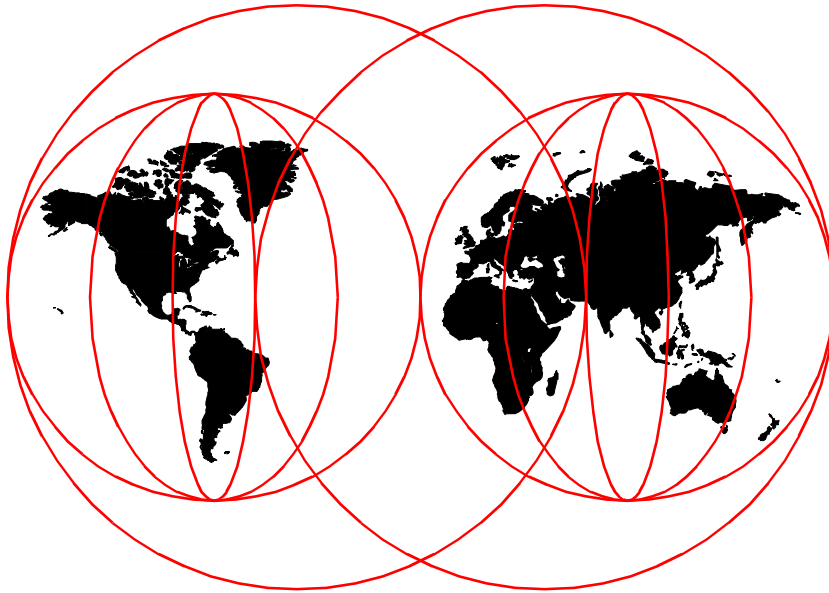


Migrating to HACMP/ES

Ole Conradsen, Mitsuyo Kikuchi, Adam Majchczyk, Shona Robertson, Olaf Siebenhaar



International Technical Support Organization

www.redbooks.ibm.com

SG24-5526-00



International Technical Support Organization

Migrating to HACMP/ES

January 2000

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix D, "Special notices" on page 133.

First Edition (January 2000)

This edition applies to HACMP version 4.3.1 and HACMP/ES version 4.3.1 for use with the AIX Version 4.3.2 Operating System.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. JN9B Building 003 Internal Zip 2834
11400 Burnet Road
Austin, Texas 78758-3493

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

© Copyright International Business Machines Corporation 2000. All rights reserved.
Note to U.S Government Users – Documentation related to restricted rights – Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Preface	vii
The team that wrote this redbook	vii
Comments welcome	viii
Chapter 1. Introduction	1
1.1 High Availability (HA)	1
1.2 Features and device support in HACMP	2
1.3 Scalability and HACMP/ES	9
1.3.1 RISC System Cluster Technology (RSCT)	10
1.3.2 Benefits of HACMP/ES	11
Chapter 2. Upgrade and migration paths	13
2.1 Why upgrade or migrate?	13
2.1.1 Fixes and back ported device support	14
2.1.2 The terms upgrade and migrate defined	15
2.2 Release levels and upgrade or migration	16
2.2.1 Version compatibility.	16
2.2.2 Classic to Classic	16
2.2.3 ES to ES.	18
2.2.4 Classic to ES	20
2.2.5 Unsupported levels of HACMP	22
Chapter 3. Planning for upgrade and migration	23
3.1 Common tasks	23
3.1.1 Current documentation	24
3.1.2 Prior tasks	27
3.1.3 Migration HACMP to HACMP/ES using node-by-node migration.	31
3.1.4 Converting from HACMP Classic to HACMP/ES directly	34
3.1.5 Upgrading HACMP/ES	36
3.1.6 Fall-back plans	37
3.1.7 Acceptance of planned outage	39
3.2 Subsequent actions	40
3.2.1 Changing cluster configuration	40
3.2.2 Saving cluster configuration	40
3.2.3 Cluster verification	40
3.2.4 Error notifications	40
3.2.5 Testing	40
3.2.6 Acceptance of the cluster following migration	41
3.2.7 Expected time to do migration	41
3.3 Sample scenarios for upgrading and migrating HACMP	42
3.3.1 Scenario 1 - A newspaper publisher	42

3.3.2 Scenario 2 - An SAP environment	43
3.3.3 Conclusion	43
3.4 Planning our test case	44
3.4.1 Configurations	44
3.4.2 Our testing list	50
3.4.3 Detailed migration plan	54
Chapter 4. Practical upgrade and migration to HACMP/ES	55
4.1 Prerequisites	55
4.2 Upgrade to HACMP 4.3.1	57
4.2.1 Preparations	57
4.2.2 Steps to upgrade to HACMP 4.3.1	59
4.2.3 Final tasks	62
4.3 Cluster characteristics during migration to HACMP/ES 4.3.1	63
4.3.1 Functionality	63
4.3.2 Daemons, processes, and kernel extensions	65
4.3.3 Log files	68
4.3.4 Cluster lock manager	68
4.4 Node-by-node migration to HACMP/ES 4.3.1	70
4.4.1 Preparations	70
4.4.2 Migration steps to HACMP/ES 4.3.1 for nodes 1..(n-1)	74
4.4.3 Migration steps to HACMP/ES 4.3.1 on the last node n	90
4.4.4 Final tasks	96
4.4.5 Fall-Back procedure	99
4.5 Differences for cluster administration after migration to HACMP/ES .	100
4.5.1 Log files	100
4.5.2 Utilities	101
4.5.3 Miscellaneous	102
Appendix A. Our environment	103
A.1 Hardware and software	103
A.2 Hardware configuration	104
A.3 Cluster screem	105
A.3.1 TCP/IP networks	105
A.3.2 TCP/IP network adapters	106
A.3.3 Cluster resource groups	107
A.4 Cluster spscreem	113
A.4.1 TCP/IP networks	113
A.4.2 TCP/IP network adapters	113
A.4.3 Cluster spscreem resource groups	114
Appendix B. Migration to HACMP/ES 4.3.1 on the RS/6000 SP	117
B.1 General preparations for upgrade and migration	117
B.1.1 Using the working collective (WCOLL) variable	117

B.1.2 Making HACMP install images available from the CWS.....	118
B.1.3 Does PSSP have to be updated?.....	118
B.2 Summary of migrations done on the RS/6000 SP	119
B.2.1 Migration of PSSP	119
Appendix C. Script utilities.....	125
C.1 Test scripts	125
C.2 Scripts used to create the test environment	127
C.3 HACMP script utilities	129
Appendix D. Special notices	133
Appendix E. Related publications	137
E.1 IBM Redbooks publications	137
E.2 IBM Redbooks collections.....	137
E.3 Other Publications.....	137
E.4 Referenced Web sites.....	138
How to get IBM Redbooks	139
IBM Redbooks fax order form	140
Glossary	141
Index	143
IBM Redbooks evaluation	147

Preface

HACMP/ES became available across the entire RS/6000 product range with release 4.3. Previously, it was only supported on the SP2 system. This redbook focuses on migration from HACMP to HACMP/ES using the node-by-node migration feature, which was introduced with HACMP 4.3.1.

This document will be especially useful for system administrators currently running both HACMP and HACMP/ES clusters who may wish to consolidate their environment and move to HACMP/ES at some point in the future. There is a detailed description of a node-by-node migration to HACMP/ES and a comprehensive discussion on how to prepare for an upgrade or migration.

This redbook contains useful information about HACMP in addition to HACMP/ES, which should help a system administrator understand the life-cycle of their high availability (HA) software and the upgrade or migration paths available.

The team that wrote this redbook

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

Ole Conradsen is a project leader at the International Technical Support Organization, Austin center. He has worked with IBM and RS/6000 related products since 1988. Before joining the ITSO in 1999, Ole worked in the AIX solution group in Denmark.

Mitsuyo Kikuchi is an I/T specialist at the IBM Japan Systems Engineering Co. Ltd., in Makuhari, Japan. She has worked at IBM for nine years, has eight years of experience with AIX and RS/6000s, and currently works in the HACMP support team in Japan.

Adam Majchczyk is an I/T specialist within Global Services in IBM Belgium. He has worked for IBM for 18 years in areas of GIS and HPC, and since 1993, with AIX, the RS/6000 SP, and HACMP. He holds an MSc degree in Aeronautics from the University of Rzeszow and is co-author of the redbook *Implementing TME 10 in HA environments*.

Shona Robertson is an AIX systems support specialist in Basingstoke, United Kingdom. She has worked with the RS/6000 brand since joining IBM two years ago and is currently the technical advisor for the HACMP team within the AIX support center in the UK.

Olaf Siebenhaar is a senior I/T Specialist within Global Service in IBM Germany. He has worked at IBM for seven years and has six years of experience in the AIX and SAP R/3 field. His areas of expertise includes HACMP and ADSM especially in conjunction with SAP R/3. He holds a degree in Computer Science from Technical University Dresden.

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

Cindy Barrett, IBM Austin

Bernhard Buehler, IBM Germany

Michael Coffey, IBM Poughkeepsie

John Easton, IBM High Availability Cluster Competency Centre (UK)

David Spurway, IBM High Availability Cluster Competency Centre (UK)

Comments welcome

Your comments are important to us!

We want our Redbooks to be as helpful as possible. Please send us your comments about this or other Redbooks in one of the following ways:

- Fax the evaluation form found in “IBM Redbooks evaluation” on page 147 to the fax number shown on the form.
- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Send your comments in an Internet note to redbook@us.ibm.com

Chapter 1. Introduction

This chapter provides an overview of the evolution of HACMP through main versions and releases of the product and mentions the new functionality and principal enhancements at each release. HACMP/ES is focused on to provide some background for the migration from HACMP to HACMP/ES.

1.1 High Availability (HA)

As an administrator of a high availability (HA) environment, you will be aware that HACMP was developed to address inadequacies of the AIX software and hardware where reliability was of primary importance. HACMP, in its simplest form, provides a mechanism to automatically takeover resources (primarily disks and IP addresses) and to restart applications. In doing so, you can avoid a possibly long outage required to fix hardware or the time to reboot a system in the event of a software failure. In this way, production can be resumed very quickly.

An important issue for setting up an HA environment is the planning phase. Considerable work has been done to make the planning process easier and to try and eliminate sources of error. In addition to the *HACMP V4.3: Planning Guide*, SC23-4277, there is an online cluster worksheet planning program. The worksheets are accessed through a Web browser from a PC based system, and the resulting configuration can be applied directly to a cluster.

Management and maintenance of the cluster, once it is up and running, is another aspect of the HACMP software that has been significantly developed to make it easier to set up and manage a complex HA environment. The redbook *HCAMP V4.3: Administration Guide*, SC23-4279, documents the common administration tasks and tools that exist to help carry out these tasks.

Among the tools that comprise HACMP and HACMP/ES, the following probably contribute most to the manageability of a multi-node environment.

- Version compatibility to support node-by-node migration for upgrade or migration of HA software.
- Monitoring tools, clstat for cluster status, event error notification, the add on, HAView, for use with the Netview product, and, of course, log files.
- Cluster single point of control (C-SPOC), which allows an operation defined on one node to be executed on all nodes in the cluster (to maintain shared logical manager components, users, and groups).

- Dynamic reconfiguration of resources and cluster topology (DARE).
- The verification tool, clverify, which can be run after any changes are made to the cluster to verify that the cluster topology and resource settings are valid.
- The cluster snapshot utility that can be used to save and restore cluster configurations.

HACMP aims, in a well designed cluster, to eliminate any single point of failure, be it disks, network adapters, power supply, processors, or other component. The *HCAMP V4.3 Concepts and Facilities Guide*, SC23-4276, shipped with the HACMP and HACMP/ES products, provides a comprehensive discussion of how high availability is achieved and what is possible in a cluster.

The focus of this redbook, however, is migration to HACMP/ES and, primarily, node-by-node migration from HACMP to HACMP/ES. To give some background to the HACMP to HACMP/ES migration path, this introduction looks at the evolution of HACMP through the enhanced scalable feature HACMP/ES.

1.2 Features and device support in HACMP

HACMP has been in the market place since 1992 when version 1.1 of the high availability product HACMP/6000 was announced. There then followed release 1.2, version 2 (release 2.1), and version 3 (release 3.1 and modification level 3.1.1), all of which were based on AIX 3.2.5. Since this time there have been three releases of HACMP Version 4 (4.1, 4.2, and 4.3), and the product line has been expanded to include HACMP enhanced scalability (ES).

The subsections that follow provide a brief look at each release of HACMP and HACMP/ES. Before this, however, it is useful to know the features that comprise HACMP and HACMP/ES:

HACMP (Classic)

- High availability subsystem (HAS)
- Concurrent resource manager (CRM)
- High availability network file system (HANFS)

HACMP/ES

- Enhanced scalability (ES)

- Enhanced scalability concurrent resource manager (ESCRM)

This simple timeline in the following sections, does not show the additional support that was added for new hardware or protocols at each release. For the administrator of a system, it is, perhaps, the functional improvements to HACMP itself and, therefore, to high availability that are of greater interest.

1.2.0.1 HACMP Version 1.1

This first version of HACMP, released June, 1992, stated that it aimed to bring together cluster processing and high availability to open system client server configurations. It included utilities for the “no single point of failure” concept that underpins HACMP.

The following statement is taken from the IBM announcement letter, ZP92-0325, as is, because it succinctly describes the purpose of the software.

HACMP/6000 software controls the cluster environment behavior. It is also a reporting and management tool for system administrators and/or programmers to use in implementing applications for HACMP/6000 cluster systems or in managing HACMP/6000 systems configurations where mission critical application fallover is a requirement.

In a simple fallover configuration, the cluster manager provides for prompt restart of an application or subsystem on a backup RISC System/6000 processor after there has been a failure of the primary processor. The backup processor can be a dedicated stand-by or actively operational in a load sharing configuration with the primary processor.

- There were three cluster configurations known as mode 1, 2, and 3, although only mode 1 and 2 were available with version 1.1.
 - Mode 1 - Standby mode, where each cluster has an active processor and a standby processor.
 - Mode 2 - Partitioned workload, where each cluster supports different users and can permit each of the two processors to be protected by the other.
- The HAS and CRM features comprised version 1.1.

1.2.0.2 HACMP Version 1.2

HACMP version 1.2 was released in May, 1993 and includes a management information base (MIB) for the simple network management protocol (SNMP) and allows for a cluster to be monitored from a network management console. This introduced the SNMP sub-agent clsmuxpd.

- Monitoring of the cluster is further improved as AIX tracing is now supported for clstrmgr, clinfo, cllockd, and clsmuxpd daemons.
- HACMP 1.2 is extended to use the standard AIX error notification facility.
- Mode 3 configurations became available in version 1.2
 - Mode 3 - Loosely coupled multi processing, which allows two processors to work on same shared workload

1.2.0.3 HACMP Version 2.1

In this version of HACMP, released in December, 1993, system management is done through smit. It is possible to install and configure a cluster from a single processor. Cluster diagnostics and verification of hardware, software, and configuration levels for all nodes in the cluster are included in the product.

- Resources can be disk, volume groups, file systems, networks or applications.
- Cluster configurations are now described in the terms used today.
 - Standby configurations. These are the traditional redundant hardware configurations where one or more standby nodes stands idle.
 - Takeover configurations. In these configurations, all cluster nodes are doing useful work.

Standby and takeover configurations can handle resources in one of the following ways:

- Rotating failover to a standby machine(s)
- Standby or takeover configurations with cascading resource groups
- Concurrent access
- Cluster customizing is enhanced by supporting user pre and post event scripts.
- Concurrent LVM and concurrent access subsystems extended for new disk subsystems (IBM 9333 and IBM 7135-110 Radiant array).

1.2.0.4 HACMP Version 3.1

HACMP Version 3.1 was released in December 1994. Even at this early stage in the product's development, scalability was an issue, and support for 8 RS/6000 nodes in a cluster was added.

- Resource groups were introduced to logically group resources, thus, improving an administrators ability to manage resources.

- Similarly, cascading failover, allowing more than one node to be configured to take over a given resource group, was an additional configuration option introduced at release 3.1.
- There were changes made to the algorithm used for the keep alive subsystem reducing network traffic and an option in smit to allow the administrator to modify the time used to detect when a heartbeat path has failed and, therefore, initiate fail over sooner.

1.2.0.5 HACMP Version 3.1.1

With version 3.1.1 released in December, 1994, HACMP became available on the RS/6000 SP, which increased the application base that might make use of high availability. The SP has a very fast communication subsystem that allows large quantities of data to be transferred quickly between nodes. Customers now have high availability on two different architectures, stand-alone RS/6000s and the RS/6000 SP.

- Support for eight nodes in a cluster on an SP and node failover using the SP high performance switch (HPS) is added.

1.2.0.6 HACMP Version 4.1

Important for cluster administrators HACMP 4.1, released July, 1995, included version compatibility to allow a cluster to be upgraded from an earlier software release without taking the whole cluster off-line (see Section 2.2.1, “Version compatibility” on page 16).

- The CRM feature optionally adds concurrent shared access management for supported disks.
- Support is added for IBMs symmetric multi processor (SMP) machines.

1.2.0.7 HACMP Version 4.1.1

- New features were announced with version 4.1.1 of HACMP in May 1996, namely HANFS, which enhanced the availability of data accessed by NFS, and as a separate product HAGEO, which support HACMP over WAN networks.
- HACMP 4.1.1 also included the introduction of the cluster snapshot utility, which is used extensively in the management of clusters, to clone or recover a cluster, and also for problem determination. Using the cluster snapshot utility, the current cluster configuration can be captured as two simple ASCII files.
- There was a further enhancement to concurrent access to support serial link and serial storage architecture (SSA).

1.2.0.8 HACMP Version 4.2.0

Dynamic Automatic reconfiguration Events (DARE) was introduced in July, 1996, and allows clusters to be modified while up and running. The cluster topology can be changed, a node removed or added to the cluster, network adapters and networks modified, even failover behavior changed while the cluster is active. Single dynamic reconfiguration events can be used to modify resource groups.

- The cluster single point of control (C-SPOC) allows resources to be modified by any node in the cluster regardless of which node currently controls the resource. At this time, C-SPOC is only available for a two node cluster.

C-SPOC commands can be used to administer user accounts and make changes to shared LVM components taking advantage of advances in the underlying logical volume manager (LVM).

Changes resulting from C-SPOC commands do not have to be manually resynchronized across the cluster. This feature is known as “lazy update” and was a significant step forward in reducing downtime for a cluster.

This small section does not do justice to the benefit system administrators derive from these features.

1.2.0.9 HACMP Version 4.2.1

In HACMP Version 4.2.1 introduced in May, 1997, HAView is now packaged with HACMP and extend the TME 10 Net View product to provide monitoring of clusters and cluster components across a network using a graphical interface.

- Enhanced security, which uses kerberos authentication, is now supported with HACMP on the RS/6000 SP.
- Support for HAGEO in version 4.2.1, HACMP Version 4.2.0 did not support HAGEO.

1.2.0.10 HACMP/ES Version 4.2.1

The first release of HACMP/ES, 4.2.1, was announced shortly after HACMP 4.2.1 in September, 1997. It could only be installed on SP nodes and was scalable beyond the eight node limit that exist in HACMP.

- Up to 16 RS/6000 SP nodes in a single cluster were supported with HACMP/ES 4.2.1.
- The heart beating used by the HACMP/ES cluster manager exploited the cluster technology in the parallel system support program (PSSP) on the

SP, specifically Event Management, group services, and topology services.

It was then possible to define new events using the facilities provided, at that time, in PSSP 2.3 and to define customer recovery programs to be run upon detection of such events.

- The ES version of HACMP now uses the same configuration features as the HACMP options; so, migration is possible from HACMP to HACMP/ES without losing any custom configurations.

There were, however, some functional limitations of HACMP/ES, which are listed as follows:

- Fast fallover is not supported.
- C-SPOC is not supported with a cluster greater than eight nodes.
- VSM is not supported with a cluster greater than eight nodes.
- Enhanced Scalability cannot execute in the same cluster as the HAS, CRM, or HANFS for AIX features.
- Concurrent access configurations are not supported.
- Cluster Lock Manager is not supported.
- Dynamic reconfiguration of topology is not supported.
- It is necessary to upgrade all nodes on the cluster at the same time.
- ATM, SOCC, and slip networks are not supported.

As this overview continues, you will see that the functional limitations of HACMP/ES have been addressed. Customers who opted for HACMP/ES identified the benefits as being derived from being able to use the facilities in PSSP.

1.2.0.11 HACMP and HACMP/ES Version 4.2.2

Fast recovery for systems with large and complex disk configurations was introduced in October, 1997. Faster failover was achieved by running recovery phases simultaneously as opposed to serially as they had been in the past. Also, the option was added to run full fsck or logredo for a file system.

- A system administrator using 4.2.2 could evaluate new event scripts with event emulation. Events are now emulated without affecting the cluster and the result reported.
- Dynamic resource movement with the introduction of the `cldare` utility.
- There are significant enhancements to the verification process and the `clverify` command to check name resolution for ip labels used in the cluster configuration, network conflicts where hardware address takeover is configured, lvm, and rotating resource groups.

In addition, `clverify` can now support custom verification methods.

- Support for NFS Version 3 is now included.
- There are still limitations in HACMP/ES as for 4.2.1.
- In March 1998, target mode SSA was announced for HACMP 4.2.2 only.

1.2.0.12 HACMP and HACMP/ES Version 4.3

Custom snapshot methods were introduced in HACMP 4.3.0 in September, 1998, allowing an administrator to capture additional information with a snapshot. The ability to capture a cluster snapshot was added to the HACMP/ES feature.

- C-SPOC support for shared volumes and users or groups was further extended. All nodes in cluster are now immediately informed of any changes; so, failover time is reduced, as “lazy update” will not usually need to be used.
- A task guide was introduced for shared volume groups that aims to reduce the time spent on some administration tasks within the cluster (`smit hacmp -> cluster system management -> task guide for creating a shared volume group`).
- Multiple pre- and post-event scripts are now supported in 4.3.0 extending once more the level to which an HACMP or HACMP/ES cluster events can be tailored.
- HACMP/ES is now available on stand-alone RS/6000s. RISC system cluster technology (RSCT) previously packaged with PSSP is now packaged with HACMP/ES.
- Scalability in HACMP/ES was increased once more to support up to 32 nodes in a SP cluster. Stand-alone system was still limited to eight nodes.
- The CRM feature was added to HACMP/ES. Eight nodes in a cluster are supported for ESCRM.

Note

In October, 1998, VSS support was announced for HACMP 4.1.1, 4.2.2 and 4.3.

1.2.0.13 HACMP and HACMP/ES Version 4.3.1

Support for target mode SSA was added to the HACMP/ES and ESCRM features with version 4.3.1 in July, 1999.

- HACMP 4.3.1 supports the AIX Fast Connect application as a highly available resource, and the clverify utility is enhanced to detect errors in the configuration of AIX Fast Connect with HACMP.
- The dynamic reconfiguration resource migration utility is now accessible through SMIT panels as well as the original command line interface, therefore, making it easier to use.
- There is a new option in smit to allow you to swap, in software, an active service or boot adapter to a standby adapter (smit hacmp -> cluster system management -> swap adapter).
- The location of HACMP log files can now be defined by the user by specifying an alternative pathname via smit (smit hacmp -> cluster system management -> cluster log management).
- Emulation of events for AIX error notification.
- Node-by-node Migration to HACMP/ES is now included to assist customers who are planning to migrate from the HAS or CRM features to the ES or ESCRM features, respectively. Prior to this, conversion from HACMP to HACMP/ES was the only option and required the cluster to be stopped.
- C-SPOC is again enhanced to take advantage of improvements to the LVM with the ability to make a disk known to all nodes and to create new shared or concurrent volume groups.
- An online cluster planning worksheet program is now available and is very easy to use.
- The limitations detailed for HACMP/ES 4.2.1 and 4.2.2 have been addressed with the exception of the following:
 - VSM is not supported with a cluster greater than eight nodes
 - Enhanced Scalability cannot execute in the same cluster as the HAS, CRM, or HANFS for AIX features
 - The forced down function is not supported.
 - SOCC and slip networks are not supported

The following sections look at HACMP/ES in a little more detail.

1.3 Scalability and HACMP/ES

Scalability, support of large clusters and, therefore, large configurations of nodes and potentially disks leads to a requirement to manage “clusters” of nodes. To address management issues and take advantage of new disk

attachment technologies, HACMP enhanced scalable (ES) was released. This was originally only available for the SP where tools were already in place with PSSP to manage larger clusters.

This section mentions the technology that underlies HACMP/ES, RSCT. It might also be of interest to have access to the HACMP/ES documentation. The HACMP/ES 4.3 manuals are available at the Web page:

http://www.rs6000.ibm.com/doc_link/en_US/a_doc_lib/aixgen/hacmp_index.html

1.3.1 RISC System Cluster Technology (RSCT)

The IBM RS/6000 Cluster Technology (RSCT) high availability services provide greater scalability, notify distributed subsystems of software failure, and coordinate recovery and synchronization among all subsystems in the software stack. Packaging these services with HACMP/ES makes it possible to run this software on all RS/6000s, not just on SP nodes.

RSCT Services include the following components:

1.3.1.1 Topology Services

This subsystem is responsible for messaging service, nodes connectivity information and for monitoring network adapter status. All this information is available to the Group Services and, furthermore, to its client's subsystems.

The Topology Services communication infrastructure is based on a heartbeat ring that consists of the node, adapter, and its neighbors on the network. Each node sends a heartbeat message to one of its neighbors and receives a heartbeat from its other neighbor. There is an independent heartbeat ring for every network defined in an HACMP configuration.

The Topology Services daemon whose network interface has the highest IP address becomes the Group Leader and controls the given ring. It receives reports from other members if the loss of heartbeats is detected and forms a new heartbeat ring.

In addition to the heartbeat messages, connectivity messages are sent around on all the heartbeat rings. Connectivity messages for each ring are forwarded to other rings, if necessary, so that all nodes can construct a connectivity graph. This graph determines node connectivity and defines a set of routes that Reliable Messaging can use to send a message between this node and any other node in the cluster.

1.3.1.2 Group Services

Group Services is a system-wide, highly-available facility for coordinating and monitoring changes to the state of an application running on a set of nodes. Group Services helps both in the design and implementation of highly available applications and in the consistent recovery of multiple applications. It accomplishes these two distinct tasks in an integrated framework.

The subsystems that depend on Group Services are called client subsystems. One of these is Event Management, which is the third element of the RSCT structure.

1.3.1.3 Event Management

The function of the Event Management subsystem is to match information about the state of system resources with information about resource conditions that are of interest to client programs, which may include applications, subsystems, and other programs. In this chapter, these client programs are referred to as *EM clients*.

Resource states are represented by resource variables. Resource conditions are represented as expressions that have a syntax that is a subset of the expression syntax of the C programming language.

Resource monitors are programs that observe the state of specific system resources and transform this state into several resource variables. The resource monitors periodically pass these variables to the Event Manager daemon. The Event Manager daemon applies expressions, which have been specified by EM clients, to each resource variable. If the expression is true, an event is generated and sent to the appropriate EM client. EM clients may also query the Event Manager daemon for the current values of resource variables.

Resource variables, resource monitors, and other related information are contained in an Event Management Configuration Database (EMCDB) for the HACMP/ES cluster. This EMCDB is included with RSCT.

1.3.2 Benefits of HACMP/ES

The benefits of HACMP/ES arise from the technologies that underlie its heart beating mechanism supplied in RSCT. HACMP/ES can support up to 32 nodes in a cluster. It can be used to create a large HA cluster and take advantage of, for example, disk technologies, such as storage area networks (SANs), that allow disks to be accessed by a large number of systems without a performance penalty.

The Event Management part of RSCT has been a tool mainly familiar to users of an SP. It is worthwhile taking a little bit of time to become familiar with its potential benefit.

The best sources of information about RSCT is the redbook *HACMP Enhanced Scalability: User-Defined Events*, SG24-5327, and the RSCT publications, *RSCT: Group Services Programming Guide and Reference*, SA22-7355, and *RSCT: Event Management Programming Guide and Reference*, SA22-7354, which are available online from the AIX documentation library:

```
http://www.rs6000.ibm.com/resource/aix_resource/Pubs/  
-> RS/6000 SP (Scalable Parallel) Hardware and Software  
-> PSSP
```

As for the future, we can be assured that HACMP will continue, but HACMP/ES and RSCT are the focus for long term future strategy. Functional limitations of HACMP/ES have been resolved and additional new features lend themselves well to the structures that underlie HACMP/ES.

Chapter 2. Upgrade and migration paths

Maintenance of the software and hardware that comprises a cluster is an integral part of a system administrator's job. This chapter discusses why it is beneficial to upgrade software and looks at the migration paths available for HACMP and HACMP/ES specifically.

2.1 Why upgrade or migrate?

There are a number of management tasks that a system administrator must undertake to make available a computer system for the purpose, ultimately, of generating revenue.

Chapter 1 of the redbook *HACMP Enhanced Scalability: User-Defined Events*, SG24-5327, gives an interesting and comprehensive overview of High Availability clusters in the real world. The author defines a number of different System Management Disciplines including Configuration Management (functions that allow for the planning, maintenance, and development of the resources that make up the computer system) and Change Management (functions that manage and control changes within the computing system).

The need to keep software at reasonably current levels arises partly from the fact that software products are continually being developed to include new features, improve existing features, and also to take advantage of new technology. For example, target mode SSA was introduced in HACMP 4.2.2 (AIX 4.2.1 or later) to allow a serial network to be defined on an SSA loop. It is often the case that a new software release is produced to provide support for new devices. The new features and device support added at each release of HACMP have been considered in detail in Section 1.2, "Features and device support in HACMP" on page 2.

It would be impossible for developers to maintain a code tree for many different releases of a given piece of software and thoroughly test each release. A migration or upgrade path is provided from existing supported release levels to the current release of the software, which will include migration of the current configuration. In this way, the time that has been invested in planning and implementing the cluster is not lost.

Administrators and consultants responsible for the implementation of a cluster should have a long term plan for the environment and be aware of the important dates in a product's life cycle, for example, the planned date for withdrawal of service (development support). It is, in general, easier to

upgrade or migrate while the running software is still supported, thus, taking advantage of a tested upgrade path.

When a new release of a piece of software is made available, a decision needs to be made as to whether it is of benefit to move to the new release at this point in time.

Ultimately, an administrator should maintain good levels of code for stability and to enable the cluster to be modified easily to meet changing business needs.

2.1.1 Fixes and back ported device support

To attain stability in a computing environment, it is usually appropriate to apply a certain amount of software maintenance. This intermediate maintenance will form part of an overall upgrade strategy; so, although it is not directly relevant to migrating HACMP, it is very closely related.

IBM has the concept of Program Temporary Fixes (or PTFs), which are built against a specific software release to fix problems that have been identified during testing or have, subsequently, been reported. Each PTF is associated with a problem description called an Authorized Program Analysis Report or APAR. The APAR database is available and searchable online at the following Web site:

```
http://www.rs6000.ibm.com/support
-> Tech Support Online
   -> Databases
       -> APAR database
```

Note

APARs are labelled either IX or IY followed by five digits.
PTFs, however, start with a U and are followed by six digits.

With the AIX operating system and also for HACMP, there are the concepts of maintenance levels and latest fixes.

- Maintenance levels

Are available for the operating system and for some licensed products (LPPs). These are known, stable levels of code.

To enable a system administrator to establish a strategy for applying fixes, it is important to understand the numbering scheme for the operating system. This is version, release, and modification level. AIX Version 4,

release 3 (4.3,) currently has three modification levels, 4.3.1, 4.3.2, and 4.3.3.

Through time, maintenance levels will be recommended for each modification level of a release of the operating system.

At the time of writing, the second maintenance level for AIX 4.3.2 has been released, 4320-02. It can be downloaded from the Web or ordered through AIX support. To check whether a given maintenance level is installed, the `instfix` command can be run. So, to check for the 4320-02 maintenance level the following command would be executed:

```
# instfix -ik 4320-02_AIX_ML
All filesets for 4320-02_AIX_ML were found.
```

The search string to be used with `instfix` can be found in the APAR description for the maintenance level and in the documentation that accompanies a fix order.

It is possible to move from one modification level of the operating system to another via PTF. So, there is a PTF path from AIX 4.3.0 to either 4.3.1, 4.3.2, or 4.3.3. The PTF supplied for this purpose is also known as a maintenance level.

- Latest fixes

IBM may also release a *latest* fixes package. These packages can contain fixes for a problem that may cause other unresolved problems. Where there are unresolved or pending APARS, it is generally an unusual set of circumstances that leads to additional problems.

It is recommended that latest fixes are installed in the applied state first of all and committed only after a period of time when the administrator is sure the system is stable. If an administrator is concerned about how fixes may impact their environment, then they should only apply maintenance levels.

Support for a new feature may be back-ported to an earlier release of software where it is deemed worthwhile and there is likely to be a take-up of the new feature. Occasionally a customer request results in code being back-ported, for example, support for asynchronous transfer mode (ATM) was back-ported to AIX 4.1.5. Back-ported code is usually supplied via PTFs.

2.1.2 The terms upgrade and migrate defined

The terms upgrade and migrate can, to some extent, be used interchangeably. For the purposes of this redbook, the term upgrade will be used when talking about a move from one release of a given version of HACMP (or HACMP/ES) to another, for example, HACMP 4.3.0 to HACMP

4.3.1 is an upgrade. A move from HACMP 4.2.2 to 4.3.0 or 4.3.1 would, however, be a migration. Similarly, the move from HACMP Classic to HACMP/ES is also a migration.

2.2 Release levels and upgrade or migration

The administrator of an HACMP cluster will be concerned with the level of the operating system in addition to the level of HACMP. Prior to an upgrade of HACMP, the operating system must be at an appropriate level to support the level of HACMP that is to be installed.

The *IBM Application Availability Guide*, which is maintained on the Web, includes compatibility details with AIX Version 4 for AIX software products, support and marketing dates, and information about the latest version or release. It can be found at the following Web site:

<http://www.ibm.com/servers/aix/products/ibmsw/list>

This Web site is intended as a quick reference guide only. Thorough planning, including reference to the release notes and the software installation guide, should always be undertaken prior to any major software update.

The concept of version compatibility is touched on briefly in this section, as it underpins an upgrade or migration with minimal planned downtime. Version compatibility is a part of both HACMP and HACMP/ES.

The remainder of this section summarizes the migration options for each release of HACMP, HACMP/ES, and HACMP to HACMP/ES.

2.2.1 Version compatibility

Version compatibility was introduced with HACMP 4.2 to allow node-by-node migration to a later release of HACMP. During the migration, the cluster will comprise nodes at two different levels of HACMP. Version compatibility supports this cluster state. The rest of the cluster nodes are migrated, one at a time, until all nodes in the cluster are running the same level of HACMP. The concept of version compatibility also applies in an HACMP/ES environment.

2.2.2 Classic to Classic

Table 1 shows, for each release of HACMP Classic, whether fixes or install media is required to upgrade or migrate to subsequent levels of HACMP.

Note

Each of the following options are referred to as HACMP Classic:

High availability subsystem (HAS)

Concurrent resource manager (CRM)

High availability network file system (HANFS)

As most clusters are likely to be upgraded or migrated to the latest available level of software, it is worth noting that it is possible to upgrade or migrate directly from HACMP 4.1.0, 4.1.1, 4.2.0, 4.2.1, 4.2.2, and 4.3.0 to the current version of HACMP 4.3.1.

An upgrade or migration from one version of HACMP Classic to another can be done while the cluster is still in production, as version compatibility supports node-by-node migration. Node-by-node migration incurs a small planned outage when the resources associated with a given node are failed over prior to updating HACMP.

Table 1. Upgrade paths or migration for HACMP Classic

From/To	4.1.1	4.2.0	4.2.1	4.2.2	4.3.0	4.3.1
4.1.0	M	M	M	M	M	M
4.1.1	/	M	M	M	M	M
4.2.0	/	/	P M	P M	M	M
4.2.1	/	/	/	P M	M	M
4.2.2	/	/	/	/	M	M
4.3.0	/	/	/	/	/	P* M

M can migrate to the given level of HACMP via install media.
P can upgrade to the given level of HACMP via PTF (fix package).
*The PTF associated with APAR IY01545 is not yet available as of Oct. 99.

In general, it is possible to move, via fixes, from one modification level to another. From Table 1, you can see that it is possible to move from 4.2.0 to either 4.2.1 or 4.2.2 via PTF. This was not the case for HACMP 4.1.0 to 4.1.1.

As mentioned in the introduction to this section, the administrator of an HACMP cluster will also be concerned with compatibility with AIX. Table 2 records for each release of HACMP compatibility with AIX Version 4. Where the release of HACMP is no longer under defect support, that is, for 4.1.0, 4.1.1, 4.2.0, and 4.2.1, compatibility with the operating system is reported up

to this point. It should also be noted that some of the devices supported by a given level of HACMP may require a later release of the operating system.

Table 2. Compatibility of HACMP with AIX version 4

HACMP Classic	Level of AIX
4.1.0	4.1.3, 4.1.4, 4.1.5
4.1.1	4.1.4, 4.1.5
4.2.0	4.1.4, 4.1.5
4.2.1	4.1.5, 4.2.1, 4.2.2
4.2.2	4.1.5, 4.2.1, 4.2.2 or 4.3.x
4.3.0	4.3.2, 4.3.3
4.3.1	4.3.2, 4.3.3

2.2.2.1 HACMP on the RS/6000 SP

HACMP Classic is supported on the RS/6000 SP. The installation guide for HACMP has an appendix for installing and configuring HACMP on the SP. There are no unique SP filesets, but there is a requirement for a minimum level of Parallel Systems Support Programs (PSSP), which is listed in the appendix. HACMP itself has no dependency on PSSP.

2.2.3 ES to ES

Table 3 on page 19 shows, for HACMP/ES, whether fixes (PTFs) or install media is required to move from a given release of HACMP/ES to another.

Note

Each of the following options are referred to as HACMP/ES:

Enhanced scalability (ES)

Enhanced scalability concurrent resource manager (ESCRM)

For HACMP/ES, there is a direct path for upgrade or migration from 4.2.1, 4.2.2, and 4.3.0 to the current level of HACMP/ES. 4.3.1 provided that a supported version of AIX is installed. Version compatibility supports node-by-node migration from one release of HACMP/ES to a later release.

As can be seen from Table 3, a move between modification levels can be done via fixes, as for HACMP/ES 4.3.0 to 4.3.1; whereas, a move between releases requires install media.

Table 3. Upgrade or migration for HACMP enhanced scalable

From/To	4.2.2	4.3.0	4.3.1
4.2.1	P M PSSP 2.3 or later	M PSSP 3.1 or later	M PSSP 3.1.1 or later
4.2.2	/	M PSSP 3.1 or later	M PSSP 3.1.1 or later
4.3.0	/	/	M* PSSP 3.1.1 or later
<p>M can migrate to the given level of HACMP via install media. P can upgrade to the given level of HACMP via PTF (fix package). *For upgrade from 4.3.0 to 4.3.1, HACMP/ES need to order a media refresh NOTE: PSSP for SP environments only.</p>			

Where HACMP/ES is installed on the SP, there is the added consideration of compatibility with the level of Parallel Systems Support Programs (PSSP) installed. The level of PSSP required to support a given version of HACMP can have implications for upgrade beyond the nodes in the HA cluster.

If the level of PSSP required is higher than that currently installed on the control workstation, it must be upgraded first; however, it is not necessary to update all nodes in the SP frame only those that are to be part of the HA cluster. Issues with PSSP must be resolved before upgrading or migrating HACMP/ES on the SP.

Information about compatibility of PSSP with AIX Version 4 can be found in the *PSSP: Installation and Migration Guide*, GA22-7347. The AIX RS/6000 documentation library can be found online, and the latest release of the PSSP installation and migration guide can be downloaded:

http://www.rs6000.ibm.com/resource/aix_resource/Pubs/
-> RS/6000 SP (Scalable Parallel) Hardware and Software
-> PSSP

The steps to upgrade PSSP are very much configuration dependent, and it is necessary to have a copy of the guide and to take time to plan the migration.

There is a description of what the redbook team did to upgrade PSSP to 3.1.1 from 3.1.0 in Appendix B, "Migration to HACMP/ES 4.3.1 on the RS/6000 SP" on page 117. This gives an example of one upgrade path for PSSP. It is not within the scope of this redbook to consider migration of PSSP in greater

depth. Table 3 does, however, report the minimum level of PSSP required for each release of HACMP/ES on the SP.

For completeness, a table of compatibility of HACMP/ES with the operating system is included below.

Table 4. Compatibility of HACMP enhanced scalable with AIX version 4

HACMP/ES	RS/6000 SP	RS/6000	Level of AIX
4.2.1	Yes	No	4.2.1
4.2.2	Yes	No	4.2.1or 4.3.x
4.3.0	Yes	Yes	4.3.2, 4.3.3
4.3.1	Yes	Yes	4.3.2, 4.3.3

HACMP/ES 4.2.2, and earlier, is not supported on stand-alone RS/6000s. Therefore, Table 3 on page 19 describes the possible migration paths for HACMP/ES assuming that the platform it will run on has been taken into account.

2.2.4 Classic to ES

When migrating from HACMP Classic to HACMP/ES, there is a difference depending on whether the system is an SP or a stand-alone RS/6000.

Let us take a look at the migration table for HACMP to HACMP/ES, Table 5. The table is complete in that it shows whether there is a path from 4.1.0 and later to all releases of HACMP/ES. It is unlikely that an HACMP cluster, which has been in production, will be converted to anything less than 4.3.0 except, perhaps, for development or testing purposes.

Where the effort is being made to move to HACMP/ES, the current release is the most likely choice for migration.

Table 5. Migration from HACMP Classic to HACMP/ES

From Classic /To ES	4.2.1	4.2.2	4.3.0	4.3.1
4.1.0	*(SP only)	*(SP only)	M	M
4.1.1	*(SP only)	*(SP only)	M	M
4.2.0	*(SP only)	*(SPonly)	M	M
4.2.1	M(SP only)	*(SPonly)	M	M

From Classic /To ES	4.2.1	4.2.2	4.3.0	4.3.1
4.2.2	/	M(SP only)	M	M
4.3.0	/	/	M	M
4.3.1	/	/	/	M

M can migrate to the given level of HACMP/ES via install media.
SPonly is reported where the level of HACMP/ES is only available for the SP system.
*Must first upgrade to a comparable level of HACMP, then conversion to HACMP/ES is supported via install media.

There are two methods that can be used to migrate from HACMP Classic to HACMP/ES listed below and described in more detail in Section 3.1.3, “Migration HACMP to HACMP/ES using node-by-node migration” on page 31 and Section 3.1.4, “Converting from HACMP Classic to HACMP/ES directly” on page 34.

- Node-by-node migration from HACMP Classic for AIX to HACMP/ES, 4.3.1 or later.
- Convert from HACMP Classic to HACMP/ES using a snapshot of the Classic cluster.

The facility to convert an HACMP cluster on an SP system to HACMP/ES has been in place for all releases of HACMP/ES and for stand-alone RS/6000s from HACMP/ES 4.3.0. There are, however, limitations on this feature at HACMP/ES 4.2. As can be seen from Table 5, conversion of an HACMP cluster to HACMP/ES 4.2 is possible from the same modification level only.

As the focus of this redbook is migration from HACMP to HACMP/ES, Figure 1 is included to summarize the migration paths available.

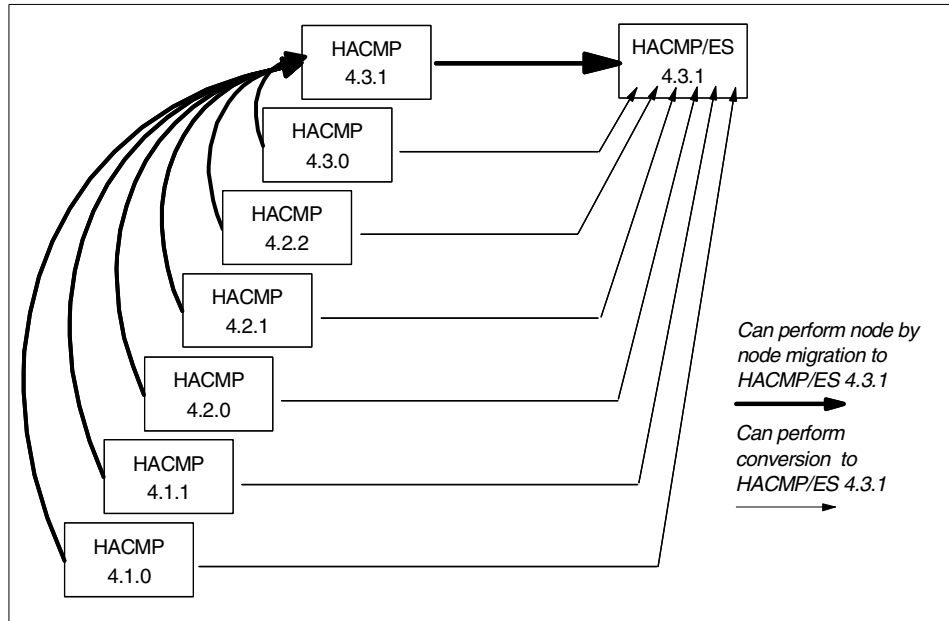


Figure 1. Migration paths from HACMP to HACMP/ES 4.3.1

2.2.5 Unsupported levels of HACMP

The earlier versions of HACMP 1.1, 1.2, 2.1, 3.1, and 3.1.1 are no longer supported, and there is no direct migration path to the current level of HACMP. These earlier versions of the code were based on AIX 3.2.5. HACMP 4.1, the first HACMP release to be based on AIX Version 4, is no longer under defect support, but a migration path exists to the latest level of HACMP and HACMP/ES. If you have a cluster at an unsupported level of HACMP with no migration path, then it must be un-installed prior to installing and configuring a later release. This implies thorough planning as per the *HACMP V4.3 AIX Planning Guide*, SC23-4277, for HACMP or HACMP/ES. Device support is backwards compatible; so, the devices that comprise the cluster will be supported under a more current version of HACMP or HACMP/ES.

Chapter 3. Planning for upgrade and migration

This chapter presents information to help plan for upgrade and migration of an HACMP cluster.

If the working cluster is stable, changing its configuration (including hardware and software) might cause unexpected system downtime or application failure. When you upgrade not only HACMP, but also AIX, this possibility may increase. It is, therefore, important to consider the migration plan carefully in advance.

The system administrator or person undertaking services has to review each cluster environment separately. There are no absolute recommendations about how an upgrade or migration should be done. User requirements for availability, the cost of planned downtime, and management policies must all be taken into account.

The planning phase prior to the upgrade should be comprehensive, and the person doing the upgrade will agree with users and management whether there will be a planned outage or if production is to be kept available throughout the migration process.

3.1 Common tasks

This section describes how to prepare for, and carry out, the upgrade and migration to HACMP/ES 4.3.1. The most important thing that system administrator should keep in mind is that, after the migration, the cluster must be stable and fail over to the backup node successfully. Otherwise, the cluster cannot give highly-available services to the customer any longer.

A checklist like Table 6 below can help you to plan for the migration and carry it out successfully.

Table 6. Helpful planning information

Helpful information	Status
What is the current HACMP version?	
Do you need to upgrade AIX?	
Do you need to upgrade PSSP?	
What applications are currently installed on the cluster? Are they supported on the level of AIX required for HACMP?	

Helpful information	Status
Are there any AIX Error Notifications configured?	
Is planned downtime acceptable?	
Do you have a current cluster diagram?	
Do you have current cluster worksheets?	
Do you need to change the cluster configuration?	
How much free disk space is available on the nodes? Is there enough to migrate?	

3.1.1 Current documentation

In order to perform a successful migration, you need to be familiar with the cluster and have to have documentation of the current HACMP environment.

- System configuration (including hardware and software)
- Cluster configuration

The prerequisite operating system level to run HACMP/ES 4.3.1 is AIX 4.3.2 or later. This means applications working on the cluster must be supported on the level of AIX that will be installed.

3.1.1.1 Cluster worksheet

Cluster worksheets, copies of which are in the *HCAMP V4.3 AIX Planning Guide*, SC23-4277, for HACMP and HACMP/ES, can be used to help understand how the cluster is configured. These worksheets should be filled out whenever an HACMP cluster is configured. If they are missing, you can get the information using the cluster snapshot utility, through smit (`smit hacmp -> Cluster Configurations`) or from HACMP commands (`cllsif`, `clshowres`, and so forth). If there are any cluster customizations or error notifications configured, you will also need to document this information.

If you want to change the cluster configuration in the migration process, you should create new cluster worksheets.

Cluster planning worksheets are provided in the planning guide, but the worksheets are also available online from HACMP 4.3.1. The worksheets are accessible through a browser. This feature is called, Online Cluster Planning Worksheet Program. The sample screen in Figure 2 shows the Online Cluster Worksheet.

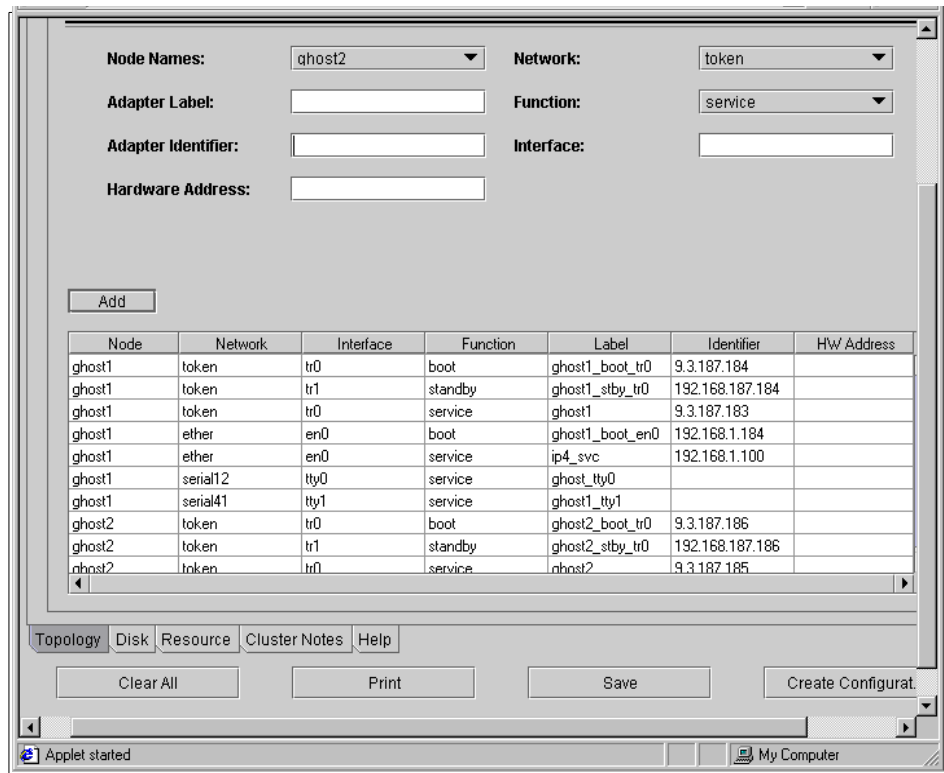


Figure 2. Sample adapter configuration

3.1.1.2 Cluster diagram (State-transition diagram)

One part of the cluster documentation could comprise a state-transition diagram. The state-transition diagram is neither a completely new method nor the ultimate solution for system documentation. It is very often used to describe the behavior of complex systems of which an HACMP cluster undoubtedly is. The behavior within an HACMP environment can be completely described in a state-transition diagram. Compared with the description in text form, state-transition gives an overview and detail at the time. It is a very compact form of visualization.

Two element types are used for the state transition diagram: Circles and arrows. A **circle** is the description of a state with particular attributes. A dark grey circle (or state) means the cluster, or better, the service delivered by the cluster, is in an highly available state. All light grey state represents limited high availability. States drawn in white are not highly available. State-transition diagrams contain only stable states. There are no temporary or intermediate ones.

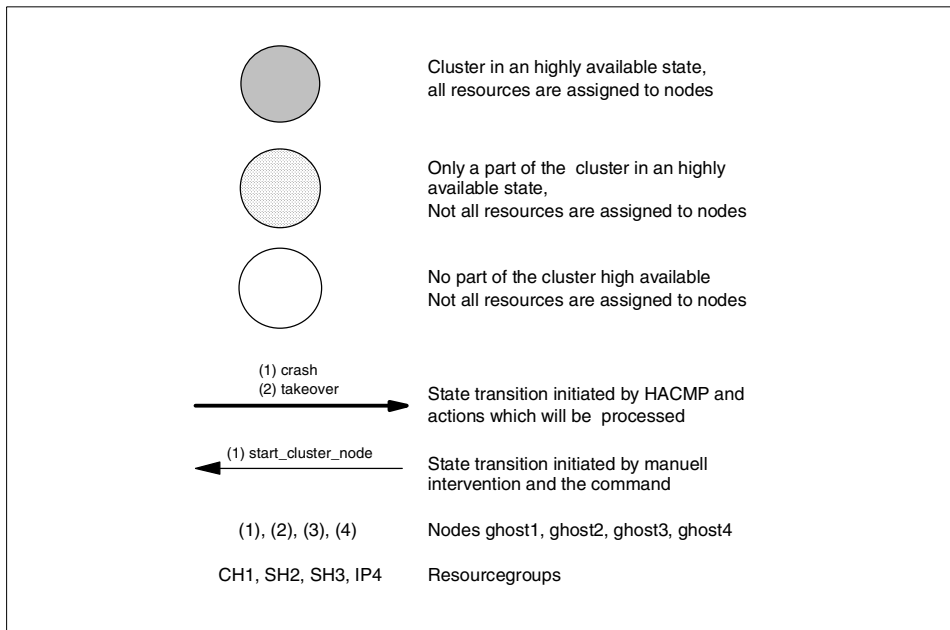


Figure 3. Elements of a state-transition diagram

The second type of element is the **arrow**, which represents a single transition from one state to another. All bold arrows are automatically initiated transitions. Normal drawn arrows are manually initiated transitions. As well as the states, the diagram should contain only relevant transitions.

Some advantages exists if a state-transition diagram is used to document a cluster. The diagram itself has some easy rules to describe behavior that can be checked without any knowledge about the cluster. The following rules apply for the state-transition diagram:

- An arrow must start and end in a circle. There should be at least one arrow from and to every state (no orphaned states).
- All states in the state-transition diagram must be valid according to the user's expectations for the level of service. All required states must be included. The content of the diagram should, therefore, be negotiated with everyone involved, for instance, what impact to the performance a state could have. This can be a basis for a service-level agreement, perhaps between system administrator and user.

A minimal set of test cases are defined in the state-transition diagram. This means that every arrow describes a transition that should be tested independent of whether it is initiated automatically or manually.

In a cluster, there normally exist more possible states and transitions than the diagram contains. But, most of them cover multiple errors at the same time. This seems to be good, but a state and the way to reach it is only highly available if the transition and the state have been tested. The diagram must cover all possible single faults and can cover selected multiple error situations. These multiple error situations are subject to negotiation.

The state-transition diagram should be created and handled intelligently to make the best use of it. Figure 11 on page 51 describes our sample cluster diagram.

3.1.1.3 Snapshots

HACMP has a cluster snapshot utility that enables you to save cluster configurations and additional information, and it allows you to re-create a cluster configuration. Cluster snapshot consists of two text files. It saves all of the HACMP ODM class.

3.1.2 Prior tasks

Prior to an upgrade or migration, there are several things that should be checked:

- If a cluster node diagram or cluster worksheets are available, they must be created before proceeding. If cluster worksheets already exist, confirm that there have been no changes to the cluster.
- Ensure that each node of the cluster has its own HACMP for AIX license.
- Ensure that the cluster is stable.
- If you have applied maintenance for HACMP, be sure that all filesets are committed.
- It is highly recommended to make a full system backup (mksysb and backup any other required data).
- Save your cluster configuration using the cluster snapshot utility.
- Save any customized scripts to a safe directory (that is, other than those associated with HACMP or HACMP/ES).
- Ensure that AIX Error Notifications are configured and save the related information and scripts.
- Ensure that appropriate free disk space is available.

- Create a test plan.

Note

If the level of HACMP in your cluster is pre-version 4.1, you can upgrade your cluster to HACMP 4.1 or greater, then follow the procedure above.

3.1.2.1 Customized scripts

If you have customized any HACMP event scripts, those scripts might be overwritten by reinstalling the new HACMP software. For this reason, it is also recommendable to save a copy of the modified scripts to a safe directory (again, other than those associated with HACMP or HACMP/ES) or media so that you can recover them, if necessary, without having to use your mksysb tape.

3.1.2.2 AIX Error Notification

If any AIX Error Notifications are configured, it is recommended to save information and scripts about them. The information about AIX Error Notifications does not exist in the HACMP ODM but in the AIX ODM. And, it is not saved into a cluster snapshot. If you upgrade AIX, entries of AIX Error Notifications in the AIX ODM might be overwritten. You can see this information through `odmget` command or `smit` (`smit hacmp -> RAS Support -> Error Notification -> Change/Show a Notify Method`).

3.1.2.3 Free disk space

When you migrate the cluster, you will need to have sufficient free disk space on each node in the cluster. This is shown in Table 7.

Table 7. Required free disk space in / and /usr

HACMP migration paths	Component	Required free disk space
Classic to classic 4.3.1	Base	50 MB in /usr
Classic 4.3.1 to ES 4.3.1 (Node-by-node migration)	Base	80 MB in /usr 950 KB in /
Classic to ES 4.3.1	Base	44 MB in /usr 500 KB in /
ES to ES 4.3.1	Base	44 MB in /usr

HACMP migration paths	Component	Required free disk space
Optional to install	HAView	3 MB in /usr
	TaskGuide	400 KB in /usr
	VSM	5.2 MB in /usr
	Japanese msg	1.6 MB in /usr

3.1.2.4 Test plan

After migration, you will need to test the behavior of the cluster. Testing is an important procedure. Imagine that a node crashes after the migration. If takeover fails, the cluster can not give high-availability services to the user any longer. The test plan is a key item to guarantee high availability.

Necessary information to confirm cluster behavior

Deciding what information to gather to test the cluster is an important task. The tests should be comprehensive and, most importantly, repeatable in a consistent manner. There are many sources of information that can be used to indicate cluster stability, some of which are listed below.

- Cluster daemons
- Cluster log files
- clstat utility
- Network addresses
- Shared volume groups
- Routing tables
- Application daemons(process)
- Application behaviors
- Application log files
- Access from the clients
- Resource group status, and so forth

Test cases

A test plan must be developed for each individual cluster. This test plan can then be used after any change to the cluster environment including changes to the operating system. If the cluster configuration is modified, any existing test plan should be reviewed. So, the question is what kind of tests do you need to do? Table 8 is a high level view of some of the things related to HACMP that can be tested. The majority of the items listed below should be tested. It is impossible to be sure that the cluster will fail over correctly and provide high availability if it hasn't been thoroughly tested. The test cases

below in the column *Necessary* should be included in a test plan as a minimum requirement.

Table 8. General testing cases

Testing case	Necessary	Recommendable
Start cluster services	X	X
Stop cluster services	X	X
Service adapter failure	X	X
Standby adapter failure		X
Node failure	X	X
Node reintegration	X	X
Serial connection failure		X
Network failure	X	X
Error notification		X
Other customizations		X
SPS adapter failure*		X
SPS failure*		X
*This is for SP only.		

3.1.3 Migration HACMP to HACMP/ES using node-by-node migration

When you migrate an existing HACMP cluster using node-by-node migration, there are two steps you have to perform:

- Upgrade to HACMP Classic 4.3.1
- Node-by-node migration from HACMP Classic to HACMP/ES.

The HACMP upgrade is required because node-by-node migration is only supported between the same levels of HACMP and HACMP/ES. If your cluster already installed HACMP Classic 4.3.1, you can perform node-by-node migration directly.

Note

The cluster should not be left in a mixed state for long.

3.1.3.1 Upgrade HACMP to version 4.3.1

If you use node-by-node migration from HACMP to HACMP/ES, you must upgrade HACMP to version 4.3.1. This means you also need to have the cluster installed AIX 4.3.2 or later (and if on SP, PSSP 3.1.1 or later). If the current level of the cluster is 4.3.0 or lower, it must be upgraded or migrated to HACMP 4.3.1. See the *HACMP V4.3 AIX: Installation Guide, SC23-4278*, for more detail.

Case1: You need to upgrade AIX

1. Back up all required information to be able to recover the system or reapply customizations and take a cluster snapshot of the original classic cluster.
2. For the node to be upgraded stop the cluster services using the graceful with takeover option and takeover resources to a backup node.
3. Upgrade AIX to release 4.3.2 or later using the Migration Installation option.
4. If using SCSI disks, ensure that adapter SCSI IDs for the shared disk are not the same on each node.
5. Install HACMP 4.3.1 on the node.
6. Start cluster services on the node and make sure it has reintegrated into the cluster successfully.
7. Repeat step 2 through 7 on remaining cluster nodes, one at a time.
8. Restore the custom scripts if you have to.

9. Synchronize and verify the cluster configuration.
10. Test fail over behavior.

Case2: You do NOT need to upgrade AIX

1. Back up all required information to be able to recover the system or reapply customizations and take a cluster snapshot of the original classic cluster.
2. For the node to be upgraded, stop the cluster services using the graceful with takeover option and takeover resources to a backup node.
3. Install HACMP 4.3.1 on the node.
4. Start cluster services on the node and make sure it has reintegrated into the cluster successfully.
5. Repeat step 2 through 5 on remaining cluster nodes, one at a time.
6. Restore the custom scripts if you have to.
7. Synchronize and verify the cluster configuration.
8. Test fail over behavior.

The `cl_convert` utility

The `cl_convert` utility converts cluster and node configurations from earlier releases of HACMP or HACMP/ES to the release that is currently being installed.

The syntax for `cl_convert` is as follows.

```
cl_convert -v [version] [-F]
```

Where there is an existing cluster configuration, `cl_convert` will be run automatically during an install of HACMP or HACMP/ES, with the exception of a migration from HACMP 4.1.0. When migrating from HACMP 4.1.0, `cl_convert` must be run manually from the command line. This is clearly documented in the installation guide at each release.

During the conversion, `cl_convert` creates new ODM data from the ODM classes previously saved during the install of HACMP or HACMP/ES. What is done depends on the level of HACMP the conversion is started from. The man pages for `cl_convert` describe how the ODM is manipulated in a little more detail.

3.1.3.2 Node-by-node migration from HACMP to HACMP/ES

As mentioned in the previous chapter, node-by-node migration is a new feature of HACMP 4.3.1, and we focus on it to migrate HACMP/ES 4.3.1 in

this redbook. The migration can be done with minimal downtime due to version compatibility. Specifically, during the fail over to a backup node before migration and upon reintegration, the cluster service will be suspended for a short period of time.

When using node-by-node migration, there are several things to be aware of:

- Node-by-node migration from HACMP to HACMP/ES is supported only between the same modification levels of HACMP and HACMP/ES software, that is, from HACMP 4.3.1. to HACMP/ES 4.3.1.
- The nodes need a minimum of 64 MB memory to run HACMP Classic and HACMP/ES simultaneously. 128 MB of RAM is recommended.
- Don't migrate or upgrade clusters that have HAGEO for AIX installed into the cluster as HACMP/ES does not support HAGEO.
- If you have more than three nodes, and they have serial connections, ensure that they are connected each other. If multiple serial networks connect to one node, you must change this to a daisy chain as you can see in Figure 4.

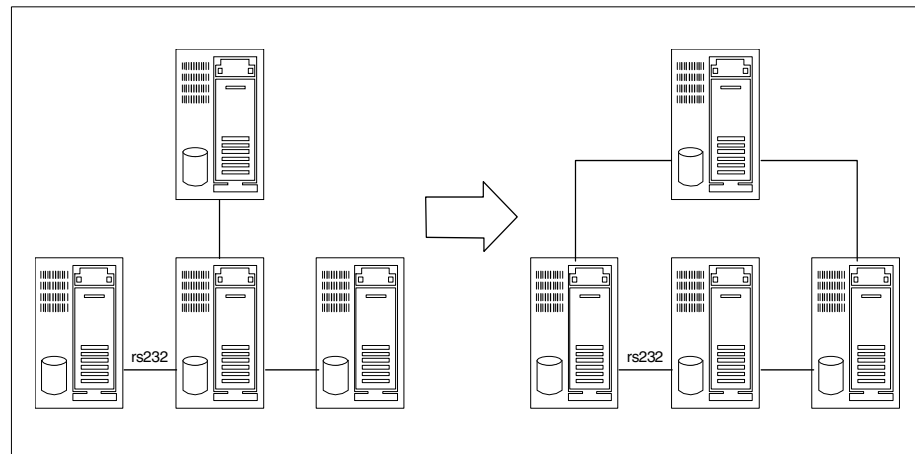


Figure 4. Serial cabling connection for HACMP and HACMP/ES

Node-by-node migration to HACMP/ES, starting from a comparable level of HACMP (4.3.1 or later), is detailed below. A practical migration using this method is described in Chapter 4, “Practical upgrade and migration to HACMP/ES” on page 55.

1. Back up all required information to be able to recover the system.
2. Save the current configuration in a snapshot to the safe directory.

3. Stop the HACMP on the node that you will migrate using the graceful with takeover option and takeover to the backup node.
4. Install HACMP/ES 4.3.1 on the node.
5. After the successful installing, reboot the node.
6. Start HACMP on the node, at which time all migrated nodes will be running both HACMP and HACMP/ES daemons. Only HACMP Classic can control the cluster events and resources at this moment.
7. Repeat steps 2 through 5 for other nodes in the cluster.

On reintegration of the last node in the cluster to be migrated, an event is automatically run that hands control to HACMP/ES from HACMP. When the migration is complete, HACMP Classic is un-installed automatically, and all the cluster nodes are up and running HACMP/ES.

Of course, it is always possible to reinstall any cluster from scratch but this is not considered a migration. Both migration paths listed above are valid. Their benefits will depend on working practices.

Note

Changes to the cluster configuration can be done once the migration has been completed.

3.1.4 Converting from HACMP Classic to HACMP/ES directly

You can convert from HACMP Classic to HACMP/ES using a snapshot from the Classic cluster. In this case, it is one step from Classic to enhanced scalable, excluding a possible operating system upgrade. Planned downtime is required for conversion to HACMP/ES.

The steps to convert from HACMP 4.1, or later, to any release of HACMP/ES are listed below. This assumes that the platform supports the release of HACMP/ES that is to be installed.

Since this procedure can't use rolling migration, the cluster service is not available during the conversion. If the cluster is allowed to have system downtime, you can choose this migration method.

Case1: You need to upgrade AIX

1. Back up all required information to be able to recover the system or reapply customizations, and take a cluster snapshot of the original classic cluster.

2. Un-install HACMP software.
3. Upgrade AIX to 4.3.2 or later using the Migration Installation option.
4. If using SCSI disks, ensure that adapter SCSI IDs for the shared disk are not the same on each node.
5. Install HACMP/ES 4.3.1.
6. Repeat 1 through 5 on remaining nodes.
7. Convert the saved snapshot from HACMP Classic to HACMP/ES format using `clconvert_snapshot` on one node.
8. Reinstall saved customized event scripts on all nodes.
9. Apply the converted snapshot on the node. The data in existing HACMP ODM classes are overwritten in this step.
10. Verify the cluster configuration from one node.
11. Reboot all cluster nodes and clients.
12. Test fail over behavior.

Case2: You do NOT need to upgrade AIX

1. Back up all required information to be able to recover the system or reapply customizations, and take a cluster snapshot of the original classic cluster.
2. Remove HACMP classic software.
3. Install HACMP/ES 4.3.1.
4. Reboot the node or refresh the inetd.
5. Repeat 1 through 4 on remaining nodes.
6. Convert the saved snapshot from HACMP classic to HACMP/ES format using `clconvert_snapshot` on one node.
7. Reinstall saved customized event scripts on all nodes.
8. Apply the converted snapshot on the node. The data existing HACMP ODM classes on all nodes are overwritten.
9. Verify the cluster configuration from one node.
10. Reboot all cluster nodes and clients.
11. Test fail over behavior.

The clconvert_snapshot utility

The `clconvert_snapshot` utility is used to convert a cluster snapshot where there is a migration from HACMP to HACMP/ES.

The syntax for `clconvert_snapshot` is as follows:

```
clconvert_snapshot [-C] -v [release] -s [snapshot file]
```

The man pages for `clconvert_snapshot` are fairly comprehensive. The `-c` flag is available with HACMP/ES only and is used to convert an HACMP snapshot. It should not be used on a HACMP/ES cluster snapshot.

As with the `cl_convert` utility, `clconvert_snapshot` populates the ODM with new entries based on those from the original cluster snapshot. It replaces the supplied snapshot file with the result of the conversion; so, the process cannot be reverted. A copy of the original snapshot should be kept safely somewhere else.

For rolling migration, HACMP 4.3.1 and later, the `clconvert_snapshot` command will be run automatically.

3.1.5 Upgrading HACMP/ES

For cluster currently installed with HACMP/ES, the migration method will be to upgrade HACMP/ES directly. A rolling upgrade can be used to upgrade an existing HACMP/ES cluster to 4.3.1. This enables you to minimize system downtime. You cannot leave the cluster in a mixed version for long periods. There is less guarantee of high availability. As we already mentioned, HACMP/ES 4.3.1 requires AIX 4.3.2, and if your cluster nodes are on an RS/6000 SP, PSSP 3.1.1 as well. So, if you need to upgrade AIX or PSSP, the cluster services are temporarily suspended.

Note

It is strongly recommended to upgrade all nodes in a cluster immediately.

Case1: You need to upgrade AIX or PSSP

1. Back up all required information to be able to recover the system or reapply customizations and take a cluster snapshot of the original classic cluster.
2. Commit your current HACMP/ES software on all nodes.
3. Stop HACMP/ES with the gracefully with takeover option on one node and takeover its resources to a backup node.
4. Upgrade AIX to 4.3.2 or later using the Migration Installation option.
5. If using SCSI disks, ensure that adapter SCSI IDs for the shared disk are not the same on each node.

6. Upgrade PSSP to V3.1.1 or later if you need to upgrade.
7. Install HACMP/ES 4.3.1 on the node.
8. The `cl_convert` utility will automatically convert the previous configuration to HACMP/ES 4.3.1 as part of the installation procedure.
9. Reboot the node.
10. Start HACMP/ES and verify the node joins the cluster.
11. Repeat 1 through 10 on remaining cluster nodes, one at a time.
12. Synchronize the cluster configuration.

3.1.6 Fall-back plans

It is advisable to have a fall-back plan in case you decide not to continue the migration process or if a problem is encountered during migration. If you updated AIX to 4.3.2 or later, the simplest method might be to restore from `mksysb`. When you upgrade or migrate only HACMP software, you can use the following methods to back-out changes to the cluster.

3.1.6.1 Fall-back during an HACMP upgrade

During the upgrade process, you can return to previous level of HACMP by reinstalling the original software, for example, if you have a two-node cluster and have completed the upgrade of one node. To recover the cluster, you could take the following actions:

1. If cluster service has started, stop the service with the graceful option.
2. De install HACMP from the node that has already upgraded.
3. Reinstall the previous level HACMP that had been installed before on the node.
4. Check that the appropriate level of HACMP is successfully installed.
5. Restore the cluster snapshot saved before upgrade.
6. Verify the cluster configuration from the node that has not upgraded.

3.1.6.2 Fall-back after a completed HACMP upgrade

After upgrading all nodes of the cluster, you can recover the cluster using the same method as Section 3.1.6.1, “Fall-back during an HACMP upgrade” on page 37, but the simplest method might be to restore from `mksysb`.

1. If cluster services has started, stop the service with the graceful option.
2. Un-install HACMP from the node.

3. Reinstall previous level HACMP that had been installed before on the node.
4. Check that the appropriate level of HACMP is successfully installed.
5. Reboot the node.
6. Repeat the 1 through 5 until all of the nodes have fallen back.
7. Restore the cluster snapshot saved before upgrade on one node.
8. Verify the cluster configuration from the node.
9. Start the cluster services on all of the nodes.

3.1.6.3 Fall-back during a node-by-node migration

The start of HACMP cluster services on the last node in the cluster is the turning point in the migration process, as it initiates the actual migration to HACMP/ES. At any time before HACMP is started on the last node in the cluster, you can use the following steps to back-out:

1. On the nodes that have both HACMP and HACMP/ES installed, if the cluster services have been started, stop the service using the graceful with takeover option and failover its resources to another node.
2. Un-install the HACMP/ES software. Be sure not to remove software shared with the HACMP classic.
 - man pages
for example,
cluster.man.en_US.client.data
cluster.man.en_US.cspoc.data
cluster.man.en_US.server.data
 - C-SPOC messages
for example,
cluster.msg.en_US.cspoc
3. Start cluster service on this node and reintegrate it into the cluster.
4. If you have other nodes to be backed out, repeat steps 1 through 3 until HACMP/ES has been removed from all nodes in the cluster.

If the cluster services have been started on the last node in the cluster, the migration to ES must be allowed to complete. The process to back out from this position is detailed in next section.

3.1.6.4 Fall-back after a completed node-by-node migration

Once you have migrated the last node of the cluster from HACMP Classic 4.3.1 to HACMP/ES 4.3.1, the fall-back process is to reinstall the HACMP and apply a saved cluster snapshot. The HACMP Classic software has already been removed in the migration process. In this case, you cannot revert to a Classic cluster node-by-node like a migration. All the cluster services must be suspended during the fall-back procedure. Similarly, restoration from a mksysb backup is an option, but cluster services must be stopped for it to be used.

1. If cluster services has started, stop the service with the graceful option.
2. Remove HACMP/ES 4.3.1 from the node.
3. Remove /usr/es directory. (This directory could be left.)
4. Reinstall HACMP Classic 4.3.1 on the node.
5. Check that the HACMP is successfully installed.
6. Reboot the node.
7. Repeat the 1 through 5 until all of the nodes have fallen back.
8. Restore the cluster snapshot saved before node-by-node migration on one node.
9. Verify the cluster configuration from the node.
10. Start the cluster services on all of the nodes.

3.1.7 Acceptance of planned outage

If planned downtime is allowed for the whole cluster, some migration steps could be done in parallel among the nodes. Otherwise, you might use node-by-node migration to reduce the system downtime.

3.2 Subsequent actions

This section describes tasks that should be done after migration.

3.2.1 Changing cluster configuration

When you want to change cluster configuration, you can do that after the migration. Please keep in mind not to change configuration during the migration process.

3.2.2 Saving cluster configuration

When you have changed the cluster configuration (including upgrade and migration), it is recommended to save a cluster snapshot using the cluster snapshot utility.

3.2.3 Cluster verification

After migration is complete, you may verify the cluster configuration. The cluster verification utility, `/usr/sbin/cluster/utilities/clverify`, verifies that HACMP-specific modifications to AIX system files are correct and that the cluster and its resources are configured correctly.

3.2.4 Error notifications

If you have configured any AIX Error Notifications on your cluster and upgraded AIX, you need to confirm whether those entries remained on each node. Because the error notification utility is not based on HACMP, but on AIX, by upgrading AIX, the entries on the AIX ODM might be deleted. You can see this information through the smit panel (`smit hacmp -> RAS Support -> Error Notification -> Change/Show a Notify Method`) or the `odmget` command. If you cannot find them, you must configure again.

3.2.5 Testing

Testing the behavior of the cluster is required whenever you change its configuration. As we already mentioned in this chapter, it is the most important thing that the cluster can give the high availability services continuously to the end-users after any configuration change. To ensure this, testing based on the testing plan is needed.

If the testing result is different from your testing plan, you need to try to find the cause and resolve it. The *HACMP troubleshooting Guide*, SC23-4280, and *HACMP Enhanced Scalability Installation and Administration Guide*, Vol.2, SC23-4306, will help you.

3.2.6 Acceptance of the cluster following migration

The final step of the migration process is to have the cluster accepted following the migration. The work to upgrade the cluster will ultimately be signed off by management. The person carrying out the migration will be expected to explain what has been done and prove that the cluster is working as it should, testing complete. This work will have been done according to the migration plan that was created prior to the migration and would also have been agreed on with management. The level of interaction with management will depend largely on whether the work being done is part of a contract requiring on site services.

3.2.7 Expected time to do migration

Prior to migration, it is helpful to estimate the total time for the migration. This will help you and the customer to schedule any planned outage and allocate the necessary resources, including people and machines. The time for migration depends on several factors, which are outlined in the next sections.

3.2.7.1 HACMP migration paths

The paths to upgrade and migrate to HACMP/ES 4.3.1 are different depending on the level of HACMP that is currently installed. For example, if the nodes in the cluster already have HACMP Classic 4.3.1 installed, you only need to migrate them to HACMP/ES 4.3.1. In this case, you can use node-by-node migration, which is a new feature of HACMP 4.3.1, and we will describe the steps in Chapter 4, “Practical upgrade and migration to HACMP/ES” on page 55.

On the other hand, if the nodes in your cluster are running HACMP Classic 4.3.0 or lower level, you must first upgrade them to HACMP Classic 4.3.1 to take advantage of node-by-node migration to ES.

Of course, you can migrate without using node-by-node migration to ES 4.3.1. Refer to Chapter 2, “Upgrade and migration paths” on page 13 about migration paths for further detail.

3.2.7.2 RS/6000 SPs

Where nodes in the cluster to be migrated are SP nodes, you may have SP specific work, such as migrating PSSP software. Section 2.2, “Release levels and upgrade or migration” on page 16 talks about when an upgrade of PSSP might be required and details an actual upgrade of PSSP for the migration process.

3.2.7.3 The number of nodes

The greater the number of nodes in the cluster, the more time will be needed for the migration process. Migration work, testing, and so on, will have to be done for each node in the cluster. When you perform node-by-node migration or rolling HACMP migration, the migration time is a multiple of the time to migrate one node.

3.3 Sample scenarios for upgrading and migrating HACMP

To illustrate some of the factors which may influence the decision over which migration route to choose let us consider two example environments.

3.3.1 Scenario 1 - A newspaper publisher

The publisher of a newspaper has an online service available via the World Wide Web. The online version is accessed by many thousands of readers worldwide. HACMP has been employed make the Web servers that provide this service highly available.

The HACMP environment comprises two nodes with HACMP configured for AIX 4.1.5 is installed, as all RS/6000 systems used by the publisher were upgraded earlier in the year to the latest release of the operating system. HACMP, however, is at release 4.2.2.

Due to fierce competition the newspaper has been experiencing lately, it has been decided that the high-profile, online version should be accessible as close to 100 percent of the time as possible for the next three months. It is also policy that the systems run supported code; so, HACMP is to be migrated to the latest level of code. All other clusters the publisher has run on the SP; so, a decision has been made to consolidate the HA products in production and move to HACMP/ES in this migration.

With these policies, the system administrator has opted for a node-by-node migration. There is a small risk that service may be interrupted during the migration, but this has been highlighted and accepted by management. Good prior planning will minimize this risk.

In this scenario, three migration steps are required:

- Upgrade AIX from version 4.1.5 to 4.3.2 or later
- Migrate from HACMP 4.2.2 to HACMP 4.3.1 (node-by-node supported by version compatibility)
- Migrate from HACMP 4.3.1 to HACMP/ES 4.3.1 (using the node-by-node migration facility)

3.3.2 Scenario 2 - An SAP environment

This scenario includes a typical two node SAP environment, which is configured for mutual takeover where there are approximately 200 shared disks, which comprise a cascading resource group for the main instance of SAP, and a second cascading resource group say only five disks in it. The cluster has AIX 4.3.1 installed and HACMP 4.2.2, and a decision has been taken to move to HACMP/ES 4.3.1 and, therefore, AIX 4.3.2.

Let us say that there are approximately 500 users in this environment located at 10 different customer sites in four countries. The cluster is utilized beyond normal working hours, as there is a time zone difference for the users of this system in the four countries. In this scenario, the main production hours are from Monday morning to late Friday evening.

The system administrator has negotiated a maintenance window over a weekend. No users will be allowed access to the machines during this period. As the maintenance window is outside of the main production hours, the administrator has decided to convert the cluster from HACMP to HACMP/ES. With the cluster unavailable, some of the work to migrate the cluster can, to some extent, be done in parallel.

Failover of the main SAP instance might take up to 20 minutes; so, if the system administrator had chosen to do a node-by-node migration there would be at least this length of downtime.

In this scenario, the system administrator must perform the following migrations or upgrades during the maintenance window:

- Upgrade AIX from 4.3.1 to 4.3.2
- Migrate from HACMP 4.2.2 to HACMP 4.3.1
- Migrate from HACMP 4.3.1 to HACMP/ES 4.3.1 (converting the Classic cluster)

3.3.3 Conclusion

The two simple scenarios described in 3.3.1 and 3.3.2 show how implementations of HACMP can vary. The HACMP and HACMP/ES products have been developed to give system administrators a great deal of control over how a cluster is managed and, therefore, support multiple approaches to upgrading and migrating. The following chapters of this redbook describe the migration experience of the redbook team.

3.4 Planning our test case

Provided here is the documentation of the test cases we tested and described in Chapter 4, “Practical upgrade and migration to HACMP/ES” on page 55. Our final aim was migrating the nodes of the cluster to HACMP/ES 4.3.1 using node-by-node migration, which has been announced from HACMP 4.3.1. We configured the test clusters below. We have tested and migrated successfully in two cases.

- 4-node cluster (standalone RS/6000) named “Scream”.
- 2-node cluster (RS/6000 SP) named “spscream”.

You can see the detailed hardware and software configurations in Appendix A, “Our environment” on page 103.

3.4.1 Configurations

In this section, we will briefly describe the cluster configurations that have been used as test cases.

3.4.1.1 4-node cluster (stand-alone RS/6000)

We configured the 4-node cluster as can be seen in Figure 5 on page 45. Generally, the cluster nodes have been divided into two groups. One is a concurrent group between ghost1 and ghost2, and the other is a cascading group between ghost3 and ghost4. We have defined four resource groups for this cluster as can be seen in the Table 9 on page 45.

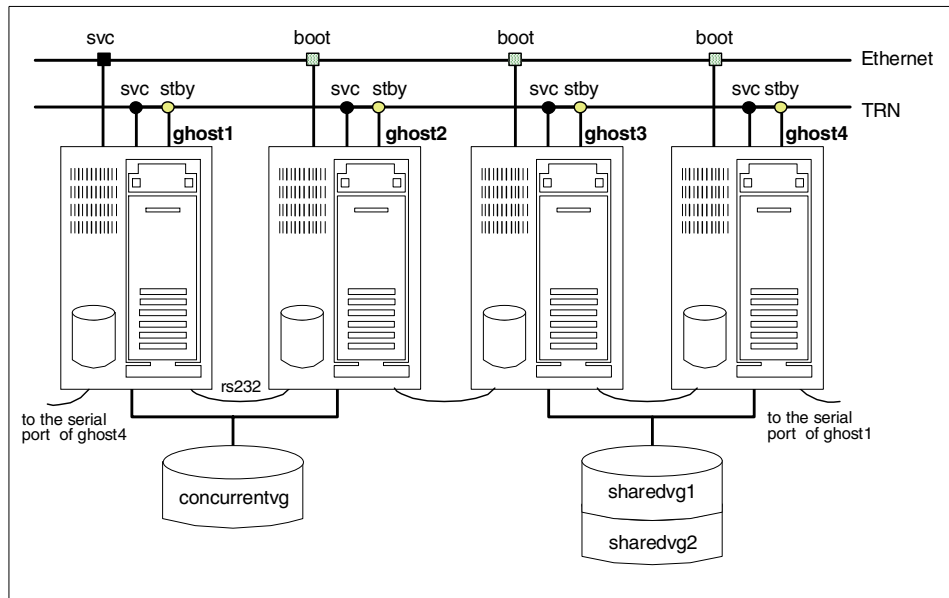


Figure 5. Hardware configuration for cluster Scream

Table 9. Resource groups defined for cluster Scream

Resource group	Node relationship	Participating node
CH1	concurrent	ghost1 ghost2
SH2	cascading	ghost3 ghost4
SH3	cascading	ghost4 ghost3
IP4	rotating	ghost1 ghost2 ghost3 ghost4

CH1 is a concurrent resource group shared between ghost1 and ghost2. SH2 and SH3 are cascading resource groups shared between ghost3 and ghost4. These resource groups control IP address and shared volume groups as resources to takeover (mutual takeover). IP4 is a rotating resource group shared with all four nodes. IP4 contains only an IP address only.

3.4.1.2 2-node cluster (RS/6000 SP)

We also configured a 2-node cluster on SP. Figure 6 and Table 10 represent the cluster configurations.

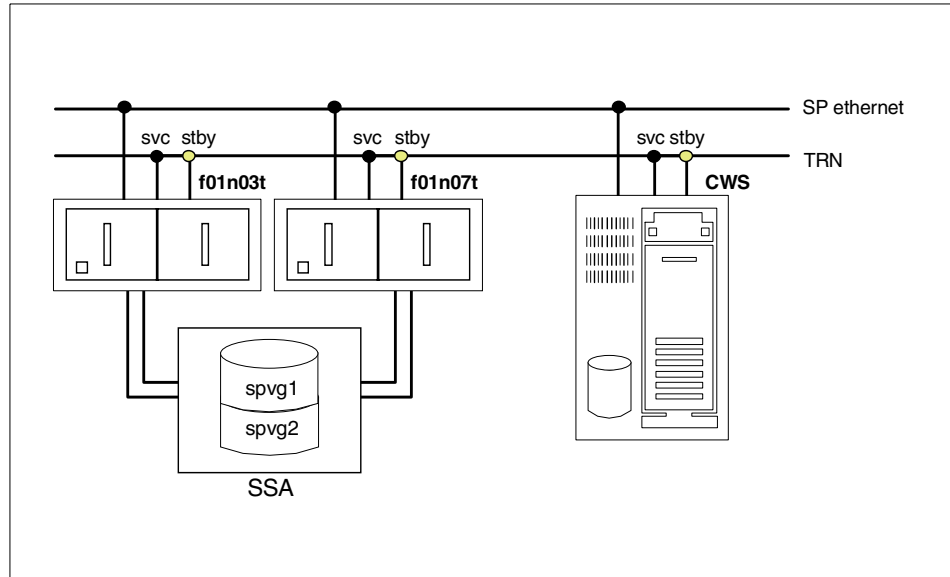


Figure 6. Hardware configuration for cluster spscream

Table 10. Resource groups defined for cluster spscream

Resource group	Node relationship	Participating node
VG1	cascading	f01n03t f01n07t
VG2	cascading	f01n07t f01n03t

To plan and configure the cluster we took advantage of the Online Cluster Worksheet Program which is one of the new features of HACMP 4.3.1. This utility, which run under Microsoft Internet Explorer, allowed us to fully define the configuration of a cluster through the worksheets and generate the final configuration file based on the input.

Note

Please note that there is no verification being performed at this stage.

The configuration file was further transferred to one of the cluster nodes and converted to the actual cluster configuration by the following command:

```
/usr/sbin/cluster/utilities/cl_opsconfig < filename >
```

```
# Upload configuration
```

This performed a cluster synchronization and verification of the configuration. The following figures (Figure 7, 8, and 9) are sample worksheets which we used in our cluster:

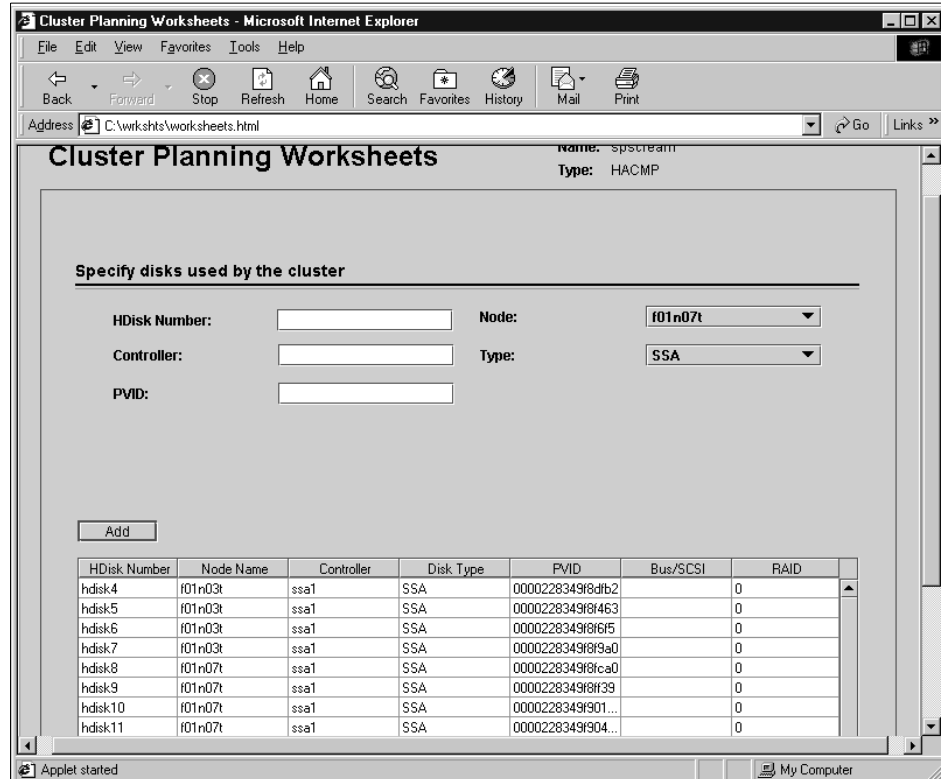


Figure 7. Disks configuration for cluster spscream

Cluster Planning Worksheets Name: spscream
Type: HACMP

Specify network adapter information for the nodes

Node Names: Network:
 Adapter Label: Function:
 Adapter Identifier: Interface:
 Hardware Address:

Node	Network	Interface	Function	Label	Identifier	HW Address
f01n03t	sptoken1	tr0	service	f01n03t	9.3.187.246	
f01n03t	sptoken2	tr1	service	f01n03t	11.3.187.246	
f01n07t	sptoken1		service	f01n07t	9.3.187.250	
f01n07t	sptoken2		service	f01n07t	11.3.187.250	

Figure 8. Adapters configuration for cluster spscream

Cluster Planning Worksheets Name: spscream
Type: HACMP

Specify logical volumes for previously specified volume groups

Logical Volume Name: Volume Group:
 Logical Volume Type: Separate Physical Volumes: Yes
 Size: No
 Logical Partition Copies:
 Filesystem Mount Point:

Name	Volume Group	Type	Size	Copies	Separate Phys. V...	Mount Point
logsplv1	spvg1	JFS Log		1	Yes	
splv1	spvg1	JFS	3276800	1	Yes	/spfs1
logsplv2	spvg2	JFS Log		1	Yes	
splv2	spvg2	JFS	3276800	1	Yes	/spfs2

Figure 9. Logical volumes for cluster spscream

3.4.1.3 Migration paths

In the migration test and examples, we used AIX 4.3.2 as the OS level and HACMP 4.3.0 as the starting level for our migration on both clusters. The reason for the choice was that AIX 4.3.2 is the newest generally available AIX level (4.3.3 is not installed at customers yet), and we assumed that most customers who are going to migrate to HACMP/ES are already on HACMP release 4.3.0.

In this case, the migration step had two phases.

1. Upgrading HACMP from 4.3.0 to 4.3.1.
 - Before upgrading HACMP, we tested cluster behavior to ensure that it worked.
2. Node-by-node migration to HACMP/ES 4.3.1.

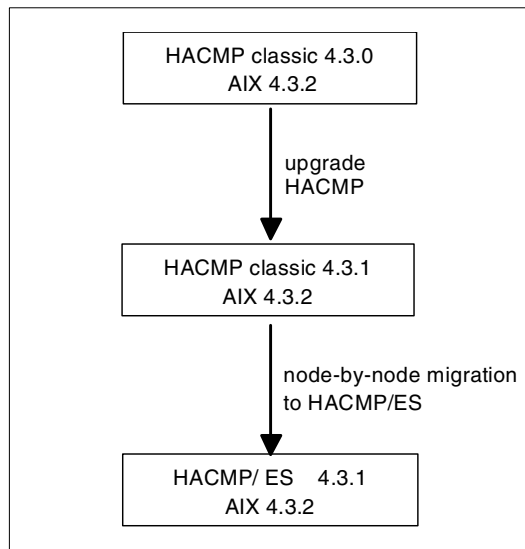


Figure 10. Migration steps of our testcases

3.4.2 Our testing list

The planned tests are listed in Table 11. For each of the listed test items we tested the function before and after the occurrence of six events.

We also checked the state of resource groups on how they should be before and after the trigger.

We tested the function after the following events:

- Start cluster services
- Stop cluster services
- Service adapter failure
- Node failure
- Node reintegration
- Serial connection failure

For each event above, we made a function test on the items listed in Table 11. Everything was working.

Table 11. Testing features of testcase

Testing feature	nonSP	SP
IPAT	OK	OK
HWAT*	OK	-
Disk takeover	OK	OK
Application server	OK	OK
Post-event scripts	OK	-
Changing NIM	OK	OK
Error notifications	-	OK
Cascading resource group	OK	-
Rotating resource group	OK	-
Concurrent resource group	OK	-
clockd	OK	-
*Hardware address takeover - Test not available		

3.4.2.1 Cluster diagram for cluster Scream

Figure 11 is the cluster diagram we used as a testcase. As we explained in the previous section, the cluster diagram describes cluster behavior pictorially. Based on this diagram, we created a testing list.

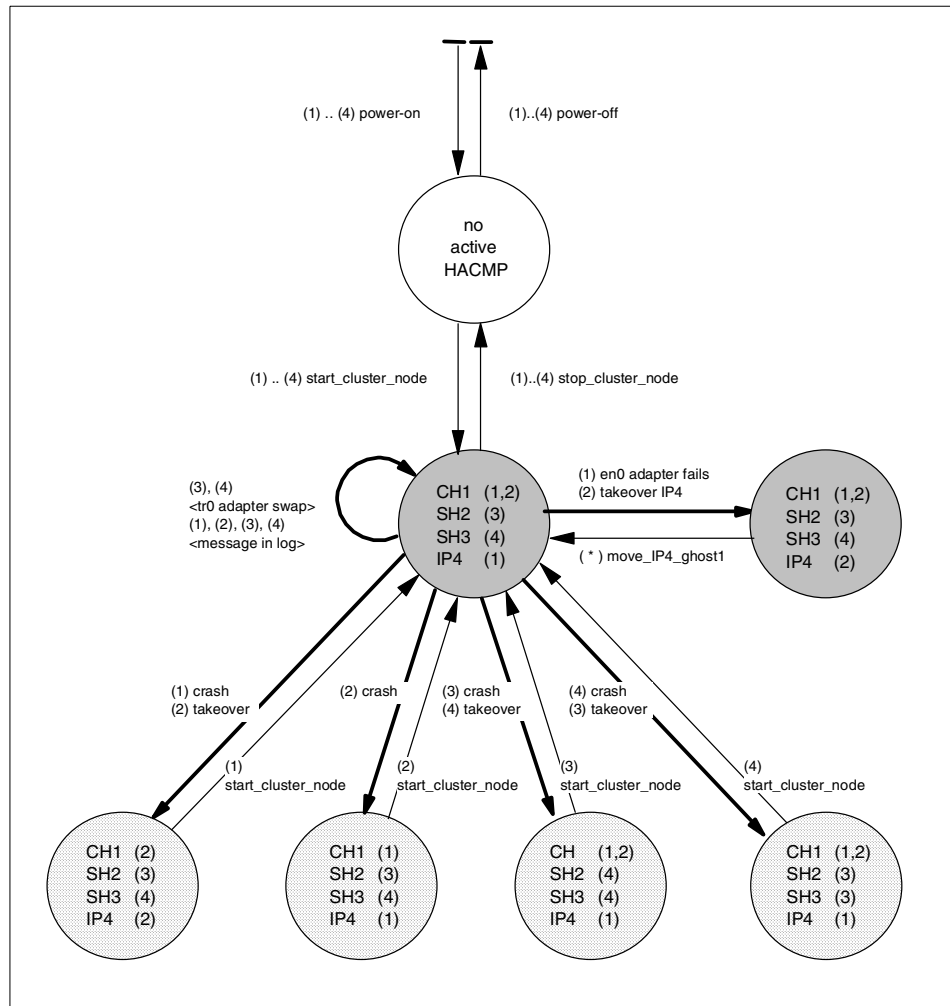


Figure 11. State transition-Diagram of our test environment

3.4.2.2 Testing list

You can see our testing list in Table 12. The columns *State before test* and *State after test* show which node should have resource groups. The column *Command, Reason, Initiator* shows the trigger of its test case.

Table 12. Testing list for cluster *Screamr*

Test case	State before test	Command, Reason; Initiator	State after test
1	no program active	start_cluster_node on all nodes; manually	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
2	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	tr0 fails on a single node (3), (4); HACMP swaps adapter	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
3	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	ghost1 crashed; HACMP takeover IP4	ghost1: --- ghost2: CH1, IP4 ghost3: SH2 ghost4: SH3
4	ghost1: --- ghost2: CH1, IP4 ghost3: SH2 ghost4: SH3	start_cluster_node on ghost1; manually	ghost1: CH1 ghost2: CH1, IP4 ghost3: SH2 ghost4: SH3
5	ghost1: CH1 ghost2: CH1, IP4 ghost3: SH2 ghost4: SH3	ghost2 crashed; HACMP takeover IP4	ghost1: CH1, IP4 ghost2: --- ghost3: SH2 ghost4: SH3
6	ghost1: CH1, IP4 ghost2: --- ghost3: SH2 ghost4: SH3	start_cluster_node on ghost2; manually	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
7	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	ghost3 crashed; HACMP takeover SH2	ghost1: CH1, IP4 ghost2: CH1 ghost3: --- ghost4: SH2, SH3, IP4
8	ghost1: CH1, IP4 ghost2: CH1 ghost3: --- ghost4: SH2, SH3	start_cluster_node on ghost3; manually	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3

Test case	State before test	Command, Reason; Initiator	State after test
9	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	ghost4 crashed; HACMP takeover SH3	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2, SH3 ghost4: ---
10	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2, SH3 ghost4: ---	start_cluster_node on ghost4; manually	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
11 ¹⁾	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	en0 fails on ghost1 (1);	ghost1: CH1 ghost2: CH1,IP4 ghost3: SH2 ghost4: SH3
12 ¹⁾	ghost1: CH1 ghost2: CH1,IP4 ghost3: SH2 ghost4: SH3	move_IP4_ghost1 (2) takeover IP4 manually	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
13	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	tty fails on a single node (1), (2), (3), (4); message in cluster log	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
14	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	tty repaired on a single node (1), (2), (3), (4); message in cluster log	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3
15	ghost1: CH1, IP4 ghost2: CH1 ghost3: SH2 ghost4: SH3	stop_cluster_node on all nodes; manually	no program active
1) Test cases 11 and 12 can't be run in a mixed cluster and in an HACMP/ES post-event script because cldare is not working there.			

3.4.3 Detailed migration plan

After the general migration path for the cluster has been selected all steps along them should also be planned carefully. A good starting point are the test cases which have hopefully been defined.

Table 13 is an example of the steps that have been used to migrate our 4-node cluster. You can see the actual steps in detail in Chapter 4, “Practical upgrade and migration to HACMP/ES” on page 55.

Table 13. Migration steps for cluster Scream

Migration step	Node	Action
1	ghost4	Take over resource group SH3 to ghost3
2	ghost4	Migrate to HACMP/ES 4.3.1
3	ghost4	Take over resource group SH3 back from ghost3
4	ghost2	Stop resource group CH1
5	ghost2	Migrate to HACMP/ES 4.3.1
6	ghost2	Start resource group CH1
7	ghost3	Take over resource group SH2 to ghost4
8	ghost3	Migrate to HACMP/ES 4.3.1
9	ghost3	Take over resource group SH2 back from ghost4
10	ghost1	Stop resource group CH1
11	ghost1	Migrate to HACMP/ES 4.3.1
12	ghost1	Start resource group CH1
13	all nodes	Run test cases 2 to 14

Chapter 4. Practical upgrade and migration to HACMP/ES

This chapter describes migration scenarios from HACMP to HACMP/ES. It includes a discussion of an HACMP upgrade, as it may be necessary to get HACMP to a level of code, which allows node-by-node migration to HACMP/ES. This chapter mentions briefly the option to reinstall HACMP/ES but not in detail, as this is not a migration or upgrade.

4.1 Prerequisites

Modifications in an HACMP cluster can have a great impact on the level of high availability; so, you must plan all actions very carefully.

Before starting the practical steps of the migration, certain prerequisites must have already been met. Which ones are dependent on the current cluster environment and the expectations of everyone involved. These topics are discussed in Chapter 3, “Planning for upgrade and migration” on page 23 in more detail. Prior to the upgrade or migration, the following should be available for inspection:

- A project plan including detail of the migration plan
- An agreement with management and users about planned outage of service
- A Fall-Back-Plan
- A list of test cases
- Required software and licenses
- Documentation of current cluster

If some items are missing, you need to prepare them before continuing.

The current cluster should be in a suitable state because we assume your cluster is working before the migration. Otherwise, it's required to review all components and the cluster configuration according to the *HCAMP V4.3 AIX:Installation Guide, SC23-4278*.

It is a good idea to carry out the test plan before migration. If all test cases have been carried out successfully, the current cluster state can be considered a safe point for recovery. If the cluster fails during a test in the test plan, it gives you the chance to correct the cluster configuration before making any further changes by updating the HA software.

HACMP 4.3.1 requires AIX version 4.3.2. If not all cluster nodes are on this software level, they must be upgraded beforehand. The actions for this task should be performed according to the *HCAMP V4.3 AIX: Installation Guide*, SC23-4278.

Additional sources of the most up-to-date information are the release notes for a piece of software. It is highly recommended that these are read carefully. For HACMP and HACMP/ES, the release notes are located as follows:

```
/usr/lpp/cluster/doc/release_notes  
/usr/es/lpp/cluster/doc/release_notes
```

The stand-alone cluster that has been used as an example in this redbook contains four resource groups, each with different behavior. Every resource group should be an equivalent for a service. To check the service availability, we have developed a simple Korn-Shell script called

`/usr/scripts/cust/itso/check_resgr`, the output of which can be seen in Figure 12. The output shows that all services are available and assigned to their primary locations.

```
> checking ressourcegroup CH1  
  iplabel ghost1  
  on node ghost1  
  logical volume lvcon01.01  
  iplabel ghost2  
  on node ghost2  
  logical volume lvcon01.01  
> checking ressourcegroup SH2  
  iplabel ghost3  
  on node ghost3  
  filesystem /shared.01/lv.01  
  iplabel ghost3  
  on node ghost3  
  application server 1  
> checking ressourcegroup SH3  
  iplabel ghost4  
  on node ghost4  
  filesystem /shared.02/lv.01  
  iplabel ghost4  
  on node ghost4  
  application server 2  
> checking ressourcegroup IP4  
  iplabel ip4_svc  
  on node ghost1
```

Figure 12. Output from `check_resgr` command

The complete code of all self-developed scripts which we have used for this IBM Redbook are contained in Appendix C, “Script utilities” on page 125.

4.2 Upgrade to HACMP 4.3.1

The migration to HACMP/ES for our stand-alone cluster started on HACMP 4.3.0 with AIX 4.3.2. If there are older releases of either software component installed in your cluster environment, other steps might be required. Refer to Chapter 2, “Upgrade and migration paths” on page 13 for information about other releases.

4.2.1 Preparations

All the following tasks must be done as user “root”.

If there are Twin-Tailed SCSI Disks used in the cluster, check the SCSI IDs for all adapters. Only one adapter and no disk should use the reserved SCSI ID 7. Usage of this ID could cause an error during specific AIX installation steps. You can use the commands below to check for a SCSI conflict.

```
lsdev -Cc adapter | grep scsi | awk '{ print $1 }' | \
xargs -i lsattr -El {} | grep id # show scsi id used
```

Any new ID assigned must be currently unused. All used SCSI IDs can be checked and the adapter ID changed if necessary. The next example shows how to get the SCSI ID for all SCSI devices and then change the ID for SCSI0 to ID 5.

```
lsdev -Cs scsi
chdev -l scsi0 -a id=5 -P # apply change only in ODM
```

A subsequent reboot of all changed nodes is needed and should be completed before further installation steps are carried out.

The installation of HACMP 4.3.1 requires at least 51 MB available free space in directory /usr on every cluster node. This available space can be checked with:

```
df -k /usr # check free space
```

If there is not enough space in /usr, the file system can be extended or some space freed. To add one logical partition, you can use the command below. It tries to add 512 bytes, but the Logical Volume Manager extends a file system with at least one logical partition.

```
chfs -a size=+1 /usr # add 1 logical partition
```

Any file set updates (maintenance) should be in a committed state on all nodes before the installation of HACMP 4.3.1 is started.

```
installp -cgX all 2>$1 | tee /tmp/commit.log # commit all file sets
```

To have the chance to go back if something fails during the following installation steps, it is necessary to have a current system backup. At least rootvg should be backed up on bootable media, for instance a tape or image on a network install server.

```
mksysb -i /dev/rmt0 # backup rootvg
```

Customer specific pre- and post-event scripts will never be overwritten if they are stored separately from the lpp directory tree; however, any HACMP scripts that have been modified directly must be saved to another safe directory. Installing a new release of HACMP will replace standard scripts. Even file set updates or fixes may result in HACMP scripts being overwritten. By having a copy of the modified scripts, you have a reference for manually reapplying modifications in the new scripts. In addition to the backup of rootvg, it is useful to have an extra backup of the cluster software, which will include all modifications to scripts.

```
cd /usr/sbin/cluster
tar -cvf /tmp/hacmp430.tar ./* # backup HACMP
```

During the migration, all nodes will be restarted. To avoid any inadmissible cluster states, it is highly recommended to remove any automatic cluster starts while the system boots.

```
/usr/sbin/cluster/utilities/clstop -y -R -gr # only modify inittab
```

This command doesn't stop HACMP on the node it modifies only /etc/inittab to prevent HACMP from starting at the next reboot.

Note

Cluster verification will fail until all nodes have been upgraded. Any forced synchronization could destroy the cluster configuration.

The preparation tasks discussed above should have been executed on all nodes before continuing with the upgrade steps.

The installable file sets for HACMP 4.3.1 are delivered on CD. It is possible for easier remote handling to copy all file sets to a file system on one server, which will be mounted from every cluster node.

```
mkdir /usr/cdrom # copy lpp to disk
mount -rv cdrfs /dev/cd0 /usr/cdrom
cp -rp /usr/cdrom/usr/sys/inst.images/* /usr/sys/inst.images
umount /usr/cdrom
cd /usr/sys/inst.images
```

```
/usr/sbin/inutoc .  
mknfsexp -d /usr/sys/inst.images -t rw -r ghost1,ghost2,ghost3,ghost4 -B
```

4.2.2 Steps to upgrade to HACMP 4.3.1

Before starting with the installation of any HACMP file sets, the node has to be restarted without HACMP resource groups. For the selection of the node(s), the migration plan should be used to avoid an unnecessary outages. The resource groups of every required node must be taken over:

```
/usr/sbin/cluster/utilities/clstop -y -N -s -gr      # shutdown with takeover  
lssrc -g cluster                                  # check daemons stopped
```

Note

Please keep in mind that any takeover causes outage of service.

How long this takes depends on the cluster configuration. The current location of cascading or rotating resources and the restart time of the application servers that belong to the resources groups will be important.

The current cluster state can be checked by issuing the following command:

```
/usr/sbin/cluster/clstat -a -c 20                  # check cluster state
```

All resources must be available as expected before continuing; otherwise, subsequent actions could cause an unplanned outage of service. In our cluster, for instance, the state after stopping HACMP on node ghost4 is shown in Figure 13 and Figure 14.

```

                                clstat - HACMP for AIX Cluster Status Monitor
                                -----
Cluster: Scream (20)                Mon Oct 11 10:38:47 CDT 1999
      State: UP                      Nodes: 4
      SubState: STABLE
      Interface: ghost3_tty1 (3)      Address: 0.0.0.0
                                       State: UP
      Interface: ghost3 (5)           Address: 9.3.187.187
                                       State: UP

      Node: ghost4                   State: DOWN
      Interface: ghost4_boot_en0 (0)  Address: 192.168.1.190
                                       State: DOWN
      Interface: ghost4_tty0 (3)      Address: 0.0.0.0
                                       State: DOWN
      Interface: ghost4_tty1 (4)      Address: 0.0.0.0
                                       State: DOWN
      Interface: ghost4_boot_tr0 (5)  Address: 9.3.187.190
                                       State: DOWN

```

Figure 13. Cluster state after a stop of HACMP on node ghost4

/usr/scripts/cust/itso/check_resgr

```

> checking ressourcegroup CH1
  iplabel ghost1
  on node ghost1
  logical volume lvconc01.01
  iplabel ghost2
  on node ghost2
  logical volume lvconc01.01
> checking ressourcegroup SH2
  iplabel ghost3
  on node ghost3
  filesystem /shared.01/lv.01
  iplabel ghost3
  on node ghost3
  application server 1
> checking ressourcegroup SH3
  iplabel ghost4
  on node ghost3                 <- after ghost4 takeover or crashed
  filesystem /shared.02/lv.01
  iplabel ghost4
  on node ghost3                 <- after ghost4 takeover or crashed
  application server 2
> checking ressourcegroup IP4
  iplabel ip4_svc
  on node ghost1

```

Figure 14. Output from check_resgr command

If the install images were copied to a file system on ghost4 and NFS exported as described in previous section, the file sets can be accessed and installed on a cluster node from ghost4 as follow:

```
mount ghost4:/usr/sys/inst.images /mnt
cd /mnt/hacmp431classic
smitty update_all
-> INPUT device: .
-> SOFTWARE to update: _update_all
-> PREVIEW only? : no
lppchk -v
```

Reboot the node to bring all new executables and libraries into effect.

```
shutdown -Fr # reboot node
```

After the node has come up some checks are useful to be sure that the node will become a stable member of the cluster. Check for example the AIX error log and physical volumes.

```
errpt | pg # check AIX error log
lspv # check physical volumes
```

Once the node has come up, and assuming there are no problems, then HACMP can be started on it.

```
/usr/sbin/cluster/etc/rc.cluster -boot -N -i # start without cllockd
/usr/sbin/cluster/etc/rc.cluster -boot -N -i -l # start with cllockd
```

Note

Please keep in mind that the start of HACMP on a node that is configured in an cascading resource group can cause outage of service.

Whether this occurs and how long it takes depends on the cluster configuration. The current location of the cascading resources and the restart time of the application servers that belong to the resources groups will be important.

After the start of HACMP has completed, there are some additional checks that are useful to make sure that the node is a stable member of the cluster. For instance, check HACMP logs and cluster state.

```
tail -f /var/adm/cluster.log | pg # check cluster log
more /tmp/hacmp.out # check event script log
/usr/sbin/cluster/clstat -a -c 20 # check cluster state
```

These upgrade steps must be repeated for every cluster node until all nodes are installed at HACMP 4.3.1.

4.2.3 Final tasks

If any HACMP event scripts had been modified during the installation or customizing of the previous HACMP level, some additional tasks are necessary. Where cluster event scripts have been modified, and you want to keep these modifications, identify, at first, all of them. Then, select the new scripts that are suitable for reapplying the modifications. If the migration was just from 4.3.0 to 4.3.1, it should not be too complicated. Customer event scripts are never subject to change during any installation or migration. Some scripts will be saved by the `installp` command to the directory `/usr/lpp/save.config/usr/sbin/cluster` during the installation process.

Because there were some changes made in the cluster, a cluster verification should be made also if the only change is to the underlying software. This is the first verification of the cluster configuration on the new HACMP level. If there have been any changes to the cluster, for instance, to resource groups, it is very important to do a system verification, and the results should be read very carefully.

```
/usr/sbin/cluster/diag/clverify software lpp          # verify software  
/usr/sbin/cluster/diag/clverify cluster config all   # verify configuration
```

4.3 Cluster characteristics during migration to HACMP/ES 4.3.1

A cluster can be described as having a type, according to the software product(s), installed on its nodes. There are three possible types:

HACMP cluster	HACMP is installed on all cluster nodes.
Mixed cluster	One or more, but not all cluster nodes have both HACMP and HACMP/ES installed. A mixed cluster exists only during a unfinished migration process.
HACMP/ES cluster	HACMP/ES is installed on all cluster nodes.

Every cluster type can be further defined in terms of functionality, daemons, processes, kernel extensions and log files. These items will be discussed in this chapter.

4.3.1 Functionality

Within a mixed cluster, the RSCT technology is installed and running on the migrated nodes but has no influence on how the cluster reacts. While the cluster is in a mixed state, prior to starting HACMP on the last node to be migrated, the HACMP cluster manager and event scripts are active with all HACMP/ES event scripts suppressed.

During node-by-node migration it is possible to maintain availability of all resource groups if the migration planning has been done carefully; however, not all administration tools will be working. Among the common administration tools is dynamic reconfiguration that does not work in a mixed cluster. This fact must be taken into consideration when the migration plan is created.

If the `cldare` utility (providing dynamic reconsideration) is used in a customer-specific event script, it could partly disable the high availability of a cluster during migration. These scripts will fail during the whole migration period. The usage of `cldare` in an HACMP/ES event script should be avoided because the `cldare` ES requires a stable cluster to be able to work.

The result of the failure is an error message that appears on the standard output of the `cldare` utility. If this stream has been redirected to `/dev/null` in the scripts, the malfunction could stay undetected (see Figure 15 and Figure 16).

```
/usr/sbin/cluster/utilities/cldare -M IP4:stop:sticky # stop resgroup  
/usr/sbin/cluster/utilities/cldare -M IP4:ghost1 # activate resgr
```

```

cldare: Migration from HACMP to HACMP/ES detected.
A DARE event cannot be run until the migration has completed.

cldare: Migration from HACMP to HACMP/ES detected.
A DARE event cannot be run until the migration has completed.

```

Figure 15. Output from cldare command in a mixed cluster

```

Oct 29 10:46:42 ghost1 HACMP for AIX: EVENT START: network_down ghost1 ether
Oct 29 10:46:44 ghost1 :
Oct 29 10:46:44 ghost1 : > Move ressourcegroup IP4 to node ghost2.
Oct 29 10:46:44 ghost1 :
Oct 29 10:49:24 ghost1 clstrmgr[10340]: Fri Oct 29 10:49:24 SetupRefresh:
Cluster must be STABLE for dare processing
Oct 29 10:49:24 ghost1 clstrmgr[10340]: Fri Oct 29 10:49:24 SrcRefresh: Error encountered
DARE is ignored
Oct 29 10:49:25 ghost1 HACMP for AIX: EVENT COMPLETED: network_down ghost1 ether
Oct 29 10:49:26 ghost1 HACMP for AIX: EVENT START: network_down_completeghost1ether
Oct 29 10:49:27 ghost1 HACMP for AIX: EVENT COMPLETED: network_down_completeghost1ether
Oct 29 10:49:40 ghost1 HACMP for AIX: EVENT START: reconfig_resource_release
Oct 29 10:50:11 ghost1 HACMP for AIX: EVENT COMPLETED: reconfig_resource_release
Oct 29 10:50:12 ghost1 HACMP for AIX: EVENT START: reconfig_resource_acquire
Oct 29 10:50:18 ghost1 HACMP for AIX: EVENT COMPLETED: reconfig_resource_acquire
Oct 29 10:50:53 ghost1 HACMP for AIX: EVENT START: reconfig_resource_complete
Oct 29 10:51:04 ghost1 HACMP for AIX: EVENT COMPLETED: reconfig_resource_complete

```

Figure 16. cldare command used in post-event script (cl_event_network_down)

Not all old and new components in a mixed cluster are fully integrated. The functionality, which is based on the new technology (RSCT), is only available on already migrated nodes.

For example, the output of the clRGinfo command can contain incomplete information in a mixed cluster state. In Figure 17, all nodes except ghost1 have been migrated. The label ip4_svc (192.168.1.100) seems to be unavailable when using clRGinfo, but it is available as can be seen in clstat.

```

/usr/es/sbin/cluster/utilities/clRGinfo IP4           # check resource group
unset DISPLAY
/usr/es/sbin/cluster/clstat -c 20                   # check cluster state

```

```

-----
Group Name                State      Location
-----
IP4                        OFFLINE   ghost1
                          OFFLINE   ghost2
                          OFFLINE   ghost3
                          OFFLINE   ghost4
-----

clstat - HACMP for AIX Cluster Status Monitor
-----

Cluster: Scream (20)      Tue Oct 19 09:45:20 CDT 1999
      State: UP          Nodes: 4
      SubState: STABLE
Node: ghost1             State: UP
  Interface: ip4_svc (0)   Address: 192.168.1.100
                          State: UP
  Interface: ghost1_tty1 (1) Address: 0.0.0.0
                          State: UP
  Interface: ghost1_tty0 (4) Address: 0.0.0.0
                          State: UP
  Interface: ghost1 (5)   Address: 9.3.187.183
                          State: UP

```

Figure 17. Incomplete resource information

Table 14. Functionality of the different cluster types

Functionality	HACMP	Mixed cluster	HACMP/ES
Event scripts	used	used	
Event scripts ES		suppressed	used
Customer event scripts	used	used	used
Dynamic reconfiguration	cldare		cldare ¹⁾
clRGinfo		incomplete	complete
clfindres		incomplete after ODMDIR= /etc/es/objrepos	complete
1) cldare works only in a STABLE cluster.			

4.3.2 Daemons, processes, and kernel extensions

The technology underlying heartbeat in HACMP/ES is different to that in HACMP. HACMP/ES uses topology and group services, which are briefly described in Section 1.3.1, “RISC System Cluster Technology (RSCT)” on page 10 and supplied via RSCT on stand-alone RS/6000. HACMP/ES has

been running on RS/6000 SP s for several years, but for people who are familiar with HACMP on stand-alone nodes, there are new daemons and processes to be aware of after migration to HACMP/ES.

Table 16 gives an short overview of both HACMP and HACMP/ES daemons and which you can expect to see running on the different cluster types. Non-migrated nodes are still running the same daemons running as in an HACMP cluster. Migration of an intermediate node (1..n-1) brings up all HACMP/ES daemons in addition to those for HACMP. Migration of the last node in the cluster leaves only the HACMP/ES daemons running.

For problem determination, it is useful to have another running system for comparisons, but, most often, you will not have access to such a system. Figure 18 and Figure 19 are process lists from a running cluster. This will give you an idea of how the actual process list should look. These figures show only HACMP relevant daemons; other processes may exist.

```

root 6010 3360 0 Oct 31 - 33:28 /usr/es/sbin/cluster/clstmgr
root 8590 3360 0 Oct 31 - 46:27 hagsd grpsvcs
root 8812 3360 0 Oct 31 - 0:06 haemd HACMP 1 Scream SECNOSUPPORT
root 9042 3360 0 Oct 31 - 0:28 hagsglsmd grpglsm
root 9304 3360 0 Oct 31 - 238:52 /usr/sbin/rsct/bin/hatsd -n 1 -o deadManSwit
root 9810 3360 0 Oct 31 - 1:53 harmad -t HACMP -n Scream
root 10102 3360 0 Oct 31 - 0:07 /usr/es/sbin/cluster/cllockd
root 10856 3360 0 Oct 31 - 10:56 /usr/es/sbin/cluster/clsmuxpd
root 11108 3360 0 Oct 31 - 9:02 /usr/es/sbin/cluster/clresmgrd
root 11354 3360 0 Oct 31 - 6:24 /usr/es/sbin/cluster/clinfo
root 18060 3360 0 Oct 31 - 0:00 /usr/sbin/clvmd

```

Figure 18. Processes in a HACMP/ES cluster with concurrent resgroups (ghost1)

```

root 7122 4136 0 Nov 03 - 1:30 hagsd grpsvcs
root 7400 4136 0 Nov 03 - 4:14 /usr/es/sbin/cluster/clresmgrd
root 7848 4136 0 Nov 03 - 0:58 harmad -t HACMP -n Scream
root 8044 4136 0 Nov 03 - 0:02 haemd HACMP 4 Scream SECNOSUPPORT
root 8804 4136 0 Nov 03 - 0:12 hagsglsmd grpglsm
root 10654 4136 0 Nov 03 - 5:15 /usr/es/sbin/cluster/clstmgr
root 11124 4136 5 Nov 03 - 134:02 /usr/sbin/rsct/bin/hatsd -n 4 -o deadManSwit
root 11350 4136 0 Nov 03 - 5:59 /usr/es/sbin/cluster/clinfo
root 12196 4136 0 Nov 03 - 0:12 /usr/es/sbin/cluster/clsmuxpd

```

Figure 19. Processes in a HACMP/ES cluster (ghost4)

The principle of a dead man switch hasn't changed in the development of HACMP/ES although it is coded in a new kernel extension. In a mixed cluster,

both kernel extensions are loaded, but only the HACMP dead man switch is active.

```
genkex | grep -i dms # view kernel extensions
```

```
11c7ee0          150c /usr/sbin/cluster/dms
11c3940          4580 /usr/sbin/rsct/bin/haDMS/haDMS_kex
```

Figure 20. Loaded Kernel extensions on a mixed cluster node.

Table 15. Dead man switch kernel extensions

Kernel extension	HACMP	Mixed cluster	HACMP/ES
/usr/sbin/cluster/dms	active	active	
/usr/sbin/rsct/./haDMS_kex		loaded	active

Table 16. Daemons and process in different cluster types

Daemon / Process	HACMP	Mixed cluster migrated node	HACMP/ES
clstrmgr, nim_*	active	active	
clstrmgrES		passive	active
clresmgrdES		active	active
clsmuxpd	active	active	
clsmuxpdES			active
cllockd ¹⁾	active		
cllockdES ¹⁾		active	active
clvmd ¹⁾	active	active	active
clinfo ²⁾	active	active	
clinfoES ³⁾			active
topsvcs (hatsd)		active	active
grpsvcs (hagsd, hagsglcmd)		active	active
emsvcs (haemd, harmad)		active	active
heartbeat	clstrmgr	clstrmgr	topsvcs
1) Only on nodes with concurrent volume groups 2) Using /usr/sbin/cluster/etc/clinfo.rc 3) Using /usr/es/sbin/cluster/etc/clinfo.rc			

4.3.3 Log files

The processes and daemons associated with HACMP and HACMP/ES write useful information to various log files. Some log files are growing very fast in error situations, for instance, if the local system time differs from the cluster nodes, the topology services log file, topsvcs, will grow rapidly. This could fill the file system and hide further messages if there is no free space to write into the log files. Therefore, it is important that there is enough free space in the /var file system.

Table 17 on page 68 give a overview about the HACMP and HACMP/ES related log files and the processes which are written into them.

Table 17. Log files in different cluster types

Logs (written by)	HACMP	Mixed cluster	HACMP/ES
/var/adm/cluster.log (always via syslogd)	clstrmgr	clstrmgr clstrmgrES	clstrmgrES
/tmp/cm.log	clstrmgr	clstrmgr	clstrmgrES
/tmp/clstrmgr.debug		clstrmgrES	clstrmgrES
/var/ha/log/dms_loads.out		ha_DMSkex	ha_DMSkex
/var/ha/log/topsvcs.*		hatsd	hatsd
/var/ha/log/grpsvcs*		hagsd	hagsd
/var/ha/log/grpqlsm*		hagsqlsmd	hagsqlsmd
/tmp/clresmgrd.log		clresmgrdES	clresmgrdES
/tmp/clinfo.rc.out	clinfod	clinfod	clinfoES
/tmp/clvmd.log	clvmd	clvmd	clvmd
/tmp/cld_debug.out		clstrmgrES	clstrmgrES
/tmp/clconvert.log		cl_convert	

4.3.4 Cluster lock manager

In a mixed version cluster, both the HACMP and the HACMP/ES cluster managers are connected to the same lock manager (HACMP/ES), but only the HACMP cluster manager is sending topology information to it in this mode. When the migration is completed and the HACMP cluster manager is stopped, the lock manager will start receiving topology information from the HACMP/ES cluster manager. There are now two ports for connection to the lock manager, and the HACMP/ES cluster manager will remain connected to

the backup port even after the migration is complete, which will use the normal port the next time the cluster software is started on that node.

```
clm_lkm          6150/tcp        # normal port
clm_mig_lkm     6151/tcp        # backup port
```

Figure 21. Cluster lock manager ports in a mixed cluster (/etc/services)

4.4 Node-by-node migration to HACMP/ES 4.3.1

The migration to HACMP/ES 4.3.1 can be initiated when the prerequisites described in Section 4.1, “Prerequisites” on page 55 are fulfilled.

The migration to HACMP/ES 4.3.1 has similar requirements as the upgrade to HACMP 4.3.1. If you have started the migration with an upgrade to HACMP 4.3.1, most preparation tasks will already be completed.

Note

Please keep in mind that any modification on the cluster configuration is not possible from the beginning of migration on the first node until the migration has been finished on the last node.

4.4.1 Preparations

All following tasks must be done as user “root”.

If there are Twin-Tailed SCSI Disks used in the cluster, check the SCSI IDs for all adapters. No adapter or disk should use the reserved SCSI ID 7. Usage of this ID could cause an error during specific AIX installation steps. You can use the commands below to check for a SCSI conflict.

```
lsdev -Cc adapter | grep scsi | awk '{ print $1 }' | \
xargs -i lsattr -El {} | grep id # show scsi ids used
```

Any new ID assigned must be currently unused. All used SCSI IDs can be checked, and the adapter ID changed if necessary. The next example shows how to get the SCSI ID for all SCSI devices and then change the ID for SCSI0 to ID 5

```
lsdev -Cs scsi
chdev -l scsi0 -a id=5 -P # apply changes only in odm
```

A subsequent reboot of all changed nodes is needed and should be done before any further installation step is started.

The nodes must have at least 64 MB of memory to run HACMP/ES 4.3.1.

```
lsdev -Cc memory # check installed memory
```

If the current cluster has more than three nodes, any serial connections must be in what is often called a daisy-chain configuration. Cabling, as shown in Figure 22 on page 71, for cluster 1 is not valid and should be changed to look

like the daisy-chain of cluster 2. The current network configuration for a cluster configuration can be viewed by running the `c11snw` command.

```
/usr/sbin/cluster/utilities/c11snw          # check serial connections
```

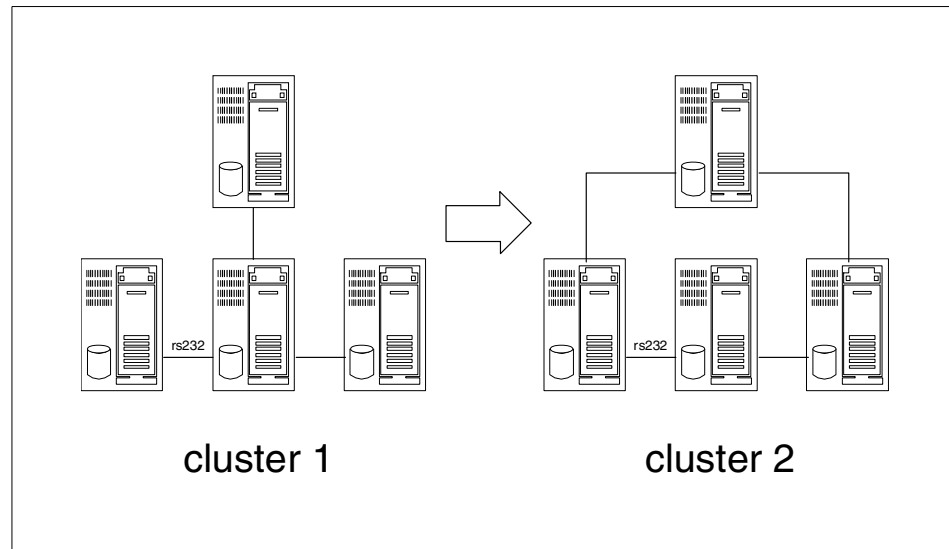


Figure 22. Cabling of serial connections

The migration process to HACMP/ES needs approximately 100 MB free space in `/usr` directory and 1 MB in the root directory `/`. Available space can be checked with:

```
df -k /usr          # check free space
df -k /
```

If there is not enough space in `/` and `/usr`, the file systems can be extended, or some space can be freed. To add one logical partition, you can use the command below. It tries to add one block = 512 bytes, but the Logical Volume Manager always extend a file system with at least one logical partition.

```
chfs -a size=+1 /usr          # add 1 logical partition
chfs -a size=+1 /
```

Any file sets updates (maintenance) should be in an committed state on all nodes before the installation of HACMP/ES 4.3.1 is started.

```
installp -cgX all 2>&1 | tee /tmp/commit.log          # commit all file sets
```

To have the chance to go back if something fails during the following installation steps, it is necessary to have a current system backup. At least

the rootvg should be backed up on a bootable media, for instance, a tape or image on a network install server.

```
mksysb -i /dev/rmt0 # backup rootvg
```

Customer specific pre- and post-event scripts will never be overwritten if they are not stored in the lpp directory tree; however, any HACMP scripts that have been modified must be saved to another safe directory. Installing a new release of HACMP will replace standard scripts. Even file set updates or fixes may result in HACMP scripts being overwritten. By taking a safe copy of modified scripts, you have a reference for reapplying modifications in the new scripts manually. In addition to the backup of rootvg, it is useful to have an extra backup of the cluster software with all modifications on their scripts.

```
cd /usr/sbin/cluster
tar -cvf /tmp/hacmp431.tar ./* # backup HACMP
```

A snapshot of the current cluster configuration is not a requirement for a node-by-node migration because these migration process reads all necessary information directly from the ODM. However, it is highly recommended to create a new snapshot as part of the fall back plan.

```
/usr/sbin/cluster/diag/clconfig -v -t -r # verify cluster config
export SNAPSHOTPATH=/usr/hacmp431
/usr/sbin/cluster/utilities/clsnapshot -cin PreMigr431 -d PreMigr431
```

During the migration, all nodes will be restarted. To avoid any inadmissible cluster states, it is highly recommended to remove any automatic cluster starts while the system boots.

```
/usr/sbin/cluster/utilities/clstop -y -R -gr # modify only inittab
```

This command does not stop HACMP on the node but prevents HACMP from starting at the next reboot by modifying */etc/inittab*.

The installable file sets for HACMP/ES 4.3.1 are delivered on CD. It is possible, for easier remote handling, to copy all file sets to a file system on a server, which can be mounted from every cluster node.

```
mkdir /usr/cdrom # copy lpp to disk
mount -rv cdrfs /dev/cd0 /usr/cdrom
cp -rp /usr/cdrom/usr/sys/inst.images/* /usr/sys/inst.images
umount /usr/cdrom
cd /usr/sys/inst.images
inutoc .
mknfsexp -d /usr/sys/inst.images -t rw -r ghost1,ghost2,ghost3,ghost4 -B
```

The preparation tasks discussed should have been carried out on all nodes before continuing with the migration.

4.4.2 Migration steps to HACMP/ES 4.3.1 for nodes 1..(n-1)

Before starting the installation of any HACMP/ES file sets on a node, the node has to be rebooted. To avoid any unnecessary outages, the migration plan should be followed.

Note

Please keep in mind that any takeover causes outage of service.

The duration of the outage depends on the cluster configuration. The current location of cascading or rotating resources, and the restart time of the application servers which belong to the resources groups will all be important for the outage time.

First, shut down cluster service on a node with takeover, and after the completion of the command, verify that the daemons actually are stopped.

```
/usr/sbin/cluster/utilities/clstop -y -N -s -gr # shutdown with takeover
lssrc -g cluster # check daemons
```

clstmgr	cluster	inoperative
clsmuxpd	cluster	inoperative
clinfo	cluster	inoperative
cllockd	lock	inoperative

Figure 23. Daemons after stop of HACMP

After verifying that all the daemons are stopped, you should check the cluster status. The best way to do this is to issue the `clstat` command as in the following:

```
/usr/sbin/cluster/clstat -a -c 20 # check cluster state
```

```
clstat - HACMP for AIX Cluster Status Monitor
-----
Cluster: Scream (20)          Thu Oct 21 17:24:31 CDT 1999
State: UP                    Nodes: 4
SubState: STABLE
Node: ghost1                 State: DOWN
  Interface: ghost1_boot_en0 (0) Address: 192.168.1.184
                                     State: DOWN
  Interface: ghost1_tty1 (1)   Address: 0.0.0.0
                                     State: DOWN
  Interface: ghost1_tty0 (4)   Address: 0.0.0.0
                                     State: DOWN
  Interface: ghost1_boot_tr0 (5) Address: 9.3.187.184
                                     State: DOWN
Node: ghost2                 State: UP
  Interface: ip4_svc (0)       Address: 192.168.1.100
```

Figure 24. Cluster state after stop of node ghost1

The current cluster state can also be checked with execution of the command `clstat`, but on another node. The output of `clstat` is shown in Figure 24.

All resources must be available, and the node you are going to migrate must have released its resources before continuing; otherwise, subsequent actions could cause an unplanned outage of service. Figure 25 is an example of the cluster state after HACMP has been stopped with the option `takeover` on `ghost4`.

```
/usr/scripts/cust/itso/check_resgr
```

```

> checking ressourcegroup CH1
  iplabel ghost1
  on node ghost1
  logical volume lvconc01.01
  iplabel ghost2
  on node ghost2
  logical volume lvconc01.01
> checking ressourcegroup SH2
  iplabel ghost3
  on node ghost3
  filesystem /shared.01/lv.01
  iplabel ghost3
  on node ghost3
  application server 1
> checking ressourcegroup SH3
  iplabel ghost4
  on node ghost3          <- after ghost4 takeover or crashed
  filesystem /shared.02/lv.01
  iplabel ghost4
  on node ghost3          <- after ghost4 takeover or crashed
  application server 2
> checking ressourcegroup IP4
  iplabel ip4_svc
  on node ghost1

```

Figure 25. Resource location after takeover

All available file sets can be installed on every cluster node from the server where they have been made accessible via nfs.

```

cd /mnt/hacmp431es
installp -aqX -d . rsct \
    cluster.adt.es \
    cluster.es.server \
    cluster.es.client \
    cluster.es.clvm \
    cluster.es.cspoc \
    cluster.es.taskguides \
    cluster.man.en_US.es | tee /tmp/install.log

lppchk -v

```

The file sets can also be explicitly selected and installed using smit.

```

mount ghost4ctl:/usr/sys/inst.images /mnt
cd /mnt/hacmp431es
smitty install
-> INPUT device: .
--> SOFTWARE to install:
    cluster.at.es

```

```

cluster.clvm
cluster.cspoc
cluster.es
cluster.es.clvm
cluster.es.cspoc
cluster.es.taskguides
clustyer.man.en_US.es
rsct.basic
rsct.clients
-> PREVIEW only? : no
lppchk -v

```

As a result of the installation of HACMP/ES just completed, and before migration completes, there are HACMP and HACMP/ES file sets installed at the same time. See Figure 26.

```

lslpp -Lc 'cluster.*' | cut -f1 -d: | sort | uniq # check installed lpp

```

```

cluster.adt
cluster.adt.es
cluster.base
cluster.clvm
cluster.cspoc
cluster.es
cluster.es.clvm
cluster.es.cspoc
cluster.es.taskguides
cluster.man.en_US
cluster.man.en_US.es
cluster.msg.en_US
cluster.msg.en_US.cspoc
cluster.msg.en_US.es

```

Figure 26. Installed file sets in a mixed cluster

Messages from the `installp` command appear on standard output and in the file `/smit.log` if the installation has been completed using `smit`. The log for `installp` is shown in Figure 27 and indicates that this installation was a valid migration from HACMP 4.3.1 to HACMP/ES 4.3.1.

```

HACMP version 4.3.1.0 is installed on the system.
HACMP to HACMP/ES migration detected.
Restoring files, please wait.
...
Adding default (loopback) server information to /usr/es/sbin/cluster/etc/clhosts file.
topsvcs6178/udp
grpsvcs6179/udp
0513-071 The topsvcs Subsystem has been added.
0513-068 The topsvcs Notify method has been added.
0513-071 The grpsvcs Subsystem has been added.
0513-071 The grpqlsm Subsystem has been added.
0513-071 The emsvcs Subsystem has been added.
0513-071 The emaixos Subsystem has been added.
0513-071 The clinfoES Subsystem has been added.
0513-068 The clinfoES Notify method has been added.
0513-071 The clstmgrES Subsystem has been added.
0513-068 The clstmgrES Notify method has been added.
0513-071 The clsmuxpdES Subsystem has been added.
0513-068 The clsmuxpdES Notify method has been added.
0513-071 The clresmgrdES Subsystem has been added.
0513-071 The cllockdES Subsystem has been added.
...

```

Figure 27. Log of `installp` command

For a better understanding and to avoid problem determination, it is recommended to make all relevant HACMP subsystems visible with the `lssrc` command, independent of whether they are active or not.

```

chssys -s clinfoES -d # modif subsystem definition
chssys -s cllockdES -d
chssys -s clsmuxpdES -d
chssys -s clvmd -d

```

The HACMP cluster log file, `/var/adm/cluster.log`, will be removed during the un-installation of these file sets. To preserve the information that will be written to this log, it is recommended to move it to a safe location.

```

mv /var/adm/cluster.log /tmp # relocate cluster.log
ln -s /var/adm/cluster.log /tmp/cluster.log
edit /etc/syslog.conf
-> change /var/adm/cluster.log to /tmp/cluster.log
refresh -s syslogd

```

Reboot the node to bring all new executables, libraries, and the new underlying services, Topology Service, Group Service, Event Management, into effect.

```

shutdown -Fr # reboot node

```


After the node has been restarted, it is recommended to do some checks to be sure that the node will become a stable member of the cluster. Check, for example, the AIX error log and physical volumes.

```
errpt | pg # check AIX error log
lspv # check physical volumes
```

Once the node has come up, and assuming there are no problems, then HACMP can be started on it.

Note

Please keep in mind that the start of HACMP on a node that is configured in an cascading resource group can cause outage of service.

Whether this occurs and how long it takes depends on the cluster configuration. The current location of the cascading resources and the restart time of the application servers that belong to the resources groups will be important.

```
/usr/sbin/cluster/etc/rc.cluster -boot -N -i # start without cllockd
/usr/sbin/cluster/etc/rc.cluster -boot -N -i -l # start with cllockd
```

This command starts all HACMP and HACMP/ES daemons. Not all daemons belonging to HACMP/ES are allowed to run in a mixed cluster. The script `rc.cluster` will automatically start the HACMP and HACMP/ES daemons based on the current cluster type; mixed or HACMP/ES. The script will determine which daemons can be started.

```

Migration from HACMP to HACMP/ES detected.
Performing startup of HACMP/ES services.

Starting execution of /usr/es/sbin/cluster/etc/rc.cluster
with parameters : -boot -N -i
0513-029 The portmap Subsystem is already active.
Multiple instances are not supported.
0513-029 The inetd Subsystem is already active.
Multiple instances are not supported.
Checking for sromstr active...complete.
 6222      - 0:00 syslogd
0513-059 The topsvcs Subsystem has been started. Subsystem PID is 10886.
0513-059 The grpsvcs Subsystem has been started. Subsystem PID is 9562.
0513-059 The grpglsm Subsystem has been started. Subsystem PID is 9812.
0513-059 The emsvcs Subsystem has been started. Subsystem PID is 8822.
0513-059 The emaixos Subsystem has been started. Subsystem PID is 10666.
0513-059 The clstmgrES Subsystem has been started. Subsystem PID is 8554.
0513-059 The clresmgrdES Subsystem has been started. Subsystem PID is 11624.

Completed execution of /usr/es/sbin/cluster/etc/rc.cluster
with parameters: -boot -N -i.
Exit Status = 0.

Starting execution of /usr/sbin/cluster/etc/rc.cluster
with parameters : -boot -N -i
0513-029 The portmap Subsystem is already active.
Multiple instances are not supported.
0513-029 The inetd Subsystem is already active.
Multiple instances are not supported.
Loaded kernel extension kmid = 18965120
dms init
Checking for sromstr active...complete.
 6222      - 0:00 syslogd
0513-059 The clstmgr Subsystem has been started. Subsystem PID is 9138.
0513-059 The snmpd Subsystem has been started. Subsystem PID is 12386.
0513-059 The clsmuxpd Subsystem has been started. Subsystem PID is 12138.
0513-059 The clinfo Subsystem has been started. Subsystem PID is 11884.

Completed execution of /usr/sbin/cluster/etc/rc.cluster
with parameters: -boot -N -i.
Exit Status = 0.

```

Figure 28. Output of rc.cluster command in a mixed cluster

After HACMP has been started, there are some additional checks required to be sure that the node is a stable member of the cluster. For instance, check the HACMP logs and state.

```

tail -f /var/adm/cluster.log           # check cluster log
tail -f /tmp/cm.log                   # check cluster manager log
tail -f /tmp/hacmp.out                # check event script log
unset DISPLAY
/usr/es/sbin/cluster/clstat -c 20     # check cluster state

```

```

Oct 18 15:24:07 ghost4 clstmgr[8892]: Mon Oct 18 15:24:07 HACMP/ES Cluster Manager
Started
Oct 18 15:24:10 ghost4 clresmgrd[8468]: Mon Oct 18 15:24:10 HACMP/ES Resource
Manager Started
Oct 18 15:24:31 ghost4 clstmgr[9090]: CLUSTER MANAGER STARTED
Oct 18 15:24:39 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[7902]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/cmd/clevlog has been suppressed by HAES
Oct 18 15:24:39 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[7902]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/node_up has been suppressed by HAES
Oct 18 15:24:39 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[7902]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/cmd/clevlog has been suppressed by HAES
Oct 18 15:24:40 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[21136]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/cmd/clevlog has been suppressed by HAES
Oct 18 15:24:40 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[21136]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/node_up_complete has been suppressed by
HAES
Oct 18 15:24:40 ghost4 /usr/es/sbin/cluster/events/cmd/clcallev[21136]: NODE BY NODE
WARNING: Event /usr/es/sbin/cluster/events/cmd/clevlog has been suppressed by HAES
Oct 18 15:25:03 ghost4 HACMP for AIX: EVENT START: node_up_ghost4
Oct 18 15:25:09 ghost4 HACMP for AIX: EVENT START: node_up_local
Oct 18 15:25:21 ghost4 HACMP for AIX: EVENT START: acquire_service_addr ghost4
Oct 18 15:25:54 ghost4 HACMP for AIX: EVENT START: acquire_aconn_service tr0 token
Oct 18 15:25:56 ghost4 HACMP for AIX: EVENT START: swap_aconn_protocols tr0 tr1
Oct 18 15:25:57 ghost4 HACMP for AIX: EVENT COMPLETED: swap_aconn_protocols tr0 tr1
Oct 18 15:25:58 ghost4 HACMP for AIX: EVENT COMPLETED: acquire_aconn_service tr0
token
Oct 18 15:25:59 ghost4 HACMP for AIX: EVENT COMPLETED: acquire_service_addr ghost4
Oct 18 15:26:00 ghost4 HACMP for AIX: EVENT START: get_disk_vg_fs /shared.02/lv.01
/shared.02/lv.02
Oct 18 15:26:22 ghost4 HACMP for AIX: EVENT COMPLETED: get_disk_vg_fs
/shared.02/lv.01
Oct 18 15:26:23 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_local
Oct 18 15:26:24 ghost4 HACMP for AIX: EVENT START: node_up_local
Oct 18 15:26:25 ghost4 HACMP for AIX: EVENT START: get_disk_vg_fs
Oct 18 15:26:27 ghost4 HACMP for AIX: EVENT COMPLETED: get_disk_vg_fs
Oct 18 15:26:27 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_local
Oct 18 15:26:28 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_ghost4
Oct 18 15:26:30 ghost4 HACMP for AIX: EVENT START: node_up_complete ghost4
Oct 18 15:26:31 ghost4 HACMP for AIX: EVENT START: node_up_local_complete
Oct 18 15:26:32 ghost4 HACMP for AIX: EVENT START: start_server Application_2
Oct 18 15:26:34 ghost4 HACMP for AIX: EVENT COMPLETED: start_server Application_2
Oct 18 15:26:39 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_local_complete
Oct 18 15:26:40 ghost4 HACMP for AIX: EVENT START: node_up_local_complete
Oct 18 15:26:41 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_local_complete
Oct 18 15:26:44 ghost4 HACMP for AIX: EVENT COMPLETED: node_up_complete ghost4

```

Figure 29. Cluster log during the start of HACMP in a mixed cluster

Both HACMP and HACMP/ES event scripts are started, but only the HACMP event scripts will actually do any work. The execution of ES scripts are suppressed.

```

*** ADDN ghost4 9.3.187.191 (config) ***
Oct 18 15:25:03 EVENT START: node_up ghost4
Oct 18 15:25:08 EVENT START: node_up_local
Oct 18 15:25:21 EVENT START: acquire_service_addr ghost4
*** ADDN ghost4 9.3.187.190 (poll) ***
*** ADUP ghost4 9.3.187.191 (poll) ***
Oct 18 15:25:54 EVENT START: acquire_aconn_service tr0 token
Oct 18 15:25:56 EVENT START: swap_aconn_protocols tr0 tr1
Oct 18 15:25:57 EVENT COMPLETED: swap_aconn_protocols tr0 tr1
Oct 18 15:25:58 EVENT COMPLETED: acquire_aconn_service tr0 token
Oct 18 15:25:59 EVENT COMPLETED: acquire_service_addr ghost4
Oct 18 15:26:00 EVENT START: get_disk_vg_fs /shared.02/lv.01 /shared.02/lv.02
Oct 18 15:26:22 EVENT COMPLETED: get_disk_vg_fs /shared.02/lv.01 /shared.02/lv.02
Oct 18 15:26:23 EVENT COMPLETED: node_up_local
Oct 18 15:26:24 EVENT START: node_up_local
Oct 18 15:26:25 EVENT START: get_disk_vg_fs
Oct 18 15:26:26 EVENT COMPLETED: get_disk_vg_fs
Oct 18 15:26:27 EVENT COMPLETED: node_up_local
Oct 18 15:26:28 EVENT COMPLETED: node_up_ghost4
Oct 18 15:26:29 EVENT START: node_up_complete ghost4
Oct 18 15:26:31 EVENT START: node_up_local_complete
Oct 18 15:26:32 EVENT START: start_server Application_2
Oct 18 15:26:34 EVENT COMPLETED: start_server Application_2
Oct 18 15:26:38 EVENT COMPLETED: node_up_local_complete
Oct 18 15:26:39 EVENT START: node_up_local_complete
Oct 18 15:26:40 EVENT COMPLETED: node_up_local_complete
Oct 18 15:26:43 EVENT COMPLETED: node_up_complete ghost4

```

Figure 30. Cluster manager log during start of HACMP in a mixed cluster

```

PROGRAMME=cl_RMupdate
+ RM_EVENT=node_up
+ RM_PARAM=ghost4
...
+ clcheck_server clresmgrdES
PROGRAMME=clcheck_server
+ SERVER=clresmgrdES
+ lssrc -s clresmgrdES
...
+ clRMupdate node_up ghost4
executing clRMupdate
clRMupdate: checking operation node_up
clRMupdate: found operation in table
clRMupdate: operating on ghost4
clRMupdate: sending operation to resource manager
clRMupdate completed successfully
...
+ exit 0
Oct 18 15:26:43 EVENT COMPLETED: node_up_complete ghost4

```

Figure 31. Script log during start of HACMP in a mixed cluster

```

clstat - HACMP for AIX Cluster Status Monitor
-----
Cluster: Scream (20)           Wed Oct 20 18:11:04 CDT 1999
      State: UP                 Nodes: 4
      SubState: STABLE
      Interface: ghost3_tty1 (3)   Address: 0.0.0.0
                                   State: UP
      Interface: ghost3 (5)       Address: 9.3.187.187
                                   State: UP

Node: ghost4                   State: UP
      Interface: ip4_svc (0)       Address: 192.168.1.100
                                   State: UP
      Interface: ghost4_tty0 (3)   Address: 0.0.0.0
                                   State: UP
      Interface: ghost4_tty1 (4)   Address: 0.0.0.0
                                   State: UP
      Interface: ghost4 (5)       Address: 9.3.187.191
                                   State: UP

```

Figure 32. Cluster state in a mixed cluster

Topology service, group service, and event management consist of at least one daemon that is essential for the correct function of HACMP/ES. After the first start of HACMP/ES, it should be checked that all required daemons are running. Table 16 on page 67 is a short summary of all expected daemons within the different cluster types. Figure 33 shows the daemons running on a stand-alone cluster after installing HACMP/ES on a cluster node, but before the migration process to HACMP/ES takes place.

```
show_cld # check cluster daemons
```

```

clstmgrES    cluster    10846  active
clstmgr      cluster    10350  active
clsmuxpd     cluster    11626  active
clinfo       cluster    12384  active
clinfoES     cluster           inoperative
clsmuxpdES   cluster           inoperative
cllockdES    lock           inoperative
clvmd        cluster           inoperative
clresmgrdES  cluster    9814  active
topsvcs      topsvcs    8298  active
grpsvcs      grpsvcs    9108  active
grpglsm      grpsvcs    9578  active
emsvcs       emsvcs    9372  active
emaixos      emsvcs    5718  active

```

Figure 33. Daemons after start of HACMP in a mixed cluster

The previous `show_cld` command is not an AIX standard command. It is an alias whose definition is included in Appendix C, “Script utilities” on page 125.

```
tail /var/ha/log/topsvcs.15.112005.Scream # check topology service log
```

```
10/15 16:31:47 hatsd[1]: My New Group ID = (9.3.187.191:0x40079d41) and is Stable.
    My Leader is (9.3.187.191:0x400754c4).
    My Crown Prince is (9.3.187.187:0x40079cf6).
    My upstream neighbor is (9.3.187.185:0x40074d22).
    My downstream neighbor is (9.3.187.187:0x40079cf6).
10/15 16:31:56 - (0) - send_node_connectivity() CurAdap 0
10/15 16:31:56 - (2) - send_node_connectivity() CurAdap 2
10/15 16:32:02 hatsd[2]: Node Connectivity Message stopped on adapter offset 2.
10/15 16:32:02 hatsd[0]: Node Connectivity Message stopped on adapter offset 0.
```

Figure 34. Log of topology services

The topology services are started on the migrated nodes after installing HACMP/ES and rebooting the nodes. They recognize each other and form groups with all these nodes. The number of migrated nodes must be correct in the topology service status; otherwise, the new heartbeat, via topology services, is not working correctly. Failing to detect a problem within topology services, and, thereby, HACMP/ES heartbeat, might lead to the cluster crashing when migration completes.

```
lssrc -s topsvcs -l # check topology service state
```

```
Network Name  Indx Defd Mbrs St Adapter ID      Group ID
ether_0       [ 0]   4   3 S 192.168.1.190 192.168.1.190
ether_0       [ 0]                   0x400b81d7 0x400b81da
HB Interval = 1 secs. Sensitivity = 4 missed beats
token_0       [ 1]   4   3 S 9.3.187.191   9.3.187.191
token_0       [ 1]                   0x400b8251 0x400b8291
HB Interval = 1 secs. Sensitivity = 4 missed beats
token_1       [ 2]   4   3 S 192.168.187.190 192.168.187.190
token_1       [ 2]                   0x400b81da 0x400b81db
HB Interval = 1 secs. Sensitivity = 4 missed beats
  2 locally connected Clients with PIDs:
haemd( 8070) hagsd( 22362)
  Dead Man Switch Enabled:
    reset interval = 1 seconds
    trip interval = 8 seconds
  Configuration Instance = 1
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats
```

Figure 35. State of the topology service in a mixed cluster

```
grep -v '^*' /var/ha/run/topsvcs.Scream/machine.20.lst
```

```
TS_Frequency=1
TS_Sensitivity=4
TS_FixedPriority=38
TS_LogLength=5000
Network Name ether_0
Network Type ether
  1 en0 192.168.1.184
  2 en0 192.168.1.186
  3 en0 192.168.1.188
  4 en0 192.168.1.190
Network Name token_0
Network Type token
  1 tr0 9.3.187.184
  2 tr0 9.3.187.186
  3 tr0 9.3.187.188
  4 tr0 9.3.187.190
Network Name token_1
Network Type token
  1 tr1 192.168.187.184
  2 tr1 192.168.187.186
  3 tr1 192.168.187.188
  4 tr1 192.168.187.190
```

Figure 36. Machine list of a node in a mixed cluster

Any serial connections configured in the cluster won't appear in the machine list file, an example of which is given in Figure 36.

Figure 36 lists all networks in the cluster. Serial (tty) connections configured in the cluster will not appear in the list. Serial ports are non-shareable, and at this point, serial ports are used by the HACMP cluster manager.

If there are any unexpected behaviors after the migration of a node, for instance, missing adapter definitions or resource groups, it could be helpful to take a view into the ODM conversion log `/tmp/clconvert.log` for problem determination.

Two cluster managers (one for HACMP and one for HACMP/ES) are running on every migrated node until the last node has been migrated and the migration to HACMP/ES is completed. In a mixed cluster, the cluster manager for HACMP/ES (`clstrmgrES`) runs in passive mode and is only using the heartbeats based on topology services for monitoring.

The HACMP heartbeats can be traced with the `clstrmgr` command. The trace output is listed in Figure 37. The HACMP/ES heart beating is based on the topology service. Figure 39 shows the trace of `topsvcs` daemon heartbeats. All HACMP decisions, which are taken based on these heartbeats are written

to the topology services log file if the tracing has been turned on (see Figure 38).


```
/usr/sbin/cluster/diag/cldiag debug clstmgr -l 3 # show HACMP heartbeat
```

```
Waiting for clstmgr to respond to SRC request...
+ executing callback, no acknowledgement
+ empty queue
SEND: Message type ARE YOU ALIVE to 1
+ sending ARE YOU ALIVE message seq 1461 immediately
+ executing callback, no acknowledgement
+ empty queue
Turning debugging OFF (Wed Oct 13 14:41:15 1999
).
Turning debugging ON (level 3) (Wed Oct 13 14:42:04 1999
).
timestamp: Wed Oct 13 14:42:05 1999
got ARE YOU ALIVE message (9355) from ghost3
SEND: Message type I AM ALIVE to 2
+ sending I AM ALIVE message seq 1495 immediately
+ executing callback, no acknowledgement
+ empty queue
SEND: Message type ARE YOU ALIVE to 0
+ sending ARE YOU ALIVE message seq 1496 immediately
+ executing callback, no acknowledgement
+ empty queue
SEND: Message type ARE YOU ALIVE to 1
+ sending ARE YOU ALIVE message seq 1497 immediately
+ executing callback, no acknowledgement
+ empty queue
^CExiting.
Waiting for clstmgr to respond to SRC request...
SRC command succeeded.
```

Figure 37. Output with trace started (*traceon -l -s topsvc*)

All HACMP decisions will be written to the topology services log file. If tracing has been turned on, the following commands list the heartbeats, and the output is listed in Figure 38.

```
/usr/sbin/rsct/bin/topsvcsctrl -t # show topsvc heartbeat
view /var/ha/log/topsvcs.15.112005.Scream
/usr/sbin/rsct/bin/topsvcsctrl -o
```

```

10/19 15:21:47 hatsd[2]: Received a GROUP CONNECTIVITY message(4) from node (19
2.168.187.190:0x400cd061) in group (192.168.187.190:0x400cd2c7) .
10/19 15:21:47 hatsd[2]: Adapter 2 is in group (192.168.187.190:0x400cd2c7) on
the following nodes.
    2 3 4
10/19 15:21:48 hatsd[0]: Received a GROUP CONNECTIVITY message(4) from node (19
2.168.1.100:0x400cd10c) in group (192.168.1.100:0x400cd2c7) .
10/19 15:21:48 hatsd[0]: Adapter 0 is in group (192.168.1.100:0x400cd2c7) on th
e following nodes.
    2 3 4
10/19 15:21:48 hatsd[1]: Received a GROUP CONNECTIVITY message(4) from node (9.
3.187.191:0x400cd0ca) in group (9.3.187.191:0x400cd2c7) .
10/19 15:21:48 hatsd[1]: Adapter 1 is in group (9.3.187.191:0x400cd2c7) on the
following nodes.
    2 3 4

```

Figure 38. Trace of topology services

Trace of the subsystem *topsvcs* is enabled with the following commands, and the output is listed in Figure 39.

```

taceson -l -s topsvcs # show topsvcs heartbeat
view /var/ha/log/topsvcs.15.112005.Scream
tracesoff -s topsvcs

```

```

10/19 15:17:21 hatsd[2]: Received a [GROUP_PROCLAIM] Daemon Message.
10/19 15:17:21 hatsd[0]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:21 hatsd[0]: Sending HEARTBEAT Message.
10/19 15:17:21 hatsd[2]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:21 hatsd[1]: Sending Group PROCLAIM messages.
10/19 15:17:21 hatsd[2]: Sending HEARTBEAT Message.
10/19 15:17:21 hatsd[1]: Received a [GROUP_PROCLAIM] Daemon Message.
10/19 15:17:21 hatsd[1]: Sending HEARTBEAT Message.
10/19 15:17:21 hatsd[1]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:21 hatsd[0]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:21 hatsd[1]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:21 hatsd[2]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:22 hatsd[0]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:22 hatsd[0]: Sending HEARTBEAT Message.
10/19 15:17:22 hatsd[2]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:22 hatsd[2]: Sending HEARTBEAT Message.
10/19 15:17:22 hatsd[1]: Sending HEARTBEAT Message.
10/19 15:17:22 hatsd[1]: Received a [HEART_BEAT] Daemon Message.
10/19 15:17:23 hatsd[0]: Received a [HEART_BEAT] Daemon Message.

```

Figure 39. Trace of topology services (Heartbeats)

The upgrade steps must be repeated for every node in the cluster one by one until only one node is waiting for upgrade. The installation of HACMP/ES on the last node, described in Section 4.4.3, finishes the migration process for the cluster.

It is supported to run a mixed cluster for a limited period of time; however, this period should be as short as possible.

4.4.3 Migration steps to HACMP/ES 4.3.1 on the last node n

The migration process for the last node in the cluster is slightly different than for nodes 1 to n-1, as the cluster type will change from HACMP to HACMP/ES with the completion of the installation.

Final migration to HACMP/ES is the result of the HACMP event migrate that is started after the event node_up_complete on the last node. The migrate event is enabled on all nodes in the cluster and is executed serially.

```
Oct 27 13:19:36 ghost1 HACMP for AIX: EVENT COMPLETED: node_up_complete ghost1
Oct 27 13:19:37 ghost1 clstmgr[10884]: cc_rpcProg : CM reply to RD : allowing WAIT event
to enqueue!
Oct 27 13:19:37 ghost1 clstmgr[10884]: WAIT: Received WAIT request...
Oct 27 13:19:37 ghost1 clstmgr[10884]: WAIT: Preparing FSMrun call...
Oct 27 13:19:37 ghost1 clstmgr[10884]: WAIT: FSMrun call successful...
Oct 27 13:19:48 ghost1 HACMP for AIX: EVENT START: migrate ghost1
Oct 27 13:20:08 ghost1 clstmgr[10884]: Cluster Manager for node name ghost1 is exiting
with code 0
Oct 27 13:20:24 ghost1 HACMP for AIX: EVENT COMPLETED: migrate ghost1
Oct 27 13:20:25 ghost1 HACMP for AIX: EVENT START: migrate_complete ghost1
Oct 27 13:25:37 ghost1 HACMP for AIX: EVENT COMPLETED: migrate_complete ghost1
Oct 27 13:25:52 ghost1 HACMP for AIX: EVENT START: network_up ghost1 serial12
Oct 27 13:25:52 ghost1 HACMP for AIX: EVENT COMPLETED: network_up ghost1 serial12
Oct 27 13:25:54 ghost1 HACMP for AIX: EVENT START: network_up_complete ghost1 serial12
Oct 27 13:25:55 ghost1 HACMP for AIX: EVENT COMPLETED: network_up_complete ghost1 serial12
...
Oct 27 13:27:25 ghost1 HACMP for AIX: EVENT START: network_up ghost4 serial41
Oct 27 13:27:26 ghost1 HACMP for AIX: EVENT COMPLETED: network_up ghost4 serial41
Oct 27 13:27:27 ghost1 HACMP for AIX: EVENT START: network_up_complete ghost4 serial41
Oct 27 13:27:28 ghost1 HACMP for AIX: EVENT COMPLETED: network_up_complete ghost4 serial41
Oct 27 13:27:41 ghost1 HACMP for AIX: EVENT START: migrate ghost2
Oct 27 13:27:42 ghost1 HACMP for AIX: EVENT COMPLETED: migrate ghost2
Oct 27 13:27:44 ghost1 HACMP for AIX: EVENT START: migrate_complete ghost2
Oct 27 13:27:44 ghost1 HACMP for AIX: EVENT COMPLETED: migrate_complete ghost2
...
Oct 27 13:28:02 ghost1 HACMP for AIX: EVENT START: migrate ghost4
Oct 27 13:28:02 ghost1 HACMP for AIX: EVENT COMPLETED: migrate ghost4
Oct 27 13:28:03 ghost1 HACMP for AIX: EVENT START: migrate_complete ghost4
Oct 27 13:28:04 ghost1 HACMP for AIX: EVENT COMPLETED: migrate_complete ghost4
```

Figure 40. Cluster log during event migrate

Depending on the complexity of the cluster configuration and the CPU speed, this could take longer than 360 seconds. In this case, a message “config_too_long” appears in the cluster log, and the cluster will become unstable, which requires a manual intervention. You must stop and restart cluster nodes to exit from this event. The config_too_long messages may occur more than once and obscure other important messages. It can be helpful to prevent these messages during the migration by modifying the cluster manager subsystem.

```
chssys -s clstrmgr -a '-u 3200000' # modify subsystem def
```

Note

During the installation of HACMP, the cluster log's file will be removed. To keep the information, you can replace the log file, on each node, with a link to another file before starting HACMP on the last node. HACMP then only removes the cluster log file after migration.

Start HACMP and HACMP/ES with the command:

```
/usr/sbin/cluster/etc/rc.cluster -boot -N -i # start without cllockd  
/usr/sbin/cluster/etc/rc.cluster -boot -N -i -l # start with cllockd  
tail -f /var/adm/cluster.log # check cluster log
```

The special event migrate runs only once per node and concludes the migration to HACMP/ES. The migrate event comprises the following actions:

- Deactivates HACMP cluster manager and heartbeat
- Activates heartbeat via topology services of HACMP/ES
- Un-installs all HACMP file sets on every cluster node
- Moves control of serial connections to topology services

After the event has finished on all nodes, the HACMP kernel extensions is the only thing remaining. The next reboot of the node will remove this last part of HACMP.

The uninstallation of HACMP on every cluster node is done by using the AIX `installp` command. Figure 41 shows the process tree as a snapshot of the process list during the event migrate_complete.

```
ps -ef | grep -E 'clstrmgr|run_rcovcmd|installp|migrate_'
```

```

root 8176 3360 1 13:16:22 - 0:03 /usr/es/sbin/cluster/clstmgr
root 12590 8176 0 13:20:24 - 0:00 run_rcovcmd -sport 1000 -result_node 1
root 14464 18622 0 13:20:44 - 0:01 installp -gu cluster.adt.client.demos
cluster.adt.client.include cluster.adt.client.samples.clinfo
cluster.adt.client.samples.clstat cluster.adt.client.samples.demos
cluster.adt.client.samples.libcl cluster.adt.server.demos
cluster.adt.server.samples.demos cluster.adt.server.samples.images
cluster.base.client.lib cluster.base.client.rte cluster.base.client.utils
cluster.base.server.diag cluster.base.server.events cluster.base.server.rte
cluster.base.server.utils cluster.cspoc.cmds cluster.cspoc.dsh cluster.cspoc.rte
cluster.msg.en_US.client cluster.msg.en_US.server
root 15502 12590 0 13:20:24 - 0:00 /usr/es/sbin/cluster/events/cmd/clcallev
root 18622 15502 0 13:20:25 - 0:00 ksh
/usr/es/sbin/cluster/events/migrate_complete

```

Figure 41. Processes during un-installation of HACMP

Only the HACMP/ES file sets are remaining in a HACMP/ES cluster. For example, the file sets listed in Figure 42.

```
lslpp -Lc 'cluster.*' | cut -f1 -d: | sort | uniq # check installed lpp
```

```

cluster.adt.es
cluster.clvm
cluster.es
cluster.es.clvm
cluster.es.cspoc
cluster.es.taskguides
cluster.man.en_US.es
cluster.msg.en_US.cspoc
cluster.msg.en_US.es

```

Figure 42. Installed file sets in a HACMP/ES cluster

Migration tasks for the last non-migrated node are:

- Stop HACMP with takeover
- Install HACMP/ES and check the cluster. This is the same procedure as for nodes 1 to n-1. The daemons that are running after the first start of HACMP/ES are quite different.

Compare Figure 43 and Figure 44 with your environment. Table 16 on page 67 gives a more detailed overview on the daemons we expect to run on the nodes.

```

Starting execution of /usr/sbin/cluster/etc/rc.cluster
with parameters: -boot -N -i -l

0513-029 The portmap Subsystem is already active.
Multiple instances are not supported.
0513-029 The inetd Subsystem is already active.
Multiple instances are not supported.
Checking for sranstr active...complete.
  4196      - 0:00 syslogd
0513-059 The topsvcs Subsystem has been started. Subsystem PID is 7490.
0513-059 The grpsvcs Subsystem has been started. Subsystem PID is 7964.
0513-059 The grpqlsm Subsystem has been started. Subsystem PID is 11658.
0513-059 The emsvcs Subsystem has been started. Subsystem PID is 10662.
0513-059 The emaixos Subsystem has been started. Subsystem PID is 14956.
0513-059 The clstmgrdES Subsystem has been started. Subsystem PID is 9452.
0513-059 The clresmgrdES Subsystem has been started. Subsystem PID is 11022.
0513-059 The cllockdES Subsystem has been started. Subsystem PID is 7764.
0513-059 The clsmuxpdES Subsystem has been started. Subsystem PID is 20854.
0513-059 The clinfoES Subsystem has been started. Subsystem PID is 7042.

Completed execution of /usr/sbin/cluster/etc/rc.cluster
with parameters: -boot -N -i -l.
Exit Status = 0.

```

Figure 43. Output of rc.cluster command in a HACMP/ES cluster

clstmgrdES	cluster	8176	active
clsmuxpdES	cluster	14932	active
clinfoES	cluster	11366	active
cllockdES	lock	9424	active
clvmd		3862	active
clresmgrdES		11234	active
topsvcs	topsvcs	8864	active
grpsvcs	grpsvcs	7580	active
grpqlsm	grpqlsm	7932	active
emsvcs	emsvcs	11622	active
emaixos	emaixos	8574	active

Figure 44. Daemons on a node in a HACMP/ES cluster

At this point in the migration of a high-availability cluster, the cluster is managed by HACMP/ES.

The serial connections via /dev/tty0 and /dev/tty1 are now under control of topology services. This is reflected in the topology service state and the machine list file as shown in Figure 45 and Figure 46.

```
lssrc -s topsvcs -l # check topology services state
```

```

topsvcs          topsvcs          13502  active
Network Name    Indx Defd Mors St Adapter ID      Group ID
ether_0         [ 0]  4   4  S 192.168.1.190  192.168.1.190
ether_0         [ 0]                0x40161354      0x40174342
HB Interval = 1 secs. Sensitivity = 4 missed beats
token_0         [ 1]  4   4  S 9.3.187.191   9.3.187.191
token_0         [ 1]                0x401613c1      0x401743c0
HB Interval = 1 secs. Sensitivity = 4 missed beats
token_1         [ 2]  4   4  S 192.168.187.190 192.168.187.190
token_1         [ 2]                0x401717c5      0x40174341
HB Interval = 1 secs. Sensitivity = 4 missed beats
rs232_0         [ 3]  4   2  S 255.255.0.7    255.255.0.7
rs232_0         [ 3]                0x8017445d      0x80174480
HB Interval = 2 secs. Sensitivity = 4 missed beats
rs232_1         [ 4]  4   2  S 255.255.0.6    255.255.0.6
rs232_1         [ 4]                0x8017445e      0x80174478
HB Interval = 2 secs. Sensitivity = 4 missed beats
  2 locally connected Clients with PIDs:
haemd( 5650) hagsd( 10528)
  Dead Man Switch Enabled:
    reset interval = 1 seconds
    trip interval = 16 seconds
  Configuration Instance = 2
  Default: HB Interval = 1 secs. Sensitivity = 4 missed beats

```

Figure 45. State of the topology services in a HACMP/ES cluster

The listing is taken from topology service state log file. To make the output more readable, you can remove all lines with a leading '*'. This can be done with the following command:

```
grep -v '^*' /var/ha/run/topsvcs.Scream/machine.20.lst
```



```

TS_Frequency=1
TS_Sensitivity=4
TS_FixedPriority=38
TS_LogLength=5000
Network Name ether_0
Network Type ether
  1 en0 192.168.1.184
  2 en0 192.168.1.186
  3 en0 192.168.1.188
  4 en0 192.168.1.190
Network Name token_0
Network Type token
  1 tr0 9.3.187.184
  2 tr0 9.3.187.186
  3 tr0 9.3.187.188
  4 tr0 9.3.187.190
Network Name token_1
Network Type token
  1 tr1 192.168.187.184
  2 tr1 192.168.187.186
  3 tr1 192.168.187.188
  4 tr1 192.168.187.190
Network Name rs232_0
Network Type rs232
  1 tty0 255.255.0.1 /dev/tty0
  4 tty1 255.255.0.7 /dev/tty1
  2 tty1 255.255.0.3 /dev/tty1
  3 tty0 255.255.0.4 /dev/tty0
Network Name rs232_1
Network Type rs232
  1 tty1 255.255.0.0 /dev/tty1
  2 tty0 255.255.0.2 /dev/tty0
  3 tty1 255.255.0.5 /dev/tty1
  4 tty0 255.255.0.6 /dev/tty0

```

Figure 46. Machine list in a HACMP/ES cluster

4.4.4 Final tasks

Cluster log file

The ES cluster log is located in `/usr/es/adm`. This is not a good place for a log file that may grow; so, we recommend moving it to `/var/adm`. This can be done with the HACMP/ES `cllog` utility.

```
/usr/es/sbin/cluster/utilities/cllog -c cluster.log -v /var/adm
/usr/es/sbin/cluster/utilities/clhare -r -n -t      # sync cluster
```

Unfortunately, the `cllog` command didn't work in our test environment. So, we relocated the log file with AIX standard commands.

```
mv /usr/es/adm/cluster.log /var/adm
vi /etc/syslog.conf
-> change /tmp/cluster.log to /var/adm/cluster.log
refresh -s syslogd
```

Customized event scripts

If any HACMP event scripts have been modified during the installation or customized at the previous HACMP level, some additional tasks are necessary where cluster event scripts have been modified, and you want to keep these modifications. First of all, identify them, then find the new scripts that are suitable for reapplying the modifications. Customer event scripts are never subject to change during any installation or migration. The `installp` command will save some scripts to the directory `/usr/lpp/save.config/usr/sbin/cluster` during the installation process.

Customized clhosts

After migration, the HACMP/ES `clhosts` becomes active because the new daemon `clinoES` is running now, but the entries in the former `clhosts` file have not been migrated automatically. The migration must be performed manually.

Customized clexit.rc

If the `clexit.rc` file has been modified in the previous level of HACMP, these changes must be reapplied to the new HACMP/ES `clexit.rc` file.

Customized clinfo.rc

If the `clinfo.rc` file has been modified in the previous level of HACMP, these changes must be reapplied to the HACMP/ES `clinfo.rc` file.

Changed grace period (config_too_long)

The default values for the cluster manager subsystem definitions should be restored on all nodes where the definition has been modified during migration. The default (360000) or previous customized value should be reapplied. The modification was suggested in Section 4.4.3, “Migration steps to HACMP/ES 4.3.1 on the last node n” on page 90. If the *grace-period* was changed in the original cluster, based on the time the cluster needed to become stable, then the default value of 360000 should be reapplied with the command:

```
chssys -S clstrmgrES -a "-u 360000" # modify HACMP/ES cluster manag
```

Figure 47 shows the HACMP and HACMP/ES subsystem definition before anything is changed.

```
lssrc -Ss clstrmgr # show HACMP cluster manager
lssrc -Ss clstrmgrES # show HACMP/ES cluster manager
```

```
[root@ghost2]/ > lssrc -Ss clstrmgr # mixed cluster
#subsysname:synonym:cmdargs:path:uid:auditid:standin:standout:standerr:action:mu
lti:contact:svrkey:svrmtpe:priority:signorm:sigforce:display:waittime:grpname:
clstrmgr: : -u 360000:/usr/sbin/cluster/clstrmgr:0:0:/dev/null:/dev/null:/dev/nul
l:-O:-Q:-K:0:0:20:0:0:-d:15:cluster:
-----
[root@ghost2]/ > lssrc -Ss clstrmgrES # HACMP/ES cluster
#subsysname:synonym:cmdargs:path:uid:auditid:standin:standout:standerr:action:mu
lti:contact:svrkey:svrmtpe:priority:signorm:sigforce:display:waittime:grpname:
clstrmgrES: : /usr/es/sbin/cluster/clstrmgr:0:0:/dev/null:/dev/null:/dev/null:-O:
-Q:-K:0:0:20:0:0:-D:15:cluster:
```

Figure 47. Cluster manager subsystems

Modify disabled dead man switch

If the dead man switch was disabled in the original cluster, you must reapply this setting. The modification in the original cluster was done by adding the `-D` option to the subsystem definition. This modification must be reapplied.

```
chssys -s clstrmgrES -a "-D" # disable dead man switch
```

Clean up previous modifications

If the automatic start of HACMP has been disabled during the preparation tasks, then enable this feature again.

```
/usr/sbin/cluster/etc/rc.cluster -boot -R # enable HACMP start on reboot
```

Cluster verification

A cluster verification should be done if there were any changes in the cluster, even if it was only to underlying software. This would be the first verification of the cluster configuration at the new HACMP level. If there have been any changes to the cluster, for instance, to resource groups, it is very important to do a verification, and the results should be read carefully.

```
/usr/es/sbin/cluster/diag/clverify software lpp          # check software  
/usr/es/sbin/cluster/diag/clverify cluster config all   # check config
```

Back up all cluster nodes

Before and after testing the system, a new system backup should be taken. Applying an old system backup to a node would result in a unsupported type of mixed cluster. A mixed cluster exists only during an unfinished migration from HACMP to HACMP/ES. This cluster type is, therefore, invalid after migration.

```
mksysb -i /dev/rmt0                                     # backup rootvg
```

Cluster test

Note

One of the most important *final* tasks is to test the cluster. The only way to be sure of high availability in a cluster is to test it, therefore, a detailed test plan for the cluster should be carried out.

The test procedure including all test cases, test steps, and their results, should be documented completely. This is the basis for the system approval by the system administrator and the management. If the cluster has been modified or expanded in any way, the cluster documentation must be refreshed.

4.4.5 Fall-Back procedure

As long as HACMP/ES has not been installed and started on the last node, the way back to HACMP is easy. Simply un-install all HACMP/ES file sets, and all nodes must be rebooted after HACMP/ES is un-installed. This could be done node-by-node with the same implications on the availability of service as the migration process has. Please refer to Section 3.1.6, "Fall-back plans" on page 37.

First, stop HACMP on the node with take over of all resource groups.

```
/usr/sbin/cluster/clstop -N -y -s -gr          # shutdown with takeover
```

After HACMP is stopped, the cluster is in a stable state again, and the un-installation of HACMP/ES can be proceed.

```
installp -ug cluster.es cluster.man.en_US.es rsct 2>&1 |  
tee /tmp/deinst.log                          # de install HACMP/ES
```

Un-installing HACMP/ES requires some steps to clean up fully. A few directories must be removed and the node rebooted to clean up the HACMP/ES dead man switch, which remains as a kernel extension.

```
shutdown -Fr                                  # reboot node  
rm -rf /var/ha                                # remove remaining parts  
rm -rf /usr/es  
rm -rf /usr/sbin/rsct
```

When the un-installation has successfully finished, HACMP can be restarted on the node, and no components of HACMP/ES will be loaded.

```
/usr/sbin/cluster/etc/rc.cluster -boot -N -i  # start without cllockd  
/usr/sbin/cluster/etc/rc.cluster -boot -N -i -l # start with cllockd
```

4.5 Differences for cluster administration after migration to HACMP/ES

As discussed in the preceding sections of this chapter, some new or changed components are built into the infrastructure of HACMP/ES. This includes new or changed commands and log files, for example, the `clRGinfo` command and `topsvcs.*` log file.

4.5.1 Log files

The following is a list of all log files that are common between HACMP and HACMP/ES. Neither their content nor their location has been changed.

<code>/tmp/hacmp.out</code>	Log file for output of event scripts
<code>/tmp/cspoc.out</code>	Log file for C-SPOC operations
<code>/tmp/emuhacmp.out</code>	Log file for emulated cluster events
<code>/tmp/clvmd.log</code>	Log file of concurrent logical volume manager daemon

The executable in the directory `/usr/sbin/cluster` has been completely replaced with links to the new HACMP/ES executables. The history files and the directory have also been moved.

<code>/var/adm/cluster.log</code>	Cluster log file for events
<code>/usr/es/sbin/cluster/history/cluster.<mmdd></code>	

The log files for the new components are very important in a problem determination procedure in an HACMP/ES cluster. The files contain a lot of useful information on which the cluster manager's decisions were made. Unfortunately, they can be quite big, especially in problem situations where a certain test case may have been replayed many times. To have a copy of these files from a well functioning cluster can be very useful for comparison.

<code>/tmp/clstrmgr.debug</code>	Cluster manager debug output
<code>/var/ha/log/topsvcs.*</code>	Log file topology services
<code>/var/ha/log/grpsvc.*</code>	Log file group services
<code>/var/ha/log/grpqlsm.*</code>	Log file group services internal
<code>/var/ha/log/dms_load.out</code>	Log file DMS kernel extension

4.5.2 Utilities

All scripts, utilities, and binaries for HACMP/ES have been moved to the directory `/usr/es/sbin`. Some links are created as substitution on the old HACMP executable. However, these links will be removed later.

This could have an impact on tools or utilities used to help administrate a cluster. If you are using HACMP utilities in shell scripts with full path, then you should change these scripts. The same problem can occur on usage of utilities in alias definitions or self-written applications. All commands and utilities ought to be checked to confirm they work as expected.

There are also some partly undocumented and unsupported utilities that exist in both products. Their name and function remain unchanged; only their location has been changed. If HACMP/ES has been extended with new features, and if these are to be used in a mixed cluster, the variable `ODMDIR` must be set to the temporary object repository `/etc/es/objrepos`.

<code>/usr/es/sbin/cluster/clstat</code>	View cluster state
<code>/usr/es/sbin/cluster/utilities/clhandle</code>	View node handle
<code>/usr/es/sbin/cluster/utilities/clmixver</code>	View cluster type
<code>/usr/es/sbin/cluster/utilities/cllsgnw</code>	View global networks

HACMP/ES 4.3 was the first release that was available on both RS/6000 SP and stand-alone nodes. Some important, partly undocumented and unsupported parts have been added because they have not been available on non-RS/6000 SP nodes before.

<code>/usr/sbin/rsct/bin/topsvcsctrl</code>	Control topology services
<code>/usr/sbin/rsct/bin/grpsvcsctrl</code>	Control group services
<code>/usr/sbin/rsct/bin/emsvcsctrl</code>	Control event management
<code>/usr/sbin/rsct/bin/hagscl</code>	View group services clients
<code>usr/sbin/rsct/bin/hagsgr</code>	View group services state
<code>/usr/sbin/rsct/bin/hagsms</code>	View group services meta groups
<code>/usr/sbin/rsct/bin/hagsns</code>	View group services name server
<code>/usr/sbin/rsct/bin/hagspbs</code>	View group service broadcasting
<code>//usr/sbin/rsct/bin/hagsvote</code>	View group services voting
<code>/usr/sbin/rsct/bin/hagsreap</code>	Log cleaning

The previously mentioned services are implemented as daemons running under control of the AIX subsystem resource controller. Therefore, all AIX standard commands according to the subsystem handling can be used.

lssrc -ls topsvcs	View topology service daemons
lssrc -ls grpsvcs	View group service daemon
lssrc -ls grpqlsm	View switch daemon
lssrc -ls emsvcs	View event management daemon
lssrc -ls emaixos	View AIX system resource monitor daemon

For further information, please refer to the redbook *HACMP Enhanced Scalability Handbook*, SG24-5328. All commands, log files, and the problem determination procedures are described in detail.

4.5.3 Miscellaneous

HACMP/ES does not support the forced down option. This is mainly because of a current requirement by topology services that each interface must be on its boot address when cluster services are started.

The HACMP/ES is based on the topology services and its heartbeat. If the time between the cluster nodes differ too much, a log entry will be written. This fills the topsvcs log file unnecessarily. Therefore, the time on the cluster nodes should be kept synchronized automatically, perhaps with the xntp daemon.

Appendix A. Our environment

This appendix describe the RS/6000 and RS/6000 SP environments used for testing. The results of the tests are reported primarily in Chapter 4 of this redbook.

A.1 Hardware and software

Hardware

- RS/6000 SP wide node
- RS/6000 model 560
- RS/6000 model 520
- SSA
- IBM 9333 serial disk subsystems

Software

- IBM HACMP Version 4 Release 3
- IBM HACMP Enhanced Scalable Version 4 Release 3
- IBM AIX Version 4 Release 3
- IBM Parallel System Support Programs for AIX Version 3 Release 1

A.2 Hardware configuration

Figure 48 shows the configuration of the RS/6000 cluster, *scream*. The cluster consists of four RS/6000 model 560 connected two by two to the shared disks. All systems are daisy chained with a serial rs232 link, and all systems are connected to two LANs, an Ethernet, and a Token Ring network.

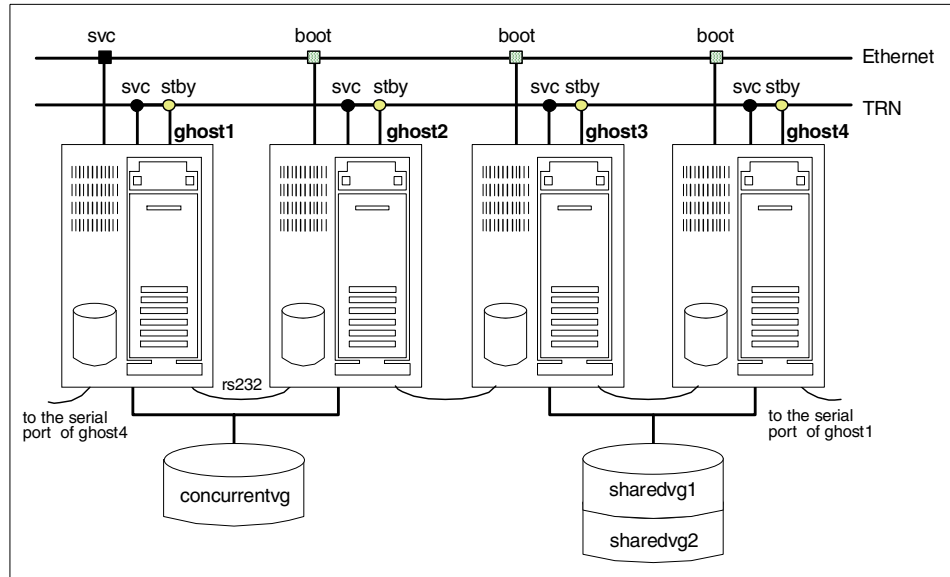


Figure 48. Hardware configuration for cluster *scream*

Figure 49 shows the configuration of the SP cluster, *scream*. The cluster consist of an RS/6000 model 520 control work station and two wide nodes in a SP system. The two nodes are connected to shared disks, and all systems are connected to two LANs, a Ethernet, and a Token Ring network.

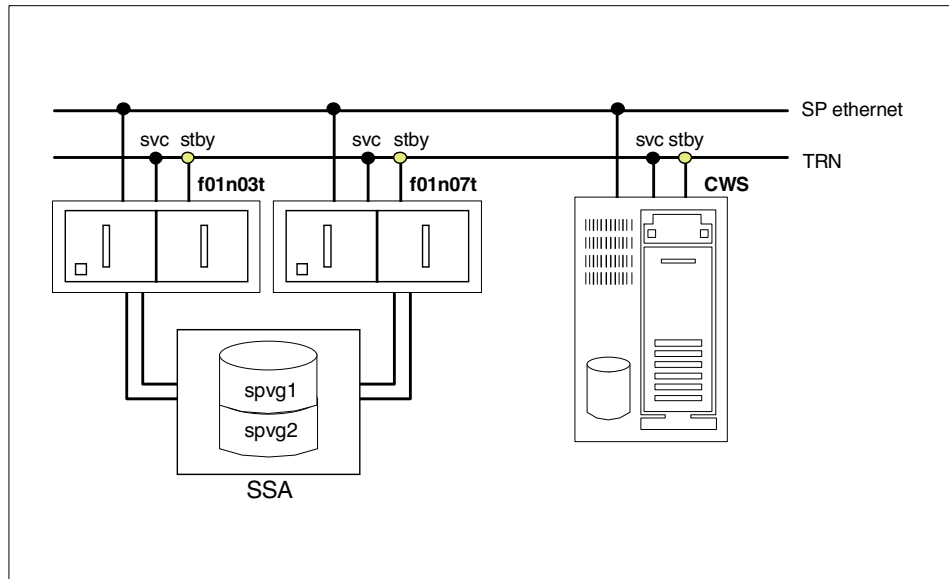


Figure 49. Hardware configuration for cluster spscream

A.3 Cluster scream

The cluster scream uses the following nodes:

- ghost1 (RS/6000 560)
- ghost2 (RS/6000 560)
- ghost3 (RS/6000 560)
- ghost4 (RS/6000 560)

A.3.1 TCP/IP networks

Cluster ID 20

Cluster Name scream

Table 18. Cluster scream TCP/IP networks

Network Name	Network Type	Network Attribute	Netmask
ether	ether	public	255.255.255.0
token	token	public	255.255.255.0

A.3.2 TCP/IP network adapters

Node Name ghost1

Table 19. Cluster scream TCP/IP network interfaces for ghost1

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	ghost1	svc	9.3.187.183	token
tr0	ghost1_boot_tr0	boot	9.3.187.184	token
tr1	ghost1_stby_tr0	stby	192.168.187.184	token
en0	ip4_svc	svc	192.168.1.100	ether
en0	ghost1_boot_en0	boot	192.168.1.184	ether

Node Name ghost2

Table 20. Cluster scream TCP/IP network interfaces for ghost2

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	ghost2	svc	9.3.187.185	token
tr0	ghost2_boot_tr0	boot	9.3.187.186	token
tr1	ghost2_stby_tr0	stby	192.168.187.186	token
en0	ip4_svc	svc	192.168.1.100	ether
en0	ghost2_boot_en0	boot	192.168.1.186	ether

Node Name ghost3

Table 21. Cluster scream TCP/IP network interfaces for ghost3

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	ghost3	svc	9.3.187.187	token
tr0	ghost3_boot_tr0	boot	9.3.187.188	token
tr1	ghost3_stby_tr0	stby	192.168.187.188	token
en0	ip4_svc	svc	192.168.1.100	ether
en0	ghost3_boot_en0	boot	192.168.1.188	ether

Node Name ghost4

Table 22. Cluster scream TCP/IP network interfaces for ghost4

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	ghost4	svc	9.3.187.191	token
tr0	ghost4_boot_tr0	boot	9.3.187.190	token
tr1	ghost4_stby_tr0	stby	192.168.187.190	token
en0	ghost4_boot_en0	boot	192.168.1.186	ether
en0	ip4_svc	svc	192.168.1.100	ether

A.3.3 Cluster resource groups

As mentioned earlier, we defined four resource groups with different characteristics, concurrent cascading and rotating. This was to make sure that we did test all combinations during the migration.

Table 23. Resource groups defined for cluster scream

Resource group	Node relationship	Participating node
CH1	concurrent	ghost1 ghost2
SH2	cascading	ghost3 ghost4
SH3	cascading	ghost4 ghost3

Resource group	Node relationship	Participating node
IP4	rotating	ghost1 ghost2 ghost3 ghost4

```

# /usr/sbin/cluster/utilities/clshowres -g CH1

Resource Group Name                CH1
Node Relationship                   concurrent
Participating Node Name(s)        ghost1 ghost2
Service IP Label
HTY Service IP Label
Filesystems
Filesystems Consistency Check      fsck
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups
Concurrent Volume Groups           concvg.01
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false

Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups
Concurrent Volume Groups           concvg.01
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Run Time Parameters:

Node Name                           ghost1
Debug Level                          high
Host uses NIS or Name Server         false

Node Name                           ghost2
Debug Level                          high
Host uses NIS or Name Server         false

```

Figure 50. Cluster scream resource group CH1

```

# /usr/sbin/cluster/clshowres/utilities -g SH2

Resource Group Name                SH2
Node Relationship                   cascading
Participating Node Name(s)        ghost3 ghost4
Service IP Label                   ghost3
HTY Service IP Label
Filesystems                        /shared.01/lv.01 /shared.01/lv.02
Filesystems Consistency Check      fsck
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups                      sharedvg.01
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers                Application_1
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                 false
SSA Disk Fencing                   false
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups                      sharedvg.01
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers                Application_1
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                 false
SSA Disk Fencing                   false
Run Time Parameters:

Node Name                          ghost3
Debug Level                        high
Host uses NIS or Name Server        false

Node Name                          ghost4
Debug Level                        high
Host uses NIS or Name Server        false

```

Figure 51. Cluster scream resource group SH2


```

# /usr/sbin/cluster/utilities/clshowres -g SH3

Resource Group Name                SH3
Node Relationship                   cascading
Participating Node Name(s)        ghost4  ghost3
Service IP Label                   ghost4
HTY Service IP Label
Filesystems                        /shared.02/lv.01 /shared.02/lv.02
Filesystems Consistency Check      fsck
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups                      sharedvg.02
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers                Application_2
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups                      sharedvg.02
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers                Application_2
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Run Time Parameters:

Node Name                          ghost4
Debug Level                        high
Host uses NIS or Name Server        false

Node Name                          ghost3
Debug Level                        high

```

Figure 52. Cluster scream resource group SH3

```

# /usr/sbin/cluster/utilities/clshowres -g IP4

Resource Group Name                IP4
Node Relationship                   rotating
Participating Node Name(s)        ghost1 ghost2 ghost3 ghost4
Service IP Label                   ip4_svc
HTY Service IP Label
Filesystems
Filesystems Consistency Check      fsck
Filesystems Recovery Method        sequential
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Standard input
Application Servers
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Run Time Parameters:

Node Name                           ghost1
Debug Level                          high
Host uses NIS or Name Server         false

Node Name                           ghost2
Debug Level                          high
Host uses NIS or Name Server         false

Node Name                           ghost3
Debug Level                          high
Host uses NIS or Name Server         false

Node Name                           ghost4
Debug Level                          high
Host uses NIS or Name Server         false

```

Figure 53. Cluster scream resource group IP4

A.4 Cluster spscream

The cluster spscream uses the following nodes

- f01n03 (RS/6000 SP wide node)
- f01n07 (RS/6000 SP wide node)

A.4.1 TCP/IP networks

Cluster ID 1
Cluster Name spscream

Table 24. Cluster spscream TCP/IP networks

Network Name	Network Type	Network Attribute	Netmask
sptoken1	token	public	255.255.255.0
sptoken2	token	private	255.255.255.0

A.4.2 TCP/IP network adapters

Node Name ff01n03

Table 25. Cluster spscream TCP/IP network interfaces for f01n03

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	f01n03t	svc	9.3.187.246	token
tr1	f01n03tt0	svc	11.3.187.246	token

Node Name ff01n07

Table 26. Cluster spscream TCP/IP network interfaces for f01n07

Interface Name	Adapter IP Label	Adapter Function	Adapter IP Address	Network name
tr0	f01n07t	svc	9.3.187.250	token
tr1	f01n07tt	svc	11.3.187.250	token

A.4.3 Cluster spscream resource groups

As mentioned earlier, we defined two resource groups in the cluster, spscream. The two resource groups are described in Table 27.

Table 27. Resource groups defined for cluster spscream

Resource group	Node relationship	Participating node
VG1	cascading	f01n03t f01n07t
VG2	cascading	f01n07t f01n03t

```

An extract from a cluster snapshot of the SP environment with HACMP 431.

Resource Group Name                VG1
Node Relationship                   cascading
Participating Node Name(s)        f01n03t f01n07t
Service IP Label
HTY Service IP Label
Filesystems                        /spfs1
Filesystems Consistency Check
Filesystems Recovery Method
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers
Highly Available Communication Links
Inactive Takeover                  false
9333 Disk Fencing                  false
SSA Disk Fencing                   false
Filesystems mounted before IP configured
Run Time Parameters:

Node Name                          f01n03t
Debug Level                         high
Host uses NIS or Name Server        false

Node Name                          f01n07t
Debug Level                         high
Host uses NIS or Name Server        false

```

Figure 54. Cluster spscream resource group VG1

An extract from a cluster snapshot of the SP environment HACMP 431.

```
Resource Group Name                VG2
Node Relationship                    cascading
Participating Node Name(s)         f01n07t f01n03t
Service IP Label
HTY Service IP Label
Filesystems                          /spfs2
Filesystems Consistency Check
Filesystems Consistency Check
Filesystems to be exported
Filesystems to be NFS mounted
Volume Groups
Concurrent Volume Groups
Disks
AIX Connections Services
AIX Fast Connect Services
Application Servers
Highly Available Communication Links
Miscellaneous Data
Inactive Takeover                    false
9333 Disk Fencing                    false
SSA Disk Fencing                      false
Filesystems mounted before IP configured
Run Time Parameters:

Node Name                            f01n07t
Debug Level                           high
Host uses NIS or Name Server          false

Node Name                            f01n03t
Debug Level                           high
Host uses NIS or Name Server          false
```

Figure 55. Cluster spscream resource group VG2

Appendix B. Migration to HACMP/ES 4.3.1 on the RS/6000 SP

This appendix focuses on the differences during migration to HACMP/ES that arise because of the HACMP management structures. The structures on an SP are compared to those on a stand-alone RS/6000. The differences include the use of the distributed shell (which, although available on stand-alone RS/6000s, is not so commonly used) and also, importantly, the level of PSSP that is installed in the environment.

The majority of an HACMP and HACMP to HACMP/ES migration is common to both SP and stand-alone environments.

B.1 General preparations for upgrade and migration

In addition to the preparations that are made prior to any installation of HACMP, there are some actions that can be completed on the SP that will make the install process smoother and easier. These include use of the working collective variable (WCOLL) and making install images available on the control workstation (CWS).

B.1.1 Using the working collective (WCOLL) variable

All nodes in an SP are controlled and managed from a control workstation (CWS). To make the install of HACMP easier, you can create a file, usually called /HACMPHOSTS, that contains a list of ip labels for the nodes that will comprise the HACMP cluster. This file can be exported for the working collective (WCOLL) variable, which is used by the distributed shell by default when `dsh` commands are called. With this variable exported, the `dsh` command can be executed just once for all the nodes in the cluster.

```
vi /HACMPHOSTS                                # create working collection file
-> insert f01n03
        f01n07
export WCOLL=/HACMPHOSTS                       # activate working collection
```

An example of how the working collective might be used is given below.

To check the available space in /usr on the nodes that comprise the working collective you could issue the following `dsh` command (or alternatively `rsh` to each node in turn).

```
dsh /usr/bin/df -k /usr                         # check free space on all nodes
```

B.1.2 Making HACMP install images available from the CWS

As mentioned in Section 2.2.2.1 “HACMP on the RS/6000 SP”, HACMP is installed on the SP using the standard filesets. The method of installation, however, differs from a stand-alone RS/6000, as the install images are usually copied to the cws and to a directory exported via nfs to each of the nodes.

`/usr/sys/inst.images` is the default directory in `smit` when opting to copy filesets to hard disk for future installation, but there is no requirement that this directory be used. In our environment, a directory called `/spdata/hacmp430` was used. Install images for HACMP were copied to hard disk on the CWS and the `.toc` file was re-created to ensure that it was up-to-date prior to the install.

A `dsh` command could then be used to mount the directory to install from on each of the nodes.

```
cd /spdata/hacmp431
inutoc . # create toc
dsh /usr/sbin/mount cws:/spdata/hacmp431 /mnt # mount on all nodes
```

It is recommended that the HACMP client image be installed on the cws to allow the cluster to be easily monitored.

B.1.3 Does PSSP have to be updated?

For the SP, the level of PSSP installed on the nodes in the cluster and the control workstation must be equal or greater than the minimum level of PSSP supported. For HACMP 4.3.1, the minimum level of PSSP is 3.1.0 as per *HACMP V4.3 AIX: Installation Guide*, SC23-4278. It should be PSSP 3.1.1 for the installation of HACMP/ES 4.3.1.

B.2 Summary of migrations done on the RS/6000 SP

HACMP is a well established product and there have been many practical migrations at the different releases in IBM laboratories and customer environments. Our experience was therefore another example of this common experience.

During testing on the SP we started the migration from two HACMP levels, HACMP 4.2.1 and release HACMP 4.3.0. In both cases, an HACMP update was undertaken to enable node-by-node migration to HACMP/ES 4.3.1.

The working environment was comprised of a cluster of two RS/6000 SP wide nodes (see Appendix A, "Our environment" on page 103 for full details). The control workstation and SP nodes were running AIX 4.3.2 and PSSP 3.1.0 plus the latest available fixes as of August, 1999.

The primary difference for migration of HACMP to HACMP/ES on an SP or stand-alone system is the importance of which PSSP level is installed. For a description of the planning, testing, and migration of HACMP to HACMP/ES, please refer to Chapter 3, "Planning for upgrade and migration" and Chapter 4, "Practical upgrade and migration to HACMP/ES". Included next is an example migration of PSSP from version 3.1.0 to 3.1.1.

B.2.1 Migration of PSSP

Migration of the control workstation and nodes in an SP is a complex task. Thorough planning should be undertaken prior to migrating PSSP, and as for any upgrade or migration, the following items are a minimum prerequisite.

- A project plan including detail of the migration plan
- An agreement with management and users about planned outage of service
- A fall back plan
- Required software and licenses

The procedure for migrating to PSSP 3.1.1, which includes a description of all possible migration paths, is described in *PSSP Installation and Migration Guide*, GA22-7347.

It is essential to have a copy of this guide, which can be downloaded from the following Web site:

http://www.rs6000.ibm.com/resource/aix_resource/Pubs/
-> RS/6000 SP (Scalable Parallel) Hardware and Software

-> PSSP

Also available from this Web site is a document called "Read this first". This document contains the most up-to-date information about PSSP. It may include references to APARs where problems in the migration process have been identified.

B.2.1.1 Preparations on the control workstation

All tasks must be done as root user.

Check for sufficient space in rootvg and also in the /spdata directory. Requirements for /spdata will depend on whether you wish to support mixed levels of PSSP. 2 GB is likely to be sufficient.

```
lsvg rootvg # check possible free space
df /spdata # check free space
```

Confirm name resolution by hostname and by IP address for the cws and all nodes that will be migrated.

```
host host_name # check name resolution
host ip_address
```

Back up the current SDR configuration using the command

```
SDRArchive PSSP3.1 # backup SDR
```

The rootvg for the control workstation should be backed up to bootable media, for example, tape.

```
mksysb -i /dev/rmt0 # backup rootvg on cws
```

Copies of critical files and file systems should also be kept, for example, /spdata.

```
backup -f /dev/rmt0 -0 /spdata # backup complete PSSP
```

Create a system backup of each node to be migrated assuming that the directory /spdata/sys1/images has been exported as an nfs directory on the control work station.

```
# backup rootvg on all nodes
mknfsexport -d /spdata/sys1/images -t rw -r f01n03,f01n07 -B
dsh /usr/sbin/mount cws:/spdata/sys1/images /mnt
dsh '/usr/bin/mksysb -i -X -p /mnt/bos.obj.sysb.$( hostname )'
```

Create the required /spdata directories to hold the PSSP 3.1.1 installp file sets.

```
mkdir /spdata/sys1/install/pssplpp/PSSP-3.1.1 # create new product dir
```

Copy the PSSP 3.1.1 file sets to this directory. Make the name changes shown below for ssp.usr and the rsct filesets, then run the `inutoc` command to generate a new table of contents.

```
bffcreate -qvX -t /spdata/sys1/install/pssplpp/PSSP-3.1.1 -d /dev/cd0 all
cd /spdata/sys1/install/pssplpp/PSSP-3.1.1
mv ssp.usr.3.1.1.0 pssp.installp
mv rsct.basic.1.1.1.0 rsct.basic
mv rsct.clients.usr.1.1.1.0 rsct.clients
inutoc .
```

4.5.3.1 Migrating the control workstation to PSSP 3.1.1

Take steps to log off all users from the nodes and CWS. Check that there are no crontab entries or jobs that might run during the migration. If there are, then these should be disabled. Any running jobs should be stopped. The system administrator should also think about preventing local applications that start on system reboot, from `/etc/inittab`, from running for the period of the migration.

The system administrator must be aware of what is running on the system and quiesce the system prior to migration.

Stop the following daemons on the CWS:

```
syspar_ctrl -G -k          # stop all subsystems
stopsrc -s sysctld         # stop sysctl daemon
stopsrc -s splogd         # stop splog daemon
stopsrc -s hardmon        # stop hardmon daemon
stopsrc -g sdr            # stop SDR daemon
lssrc -a | pg             # check if all inactive
```

Install PSSP 3.1.1 on the CWS.

```
installp -aXd /spdata/sys1/install/pssplpp/PSSP-3.1.1 all
k4init root.admin          # authenticate
install_cw                 # complete PSSP install
```

Verify the SDR and run system monitor verification test.

```
SDR_test                   # check SDR accessibility
spmon_itest                # check sys monitor install
```

Check status of frame supervisor microcode. Action was required in our environment and the command to update all microcode at once was issued. The `spsvmgr` status following this update can be seen in Figure 56.

```
spsvmgr -G -r status all   # show microcode status
spsvmgr -G -u all          # updates microcodes
```

spsvnmgr:	Frame	Slot	Supervisor State	Media Versions	Installed Version	Required Action
	1	0	Active	u_10.3c.070c u_10.3c.0720	u_10.3c.0720	None
		17	Active	u_80.19.060b	u_80.19.060b	None

Figure 56. Output of frame supervisor status after updating microcode

Set authentication `k4` and `std` via `smit`.

```
smittty spauth_rcmd # set authentication method
```

Remove old subsystem and add new subsystem.

```
syspar_ctrl -c # cleanup all subsystems
dsh -avG stopsrc -s pman # stop pman daemon
syspar_ctrl -A -G # add and start all subsystems
dsh -avG startsrc -s pman # start pman daemon
```

Run relevant verification checks on the CWS.

```
SYSMAN_test # check systems management
smon_ctest # check sys monitor config
```

4.5.3.2 Migrating nodes to PSSP 3.1.1

To migrate nodes to PSSP 3.1.1, change the node configuration data to PSSP 3.1.1 and customize, then run the `setup_server` command.

```
spchvgobj -r rootvg -p PSSP-3.1.1 1 1 2 # change node attributes
spbootins -s yes -r customize 1 1 2 # setup node install
setup_server 2>&1 | tee /tmp/setup_server.out
```

Refresh system subsystems on the cws and the nodes.

```
syspar_ctrl -r -G # refresh all subsystems
```

Copy the script, `pssp_script`, to `/tmp` on the nodes and start it there.

Shutdown and reboot the nodes after the scripts have finished.

```
pcp -w f01n03,f01n07 /spdata/sys1/install/pssp/pssp_script /tmp/pssp_script
dsh -a /tmp/pssp_script
cshutdowm -rFN f01n03, f01n07 # reboot all nodes
```

Remove all old subsystems and add new subsystems.

```
dsh -a /usr/lpp/ssp/bin/syspar_ctrl -c # cleanup all subsystems
dsh -a /usr/lpp/ssp/bin/syspar_ctrl -A # add and start all subsystems
```

Run relevant verification checks.

`SYSMAN_test`

`# check systems management`

Appendix C. Script utilities

We are using different types of self developed scripts in this redbook:

- Scripts that have been developed to test the specific test environment
- Scripts that have been developed to create the test environment
- General purpose scripts that have been developed for a more comfortable handling of HACMP and HACMP/ES

The scripts are listed in this order in this appendix.

C.1 Test scripts

Checks availability of services (check_resgr)

```
#!/usr/bin/ksh
#set -x
#####
#
#. This script checks availability of resource groups
#
#. Name      :  check_resgr
#. Parameters :
#
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
# This productivity aid is provided on an "AS-IS" basis
# WITHOUT ANY WARRANTY OF ANY KIND
#
#####

function check_iplabel
{
    /usr/bin/echo "    iplabel $1 "
    /usr/sbin/ping -qc 1 -il $1 >/dev/null
    if [ $? -ne 0 ]; then
        /usr/bin/echo " ERROR: iplabel '$1' is not available."
        return 1
    else
        /usr/bin/echo "    on node $(/usr/bin/rsh $1 /usr/bin/hostname )"
        return 0
    fi
    return
}

function check_fs
{
    if check_iplabel $1 ; then
        /usr/bin/echo "    filesystem $2"
        TAG=$( /usr/bin/date )
        /usr/bin/echo "${TAG}" |
        /usr/bin/rsh $1 "/usr/bin/dd of=$2/testfile > /dev/null 2>&1"
        REC=$( /usr/bin/rsh $1 "/usr/bin/cat $2/testfile" )
        if [[ ${TAG} != ${REC:-#} ]]; then
            /usr/bin/echo " ERROR: filesystem '$2' on '$1' is not available."
        fi
    fi
}
```

```

    return
}

function check_concvg
{
    if check_iplabel $1 ; then
        /usr/bin/echo "    logical volume $2"
        TAG=$( /usr/bin/date )
        /usr/bin/echo "${TAG}" |
        /usr/bin/rsh $1 "/usr/bin/dd of=/dev/$2 2>/dev/null"
        REC=$( /usr/bin/rsh $1 "/usr/bin/dd if=/dev/$2 2>/dev/null |
        /usr/bin/head -1"
        )
        if [[ "${TAG}" != "${REC}" ]]; then
            /usr/bin/echo "    ERROR: logical volume '$2' on '$1' is not available."
        fi
    fi
    return
}

function check_appl_server
{
    if check_iplabel $1 ; then
        /usr/bin/echo "    application server $2"
        PROC=$( /usr/bin/rsh $1 "/usr/bin/ps -eF%c%a"
        /usr/bin/grep -v grep
        /usr/bin/grep "dd if=/shared..$2/lv.../appl$2" )
        if [ -z "${PROC}" ]; then
            /usr/bin/echo "    ERROR: application server$2 is not running."
        fi
    fi
    return
}

# ----- main -----

/usr/bin/echo "> checking ressourcegroup CH1"
check_concvg ghost1 lvconc01.01
check_concvg ghost2 lvconc01.01

/usr/bin/echo "> checking ressourcegroup SH2"
check_fs ghost3 /shared.01/lv.01
check_appl_server ghost3 1

/usr/bin/echo "> checking ressourcegroup SH3"
check_fs ghost4 /shared.02/lv.01
check_appl_server ghost4 2

/usr/bin/echo "> checking ressourcegroup IP4"
check_iplabel ip4_svc

/usr/bin/echo

exit 0

```

C.2 Scripts used to create the test environment

Start application server script (cl_start_appl1.ITSO)

```
#!/usr/bin/ksh
#set -x
#####
#
#   Script for the 'start_application' cluster event
#
#   Name       :   cl_start_appl1.ITSO
#   Parameters :
#
#   Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
#   This productivity aid is provided on an "AS-IS" basis
#   WITHOUT ANY WARRANTY OF ANY KIND
#
#####
# create named pipe before
# mknod /shared.01/lv.01/appl1 p
# ----- main -----

{
  /usr/bin/echo "-----"
  /usr/bin/echo " Applicationserver 1 started. "
  /usr/bin/echo "-----"

  /usr/bin/echo "/usr/bin/dd if=/shared.01/lv.01/appl1 of=/dev/console" |
  /usr/bin/at now

} > /dev/console 2>&1

exit 0
```

Stop application server script (cl_stop_appl1.ITSO)

```
#!/usr/bin/ksh
#set -x
#####
#
#   Script for the 'stop_application' cluster event
#
#   Name       :   cl_stop_appl1.ITSO
#   Parameters :
#
#   Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
#   This productivity aid is provided on an "AS-IS" basis
#   WITHOUT ANY WARRANTY OF ANY KIND
#
#####
# ----- main -----

{
  /usr/bin/echo "-----"
  /usr/bin/echo " Applicationserver 1 stopped. "
  /usr/bin/echo "-----"

  /usr/bin/echo "appl1 terminated." > /shared.01/lv.01/appl1

} > /dev/console 2>&1

exit 0
```

Post-event script (cl_event_network_down.ITSO)

```
#!/usr/bin/ksh
#set -x
#####
#
#. Script for the 'network_down' HACMP cluster event notification
#
#. Name      : cl_event_network_down.ITSO
#. Parameters : <Event> <State> <NodeID> <NodeName> <NetworkName>
#
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
# This productivity aid is provided on an "AS-IS" basis
# WITHOUT ANY WARRANTY OF ANY KIND
#
#####

function exists_local_label
{
    ADR=$( /usr/bin/host $1 2>/dev/null |
        /usr/bin/awk '{ print $3 }' )
    ADR=${ADR%,}
    LBL=$( /usr/bin/netstat -in |
        /usr/bin/grep ${ADR}:-### )
    if [ -z "${LBL}" ] ; then
        return 1
    fi
    return 0
}

function move_IP4
{
    NODES="ghost2 ghost3 ghost1"
    REF_ADR="ghost4_boot_en0"
    for NODE in ${NODES} ; do
        if /usr/sbin/ping -qc 1 -i1 ${NODE} ; then
            CMD="/usr/sbin/ping -qc 1 -i1 ${REF_ADR} >/dev/null 2>&1 ; echo \${?}"
            if [[ $( /usr/bin/rsh ${NODE} "${CMD}" ) = 0 ]] ; then
                /usr/bin/logger -t ' '
                /usr/bin/logger -t ' ' "> Move ressourcegroup IP4 to node ${NODE}."
                /usr/bin/logger -t ' '
                /usr/sbin/cluster/utilities/cldare -vNM IP4:${NODE}
                return 0
            fi
        fi
    done
    /usr/bin/logger -t ' '
    /usr/bin/logger -t ' ' " ERROR: No en0 adapter on any node available."
    /usr/bin/logger -t ' '
    return 99
}

# ----- main -----
NETWORK=${4:-"***unknown***"}
{
    if exists_local_label ip4_svc ; then
        case "${NETWORK}" in
            ether ) move_IP4 ;;
            * ) ;;
        esac
    fi
} > /dev/console 2>&1
exit 0
```

C.3 HACMP script utilities

Common alias definitions (.define_alias.common)

```
#####  
#  
# Definition of some usefull common alias  
#  
# Name      : .define_alias.common  
#  
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999  
# This productivity aid is provided on an "AS-IS" basis  
# WITHOUT ANY WARRANTY OF ANY KIND  
#  
#####  
  
alias clear_log='cat /dev/null > '  
alias cls='tput clear'  
alias df='/usr/bin/df -k -I'  
alias dsh_all="export WCOLL=/wcoll_all; dsh"  
alias dsh_prod="export WCOLL=/wcoll_prod; dsh"  
alias dfp='/usr/bin/df -k -I | pg'  
alias halt='/usr/bin/echo "Please use /usr/sbin/halt !!"  
alias l='ls -l'  
alias ld='ls -al | /usr/bin/grep "^d" '  
alias ll='ls -al | pg'  
alias pg='pg -cnspl%d]  
alias psp='/usr/bin/ps -ef | pg'  
alias rm='/usr/bin/rm -i'  
alias smon='/usr/local/bin/monitor -s 1 -H -D'
```

HACMP alias definitions (.define_alias.hacmp)

```
#####  
#  
# Definition of some usefull hacmp alias  
#  
# Name      : .define_alias.hacmp  
#  
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999  
# This productivity aid is provided on an "AS-IS" basis  
# WITHOUT ANY WARRANTY OF ANY KIND  
#  
#####  
  
alias cldiag='/usr/sbin/cluster/diag/cldiag debug clstrmgr -l'  
  
alias show_cld='( /usr/bin/lssrc -g cluster      ;  
                 /usr/bin/lssrc -g lock        ;  
                 /usr/bin/lssrc -s clvmd       ;  
                 /usr/bin/lssrc -s clresmgrdES ;  
                 /usr/bin/lssrc -g topsvcs     ;  
                 /usr/bin/lssrc -g grpsvcs     ;  
                 /usr/bin/lssrc -g emsvcs     ) |  
                 /usr/bin/grep -vE "Sub|is not on file"'
```

```

alias show_cldl='( /usr/bin/echo ">--- clresmgrdES ---\n" ;
                  /usr/bin/lssrc -s clresmgrdES -l ;
                  /usr/bin/echo "\n>--- topsvcs -----\n" ;
                  /usr/bin/lssrc -s topsvcs -l ;
                  /usr/bin/echo "\n>--- grpsvcs -----\n" ;
                  /usr/bin/lssrc -s grpsvcs -l ;
                  /usr/bin/echo "\n>--- grpglsm -----\n" ;
                  /usr/bin/lssrc -s grpglsm -l ;
                  /usr/bin/echo "\n>--- emaixos -----\n" ;
                  /usr/bin/lssrc -s emaixos -l ;
                  /usr/bin/echo "\n>--- emsvcs -----\n" ;
                  /usr/bin/lssrc -s emsvcs -l ) |
                  /usr/bin/grep -vE "Sub|is not on file" |
                  /usr/bin/more'

alias show_tlog='/usr/bin/tail -f $( ls -tl /var/ha/log/topsvcs.* |
                                   /usr/bin/head -1 )'

alias show_mlist='/usr/bin/grep -v "^*" /var/ha/run/topsvcs.*/machines.*.lst'

alias stabilize_cluster='/usr/sbin/cluster/utilities/clruncmd $( hostname )'

alias takeover_resgr="/usr/bin/logger -t '
'>+++++
>/dev/null 2>&l ;
/usr/sbin/cluster/utilities/clstop -y -N -s -gr " \

alias start_cluster_node="/usr/bin/logger -t '
'>+++++
>/dev/null 2>&l ;
/usr/sbin/cluster/etc/rc.cluster -boot -N -i " \

alias start_cluster_cnode="/usr/bin/logger -t '
'>+++++
>/dev/null 2>&l ;
/usr/sbin/cluster/etc/rc.cluster -boot -N -i -l " \

alias stop_cluster_node="/usr/bin/logger -t '
'>+++++
>/dev/null 2>&l ;
/usr/sbin/cluster/utilities/clstop -y -N -s -g " \

# customer spezific alias -----

alias move_IP4_ghost1="/usr/sbin/cluster/utilities/cldare -M IP4:stop:sticky; \
/usr/sbin/cluster/utilities/cldare -M IP4:ghost1 ; "

```

Script to show the cluster log (show_clog)

```
#!/usr/bin/ksh
#set -x
#####
#
#. This script displays the HACMP cluster log on a node
#
#. Name      : show_clog
#. Parameters : [ <hostname> ]
#.          |
#. Default   : <local_host>
#
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
# This productivity aid is provided on an "AS-IS" basis
# WITHOUT ANY WARRANTY OF ANY KIND
#
#####

function show_tail
{
    LOG=$1
    NODE=$2
    typeset -L78 HDR
    N="${NODE:-$( /usr/bin/hostname )}:${LOG}"
    HDR="----- ${N} -----"
    /usr/bin/echo ${HDR}

    if [ -n "${NODE}" ]; then
        /usr/bin/rsh ${NODE} "/usr/bin/tail -f -21 ${LOG}"
    else
        /usr/bin/tail -f -21 ${LOG}
    fi
    return
}

# ----- main -----
NODE=$1

trap "exit 2" 2

show_tail /var/adm/cluster.log ${NODE}

exit 0
```

Script to show the script log (show_ctrace)

```
#!/usr/bin/ksh
#set -x
#####
#
#. This script displays the HACMP script log on a node
#
#. Name      : show_ctrace
#. Parameters : [ <hostname> ]
#.          |
#. Default   : <local_host>
#
# Copyright IBM Deutschland Informationssysteme GmbH 1998, 1999
# This productivity aid is provided on an "AS-IS" basis
# WITHOUT ANY WARRANTY OF ANY KIND
#
#####

function show_tail
{
    LOG=$1
    NODE=$2
    typeset -L78 HDR
    N="${NODE:-$( /usr/bin/hostname )}:${LOG}"
    HDR="----- ${N} -----"
    /usr/bin/echo ${HDR}

    if [ -n "${NODE}" ]; then
        /usr/bin/rsh ${NODE} "/usr/bin/tail -f -21 ${LOG}"
    else
        /usr/bin/tail -f -21 ${LOG}
    fi
    return
}

# ----- main -----
NODE=$1

trap "exit 2" 2

show_tail /tmp/hacmp.out ${NODE}

exit 0
```

Appendix D. Special notices

This publication is intended to help systems engineers, service providers, and customers to successfully plan and carry through migration from HACMP to HACMP/ES on RS/6000 systems. The information in this publication is not intended as the specification of any programming interfaces that are provided by HACMP and HACM/ES Version 4.3.1. See the PUBLICATIONS section of the IBM Programming Announcement for HACMP Version 4.3.1 and HACMP/ES Version 4.3.1 for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate

them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

This document contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples contain 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.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

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

AIX	AS/400
AT	CT
HACMP/6000	IBM ®
Netfinity	NetView
RISC System/6000	SP
SP2	System/390
400	

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere., The Power To Manage., Anything. Anywhere., TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli

A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

UNIX is a registered trademark in the United States and/or other countries licensed exclusively through X/Open Company Limited.

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix E. Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

E.1 IBM Redbooks publications

For information on ordering these publications see “How to get IBM Redbooks” on page 139.

- *HACMP Enhanced Scalability*, SG24-2081 (online version available)
- *HACMP Enhanced Scalability: User-Defined Events*, SG24-5327
- *Disaster Recovery with HAGEO: An Installer's Companion*, SG24-2018 (online version available)
- *HACMP/6000 Customization Examples*, SG24-4498
- *HACMP Enhanced Scalability Handbook*, SG24-5328

E.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at <http://www.redbooks.ibm.com/> for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

E.3 Other Publications

- *AIX Performance Monitoring and Tuning Guide*, SC23-2365

- *HACMP V4.3 Concepts and Facilities*, SC23-4276
- *HACMP V4.3 AIX Planning Guide*, SC23-4277
- *HACMP V4.3 AIX: Installation Guide*, SC23-4278
- *HACMP V4.3 AIX: Administration Guide*, SC23-4279
- *HACMP V4.3 AIX: Troubleshooting Guide*, SC23-4280
- *HACMP V4.3 AIX: Programming Locking Applications*, SC23-4281
- *HACMP V4.3 AIX: Program Client Applications*, SC23-4282
- *HACMP V4.3 AIX: Master Index and Glossary*, SC23-4285
- *HACMP V4.3 AIX: Enhanced Scalability Installation and Administration Guide*, SC23-4284
- *RSCT: Event Management Programming Guide and Reference*, SA22-7354
- *RSCT: Group Services Programming, Guide and Reference*, SA22-7355
- *PSSP: Administration Guide*, SA22-7348
- *PSSP: Installation and Migration Guide*, GA22-7347

E.4 Referenced Web sites

These Web sites are also relevant as further information sources:

- <http://hacmp.aix.dfw.ibm.com/htmls/howto/>
- http://www.rs6000.ibm.com/doc_link/en_US/a_doc_lib/aixgen/hacmp_index.html
- http://www.rs6000.ibm.com/resource/aix_resource/Pubs
- <http://www.rs6000.ibm.com/support>
- <http://www.ibm.com/servers/aix/products/ibmsw/list>

How to get IBM Redbooks

This section explains how both customers and IBM employees can find out about IBM Redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** <http://www.redbooks.ibm.com/>

Search for, view, download, or order hardcopy/CD-ROM Redbooks from the Redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

Redpieces are Redbooks in progress; not all Redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

Send orders by e-mail including information from the IBM Redbooks fax order form to:

	e-mail address
In United States	usib6fpl@ibmmail.com
Outside North America	Contact information is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the Redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and Redbooks by accessing the IBM Intranet Web site at <http://w3.itso.ibm.com/> and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at <http://w3.ibm.com/> for redbook, residency, and workshop announcements.

IBM Redbooks fax order form

Please send me the following:

Title	Order Number	Quantity

First name	Last name
------------	-----------

Company

Address

City	Postal code	Country
------	-------------	---------

Telephone number	Telefax number	VAT number
------------------	----------------	------------

Invoice to customer number

Credit card number

Credit card expiration date	Card issued to	Signature
-----------------------------	----------------	-----------

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

Glossary

- ACK.** Acknowledgment.
- AIX.** Advanced Interactive Executive.
- API.** Application Program Interface.
- ARP.** Address Resolution Protocol.
- ATM.** Asynchronous Transfer Mode.
- C-SPOC.** Cluster Single Point of Control facility.
- Clinfo.** Client Information Program.
- clsmuxpd.** Cluster Smux Peer Daemon.
- CNN.** Cluster Node Number.
- CPU.** Central Processing Unit.
- CWS.** Control Workstation.
- DARE.** Dynamic Automatic Reconfiguration Event.
- DMS.** Deadman Switch.
- DNS.** Domain Name Server.
- EM.** Event Management.
- FDDI.** Fiber Distributed Data Interface.
- FSM.** Finite State Machine.
- GID.** Group Identifier.
- GL.** Group Leader.
- GODM.** Global Object Data Manager.
- GS.** Group Services.
- GSAPI.** Group Services Application Programming Interface.
- HACMP.** High Availability Cluster Multi-Processing for AIX.
- HACMP/ES.** High Availability Cluster Multi-Processing for AIX Enhanced Scalability.
- HANFS.** High Availability Network File System.
- HB.** Heartbeat.
- HWAT.** Hardware Address Takeover.
- HWM.** High Water Mark.
- IBM.** International Business Machines Corporation.
- ICMP.** Internet Control Message Protocol.
- IP.** Interface Protocol.
- IPAT.** IP Address Takeover.
- ITSO.** International Technical Support Organization.
- JFS.** Journaled File System.
- KA.** Keep Alive.
- LAN.** Local Area Network.
- LPP.** Licensed Program Product.
- LVCB.** Logical Volume Control Block.
- LVM.** Logical Volume.

Index

A

Administration Guide 1
AIX Fast Connect 9
APAR 14

C

cascading resource 43
cascading resource group 45
cl_convert 32, 36
cl_opsconfig 47
clconvert.log 85
clconvert_snapshot 35
cldare 7, 63
clexit.rc 96
clhandle 101
clinfo 4
clinfo.rc 96
cllockd 4
cllsgnw 101
clmixver 101
clsmuxpd 3, 4
clstat 75, 101
clstrmgr 4
clstrmgr.debug 100
cluster administration 100
cluster log 96
cluster manager 68
cluster verification 40, 62
Cluster Worksheet 46
Cluster worksheet 24
cluster worksheet 1
cluster.log 100
cluster.man.en_US.client.data 38
clverify 2, 7, 40
clvmd.log 100
compatibility of HACMP/ES 20
Concurrent LVM 4
concurrent resource group 45
config_too_long 96
configurations 44
connectivity graph 10
convert HACMP to HACMP/ES 34
CRM 17
C-SPOC 1, 6, 8
cspoc.out 100

Customized event scripts 96
customized event scripts 28
CWS 117

D

daemons 66, 83
DARE 2, 6
dead man switch 97
dms_load.out 100
dsh 117
Dynamic reconfiguration 2

E

EMCDB 11
emsvcsctrl 101
emuhacmp.out 100
Error Notification 40
Error Notifications 28
ESCRM 8
event emulation 7
Event Management 11

F

fall-back plan 37
forced down 102
fsck 7

G

grace period 96
group service 83
Group Services 11
grpsvcctrl 101

H

HACMP and HACMP/ES Version 4.2.2 7
HACMP and HACMP/ES Version 4.3 8
HACMP and HACMP/ES Version 4.3.1 8
HACMP Version 1.1 3
HACMP Version 2.1 4
HACMP Version 3.1 4
HACMP version 3.1.1 5
HACMP version 4.1 5
HACMP version 4.1.1 5
HACMP version 4.2.0 6
HACMP Version 4.2.1 6

hacmp.out 100
HACMP/ES Version 4.2.1 6
HAGEO 6
hagscl 101
hagsgr 101
hagsms 101
hagsns 101
hagspbs 101
hagsreap 101
hagsvote 101
HANFS 17
HAS 17
heartbeats 85
high availability 1
HPS 5

I

IBM 7135-110 4
IBM 9333 4
install images 118

K

kerberos authentication 6

L

lock manager 68
log file 96
log files 68, 100

M

Maintenance levels 14
MIB 3
migrate 15
migration scenarios 55
mission critical applicatio 3
Mode 1 3
Mode 2 3
mutual takeover 45

N

Net View 6
NFS Version 3 8
Node-by-node migration 33
node-by-node migration 31

O

ODM conversion log 85

P

Planning 23
Planning Guide 1
post-event scripts 72
pre-event scripts 72
prerequisites 55
PSSP 6, 19, 118
PTF 14

R

Resource groups 4
resource groups 107
Resource monitors 11
RS/6000 SP 117
RSCT 8, 10

S

SAP environment 43
scripts 125
SCSI ID 57
SDR 120
show_cld 84
SMP 5
snapshot utility 27
SNMP 3
SP 117
state-transition 25
State-transition diagrams 25
symmetric multi processor 5

T

takeover 45
target mode SSA 13
test cases 29
test plan 29
TME 6
Topology service 83
Topology Services 10
topsvcsctrl 101

U

Unsupported levels of HACMP 22
upgrade 15

Upgrade paths 17
User-Defined Events 13

V

Version compatibility 1

W

WCOLL 117

IBM Redbooks evaluation

Migrating to HACMP/ES
SG24-5526-00

Your feedback is very important to help us maintain the quality of IBM Redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Which of the following best describes you?

Customer **Business Partner** **Solution Developer** **IBM employee**
 None of the above

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction _____

Please answer the following questions:

Was this redbook published in time for your needs? Yes___ No___

If no, please explain:

What other Redbooks would you like to see published?

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

SG24-5526-00
Printed in the U.S.A.

Migrating to HACMP/ES

SG24-5526-00

