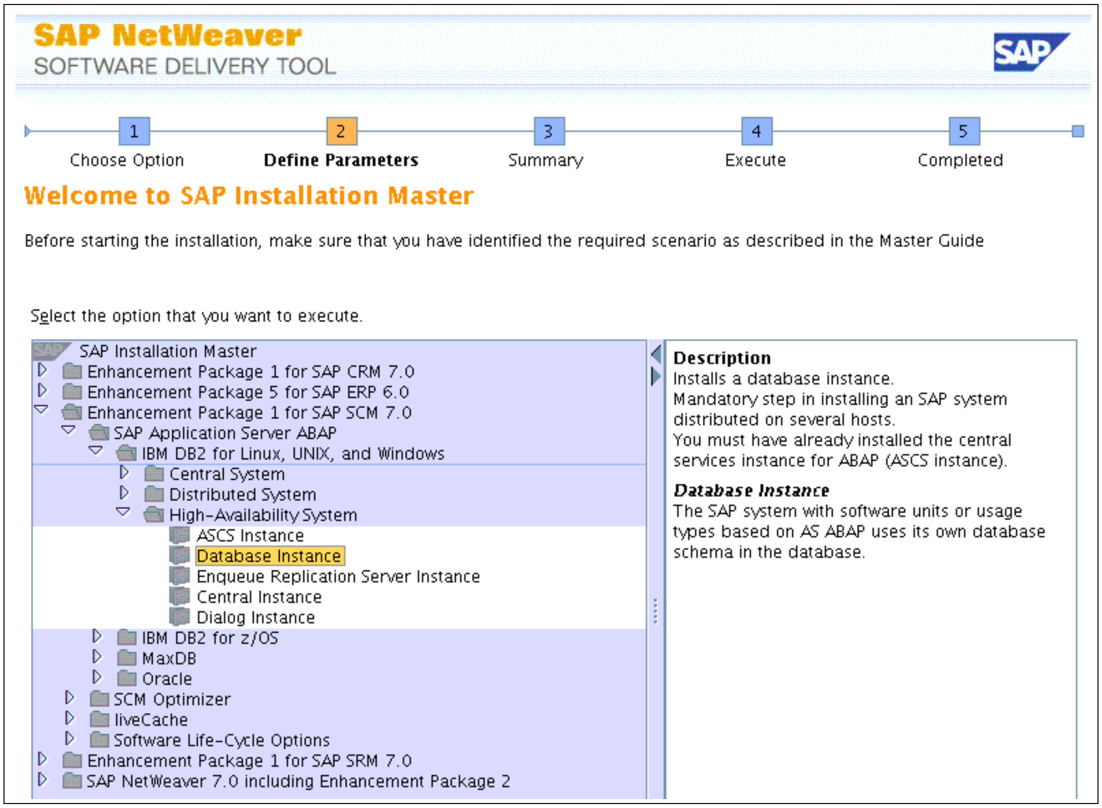



*Figure E-11 SAPinst database installation screen capture 1: Start installation*

2. The SAP system mount directory on an AIX system must always be the default, `/sapmnt/TC1/profile`, as shown in Figure E-12.

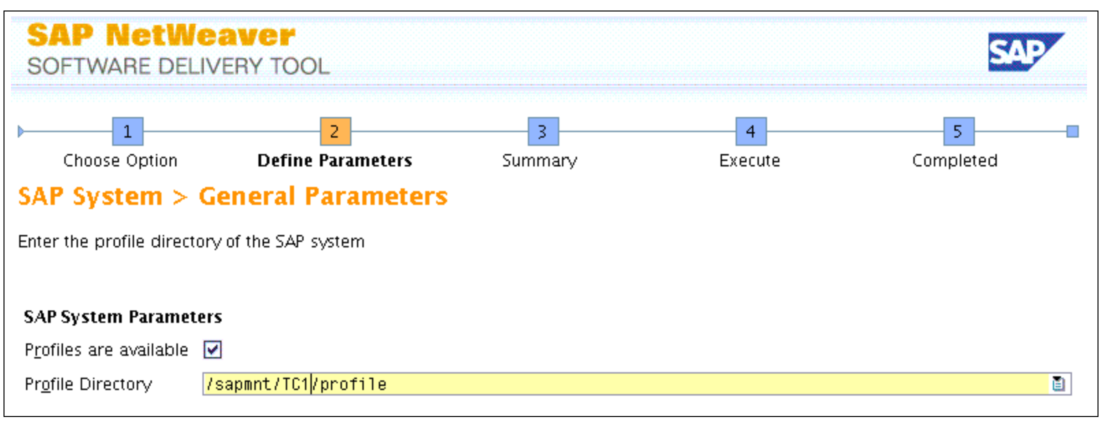



*Figure E-12 SAPinst database installation screen capture 2: Profile directory*

3. Set up a master password. This password is used by all users that the SAPinst creates as shown in Figure E-13 on page 616.

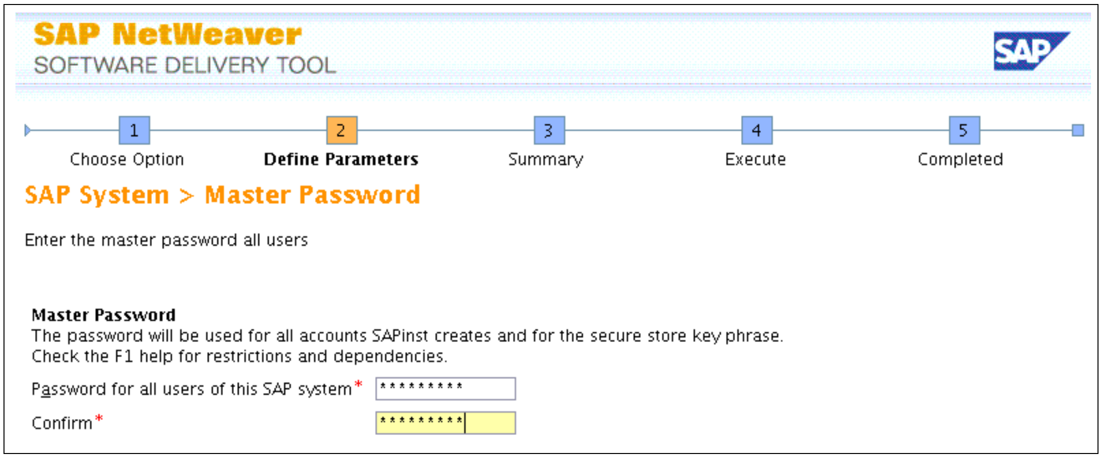



*Figure E-13   SAPinst database installation screen capture 3: Master Password*

4. Insert the database ID. We used SID as shown in Figure E-14.

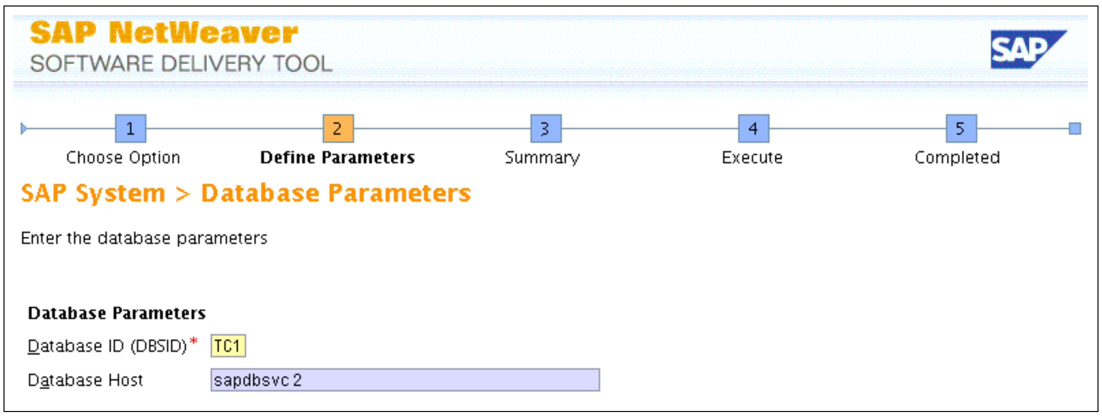



*Figure E-14   SAPinst database installation screen capture 4: Database Parameters*

5. Define the database connect user. The name that is used is `sapsr3` as shown in Figure E-15.

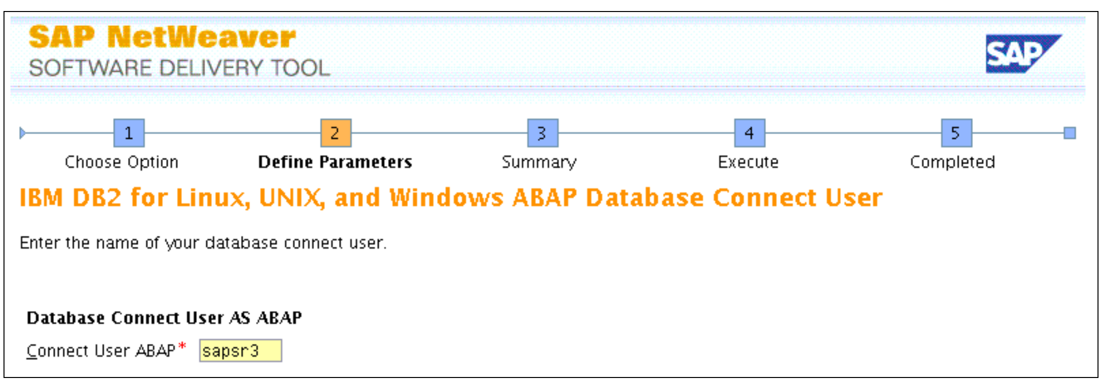



*Figure E-15   SAPinst database installation screen capture 5: Database Connect User*

6. Insert the installation path of the DB2 database software. The default is <*HOME db2sid user*>`/db2_software` as shown in Figure E-16 on page 617.

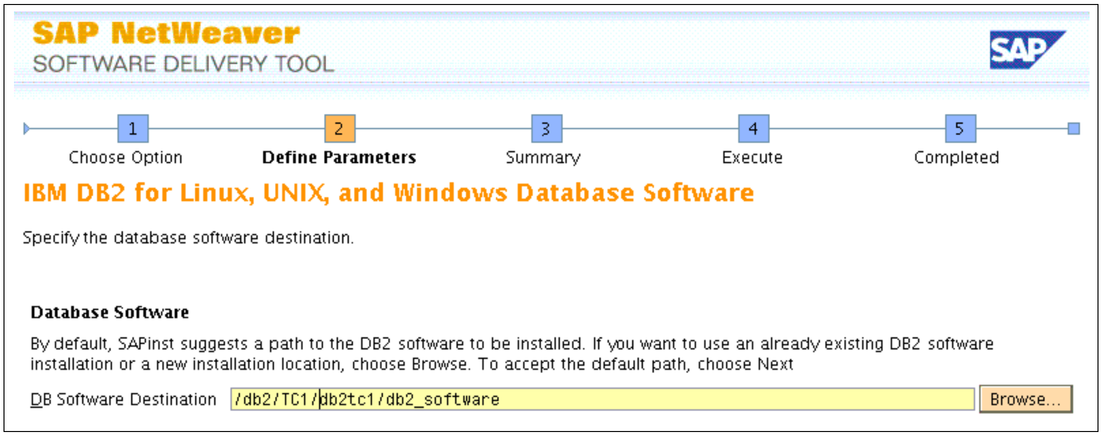



*Figure E-16 SAPinst database installation screen capture 6: Database software*

7. Enter the passwords for the already existing users as shown in Figure E-17, Figure E-18, and Figure E-19 on page 618.

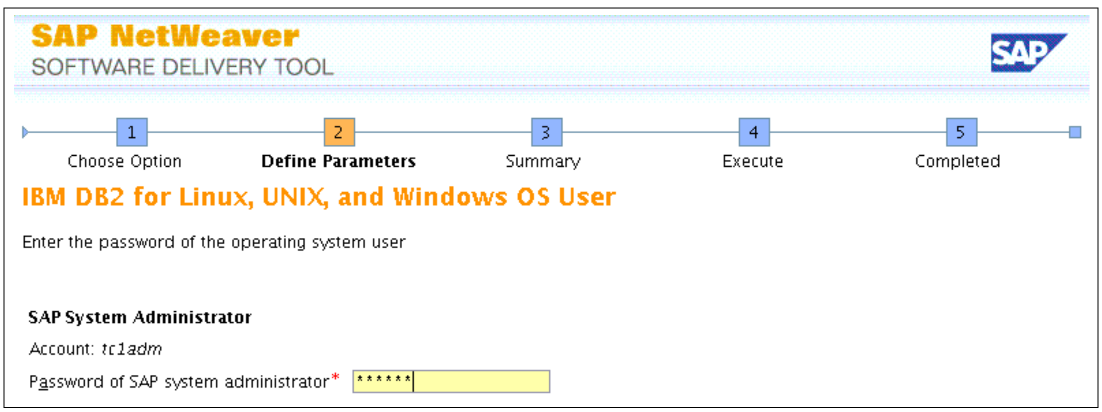



*Figure E-17 SAPinst database installation screen capture 7: tc1adm password*

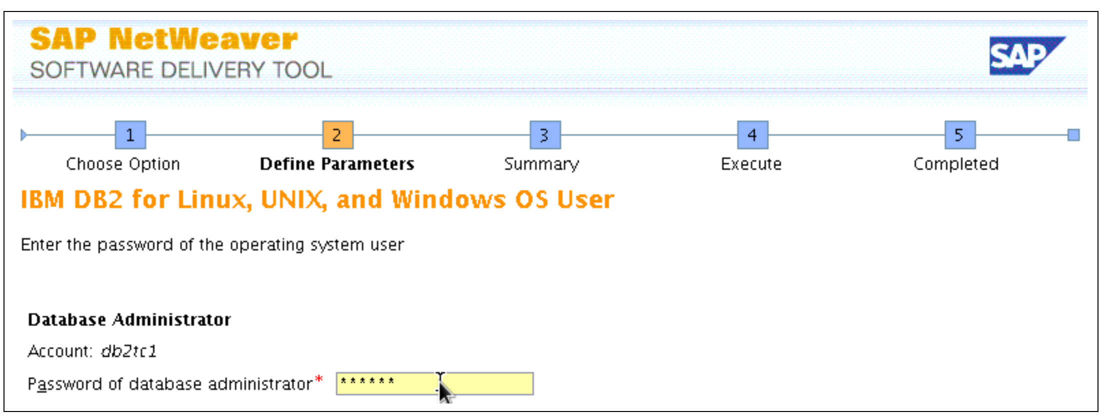



*Figure E-18 SAPinst database installation screen capture 8: db2tc1 password*

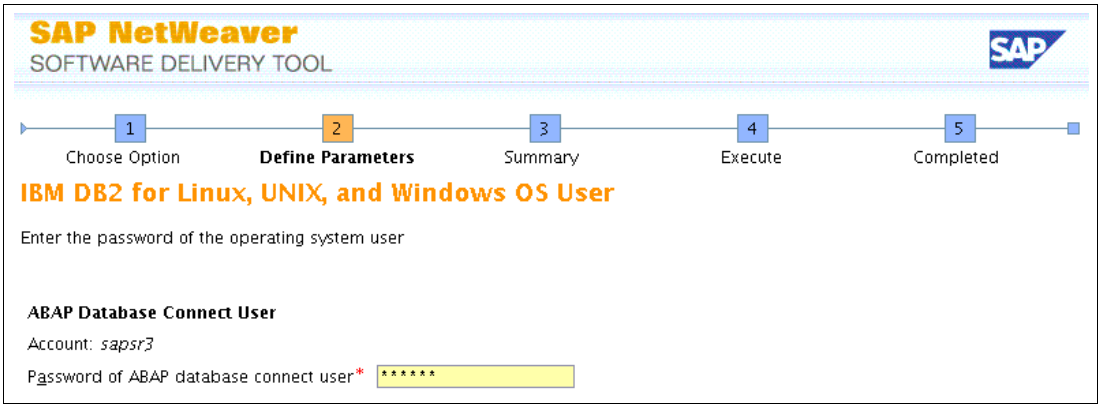



*Figure E-19 SAPinst database installation screen capture 9: sapsr3 password*

8. SAPinst now asks for the location of the UC Kernel NetWeaver 7.20 as shown in Figure E-20.

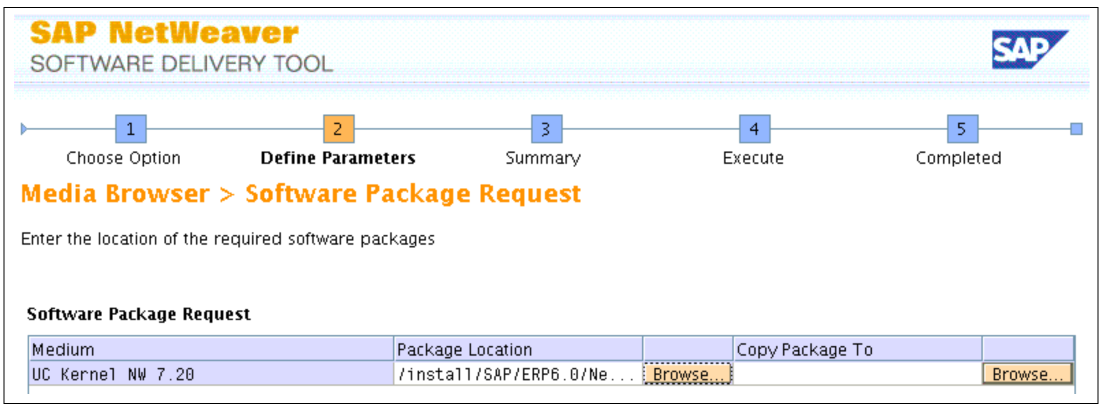



*Figure E-20 SAPinst database installation screen capture 10: UC Kernel NW 7.20 software package*

9. SAPinst asks for the location of the SAP installation export Supply Chain Management (SCM) directory as shown in Figure E-21.

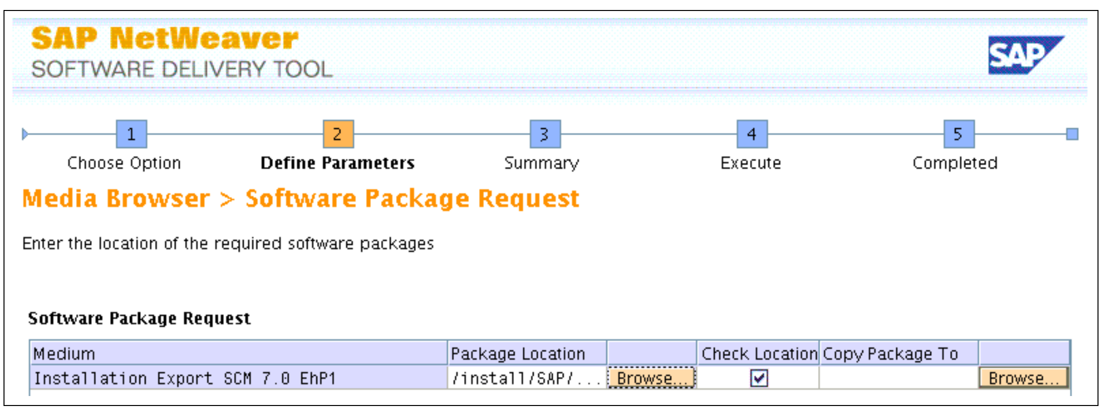



*Figure E-21 SAPinst database installation screen capture 11: Installation export SCM*

10. Enter the location of the installation files for DB2. See Figure E-22 on page 619.

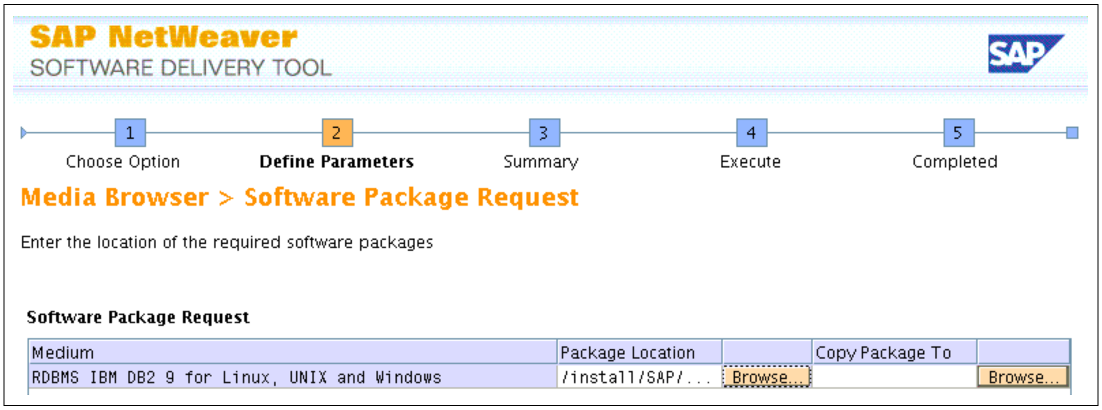



*Figure E-22 SAPinst database installation screen capture 12: DB2 software package*

11. SAPinst needs to know the DB2 database communication ports. We use the defaults. If you want to install more than one DB2 instance, you must define different ports for every database installation. See Figure E-23.

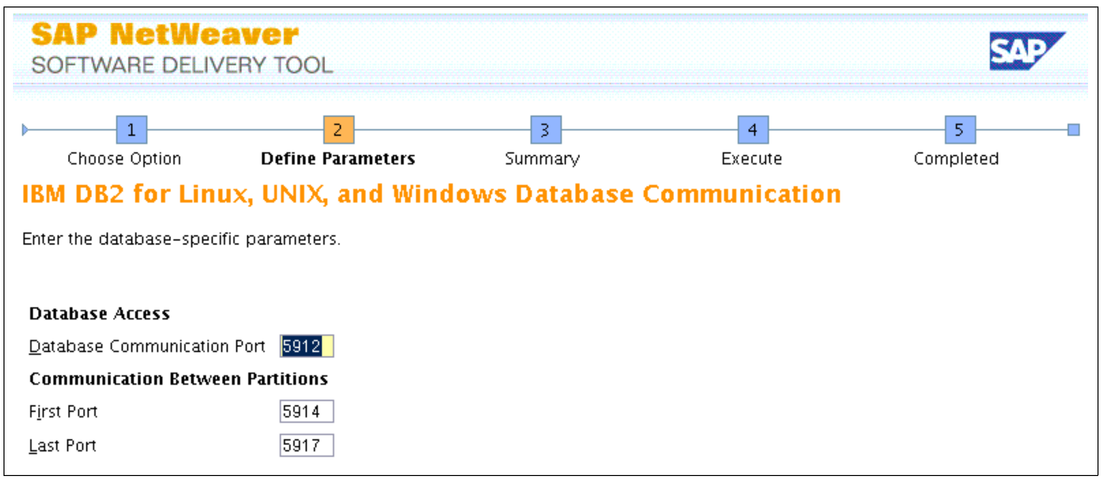



*Figure E-23 SAPinst database installation screen capture 13: DB2 communication ports*

12. Insert the location of the DB2 client as shown in Figure E-24.

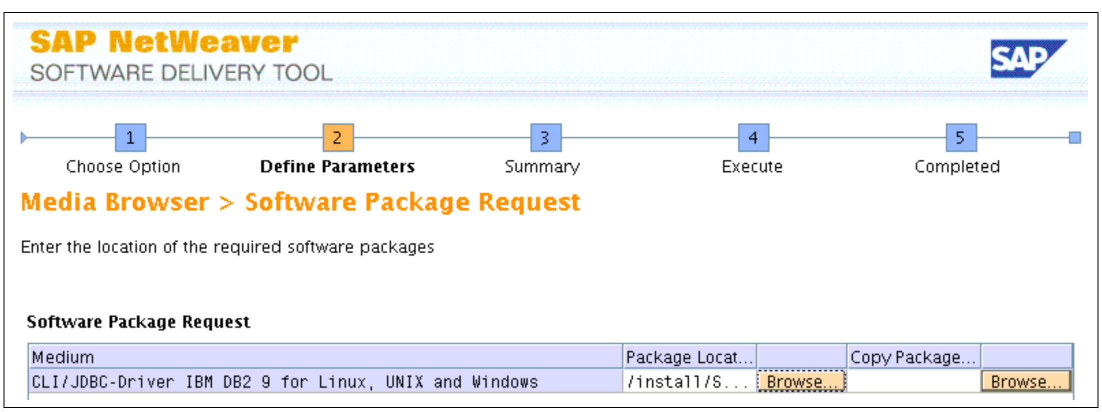



*Figure E-24 SAPinst database installation screen capture 14: DB2 CLI/JDBC driver*

13. Insert the instance memory for the DB2 instance. We use the default value as shown in Figure E-25 on page 620.

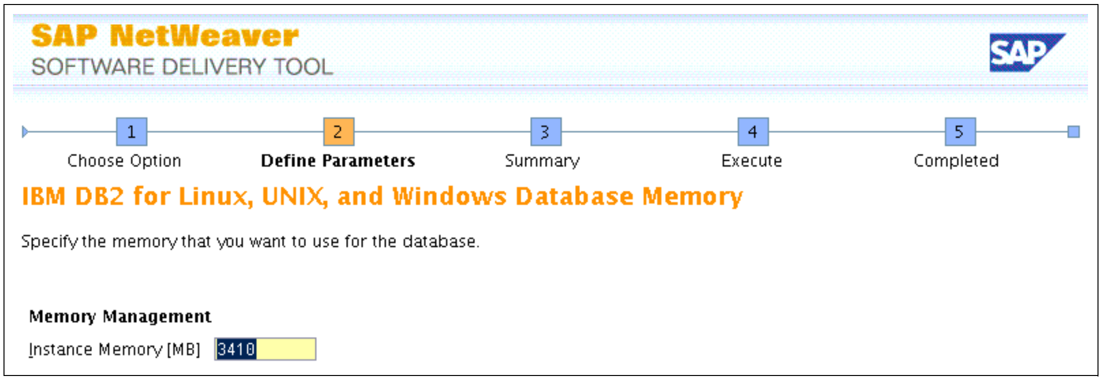



*Figure E-25 SAPinst database installation screen capture 15: Instance Memory*

14. The SAPinst asks for the destination path for the liveCache software as shown in Figure E-26.

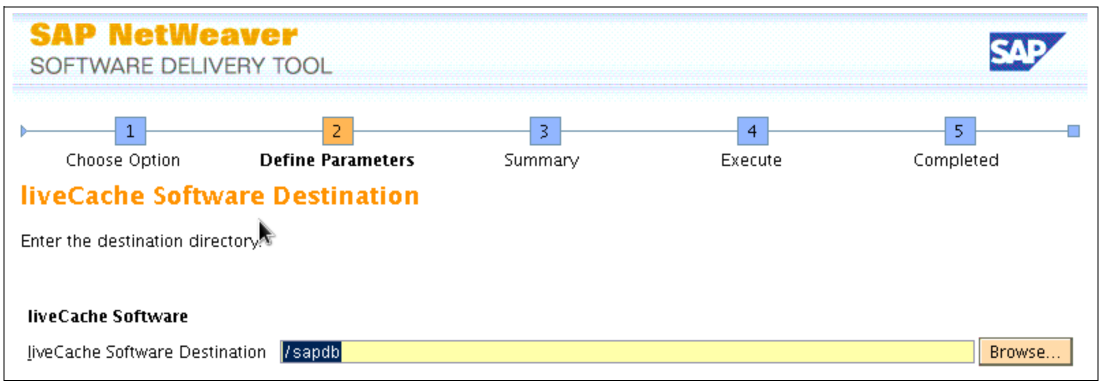



*Figure E-26 SAPinst database installation screen capture 16: liveCache software destination*

15. We use the defaults to create the DB2 database instance. See Figure E-27 and Figure E-28 on page 621.

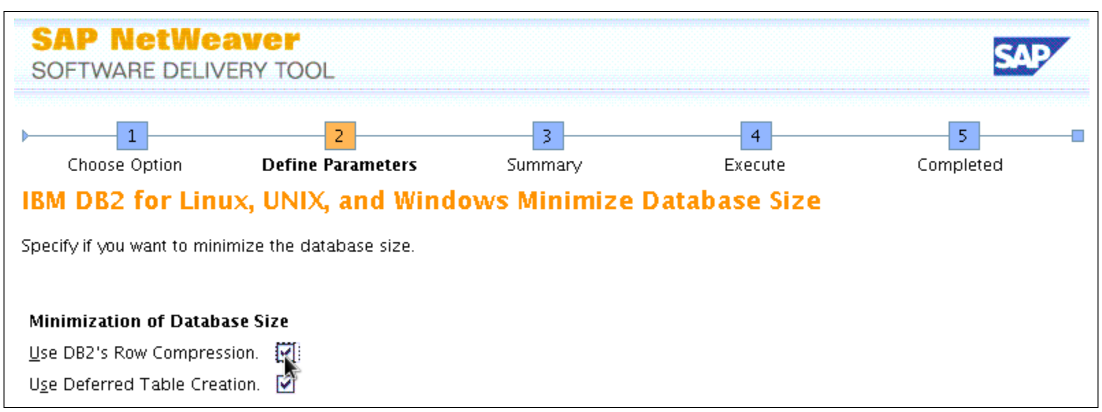



*Figure E-27 SAPinst database installation screen capture 17: DB2 database settings*

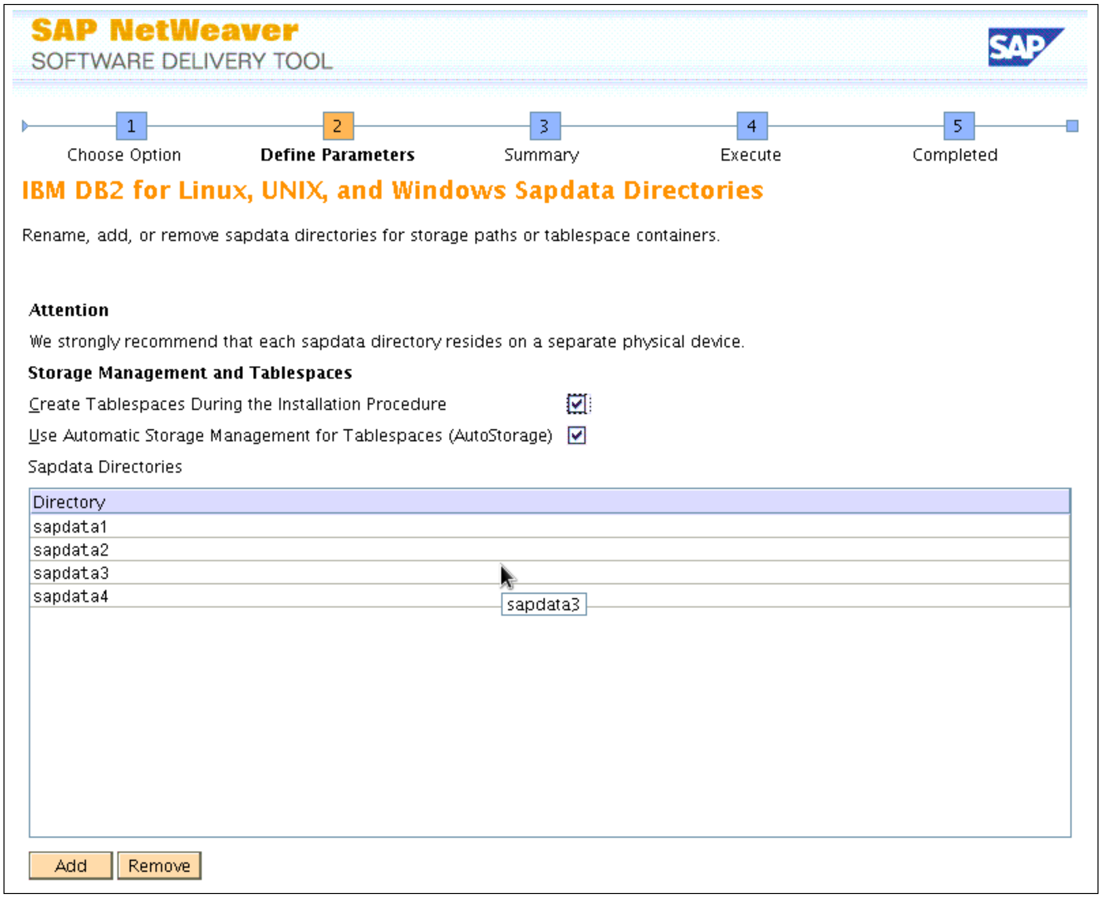



Figure E-28 SAPinst database installation screen capture 18: DB2 database settings

16. We want the SAPinst to use nine parallel load processes. The number depends on the allocatable CPUs that you have in the LPAR. We enter `9`. See Figure E-29.

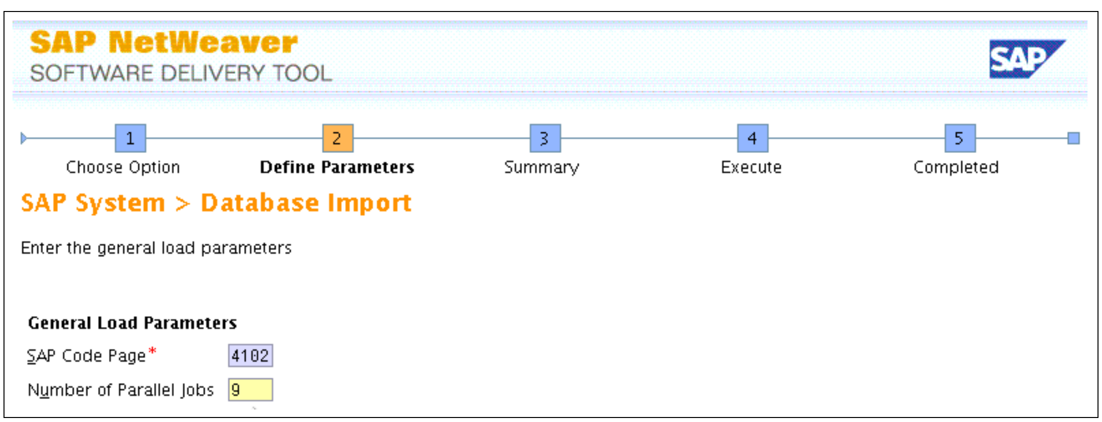



Figure E-29 SAPinst database installation screen capture 19: SAP import settings

17. We select **Unpack** to let SAPinst unpack the necessary SAR files as shown in Figure E-30 on page 622.

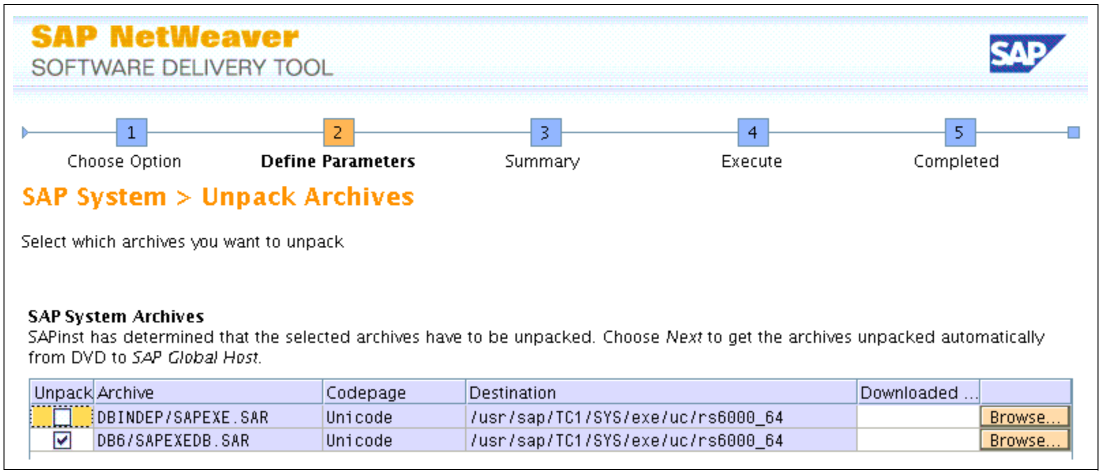



*Figure E-30   SAPinst database installation screen capture 20: SAP archives to unpack*

18. Insert the liveCache database software owner as shown in Figure E-31.

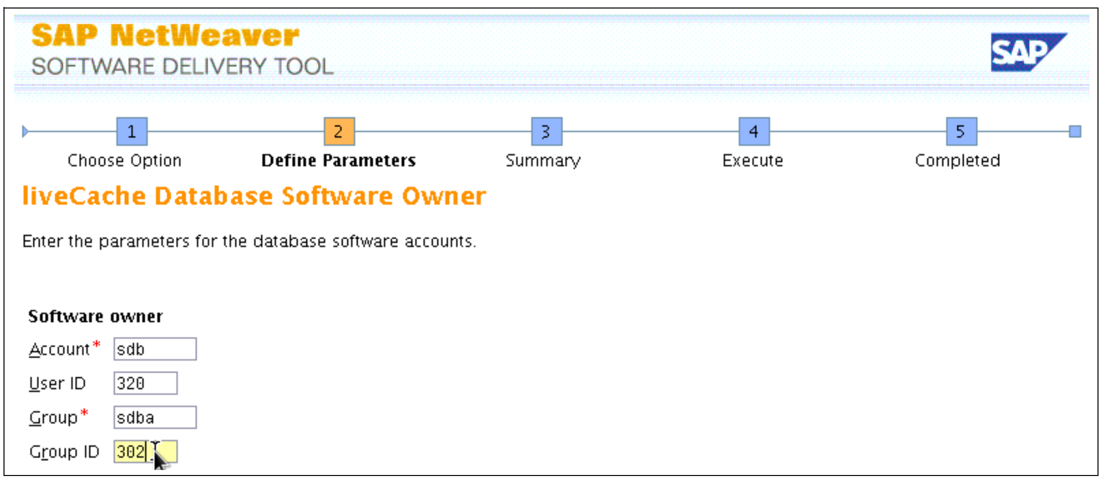



*Figure E-31   SAPinst database installation screen capture 21: liveCache database software owner*

19. Select the location of the liveCache software as shown in Figure E-32.

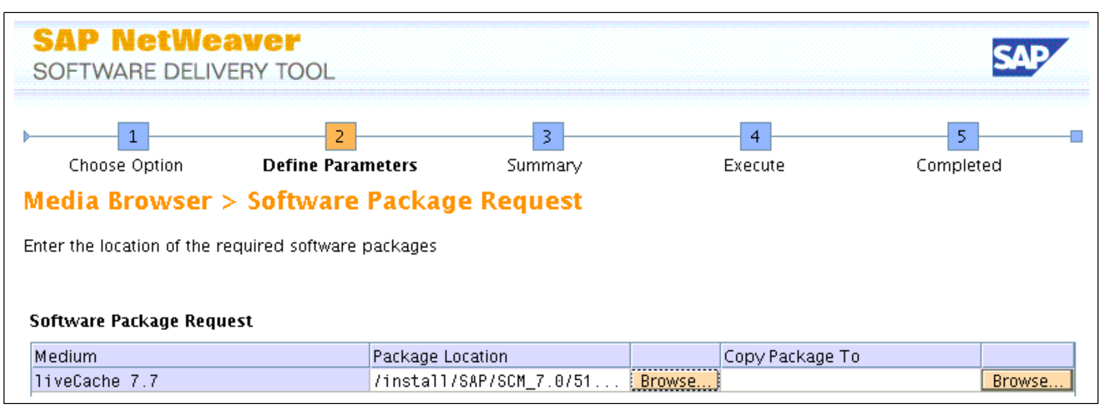



*Figure E-32   SAPinst database installation screen capture 22: liveCache software package location*

## sapci_cluster: Installing the SAP Enqueue Replication Service (ERS) instance on the second node

For the ERS installation with SAPinst, a virtual SAP host name (`sapcisvc3`) is used. The advantage is that the ERS processes can be monitored from outside if there is a service address configured. This option is not the default for the Smart Assist; it sets up the ERS without a service IP.

The usual SAP parameters are set as seen in Example E-12.

*Example E-12 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc3
```

Figure E-33 shows the start of the SAPinst ERS installation.

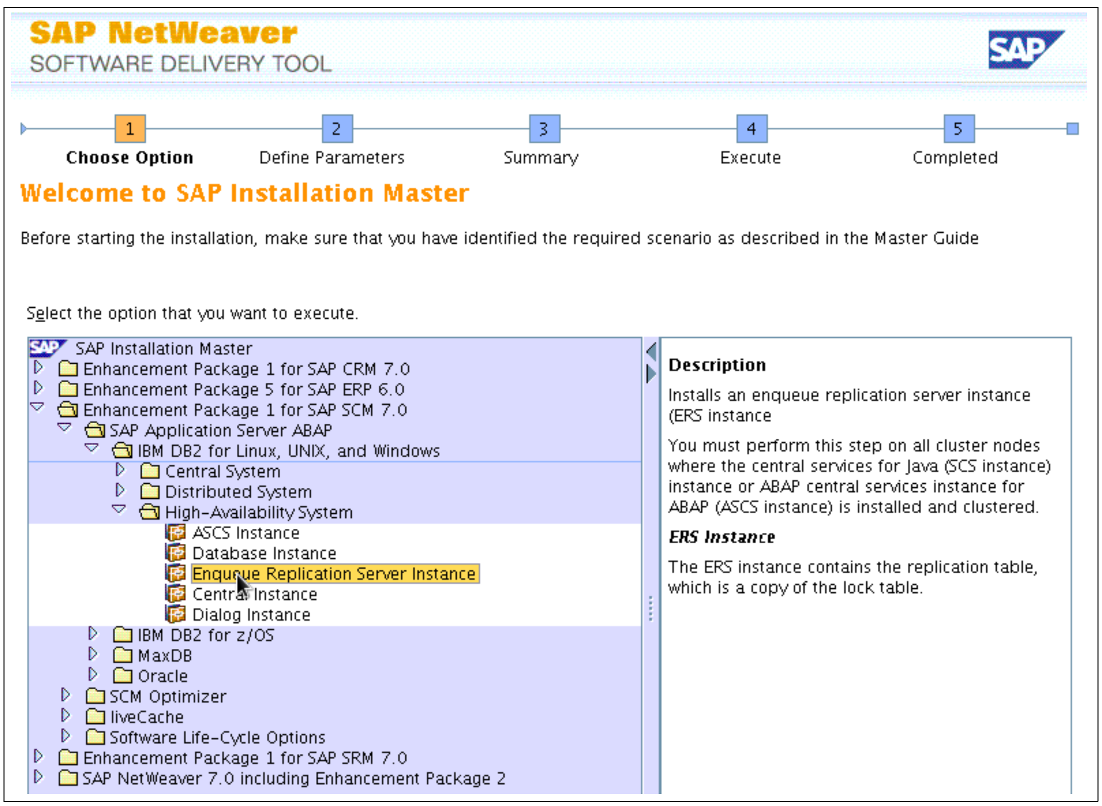



*Figure E-33 SAPinst ERS installation screen capture 1: Start installation*

Follow these steps:

1. Enter the profile path of the SAP system as shown in Figure E-34 on page 624.

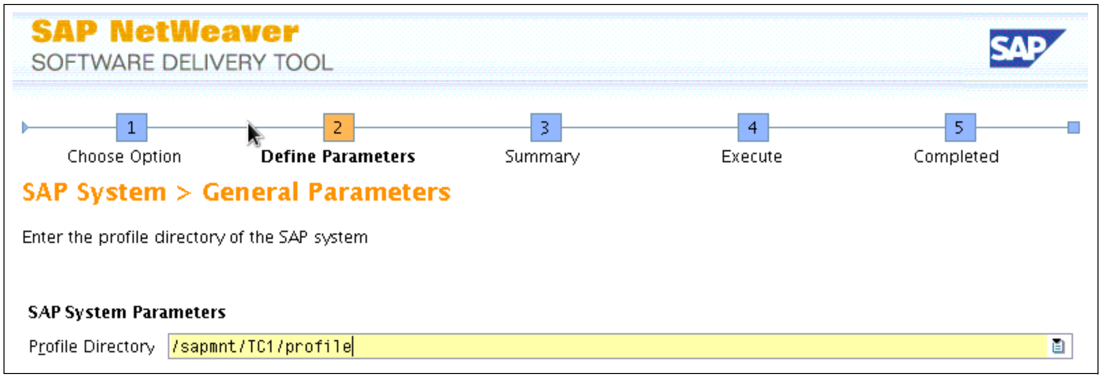



*Figure E-34 SAPinst ERS installation screen capture 2: /sapmnt directory*

2. Select the enqueue server for which you want to install the ERS. See Figure E-35.

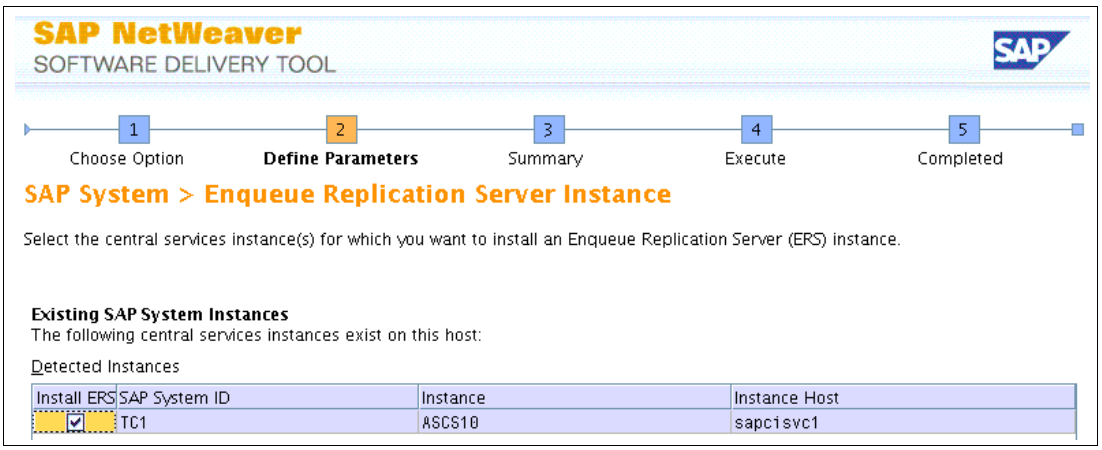



*Figure E-35 SAPinst ERS installation screen capture 3: Select the appropriate ASCS*

3. The SAPinst asks for the location of the UC Kernel NetWeaver 7.20 as shown in Figure E-36.

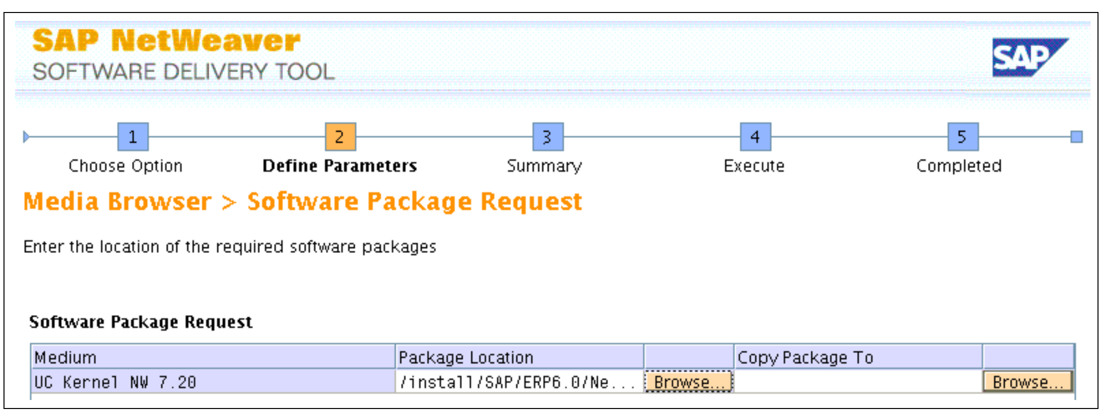



*Figure E-36 SAPinst ERS installation screen capture 4: UC Kernel NW 7.20 location*

4. Enter the instance number that you want to use for the ERS as shown in Figure E-37 on page 625.

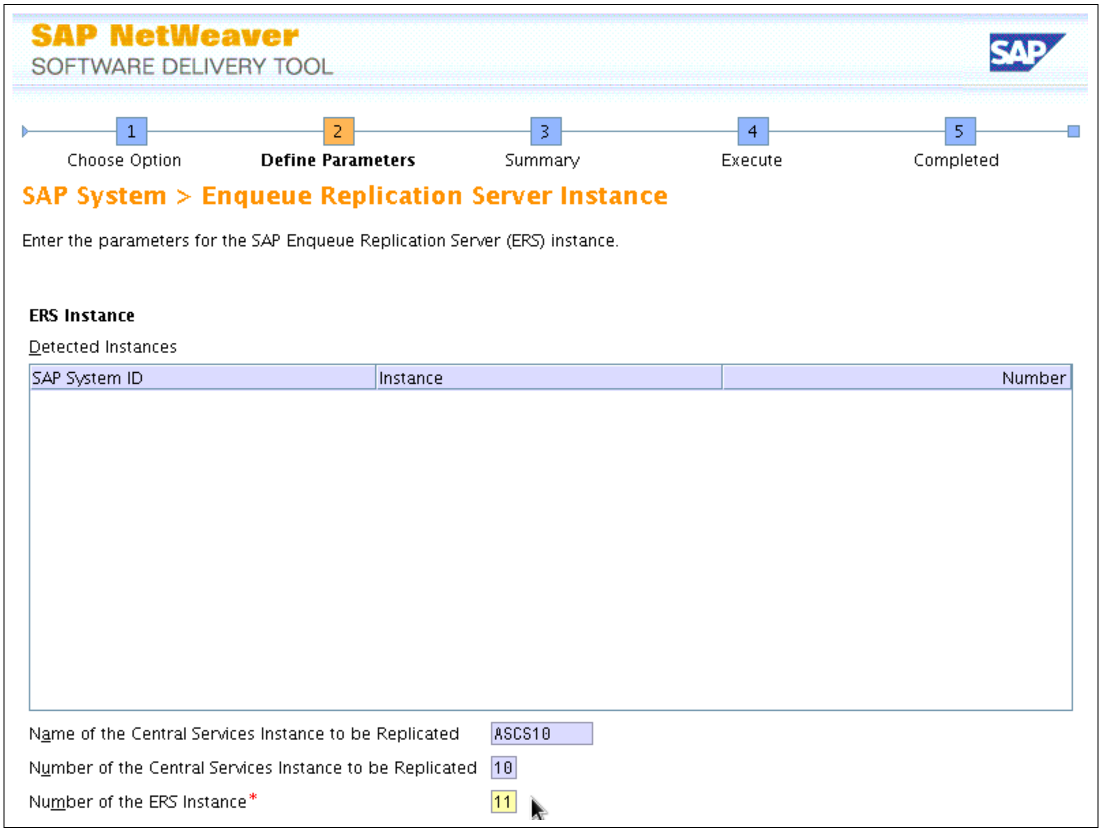



*Figure E-37 SAPinst ERS installation screen capture 5: ERS instance number*

5. Select **Get the ASCS Instance restarted** so that SAPinst automatically restarts the ASCS instance as shown in Figure E-38.

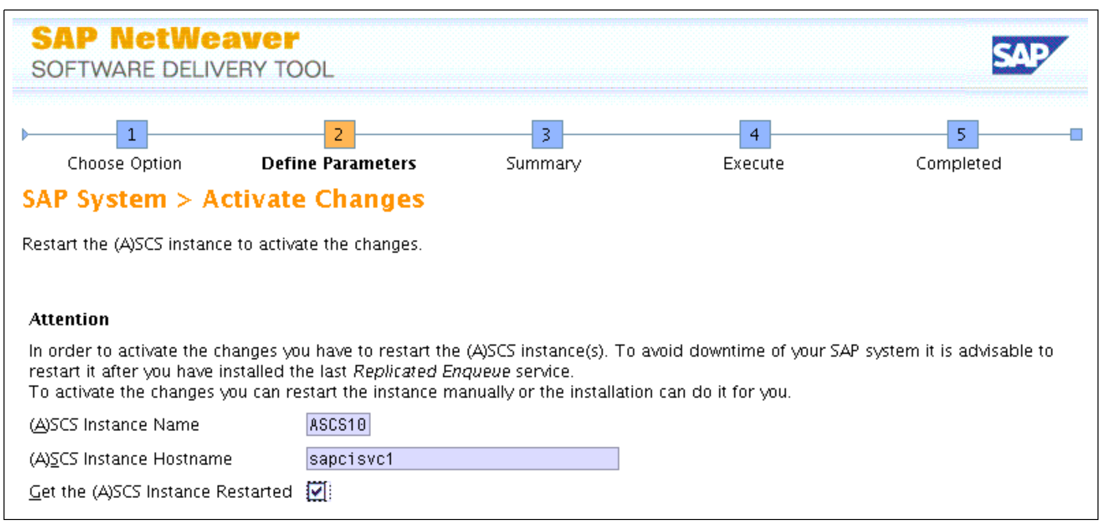



*Figure E-38 SAPinst ERS installation screen capture 6: SAPinst restarts ASCS*

SAPinst installs the ERS.

## sapci_cluster: Installing the SAP central instance on the first node

For the ERS installation with the SAPinst, a virtual SAP host name (`sapcisvc4`) is used. The normal SAP parameters are set as shown in Example E-13.

*Example E-13 SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=sapcisvc4
```

Figure E-39 shows the SAPinst CI installation start.

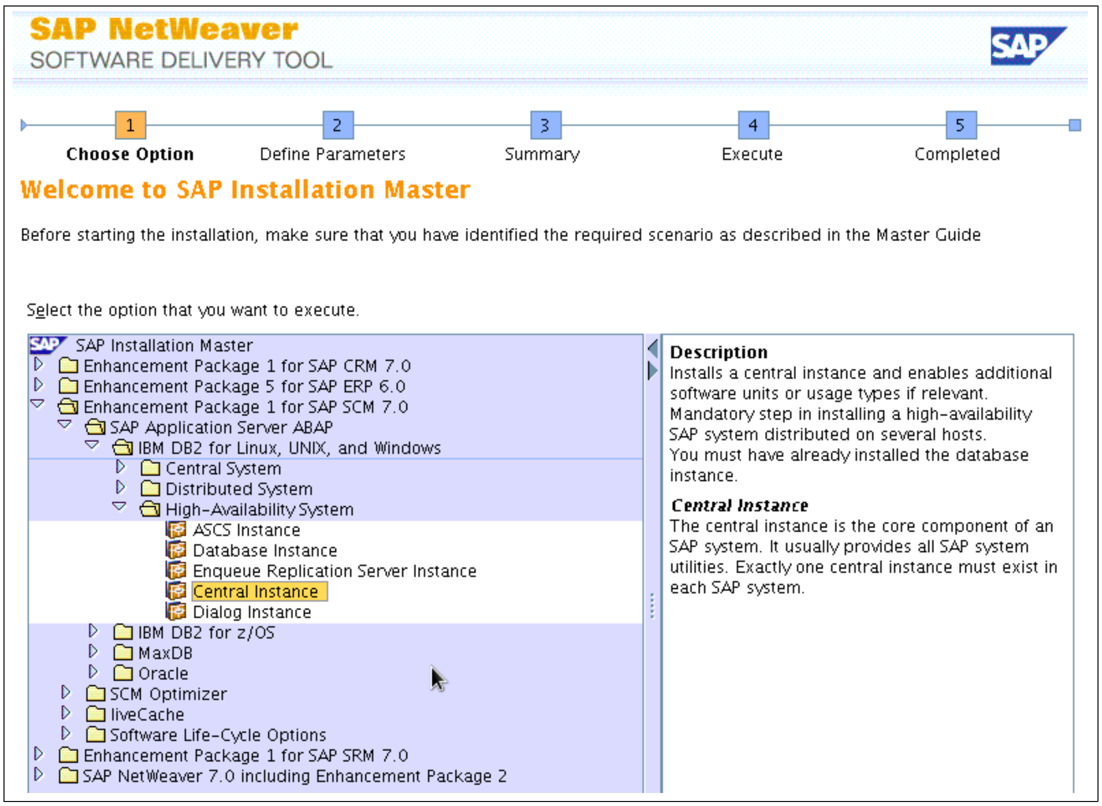



*Figure E-39 SAPinst CI installation screen capture 1: Start installation*

Follow these steps:

1. Enter the SAP system profile directory as shown in Figure E-40 on page 627.

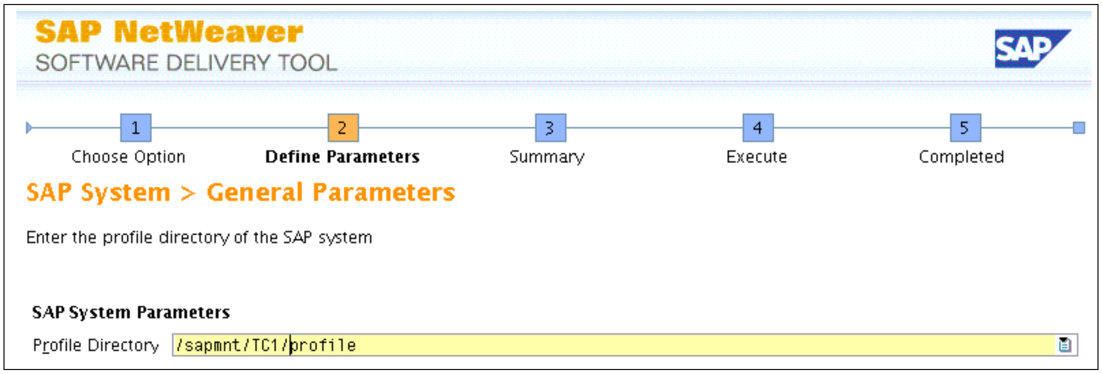



*Figure E-40 SAPinst CI installation screen capture 2: /sapmnt directory*

2. You have to set up a master password. This password is used by all users that SAPinst creates as shown in Figure E-41.

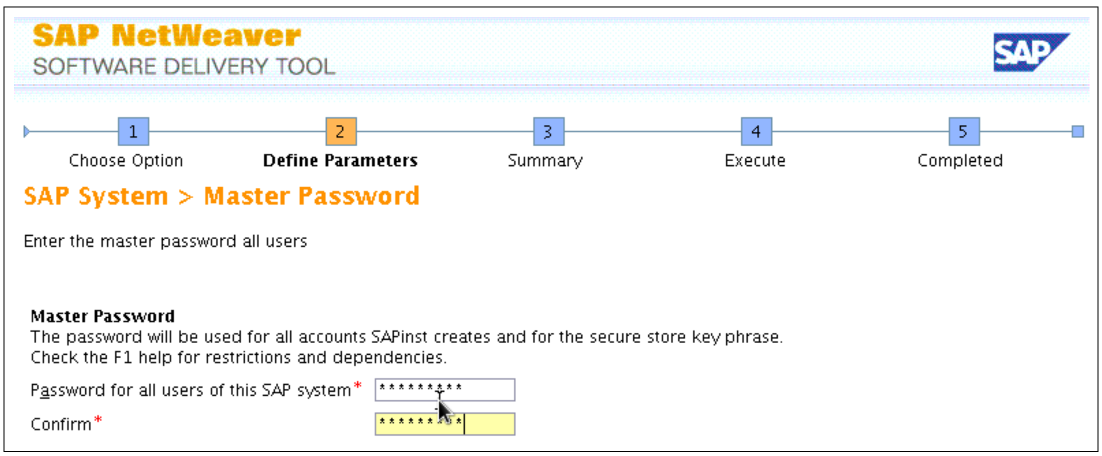



*Figure E-41 SAPinst CI installation screen capture 3: Master password*

3. You have to enter the database connect user as shown in Figure E-42. In the past, SAP uses `sapr3` and in the newer releases `sap<sid>`, but we suggest `sapsr3`. The user name `sapr3` is no longer valid because the user name must be a minimum of six characters.

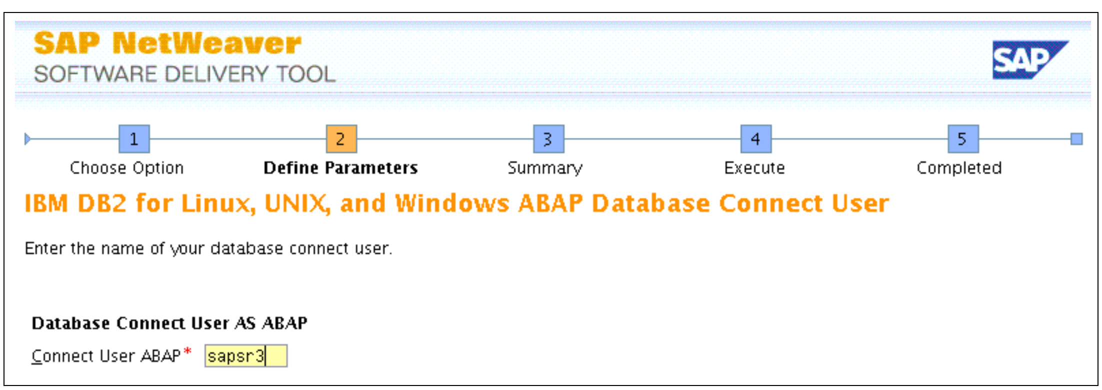



*Figure E-42 SAPinst CI installation screen capture 4: Database connect user*

4. SAPinst asks for the location of the UC Kernel NetWeaver 7.20 as shown in Figure E-43 on page 628.

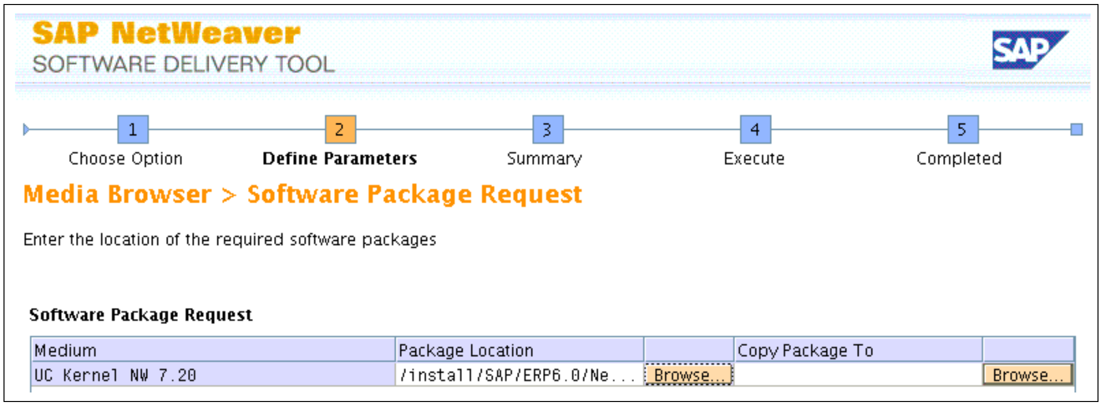



*Figure E-43 SAPinst CI installation screen capture 5: Enter the UC Kernel NW 7.20 location*

5. Figure E-44 shows the liveCache database software owner.

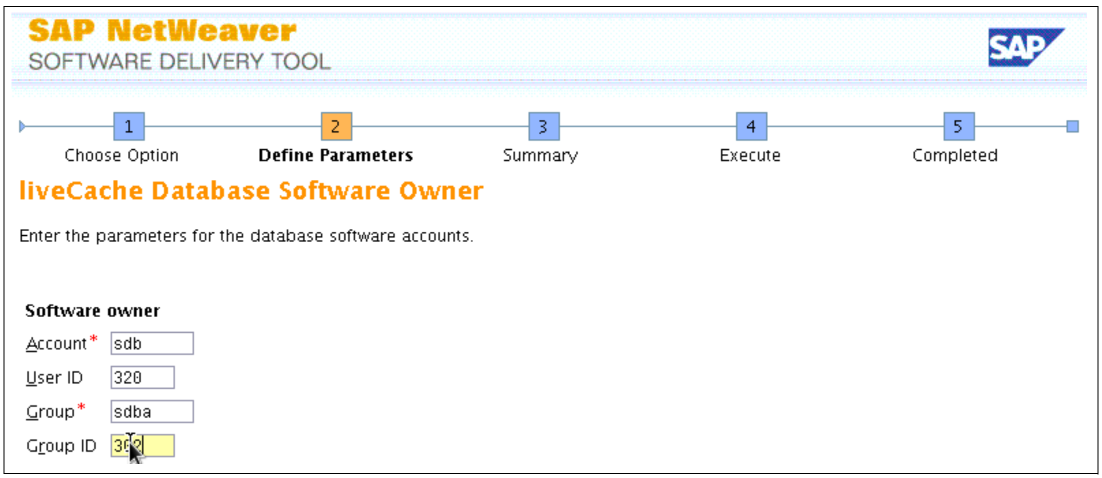



*Figure E-44 SAPinst CI installation screen capture 6: liveCache database software owner*

6. SAPinst asks for the location of the liveCache installation media as shown in Figure E-45.

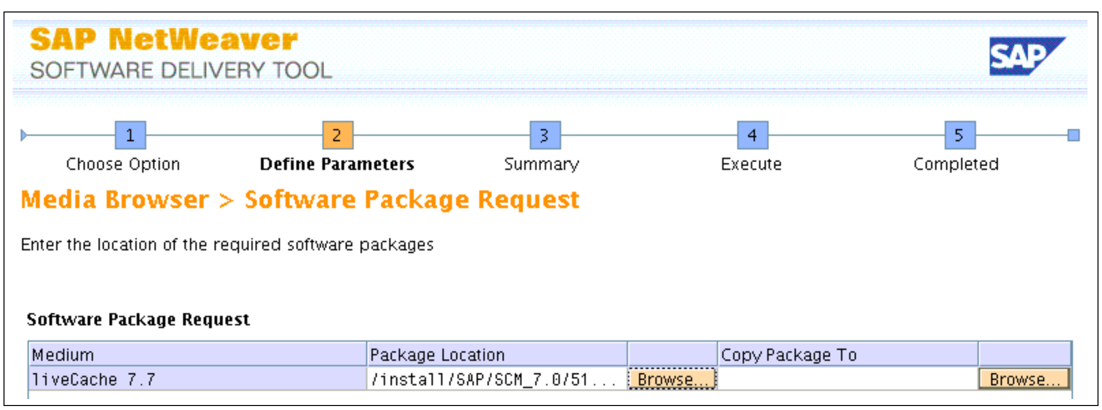



*Figure E-45 SAPinst CI installation screen capture 7: Location of liveCache installation*

The liveCache software directory is local, not shared. See Figure E-46 on page 629.

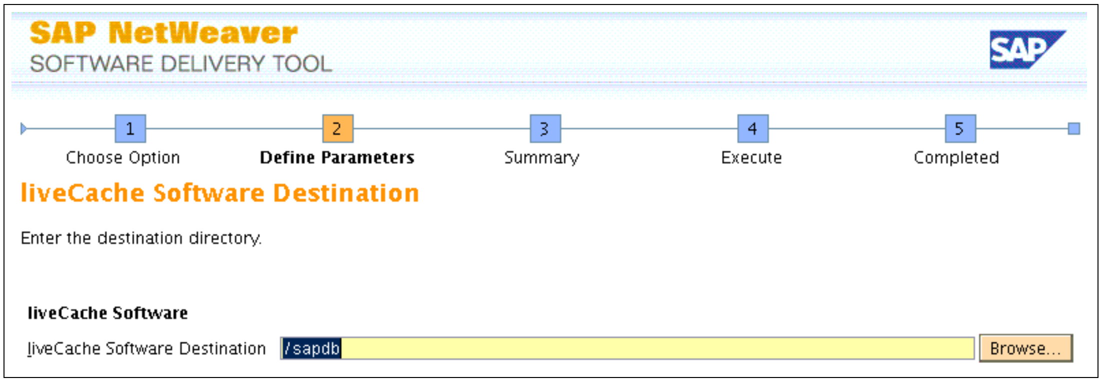



*Figure E-46 SAPinst CI installation screen capture 8: sapdb binary installation directory*

7. Enter the instance number that you want to use for the central instance as shown in Figure E-47.


*Figure E-47 SAPinst CI installation screen capture 9: CI instance number*

8. We used the master password for the Data Dictionary (DDIC) user as shown in Figure E-48 on page 630.

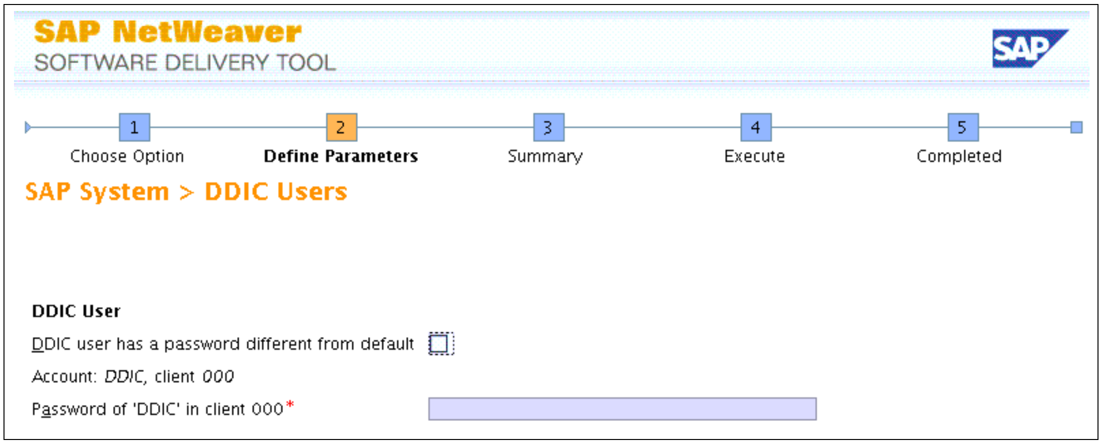



*Figure E-48 SAPinst CI installation screen capture 10: DDIC password*

9. This SAP kernel DVD differs from the first UC kernel DVD. In our case, it is labeled `SAP:AKK:701:DVD_KERNEL:SAP Kernel 7.01:D51035686`. See Figure E-49.

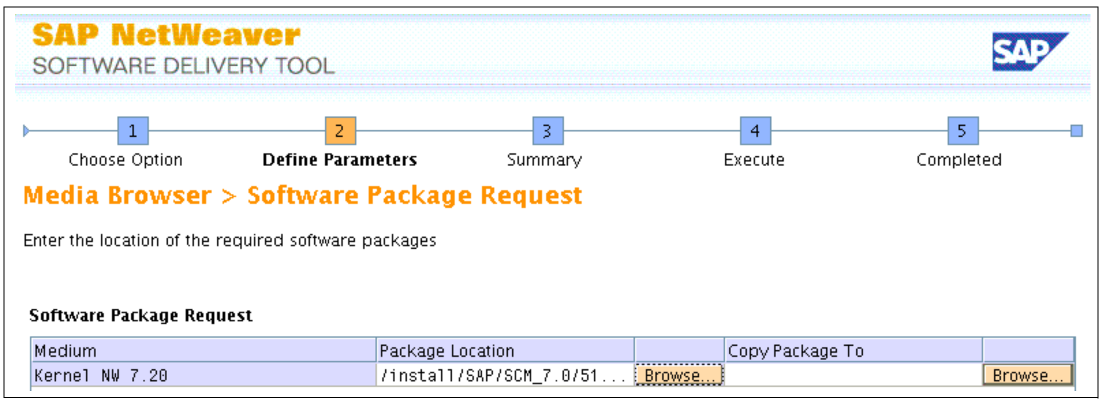



*Figure E-49 SAPinst CI installation screen capture 11: Kernel NW 7.20*

10. Click **Next** for SAPinst to unpack the necessary SAR files as shown in Figure E-50.

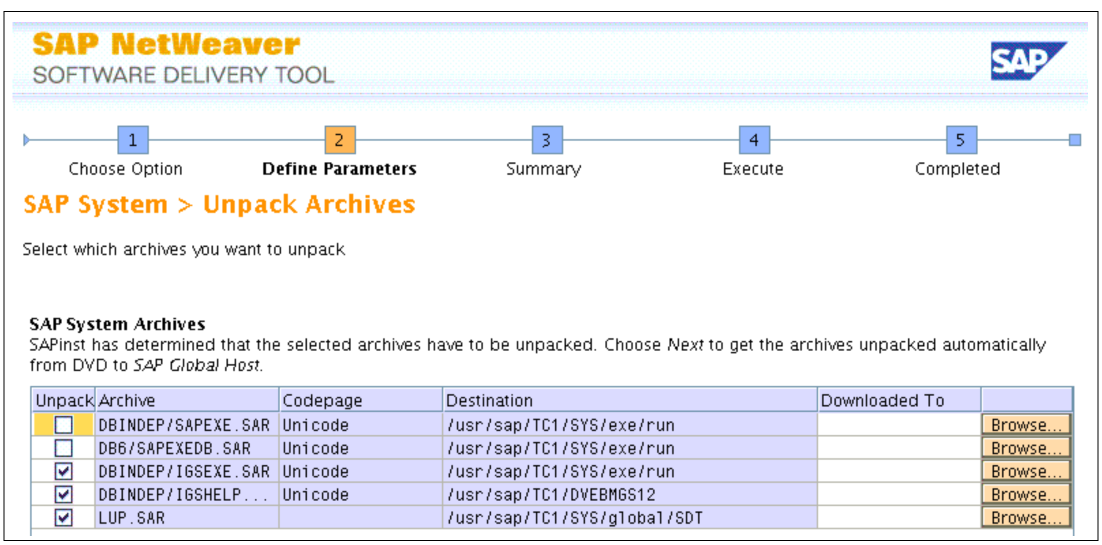



*Figure E-50 SAPinst CI installation screen capture 12: SAR archives to unpack*

11. Enter the liveCache Server Connection information as shown in Figure E-51 on page 631.

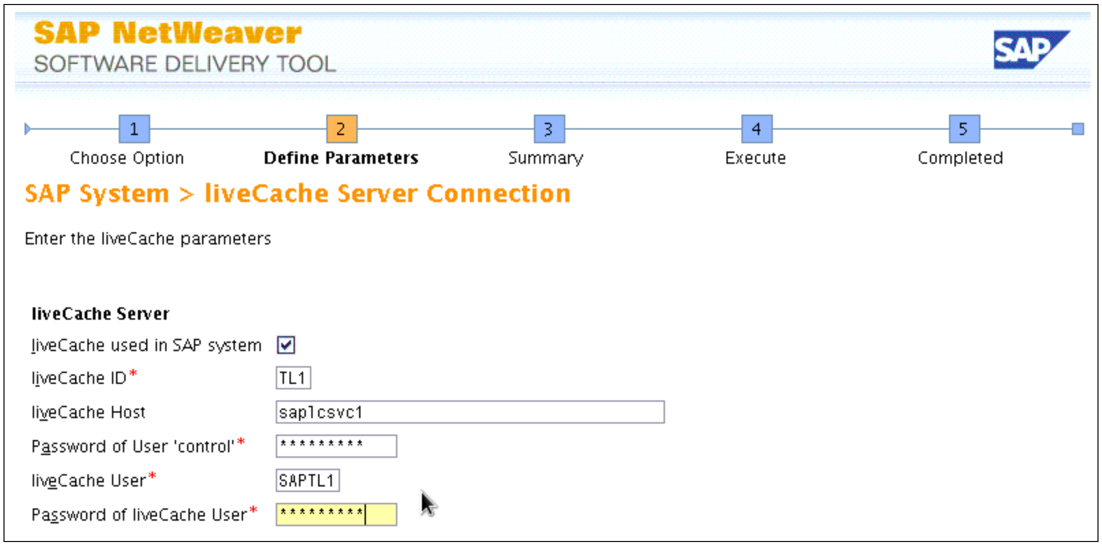



*Figure E-51 SAPinst CI installation screen capture 13: liveCache server connection*

12. We enter the SAPinst CI installation diagnostics agent system ID (Figure E-52).

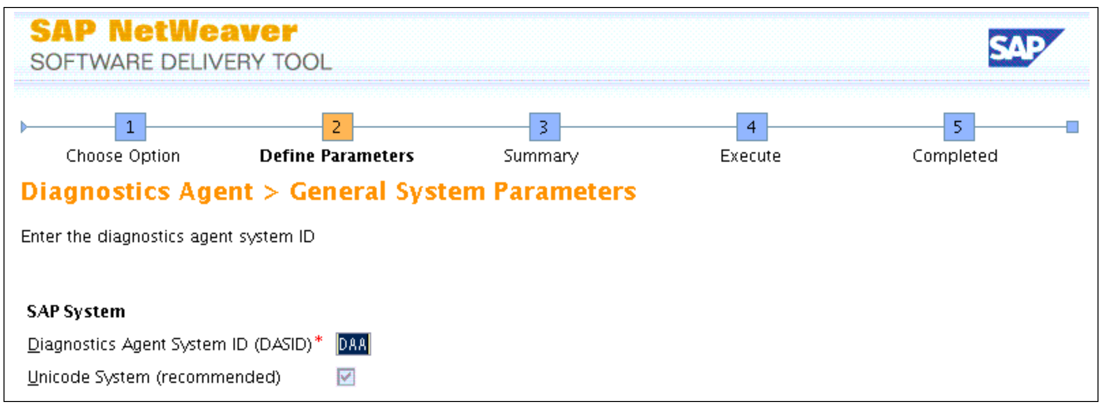



*Figure E-52 SAPinst CI installation screen capture 14: Diagnostics agent system ID*

13. Figure E-53 shows the path to the Java Cryptography Extension (JCE) policy file. In this version, we need the file for JAVA 6.

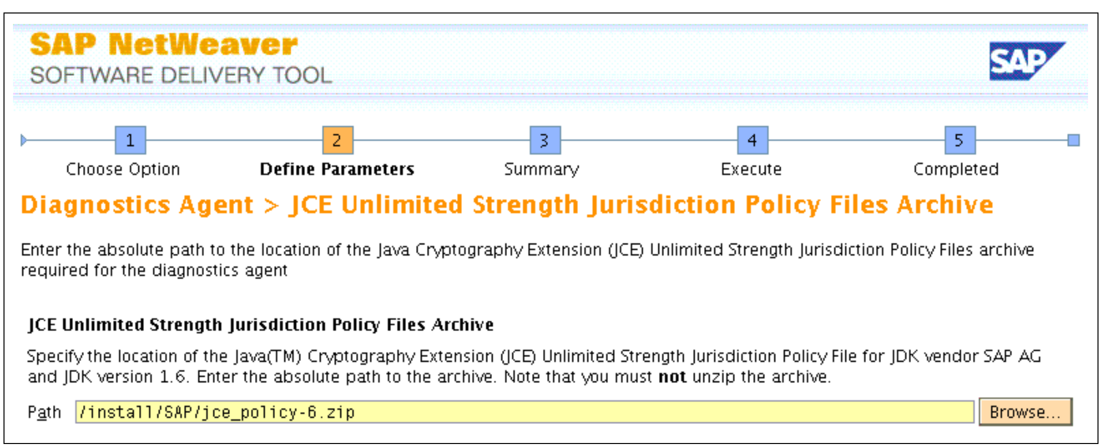



*Figure E-53 SAPinst CI installation screen capture 15: JCE policy file*

14.The user ID for the diagnostic agent is created if it does not exist. See Figure E-54.

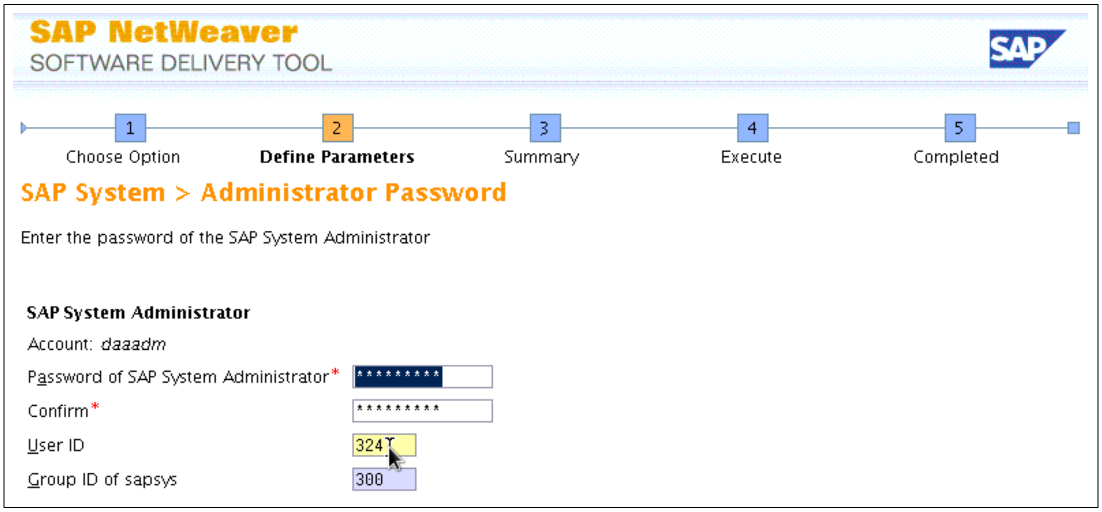



*Figure E-54 SAPinst CI installation screen capture 16: Diagnostics agent user*

15.Figure E-55 shows the instance number for the diagnostics agent.

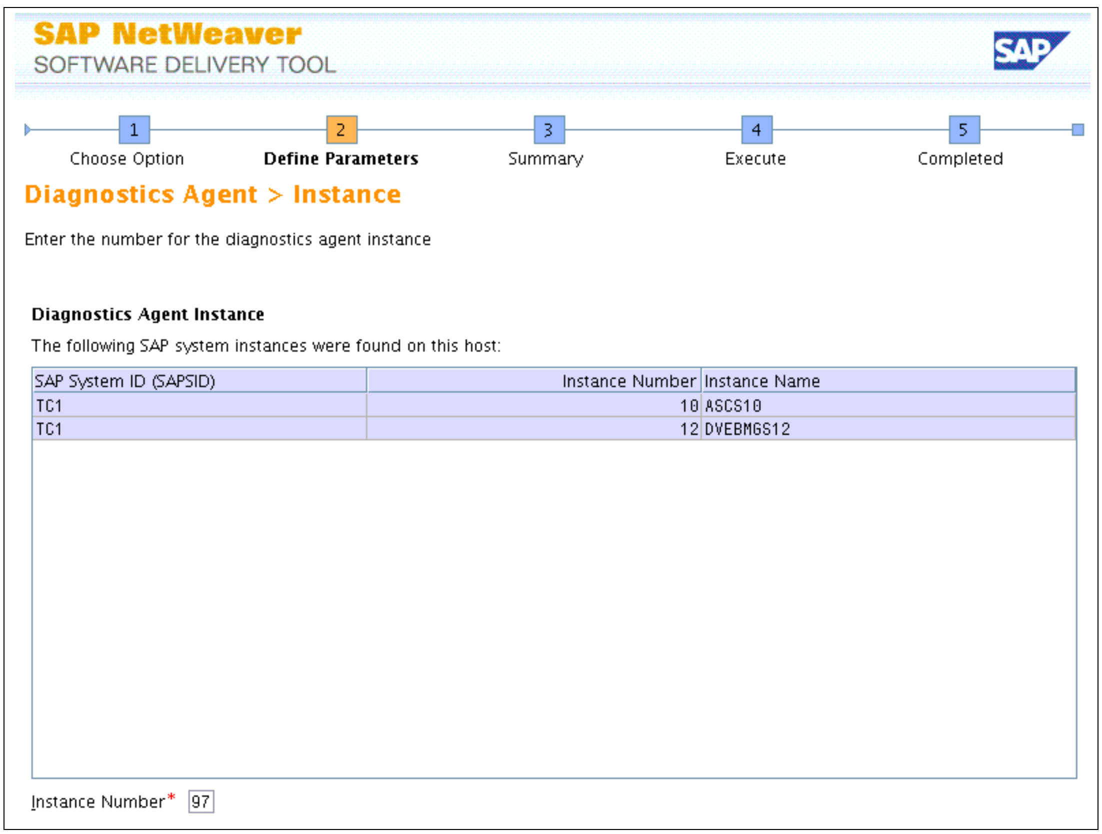



*Figure E-55 SAPinst CI installation screen capture 17: CI instance number*

16.We have no System Landscape Directory (SLD) that is configured in our demonstration environment, so we did not configure it. See Figure E-56 on page 633.

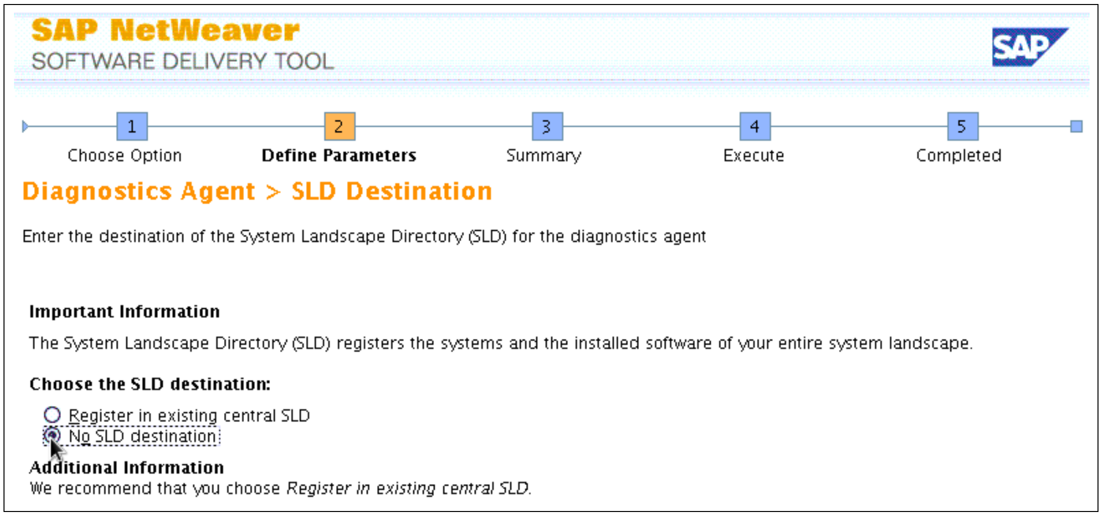



*Figure E-56   SAPinst CI installation screen capture 18: SLD configuration*

17. Click **Next** for SAPinst to unpack the necessary SAR archives as shown in Figure E-57.

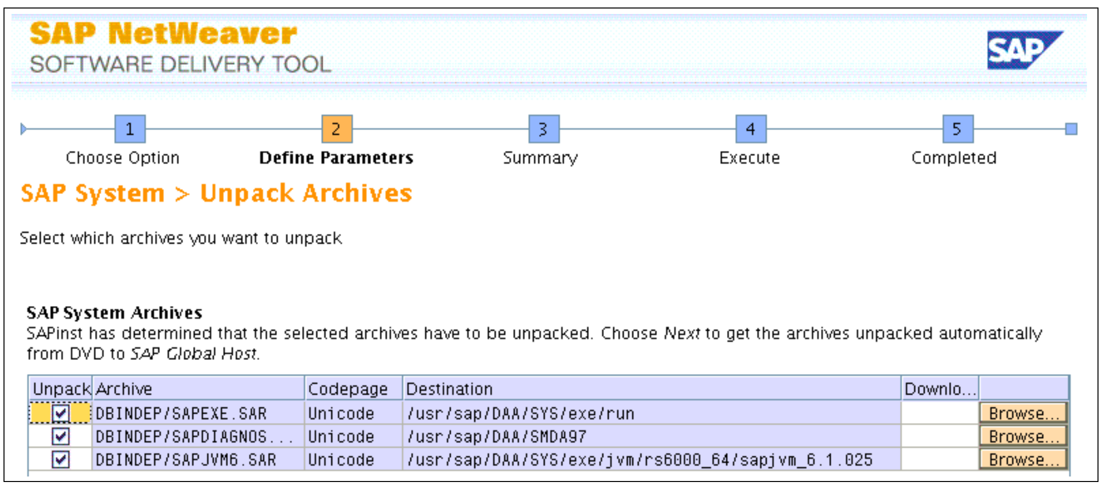



*Figure E-57   SAPinst CI installation screen capture 19: SAR archives to unpack*

## saplc_cluster: Installing the SAP liveCache in the first node

We performed an SAP installation with a virtual SAP host name. Only the normal SAP parameters are set as shown in Example E-14.

*Example E-14   SAP sapinst environment parameters*

```
export TMPDIR=/tmp/TC1
mkdir $TMPDIR
export JAVA_HOME=/usr/java14_64
./sapinst SAPINST_USE_HOSTNAME=saplcsvc1
```

SAPinst in the installation DVD provides options when you create a liveCache instance. Important entries are highlighted with the screen captures. You can install the MAXDB database first, but it is easier for SAPinst to install it for you.

Follow these steps:

1. Start the liveCache installation as shown in Figure E-58. Select **liveCache Server Installation**. The following panels might appear in a different order. It depends on the SAPinst version that you use.

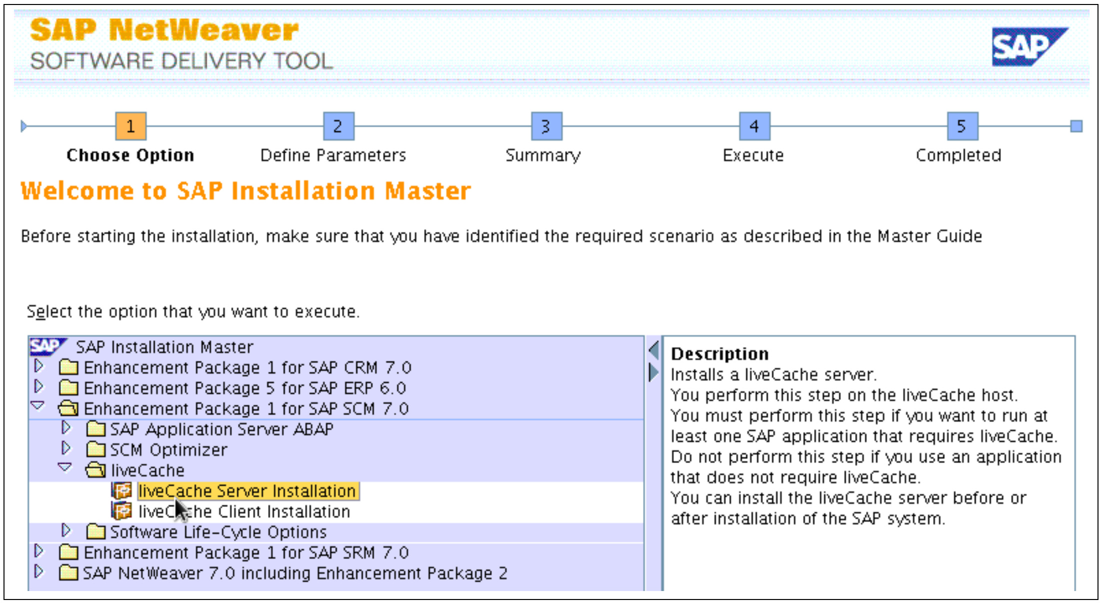



*Figure E-58 SAPinst liveCache installation screen capture 1: Start the database installation*

2. Always select **Custom** installation because you can change the default values. See Figure E-59.

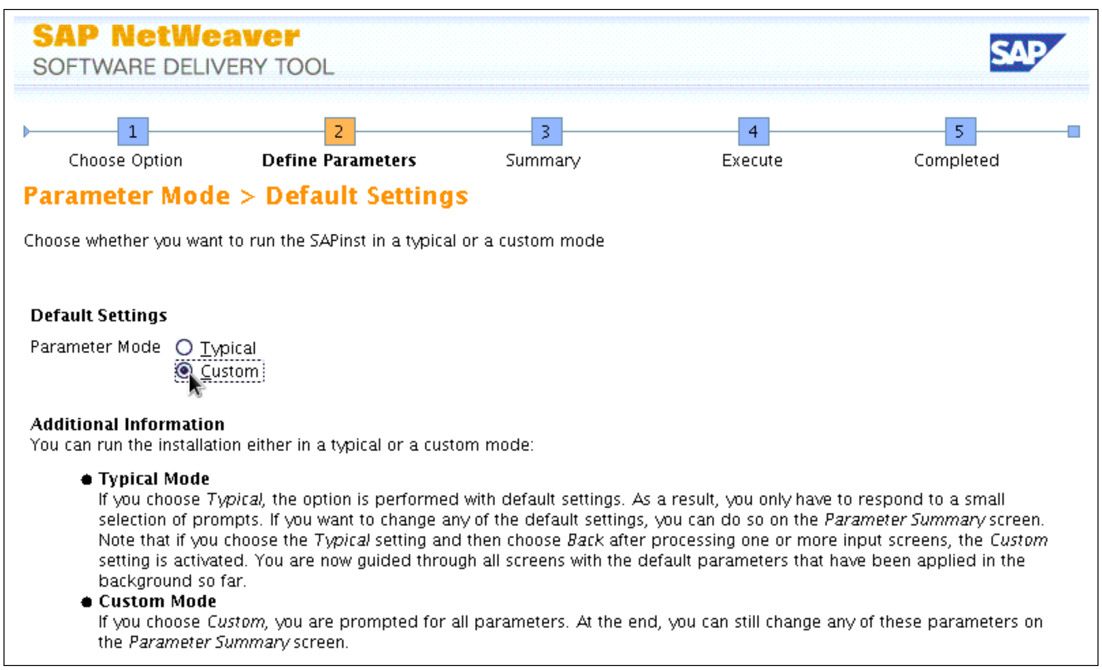



*Figure E-59 SAPinst liveCache installation screen capture 2: Custom installation*

3. In the next window, ensure that liveCache knows whether the Advanced Planning and Optimization (APO) system is Unicode or non-Unicode. Newer installations are only Unicode. If this value is left to the default value, you must manually correct it later. The liveCache ID is the SID that you selected. See Figure E-60.

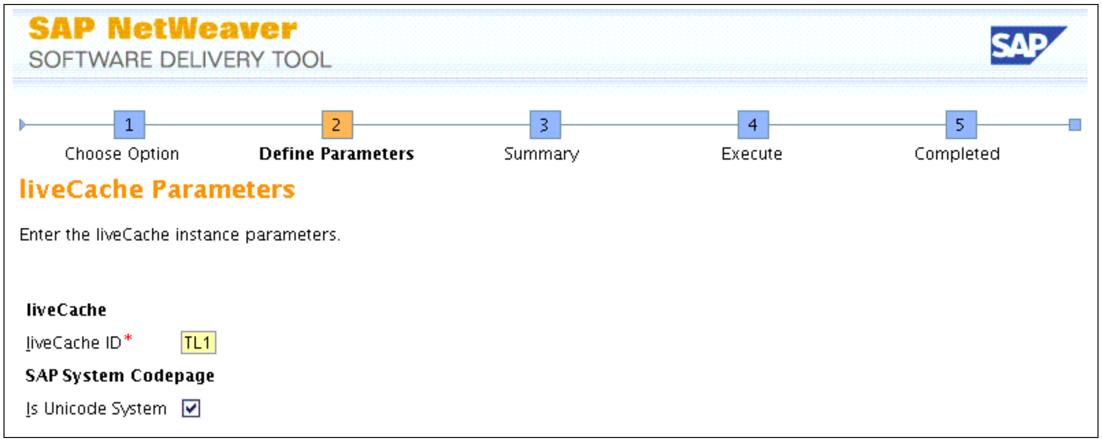



*Figure E-60 SAPinst liveCache installation screen capture 3: liveCache ID*

4. You must set up a master password as shown in Figure E-61. This password is used by all users that the SAPinst creates.

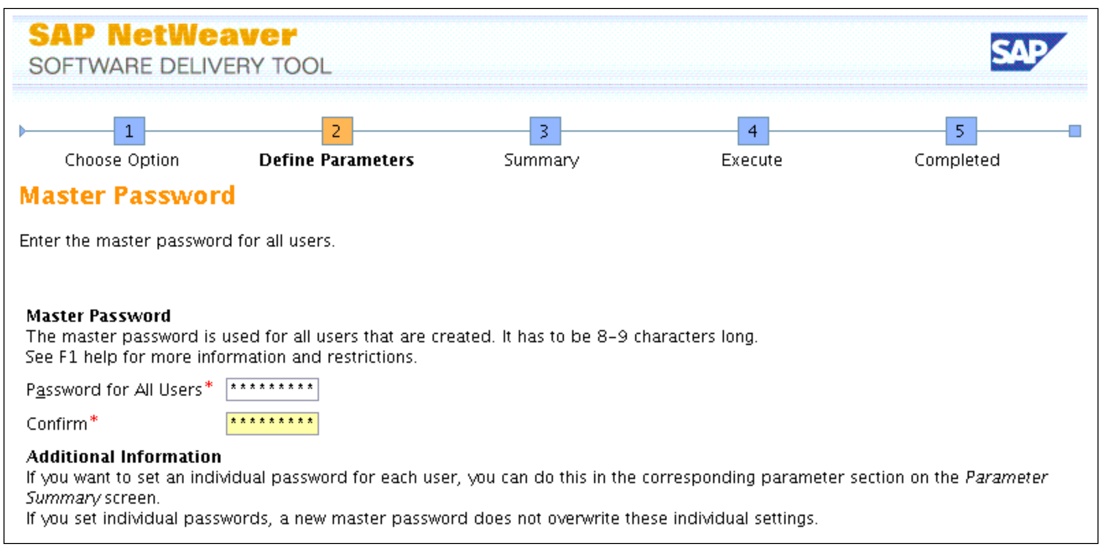



*Figure E-61 SAPinst liveCache installation screen capture 4: Master password*

5. If you did not create the users before, SAPinst prompts for the user ID and group ID. This user is the owner of the MaxDB software. These IDs must be the same on both nodes. See Figure E-62 on page 636.

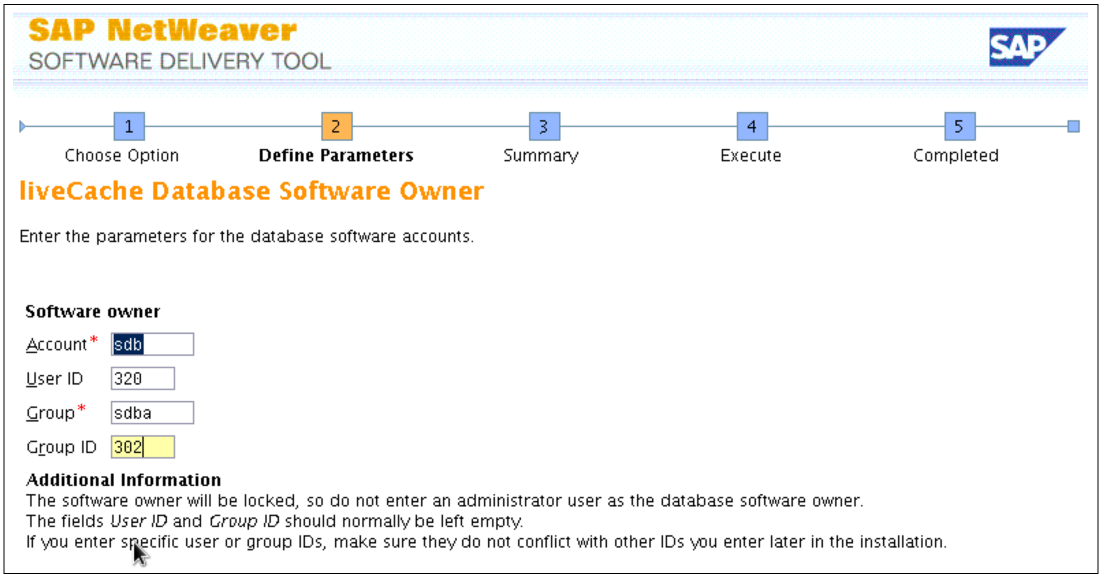



*Figure E-62 SAPinst liveCache installation screen capture 5: Database user ID and group ID*

6. This step creates a liveCache administration user at the operating systems level if it does not exist. The profiles of this user are then extended to include the program path that it can access with fundamental commands, such as **dbmcli** and **x_server**. Define the user ID and group ID, which are the same on both nodes. See Figure E-63.

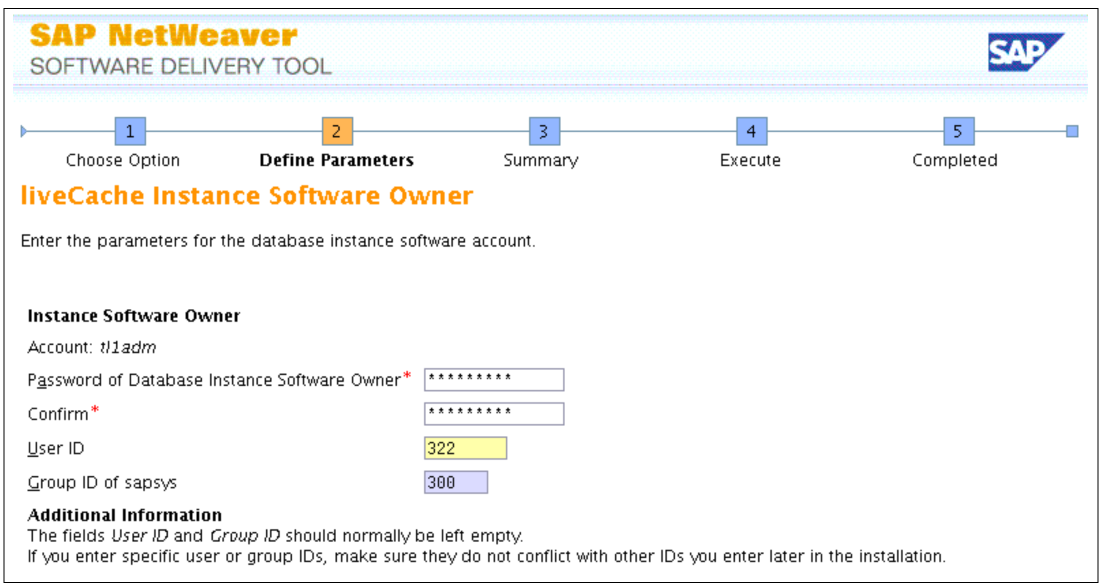



*Figure E-63 SAPinst liveCache installation screen capture 6: Password for <sid>adm user*

7. SAPinst asks for the location of the liveCache installation DVD as shown in Figure E-64 on page 637.

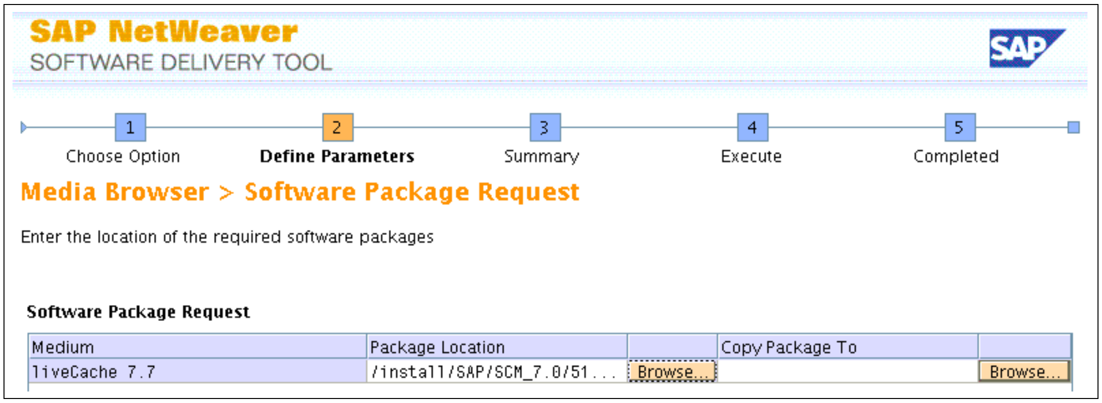



*Figure E-64 SAPinst liveCache installation screen capture 7: Location of liveCache installation DVD*

8. Set the passwords for the superdba and control user as shown in Figure E-65. If you do not change the passwords, SAPinst uses the master password.

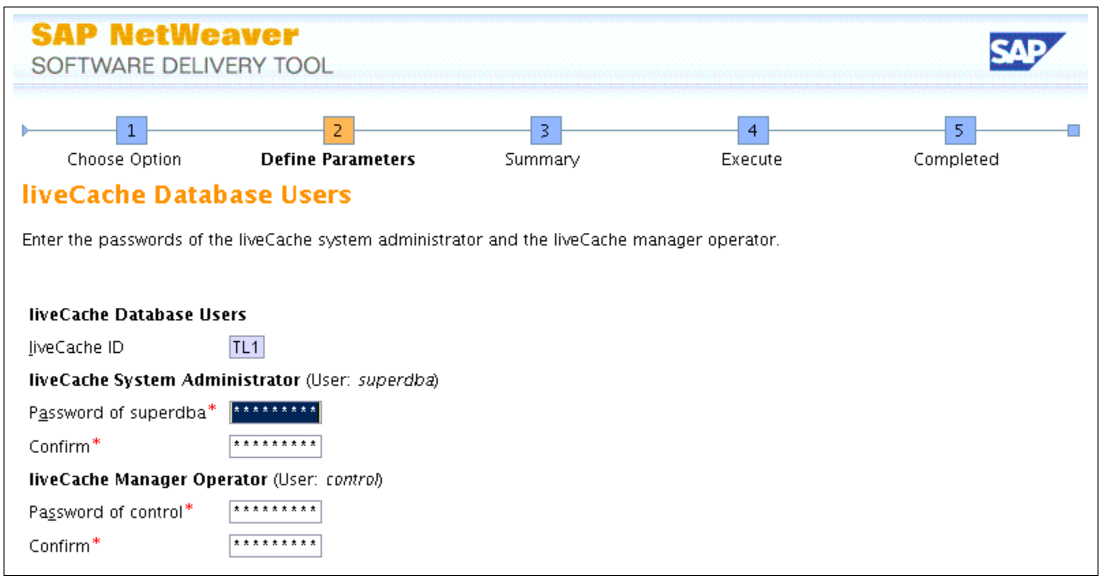



*Figure E-65 SAPinst liveCache installation screen capture 8: Password for superdba and control user*

9. Set the password for the liveCache access user. This user is the owner of the liveCache database schema. This user and password must be supplied to the APO administrator for integration with APO. APO connects to liveCache through two users. The DBM operator user performs liveCache administration, such as starting, stopping, reinitializing, and tracing. The sql user (this liveCache user) accesses the actual application data.

   If the password is not set to a different value, SAPinst uses the master password. See Figure E-66 on page 638.

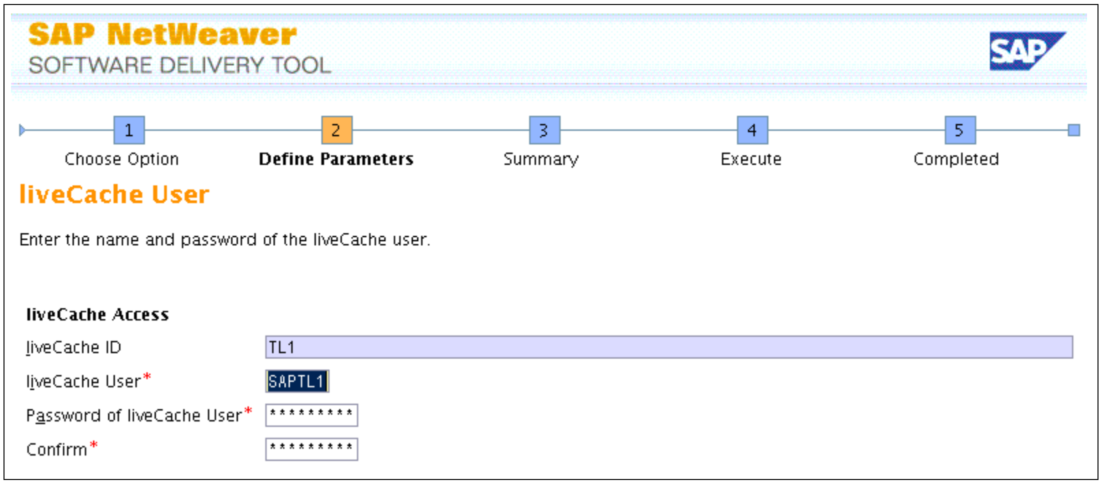



*Figure E-66 SAPinst liveCache installation screen capture 9: Password for liveCache access user*

10. In an SAP liveCache hot standby solution, we must select **Raw Devices** for the MaxDB/liveCache logging as shown in Figure E-67.

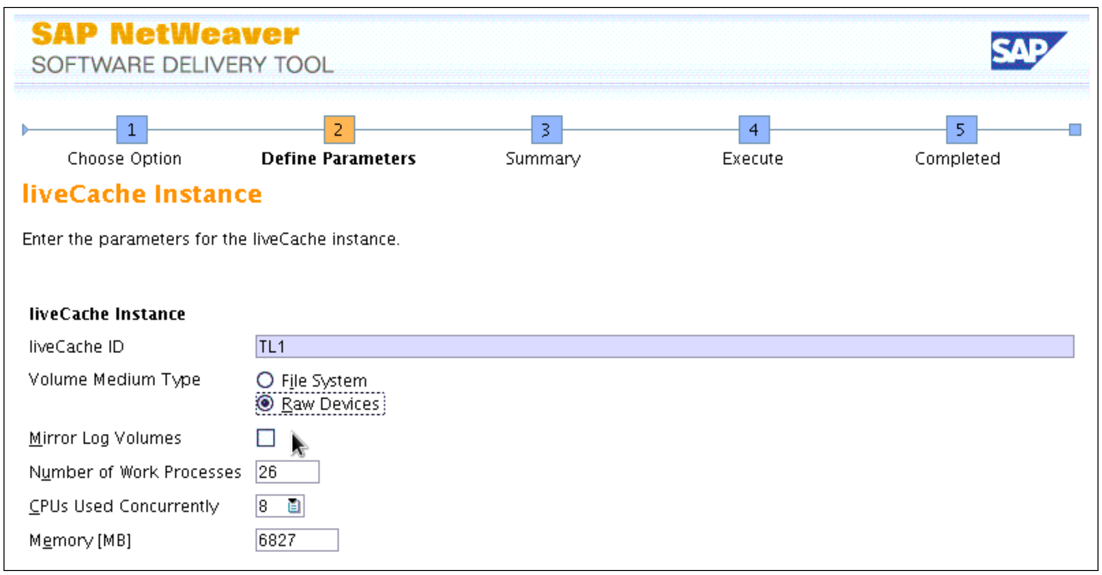



*Figure E-67 SAPinst liveCache installation screen capture 10: Use raw devices for liveCache instance*

11. Select the raw devices for liveCache logging as shown in Figure E-68 on page 639.

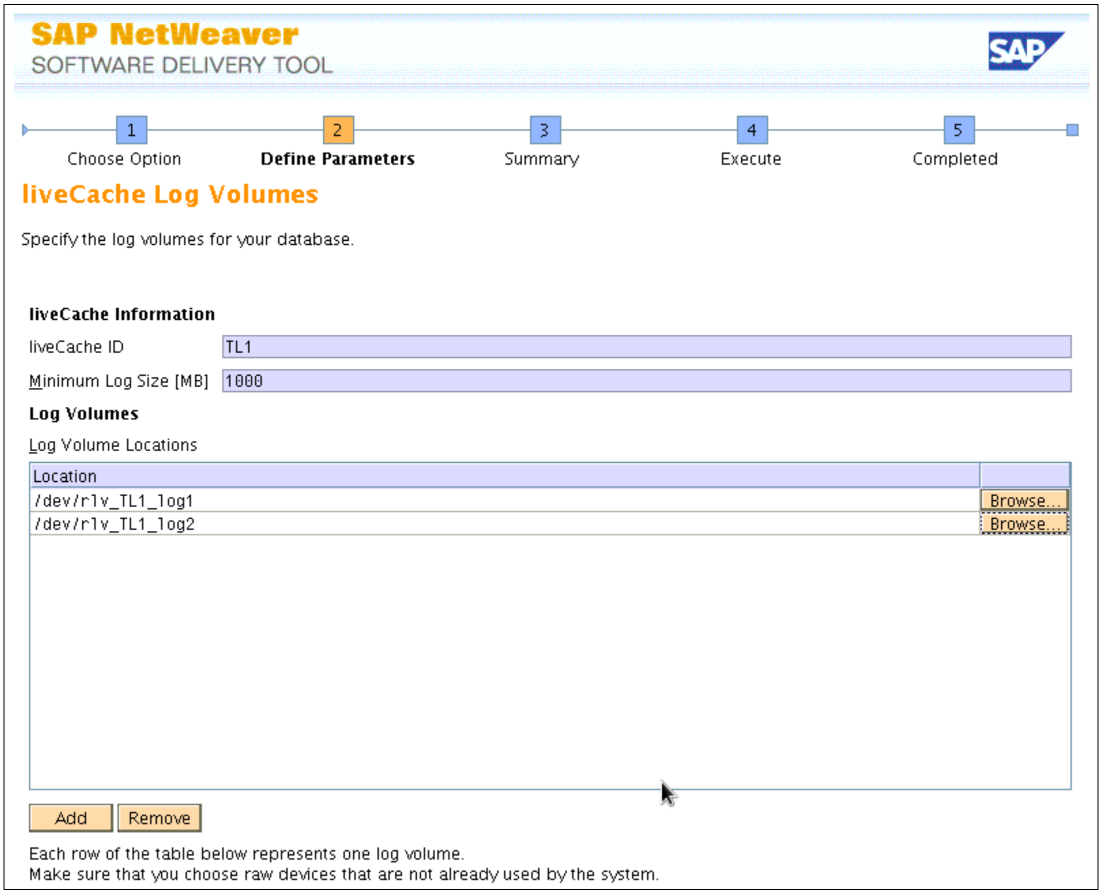



*Figure E-68 SAPinst liveCache installation screen capture 11: Raw devices for logging*

12. Select the raw devices for the liveCache database. In a production environment, more than one logical volume can exist. Remember that liveCache supports up to 256 logical volumes (for logging and data) and each logical volume can be up to (but not exceed) 128 GB in size. See Figure E-69 on page 640.

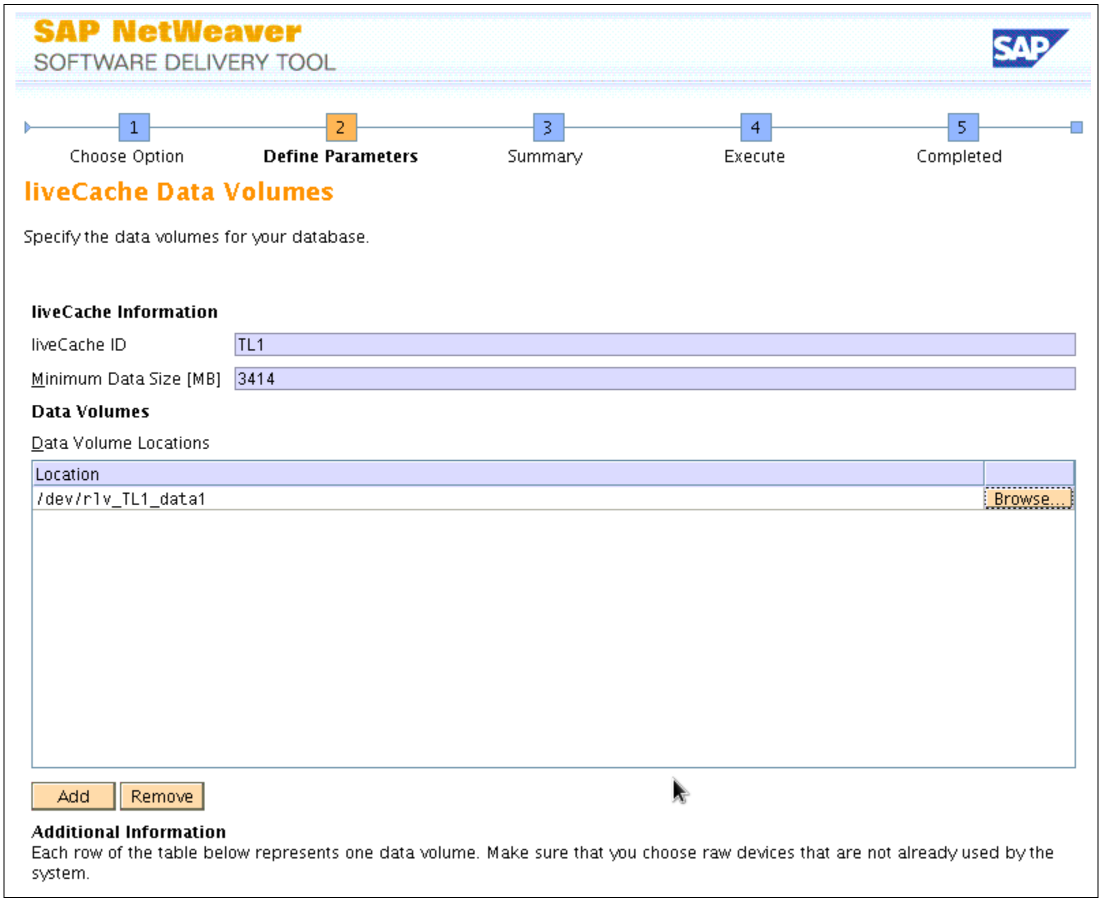



*Figure E-69 SAPinst liveCache installation screen capture 12: Raw devices for data volumes*

## sapdbhr_cluster: Installing DB2 high availability disaster recovery

The section describes the steps to install DB2 high availability disaster recovery (HADR). Follow these steps:

1. Install the SAP database instance on the first node of this cluster.

   We performed a normal SAP installation without the virtual SAP host name. Only the normal SAP parameters are set as shown in Example E-15.

   *Example E-15 SAPinst environment parameters*

   ```
   export TMPDIR=/tmp/TC1
   mkdir $TMPDIR
   export JAVA_HOME=/usr/java14_64
   ./sapinst
   ```

   You must perform the same installation steps as described in Step 6. Install the SAP database instance on the first node of this cluster in 7.6.2, “Installation and configuration steps before you use Smart Assist for SAP” on page 431.

   DB2 version 9.7 supports this script set. For the example in this appendix, we installed NW 7.0 with EHP 2 and DB2 9.7 SP4 in this test environment as shown in Example E-16 on page 641.

*Example E-16 DB2 version installed*

```
sapdbhr3:db2tc1 1> db2pd -v

Instance db2tc1 uses 64 bits and DB2 code release SQL09074
with level identifier 08050107
Informational tokens are DB2 v9.7.0.4, s110330, IP23236, Fix Pack 4.
```

2. Stop SAP and DB2.

   To set up the DB2 HADR feature, you must stop DB2 (Example E-17) and the SAP system.

*Example E-17 Stop DB2*

```
sapdbhr3:db2tc1 9> db2stop
04/06/2012 12:10:33     0   0   SQL1064N  DB2STOP processing was successful.
SQL1064N  DB2STOP processing was successful.
```

3. Copy the DB2 binaries:

```
root@sapdbhr4 / # cd /db2/TC1
root@sapdbhr4 / # ssh sapdbhr3 tar -cvf - -C /db2/TC1 ./db2tc1 | tar -xvf -
```

4. Copy additional DB2 files:

```
root@sapdbhr4 /db2/TC1/db2tc1 # scp -pr sapdbhr3:/var/db2 /var
```

5. Rename the environment files of the user `db2tc1`:

```
root@sapdbhr4 / # cd /db2/TC1/db2tc1
root@sapdbhr4 /db2/TC1/db2tc1 # mv .dbenv_sapdbhr3.csh .dbenv_sapdbhr4.csh
root@sapdbhr4 /db2/TC1/db2tc1 # mv .dbenv_sapdbhr3.sh .dbenv_sapdbhr4.sh
root@sapdbhr4 /db2/TC1/db2tc1 # mv .sapenv_sapdbhr3.csh .sapenv_sapdbhr4.csh
root@sapdbhr4 /db2/TC1/db2tc1 # mv .sapenv_sapdbhr3.sh .sapenv_sapdbhr4.sh
```

6. Modify the `db2nodes.cfg` in the local node as shown in Figure E-70:

```
vi /db2/TC1/db2tc1/sqllib/db2nodes.cfg
```

```
0 sapdbhr4 0
```

*Figure E-70 Edit /db2/TC1/db2tc1/sqllib/db2nodes.cfg*

7. Update the `/etc/services` file on the secondary node.

   When the instance is created in the primary node, the `/etc/services` file is updated with information for SAP use. You must add the lines that the SAPinst created on the installing node in the `/etc/services` file on the secondary node. There are two methods to make these entries available on the second node:

   – The first and the easiest way is to copy the complete file to the second node (with the same timestamp). With this method, you can use it only if both files are the same before the SAPinst starts, and after the SAPinst, only the file on `sapsma1` has additional files:

```
root@sapdbhr3 / # clrcp /etc/services sapdbhr4:/etc/services
```

   This command is better:

```
root@sapdbhr3 / # scp -p /etc/services sapdbhr4:/etc/services
```

– The second method is to add every necessary line on the second node with an AIX command, for example:

```
root@sapdbhr4 / # chservices -a -v sapgw99s -p tcp -n 4899 -u "SAP System
Gateway Security Port"
```

8. Create the DB2 backup on the first node:

```
su - db2tc1
db2start
db2 backup db TC1 to /install/backup compress
```

9. Restore it on the second node:

```
su - db2tc1
db2start
db2 uncatalog db TC1
db2 restore db TC1 from /install/backup
```

10. Configure the DB2 HADR on the secondary node `sapdbhr4`:

```
su - db2tc1
db2 update db cfg for TC1 using HADR_LOCAL_HOST sapdbhr4
db2 update db cfg for TC1 using HADR_LOCAL_SVC 61000
db2 update db cfg for TC1 using HADR_REMOTE_HOST sapdbhr3
db2 update db cfg for TC1 using HADR_REMOTE_SVC 61000
db2 update db cfg for TC1 using HADR_REMOTE_INST db2tc1
db2 update db cfg for TC1 using HADR_SYNCMODE NEARSYNC
db2 update db cfg for TC1 using HADR_TIMEOUT 120
db2 update db cfg for TC1 using INDEXREC RESTART logindexbuild ON
db2 update db cfg for TC1 using HADR_PEER_WINDOW 300
```

11. Configure the DB2 HADR configuration on the primary node `sapdbhr3`:

```
su - db2tc1
db2 update db cfg for TC1 using HADR_LOCAL_HOST sapdbhr3
db2 update db cfg for TC1 using HADR_LOCAL_SVC 61000
db2 update db cfg for TC1 using HADR_REMOTE_HOST sapdbhr4
db2 update db cfg for TC1 using HADR_REMOTE_SVC 61000
db2 update db cfg for TC1 using HADR_REMOTE_INST db2tc1
db2 update db cfg for TC1 using HADR_SYNCMODE NEARSYNC
db2 update db cfg for TC1 using HADR_TIMEOUT 120
db2 update db cfg for TC1 using INDEXREC RESTART logindexbuild ON
db2 update db cfg for TC1 using HADR_PEER_WINDOW 300
```

12. Start DB2 HADR on the secondary node `sapdbhr4`:

```
su – db2tc1
db2 deactivate DB TC1
db2 START HADR ON DB TC1 AS STANDBY
```

13. Start DB2 HADR on the primary node `sapdbhr3`:

```
su – db2tc1
db2 deactivate DB TC1
db2 START HADR ON DB TC1 AS PRIMARY
```

14. Check whether the database works in HADR mode as shown in Example E-18.

*Example E-18 db2 GET SNAPSHOT FOR DB ON TC1*

```
sapdbhr3:db2tc1 15> db2 GET SNAPSHOT FOR DB ON TC1

              Database Snapshot
```

```
Database name                              = TC1
Database path                              = /db2/TC1/db2tc1/NODE0000/SQL00001/
Input database alias                       = TC1
Database status                            = Active
Catalog database partition number          = 0
Catalog network node name                  =
Operating system running at database server= AIX 64BIT
Location of the database                   = Local
First database connect timestamp           = 04/18/2012 14:56:48.079258
Last reset timestamp                       =
Last backup timestamp                      = 04/18/2012 14:20:54.000000
Snapshot timestamp                         = 04/18/2012 14:58:21.599595

...

HADR Status
  Role                   = Primary
  State                  = Peer
  Synchronization mode   = Nearsync
  Connection status      = Connected, 04/18/2012 14:56:49.840832
  Peer window end        = 04/18/2012 15:03:20.000000 (1334775800)
  Peer window (seconds)  = 300
  Heartbeats missed      = 0
  Local host             = sapdbhr3
  Local service          = 61000
  Remote host            = sapdbhr4
  Remote service         = 61000
  Remote instance        = db2tc1
  timeout(seconds)       = 120
  Primary log position(file, page, LSN) = S0000000.LOG, 0, 000000019E58C010
  Standby log position(file, page, LSN) = S0000000.LOG, 0, 000000019E58C010
  Log gap running average(bytes) = 0

...
```

15. Example E-19 shows `sapdbhr3` as the primary.

*Example E-19 db2pd -db TC1 -hadr on sapdbhr3 as primary*

```
sapdbhr3:db2tc1 16> db2pd -db TC1 -hadr

Database Partition 0 -- Database TC1 -- Active -- Up 0 days 00:05:26 -- Date
04/18/2012 15:02:14

HADR Information:
Role    State                SyncMode HeartBeatsMissed   LogGapRunAvg (bytes)
Primary Peer                 Nearsync     0                  0

ConnectStatus ConnectTime                           Timeout
Connected     Wed Apr 18 14:56:49 2012 (1334775409) 120

PeerWindowEnd                         PeerWindow
Wed Apr 18 15:06:50 2012 (1334776010) 300

LocalHost                                LocalService
```

```
sapdbhr3                                 61000

RemoteHost                               RemoteService      RemoteInstance
sapdbhr4                                 61000              db2tc1

PrimaryFile  PrimaryPg  PrimaryLSN
S0000000.LOG 0          0x000000019E58C010

StandByFile  StandByPg  StandByLSN
S0000000.LOG 0          0x000000019E58C010
```

16. Try a manual switch on `sapdbhr4` and check as shown in Example E-20.

```
su - db2tc1
db2 takeover HADR ON DB TC1
```

*Example E-20 db2pd -db TC1 -HADR on sapdbhr4 as primary*

```
apdbhr4:db2tc1 24> db2pd -db TC1 -hadr

Database Partition 0 -- Database TC1 -- Active -- Up 0 days 00:15:04 -- Date
04/18/2012 15:10:32

HADR Information:
Role    State                SyncMode HeartBeatsMissed   LogGapRunAvg (bytes)
Primary Peer                Nearsync     0                  0

ConnectStatus ConnectTime                           Timeout
Connected     Wed Apr 18 14:56:49 2012 (1334775409) 120

PeerWindowEnd                         PeerWindow
Wed Apr 18 15:15:15 2012 (1334776515) 300

LocalHost                                LocalService
sapdbhr4                                 61000

RemoteHost                               RemoteService      RemoteInstance
sapdbhr3                                 61000              db2tc1

PrimaryFile  PrimaryPg  PrimaryLSN
S0000000.LOG 0          0x000000019E58C010

StandByFile  StandByPg  StandByLSN
S0000000.LOG 0          0x000000019E58C010
```

17. On the second node, you can see its status as shown in Example E-21.

*Example E-21 db2pd -db TC1 -HADR on sapdbhr3 as standby*

```
sapdbhr3:db2tc1 17> db2pd -db TC1 -HADR

Database Partition 0 -- Database TC1 -- Standby -- Up 0 days 00:20:26 -- Date
04/18/2012 15:17:14

HADR Information:
Role    State                SyncMode HeartBeatsMissed   LogGapRunAvg (bytes)
Standby Peer                Nearsync     0                  0
```

```
ConnectStatus ConnectTime                              Timeout
Connected     Wed Apr 18 14:56:49 2012 (1334775409) 120

PeerWindowEnd                          PeerWindow
Wed Apr 18 15:21:45 2012 (1334776905) 300

LocalHost                                  LocalService
sapdbhr3                                   61000

RemoteHost                                 RemoteService      RemoteInstance
sapdbhr4                                   61000              db2tc1

PrimaryFile  PrimaryPg  PrimaryLSN
S0000000.LOG 0          0x000000019E58C010

StandByFile  StandByPg  StandByLSN         StandByRcvBufUsed
S0000000.LOG 0          0x000000019E58C010 0%
```

18. Switch back from `sapdbhr4` to `sapdbhr3`:

```
sapdbhr3 > su - db2tc1
db2 takeover HADR ON DB TC1
```

# Cluster scripts for DB2 HADR

We list the scripts that we use in 7.8, “DB2 HADR cluster solution” on page 463 for the DB2 HADR solution:

- Readme.txt
- License.txt
- sapha_env
- sapha_TL1_cfg
- cl_db2_start_local
- cl_db2_stop_local
- cl_db2_start_hadr
- cl_db2_stop_hadr
- cl_db2_monitor_hadr
- lib/DButil.db2
- lib/log
- lib/SAPutil
- lib/util

You can get the latest version of these scripts by sending a note to this email address:

`isicc@de.ibm.com`

## Readme.txt

Example E-22 shows the `readme.txt` file.

*Example E-22 readme.txt*

```
Disclaimers: We do not guarantee anything about these scripts;
```

```
you are officially "yoyo" (You'reOnYourOwn). We do encourage/welcome
your thoughts, feedback and contributions.

Author:
   Walter Orb
   IBM/SAP International Competence Center
   Altrottstr. 31
   69190 Walldorf
   Germany
   e-mail: walter.orb@de.ibm.com

   Katharina Probst
   IBM Labor Boeblingen
   Schoenaicherstrasse 220
   71032 Boeblingen
   Germany
   e-mail: probs@de.ibm.com

Legal Stuff:
LICENSE: LICENSE.txt

No Warranty
Subject to any statutory warranties which can not be excluded, IBM makes no
warranties or conditions either express or implied, including without limitation,
the warranty of non-infringement and the implied warranties of merchantability and
fitness for a particular purpose, regarding the program or technical support, if
any.

Limitation of Liability
Neither IBM nor its suppliers will be liable for any direct or indirect damages,
including without limitation, lost profits, lost savings, or any incidental,
special, or other economic consequential damages, even if IBM is informed of their
possibility. Some jurisdictions do not allow the exclusion or limitation of
incidental or consequential damages, so the above exclusion or limitation may not
apply to you.
```

## License.txt

Example E-23 shows the `license.txt` file.

*Example E-23 License.txt*

```
International License Agreement for Non-Warranted Programs

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR
OTHERWISE USING THE PROGRAM, LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF
YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT AND WARRANT
THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE
TO THESE TERMS,

* DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN "ACCEPT" BUTTON, OR USE THE
PROGRAM; AND
```

* PROMPTLY RETURN THE UNUSED MEDIA AND DOCUMENTATION TO THE PARTY FROM WHOM IT WAS OBTAINED FOR A REFUND OF THE AMOUNT PAID. IF THE PROGRAM WAS DOWNLOADED, DESTROY ALL COPIES OF THE PROGRAM.

1. Definitions

"Authorized Use" - the specified level at which Licensee is authorized to execute or run the Program. That level may be measured by number of users, millions of service units ("MSUs"), Processor Value Units ("PVUs"), or other level of use specified by IBM.

"IBM" - International Business Machines Corporation or one of its subsidiaries.

"License Information" ("LI") - a document that provides information and any additional terms specific to a Program. The Program's LI is available at www.ibm.com/software/sla. The LI can also be found in the Program's directory, by the use of a system command, or as a booklet included with the Program.

"Program" - the following, including the original and all whole or partial copies: 1) machine-readable instructions and data, 2) components, files, and modules, 3) audio-visual content (such as images, text, recordings, or pictures), and 4) related licensed materials (such as keys and documentation).

2. Agreement Structure

This Agreement includes Part 1 - General Terms, Part 2 - Country-unique Terms (if any) and the LI and is the complete agreement between Licensee and IBM regarding the use of the Program. It replaces any prior oral or written communications between Licensee and IBM concerning Licensee's use of the Program. The terms of Part 2 may replace or modify those of Part 1. To the extent of any conflict, the LI prevails over both Parts.

3. License Grant

The Program is owned by IBM or an IBM supplier, and is copyrighted and licensed, not sold.

IBM grants Licensee a nonexclusive license to 1) use the Program up to the Authorized Use specified in the invoice, 2) make and install copies to support such Authorized Use, and 3) make a backup copy, all provided that

a. Licensee has lawfully obtained the Program and complies with the terms of this Agreement;

b. the backup copy does not execute unless the backed-up Program cannot execute;

c. Licensee reproduces all copyright notices and other legends of ownership on each copy, or partial copy, of the Program;

d. Licensee ensures that anyone who uses the Program (accessed either locally or remotely) 1) does so only on Licensee's behalf and 2) complies with the terms of this Agreement;

e. Licensee does not 1) use, copy, modify, or distribute the Program except as expressly permitted in this Agreement; 2) reverse assemble, reverse compile, otherwise translate, or reverse engineer the Program, except as expressly permitted by law without the possibility of contractual waiver; 3) use any of the Program's components, files, modules, audio-visual content, or related licensed materials separately from that Program; or 4) sublicense, rent, or lease the Program; and

f. if Licensee obtains this Program as a Supporting Program, Licensee uses this Program only to support the Principal Program and subject to any limitations in the license to the Principal Program, or, if Licensee obtains this Program as a Principal Program, Licensee uses all Supporting Programs only to support this Program, and subject to any limitations in this Agreement. For purposes of this Item "f," a "Supporting Program" is a Program that is part of another IBM Program ("Principal Program") and identified as a Supporting Program in the Principal Program's LI. (To obtain a separate license to a Supporting Program without these restrictions, Licensee should contact the party from whom Licensee obtained the Supporting Program.)

This license applies to each copy of the Program that Licensee makes.

3.1 Trade-ups, Updates, Fixes, and Patches

3.1.1 Trade-ups

If the Program is replaced by a trade-up Program, the replaced Program's license is promptly terminated.

3.1.2 Updates, Fixes, and Patches

When Licensee obtains an update, fix, or patch to a Program, Licensee accepts any additional or different terms that are applicable to such update, fix, or patch that are specified in its LI. If no additional or different terms are provided, then the update, fix, or patch is subject solely to this Agreement. If the Program is replaced by an update, Licensee agrees to promptly discontinue use of the replaced Program.

3.2 Fixed Term Licenses

If IBM licenses the Program for a fixed term, Licensee's license is terminated at the end of the fixed term, unless Licensee and IBM agree to renew it.

3.3 Term and Termination

This Agreement is effective until terminated.

IBM may terminate Licensee's license if Licensee fails to comply with the terms of this Agreement.

If the license is terminated for any reason by either party, Licensee agrees to promptly discontinue use of and destroy all of Licensee's copies of the Program. Any terms of this Agreement that by their nature extend beyond termination of this Agreement remain in effect until fulfilled, and apply to both parties' respective successors and assignees.

4. Charges

Charges, if any, are based on Authorized Use obtained, which is specified in the invoice. IBM does not give credits or refunds for charges already due or paid, except as specified elsewhere in this Agreement.

If Licensee wishes to increase its Authorized Use, Licensee must notify IBM or an authorized IBM reseller in advance and pay any applicable charges.

5. Taxes

If any authority imposes on the Program a duty, tax, levy, or fee, excluding those based on IBM's net income, then Licensee agrees to pay that amount, as specified in an invoice, or supply exemption documentation. Licensee is responsible for any personal property taxes for the Program from the date that Licensee obtains it. If any authority imposes a customs duty, tax, levy, or fee for the import into or the export, transfer, access, or use of the Program outside the country in which the original Licensee was granted the license, then Licensee agrees that it is responsible for, and will pay, any amount imposed.

6. Money-back Guarantee

If Licensee is dissatisfied with the Program for any reason and is the original Licensee, Licensee may terminate the license and obtain a refund of the amount Licensee paid, if any, for the Program, provided that Licensee returns the Program to the party from whom Licensee obtained it within 30 days of the invoice date. If the license is for a fixed term that is subject to renewal, then Licensee may obtain a refund only if the Program is returned within the first 30 days of the initial term. If Licensee downloaded the Program, Licensee should contact the party from whom Licensee obtained it for instructions on how to obtain the refund.

7. Program Transfer

Licensee may transfer the Program and all of Licensee's license rights and obligations to another party only if that party agrees to the terms of this Agreement. If the license is terminated for any reason by either party, Licensee is prohibited from transferring the Program to another party. Licensee may not transfer a portion of 1) the Program or 2) the Program's Authorized Use. When Licensee transfers the Program, Licensee must also transfer a hard copy of this Agreement, including the LI. Immediately after the transfer, Licensee's license terminates.

8. No Warranties

SUBJECT TO ANY STATUTORY WARRANTIES THAT CANNOT BE EXCLUDED, IBM MAKES NO WARRANTIES OR CONDITIONS, EXPRESS OR IMPLIED, REGARDING THE PROGRAM, INCLUDING, BUT NOT LIMITED TO, ANY IMPLIED WARRANTIES OR CONDITIONS OF MERCHANTABILITY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, AND TITLE, AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

SOME STATES OR JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF EXPRESS OR IMPLIED WARRANTIES, SO THE ABOVE EXCLUSION MAY NOT APPLY TO LICENSEE. IN THAT EVENT, SUCH WARRANTIES ARE LIMITED IN DURATION TO THE MINIMUM PERIOD REQUIRED BY LAW. NO WARRANTIES APPLY AFTER THAT PERIOD. SOME STATES OR JURISDICTIONS DO NOT ALLOW LIMITATIONS ON HOW LONG AN IMPLIED WARRANTY LASTS, SO THE ABOVE LIMITATION MAY NOT APPLY TO LICENSEE. LICENSEE MAY HAVE OTHER RIGHTS THAT VARY FROM STATE TO STATE OR JURISDICTION TO JURISDICTION.

THE DISCLAIMERS AND EXCLUSIONS IN THIS SECTION 8 ALSO APPLY TO ANY OF IBM'S PROGRAM DEVELOPERS AND SUPPLIERS.

MANUFACTURERS, SUPPLIERS, OR PUBLISHERS OF NON-IBM PROGRAMS MAY PROVIDE THEIR OWN WARRANTIES.

IBM DOES NOT PROVIDE SUPPORT OF ANY KIND, UNLESS IBM SPECIFIES OTHERWISE. IN SUCH EVENT, ANY SUPPORT PROVIDED BY IBM IS SUBJECT TO THE DISCLAIMERS AND EXCLUSIONS IN THIS SECTION 8.

9. Licensee Data and Databases

To assist Licensee in isolating the cause of a problem with the Program, IBM may request that Licensee 1) allow IBM to remotely access Licensee's system or 2) send Licensee information or system data to IBM. However, IBM is not obligated to provide such assistance unless IBM and Licensee enter a separate written agreement under which IBM agrees to provide to Licensee that type of support, which is beyond IBM's obligations in this Agreement. In any event, IBM uses information about errors and problems to improve its products and services, and assist with its provision of related support offerings. For these purposes, IBM may use IBM entities and subcontractors (including in one or more countries other than the one in which Licensee is located), and Licensee authorizes IBM to do so.

Licensee remains responsible for 1) any data and the content of any database Licensee makes available to IBM, 2) the selection and implementation of procedures and controls regarding access, security, encryption, use, and transmission of data (including any personally-identifiable data), and 3) backup and recovery of any database and any stored data. Licensee will not send or provide IBM access to any personally-identifiable information, whether in data or any other form, and will be responsible for reasonable costs and other amounts that IBM may incur relating to any such information mistakenly provided to IBM or the loss or disclosure of such information by IBM, including those arising out of any third party claims.

10. Limitation of Liability

The limitations and exclusions in this Section 10 (Limitation of Liability) apply to the full extent they are not prohibited by applicable law without the possibility of contractual waiver.

10.1 Items for Which IBM May Be Liable

Circumstances may arise where, because of a default on IBM's part or other liability, Licensee is entitled to recover damages from IBM. Regardless of the basis on which Licensee is entitled to claim damages from IBM (including fundamental breach, negligence, misrepresentation, or other contract or tort claim), IBM's entire liability for all claims in the aggregate arising from or related to each Program or otherwise arising under this Agreement will not exceed the amount of any 1) damages for bodily injury (including death) and damage to real property and tangible personal property and 2) other actual direct damages up to the charges (if the Program is subject to fixed term charges, up to twelve months' charges) Licensee paid for the Program that is the subject of the claim.

This limit also applies to any of IBM's Program developers and suppliers. It is the maximum for which IBM and its Program developers and suppliers are collectively responsible.

10.2 Items for Which IBM Is Not Liable

UNDER NO CIRCUMSTANCES IS IBM, ITS PROGRAM DEVELOPERS OR SUPPLIERS LIABLE FOR ANY OF THE FOLLOWING, EVEN IF INFORMED OF THEIR POSSIBILITY:

a. LOSS OF, OR DAMAGE TO, DATA;

b. SPECIAL, INCIDENTAL, EXEMPLARY, OR INDIRECT DAMAGES, OR FOR ANY ECONOMIC CONSEQUENTIAL DAMAGES; OR

c. LOST PROFITS, BUSINESS, REVENUE, GOODWILL, OR ANTICIPATED SAVINGS.

11. Compliance Verification

For purposes of this Section 11 (Compliance Verification), "ILAN Program Terms" means 1) this Agreement and applicable amendments and transaction documents provided by IBM, and 2) IBM software policies that may be found at the IBM Software Policy website (www.ibm.com/softwarepolicies), including but not limited to those policies concerning backup, sub-capacity pricing, and migration.

The rights and obligations set forth in this Section 11 remain in effect during the period the Program is licensed to Licensee, and for two years thereafter.

11.1 Verification Process

Licensee agrees to create, retain, and provide to IBM and its auditors accurate written records, system tool outputs, and other system information sufficient to provide auditable verification that Licensee's use of all Programs is in compliance with the ILAN Program Terms, including, without limitation, all of IBM's applicable licensing and pricing qualification terms. Licensee is responsible for 1) ensuring that it does not exceed its Authorized Use, and 2) remaining in compliance with ILAN Program Terms.

Upon reasonable notice, IBM may verify Licensee's compliance with ILAN Program Terms at all sites and for all environments in which Licensee uses (for any purpose) Programs subject to ILAN Program Terms. Such verification will be conducted in a manner that minimizes disruption to Licensee's business, and may be conducted on Licensee's premises, during normal business hours. IBM may use an independent auditor to assist with such verification, provided IBM has a written confidentiality agreement in place with such auditor.

11.2 Resolution

IBM will notify Licensee in writing if any such verification indicates that Licensee has used any Program in excess of its Authorized Use or is otherwise not in compliance with the ILAN Program Terms. Licensee agrees to promptly pay directly to IBM the charges that IBM specifies in an invoice for 1) any such excess use, 2) support for such excess use for the lesser of the duration of such excess use or two years, and 3) any additional charges and other liabilities determined as a result of such verification.

12. Third Party Notices

The Program may include third party code that IBM, not the third party, licenses to Licensee under this Agreement. Notices, if any, for the third party code ("Third Party Notices") are included for Licensee's information only. These notices can be found in the Program's NOTICES file(s). Information on how to obtain source code for certain third party code can be found in the Third Party Notices. If in the Third Party Notices IBM identifies third party code as "Modifiable Third Party Code," IBM authorizes Licensee to 1) modify the Modifiable Third Party Code and 2) reverse engineer the Program modules that directly interface with the Modifiable Third Party Code provided that it is only for the purpose of debugging Licensee's modifications to such third party code. IBM's service and support obligations, if any, apply only to the unmodified Program.

13. General

a. Nothing in this Agreement affects any statutory rights of consumers that cannot be waived or limited by contract.

b. For Programs IBM provides to Licensee in tangible form, IBM fulfills its shipping and delivery obligations upon the delivery of such Programs to the IBM-designated carrier, unless otherwise agreed to in writing by Licensee and IBM.

c. If any provision of this Agreement is held to be invalid or unenforceable, the remaining provisions of this Agreement remain in full force and effect.

d. Licensee agrees to comply with all applicable export and import laws and regulations, including U.S. embargo and sanctions regulations and prohibitions on export for certain end uses or to certain users.

e. Licensee authorizes International Business Machines Corporation and its subsidiaries (and their successors and assigns, contractors and IBM Business Partners) to store and use Licensee's business contact information wherever they do business, in connection with IBM products and services, or in furtherance of IBM's business relationship with Licensee.

f. Each party will allow the other reasonable opportunity to comply before it claims that the other has not met its obligations under this Agreement. The parties will attempt in good faith to resolve all disputes, disagreements, or claims between the parties relating to this Agreement.

g. Unless otherwise required by applicable law without the possibility of contractual waiver or limitation: 1) neither party will bring a legal action, regardless of form, for any claim arising out of or related to this Agreement more than two years after the cause of action arose; and 2) upon the expiration of such time limit, any such claim and all respective rights related to the claim lapse.

h. Neither Licensee nor IBM is responsible for failure to fulfill any obligations due to causes beyond its control.

i. No right or cause of action for any third party is created by this Agreement, nor is IBM responsible for any third party claims against Licensee, except as permitted in Subsection 10.1 (Items for Which IBM May Be Liable) above for bodily injury (including death) or damage to real or tangible personal property for which IBM is legally liable to that third party.

j. In entering into this Agreement, neither party is relying on any representation not specified in this Agreement, including but not limited to any representation concerning: 1) the performance or function of the Program; 2) the experiences or recommendations of other parties; or 3) any results or savings that Licensee may achieve.

k. IBM has signed agreements with certain organizations (called "IBM Business Partners") to promote, market, and support certain Programs. IBM Business Partners remain independent and separate from IBM. IBM is not responsible for the actions or statements of IBM Business Partners or obligations they have to Licensee.

l. The license and intellectual property indemnification terms of Licensee's other agreements with IBM (such as the IBM Customer Agreement) do not apply to Program licenses granted under this Agreement.

m. Both parties agree that all information exchanged is nonconfidential. If either party requires the exchange of confidential information, it will be made under a signed confidentiality agreement;

14. Geographic Scope and Governing Law

14.1 Governing Law

Both parties agree to the application of the laws of the country in which Licensee obtained the Program license to govern, interpret, and enforce all of Licensee's and IBM's respective rights, duties, and obligations arising from, or relating in any manner to, the subject matter of this Agreement, without regard to conflict of law principles.

The United Nations Convention on Contracts for the International Sale of Goods does not apply.

14.2 Jurisdiction

All rights, duties, and obligations are subject to the courts of the country in which Licensee obtained the Program license.

Part 2 - Country-unique Terms

For licenses granted in the countries specified below, the following terms replace or modify the referenced terms in Part 1. All terms in Part 1 that are not changed by these amendments remain unchanged and in effect. This Part 2 is organized as follows:

* Multiple country amendments to Part 1, Section 14 (Governing Law and Jurisdiction);

* Americas country amendments to other Agreement terms;

* Asia Pacific country amendments to other Agreement terms; and

* Europe, Middle East, and Africa country amendments to other Agreement terms.

Multiple country amendments to Part 1, Section 14 (Governing Law and Jurisdiction)

14.1 Governing Law

The phrase "the laws of the country in which Licensee obtained the Program license" in the first paragraph of 14.1 Governing Law is replaced by the following phrases in the countries below:

AMERICAS

(1) In Canada: the laws in the Province of Ontario;

(2) in Mexico: the federal laws of the Republic of Mexico;

(3) in the United States, Anguilla, Antigua/Barbuda, Aruba, British Virgin Islands, Cayman Islands, Dominica, Grenada, Guyana, Saint Kitts and Nevis, Saint Lucia, Saint Maarten, and Saint Vincent and the Grenadines: the laws of the State of New York, United States;

(4) in Venezuela: the laws of the Bolivarian Republic of Venezuela;

ASIA PACIFIC

(5) in Cambodia and Laos: the laws of the State of New York, United States;

(6) in Australia: the laws of the State or Territory in which the transaction is performed;

(7) in Hong Kong SAR and Macau SAR: the laws of Hong Kong Special Administrative Region ("SAR");

(8) in Taiwan: the laws of Taiwan;

EUROPE, MIDDLE EAST, AND AFRICA

(9) in Albania, Armenia, Azerbaijan, Belarus, Bosnia-Herzegovina, Bulgaria, Croatia, Former Yugoslav Republic of Macedonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Moldova, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan: the laws of Austria;

(10) in Algeria, Andorra, Benin, Burkina Faso, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo Republic, Djibouti, Democratic Republic of Congo, Equatorial Guinea, French Guiana, French Polynesia, Gabon, Gambia, Guinea, Guinea-Bissau, Ivory Coast, Lebanon, Madagascar, Mali, Mauritania, Mauritius, Mayotte, Morocco, New Caledonia, Niger, Reunion, Senegal, Seychelles, Togo, Tunisia, Vanuatu, and Wallis and Futuna: the laws of France;

(11) in Estonia, Latvia, and Lithuania: the laws of Finland;

(12) in Angola, Bahrain, Botswana, Burundi, Egypt, Eritrea, Ethiopia, Ghana, Jordan, Kenya, Kuwait, Liberia, Malawi, Malta, Mozambique, Nigeria, Oman, Pakistan, Qatar, Rwanda, Sao Tome and Principe, Saudi Arabia, Sierra Leone, Somalia, Tanzania, Uganda, United Arab Emirates, the United Kingdom, West Bank/Gaza, Yemen, Zambia, and Zimbabwe: the laws of England; and

(13) in South Africa, Namibia, Lesotho, and Swaziland: the laws of the Republic of South Africa.

14.2 Jurisdiction

The following paragraph pertains to jurisdiction and replaces Subsection 14.2 (Jurisdiction) as it applies for those countries identified below:

All rights, duties, and obligations are subject to the courts of the country in which Licensee obtained the Program license except that in the countries identified below all disputes arising out of or related to this Agreement, including summary proceedings, will be brought before and subject to the exclusive jurisdiction of the following courts of competent jurisdiction:

AMERICAS

(1) In Argentina: the Ordinary Commercial Court of the city of Buenos Aires,

(2) in Brazil: the court of Rio de Janeiro, RJ;

(3) in Chile: the Civil Courts of Justice of Santiago;

(4) in Ecuador: the civil judges of Quito for executory or summary proceedings (as applicable);

(5) in Mexico: the courts located in Mexico City, Federal District;

(6) in Peru: the judges and tribunals of the judicial district of Lima, Cercado;

(7) in Uruguay: the courts of the city of Montevideo;

(8) in Venezuela: the courts of the metropolitan area of the city of Caracas;

EUROPE, MIDDLE EAST, AND AFRICA

(9) in Austria: the court of law in Vienna, Austria (Inner-City);

(10) in Algeria, Andorra, Benin, Burkina Faso, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo Republic, Djibouti, Democratic Republic of Congo, Equatorial Guinea, France, French Guiana, French Polynesia, Gabon, Gambia, Guinea, Guinea-Bissau, Ivory Coast, Lebanon, Madagascar, Mali, Mauritania, Mauritius, Mayotte, Monaco, Morocco, New Caledonia, Niger, Reunion, Senegal, Seychelles, Togo, Tunisia, Vanuatu, and Wallis and Futuna: the Commercial Court of Paris;

(11) in Angola, Bahrain, Botswana, Burundi, Egypt, Eritrea, Ethiopia, Ghana, Jordan, Kenya, Kuwait, Liberia, Malawi, Malta, Mozambique, Nigeria, Oman, Pakistan, Qatar, Rwanda, Sao Tome and Principe, Saudi Arabia, Sierra Leone, Somalia, Tanzania, Uganda, United Arab Emirates, the United Kingdom, West Bank/Gaza, Yemen, Zambia, and Zimbabwe: the English courts;

(12) in South Africa, Namibia, Lesotho, and Swaziland: the High Court in Johannesburg;

(13) in Greece: the competent court of Athens;

(14) in Israel: the courts of Tel Aviv-Jaffa;

(15) in Italy: the courts of Milan;

(16) in Portugal: the courts of Lisbon;

(17) in Spain: the courts of Madrid; and

(18) in Turkey: the Istanbul Central Courts and Execution Directorates of Istanbul, the Republic of Turkey.

14.3 Arbitration

The following paragraph is added as a new Subsection 14.3 (Arbitration) as it applies for those countries identified below. The provisions of this Subsection 14.3 prevail over those of Subsection 14.2 (Jurisdiction) to the extent permitted by the applicable governing law and rules of procedure:

ASIA PACIFIC

(1) In Cambodia, India, Indonesia, Laos, Philippines, and Vietnam:

Disputes arising out of or in connection with this Agreement will be finally settled by arbitration which will be held in Singapore in accordance with the Arbitration Rules of Singapore International Arbitration Center ("SIAC Rules") then in effect. The arbitration award will be final and binding for the parties without appeal and will be in writing and set forth the findings of fact and the conclusions of law.

The number of arbitrators will be three, with each side to the dispute being entitled to appoint one arbitrator. The two arbitrators appointed by the parties will appoint a third arbitrator who will act as chairman of the proceedings. Vacancies in the post of chairman will be filled by the president of the SIAC. Other vacancies will be filled by the respective nominating party. Proceedings will continue from the stage they were at when the vacancy occurred.

If one of the parties refuses or otherwise fails to appoint an arbitrator within 30 days of the date the other party appoints its, the first appointed arbitrator will be the sole arbitrator, provided that the arbitrator was validly and properly appointed. All proceedings will be conducted, including all documents presented in such proceedings, in the English language. The English language version of this Agreement prevails over any other language version.

(2) In the People's Republic of China:

In case no settlement can be reached, the disputes will be submitted to China International Economic and Trade Arbitration Commission for arbitration according to the then effective rules of the said Arbitration Commission. The arbitration will take place in Beijing and be conducted in Chinese. The arbitration award will be final and binding on both parties. During the course of arbitration, this agreement will continue to be performed except for the part which the parties are disputing and which is undergoing arbitration.

EUROPE, MIDDLE EAST, AND AFRICA

(3) In Albania, Armenia, Azerbaijan, Belarus, Bosnia-Herzegovina, Bulgaria, Croatia, Former Yugoslav Republic of Macedonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Moldova, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan:

All disputes arising out of this Agreement or related to its violation, termination or nullity will be finally settled under the Rules of Arbitration and Conciliation of the International Arbitral Center of the Federal Economic Chamber in Vienna (Vienna Rules) by three arbitrators appointed in accordance with these rules. The arbitration will be held in Vienna, Austria, and the official language of the proceedings will be English. The decision of the arbitrators will be final and binding upon both parties. Therefore, pursuant to paragraph 598 (2) of the Austrian Code of Civil Procedure, the parties expressly waive the application of paragraph 595 (1) figure 7 of the Code. IBM may, however, institute proceedings in a competent court in the country of installation.

(4) In Estonia, Latvia, and Lithuania:

All disputes arising in connection with this Agreement will be finally settled in arbitration that will be held in Helsinki, Finland in accordance with the arbitration laws of Finland then in effect. Each party will appoint one arbitrator. The arbitrators will then jointly appoint the chairman. If arbitrators cannot agree on the chairman, then the Central Chamber of Commerce in Helsinki will appoint the chairman.

AMERICAS COUNTRY AMENDMENTS

CANADA

10.1 Items for Which IBM May Be Liable

The following replaces Item 1 in the first paragraph of this Subsection 10.1 (Items for Which IBM May Be Liable):

1) damages for bodily injury (including death) and physical harm to real property and tangible personal property caused by IBM's negligence; and

13. General

The following replaces Item 13.d:

d. Licensee agrees to comply with all applicable export and import laws and regulations, including those of that apply to goods of United States origin and that prohibit or limit export for certain uses or to certain users.

The following replaces Item 13.i:

i. No right or cause of action for any third party is created by this Agreement or any transaction under it, nor is IBM responsible for any third party claims against Licensee except as permitted by the Limitation of Liability section above for bodily injury (including death) or physical harm to real or tangible personal property caused by IBM's negligence for which IBM is legally liable to that third party.

The following is added as Item 13.n:
n. For purposes of this Item 13.n, "Personal Data" refers to information relating to an identified or identifiable individual made available by one of the parties, its personnel or any other individual to the other in connection with this Agreement. The following provisions apply in the event that one party makes Personal Data available to the other:

(1) General

(a) Each party is responsible for complying with any obligations applying to it under applicable Canadian data privacy laws and regulations ("Laws").

(b) Neither party will request Personal Data beyond what is necessary to fulfill the purpose(s) for which it is requested. The purpose(s) for requesting Personal Data must be reasonable. Each party will agree in advance as to the type of Personal Data that is required to be made available.

(2) Security Safeguards

(a) Each party acknowledges that it is solely responsible for determining and communicating to the other the appropriate technological, physical and organizational security measures required to protect Personal Data.

(b) Each party will ensure that Personal Data is protected in accordance with the security safeguards communicated and agreed to by the other.

(c) Each party will ensure that any third party to whom Personal Data is transferred is bound by the applicable terms of this section.

(d) Additional or different services required to comply with the Laws will be deemed a request for new services.

(3) Use

Each party agrees that Personal Data will only be used, accessed, managed, transferred, disclosed to third parties or otherwise processed to fulfill the purpose(s) for which it was made available.

(4) Access Requests

(a) Each party agrees to reasonably cooperate with the other in connection with requests to access or amend Personal Data.

(b) Each party agrees to reimburse the other for any reasonable charges incurred in providing each other assistance.

(c) Each party agrees to amend Personal Data only upon receiving instructions to do so from the other party or its personnel.

(5) Retention

Each party will promptly return to the other or destroy all Personal Data that is no longer necessary to fulfill the purpose(s) for which it was made available, unless otherwise instructed by the other or its personnel or required by law.

(6) Public Bodies Who Are Subject to Public Sector Privacy Legislation

For Licensees who are public bodies subject to public sector privacy legislation, this Item 13.n applies only to Personal Data made available to Licensee in connection with this Agreement, and the obligations in this section apply only to Licensee, except that: 1) section (2)(a) applies only to IBM; 2) sections (1)(a) and (4)(a) apply to both parties; and 3) section (4)(b) and the last sentence in (1)(b) do not apply.

PERU

10. Limitation of Liability

The following is added to the end of this Section 10 (Limitation of Liability):

Except as expressly required by law without the possibility of contractual waiver, Licensee and IBM intend that the limitation of liability in this Limitation of Liability section applies to damages caused by all types of claims and causes of action. If any limitation on or exclusion from liability in this section is held by a court of competent jurisdiction to be unenforceable with respect to a particular claim or cause of action, the parties intend that it nonetheless apply to the maximum extent permitted by applicable law to all other claims and causes of action.

10.1 Items for Which IBM May Be Liable

The following is added to the end of this Subsection 10.1:

In accordance with Article 1328 of the Peruvian Civil Code, the limitations and exclusions specified in this section will not apply to damages caused by IBM's willful misconduct ("dolo") or gross negligence ("culpa inexcusable").

UNITED STATES OF AMERICA

5. Taxes

The following is added to the end of this Section 5 (Taxes):

For Programs delivered electronically in the United States for which Licensee claims a state sales and use tax exemption, Licensee agrees not to receive any tangible personal property (e.g., media and publications) associated with the electronic program.

Licensee agrees to be responsible for any sales and use tax liabilities that may arise as a result of Licensee's subsequent redistribution of Programs after delivery by IBM.

13. General

The following is added to Section 13 as Item 13.n:

n. U.S. Government Users Restricted Rights - Use, duplication or disclosure is restricted by the GSA IT Schedule 70 Contract with the IBM Corporation.

The following is added to Item 13.f:

Each party waives any right to a jury trial in any proceeding arising out of or related to this Agreement.

ASIA PACIFIC COUNTRY AMENDMENTS

AUSTRALIA

5. Taxes

The following sentences replace the first two sentences of Section 5 (Taxes):

If any government or authority imposes a duty, tax (other than income tax), levy, or fee, on this Agreement or on the Program itself, that is not otherwise provided for in the amount payable, Licensee agrees to pay it when IBM invoices Licensee. If the rate of GST changes, IBM may adjust the charge or other amount payable to take into account that change from the date the change becomes effective.

8. No Warranties

The following is added to the first paragraph of Section 8 (No Warranties):

Although IBM specifies that there are no warranties Licensee may have certain rights under the Trade Practices Act 1974 or other legislation and are only limited to the extent permitted by the applicable legislation.

10.1 Items for Which IBM May Be Liable

The following is added to Subsection 10.1 (Items for Which IBM Maybe Liable):

Where IBM is in breach of a condition or warranty implied by the Trade Practices Act 1974, IBM's liability is limited to the repair or replacement of the goods, or the supply of equivalent goods. Where that condition or warranty relates to right to sell, quiet possession or clear title, or the goods are of a kind ordinarily obtained for personal, domestic or household use or consumption, then none of the limitations in this paragraph apply.

HONG KONG SAR, MACAU SAR, AND TAIWAN

As applies to licenses obtained in Taiwan and the special administrative regions, phrases throughout this Agreement containing the word "country" (for example, "the country in which the original Licensee was granted the license" and "the country in which Licensee obtained the Program license") are replaced with the following:

(1) In Hong Kong SAR: "Hong Kong SAR"

(2) In Macau SAR: "Macau SAR" except in the Governing Law clause (Section 14.1)

(3) In Taiwan: "Taiwan."

INDIA

10.1 Items for Which IBM May Be Liable

The following replaces the terms of Items 1 and 2 of the first paragraph:

1) liability for bodily injury (including death) or damage to real property and tangible personal property will be limited to that caused by IBM's negligence; and 2) as to any other actual damage arising in any situation involving nonperformance by IBM pursuant to, or in any way related to the subject of this Agreement, IBM's liability will be limited to the charge paid by Licensee for the individual Program that is the subject of the claim.

13. General

The following replaces the terms of Item 13.g:

g. If no suit or other legal action is brought, within three years after the cause of action arose, in respect of any claim that either party may have against the other, the rights of the concerned party in respect of such claim will be forfeited and the other party will stand released from its obligations in respect of such claim.

INDONESIA

3.3 Term and Termination

The following is added to the last paragraph:

Both parties waive the provision of article 1266 of the Indonesian Civil Code, to the extent the article provision requires such court decree for the termination of an agreement creating mutual obligations.

JAPAN

13. General

The following is inserted as Item 13.n:

n. Any doubts concerning this Agreement will be initially resolved between us in good faith and in accordance with the principle of mutual trust.

MALAYSIA

10.2 Items for Which IBM Is Not Liable

The word "SPECIAL" in Item 10.2b is deleted.

NEW ZEALAND

8. No Warranties

The following is added to the first paragraph of this Section 8 (No Warranties):

Although IBM specifies that there are no warranties Licensee may have certain rights under the Consumer Guarantees Act 1993 or other legislation which cannot be excluded or limited. The Consumer Guarantees Act 1993 will not apply in respect of any goods which IBM provides, if Licensee requires the goods for the purposes of a business as defined in that Act.

10. Limitation of Liability

The following is added:

Where Programs are not obtained for the purposes of a business as defined in the Consumer Guarantees Act 1993, the limitations in this Section are subject to the limitations in that Act.

PEOPLE'S REPUBLIC OF CHINA

4. Charges

The following is added:

All banking charges incurred in the People's Republic of China will be borne by Licensee and those incurred outside the People's Republic of China will be borne by IBM.

PHILIPPINES

10.2 Items for Which IBM Is Not Liable

The following replaces the terms of Item 10.2b:

b. special (including nominal and exemplary damages), moral, incidental, or indirect damages or for any economic consequential damages; or

SINGAPORE

10.2 Items for Which IBM Is Not Liable

The words "SPECIAL" and "ECONOMIC" are deleted from Item 10.2b.

13. General

The following replaces the terms of Item 13.i:

i. Subject to the rights provided to IBM's suppliers and Program developers as provided in Section 10 above (Limitation of Liability), a person who is not a party to this Agreement will have no right under the Contracts (Right of Third Parties) Act to enforce any of its terms.

TAIWAN

10.1 Items for Which IBM May Be Liable

The following sentences are deleted:

This limit also applies to any of IBM's subcontractors and Program developers. It is the maximum for which IBM and its subcontractors and Program developers are collectively responsible.

EUROPE, MIDDLE EAST, AFRICA (EMEA) COUNTRY AMENDMENTS

EUROPEAN UNION MEMBER STATES

8. No Warranties

The following is added to Section 8 (No Warranties):

In the European Union ("EU"), consumers have legal rights under applicable national legislation governing the sale of consumer goods. Such rights are not affected by the provisions set out in this Section 8 (No Warranties).

EU MEMBER STATES AND THE COUNTRIES IDENTIFIED BELOW

Iceland, Liechtenstein, Norway, Switzerland, Turkey, and any other European country that has enacted local data privacy or protection legislation similar to the EU model.

13. General

The following replaces Item 13.e:

(1) Definitions - For the purposes of this Item 13.e, the following additional definitions apply:

(a) Business Contact Information - business-related contact information disclosed by Licensee to IBM, including names, job titles, business addresses, telephone numbers and email addresses of Licensee's employees and contractors. For Austria, Italy and Switzerland, Business Contact Information also includes information about Licensee and its contractors as legal entities (for example, Licensee's revenue data and other transactional information)

(b) Business Contact Personnel - Licensee employees and contractors to whom the Business Contact Information relates.

(c) Data Protection Authority - the authority established by the Data Protection and Electronic Communications Legislation in the applicable country or, for non-EU countries, the authority responsible for supervising the protection of personal data in that country, or (for any of the foregoing) any duly appointed successor entity thereto.

(d) Data Protection & Electronic Communications Legislation - (i) the applicable local legislation and regulations in force implementing the requirements of EU Directive 95/46/EC (on the protection of individuals with regard to the processing of personal data and on the free movement of such data) and of EU Directive 2002/58/EC (concerning the processing of personal data and the protection of privacy in the electronic communications sector); or (ii) for non-EU countries, the legislation and/or regulations passed in the applicable country relating to the protection of personal data and the regulation of electronic communications involving personal data, including (for any of the foregoing) any statutory replacement or modification thereof.

(e) IBM Group - International Business Machines Corporation of Armonk, New York, USA, its subsidiaries, and their respective Business Partners and subcontractors.

(2) Licensee authorizes IBM:

(a) to process and use Business Contact Information within IBM Group in support of Licensee including the provision of support services, and for the purpose of furthering the business relationship between Licensee and IBM Group, including, without limitation, contacting Business Contact Personnel (by email or otherwise) and marketing IBM Group products and services (the "Specified Purpose"); and

(b) to disclose Business Contact Information to other members of IBM Group in pursuit of the Specified Purpose only.

(3) IBM agrees that all Business Contact Information will be processed in accordance with the Data Protection & Electronic Communications Legislation and will be used only for the Specified Purpose.

(4) To the extent required by the Data Protection & Electronic Communications Legislation, Licensee represents that (a) it has obtained (or will obtain) any consents from (and has issued (or will issue) any notices to) the Business Contact Personnel as are necessary in order to enable IBM Group to process and use the Business Contact Information for the Specified Purpose.

(5) Licensee authorizes IBM to transfer Business Contact Information outside the European Economic Area, provided that the transfer is made on contractual terms approved by the Data Protection Authority or the transfer is otherwise permitted under the Data Protection & Electronic Communications Legislation.

AUSTRIA

8. No Warranties

In Austria (and Germany) the following replaces Section 8 (No Warranties) in its entirety, including its title, if Licensee paid a charge to obtain the Program.

8. Warranties and Exclusions

The warranty period is twelve months from the date of delivery. The limitation period for consumers in action for breach of warranty is the statutory period as a minimum.

The warranty for an IBM Program covers the functionality of the Program for its normal use and the Program's conformity to its specifications.

IBM warrants that when the Program is used in the specified operating environment it will conform to its specifications. IBM does not warrant uninterrupted or error-free operation of the Program or that IBM will correct all Program defects. Licensee is responsible for the results obtained from the use of the Program.

The warranty applies only to the unmodified portion of the Program.

If the Program does not function as warranted during the warranty period and the problem cannot be resolved with information available, Licensee may return the Program to the party from whom Licensee acquired it and receive a refund of the amount Licensee paid. If Licensee down loaded the Program, Licensee may contact the party from whom Licensee acquired it to obtain the refund.

This is IBM's sole obligation to Licensee, except as otherwise required by applicable statutory law.

10. Limitation of Liability

The following is added:

The following limitations and exclusions of IBM's liability do not apply for damages caused by gross negligence or willful misconduct.

10.1 Items for Which IBM May Be Liable

The following replaces the first sentence in the first paragraph:

Circumstances may arise where, because of a default by IBM in the performance of its obligations under this Agreement or other liability, Licensee is entitled to recover damages from IBM.

In the second sentence of the first paragraph, delete entirely the parenthetical phrase:

"(including fundamental breach, negligence, misrepresentation, or other contract or tort claim)".

10.2 Items for Which IBM Is Not Liable

The following replaces Item 10.2b:

b. indirect damages or consequential damages; or

BELGIUM, FRANCE, ITALY, AND LUXEMBOURG

10. Limitation of Liability

The following replaces the terms of Section 10 (Limitation of Liability) in its entirety:

Except as otherwise provided by mandatory law:

10.1 Items for Which IBM May Be Liable

IBM's entire liability for all claims in the aggregate for any damages and losses that may arise as a consequence of the fulfillment of its obligations under or in connection with this Agreement or due to any other cause related to this Agreement is limited to the compensation of only those damages and losses proved and actually arising as an immediate and direct consequence of the non-fulfillment of such obligations (if IBM is at fault) or of such cause, for a maximum amount equal to the charges (if the Program is subject to fixed term charges, up to twelve months' charges) Licensee paid for the Program that has caused the damages.

The above limitation will not apply to damages for bodily injuries (including death) and damages to real property and tangible personal property for which IBM is legally liable.

10.2 Items for Which IBM Is Not Liable

UNDER NO CIRCUMSTANCES IS IBM OR ANY OF ITS PROGRAM DEVELOPERS LIABLE FOR ANY OF THE FOLLOWING, EVEN IF INFORMED OF THEIR POSSIBILITY: 1) LOSS OF, OR DAMAGE TO, DATA; 2) INCIDENTAL, EXEMPLARY OR INDIRECT DAMAGES, OR FOR ANY ECONOMIC CONSEQUENTIAL DAMAGES; AND / OR 3) LOST PROFITS, BUSINESS, REVENUE, GOODWILL, OR ANTICIPATED SAVINGS, EVEN IF THEY ARISE AS AN IMMEDIATE CONSEQUENCE OF THE EVENT THAT GENERATED THE DAMAGES.

10.3 Suppliers and Program Developers

The limitation and exclusion of liability herein agreed applies not only to the activities performed by IBM but also to the activities performed by its suppliers and Program developers, and represents the maximum amount for which IBM as well as its suppliers and Program developers are collectively responsible.

GERMANY

8. No Warranties

This Section 8 (No Warranties) is amended as specified for AUSTRIA.

10. Limitation of Liability

The following replaces this Section 10 (Limitation of Liability) in its entirety:

a. IBM will be liable without limit for 1) loss or damage caused by a breach of an express guarantee; 2) damages or losses resulting in bodily injury (including death); and 3) damages caused intentionally or by gross negligence.

b. In the event of loss, damage and frustrated expenditures caused by slight negligence or in breach of essential contractual obligations, IBM will be liable, regardless of the basis on which Licensee is entitled to claim damages from IBM (including fundamental breach, negligence, misrepresentation, or other contract or tort claim), per claim only up to the greater of 500,000 euro or the charges (if the Program is subject to fixed term charges, up to 12 months' charges) Licensee paid for the Program that caused the loss or damage. A number of defaults which together result in, or contribute to, substantially the same loss or damage will be treated as one default.

c. In the event of loss, damage and frustrated expenditures caused by slight negligence, IBM will not be liable for indirect or consequential damages, even if IBM was informed about the possibility of such loss or damage.

d. In case of delay on IBM's part: 1) IBM will pay to Licensee an amount not exceeding the loss or damage caused by IBM's delay and 2) IBM will be liable only in respect of the resulting damages that Licensee suffers, subject to the provisions of Items a and b above.

13. General

The following replaces the provisions of 13.g:

g. Any claims resulting from this Agreement are subject to a limitation period of three years, except as stated in Section 8 (No Warranties) of this Agreement.

The following replaces the provisions of 13.i:

i. No right or cause of action for any third party is created by this Agreement, nor is IBM responsible for any third party claims against Licensee, except (to the extent permitted in Section 10 (Limitation of Liability)) for: i) bodily injury (including death); or ii) damage to real or tangible personal property for which (in either case) IBM is legally liable to that third party.

IRELAND

8. No Warranties

The following paragraph is added to the second paragraph of this Section 8 (No Warranties):

Except as expressly provided in these terms and conditions, or Section 12 of the Sale of Goods Act 1893 as amended by the Sale of Goods and Supply of Services Act, 1980 (the "1980 Act"), all conditions or warranties (express or implied, statutory or otherwise) are hereby excluded including, without limitation, any warranties implied by the Sale of Goods Act 1893 as amended by the 1980 Act (including, for the avoidance of doubt, Section 39 of the 1980 Act).

IRELAND AND UNITED KINGDOM

2. Agreement Structure

The following sentence is added:

Nothing in this paragraph shall have the effect of excluding or limiting liability for fraud.

10.1 Items for Which IBM May Be Liable

The following replaces the first paragraph of the Subsection:

For the purposes of this section, a "Default" means any act, statement, omission or negligence on the part of IBM in connection with, or in relation to, the subject matter of an Agreement in respect of which IBM is legally liable to Licensee, whether in contract or in tort. A number of Defaults which together result in, or contribute to, substantially the same loss or damage will be treated as one Default.

Circumstances may arise where, because of a Default by IBM in the performance of its obligations under this Agreement or other liability, Licensee is entitled to recover damages from IBM. Regardless of the basis on which Licensee is entitled to claim damages from IBM and except as expressly required by law without the possibility of contractual waiver, IBM's entire liability for any one Default will not exceed the amount of any direct damages, to the extent actually suffered by Licensee as an immediate and direct consequence of the Default, up to the greater of (1) 500,000 euro (or the equivalent in local currency) or (2) 125% of the charges (if the Program is subject to fixed term charges, up to 12 months' charges) for the Program that is the subject of the claim. Notwithstanding the foregoing, the amount of any damages for bodily injury (including death) and damage to real property and tangible personal property for which IBM is legally liable is not subject to such limitation.

```
10.2 Items for Which IBM Is Not Liable

The following replaces Items 10.2b and 10.2c:

b. special, incidental, exemplary, or indirect damages or consequential damages;
or

c. wasted management time or lost profits, business, revenue, goodwill, or
anticipated savings.
```

## sapha_env

Example E-24 shows the `sapha_env` script.

*Example E-24 sapha_env*

```
#--------------------------------------------------------------------------------+
# FUNCTION:
#          Helper function to source instance and database specific configuration
#          files and libraries
#
# AUTHOR:  Katharina Probst, probst@de.ibm.com
#          Walter Orb, walter.orb@de.ibm.com
# VERSION: 1.9.1
#          Use as template only
# CONFIGURATION:
#          SAP System ID
# LICENSE: LICENSE.txt
#--------------------------------------------------------------------------------+
typeset L_SID=$1
###Build environment ###
export
PATH=/bin:/usr/bin:/sbin:/usr/sbin:/usr/sbin/rsct/bin:/usr/es/sbin/cluster/utiliti
es:/usr/es/sbin/cluster/sap:/usr/es/sbin/cluster/cspoc

### Read Global Variables ###
. /usr/es/sbin/cluster/sap/sapha_${L_SID}.cfg

[[ -e $LOGFILE_DIR ]] || {
   mkdir -p $LOGFILE_DIR
}

LIB_LOGFILE="${LOGFILE_DIR}/sapha_$( print ${STEP}| tr 'A-Z' 'a-z' ).log"

case $LOG_LEVEL in
   0) exec  > $DEV_NULL    2>&1
      ;;
   1) exec >> $LIB_LOGFILE 2>&1
      ;;
   2) exec > $DEV_NULL 2>&1
      LIB_LOGFILE="${LOGFILE}"
      ;;
   3) LIB_LOGFILE="${LOGFILE}"
     exec >> $LIB_LOGFILE 2>&1
      ;;
```

```
esac

### Includes ###   #SAPutil required to be included. call nfs_service check!
. $LIB_DIR/log
. $LIB_DIR/util
. $LIB_DIR/SAPutil
. $LIB_DIR/DButil.db2
```

## sapha_TL1_cfg

Every line that ends with **`### <- edit`** needs to be adapted to the local environment and installation.

Example E-25 shows the `sapha_TL1_cfg` script.

*Example E-25 sapha_TL1_cfg*

```
#--------------------------------------------------------------------------------+
# FUNCTION:
#       Configuration file for SAP start/stop scripts in a PowerHA cluster
# AUTHOR:  Katharina Probst, probst@de.ibm.com
#          Walter Orb, walter.orb@de.ibm.com
# VERSION: 1.9.2
#          Use as template only
#          Sample file for SAP ABAP, Java, DB2 and Oracle
# Note:    ksh requires to move all lines only containing comments within an array
to work
# LEGAL:   see README.txt ILAN not applicable. you are entitled to apply changes
in here
#--------------------------------------------------------------------------------+
### Variables ###
### General Variables ###

typeset -l LC_SID=$SID
typeset -l DBTYPE="db2"                                                   ### <- edit

### User IDs and Versions###
         SAPADM="${LC_SID}adm"
          DBADM="${DBTYPE}${LC_SID}"
         DBVERS="102_64"                                                  ### <- edit
#MIN_SAP_RELEASE="700" # min SAP release the script is made for. Change to
suppress Warnings.
VERBOSE_LOGGING=""          # disable verbose logging for calls to PowerHA
utilities
         DEBUG="set -vx"    # enable for debugging with -x or -xv

### FLAGS ###
#DUAL_STACK=0 # should be 0, set it to 1 only for double stack systems like PI
# ERS_FLAG=1  # 1 = ERS should be handled by the cluster. If not set it to 0.
 NFS_FLAG=0  # 0= disable 1 = enable
LOG_LEVEL=3  # 0-3
 TIME_LOG=0  # prints a special logfile (TIME_LOGFILE) to estimate the cluster
tuning if set to 1
             # The information is printed using the timeout function in the
function library
```

```
             # Values
             # 1: write log message at the end of timeout function
             # 2: write log message at each iteration of command retry
             # 3: write log message at the end of timeout function and at each
iteration of command retry
#MONITOR="PROCESS_SERVICE" #Values:
                    # PROCESS: The monitor will only check for processes. (low
network traffic, reduced false failover during high network load due to timeouts)
                   # SERVICE: The monitor tests for the service availability (not
recommended for automated failovers)
                   # PROCESS_SERVICE: The monitor alerts if the process cannot be
found and prints logmessages if the Service is not accessible

       OK=0
    ERROR=1
  WARNING=2
   ONLINE=0
  OFFLINE=2
 DEV_NULL="/dev/null"

### Script Directories ###
     SCRIPT_DIR="/usr/es/sbin/cluster/sap"                          ### <- edit
        LIB_DIR="/usr/es/sbin/cluster/sap/lib"                      ### <- edit
    LOGFILE_DIR="/var/hacmp/log/sap_${SID}"                         ### <- edit
        LOGFILE="${LOGFILE_DIR}/sapha.log"                          ### <- edit
MONITOR_LOGFILE="${LOGFILE_DIR}/sapha_monitor.log"                  ### <- edit
   TIME_LOGFILE="${LOGFILE_DIR}/sapha_time.log                      ### <- edit

### PowerHA resource groups ###
typeset -A RG_SAP=(
        [GFS]="rg_sap_${LC_SID}_gfs"                                ### <- edit
        [NFS]="rg_sap_${LC_SID}_nfs"                                ### <- edit
        [DB2]="rg_sap_${LC_SID}_db2"                                ### <- edit
[DB2_PRIMARY]="rg_db2hadr_primary"                                  ### <- edit
#        [ORA]="rg_sap_${LC_SID}_ora"                               ### <- edit
#     [ASCS10]="rg_sap_${LC_SID}_ascs"                              ### <- edit
#      [SCS11]="rg_sap_${LC_SID}_scs"                               ### <- edit
#      [ERS20]="rg_sap_${LC_SID}_aers"                              ### <- edit
#      [ERS21]="rg_sap_${LC_SID}_ers"                               ### <- edit
#  [DVEBMGS12]="rg_sap_${LC_SID}_ci01"                              ### <- edit
#        [D02]="rg_sap_${LC_SID}_as11"                              ### <- edit
)

### SAP variables ###
SAPMNT_NFS="/export/sapmnt/${SID}" ### <- edit, if NFS_FLAG is set to 1 this is
mandatory!

typeset -A SAP_VHOST=(
#    [ASCS10]="${LC_SID}ascs"                                       ### <- edit
#     [SCS11]="${LC_SID}scs"                                        ### <- edit
# [DVEBMGS12]="${LC_SID}ci01"                                       ### <- edit
       [DB2]="${LC_SID}db2"                                         ### <- edit
#       [ORA]="${LC_SID}db01"                                       ### <- edit
#       [D11]="is3${LC_SID}as11"                                    ### <- edit
#     [ERS20]="${LC_SID}aers"                                       ### <- edit
```

```
#     [ERS21]="${LC_SID}ers"                                        ### <- edit
#      [JC10]="is3${LC_SID}ci10"                                    ### <- edit
#       [J11]="is3${LC_SID}as11"                                    ### <- edit
       [GFS]="xyz"       ### <- edit, if NFS_FLAG is set to 1 this is mandatory!
)

#SAP_SCS=(
#   ABAP="ASCS10"                                                   ### <- edit
#   JAVA="SCS11"                                                    ### <- edit
#)

#typeset -A SAP_ERS_MAP=(
#   [ASCS10]="ERS20"                                                ### <- edit
#    [SCS11]="ERS21"                                                ### <- edit
#    [ERS20]="ASCS10"                                               ### <- edit
#    [ERS21]="SCS11"                                                ### <- edit
#)

INSTANCE_TYPE=${INSTANCE%[0-9][0-9]}
  INSTANCE_NR=${INSTANCE#${INSTANCE_TYPE}}

        VHOST=${SAP_VHOST[$INSTANCE]}

  SAP_PROFILE="/usr/sap/${SID}/SYS/profile/${SID}_${INSTANCE}_${VHOST}" # a weak 
prereque is now to maintain the profiles locally as a copy

#typeset -A SAP_EXE_DIR=( # edit: adjust accordingly to your sapcpe settings. 
prefer instance directories /usr/sap/${SID}/${INSTANCE}/exe
#      [ASCS]="/usr/sap/${SID}/SYS/exe/run"
#       [SCS]="/usr/sap/${SID}/SYS/exe/run"
#   [DVEBMGS]="/usr/sap/${SID}/${INSTANCE}/exe"
#         [D]="/usr/sap/${SID}/${INSTANCE}/exe"
#        [JC]="/usr/sap/${SID}/SYS/exe/run"
#         [J]="/usr/sap/${SID}/SYS/exe/run"
#       [SYS]="/usr/sap/${SID}/SYS"
#       [ERS]="/usr/sap/${SID}/SYS/exe/run"
#)

typeset -A SAP_EXE=(
   [startsap]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/startsap"
    [stopsap]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/stopsap"
    [r3trans]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/R3trans"
         [dw]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/disp+work"
     [ensmon]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/ensmon"
 [getwebpage]="${SCRIPT_DIR}/GetWebPage.class"
  [brconnect]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/brconnect"
#    [rfcping]="${SAP_EXE_DIR[$INSTANCE_TYPE]}/rfcping"       SAP removed this 
from the kernel due to security issues
)

### SAP Commands ###
SAP_STARTSAP="su - ${SAPADM} -c ${SAP_EXE[startsap]}"
 SAP_STOPSAP="su - ${SAPADM} -c ${SAP_EXE[stopsap]}"
 SAP_R3TRANS="su - ${SAPADM} -c ${SAP_EXE[r3trans]}"
      SAP_DW="su - ${SAPADM} -c ${SAP_EXE[dw]}"
```

```
  SAP_ENSMON="su - ${SAPADM} -c ${SAP_EXE[ensmon]}" # pass 1 fore ens and 2 for
ers check
  SAP_BRCONNECT="su - ${SAPADM} -c ${SAP_EXE[brconnect]}"
# SAP_RFCPING="su - ${SAPADM} -c ${SAP_EXE[rfcping]}"  SAP removed this from the
kernel due to security issues
   SAP_ENREP="su - ${SAPADM} -c cd ${SAP_EXE_DIR[$INSTANCE]} &&
er.sap${SID}_${INSTANCE}"
SAP_GETWEBPAGE="java ${SAP_EXE[getwebpage]}  ${INSTANCE}
http://${VHOST}.wdf.sap.corp:5${INSTANCE_NR}00
${LOGFILE_DIR}/GetWebPage${INSTANCE}.html 30"                          ### <- edit

#  CLUSTER_NODES_UP="cldump | grep "Node Name:" | grep -c  "State: UP" " # returns
the number of nodes in the state UP
#  SAP_MSPROT="su - $SAPADM -c ${SAP_EXE[msprot]}"


### DB2 Variables ###
DB2INSTANCE="${DBADM}"                                                 ### <- edit
DB2_HOME="/db2/${DB2INSTANCE}"                                         ### <- edit
DB2NODES_CFG="${DB2_HOME?}/sqllib/db2nodes.cfg"

PEER_WINDOW="peer window only" # edit change to empty may result in data loss!
APPLIACTION_INTEGRATION=1                                              ### <- edit
             # 0 stopping the instance will result in failover behaviour,
                         # 1 will deactivate the application monitors if the
instance is stopped.
                         # For a restart it is recommended to first turnoff
Application monitoring to give time to start the instance incl. the activation of
the DB HADR

typeset -A DB2_EXE=(
    [db2gcf]="${DB2_HOME}/sqllib/bin/db2gcf"
    [db2fm]="${DB2_HOME}/sqllib/bin/db2fm"
   [db2stop]="${DB2_HOME}/sqllib/adm/db2stop"
  [db2_kill]="${DB2_HOME}/sqllib/bin/db2_kill"
)

    DB2_DB2GCF="su - ${DBADM} -c ${DB2_EXE[db2gcf]}"
DB2_STOP_FORCE="su ${DBADM} -c ${DB2_EXE[db2stop]} force"
      DB2_KILL="su ${DBADM} -c ${DB2_EXE[db2_kill]}"

  DB2_SATO=60

### Sleep and retry parameters for timeout function
typeset -A TIMEOUT=(
                        [sap_check_nfs_avail]="2 2"       #important for
controlling startup sequence. Gives NFS time to start.          ### <- edit
  [sap_ers_move_wait_for_clusterstate_stable]="2 2"            ### <- edit
)
```

## cl_db2_start_local

Example E-26 on page 673 shows the `cl_db2_start_local` script to start a DB2 partition in a PowerHA cluster.

*Example E-26 cl_db2_start_local*

```
#!/bin/ksh93
##################################################################################
#
# FUNCTIONALITY:
#          Start a DB2 partition in a PowerHA cluster
#        It does not support cold standby DB2 handling!!! For that purpose please
use the other
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 0.0
#          Use as template only
#          requires db2gcf command to work
# ARGUMENTS:
#          SID=$1
#          PARTITION=$2  /0, 1, 2
# CONFIGURATION:
#          sapha_{SID}.cfg
# PREREQ: usage of default installation path /usr/es/sbin/cluster/sap/SID
# LICENSE: LICENSE.txt
##################################################################################
# set environment variable DEBUG='set -x' for verbose logging
$DEBUG

### Read global variables and functions
typeset -u  INSTANCE=$1      #DB2
typeset -u       SID=$2
typeset    PARTITION=${3:-0} # This option has only tested with value "0"

. /usr/es/sbin/cluster/sap/sapha_env ${SID}
STEP="${DB2INSTANCE}_INSTANCE_START"

#########################################################
# MAIN: Start DB                                        #
#########################################################

log "[INFO]: Start DB2 instance ${DB2INSTANCE} partition ${PARTITION}"

# TODO: This is a temporary fix. db2fmcd does not start via the inittab entry.
ps -ef | grep -v "grep" | grep "db2fmcd" || nohup
"/db2/db2sc1/db2_software/bin/db2fmcd" &

# NFS Check. If NFS unavailable su - db2sc1 might not work
sap_check_nfs_service || {
   log "[ERROR]: NFS is not acessable. Stop processing! If you do not want this
check to be performed disable NFS_FLAG."
   exit 1
}

#Start the DB2 fault monitor, not the service (here we want to ensure the
environment is cleaned up)
log "[INFO]: Start the fault monitor if not already started."
su - $DBADM -c $DB2_HOME/sqllib/bin/db2fm  -S -i $DB2INSTANCE | grep "state is
AVAILABLE" || {
   su - $DBADM -c $DB2_HOME/sqllib/bin/db2fm -U -i $DB2INSTANCE || {
```

```
      log "[ERROR]: Can not start the db2fm. Since there is no other instance to
ensure the db2 instance stays alive we stop cluster processing."
      exit 1
   }
   log "[OK]: Start of fault monitor completed successfully. Continue instance
startup."
}

#check for start allowance
if [[ -r /tmp/do_not_start.${DB2INSTANCE?}.${PARTITION?} ]]; then
   log "[OK]: Bypassing startup for ${DB2INSTANCE},${PARTITION}. Found the
following lock: /tmp/do_not_start.${DB2INSTANCE?}.${PARTITION?}"
   exit 0
fi

#if the db is up we are done
check_db2_partition_status  $PARTITION && {
  log  "[OK]: Partition ${PARTITION} was already started."
  exit 0
}

db2_start_database || {
   log "[ERROR]: Could not start DB2 Partition ${PARTITION} of Instance ${DBADM}"
    exit 1
}

log "[OK]: Start DB2 Partition ${PARTITION} of Instance ${DBADM} successful."
exit 0
```

## cl_db2_stop_local

Example E-27 shows the script to stop a DB2 partition in a PowerHA cluster.

*Example E-27   cl_db2_stop_local*

```
#!/bin/ksh93
##################################################################################
#
# FUNCTIONALITY:
#          Stop a DB2 partition in a PowerHA cluster
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 0.0
#          Use as template only
# ARGUMENTS:
#          INSTANCE=$1
#          PARTITION=$2
#          SID=$3
# CONFIGURATION:
#          sapha_{SID}.cfg
# PREREQ: usage of default installation path /usr/es/sbin/cluster/sap/SID
# LICENSE: LICENSE.txt
##################################################################################
# set environment variable DEBUG='set -x' for verbose logging
$DEBUG
```

```
### Read global variables and functions
typeset -u  INSTANCE=$1
typeset -u       SID=$2
typeset    PARTITION=${3:-0}

. /usr/es/sbin/cluster/sap/sapha_env ${SID}
STEP="${DB2INSTANCE}_INSTANCE_STOP"


#########################################################
# MAIN: Stop DB                                         #
#########################################################

log "[INFO]: Stop DB instance ${DBADM} partition ${PARTITION}"

#here intentionally without su - db2adm -c just in case the NFS is not up and
running. If not up the script will get hang in su command.
#the fault monitor is intentionally stopped. In case we need to do a hard kill it
would restart the instance.
$DB2_HOME/sqllib/bin/db2fm -D -i $DB2INSTANCE

db2_stop_database || {
  log "[ERROR]: Could not stop DB2 Partition ${PARTITION} of Instance ${DBADM}"
  exit 0
}

log "[OK]: Stop DB2 Partition ${PARTITION} of Instance ${DBADM} successful."
exit 0
```

## cl_db2_start_hadr

Example E-28 shows the script to activate a DB for hot standby after you start the instance.

*Example E-28 cl_db2_start_hadr*

```
#!/bin/ksh93
##################################################################################
#
# FUNCTIONALITY:
#          Activate a DB for hotstandby after starting the instance
#          This script does not support partitioned db2 feature!
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 0.0
#          Use as template only
#          requires db2gcf command to work
# ARGUMENTS:
#          SID=$1
#     ROLE=$2
# CONFIGURATION:
#          sapha_{SID}.cfg
# PREREQ: usage of default installation path /usr/es/sbin/cluster/sap/SID
# LICENSE: LICENSE.txt
##################################################################################

### Read global variables and functions
```

```
typeset -u       ROLE=$1
typeset -u       SID=$2

. /usr/es/sbin/cluster/sap/sapha_env ${SID}

$DEBUG
STEP="ACTIVATE_HADR_${ROLE}"


#########################################################
# MAIN: Activate                                        #
#########################################################

log "[INFO]: Enter start sequence of a ${ROLE}."

   get_role=$(su - ${DBADM} -c db2 get db cfg for ${LC_SID} | grep "HADR database
role")
   has_role=$(echo ${get_role} | egrep -c "STANDBY|PRIMARY")
   is_primary=$(echo ${get_role} | grep -c "PRIMARY")

#Ensure roles are assigned
[[ $has_role == 0 ]] && {
   log "[INFO]: Instance has no role. Start the instance as ${ROLE}."
      startAs $ROLE && {
         log "[OK]: Start database as ${ROLE} completed successfully."
         exit 0
      }
      log "[ERROR]: Start instance as ${ROLE} failed or the DB was already started
as HADR (returncode 4 from startAS)"
      [[ $ROLE == "STANDBY" ]]  & exit 0
      exit 1 #here we must not provide the option to failover. the risk is to
crash the content of the DB
}

log "[INFO]: Check if already active."
   check_db2_hadr_status ${ROLE} && {
      log "[OK]: Already active. Exit startup"
      exit 0
   }
log "[INFO]: Instance is not an active HADR member."

[[ $ROLE == "PRIMARY" ]] && {
   deact=$(su - ${DBADM} -c "db2pd -db ${LC_SID}" | grep -c "not activated" )
   [[ $is_primary == 1 ]] && {
      [[ ${deact} != 0 ]] && {
         log "[INFO]: Activate DB for HADR. ${get_role}"
         activateDatabase && {
            log "[OK]: Activation of DB was required and completed successfully."
            exit 0
         }
         log "[ERROR]: Activation of ${ROLE} failed."
         exit 0
      }
   }
```

```
   # has role but is not primary then it is a takeover. Only possible with an
activated Standby
   [[ ${deact} == 0 ]] && {
      log "[INFO]: Detected a running standby. Initiate takeover now."
      takeoverDB && {
         log "[OK]: Takeover completed successfully."
         exit 0
      }
      log "[ERROR]: Takeover failed."
      exit 1 #here we must not provide the option to failover. the risk is to
crash the content of the DB
   }
   log "[INFO]: Will not force a deactivated standby up as Primary. Stop Cluster
Processing and wait for manual intervention."
   exit 1
}

[[ $ROLE == "STANDBY" ]] && {
   [[ $is_primary == 1 ]] && {
      log "[INFO]: Instance was a Primary and the Standby needs to be
reintegrated."
      reintegrateStandby && {
         log "[OK]: STANDBY reintegrated"
         exit 0
      }
      log "[ERROR]: failed to integrate instance into DB2 HADR. Try to activate
and integrate again."
      exit 0
   }

   deact=$(su - ${DBADM} -c "db2pd -db ${LC_SID}" | grep -c "not activated" )
   [[ ${deact} != 0 ]] && {
      log "[INFO]: Activate DB for HADR. ${get_role}"
      activateDatabase && {
         log "[OK]: Activation of DB was required and completed successfully."
         exit 0
      }
      log "[ERROR]: Activation of ${ROLE} failed or DB was already activated. Try
it differently"
   }

   startAs ${ROLE} && {
      log "[OK]: Standby is up."
      exit 0
   }
   log "[ERROR]: Bringup of STANDBY failed."
   exit 0
}

log "[ERROR]: Script is not configured correctly. Parameters are: SID=<SID> and
ROLE=[STANDBY|PRIMARY]."
exit 1
```

## cl_db2_stop_hadr

Example E-29 shows the script to stop a DB2 partition in a PowerHA cluster.

*Example E-29 cl_db2_stop_hadr*

```
#!/bin/ksh93
#################################################################################
#
# FUNCTIONALITY:
#          Stop a DB2 partition in a PowerHA cluster
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 0.0
#          Use as template only
# ARGUMENTS:
#          INSTANCE=$1
#          PARTITION=$2
#          SID=$3
# CONFIGURATION:
#          sapha_{SID}.cfg
# PREREQ: usage of default installation path /usr/es/sbin/cluster/sap/SID
# LICENSE: LICENSE.txt
#################################################################################


### Read global variables and functions
typeset -u       ROLE=$1
typeset -u       SID=$2

. /usr/es/sbin/cluster/sap/sapha_env ${SID}
$DEBUG
STEP="DEACTIVATE_HADR_${ROLE}"


#########################################################
# MAIN: Stop DB                                         #
#########################################################
#Alternative stop command:   su - ${candidate_P_instance?} -c "db2gcf -t 3600 -d
-i ${candidate_P_instance?} -i ${candidate_S_instance?} -h ${DB2HADRDBNAME?} -L"

log "[INFO]: Stop DB ${ROLE} instance ${DBADM}"
stopDB ${ROLE} && {
  log "[OK]: Stop DB2 ${ROLE} of Instance ${DBADM} successful."
  exit 0
}

log "[WARNING]: Could not stop DB2 ${ROLE} of Instance ${DBADM}. Leave as is and
continue cluster activity."
exit 0
```

# cl_db2_monitor_hadr

Example E-30 on page 679 shows the script to monitor a DB2 partition in a PowerHA cluster.

*Example E-30   cl_db2_monitor_hadr*

```
#!/bin/ksh93
################################################################################
#
# FUNCTIONALITY:
#          Monitor a DB2 partition in a PowerHA cluster
#          The monitor only writes a message in case the test failed
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 0.0
#          use as template only
# CONFIGURATION:
#          SID
#          Role
#          sapha_SID.cfg
# PREREQ: usage of default installation path /usr/es/sbin/cluster/sap/
# LICENSE: LICENSE.txt
################################################################################


### Read global variables and functions
typeset -u   ROLE=$1
typeset -u     SID=$2

. /usr/es/sbin/cluster/sap/sapha_env ${SID}
$DEBUG
STEP="${ROLE}_${DB2INSTANCE}_MONITOR"


############################################################
# MAIN: Monitor DB                                         #
############################################################
# check if db2start or db2stop is the state to set the application to using db2fm

# if we nly allow to start the standby if th eprimary is onlie we have a hard time
to recover
#[[ $ROLE == STANDBY ]] && {
#  clRGinfo -s | grep ${RG_SAP[DB2_PRIMARY]} | grep "ONLINE" || exit 0 #primary RG
is not online we do not monitor the standby
#}
[[ $APPLIACTION_INTEGRATION == 1 ]] && {
   isInProgress
   [[ $? != $ERROR ]] && exit 0
}

# Monitor instance as a service
check_db2_hadr_status ${ROLE} && exit 0

[[ $APPLIACTION_INTEGRATION == 1 ]] && {
   #this is to avoid a racecondition between db2gcf -d command and the monitor
   isInProgress
   [[ $? != $ERROR ]] && exit 0
}
exit 1
```

## lib/DButil.db2

Example E-31 shows a function library for DB2 start and stop scripts in a PowerHA DB2HADR cluster.

*Example E-31 lib/DButil.db2*

```
#--------------------------------------------------------------------------------+
# FUNCTION:
#          Function library for DB2 start and stop scripts in a PowerHA DB2HADR
#          cluster
# AUTHOR:  Katharina Probst, probst@de.ibm.com
# VERSION: 1.9.2
#          Use as template only
# CONFIGURATION:
#          sapha_${SID}.cfg
# PREREQ:  usage of default installation path /usr/es/sbin/cluster/sap
# License: ../LICENSE.txt
#--------------------------------------------------------------------------------+

####################################################
# FUNCTION:   isInProgress                         #
# PURPOSE:    return if db2 is in startup or shutdown #
# RETURNCODE: OK = stable, continue with monitoring
#             WARNING = startup, maintenance or shutdown in progress
#             ERROR= db2fm not enabled               #
####################################################
function isInProgress {
   $DEBUG

   typeset PROCS=$(ps -ef | egrep "sapdbctrl start|sapdbctrl
stop|startdb|stopdb|db2start|db2stop|db2gcf|db2fm|db2 activate|db2 deactivate" |
egrep "${LC_SID}|${SID}")
   typeset FM_STATE=$(su - ${DBADM} -c ${DB2_EXE[db2fm]} -S -s)

   [[ $(echo ${FM_STATE} | grep "'fault monitor' state is AVAILABLE") ]] || {
      monitor_log "[WARNING]: The fault monitor state is not set to AVAILABLE. We
cannot determine correct state. Monitoring will be disabled to avoid false
failovers."
      return $ERROR
   }

   [[ $(echo ${FM_STATE} | grep "lib64/libdb2gcf.a' state is AVAILABLE") ]] || {
      monitor_log "[INFO]: The DB is intentionally put into offline state for
maintenance or other purposes by the DB administrator. No active monitoring."
      return $OK
   }

   [[ $PROCS ]] || {
      monitor_log "[WARNING]: The following processes indicate a DB2 instance in
startup or shutdown."
      return $WARNING
   }
   monitor_log "[OK]: The DB2 instance is up. Exit with ERROR to continue with the
monitoring of HADR."
   return $ERROR
```

```
}

function check_db2_prerequisites {
##########################################################################
# FUNCTION:   check_db2_prerequisites                                    #
# PURPOSE:    check db2 prerequisites before attempt to start/stop      #
# RETURNCODE: OK, ERROR                                                  #
##########################################################################
  $DEBUG

  check_executable_status ${DB2_EXE[db2gcf]} DB2_EXE || {
      lib_log "[INFO] check_db2_prerequisites: check executable status for db2gcf
failed."
      return $ERROR
  }
  return $OK
}

function check_db2_partition_status {
####################################################
# FUNCTION:   check_db2_partition_status          #
# PURPOSE:    return status of DB2 Partition      #
# PARAMETER:  PARTITION= Number of Partition      #
# RETURNCODE: OK, ERROR, WARNING                  #
# Note: db2gcf check can not be done remotley using#
#       clrsh. The db2<SID> user is no login user. #
#       This will cause the call to fail.         #
####################################################
 $DEBUG
 typeset PARTITION=$1

   lib_log "[INFO]: Check DB2 instance status for Partition ${PARTITION}."
   [[ -x $DB2_HOME/sqllib/bin/db2gcf ]] && {
      $DB2_DB2GCF -s -p $PARTITION -i $DB2INSTANCE | grep "Available" && return
$OK
      $DB2_DB2GCF -s -p $PARTITION -i $DB2INSTANCE | grep "Not Operable" &&
lib_log "[INFO] check_db2_partition_status: DB2 Partition is not operable" &&
return $ERROR
      lib_log "[INFO] check_db2_partition_status: DB2 Partition is not available,
but can be started"
      return $WARNING
   }

   lib_log "[INFO]: db2gcf could not be found. Use process monitor."
   p_pid=$(ps -u ${DB2INSTANCE?} -o args | grep -c "^db2sysc ${NN?}[ ]*$")
   if [[ $p_pid == 0 && $PARTITION -eq 0 ]]; then
      p_pid=$(ps -u ${DB2INSTANCE?} -o args | grep -v "^db2sysc [0-9]" | grep -c
"^db2sysc")
   fi
   [[ $p_pid == 0 ]] && return $ERROR
   return $OK
}

################################################################
# FUNCTION:   check_db2_hadr_status                           #
```

```
# PURPOSE:    check if the Peers are activated and connected    #
# RETURNCODE: OK, ERROR                                         #
#################################################################
function check_db2_hadr_status {
  $DEBUG
  typeset -u ROLE=$1

   lib_log "[INFO]: Start to check DB2 HADR status for ${ROLE}."
   is_hadr_active=$(su - ${DBADM} -c db2pd -hadr -db ${LC_SID} | grep -c "HADR is
not active")

   [[ $is_hadr_active == 0 ]] && { # HADR is active
      lib_log "[INFO]: HADR is activated. Start to check Service"
      # Handle if Primary
      [[ $ROLE == PRIMARY ]] && {
         is_active_primary=$(su - ${DBADM} -c db2pd -hadr -db ${LC_SID} | grep -c
"Active")
         # Handle if Primary is active
         [[ $is_active_primary == 1 ]] && {
            get_state_primary=$(su - ${DBADM} -c db2pd -hadr -db ${LC_SID} | awk
'/Primary / {print $2}')
            lib_log "[INFO]: The primary instance is active and is in state
${get_state_primary}."
            return $OK
         }
         # Handle if Primary is NOT active
         lib_log "[INFO]: The primary instance is NOT active"
         return $ERROR
      }
      # Handle if Standby
      is_active_standby=$(su - ${DBADM} -c db2pd -hadr -db ${LC_SID} | grep -c
"Standby -- Up")
      [[ $is_active_standby == 1 ]] && {
         lib_log "[INFO]: The standby instance is up. Start to validate connection
state."
         get_state_secondary=$(su - ${DBADM} -c db2pd -hadr -db ${LC_SID} | awk
'/Standby / {print $2}')
         lib_log "[INFO]: The standby instance has a connection status of:
${get_state_secondary}."
         return $OK
      }
   }
   lib_log "[ERROR]: HADR is not active."
   return $ERROR
}

function activateDatabase {
  $DEBUG
   typeset -i RC=$OK
   lib_log "[INFO]: Activate database ${DB2INSTANCE}."
    su - ${DBADM} -c "db2 activate database ${LC_SID}"
    [[ $? = @(0|2) ]] && return $OK #2 indicates
    return $?
}
```

```
function startAs {
  $DEBUG
  typeset -u ROLE=$1

   lib_log "[INFO]: Start database as ${ROLE}."
   su - ${DBADM} -c "db2 start hadr on db ${LC_SID} as ${ROLE}"
   return $?
}

function takeoverDB {
  $DEBUG

   lib_log "[INFO]: Start to takeover STANDBY. First without force."
   su - ${DBADM} -c "db2 takeover hadr on db ${LC_SID}" && {
      lib_log "[OK]: Normal takeover, no force required."
      return $OK
   }

   lib_log "[Info]: Call takeover by force."
   su - ${DBADM} -c "db2 takeover hadr on db ${LC_SID} by force ${PEER_WINDOW}" &&
{
      lib_log "[OK]: Takeover by force ${PEER_WINDOW} completed."
      return $OK
   }
   return $ERROR #takeover by force failed
}

function reintegrateStandby {
  $DEBUG

    lib_log "[INFO]: Reintegrate Standby."
   get_role=$(su - ${DBADM} -c db2 get db cfg for ${LC_SID} | grep "HADR database
role")
   has_role=$(echo ${get_role} | egrep -c "STANDBY|PRIMARY")
   is_primary=$(echo ${get_role} | grep -c "PRIMARY")

    lib_log "[INFO]: Current role of instance before reintegrate as standby:
${get_role}."
   # if role == PRIMARY we force all connections offline
   [[ $is_primary == 1 ]] && {
      lib_log "[INFO]: Forcing apps off before reintegration. This was a primary."
      su - ${DBADM} -c db2pd -applications -db ${LC_SID}  | grep -v "Restoring" |
awk '{print $2}' | grep ^"[0-9]" | while read applid; do
         su - ${DBADM} -c "db2 force application \( $applid  \)"
         lib_log "[INFO]: Force applid ${applid} connected to ${DB2INSTANCE}"
      done
      lib_log "[OK]: All connections are forced offline."
   }
   lib_log "[INFO]: Try to start instance as STANDBY now."

   startAs STANDBY && return $OK
   lib_log "[WARNING]: Start instance did not complete as expected."
   return $ERROR
}
```

```
function stopDB {
  $DEBUG
  typeset -u ROLE=$1

   lib_log "[INFO]: Stop ${ROLE} database."
   [[ $ROLE == "PRIMARY" ]] && {
      lib_log "[INFO]: Will stop ${ROLE}."
      su - ${DBADM} -c "db2 deactivate db ${LC_SID}"
      return $?
   }

   lib_log "[INFO]: Will NOT deactivate a standby."
   return $OK
}

function db2_start_database {
#########################################################################
# FUNCTION:   db2_start_database                                        #
# PURPOSE:    startup DB2 database                                      #
# RETURNCODE: OK, ERROR                                                 #
#########################################################################
   $DEBUG

   typeset RC=$OK

#comment code to use if running multible partitions
   #check if we run a single partition: If single partition, clean IPCs first
#  if test -r $DB2NODES_CFG ; then
#    nln=$(wc -l $DB2NODES_CFG | awk '{print $1}')
#  else
#     nln=1
#  fi

#  # we are a single partition, clean IPC!
#  log "[INFO]: Validate if we run in single partition mode. before start to
cleanup."
#  if [ $nln -eq 1 ]; then
   lib_log "[INFO]: Clean up resources before starting the DB2 Partition calling
db2gcf followed by an ipclean "
   $DB2_DB2GCF -k -i $DBADM -L || kill -9 $(ps -o pid,comm -u $DB2INSTANCE | grep
"db2[a-z]" | grep -v "db2haicu" | grep -v \.ksh | awk '{print $1}') 2> /dev/null
   su - ${DB2INSTANCE} -c "ipclean -a"
#  fi

   log "[INFO]: Ready to start the DB2 instance."
#  (($DB2_SATO=nln*2+60))
   $DB2_DB2GCF -t $DB2_SATO -u -i $DB2INSTANCE -L || RC=$ERROR

   return $RC
}

function db2_stop_database {
#########################################################################
# FUNCTION:   db2_stop_database                                         #
# PURPOSE:    stop DB2 database                                         #
```

```
# RETURNCODE: OK, ERROR                                                       #
##############################################################################
   $DEBUG

   typeset RC=$OK


   # Ensure home directory is accessible ...
   /bin/ksh -c "cd $DB2_HOME/sqllib/bin; touch
$DB2_HOME/sqllib/tmp/.tmp.$PARTITION; rm -f $DB2_HOME/sqllib/tmp/.tmp.$PARTITION"
&
   ProcNum=$!
   sleep 2
   kill -0 ${ProcNum} 2> /dev/null
   ret=$?
   kill -9 ${ProcNum} 2> /dev/null

   #NFS and Home is accessible, do a normal stop
   [[ sap_check_nfs_service ]] && [[ ${ret} != 0 ]] && {
      $DB2_DB2GCF -t $DB2_SATO -d -p $PARTITION -i $DBADM -L || RC=$ERROR      #
call normal shutdown when NFS is available
      return $RC
   }

   # Home is not accessible, so th eonly option is to kill the processes
   [[ ${ret} == 0 ]] && {
      log "[ERROR]: ${DB2_HOME} may not be accessible, Will do a hard kill of
${DB2INSTANCE} processes using kill -9."
      kill -9 $(ps -o pid,comm -u $DB2INSTANCE | grep "db2[a-z]" | awk '{print
$1}') 2> /dev/null > /dev/null
      exit 0
   }

   # NFS not available
    log "[WARNING] NFS Service not available, using db2stop force to shutdown
database"
    INSTHOME=$DB2_HOME

PATH=$PATH:${DB2_HOME}/sqllib/bin:${DB2_HOME}/sqllib/adm:${DB2_HOME}/sqllib/misc
    export INSTHOME DB2INSTANCE PATH                                         #
export environment variables for db2 utilities

    $DB2_STOP_FORCE                                                          #
call db2stop force
    $DB2_KILL


   return $RC
}
```

## lib/log

Example E-32 on page 686 shows a function library for writing log files in a PowerHA cluster.

*Example E-32 lib/log*

```
#--------------------------------------------------------------------------------+
# FUNCTION:
#          Function library for writing log files in a PowerHA cluster
#
# AUTHOR:  Katharina Probst, probst@de.ibm.com
#          Walter Orb, walter.orb@de.ibm.com
# VERSION: 1.9.2
#          Use as template only
# CONFIGURATION:
#          sapha_${SID}.cfg
# LICENSE: ../LICENSE.txt
#--------------------------------------------------------------------------------+
####################################################
# FUNCTION:   log                                  #
# PURPOSE:    write standardized logfile           #
# PARAMETER:  string                               #
# RETURNCODE: -                                    #
####################################################
function log {
  typeset DATE="$(date +"%y%m%d %H:%M:%S") "
  print "${DATE} ${STEP} $*" >> $LOGFILE
  (( $LOG_LEVEL < 2 )) &&  print "${DATE} ${STEP} $*" >> $LIB_LOGFILE
}

function monitor_log {
  typeset DATE="$(date +"%y%m%d %H:%M:%S") "
  print "${DATE} ${STEP} $*" >> $MONITOR_LOGFILE
}

function time_log {
 typeset DATE="$(date +"%y%m%d %H:%M:%S") "
 print "${DATE} ${STEP} $*" >> $TIME_LOGFILE
}

function lib_log {
 typeset DATE="$(date +"%y%m%d %H:%M:%S") "
 print "${DATE} ${STEP} $*" >> $LIB_LOGFILE
}
```

## lib/SAPutil

Example E-33 shows a function library for SAP start and stop scripts in a PowerHA cluster.

*Example E-33 lib/SAPutil*

```
#--------------------------------------------------------------------------------+
# FUNCTION:
#          Function library for SAP start and stop scripts in a PowerHA cluster
#
# AUTHOR:  Katharina Probst, probst@de.ibm.com
#          Walter Orb, walter.orb@de.ibm.com
# VERSION: 1.9.2
#          Use as template only
```

```
# CONFIGURATION:
#          sapha_${SID}.cfg
# LICENSE: ../LICENSE.txt
#--------------------------------------------------------------------------------+
function sap_check_service_status {
####################################################################
# FUNCTION:   sap_check_service_status                             #
# PURPOSE:    return status of ASCS or Application Server instance #
# PARAMETER:  INSTANCE = full name of Instance                     #
# RETURNCODE: 0=OK, or accumulated  ERROR codes                    #
####################################################################
   $DEBUG

   typeset    INSTANCE=$1
   typeset INSTANCE_NR=${INSTANCE:${#INSTANCE}-2}
   typeset       VHOST=${SAP_VHOST[$INSTANCE]}


   case "$INSTANCE" in
   ASCS*|SCS* )
          #if [[ -x  ${SAP_EXE[ensmon]} ]] ; thenensmon hangs if we test a network
down instead of returning 8 what causes a false failover -> ask for SAP fix
          #   $SAP_ENSMON -H $VHOST -I $INSTANCE_NR 1
          #else
             sap_check_process_status $INSTANCE
          #fi
          ;;
   D* )
          #if [[ -x ${SAP_EXE[rfcping]} ]] ; then SAP stops to ship this function
due to security issues
          #   $SAP_RFCPING ashost=$VHOST sysnr=$INSTANCE_NR     No alternative
(except niping) is delivered up to 7.20 as of 7/2010
          #else
             sap_check_process_status $INSTANCE
          #fi
          ;;
   J* )
          if [[ -x ${SAP_EXE[getwebpage]} ]] ; then
             $SAP_GETWEBPAGE
          else
            sap_check_process_status $INSTANCE
          fi
          ;;
   ERS* )
        #if [[ -x  ${SAP_EXE[ensmon]} ]] ; then         ensmon hangs if we test a
network down instead of returning 8 what causes a false failover -> ask for SAP
fix
          #   $SAP_ENSMON pf=$SAP_PROFILE 2
          #else
             sap_check_process_status $INSTANCE
          #fi
          ;;
   esac
   return $?
}
```

```
function sap_check_nfs_service {
#########################################################################
# FUNCTION:   sap_check_nfs_service                                     #
# PURPOSE:    check NFS server availability before attempt to           #
#             start/stop                                                #
# PARAMETER:  -                                                         #
# RETURNCODE: accumulated returncodes. 0 = OK                           #
#########################################################################
   $DEBUG

   NFS_SERVER=${SAP_VHOST[GFS]}

   (( NFS_FLAG == 1 )) || return $OK                                     # exit
if NFS is not used

   /usr/sbin/ping -c 1 -w 3 $NFS_SERVER || {                             # check
network connectivity to NFS server
      log "[WARNING] Cannot ping NFS server $NFS_SERVER"
      return $ERROR
   }
   /usr/bin/rpcinfo -u $NFS_SERVER nfs  || {                             # check
that NFS services are running on NFS server
     log "[WARNING] NFS services not running on $NFS_SERVER"
     return $ERROR
   }
   /usr/bin/showmount -e $NFS_SERVER | /usr/bin/grep -c $SAPMNT_NFS || {  # check
that sapmnt is exported on NFS server
     log "[WARNING] $SAPMNT_NFS is not exported on NFS server $NFS_SERVER"
     return $ERROR
   }
   /usr/sbin/mount | /usr/bin/grep -c $SAPMNT_NFS || {                   # check
that sapmnt is mounted on local host
      log "[WARNING] $SAPMNT_NFS is currently not mounted on local host"
      return $ERROR
   }
   return $OK
}

function sap_check_service_prerequisites {
#########################################################################
# FUNCTION:   sap_check_service_prerequisites                           #
# PURPOSE:    check service prerequisites before attempt to start/stop  #
# PARAMETER:  -                                                         #
# RETURNCODE: accumulated returncodes. 0 = OK                           #
#########################################################################
   $DEBUG

   sap_check_nfs_service                      || return $ERROR   # first check
that the NFS server is available
   check_executable_status ${SAP_EXE[startsap]} || return $ERROR   # startsap and
stopsap will let us now in case we did not configure the sapcpe right
   check_executable_status ${SAP_EXE[stopsap]}  || return $ERROR   # ... at least
during runtime :-(
```

```
   [[ -e $SAP_PROFILE ]]  || return $ERROR                         # This covers
the check if the profiles are copied to SYS/profile as well

   return $OK
}

function sap_check_process_status {
####################################################################
# FUNCTION:   sap_check_process_status                             #
# PURPOSE:    check for the most important processes of ASCS or    #
#             Application Server, ... instances                    #
# PARAMETER:  INSTANCE = full name of Instance                     #
# RETURNCODE: 0=OK, or accumulated  ERROR codes                    #
####################################################################
   $DEBUG

   typeset INSTANCE=$1
   PROCESS_NAME="sap${SID}_${INSTANCE}"

   case "$INSTANCE" in
   ASCS*|SCS* )
            check_global_process_status ${SAPADM} "en.${PROCESS_NAME}"
          ;;
   D* )
             check_global_process_status ${SAPADM} "dw.${PROCESS_NAME}" \
              && check_global_process_status ${SAPADM}  "icman"
          ;;
   J* )
            check_global_process_status ${SAPADM} "ig.${PROCESS_NAME}" \
              && check_global_process_status ${SAPADM} "jc.${PROCESS_NAME}" \
              && check_global_process_status ${SAPADM}
"usr/sap/${SID}/${INSTANCE}/exe/jlaunch" # TODO: after 7.10 one jlaunch process is
replaced
          ;;                                             # by an icm process
belonging to the dispatcher <-keep this as a remark
   ERS* )
            check_local_process_status ${SAPADM} "er.${PROCESS_NAME}"
          ;;
   esac
   return $?
}

function sap_start_service {
#########################################################################
# FUNCTION:   sap_start_service                                         #
# PURPOSE:    generic function to start ASCS, SCS, CI or AS             #
# PARAMETER:  SERVICE = r3 or sapstartsrv                               #
# RETURNCODE: Returncode of startsap                                    #
#########################################################################
   $DEBUG
   typeset TASK="R3"


   [[ $INSTANCE == @(J*|SCS*) ]] && TASK=J2EE
```

```
   [[ ( $INSTANCE == @(ERS*) ) && ( ${SAP_ERS_MAP[$INSTANCE]} == @(SCS*) ) ]] &&
TASK=J2EE

   [[ $INSTANCE == @(ERS*) ]] && {
      #If SAP central Services run on this host we do not start the ERS
      check_local_process_status ${SAPADM}
"en.sap${SID}_${SAP_ERS_MAP[$INSTANCE]}" && {
        log "[WARNING]: ERS will not be started. On this node Central Services of
${SAP_ERS_MAP[$INSTANCE]} are running. Start of ERS will be in monitoring mode for
a new cluster node."
        return 0
      }
      # To ensure to get a fresh socket we kill the ers process (NO STOPSAP MUST
BE CALLED)
      # The difference: shutdown.sap= er.* process <-> kill.sap= startsap process
killing itself and all children

      HOST=$(hostname)
      if [[ -e /usr/sap/${SID}/${INSTANCE}/work/kill.sap.${HOST} ]] ; then
        KILL_CMD=$(< /usr/sap/${SID}/${INSTANCE}/work/kill.sap.${HOST} )
        eval $KILL_CMD
        rm /usr/sap/${SID}/${INSTANCE}/work/kill.sap.${HOST}
      fi
      if [[ -e /usr/sap/${SID}/${INSTANCE}/work/shutdown.sap.${HOST} ]] ; then
        KILL_CMD=$(< /usr/sap/${SID}/${INSTANCE}/work/shutdown.sap.${HOST} )
        eval $KILL_CMD
        rm /usr/sap/${SID}/${INSTANCE}/work/shutdown.sap.${HOST}
      fi

      #this kills the ers in case the kill -2 hangs in the kernel and does not
stop the process.
      #It is required to have a working ERS at the end.
      TEMP="er.sap${SID}_${INSTANCE}"
      LIST=$(ps -ef | grep $TEMP | awk '/er.sap/ {print $2}') > $DEV_NULL
      if [[ $LIST ]] ; then
        for pid in $LIST; do
          lib_log "[INFO]: old ERS process er.sap${SID}_${INSTANCE} with ${pid} is
killed before new ERS is started."
          kill -9 $pid
        done
      fi
   }

   $SAP_STARTSAP $TASK $INSTANCE ${SAP_VHOST[$INSTANCE]} >> $LIB_LOGFILE
   RET=$?

   [[ $INSTANCE == @(ERS*) ]] && {
    cp /usr/sap/${SID}/${INSTANCE}/work/kill.sap
/usr/sap/${SID}/${INSTANCE}/work/kill.sap.${HOST}
     cp /usr/sap/${SID}/${INSTANCE}/work/shutdown.sap
/usr/sap/${SID}/${INSTANCE}/work/shutdown.sap.${HOST}
   }
   return $RET
}
```

```
function sap_stop_service {
#########################################################################
# FUNCTION:   sap_stop_service                                          #
# PURPOSE:    generic function to stop ASCS, SCS, CI or AS              #
# PARAMETER:  TASK = j2ee, r3 or sapstartsrv                            #
# RETURNCODE: Returncode of stopsap                                     #
#########################################################################

   $DEBUG
   typeset TASK=${1:-R3}

   [[ $INSTANCE == @(ERS*) ]] && return 0
   [[ ( $TASK == "R3" ) && ( $INSTANCE == @(J*|SCS*) ) ]] && TASK=J2EE

   $SAP_STOPSAP $TASK $INSTANCE ${SAP_VHOST[$INSTANCE]} >> $LIB_LOGFILE
   return $?
}

function sap_kill_instance {
#########################################################################
# FUNCTION:   sap_kill_instance                                         #
# PURPOSE:    generic function to kill ASCS, SCS, CI or AS              #
# PARAMETER:  INSTANCE                                                  #
# RETURNCODE: 0                                                         #
#########################################################################
   $DEBUG
   typeset INSTANCE=${1}

   [[ -r /usr/sap/${SID}/${INSTANCE}/work/kill.sap ]] || {       # check that
kill.sap file is available
     log "[WARNING] Cannot find kill.sap"
     return 0
   }

       KILL_CMD=$(< /usr/sap/${SID}/${INSTANCE}/work/kill.sap )  # read kill
command from kill.sap
    PROCESS_PID=$(print $KILL_CMD | /usr/bin/cut -d" " -f 3)     # get sapstart
process id from kill.sap
   PROCESS_NAME=$(/usr/bin/ps -p $PROCESS_PID -o comm=)          # get active
process name for this PID

   [[ $PROCESS_NAME == "sapstart" ]] && eval $KILL_CMD           # call kill
command only if this is a sapstart process

   # now kill any remaining processes of this instance
   /usr/bin/ps -u $SAPADM -o pid=,args= | /usr/bin/grep $INSTANCE | while read PID
CMD
   do
      log "[WARNING] Process ${CMD} with process id ${PID} still running, send
kill -9 signal"
      /usr/bin/kill -9 $PID
   done

   return 0
}
```

```
function sap_check_database_connect {
########################################################################
# FUNCTION:   sap_check_database_connect                               #
# PURPOSE:    issue R3trans -d to check DB connectivity                #
# RETURNCODE: OK, ERROR                                                #
########################################################################
   $DEBUG

   typeset STARTSAP_CHECK_RC=(
        DB=1
       JDB=2
       SCS=4
      ASCS=8
      ABAP=16
      JAVA=32
   )

   if [[ $INSTANCE == @(J*) ]] ; then
#          $SAP_STARTSAP check $INSTANCE ${SAP_VHOST[$INSTANCE]} #startsap check
seems not perfectly working for 7.01 to keep the scripts working from 7.00 onward
we first need a fix
#          (( $? & ${STARTSAP_CHECK_RC.JDB} )) && return $OK
      [[ $DBTYPE == db2 ]] && {
         check_global_process_status ${DBADM} db2sysc && return $OK
      }
      [[ $DBTYPE == ora ]] && {
         $SAP_BRCONNECT -u / -f dbstate > $DEV_NULL 2>&1 && return $OK
      }
   elif [[ ( $INSTANCE == @(SCS*) ) && ( $DUAL_STACK != 1 ) ]]; then  # DB check
for SCS only doesn't work correctly, skip for now
      return $OK                                                     # Dual stack
SCS systems can be checked with R3trans
   else
       $SAP_R3TRANS -d -w $DEV_NULL  > $DEV_NULL
       [[ $? == @(0|4|8) ]]  && return $OK
   fi
   return $ERROR
}

function sap_check_release {
########################################################################
# FUNCTION:   sap_check_release                                        #
# PURPOSE:    the provided scripts require a specific SAP Version. This #
#             function extracts the SAp release out of disp+work and    #
#             compares it to the Required minimum release              #
# PARAMETER:  -                                                        #
# RETURNCODE: OK, ERROR                                                #
########################################################################

   $DEBUG
   [[ $INSTANCE == @(ERS*) ]] && return 0
   RELEASE=$( $SAP_DW |  awk '/kernel release/ {print $3}') > $DEV_NULL
   (( $RELEASE >= $MIN_SAP_RELEASE ))
   return $?
```

```
}

function sap_ers_move {
#####################################################################
# FUNCTION:   sap_ers_move                                          #
# PURPOSE:    move the ERS accordingly to the amount of online nodes    #
#             and where the SAP Central Services are located        #
#####################################################################
$DEBUG

case "$INSTANCE" in
        ASCS*|SCS*)
                typeset       CS="${INSTANCE}"
                typeset
NODENAME=$(/usr/es/sbin/cluster/utilities/get_local_nodename)
                typeset      ERS="${SAP_ERS_MAP[$INSTANCE]}"
                typeset     NODE=$(clRGmove -l nodes_acquire_for_rg_or_set -g
${RG_SAP[$INSTANCE]} ONLINE $NODENAME | awk '!/^#/  {node=$1}  END {print node}')
                typeset   ERS_RG="${RG_SAP[${SAP_ERS_MAP[${INSTANCE}]}]}"
                typeset      LOG="log"
                 ;;
        ERS*)
                typeset       CS="${SAP_ERS_MAP[$INSTANCE]}"
               typeset NODENAME=$(clRGinfo -s | awk -v RG=${RG_SAP[$CS]} -F : '$1
== RG && $2 == "ONLINE" {print $3}')
                typeset      ERS="${INSTANCE}"
                typeset     NODE=$(clRGmove -l nodes_acquire_for_rg_or_set -g
${RG_SAP[${SAP_ERS_MAP[${INSTANCE}]}]} ONLINE $NODENAME | awk '!/^#/  {node=$1}
END {print node}')
                typeset   ERS_RG="${RG_SAP[$INSTANCE]}"
                typeset      LOG="monitor_log"
                 ;;
esac
   # wait until the cluster is stable. not waiting and then issue a move will
bring the ERS resource group into an ERROR state requiring manual steps to recover
the cluster

   timeout "get_cluster_state_stable"
${TIMEOUT[sap_ers_move_wait_for_clusterstate_stable]} || {
      $LOG "[INFO]: move of ERS stopped in function sap_ers_move in SAPutil. The
cluster is unstable"
      exit 0
   }

   #in case we have less than 2 nodes up ERS can not be started
   [[ $(cldump | grep "Node Name:" | grep -c  "State: UP") < 2 ]] && {
          $LOG "[INFO]: Cluster has only one node up. ${ERS} cannot be started."
          return 0
   }
# IF there is a node to move/start the ERS determine if it needs to be started or
moved.
   [[ -n $NODE ]] && {
      [[ $(clRGinfo -s | grep "${ERS_RG}" | grep -c "ONLINE") > 0 ]] && {
       ERS_NODE=$(clRGinfo -s | awk -v RG=${ERS_RG} -F : '$1 == RG && $2 ==
"ONLINE" {print $3}')
```

```
        if [[ $ERS_NODE == $NODE ]] ; then  #we restart and do not move
           $LOG "[INFO]: Instance ${INSTANCE} restarted on ${NODE}."
         sap_start_service
        else# in case we really have to change node we call move
           $LOG "[INFO]: Instance ${INSTANCE} initiated the move of ${ERS} to
${NODE}."
        clRGmove -g ${ERS_RG} -n $NODE -m  > $DEV_NULL
        fi
        return 0
      }
      clRGmove -g ${ERS_RG} -n $NODE -u  > $DEV_NULL
      $LOG "[INFO]: Instance ${INSTANCE} initiated the start of ${ERS} on
${NODE}."
   }
}

function sap_set_scs_instance {
#########################################################################
# FUNCTION:   sap_set_scs_instance                                     #
# PURPOSE:    copy variable SAP_SCS from variable ABAP_SCS or          #
#             JAVA_SCS dependent on the instance type                  #
#########################################################################
$DEBUG

   case "$INSTANCE" in
      ASCS*|D*)  SAP_SCS_INSTANCE=${SAP_SCS.ABAP}
                 ;;
      SCS*|J*)   SAP_SCS_INSTANCE=${SAP_SCS.JAVA}
                 ;;
      ERS*)      if [[ ${SAP_ERS_MAP[$INSTANCE]} == @(SCS*) ]]; then
                    SAP_SCS_INSTANCE=${SAP_SCS.JAVA}
                 else
                    SAP_SCS_INSTANCE=${SAP_SCS.ABAP}
                 fi
                 ;;
   esac
}
```

## lib/util

Example E-34 shows a function library for various utility functions in a PowerHA cluster.

*Example E-34 lib/util*

```
#---------------------------------------------------------------------------+
# FUNCTION:
#          Function library for various utility functions in a PowerHA cluster
#
# AUTHOR:  Katharina Probst, probst@de.ibm.com
#          Walter Orb, walter.orb@de.ibm.com
# VERSION: 1.9.2
#          Use as template only
# CONFIGURATION:
#          sapha_${SID}.cfg
# LICENSE: ../LICENSE.txt
```

```
#-----------------------------------------------------------------------------+
################################################
# FUNCTION:   timeout                          #
# PURPOSE:    implements a while loop with timing  #
#              add-on                          #
# PARAMETER:  CMD = function to be tested      #
#             RETRY = retry the test           #
#             SLEEP = sleep between the tests  #
# RETURNCODE: OK, ERROR                        #
################################################
function timeout {
  $DEBUG

  typeset CMD=$1  RETRY=${2-10}  SLEEP=${3-10}
  typeset -i COUNT

   for (( COUNT=1;; COUNT++ ))
   do
     $CMD && {
         (( TIME_LOG & 1 )) && time_log "${INSTANCE} [SUCCESS]: ${CMD} was
successful after ${COUNT} out of ${RETRY} retries with a sleep of ${SLEEP}
seconds."
         return $OK
      }
      (( TIME_LOG & 2 ))   && time_log "${INSTANCE} [WARNING]: ${CMD} failed at
${COUNT} out of ${RETRY} retries with a sleep of ${SLEEP} seconds."
      (( COUNT == RETRY )) && break
      sleep $SLEEP
   done
   (( TIME_LOG & 1 )) && time_log "${INSTANCE} [ERROR]: ${CMD} failed after
${RETRY} retries with a sleep of ${SLEEP} seconds."
   return $ERROR
}

############################################################################
# FUNCTION:   check_global_process_status                                  #
# PURPOSE:    return status of process                                     #
# PARAMETER:  USER = SIDADM or DB2 user                                    #
#             PROCESS = Name to look for in the CMD column in the ps output #
# RETURNCODE: OK, ERROR                                                    #
############################################################################
function check_global_process_status {
   $DEBUG
   typeset USER=$1 PROCESS=$2

   /usr/bin/ps -u ${USER} -o args | grep -q "${PROCESS}" && return $OK # clrsh
fails if cluster is not stable. This causes an Application monitor to hang.

   for i in `/usr/es/sbin/cluster/utilities/clnodename`
   do
      clrsh $i "ps -u ${USER} -o args" | grep -q "${PROCESS}" && return $OK #ToDo:
clrsh call is not valid starting with PowerHA7.1 clusters
   done

   return $ERROR
```

```
}
###########################################################################
# FUNCTION:   check_local_process_status                                  #
# PURPOSE:    return status of process                                    #
# PARAMETER:  USER = SIDADM or DB2 user                                   #
#             PROCESS = Name to look for in the CMD column in the ps output #
#          NODE = node to check for the process         #
# RETURNCODE: OK, ERROR                                                   #
###########################################################################
function check_local_process_status {
   $DEBUG
   typeset USER=$1  PROCESS=$2   #NODE=$3

   #clrsh $NODE  "ps -u ${USER} -o args" | grep "${PROCESS}" && return $OK
   /usr/bin/ps -u ${USER} -o args | grep -q "${PROCESS}" && return $OK

   return $ERROR
}

#########################################################
# FUNCTION:   check_fs_status                           #
# PURPOSE:    return status of fs                       #
# PARAMETER:  FS = name to grep for in the mount output #
# RETURNCODE: OK, ERROR                                 #
#########################################################
function check_fs_status {
   $DEBUG

   typeset FS=$1

   mount | grep $FS && return $OK

   lib_log "[INFO] check_fs_status: Filesystem ${FS} is not mounted"
   return $ERROR
}

####################################################
# FUNCTION:   check_executable_status              #
# PURPOSE:    check if executable can be used      #
# PARAMETER:  EXECUTABLE = full path to executable #
# RETURNCODE: OK, ERROR                            #
####################################################
function check_executable_status {
   $DEBUG

   typeset EXECUTABLE=$1

   [[ -x $EXECUTABLE ]] && return $OK

   if [[ -e $EXECUTABLE ]] ; then
      lib_log "[INFO] check_executable_status: can not execute ${EXECUTABLE}."
   else
      lib_log "[INFO] check_executable_status: ${EXECUTABLE} not found."
   fi
```

```
   return $ERROR
}

function get_cluster_state_stable {
  lssrc -ls clstrmgrES | grep "ST_STABLE"
  return $?
}
```

# F

# PowerHA and workload partition examples

This appendix contains the full screen captures and examples that we used to test the features of workload partitions (WPARs) under PowerHA.

# PowerHA smit examples

We list the full smit windows in addition to the excerpts that we used in Chapter 8, “Workload partition and PowerHA scenario” on page 487.

## Configuring the resource group in PowerHA

Example F-1 shows the full smit window for the definition of the resource group that we used for “Configuring the resource group in PowerHA” on page 491.

*Example F-1 Resource group settings for WPAR (testwpar)*

```
Change/Show All Resources and Attributes for a Resource Group

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

                                                        [Entry Fields]
  Resource Group Name                                 testwpar
  Participating Nodes (Default Node Priority)         sys5lpar3 sys5lpar4

  Startup Policy                                      Online On Home Node Only
  Fallover Policy                                     Fallover To Next Priority
Node In The List
  Fallback Policy                                     Fallback To Higher Priority
Node In The List
  Fallback Timer Policy (empty is immediate)         []                        +

  Service IP Labels/Addresses                        [localwpar]               +
  Application Controllers                            [ApplicationB]            +

  Volume Groups                                      []                        +
  Use forced varyon of volume groups, if necessary    false                    +
  Automatically Import Volume Groups                  false                    +

  Filesystems (empty is ALL for VGs specified)       []                        +
  Filesystems Consistency Check                       fsck                     +
  Filesystems Recovery Method                         sequential               +
  Filesystems mounted before IP configured            false                    +

  Filesystems/Directories to Export (NFSv2/3)        []                        +
  Filesystems/Directories to Export (NFSv4)          []                        +
  Stable Storage Path (NFSv4)                        []                        +
  Filesystems/Directories to NFS Mount               []
  Network For NFS Mount                              []                        +

  Tape Resources                                     []                        +
  Raw Disk PVIDs                                     []                        +

  Primary Workload Manager Class                     []                        +
  Secondary Workload Manager Class                   []                        +

  Miscellaneous Data                                 []
  WPAR Name                                          [testwpar]                +
  User Defined Resources                             [ ]
```

## Scripts that we used for WPARs in PowerHA

We list the scripts that we used for our application named `ApplA`.

Example F-2 shows the application. Due to the settings that we used in the start and stop script, this script must be placed in the directory `/usr/local/bin`. The application is a simple loop that creates some **echo** output. In the start and stop script, we create or remove the control files that keep the loop running or end it.

*Example F-2 Application script (ApplA)*

```
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x

echo "$(date) \"Program ApplA initiated \" " >> /var/hacmp/log/applA.log
echo $(ls -l /var/hacmp/adm/AppAup) >> /var/hacmp/log/applA.log

while [ -a /var/hacmp/adm/AppAup ]
do
  echo "$(date) \"Application A is running\" " >> /var/hacmp/log/applA.log
  echo "on Hostname: $(hostname) \n   uname: $(uname -n) " >>
/var/hacmp/log/applA.log
  sleep 30
done
```

Example F-3 shows the start and stop script that we used. We placed it under `/usr/local/ha`. There, we created a hard link of the file name `StartA` to `StopA`. This way, we only maintain one file. At the end of this example, we used exit code 888. It can be any number equal to or greater than 1. We used 888 to identify it in the log files more easily.

*Example F-3 Start and stop script for ApplA*

```
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#
Name=$(basename $0 )

if [ "$Name" = "StartA" ]
then
  echo "$(date) \"Application A started\" " >> /var/hacmp/log/applA.log
  touch /var/hacmp/adm/AppAup
  nohup /usr/local/bin/ApplA &
  exit 0
elif [ "$Name" = "StopA" ]
then
  rm -f /var/hacmp/adm/AppAup
  echo "$(date) \"Application A stopped\" " >> /var/hacmp/log/applA.log
  exit 0
else
  echo "$(date) \"ERROR - Application A start/stop script called with wrong name\"
" >> /var/hacmp/log/applA.log
  exit 888
fi
```

Example F-4 shows the monitoring script that we used. If the application is running, the exit code must be 0. If the application is not running, the exit code must be unequal to 0. In our example, we used 777, which is easier to identify in the log files.

*Example F-4 Monitoring script for application (ApplA)*

```
#!/usr/bin/ksh
#
[[ "$VERBOSE_LOGGING" = "high" ]] && set -x
#

if ( $(ps -ef | grep -w ApplA | grep -vq grep) )
then
  echo "$(date) : ApplA is running \n" >>/var/hacmp/log/monA.log
  exit 0
else
  echo "$(date) : ApplA is NOT running \n" >>/var/hacmp/log/monA.log
  exit 777
fi
```

# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

## IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.


- *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940
- *Exploiting IBM AIX Workload Partitions*, SG24-7955
- *IBM System Storage Solutions Handbook*, SG24-5250


You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

**ibm.com**/redbooks

## Other publications

These publications are also relevant as further information sources:


- *RSCT Version 3.1.2.0 Administration Guide*, SA22-7889
- "Invincible Supply Chain - Implementation Guide for SAP HotStandby liveCache with PowerHA 7.1.1" on techdocs, which is published by the International SAP IBM Competence Center (ISICC):

  http://w3-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100677


## Online resources

These websites are also relevant as further information sources:


- IBM Information Center for PowerHA SystemMirror

  http://publib.boulder.ibm.com/infocenter/aix/v7r1/topic/com.ibm.aix.powerha.navigation/powerha_main.htm
- PowerHA SystemMirror Concepts Guide

  http://public.boulder.ibm.com/infocenter/aix/v7r1/topic/com.ibm.aix.powerha.concepts/hacmpconcepts_pdf.pdf
- Compatibility between IBM AIX, IBM PowerHA, and IBM disk subsystems

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105638

# Help from IBM

IBM Support and downloads

**ibm.com**/support

IBM Global Services

**ibm.com**/services

IBM PowerHA SystemMirror Standard Edition 7.1.1 for AIX Update

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

# IBM PowerHA SystemMirror Standard Edition 7.1.1 for AIX Update

**Introduces the latest features of IBM PowerHA SystemMirror**

**Provides application sample scenarios**

**Implements a high availability environment**

This IBM Redbooks publication helps you install, tailor, and configure the new IBM PowerHA SystemMirror for AIX 7.1.1 Standard Edition. This book gives an understanding of the Cluster Aware AIX (CAA). This book helps you design a solution to migrate from the previous version of the IBM PowerHA.

This IBM Redbooks publication is targeted toward technical professionals (consultants, technical support staff, IT architects, and IT specialists) responsible for providing continuous availability solutions and support.

**INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION**

**BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE**

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

**For more information:**
**ibm.com**/redbooks

SG24-8030-00 ISBN 0738437379