Hewlett Packard Enterprise

HPE Performance Cluster Manager Getting Started Guide

Abstract

This publication describes how to deploy an HPE cluster with the HPE Performance Cluster Manager 1.6 software. Use this publication to configure a factory-installed HPE cluster on your site network.

Part Number: 007-6500-011
Published: September 2021
Edition: 1

Notices

The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Confidential computer software. Valid license from Hewlett Packard Enterprise required for possession, use, or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Links to third-party websites take you outside the Hewlett Packard Enterprise website. Hewlett Packard Enterprise has no control over and is not responsible for information outside the Hewlett Packard Enterprise website.

Acknowledgments

Intel®, Itanium®, Optane™, Pentium®, Xeon®, Intel Inside®, and the Intel Inside logo are trademarks of Intel Corporation or its subsidiaries.

AMD and the AMD EPYC™ and combinations thereof are trademarks of Advanced Micro Devices, Inc.

Microsoft® and Windows® are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Adobe® and Acrobat® are trademarks of Adobe Systems Incorporated.

Java® and Oracle® are registered trademarks of Oracle and/or its affiliates.

UNIX® is a registered trademark of The Open Group.

All third-party marks are property of their respective owners.

Product-specific acknowledgments

ARM® is a registered trademark of ARM Limited.

Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries.

Red Hat® and Red Hat Enterprise Linux® are registered trademarks of Red Hat, Inc., in the United States and other countries.

Revision history

Part number	Publication date	Edition	Summary of changes
007-6500-011	September 2021	1	Supports the HPE Performance Cluster Manager 1.6 release.
007-6500-010	March 2021	1	Supports the HPE Performance Cluster Manager 1.5 release.
007–6500-009	September 2020	1	Supports the HPE Performance Cluster Manager 1.4 release.

Table Continued

Part number	Publication date	Edition	Summary of changes
007–6500-008	April 2020	1	Supports the HPE Performance Cluster Manager 1.3.1 release.
007–6500-007	December 2019	1	Supports the HPE Performance Cluster Manager 1.3 release.
007–6500-006	June 2019	1	Supports the HPE Performance Cluster Manager 1.2 release.
007–6500-005	April 2019	1	Supports the HPE Performance Cluster Manager 1.1 release. Includes SLES 15 and CentOS 7.6 support.
007–6500-004	December 2018	1	Supports the HPE Performance Cluster Manager 1.1 release.
007–6500-003	September 2018	1	Supports the HPE Performance Cluster Manager 1.0 release, patch 3.
007–6500-002	June 2018	1	Supports the HPE Performance Cluster Manager 1.0 release, patch 1.
007–6500-001	June 2018	1	Original publication. Supports the HPE Performance Cluster Manager 1.0 release.

Contents

Getting started with HPE Performance Cluster Manager

This guide explains how to deploy a factory-installed HPE cluster with the HPE Performance Cluster Manager software.

Hewlett Packard Enterprise installs the cluster manager on many clusters before they leave the factory. A factory installation includes the following:

- A factory-specified root password. One of the first steps in the configuration procedure is to change this root password on the admin node.
- One or two operating system slots.

  On clusters with leader nodes, Hewlett Packard Enterprise configures two operating system slots by default.

  On clusters without leader nodes, Hewlett Packard Enterprise configures one operating system slot.

  Hewlett Packard Enterprise configured the operating system that you ordered on operating system slot 1. If the cluster has leader nodes, operating system slot 2 is present, and operating system slot 2 is blank.

  The cluster supports a maximum of 10 operating system slots. If you need more than the factory-configured number of operating system slots, reinstall the cluster manager.
- One of the following operating systems:
  - Red Hat Enterprise Linux (RHEL) 8.X
  - RHEL 7.X
  - SLES 15 SPX
  - SLES 12 SPX
- A serial-over-LAN connection on the admin node. This connection allows you to use the admin node as the system console. You can access the console by using the IPMItool serial-over-LAN function.

Your first steps with the cluster depend on whether you want to keep the factory configuration. If you keep the factory configuration, use this guide to add networking and other site-specific characteristics. The following table shows various scenarios and links to more information.

Scenario	Instructions
HPE configured the cluster with an operating system and the cluster manager at the HPE factory. The admin node is a single physical admin node, not an HA admin node. You want to attach the cluster to your site network and put the cluster into production.	Use this manual. Proceed to the following: **Attaching a cluster without a high availability (HA) admin node to your site network**
HPE configured the cluster with an operating system and the cluster manager at the HPE factory. The admin node is an HA admin node. You want to attach the cluster to your site network and put the cluster into production.	Use this manual. Proceed to the following: **Attaching a cluster with a high availability (HA) admin node to your site network**

Table Continued

Scenario	Instructions
HPE configured the cluster with an operating system and the cluster manager at the HPE factory. Your admin node might, or might not, be an HA admin node. You want to attach the cluster to your site network, reinstall the operating system, and reinstall the cluster manager. You want to modify the configuration. For example, you might want to change the factory-specified passwords for the compute node controllers and the leader node controllers. After you complete the software reinstallation, you plan to put the cluster into production.	Your task has three parts: Part 1: • Back up the admin node. See the following: **Backing up the admin node** • Back up the cluster configuration files. See the following: **Backing up the cluster configuration files** Part 2: • If you have an HA admin node, proceed to the following: **Attaching a cluster with a high availability (HA) admin node to your site network** • If you do not have an HA admin node, proceed to the following: **Attaching a cluster without a high availability (HA) admin node to your site network** Part 3: Reinstall the operating system and cluster manager. The installation instructions are specific to each cluster platform. For links to the installation guides, see the following: **Cluster manager documentation**
HPE did not install an operating system or the cluster manager.	Install the operating system and cluster manager. The installation instructions are specific to each cluster platform. For links to the installation guides, see the following: **Cluster manager documentation**

Cluster manager documentation

The following list shows the HPE Performance Cluster Manager documentation:

- The release notes contain feature information, platform requirements, and other release-specific guidance. To access the release notes, follow the links on the following website:

  https://www.hpe.com/software/hpcm

  On the product media, the release notes appear in a text file in the following directory:

  `/docs`

Hewlett Packard Enterprise strongly recommends that you read the release notes, particularly the Known Issues section and the Workarounds section.

- The following guide presents an overview of the cluster manager and explains how to attach a factory-installed cluster to your site network:

  HPE Performance Cluster Manager Getting Started Guide

- The bare-metal installation documentation is specific to each platform. These guides are as follows:

  - **HPE Performance Cluster Manager Installation Guide for Clusters With ICE Leader Nodes**
  - **HPE Performance Cluster Manager Installation Guide for Clusters With Scalable Unit (SU) Leader Nodes**
  - **HPE Performance Cluster Manager Installation Guide for Clusters Without Leader Nodes**

- The following guide explains the power management features included in the cluster manager:

  HPE Performance Cluster Manager Power Management Guide

- The following guide includes procedures and information about system-wide administration features:

  HPE Performance Cluster Administration Guide

- The following quick-start guide presents an overview of the installation process:

  HPE Performance Cluster Manager Installation Quick Start

- The following command reference shows the cluster manager commands and compares them with the commands used in the SGI Management Suite and in the HPE Insight Cluster Manager Utility:

  HPE Performance Cluster Manager Command Reference

After installation, the documentation reside on the system in the following directories:

- Release notes and user guides: `/opt/clmgr/doc`
- Manpages: `/opt/clmgr/man`

Feature descriptions for the HPE SGI 8600 system also apply to SGI ICE XA and SGI ICE systems.

`cm` command information

Many cluster manager commands are of the following form:

`cm` *`topic [subtopic ...] action parameters`*

The `cm` commands support tab completion for each *topic*, each *subtopic*, each *action*, and many parameters.

The cluster manager implements tab completion for the `-i` *`image`* and the `--image` *`image`* options by comparing command input against the image names stored in the HPE Performance Cluster Manager database.

You can use wildcard characters in the cluster manager `cm` commands. If you use wildcards in the `cm` commands, enclose your specification in apostrophes (`' '`). The following table shows the most commonly used wildcard characters.

Table 1: Wildcard characters

Wildcard	Effect
*	Matches one or more characters. For example, the following specifies all nodes in rack 1, chassis 1, tray 1 on an HPE Apollo 9000 cluster: `'r1c1t1n*'`
?	Matches exactly one character. For example, the following specifies all nodes in rack 1 that have a single-character chassis: • On an HPE Apollo 9000 cluster: `'r1c?t*n*'` • On an HPE SGI 8600 cluster: `'r1i?n*'`
[ ]	Matches any of the range of characters specified within brackets. For example, the following specifies racks `11`, `12`, `13`, and `14`: `'rack1[1-4]'`

Node identification

The cluster manager recognizes distinct node hostnames for each type of cluster that it supports.

NOTE: The information in this topic shows the compute node names that the cluster manager assigns to nodes by default. This naming scheme identifies components by their location in the cluster. These names are assigned automatically when the compute nodes are configured into the cluster.

HPE Cray EX node identification

On HPE Cray EX supercomputers, the node name is in the following format:

`x`*`CABINET`*`c`*`CHASSIS`*`s`*`SLOT`*`b`*`BLADE`*`n`*`NODE`*

The variables are as follows:

Variable	Specification
CABINET	A 4-digit cabinet identifier in the range 1 <= *CABINET* <= 9999. Specific cabinet identifiers are as follows: • HPE Cray EX fluid-cooled compute: x1000 - x2999 • HPE Cray EX air-cooled I/O: x3000 - x4999 • HPE Cray EX air-cooled compute: x5000 - x5999 • HPE Cray EX TDS: x9000 Examples: `x1004`, `x3001`.
CHASSIS	A 1-digit chassis identifier in the range 0 <= *CHASSIS* <= 7. Examples: `c1`, `c7`.

Table Continued

Variable	Specification
SLOT	A 1-digit slot identifier in the range 0 <= *SLOT* <= 7. Examples: `s1`, `s4`.
BLADE	A 1-digit blade identifier in the range 0 <= *BLADE* <= 1. Examples: `b0`, `b1`.
NODE	A 1-digit node identifier in the range 0 <= *NODE* <= 1. Examples: `n0`, `n1`.

The following are node identification examples:

- `x9000c1s2b0n0` is a compute node.
- `fmn01` and `fmn02` are HPE Slingshot fabric management nodes.

HPE Cray EX switch identification

The default switch naming conventions are similar to the default node naming conventions. On HPE Cray EX supercomputers, the switch names are in the following format:

`x`*`CABINET`*`c`*`CHASSIS`*`r`*`SWITCH`*`b`*`BMC`*

The variables are as follows:

Variable	Specification
CABINET	A 4-digit rack identifier in the range 1 <= *CABINET* <= 9999. Examples: `x0046`, `x0178`.
CHASSIS	A 1-digit chassis identifier in the range 1 <= *CHASSIS* <= 4. Examples: `c1`, `c2`.
SWITCH	A 1-digit tray identifier in the range 0 <= *SWITCH* <= 7. Examples: `r5`, `r7`.
BMC	A 1-digit switch identifier in the range 0 <= *BMC* <= 1. Examples: `b0`, `b1`.

For example: `x1203c0r5b0` is a hostname for an HPE Cray EX switch controller.

HPE Apollo 9000 node identification

On HPE Apollo 9000 clusters, the node name is in one of the following formats:

`r`*`RACK`*`c`*`CHASSIS`*`t`*`TRAY`*`n`*`NODE`*

The variables are as follows:

Variable	Specification
RACK	A 3-digit rack identifier in the range 1 <= *RACK* <= 999. Examples: `r46`, `r178`.
CHASSIS	A 1-digit chassis identifier in the range 1 <= *CHASSIS* <= 4. Examples: `c1`, `c2`.

Table Continued

Variable	Specification
TRAY	A 1-digit tray identifier in the range 1 <= *TRAY* <= 8. Examples: `t5`, `t8`.
NODE	A 1-digit node identifier in the range 1 <= *NODE* <= 4. Examples: `n1`, `n4`.

For example: `r100c3t5n1`

HPE Apollo 9000 switch identification

The default switch naming conventions are similar to the default node naming conventions. On HPE Apollo 9000 clusters, the switch names are in the following format:

`r`*RACK*`c`*CHASSIS*`t`*TRAY*`s`*SWITCH*

The variables are as follows:

Variable	Specification
RACK	A 3-digit rack identifier in the range 1 <= *RACK* <= 999. Examples: `r46`, `r178`.
CHASSIS	A 1-digit chassis identifier in the range 1 <= *CHASSIS* <= 4. Examples: `c1`, `i2`.
TRAY	A 1-digit tray identifier in the range 1 <= *TRAY* <= 8. Examples: `t5`, `t8`.
SWITCH	A 1-digit switch identifier in the range 1 <= *SWITCH* <= 4. Examples: `s2`, `s3`.

HPE SGI 8600 node identification

On HPE SGI 8600 clusters, the node name is in the following format:

`r`*RACK*`i`*CHASSIS*`n`*NODE*

The variables are as follows:

Variable	Specification
RACK	A 3-digit rack identifier in the range 1 <= *RACK* <= 999. Examples: `r46`, `r178`.
CHASSIS	A 1-digit chassis identifier in the range 0 <= *CHASSIS* <= 3. Examples: `i1`, `i2`.
NODE	A 2-digit node identifier in the range 0 <= *NODE* <= 35. Examples: `n2`, `n33`.

For example: `r10i0n31`

HPE SGI 8600 switch identification

The default switch naming conventions are similar to the default node naming conventions. On HPE SGI 8600 clusters, the switch names are in the following format:

`r`*RACK*`i`*CHASSIS*`s`*SWITCH*

The variables are as follows:

Variable	Specification
RACK	A 3-digit rack identifier in the range 1 <= *RACK* <= 999. Examples: `r46`, `r178`.
CHASSIS	A 1-digit chassis controller identifier in the range 0 <= *CHASSIS* <= 3. Examples: `i1`, `i2`.
SWITCH	A 2-digit switch identifier in the range 0 <= *SWITCH* <= 1. Examples: `s0`, `s1`.

HPE cluster computing systems

The cluster manager components are as follows:

- An admin node. All clusters include an admin node.

  On HPE Cray EX supercomputers, the admin node is an HPE ProLiant DL325.

  On HPE Apollo 9000 systems and HPE SGI 8600 systems, the admin node is an HPE ProLiant DL360 server.

- Leader nodes. Depending on the number of compute nodes in the cluster, HPE might recommend leader nodes. The admin node delegates boot, monitoring, and logging tasks to leader nodes in large clusters.

  The following figures show examples of clusters with leader nodes.

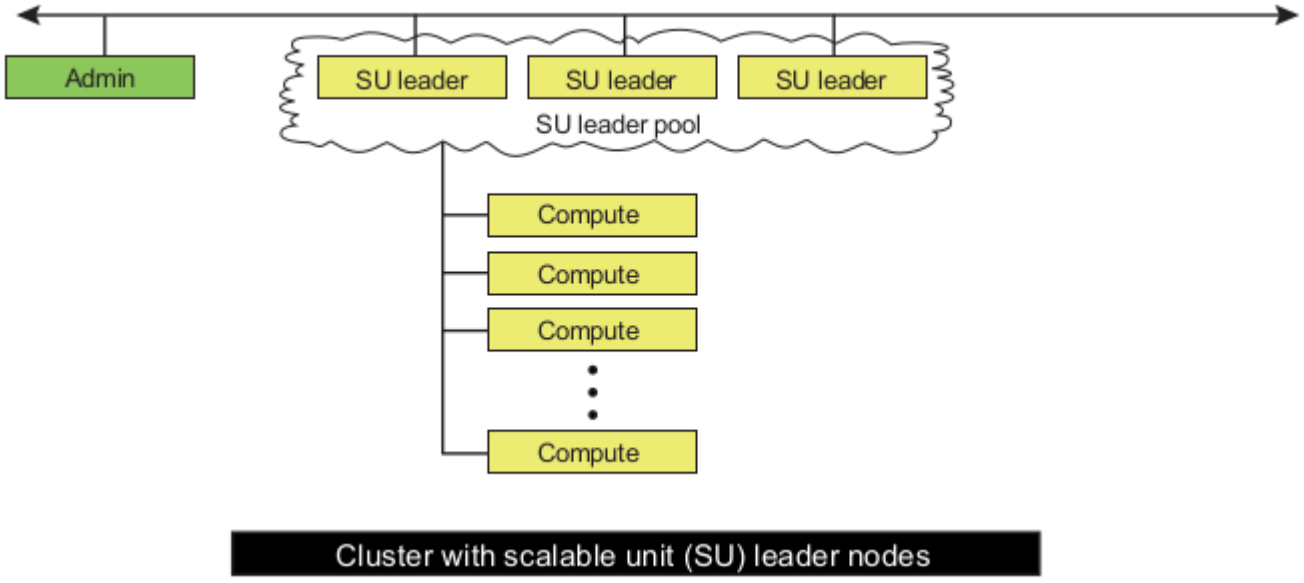

Figure 1: Cluster with scalable unit (SU) leader nodes

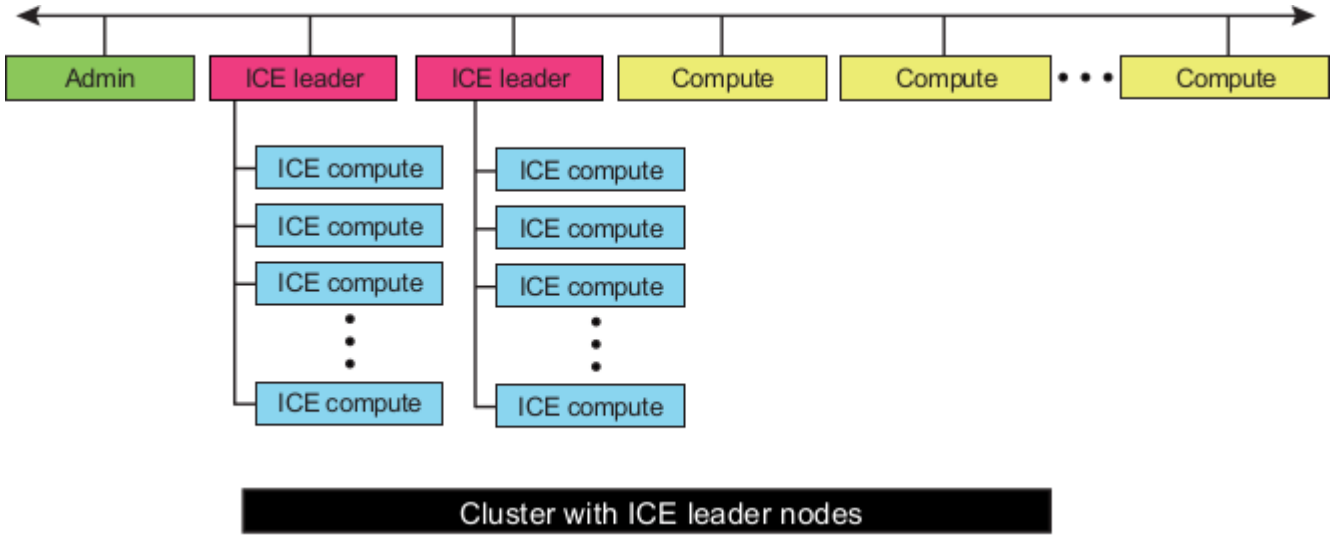

Figure 2: Cluster with ICE leader nodes

- ICE compute nodes. These compute nodes, typically diskless, are under the control of an ICE leader node. Only clusters with ICE leader nodes have ICE compute nodes.
- Compute nodes. All clusters can include compute nodes, and some clusters consist of only an admin node and many compute nodes.

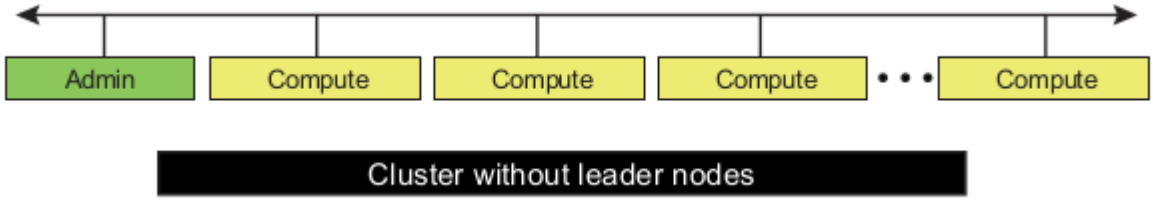

Figure 3: Cluster without leader nodes

On clusters with SU leader nodes, you can configure compute nodes to be under the control of an SU leader node. On these clusters, you can attach additional compute nodes to the admin node and use those compute nodes for user services.

On clusters with ICE leader nodes, compute nodes are often deployed with user services. For example, compute nodes can be configured as login nodes or gateway nodes. These compute nodes cannot be configured to be under the control of an ICE leader node.

NOTE: The term **compute node** describes nodes that can be connected directly to the admin node or nodes that run under the control of an SU leader node.

The term **ICE compute node** describes nodes that run under the control of an ICE leader node.

In several ways, compute nodes are different from the ICE compute nodes that exist in clusters with ICE leader nodes. In a cluster with ICE leader nodes, the compute nodes are not under the control of any kind of leader node.

In a cluster without leader nodes, the admin node and the compute nodes attach directly to the management network. This cluster includes the admin node and several compute nodes. All nodes attach directly to the management network, and all compute nodes can communicate with the admin node. Several HPE Apollo models are clusters without leader nodes.

In a cluster with leader nodes, the leader nodes perform an administrative role for the compute nodes. Leader nodes improve installation performance, collate log files, and perform other tasks for the compute nodes that they control.

The following are additional cluster characteristics:

- Operating system slots. A slot includes all the partitions related to a single, Linux installation. You can configure from 1 to 10 slots. Typically, the factory configures two slots. When you have two slots, the first slot is dedicated to the operating system you ordered, and the second slot is blank. To alter the number of slots, you have to reinstall the cluster manager.
- On a cluster without leader nodes, the admin node and the compute nodes communicate with each other directly for provisioning, reporting, and other actions. An HPE Apollo 70 cluster is an example of a cluster without leader nodes.
- On a cluster with leader nodes, the leader nodes create a network and communication hierarchy for the compute nodes that are attached to the leader node.

  On a cluster with ICE leader nodes, the following bidirectional communication takes place:

  - The admin node and the compute nodes that are deployed with user services communicate with each other.
  - The admin node and the ICE leader nodes communicate with each other.
  - The ICE leader node and the ICE compute nodes under each ICE leader node communicate with each other.

On a cluster with SU leader nodes, the SU leader nodes and all the compute nodes communicate with each other. The communication is bidirectional, and occurs regardless of the compute node role.

Output from the `cm node show` command can help you to determine the type of cluster you have. For example, if you log into the cluster remotely, you can enter this command on the admin node and examine its output. The following output was obtained on an ICE cluster:

```
ICE_cluster# cm node show
r1i0n0
r1i0n1
r1i0n3
r1i0n4
r1lead
n0
```

The preceding output includes `r1lead`. By convention, many clusters have ICE leader nodes of the form `r*lead`. On clusters with leader nodes, the `cm node show` output often includes nodes with `lead` in the node name.

Clusters without leader nodes

In clusters without leader nodes, the nodes have the following roles:

- The admin node is the cluster administrative node for the cluster. The cluster manager enables you to install, provision, configure, and manage the cluster computing system. After you install the cluster manager, the admin node hosts the original, factory-installed copies of the software images for each component. System administrators log into the admin node for the following reasons:
  - To manage the cluster configuration
  - To run system management commands
  - To modify component images
  - To perform system-wide operations

  The cluster manager software distribution includes the master system image for the admin node. During the installation and configuration process, the installation software creates the master system images for the other components in the cluster. As you customize the system for your site, you modify the component-specific system images on the admin node and push updated images to the other nodes.
- The compute nodes perform computing functions or are configured as user services nodes.

  In clusters without leader nodes, you can use the compute nodes in any of the following ways:
  - As computing nodes. An admin node can manage thousands of compute nodes, depending on the cluster workload. The compute nodes all receive a hostname and an IP address during the configuration process.
  - As user services nodes. You can configure the following types of user services on the compute nodes:
    - Login services. These services allow an end user to log in and then, for example, run or monitor HPE Message Passing Interface (MPI) jobs.
    - Batch scheduling services. You can install workload schedulers such as Altair PBS Professional, Adaptive Computing Moab, Slurm, or TORQUE.
    - I/O gateway. On a small system, you can combine the I/O gateway, login services, and batch scheduling on the same compute node.

      The I/O gateway services connect the cluster to your site network. You can configure one or more of the following protocols on the node: network file system (NFS), network address translation (NAT), or network information service (NIS).

- Storage. A compute node with storage is a network attached (NAS) appliance bundle that provides fabric-attached storage for the cluster.
- Object storage server. This server is used in Lustre File Storage configurations.
- Metadata server. This server is used in Lustre File Storage configurations.

Hewlett Packard Enterprise recommends the following login practices:

- Restrict admin node login privileges to the system administrator only. Hewlett Packard Enterprise recommends that sites prohibit end-user access to the admin node.
- Configure user services on the compute nodes. End users can have access to these compute nodes to log in, for example.

Clusters with scalable unit (SU) leader nodes

HPE Cray EX clusters and HPE Apollo clusters can have SU leader nodes. On a cluster with SU leader nodes, the admin node supports SU leader nodes in an **SU leader pool**. In the SU leader pool, a single SU leader node is configured with two other SU leader nodes into a computing trio. This trio of SU leader nodes forms a resilient unit that administers computing functions. You are required to configure SU leader nodes in multiples of three. For example, you can configure 3, 6, 9, or any number of SU leader nodes that is a multiple of three.

The cluster manager documentation uses the term **compute nodes** to refer to all the compute nodes in a cluster with SU leader nodes. The following figure shows a cluster with SU leader nodes.

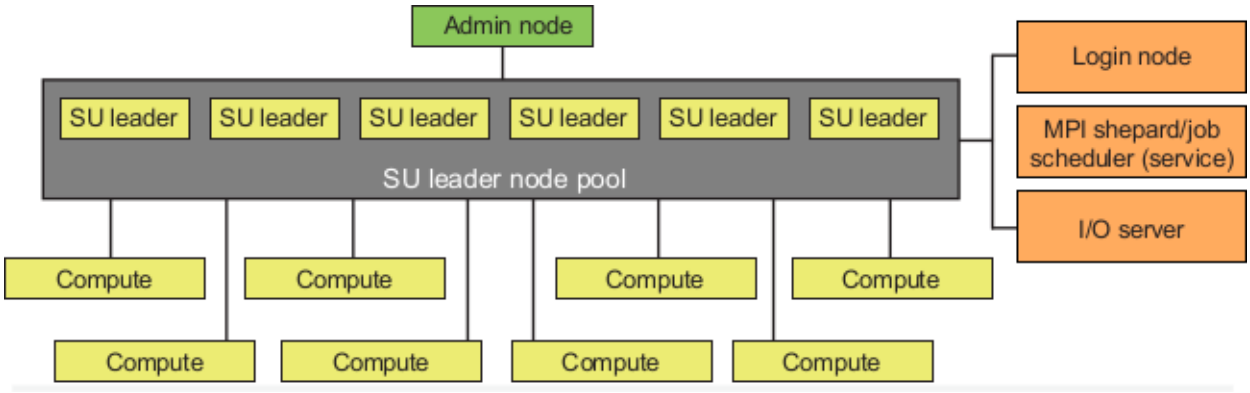

Figure 4: Cluster with SU leader nodes

On large clusters, SU leader nodes help the admin node manage the compute nodes. SU leader nodes help manage the following actions:

- Network booting
- Node installation
- Diskless NFS services
- System and console log consolidation

The cluster manager uses the CTDB service to provide a level of high availability to services. NFS and network boot are the primary HA services managed. CTDB assigns IP aliases to each SU leader node. These IP aliases are what the actual compute nodes use for booting and NFS services. When an SU leader node goes down, CTDB detects the failure and automatically moves the IP alias of the bad node to one of the remaining nodes. Other than a momentary pause in NFS access, activity on the compute nodes is not disrupted.

You can consider a cluster with SU leaders from the following perspectives:

- The Gluster file system.

  SU leader nodes use shared storage to synchronize image management and system boot configuration. The underlying storage technology is Gluster, which is a scalable network file system. The following figure shows an example cluster with SU leader nodes and a Gluster storage system.

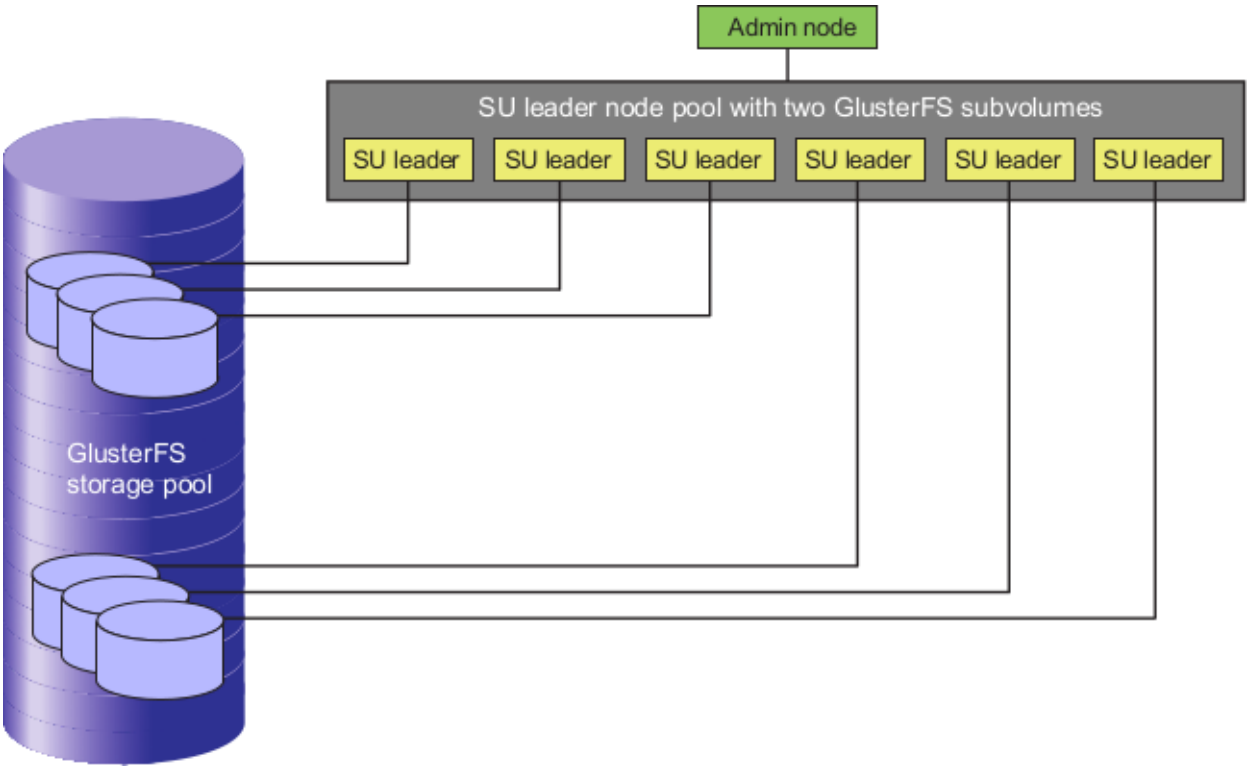

Figure 5: SU leader node pool with subvolumes

  To provide both resiliency and high availability, the SU leader nodes are configured into groups of 3. Each group of 3 forms a Gluster subvolume. For systems with 6 or more SU leader nodes, the data is distributed across the subvolumes. The system can sustain a failure in any single subvolume (group of 3) without losing access to storage.

  The Gluster file system is replicated across the SU leader nodes. Every file is replicated on at least 3 leader nodes. On larger clusters with 6 or more SU leader nodes, files are spread across (distributed) on the larger number of nodes.

- Network booting and NFS root file systems.

  The compute nodes view the SU leader nodes as one giant resource. The SU leader node to which a compute node ultimately connects does not matter. If a compute node is connected to an SU leader node that goes down, the IP address (alias) of the SU leader node automatically moves, and the compute is unaware of this action. Any SU leader node can take over for any other SU leader node.

- System console or serial console and system logs.

  System/Serial Console and system logs.

  The cluster manager does not use the **pool** concept in the case of console and log consolidation. Instead, a designated SU leader node handles that task for a given node. If that SU leader goes down, these services are disrupted until the SU Leader returns to service. These services are not yet designed for high availability.

In a cluster with SU leader nodes, you can configure the following types of user services on the compute nodes:

- Logging in
- Batch computing
- I/O
- Object Storage Server (OSS) for Lustre
- Metadata Server (MDS) for Lustre
- Gateway
- Storage

SU leader nodes enable cluster to be provisioned quickly. Master software images for each node type reside on the admin node. During configuration, the admin node pushes the software images to the SU leader nodes. Each SU leader node provisions the compute node images to the compute nodes that are associated with it.

The cluster manager can support many SU leader nodes in a cluster. Each SU leader node can manage hundreds of compute nodes. The following figure shows the cluster manager software layer that supports the SU leader node pool and the compute nodes.

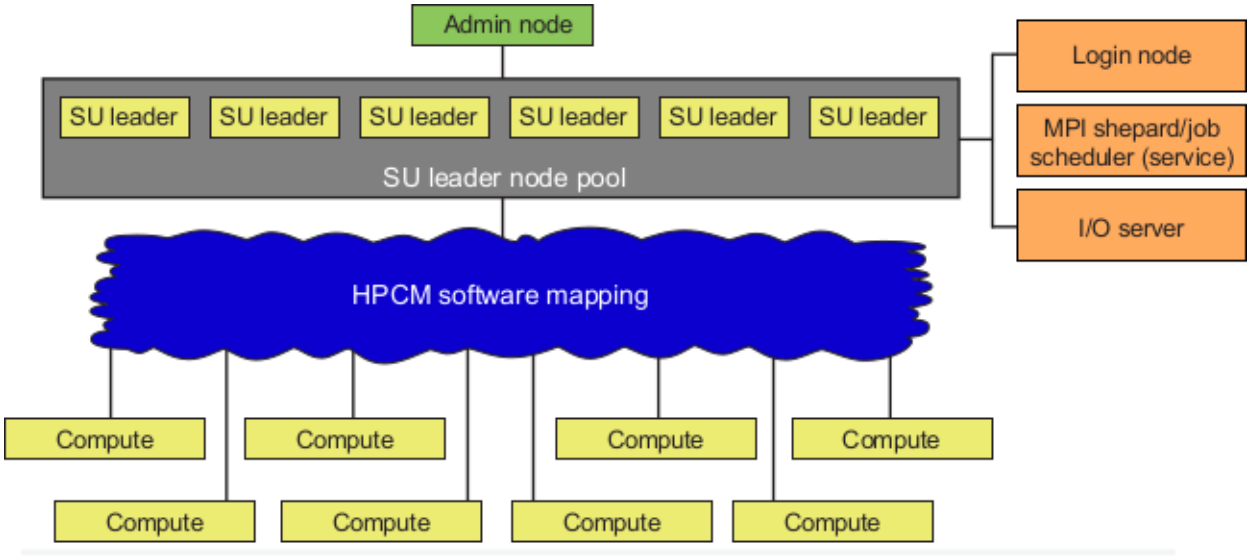

Figure 6: Cluster manager role in a cluster with SU leader nodes

For more information about the compute nodes, see the following:

Clusters without leader nodes

Clusters with ICE leader nodes

The HPE SGI 8600 clusters and the SGI ICE clusters include ICE leader nodes, ICE compute nodes, and compute nodes.

In a cluster with ICE leader nodes, user services are installed on the compute nodes. The admin node can support compute nodes configured for services such as the following:

- Logging in
- Batch computing
- I/O
- Object Storage Server (OSS) for Lustre

- Metadata Server (MDS) for Lustre
- Gateway
- Storage

ICE leader nodes enable these computing systems to be provisioned quickly. Master software images for each node type in the cluster reside on the admin node. During configuration, the admin node pushes the software images to the ICE leader nodes and to the compute nodes. Each ICE leader node pushes the ICE compute node images to the ICE compute nodes that reside in its rack.

An admin node can support many ICE leader nodes, each of which can manage hundreds of ICE compute nodes. The characteristics of the leader nodes and ICE compute nodes are as follows:

- The role of the ICE leader node is to manage a set of ICE compute nodes in one or two racks.
- The ICE compute nodes are simplified compute nodes. Typically, they are diskless. The ICE compute nodes reside in a rack supported by a leader node. These ICE compute nodes require a leader node for services, infrastructure, and support. An ICE leader node can manage up to 288 ICE compute nodes, each running its own Linux operating system. The exact number of ICE compute nodes that an ICE leader node can manage depends on the specific hardware model.

For more information about the compute nodes, see the following:

Clusters without leader nodes

Cluster manager high availability (HA) nodes

You can implement HA capabilities on the admin node and on leader nodes. If you need HA, the cluster manager enables you to configure the following:

Capability	Platform availability examples	Physical description	Software implementation
Quorum HA admin nodes	Any cluster.	Three physical nodes.	One of the three physical nodes hosts a virtual machine (VM). The VM acts as the functioning admin node. If the physical node that hosts the VM fails, the VM admin node passes to one of the other nodes. The VM uses the Gluster file system storage pool for its system disk image.
System admin controller high availability (SAC HA) admin nodes	HPE Apollo 9000 HPE SGI 8600	Two physical nodes attached to an external, shared storage disk.	One of the two physical nodes hosts a virtual machine (VM). The VM acts as the functioning admin node. If the physical node that hosts the VM fails, the VM admin node passes to the other node. The VM uses the shared storage for its system disk image. The physical HA leader nodes host HA software such as Pacemaker and Corosync.
Scalable unit (SU) leader nodes	HPE Cray EX Any HPE Apollo cluster	Multiples of three physical nodes connected to the admin node, compute nodes, and service nodes.	The physical SU leader nodes use a shared Gluster file system.
HA ICE leader nodes	HPE SGI 8600	Two physical nodes connected to the admin node, compute nodes, and service nodes.	The physical HA leader nodes host HA software such as Pacemaker and Corosync.

The following additional information pertains to HA admin nodes and HA ICE leader nodes:

- For a quorum HA admin node or a SAC HA admin node, the host upon which the VM resides is the **active node**. The other nodes are the **passive nodes**. If the active node fails, one of the passive nodes assumes the active role and starts up the admin node VM. The Gluster file system requires at least two of the physical admin nodes to be up and running.

  In a quorum HA configuration, there are two passive nodes. In a SAC HA configuration, there is one passive node.
- An HA ICE leader node has an active/passive configuration. It uses tools that reside in the HA repository of the operating system distribution. The cluster manager provides additional packages. The HA ICE leader node includes the following software:
  - Pacemaker. This software provides a cluster resource management framework. This framework automatically starts, stops, monitors, and migrates resources.
  - Corosync, which is a cluster messaging layer. Pacemaker uses Corosync for messaging and node membership.
  - A distributed replicated block device (DRBD) partition, which is similar to a RAID mirror shared between two systems.

The DRBD software includes the `drbd` and `drbd-kmp` packages. Both software packages belong to DRBD, the kernel block-level synchronous replication facility that serves as an imported, shared-nothing cluster building block.

At any time, the active node is the node that has the DRBD partition mounted. If the active node fails, the passive node takes over the IP addresses and all active node duties.

- Linux Logical Volume Manager, version 2 (LVM2), which enables easy and flexible data management.
- The admin node requires the `ha-rlc-admin` packages and its dependent packages. The HA leader nodes use the `ha-rlc-lead` package and its dependent packages.

The cluster manager permits HA configurations on nodes that run RHEL or SLES. The cluster manager does not support HA capabilities on nodes that run CentOS.

For more information about clusters with scalable unit (SU) leader nodes, see the following:

Clusters with scalable unit (SU) leader nodes

High Availability (HA) admin nodes

An HA admin node consists of physical admin nodes and a virtual machine (VM) admin node. One of the physical admin nodes host the VM, and the VM is the functioning admin node. The cluster manager supports the following types of HA admin nodes:

- A quorum admin node, which consists of three physical admin nodes and one virtual machine.
- A system admin controller high availability (SAC HA) admin node, which consists of two physical admin nodes and one virtual machine.

The following figure shows a quorum HA admin node:

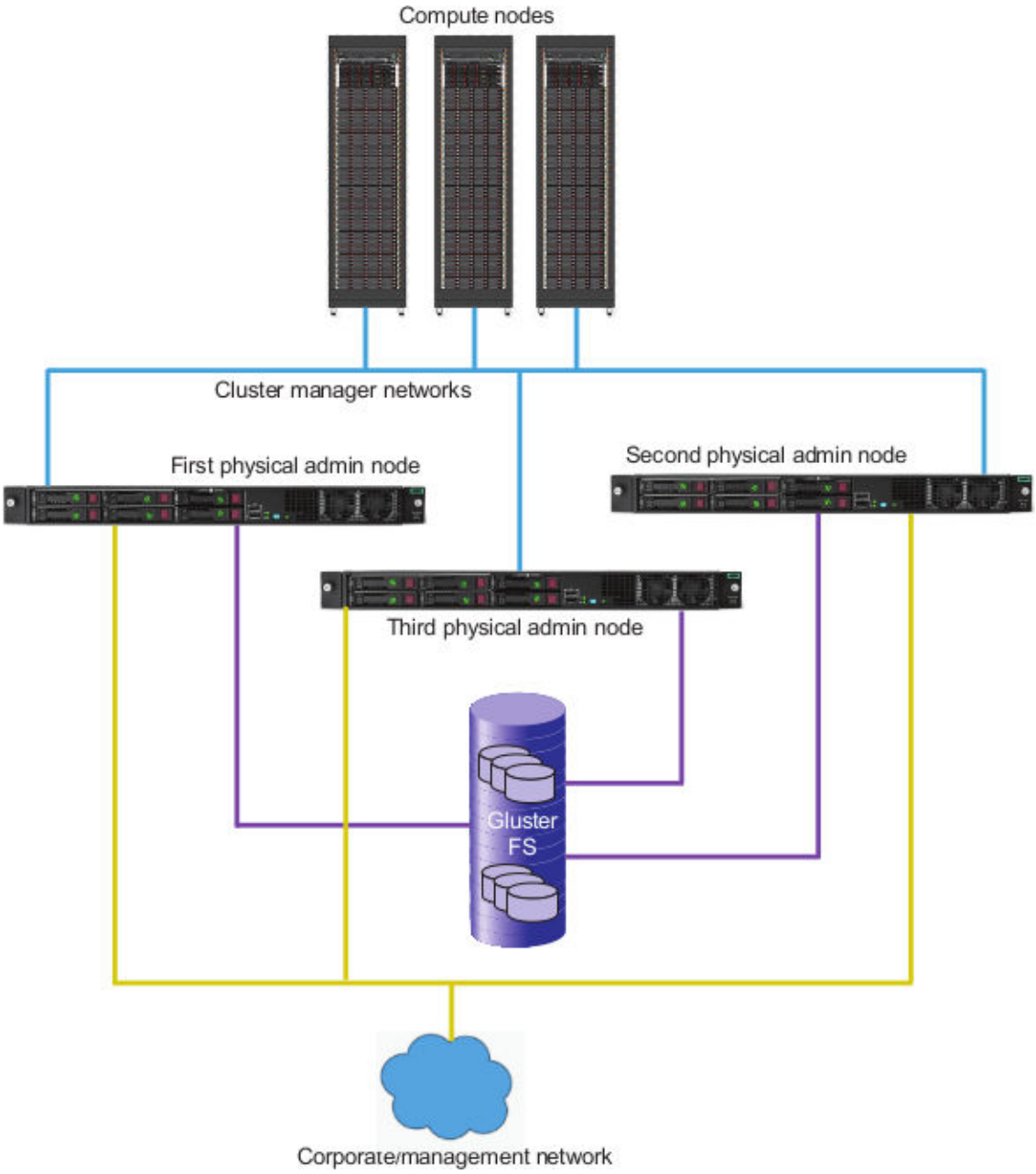

Figure 7: Quorum HA admin node

The following figure shows a SAC HA admin node:

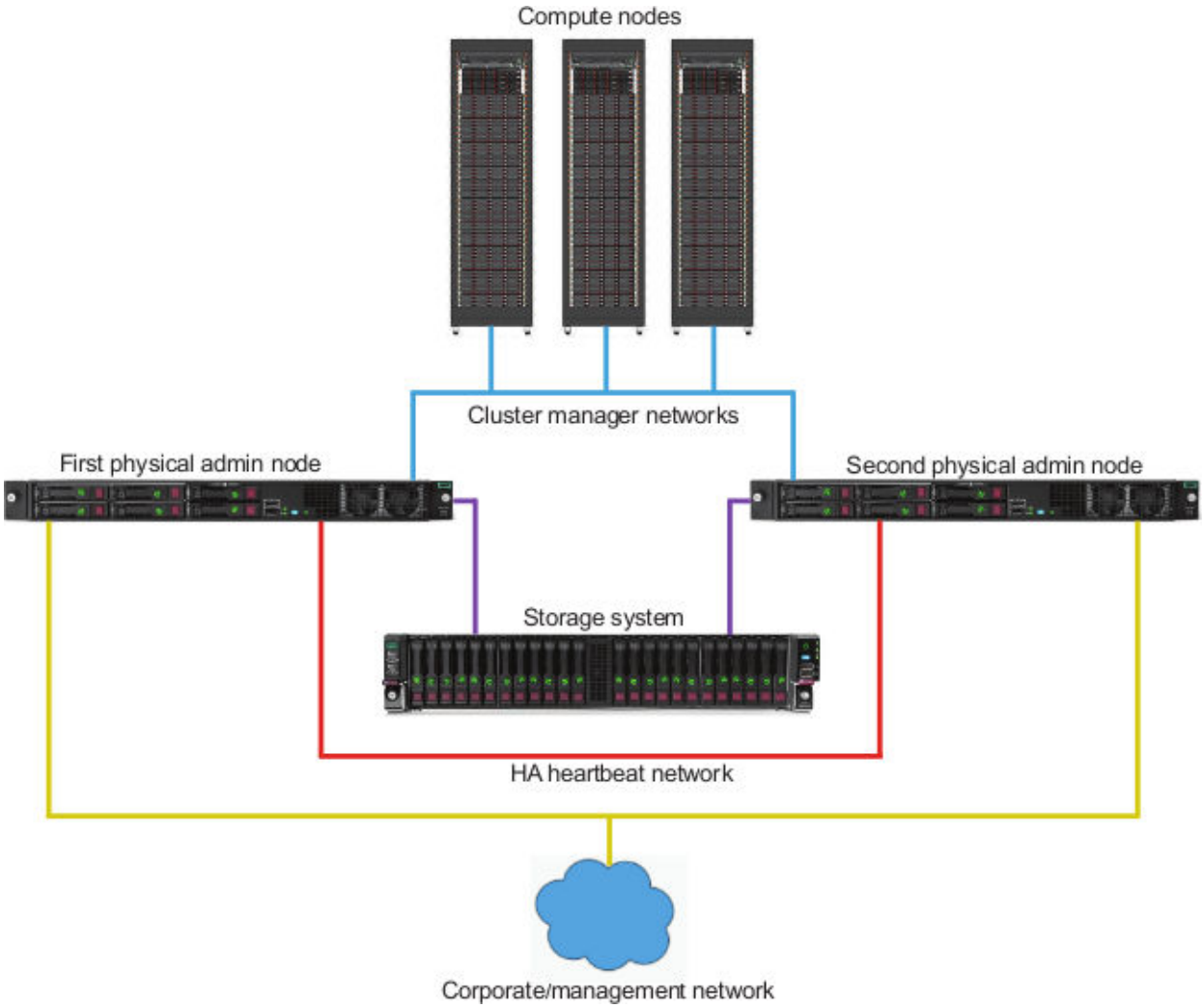

Figure 8: SAC HA admin node

Managed compute nodes and unmanaged compute nodes

Some clusters include compute nodes that are not under cluster manager control. These unmanaged nodes receive an IP address and a hostname from the admin node when you include them on the `discover` command. This action reserves the following IP addresses for the unmanaged nodes

- Node controller IP addresses
- Management network IP addresses
- Data network IP addresses

The `discover` command reserves the preceding addresses from the cluster IP ranges and configures name resolution. However, all other operations occur outside the cluster manager. For example, the software installation and configuration for these nodes occurs outside of cluster manager workflows.

The cluster manager documentation addresses only the nodes that the cluster manager actively manages. The documentation does not pertain to nodes that are attached to the cluster but are managed separately.

Node images

The following figure shows three example cluster types. The figure shows the default image name for each node.

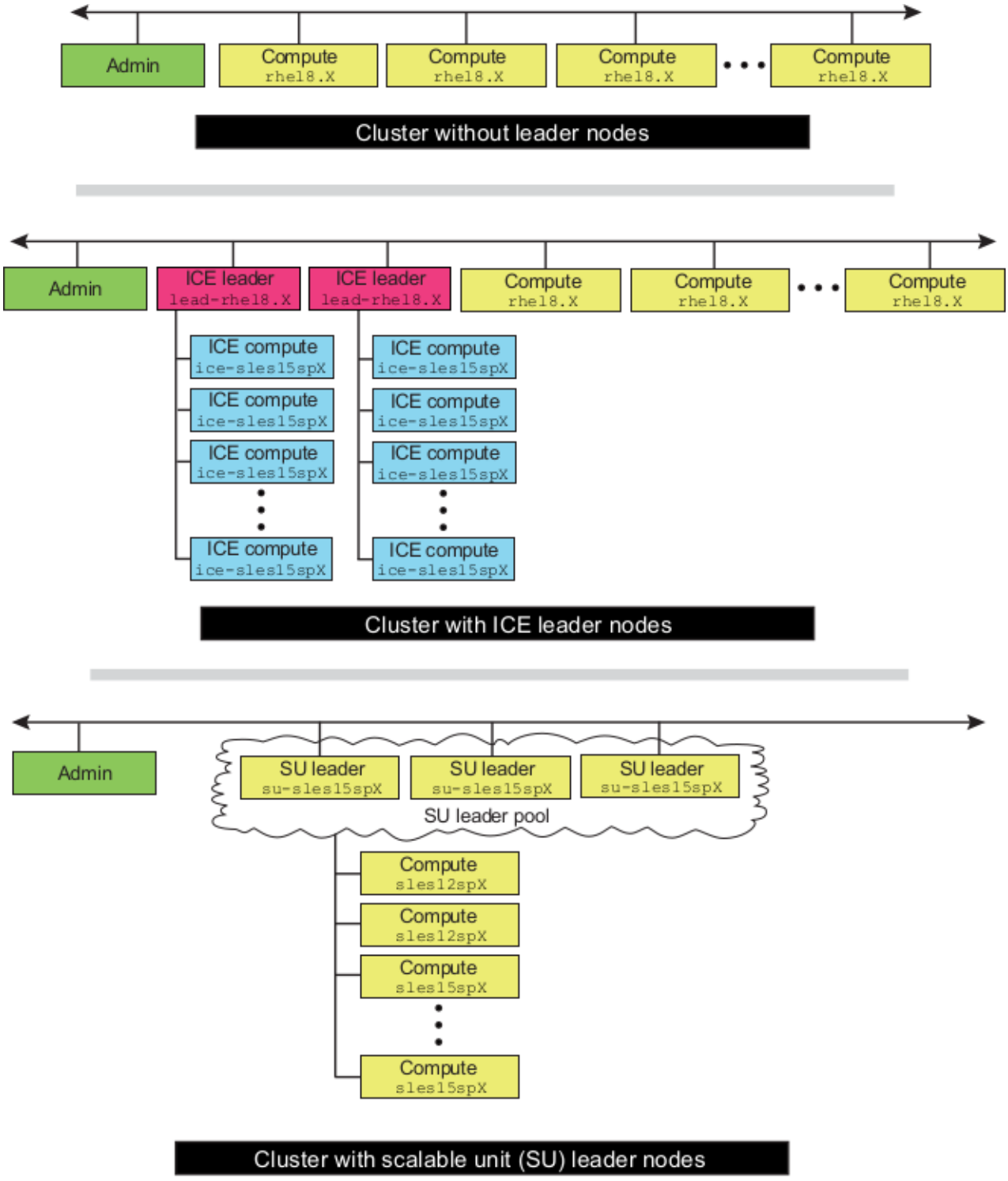

Figure 9: Nodes and default image names

The figures show the default image names for the software images on each node. As the example shows, you can configure nodes with different operating system images within the same cluster. If the cluster has leader nodes, the admin node and the leader nodes must have the same operating system. On clusters with ICE leader nodes, an ICE compute node can have an operating system that differs from the ICE leader node to which it is attached.

The image names shown in the figure are the default names for the factory-installed system images. The system images for the cluster nodes are unique to each type of node.

The following list shows the nodes in each type of cluster and shows the factory-default names for the images that reside on each node:

- ICE leader node images are `lead-os_name`. For example, `lead-sles15spX`.
- SU leader node images are `su-os_name`. For example, `su-sles15spX`.
- ICE compute node images are `ice-os_name`. For example, `ice-sles15spX`.
- X86_64 compute node images are `os_name`. For example, `sles15spX`.
- Arm (AArch64) compute node images are `os_name-aarch64`. For example, `sles15spX-aarch64`.

If you modify the image to include site-specific software, it is typical to back up the existing image to the version control system (VCS), give the modified image a new name, and back up the new image. When you use a new image name for your changes, you preserve the ability to revert to the original image.

The version control software can help you manage multiple versions of that image. The version control system facilitates the following:

- Storage. You can have many versions of each individual software image. When you modify an image, the first modification is the same size as the original image. Subsequent images are smaller because they include only the changes you made. Each version is easily retrieved.
- Experimentation. Each software image is tagged with a version number, so you can easily enable and disable specific versions of the software images.

Cluster networks

NOTE: The cluster manager requires each defined subnet address to be unique. That is, the head network and the BMC network cannot be the same.

External management network connectivity

By default, the cluster manager management network design dictates that external connectivity is first accessed on a public-facing network interface card (NIC) in one of the following locations:

- In the cluster manager admin node
- In one or more designated login nodes

The design isolates network vulnerabilities from external threats by minimizing the number of access points that you need to secure. When you restrict access to only the admin node, you only have to secure one access point. When you configure dedicated login nodes, you need to secure only those nodes.

After a user connects to one of these nodes, the user can access private, management network resources over a dedicated management NIC.

If your users require direct access to the management network, the cluster manager does not inherently prevent that access, but you must configure the management switches and/or nodes manually. You must also consider the implications associated with a different design.

Below are three types of management network designs that have been considered in the past. These do not include all forms of network design, as there are many types of unique requirements and designs across all types of users.

Standard cluster manger management network design

The following diagram shows the standard cluster manager management network.

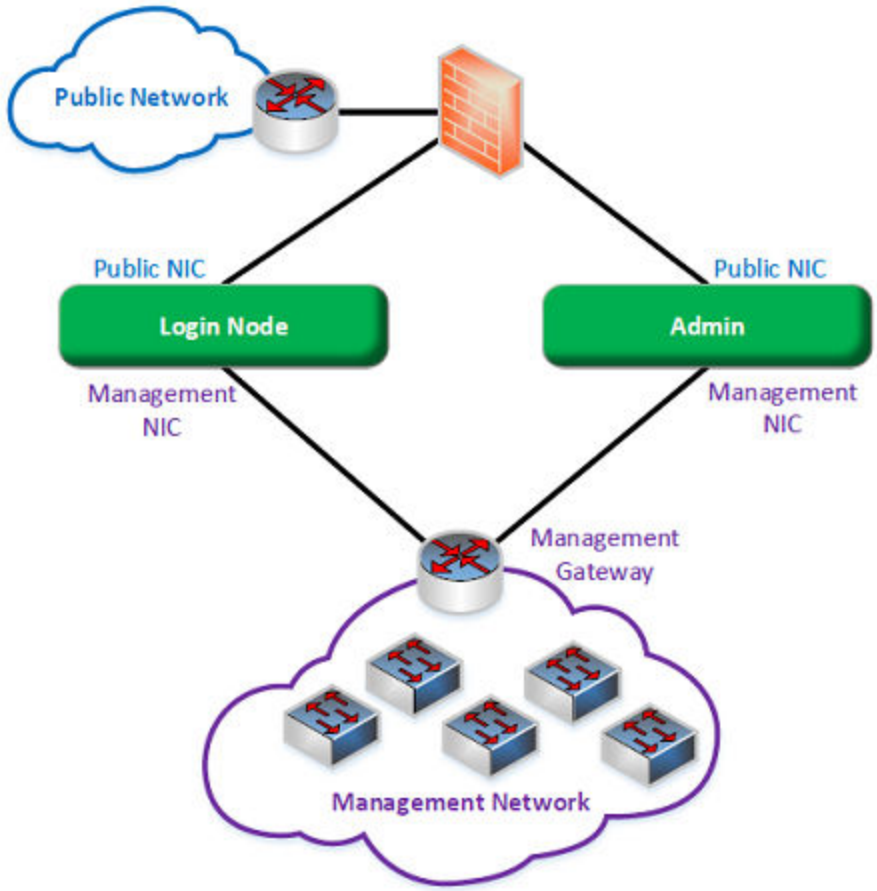

Figure 10: Standard management network design

From a public network, users log in either to the admin node or to designated login nodes directly. They access the management network resources from these nodes.

Alternative network design: management network with external routing

In this topology, users can connect directly to one or more L3 routers and then to the management network routers. Considerations include inter-vendor optics (QSFP+/QSFP28/SFP+/SFP28), especially at the newest developments in link-speed technology. Users can use static routing or routing protocols such as OSPF or BGP to share network information across networks in order to automate access.

If you implement this design, Hewlett Packard Enterprise recommends the following:

- That you read through all vendor-specific documentation regarding best practices in IP routing and security.
- That users use authentication on routing protocols. That is, they must change all usernames and passwords from default values and implement strong firewall rules and access control lists (ACLs) where appropriate.

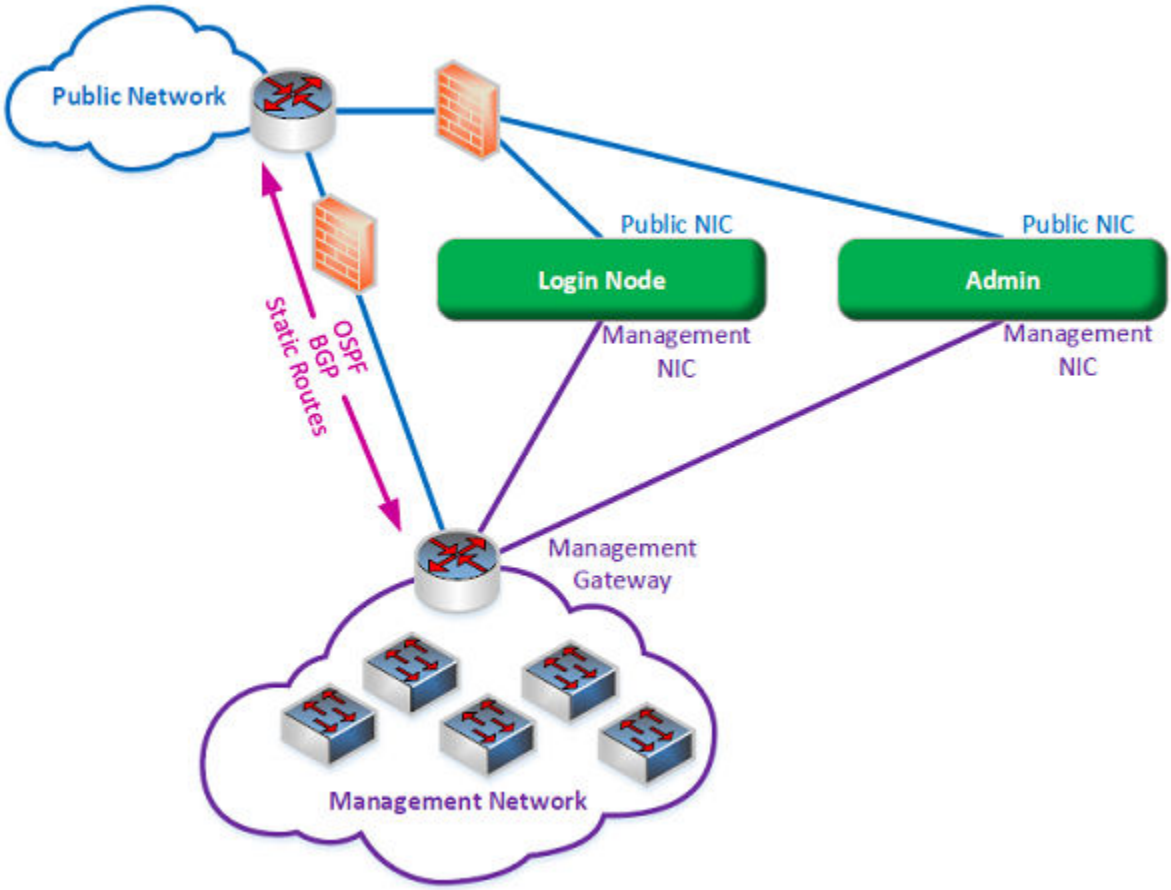

Figure 11: Alternative management network design - management network with external routing

Alternative network design: gateway node network address translation (NAT)

In this topology, a standard Linux server acts as a NAT gateway that negotiates external access for the management network. This is generally configured by using standard Linux packages such as `iptables` and the built in routing engine in Linux distributions. This design preserves some of the security risks of direct connection, but it can suffer from the performance limitations of the gateway node.

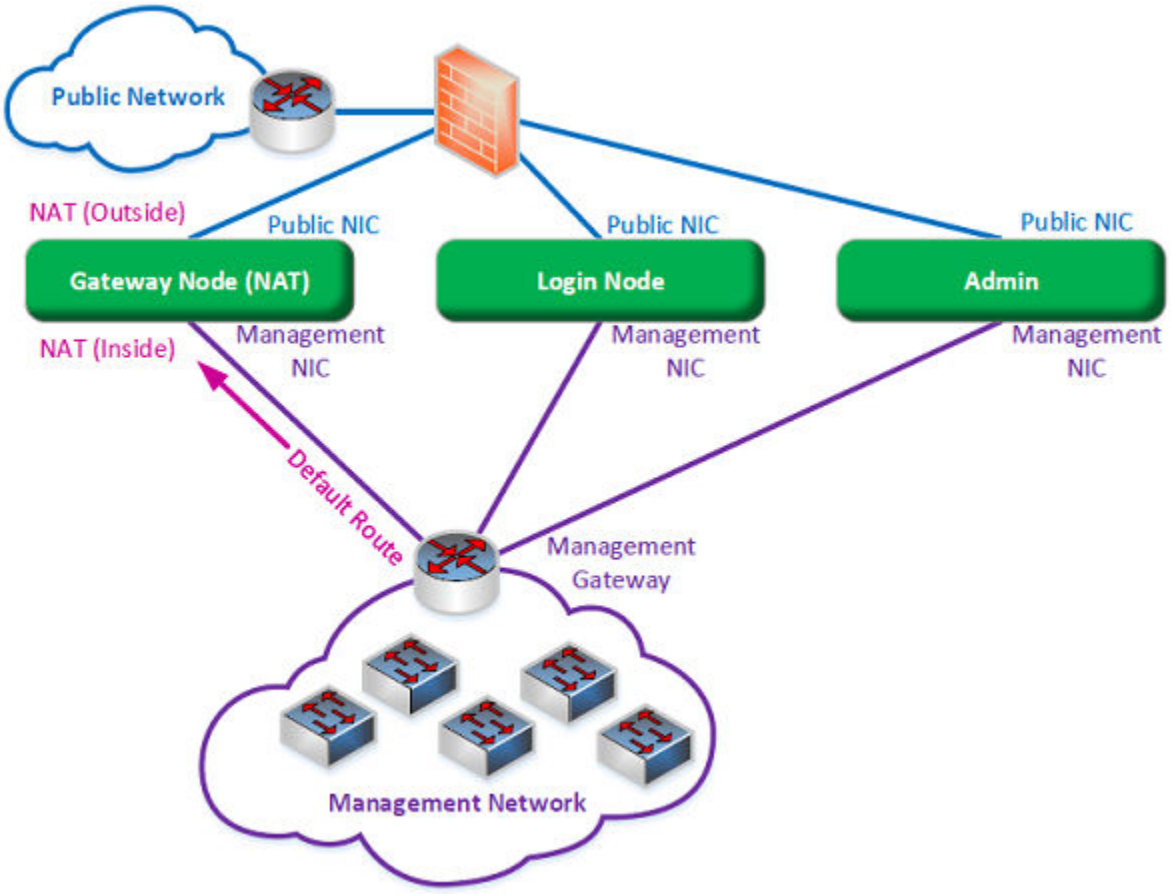

Figure 12: Alternative management network design - gateway node with NAT

Management networks in clusters with scalable unit (SU) leader nodes and clusters without leader nodes

For clusters with SU leader nodes and clusters without any leader nodes, the cluster manager supports the following types of management network topologies:

- Shared Gigabit Ethernet (GbE) and iLO management network.

  Both the `head` and `head-bmc` networks share the same set of physical management switches. The interface `bond0` on the admin node has two different IP addresses to handle traffic for each network/subnet.

- Separate GbE and iLO management network.

  The `head` and `head-bmc` networks each have their own set of physically separated management switches. These two networks are completely separated, so the admin node requires a dedicated interface that is not part of `bond0` to handle the `head-bmc` network traffic.

Compute nodes and leader nodes can connect to the management network in a few different ways. The following describes the most common types of cabling found in a cluster:

- Redundant management + dedicated iLO.

  Two connections dedicated to the GbE traffic to a node. Typically, these connections are bonded under an active-active bonding protocol such as IE: 802.3ad or LACP. In addition, the node has a third connection that is used for dedicated iLO traffic. This third connection plugs into the separate iLO port located on the server.

- Non-redundant management + dedicated iLO.

This is exactly the same as the preceding connection type, with the major difference being that there is only a single connection used for the GbE traffic for the node. With a single connection, bonding is still configured, but the bonding mode is `active-backup`. In this case, no management switch configuration is required because `active-backup` bonding with a single link is functionally equivalent to no bonding. The dedicated iLO connection is present.

- Non-redundant management + shared iLO.

  This method uses only a single connection to a node. This single connection shares both GbE and iLO traffic. Different motherboard vendors implement this functionality in various ways. See your hardware manufacturer's guide for specific details. The GbE traffic uses bonding, but just as for the preceding cabling type, bonding is configured to use `active-backup` bonding. No management switch configuration is needed.

NOTE: If your cluster does not have SU leader nodes, you can ignore the information in the following figures that describes SU leader nodes.

The following figures show examples of how compute nodes or SU leader nodes can connect to a management network.

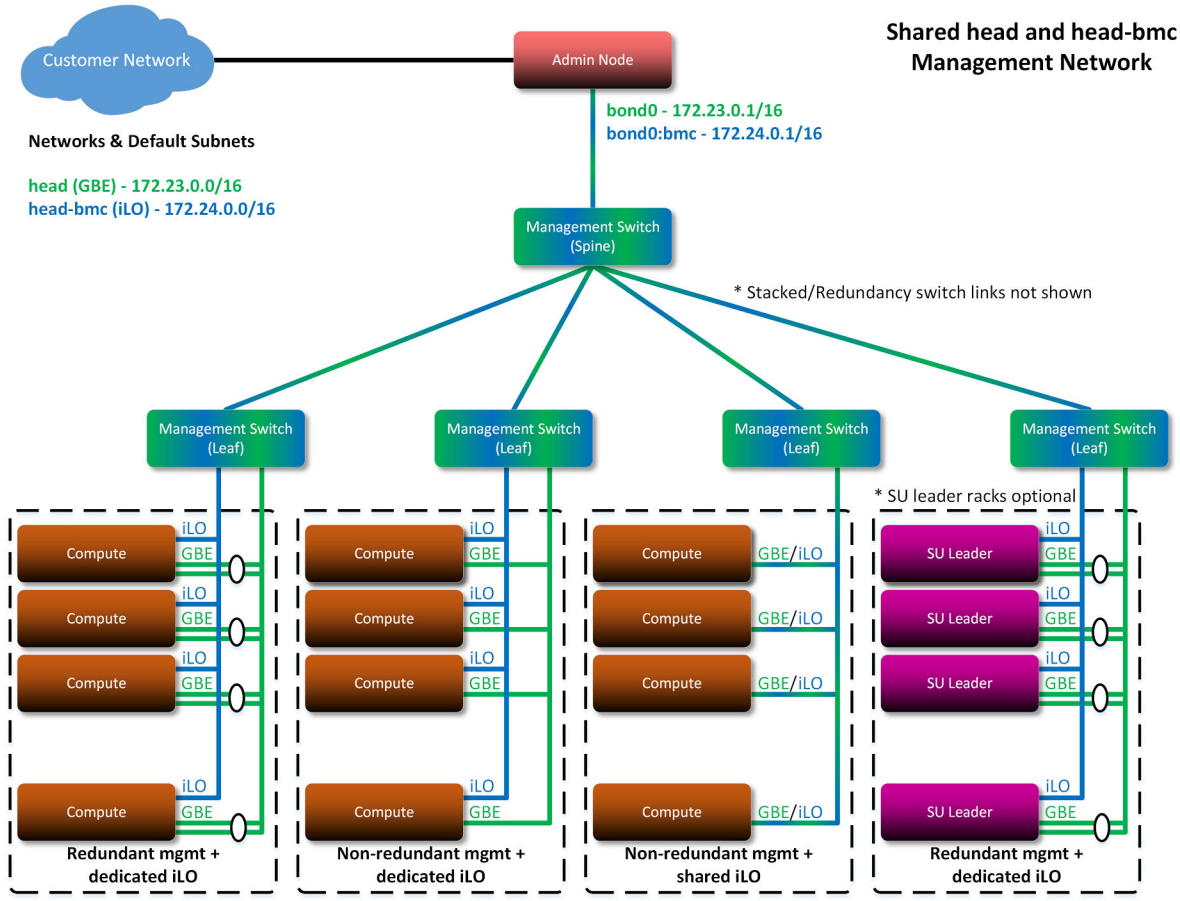

Figure 13: Shared head and head-bmc management network for clusters with SU leader nodes

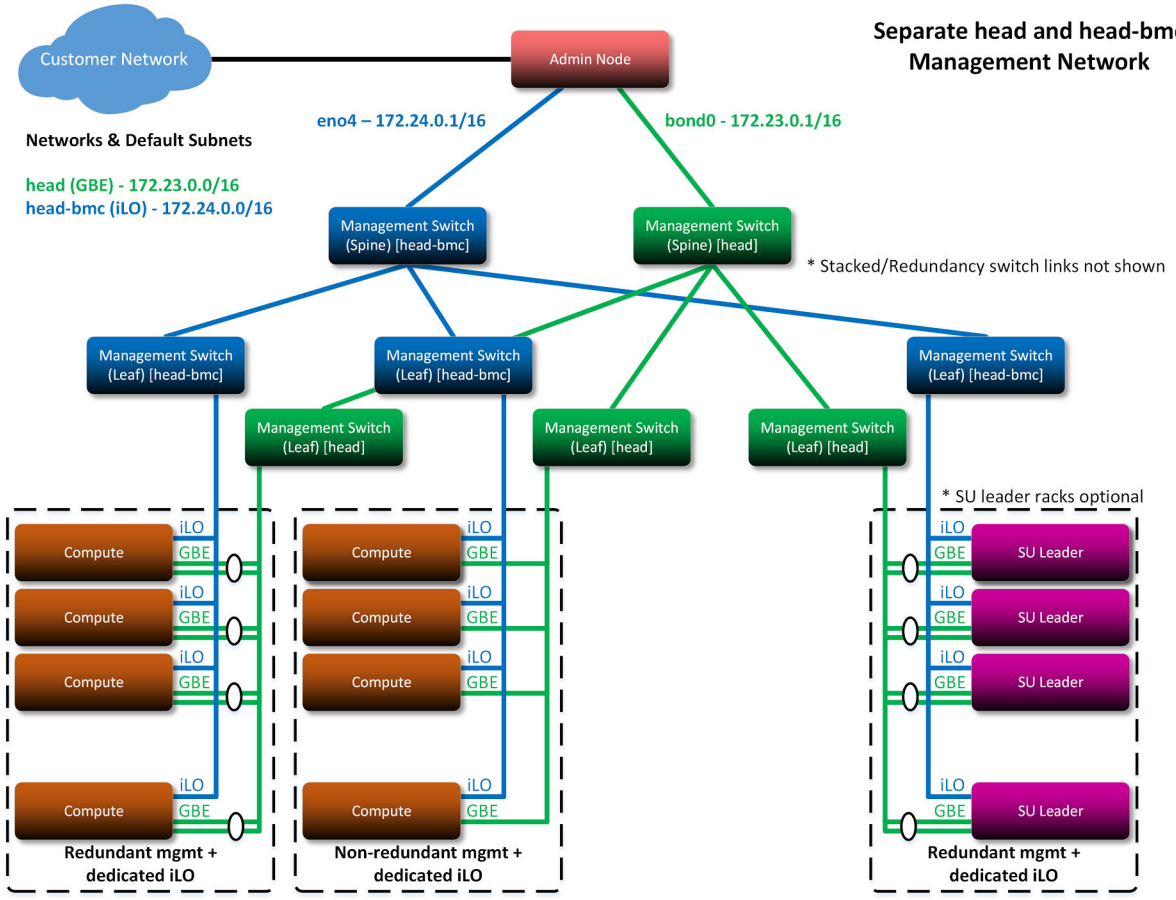

Figure 14: Separate head and head-bmc management network for clusters with SU leader nodes

The following figure shows the management network for an HPE Cray EX supercomputer.

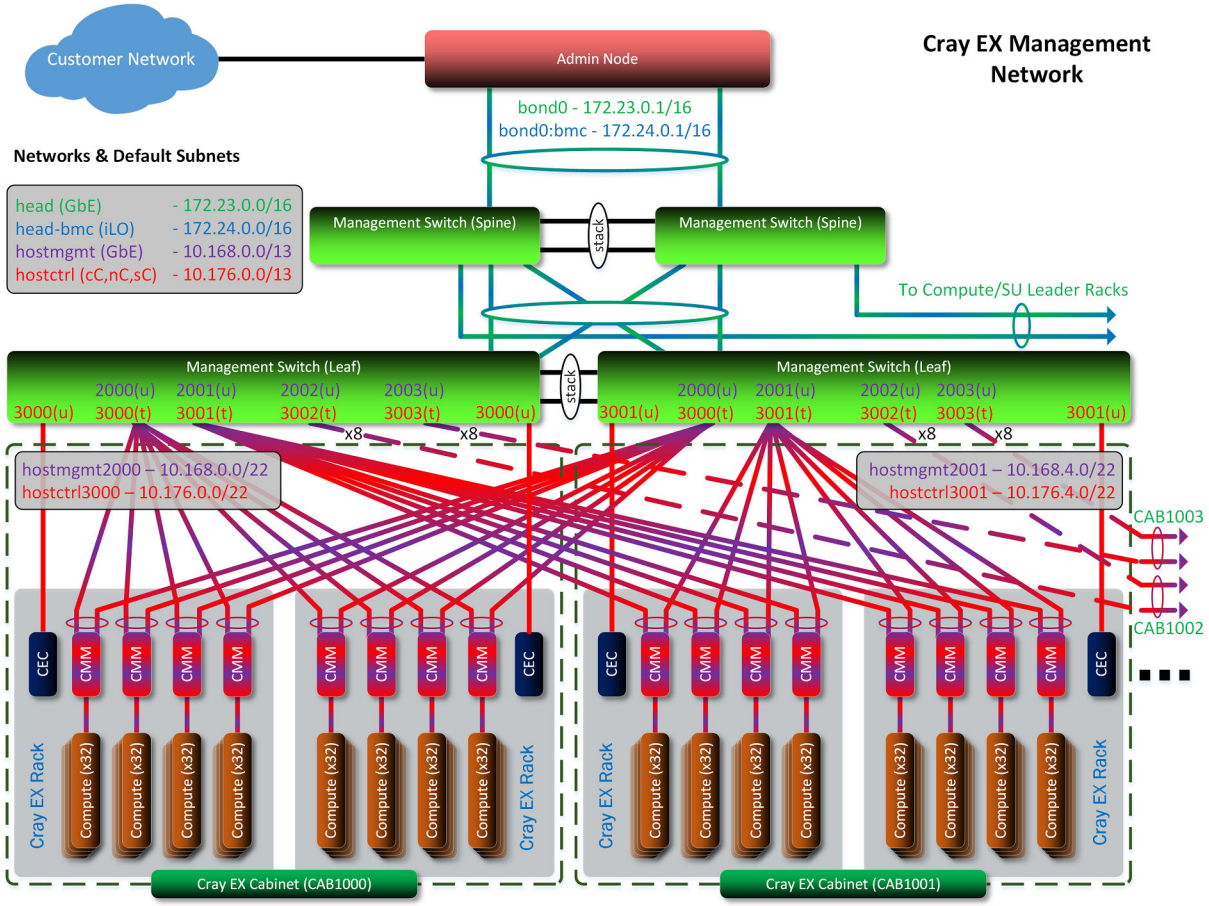

Figure 15: HPE Cray EX management network

Data networks in clusters with ICE leader nodes

The system components in a cluster with ICE leader nodes are attached to a management network and one or more data networks.

The data network is designed for high-performance computing and bandwidth-intensive applications. This high-speed data fabric network is based on either InfiniBand (IB) technology or Omni-Path technology. The data network facilitates communication to all ICE compute nodes from the OB or Omni-Path fabric.

The fabric network connects the ICE compute nodes to each other. The ICE compute node InfiniBand connections are not part of the head network. The fabric network also connects the compute nodes to the ICE compute nodes.

In addition, one or two separate fabric networks segregate traffic within a cluster with ICE compute nodes in a way that optimizes computing performance.

For example, when there are two fabric networks, each interface segregates communication, as follows:

- `ib0`, which is used typically for HPE Message Passing Interface (MPI) communication
- `ib1`, which is typically used for storage traffic

Management networks in clusters with ICE leader nodes

The system components in a cluster with ICE leader nodes are attached to a management network and one or more data networks.

The management network, also known as the head network, is designed for monitoring, provisioning, and other functions not covered by the data network. This Ethernet network facilitates communication between the admin node, the ICE

leader nodes, and the compute nodes. These components communicate to each other directly within the head network. The head network connects the following nodes directly into the Ethernet switches:

- Admin node
- ICE leader nodes
- Compute nodes
- More Ethernet switches

The site network interface connects the cluster to the site network infrastructure by way of a site-specific IP address.

The head network also includes several additional virtual local area networks (VLANs). The following figure is a logical representation of the Ethernet networks on a cluster with ICE leader nodes. The figure shows how the VLAN boundaries logically separate components.

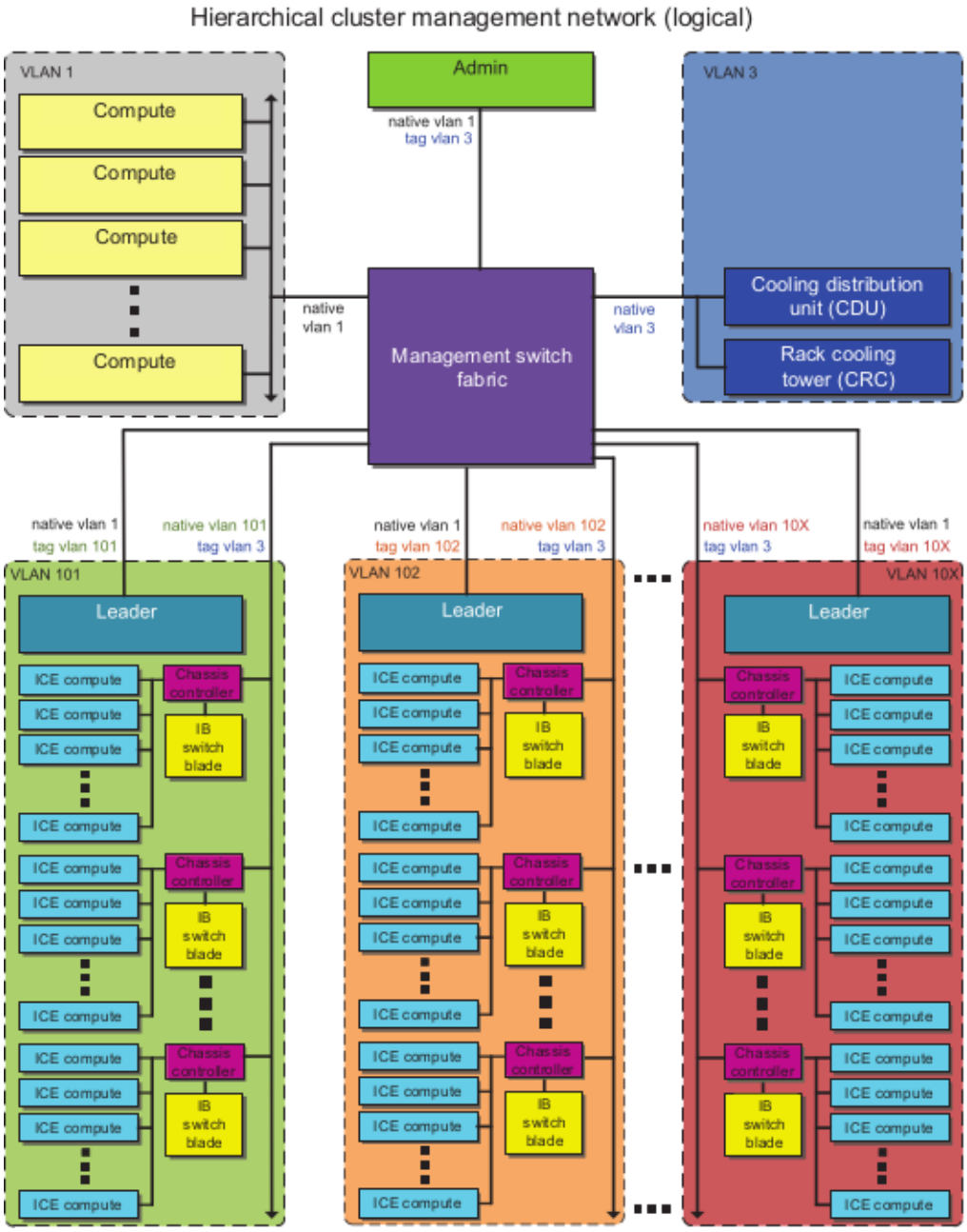

Figure 16: HPE SGI 8600 hierarchical cluster management network (logical) - includes ICE leader nodes

A cluster with ICE leader nodes includes several virtual local area networks (VLANs). The following list shows the typical VLAN names and numbers:

- Head network. VLAN tag 1.
- Cooling network. VLAN tag 3. Clusters with liquid cooling cells only.
- Rack 1 network. VLAN tag 101.
- Rack 2 network. VLAN tag 102.
- Rack *x* network. VLAN tag 10*x*.

NOTE: The head VLAN network must always be VLAN 1. Do not attempt to change the VLAN number of the head VLAN network. You can change the other VLAN numbers.

In the preceding figure, the head network is VLAN 1. The ports connected to the admin node and the compute nodes are in VLAN 1 natively.

The Ethernet Switches are configured with a VLAN for each ICE leader node. This VLAN segregates management traffic. In this VLAN, communications between the ICE compute nodes and the corresponding ICE leader node are contained within that VLAN. Physically, the ICE compute nodes are connected to a chassis controller and do not directly connect to the Ethernet switch. Instead, the chassis controllers connect directly to the Ethernet Switch. Only the ICE leader node can communicate with the ICE compute nodes and chassis controllers in its own logical rack.

Users can log into the admin node and into the compute nodes directly. If access to the ICE leader node is required, users can log in directly to the admin node and then use the `ssh` command to log into an ICE leader node. The typical VLAN mapping for the Ethernet switches on each node is as follows:

Node type	VLANs
Admin	Native VLAN 1 Tagged VLAN 3
Compute	Native VLAN 1
ICE leader node	Native VLAN 1 Tagged VLAN 10*x*. VLAN created for each ICE leader node.
Chassis controllers	Native VLAN 10*x*. VLAN created for their corresponding ICE eader. Tagged VLAN 3.
Cooling equipment, cooling distribution units (CDUs) and cooling rack controllers (CRCs)	Native VLAN 3. Clusters with liquid cooling cells only.

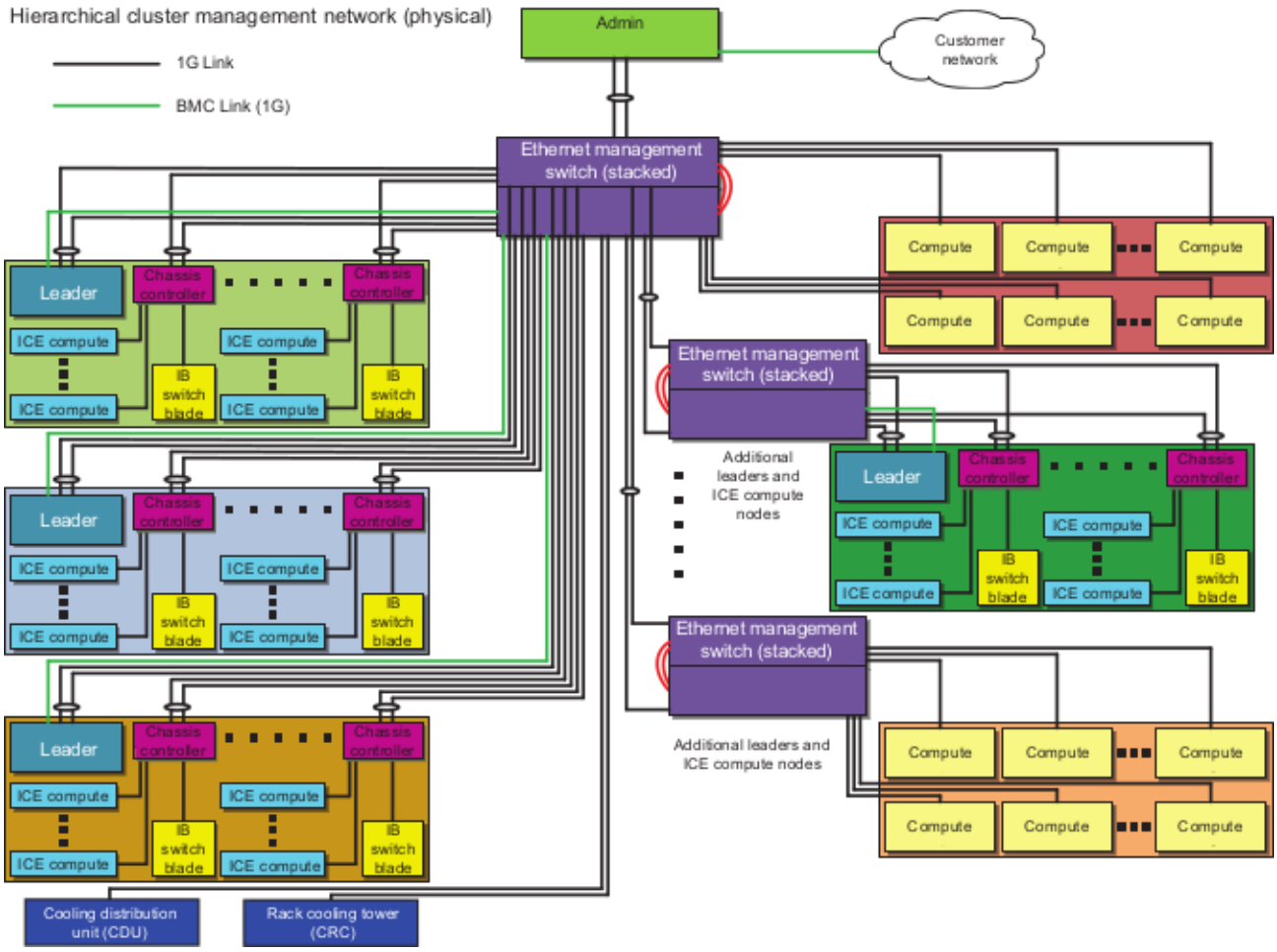

Figure 17: HPE SGI 8600 hierarchical cluster management network (physical) - includes ICE leader nodes

Attaching a cluster without a high availability (HA) admin node to your site network

The process of attaching a cluster without an HA admin node to your site network assumes that you want to keep the factory configuration and that the cluster has a single physical admin node. That is, the cluster does **not** include one of the following high availability admin node implementations:

- Quorum high availability (Quorum HA)
- System admin controller with high availability (SAC HA)

The following procedures explain how to attach the cluster to your network and retain the factory-installed configuration:

Procedure

1. **Obtaining information for a RHEL 8, RHEL 7, SLES 15, or SLES 12 admin node**
2. **Specifying network information and changing passwords**
3. **Completing the configuration**

Obtaining information for a RHEL 8, RHEL 7, SLES 15, or SLES 12 admin node

If you gather information before you start, your configuration session can proceed more quickly. When you perform the configuration, you update the factory-installed, system-wide root password, and you update the time zone. In addition, you provide information about your site network for the network interface card (NIC) in the admin node.

Procedure

Gather information in the following table for the admin node:

Information needed	Specifics for the admin node
Factory-installed password	_____
Password for this system at your site	_____
IP address	_____
Hostname	_____
Fully qualified domain name (FQDN)	_____
Time zone	_____

Table Continued

Information needed	Specifics for the admin node
Site (house) NTP server (or servers) IP address	_____
Cluster subdomain name	_____

The preceding table shows most of the information you need to supply when you attach the cluster to your site network. When you configure the network, you also need to supply information that is specific to the admin node operating system.

Specifying network information and changing passwords

Prerequisites

Obtaining information for a RHEL 8, RHEL 7, SLES 15, or SLES 12 admin node

Procedure

1. Use the console attached to the admin node, and log into the admin node as the root user.
2. Enter the following command to retrieve the site network interface name from the cluster configuration tool:

   ```
   # configure-cluster
   ```

   In the popup window that appears, the first line contains the site (house) network interface name.

   For example, the name is `ens20f0` in the following message:

   ```
   Used cached values for house network (ens20f0) and/or Management
   Network Interfaces (ens3f1,ens3f2). Remove
   /etc/opt/sgi/configure-cluster-ethernets to be prompted again or
   use the cluster definition file.
   ```

   The preceding command shows example output. For this example admin node, the Ethernet interface name is `ens20f0`. This admin node uses predictable network names.

   Note the Ethernet interface name that this command displays. You need this information later in this procedure.
3. Click **Quit** to close the cluster configuration tool.
4. Enter the following command to retrieve the MAC address of the site network interface:

   ```
   # ip addr show
   ```

   In the output, look for the `link/ether` line for the site network interface. For example:

   ```
    # ip addr show
   1: lo:  mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1
       link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
       inet 127.0.0.1/8 scope host lo
          valid_lft forever preferred_lft forever
       inet6 ::1/128 scope host
          valid_lft forever preferred_lft forever
   2: ens20f0:  mtu 1500 qdisc mq state UP group default qlen 1000
       link/ether 00:25:90:fd:3d:a8 brd ff:ff:ff:ff:ff:ff
       inet 128.162.243.106/24 brd 128.162.243.255 scope global eth0
          valid_lft forever preferred_lft forever
       inet6 fe80::225:90ff:fefd:3da8/64 scope link
   ```

```
       valid_lft forever preferred_lft forever
3: ens20f1:  mtu 1500 qdisc mq master bond0 state UP group default qlen
1000
    link/ether 00:25:90:fd:3d:a9 brd ff:ff:ff:ff:ff:ff
4: ens20f2:  mtu 1500 qdisc mq state DOWN group default qlen 1000
    link/ether 00:25:90:fd:3d:aa brd ff:ff:ff:ff:ff:ff
```

Note the MAC address that this command displays. You need this information later in this procedure.

5. Update the configuration file for the Ethernet interface and the network.

   Use the instructions in one of the following tables to complete this step:

   - For RHEL installations, see **Table 2: Updating network information (RHEL)**. For more information about the inputs to this configuration file, see one of the following:
     - Your RHEL documentation
     - The **nm-settings-ifcfg-rh** manpage
   - For SLES installations, see **Table 3: Updating network information (SLES)**.

Table 2: Updating network information (RHEL)

Step	Action
A.	Open file `/etc/sysconfig/network-scripts/ifcfg-`*name*. For example, this file might be `ifcfg-ens20f0`.

B.	Update the following fields in the `ifcfg-`*name* file:
	
	`NAME=`*name* `# Add the name of the house NIC`
	`DEVICE=`*name* `# Add the name of the house NIC`
	`IPADDR= # Add the IP address of this admin node`
	`PREFIX= # Add your site netmask setting. For example: 24`
	`GATEWAY= # Add your site gateway`
	`DNS1= # Add your site primary DNS IP address`
	`DNS2= # Add your site secondary DNS IP address`
	`DOMAIN= # Add your site domain`
	`HWADDR= # Add the NIC MAC address`
	`BOOTPROTO= # Set to "none"`
	`ONBOOT= # Set to "yes"`
	`DEFROUTE= # Set to "yes"`
	`TYPE= # Set to "Ethernet"`
	`UUID= # Enter "nmcli connection show" to get the UUID value`

C.	Save and close file `/etc/sysconfig/network-scripts/ifcfg-`*name*.
D.	Open file `/etc/sysconfig/network`.

Table Continued

Step	Action
E.	Specify the global default gateway by adding a line that begins with `GATEWAY=` and specifying the IP address of your site default gateway. For example, `GATEWAY=135.135.135.135`.
F.	Save and close file `/etc/sysconfig/network`.

Table 3: Updating network information (SLES)

Step	Action
A.	Open file `/etc/sysconfig/network/ifcfg-`*name*.

B.	Edit the file to add the admin node IP address, and update the following field:
	
	`IPADDR=''  # replace with admin node's IP address and prefix`
	
	The following is an example of the `IPADDR` field:
	
	`IPADDR='150.166.33.110/24'`

C.	Save and close file `/etc/sysconfig/network/ifcfg-`*name*.
D.	Open file `/etc/sysconfig/network/routes`.
E.	Search for the line that begins with `default`.
F.	Edit the `default ...` line to contain the following sequence: • The keyword `default` • A space • Your site gateway IP address • A space • A dash • A space • A dash For example: `default 100.100.100.100 - -`
G.	Save and close file `/etc/sysconfig/network/routes`.

6. Familiarize yourself with the cluster networks.

The cluster networks determine the DNS search path or search paths that you must specify. Complete the following steps:

a. Enter the following command to display the cluster configuration file:

```
discover --show-configfile
```

b. Search for the `[attributes]` section and the `[networks]` section, and examine the contents of both sections.

For example:

```
[attributes]
admin_house_interface=eno1
admin_mgmt_interfaces="ens2f0,ens2f1"
admin_mgmt_bmc_interfaces="ens2f0,ens2f1"
admin_udpcast_ttl=2
admin_udpcast_mcast_rdv_addr=239.255.255.1
admin_mgmt_bonding_mode=802.3ad
blademond_scan_interval=120
cmcs_per_mgmt_vlan=8
cmcs_per_rack=4
cmms_per_rack=8
conserver_logging=yes
conserver_ondemand=no
conserver_timestamp=no
copy_admin_ssh_config=yes
dhcp_bootfile=grub2
discover_skip_switchconfig=no
domain_search_path=head.cm.clusterdomain.com,hostmgmt.cm.clusterdomain.com,
head-bmc.cm.clusterdomain.com,hostctrl.cm.clusterdomain.com,
cm.clusterdomain.com,clusterdomain.com
head_vlan=1
ipv6_local_site_ula=fdee:1c3e:8af1::/48
max_rack_irus=16
mcell_network=no
mcell_vlan=3
mgmt_ctrl_vlan_end=3999
mgmt_ctrl_vlan_start=3001
mgmt_net_routing_protocol=ospf
mgmt_net_subnet_selection=rack-based
mgmt_vlan_end=2999
mgmt_vlan_start=2001
predictable_net_names=yes
rack_start_number=1
rack_vlan_end=1100
rack_vlan_start=101
redundant_mgmt_network=yes
switch_mgmt_network=yes
udpcast_max_bitrate=900m
udpcast_max_wait=10
udpcast_mcast_rdv_addr=224.0.0.1
udpcast_min_receivers=1
udpcast_min_wait=10
udpcast_rexmit_hello_interval=0
monitoring_kafka_elk_alerta_enabled=no
monitoring_native_enabled=no

[networks]
name=public, subnet=129.111.3.0, netmask=255.255.255.0, gateway=129.111.3.1
name=head, type=mgmt, vlan=1, subnet=172.23.0.0, netmask=255.255.0.0, gateway=172.23.255.254
name=head-bmc, type=mgmt-bmc, vlan=1, subnet=172.24.0.0, netmask=255.255.0.0
name=hostctrl, type=mgmt-bmc, subnet=10.176.0.0, netmask=255.248.0.0,
rack_netmask=255.255.252.0
name=hostmgmt, type=mgmt, subnet=10.168.0.0, netmask=255.248.0.0, rack_netmask=255.255.252.0
name=ib0, type=ib, subnet=10.148.0.0, netmask=255.255.0.0
name=ib1, type=ib, subnet=10.149.0.0, netmask=255.255.0.0
```

NOTE: Some text in the preceding example has been wrapped for inclusion in this documentation.

7. Use the `cadmin` command, in the following format, to set the admin node domain name:

   `# cadmin --set-admin-domain site_domain`

   For *site_domain*, specify the full name of your site domain. For example, `clusterdomain.com`.

8. Use the `cadmin` command, in the following format, to change the domain name for the cluster:

   `# cadmin --set-cluster-domain cluster_name`

   For *cluster_name*, specify the name of the cluster. For example, `cm.clusterdomain.com`.

   For more information about the `cadmin` command, type `cadmin -h` at the system prompt.

9. Use the `cadmin` command, in the following format, to set the correct DNS paths:

   `cadmin --set-domain-search-path path`

   For *path*, specify the search paths you need. The maximum number of paths is six. The maximum number of characters that you can specify in the *path* argument is 256.

   The following table shows example search paths.

Cluster type	Pathname
Clusters that include an InfiniBand data network or an Omni-Path data network	This path name typically starts with `ib0` followed by the cluster domain. For example, modify the search paths so that they appear as follows: `ib0.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`.
Clusters that include an HPE Slingshot data network	This path name typically starts with `hsn0` followed by the cluster domain. For example, modify the search paths so that they appear as follows: `hsn0.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`
Clusters that include an Ethernet data network	Modify the search paths to omit the `ib0` entries. For example, modify the search paths so that only the following appears: `datanet.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`

10. Use a text editor to open file `/etc/hosts`.

11. Add a line in the following format to file `/etc/hosts`:

    admin_node_IP admin_node_FQDN admin_node_hostname

    The variables are as follows:

Variable	Specification
admin_node_IP	The IP address of the admin node.
admin_node_FQDN	The fully qualified domain name (FQDN) of the admin node.
admin_node_hostname	The hostname of the admin node.

For example, add the following line:

```
100.162.244.88 admin.cm.clusterdomain.com admin
```

12. Save and close file `/etc/hosts`.
13. Enter the following command to verify the hostname of the cluster:

    ```
    # cm node show -n admin
    ```

14. Use the `cm node set` command in the following format to set the hostname of the admin node:

    ```
    cm node set -n admin --name admin_node_hostname
    ```

    For *admin_node_hostname*, enter the hostname you want to assign to the admin node. Make sure to enter the hostname, which is the short name. Do not enter the FQDN of the admin node, which is the longer name.

15. Restart the network daemon.

    On RHEL 8.X clusters, enter the following command:

    ```
    # systemctl restart NetworkManager
    ```

    On RHEL 7.X, SLES 15 SPX, and SLES 12 SPX, clusters, enter the following command:

    ```
    # systemctl restart network
    ```

16. Enter the following command to restart the name service cache daemon, `ncsd`:

    ```
    # systemctl restart nscd
    ```

17. Complete the following steps to change the system time zone:

    a. Enter the following command to display a list of time zones:

    ```
    # timedatectl list-timezones
    ```

    b. Enter the following command to set the time zone:

    ```
    # timedatectl set-timezone time_zone
    ```

    For *time_zone*, specify one of the time zones from the `timedatectl list-timezones` command output.

    c. Use the `timedatectl` command to display the time zone information you configured. For example:

    ```
    # timedatectl
          Local time: Fri 2021-04-15 14:55:33 PDT
      Universal time: Fri 2021-04-15 21:55:33 UTC
            RTC time: Fri 2021-04-15 21:55:33
           Time zone: America/Los_Angeles (PDT, -0700)
         NTP enabled: yes
    NTP synchronized: yes
    ```

```
RTC in local TZ: no
     DST active: yes
Last DST change: DST began at
                 Sun 2021-03-13 01:59:59 PST
                 Sun 2021-03-13 03:00:00 PDT
Next DST change: DST ends (the clock jumps one hour backwards) at
                 Sun 2021-11-06 01:59:59 PDT
                 Sun 2021-11-06 01:00:00 PST
```

18. Enter the following command to synchronize time:

```
# cm node update config --sync ntp -n "*"
```

19. Enter the following command, and examine the output:

```
# cm power status -t system
```

Make sure that all cluster nodes are up and running.

If some nodes are not up, enter the following command:

```
# cm power on -t system
```

20. (Conditional) Start the scalable unit (SU) leader nodes.

Complete this step if there are SU leader nodes on the cluster.

Complete the following steps:

a. Enter the following command to power on the SU leader nodes:

```
cm power on -t 'leader_hostname*'
```

For *leader_hostname*, specify the portion of the SU leader node hostname that all SU leader nodes share.

For example:

```
# cm power on -t 'leader*'
```

b. For each SU leader node, `ssh` to the leader node, retrieve the status of SU leader nodes in the cluster database, and examine the status messages.

Use one of the following methods:

- Method 1 - Use the `clush` command. For example:

  ```
  # clush -g su-leader ctdb status
  ```

- Method 2 - Log into each SU leader node individually, retrieve the status of SU leader nodes in the cluster database, and examine the status messages.

  For example, the following command logs into one SU leader node:

  ```
  # ssh leader1 ctdb status
  Number of nodes:3
  pnn:0 172.23.0.2 OK (THIS NODE)
  pnn:1 172.23.0.3 OK
  pnn:2 172.23.0.4 OK
  Generation:370917201
  ```

```
Size:3
hash:0 lmaster:0
hash:1 lmaster:1
hash:2 lmaster:2
Recovery mode:NORMAL (0)
Recovery master:1
```

If you use Method 2, make sure to enter the command on each SU leader node.

If the `ctdb` status shows anything other than `OK` or if the Gluster volumes do not look to be in the right state, log into the admin node and run the following command:

```
# su-leader-setup --bring-online
```

The **`--bring-online`** parameter ensures that the volumes and `ctdb` status are correct. When this command finishes, `ssh` to the SU leader node again, and run the verification steps again on that node.

NOTE: You can run the `su-leader-setup --bring-online` command at any time the volume or `ctdb` status is not correct. For example, you can run this command if an SU leader node goes down and the SU leader node does not register itself properly with `ctdb` after it comes back up.

---

c. From the admin node, enter the following command to mount the Gluster file system on the admin node:

```
# enable-su-leader --skip-sync
```

d. Power on the compute nodes.

For example:

```
# cm power on -t node '*'
```

21. Change the root password on the admin node and all other managed nodes.

Enter the following command, and follow the prompts:

```
# cpasswd
```

For example:

```
admin node:~ # cpasswd
Enter new password:
Enter new password (again):
admin: updating /etc/shadow
r1lead: updating /etc/shadow
n0: updating /etc/shadow
admin node:~ #
```

If you need help, enter the following command:

```
# cpasswd -h
```

The `cpasswd` command changes the password on all nodes that are booted at the time the command is run.

To change the root password on only one node, enter the `cpasswd` command in the following format:

```
cpasswd -N node_hostname
```

If you do not have the current password, you can obtain the factory-installed password from your technical support representative.

22. Use the `switchconfig` command to change the management switch password for the `admin` account.

The format for this command is as follows:

`switchconfig change_password --switches` *`hostname`* `--new` *`new_password`*

The variables are as follows:

Variable	Specification
hostname	The hostname of the management switch or the keyword `all`. To specify a different password for each switch, enter a `switchconfig change_password` command for each switch. To specify the same password for each switch, use the keyword `all`.
new_password	A new password for the switch.

For example:

```
# switchconfig change_password --switches mgmtsw0 --new Hp3@dm!n2o20
```

NOTE: Hewlett Packard Enterprise recommends that you implement standard and secure practices to store all passwords at your site. Do not lose this information.

Attaching a cluster with a high availability (HA) admin node to your site network

The process of attaching a cluster with an HA admin node to your site network assumes that you want to keep the factory configuration and that the cluster has an HA admin node. An HA admin node consists of one of the following:

The following procedures explain how to attach the cluster to your network and retain the factory-installed configuration:

Procedure

1. **Obtaining information for a high availability (HA) admin node**
2. **Setting the hostname and the time zone on the physical admin nodes**
3. **Updating the network configuration file on the physical admin nodes**
4. **Bringing the virtual machine (VM) admin node online**
5. **Setting the hostname and the time zone on the virtual machine (VM) admin node**
6. **Updating the network configuration file on the virtual machine (VM) admin node**
7. **Specifying domain information**
8. **Updating the `/etc/hosts` file on the virtual machine (VM) admin node**
9. **Checking the status of the cluster nodes and setting the cluster password**
10. **Completing the configuration**

Obtaining information for a high availability (HA) admin node

If you gather information before you start, your configuration session can proceed more quickly. When you perform the configuration, you update the factory-installed, system-wide root password, and you update the time zone. In addition, you provide information about your site network for the network interface card (NIC) in the admin node.

The admin node is actually a virtual machine (VM) hosted by one of the physical admin nodes. When performing administrative tasks, the administrator must log into the VM admin node as the root user. The procedure in this topic explains the information to gather for the physical admin nodes and the VM HA admin node. When your cluster includes an HA admin node, you must attach the physical admin nodes and the VM admin node to your network. When discussing failover activities, the documentation refers to the physical admin nodes as the **primary node**, the **secondary node**, and the **tertiary node**. Only a quorum HA configuration has a third, tertiary node.

The following values are required to be the same for the physical admin nodes and the VM admin node:

- Time zone
- Default route/Gateway
- Netmask/Prefix
- Site (house) NTP server or servers
- First site DNS resolver IP address
- Second site DNS resolver IP address
- Third site DNS resolver IP address

- Site domain
- Cluster subdomain name

The following procedure explains the information that you need for this process.

Procedure

1. Gather the information in the following table for the primary physical admin node:

Information needed	Specifics for the primary physical admin node
Factory-installed password	_____
Password for this system at your site	_____
IP address	_____
Hostname	_____
Fully qualified domain name (FQDN)	_____
Time zone	_____
Site (house) NTP server or servers IP address	_____
Cluster subdomain name	_____

   The preceding table shows most of the information you need to supply when you attach the cluster to your site network. When you configure the network, you also need to supply information that is specific to the admin node operating system.

2. Gather the information in the following table for the secondary physical admin node.

Information needed	Specifics for the secondary physical admin node
Factory-installed password	_____
Password for this node at your site	_____
IP address	_____
Hostname	_____
FQDN	_____

3. (Conditional) Gather information for the tertiary physical admin node.

   Complete this step only for quorum HA admin nodes. The information you need to gather is as follows.

Information needed	Specifics for the tertiary physical admin node
Factory-installed password	_____
Password for this node at your site	_____
IP address	_____
Hostname	_____
FQDN	_____

4. Gather information in the following table for the VM admin node:

Table 4: VM admin node information

Information needed	Specifics for the VM admin node
Factory-installed password	_____
Password for the VM admin node at your site	_____
IP address	_____
Hostname	_____
FQDN	_____

After the cluster is configured on the site network, log into the VM admin node to perform system administration tasks.

Setting the hostname and the time zone on the physical admin nodes

Prerequisites

Obtaining information for a high availability (HA) admin node

Procedure

1. Log into one of the physical admin nodes as the root user, and enter the following command to set the hostname of the node:.

   ```
   hostnamectl set-hostname hostname_of_this_node
   ```

For example:

```
# hostnamectl set-hostname phys1admin
```

2. Log into another physical admin node as the root user, and enter the following command to set the hostname of the node:

   ```
   hostnamectl set-hostname hostname_of_this_node
   ```

   For example:

   ```
   # hostnamectl set-hostname phys2admin
   ```

3. (Conditional) Log into the last physical admin node as the root user, and set the hostname of the node.

   Complete this step on the third physical admin node of a quorum HA admin node.

   Enter the following command:

   ```
   hostnamectl set-hostname hostname_of_this_node
   ```

   For example:

   ```
   # hostnamectl set-hostname phys3admin
   ```

4. Complete the following steps to set the time zone of the node:

   a. Enter the following command to display a list of time zones:

      ```
      # timedatectl list-timezones
      ```

   b. Enter the following command to set the time zone:

      ```
      # timedatectl set-timezone time_zone
      ```

      For *time_zone*, specify one of the time zones from the `timedatectl list-timezones` command output.

   c. Use the `timedatectl` command to display the time zone information you configured.

      For example, for RHEL nodes, the output is as follows:

      ```
      # timedatectl
            Local time: Fri 2021-09-15 14:55:33 PDT
        Universal time: Fri 2021-09-15 21:55:33 UTC
              RTC time: Fri 2021-09-15 21:55:33
             Time zone: America/Los_Angeles (PDT, -0700)
          NTP enabled: yes
      NTP synchronized: yes
       RTC in local TZ: no
            DST active: yes
       Last DST change: DST began at
                        Sun 2021-03-13 01:59:59 PST
                        Sun 2021-03-13 03:00:00 PDT
       Next DST change: DST ends (the clock jumps one hour backwards) at
                        Sun 2021-11-06 01:59:59 PDT
                        Sun 2021-11-06 01:00:00 PST
      ```

5. Complete the following steps to direct network time protocol (NTP) server requests to the server at your site rather than the public time servers of the `pool.ntp.org` project:

a. Use a text editor to open file `/etc/chrony.conf`.

b. Insert a pound character (#) into column 1 of each line that includes `rhel.pool.ntp.org`.

NOTE: Do not edit or remove entries that serve the cluster networks.

c. At the end of the file, add a line that points to your site NTP server.

The following is an example of a correctly edited file:

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
# server 0.rhel.pool.ntp.org
# server 1.rhel.pool.ntp.org
# server 2.rhel.pool.ntp.org
server ntp.mycompany.com iburst
```

d. Enter the following line to allow clients on the management network to query the NTP server:

```
allow 172.23
```

e. Save and close the file.

f. Enter the following command to restart the NTP server:

```
# systemctl restart chronyd
```

6. Enter the following command to synchronize time:

```
# cm node update config --sync ntp -n '*'
```

Updating the network configuration file on the physical admin nodes

The procedure in this topic explains how to update the network configuration file with your networking information for your site. Complete this procedure once on each physical admin node.

Prerequisites

Setting the hostname and the time zone on the physical admin nodes

Procedure

1. Use the console attached to one of the physical admin nodes to log into that admin node as the root user.
2. (Conditional) Disable fencing.

   Complete this step on the first physical admin node of a system admin controller high availability (SAC HA) admin node. This step ensures that no failover occurs when you attach the nodes to your network. Do not complete this step on any of the following node types:

   - On the second or third physical node of a SAC HA admin node.
   - On any node of a quorum HA admin node.

   To complete this step,.enter one of the following command sets:

- On RHEL systems, enter the following commands:

  ```
  # pcs resource disable p_ipmi_fencing_1
  # pcs resource disable p_ipmi_fencing_2
  ```

- On SLES systems, enter the following commands:

  ```
  # crm resource stop p_ipmi_fencing_1
  # crm resource stop p_ipmi_fencing_2
  ```

3. Enter the following command to retrieve the MAC address of `br0` on this admin node:

   ```
   # ip addr show
   ```

   In the output, note the MAC address of `br0`. You need this MAC address later in this procedure.

   For example, in the following output, the MAC address is `0c:c4:7a:43:8e:d0`:

   ```
   # ip addr show
   .
   .
   .
   8: br0:  mtu 1500 qdisc noqueue state UP group default qlen 1000
       link/ether 0c:c4:7a:43:8e:d0 brd ff:ff:ff:ff:ff:ff
       inet 150.166.33.110/24 brd 150.166.33.255 scope global br0
          valid_lft forever preferred_lft forever
       inet6 fe80::ec4:7aff:fe43:8ed0/64 scope link
          valid_lft forever preferred_lft forever
   .
   .
   .
   ```

4. Update the configuration file for `br0`.

   Use the instructions in one of the following tables to complete this step:

   - For RHEL installations, see **Table 5: Updating network information on a physical admin node (RHEL)**. For more information about the inputs to this configuration file, see one of the following:

     - Your RHEL documentation
     - The **nm-settings-ifcfg-rh** manpage

   - For SLES installations, see **Table 6: Updating network information on a physical admin node (SLES)**.

Table 5: Updating network information on a physical admin node (RHEL)

Step	Action
A.	Open file `/etc/sysconfig/network-scripts/ifcfg-br0`.

B.	Update the contents of the following fields in the `ifcfg-br0` file:
	`IPADDR=  # Add the IP address of this admin node`
	`PREFIX=  # Add your site netmask setting.  For example:  24`
	`GATEWAY=  # Add your site gateway`
	`DNS1=  # Add your site primary DNS IP address`
	`DNS2=  # Add your site secondary DNS IP address`
	`DOMAIN=  # Add your site domain`
	`HWADDR=  # Add the NIC MAC address`
	`BOOTPROTO=  # Set to "none"`
	`ONBOOT=  # Set to "yes"`
	`DEFROUTE=  # Set to "yes"`
	`TYPE=  # Set to "Bridge"`
	`UUID=  # Enter "nmcli connection show" to get the UUID value`

C.	Save and close file `/etc/sysconfig/network-scripts/ifcfg-br0`.

Table 6: Updating network information on a physical admin node (SLES)

Step	Action
A.	Open file `/etc/sysconfig/network/ifcfg-br0`
B.	Edit the file to add the admin node IP address: `IPADDR=''  # replace with admin node's IP address and prefix` The following is an example of the `IPADDR` field: `IPADDR='150.166.33.110/24'`
C.	Save and close file `/etc/sysconfig/network/ifcfg-br0`.
D.	Open file `/etc/sysconfig/network/routes`.
E.	Search for the line that begins with `default`.

Table Continued

Step	Action
F.	Edit the `default ...` line to contain the following sequence: • The keyword `default` • A space • Your site gateway IP address • A space • A dash • A space • A dash For example: `default 128.162.243.1 - -`
G.	Save and close file `/etc/sysconfig/network/routes`.

5. Open file `/etc/resolv.conf`, and update the domain name service (DNS) information.

   For example:

   ```
   search public.clusterdomain.com clusterdomain.com
   nameserver 93.184.216.37
   nameserver 93.184.216.38
   ```

6. Repeat the preceding steps on another physical admin node.

Bringing the virtual machine (VM) admin node online

Prerequisites

Updating the network configuration file on the physical admin nodes

Procedure

1. On each physical admin node, enter the following command to reboot the node:

   ```
   # reboot
   ```

2. Enable fencing.

   Complete this step on a cluster with a system admin controller high availability (SAC HA) admin node. Do not complete this step on a cluster with a quorum HA admin node.

   Enter one of the following command sets:

- On RHEL systems, enter the following commands:

  ```
  # pcs resource enable p_ipmi_fencing_1
  # pcs resource enable p_ipmi_fencing_2
  ```

- On SLES systems, enter the following commands:

  ```
  # crm resource start p_ipmi_fencing_1
  # crm resource start p_ipmi_fencing_2
  ```

3. After the `login:` prompt appears on each node, wait at least one minute before proceeding to the next step in this procedure.
4. Verify that these physical nodes are working as you configured them in the preceding steps.

   For example, verify that the IP address and hostname of each node is correct.
5. On each physical admin node, enter the following command to verify that the `/images` directory is mounted:

   ```
   # df -h | grep "/images"
   ```
6. On the first physical admin node, enter the following commands to restart the VM admin node:

   ```
   # pcs resource manage virt
   # pcs resource enable virt
   # virsh console sac
   ```

Setting the hostname and the time zone on the virtual machine (VM) admin node

Prerequisites

Bringing the virtual machine (VM) admin node online

Procedure

1. Log into the VM admin node as the root user.
2. Enter the following command to verify the hostname of the cluster:

   ```
   # cm node show -n admin
   ```
3. Use the `cm node set` command in the following format to set the hostname of the admin node:

   `cm node set -n admin --name` ***`admin_node_hostname`***

   For *admin_node_hostname*, enter the hostname you want to assign to the admin node. Make sure to enter the hostname, which is the short name. Do not enter the FQDN of the admin node, which is the longer name.
4. Use the `ip addr show` command to retrieve the MAC address of the site NIC.

   For example:

   ```
   # ip addr show
   1: lo:  mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1
       link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
       inet 127.0.0.1/8 scope host lo
          valid_lft forever preferred_lft forever
   ```

```
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: ens3:  mtu 1500 qdisc mq state UP group default qlen 1000
    link/ether 52:54:00:B8:A8:72 brd ff:ff:ff:ff:ff:ff
    inet 128.162.243.106/24 brd 128.162.243.255 scope global ens20f0
       valid_lft forever preferred_lft forever
    inet6 fe80::225:90ff:fefd:3da8/64 scope link
       valid_lft forever preferred_lft forever
.
.
.
```

In this example, the MAC address of `ens3` is `52:54:00:B8:A8:72`.

5. Update the configuration file for the Ethernet interface.

   Use the instructions in one of the following tables to complete this step:

   - For RHEL installations, see **Table 7: Updating network information on the VM admin node (RHEL)**.
   - For SLES installations, see **Table 8: Updating network information on the VM admin node (SLES)**.

Table 7: Updating network information on the VM admin node (RHEL)

Step	Action
A.	Open file `/etc/sysconfig/network-scripts/ifcfg-`*name*. For *name*, specify the name returned in the message from the `configure-cluster` command. For example, this file might be `ifcfg-ens3`.

B.	Update the contents of the `ifcfg-`*name* file so it contains data in the following fields:
	
	`IPADDR=  # Add the IP address of this admin node`
	`PREFIX=  # Add your site netmask setting.  For example:  24`
	`GATEWAY=  # Add your site gateway`
	`DNS1=  # Add your site primary DNS IP address`
	`DNS2=  # Add your site secondary DNS IP address`
	`DOMAIN=  # Add your site domain`
	`HWADDR=  # Add the NIC MAC address`
	`BOOTPROTO=  # Set to "none"`
	`ONBOOT=  # Set to "yes"`
	`DEFROUTE=  # Set to "yes"`
	`TYPE=  # Set to "Ethernet"`
	`UUID=  # Enter "nmcli connection show" to get the UUID value`

C.	Save and close file `/etc/sysconfig/network-scripts/ifcfg-`*name*.

Table 8: Updating network information on the VM admin node (SLES)

Step	Action
A.	Open file `/etc/sysconfig/network/ifcfg-`*`name`*
B.	Edit the file to add the admin node IP address: `IPADDR=''  # Add the IP address of the VM admin node` The following is an example of the `IPADDR` field: `IPADDR='150.166.33.110/24'`.
C.	Save and close file `/etc/sysconfig/network/ifcfg-`*`name`*.

6. (Conditional - SLES only) Open file `/etc/sysconfig/network/routes` and update the default route information.

   Complete the following steps for SLES admin nodes:

   a. Within the file, search for the line that begins with `default`.

   b. Edit the `default ...` line to contain the following sequence:

   - The keyword `default`
   - A space
   - Your site gateway IP address
   - A space
   - A dash
   - A space
   - A dash

   For example:

   ```
   default 128.162.243.1 - -
   ```

   c. Save and close file `/etc/sysconfig/network/routes`.

7. Restart the network daemon.

   On RHEL 8.X clusters, enter the following command:

   ```
   # systemctl restart NetworkManager
   ```

   On RHEL 7.X, SLES 15 SPX, and SLES 12 SPX, clusters, enter the following command:

   ```
   # systemctl restart network
   ```

8. Enter the following command to restart the name service cache daemon, `ncsd`:

   ```
   # systemctl restart nscd
   ```

9. Complete the following steps to set the time zone of the node:

a. Enter the following command to display a list of time zones:

```
# timedatectl list-timezones
```

b. Enter the following command to set the time zone:

```
# timedatectl set-timezone time_zone
```

For *time_zone*, specify one of the time zones from the `timedatectl list-timezones` command output.

c. Use the `timedatectl` command to display the time zone information you configured. For example:

```
# timedatectl
      Local time: Fri 2021-09-15 14:55:33 PDT
  Universal time: Fri 2021-09-15 21:55:33 UTC
        RTC time: Fri 2021-09-15 21:55:33
       Time zone: America/Los_Angeles (PDT, -0700)
     NTP enabled: yes
NTP synchronized: yes
 RTC in local TZ: no
      DST active: yes
 Last DST change: DST began at
                  Sun 2021-03-13 01:59:59 PST
                  Sun 2021-03-13 03:00:00 PDT
 Next DST change: DST ends (the clock jumps one hour backwards) at
                  Sun 2021-11-06 01:59:59 PDT
                  Sun 2021-11-06 01:00:00 PST
```

10. Complete the following steps to direct network time protocol (NTP) server requests to the server at your site rather than the public time servers of the `pool.ntp.org` project:

a. Use a text editor to open file `/etc/chrony.conf`.

b. Insert a pound character (#) into column 1 of each line that includes `rhel.pool.ntp.org`.

NOTE: Do not edit or remove entries that serve the cluster networks.

c. At the end of the file, add a line that points to your site NTP server.

The following is an example of a correctly edited file:

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
# server 0.rhel.pool.ntp.org
# server 1.rhel.pool.ntp.org
# server 2.rhel.pool.ntp.org
server ntp.mycompany.com iburst
```

d. Enter the following line to allow clients on the management network to query the NTP server:

```
allow 172.23
```

e. Save and close the file.

f. Enter the following command to restart the NTP server:

```
# systemctl restart chronyd
```

11. Enter the following command to synchronize time:

```
# cm node update config --sync ntp -n '*'
```

Updating the network configuration file on the virtual machine (VM) admin node

Prerequisites

Setting the hostname and the time zone on the virtual machine (VM) admin node

Procedure

1. Log into the VM admin node as the root user.
2. Enter the following command to retrieve the site network interface name from the cluster configuration tool:

   ```
   # configure-cluster
   ```

   In the popup window that appears, notice that the first line contains the site (house) network interface name.

   For example, the name is `ens3` in the following message:

   ```
   Used cached values for house network (ens3) and/or Management
   Network Interfaces (ens3f1,ens3f2). Remove
   /etc/opt/sgi/configure-cluster-ethernets to be prompted again or
   use the cluster definition file.
   ```

3. Click **Quit** to close the cluster configuration tool.

Specifying domain information

The following procedure explains how to specify additional domain information on the virtual machine (VM) admin node.

NOTE: Do not complete this procedure on the physical admin nodes.

Prerequisites

Updating the network configuration file on the virtual machine (VM) admin node

Procedure

1. As the root user, logged into the VM admin node, familiarize yourself with the cluster networks.

   The cluster networks determine the DNS search path or search paths to specify. Complete the following steps:

a. Enter the following command to display the cluster configuration file:

```
# discover --show-configfile
```

b. Search for the `[attributes]` section and the `[networks]` section, and examine the contents of both sections.

For example:

```
[attributes]
admin_house_interface=eno1
admin_mgmt_interfaces="ens2f0,ens2f1"
admin_mgmt_bmc_interfaces="ens2f0,ens2f1"
admin_udpcast_ttl=2
admin_udpcast_mcast_rdv_addr=239.255.255.1
admin_mgmt_bonding_mode=802.3ad
blademond_scan_interval=120
cmcs_per_mgmt_vlan=8
cmcs_per_rack=4
cmms_per_rack=8
conserver_logging=yes
conserver_ondemand=no
conserver_timestamp=no
copy_admin_ssh_config=yes
dhcp_bootfile=grub2
discover_skip_switchconfig=no
domain_search_path=head.cm.clusterdomain.com,hostmgmt.cm.clusterdomain.com,
head-bmc.cm.clusterdomain.com,hostctrl.cm.clusterdomain.com,
cm.clusterdomain.com,clusterdomain.com
head_vlan=1
ipv6_local_site_ula=fdee:1c3e:8af1::/48
max_rack_irus=16
mcell_network=no
mcell_vlan=3
mgmt_ctrl_vlan_end=3999
mgmt_ctrl_vlan_start=3001
mgmt_net_routing_protocol=ospf
mgmt_net_subnet_selection=rack-based
mgmt_vlan_end=2999
mgmt_vlan_start=2001
predictable_net_names=yes
rack_start_number=1
rack_vlan_end=1100
rack_vlan_start=101
redundant_mgmt_network=yes
switch_mgmt_network=yes
udpcast_max_bitrate=900m
udpcast_max_wait=10
udpcast_mcast_rdv_addr=224.0.0.1
udpcast_min_receivers=1
udpcast_min_wait=10
udpcast_rexmit_hello_interval=0
monitoring_kafka_elk_alerta_enabled=no
monitoring_native_enabled=no

[networks]
name=public, subnet=129.111.3.0, netmask=255.255.255.0, gateway=129.111.3.1
name=head, type=mgmt, vlan=1, subnet=172.23.0.0, netmask=255.255.0.0, gateway=172.23.255.254
name=head-bmc, type=mgmt-bmc, vlan=1, subnet=172.24.0.0, netmask=255.255.0.0
name=hostctrl, type=mgmt-bmc, subnet=10.176.0.0, netmask=255.248.0.0, rack_netmask=255.255.252.0
name=hostmgmt, type=mgmt, subnet=10.168.0.0, netmask=255.248.0.0, rack_netmask=255.255.252.0
name=ib0, type=ib, subnet=10.148.0.0, netmask=255.255.0.0
name=ib1, type=ib, subnet=10.149.0.0, netmask=255.255.0.0
```

NOTE: Some text in the preceding example has been wrapped for inclusion in this documentation.

2. Use the `cadmin` command, in the following format, to set the admin node domain name:

   `# cadmin --set-admin-domain site_domain`

   For *site_domain*, specify the full name of your site domain. For example, `clusterdomain.com`.

3. Use the `cadmin` command, in the following format, to change the domain name for the cluster:

   `# cadmin --set-cluster-domain cluster_name`

   For *cluster_name*, specify the name of the cluster. For example, `cm.clusterdomain.com`.

   For more information about the `cadmin` command, enter `cadmin -h` at the system prompt.

4. Use the `cadmin` command, in the following format, to set the correct DNS paths:

   `cadmin --set-domain-search-path path`

   For *path*, specify the search paths you need. The maximum number of paths is six. The maximum number of characters that you can specify in the *path* argument is 256.

   The following table shows example search paths.

Cluster type	Pathname
Clusters that include an InfiniBand data network or an Omni-Path data network	This path name typically starts with `ib0` followed by the cluster domain. For example, modify the search paths so that they appear as follows: `ib0.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`.
Clusters that include an HPE Slingshot data network	This path name typically starts with `hsn0` followed by the cluster domain. For example, modify the search paths so that they appear as follows: `hsn0.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`
Clusters that include an Ethernet data network	Modify the search paths to omit the `ib0` entries. For example, modify the search paths so that only the following appears: `datanet.cm.clusterdomain.com head.cm.clusterdomain.com cm.clusterdomain.com clusterdomain.com`

Updating the `/etc/hosts` file on the virtual machine (VM) admin node

Prerequisites

Specifying domain information

Procedure

1. As the root user, log into the VM admin node.
2. Enter the following commands to update the `/etc/hosts` file:

   - `# discover --node 501,generic,mgmt_net_name=head,hostname1=admin1, \`
     `mgmt_net_ip=172.23.255.101,mgmt_net_macs=14:02:ec:d9:78:08`
   - `# discover --node 502,generic,mgmt_net_name=head,hostname1=admin2,\`
     `mgmt_net_ip=172.23.255.102,mgmt_net_macs=14:02:ec:d9:7c:a0`

3. (Conditional) Update the `/etc/hosts` file for the third physical admin node:

   Complete this step if the admin node is a three-node quorum admin node. Enter the following command:

   `# discover --node 503,generic,mgmt_net_name=head,hostname1=admin3,\`
   `mgmt_net_ip=172.23.255.103,mgmt_net_macs=14:02:ec:df:ac:40`

4. Enter the following command to verify the `/etc/hosts` file:

   `# cat /etc/hosts | grep admin`

   On a SAC HA admin node, the preceding commands create the following entries in the `/etc/hosts` file:

   ```
   172.23.255.101  admin1.head.cm.milp.rdlabs.hpecorp.net admin1
   172.23.255.102  admin2.head.cm.milp.rdlabs.hpecorp.net admin2
   ```

   On a quorum HA admin node, the preceding commands creates the following entries in the `/etc/hosts` file:

   ```
   172.23.255.101  admin1.head.cm.milp.rdlabs.hpecorp.net admin1
   172.23.255.102  admin2.head.cm.milp.rdlabs.hpecorp.net admin2
   172.23.255.103  admin3.head.cm.milp.rdlabs.hpecorp.net admin3
   ```

Checking the status of the cluster nodes and setting the cluster password

Prerequisites

Updating the `/etc/hosts` file on the virtual machine (VM) admin node

Procedure

1. Log into the VM admin node as the root user.
2. Enter the following command, and examine the output:

   `# cm power status -t system`

   If some nodes are not up, enter the following command:

   `# cm power on -t system`

3. (Conditional) Start the scalable unit (SU) leader nodes.

   Complete the following steps if there are SU leader nodes on the cluster:

a. Enter the following command to power on the SU leader nodes:

```
cm power on -t 'leader_hostname*'
```

For *leader_hostname*, specify the portion of the SU leader node hostname that all SU leader nodes share.

For example:

```
# cm power on -t 'leader*'
```

b. For each SU leader node, `ssh` to the leader node, retrieve the status of SU leader nodes in the cluster database, and examine the status messages.

Use one of the following methods:

- Method 1 - Use the `clush` command. For example:

  ```
  # clush -g su-leader ctdb status
  ```

- Method 2 - Log into each SU leader node individually, retrieve the status of SU leader nodes in the cluster database, and examine the status messages.

  For example, the following command logs into one SU leader node:

  ```
  # ssh leader1 ctdb status
  Number of nodes:3
  pnn:0 172.23.0.2 OK (THIS NODE)
  pnn:1 172.23.0.3 OK
  pnn:2 172.23.0.4 OK
  Generation:370917201
  Size:3
  hash:0 lmaster:0
  hash:1 lmaster:1
  hash:2 lmaster:2
  Recovery mode:NORMAL (0)
  Recovery master:1
  ```

  If you use Method 2, make sure to enter the command on each SU leader node.

If the `ctdb` status shows anything other than `OK` or if the Gluster volumes do not look to be in the right state, log into the admin node and run the following command:

```
# su-leader-setup --bring-online
```

The **`--bring-online`** parameter ensures that the volumes and `ctdb` status are correct. When this command finishes, `ssh` to the SU leader node again, and run the verification steps again on that node.

NOTE: You can run the `su-leader-setup --bring-online` command at any time the volume or `ctdb` status is not correct. For example, you can run this command if an SU leader node goes down and the SU leader node does not register itself properly with `ctdb` after it comes back up.

---

c. From the admin node, enter the following command to mount the Gluster file system on the admin node:

```
# enable-su-leader --skip-sync
```

d. Power on the compute nodes.

For example:

```
# cm power on -t node '*'
```

4. Change the root password on the VM admin node and all other managed nodes.

   Enter the following command, and follow the prompts:

   ```
   # cpasswd
   ```

   For example:

   ```
   admin node:~ # cpasswd
   Enter new password:
   Enter new password (again):
   admin: updating /etc/shadow
   r1lead: updating /etc/shadow
   n0: updating /etc/shadow
   admin node:~ #
   ```

   If you need help, enter the following command:

   ```
   # cpasswd -h
   ```

   The `cpasswd` command changes the password on all nodes that are booted at the time the command is run.

   To change the root password on only one node, enter the `cpasswd` command in the following format:

   `cpasswd -N node_hostname`

   If you do not have the current password, you can obtain the factory-installed password from your technical support representative.

5. Use the `switchconfig` command to change the management switch password for the `admin` account.

   The format for this command is as follows:

   `switchconfig change_password --switches hostname --new new_password`

   The variables are as follows:

Variable	Specification
hostname	The hostname of the management switch or the keyword `all`. To specify a different password for each switch, enter a `switchconfig change_password` command for each switch. To specify the same password for each switch, use the keyword `all`.
new_password	A new password for the switch.

   For example:

   ```
   # switchconfig change_password --switches mgmtsw0 --new Hp3@dm!n2o20
   ```

   NOTE: Hewlett Packard Enterprise recommends that you implement standard and secure practices to store all passwords at your site. Do not lose this information.

6. Log into the first physical admin node as the root user, and use the `passwd` command to change the password.
7. Log into the second physical admin node as the root user, and use the `passwd` command to change the password.
8. (Conditional) Log into the third physical admin node as the root user, and use the `passwd` command to change the password.

   Complete this step if the admin node is a quorum HA admin node.

Completing the configuration

Prerequisites

One of the following procedures:

- **Attaching a cluster without a high availability (HA) admin node to your site network**

  Or

- **Attaching a cluster with a high availability (HA) admin node to your site network**

Procedure

1. **Setting the domain search path for admin nodes and leader nodes**
2. **Specifying additional network information**
3. **Backing up the admin node**
4. **Backing up the cluster configuration files**
5. **Configuring additional features**

Setting the domain search path for admin nodes and leader nodes

The domain name system (DNS) is a service. On a cluster, the DNS service converts a fully qualified domain name (FQDN) or hostname into an IP address.

A cluster has two or more subdomains, as follows:

- One subdomain is dedicated to the high-speed fabric.

  The following is an InfiniBand fabric subdomain example:

  `ib0.cm.clusterdomain.com`.

- On HPE SGI 8600 clusters, one subdomain is dedicated to the rack-specific management fabric. For example: `gbe.cm.clusterdomain.com`.

You can configure more than these two subdomains in the cluster.

When you run commands such as `ssh` *`hostname`* or `ping` *`hostname`* from the admin node, the domain search path determines the IP address used to complete the command.

By default, only the global domain search path is configured. If a per-node value is configured, it takes precedence over the global domain search path.

The following `nslookup` command example shows that the admin node uses the IP address in the `ib0` subdomain when attempting to reach ICE compute node `r1i0n1`:

```
admin:~ # nslookup r1i0n1
Server:           127.0.0.1
Address:        127.0.0.1#53

Name:   r1i0n1.ib0.cm.clusterdomain.com
Address: 10.148.0.2
```

The following procedure explains how to change the domain search path globally or on a per-node basis.

Prerequisites

One of the following procedures:

- **Attaching a cluster without a high availability (HA) admin node to your site network**

  Or

- **Attaching a cluster with a high availability (HA) admin node to your site network**

Procedure

1. Log into the admin node as the root user.

   If the cluster has a high availability (HA) admin node, log into the virtual machine (VM) admin node.

2. (Optional) Display the global domain search path value:

   ```
   admin:~ # cm node show -g --domain-search-path
   ib0.cm.clusterdomain.com,cm.clusterdomain.com,public.clusterdomain.com
   ```

3. To set a global domain search path, use the `cadmin` command in the following format:

   ```
   cadmin --set-domain-search-path subdomain_search_order
   ```

   For *subdomain_search_order*, specify one or more subdomains. If you specify more than one, use a comma to separate the subdomains.

4. To set the domain search path for a specific node, use the `cm node set` command in the following format:

   ```
   cm node set -n node --set-domain-search-path subdomain_search_order
   ```

   The variables are as follows:

Variable	Specification
node	The node hostname.
subdomain_search_order	One or more subdomains. If you specify more than one, use a comma to separate the subdomains.

   For example, to use the management fabric IP addresses for ICE compute nodes instead of the InfiniBand fabric on the admin node, enter the following:

   ```
   # cm node set -n admin --set-domain-search-path gbe.cm.clusterdomain.com,ib0.cm.clusterdomain.com,public.clusterdomain.com
   ```

5. Enter the following command to display the values you set in the previous steps:

   ```
   admin:~ # nslookup r1i0n1
   Server:           127.0.0.1
   Address:        127.0.0.1#53

   Name:   r1i0n1.gbe.cm.clusterdomain.com
   Address: 10.159.0.2
   ```

Specifying additional network information

The following topic explains how to use the cluster configuration tool to add information about your site network to the cluster database.

Prerequisites

Setting the domain search path for admin nodes and leader nodes

Procedure

1. Log into the admin node as the root user.

   If the cluster has a high availability (HA) admin node, log into the virtual machine (VM) admin node.

2. Back up the cluster definition file and the cluster database.

   Use the information in the following:

   Backing up the cluster configuration files

   Complete the following steps only and in the following order:

   - Step **2**.
   - Step **6**.
   - Step **7**.
   - Step **9**.
   - Step **8**. Return here, to this procedure, after you complete step **8**.

3. (Conditional) Start the cluster configuration tool.

   Complete this step if you do not have a window to the cluster configuration tool open at this time.

   Enter the following command to start the cluster configuration tool:

   ```
   # configure-cluster
   ```

   For general information about the cluster configuration tool, see the installation guide for your platform. The installation instructions are specific to each cluster platform. For links to the installation guides, see the following:

   Cluster manager documentation

4. On the cluster configuration tool main menu, select **D Configure Domain Name System (DNS)** and select **OK**.

5. On the **Domain Name System (DNS) Menu**, select **H Configure House DNS Resolvers**, and follow the prompts.

   You can specify up to three site (house) DNS resolvers.

6. Select **Quit** and select **OK** to log out from the cluster configuration tool.

7. Enter the following command to verify the name of the admin node:

   ```
   # cm node show -n admin
   ```

8. Use the `cm node set` command, in the following format, to customize the admin node name:

   `cm node set -n admin` ***`admin_node_hostname`***

   For *admin_node_name*, specify the hostname you want to use for the admin node.

9. Enter the following command to display the node images on the admin node at this time:

   ```
   # cm image show -d
   ```

10. Use the `cp` command to set the time zone in the system images.

    The format of this command is as follows:

    ```
    cp /etc/localtime /opt/clmgr/image/images/image_name/etc
    ```

For *image_name*, type the name of one of the system images you retrieved with the `cm image` command. Enter one `cp` command for each system image. For example:

- On a cluster with leader nodes, enter `cp` commands as follows:
  - One for each type of leader node present.

    That is, one for the scalable unit (SU) leader nodes and/or one for the ICE leader nodes.
  - One for each type of compute node present.

    That is, one for the ICE compute nodes and/or one for the compute nodes.
- On a cluster without leader nodes, you need one `cp` command for the compute nodes.

Example 1. On a cluster with ICE leader nodes and RHEL 8.4, enter the following commands, one for each master image:

```
# cp /etc/localtime /opt/clmgr/image/images/ice-rhel8.4/etc
# cp /etc/localtime /opt/clmgr/image/images/rhel8.4/etc
# cp /etc/localtime /opt/clmgr/image/images/lead-rhel8.4/etc
```

Example 2. On a cluster without leader nodes and SLES 15 SP3, enter the following command to copy the compute node master image:

```
# cp /etc/localtime /opt/clmgr/image/images/sles15sp3/etc
```

11. Update the system images with your time zone information.

Use the following command:

```
# cp /etc/localtime /opt/clmgr/image/images/image_name/etc/localtime
```

For *image_name*, specify the image name.

For example, the following commands change all three images on a cluster with leader nodes and RHEL 8.4:

```
# cp /etc/localtime \
/opt/clmgr/image/images/lead-rhel8.4/etc/localtime
cp: overwrite '/opt/clmgr/image/images/lead-rhel8.4/etc/localtime'? y
# cp /etc/localtime /opt/clmgr/image/images/rhel8.4/etc/localtime
cp: overwrite '/opt/clmgr/image/images/rhel8.4/etc/localtime'? y
# cp /etc/localtime /opt/clmgr/image/images/ice-rhel8.4/etc/localtime
cp: overwrite '/opt/clmgr/image/images/ice-rhel8.4/etc/localtime'? y
```

Backing up the admin node

Use a backup program at your site to back up the admin node. Completing this procedure now, before you put your cluster into production. In this way, you ensure that you have a copy of the admin node that you can use in case a disaster occurs. Make sure to write the backup copies to a safe location on a computer that resides outside the cluster. An admin node backup protects the following:

- The cluster database
- The cluster definition file
- The node images
- The VCS source control system

NOTE: Make sure to back up the admin node regularly. Backing up the admin node protects your cluster configuration if a disk failure or other disaster occurs.

The following procedure explains how to back up the admin node.

Prerequisites

Specifying additional network information

Procedure

1. Use your site practices and site backup program to back up the cluster admin node.

   To restore the admin node, use the restore procedure for your site backup program.
2. (Optional) Use your site file restore practices to test a restore of the admin node.

Backing up the cluster configuration files

Complete this procedure now and at any other time you significantly modify the cluster. For example, repeat this procedure in the following situations:

- Changing cluster attributes.
- Adding nodes to the cluster.
- Deleting nodes from the cluster.
- Changing the software image on a node.
- Changing the kernel on a node.
- Changing the hostname of a node.
- Changing the IP address of a node.

If you have more than one slot, remember that backing up a slot by cloning the slot is not equal to backing up the admin node. Disk failures can occur.

The following procedure explains how to back up the cluster definition file and the cluster database.

Prerequisites

Procedure

1. Log into the admin node as the root user.

   If the cluster has a high availability (HA) admin node, log into the virtual machine (VM) admin node.
2. Enter the following command to back up the cluster definition file:

   ```
   # discover --show-configfile --images --kernel --bmc-info \
   --kernel-parameters --ips > filename
   ```

   For *filename*, specify a file name.

For example:

```
# discover --show-configfile --images --kernel --bmc-info \
--kernel-parameters --ips > my.config.file
```

NOTE: Hewlett Packard Enterprise recommends that you keep a copy of the cluster definition file on another computer system at your site.

You need the cluster definition file in case you have to reconfigure one or more nodes. This file can be useful when troubleshooting. You also need the cluster definition file for disaster recovery. You can supply the cluster definition file as input to the `cm node add` command, the `cm node discover add` command, the `discover` command, and the `configure-cluster` command. The cluster definition file supplies the information that you would typically define by using the menus in the cluster configuration tool. When you specify a cluster configuration file as input to these commands, the commands read in the options from the file and implement them in the cluster.

Save a new copy of the cluster definition file anytime you modify your cluster. Without a cluster definition file, to reconfigure any aspect of your cluster, you have to power on and power off each component during the configuration process. To restore the cluster definition file, copy the file from its backup location to the admin node.

3. Copy the cluster definition file to another server at your site.
4. (Conditional) Back up the custom partitioning file.

   Complete this step if you configured custom partitioning.

   The custom partitioning file resides in the following location:

   `/opt/clmgr/image/scripts/pre-install/custom_partitions.cfg`
5. (Conditional) Back up the scalable unit (SU) leader node configuration files.

   Complete this step if the cluster includes SU leader nodes.

   Back up the following files:

   - `/opt/clmgr/etc/su-leader-setup.conf`
   - `/opt/clmgr/etc/su-leader-nodes.lst`
6. Enter the following commands to stop the cluster manager:

   ```
   # systemctl stop config_manager.service
   # systemctl stop clmgr-power.service
   # systemctl stop cmdb.service
   ```
7. Enter the following command to back up the cluster database:

   ```
   # sqlite3 /opt/clmgr/database/db/cmu.sqlite3 ".backup file"
   ```

   For *file*, specify a name for the backup file. The cluster manager writes the backup file to the current working directory.

   For example:

   ```
   # sqlite3 /opt/clmgr/database/db/cmu.sqlite3 ".backup cmu.backup.sqlite3"
   ```
8. Enter the following commands to start the cluster manager:

   ```
   # systemctl start cmdb.service
   # systemctl start clmgr-power.service
   # systemctl start config_manager.service
   ```

NOTE: In the future, to restore the cluster database and start the cluster manager, see the procedure in the following:

HPE Performance Cluster Administration Guide

9. Copy the database backup file to another server at your site.

   The cluster database is the internal database that hosts information about each cluster component. A copy of the original cluster database can be valuable when performing a disaster recovery. Make sure to take additional, periodic database backups in the future as you modify your system.

10. Enter the following command to save the changed configuration to the nonvolatile memory (NVM) on the switches:

    ```
    # switchconfig config -s all --save
    ```

11. Enter the following command to back up all the switch configuration information:

    ```
    # switchconfig config -s all --pull
    configuration file 'startup-config' copied from mgmtsw0 to 172.23.0.1
    at /opt/clmgr/tftpboot/mgmtsw_config_files/mgmtsw0/startup-config
    configuration file 'startup-config' copied from mgmtsw1 to 172.23.0.1
    at /opt/clmgr/tftpboot/mgmtsw_config_files/mgmtsw1/startup-config
    configuration file 'startup.cfg' copied from mgmtsw2 to 172.23.0.1
    at /opt/clmgr/tftpboot/mgmtsw_config_files/mgmtsw2/startup.cfg
    configuration file 'primary.cfg' copied from mgmtsw3 to 172.23.0.1
    at /opt/clmgr/tftpboot/mgmtsw_config_files/mgmtsw3/primary.cfg
    ```

    Observe the message that this command issues upon completion. This message contains the location of the backup files. By default, the message points to the files in the following directory on the admin node:

    ```
    /opt/clmgr/tftpboot/mgmtsw_config_files
    ```

12. Note the file name or names from the preceding command output. Copy each backup file from the admin node to a safe storage space at your site.

Configuring additional features

The following procedure explains where you can obtain information about how to configure additional features.

Prerequisites

Backing up the cluster configuration files

Procedure

1. Configure one or more of the following recommended features:

   - The GUI. You can configure the GUI on a client computer outside of the cluster system. The client computer is required to have Java 8+ installed. For example, you can install the client software on a laptop computer. Complete the following steps:
     - Open a browser, and enter one of the following addresses for the admin node:
       - The IP address

or

- The fully qualified domain name (FQDN)

- Follow the instructions on the cluster manager splash page to download and install the GUI client.

- Firewalls. The cluster manager implements security features over the entire cluster.

  If you have an HA admin node, realize that the physical admin nodes have the firewall turned on by default to help lock down the public/site network. The following services are open in the firewall:

  - VNC server
  - High availability
  - The `ssh` command.

  The point to point (PTP) network for high availability, usually configured on `eth3`, is in the Trusted zone.

  For information about how to create or configure firewalls, see your operating system documentation.

- Functional node hostnames. At the factory, HPE configured default hostnames on each node. For example, on a cluster with leader nodes, `n0` is a typical default hostname for a compute node upon which user services are installed. You can see these default names in the cluster configuration file. You can set a site-specific hostname for these nodes. If you set a site-specific hostname, most cluster command output displays the hostname you set, not the default hostname. To retrieve a list of all node names, enter the following command:

  ```
  # cm node show
  ```

  If you plan to install additional functional services on a node, Hewlett Packard Enterprise recommends that you rename the node and give the node a unique hostname. For example, you might want to install login services on a node and change the hostname to `login_node`.

  You can use the `cm node set` command to give meaningful hostnames to compute nodes that you want to install with login services, storage services, and so on.

  If you plan to install Array Services, also plan to configure functional hostnames before you configure Array Services. If you rename the nodes after you configure Array Services, you will need to configure Array Services again.

  NOTE: If you change any hostnames or any other aspect of system configuration, complete the following procedures:

  - **Backing up the admin node**
  - **Backing up the cluster configuration files**

  For information about the cluster manager security features, the `cm node` command, and functional hostnames, see the following:

  HPE Performance Cluster Administration Guide

2. Network groups of compute nodes.

   If you want to use native monitoring, configure the compute nodes into network groups. It is most common to configure all the compute nodes under a common switch into a network group.

   For information about how to configure network groups, see the following:

   HPE Performance Cluster Administration Guide

3. Configure one or more of the following additional features, some of might be required for your installation:

- Networking features, such as network address translation (NAT).
- Array Services. The Array Services software product is required for HPE Message Passing Interface (MPI) programming.

For information about additional features, see the following:

- The installation guide for your platform. For links to the installation guides, see the following:

  Cluster manager documentation
- **HPE Performance Cluster Administration Guide**
- **HPE Performance Cluster Manager Power Management Guide**

HPE Performance Cluster Manager command reference

The following information summarized the cluster manager commands.

For more information on any command, enter the command name followed by `-h` at the system prompt, as follows:

```
# commandname -h
```

Power commands

General format for power commands

```
cm power action -t target hostname
```

Obtain power status for a node

```
cm power status -t node node_hostname
```

Power on the cluster

```
cm power on -t system
```

Power off the cluster

```
cm power off -t system
```

Identify node

```
cm power identify -t node node_hostname -i 60
```

Network interface card (NIC) operations

Create a NIC and add it to a node on a network

```
cm node nic add
```

Delete a NIC from a node

```
cm node nic delete
```

Set NIC properties

```
cm node nic set
```

Display NICs

```
cm node nic show
```

Convey custom configuration information for nodes in an HPE Cray EX chassis

Specify a template file

```
cm node template
```

Modify the cluster network

The `cm network` commands modify or add networks without entering the `configure-cluster` command and going through the menu system.

Add a network

```
cm network add
```

Delete a network

```
cm network delete
```

Display network properties

```
cm network show
```

Set network properties

```
cm network set
```

Clear one or more network attributes

```
cm network unset
```

Display node hostnames

Display hostnames for all nodes

```
cm node show
```

Display leader node hostnames

```
cm node show -t system leader
```

Display compute node hostnames

```
cm node show -t system compute
```

Display ICE compute node hostnames

```
cm node show -t system ice_compute
```

Display chassis manager controller (CMC) hostnames

```
cm node show -t system cmc
```

Display compute nodes that are offline

```
cm node show --offline
```

Multinode queries

For a multinode query, use one of the following commands:

- `clush`

  This command displays line-by-line output. As a node responds to the query, it appears in the output.

  Or

- `cm node run`

  This command displays consolidated output in a condensed visual format.

Query compute nodes

```
clush -g compute query
```

Or

```
cm node run -t system compute query
```

Query ICE compute nodes

```
clush -g ice-compute query
```

Or

```
cm node run -t system ice-compute query
```

Query leader nodes

```
clush -g leader query
```

Or

```
cm node run -t role su-leader query
```

Display the number of compute nodes that currently report a certain speed

```
cm node run -t system compute 'ethtool device_name | grep -i speed'
```

For example:

```
cm node run -t system compute 'ethtool eno1 | grep -i speed'
```

Display compute nodes with unexpected memory size

```
cm node run -t system compute 'grep MemTotal /proc/meminfo'
```

Display compute nodes with unexpected load averages

```
cm node run -t system compute 'uptime'
```

Display the boot file systems that are mounted

```
cm node run -t system compute 'df -h | grep -i boot'
```

Display cluster information

Display monitoring data

```
cm  monitoring
```

Display system health

```
cm health check
```

Display alerts

```
cm health alert
```

Display system infrastructure monitoring information

```
cm sim
```

Change slot

Copy one slot to a new slot

```
cm node slot
```

Repository management commands

Display all repositories

```
cm repo show
```

Display selected repositories

```
cm repo show | grep ^\*
```

Display unselected repositories

```
cm repo show | grep -v ^\*
```

Make repository available to provisioning commands

```
cm repo select repository_name
```

Make repository unavailable to provisioning commands

```
cm repo unselect repository_name
```

Refresh custom repository metadata

```
cm repo refresh repository_name
```

Add custom repository

```
cm repo add /opt/clmgr/repos/mypkgs --custom mypkgs
```

Configure the cluster

Add nodes using a cluster definition file

```
cm node add -c cluster_definition_file
```

For more information, enter the following:

```
# man cluster-configfile
```

Add nodes by detecting the MAC address and other info and adding the nodes into the cluster

```
cm node discover add
```

Display path to cluster definition file

```
discover --show-configfile
```

Manage images and provision

Copy files into an image

```
cp /etc/myfile /opt/clmgr/image/images/image_name/etc
```

Display managed images and image kernel versions

```
cm image show -I
```

Display image and kernel to be installed at next image operation

```
cm node show -I
```

From a flat compute node, display the image that was used on boot

```
cat /proc/cmdline
```

From the admin node, display the image that was used to boot all compute nodes

```
clush -g compute cat /proc/cmdline
```

Assign an image to a node for provisioning the next time the node boots

```
cm node provision -n hostname -i image_name -s
```

Assign an image to a node, reboot, and provision with an operating system

```
cm node provision -n hostname -i image_name
```

Specify that a node boot from its Ethernet device the next time it boots

```
ipmiwrapper hostname chassis bootdev pxe
```

Create an image from selected repositories and an RPM list

```
cm image create -i new_image -l /opt/clmgr/image/rpmlists/my-rhel8.X.rpmlist
```

Create a new image from an existing image

```
cm image copy -o existing_image -i new_image
```

Capture an image from a running compute node

```
cm image capture --image image_name --node number
```

Capture an image from a running compute node and exclude some structures

```
cm image capture --image image_name --node hostname --exclude /var/
spool/BBS/\\*_logs/\\* --exclude /var/spool/PBS/server_priv/accounting/\\*
```

Installing RPMs into an image

RHEL 8:

```
cm image dnf --duk -i image_name install RPM RPM ...
```

RHEL 7:

```
cm image yum --duk -i image_name install RPM RPM ...
```

SLES:

```
cm image zypper --duk -i image_name install RPM RPM ...
```

Removing RPMs from an image

RHEL 8:

```
cm image dnf --duk -i image_name remove RPM RPM ...
```

RHEL 7:

```
cm image yum --duk -i image_name remove RPM RPM ...
```

SLES:

```
cm image zypper --duk -i image_name remove RPM RPM ...
```

Installing RLMs into an image on a running compute node

RHEL 8:

```
cm node dnf -n hostname install RPM RPM ...
```

RHEL 7:

```
cm node yum -n hostname install RPM RPM ...
```

SLES:

```
cm node zypper -n hostname install RPM RPM ...
```

Removing RPMs from an image on a running compute node

RHEL 8:

```
cm image dnf --duk -n hostname remove RPM RPM ...
```

RHEL 7:

```
cm image yum --duk -n hostname remove RPM RPM ...
```

SLES:

```
cm image zypper --duk -n hostname remove RPM RPM ...
```

Commands for ICE compute nodes

Pushing ICE compute node images to a rack

```
cimage --push-rack image_name rack_number
```

For *rack_number*, specify a rack number in the `rX` format.

Display ICE compute node images and kernels

```
cimage --show-images
```

Display image, kernel, and file system mode of ICE compute nodes

```
cimage --show-nodes rack_number
```

For *rack_number*, specify a rack number in the `rX` format.

Change the image, kernel, or file system mode of ICE compute nodes

```
cimage --set [option] image_name kernel node
```

Update existing RPMs in an image

Update existing RPMs on a node

```
cm node update rpms -n hostname
```

Update existing RPMs in an image in the admin node repository

```
cm image update -i image_name
```

VCS commands

Check in an image with changes or updates into version control

```
cm image revision commit -i image_name -m 'checking in'
```

Review version control history for an image

```
cm image revision history -i image_name
```

Revert to a previous image version

```
cm image revision revert -i image_name  --rev version_number
```

Directories

Review post-installation scripts

```
ls /opt/clmgr/image/scripts/post-install
```

Configuration scripts directory

```
/opt/clmgr/image/images/image_name/etc/opt/sgi/conf.d
```

Set up one-time PXE boot for HPE ProLiant DL360 Gen10, HPE ProLiant DL380 Gen10, and HPE ProLiant DL380 Gen9 server installation

Use one of the following methods to configure a one-time PXE boot on nodes with iLO node controllers:

- From the admin node, use the following `ilorest` command:

  ```
  ilorest bootorder --onetimeboot=pxe --url=iLO_IP -u ADMIN -p iLO_password --commit
  ```

  For *iLO_IP*, *ADMIN*, and *iLO_password*, enter the values for this cluster.

- Use `ssh` to log into a node, and then run one of the following commands:

  ```
  hpbootcfg -P -b
  ```

  Or

  ```
  efibootmgr options ...
  ```

- On the node console, press F12 or Esc+@ (serial).

You can download `ilorest` and `hpbootcfg` from the HPE support website.

Websites

General websites

Single Point of Connectivity Knowledge (SPOCK) Storage compatibility matrix

https://www.hpe.com/storage/spock

Storage white papers and analyst reports

https://www.hpe.com/storage/whitepapers

For additional websites, see **Support and other resources**.

Support and other resources

Accessing Hewlett Packard Enterprise Support

- For live assistance, go to the Contact Hewlett Packard Enterprise Worldwide website:

  https://www.hpe.com/info/assistance

- To access documentation and support services, go to the Hewlett Packard Enterprise Support Center website:

  https://www.hpe.com/support/hpesc

Information to collect

- Technical support registration number (if applicable)
- Product name, model or version, and serial number
- Operating system name and version
- Firmware version
- Error messages
- Product-specific reports and logs
- Add-on products or components
- Third-party products or components

Accessing updates

- Some software products provide a mechanism for accessing software updates through the product interface. Review your product documentation to identify the recommended software update method.
- To download product updates:

  Hewlett Packard Enterprise Support Center

  https://www.hpe.com/support/hpesc

  Hewlett Packard Enterprise Support Center: Software downloads

  https://www.hpe.com/support/downloads

  My HPE Software Center

  https://www.hpe.com/software/hpesoftwarecenter

- To subscribe to eNewsletters and alerts:

  https://www.hpe.com/support/e-updates

- To view and update your entitlements, and to link your contracts and warranties with your profile, go to the Hewlett Packard Enterprise Support Center **More Information on Access to Support Materials** page:

  https://www.hpe.com/support/AccessToSupportMaterials

IMPORTANT: Access to some updates might require product entitlement when accessed through the Hewlett Packard Enterprise Support Center. You must have an HPE Passport set up with relevant entitlements.

Remote support

Remote support is available with supported devices as part of your warranty or contractual support agreement. It provides intelligent event diagnosis, and automatic, secure submission of hardware event notifications to Hewlett Packard Enterprise, which initiates a fast and accurate resolution based on the service level of your product. Hewlett Packard Enterprise strongly recommends that you register your device for remote support.

If your product includes additional remote support details, use search to locate that information.

HPE Get Connected

https://www.hpe.com/services/getconnected

HPE Pointnext Tech Care

https://www.hpe.com/services/techcare

HPE Complete Care

https://www.hpe.com/services/completecare

Warranty information

To view the warranty information for your product, see the links provided below:

HPE ProLiant and IA-32 Servers and Options

https://www.hpe.com/support/ProLiantServers-Warranties

HPE Enterprise and Cloudline Servers

https://www.hpe.com/support/EnterpriseServers-Warranties

HPE Storage Products

https://www.hpe.com/support/Storage-Warranties

HPE Networking Products

https://www.hpe.com/support/Networking-Warranties

Regulatory information

To view the regulatory information for your product, view the *Safety and Compliance Information for Server, Storage, Power, Networking, and Rack Products*, available at the Hewlett Packard Enterprise Support Center:

https://www.hpe.com/support/Safety-Compliance-EnterpriseProducts

Additional regulatory information

Hewlett Packard Enterprise is committed to providing our customers with information about the chemical substances in our products as needed to comply with legal requirements such as REACH (Regulation EC No 1907/2006 of the European Parliament and the Council). A chemical information report for this product can be found at:

https://www.hpe.com/info/reach

For Hewlett Packard Enterprise product environmental and safety information and compliance data, including RoHS and REACH, see:

https://www.hpe.com/info/ecodata

For Hewlett Packard Enterprise environmental information, including company programs, product recycling, and energy efficiency, see:

https://www.hpe.com/info/environment

Documentation feedback

Hewlett Packard Enterprise is committed to providing documentation that meets your needs. To help us improve the documentation, use the **Feedback** button and icons (located at the bottom of an opened document) on the Hewlett Packard Enterprise Support Center portal (**https://www.hpe.com/support/hpesc**) to send any errors, suggestions, or comments. All document information is captured by the process.

YaST navigation

The following table shows SLES YaST navigation key sequences.

Key	Action
Tab **Alt** + **Tab** **Esc** + **Tab** **Shift** + **Tab**	Moves you from label to label or from list to list.
Ctrl + **L**	Refreshes the screen.
Enter	Starts a module from a selected category, runs an action, or activates a menu item.
Up arrow	Changes the category. Selects the next category up.
Down arrow	Changes the category. Selects the next category down.
Right arrow	Starts a module from the selected category.
Shift + **right arrow** **Ctrl** + **A**	Scrolls horizontally to the right. Useful in screens if use of the **left arrow** key would otherwise change the active pane or current selection list.
Alt + *letter* **Esc** + *letter*	Selects the label or action that begins with the *letter* you select. Labels and selected fields in the display contain a highlighted *letter*.
Exit	Quits the YaST interface.

Glossary

admin node

The administrative node on an HPE cluster computer system. This is the node you log into when you want to install software or manage the cluster. You also log into this node to install the cluster manager software.

The keyword `admin` appears on many command lines as a shortened form of the term `admin node`. For example, the following command retrieves the hostname of the admin node:

```
# cm node show -n admin
```

In a cluster without leader nodes, the admin node provisions all the compute nodes.

In a cluster with leader nodes, the admin node provisions the leader nodes. The leader nodes provision the nodes that they control.

See **infrastructure node**.

batch node

A compute node that has batch or scheduler services configured upon it.

See **compute node**.

blade enclosure

See **individual rack unit**.

BMC

Baseboard management controller. A type of node controller.

See **node controller**.

cabinet

A hardware structure on an HPE Cray EX supercomputer. A cabinet includes the following types of controllers:

- Chassis environment controller (CEC).

  This controller is also referred to as a cabinet environment controller or an eC. Controls the chassis, the CEC, and the cooling distribution unit (CDU).
- Chassis controller (cC). This controller is also referred to as the chassis management module (CMM).
- Node controller (nC).
- Switch controller (sC).

CDU

Cooling distribution unit.

CEC

Chassis environment controller. A controller found on HPE Cray EX supercomputer.

chassis

A **chassis** contains cluster compute nodes.

On an HPE SGI 8600 cluster, the term **individual rack unit (IRU)** is deprecated. An IRU is now known as a **chassis**.

For information about chassis on various clusters, see the following:

`cm` command information

chassis controller

Depending on the cluster hardware, a chassis controller can be either a chassis management module (CMM) or chassis management controller (CMC). These devices are as follows:

- On HPE Cray EX supercomputers, the CMM controls the chassis, switch controllers, and a chassis environment controller.
- On HPE Apollo clusters and HPE SGI 8600 clusters, CMCs provide the entry point to the following:
  - On HPE Apollo 9000 clusters, the CMC provides access to the chassis. On HPE Apollo 9000 systems, the CMC includes features from the HPE Apollo Platform Manager. For more information, see the following:

    HPE Apollo Platform Manager User Guide
  - On HPE SGI 8600 clusters, the CMC provides access to the system control network for one chassis.

  The CMC provides power control, environment monitoring, and console access to compute nodes and ICE compute nodes. Each ICE compute chassis includes one or two CMCs. The CMC also provides blade inventory functions for compute blades and InfiniBand switch blades.

  The CMC supports the intelligent platform management interface (IPMI).

chassis management controller (CMC)

See **chassis controller** and **CMC**.

chassis management module (CMM)

See **chassis controller** and **CMM**.

`cimage`

The `cimage` command manages ICE compute node images.

cluster

A set of computers interconnected with both management and data networks.

A cluster refers to all the physical elements of the cluster computing solution. The term includes the following components:

- Cluster manager master host
- Compute hosts
- Cluster manager
- UPS
- High-speed network
- Storage
- Cabinet

Clustering is a method of linking multiple computers or compute hosts together to form a unified and more powerful system. Cluster systems divide computations among all the processors in the cluster. They gather the data after the computations are complete.

cluster definition file

The master configuration file for the cluster. Defines switches, boot parameters, nodes, and many other aspects of the cluster.

cluster domain

The cluster domain name. The HPE factory-default fully qualified domain name is `cm.americas.sgi.com`. The cluster domain in this example is `cm`. The corporate subdomain is `americas`. The corporate domain is `sgi.com`.

Hewlett Packard Enterprise recommends that you change the default to a site-specific value. This value is used in the initial cluster configuration procedure and in the cluster definition file.

The cluster domain is a subdomain in your site network. The cluster domain name distinguishes the cluster domain from the site domain.

clusters with leader nodes

A cluster that includes either scalable unit (SU) leader nodes or ICE leader nodes. SU leader nodes and ICE leader nodes manage communication between the admin node and the compute nodes under their control.

On a cluster with SU leader nodes, the SU leader nodes manage many compute nodes. These nodes can be deployed exclusively for use as compute nodes, or they can be deployed with user services. HPE Apollo clusters can have SU leader nodes.

On a cluster with ICE leader nodes, the ICE leader nodes manage ICE compute nodes. ICE clusters can also contain compute nodes that are deployed with user services. These compute nodes are under the control of the admin node, not an ICE leader node. By default, ICE clusters include at least one compute node designated as a login node. HPE SGI 8600 clusters are an example of a cluster with ICE leader nodes.

clusters without leader nodes

A cluster that consists of one admin node and several compute nodes. Among the compute nodes, there is no node hierarchy.

CMC

A chassis management controller.

HPE Apollo 9000 clusters and HPE SGI 8600 clusters with leader nodes have CMCs.

HPE Apollo 9000 clusters and HPE SGI 8600 clusters that do not include leader nodes do not have CMCs.

Typically, the documentation uses the term **chassis controller** to refer to a CMC.

Also see **chassis controller**.

CMM

A chassis management module.

HPE Cray EX supercomputers have CMMs. CMMs control the following:

- The cabinet environment controllers (CECs). There can be up to four.
- The in-rack switch controller.
- The controller for each node.
- The chassis. There can be up to eight.

Typically, the documentation uses the term **chassis controller** to refer to a CMM.

Also see **chassis controller**.

compute node

A dedicated computation node that is not under the control of an ICE leader node. In HPE Performance Cluster Manager releases before the 1.5 release, these nodes were referred to as **non-ICE compute nodes**.

You can designate a compute node as a host for a specific service. This designation transforms the node into one of the following types of dedicated nodes:

- Batch node
- Gateway node
- Login node
- Storage node

The admin node and the compute nodes attach to the management network directly.

An ICE leader node manages the ICE compute nodes. Also see **ICE compute node**.

configfile

A shortened term for **cluster definition file**.

See **cluster definition file**.

controller

Various hardware platforms have controllers.

In the Ganglia and Nagios user interfaces, a controller is a construct used to group admin nodes and rack leader controller nodes.

NOTE: Cluster manager support for Ganglia and Nagios will be removed in a future release.

CRC

Cooling rack controller.

Applicable on clusters with ICE leader nodes.

custom group

A group of nodes assigned to a specific user.

domain

See **cluster domain**.

domain name system (DNS)

A hierarchical naming system for all nodes contained within the cluster. The actual service converts a domain name into a numerical IP address. Several nodes provide DNS services to other nodes. For example, ICE compute nodes use their respective leader node as the primary DNS server for name resolution. A compute node uses the admin node as the primary DNS server for name resolution.

E-cell

HPE supports E-cells on HPE SGI 8600 clusters and SGI ICE XA clusters. E-cells provide power and cooling efficiency.

environment controller

See **CEC**.

fabric management node (FMN)

A node associated with the HPE Slingshot fabric manager. Typically, a cluster has two FMNs. These nodes appear in the `/etc/hosts` file.

gateway, gateway node

A compute node with corporate or site gateway services configured upon the node.

GPU

Graphics processing unit.

GUI client workstation

You can access the cluster manager GUI from any client workstation that can reach the admin node over the network. Several users on different workstations can use the GUI simultaneously. The GUI client workstation can run Linux or Windows.

head network

The **head network** is also known as the **cluster management network**. The head network has the following characteristics:

- It facilitates monitoring, provisioning, and other functions not covered by the data network.
- It is an Ethernet network. It is designed for communication between the admin node and the compute nodes.

On clusters with leader nodes, the following reside on the head network:

- The admin node
- The leader nodes
- The compute nodes
- Managed switches
- Chassis management controllers

On clusters without leader nodes, the following reside on the head network:

- The admin node
- The compute nodes
- Managed switches
- Chassis management controllers

high availability (HA)

HPE offers cluster systems with admin nodes or ICE leader nodes designed for high availability.

In a cluster with an HA admin node, there are two physical admin nodes attached to a storage unit. One HA admin node always hosts the administrative admin node. The administrative admin node is deployed as a virtual machine.

ICE leader nodes can also be configured as HA ICE leader nodes.

See the following for information about high availability in clusters with scalable unit (SU) leader nodes:

Clusters with scalable unit (SU) leader nodes

house network

See **site network**.

ICE leader node

ICE is an acronym for **integrated compute environment**. The ICE acronym can refer to an ICE leader node and to the ICE compute nodes under the control of the ICE leader node.

HPE SGI 8600 clusters, for example, include ICE leader nodes and ICE compute nodes. HPE SGI 8600 clusters can also include compute nodes that are not under the control of an ICE leader node; typically, these compute nodes are deployed with user services.

The ICE leader nodes provision the ICE compute nodes under their control. In addition, the ICE leader nodes enable scaling and facilitate communication.

This term is not applicable to clusters without ICE leader nodes or to clusters with scalable unit (SU) leader nodes.

ICE compute node

A type of compute node on a cluster with ICE leader nodes. For example, an HPE SGI 8600 cluster includes one or more ICE leader nodes. The admin node manages the ICE leader nodes, and the ICE leader nodes manage the ICE compute nodes. The admin node also manages any compute nodes that are configured with user services.

iLO

Integrated Lights-Out. A self-contained hardware technology that enables remote management of any node within a system. This autonomous management processor is built into the server and is a type of node controller. It performs or facilitates the following functions:

- Server setup
- Server health monitoring
- Power and thermal optimization
- Remote server administration

This wide latitude of control is available independently of the operating system and independently of the server hardware.

See **node controller**.

iLO chassis managers (iLOCMs)

The HPE Moonshot iLO Chassis Management Firmware is the gateway for aggregated chassis management on the HPE Moonshot System. As a single point of access to the chassis, Moonshot iLO CM firmware allows you to configure, update, and operate Moonshot through CLI, IPMI, and remote serial console access.

image

The set of files on a node that forms the operating system, its configuration, and the applications distributed with the operating system. When you provision a node, you install an image on the node.

image group

A group of similar machines running the same image. A node can belong to several image groups.

imaging

In the cluster manager documentation, **imaging** and **provisioning** are synonyms for the same activity.

See **provisioning**.

individual rack unit (IRU)

A term used when referring to clusters with ICE leader nodes, such as HPE SGI 8600 clusters. The term **IRU** is deprecated. The preferred term is **chassis**.

See **chassis**.

infrastructure node

A node that provides support for the cluster compute nodes. For example, the admin node is an infrastructure node, and leader nodes are infrastructure nodes.

See **admin node**.

See **leader node**.

internal name

The internal name for a node. If you do not supply a custom name, the cluster manager uses the internal name as the hostname for a node. All tools let you specify a custom host name for a node, and the hostname becomes the primary name of the node. In the cluster definition file, this name appears in the `internal_name=` field.

Some tools include the internal name of the node in output even if you specify a custom hostname.

iPXE

The open source implementation of the preboot execution environment. This standard client/server interface enables uninstalled network computers with an operating system to be configured and booted remotely. iPXE booting is configured at the BIOS level.

IRU

See **chassis**.

This term is deprecated.

kernel

When **kernel** appears as a command parameter, specify a specific operating system kernel version from the available kernel versions installed in an image.

leader node

A type of infrastructure node that manages other nodes. The admin node and all the leader nodes attach to the management network directly.

The following are the two types of leader nodes:

- ICE leader nodes. In a cluster with ICE leader nodes, the ICE compute nodes attach to the ICE leader node, not to the management network.
- Scalable unit (SU) leader nodes. SU leader nodes are used in multiples of three and constitute a leader node pool. Compute nodes reside under the control of the SU leader nodes.

Sometimes shortened to **leader**.

leader node pool

A group of scalable unit (SU) leader nodes. The number of nodes in the pool must be a multiple of 3. For example, you can have 3, 6, 9, or more individual nodes in the leader node pool.

login node

A compute node with login services configured upon the node.

By default, all clusters with leader nodes include at least one compute node that is configured as a login node.

M-cell

The cluster manager supports M-cells on clusters with leader nodes that include dedicated cooling units. M-cells provide power and cooling efficiency.

This term is not applicable on clusters without leader nodes.

network group

A group of machines. Typically, these machines are connected to the same Ethernet switch. For example, this group is a standalone rack or standalone enclosure. Each node belongs to a unique network group. In each network group, one of the nodes is temporarily dedicated as the secondary server, which means that it has the same role as the image server.

For example, the cluster itself is a network group that includes the admin node and all other nodes in the cluster. On clusters with ICE leader nodes, the leader node and all the ICE compute nodes behind the leader node are a network group. You can create your own customer network groups with selected groups of node.

node

An individual computer in a cluster system.

node controller

A hardware device on a server that enhances remote manageability. For example:

- The iLO node controllers.
- The baseboard management controller (BMC) on a compute node. This component is typically a type of node controller.

node discovery commands

One of the commands that you can use to configure a component into the cluster database. Commands include `cm node add`, `cm node discover add`, and `discover`.

non-ICE compute node

Deprecated term. The cluster manager now uses the term **compute node** to refer to compute nodes that are not under the control of an ICE leader node.

per-host customization script

On a cluster with ICE leader nodes, this script forms the infrastructure for the ICE compute nodes.

provisioning

The process of installing a software image on a node. In this operation, the next time the node boots, the node installs its operating system.

In the documentation, **imaging** and **provisioning** are synonyms for the same activity.

rack leader

See **ICE leader node**.

rack leader controller (RLC)

An alternative term for **ICE leader node**.

See **ICE leader node**.

RLC

See **ICE leader node**.

scalable unit (SU) leader node

A leader node on an HPE Apollo cluster that works in concert with other SU leader nodes to form a leader node pool. A leader node pool improves the data resiliency in the cluster.

Assign SU nodes to groups of three. The number of SU leader nodes in the cluster must be a multiple of 3.

service node

A compute node with one or more services installed upon it. For example, these services can be batch scheduling services, login services, or other services.

You might see this term used in product messages or other cluster system information. Typically, the cluster manager does not use this term. The product documentation identifies a node that hosts user services as a **compute node with services installed upon it**.

See **compute node**.

site network

The network that exists at your site. This network is the network to which you want to attach the cluster. For example, if you want to attach the cluster to your corporation's network, the corporate network is the site network.

In the cluster configuration tool, this term appears as *house network*.

slot

The cluster manager uses the term **slot** in the following ways:

1. All the partitions related to a Linux installation. A multiple-slot disk layout is also called a **cascading multiroot layout** or a **cascading multiboot layout**. A multiple-slot disk layout enables the following:

   - You can install different operating system releases or distributions.
   - You can choose the slot you want to boot.

2. The physical location of a blade within the blade enclosure. Information regarding the physical slot resides in configuration files on admin and ICE leader nodes. For an example, see `/var/opt/sgi/lib/blademond/slot_map`.

storage node

A compute node with storage services installed upon the node. The storage node could be a Cisco MDS node, an object storage server (OSS) node, or another type of service node.

switch controller

A device that enables you to power up or power down a switch.. A switch controller is similar to a node controller on a compute node.

system group

Nodes that the cluster manager addresses in a similar way. These nodes are grouped according to their function. The groups are as follows:

- `admin`
- `chassis`
- `cmc`
- `compute`
- `cooldev`
- `ib_switch`
- `ice_compute`
- `leader`
- `leader_alias`
- `mgmt_switch`
- `pdu`
- `switch_blade`

temponame

This term is deprecated.

See **internal name**.