

Hitachi Content Platform for Cloud Scale

2.6

Installing HCP for Cloud Scale

This document contains instructions for installing or updating the Hitachi Content Platform for cloud scale (HCP for cloud scale) software. It describes how to install single- and multi-instance HCP for cloud scale systems.

© 2020, 2023 Hitachi Vantara LLC. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including copying and recording, or stored in a database or retrieval system for commercial purposes without the express written permission of Hitachi, Ltd., or Hitachi Vantara LLC (collectively "Hitachi"). Licensee may make copies of the Materials provided that any such copy is: (i) created as an essential step in utilization of the Software as licensed and is used in no other manner; or (ii) used for archival purposes. Licensee may not make any other copies of the Materials. "Materials" mean text, data, photographs, graphics, audio, video and documents.

Hitachi reserves the right to make changes to this Material at any time without notice and assumes no responsibility for its use. The Materials contain the most current information available at the time of publication.

Some of the features described in the Materials might not be currently available. Refer to the most recent product announcement for information about feature and product availability, or contact Hitachi Vantara LLC at https://support.hitachivantara.com/en_us/contact-us.html.

Notice: Hitachi products and services can be ordered only under the terms and conditions of the applicable Hitachi agreements. The use of Hitachi products is governed by the terms of your agreements with Hitachi Vantara LLC.

By using this software, you agree that you are responsible for:

1. Acquiring the relevant consents as may be required under local privacy laws or otherwise from authorized employees and other individuals; and
2. Verifying that your data continues to be held, retrieved, deleted, or otherwise processed in accordance with relevant laws.

Notice on Export Controls. The technical data and technology inherent in this Document may be subject to U.S. export control laws, including the U.S. Export Administration Act and its associated regulations, and may be subject to export or import regulations in other countries. Reader agrees to comply strictly with all such regulations and acknowledges that Reader has the responsibility to obtain licenses to export, re-export, or import the Document and any Compliant Products.

Hitachi and Lumada are trademarks or registered trademarks of Hitachi, Ltd., in the United States and other countries.

AIX, AS/400e, DB2, Domino, DS6000, DS8000, Enterprise Storage Server, eServer, FICON, FlashCopy, GDPS, HyperSwap, IBM, Lotus, MVS, OS/390, PowerHA, PowerPC, RS/6000, S/390, System z9, System z10, Tivoli, z/OS, z9, z10, z13, z14, z15, z16, z/VM, and z/VSE are registered trademarks or trademarks of International Business Machines Corporation.

Active Directory, ActiveX, Bing, Excel, Hyper-V, Internet Explorer, the Internet Explorer logo, Microsoft, Microsoft Edge, the Microsoft corporate logo, the Microsoft Edge logo, MS-DOS, Outlook, PowerPoint, SharePoint, Silverlight, SmartScreen, SQL Server, Visual Basic, Visual C++, Visual Studio, Windows, the Windows logo, Windows Azure, Windows PowerShell, Windows Server, the Windows start button, and Windows Vista are registered trademarks or trademarks of Microsoft Corporation. Microsoft product screen shots are reprinted with permission from Microsoft Corporation.

All other trademarks, service marks, and company names in this document or website are properties of their respective owners.

Copyright and license information for third-party and open source software used in Hitachi Vantara products can be found in the product documentation, at <https://www.hitachivantara.com/en-us/company/legal.html> or https://knowledge.hitachivantara.com/Documents/Open_Source_Software.

Contents

Preface	6
About this document.....	6
Intended audience.....	6
Product version.....	6
Release notes.....	6
Related documents.....	6
Document conventions.....	7
Accessing product documentation.....	8
Getting help.....	9
Comments.....	9
Chapter 1: Hitachi Content Platform for cloud scale overview	10
About HCP for cloud scale.....	10
System scaling.....	10
Single-instance systems vs. multi-instance systems.....	11
About master and worker instances.....	12
Services.....	12
Volumes.....	13
Chapter 2: System requirements and sizing	15
Hardware requirements.....	15
Software requirements.....	15
Operating system and Docker minimum requirements.....	16
Operating system and Docker qualified versions.....	16
Docker considerations.....	17
SELinux considerations.....	17
Virtual machine host requirements.....	18
Networking.....	18
Time source.....	20
Supported browsers.....	20
Chapter 3: Installing HCP for cloud scale	21
Installing HCP for cloud scale using the cluster deployment tool.....	21
Items and information you need.....	21
Configuring boot settings.....	22

Preparing a bootable USB drive and installing OS.....	23
Preparing to install using cluster deployment tool	24
Installing using GUI.....	30
Installing using the configuration file	32
Deploying HCP for cloud scale software.....	37
Adding a new node.....	39
Restarting a failed installation.....	40
Installing HCP for cloud scale manually.....	40
Items and information you need.....	40
Installation workflow.....	41
Decide how many instances to deploy.....	41
Configure your networking environment.....	42
Optional: Select master instances.....	42
Install Docker on each server or virtual machine	42
Configure Docker on each server or virtual machine.....	43
Optional: Install Docker volume drivers.....	43
Optional: Enable or disable SELinux on each server or virtual machine.....	44
Configure maximum map count setting.....	44
Configure the firewall rules on each server or virtual machine.....	44
Run Docker on each server or virtual machine.....	44
Unpack the installation package.....	45
(Optional) Reconfigure network.config on each server or virtual machine...	46
(Optional) Reconfigure volume.config on each server or virtual machine....	47
Run the setup script on each server or virtual machine.....	50
Start the application on each server or virtual machine.....	53
Optional: Configure NTP.....	54
Use the service deployment wizard.....	54
Optional: Configure networks for services.....	57
Optional: Configure volumes for services.....	57
Deploying the system using CLI commands.....	61
Create an owner for new files.....	62
Optional: Verify the created volumes.....	63
Optional: Distribute services among system instances.....	63
Moving and scaling floating services.....	63
Moving and scaling services with multiple types.....	63
Best practices.....	64
Considerations.....	64
Troubleshooting.....	64
Relocating services.....	64
Configure the system for your users.....	66

Chapter 4: Updating HCP for cloud scale.....	68
Items and information you need.....	68
Verify and place the update archive.....	68
Use the System Management application to update the system.....	69
Configure the system for your users.....	71
Troubleshooting.....	71
Appendix A: Logs and diagnostic information.....	73
Log levels.....	73
Log management.....	73
Retrieving logs and diagnostic information.....	74
Default log locations.....	75
Appendix B: Services list.....	79
HCP for cloud scale services.....	79
Appendix C: Handling network changes.....	98
Safely changing an instance IP address.....	98
After a network change.....	98

Preface

About this document

This document contains instructions for installing or updating the Hitachi Content Platform for cloud scale (HCP for cloud scale) software. It describes how to install single- and multi-instance HCP for cloud scale systems.

Intended audience

This document is intended for people who are installing or updating HCP for cloud scale systems. It assumes you have some experience creating Docker configurations and installing computer software.

Product version

This document applies to v2.6 of Hitachi Content Platform for cloud scale.

Release notes

Read the release notes before installing and using this product. They may contain requirements or restrictions that are not fully described in this document or updates or corrections to this document. Release notes are available on the Hitachi Vantara Support Website: <https://knowledge.hitachivantara.com/Documents>.

Related documents

The following list describes documents containing information about v2.6 of HCP for cloud scale. You should have these documents available before using the product. Refer to the latest version of the *Hitachi Content Platform for Cloud Scale Release Notes* for information on document version numbers.

- *Hitachi Content Platform for Cloud Scale Release Notes* (RN-HCPCS004): This document is for customers and describes new features, product documentation, and resolved and known issues, and gives other useful information about this release of the product.
- *Installing Hitachi Content Platform for Cloud Scale* (MK-HCPCS002): This document gives you the information required to install or update the HCP for cloud scale software.

- *Hitachi Content Platform for Cloud Scale Administration Guide (MK-HCPCS008)*: This document explains how to use the HCP for cloud scale Object Storage Management and System Management applications to configure and operate a common object storage interface for clients to interact with; configure HCP for cloud scale for users; enable and disable system features; and monitor the system and its connections.
- *Hitachi Content Platform for Cloud Scale S3 Console Guide (MK-HCPCS009)*: This document is for end users and explains how to use the HCP for cloud scale S3 Console application to use S3 credentials and to simplify the process of creating, monitoring, and maintaining S3 buckets and the objects they contain.
- *Hitachi Content Platform for Cloud Scale Management API Reference (MK-HCPCS007)*: This document is for customers and describes the object storage management application programming interface (API) methods available for customer use.







Document conventions

This document uses the following typographic conventions:

Convention	Description
Bold	<ul style="list-style-type: none"> ▪ Indicates text in a window, including window titles, menus, menu options, buttons, fields, and labels. Example: Click OK. ▪ Indicates emphasized words in list items.
<i>Italic</i>	<ul style="list-style-type: none"> ▪ Indicates a document title or emphasized words in text. ▪ Indicates a variable, which is a placeholder for actual text provided by the user or for output by the system. Example: <pre>pairdisplay -g group</pre> <p>(For exceptions to this convention for variables, see the entry for angle brackets.)</p>
Monospace	Indicates text that is displayed on screen or entered by the user. Example: <code>pairdisplay -g oradb</code>
< > angle brackets	Indicates variables in the following scenarios: <ul style="list-style-type: none"> ▪ Variables are not clearly separated from the surrounding text or from other variables. Example: <pre>Status-<report-name><file-version>.csv</pre> ▪ Variables in headings.
[] square brackets	Indicates optional values. Example: [a b] indicates that you can choose a, b, or nothing.

Convention	Description
{ } braces	Indicates required or expected values. Example: { a b } indicates that you must choose either a or b.
vertical bar	Indicates that you have a choice between two or more options or arguments. Examples: [a b] indicates that you can choose a, b, or nothing. { a b } indicates that you must choose either a or b.

This document uses the following icons to draw attention to information:

Icon	Label	Description
	Note	Calls attention to additional information.
	Tip	Provides helpful information, guidelines, or suggestions for performing tasks more effectively.
	Important	Highlights information that is essential to the completion of a task.
	Caution	Warns the user of adverse conditions and/or consequences (for example, disruptive operations, data loss, or a system crash).
	CAUTION	Warns the user of a hazardous situation that, if not avoided, could result in major or minor injury.
	WARNING	Warns the user of a hazardous situation which, if not avoided, could result in death or serious injury.

Accessing product documentation

Product user documentation is available on the Hitachi Vantara Support Website: <https://knowledge.hitachivantara.com/Documents>. Check this site for the most current documentation, including important updates that may have been made after the release of the product.

Getting help

The [Hitachi Vantara Support Website](https://support.hitachivantara.com/en_us/contact-us.html) is the destination for technical support of products and solutions sold by Hitachi Vantara. To contact technical support, log on to the Hitachi Vantara Support Website for contact information: https://support.hitachivantara.com/en_us/contact-us.html.

[Hitachi Vantara Community](https://community.hitachivantara.com) is a global online community for Hitachi Vantara customers, partners, independent software vendors, employees, and prospects. It is the destination to get answers, discover insights, and make connections. **Join the conversation today!** Go to community.hitachivantara.com, register, and complete your profile.

Comments

Please send comments to doc.comments@hitachivantara.com. Include the document title and number, including the revision level (for example, -07), and refer to specific sections and paragraphs whenever possible. All comments become the property of Hitachi Vantara LLC.

Thank you!

Chapter 1: Hitachi Content Platform for cloud scale overview

This module describes the Hitachi Content Platform for cloud scale (HCP for cloud scale) system and its main use cases.

About HCP for cloud scale

HCP for cloud scale is a software-only data storage platform that rests on top of physical or cloud-based data storage systems, such as Hitachi Content Platform (HCP) and Amazon Web Services (AWS). HCP for cloud scale acts as a common interface between the storage systems that manages all storage objects, including buckets, objects, and metadata. HCP for cloud scale can scale to accommodate for any number of storage systems, and its data storage limitations are defined only by its underlying technologies.

System scaling

You manage how the system scales by adding or removing instances to the system and also by specifying which services run on those instances.

Instances

An instance is a server or virtual machine on which the software is running. A system can have either a single instance or multiple instances. Multi-instance systems have a minimum of four instances.

A system with multiple instances maintains higher availability in the event of instance failures. Additionally, a system with more instances can run tasks concurrently and can typically process tasks faster than a system with fewer or only one instance.

A multi-instance system has two types of instances: master instances, which run an essential set of services, and non-master instances, which are called workers.

Services

Each instance runs a configurable set of services, each of which performs a specific function. For example, the Metadata Gateway service stores metadata persistently.

In a single-instance system, that instance runs all services. In a multi-instance system, services can be distributed across all instances.

Single-instance systems vs. multi-instance systems

An HCP for cloud scale system can have a single instance or can have multiple instances (four or more).



Note: Every instance must meet RAM, CPU, and disk space minimums.

One instance

A **single-instance system** is useful for testing and demonstration purposes. It needs only a single server or virtual machine and can perform all product functionality.

However, a single-instance system has these drawbacks:

- A single-instance system has a single point of failure. If the instance hardware fails, you lose access to the system.
- With no additional instances, you cannot choose where to run services. All services run on the single instance.

Multiple instances

A multi-instance system is suitable for use in a production environment because it offers these advantages over a single-instance system:

- You can control how services are distributed across the multiple instances, providing improved service redundancy, scale out, and availability.
For information on services, see [Services list \(on page 79\)](#).
- A multi-instance system can survive instance outages. For example, with a four-instance system running the default distribution of services, the system can lose one instance and still remain available.
- Performance is improved as work can be performed in parallel across instances.
- You can add additional instances to the system at any time.



Note: You cannot change a single-instance system into a production-ready multi-instance system by adding new instances. This is because you cannot add master instances. **Master instances** are special instances that run a particular set of HCP for cloud scale services. Single-instance systems have one master instance. Multi-instance systems have at least three.

By adding additional instances to a single-instance system, your system still has only one master instance, meaning there is still a single point of failure for the essential services that only a master instance can run.

For information about adding instances to an existing HCP for cloud scale system, see the HCP for cloud scale online help.

Four-instance system considerations

The minimum HCP for cloud scale configuration has four instances. Four-instance systems should have three master instances.

For information about master instances, see [About master and worker instances \(on page 12\)](#).

About master and worker instances

Master instances are special instances that run an essential set of services, including:

- Admin-App service
- Cluster-Coordination service
- Synchronization service
- Service-Deployment service

Non-master instances are called workers. Workers can run any services except for those listed previously.

Single-instance systems have one master instance while multi-instance systems have either one or three master instances.



Important: You cannot add master instances to a system after it's installed. However, you can add worker instances.

Services

Services perform functions essential to the health or functionality of the system. For example, the Cluster Coordination service manages hardware resource allocation, while the Policy Engine service runs synchronous and asynchronous policies triggered by S3 API requests. Internally, services run in Docker containers on the instances in the system.

Service categories

Depending on what actions they perform, services are grouped into these categories:

- *Services*: Enable product functionality. You can scale, move, and reconfigure these services.
- *System services*: Maintain the health and availability of the system. You cannot scale, move, or reconfigure these services.

Some system services run only on master instances.

For a complete list of services, see [Services list \(on page 79\)](#).

Applications

Some services are classified as *applications*. These are the services with which users interact. Services that are not applications typically interact only with other services.

Service instances

Services run on instances in the system. Most services can run simultaneously on multiple instances. That is, you can have multiple instances of a service running on multiple instances in the system. Some services run on only one instance.

Each service has a best and required number of instances on which it should run.

You can configure where Hitachi Content Platform for cloud scale services run, but not system services.

Floating services

If a service supports *floating*, you have flexibility in configuring where new instances of that service are started when service instances fail.

Non-floating (or *persistent*) services run on the specific instances that you specify. If one of those service instances fails, the system does not automatically bring up a new instance of that service on another system instance.

With a service that supports floating, you specify a pool of eligible system instances and the number of service instances that should be running at any time. If a service instance fails, the system brings up another one on one of the system instances in the pool that doesn't already have an instance of that service running.

For services with multiple types, the ability to float can be supported on a per-type basis.



Note: HCP for cloud scale has no services with multiple types.

Networking

Each service binds to a number of ports and to one type of network, either internal or external. Networking for each service is configured during system installation and cannot be changed after a system is running.

For information on configuration, see [Networking \(on page 18\)](#).

Storage for services

Services can use volumes for storing data.

For information on volumes, see [Volumes \(on page 13\)](#).

Volumes

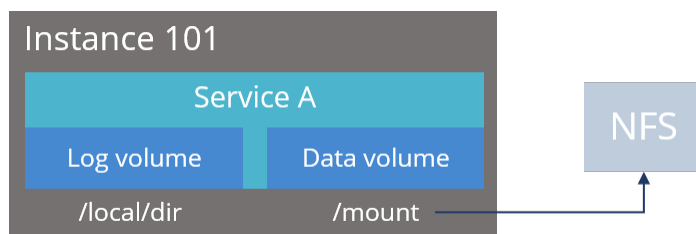
Volumes are properties of services that specify where and how a service stores its data.

You can use volumes to configure services to store their data in external storage systems, outside of the system instances. This allows data to be more easily backed up or migrated.

Volumes can also allow services to store different types of data in different locations. For example, a service might use two separate volumes, one for storing its logs and the other for storing all other data.

Example

In this example, service A runs on instance 101. The service's Log volume stores data in a folder on the system instance and the service's Data volume stores data in an NFS mount.



Creating and managing volumes

Depending on how they are created and managed, volumes are separated into these groups:

- System-managed volumes are created and managed by the system. When you deploy the system, you can specify the volume driver and options that the system should use when creating these volumes.

After the system is deployed, you cannot change the configuration settings for these volumes.

- User-managed volumes can be added to services and job types after the system has been deployed. These are volumes that you manage; you need to create them on your system instances before you can configure a service or job to use them.



Note: As of release 1.3.0, none of the built-in services support adding user-managed volumes.

Volume drivers

When configuring a volume, you specify the volume driver that it should use. The volume driver determines how and where data is stored.

Because services run in Docker containers on instances in the system, volume drivers are provided by Docker and other third-party developers, not by the system itself. For information about volume drivers you can use, see the applicable Docker or third-party developer's documentation.

By default, all services do not use volume drivers but instead use the bind-mount setting. With this setting, data for each service is stored within the system installation folder on each instance where the service runs.

For more information on volume drivers, see the Docker documentation.

For information about services, see [Services \(on page 12\)](#).


Chapter 2: System requirements and sizing

This module describes the hardware, networking, and operating system requirements for running an HCP for cloud scale system with one or more instances.

Hardware requirements

To install HCP for cloud scale on on-premises hardware for production use, you must provision at least four instances (nodes) with sufficient CPU, RAM, disk space, and networking capabilities. This table shows the hardware resources required for each instance of an HCP for cloud scale system for a minimum qualified configuration and a standard qualified configuration.

Resource	Minimum configuration	Standard configuration
CPU	Single CPU, 10-core	Dual CPU, 20+-core
RAM	128 GB	256 GB
Available disk space	(4) 1.92 TB SSD, RAID10	(8) 1.92 TB SSD, RAID10
Network interface controller (NIC)	(2) 10 GB Ethernet NICs	(2) 25 GB Ethernet NICs or (4) 10 GB Ethernet NICs

 **Important:** Each instance uses all available RAM and CPU resources on the server or virtual machine on which it's installed.

Software requirements

The following table shows the minimum requirements and best-practice software configurations for each instance in an HCP for cloud scale system.

Resource	Minimum	Best
IP addresses	(1) static	(2) static

Resource	Minimum	Best
Firewall Port Access	Port 443 for SSL traffic Port 8000 for System Management App GUI Port 8888 for Content Search App GUI	Same
Network Time	IP address of time service (NTP)	Same

Operating system and Docker minimum requirements

Each server or virtual machine you provide must have the following:

- 64-bit Linux distribution
- Docker version installed: Docker Community Edition 18.09.0 or later
- IP and DNS addresses configured

Additionally, you should install all relevant patches on the operating system and perform appropriate security hardening tasks.



Important: The system cannot run with Docker versions before 1.13.1.

To execute scripts provided with the product on RHEL, you should install Python.

Operating system and Docker qualified versions

This table shows the operating system, Docker, and SELinux configurations with which the HCP for cloud scale system has been qualified.



Important: An issue in Docker Enterprise Edition 19.03.15 and resolved in 20.10.5 prevented HCP for cloud scale deployment. Do not install any version of Docker Enterprise Edition above 19.03.14 and below 20.10.5.

Operating system	Docker version	Docker storage configuration	SELinux setting
Red Hat Enterprise Linux 8.6	Docker Community Edition 19.03.12 or later	overlay2	Enforcing

If you are installing on Amazon Linux, before deployment, edit the file `/etc/security/limits.conf` on every node to add the following two lines:

```
* hard nofile 65535
* soft nofile 65535
```

Docker considerations

The Docker installation folder on each instance must have at least 20 GB available for storing the Docker images.

Make sure that the Docker storage driver is configured correctly on each instance before installing the product. After you install the product, to change the Docker storage driver you must reinstall the product. To view the current Docker storage driver on an instance, run:

```
docker info
```

Core dumps can fill a host's file system, which can result in host or container instability. Also, if your system uses the data at rest encryption (DARE) feature, encryption keys are written to the dump file. It's best to disable core dumps.

To enable SELinux on the system instances, you need to use a Docker storage driver that SELinux supports. The storage drivers that SELinux supports differ depending on the Linux distribution you're using. For more information, see the Docker documentation.

If you are using the Docker `devicemapper` storage driver:

- Make sure that there's at least 40 GB of Docker metadata storage space available on each instance. The product needs 20 GB to install successfully and an additional 20 GB to successfully update to a later version.

To view Docker metadata storage usage on an instance, run:

```
docker info
```

- On a production system, do not run `devicemapper` in `loop-lvm` mode. This can cause slow performance or, on certain Linux distributions, the product might not have enough space to run.

SELinux considerations

- You should decide whether you want to run SELinux on system instances and enable or disable it before installing additional software on the instance.

Enabling or disabling SELinux on an instance needs a restart of the instance.

To view whether SELinux is enabled on an instance, run: `sestatus`

- To enable SELinux on the system instances, you need to use a Docker storage driver that SELinux supports.

The storage drivers that SELinux supports differ depending on the Linux distribution you're using. For more information, see the Docker documentation.

Virtual machine host requirements

You can deploy HCP for cloud scale on virtual machines from a .tgz file. Instances deploying HCP for cloud scale cannot run other software and multiple HCP for cloud scale nodes cannot be deployed on a single instance. To deploy multiple instance on the same hardware, use a hypervisor and guest VMs and deploy each HCP for cloud scale instance on a separate guest.

The HCP for cloud scale software has been qualified on these virtual machine host platforms:

- Hitachi Unified Compute Platform (UCP) 4.1.0

Networking

This topic describes the network usage by, and requirements for, both system instances and services.



Note:

- You can configure the network settings for each service when you install the system. You cannot change these settings after the system is up and running.
- If the networking environment changes such that the system can no longer function with its current networking configuration, you must reinstall the system.

Cluster host name

The HCP for cloud scale cluster host name is configured during installation. The cluster host name is required because it's needed for access to both the HCP for cloud scale user interface and the S3 API.

Instance IP address requirements

All instance IP addresses must be static, including both internal and external network IP addresses if applicable to the system. If you replace an instance, you can reuse its IP address. By doing so you don't have to change DNS entries and you conserve the address.

Network types

Each of the HCP for cloud scale services can bind to one type of network, either **internal** or **external**, for receiving incoming traffic. If the network infrastructure supports having two networks, you might want to isolate the traffic for most system services to a secured internal network that has limited access. You can then leave the following services on the external network for user access:

- Admin-App
- Grafana
- Message Queue
- Metadata-Cache

- Metadata-Coordination
- Metadata-Gateway
- Policy-Engine
- Metrics
- S3-Gateway
- Tracing-Agent
- Tracing-Collector
- Tracing-Query
- MAPI-Gateway

You can use either a single network type for all services or a mix of both types. To use both types, every instance in the system must be addressable by two IP addresses, one on the internal network and one on the external network. If you use only one network type, each instance needs only one IP address.

Allowing access to external resources

Regardless of whether you're using a single network type or a mix of types, you must configure the network environment to ensure that all instances have outgoing access to the external resources you want to use, such as:

- The storage components where the object data is stored
- Identity providers for user authentication
- Email servers that you want to use for sending email notifications

Ports

Each service binds to a number of ports for receiving incoming traffic. Port mapping is visible from the Network tab for each service.

Before installing HCP for cloud scale, you can configure services to use different ports, or use the default values shown in the following tables.

The following services must be deployed with their default port values:

- Message Queue
- Metadata Cache
- Tracing Agent
- Tracing Collector
- Tracing Query

External ports

The following table contains information about the service ports that users use to interact with the system.

On every instance in the system, each of these ports:

- Must be accessible from any network that needs administrative or data access to the system
- Must be accessible from every other instance in the system

Default Port Value	Used by Service	Purpose
80 (S3 HTTP port, if enabled)	S3 Gateway	Object persistence and access
443 (S3 HTTPS port)	S3 Gateway S3 Console application	Object persistence and access Proxied by Network Proxy
3000	Grafana	Dashboards
8000	Admin App	System Management application GUI
8443 (S3 HTTPS port)	S3 Gateway	Object persistence and access Not proxied by Network Proxy, used by external load balancer
9099	MAPI Gateway	Object Storage Management application GUI

Time source

If you are installing a multi-instance system, each instance should run NTP (network time protocol) and use the same external time source. For information, see support.ntp.org.

Supported browsers

The HCP for cloud scale web applications support these web browsers:

- Google Chrome latest
- Mozilla Firefox latest

Chapter 3: Installing HCP for cloud scale

The following procedures describe how to install the HCP for cloud scale software.

This module describes how to prepare for and install the HCP for cloud scale software.

After you install the software, log in and deploy the system.

Installing HCP for cloud scale using the cluster deployment tool

The cluster deployment tool automates the installation of HCP for cloud scale on one or more physical nodes in a cluster. Installation includes the operating system, Docker, and the HCP for cloud scale software. Before modifying the cluster configuration, you must change the boot configuration settings as described in the section [Configuring boot settings \(on page 22\)](#).

This section describes how to deploy HCP for cloud scale on bare metal servers using the cluster deployment tool.

Items and information you need

To use the cluster deployment tool, the cluster must be configured as follows:

- Preserve the default settings on the Hitachi Advanced Server DS120 before deployment.
- Ensure that the network interface names begin with *ens*, *eth*, or *enp*.
- The Hitachi Advanced Server DS120 servers should have a RAID controller that treats all the hard drives as a single virtual drive with at least 500 GB of free space.
- Ensure that the hard drives included in the RAID1 and/or RAID10 servers contain only 2 out of 4 hard drives in the RAID. If you require more hard drives, delete RAID1 first and then create a new RAID10 with two spans.
- Ensure that the jump server has access to internet during the throughout deployment.
- Ensure that the switch side uplinks of all the remaining nodes are configured as regular access ports (not as port channels, LACP or LAG) until you complete the installation.
- Keep the USB drive connected to the jump server during the time of deployment.
- After you complete the installation, disable PXE boot on all nodes in the cluster.
- Use Baseboard Management Server (BMC) firmware version 4.23.06 or later and BIOS version 2.4 or later.
- Configure the following settings only if you are using **4-port method**. For more information about 4-port method, see [Preparing to install using cluster deployment tool \(on page 24\)](#).
 - Configure separate networks for front-end and back-end communication.
 - All the nodes must have the same front-end and back-end subnetwork.
 - Do not use DHCP server on the back-end subnetwork.
 - Use DHCP server on the front-end subnetwork.

Configuring boot settings

You must perform the following boot configuration changes on all the nodes you want to install HCP for cloud scale.



Important: For the procedure specific to your device, see the appropriate product documentation.

Change the BIOS and boot configuration settings as follows:

- Under Advanced, select Network Configuration:
 - Disable all the IPv6 and IPv4 HTTP devices, and then save the configuration.
 - Configure the BIOS settings to boot in UEFI (Unified Extensible Firmware Interface) mode.
 - Configure the jump node to boot first from the USB drive and then from hard drive.
 - Set the rest of the nodes to boot in UEFI mode from the back-end network.
- Under Boot select UEFI Network Drive BBS Priorities:
 - Disable all boot options that are not in use except the Internal network port that is used for (PXE) Preboot Execution Environment boot.
 - Change the boot order (option 1) to hard drive.

Preparing a bootable USB drive and installing OS

Before you begin

If you are installing a multi-instance system, the system must have either one or three master instances, regardless of the total number of instances it includes. You must first determine the master node on which you want to mount the installation volume and the run the scripts from.



Important: Keep the USB drive connected to the jump server during the time of deployment.

The following procedure describe how to prepare a bootable USB drive with the required images to deploy the HCP for cloud scale software using cluster deployment tool.

Procedure

1. Plug-in the USB drive and run this command to retrieve the device name:

```
lsscsi -g
```

2. Run this command to copy the image file (ISO) to the USB drive:

```
sudo dd if=./<ISO Name>.iso of=<SD Name> bs=1048576  
status=progress
```

3. After copying the image file (ISO) to the USB drive, run this command to unmount the device

```
sync;sync; sudo eject <SD Name>
```

4. Mount the installation media on the master node that you selected as installation node and boot the node from the installation media.
5. On the **Installation Summary** screen, navigate to **User Settings** and then click **Root Password**.

Creating a user password is a mandatory step before you can start the installation process. For optimal security, a password should be a minimum of 12 characters long and contain a combination of uppercase and lowercase letters, numbers, and special characters.



Caution: It is your responsibility to remember your custom passwords. HCP for cloud scale does not have the ability to retrieve lost passwords. If you lose your password, it will be permanently unrecoverable.

- a. In the **Password** field, type your password.
- b. In the **Confirm** field, re-enter your password.
- c. Click **Done**.

6. Click **Begin Installation.**

It takes approximately 10-15 minutes for the installation to complete. The installer restarts the device after the installation is complete. However, you must manually choose to boot from hard drive when the system restarts.



Important: Before you log in, ensure that you are using the Standard (X11 display server). The XWayland display server may not display the user interface configuration settings correctly. To change the display settings to Standard (X11 display server), select the configuration icon and then select Standard (X11 display server) on Xorg.

7. Log in with your credentials:

- a. In the **Username** field, type your username (*root*).
- a. In the **Password** field, type your password.

Preparing to install using cluster deployment tool

Before you begin

Determine the master node to mount the installation volume and to run all the scripts. Ensure that you wipe the hard drives clean on all the servers using the command `dd if=/dev/zero of=/dev/sdX bs=1G count=10 - replace /dev/sdX`.



Note: `/dev/sdX` represents the destination drive of your computer. If you are booting from a USB drive, then your destination drive (`/dev/sdX`) is that USB drive.

You must also identify three additional master nodes and the following information:

- Front-end and back-end IP addresses of all nodes. It can be an address range or a comma-separated address list.
- Hostnames for all nodes.
- Classless Inter-Domain Routing (CIDR) block size (for example, 24) for the customer subnetwork.
- IP address of site gateway.

- Front-end network interface names.
- Domain name server (DNS) name and IP address.
- Network time protocol (NTP) server name or IP address.
- Front-end and back-end IP addresses of HCP for cloud scale master nodes.
- Ensure that the network interface is properly connected and configured.

Networking Options

HCP for cloud scale offer the following three networking options:

- **4-port method:** Designed for systems with two NICs, each containing a total of four ports. This method combines the one set of two ports to create one external port and the other set of two ports to create one internal port.
- **2-port method:** Designed for systems equipped with a single NIC containing two ports. In this method, the two ports are combined to form a unified bonded external port. In order to ensure seamless communication between the system and the network, you must assign same IP addresses for both the frontend and backend IP addresses.
- **Custom Networking:** Designed for systems with custom networks, such as tagged VLANs, VLANs, and Virtual Machine networking.

The following procedures describe how to automatically deploy the HCP for cloud scale software using cluster deployment tool.

Procedure

1. Copy the installation files and store them in a folder on the server.
2. Create a folder named `install_path/hcpcs` on the server, using the following command:

```
mkdir /opt/hcpcs
```

3. Copy the installation package from the folder where you stored it to `install_path/hcpcs` using this command:

```
cp -r /run/media/root/RHEL-8-4-0-BaseOS-x86_64/hcpcsInstaller /opt/hcpcs/.
```

4. Run the `start.sh` script to set up the node for the deployment process using this command:

```
/opt/hcpcs/hcpcsInstaller/csinstaller/start.sh
```



Note: The node reboots after executing the `start.sh` script. During the reboot, ensure that the node boots from the hard disk.

5. Navigate to the installation folder. For example:
`cd /opt/hcpcs`
6. Select the appropriate course of action based on the networking option you have chosen.

- a. If you have selected the **4-port method** option, run the server network configuration script `jump_server_network_config.sh`

For example: `/opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_network_config.sh -i <frontend_ip> -I <backend_ip> -p <frontend_cidr_prefix> -P <backend_cidr_prefix> -d <dns_ip> -g <gateway_ip> -n <frontend_interface_name> -N <backend_interface_name> -b <additional_frontend_interface_name> -B <additional_backend_interface_name>`

```
./opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_network_config.sh -i
172.168.10.1 -I 172.168.20.1 -p 24 -P 24 -d 172.168.100.50 -g
172.168.10.200 -n ens17f0 -N ens17f1 -b ens49f0 -B ens49f1
```

Use the following table to determine which options to use:

Option	Description
<code>-i <frontend_ip></code>	The front-end IP address of the installation master node.
<code>-I <backend_ip></code>	The back-end IP address of the installation master node.
<code>-p <frontend_cidr_prefix></code>	The front-end CIDR block size for the customer subnetwork.
<code>-P <backend_cidr_prefix></code>	The back-end CIDR block size of the customer subnetwork.
<code>-d <dns_ip></code>	The front-end IP address of the DNS node.
<code>-g <gateway_ip></code>	The IP address of the gateway through which all requests are routed. This supports customer sites, such as labs, with a gateway server.
<code>-n <frontend_interface_name></code>	Name of the front-end network interface.
<code>-N <backend_interface_name></code>	Name of the back-end network interface.
<code>-b <additional_frontend_interface_name></code>	Name of the additional front-end interface.

Option	Description
-B <additional_backend_interface_name>	Name of the additional back-end interface.



Note:

The script configures the network on the installation master node, allows SSH connections and download all the packages required to run the installer.

- b. If you have selected the **2-port method** option, check the `config.ini` file to verify if the Single NIC status tag is set to true or false and then run the server network configuration script `jump_server_network_config.sh`
 - Verify the Single NIC status. The Single NIC status tag requires one of the following arguments:
 - `true` = The system has a single network interface controller (NIC) with two ports.
 - `false` = The system has more than one NIC.



Note: The default value is `false`.

- Run the server network configuration script `jump_server_network_config.sh`.

For example: `/opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_network_config.sh -i <frontend_ip> -I <backend_ip> -p <frontend_cidr_prefix> -P <backend_cidr_prefix> -d <dns_ip> -g <gateway_ip> -n <frontend_interface_name> -N <backend_interface_name> -b <additional_frontend_interface_name> -B <additional_backend_interface_name> -s <true/false>`

```
./opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_network_config.sh -i
172.168.10.1 -I 172.168.20.1 -p 24 -P 24 -d 172.168.100.50 -g
172.168.10.200 -n ens17f0 -N ens17f1 -b ens49f0 -B ens49f1 -s true
```

Use the following table to determine which options to use:

Option	Description
-s	Verifies whether the Single NIC status tag is true or false.

Option	Description
<code>-i <frontend_ip></code>	The front-end IP address of the installation master node.
<code>-I <backend_ip></code>	The back-end IP address of the installation master node. If the value of the Single NIC status tag is set to 'true,' enter the front-end IP address of the installation master node.
<code>-p <frontend_cidr_prefix></code>	The front-end CIDR block size for the customer subnetwork.
<code>-P <backend_cidr_prefix></code>	The back-end CIDR block size of the customer subnetwork. If the value of the Single NIC status tag is set to 'true,' enter the front-end CIDR block size for the customer subnetwork.
<code>-d <dns_ip></code>	The front-end IP address of the DNS node.
<code>-g <gateway_ip></code>	The IP address of the gateway through which all requests are routed. This supports customer sites, such as labs, with a gateway server.
<code>-n <frontend_interface_name></code>	Name of the front-end network interface.
<code>-N <backend_interface_name></code>	Name of the back-end network interface. If the value of the Single NIC status tag is set to 'true', enter the second interface name.
<code>-b <additional_frontend_interface_name></code>	Name of the additional front-end interface.
<code>-B <additional_backend_interface_name></code>	Name of the additional back-end interface.

**Note:**

The script configures the network on the installation master node, allows SSH connections and download all the packages required to run the installer.

- c. If you have selected the **custom networking** option, you must manually configure your custom network settings on the jump node before proceeding.
 - i. You must verify the `CustomNetwork` setting in the `config.ini` file. The `CustomNetwork` tag requires one of the following arguments:
 - `true` = The custom network is enabled.
 - `false` = The custom network is disabled.



Note: The default value is `true`.

- ii. Navigate to the scripts folder (`/etc/sysconfig/network-scripts/`) and modify the jumpserver network configuration file according to your preferred settings.



Caution: The settings and the configuration provided here is just an example. Actual configuration may vary based on your specific network setup. Consider the unique characteristics of your network when implementing any changes. Consult your network administrator for customized guidance and adjustments to align with your networking requirements.

The following is an example of the configuration file:

```
TYPE-ethernet
PROXY_METHOD=none
BROWSER_ONLY=no
BOOTPROTO=static
DEFROUTE=yes
IPADDR=173.18.24B 190
PREFIX=24
NAME=eosis192
GATEWAY=173.18.252.252
UUID=9a861c44-aa23-285e.8008-f2833541cd69
DEVICE=ens192
ONBOOT=yes

ifcfg.eosis192 229C
```

- iii. Ensure that you can successfully SSH into the jump node.

7. Run this script to launch the installer tool:

```
/opt/hcps/hcpsInstaller/csinstaller/csinstaller.sh
```

You see the HCP for cloud scale installation wizard prompting you to set up a password.

8. Enter a unique password:

This password is for the RHEL operating system and PXE nodes.

**Note:**

Your password must meet all the following conditions:

- It should be at least 12 characters long.
- It must contain at least one uppercase letter, one lowercase letter, and one digit.
- We recommend to use the same password that was used for the jump server in one of the earlier steps.

9. Re-enter the password:

When confirming your password, ensure that you enter it exactly as you did the first time.

Next steps

During the installation, you have the option of providing the required configuration information either automatically or manually using a configuration file.

- If you choose to install using configuration file, see [Installing using the configuration file \(on page 32\)](#).
- If you choose to install using GUI, see [Installing using GUI \(on page 30\)](#).

Installing using GUI

Before you begin

Before running the installer script, you must:

- Complete all the steps mentioned in the procedure [Preparing to install using cluster deployment tool](#) (on page 24).
- We recommend that you turn off all the nodes except the jump server, before running the `csinstaller.sh` script. You can turn on the nodes after clicking OK at the GUI installer prompt. The nodes will then automatically start to install RHEL.
- Collect the following information:
 - Front-end and back-end IP addresses of all nodes. It can be an address range or a comma-separated address list.
 - Hostnames for all nodes.
 - Classless Inter-Domain Routing (CIDR) block size (for example, 24) for the customer subnetwork.
 - IP address of site gateway.
 - Front-end network interface names.
 - Domain name server (DNS) name and IP address.
 - Network time protocol (NTP) server name or IP address.
 - Front-end and back-end IP addresses of HCP for cloud scale master nodes.

The following procedures describe how to manually deploy the HCP for cloud scale software using cluster deployment tool.

Procedure

1. Run this script to launch the installer tool:

```
/opt/hcpcs/hcpcsInstaller/csinstaller/csinstaller.sh
```

You see the HCP for cloud scale installation wizard.

2. Enter the information for all the fields:

In the **Frontend IP Addresses of all nodes** section:

- a. In the **Address Range Start** field, type the starting IP address.
- b. In the **Address Range End** field, type the ending IP address.


In the **Backend IP Addresses of all nodes** section:

- a. In the **Address Range Start** field, type the starting IP address.
- b. In the **Address Range End** field, type the ending IP address.

In the **Number of network Interfaces** section, select the number of interfaces, and then click **Next**.

3. The Settings area contains the fields and entries that make up the configuration. The following table describes the fields on this page:

Settings	Description
Frontend IPv4	Enter the frontend IPv4 address

Settings	Description
Backend IPv4	Enter the backend IPv4 address
Frontend IPv4 Gateway	Enter the frontend IPv4 gateway IP address.
Enter System DNS Name	Enter the system DNS server IP address.
Enter NTP Sever IP	NTP server IP addresses.  Note: If you do not have the IP address of the NTP server, enter 0.0.0.0
Enter DNS Server IP list	DNS Server IP addresses.
Select OS Version	Choose the operating system. The choices are
Select Frontend Master Nodes	Enter the frontend IP address list.
Select Backend Master Nodes	Enter the backend IP address list.

- Click **Next**.
- You see the **Please enter Hostnames for all nodes** page. Enter the host name against each MAC address in the page.



Note: You must provide the MAC address list to the application either through the user interface or through a file. The interface supports only up to 10 MAC addresses. If your scale of installation is large and have more than 10 nodes, use the `mac_list.txt` file. The script points to this file during the installation process and populates the section.

- In the **Summary of the Input** page, click **OK**.

It takes about 10 minutes to power up all the member nodes in the system. All the nodes PXE (Preboot Execution Environment) boot from the jump node, receiving IP assignments from the range you specified.



Important:

- Ensure that the PXE boot is completed on all the nodes.
- Ensure that the Boot Option #1 is set to [Network:UEFI] in the BIOS of all the member nodes.

Next steps

Depending on the Networking Options method you choose at the time of installation, follow the instructions in the [Deploying HCP for cloud scale software \(on page 37\)](#) section.

Installing using the configuration file

Before modifying the `config.ini` file and running the installer script, you must:

- Complete all the steps mentioned in the procedure [Preparing to install using cluster deployment tool \(on page 24\)](#).
- Collect the following information:
 - Front-end and back-end IP addresses of all nodes. It can be an address range or a comma-separated address list.
 - Hostnames for all nodes.
 - Classless Inter-Domain Routing (CIDR) block size (for example, 24) for the customer subnetwork.
 - IP address of site gateway.
 - Front-end network interface names.
 - Domain name server (DNS) name and IP address.
 - Network time protocol (NTP) server name or IP address.
 - Front-end and back-end IP addresses of HCP for cloud scale master nodes.

The following procedures describe how to install HCP for cloud scale software using a configuration file (`config.ini`).

A configuration file (`config.ini`) is a text file that you can create before starting `csinstaller.sh`. This file provides the installer with all the necessary information such as IP addresses, CIDR, DNS name, DNS server IP, NTP server IP, time zone, MAC addresses, and master node IPs that are required to configure the cluster. It is best to use configuration file if you intend to perform repeated or large-scale deployments.

The configuration file is available in the directory: `/opt/hcpcs/hcpcsInstaller/csappliance/csinstaller/`


Procedure

1. Run this command to edit the `config.ini` file and then press "i" on your keyboard to enable editing.

```
vi /opt/hcpcs/hcpcsInstaller/csappliance/csinstaller/config.ini/config.ini
```

Use the following table to determine which sections in the configuration file to update:

Section	Description
IPINFO	Add IPs in this section only if you have a specific range of IP addresses. Do not add non sequential IP addresses.
NONSEQINFO	Add non-sequential IP addresses separated by a comma.

Section	Description
MACADDRESSINFO	Enter Mac addresses of all the nodes. You need not provide the IP address of the first node (master node). If your scale of installation is large and have more than 10 nodes, use the <code>mac_list.txt</code> file. The script points to this file during the installation process and populates the section.
TIMEZONE	Enter time zone information.  Note: You must make sure that the time zone information is correct. The script stops executing with an error otherwise.

The following is an example of the configuration file:

```
#This flag remains true in case of custom networking and changes to 'false' in
case of standard networking
[NETWORKFLAG]
CustomNetwork=true

#Set this to 'true' in case of Single NIC to get ports forwarded externally.
Else it remains as 'false'
[NICTYPE]
SingleNIC=false

#In case of sequential IP Addresses, add the IP Addresses below and leave
"NONSEQINFO" blank
[IPINFO] #eg: 172.18.1.1
FrontIPRangeStart = 172.19.244.190
FrontIPRangeEnd = 172.19.244.94
BackIPRangeStart = 172.248.190.190
BackIPRangeEnd = 172.248.190.194

#In case of non sequential IP Addresses, add the IP Addresses below and leave
"IPINFO" blank
[NONSEQIPINFO] #Add Comma separated values eg: 172.18.1.7,172.18.1.2,172.18.1.5
FrontIPAddressList = 172.18.2.9,172.18.2.2,172.18.2.1
BackIPAddressList =

#Add Network CIDR details
[NETWORKINFO] #eg: 172.18.1.0/24
FrontIPNetwork = 172.19.244.0/24
BackIPNetwork = 172.248.190.0/24

#Add Network Gateway info
[GATEWAYINFO] #eg: 172.18.1.254
Gateway = 172.18.251.254

#Add the DNS Name and DNS IP Address
[DNSINFO] #eg: test and 172.18.4.45
DNSName = lab.archivas.com
DNSIP = 172.18.4.45

#Enter the frontend and backend IP Addresses of the 3 master nodes
[MASTERNODEINFO]
Front_masters_IP1 = 172.19.244.190
Front_masters_IP2 = 172.19.244.191
Front_masters_IP3 = 172.19.244.192
Back_masters_IP1 = 172.248.190.190
Back_masters_IP2 = 172.248.190.191
Back_masters_IP3 = 172.248.190.192

#For Total nodes < 10 = Enter the Mac Addresses of all the nodes which will be
```

```

PXE booted (Exclude the jump node mac id)
#For Total nodes > 10 = In the "mac_list.txt" file, Enter the Mac Addresses of
all the nodes which will be PXE booted (Exclude the jump node mac id)
[MACADDRESSINFO]
MacAddress1 = 42:g0:9c:30:02:01
MacAddress2 = 42:g0:9c:30:03:01
MacAddress3 = 42:g0:9c:30:04:01
MacAddress4 =
MacAddress5 =
MacAddress6 =
MacAddress7 =
MacAddress8 =
MacAddress9 =

#Add the NTP Server IP address, NTP Peer Server IP address and the Timezone
[NTPINFO]
NTPServerIP = 172.22.355.190
NTPPeerServerIP = 0.0.0.0
TimeZone = America/St_Thomas

```

2. Press **exit** to save the changes and then enter: `wq`.
3. Navigate to the installation folder. For example:

```
cd /opt/hcpcs/hcpcsInstaller/csinstaller/
```

4. Run the installer script `csinstaller.sh`.
For example:

```
/opt/hcpcs/hcpcsInstaller/csinstaller/csinstaller.sh
```

The script imports a set of Python packages as well as Trivial File Transfer Protocol (TFTP), Network File System (NFS), and DHCP software. After installing these packages, the script starts the wizard, and the HCP for Cloud Scale Installer window opens.

5. In the **Summary of the Input** page, verify all the information and then click **OK**.
It takes about 10 minutes to power up all the member nodes in the system. All the nodes PXE (Preboot Execution Environment) boot from the jump node, receiving IP assignments from the range you specified.



Important:

- Ensure that the PXE boot is completed on all the nodes.
- Ensure that the Boot Option #1 is set to [Network:UEFI] in the BIOS of all the member nodes.

Next steps

Depending on the Networking Options method you choose at the time of installation, follow the instructions in the [Deploying HCP for cloud scale software \(on page 37\)](#) section.

Deploying HCP for cloud scale software

Before you begin

- Ensure that PXE boot has completed on all nodes and that Boot Option #1 is set to [Network:UEFI] in the BIOS of each member node.
- Follow the instructions based on the method you selected at the time of installation: the 4-port method, the 2-port method, or custom networking.
 - If you are using the 4-port or 2-port method, skip to **step 3** in the following procedure.
 - The first two steps are only applicable for custom networking configurations.

The following procedures describe how to deploy the HCP for cloud scale software.

Procedure

1. Establish a connection with the PXE nodes and manually configure networking on each node. Ensure that you can successfully SSH into all the nodes.



Important: This step is applicable only for custom networking configuration. If you are using the 4-port or 2-port method, skip this step and go to **step 3**.

2. To initiate HCP for cloud scale on the PXE nodes, execute this script on the jump node. The script is located at `/opt/hcpcs/hcpcsInstaller/csinstaller/start_hcpcs.sh`.

This process may take anywhere between 30 seconds to 5 minutes.

After the script has successfully executed on all nodes, use SSH to log in to each individual node and perform the following verifications:

- a. Run `docker ps` and ensure that the watchdog service is running.
- b. Run `systemctl status hcpcs.service`.
- c. If the `systemctl status hcpcs.service` either shows as inactive or does not return anything and the watchdog service is not running, check the networking settings on that node.
- d. After you have resolved any networking issues, manually execute the `custom_post_deploy.sh` script. You can find this script in the root folder.



Important: This step is applicable only for custom networking configuration. If you are using the 4-port or 2-port method, skip this step and go to **step 3**.

3. Run this script to deploy HCP for cloud scale software:

```
/opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_post_deploy.sh
[-h] -i <frontend_ip> -I <backend_ip> -m <master_frontend_ips> -M
<master_backend_ips>
```

For example:

```
/opt/hcpcs/hcpcsInstaller/csinstaller/jump_server_post_deploy.sh -i 172.168.10.1
-I 172.168.20.1 -m 172.168.10.1,172.168.10.2,172.168.10.3 -M 172.168.20.1,
172.168.20.2,172.168.20.3 -c 172.168.10.0/24 -C 172.168.20.0/24
```



Note: The script installs Docker-CE, applies security updates, configures the firewall, and deploys HCP for cloud scale software. For the rest of the nodes, these steps are automated as part of the boot process.

Option	Description
-h	(Optional) Displays command syntax.
-i <frontend_ip>	The front-end IP address of the installation master node.
-I <backend_ip>	The back-end IP address of the installation master node.
-m <master_frontend_ips>	The front-end IP addresses of the HCP for cloud-scale master nodes as a comma-separated list.
-M <master_backend_ips>	The back-end IP addresses of the HCP for cloud-scale master nodes as a comma-separated list.
-c <frontend_cidr>	The front-end network CIDR of the cluster.
-C <backend_cidr>	The back-end network CIDR of the cluster.

After successful completion, you see the following message and the microservices admin application is launched:

```
Successfully ran security hardening script
Successfully completed all post deploy actions
```

4. Connect to the external IP address of the server.

For example:

```
https://10.0.0.1:8000
```

You see the HCP for cloud scale application's welcome page. Accept the certificate and enter a new password for the administrator.



Note: Administrator is the only local user. You must use either Active Directory or LDAP services to create additional users.

5. Click **Continue**.

6. In the **Cluster hostname/IP Address** field, enter the host name and then click **Continue**.

You see the **Instance Discovery** page.

7. Accept the default service port and network assignments and then click **Continue**.
Service deployment takes several minutes to complete. You see a progress bar showing the progress. When the deployment is complete, you see a **Set Up** completed message.
8. Click **Finish**.
9. Run this script to remove residual files: `/opt/hcpcs/hcpcsInstaller/csinstaller/cleanup_artifacts.sh`.

This script removes all installer packages, files, and directories, and disables and removes the TFTP service. It also disables and stops the NFS and DHCP servers but does not remove them.

If the installation fails, you can restart it. To restart a failed installation, see [Restarting a failed installation \(on page 40\)](#).

Adding a new node

With the cluster deployment tool, you can add a new node to a cluster, replace an existing node, or replace both master and worker nodes. The cluster deployment tool allows you to install the RHEL OS, Docker, HCP for cloud scale, and other required third-party software on the new node that you want to add to the cluster.



WARNING: If you're replacing a node, make sure to take these steps:

- Have system administrator privileges to perform this operation or inform your system administrator.
- Decommission the node(s) before adding the new node to the cluster.

Before you begin

Before you perform this operation, review the following installation procedures. Completing one or all of the required steps is a prerequisite:

1. [Preparing to install using cluster deployment tool \(on page 24\)](#)
2. [Installing using GUI \(on page 30\)](#)
3. [Installing using the configuration file \(on page 32\)](#)
4. [Deploying HCP for cloud scale software \(on page 37\)](#)

To add a node, in the Admin App:

Procedure

1. Select **System Management**.
The **System Management** page opens, displaying the system management services.
2. Select **Instances**.
A list of instances is displayed. Check whether the new instance is listed here.
3. Select **Services**.
The **Services** page opens, displaying the services and system services.

4. Select the service on the master node or worker node that you want to scale.
 - **Master node:** Services balance on the new master node based on the cluster load. You can scale up any service you need.
 - **Worker node:** By default, only the watchdog service starts. You can scale up other services as needed.

Based on your selection, configuration information for the service is displayed.

5. Click **Scale**, and if the service has more than one type, select the instance type that you want to scale.
6. Run this script to remove residual files: `/opt/hcpcs/hcpcsInstaller/csinstaller/cleanup_artifacts.sh`.

This script removes all installer packages, files, and directories, and disables and removes the TFTP service. It also disables and stops the NFS and DHCP servers but does not remove them.

If the installation fails, you can restart it. To restart a failed installation, see [Restarting a failed installation \(on page 40\)](#).

Restarting a failed installation

Use this step only to restart an installation if your earlier installation has failed. If one or more nodes fails to install, we recommend that you wipe the hard drives (RAID of all nodes including the jump server) before restarting the installation process.

Procedure

1. Run this script to restart the installation process: `/opt/hcpcs/hcpcsInstaller/csinstaller/clean_reset.sh`.

This script invokes the `cleanup_artifacts.sh`, preserves the old configuration (`config.ini`) file, restarts the DHCP, TFTP and other services. It deletes the network bonds and runs the `start.sh` script.

Installing HCP for cloud scale manually

The following procedures describe how to install the HCP for cloud scale software.

This module describes how to prepare for and install the HCP for cloud scale software.

After you install the software, log in and deploy the system.

Items and information you need

To install an HCP for cloud scale system, you need the appropriate installation package containing the product installation tarball (archive) file `hcpcs-version_number.tgz`.

This document shows the path to the HCP for cloud scale folder as `install_path`. The best folder path is `/opt`.

You need to determine the IP addresses of instances (nodes). It's best to use static IP addresses because if an IP address changes you must reinstall the system.

It's best to create an owner for the new files created during installation.

Installation workflow

The installation workflow for HCP for cloud scale consists of these steps. Some steps are required and some are optional, but you must do them in the order shown.

1. [Decide how many instances to deploy \(on page 41\).](#)
2. [Configure your networking environment \(on page 42\).](#)
3. [Optional: Select master instances \(on page 42\).](#)
4. [Install Docker on each server or virtual machine \(on page 42\).](#)
5. [Configure Docker on each server or virtual machine \(on page 43\).](#)
6. [Optional: Install Docker volume drivers \(on page 43\).](#)
7. [Optional: Enable or disable SELinux on each server or virtual machine \(on page 44\).](#)
8. [Configure maximum map count setting \(on page 44\).](#)
9. [Configure the firewall rules on each server or virtual machine \(on page 44\).](#)
10. [Run Docker on each server or virtual machine \(on page 44\).](#)
11. [Unpack the installation package \(on page 45\).](#)
12. [\(Optional\) Reconfigure network.config on each server or virtual machine \(on page 46\).](#)
13. [\(Optional\) Reconfigure volume.config on each server or virtual machine \(on page 47\).](#)
14. [Run the setup script on each server or virtual machine \(on page 50\).](#)
15. [Start the application on each server or virtual machine \(on page 53\).](#)
16. [Optional: Configure NTP \(on page 54\).](#)
17. [Use the service deployment wizard \(on page 54\).](#)
 - a. [Optional: Configure networks for services \(on page 57\).](#)
 - b. [Optional: Configure volumes for services \(on page 57\).](#)
18. [Create an owner for new files \(on page 62\).](#)
19. [Optional: Verify the created volumes \(on page 63\).](#)
20. [Optional: Distribute services among system instances \(on page 63\).](#)
21. [Configure the system for your users \(on page 66\).](#)

Decide how many instances to deploy

Before installing a system, you need to decide how many instances the system will have.

The minimum for a production system is four instances.

Procedure

1. Decide how many instances you need.
2. Select the servers or virtual machines in your environment that you intend to use as HCP for cloud scale instances.

Configure your networking environment

Before installing the system, you need to determine the networks and ports each HCP for cloud scale service will use.

Procedure

1. Determine what ports each HCP for cloud scale service should use. You can use the default ports for each service or specify different ones.
In either case, these restrictions apply:
 - Every port must be accessible from all instances in the system.
 - Some ports must be accessible from outside the system.
 - All port values must be unique; no two services, whether System services or HCP for cloud scale services, can share the same port.
2. Determine what types of networks, either internal or external, to use for each service.
If you're using both internal and external networks, each instance in the system must have IP addresses on both your internal and external networks.

Optional: Select master instances

If you are installing a multi-instance system, the system must have either one or three master instances, regardless of the total number of instances it includes.

You need to select which of the instances in your system will be master instances.

If you are installing a multi-instance system, the system must have either one or three master instances, regardless of the total number of instances it includes.



Important:

- For a production system, use three master instances.
- You cannot add master instances to a system after it's installed. You can, however, add any number of worker instances.

If you are deploying a single-instance system, that instance will automatically be configured as a master instance and run all services for the system.

Procedure

1. Select which of the instances in your system are intended as master instances.
2. Make note of the master instance IP addresses.



Note: To ensure system availability, run master instances on separate physical hardware from each other, if possible.

Install Docker on each server or virtual machine

On each server or virtual machine that is to be an HCP for cloud scale instance:

Procedure

1. In a terminal window, verify whether Docker 1.13.1 or later is installed:
`docker --version`
2. If Docker is not installed or if you have a version before 1.13.1, install the current Docker version suggested by your operating system.

The installation method you use depends on your operating system. See the [Docker website](#) for instructions.

Configure Docker on each server or virtual machine

Before installing the product, configure Docker with settings suitable for your environment. For guidance on configuring and running Docker, see the applicable Docker documentation.

Procedure

1. Ensure that the Docker installation folder on each instance has at least 20 GB available for storing the product Docker images.
2. Ensure that the Docker storage driver is configured correctly on each instance. After installation, changing the Docker storage driver needs reinstallation of the product.
To view the current Docker storage driver on an instance, run: `docker info`.
3. To enable SELinux on the system instances, use a Docker storage driver that SELinux supports.

The storage drivers that SELinux supports differ depending on the Linux distribution you're using. For more information, see the Docker documentation.

4. If you are using the Docker `devicemapper` storage driver, ensure that there's at least 40 GB of Docker metadata storage space available on each instance.

The product needs 20 GB to install successfully and an additional 20 GB to successfully update to a later version.

To view Docker metadata storage usage on an instance, run: `docker info`

Next steps

On a production system, do not run `devicemapper` in `loop-lvm` mode. This can cause slow performance or, on certain Linux distributions, the product might not have enough space to run.

Optional: Install Docker volume drivers

Volume drivers are provided by Docker and other third-party developers, not by the HCP for cloud scale system itself. For information on volume drivers, their capabilities, and their valid configuration settings, see the applicable Docker or third-party developer's documentation.

Procedure

1. If any services on your system are using Docker volume drivers (not the `bind-mount` setting) for storing data, install those volume drivers on the new instance that you are adding.

If you don't, services might fail to run on the new instance.

2. If any services on your system use Docker volume drivers for storing data (instead of using the default bind-mount setting), install those volume drivers on all instances in the system.

Optional: Enable or disable SELinux on each server or virtual machine

You should decide whether you want to run SELinux on system instances before installation.

Procedure

1. Enable or disable SELinux on each instance.
2. Restart the instance.

Configure maximum map count setting

You need to configure a value in the file `sysctl.conf`.

Procedure

1. On each server or virtual machine that is to be a system instance, open the file `/etc/sysctl.conf`.
2. Append this line: `vm.max_map_count = 262144`
If the line already exists, ensure that the value is greater than or equal to 262144.
3. Save and close the file.

Configure the firewall rules on each server or virtual machine

Before you begin

Determine the port values currently used by your system. To do this, on any instance, view the file `install_path/config/network.config`.

On each server or virtual machine that is to be a system instance:

Procedure

1. Edit the firewall rules to allow communication over all network ports that you want your system to use. You do this using a firewall management tool such as `firewalld`.
2. Restart the server or virtual machine.

Run Docker on each server or virtual machine

On each server or virtual machine that is to be a system instance, you need to start Docker and keep it running. You can use whatever tools you typically use for keeping services running in your environment.

For example, to run Docker using `systemd`:

Procedure

1. Verify that Docker is running:
`systemctl status docker`
2. If Docker is not running, start the `docker` service:
`sudo systemctl start docker`
3. (Optional) Configure the Docker service to start automatically when you restart the server or virtual machine:
`sudo systemctl enable docker`

Unpack the installation package

On each server or virtual machine that is to be a system instance:

Procedure

1. Download the installation package `hcpcs-version_number.tgz` and the MD5 checksum file `hcpcs-version_number.tgz.md5` and store them in a folder on the server or virtual machine.
2. Verify the integrity of the installation package. For example:
`md5sum -c hcpcs-version_number.tgz.md5`
If the package integrity is verified, the command displays `OK`.
3. In the largest disk partition on the server or virtual machine, create a folder named `install_path/hcpcs`. For example:
`mkdir /opt/hcpcs`
4. Move the installation package from the folder where you stored it to `install_path/hcpcs`. For example:
`mv hcpcs-version_number.tgz /opt/hcpcs/hcpcs-version_number.tgz`
5. Navigate to the installation folder. For example:
`cd /opt/hcpcs`
6. Unpack the installation package. For example:
`tar -zxf hcpcs-version_number.tgz`
A number of directories are created within the installation folder.

**Note:**

If you encounter problems unpacking the installation file (for example, the error message "tar: This does not look like a tar archive"), the file might have been packed multiple times during download. Use the following commands to fully extract the file:

```
$ gunzip hcpcs-version_number.tgz
$ mv hcpcs-version_number.tar hcpcs-version_number.tgz
$ tar -zxf hcpcs-version_number.tgz
```

7. Run the installation script `install`:

```
./install
```



Note:

- Don't change directories after running the installation script. The following tasks are performed in your current folder.
- The installation script can be run only one time on each instance. You cannot rerun this script to try to repair or upgrade a system instance.

(Optional) Reconfigure `network.config` on each server or virtual machine

Before you begin



Important: To reconfigure networking for the System services, you must complete this step before you run the setup script on each server or virtual machine.

You cannot change networking for System services after running the script `run` or after starting `HCI.service` using `systemd`.

You can change these networking settings for each service in your product:

- The ports that the service uses.
- The network to listen on for incoming traffic, either internal or external.

To configure networking for the System services:

Procedure

1. On each server or virtual machine that is to be an HCP for cloud scale instance, use a text editor to open the file `install_path/hcpcs/config/network.config`.

The file contains two types of lines for each service:

- **Network type assignments:** For example:

```
com.hds.ensemble.plugins.service.service_name_interface=[internal|external]
```

```
com.hds.ensemble.plugins.service.zookeeper_interface=internal
```

- **Port number assignments:** For example:

```
com.hds.ensemble.plugins.service.service_name.port.port_name=port_number
```

```
com.hds.ensemble.plugins.service.zookeeper.port.PRIMARY_PORT=2181
```

2. Type new port values for the services you want to configure.



Note: If you reconfigure service ports, make sure that each port value you assign is unique across all services, both System services and HCI services.



Note: By default, all System services are set to `internal`.

If you're only using a single network, you can leave these settings as they are. This is because all system instances are assigned both internal and external IP addresses in HCI; if you're only using a single network type, the internal and external IP addresses for each instance are identical.

3. On the lines containing `_interface`, specify the network that the service should use. Valid values are **internal** and **external**.
4. Save your changes and exit the text editor.

Next steps



Important: Ensure that the file `network.config` is identical on all HCI instances.

(Optional) Reconfigure `volume.config` on each server or virtual machine

Before you begin



Important: To reconfigure volumes for the System services, you must complete this step before you run the setup script on each server or virtual machine.

You cannot change volumes for System services after using the `run` script or after starting `HCI.service` with `systemd`.

By default, each of the System services is configured not to use volumes for storage (each service uses the bind-mount option). To change this configuration, you can do that now in this step, before running the product startup scripts.



Tip: System services typically do not store a lot of data, so you should favor keeping the default bind-mount setting for them.

You configure volumes for HCI services later when using the deployment wizard.

To configure volumes for the System services:

Procedure

1. On each server or virtual machine that is to be an HCI instance, use a text editor to open the file `install_path/hci/config/volume.config`.

This file contains information about the volumes used by the System services. For each volume, the file contains lines that specify the following:

- The name of the volume:

```
com.hds.ensemble.plugins.service.service_name.volume_name=volume_name
```



Note: Do not edit the volume names. The default volume name values contain variables (SERVICE_PLUGIN_NAME and INSTANCE_UUID) that ensure that each volume gets a unique name.

- The volume driver that the volume uses:

```
com.hds.ensemble.plugins.service.service_name.volume_driver=[volume_driver_name | bind-mount]
```

- The configuration options used by the volume driver. Each option is listed on its own line: For example, these lines describe the volume that the Admin-App service uses for storing its logs:

```
com.hds.ensemble.plugins.service.service_name.volume_driver_opt_option_number=volume_driver_option_and_value
```

```
com.hds.ensemble.plugins.service.adminApp.log_volume_name=SERVICE_PLUGIN_NAME.
INSTANCE_UUID.log
com.hds.ensemble.plugins.service.adminApp.log_volume_driver=bind-mount
com.hds.ensemble.plugins.service.adminApp.log_volume_driver_opt_1=hostpath=/
home/hcpcs/log/com.hds.ensemble.plugins.service.adminApp/
```


2. For each volume that you want to configure, you can edit the following:

- The volume driver for the volume to use. To do this, replace `bind-mount` with the name of the volume driver you want.

Volume drivers are provided by Docker and other third-party developers, not by the HCI system itself. For information on volume drivers, their capabilities, and their valid configuration settings, see the applicable Docker or third-party developer's documentation.

- On the line that contains `_opt`, the options for the volume driver.

For information about the options you can configure, see the documentation for the volume driver that you're using.



Caution: Option/value pairs can specify where data is written in each volume. These considerations apply:

- Each volume that you can configure here must write data to a unique location.
- The `SERVICE_PLUGIN` and `INSTANCE_UUID` variables cannot be used in option/value pairs.
- Make sure the options and values you specify are valid. Incorrect options or values can cause system deployment to fail or volumes to be set up incorrectly. For information on configuration, see the volume driver's documentation.



Tip: Create test volumes using the command `docker volume create` with your option/value pairs. Then, to test the volumes you've created, run the command `docker run hello-world` with the option `--volume`.

Example

These lines show a service that has been configured to use the `local-persist` volume driver to store data:

```
com.hds.ensemble.plugins.service.marathon.data_volume_name=SERVICE_PLUGIN_NAME.INSTANCE_UUID.data
com.hds.ensemble.plugins.service.marathon.data_volume_driver=local-persist
com.hds.ensemble.plugins.service.marathon.data_volume_driver_opt_1=mountpoint=/home/hcpcs/data/com.hds.ensemble.plugins.service.marathon/
```

Run the setup script on each server or virtual machine

Before you begin



Note:


- When installing a multi-instance system, make sure you specify the same list of master instance IP addresses on every instance that you are installing.
- When entering IP address lists, do not separate IP addresses with spaces. For example, the following is correct:








```
sudo install_path/hcps/bin/setup -i 192.0.2.4
-m 192.0.2.0,192.0.2.1,192.0.2.3
```

On each server or virtual machine that is to be a system instance:

Procedure

1. Run the script `setup` with the applicable options:

Option	Description
-i	The external network IP address for the instance on which you're running the script.
-I	The internal network IP address for the instance on which you're running the script.
-m	Comma-separated list of external network IP addresses of each master instance.
-M	Comma-separated list of internal network IP addresses of each master instance.
-i IPADDRESS	Displays the external instance IP address. If not specified, this value is discovered automatically.
-I IPADDRESS	Displays the internal instance IP address. If not specified, this value is the same as the external IP address.
-d	Attempts to automatically discover the real master list from the provided masters.
--hci_uid UID	Allows you to set the desired user ID (UID) for the HCI USER at <i>install time only</i> . <div data-bbox="678 1644 1393 1753" style="border: 1px solid black; background-color: #e0f2f1; padding: 5px; margin-top: 10px;">  Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster. </div>

Option	Description
<code>--hci_gid</code> GID	<p>Allows you to set the desired group ID (GID) for the HCI GROUP at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--mesos_uid</code> UID	<p>Allows you to set the desired user UID for the MESOS USER at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--mesos_gid</code> GID	<p>Allows you to set the desired GID for the MESOS GROUP at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--haproxy_uid</code> UID	<p>Allows you to set the desired UID for the HAPROXY USER at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--haproxy_gid</code> GID	<p>Allows you to set the desired GID for the HAPROXY GROUP at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--zk_uid</code> UID	<p>Allows you to set the desired UID for the ZOOKEEPER USER at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>
<code>--zk_gid</code> GID	<p>Allows you to set the desired GID for the ZOOKEEPER GROUP at <i>install time only</i>.</p> <p> Important: This value needs to be greater than 1000, less than or equal to 65533, and the same on all nodes in a cluster.</p>

Use the following table to determine which options to use:

Number of instances in the system	Network type usage	Options to use
Multiple	Single network type for all services	Either: -i and -m or -I and -M
Multiple	Internal for some services, external for others	All of these: -i, -I, -m, -M
Single	Single network type for all services	Either -i or -I
Single	Internal for some services, external for others	Both -i and -I

Result



Note: If the terminal displays Docker errors when you run the `setup` script, ensure that Docker is running.

Example

The following example sets up a single-instance system that uses only one network type for all services:

```
sudo install_path/hcpcs/bin/setup -i 192.0.2.4
```

To set up a multi-instance system that uses both internal and external networks, type the command in this format:

```
sudo install_path/hcpcs/bin/setup -i external_instance_ip -I
internal_instance_ip -m external_master_ips_list -M
internal_master_ips_list
```

For example:

```
sudo install_path/hcpcs/bin/setup -i 192.0.2.4 -I 10.236.1.0 -m
192.0.2.0,192.0.2.1,192.0.2.3 -M 10.236.1.1,10.236.1.2,10.236.1.3
```

The following table shows sample commands to create a four-instance system. Each command is entered on a different server or virtual machine that is to be a system instance. The resulting system contains three master instances and one worker instance and uses both internal and external networks.

Instance internal IP	Instance external IP	Master or worker	Command
192.0.2.1	10.236.1.1	Master	<code>sudo install_path/hcpcs/bin/setup -I 192.0.2.1 -i 10.236.1.1 -M 192.0.2.1,192.0.2.2,192.0.2.3 -m 10.236.1.1,10.236.1.2,10.236.1.3</code>
192.0.2.2	10.236.1.2	Master	<code>sudo install_path/hcpcs/bin/setup -I 192.0.2.2 -i 10.236.1.2 -M 192.0.2.1,192.0.2.2,192.0.2.3 -m 10.236.1.1,10.236.1.2,10.236.1.3</code>
192.0.2.3	10.236.1.3	Master	<code>sudo install_path/hcpcs/bin/setup -I 192.0.2.3 -i 10.236.1.3 -M 192.0.2.1,192.0.2.2,192.0.2.3 -m 10.236.1.1,10.236.1.2,10.236.1.3</code>
192.0.2.4	10.236.1.4	Worker	<code>sudo install_path/hcpcs/bin/setup -I 192.0.2.4 -i 10.236.1.4 -M 192.0.2.1,192.0.2.2,192.0.2.3 -m 10.236.1.1,10.236.1.2,10.236.1.3</code>

Start the application on each server or virtual machine

On each server or virtual machine that is to be a system instance:

Procedure

1. Start the application script `run` using whatever methods you usually use to run scripts.



Important: Ensure that the method you use can keep the `run` script running and can automatically restart it in the event of a server restart or other availability event.

Result

After the service starts, the server or virtual machine automatically joins the system as a new instance.

Example

Here are some examples of how you can start the script:

- You can run the script in the foreground:

```
sudo install_path/product/bin/run
```

When you run the `run` script this way, the script does not automatically complete, but instead remains running in the foreground.

- You can run the script as a service using `systemd`:
 1. Copy the product `.service` file to the appropriate location for your OS. For example:

```
cp install_path/product/bin/product.service /etc/systemd/system
```

2. Enable and start the `product.service` service:

```
sudo systemctl enable product.service
sudo systemctl start product.service
```

Optional: Configure NTP

If you are installing a multi-instance system:

Procedure

1. Configure NTP (network time protocol) so that each instance uses the same time source.

For information on NTP, see <http://support.ntp.org/>.

Use the service deployment wizard

After creating all of your instances and starting HCP for cloud scale, use the service deployment wizard. This wizard runs the first time you log in to the system.

To run the service deployment wizard:

Procedure

1. Open a web browser and go to `https://instance_ip_address:8000`. The Deployment Wizard starts.
2. Set and confirm the password for the main **admin** account.



Important: Do not lose or forget this password.

When you have defined the password, click **Continue**.

3. On the next page of the deployment wizard, type the cluster host name (as a fully qualified domain name in lowercase ASCII letters) in the **Cluster Hostname/IP Address** field, then click **Continue**.
Omitting this can cause links in the System Management application to function incorrectly.
4. On the next page of the deployment wizard, confirm the cluster topology. Verify that all the instances that you expect to see are listed and that their type (**Master** or **Worker**) is as you expect.
If some instances are not displayed, in the **Instance Discovery** section, click **Refresh Instances** until they appear.
When you have confirmed the cluster topology, click **Continue**.
5. On the next page of the deployment wizard, confirm the advanced configuration settings of services.



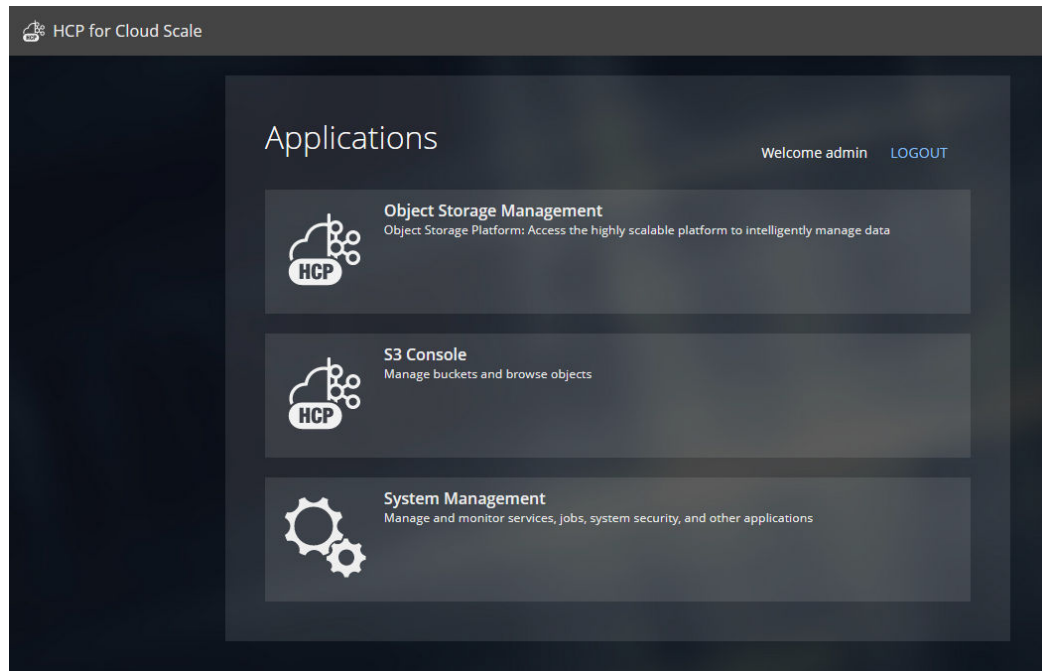
Important: If you decide to reconfigure networking or volume usage for services, you must do so now, before deploying the system.

For information on configuration, see [Networking \(on page 18\)](#).

- For information on networking settings for the HCP for cloud scale services, see [Optional: Configure networks for services \(on page 57\)](#).
- For information on storage volumes used for services, see [Optional: Configure volumes for services \(on page 57\)](#).

When you have confirmed the configuration settings, click **Continue**.

6. On the last page of the deployment wizard, to deploy the cluster, click **Deploy Cluster**.
If your network configuration results in a port collision, deployment stops and the deployment wizard notifies you which port is at issue. If this happens, edit the port numbers and try again.
After a brief delay, the deployment wizard displays the message "Starting deployment" and instances of services are started.
7. When the deployment wizard is finished, it displays the message "Setup Complete."
Click **Finish**.
The HCP for cloud scale **Applications** page opens.



Result

Service instances are deployed and you can now configure storage components.



Note: If you configured the System services networking incorrectly, the System Management application might not appear as an option on the **Applications** page. This can happen, for example, if the `network.config` file is not identical on all instances. For error information, view the file `install_path/hcpcs/config/cluster.config` or the output information logged by the script `run`.

To fix this issue, do the following:

1. Stop the script `run`. You can do this using whatever method you're currently using to run the script.
2. Run this command to stop all HCP for cloud scale Docker containers on the instance:

```
sudo install_path/hcpcs/bin/stop
```

3. Delete the contents of the folder `install_path/hcpcs` from all instances.
4. Delete any Docker volumes created during the installation:

```
docker volume rm volume-name
```

5. Begin the installation again from the step where you unpack the installation package.



Note: The following messages indicate that the deployment process failed to initialize a Metadata Gateway service instance:

- If the deployment process repeatedly tries and fails to reach a node, it displays this message: "Failed to initialize all MetadataGateway instances. Please re-deploy the system."
- If the deployment process detects an existing Metadata Gateway partition on a node, it displays this message: "Found existing metadata partitions on nodes, please re-deploy the system."

If you see either message, you can't resolve the issue by clicking Retry. Instead, you must reinstall the HCP for cloud scale software.

Optional: Configure networks for services

To change networking settings for the HCP for cloud scale services:

Procedure

1. On the **Advanced Configuration** page, select the service to configure.
2. On the **Network** tab:
 - a. Configure the ports that the service should use.



Note: If you reconfigure service ports, make sure that each port value you assign is unique across all services, both System services and HCP for cloud scale services.

- b. For each service, specify the network, either **Internal** or **External**, to which the service should bind.



Note: By default, the HCP for cloud scale services have the **External** network selected and the System services have the **Internal** network selected.

If you're only using a single network, you can leave these settings as they are. This is because all system instances are assigned both internal and external IP addresses in HCP for cloud scale; if you're only using a single network type, the internal and external IP addresses for each instance are identical.

Optional: Configure volumes for services

To change volume usage:

Procedure

1. On the **Advanced Configuration** page, select a service to configure.
2. Click the **Volumes** tab. This tab displays the system-managed volumes that the service supports. By default, each built-in service has both Data and Log volumes.
3. For each volume, provide Docker volume creation information:

- a. In the **Volume Driver** field, specify the name of the volume driver that the volume should use. To configure the volume not to use any volume driver, specify **bind-mount**, which is the default setting.



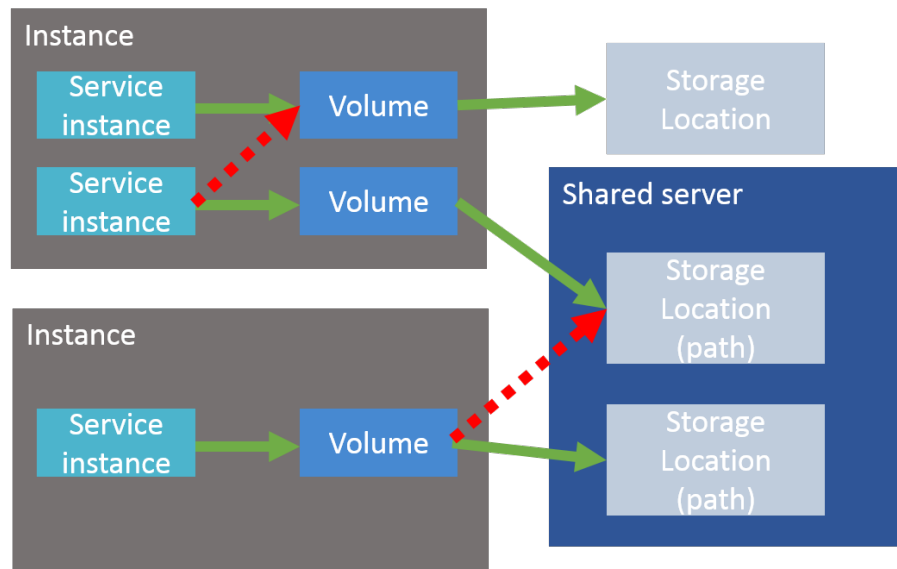
Note: Volume drivers are provided by Docker and other third-party developers, not by the HCP for cloud scale system itself. For information on volume drivers, their capabilities, and their valid configuration settings, see the applicable Docker or third-party developer's documentation.

- b. In the **Volume Driver Options** section, in the **Option** and **Value** fields, specify any optional parameters and their corresponding values for the volume driver:
 - If you're using the **bind-mount** setting, you can edit the value for the `hostpath` option to change the path where the volume's data is stored on each system instance. However, this must be a path within the HCP for cloud scale installation folder.
 - If you're using a volume driver:
 - i. Click the trashcan icon to remove the default **hostpath** option. This option applies only when you are using the **bind-mount** setting.
 - ii. Type the name of a volume driver option in the **Option** field. Then type the corresponding parameter for that option in the **Value** field.
 - iii. Click the plus-sign icon to add the option/value pair.
 - iv. Repeat this procedure for each option/value pair you want to add.

Option/value pairs can specify where data is written to in each volume. These considerations apply:

- Each service instance must write its data to a unique location. A unique location can be a file system or a unique path on a shared external storage server.

In this illustration, green arrows show acceptable configurations and red arrows show unacceptable configurations where multiple service instances are writing to the same volume, or multiple volumes are backed by the same storage location:



- For persistent (that is, non-floating) services, favor using the `$(container_inst_uuid)` variable in your option/value pairs. For persistent services, this variable resolves to a value that's unique to each service instance.

This is especially useful if the volume driver you're using is backed by a shared server. By providing a variable that resolves to a unique value, the volume driver can use the resolved variable to create unique directories on the shared server.

However, some volume drivers, such as Docker's local volume driver, do not support automatic Folder creation. If you're using such a volume driver, you need to create volume folders yourself. For an example of how to handle this, see the following Docker local volume driver example.

- Floating services do not support volumes that are backed by shared servers, because floating services do not have access to variables that resolve to unique values per service instance.
- Make sure the options and values you specify are valid. Options or values that are not valid can cause system deployment to fail or volumes to be set up incorrectly. For information on volumes, see the volume driver's documentation.



Tip: Create test volumes by use the command `docker volume create` with your option/value pairs. Then, to test the volumes you created, use the command `docker run hello-world --volume`.

You can include these variables when configuring volume options:

- `${install_dir}` is the product installation folder.
- `${data_dir}` is equal to `${install_dir}/data`
- `${log_dir}` is equal to `${install_dir}/log`
- `${volume_def_name}` is the name of the volume you are configuring.
- `${plugin_name}` is the name of the underlying service plugin.
- `${container_inst_uuid}` is the UUID for the Docker container in which the service instance runs. For floating services, this is the same value for all instances of the service.
- `${node_ip}` is the IP address for the system instance on which the service is running. This cannot be used for floating services.
- `${instance_uuid}` is the UUID for the system instance. This cannot be used for floating services. For services with multiple types, this variable resolves to the same value for all instances of the service, regardless of their types.

4. Repeat this procedure for each service that you want to configure.

bind-mount configuration for Database service log volume

The built-in Database service has a volume called log, which stores the service's logs. The log volume has this default configuration:

- Volume driver: bind-mount
- Option: hostname, Value: `${log_dir}/${plugin_name}/${container_inst_uuid}`

With this configuration, after the system is deployed, logs for the Database service are stored at a unique path on each system instance that runs the Database service:

```
install_path/hcpcs/log/com.hds.ensemble.plugins.service.cassandra/  
service-instance-uuid
```

Docker local volume driver for Database service log volume

Alternatively, you can configure the Database service to use Docker's built-in local volume driver to store logs on an NFS server. To do this:

1. Log in to your NFS server.
2. Create a folder.
3. Within that folder, create one folder for each of the instances in your system. Name each one using the instance IP address.



Note: In this example, you need to create these folders yourself because the local storage driver will not create them automatically.

4. Back in the system deployment wizard, in the Volume Driver field, specify local
5. Specify these options and values:

Option	Value
type	nfs

Option	Value
o	addr= <i>nfs-server-ip</i>,rw
device	:/path-to-folder-from-step-iii/\${node_ip}

With this configuration, each instance of the Database service stores its logs in a different folder on your NFS server.

Deploying the system using CLI commands

As an alternative to using the service deployment wizard, you can use CLI commands to deploy service instances onto all instances of the system.

Before you begin

These procedures require local access or the ability to establish an SSH session to the system.

To deploy the HCP for cloud scale system:

Procedure

1. Log in to an HCP for cloud scale instance.
2. Go to the `install_path/cli/admin` folder.
`cd /opt/hcpcs/cli/admin`
3. Use the command `setupAdminUser` to set the password for the main **admin** account:
`./admincli -k false -c setupAdminUser --pm-password password`



Important: Do not lose or forget this password.

4. Use the command `editSecuritySettings` to set the cluster host name.
`./admincli -c editSecuritySettings --ssm-cluster-hostname=cluster_name -u admin -p password`
Type a lowercase ASCII FQDN.
Omitting this step can cause links in the System Management application to function incorrectly.
5. Use the command `queryServices` to display the default configuration values, and save the output to a file:
`./admincli -c queryServices --sqrms-is-recommend true --sqrms-requested-details serviceInstances, config --sqrms-service-types product -u admin -p password > /file_path/config_filename.txt`
An example of a configuration file location and name is `/tmp/default_config.txt`.
6. Optional: If needed, use a text editor to modify the configuration file `config_filename.txt`.
7. Use the command `updateServiceConfig` to start deployment using the values in the configuration file:

```
./admincli -c updateServiceConfig --service-update-model /
file_path/config_filename.txt -u admin -p password
```



Note: If a port is already in use this step fails and an error message is displayed listing the ports in use. Edit the configuration file to change the port and repeat this step.

- Use the command `listScaleTasks` to monitor the progress of deployment until all services are deployed ("status" is "Complete"):

```
./admincli -c listScaleTasks -u admin -p password
```



Tip: You can focus on the status messages with a command such as this:

```
./admincli -c listScaleTasks -u admin -p password | grep
status
```



Note: If this step fails, log in to the HCP for cloud scale system using a browser; the service deployment wizard is displayed. Click **Retry**.

- Use the command `setupComplete` to finalize deployment:

```
./admincli -c setupComplete -u admin -p password
```



Note: If this step fails with the message `Must be in state "setup"` to complete setup, wait for a few seconds and repeat this step.

Create an owner for new files

After installation, create a user as owner of the newly installed files.

The files installed for HCP for cloud scale are created with an owner universally unique ID (UUID) of 10001. It's best for all files to have a valid owner, so you should create a user account (such as `hcpcs`) with a UUID of 10001 to own the files.



Caution: Do not try to change the file owner to the UUID of an existing user.

To create a file owner:

Procedure

- Create the user account by typing the command `sudo useradd -u 10001 account` where `account` is the name of the user account (for example, `hcpcs`).
- Verify the user account by typing the command `id -u account`. The system displays the user account UUID.
- Add a password to the user account by typing the command `sudo passwd account`. It's best to use a strong password.
 - When prompted, type the user account password.
 - When prompted, confirm the user account password.

Result

You have created a user account that owns the HCP for cloud scale files.

Optional: Verify the created volumes

Before you begin

If you configured the service volumes to use volume drivers, use these commands to list and view the Docker volumes created on all instances in the system:

```
docker volume ls
```

```
docker volume inspect volume_name
```

If volumes were created incorrectly, you need to redo the system installation:

Procedure

1. Stop the `run` script from running. You do this using whatever method you're currently using to run the script.
2. Stop all HCP for cloud scale Docker containers on the instance:

```
sudo install_path/hcpcs/bin/stop
```
3. Delete the contents of the folder `install_path/hcpcs` from all instances.
4. Delete any Docker volumes created during the installation:

```
docker volume rm volume_name
```
5. Begin the installation again from the point where you unpack the installation package.

Optional: Distribute services among system instances

By default, when you install and deploy a multi-instance system, the system automatically runs each service (except Dashboard) on its normal number of instances.

However, if you've installed more than four instances, some instances may not be running any services at all. As a result, these instances are under-used. You should manually distribute services to run across all instances in your system.

Moving and scaling floating services

For floating services, instead of specifying the specific instances on which the service runs, you can specify a pool of eligible instances, any of which can run the service.

Moving and scaling services with multiple types

When moving or scaling a service that has multiple types, you can simultaneously configure separate rebalancing for each type.

Best practices

Here are some guidelines for distributing services across instances:

- Avoid running multiple services with high service unit costs together on the same instance.
- On master instances, avoid running any services besides those classified as System services.

Considerations

- Instance requirements vary from service to service. Each service defines the minimum and maximum number of instances on which it can run.
- You cannot remove a service from an instance if doing so causes or risks causing data loss.
- Service relocation might take a long time to complete and can impact system performance.

Troubleshooting

You might encounter these issues during installation.

Service doesn't start

Rarely, a system deployment, service management action, or system update fails because a service fails to start. When this happens, the System Management application is inaccessible from the instance where the failure occurred.

The logs in the `watchdog-service` log folder contain this error:

```
Error response from daemon: Conflict. The name "service-name" is
already in use by container Docker-container-id. You have to remove
(or rename) that container to be able to reuse that name.
```

To resolve this issue, restart the Docker service on the instance where the service failed to start. For example, if you are using `systemd` to run Docker, run:

```
systemctl restart docker
```

After restarting Docker, try the system deployment, service management action, or system update again.

Relocating services

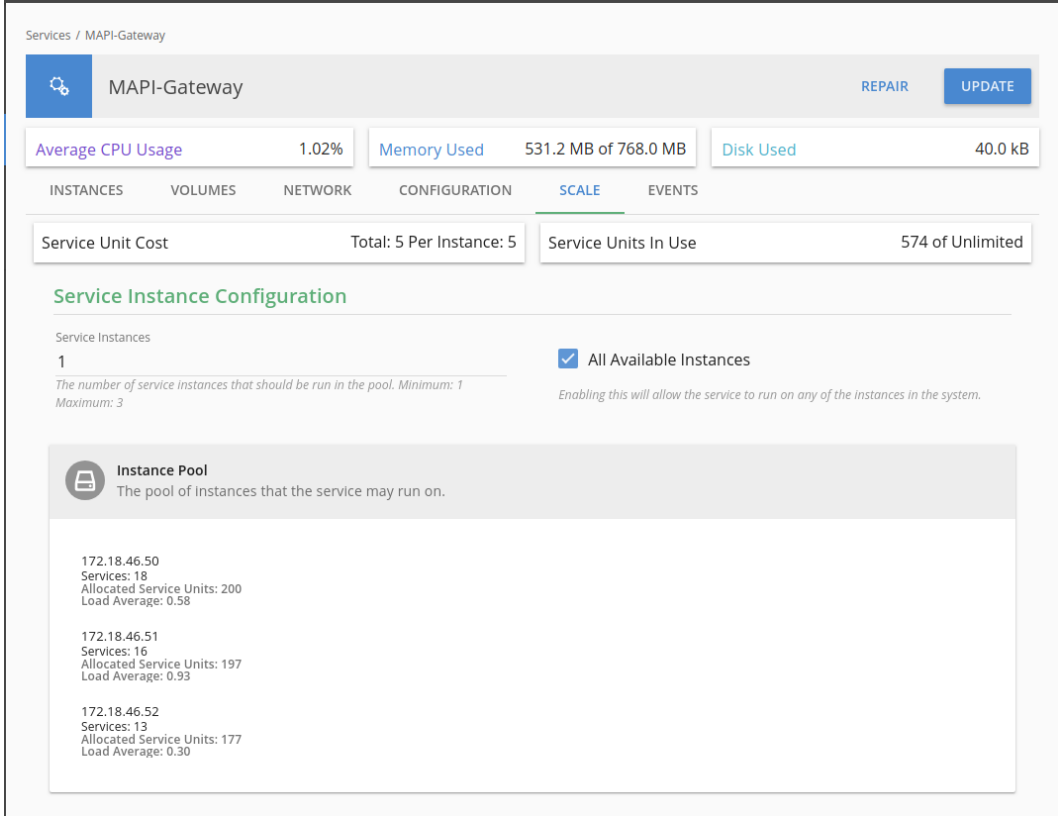
To manually relocate a service, in the Admin App:

Procedure

1. Select **Services**.
The **Services** page opens, displaying the services and system services.
2. Select the service that you want to scale or move.
Configuration information for the service is displayed.
3. Click **Scale**, and if the service has more than one type, select the instance type that you want to scale.

The next step depends on whether the service is floating or persistent (non-floating).

4. If the service is a floating service, you are presented with options for configuring an instance pool. For example:



Services / MAPI-Gateway

MAPI-Gateway REPAIR UPDATE

Average CPU Usage 1.02% Memory Used 531.2 MB of 768.0 MB Disk Used 40.0 kB

INSTANCES VOLUMES NETWORK CONFIGURATION **SCALE** EVENTS

Service Unit Cost Total: 5 Per Instance: 5 Service Units In Use 574 of Unlimited

Service Instance Configuration

Service Instances

1 All Available Instances

The number of service instances that should be run in the pool. Minimum: 1 Maximum: 3 Enabling this will allow the service to run on any of the instances in the system.

Instance Pool

The pool of instances that the service may run on.

- 172.18.46.50
Services: 18
Allocated Service Units: 200
Load Average: 0.58
- 172.18.46.51
Services: 16
Allocated Service Units: 197
Load Average: 0.93
- 172.18.46.52
Services: 13
Allocated Service Units: 177
Load Average: 0.30

- a. In the box **Service Instances**, specify the number of instances on which the service should be running at any time.
- b. Configure the instance pool:
 - For the service to run on any instance in the system, select **All Available Instances**.
With this option, the service can be restarted on any instance in the instance pool, including instances that were added to the system after the service was configured.
 - For the service to run on a specific set of instances, clear **All Available Instances**. Then:
 - To remove an instance from the pool, select it from the list **Instance Pool**, on the left, and then click **Remove Instances**.
 - To add an instance to the pool, select it from the list **Available Instances**, on the right, and then click **Add Instances**.
5. If the service is a persistent (non-floating) service, you are presented with options for selecting the specific instances that the service should run on. Do one or both of these,

then click **Next**:

- To remove the service from the instances it's currently on, select one or more instances from the list **Selected Instances**, on the left, and then click **Remove Instances**.
- To add the service to other instances, select one or more instances from the list **Available Instances**, on the right, and then click **Add Instances**.

6. Click **Update**.

The **Processes** page opens, and the **Service Operations** tab displays the progress of the service update as "Running." When the update finishes, the service shows "Complete."

Next steps

After reconfiguration, the service might take a few minutes to appear on the **Services** page.

Configure the system for your users

After your system is up and running, you can begin configuring it for your users.

For information about these procedures, see the *Administration Guide* or the online help that's available from the HCP for cloud scale application.

The overview of tasks is:

Procedure

1. Configure the connection to an IdP and create user accounts.

2. Define storage components.
3. Assign a name for your HCP for cloud scale cluster.
The host name is required for access to the System Management application and the S3 API.
4. Configure DNS servers to resolve both the fully qualified domain name for your cluster and the wildcard **.hcpcs_cluster_name*.
5. Update Secure Socket Layer (SSL) certificates for the system, storage components, or synchronized buckets.
6. If your system uses encryption, enable it.
7. Obtain S3 authorization credentials.

Chapter 4: Updating HCP for cloud scale

The following procedures describe how to update the HCP for cloud scale software.

Updates are managed by the System Management (Admin) application. Instances are shut down, updated, and restarted one at a time automatically, so you can update the HCP for cloud scale software to a newer version without interrupting availability or reingesting data. S3 API methods remain available, so that users can continue to read and write data and create and configure buckets.

During an update, management API methods that don't change the configuration remain available, so you can continue to create, monitor, and manage storage components. Tracing and the collection of metrics aren't affected.

When updating from a version of HCP for cloud scale prior to 2.3 to a version that implements consistent listing, event messages appear in the Object Storage Management app to inform you of the migration progress: when it has started, when it is in progress, when cleanup has started, and when it has completed. Once completed, the alerts clear.

Restrictions

You cannot downgrade HCP for cloud scale to a previous version.

You cannot upgrade to v1.3.0 from any previous version.

During an update, you cannot make changes to the configuration. After an update, you might need to reconfigure services.

Items and information you need

To update an HCP for cloud scale system, you need the appropriate update archive file `hcpcs-version_number.update` and, for verification purposes, its MD5 checksum file `hcpcs-version_number.md5`.

This document shows the path to the HCP for cloud scale folder as *install_path*. The best folder path is `/opt`.

This document shows the HCP for cloud scale folder as *product*. The best folder name is `hcpcs`.

Verify and place the update archive

On the server or virtual machine from which you want to start the update:

Procedure

1. Download the update archive and MD5 checksum file and store the files in a folder on the server or virtual machine.
2. The best practice is to verify the integrity of the update archive. For example:

```
md5sum -c hcpcs-version_number.update.md5
```

If the archive integrity is verified, the command displays *product-version_number.update: OK*.
3. In the largest disk partition on the server or virtual machine, create a product update folder. For example:

```
mkdir /opt/hcpcs
```
4. Move the update archive from the folder where you stored it to the product update folder. For example:

```
mv hcpcs-version_number.update /opt/hcpcs/hcpcs-version_number.update
```
5. Navigate to the update folder. For example:

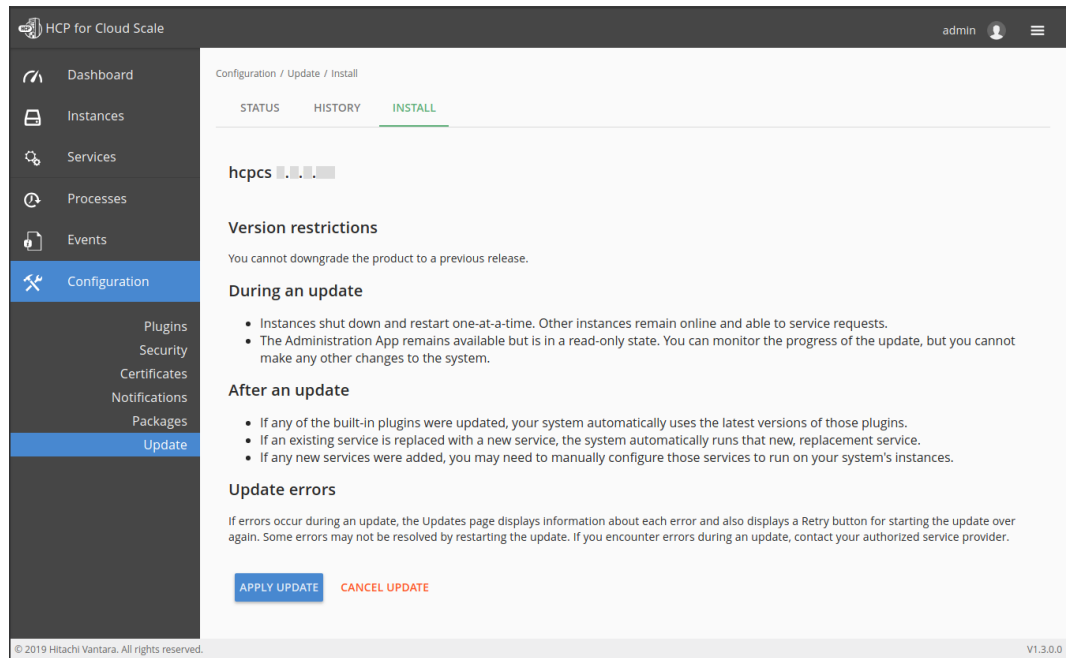
```
cd /opt/hcpcs
```

Use the System Management application to update the system

The update process is controlled by the System Management application (the Admin App).

Procedure

1. From the System Management application, select **Configuration**.
The **Configuration** page opens.
2. Click **Update**.
The **Update** page opens.
3. Select the **Install** tab.
The **Upload** area opens.
4. Click and drag the update file into the **Upload** area, or click **Click to Upload**, select the update archive, and click **Open**.
The update archive is uploaded and verified. These processes will take several minutes. When the archive is verified, the page displays information about the update process. If the update contains a new service, you can optionally configure network, volume, or log file settings.



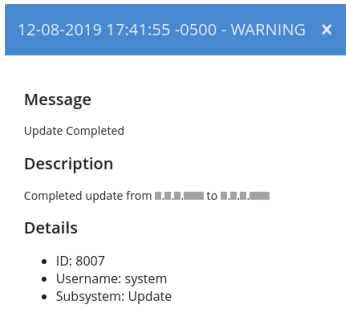
- When you're ready to begin, click **Apply Update**. The System Management application displays the message `Successfully started update package installation` and then applies the update to the cluster (each server or virtual machine in the HCP for cloud scale system).

Note: The update process can trigger alarms on the **Status** tab, or emails if notification is enabled, that services have gone down, exited abnormally, or become underprotected. These notifications are expected during an update and you can safely ignore them. Update alarms should clear automatically when the update finishes.

Result

When the system is updated:

- The `Successfully started` banner message is removed.
- In the **Status** tab, the **Update Status** displays `None` and the event `Update Completed` appears.



- Update-related alarms are cleared from the **Status** tab.
- The update version appears in the lower right corner of the **Login** page.



Note: If the update process encounters an unexpected error, update stops and the **Install** tab displays the message `The update has failed!` Click `retry` to attempt the update again. On the **Status** tab, click on the error message to get more details about the error. Click **Retry**, or click **Cancel**, correct the problem, and then restart the update. Some errors might not be resolvable by restarting the update. If a retry doesn't work, contact Support.

Configure the system for your users

After your system is up and running, you can begin configuring it for your users.

For information about these procedures, see the *Administration Guide* or the online help that's available from the HCP for cloud scale application.

The overview of tasks is:

Procedure

1. Configure the connection to an IdP and create user accounts.
2. Define storage components.
3. Assign a name for your HCP for cloud scale cluster.
The host name is required for access to the System Management application and the S3 API.
4. Configure DNS servers to resolve both the fully qualified domain name for your cluster and the wildcard `*.hcpcs_cluster_name`.
5. Update Secure Socket Layer (SSL) certificates for the system, storage components, or synchronized buckets.
6. If your system uses encryption, enable it.
7. Obtain S3 authorization credentials.

Troubleshooting

You might encounter these issues during an update.

Service doesn't start

Rarely, a system deployment, service management action, or system update fails because a service fails to start. When this happens, the System Management application is inaccessible from the instance where the failure occurred.

The logs in the `watchdog-service` log folder contain this error:

```
Error response from daemon: Conflict. The name "service-name" is
already in use by container Docker-container-id. You have to remove
(or rename) that container to be able to reuse that name.
```

To resolve this issue, restart the Docker service on the instance where the service failed to start. For example, if you are using `systemd` to run Docker, run:

```
systemctl restart docker
```

After restarting Docker, try the system deployment, service management action, or system update again.

PUT and GET calls on encrypted objects fail

After you turn on encryption, PUTS and GETS of objects require the key management server (KMS) to be up and unsealed. During an update the KMS can restart multiple times, including when the master nodes and services are upgraded and when the Vault service is updated. When the KMS service restarts, it is sealed, which can interrupt service.

If you are using encryption, monitor the Vault service closely during an update to prevent interruptions. Whenever the service restarts and gets sealed, unseal it.

If you have access to the Aspen administration app, you can monitor the health of the KMS by checking for the alert "Failed to connect to KMS server." When you see this alert, you know that the KMS is either down or sealed.

Another approach is to load the KMS page, which is at port 8200 of the system. The status of the KMS is displayed in the upper right corner. A red dot indicates that it is sealed.

Appendix A: Logs and diagnostic information

Each service maintains its own set of logs. By default, log files are maintained in the folder `install_path/hcps/log` on each instance in the system. During installation, you can configure each service to store its logs in a different (that is, non-default) location.

Log levels

The following table lists the available log levels.



Note: Raising the log level (for example, from WARN to INFO) results in writing more data to the log file, but the file size increases more rapidly. Lowering the log level (for example, from WARN to ERROR) results in the file size increasing more slowly, but results in writing less data to the log file.

Level	Levels included
ALL	FATAL, ERROR, WARN, INFO, DEBUG, TRACE
TRACE	FATAL, ERROR, WARN, INFO, DEBUG, TRACE
DEBUG	FATAL, ERROR, WARN, INFO, DEBUG
INFO	FATAL, ERROR, WARN, INFO
WARN	FATAL, ERROR, WARN (default)
ERROR	FATAL, ERROR
FATAL	FATAL
OFF	None

Log management

You can manage any of the log files yourself. That is, you can delete or archive them as necessary.



Caution: Deleting log files can make it more difficult for support personnel to resolve issues you might encounter.

System logs are managed automatically in these ways:

- **Retirement:** All log files are periodically added to a compressed file and moved to `install_path/hcpcs/retired/`. This occurs at least once a day, but can also occur:
 - Whenever you run the `log_download` script.
 - Hourly, if the system instance's disk space is more than 60% full.
 - At the optimum time for a specific service.
- **Rotation:** When a log file grows larger than 10MB in size, the system stops writing to that file, renames it, and begins writing to a new file. For example, if the file `exampleService.log.0` grows to 10 MB, it is renamed to `exampleService.log.1` and the system creates a new file named `exampleService.log.0` to write to.
- **Removal:** When a log file becomes older than 90 days, it is removed. If the system instance's disk space is more than 70% full, log files are deleted when they become older than one day.
- When an optimum number of log files for a specific service is reached, the system can overwrite the oldest file. For example, if a service is limited to 20 log files, when the file `exampleService.log.19` is filled, the system overwrites the file named `exampleService.log.0`.

Retrieving logs and diagnostic information

The tool `log_download` lets you easily retrieve logs and diagnostic information from all instances in the system. This tool is located at this path on each instance:

```
install_path/hcpcs/bin/log_download
```

For information about running the tool, use this command:

```
install_path/hcpcs/bin/log_download -h
```



Note:

- When using the tool `log_download`, if you specify the option `--output`, do not specify an output path that contains colons, spaces, or symbolic links. If you omit the option `--output`, you cannot run the script from within a folder path that contains colons, spaces, or symbolic links.
- When you run the script `log_download`, all log files are automatically compressed and moved to the folder `install_path/hcpcs/retired/`.
- If an instance is down, you need to specify the option `--offline` to collect the logs from that instance. If your whole system is down, you need to run the script `log_download` with the option `--offline` on each instance.

Default log locations

Default log locations

By default, each service stores its logs on each instance on which the service instance runs, in its own folder at this path:

```
install_path/hcpcs/log
```

This table shows the default log folder names for each service. Depending on how your system was configured when first deployed, your system's logs might not be stored in these folders.

For information about services, see [HCP for cloud scale services \(on page 79\)](#).

Service	Default log folder name	Contains information about
Admin-App	com.hds.ensemble.plugins.service.adminApp	The System Management application.
Database	com.hds.ensemble.plugins.service.cassandra	<ul style="list-style-type: none"> ▪ System configuration data. ▪ Document fields and values.
Scheduling	com.hds.ensemble.plugins.service.chronos	Workflow task scheduling.
N/A	com.hds.ensemble.plugins.service.containerAction	Created by custom actions run by service plugins.
Metrics	com.hds.ensemble.plugins.service.elasticsearch	The storage and indexing of: <ul style="list-style-type: none"> ▪ System events ▪ Performance and failure metrics for workflow tasks

Service	Default log folder name	Contains information about
Network-Proxy	com.hds.ensemble.plugins.service.haproxy	Network requests between instances.
Message Queue	com.hds.ensemble.plugins.service.kafka	The transmission of data between instances.
Logging	com.hds.ensemble.plugins.service.logstash	The transport of system events and workflow task metrics to the Metrics service.
Service-Deployment	com.hds.ensemble.plugins.service.marathon	The deployment of high-level services across system instances. High-level services are the ones that you can move and configure, not the services grouped under System Services.
Cluster-Worker	com.hds.ensemble.plugins.service.mesosAgent	The work ordered by the Cluster-Coordination service.
Cluster-Coordination	com.hds.ensemble.plugins.service.mesosMaster	Hardware resource allocation.
Watchdog	com.hds.ensemble.plugins.service.remoteAction	Internal system processes.

Service	Default log folder name	Contains information about
Sentinel	com.hds.ensemble.plugins.service.sentinel	The internal system processes.
Watchdog	com.hds.ensemble.plugins.service.watchdog	General diagnostic information.
Synchronization	com.hds.ensemble.plugins.service.zookeeper	The coordination of actions and database activities across instances.
S3-Gateway	com.hitachi.aspen.foundry.service.clientaccess.data	The client access data service.
Data-Lifecycle	com.hitachi.aspen.foundry.service.data-lifecycle.service	The data lifecycle service.
Tracing-Agent	com.hitachi.aspen.foundry.service.jaeger.agent	The tracing agent service.
Tracing-Collector	com.hitachi.aspen.foundry.service.jaeger.collector	The tracing collector service.
Tracing-Query	com.hitachi.aspen.foundry.service.jaeger.query	The tracing query service.
MAPI-Gateway	com.hitachi.aspen.foundry.service.mapi.gateway	The management API gateway.
Metadata-Policy-Engine	com.hitachi.aspen.foundry.service.metadata.async.policy.engine	The metadata asynchronous policy engine.
Grafana	com.hitachi.aspen.foundry.service.metrics.grafana	The dashboard service.
Mirror-Out-Policy	com.hitachi.aspen.foundry.service.policy.mirror.out	The mirror out (synch-to) service.

Service	Default log folder name	Contains information about
Mirror-In-Policy	com.hitachi.aspen.foundry.service.policy.mirror.in	The mirror in (synch-from) service.
S3-Notifications	com.hitachi.aspen.foundry.service.policy.s3.notifications	The S3 notifications service.
Metadata-Cache	com.hitachi.aspen.foundry.service.metadata.cache	The metadata cache. Note: This service is deprecated.
Metadata-Coordination	com.hitachi.aspen.foundry.service.metadata.coordination	Metadata coordination.
Metadata-Gateway	com.hitachi.aspen.foundry.service.metadata.gateway	The metadata gateway.
Telemetry-Service	com.hitachi.aspen.foundry.service.metrics.prometheus	Telemetry.
Message-Queue	com.hitachi.aspen.foundry.service.rabbitmq.server	The message broker.
Key-Management-Server	com.hitachi.aspen.foundry.service.vault.vault	The key management server.

Appendix B: Services list

This module describes the HCP for cloud scale services and how to configure them.

HCP for cloud scale services

The following table describes the services that HCP for cloud scale runs. Each service runs within its own Docker container. For each service, the table lists:

- **Configuration settings:** The settings you can configure for the service.
- **RAM needed per instance:** The amount of RAM that, by default, the service needs on each instance on which it's deployed. For all services except for System services, this value is also the default Docker value of **Container Memory** for the service.
- **Number of instances:** Shows both:
 - The minimum number of instances on which a service must run to function properly.
 - The best number of instances on which a service should run. If the system includes more than the minimum number of instances, you should take advantage of the instances by running services on them.



Note: Unused services do not need to be scaled.

- **Service unit cost:** For HCP for cloud scale, you can safely ignore these values.
- Whether the service is **stateful** (that is, it saves data permanently to disk) or **stateless** (that is, it does not save data to disk).
- Whether the service is **persistent** (that is, it must run on a specific instance) or supports **floating** (that is, it can run on any instance).
- Whether the service is **scalable** or not.



Note: For HCP for cloud scale services, you cannot set the size of **Max Heap Size** larger than the value of the setting **Container Memory**. For other services, you should not set the size of **Max Heap Size** larger than the value of the setting **Container Memory**.

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
Product services: These services perform HCP for cloud scale functions. You can move and reconfigure these services.		

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Cassandra</p> <p>Decentralized database, used to stores some configuration and system update packages</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2400.0 (2.4 GB). ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1200m. ▪ Heap new size: Heap size for the Java virtual machine. Valid values are integers representing a number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 512 MB. <p>Advanced Options</p> <p>Compaction Frequency: How often the database is compacted. The options are Weekly (default) and Daily.</p> <p>Caution: Changing this setting can negatively affect the service. Use with caution.</p>	<p>RAM needed per instance: 2.4 GB</p> <p>Number of instances: minimum 3, best All</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Chronos Job scheduler</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 712 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Heap settings: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 356 MB. 	<p>RAM needed per instance: 712 MB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 1</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>
<p>Data Lifecycle Processes lifecycle policies</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 4 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 1; best depends on system load, number of active client objects, and daily rate of object deletion</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 2 GB. 	
<p>Elasticsearch</p> <p>Indexes metrics and event logs</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2000.0 (2 GB); a good initial value is 10 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 10 GB</p> <p>Number of instances: minimum 3, best All</p> <p>Service unit cost: 25</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>


Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Heap settings: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1024 MB; a good initial value is 8 GB. ▪ Days to keep logs: The number of days to keep service logs, including access and metrics indexes. The default is 30 days. ▪ Index Protection Level: The number of additional replicas (copies) to keep of each index file (shard). Replicas are kept on separate instances. You can set this value for every shard. The default is 1 replica (which means that two copies are kept). The maximum is the number of instances less one. 	
<p>Grafana</p> <p>Collects data and displays dashboard metrics</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 768 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 768 MB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Grafana Scrape Interval: How often the service collects data. The default is 20 seconds. ▪ Grafana Database Path: The location of the local time-series database. The default is <code>grafana_db_path</code>. ▪ Grafana Database Retention: How long to retain data. The default is 15 days. 	
<p>Kafka</p> <p>Handles metrics and event logs</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2000.0 (2 GB). ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Heap settings: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1 GB. 	<p>RAM needed per instance: 2 GB</p> <p>Number of instances: minimum 3, best All</p> <p>Service unit cost: 5</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>
<p>Key Management Server</p> <p>Manages storage component encryption keys</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2000.0 (2 GB). ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.5. 	<p>RAM needed per instance: 2 GB</p> <p>Number of instances: minimum 1, best 2 or more</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <p>None.</p>	Scalable? Yes
<p>Logstash</p> <p>Handles metrics and event logs</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 700 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 350 MB. 	<p>RAM needed per instance: 700 MB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>
<p>MAPI Gateway</p> <p>Serves MAPI endpoints</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 2 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 2 GB</p> <p>Number of instances: minimum 1, max 1</p> <p>Service unit cost: 5</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 1 GB. ▪ Total Storage Capacity Alert Threshold: Display an alert when the total storage capacity free to store object data in the system goes below this value. Type a threshold value. You must specify the suffix % (percent of total), K (kilobytes), M (megabytes), G (gigabytes), or T (terabytes). If blank, alerts are disabled. The default is 30%. ▪ Per Storage Component Capacity Alert Threshold: Display an alert when the storage capacity free to store object data in any storage component goes below this value. Type a threshold value. You must specify the suffix % (percent of total), K (kilobytes), M (megabytes), G (gigabytes), or T (terabytes) If blank, alerts are disabled. The default is 250GB. 	
<p>Message Queue</p> <p>Coordinates and distributes messages to other services</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2048 MB; a good initial value is 8 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.5. 	<p>RAM needed per instance: 8 GB</p> <p>Number of instances: minimum 3, best 3</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Consumer Timeout: The time in milliseconds that the service waits before timing out an unacknowledged message, closing the consumer's channel, and returning the message to the queue. Type an integer number of milliseconds. The default is 172800000 ms (48 hours). If messages (for example, for synchronization of large files) take longer than that to be acknowledged by the consumer, increase this value. 	
<p>Metadata Cache</p> <p>Cache for HCP for cloud scale metadata</p> <p>Note: This service is deprecated but cannot be removed.</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 1024 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB. 	<p>RAM needed per instance: 1024 MB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Metadata Coordination</p> <p>Coordinates Metadata Gateway service instances and coordinates scaling and balancing of metadata partitions</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 4 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 2 GB. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 5</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p> <div style="border: 1px solid black; background-color: #e0f2f1; padding: 5px; margin-top: 10px;"> <p> Important: IMPORTANT: Your partition count must be kept under 1500 partitions per node. Once you reach this limit, system performance may be affected. If you are encountering issues, contact your Hitachi Vantara representative for more assistance.</p> </div>
<p>Metadata Gateway</p> <p>Stores and protects metadata and serves it to other services</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 4096 MB; a good initial value is 64 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 64 GB</p> <p>Number of instances: minimum 3, best All</p> <p>Service unit cost: 50</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 32 GB. 	
<p>Metrics</p> <p>Gathers metrics from all services and instances and supplies them to GUI and API</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 6 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Prometheus Scrape Interval: The time interval between runs of the metrics collection task. Type an integer number of seconds. You can optionally specify the suffix s (seconds). The default is 10 seconds. ▪ Prometheus Database Path: Storage location for prometheus local time-series db. Type a path. The default is <code>tsdb/</code>. ▪ Prometheus Database Retention: The number of days to retain files. Type an integer number of days. You can optionally specify the suffix d (days). The default is 15 days. 	<p>RAM needed per instance: 6 GB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateful</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Mirror In</p> <p>Executes synchronic policies</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 8 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 384 MB. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>
<p>Mirror Out</p> <p>Executes system synchronic policies</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 8 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 384 MB. 	
<p>Policy Engine</p> <p>Executes system policies</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 2048 MB; a good initial value is 4 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1024 MB; a good initial value is 2 GB. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 3, best 3</p> <p>Service unit cost: 25</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>S3 Gateway</p> <p>Serves S3 API methods and communicates with storage components</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 16 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 8 GB. <p>HTTP Options</p> <ul style="list-style-type: none"> ▪ Enable HTTP: Select to enable HTTP connections. ▪ Max Http Request Headers: The maximum number of HTTP request headers to allow. Type an integer. The default is 100 request headers. <p>HTTPS Options</p> <ul style="list-style-type: none"> ▪ SSL Ciphers: A comma-separated list of ciphers used to encode SSL traffic. Changing the list causes the service to redeploy. 	<p>RAM needed per instance: 16 GB</p> <p>Number of instances: minimum 1, best All</p> <p>Service unit cost: 25</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>S3 Notifications</p> <p>Executes S3 notifications</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB; a good initial value is 8 GB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Max Heap Size: Maximum amount of memory to allocate to the Java heap for each instance of the service. It's best to set this value to half the size of the container memory. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 384 MB; a good initial value is 384 MB. 	<p>RAM needed per instance: 4 GB</p> <p>Number of instances: minimum 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes (but not recommended on master instances)</p>
<p>Tracing Agent</p> <p>Listens for incoming tracing of S3 API and MAPI calls, batches them, and sends them to Tracing Collector service</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ Collector TChannel Hostname: Type a host name. The default is localhost. ▪ Collector TChannel Port: Type a port number. The default is 14267. 	<p>RAM needed per instance: 2 GB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 1</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Tracing Collector</p> <p>Collects traces from Tracing Agent service instances and stores them in tracing database</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. <p>Service Options</p> <ul style="list-style-type: none"> ▪ ElasticSearch Hostname: Type a host name. The default is localhost. ▪ ElasticSearch Port: Type a port number. The default is 9200. ▪ Sampling Rate: The sampling rate for all clients implementing remote sampling. Type a number between 0 and 1 inclusive. The default is 1. ▪ Max open index age: How long to keep tracing indexes open in the database, in days. Type a value from 1 to 365 days inclusive. The default is 30 days. ▪ Max index age: How long to keep tracing indexes in the database, in days. Type a value from 1 to 365 days inclusive. The default is 60 days. 	<p>RAM needed per instance: 8 GB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 10</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p> <p>Scalable? Yes</p>
<p>Tracing Query</p> <p>UI and API endpoint access for distributed tracing for S3 API and MAPI calls</p>	<p>Container Options: Default</p> <ul style="list-style-type: none"> ▪ Container Memory: The hard memory limit for the service's Docker container, in MB. The default is 768 MB. ▪ CPU: The relative CPU usage weight for the service's Docker container. Generally, a higher value means that the container receives more CPU resources than other processes (including other service Docker containers) running on the instance. Type a decimal number. The default is 0.1. 	<p>RAM needed per instance: 768 MB</p> <p>Number of instances: minimum 1, best 1</p> <p>Service unit cost: 5</p> <p>Stateful or stateless? Stateless</p> <p>Persistent or floating? Floating</p> <p>Supports volume configuration? No</p> <p>Single or multiple types? Single</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	Service Options <ul style="list-style-type: none"> ▪ ElasticSearch Hostname: Type a host name. The default is localhost. ▪ ElasticSearch Port: Type a port number. The default is 9200. 	Scalable? Yes
System services: These services manage system resources and ensure that the HCP for cloud scale system remains available and accessible. These services are persistent and cannot be moved, scaled, or reconfigured.		
Admin App The System Management application	Service Options <ul style="list-style-type: none"> ▪ Max heap size: Maximum amount of memory to allocate to the Java heap for each instance of the service. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 512 MB. 	RAM needed per instance: N/A Number of instances: N/A Persistent or floating? Persistent Supports volume configuration? Yes Single or multiple types? Single Scalable? No
Cluster Coordination Manages hardware resource allocation	None.	RAM needed per instance: N/A Number of instances: N/A Persistent or floating? Persistent Supports volume configuration? No Single or multiple types? Single Scalable? No
Cluster Worker Agent for Cluster Coordination on each instance; reports on resource utilization and availability, deploys services	None.	RAM needed per instance: N/A Number of instances: N/A Service unit cost: 5 Persistent or floating? Persistent Supports volume configuration? Yes Single or multiple types? Single Scalable? No
Network Proxy Network request load balancer	Security Protocol: Select which Transport Layer Security (TLS) versions to use: <ul style="list-style-type: none"> ▪ TLS 1.2 ▪ TLS 1.3 	RAM needed per instance: N/A Number of instances: N/A Service unit cost: 1

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
	<p>SSL Ciphers: To use another cipher suite, type it here.</p> <p>Custom Global Configuration: Select Enable Advanced Global Configuration to enable adding custom parameters to the HAProxy "global" section.</p> <p>Custom Defaults Configuration: Select Enable Defaults Configuration to enable adding custom parameters to the HAProxy "global" section.</p>	<p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>
<p>Sentinel</p> <p>Runs internal system processes and monitors the health of other services</p>	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap size: Maximum amount of memory to allocate to the Java heap for each instance of the service. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 8 GB; a good initial value is 8.6 GB. 	<p>RAM needed per instance: N/A</p> <p>Number of instances: N/A</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>
<p>Service Deployment</p> <p>Handles deployment of high-level services (that is, the services that you can configure)</p>	<p>None.</p>	<p>RAM needed per instance: N/A</p> <p>Number of instances: N/A</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>
<p>Synchronization</p> <p>Coordinates service configuration settings and other information across service instances</p>	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap size: Maximum amount of memory to allocate to the Java heap for each instance of the service. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1024 MB. 	<p>RAM needed per instance: N/A</p> <p>Number of instances: N/A</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>

Service name and description	Configuration settings (changes cause the service to redeploy)	Properties
<p>Watchdog</p> <p>Responsible for initial system startup; monitors other System services and restarts them if necessary</p>	<p>Service Options</p> <ul style="list-style-type: none"> ▪ Max heap size: Maximum amount of memory to allocate to the Java heap for each instance of the service. Type an integer number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 1024 MB. ▪ Heap new size: Heap size for the Java virtual machine. Valid values are integers representing a number of bytes. You can optionally specify the suffixes k (kilobytes), m (megabytes), or g (gigabytes). The default is 512 MB. 	<p>RAM needed per instance: N/A</p> <p>Number of instances: N/A</p> <p>Service unit cost: 5</p> <p>Persistent or floating? Persistent</p> <p>Supports volume configuration? Yes</p> <p>Single or multiple types? Single</p> <p>Scalable? No</p>

Appendix C: Handling network changes

After your system is deployed, its network infrastructure and configuration should not change. Specifically:

- All instance IP addresses should not change. It's best to use static IP addresses.
- All services should continue to use the same ports.
- All services and instances should continue to use the same network types.

If any of these values change, you will need to reinstall the system.

Safely changing an instance IP address

If you need to change the IP addresses for one or more instances in the system, use this procedure to manually change the IP addresses without risk of data loss.



Note: You can reuse the IP addresses of retired nodes for new nodes.

For each instance whose IP address you need to change:

Procedure

1. Move all services off of the instance. Distribute those services among all the other instances.
2. On the instance from step 1, stop the script `run` using whatever tool or process you used to run it.
For example, with `systemd`, run: `systemctl stop hcpcs.service`
3. Remove the instance from the system.
4. Delete the installation folder from the instance.
5. Add the instance back to the system.

After a network change

If a network infrastructure or configuration change occurs that prevents your system from functioning with its current network settings, you need to reinstall all instances in the system.

Procedure

1. If the Admin App is accessible, back up your system components by exporting a package.
2. On each instance in the system:

- a. Navigate to the installation folder.
- b. Stop the run script using whatever tool or process you used to run it. For example, with systemd, run:
`systemctl stop <service-name>`
- c. Run `bin/stop`
- d. Run the setup script, including the list of master instances:

```
sudo bin/setup -i <ip-address-for-this-instance> -m  
    <comma-separated-list-of-master-instance-IP-addresses>
```

- e. Run the run script using whatever methods you usually use to run scripts.
3. Log into Admin App and use the wizard to set up the system.
 4. After the system has been set up, upload your package.

Hitachi Vantara

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA



HitachiVantara.com/contact