High Availability Add-On for Red Hat Enterprise Linux on HP ProLiant Servers

Linux References and Best Practices

Technical white paper

Table of contents

Introduction

The High Availability Add-On for Red Hat Enterprise Linux and its associated products are used to construct clusters for:

- High availability for applications
- Network load balancing across multiple servers
- File system sharing across multiple servers
- High performance computing

HP ProLiant servers are an excellent choice for any of these cluster use cases. Red Hat and HP provide extensive documentation to help you achieve optimum results when running the High Availability Add-On and other associated products on ProLiant servers. These documents are found in a variety of locations. The purpose of this document is to help you find the information you need quickly – it is a single source directing you to reference and best practice information pertaining to both Red Hat cluster products and HP ProLiant servers.

A Red Hat cluster may contain a number of Red Hat products, including Red Hat Enterprise Linux (RHEL), the High Availability Add-On for Red Hat Enterprise Linux, and other associated Red Hat products. This document uses the term "RHEL cluster" to represent these products.

Intended Audience

This white paper is intended for HP and Red Hat customers, and HP and Red Hat personnel who design or manage RHEL clusters running on HP ProLiant servers. Readers are expected to have a basic understanding of Red Hat Enterprise Linux, the High Availability Add-On and associated products, and HP ProLiant server concepts.

Pointers to Red Hat documentation

Red Hat provides product documentation in the form of both manuals and Red Hat Knowledgebase articles.

Red Hat manuals

Manuals for both RHEL5 and RHEL6 clusters can be found on the following publicly-available Red Hat product documentation page, in both HTML and PDF formats:
https://access.redhat.com/knowledge/docs/Red_Hat_Enterprise_Linux

The RHEL5 titles include:

- *Cluster Suite Overview* – the place to start for an introduction to RHEL clustering on RHEL5.

- *Cluster Administration* – provides general configuration and routine management information.

- *Virtual Server Administration* – provides a functional overview and configuration information for this network load-balancing software.

- *Global File System 2* – documents the features of the Red Hat second-generation cluster file system (GFS2), and provides configuration, management, and troubleshooting information. Discusses the differences between the first generation GFS and GFS2, and provides conversion procedures.   GFS2 offers significant advantages over GFS on RHEL5.

- *Global File System* – documents the features of the Red Hat first generation cluster file system (GFS), and provides configuration and management information. GFS2 offers significant advantages over GFS on RHEL5.

- *Global Network Block Device* – explains the use of Red Hat's GNBD with the GFS cluster file system.

- *Logical Volume Manager Administration* – not exclusively for RHEL clusters, but contains documentation for the Clustered Logical Volume Manager (CLVM) in addition to standard LVM. Depending on the application, either CLVM or LVM can be used in a RHEL deployment.

- *DM Multipath* – not exclusively for RHEL clusters, but describes a method for making multipath device names consistent within a cluster and provides general information about RHEL multipathing features. Multipathing for storage devices is usually an essential part of any cluster configuration.

- *Configuration Example - Fence Devices* – covers APC switch, IPMI management interface, HP iLO, and SCSI persistent reservation examples. For updated information for HP iLO 3 and 4 configuration, see section 4.1.

- *Configuration Example - NFS over GFS* – provides configuration, testing, and troubleshooting information for a highly available single-instance NFS service running on a Red Hat GFS cluster file system.

- *Configuration Example - Oracle HA on Cluster Suite* – describes an installation scenario for Oracle in single-instance non-shared operation and another installation scenario for Oracle RAC with GFS.

The RHEL6 titles include:

- *High Availability Add-On Overview* – the place to start for an introduction to the High Availability Add-On for RHEL6.

- *Cluster Administration* – provides general configuration, management, and troubleshooting information.

- *Global File System* 2 – documents the features of the Red Hat second-generation cluster file system (GFS2), and provides configuration, management, and troubleshooting information. Discusses the differences between the first generation GFS and GFS2, and provides conversion procedures.   GFS2 has replaced GFS in RHEL6.

- *Load Balancer Administration* – provides configuration information for Linux Virtual Server (LVS) functionality on RHEL6. LVS is used for distributing client IP load across a set of servers while maintaining high availability.

- *Logical Volume Manager Administration* – not exclusively for RHEL clusters, but contains documentation for the Clustered Logical Volume Manager (CLVM) in addition to standard LVM. Depending on the application, either CLVM or LVM can be used in RHEL clusters.

- *DM Multipath* – not exclusively for RHEL clusters, but describes a method for making multipath device names consistent within a cluster and provides general information about RHEL multipathing features. Multipathing for storage devices is usually an essential part of any cluster configuration.

Additionally, refer to the release and technical notes for individual RHEL releases available on the product documentation page. These notes contain valuable information of interest to RHEL cluster users.

Knowledgebase articles

Red Hat often documents specific RHEL cluster hardware interactions in Knowledgebase articles rather than in the product manuals. Several Red Hat Knowledgebase articles listed in this document pertain to RHEL clusters running on HP ProLiant systems. Some of the URLs mentioned here are publicly accessible, but others require a Red Hat Customer Portal account, as does general access to the Knowledgebase. As new Knowledgebase articles pertaining to RHEL clusters on ProLiant might be added from time to time, it can be worthwhile to search the Knowledgebase whenever considering a configuration or support question.

A searchable interface to Red Hat's Knowledgebase is located at:

Red Hat has documented a set of best practices for the High Availability Add-On and related products in the following Knowledgebase article:

Red Hat Enterprise Linux Cluster, High Availability, and GFS Deployment Best Practices
https://access.redhat.com/knowledge/articles/40051

The topics discussed in this article include:

- Cluster node hardware selection

- Storage and network configuration practices

- File systems

- Quorum disk usage

- Cluster membership timing customization

- Disaster tolerant multi-cluster and stretched cluster configurations

The article includes pointers to other Knowledgebase articles documenting best practices and other useful information for cluster subsystems or other related products. The article also describes configurations and use cases that are explicitly unsupported. RHEL clustering is very flexible, and as a result, you could construct a cluster configuration that is not supported by Red Hat. Consequently, prior to deployment, Red Hat requires an architecture review of proposed clusters configured with certain features or addressing certain use cases. The best practices document identifies the features and use cases that require Red Hat architecture review. If you plan to design and deploy RHEL clusters, you should regard this document as mandatory reading.

The following Knowledgebase article documents the cluster architecture review process:

Architecture Review Process for Red Hat Enterprise Linux High Availability, Clustering, and GFS/GFS2
https://access.redhat.com/knowledge/articles/53347

Reference configurations

No tested reference configurations of the High Availability Add-On for Red Hat Enterprise Linux and its associated products on HP ProLiant servers are available at this time. Although all ProLiant server models can be used as RHEL cluster nodes, for best results HP recommends that you select products other than the Microserver and WS460c blade workstation.

Configuration considerations

This section presents considerations that will help you get the best results when configuring and running the High Availability Add-On and its associated products on HP ProLiant servers.

Selection of fence agents for HP Integrated Lights Out management processors

The `fence_ilo` fence agent is commonly used when configuring older HP iLO and iLO 2 management processors as fence devices for RHEL clusters. (When an iLO 2 management processor is running firmware revision 2.00 or greater, Red Hat prefers `fence_ipmilan`.) However, the `fence_ilo` fence agent is not compatible with the newer iLO 3 and iLO 4 management processors. Using `fence_ilo` with iLO 3 or iLO 4 results in failures. The iLO 3 and iLO 4 management processors require using the `fence_ipmilan` fence agent and specific configuration options described in two Red Hat Knowledgebase articles (subscription required):

How do you configure an iLO 3 fence device for RHEL Clustering?
https://access.redhat.com/kb/docs/DOC-56880

How do you configure the fence device agent information option for the HP iLO 3?
https://access.redhat.com/kb/docs/DOC-55705

Follow the configuration directions provided by these documents. Note that you must also configure an iLO account for the fence agent. When you use `fence_ilo` for iLO or iLO 2, you may configure an account with only power control access privileges. However, when using `fence_ipmilan` for either iLO 3 or iLO 4, it was necessary until recently to configure an iLO account with full administrator privileges. Some users prefer tighter privileges for iLO accounts. Recent RHEL5 and RHEL6 changes to `fence_ipmilan` allow you to configure an iLO account with user privileges rather than administrator privileges. For more information, see the following Red Hat Knowledgebase article (subscription required):

IPMI fence on HP iLO3 needs more privileges than acceptable
https://access.redhat.com/kb/docs/DOC-66997

HP Fibre Channel switches as fence devices

The rigorous product testing process established by Red Hat requires testing the exact components used for critical system functions such as fencing. A list of Fiber Channel (SAN-based) switches supported as fence devices for the High Availability Add-On is provided in the following Knowledgebase article:

Fence Device and Agent Information for Red Hat Enterprise Linux
access.redhat.com/kb/docs/DOC-30004

In addition, selected HP Fibre Channel switches have been tested by HP for Red Hat and are also supported as fence devices. These include the following switches that are not listed in the Knowledgebase article:

- HP 4/32 SAN Switch
- HP 4/64 SAN Switch
- HP 8/40 SAN Switch
- Brocade 8 Gb SAN Switch for HP Bladesystem c-Class

HP storage as fence devices

Shared storage devices supporting the SCSI-3 protocol can be used for fast and reliable fencing in RHEL cluster configurations. To preserve data integrity, the SCSI-3 persistent reservation feature is designed expressly to deny access to shared storage by non-member cluster nodes. It is the mechanism underlying the `fence_scsi` fence agent. HP Fibre Channel and iSCSI storage systems that comply with the SCSI-3 standard support persistent reservations. HP thoroughly tests this feature for robustness and reliability in shared storage configurations. The following HP Storage products support SCSI-3 persistent reservations:

- HP 3PAR P10000 Storage Systems (FC and iSCSI)
- HP 3PAR T-Class Storage Systems (FC and iSCSI)
- HP 3PAR F-Class Storage Systems (FC and iSCSI)
- HP Lefthand P4000 family
- XP 24000
- XP 20000
- P9500
- EVA 4400
- EVA 6400

- EVA 8400

- P6300 (FC and iSCSI)

- P6500 (FC and iSCSI)

- MSA P2000 G3 FC

- MSA P2000 G3 (FC and iSCSI)

- MSA P2000 G3 SAS

- MSA P2000 G3 10GbE iSCSI

- MSA P2000 G3 1GbE iSCSI

- MSA 2300/2400 series G2 (FC and iSCSI)

- MSA 2300/2400 series SAS G2

The following technical brief addressing SCSI-3 persistent reservations is accessible from the Red Hat Customer Portal (subscription required):

> *How to Control Access to Shared Storage Devices Using SCSI Persistent Reservations with Red Hat Enterprise Linux Clustering and High Availability*
> https://access.redhat.com/knowledge/techbriefs/how-control-access-shared-storage-devices-using-scsi-persistent-reservations-re

Disaster tolerant configurations

Cluster technology is commonly used for maintaining application availability during loss or failure of a physical site, fire cell, or other localized collection of computing resources. This capability is often referred to as disaster tolerance or disaster recovery. Disaster tolerant systems can be constructed using the High Availability Add-on and its associated products, and Red Hat supports two different general system topologies for this purpose. HP offers storage array products designed to make storage volumes disaster tolerant, and these products can be used in RHEL clusters in specific circumstances.

The following Knowledgebase article describes the supported disaster-tolerant RHEL cluster configurations:

> *Support for Red Hat Enterprise Linux Cluster and High Availability Stretch Architectures*
> https://access.redhat.com/kb/docs/DOC-58412

Multi-site disaster recovery cluster configurations

The preferred Red Hat disaster tolerant or multi-site configuration consists of two independent and essentially identical clusters, each located at a physically separate location (this is sometimes referred to as a multi-cluster).   This configuration is active/passive, meaning that only one cluster (the active cluster) is able to run services at a time; the other cluster is inactive. If shared cluster storage is used, it must be replicated from the active to the passive site synchronously, using an array-based method. Storage replicas are only accessible on the active site;   they are not accessible on the passive site.

Red Hat does not impose a distance limitation between sites, though the storage array used will typically impose a limit.

Site-level management of RHEL multi-cluster system configurations is largely manual, though selective scripting and use of available tools can reduce workload and improve reliability and repeatability. The cluster administrator must duplicate resource and service definitions and applications in the clusters on each site and maintain them in synchrony. Site failover must be managed manually by the cluster administrator; this includes powering up the passive cluster, storage array failover, and resynchronization — as needed. The cluster administrator must also manually manage recovery from

failover, including storage array resynchronization. No automation of site failover or recovery is supported or available.

Red Hat fully supports the use of multi-site disaster recovery clusters. However, a formal architecture review and approval by Red Hat is required to obtain software support.

Stretch cluster configurations

Some users prefer to implement a disaster tolerant system by stretching a single cluster across multiple physically separate sites. For these users, Red Hat supports four specific RHEL stretched cluster system configurations (Red Hat uses the term **stretch cluster**). These configurations are limited to two sites containing cluster nodes and optional storage replicas;   optionally, a third site can be used to host a quorum device.

There is a limit on the maximum allowable distance between sites. This limit is expressed in terms of the maximum packet latency on networks connecting the two sites. It must not exceed a 2 ms round trip time under moderate load conditions, a value Red Hat considers characteristic of a LAN. Red Hat does not provide guidelines for determining acceptable physical distances between sites.

Cluster network traffic between the two sites must not be routed. When using RHEL5 cluster software, the cluster network must support either multicasting or broadcasting. RHEL6.2 supports UDP unicast in addition to multicasting and broadcasting.

If a stretched cluster's services require storage, the supported configurations permit use of three different storage replication methods:

- LVM (software-based) mirroring over a shared SAN

- Synchronous coherent array-based replication

- Only one form of active/passive array-based replication

In the latter case, one site's array is active and accessible in normal operation to all nodes in the stretched cluster through a shared SAN; the other site's array is passive and typically inaccessible until site failover. (The Red Hat term for this is **disaster recovery mirroring**.) Array-based replication in which the arrays at both sites are simultaneously accessible is explicitly unsupported. (The Red Hat term for this is **asynchronous** or **active/passive array-based replication**.)

Red Hat also supports a configuration that does not contain any replicated storage.

Site failover and recovery from failover (failback) are performed manually. This includes those operations required to perform failover and recovery of replicated storage, if included in the configuration.

All stretched clusters that are to receive software support from Red Hat require a formal architecture review. Designing your proposed stretched cluster to match one of the four supported configurations is strongly recommended. Red Hat is unlikely to accept a proposed configuration that does not closely match one of these four alternatives.

The *Support for Red Hat Enterprise Linux Cluster and High Availability Stretch Architectures* Knowledgebase article mentioned previously discusses each of the four supported configurations and their individual requirements, restrictions, and limitations. It also describes use cases and configurations that are explicitly unsupported.

HP CLX and RHEL stretched clusters

Although RHEL5 and RHEL6 clusters generally support HP's Continuous Access storage replication technology, Red Hat does not directly support HP's associated Cluster Extension (CLX) software in stretched cluster configurations. However, HP provides direct support for CLX in selected stretched cluster configurations, and works with Red Hat to assure interoperability and to resolve RHEL cluster support issues should they arise.

All RHEL cluster deployments with CLX require an architecture review performed jointly by HP and Red Hat. All questions regarding architectural concepts, stretched cluster configurations, architecture reviews, and support issues should be directed to HP through your HP sales or support contacts, and should not be directed to Red Hat.

For additional information pertaining to Red Hat, see the following Red Hat Knowledgebase article:

Is HP StorageWorks CLX (Continuous Access) supported with Red Hat Enterprise Linux Cluster and High Availability?
https://access.redhat.com/kb/docs/DOC-62648

For additional information pertaining to HP, see the following HP web page:

HP Cluster Extension Software - Overview
and Features hp.com/go/clx

HP Serviceguard to RHEL cluster migration

Serviceguard to RHEL cluster migration guides are available on Red Hat's Knowledgebase through the Red Hat Customer Portal. Access to the guides is password protected and requires a Customer Portal account. They are intended to help Serviceguard users who want to move to RHEL 5.3 clusters. (The main guide also references RHEL 5.2, which has some serious limitations for fencing with SCSI-3 persistent reservations; RHEL 5.3 or later is preferred.) Red Hat is not aware of any changes through RHEL 5.7 that would make content in these guides obsolete, so they are useful for migration to more recent versions of RHEL5 clusters as well.

Although not written with RHEL6 clusters in mind, large portions of these guides might still be useful when migrating to RHEL6. However, some product enhancements have rendered some content obsolete. Here are a few of the more notable changes:

- As of RHEL 6.0, both IPv4 and IPv6 are fully supported for cluster networking, including cluster heartbeats.

- As of RHEL 6.0, freeze and unfreeze operations are supported; these can be used to perform maintenance on a service (see section 8.3.2 of the *Cluster Administration* guide). Also, freeze and unfreeze operations were added to RHEL5 in the 5.3 release.

- As of RHEL 6.1, additional watchdog integration allows a node to reboot itself when it loses quorum.

- As of RHEL 6.2, the RHEL watchdog code has been integrated with `fence_scsi` so the watchdog service can automatically reboot a node after it has been fenced.

The migration guides include:

Migrating an HP Serviceguard for Linux (SGLX) Cluster to Red Hat Cluster Suite (RHCS)
access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-cluster-red-hat-cluster-suite-rhcs

This provides the most general information and should be read first if you are not familiar with the migration process. It describes a procedure to migrate an HP Serviceguard for Linux (SGLX) cluster to a cluster running Red Hat Cluster Suite software. (Red Hat Cluster Suite is the RHEL5 term for the RHEL6 High Availability Add-On and its associated products.) It discusses differences between the two clusters, including differences pertaining to membership, quorum, fencing to prevent data corruption, and application failover control. A step-by-step process describes how to use configuration information from an existing SGLX cluster to quickly create a cluster from RHCS components. The guide lists features that exist in SGLX that have no comparable Red Hat Cluster Suite counterpart. Many of those differences can be managed by users through custom scripts, and examples are provided.

Migrating an HP Serviceguard for Linux (SGLX) Oracle Toolkit to Red Hat Cluster Suite (RHCS) for Red Hat Enterprise Linux 5

access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-oracle-toolkit-red-hat-cluster-suite-rhcs-red

Migrating HP Serviceguard for Linux (SGLX) MySQL and PostgreSQL Toolkits to Red Hat Cluster Suite (RHCS) for Red Hat Enterprise Linux 5
access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-mysql-and-postgresql-toolkits-red-hat-cluster

Migrating HP Serviceguard for Linux (SGLX) NFS and Samba Toolkits to Red Hat Cluster Suite (RHCS) for Red Hat Enterprise Linux 5
access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-nfs-and-samba-toolkits-red-hat-cluster-suite-

Migrating HP Serviceguard for Linux (SGLX) Apache and Tomcat Toolkits to Red Hat Cluster Suite (RHCS) for Red Hat Enterprise Linux 5
access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-apache-and-tomcat-toolkits-red-hat-cluster-su

Migrating HP Serviceguard for Linux (SGLX) Xen Toolkit to Red Hat Cluster Suite (RHCS) for Red Hat Enterprise Linux 5
access.redhat.com/knowledge/refarch/2009-migrating-hp-serviceguard-linux-sglx-xen-toolkit-red-hat-cluster-suite-rhcs-red-ha

Note that RHEL6 is not supported as a Xen host, so this migration guide is not applicable for Serviceguard on Linux to RHEL6 cluster migrations.

Share with colleagues

Get connected
www.hp.com/go/getconnected
Current HP driver, support, and security alerts delivered directly to your desktop

Become a fan on

Follow on