

IBM Rational ClearCase[®] support for AIX PowerHA

This document describes details of IBM Rational ClearCase configuration on AIX High Availability Cluster Multi-Processing (HACMP). AIX HACMP has been rebranded as PowerHA. This document includes brief description of PowerHA configuration to support ClearCase.

Table of Contents

| | | |
|----------|--------------------------------------------------------|-----------|
| 1 | INTRODUCTION | 4 |
| 1.1 | PURPOSE | 4 |
| 1.1.1 | Scope | 4 |
| 1.2 | DEFINITIONS, ACRONYMS AND ABBREVIATION | 4 |
| 1.3 | REFERENCES | 5 |
| 2 | AIX HACMP OVERVIEW | 6 |
| 2.1 | HIGH AVAILABILITY V/S HARDWARE AVAILABILITY | 6 |
| 2.2 | HIGH AVAILABILITY V/S FAULT TOLERANCE | 6 |
| 2.3 | CLUSTER RESOURCES, RESOURCE GROUPS AND POLICIES | 6 |
| 2.4 | SERVER NODES AND CLIENT NODES | 7 |
| 2.5 | SUBNET ROUTING REQUIREMENTS IN POWERHA | 7 |
| 2.6 | IP ADDRESS TAKEOVER | 7 |
| 2.6.1 | IPAT via IP aliases | 7 |
| 2.6.2 | IPAT via IP replacement | 7 |
| 3 | AIX POWERHA CONFIGURATION | 8 |
| 3.1 | PLANNING | 8 |
| 3.1.1 | Cluster Overview Planning | 8 |
| 3.1.2 | Hardware Planning | 8 |
| 3.1.3 | Planning Cluster Software | 8 |
| 3.1.4 | PowerHA Software Bundle | 9 |
| 3.1.5 | Application Software | 9 |
| 3.1.6 | Licensing | 9 |
| 3.2 | CONFIGURATION | 9 |
| 3.2.1 | Define Cluster name and nodes | 10 |
| 3.2.2 | Configure PowerHA Networks | 10 |
| 3.2.3 | Configure PowerHA communication interfaces | 10 |
| 3.2.4 | Configure PowerHA service IP labels/addresses | 11 |
| 3.2.5 | Start cluster services | 11 |
| 3.2.6 | Configure shared volume group | 11 |
| 3.2.7 | Configure shared filesystem | 11 |
| 3.2.8 | Monitoring status of nodes in the cluster | 12 |
| 4 | CLEARCASE CONFIGURATION FOR POWERHA | 13 |
| 4.1 | GENERAL CLEARCASE CONFIGURATION CONSIDERATIONS | 13 |
| 4.2 | TYPES OF POWERHA CONFIGURATION | 13 |
| 4.2.1 | Standby Configuration – Standard | 14 |
| 4.3 | CLEARCASE RESOURCE GROUP | 14 |
| 4.3.1 | PowerHA ClearCase resource group configuration | 14 |
| 4.4 | APPLICATION SERVER | 14 |
| 4.4.1 | Start Script – cconline | 15 |
| 4.4.2 | Stop Script – ccoffline | 15 |
| 4.4.3 | Cleanup Method – cclean | 16 |
| 4.4.4 | Monitor Method – status.sh | 17 |
| 4.4.5 | PowerHA Application Server configuration | 17 |
| 4.4.6 | Configuration for PowerHA Application monitoring | 18 |
| 4.4.7 | Defining PowerHA resource group attributes | 18 |
| 4.5 | CLEARCASE STARTUP CONFIGURATION | 19 |
| 4.6 | ALTERNATE HOSTNAMES | 19 |
| 4.7 | REGISTRY SERVER | 19 |



| | | |
|----------|----------------------------------------------------------|-----------|
| 4.8 | LICENSE SERVER | 19 |
| 4.9 | VIEW AND VOB SERVERS | 20 |
| 4.9.1 | Shared storage on SAN | 20 |
| 4.9.2 | Shared storage on NAS or NFS mount..... | 20 |
| 5 | CLEARCASE CLIENT RECOVERY ON SERVER FAILOVER..... | 21 |



1 Introduction

This document describes details of ClearCase configuration for AIX PowerHA, previously known as AIX HACMP. AIX HACMP has been rebranded as AIX PowerHA. The document also provides brief description of PowerHA configuration and some considerations regarding ClearCase fail over on PowerHA.

1.1 Purpose

AIX PowerHA is a tool for building UNIX based, mission critical computing platforms. PowerHA provides a low-cost commercial computing environment that ensures quick recovery of mission critical applications from hardware and software failures. In PowerHA cluster, PowerHA software takes full control of the critical resources and applications that it is responsible for. When one node in the cluster fails PowerHA takes responsibility of moving the application, along with all the resources that are required by the application for its working, to another node in the cluster.

This document outlines the general concepts of PowerHA, a brief description of PowerHA configuration and details of the ClearCase configuration required to host it on AIX PowerHA. This document also provides similarities and differences between AIX PowerHA and Veritas Clustering System (VCS) with respect to ClearCase configuration. The document outlines what can be expected from a performance stand point when ClearCase fails over in the PowerHA environment.

1.1.1 Scope

The scope of the validation effort was to determine all the ClearCase 7.0.1 and ClearCase 7.1 configuration requirements to support PowerHA, investigate any differences in the configuration as compared to that on Veritas Clustering System (VCS) and to identify all the required modifications in the product to support PowerHA. The validation effort was successful in outlining all aspects of ClearCase configuration and its working in PowerHA environment and it was determined that there is no need to modify any component of the product for it to work with PowerHA. It was also found that the ClearCase 7.0.1 and ClearCase 7.1 configuration on PowerHA is very similar to that on Veritas Cluster System. Only ClearCase 7.0.1 and ClearCase 7.1 have been validated to support AIX PowerHA and this document details all the key observations and learning. PowerHA version 5.5 on AIX 6.1 was used for the validations. Some of the older AIX PowerHA documentation referenced here and the tools used for configuration still use the older brand name, AIX HACMP. In either case they refer to the same High Availability technology.

1.2 Definitions, Acronyms and Abbreviation

HACMP: High Availability (HA) and Cluster multi-processing (CMP).

HA: High Availability system is one which combines custom software with industry standard hardware to minimize downtime by quickly restoring services when a system, component or application fails.

CMP: Cluster multi-processing is a group of loosely coupled machines networked together, sharing disk resources.

Nodes: Nodes form the core of an PowerHA cluster. A node is a processor that runs AIX, the PowerHA software and the application software.

Client: A client is a node that can access the nodes in the cluster over a local area network. Client nodes each run a front end or a client application that queries the server application running on the cluster node.

Boot IP: The IP that has been assigned to the node when the node boots up.



Service IP/Virtual IP: An IP address over which services such as an application are provided and over which client nodes communicate.

Service IP label/Virtual Hostname: A label or a hostname that maps to the service IP address. This is the hostname the cluster would be identified with and is similar to the hostname allocated to any node in a network. All the clients communicate with the cluster using the service IP label and thus would never need to know which node in the cluster they are actually interacting with.

IPAT: IP Address Takeover (IPAT) is a networking mechanism that allows a node in the cluster to acquire a service IP from the failing node.

Heartbeat: It is a type of communication packet that is sent between nodes to monitor their health, networks and network interfaces and to prevent cluster partitioning.

Resource group: A set of interdependent resources such as disks, volumes, file systems, IP addresses and applications configured into a single logical entity for easier configuration and management.

Startup: Startup refers to the activation of a resource group on a node.

Fallover/Failover: Fallover or failover refers to the movement of the resource group that currently owns it to another active node in the cluster after the current node experiences a failure. The terms fallover and failover are synonymous in the context of this document and are used interchangeably.

Fallback: Refers to the movement of a resource group from the node that it currently resides on to a node that has joined the cluster when the current node is not the higher priority or home node.

Hardware Availability: Fault tolerant, redundant systems (such as RAID) and dynamic switching technologies (such as DLPAR) which provide recovery of certain hardware failures.

Fault Tolerance: A fault tolerant environment is one in which there is no service interruption. This requires specialized hardware to detect a hardware fault and instantaneously and seamlessly switch to a redundant hardware component.

1.3 References

- AIX HACMP, Administration and installation guide:
http://publib.boulder.ibm.com/infocenter/systems/index.jsp?topic=/com.ibm.aix.hacmp.concepts/ha_concepts_ha_cluster.htm&tocNode=toc:front/front.cmb/0/0/16/1/0/0/
- AIX HACMP Concepts and facilities guide:
<http://publib.boulder.ibm.com/infocenter/systems/topic/com.ibm.aix.hacmp.concepts/hacmpconcepts.pdf>
- Configuring ClearCase for High Availability:
http://publib.boulder.ibm.com/infocenter/cchelp/v7r0m1/index.jsp?topic=/com.ibm.rational.clearcase.admin.doc/topics/r_vcs_config_ha.htm
- Veritas Cluster Server Configuration:
http://publib.boulder.ibm.com/infocenter/cchelp/v7r0m1/index.jsp?topic=/com.ibm.rational.clearcase.admin.doc/topics/r_vcs_config_ha.htm

2 AIX HACMP Overview

AIX PowerHA provides a low cost commercial computing environment that ensures quick recovery of mission critical applications and resources from hardware and software failures. For example, in a client server application, the server and the resources it depends on can be single point of failure. In a HA system, the PowerHA cluster hosts the server with all its dependent files on a storage device that can be shared with other nodes in the cluster. If there is a failure at the server, PowerHA ensures that the server application and all its dependent resources are moved to a standby node in the cluster. This ensures high availability of the resources to the client with minimum downtime. Although the restoration of the service is rapid, this failover is not instantaneous. Typically the failover would take 30 to 300 seconds. This down time of the service would typically be perceived by a client-server application as a network failure. A failover in such an environment requires some special considerations regarding the behavior of the client and the continuity of the transaction during the failure. These would be discussed in detail in the context of ClearCase configuration in PowerHA.

2.1 High Availability v/s Hardware availability

High availability is sometimes confused with hardware availability. Though Hardware availability does provide recovery from certain hardware failures, it does not include a complete set of error detection and recovery systems to ensure critical and complex application availability. For example, any complex application would require access to processors, memory, storage devices, network interfaces, means of communication with other processes, etc. Failures in an application are typically due to operator errors, environmental problems, application and operating system errors. Reliable and recoverable hardware cannot protect against failures of this nature in the configuration.

2.2 High Availability v/s fault tolerance

High availability is viewed not as a series of replicated physical components (as is typical of fault tolerant systems), but as a set of system-wide, shared resources that cooperate to guarantee critical services. Fault tolerance has a high cost of hardware and performance associated with it including lower utilization of resources.

HA typically does not mean zero down time as a fault tolerant system might promise and is applicable in an environment where a short service interruption is acceptable if a failure does occur.

2.3 Cluster resources, resource groups and policies

PowerHA provides interfaces to identify cluster resources that are essential to the operation of an application and combining them into a resource group. PowerHA also enables resource group management during the typical cluster operations of startup, fallover and fallback by defining the resource group policies and attributes. Cluster resources could be storage disks, volume groups, logical volumes, file systems, service IP, service IP label, applications and any of its dependencies. Cluster resource groups can be concurrent or non-concurrent. Concurrent resource groups are ones which are available simultaneously on multiple nodes in the cluster. Non-concurrent resource groups are ones which are available only on one node in the cluster at any point in time.

A home node is the first node in the node list for a particular resource group. The home node typically owns the non-concurrent resource group. However, a non-concurrent resource group may or may not have a home node associated with it depending on the resource group policies of start, fallover and fallback. Fallover or failover is a term used to refer to the movement of the resource group from the node that currently owns to another node in the cluster when the current node experiences a failure. A resource group movement from the node where it currently resides to the one that is rejoining the cluster is called a fallback operation. For example, if a home node is associated with a non-concurrent resource group and it is rejoining the cluster after being out of it due to a failure, then the resource group movement from the

node that it currently resides on, to the home node is called a fallback. Fallback can happen even when a higher priority node as compared to the one where the resource group resides rejoins the cluster. All this can be controlled via the resource group policies. For example, the fallover (or failover) policy could be “Fallover To Next Priority Node In The List” and the fallback policy could be “Fallback To Higher Priority Node In The List”.

2.4 Server nodes and client nodes

A cluster server node typically runs a critical application or service that accesses data on a shared external storage. These server nodes also run PowerHA daemons which constantly monitor the health of the resource groups configured on the nodes. The storage can be either shared between the server nodes, in other words it can be configured to be a concurrent resource, for example, a NAS storage configuration. The storage can be a non-concurrent resource and might be available to only one node at any point in time, for example, a SAN storage configuration. In either case, PowerHA can be delegated the responsibility of ensuring the availability of the resource.

A client node is one which is utilizing the services offered by the server. Client nodes can be categorized as naive or intelligent. Naive clients are not aware of the cluster and view it as a single entity. Thus, if a server fails, the client must be restarted or at the least it must reconnect to the server. An intelligent client on the other hand is cluster aware and reacts appropriately in case of a server failure. The intelligent client could mask the failure from the user while it connects to the alternate server and in order to achieve that the client must have knowledge of the cluster status.

2.5 Subnet routing requirements in PowerHA

AIX PowerHA recommends that each communication interface on the node belongs to a different subnet, so that PowerHA can monitor the health of each interface. In some cases PowerHA might enforce such a requirement. In general, PowerHA does not distinguish between logical network interfaces that share the same subnet route. If a logical network interface shares a route with another interface, PowerHA has no way to monitor its health. For example, an IP that shares the same subnet with the boot IP configured on the existing network interface of the node cannot be configured as service IP by creating an alias. In networks that restrict subnet tunneling, a second physical network interface is required on the nodes for the service IP configuration.

2.6 IP address takeover

IP address takeover (IPAT) is a mechanism through which a service IP and service label (virtual hostname) can be moved from the failing node to another node in the cluster. Through IPAT the virtual hostname over which the clients are provided the services can remain available after the failover. PowerHA provides two methods to perform IPAT, namely, IPAT via IP aliases and IPAT via IP replacement.

2.6.1 IPAT via IP aliases

This is the default configuration in IPAT. The service IP is aliased onto the same network interface as an existing communications interface. That is, multiple IP addresses are configured on the same network interface at the same time. In this configuration, all IP addresses that are configured must be from different subnets. If the IP configuration for a PowerHA network is via IP aliases, the communicating interfaces for that PowerHA network must use routes that are different from the one used by the service IP. Running IP takeover via IP aliasing is faster than running IPAT via IP replacement, because moving the IP addresses and the hardware addresses (IP replacement) takes considerably longer than just moving the IP address (IP aliases).

2.6.2 IPAT via IP replacement

The service IP address replaces the existing IP address on the network interface. That is, only one IP

address is configured on an interface at any point in time. In this configuration, two IP addresses on a node can share a subnet, while a backup IP address on the node must be on a different subnet. This method can save subnets but requires additional hardware. If the IP configuration for a PowerHA network is via IP replacement, only one communication interface for that PowerHA network must use a route that is the same as the one used by the service IP address.

3 AIX PowerHA Configuration

3.1 Planning

3.1.1 Cluster Overview Planning

PowerHA configuration requires details such as virtual host name, names for each of the nodes in the cluster, etc. A simple planning worksheet is provided below for a cluster with two nodes.

| | |
|----------------------------|-------------|
| CLUSTER NAME | testcluster |
| NODE 1 HOSTNAME | crmha1 |
| NODE 2 HOSTNAME | crmha2 |
| NODE 1 PowerHA NAME | nodeA |
| NODE 2 PowerHA NAME | nodeB |
| VIRTUAL HOST NAME | crmha |

3.1.2 Hardware Planning

PowerHA supports virtually any AIX supported node, from desktop systems to high end servers. When choosing a type of node, consider the following:

- Ensure that all the nodes in the cluster have sufficient system resources to sustain the demands during application failover.
- The nodes in the cluster should be of similar configuration so that resource distribution and administration are simpler and resource demands during failover are satisfied.

3.1.3 Planning Cluster Software

Review all software components to be used in the cluster to ensure compatibility. Items to consider are AIX, RSCT, PowerHA, application, and storage software. This section discusses the various software levels and compatibilities.

3.1.3.1 AIX and RSCT levels

The IBM® Reliable Scalable 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. RSCT handles the heartbeats and network failure detection. The PowerHA™ and RSCT software stack runs on each cluster node. Need to install RSCT filesets specific to the AIX version installed on the nodes. The following filesets are required at the minimum:

- rsct.compat.basic.hacmp
- rsct.compat.clients.hacmp
- rsct.core.rmc

- rsct.core.sec

3.1.3.2 Required AIX Filesets

The following AIX filesets are required for PowerHA. They must be installed with the latest version of fixes for the appropriate AIX level before PowerHA is installed.

- bos.adt.lib
- bos.adt.libm
- bos.adt.syscalls
- bos.net.tcp.client
- bos.net.tcp.server
- bos.rte.SRC
- bos.rte.libc
- bos.rte.libcfg
- bos.rte.libcur
- bos.rte.libpthreads
- bos.rte.odm
- bos.rte.lvm.rte (required only using existing or legacy 32-bit Concurrent Logical Volume Manager for concurrent access)
- bos.clvm.enh (required for Enhanced Concurrent Logical Volumes – disk heartbeating makes use of this feature)

3.1.4 PowerHA Software Bundle

Install PowerHA software bundle from the install media procured from IBM.

3.1.5 Application Software

Typically applications are not dependant on PowerHA versions as they are not aware of the underlying PowerHA functionality. That is, PowerHA simply starts and stops them (PowerHA can also monitor applications, but generally using an application dependent method).

Install the version of ClearCase that is intended to be used on all the nodes in the cluster. Refer to section 4 for details regarding ClearCase installation and configuration for PowerHA.

3.1.6 Licensing

PowerHA is not bundled with AIX. PowerHA software package needs to be purchased separately. Licenses to run ClearCase on the nodes have to be procured from IBM. Refer to section 4.8 for special considerations in ClearCase license procurement.

3.2 Configuration

The following steps detail the basic configuration in PowerHA required for enabling ClearCase for High Availability. “smit” is the tool used for all these configurations.



3.2.1 Define Cluster name and nodes

1. Start **smit hacmp** on one of the nodes in the cluster. This is the node that will be used for all further PowerHA configurations except when stated otherwise.
2. Select **Extended Configuration**.
3. Select **Extended Topology Configuration**.
4. Select **Configure an HACMP cluster**.
5. Select **Add/Change/Show an HACMP Cluster**.
6. Enter the desired cluster name and press enter.
7. Go back to step 4.
8. Select **Configure HACMP Nodes**.
9. Select **Add a Node to the HACMP Cluster**
10. Enter the desired name for the first node that is being added. For example, nodeA (as denoted in the planning sheet in section 3.1.1).
11. Enter the communication path to the new node or press F4 to get a list of IP labels from /etc/hosts file. Add nodes using one of their base addresses as a communication path to the node.
12. Repeat the step 8 to step 10 for every node in the cluster. For example, nodeB for the second node in the cluster, etc.

3.2.2 Configure PowerHA Networks

1. Start **smit hacmp**
2. Select **Extended Configuration**
3. Select **Extended Topology Configuration**
4. Select **Configure HACMP Networks**
5. Select **Add a Network to the HACMP Cluster**
6. Select “ether” from the network type list. Enter the desired network name (defaults to net_ether_01)

3.2.3 Configure PowerHA communication interfaces

1. Start **smit hacmp**.
2. Select **Extended Configuration**
3. Select **Extended Topology Configuration**
4. Select **Configure HACMP Communication Interfaces/Devices**
5. Select **Add Communication Interfaces/Devices**
6. Select **Add Pre-defined Communication Interfaces and Devices**
7. Select **Communication Interfaces**
8. Choose the network name created in the previous step (net_ether_01) and press enter.
9. Go to **IP Label/Address** tab, press F4 and select corresponding IP label of the host for which this communication interface is being created. Service IP label is the virtual hostname that will be used by clients to communicate with the cluster. The Service IP and Service Label must be added to the /etc/hosts file on node.

10. Repeat steps 5 to 9 for all nodes in the cluster.

3.2.4 Configure PowerHA service IP labels/addresses

1. Start **smit hacmp**.
2. Select **Extended Configuration**
3. Select **Extended Resource Configuration**
4. Select **HACMP Extended Resources Configuration**
5. Select **Configure HACMP Service IP Labels/Addresses**
6. Select **Add a Service IP Label/Address**
7. Select **Configurable on Multiple Nodes**
8. Select the network name that was created in the section 3.2.2.
9. Press F4 in the IP Label/Address field and select the service IP Label name that was added to /etc/hosts file.

3.2.5 Start cluster services

1. Select **smit hacmp**.
2. Select **System Management (C-SPOC)**.
3. Select **Manage HACMP Services**.
4. Select **Start Cluster Services**.

3.2.6 Configure shared volume group

1. Select **smit hacmp**.
2. Select **System Management (C-SPOC)**.
3. Select **HACMP Logical Volume Management**.
4. Select **Shared Volume Group**.
5. Select **Create a Shared Volume Group**.
6. Select the names of the nodes of the cluster between which the shared volume group has to be shared.
7. Now PVIDs of shared SAN disks are displayed. Select PVIDs of disks which you want to be part of volume group by pressing F7. Press Enter.

3.2.7 Configure shared filesystem

1. Select **smit hacmp**.
2. Select **System Management (C-SPOC)**.
3. Select **HACMP Logical Volume Management**.
4. Select **Shared File Systems**.
5. Select **Enhanced Journaled File Systems**.
6. Select **Add an Enhanced Journaled File System**.
7. From the list of shared volume groups, select the one in which the shared filesystem has to be created. Press enter.

8. Enter "SIZE of file system", "MOUNT POINT" and other field if required. Press enter.

NOTE:

File system size can be increased/decreased using OS commands any time during the operation of HACMP. However, after changing the size of the filesystem, the storage attributes in PowerHA on all the nodes in the cluster have to be synchronized for proper storage migration between the nodes during failover. The synchronization can be done by running: **Verify and Sync** from **smit hacmp** -> **Initialization and Standard Configuration** -> **Verify and Synchronize HACMP Configuration** -> Press Enter.

3.2.8 Monitoring status of nodes in the cluster

1. Export the path /usr/es/sbin/cluster/utilities and /usr/sbin/cluster using the following command:
export PATH=/usr/es/sbin/cluster/utilities:/usr/sbin/cluster:\$PATH
2. Run the command **clRGinfo** (command is case sensitive) to get the current status of the nodes in the cluster. A sample output is provided below:

```
# clRGinfo
```

```
-----
Group Name      Group State      Node
-----
ClearCase_RG    ONLINE           nodeA
                  OFFLINE          nodeB
```

In the above example, "ClearCase_RG" is the resource group that has been created for ClearCase (Refer to section 4.3 for details). The Cluster has two nodes nodeA and nodeB. The output shows that the ClearCase resource group is online on nodeA. nodeB is on standby.

3. To get detailed information of the nodes in the cluster and their status use **clstat**. A sample output is provided below:

```
# clstat
```

```
clstat - HACMP Cluster Status Monitor
```

```
-----
Cluster: testcluster (1236081144)
```

```
Tue Apr 7 11:55:50 2009
```

```
State: UP          Nodes: 2
```

```
SubState: STABLE
```

```
Node: nodeA        State: UP
```

```
Interface: service1 (0) Address: 10.100.100.1
```

```
State: UP
```

```
Interface: crmha (0) Address: 9.124.36.189
```

```
State: UP
```

```
Resource Group: ClearCase_RG State: On line
```

```
Node: nodeB        State: UP
```



Interface: service2 (0) Address: 10.100.100.2
State: UP

In the above example, crmha is the virtual hostname or the service IP label and 9.124.36.189 is the service IP. The service IP is currently configured on nodeA and ClearCase_RG resource group is online on nodeA.

4 ClearCase configuration for PowerHA

Rational ClearCase is supported on AIX PowerHA and all the components of ClearCase namely, view servers, VOB servers, license servers, registry servers, VOB RPC servers can be run in the HA environment. In order to support PowerHA, ClearCase installations on all the nodes in the cluster have to be configured manually. There is no out of the box installation and configuration available at this point in time to have ClearCase up and running in the HA environment. There are some special considerations and requirements stipulated by PowerHA for ClearCase to work in the HA system. The details are explained in the following sections. The ClearCase configuration on AIX PowerHA is very similar to that on Veritas Cluster System. The differences are highlighted as and when applicable. The ClearCase product documentation includes Veritas Configuration instructions.

4.1 General ClearCase configuration considerations

The following are the considerations while configuring ClearCase on all the nodes in an AIX PowerHA cluster:

- All nodes in the cluster must have the same version and patch level of ClearCase installed.
- Failover of view servers with views exported for non-ClearCase access are not supported in HA environment.
- All the ClearCase clients refer to the ClearCase servers via the virtual hostname only. For example, crmha (as denoted in the planning sheet in section 3.1.1)
- A shared storage needs to be configured for use by ClearCase. The ClearCase registry database, view storage and VOB storage directories need to be shared between ClearCase installed on all nodes in the cluster for high availability of ClearCase to work. This shared storage could be a NAS or SAN device or any NFS mount. The storage needs to be accessible by all nodes in the cluster.
- Restoration of normal ClearCase service on the node where the application failed might need administrator intervention similar to the one needed for ClearCase failure on any machine independent of a HA system. ClearCase as an application is not PowerHA aware. Any administration in case of a ClearCase failure on a node in the cluster would be specific to ClearCase.

4.2 Types of PowerHA configuration

There are several PowerHA configurations that can be considered to host ClearCase on. Some of them are listed below:

- Standby Configuration - Standard.
- Standby Configuration - Extended.
- Two nodes mutual take over configuration for concurrent access.
- Eight nodes mutual take over configuration for concurrent access.

The standby configuration is similar to the asymmetric (active/standby) configuration in Veritas Clustering

System. The Standby Configuration – Standard was used for the support validations.

4.2.1 Standby Configuration – Standard

In this configuration the resource groups are online only on the home node and the other node is configured as a standby. The resource group policy settings will be as follows:

- Startup Policy: Online On Home Node Only
- Fallover Policy: Fallover To Next Priority Node In The List
- Fallback Policy: Fallback To Higher Priority Node In The List

All the cluster resources are configured in a non-concurrent resource group. If the server node leaves the cluster or the application fails then the standby node assumes control of the resource group and starts the critical applications as configured in the resource group.

4.3 ClearCase resource group

All the resources required by ClearCase have to be grouped together in one non-concurrent resource group. Typically this resource group would contain:

- The service IP and Service IP label (virtual hostname).
- The application server (ClearCase albd server)

ClearCase registry server, view servers and VOB servers depend on the storage. For a failover to work this storage has to be shared and accessible to all nodes. Storage required for ClearCase can be configured on NAS. If so it will not be part of the non-concurrent resource group mentioned above as it would be accessible by all nodes in the cluster concurrently. If the storage is on a SAN device, it has to be included in the non-concurrent resource group mentioned above. When ClearCase fails over, PowerHA would take care of moving the SAN disk from the master to the standby.

4.3.1 PowerHA ClearCase resource group configuration

1. Select **smit hacmp**
2. Select **Extended Configuration.**
3. Select **Extended Resource Configuration.**
4. Select **HACMP Extended Resource Group Configuration.**
5. Select **Add a Resource Group.**
6. Enter the name of the resource group, for example, “ClearCase_RG”. Select the participating node names (press F4 to pop-up the node names). Set the startup, fallover and fallback policy (press F4 to get available options).
7. Verify and Synchronize HACMP Configuration. To do this, **smit hacmp -> Initialization and Standard Configuration -> Verify and Synchronize HACMP Configuration -> Press Enter**

4.4 Application server

The purpose of a HA system is to ensure that the critical applications and resources are available to clients. The ClearCase server to be monitored has to be configured in PowerHA. “albd” is the server that is configured to be monitored by PowerHA. That is because albd acts as a monitor process for the other servers in ClearCase. We register a monitor script with PowerHA for this operation. Similarly, in order to start ClearCase on a node and stop ClearCase, scripts are registered. A separate script takes care of clean up in case of an abnormal shutdown of ClearCase. These scripts could be physically located in any directory accessible by root. These scripts are not installed with Rational ClearCase and are provided here as an example. Refer to section 4.4.5 and 4.4.6 for steps to register these scripts with PowerHA.

4.4.1 Start Script – cconline

```
#!/bin/ksh
# cconline - Start ClearCase
#
if [ ! -f /opt/rational/clearcase/bin/cleartool ] ; then
    echo "this is not a ClearCase host"
    exit
fi

# Start ClearCase
/opt/rational/clearcase/etc/clearcase start

# Make a registry master server
# (Only needed if the registry server is part of the VCS service group.)
/opt/rational/clearcase/bin/cleartool setrgysvrtype -master

# unexport all the NFS exports and re-export them.
/usr/sbin/exportfs -ua
/usr/sbin/exportfs -a
```

4.4.2 Stop Script – ccoffline

```
#!/bin/ksh
# ccoffline - Stop ClearCase cleanly, wait for all ClearCase processes to exit
#
PSOPTS="-ef"

if [ ! -f /opt/rational/clearcase/bin/cleartool ] ; then
    echo "this is not a ClearCase host"
    exit
fi

# Demote the registry master to standby
# (Only needed if the registry server is part of the VCS service group.)
/opt/rational/clearcase/bin/cleartool setrgysvrtype -standby

# Stop ClearCase
/opt/rational/clearcase/etc/clearcase stop
```



```

# Wait for all ClearCase processes to exit
ccend="0"
while [ "$ccend" == 0 ]
do
    PID=`ps $PSOPTS | egrep
'(albd_server|admin_server|ccfs_server|credmap_server|db_server|mntrpc_server|msadm_server|promote_se
rver|shipping_server|syncmgr_server|view_server|vob_server|vobrpc_server)' | grep -v grep`
    if [ "${PID}" == "" ] ; then
        echo "ClearCase servers have exited."
        ccend="1"
    else
        sleep 1
    fi
done

```

4.4.3 Cleanup Method – ccclean

```

#!/bin/ksh
PSOPTS="-ef"

if [ ! -f /opt/rational/clearcase/bin/cleartool ] ; then
    echo this is not a ClearCase host
    exit
fi

# Attempt to stop ClearCase cleanly
PROC=`ps $PSOPTS | grep albd_server | grep -v grep`
if [ "${PROC}" != "" ] ; then
    /opt/rational/clearcase/etc/clearcase stop
fi

# Make sure all ClearCase processes have stopped
PID=`ps $PSOPTS | egrep
'(albd_server|admin_server|ccfs_server|credmap_server|db_server|mntrpc_server|msadm_server|promote_se
rver|shipping_server|syncmgr_server|view_server|vob_server|vobrpc_server)' | grep -v grep | awk '{print
$2;}'`
if [ "${PID}" != "" ] ; then
    echo "Attempting to kill ClearCase servers."
    for i in ${PID}

```

```

do
    kill $i
    sleep 1
done
else
    echo "ClearCase servers have exited."
fi

```

4.4.4 Monitor Method – status.sh

```

#!/bin/ksh
OPSTATE_ONLINE=0
OPSTATE_OFFLINE=1
Action=${1}
case ${Action} in
    status)
        mypid=$(ps -aef | grep "\opt\rational\clearcase\etc\albd_server"|tr -s " "|awk '{if ($8=="\opt\rational\clearcase\etc\albd_server") print $2;}')
        if [ ! -z $mypid ]
        then
            RC=${OPSTATE_ONLINE}
        else
            RC=${OPSTATE_OFFLINE}
        fi
        ;;
esac
exit $RC

```

4.4.5 PowerHA Application Server configuration

1. Select **smit hacmp**.
2. Select **Extended Configuration**.
3. Select **Extended Resource Configuration**.
4. Select **HACMP Extended Resources Configuration**.
5. Select **Configure HACMP Application Servers**.
6. Select **Configure HACMP Application Servers**.
7. Select **Add an Application Server**.
8. Enter the desired name for the Application Server. This is just an identifier for the application that is being configured for HA, for example, "ClearCaseServer". This is not the actual name of the process or daemon that needs to be monitored.
9. Enter the name of start script and stop script and the absolute path of the scripts. Refer to section

4.4.1 and 4.4.2 for an example of start and stop scripts.

4.4.6 Configuration for PowerHA Application monitoring

1. Select **smit hacmp**.
2. Select **Extended Configuration**.
3. Select **Extended Resource Configuration**.
4. Select **HACMP Extended Resources Configuration**.
5. Select **Configure HACMP Application Servers**.
6. Select **Configure HACMP Application Monitoring**.
7. Select **Configure Custom Application Monitors**.
8. Select **Add a Custom Application Monitor**.
9. Select **Application Server name to monitor**.
10. Enter **Monitor Name**. This is an identifier for the component of the application to be monitored, for example “ALBDserver”. In case of ClearCase, ALBD is the daemon that PowerHA is configured to monitor.
11. Select **Monitor Mode** from the list. The options are “Startup Monitoring”, “Long-running monitoring” and “Both”. PowerHA would monitor the application only during system startup if Startup Monitoring is set. Long-running monitoring is for PowerHA to monitor the application continuously when the system is up and running. For ClearCase choose Long-running monitoring.
12. Enter absolute path of the monitor script for **Monitor Method** along with any arguments that might be required by the script. This is the exact command line used to invoke the script. Refer to section 4.4.4 for an example of monitor script. In the example in section 4.4.4 the argument to the script would be “status”.
13. Enter absolute path of the cleanup script for **Cleanup Method** along with any arguments required by the script. Refer to section 4.4.3 for an example of cleanup script.
14. Enter the absolute path of the script along with its arguments to start and restart ClearCase for **Restart Method**. Refer to section 4.4.2 for an example of start script.
15. Enter the other mandatory fields such as **Stabilization Interval**.
16. Enter a **Restart Interval**. This is the time interval PowerHA waits before restarting the application on the standby node when the master node fails. For ClearCase this could be as low as 5 units.
17. Select “fallover” action from the list for **Action on Application Failure**.

4.4.7 Defining PowerHA resource group attributes

1. Select **smit hacmp**.
2. Select **Extended Configuration**.
3. Select **Extended Resource Configuration**.
4. Select **Hacmp Extended Resource Group Configuration**.
5. Select **Change/Show Resources and Attributes for a Resource Group**.
6. Select the ClearCase resource group name that was created in section 4.3.1 and press enter.
7. Add Application server name that was created in section 4.4.5.
8. Enter **Volume Groups** name and the **Filesystems (empty is ALL for VGs specified)** with those created as per sections 3.2.6, 3.2.7 and 4.4.5.

9. Enter the name of the shared volume group that has been created. Adding the shared volume group to the resource group attribute is needed if it is created using SAN. Refer to section 3.2.6 for steps to create a shared volume group on SAN. If the storage is on NAS, then the setup should be such that all the nodes in the cluster have concurrent access to it. Thus, it would not be part of the ClearCase non-concurrent resource group. Refer to section 4.3.1 for details.
10. Verify and Sync. This can be done using: `smit hacmp -> Initialization and Standard Configuration -> Verify and Synchronize HACMP Configuration -> Press Enter`

4.5 ClearCase startup configuration

When ClearCase is installed on AIX, an entry is added to `/etc/inittab` to start ClearCase at boot time. The entry would be in the following format: `"clearcase:2:wait:/etc/rc.clearcase"`. This entry means, when run level 2 is reached during system boot, the script `/etc/rc.clearcase` is invoked. `/etc/rc.clearcase` in turn invokes the ClearCase startup script (`/opt/rational/clearcase/etc/clearcase`). In the PowerHA environment, the complete control of starting and stopping ClearCase needs to be with PowerHA. Refer to section 4.4 for the PowerHA application server configurations involving `"cconline"` and `"ccoffline"` scripts. As the responsibility of ClearCase startup and stop would be delegated to PowerHA, ClearCase autostart needs to be disabled. This can be done by commenting out the `"clearcase:2:wait:/etc/rc.clearcase"` entry in `/etc/inittab` by placing `":"` at the beginning of the line.

4.6 Alternate Hostnames

As explained in the previous section, all the ClearCase clients refer to the cluster via the virtual hostname or service IP label only. At no point in time the ClearCase client should need to be aware of the physical hostname of the node in the cluster that it is communicating with. Thereby, if the view, VOB, license and registry servers are running in a cluster, a client would refer to all of them using the virtual hostname of the respective cluster they are hosted on. ClearCase on every node in the cluster should also be made aware that clients could be referring to it using a hostname other than the physical one. This is done by adding the virtual hostname (service IP label) to the `"/var/adm/rational/clearcase/config/alternate_hostnames"` file on each cluster node.

4.7 Registry Server

If the registry server is intended to be on a PowerHA cluster then ClearCase registry files must be moved to the shared storage. As mentioned in the previous sections, storage can be on NAS, SAN or an NFS mount. The shared location has to be accessible by all nodes in the cluster. The root account on all cluster nodes must have read and write permissions to the shared volume. The following are the steps to move the registry files to the shared storage:

- Login to the node that will host the registry server as root.
- Stop ClearCase on the node if it is running: `/opt/rational/clearcase/etc/clearcase stop`.
- Copy the registry directory to the shared storage:
`cp -r /var/adm/rational/clearcase/rgy SHARED_STORAGE`
- Rename the old registry directory:
`mv /var/adm/rational/clearcase/rgy /var/adm/rational/clearcase/rgy.old`
- Create a symlink for the registry directory on the shared storage in place of the original directory:
`ln -s SHARED_STORAGE /var/adm/rational/clearcase/rgy`.

4.8 License Server

The Rational ClearCase license keys are stored in the license database `"/var/adm/rational/clearcase/license.db"`. License keys are tied to a host ID and thus are unique to the



physical machine. If ClearCase License Server is intended to be hosted on PowerHA, then duplicate set of licenses need to be procured for every node and stored locally on the node. Customers intending to host license servers on PowerHA should get in touch with IBM sales representatives to obtain duplicate license keys.

FLEXIm, the third party license server has its own form of failover mechanism independent of PowerHA or any other HA technology. Customers who intend to use FLEXIm need to get in touch with Acreso Software (<http://www.acresso.com>) for details.

4.9 View and VOB servers

The view and VOB storage must be shared with all the nodes in the cluster. Depending on the choice of storage, view and VOB creation could be slightly different.

4.9.1 Shared storage on SAN

The shared storage for ClearCase can be on a SAN device. Refer to section 3.2.6 for steps to create shared volume groups using SAN. A SAN device would appear as any other local device to the operating system. Thus in order to use SAN device with ClearCase it has to be exported via NFS. This can be done by adding it to the `/etc/exports` file and running “`exportfs -a`”.

Views and VOBs can then be created as in the normal case.

For example:

To create a view on the virtual host “`ccha`” and the view storage to be NFS exported SAN disk, “`ccsan`” attached to it:

```
cleartool mkview -tag hacmp_test /net/ccha/ccsan/viewstore/hacmp_test.vws
```

Note: Always use the virtual hostname in the `gpaths` and not the physical hostname.

If SAN is the preferred choice of storage, then the ClearCase resource group configuration must include it. PowerHA in case of a failure would then take care of unmounting the SAN disk from the current node and attaching it to the standby node before bringing up ClearCase on the standby.

4.9.2 Shared storage on NAS or NFS mount

All views and VOBs managed by the PowerHA cluster should be created using the tuple “`-host`”, “`-hpath`” and “`-gpath`” if the storage is an NFS mount or a NAS device. In the HA environment the tuple is required to derive the hosting server unlike the normal case where just the “`-gpath`” is sufficient.

For example:

To create a view on the virtual host `ccha`, with the storage on NAS, `ccnas`:

```
cleartool mkview -tag hacmp_test -host ccha -hpath /net/ccnas/viewstore/hacmp_test.vws -gpath /net/ccnas/viewstore/hacmp_test.vws /net/ccnas/viewstore/hacmp_test.vws
```

The NFS mount should be accessible from all the nodes in the cluster. The view and VOB owners must also have read and write permissions to the NAS volume. In case of NAS or NFS mount, the storage need not be part of the ClearCase non-concurrent resource group as it would necessarily be accessible from all nodes in the cluster even when the node in question is not the current master.

5 ClearCase client recovery on server failover

There are some considerations with regard to client recovery in the wake of a server failover. They are listed below:

- From the ClearCase client's perspective the server failover is seen as a temporary network disruption.
- Some ClearCase client operations might recover before the RPC timeouts but some might not.
- If an RPC time out occurs, the ClearCase client would report a timeout or connection error.
- In cases where ClearCase client operations timeout or connection error is encountered, the operations need to be tried again.
- As explained in section 2 of the document, the application failover is not instantaneous. The ClearCase clients could encounter operation failures during certain server failure conditions and might have to be restarted.
- Certain commands such as "cleartool lsvob", "clearfsimport" report server connection errors and might face RPC timeouts as well. These commands pause for a brief period before retrying. Considering that in this duration the server failover was successful and ClearCase is up and running on the standby node, when the command tries to establish the connection again, albd spawns the necessary servers on the standby node of the cluster and the client operation resumes. This entire cycle from the time server fails to the time the client resumes operation might take anywhere between 2 to 5 mins. This could vary based on certain settings in PowerHA and certain ClearCase operations.
- Some ClearCase client operations might fail when the server node fails. For example, while a client is importing a set of files to the VOB using "clearfsimport", if the server node crashes or is shutdown, "clearfsimport" operation fails although the server failover was successful. Not all ClearCase client operations are expected to work when the server node experiences complete system failure. Once the standby is online, the client operation can be restarted.