HITACHI
Inspire the Next

ARCHITECTURE GUIDE

Deploying Red Hat OpenShift Container Platform on Hitachi Unified Compute Platform CI

Reference Architecture Guide

By Jeff Chen
Contributors: Keisuke Matsumoto, Suresh Mikkilineni (Red Hat)

September 2018

Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@HitachiVantara.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Revision History

Revision	Changes	Date
MK-SL-108-00	Initial release	September 24, 2018

Table of Contents

Deploying Red Hat OpenShift Container Platform on Hitachi Unified Compute Platform CI

Reference Architecture Guide

Introduction

Container applications are gaining popularity and the adoption rate is on the rise. As demand for the container platform increases, Hitachi introduced Hitachi Storage Plug-in for Containers (HSPC) as a storage adapter to make a more reliable container infrastructure. It enables the containers to store and manage persistent data in Hitachi's robust storage systems. HSPC supports the container orchestrators including Kubernetes and Red Hat OpenShift Container Platform.

This reference architecture document describes an example deployment of Red Hat OpenShift Container Platform on Hitachi Unified Compute Platform (UCP) CI. UCP CI is a highly configurable integrated infrastructure in which server, network, and storage can be scaled independently, to optimize performance and eliminate overprovisioning costs. UCP CI architecture consists of the following hardware components:

- Hitachi VSP Gx00/VSP Fx00 for storage
- Hitachi Advanced Server DSxx0 for compute
- Hitachi UCP Advisor for end to end management
- Cisco Nexus 3000 and 9000 for Ethernet networking
- Brocade G620 for Fibre Channel SAN

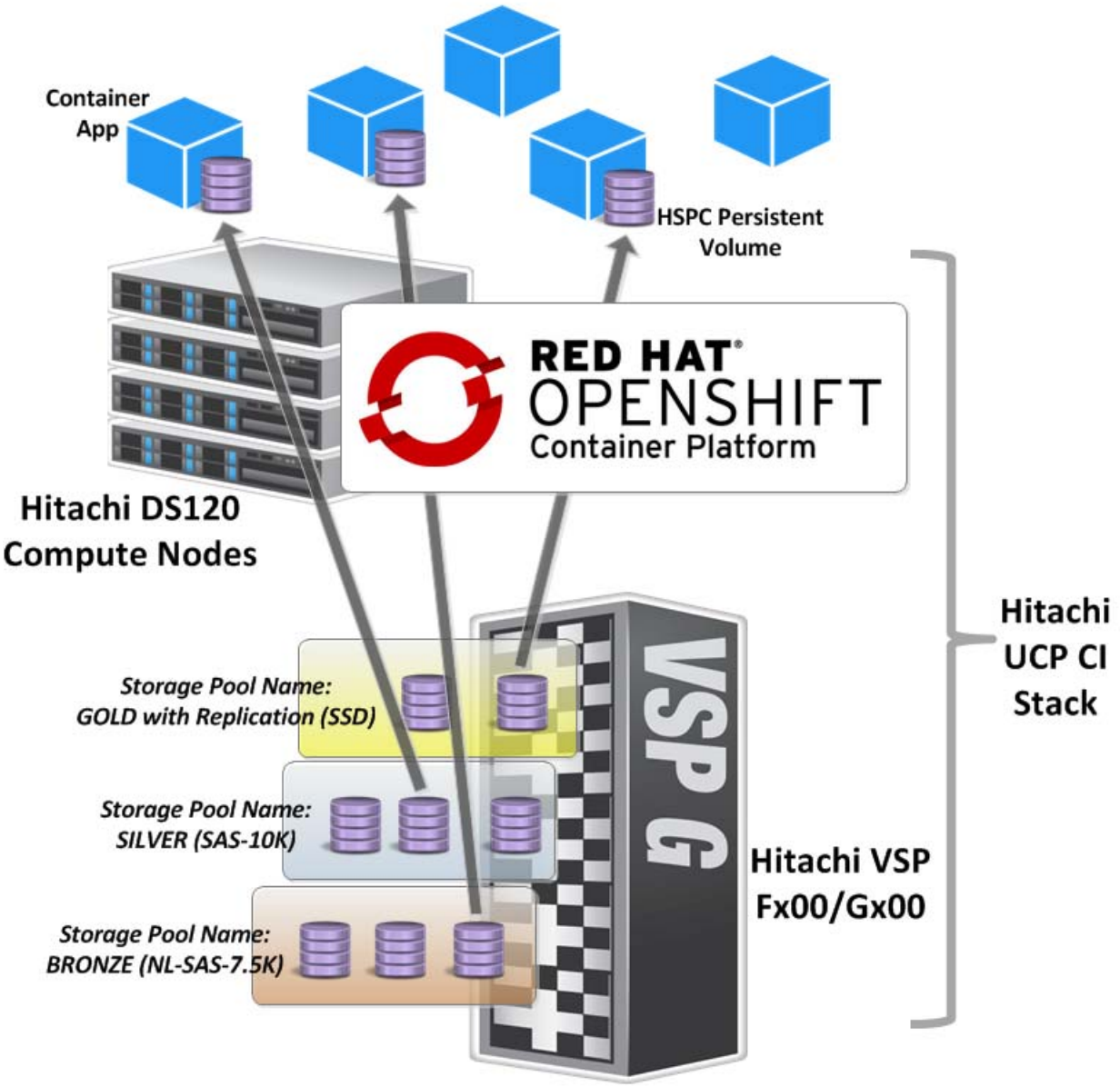
The intended audience of this document is IT administrators, system architects, consultants, and sales engineers to assist in planning, designing, and implementing UCP CI with container solutions.

Solution Overview

Hitachi Unified Compute Platform (UCP) CI is an optimized and preconfigured converged infrastructure platform. UCP CI offers a broad range of compute and storage components that can be scaled and configured independently to eliminate overprovisioning. With UCP CI, you can optimize your data center to run any container application workload, at any scale.

Red Hat OpenShift Container Platform is a powerful container cluster management and orchestration system with natively integrated technologies like Docker and Kubernetes. Providing OpenShift platform is one of the solutions for UCP CI. Figure 1 shows a high-level diagram of OpenShift managing containers and persistent volumes on the UCI CI stack.

Figure 1


Container App
HSPC Persistent Volume
RED HAT® OPENSHIFT Container Platform
Hitachi DS120 Compute Nodes
Storage Pool Name: GOLD with Replication (SSD)
Storage Pool Name: SILVER (SAS-10K)
Storage Pool Name: BRONZE (NL-SAS-7.5K)
VSP G
Hitachi VSP Fx00/Gx00
Hitachi UCP CI Stack

The following explain the example:

- Red Hat OpenShift is installed in Hitachi UCP CI DS120 compute nodes
- VSP Gx00 storage provides different storage pools with different level of capabilities as shown in Table 1
- Container applications require different storage requirements
- Through OpenShift, container administrators create volumes that fit the container application requirements

TABLE 1. VSP GX00 STORAGE POOLS

Storage Pool Name	Disk Type	Usage and Capability
GOLD with Replication	SSD	High performance with high data protection
SILVER	SAS 10K	Mid performance. General use
BRONZE	NL-SAS-7.2K	High capacity and lower performance

Solution Components

This section lists the Unified Compute Platform (UCP) CI configuration used for this reference architecture.

Hardware Components

These are the key hardware components.

TABLE 2. HARDWARE COMPONENTS

Hardware	Description	Version	Quantity
Hitachi Advanced Server DS120 (compute)	■ 2 × Intel Xeon 4110 8-core 2.10GHz v5 processor ■ 8 × 16 GB DIMM, 128 GB memory ■ 32 Gb SATADOM (boot) ■ Emulex LPe31002-M6 16 Gb/sec dual port PCIe HBA ■ Intel X527-DA4 10G quad-port NIC	BMC: 4.22.06 BIOS: 3A10.H7 v11.2.156.27	7
VSP G900	■ Dual controller ■ 512GB cache ■ 8 × 1.9 TB SSD ■ 8 × 1.8 TB SAS 10k ■ 16 × 32 Gbps FC ports	88-03-00-60/00	1

TABLE 2. HARDWARE COMPONENTS (CONTINUED)

Hardware	Description	Version	Quantity
Cisco Nexus 93180LC-EX	▪ 18-port 100 GbE ▪ 24-port 40 GbE	7.0(3)I7(3)	2
Cisco Nexus 93180YC-EX	▪ 48-port 25 GbE ▪ 6-port 100 GbE	7.0(3)I4(7)	2
Cisco Nexus 3048TP	▪ 48-port 1 GbE ▪ 4-port 10 GbE	7.0(3)I4(7)	1
Brocade G620	▪ 48-port 16/32 Gbps Fibre Channel switch	8.2.0b	2

The following hardware components are available for UCP CI configuration.

Hitachi Advanced Server DS120

Hitachi Advanced Server DS120 is a flexible and scalable 2-socket 1U server for converged datacenters.

- Intel Xeon Skylake CPU family
  - 4110 – 8-core 2.1GHz “Silver”
  - 6128/6140 – 6/18-core 3.4/2.3GHz “Gold”
  - 8160/8176M – 24/28-core 2.1GHz “Platinum”
- 24 × DIMM slots (up to 768 GB)
- 1 × SAS mezzanine slot
- 1 × OCP mezzanine slot or PHY card
- 1 × PCIe x8 FHHL slot
- 1 × PCIe x16 LP MD-2 slot
- 32 GB SLC SATADOM Boot Device
- Networking Options:
  - Intel X527-DA4 10G quad-port NIC (default NIC, PHY Mezz, 2 ports cabled by default)

- Fibre Channel Options:
  - Emulex LPe31002 16/32G (PCIe)
  - Emulex LPe32002 32G (PCIe)

Hitachi Advanced Server DS220

Hitachi Advanced Server DS220 is a flexible and scalable 2-socket 2U server for converged datacenters.

- Intel Xeon Skylake CPU family
  - 4110 – 8-core 2.1GHz “Silver”
  - 6128/6140 – 6/18-core 3.4/2.3GHz “Gold”
  - 8160/8176M – 24/28-core 2.1GHz “Platinum”
- 24 × DIMM slots (up to 768 GB)
- 1 × SAS mezzanine slot
- 1 × OCP mezzanine slot or PHY card
- Up to 5 × PCIe Gen3 FHHL slots
- 1 × PCIe LP MD-2 slot
- 32 GB SLC SATADOM Boot Device
- Networking Options:

Intel X527-DA4 10G quad-port NIC (default NIC, PHY Mezz, 2 ports cabled by default)

  - Intel XXV710 10/25G dual-port NIC (PCIe)
- Fibre Channel Options:
  - Emulex LPe31002 16/32G (PCIe)
  - Emulex LPe32002 32G (PCIe)

Hitachi Advanced Server DS240

Hitachi Advanced Server DS240 is a flexible and scalable 4-socket 2U server for converged datacenters.

- Intel Xeon Skylake CPU family
  - 6128/6140 – 6/18-core 3.4/2.3GHz “Gold”
  - 8160/8176M – 24/28-core 2.1GHz “Platinum”
- 48 × DIMM slots (up to 1.5 TB)
- 1 × SAS mezzanine slot
- 1 × OCP mezzanine slot
- Up to 4 × PCIe Gen3 FHHL slots
- 2 × PCIe HHHL slots
- 2 × 480 GB SATA SSD Boot Device (RAID1, LSI SAS3516)
- Networking Options:
  - Intel X527-DA4 10G quad-port NIC (default NIC, PHY Mezz, 2 ports cabled by default)
  - Intel XXV710 10/25G dual-port NIC (PCIe)
- Fibre Channel Options:
  - Emulex LPe31002 16/32G (PCIe)
  - Emulex LPe32002 32G (PCIe)

Hitachi Virtual Storage Platform G Series

Hitachi Virtual Storage Platform G Series is a flash-optimized storage solution for speed, resiliency and direct cloud connect.

- Improve IT agility to innovate and achieve business outcomes faster with leading application and virtualization integration
- Eliminate outages that risk revenue and your reputation with the industry's only 100% data-availability guarantee
- Transform to new levels of availability, automation and agility with a system that enables advanced storage system functions and management
- Provide proven data services for your most critical enterprise data, ensuring data protection, tiering and control
- Automate data movement for the best ROI and performance for your changing workloads
- Simplify the data tiering process to improve application performance, speed administration time and control costs
- Simplify operations and better align storage resources with dynamic business requirements
- Reduce operations management costs and accelerate time to delivery of services

Hitachi Virtual Storage Platform F Series

Hitachi Virtual Storage Platform F Series deliver superior all-flash performance for business-critical applications, with continuous data availability.

- Delivers up to 4.8M IOPS With Sub-Millisecond Response
- Accelerate IOPS and reduce latency with patented express I/O algorithms and flash-optimized processing
- Optimized metadata analysis reduces latency and improves speed by up to 240%.
- Ensures consistent performance that is only activated when needed, so resources are not wasted
- Adaptive SVOS deduplication and compression, and FMD inline, hardware-accelerated compression deliver 5:1 or greater capacity savings
- Transparently moves data to the cloud and dramatically reduces onsite storage costs
- Delivers rapid provisioning of all-flash resources with advanced SVOS RF storage management and software solutions
- Eliminates complexity and potential human error with whiteboard-style, drag-and-deliver data protection
- Optimize IT performance, resolve issues faster and predict resource requirements with advanced infrastructure analytics

Cisco Switches

Cisco Nexus 93180LC-EX is a 1U 14-port 100 GbE (downlink), 28-port 40 GbE (downlink) and six-port 100 GbE (uplink) spine switch for multiple-rack solutions.

Cisco Nexus 93180YC-EX is a 1U 48-port 10/25 GbE (downlink) and 6-port 40/100 GbE (uplink) top-of-rack or leaf switch for single-rack solutions.

Cisco Nexus 3048TP is a 1U 48-port 1 GbE management switch.

Brocade Switches

Brocade G620 is a 1U 48-port 16/32 Gb/sec Fibre Channel switch.

Software Components

These are the key software components.

TABLE 3. SOFTWARE COMPONENTS

Software	Version
Hitachi Storage Virtualization Operating System with Hitachi Dynamic Provisioning	88-03-00-60/00
Hitachi Device Manager - Storage Navigator	88-03-00-60/00
Hitachi Thin Image (volume snapshot and cloning)	88-03-00-60/00
Red Hat Enterprise Linux	7.5
Red Hat OpenShift Container Platform	3.9

Hitachi Storage Virtualization Operating System

Hitachi Storage Virtualization Operating System (SVOS) is the standard operating system for the Hitachi Virtual Storage Platform G and F series storage systems. SVOS delivers the foundation for digital transformation, enabling IT leaders to accelerate business insights and deliver a superior customer experience while realizing greater operational efficiency.

Red Hat Enterprise Linux

Red Hat Enterprise Linux delivers military-grade security, 99.999% uptime, support for business-critical workloads, and so much more. Ultimately, the platform helps you reallocate resources from maintaining the status quo to tackling new challenges.

Red Hat OpenShift Container Platform

- Red Hat OpenShift is a complete container application platform that natively integrates technologies like Docker and Kubernetes—a powerful container cluster management and orchestration system—and includes an enterprise foundation in Red Hat Enterprise Linux.
- Red Hat OpenShift integrates the architecture, processes, platforms, and services needed to empower development and operations teams. It deploys reliably across environments, and it lets you meet customer demand while reducing infrastructure costs.

Solution Design

This is the detailed solution example of a UCP CI configured in a spine-leaf network topology.

Figure 2 shows a spine-leaf network topology UCP CI.

Figure 2

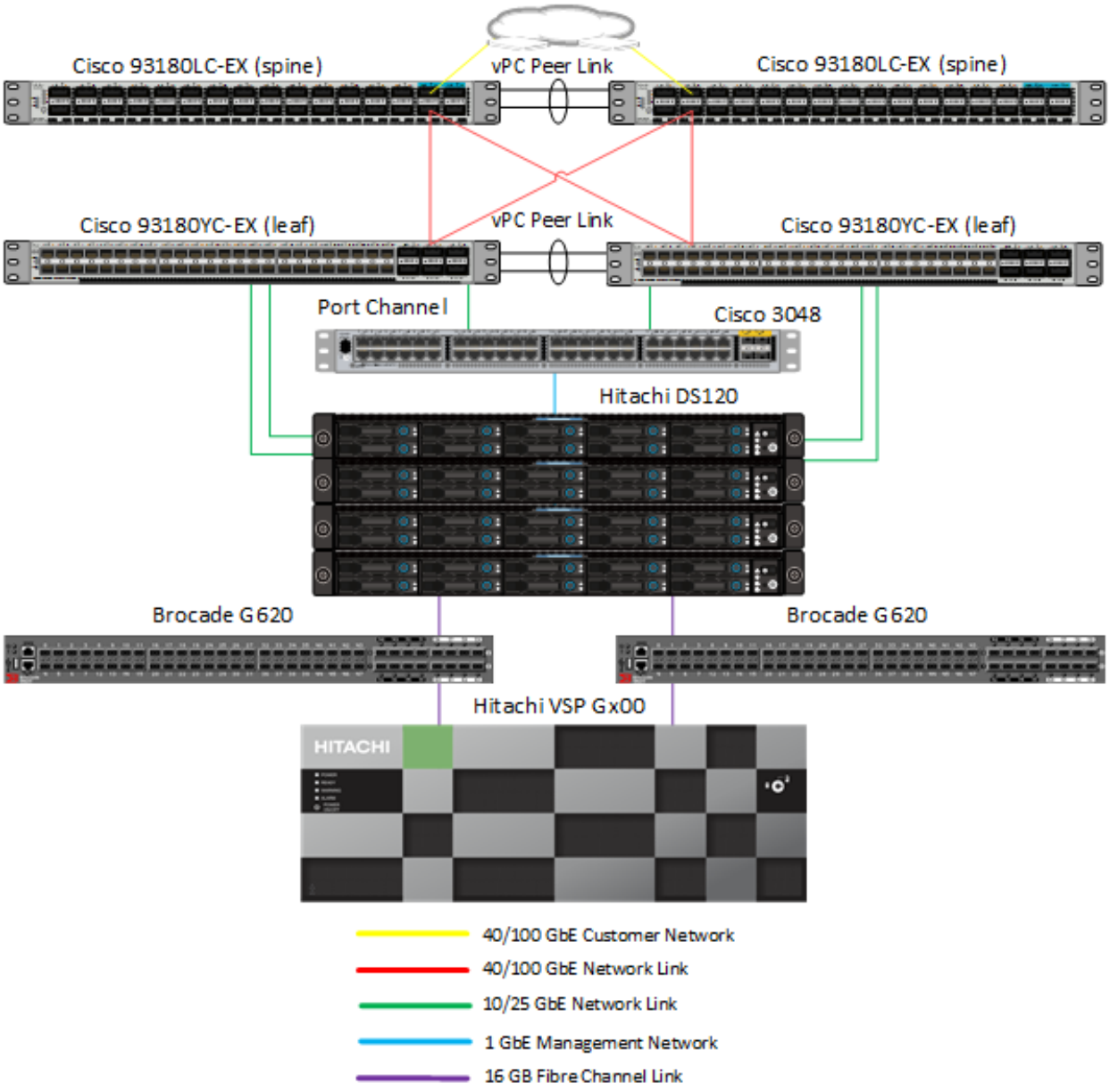
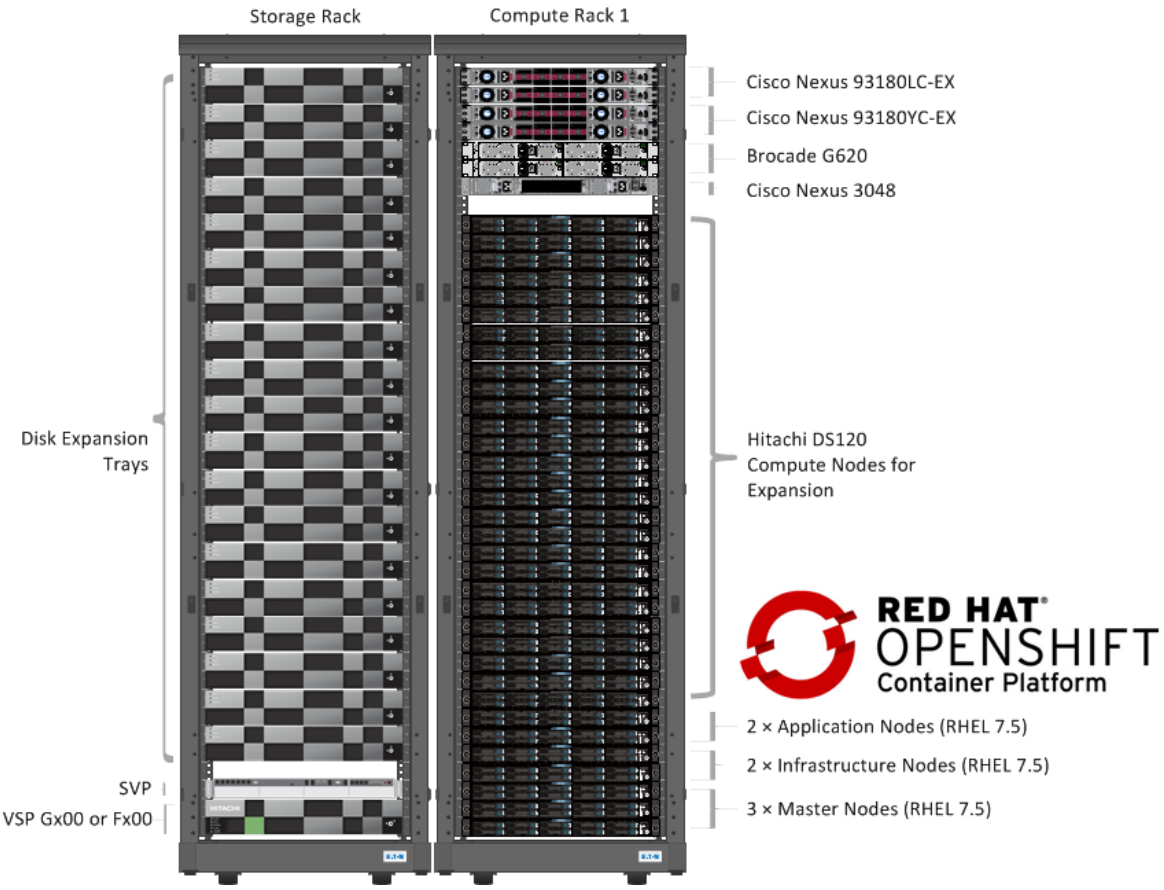


Figure 3 shows the rack diagram of a minimum OpenShift node allocation on the UCP CI.

Figure 3



The OpenShift cluster was installed on seven compute nodes to start with, and the application nodes can be added later as needed. Red Hat Enterprise Linux (RHEL) 7.5 was installed on all the compute nodes. Table 4 lists the sample OpenShift cluster node allocation and IP Addresses.

TABLE 4. OPENSHIFT NODES AND IP ADDRESSES

Node Name	Node Role	Device Type and Location	IP Address
lb1.oc.local	Load Balancer	VM resides outside of UCP CI	10.76.46.99
bastion.oc.local	Bastion Node	VM resides outside of UCP CI	10.76.46.100
master1.oc.local	Master Node 1	DS120 Rack1 U1	10.76.46.101
Master2.oc.local	Master Node 2	DS120 Rack1 U2	10.76.46.102
Master3.oc.local	Master Node 3	DS120 Rack1 U3	10.76.46.103
infra1.oc.local	Infra Node 1	DS120 Rack1 U4	10.76.46.104
Infra2.oc.local	Infra Node 2	DS120 Rack1 U5	10.76.46.105
node1.oc.local	Application Node 1	DS120 Rack1 U6	10.76.46.106
node2.oc.local	Application Node 2	DS120 Rack1 U7	10.76.46.107

- For detailed procedures of OpenShift installation, refer to the link below:

  https://docs.openshift.com/container-platform/3.9/welcome/index.html

- For detailed procedures of Hitachi Storage Plug-in for Containers installation, refer to the link below:

  https://knowledge.hitachivantara.com/Documents/Adapters_and_Drivers/Storage_Adapters_and_Drivers/Containers/1.1/Storage_Plug-in_for_Containers_Quick_Reference_Guide_v1.1.0

Solution Validation

The Red Hat OpenShift Container Platform was installed on UCP CI successfully. The Hitachi Storage Plugin for Containers (HSPC) was also installed on the OpenShift for persistent volume provisioning. The following container volume operations described in this section were performed to validate this solution.

Create a Persistent Volume from Desired Storage Tier

One of the common methods to manage the storage volume is to use policy-based storage management. Each application requires different storage requirements, including IOPS, I/O latency, and date protection level. On the storage side, storage pools can be created to meet these requirements.

For this validation the Gold and Bronze storage pools were created in the VSP G900 as shown in Figure 4

Figure 4

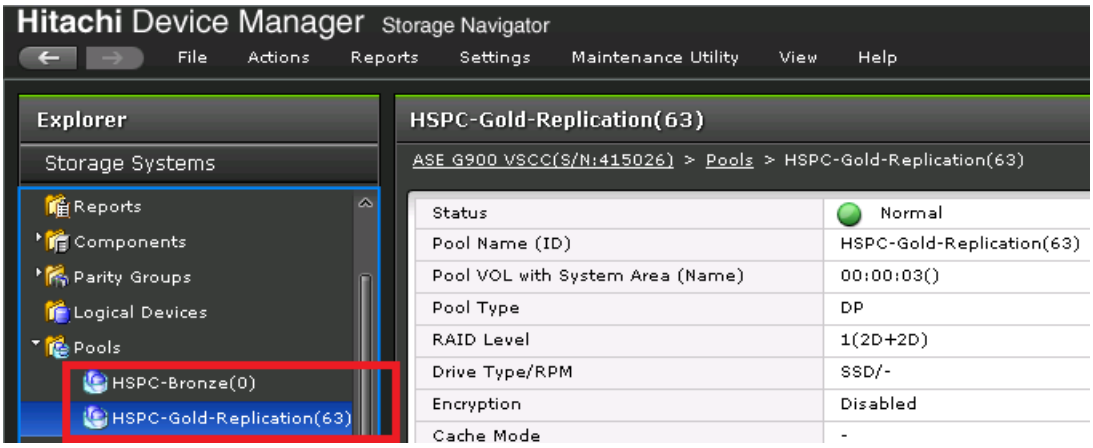


Table 5 lists the details of the storage pools.

TABLE 5. STORAGE POOL CONFIGURATION

Storage Pool Name	Pool ID	Disk Type	Data Protection	Usage
HSPC-Bronze	0	NL-SAS-7.5K	RAID	High capacity and general usage
HSPC-Gold-Replication	63	SSD	RAID + Data Replication by Global-Active-Device (GAD)	High I/O performance with data replicated to another site. For tier 1 applications.

On the OpenShift, a storage class needs to be created for each corresponding storage pool. For this test, following two storage classes were created as shown in Table 6.

TABLE 6. STORAGE CLASS NAME MAPPING IN OPENSHIFT/KUBERNETES

Storage Class Name (in OpenShift)	VSP G900 Storage Pool ID	VSP G900 Storage Pool Name
sc-bronze-vsp-g900	0	HSPC-Bronze
sc-gold-rep-vsp-g900	63	HSPC-Gold-Replication

Figure 5 shows the content of `sc-gold-rep-vsp-g900.yaml` file. In this yaml file, set the storage pool ID in `poolID` and `snapshotPoolID` fields. The `snapshotPoolID` is used for the volume cloning feature. To create the `sc-gold-rep-vsp-g900` storage class, use the following command.

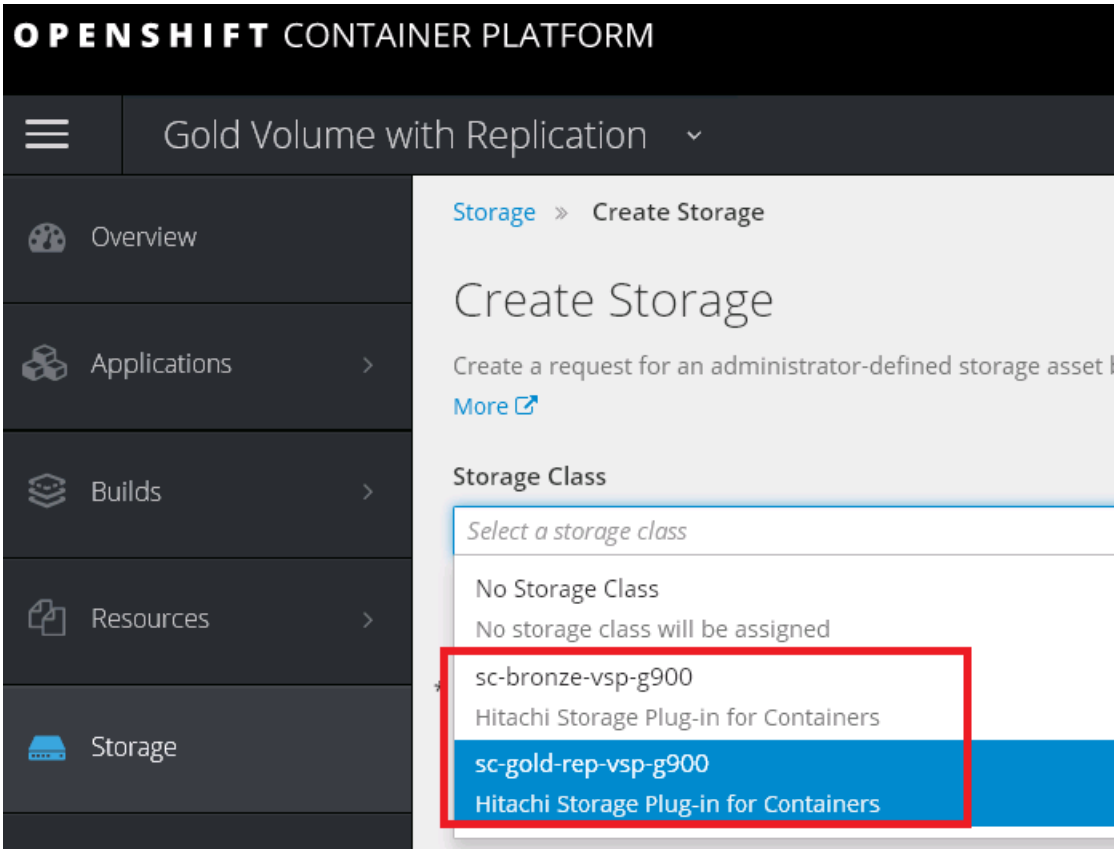
```
# oc create -f sc-gold-rep-vsp-g900.yaml
```

Figure 5

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: sc-gold-rep-vsp-g900
  annotations:
      kubernetes.io/description: Hitachi Storage Plug-in for Containers
provisioner: hitachi.io/hspc
parameters:
  serialNumber: "415026"
  url: http://172.17.43.176
  user: ·····
  password:
  poolID: "63"
  snapshotPoolID: "63"
  scsiTargetID: CL1-B-4,CL4-B-1
```

Once storage classes were created, these options showed up in the OpenShift web UI as shown in Figure 6

Figure 6



A 500 GB persistent volume was created from the OpenShift web UI. Within a few seconds after clicking the **Create** button, the newly created volume was ready to be attached to a container.

Clone a Persistent Volume

With Hitachi VSP Fx00/VSP Gx00 storage family, a container volume can be fully cloned inside of the storage system. Clone volumes can be used as testing and development without affecting original volumes. They can be used for backup as well.

For this validation test, Kibana and Elasticsearch containers were deployed with 10 GB persistent volume from VSP G900, and some mobile data set and random data were ingested to Elasticsearch. To clone this volume with newly ingested data, the following pvc-clone-mobile-vol-g900.yaml file was used as shown in Figure 7

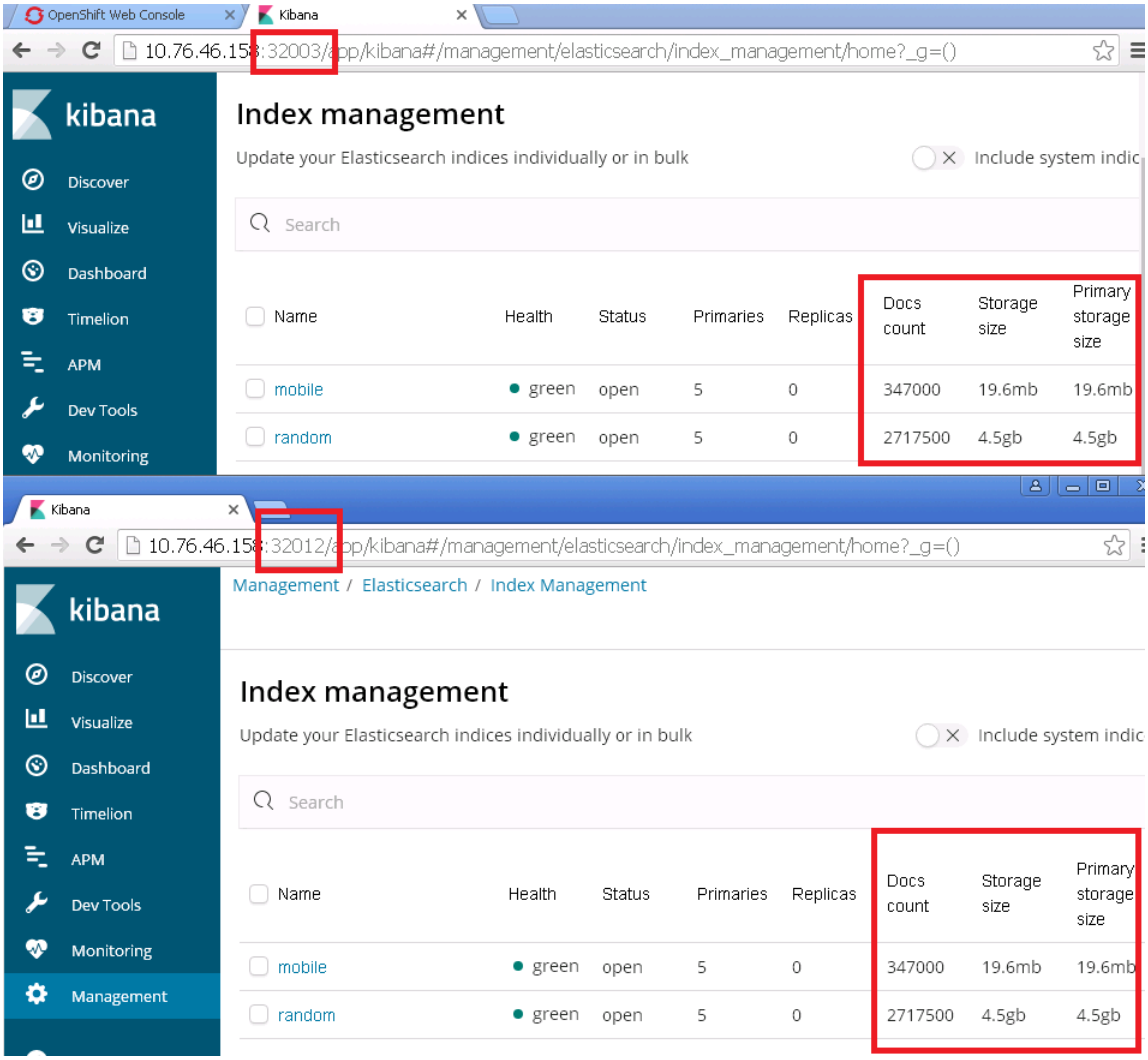
Figure 7

```
[jeff@bastion ~]$ cat pvc-clone-mobile-vol-g900.yaml
kind: PersistentVolumeClaim
apiVersion: v1
metadata:
  name: pvc-clone-mobile-vol-g900-2
  annotations:
    hitachi.io/mode: "clone"
    hitachi.io/sourceVolID: "20"
spec:
  storageClassName: sc-bronze-vsp-g900
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 10G

[jeff@bastion ~]$
[jeff@bastion ~]$
[jeff@bastion ~]$ oc create -f pvc-clone-mobile-vol-g900.yaml
```

The cloned volume is ready to be used immediately after the creation, and the full data copy is performed in the backend storage system. Figure 8 shows the original Kibana/Elasticsearch container application on the top and the new set of container applications with the cloned volume on the bottom.

Figure 8



Please note that the port numbers to access the Kibana web application are different, but since the data is loaded from the clone volume, the `Docs count` and data usage are exactly the same between these two container applications. This validates the following:

- A persistent volume was cloned successfully
- Clone volume was ready for immediate use after the creation
- The data integrity of the cloned volume was validated by attaching the cloned volume to the new set of container applications and reading from them.

Conclusion

Hitachi UCP CI with Red Hat OpenShift gives container platforms enterprise scale, efficiency, flexibility, performance, and resiliency. Hitachi VSP G/F series meets and exceeds any persistent volume needs for any container workloads. This solution provides the foundation a modern IT Infrastructure needs to support a private cloud environment.

Appendix

This section describes some of the configurations used in the lab environment.

Adding Local Storage Volume for Compute Nodes

If SATADOM is used as a boot device for the UCP CI compute nodes, the local storage capacity might not be enough for the OpenShift installation and container deployments. In this case, provision the extra volume from the VSP Fx00/VSP Gx00 storage, and follow the steps below to set it up.

1. Use the command below to find the target the volume:

   ```
   # multipath -ll
   ```

2. Use the command below to create a PV and VG

   ```
   # pvcreate /dev/mapper/<mpathX>
   ```

   ```
   # vgcreate <nodeX-vg> /dev/mapper/<mpathX>
   ```

3. Add the following to the `/etc/sysconfig/docker-storage-setup` file:

   ```
   VG=<nodeX-vg>
   ```

4. Run the following command:

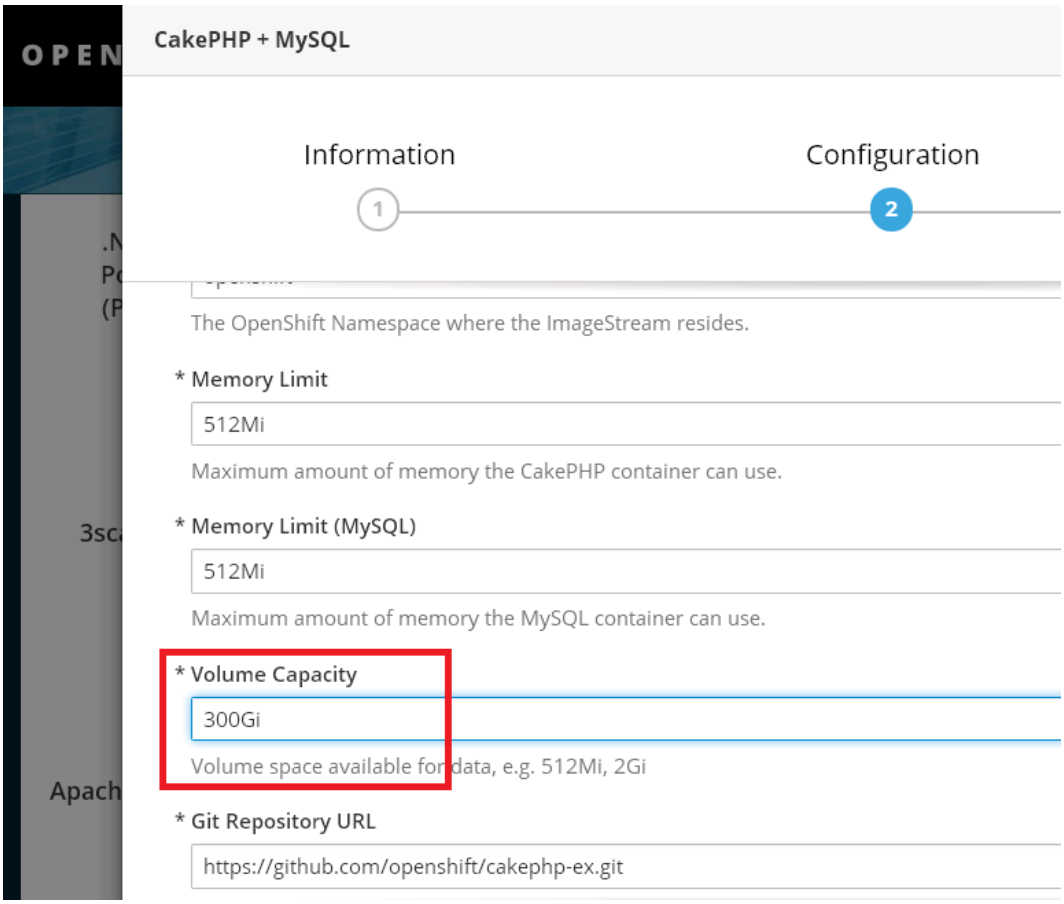
   ```
   # docker-storage-setup
   ```

For more information, refer to https://docs.openshift.com/container-platform/3.9/install_config/install/host_preparation.html#configuring-docker-thin-pool: Option B.

Setting Default Storage Class for Persistent Volumes

Setting a default storage class in OpenShift simplifies container deployments. The example in Figure 9 shows a MySQL container deployment from OpenShift Catalog. Users enter the volume capacity requirement and a persistent volume will be provisioned by the default storage class.

Figure 9



Use the following command to set the default storage class in OpenShift.

```
# oc patch storageclass <storage-class-name> -p '{"metadata": {"annotations":
{"storageclass.kubernetes.io/is-default-class": "true"}}}'
```

Hitachi Vantara

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 96050-2639 USA
HitachiVantara.com | community.HitachiVantara.com

Contact Information
USA: 1-8000446-0744
Global: 1-858-547-4526
HitachiVantara.com/contact

© Hitachi Vantara Corporation, 2018. All rights reserved. HITACHI is a trademark or registered trademark of Hitachi, Ltd. All other trademarks, service marks, and company names are properties of their respective owners

Notice: This document is for informational purposes only, and does not set forth any warranty, expressed or implied, concerning any equipment or service offered or to be offered by Hitachi Vantara Corporation.

MK-SL-108-00, September 2018