Reference Architecture

HPE Reference Configuration for Red Hat OpenShift Container Platform on HPE Synergy and HPE Nimble Storage

Contents

Executive summary

In today's digital world, organizations are under increasing pressure to deliver applications faster while reducing costs. As these applications grow more complex, this puts stress on IT infrastructure, IT teams, and processes. To remain competitive, organizations must adapt quickly, and developers need to be more effective, efficient, and agile. Container technology provides the right application platform to help organizations become more responsive and iterate across multiple IT environments as well as develop, deploy, and manage applications faster. But implementing a containerized environment across existing infrastructure is a complex undertaking that can require weeks or months to mobilize, particularly for enterprises. To help accelerate container application delivery, Hewlett Packard Enterprise and Red Hat® are collaborating to optimize Red Hat OpenShift Container Platform on HPE platforms, including HPE Synergy, the industry's first composable infrastructure, and HPE Nimble Storage.

Red Hat OpenShift Container Platform on HPE Synergy provides an end-to-end fully integrated container solution that, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale. Red Hat OpenShift Container Platform provides organizations with a reliable platform for deploying and scaling container-based applications and HPE Synergy provides the flexible infrastructure you need to run that container platform to dynamically provision and scale applications, whether they run as VMs or containers, or are hosted on-premises, in the cloud, or in a hybrid environment.

This Reference Configuration provides architectural guidance for deploying, scaling, and managing a Red Hat OpenShift environment on HPE Synergy Composable Infrastructure along with HPE Nimble Storage.

This Reference Configuration describes how to:

- Efficiently lay out an OpenShift configuration using a mix of virtual machines and bare metal hosts.
- Configure persistent storage for containers using HPE Nimble Storage.

This Reference Configuration demonstrates the following benefits of utilizing HPE Synergy for Red Hat OpenShift Container Platform:

- Automated initial installation and configuration of highly available vSphere hosts and the management virtual machine for the Red Hat OpenShift Container Platform, is reduced from more than 3 hours to approximately one hour, and complexity of the manual operation is reduced from needing to perform more than 500 steps, to running two Ansible play books.
- Automated deployment and configuration of physical worker nodes is reduced from 8 hours to under 20 minutes, and complexity of manual operation is reduced from needing to perform close to 300 steps, to running two Ansible play books.
- Automated host preparation of the Red Hat OpenShift Container Platform Nodes is reduced from up to 8 hours to as little as 20 minutes, and the complexity of the manual operation is reduced from performing more than three hundred steps, to one Ansible playbook.
- Deployment of the core management functions on VMs, to optimize resource usage while keeping the worker nodes on either bare metal or virtual, according to the capacity requirements of the pods that will be deployed.
- Disconnected installation ensures the OpenShift Container Platform software is made available to the relevant servers, then follows the same installation process as a standard connected installation.
- Using Enterprise grade storage solution, for example HPE Nimble Storage for persistent container storage, enables speed, portability, and agility for traditional enterprise applications and data.
- The HPE Synergy Composable Infrastructure solution provides a layered view of security controls. The objective of choosing this layered security view is to ensure that consumers become aware of the depth of security risk that an infrastructure can have and also make them aware of the depth of defense that is built in to the HPE Synergy Composable Infrastructure design.
- Container platform security provided by Sysdig that is installed as part of Sysdig Cloud-Native Intelligence Platform. Sysdig Secure and Sysdig Monitor offer unified container security, monitoring, and forensics for container and Kubernetes based environments.
- HPE Synergy for Red Hat OpenShift Container Platform is a business-driven container application data protection and OpenShift configuration protection automation using Ansible playbooks.
- Value add services like service mesh, cluster console, Prometheus monitoring, and Grafana dashboard are also provided.

Target audience: This document is for Chief Information Officers (CIOs), Chief Technology Officers (CTOs), data center managers, enterprise architects, and implementation personnel wishing to learn more about Red Hat OpenShift Container Platform on HPE Synergy Composable Infrastructure. Familiarity with HPE Synergy, HPE Nimble Storage, Red Hat OpenShift Container Platform, container-based solutions, Ansible Engine, and core networking knowledge is assumed.

Document purpose: The purpose of this document is to describe a Reference Configuration that describes the benefits and technical details of deploying Red Hat OpenShift Container Platform on HPE Synergy Composable Infrastructure with HPE Nimble Storage. The implementation details and plays discussed in this document are available from HPE at https://github.com/hewlettpackard/hpe-solutions-openshift.

Introduction

This Reference Configuration describes a highly available and secure Red Hat OpenShift Container Platform deployment on HPE Synergy Composable Infrastructure and includes details on how the environment is connected and configured. When combined with the accompanying deployment guide (https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere), it provides a comprehensive example of how Red Hat OpenShift Container Platform can be set up to take advantage of the HPE Synergy Composable Infrastructure and HPE Nimble Storage. The configuration used for this solution consists of three (3) OpenShift Container Platform master instances and six (6) OpenShift Container Platform worker instances. This configuration is housed on a three (3) frame HPE Synergy with nine (9) HPE Synergy 480 Gen 10 Compute Modules installed with a mix of VMware vSphere and Red Hat® Enterprise Linux®. HPE Nimble Storage is used to provide persistent storage for containers and registry, virtual machine storage, and data management. This architecture can scale between three (3) and n worker nodes.

Due to the ephemeral nature of containers, protecting persistent data associated with the containers becomes a crucial task. In this Reference Configuration, the Red Hat OpenShift Container Platform's persistent volumes can optionally be protected using replication to a remote HPE Nimble Storage array. Red Hat OpenShift Container Platform configurations are protected using Ansible playbooks to the HPE Nimble Storage array thus providing end-to-end data protection and disaster recovery capability.

Security is a required component of any production IT infrastructure solution. The overall containerized ecosystem solution discussed in this document is secured using Sysdig Cloud-Native Intelligence Platform. Sysdig Secure performs container image scanning, run-time protection, and forensics to identify vulnerabilities, blocks threats, enforces compliance, and performs audit activity across enterprise cloud-native environments at scale. Sysdig Monitor is a powerful application for monitoring and managing the risk, health, and performance of your microservices.

The HPE Synergy platform is designed to bridge traditional and cloud-native applications with the implementation of HPE Synergy Composable Infrastructure. HPE Synergy Composable Infrastructure combines the use of fluid resource pools made up of compute, storage, and fabric with software-defined intelligence. Composable pools of compute, storage, and fabric can be intelligently and automatically combined to support any workload. The resource pools can be flexed to meet the needs of any business app. HPE Synergy platform provides the agility and scalability on the hardware layer to the overall Red Hat OpenShift Container Platform solution.

Solution overview

This Reference Configuration provides an overview of the Red Hat OpenShift Container Platform on HPE Synergy and HPE Nimble Storage solution as described in greater detail at, https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere. Figure 1 provides an overview of the solution components.

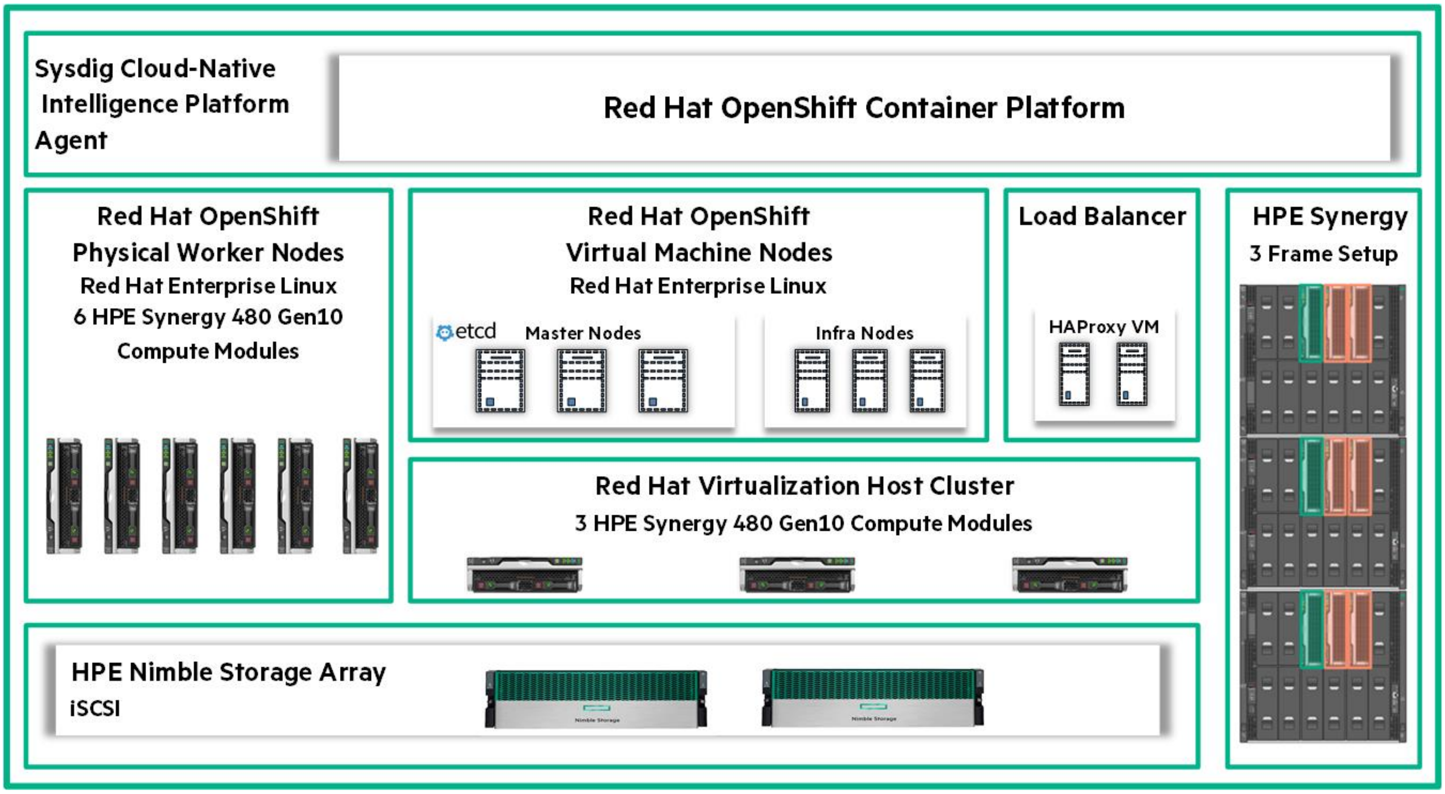

Sysdig Cloud-Native Intelligence Platform Agent
Red Hat OpenShift Container Platform
Red Hat OpenShift Physical Worker Nodes
Red Hat Enterprise Linux
6 HPE Synergy 480 Gen10 Compute Modules
Red Hat OpenShift Virtual Machine Nodes
Red Hat Enterprise Linux
etcd
Master Nodes
Infra Nodes
Load Balancer
HAProxy VM
HPE Synergy
3 Frame Setup
Red Hat Virtualization Host Cluster
3 HPE Synergy 480 Gen10 Compute Modules
HPE Nimble Storage Array
iSCSI

Figure 1. Solution layout

This Reference Configuration deploys Red Hat OpenShift Container Platform 3.11 as a combination of virtual and physical resources. The OpenShift master, and infrastructure nodes are deployed as virtual machines running on three (3) HPE Synergy 480 Gen10 Compute Modules running on VMware vSphere Virtualization host and managed by VMware vCenter Server appliance. VMware vSphere hosts are booted from HPE Image Streamer and are installed and configured using Ansible playbooks. Load balancing can be deployed as a virtual machine or as physical appliances. Red Hat OpenShift worker nodes are deployed on bare metal or as virtual machines on six (6) HPE Synergy 480 Gen10 Compute Modules running Red Hat Enterprise Linux 7.6. The operating systems or hypervisors for the Red Hat OpenShift worker node hosts are booted from HPE Image Streamer and post-installation configuration steps are performed, in part, using Ansible playbooks. HPE Nimble Storage provides support for both ephemeral and persistent container volumes.

The HPE Converged Architecture 750 (CA750) was used as the reference platform for the Red Hat OpenShift deployment. Customers can also customize their OpenShift configuration based on their workload needs, without using or just leveraging parts of the CA750 design. The CA750 approach provides pre-integrated, modular, scalable converged systems that reduce deployment risk. To implement this solution as an HPE CA750, work with your HPE Authorized Channel Partner. For more information about flexible HPE Converged Systems, refer to https://www.hpe.com/us/en/integrated-systems/converged-architecture.html.

Security

To address the security challenges that exist in containerized environments, this solution leverages Sysdig SaaS Platform to secure and monitor Red Hat OpenShift Container Platform, an enterprise-ready Kubernetes platform installed and configured on HPE Synergy Composable Infrastructure. Once the configuration is deployed, access to the Red Hat OpenShift Cluster is granted to the Sysdig SaaS Platform. The Sysdig SaaS Platform is a cloud-based service where the security and monitoring services will be available to the user based on the subscription they have chosen. For security and monitoring of OpenShift containers, it is required to install the Sysdig Agent on the OpenShift Cluster, which means Sysdig Agents that are light-weight entities will be installed within each node in the OpenShift Cluster. These agents run as a daemon to enable Sysdig Monitor and Sysdig Secure functionality. Sysdig Monitor provides deep, process-level visibility into dynamic, distributed production environments. Sysdig Secure provides image scanning, run-time protection, and forensics to identify vulnerabilities, blocks threats, enforces compliance, and performs audit activity across OpenShift Cluster.

The key security benefits that this solution provides are:

- Faster incident resolution using Sysdig Monitor for OpenShift Cluster
- Simplified compliance for the entire solution
- Service-based access control for container security and monitoring
- Less time spent on managing platforms, containers, and vulnerabilities.

The implementation of Sysdig in this solution uses the Software as a Service (SaaS) deployment method. The playbooks deploy Sysdig Agent software on every OpenShift node and captured data is relayed back to your Sysdig SaaS Cloud portal. The deployment provides access to a 90 day try-and-buy, fully featured version of the Sysdig software.

Note

The Sysdig functionality is not turned on by default in this solution. Refer to the Sysdig configuration available at, https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere for more information on how to enable Sysdig. For more information on how to access the 90 day try-and-buy version, see the GitHub repository at, https://github.com/HewlettPackard/Docker-Synergy.

Figure 2 shows the security overview for Red Hat OpenShift cluster platform on HPE Synergy.

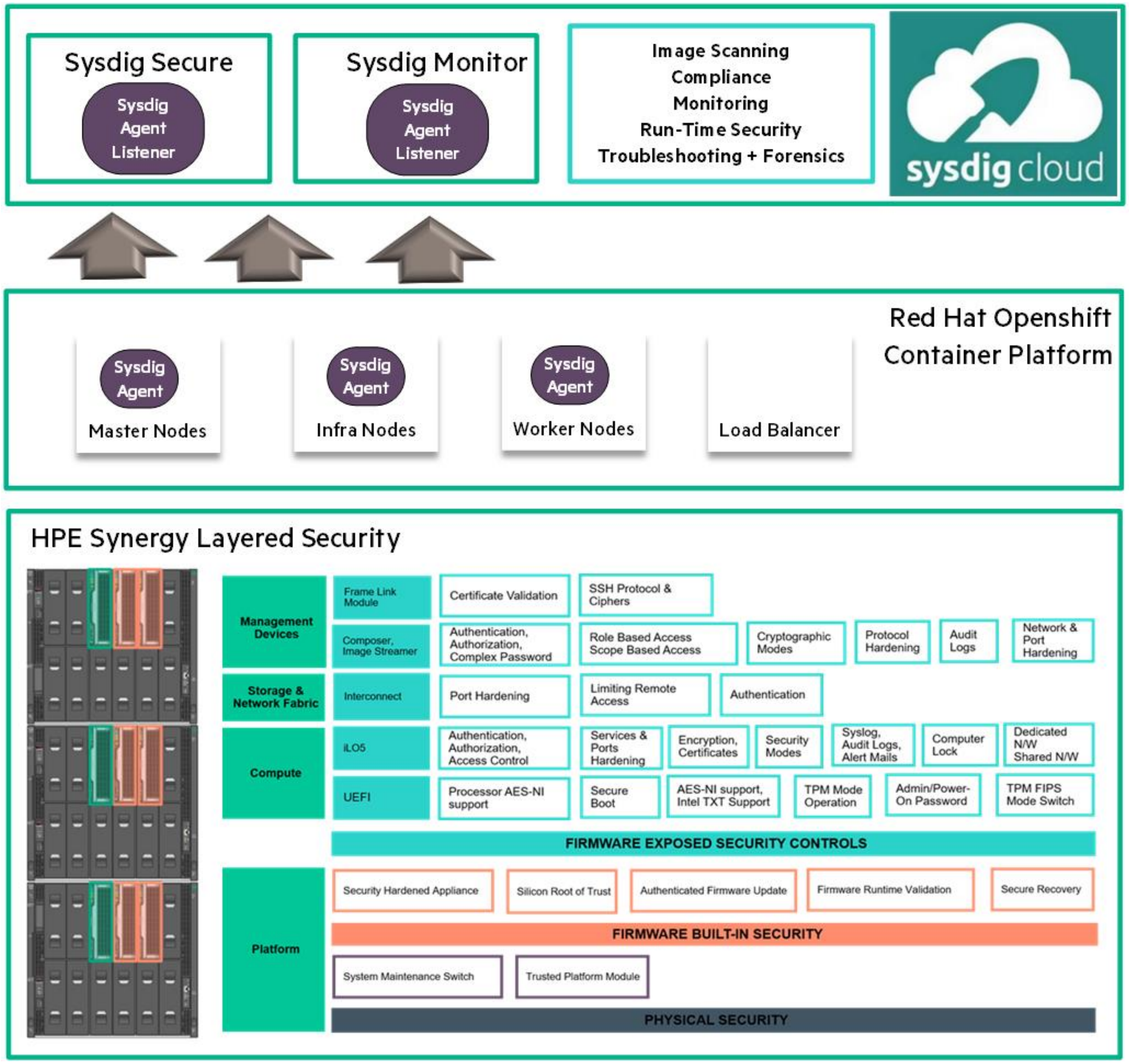


Figure 2. Security overview for Red Hat OpenShift Cluster platform on HPE Synergy platform

For more information on Sysdig security for Red Hat OpenShift Container Platform, refer to https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere.

HPE Synergy Composable Infrastructure security controls

Hewlett Packard Enterprise has security features and functionalities built into servers from the hardware level to the firmware. Customers require a holistic view of the security controls available to them in the composable infrastructure to make the most of their investment. HPE Synergy Composable Infrastructure enables IT organizations to accelerate application and service delivery through the use of fluid resource pools, made up of compute, storage, and fabric with software-defined intelligence. Each resource within the composable infrastructure is in turn made up of multiple products such as compute modules which in turn has multiple components such as iLO, UEFI, etc. Another example is the management devices like HPE Synergy Composer, which exposes its functions using HPE OneView and the HPE Synergy Frame Link Modules. With so many products available within the composable infrastructure, it is important to understand the security controls available within each of them and how they can be used to help avoid potential security breaches.

This solution provides a layered view of security controls that are available to Hewlett Packard Enterprise customers. Figure 2 shows the layered security view across various composable infrastructure component.

Table 1 describes the security control layers of Figure 2, physical security controls layer, firmware/hardware built-in security controls layer, and firmware exposed security functionality controls layer. The objective of choosing this layered security view is to ensure that the customers will be aware of the depth of security risk that an infrastructure can have and also to know the defense in depth that is built-in to the HPE Composable Synergy Infrastructure design. Each security control at each layer is designed to comply with the requirements of some security tenets. The security tenets are a set of security principles that ensure the security within the information systems, example: Authentication, Authorization, Access Control, Password Policies, Cryptographic Ciphers, Secure Protocols, Forensic Analytics – Logs, Alerts, Threat Modelling, Security Certifications and Standards.

Table 1. Physical and Firmware based security controls within HPE Synergy Composable Infrastructure

Security controls category	Description
Physical security controls	Physical security describes measures designed to ensure the physical protection and detection of threat event in the infrastructure.
Firmware/hardware built-in security controls	This cover: • The security technologies built in the firmware to make it more secure for any communication with the underlying hardware and safe for user data at rest/transit. • The threat modelling followed within HPE to security harden the infra components.
Firmware exposed security functionality controls	This is the exhaustive list of security controls that let the customers: • Define the boundaries for accessing various infra components. • Set quantum safe ciphers for encryption. • Generate alert and log changes to infra.

Firmware built-in security controls

Hewlett Packard Enterprise has used a variety of technologies to ensure that the built-in firmware security controls provide the highest level of infrastructure security. This section provides a brief overview of the security controls that HPE has built into the firmware that is used by HPE Synergy and how these security controls offers an added advantage for HPE Synergy customers.

Silicon Root of Trust: The iLO5 chipset contains a first of its kind Silicon Root of Trust for the HPE Synergy Gen 10 Compute platform which is included with the iLO Standard license. Silicon Root of Trust provides an inextricably tied link between the silicon and firmware—making it impossible to insert any malware, virus, or compromised code that would corrupt the boot process. The Silicon Root of Trust enables the boot process to provide a Secure Start. When the system boots, the iLO5 chip validates and boots its own firmware first, then validates the system BIOS. Because the Silicon Root of Trust is inextricably tied into the iLO5 hardware, every validated signature throughout the boot process can be trusted. However, in the unlikely event that iLO5 finds tampering or corruption at any point in the process, trusted firmware is immediately available for Secure Recovery. On startup, if iLO5 finds that its own firmware has been compromised, it will load its own authenticated firmware from an integrated backup. The iLO5 firmware recovery is always available and always automatic—regardless of license. Remember that the Silicon Root of Trust in hardware is how the iLO5 firmware is verified, so it can always be trusted. Second, if iLO5 finds that the system BIOS has been compromised, customers can connect to iLO5 and manually recover to authenticated firmware.

Because the Silicon Root of Trust is embedded in the hardware itself, iLO5 is able to detect any compromised firmware—as far back as the supply chain process. HPE can address platform security all the way back to the supply chain because HPE designs the iLO5 entirely—hardware and firmware—and controls the iLO5 production process. Unlike other companies, Hewlett Packard Enterprise does not outsource the server management controller. Hewlett Packard Enterprise also has strict internal processes that dictate the firmware approval process. This gives customers an unprecedented level of assurance that no hackers have compromised the firmware before the server is received.

Secure Recovery: Secure Recovery is included in the iLO Advanced Premium Security Edition license and works alongside Silicon Root of Trust to automatically recover firmware back to a known good state in the unlikely event that it is compromised. As described previously, the Silicon Root of Trust enables the secure start process. As the system boots and iLO5 verifies the series of digital signatures, iLO5 can access trusted firmware immediately and recover to a known good state, if it finds tampering or corruption in its own firmware or the system BIOS. First, if iLO5 finds that its own firmware has been compromised, it will load its own authenticated firmware from an integrated backup. The Secure Recovery of iLO5 firmware is always available and always automatic—regardless of license. Second, if iLO5 finds that the system BIOS has been compromised, iLO5 will try to recover from a backup copy. If the backup copy is also compromised and the customer has upgraded to the iLO Advanced Premium Security Edition license, iLO5 can automatically recover authentic firmware. The standard license provides the opportunity for manual recovery. The Silicon Root of Trust is the foundation for the entire Secure State and Secure Recovery process, enabling HPE Gen10 servers to be the world's most secure industry standard servers and providing the extraordinary ability to not only verify the digital signatures up through the entire boot process but also to recover securely if any firmware is compromised.

Firmware runtime validation: With the iLO Advanced Premium Security License, the iLO5 chipset enables runtime validation of firmware. With firmware runtime verification, the iLO5 chipset performs the same checking process that happens during the boot process on a continual basis while the server is running. As frequently as once a day, iLO5—with its Silicon Root of Trust—runs a background verification check on the iLO5 firmware, UEFI, and other firmware loaded after including the SPLD, IE, and ME.

Authenticated firmware updates: The iLO5 chipset expands the number of firmware items that customers can update directly and securely in the Gen10 servers. This is a standard feature on the iLO5. Firmware items that can be securely validated and updated from the iLO now includes system programmable logic devices (SPLDs), HPE ProLiant power interface control utility (PowerPIC) firmware, the Intel® innovation Engine and Management Engine, and other low-level system components. The iLO5 contains a firmware repository stored on non-volatile flash memory (NAND), which allows components such as the Service Pack for ProLiant (SPP) and other firmware updates to be applied and installed offline through iLO5.

Best practices followed by HPE to deliver security hardened Synergy Composer Appliance: Hewlett Packard Enterprise follows Secure Development Lifecycle and used a security assessment tool called Comprehensive Applications Threat Analysis (CATA) to identify and remediate security defects in the appliance operating system.

Note
The design of the appliance is based on CATA fundamentals and underwent CATA review.

The factors that contribute to appliance security hardening are listed below:

- Appliance is hardened to enforce mandatory access control. This means users of HPE Synergy are provided the role-based access control (RBAC) that enables an administrator to establish access control and authorization for users based on their responsibilities.
- Important services of the appliance run with required privileges. This implies HPE Synergy Composer is governed by scope-based access control that enables an administrator to establish access control for users by allowing a role to be restricted to a subset of resources managed by the appliance.
- The appliance is configured and maintains a firewall that blocks unused ports. Restricting the usage of all non-essential ports reduces the attack surface for HPE Synergy Composer.
- The appliance operating system bootloader is password protected. This means HPE Synergy Composer cannot be compromised by someone attempting to boot in single-user mode.
- The appliance is designed to operate in an isolated management LAN. Hewlett Packard Enterprise recommends creating a private management LAN and keeping that separate, known as air-gapped, from production LANs, using VLAN or firewall technology (or both).
- HPE supports digital signing of all software/firmware updates to ensure their integrity and authenticity. This implies that when the customer is re-imaging the composer in order to quickly bring it to a specific firmware revision level, the digital signature is verified by the reimaging process.
- Operating-system-level users are not allowed to access the appliance, with the following exceptions:
  - A special preset command used only if the Infrastructure administrator password is lost or forgotten. This command requires that you contact your authorized support representative to obtain a one-time password.
  - A setting that enables an authorized support representative to obtain a one-time password so that they can log in to the appliance console (and only the console) to perform advanced diagnostics. Customer can either enable or disable access with this setting.
- HPE closely monitors security bulletins for threats to appliance software components and, if necessary, issues software updates.

Data protection for Red Hat OpenShift

Containers have dramatically increased in popularity as organizations recognize the benefits with respect to both time and resource efficiency. This explosive growth of container applications overwhelms traditional data protection approaches. Applying traditional data protection strategies to containerized applications will not work. The goals of this solution with regards to data protection are to:

- Highlight the importance of protecting each component within an OpenShift cluster including persistent volumes in order to restore in case of corruption or system failures.
- Demonstrate Hewlett Packard Enterprise's approach to protecting and restoring critical data using HPE Nimble Storage.

OpenShift Cluster components

An OpenShift cluster is made up of several nodes and each node type has different roles. To protect the environment, it is very important to understand how these components fit together and the services provided by each component. A successful backup and recovery solution is highly unlikely without this understanding in place. This section provides details of each OpenShift node and what components require protection within the environment.

High availability is achieved using three (3) master nodes and three (3) infrastructure nodes. Worker nodes can be sized in variable quantities according to the capacity requirements of the pods that will be deployed. It is necessary to create a backup of the important components within the OCP cluster in order to recreate the nodes in the event of a failure. Figure 3 describes the major components involved in the deployment of Red Hat OCP.

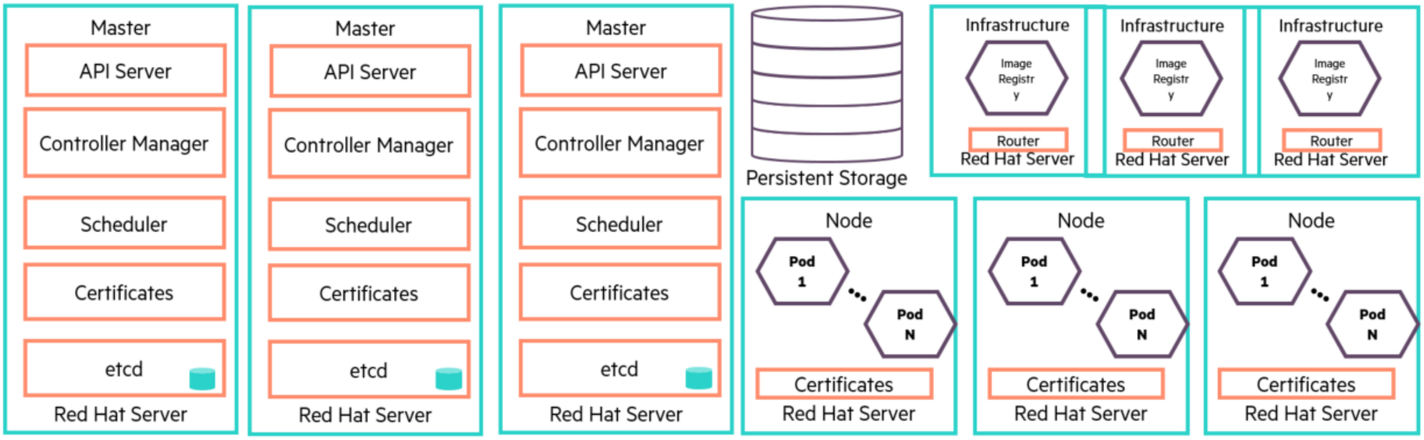


Figure 3. Solution diagram

OpenShift master node

The OpenShift master nodes are comprised of a set of core components including the API server, Controller Manager Server, Scheduler and Certificates.

The master nodes maintain the cluster's configuration, manage nodes in the OpenShift cluster, and schedule pods to run on nodes. If the OpenShift master nodes are unable to function, this will not impact the end users as the container application traffic will remain functional. However, administrators and users will not be able to make any new adjustments to the OpenShift cluster.

API server

The API server provides the management entry point of the OpenShift cluster. It mediates the interactions between the OpenShift master node components via RESTful API calls. It is responsible for storing API objects into the persistent etcd store. API server high availability is built on the persistent etcd store and deploys multiple instances of API server roles on the OpenShift cluster.

Controller Manager

The Controller Manager monitors the state of the cluster through the API Server watch feature. When a state change notification is received, it makes the necessary changes attempting to move the current state towards the desired state to keep the OpenShift cluster functioning correctly. Multiple controller manager roles are configured on OpenShift master nodes to provide high availability.

Scheduler

The scheduler ensures that container applications are scheduled to run on worker nodes within the OpenShift cluster. The scheduler reads data from the pod and attempts to find a node that is a good fit based on configured policies. To ensure high availability, more than one OpenShift master node must be configured for the scheduler roles.

Certificates

Certificates are used by the API server when securing inbound requests, authenticating users, making outbound requests, and for mutual TLS between the API server and all the other API objects in OpenShift Cluster. Certificates are copied to all the master nodes during the deployment. If more than one master host is deployed on an OpenShift cluster, the certificates are considered highly available.

etcd

etcd stores the persistent master state while other components watch etcd for changes to bring themselves into the desired state. It implements the key-value stores where all of the objects in OpenShift cluster master node components are stored. The etcd store implements a distributed consensus algorithm to ensure that even if one of the storage nodes fail, there is sufficient replication to maintain data availability. Optionally etcd role can be configured outside the master node.

OpenShift node

An OpenShift node, or worker node, provides the runtime environment for containers. Each node in an OpenShift cluster has the required services to be managed by the master. The master uses information from nodes to validate nodes with health checks. A node is ignored until it passes the health checks, and the master continues checking nodes until they are valid. Other than running pods, worker nodes contain certificates, services and authorization files. Large numbers of OpenShift nodes may be deployed in a cluster and if one node fails, it can be easily replaced without losing valuable data. However, certificates needs to be deployed on the new node. Typically, certificates and authorization files are redeployed using Ansible playbooks and the Ansible Engine/Tower will hold the files. As a result, the Ansible Engine/Tower must be protected to ensure the file is highly available.

OpenShift infrastructure node

OpenShift makes use of its local registry for storing container images. In a highly available deployment such as the one Hewlett Packard Enterprise has created, the infrastructure nodes are responsible for hosting these registry pods and this is the place where the local container images are stored. Registry pods are assigned with a Persistent Volume (PV) from external storage. In order to protect the data, it is recommended to take a snapshot or clone the volume or replicate it to a disaster recovery site.

Pods inside of an OpenShift cluster are only available via their IP addresses on the cluster network. An edge load balancer can be used to accept traffic from outside networks and route the traffic to pods inside the OpenShift cluster. An OpenShift administrator can deploy routers in an OpenShift cluster through infrastructure nodes. These enable routes created by developers to be used by external clients. OpenShift routers provide external hostname mapping and load balancing to services over protocols that pass information directly to the router. The hostname must be present in the protocol in order for the router to determine where to send traffic. For high availability, an external load balancer such as F5 BIG-IP can be used along with multiple infrastructure nodes.

Persistent storage

Containers were originally designed to run stateless applications so there was no need for persistent storage. When enterprises began adopting containers and they wanted to run stateful applications persistent storage became necessary to meet the demands of the application data. HPE Nimble Storage provides persistent storage capabilities to an OpenShift cluster using plugins. Persistent storage and the data it houses need to be protected for business continuity and disaster recovery purpose. Data protection for the persistent volume is inbuilt with HPE Nimble Storage protection plan, utilizing snapshot and replication capabilities. In this solution, data protection is configured to replicate the persistent volumes to the remote HPE Nimble Storage array, which can help reduce the RPO/RTO requirements. Optionally it can be replicated to HPE Cloud Volumes.

Figure 4 highlights the recommended data protection process for Red Hat OpenShift Container Platform using HPE Nimble Storage.

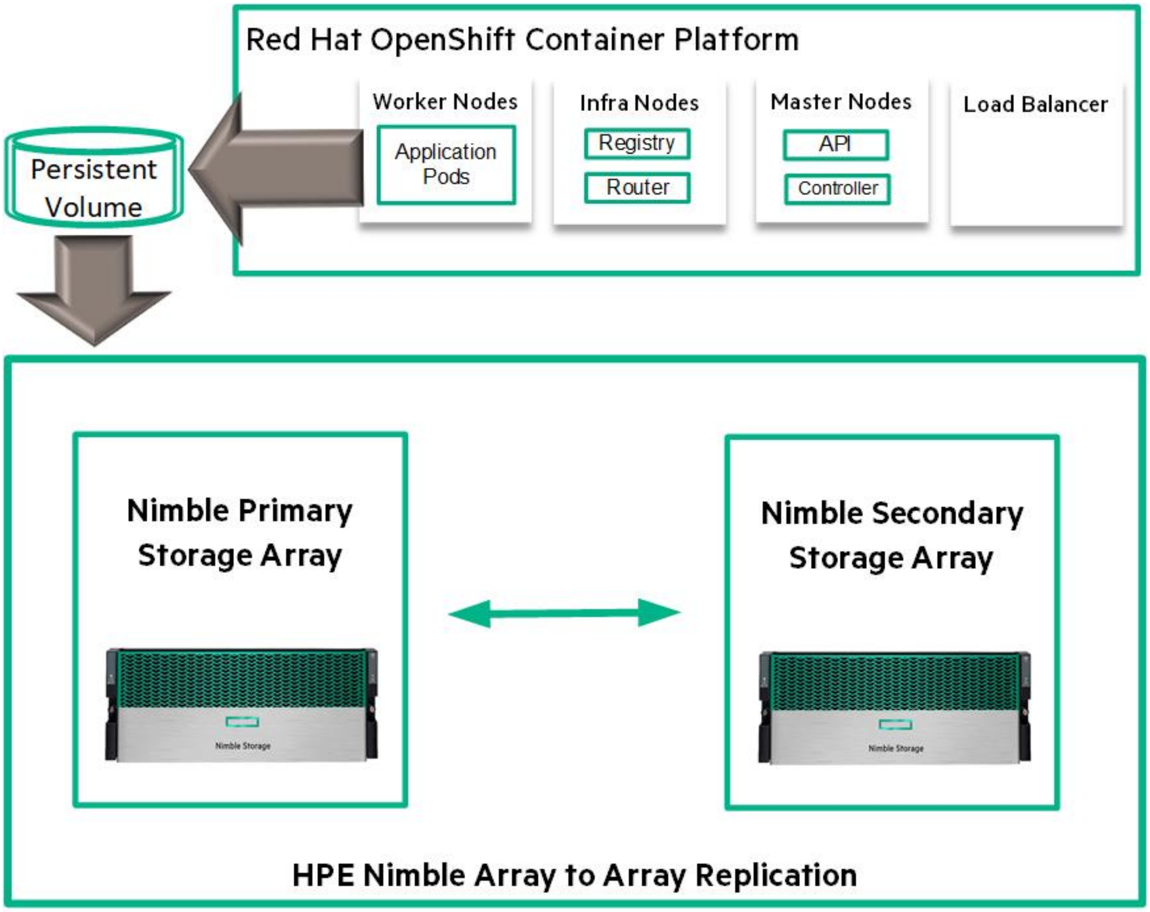


Figure 4. Data protection overview for Red Hat OpenShift cluster platform using HPE Nimble Storage Array

For more information on HPE Nimble Storage data protection for Red Hat OpenShift Container Platform, refer to the backup and recovery document at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere.

Node labeling in OpenShift

Discovering the node properties and advertising them through node labels can be used to control workload placement in an OpenShift cluster. But OpenShift does not by default label nodes with any hardware configuration information. If IT wants to use hardware configuration to optimize scheduling, the capabilities of the underlying platform must be manually uncovered and labeled by administrators in order to use the hardware configuration in scheduling decisions. An OpenShift cluster can have many nodes and each node in turn can run multiple pods which, at scale, means that this process is both tedious and error prone. With OpenShift running on HPE server platforms, organizations can automate the discovery of hardware properties and use that information to schedule workloads that benefit from the different capabilities that the underlying hardware provides. Using HPE iLO and its REST/Redfish API based discovery capabilities (proliantutils) the following properties can be discovered about the nodes:

- Presence of GPUs
- Underlying RAID configurations
- Presence of disks by type
- Persistent-Memory availability
- Status of CPU virtualization features

- SR-IOV capabilities
- CPU architecture
- CPU core count
- Platform information including model, iLO and BIOS versions
- Memory capacity
- Status of secure boot

Once these properties are discovered for the physical worker nodes, OpenShift node labeling can be applied to group nodes based on the underlying features of the hosts. By default, every node will at least have its node name as a label. Node labels can be targeted for deployments using node selectors which can be set at either a project (can be used to restrict which nodes a project gets access to) or pod level. The node labeling system is completely open ended though so administrator can choose whatever groupings make sense for the organization use cases.

Labels do not provide uniqueness. In general, it is expected that many objects will carry the same label(s). Using a label selector, the administrator can identify a set of objects with similar properties. This labelling can be used as either a hard or soft constraint for scheduling of application pods on desired node based on application requirements. For example, if the compute module in the HPE Synergy Composable Infrastructure has support for Intel TXT, which is specifically designed to harden platforms from the emerging threats of hypervisor attacks, malicious root kit installations, or other software-based attacks, administrators can use this information to restrict confidential data or sensitive workloads to nodes that are better controlled and have had their configurations more thoroughly evaluated through the use of Intel TXT-enabled platform.

Figure 5 shows the hardware properties that are used to label the physical worker nodes.

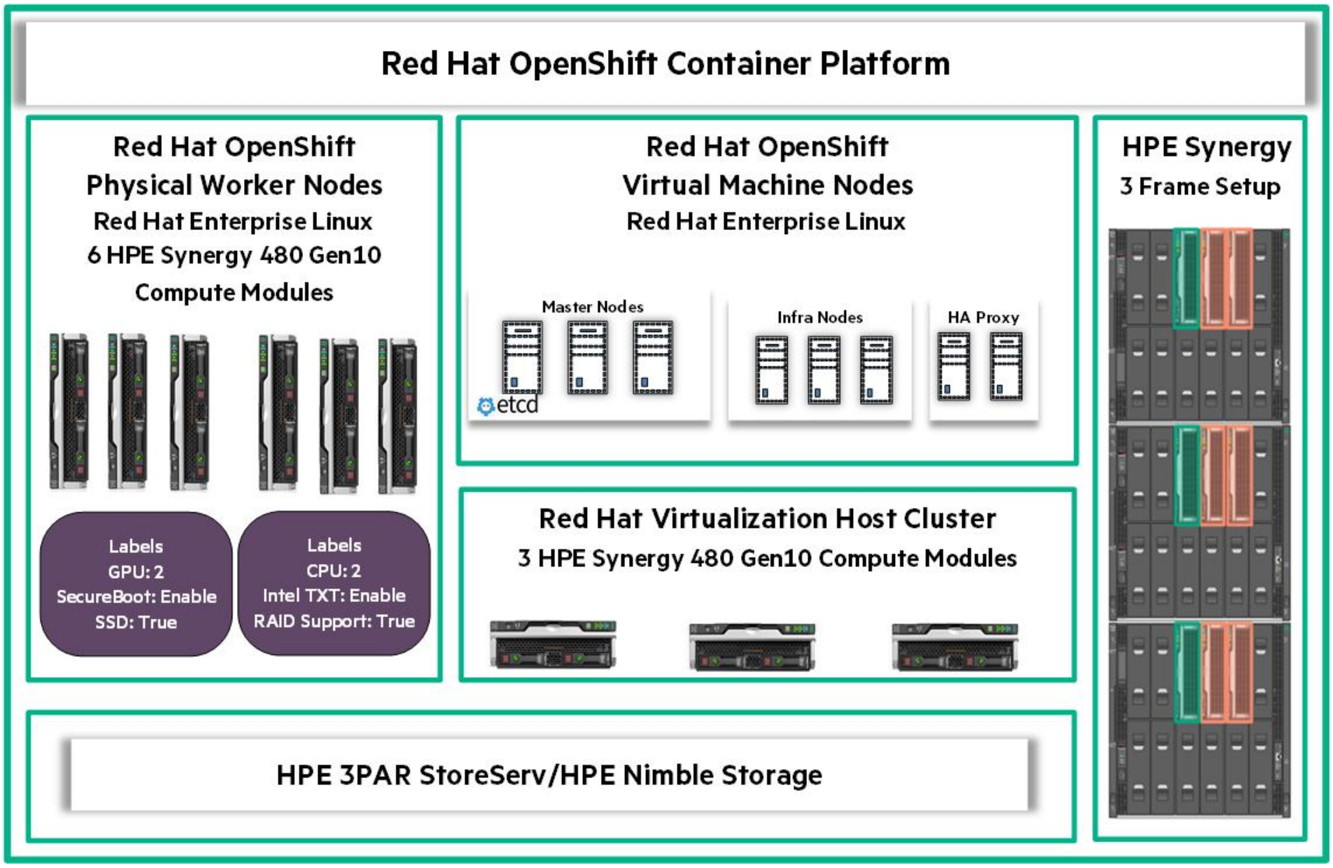


Figure 5. Node labeling within the solution

Disconnected Installation

Some data centers may not have access to the internet, even via proxy servers. Installing OpenShift Container Platform in these environments is considered a disconnected installation. In these air gapped environments, OpenShift Container Platform software channels and Red Hat's Docker registry and repositories are not available via Red Hat's content distribution network. A disconnected installation ensures the OpenShift Container Platform software is made available to the relevant server's offline, then follows the same installation process as a standard connected installation. Follow the deployment guide at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere for more detailed configuration steps.

Solution layout

Figure 6 highlights the solution at a high level. This includes a reflection of the relationship between hosts, OS/hypervisor, boot volumes, and SAN storage.

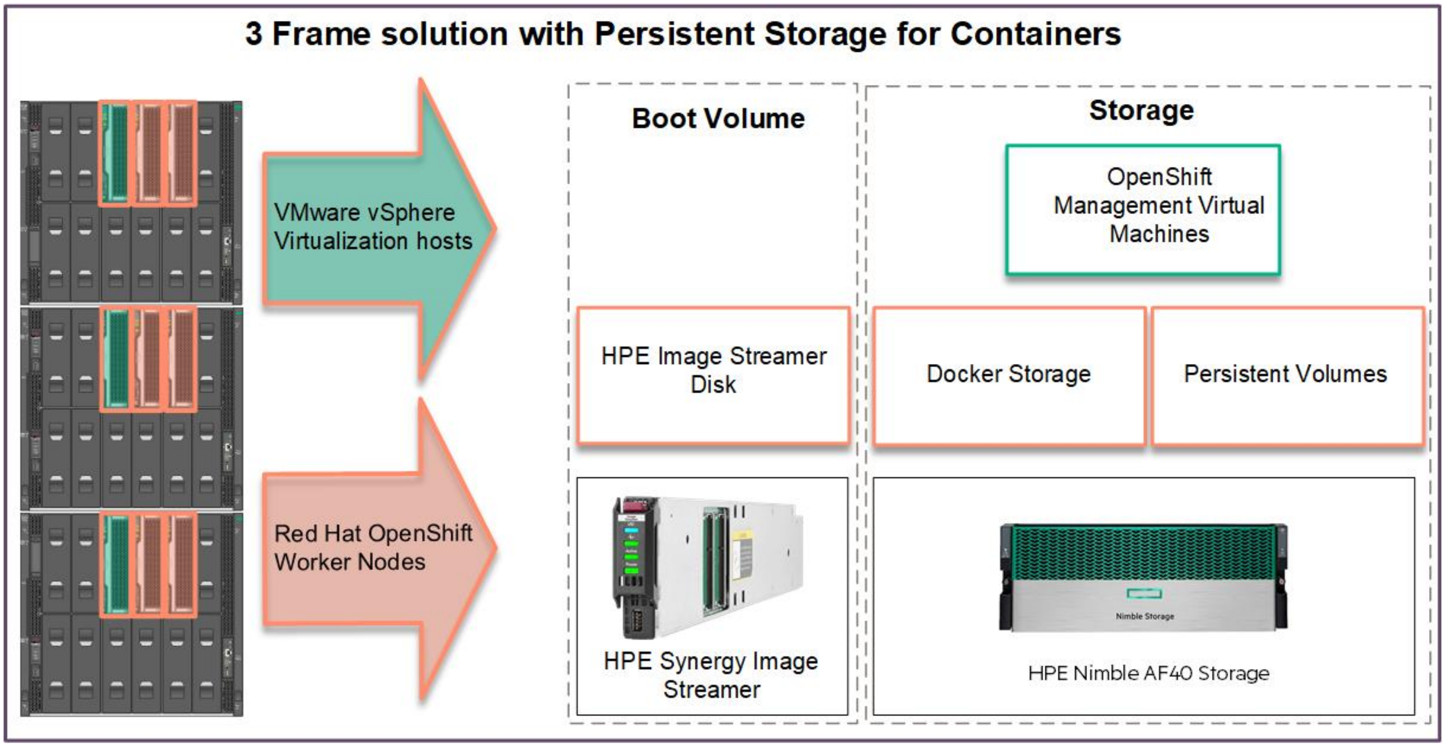

3 Frame solution with Persistent Storage for Containers
VMware vSphere Virtualization hosts
Red Hat OpenShift Worker Nodes
Boot Volume
HPE Image Streamer Disk
HPE Synergy Image Streamer
Storage
OpenShift Management Virtual Machines
Docker Storage
Persistent Volumes
Nimble Storage
HPE Nimble AF40 Storage

Figure 6. High level solution layout of storage resources

The solution environment assumes the presence of certain products and services to insure proper functionality:

- Ansible Engine
- LDAP/Active Directory
- DHCP
- DNS
- NTP
- TFTP/PXE

Figure 7 shows the configuration of the racks used to build this solution. For simplicity, the master and remote HPE Nimble Storage arrays are shown in the same rack. In a production environment, storage would be separated into separate physical racks with physical location being determined by the level of protection desired.

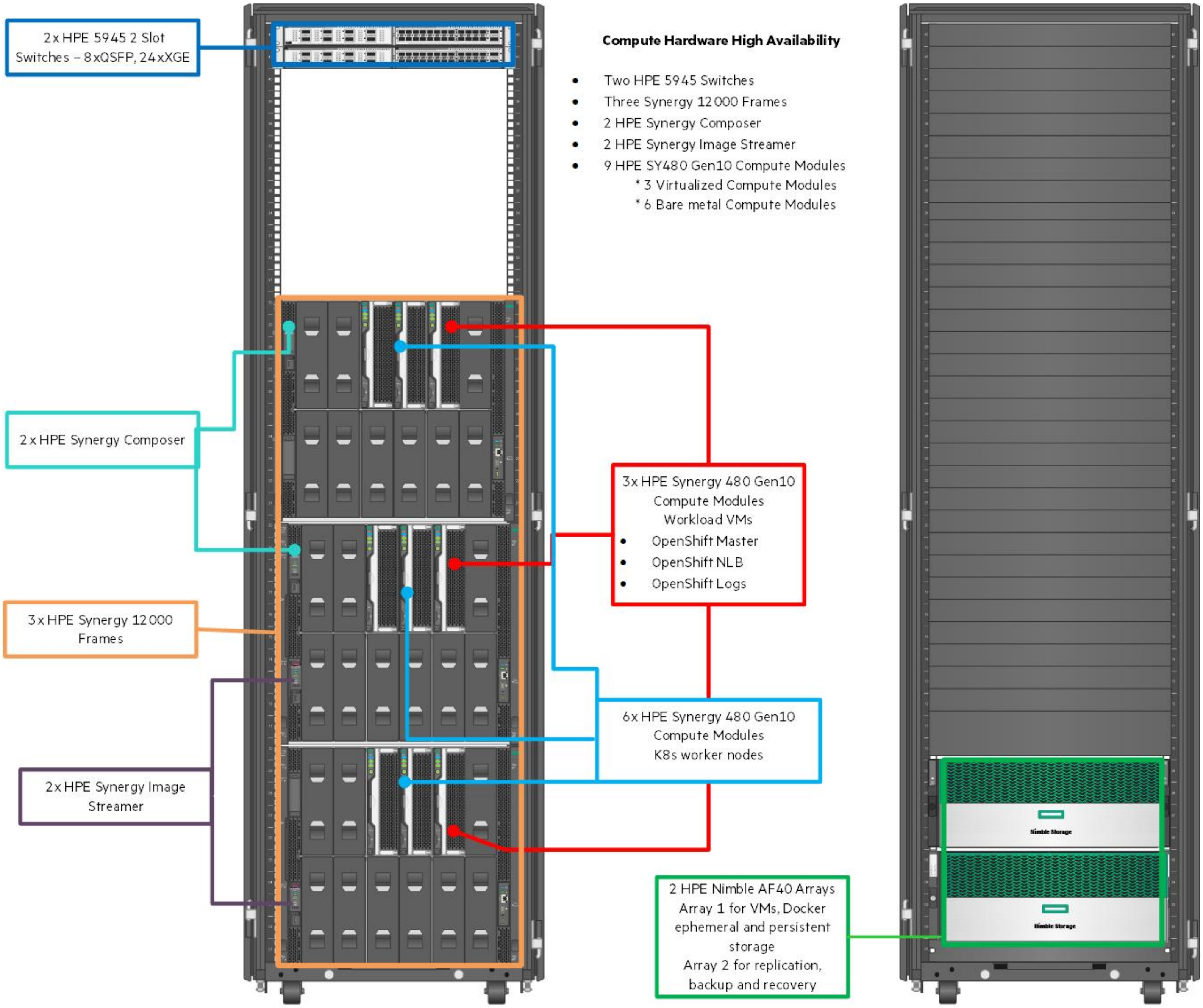

2 x HPE 5945 2 Slot Switches – 8 xQSFP, 24 xXGE
Compute Hardware High Availability
• Two HPE 5945 Switches
• Three Synergy 12000 Frames
• 2 HPE Synergy Composer
• 2 HPE Synergy Image Streamer
• 9 HPE SY480 Gen10 Compute Modules
* 3 Virtualized Compute Modules
* 6 Bare metal Compute Modules
2 x HPE Synergy Composer
3x HPE Synergy 480 Gen10 Compute Modules Workload VMs
• OpenShift Master
• OpenShift NLB
• OpenShift Logs
3 x HPE Synergy 12000 Frames
6x HPE Synergy 480 Gen10 Compute Modules K8s worker nodes
2 x HPE Synergy Image Streamer
2 HPE Nimble AF40 Arrays Array 1 for VMs, Docker ephemeral and persistent storage Array 2 for replication, backup and recovery
Nimble Storage
Nimble Storage

Figure 7. Front view of the solution with individual components highlighted

Solution components

Hardware

The following hardware components were utilized in this Reference Configuration as listed in Table 2.

Table 2. Components utilized in the creation of this solution

Component	Qty	Description
HPE Synergy 12000 Frame	3	Three (3) HPE Synergy 12000 Frames house the infrastructure used for the solution
HPE Synergy Composer	2	Two (2) HPE Synergy Composers for core configuration and lifecycle management for the Synergy components
HPE Synergy Image Streamer	2	A total of two (2) HPE Synergy Image Streamers provide Red Hat Enterprise Linux to the worker nodes in the solution
HPE Virtual Connect SE 40Gb F8 Module	2	A total of two (2) HPE Virtual Connect SE 40Gb F8 Modules provide network connectivity into and out of the frames
HPE Synergy 480 Gen10 Compute Module	9	Three (3) virtualized and six (6) bare metal hosts as described later in this document
HPE Nimble Storage AF40	2	One array for virtual machines, Docker storage and persistent volumes, and One array for replication, backup and recovery of configuration files and persistent data
HPE FlexFabric 5945 2-Slot Switch	2	Each switch contains one (1) each of the HPE 5930 modules listed below:
HPE 5930 24p SFP+ and 2p QSFP+ Module	2	• One module per HPE FlexFabric 2-Slot Switch
HPE 5930 8p QSFP+ Module	2	• One module per HPE FlexFabric 2-Slot Switch

HPE Synergy

HPE Synergy, the first platform built from the ground up for composable infrastructure, empowers IT to create and deliver new value instantly and continuously. This single infrastructure reduces operational complexity for traditional workloads and increases operational velocity for the new breed of applications and services. Through a single interface, HPE Synergy composes compute, storage, and fabric pools into any configuration for any application. It also enables a broad range of applications from bare metal to virtual machines to containers, and operational models like hybrid cloud and DevOps. HPE Synergy enables IT to rapidly react to new business demands.

HPE Synergy Frames contain a management appliance called the HPE Synergy Composer which hosts HPE OneView. HPE Synergy Composer manages the composable infrastructure and delivers:

- Fluid pools of resources, where a single infrastructure of compute, storage and fabric boots up ready for workloads and demonstrates self-assimilating capacity.
- Software-defined intelligence, with a single interface that precisely composes logical infrastructures at near-instant speeds; and demonstrates template-driven, frictionless operations.
- Unified API access, which enables simple line-of-code programming of every infrastructure element; easily automates IT operational processes; and effortlessly automates applications through infrastructure deployment.

HPE Synergy Composer provides the enterprise-level management to compose and deploy system resources to meet your application needs. This management appliance uses software-defined intelligence to aggregate compute, storage, and fabric resources in a manner that scales to your application needs, instead of being restricted to the fixed ratios of traditional resource offerings. HPE Synergy template-based provisioning enables fast time to service with a single point for defining compute module state, pooled storage, network connectivity, and boot image.

HPE OneView is a comprehensive unifying platform designed from the ground up for converged infrastructure management. A unifying platform increases the productivity of every member of the internal IT team across servers, storage, and networking. By streamlining processes, incorporating best practices, and creating a new holistic way to work, HPE OneView provides organizations with a more efficient way to work. It is designed for open integration with existing tools and processes to extend these efficiencies.

HPE OneView is instrumental for the deployment and management of HPE servers and enclosure networking. It collapses infrastructure management tools into a single resource-oriented architecture that provides direct access to all logical and physical resources of the solution.

Logical resources include server profiles and server profile templates, enclosures and enclosure groups, and logical interconnects and logical interconnect groups. Physical resources include compute modules, interconnects and storage modules.

The HPE OneView converged infrastructure platform offers a uniform way for administrators to interact with resources by providing a RESTful API foundation. The RESTful APIs enable administrators to utilize a growing ecosystem of integrations to further expand the advantages of the integrated resource model that removes the need for the administrator to enter and maintain the same configuration data more than once and keep all versions up to date. It encapsulates and abstracts many underlying tools behind the integrated resource model, so the administrator can operate with new levels of simplicity, speed, and agility to provision, monitor, and maintain the solution.

Within the context of the solution, HPE OneView for Synergy is utilized to:

- Configure the profiles of the HPE Synergy Compute Modules, resulting in a complexity reduction of needing to manually perform more than 100 manual steps to running one Ansible playbook, and a time reduction from over an hour, to 15 minutes.
- Apply and maintain compliance for firmware across the HPE Synergy infrastructure.
- Configure networking from the HPE Synergy Compute Modules to internal and outbound destinations.

HPE Synergy Image Streamer

HPE Synergy Image Streamer implements rapid image/application changes to multiple compute modules in an automated manner. HPE Synergy Image Streamer works with HPE Synergy Composer to rapidly deploy and update multiple physical compute modules. Operating environment images for bare-metal use might boot directly into a running OS, or VM hosts might perform quick image changeovers. “Infrastructure-as-code” capability enables fast delivery of applications and services, including the ability to perform rapid workload switching (using Linux, VMware® ESX, or Microsoft® Windows®). Enhanced profiles provide true stateless images, which integrate the server hardware configuration with operating environment images. Enhanced profiles are stored in redundant image repositories and are automatically integrated for simplicity of use. The unified API enables integration, automation, and customization of operations and applications with HPE Synergy Image Streamer.

HPE Synergy Image Streamer was used in this solution to provide Red Hat Enterprise Linux images to the physical OpenShift Container Platform worker nodes and to provide VMware vSphere images for virtualized hosts.

HPE Synergy 480 Gen10 Compute Module

The HPE Synergy 480 Gen10 Compute Module delivers an efficient and flexible two-socket server to support the most demanding workloads. Powered by Intel® Xeon® Scalable Family of processors, up to 3TB DDR4, and large storage capacity within a composable architecture. HPE Synergy 480 Gen10 Compute Module:

- Is the most secure server with exclusive HPE Silicon Root of Trust. Protect your applications and assets against downtime associated with hacks and viruses.
- Offers customer choice for greater performance and flexibility with Intel Xeon Scalable Family of processors on the Synergy 480 Gen10 architecture
- Offers Intelligent System Tuning with processor smoothing and workload matching to improve processor throughput/overall performance up to 8% over previous generation.
- Features a maximum memory footprint of 3TB for large in-memory database and analytic applications.
- Features a hybrid HPE Smart Array for both RAID and HBA zoning in a single controller.

The HPE Synergy 480 Gen10 provides the needed compute to power this solution running both Red Hat Virtualization for the core management pieces of Red Hat OpenShift and Red Hat Enterprise Linux to host the worker nodes.

The bill of materials found in Appendix A of this document outlines the configuration of the HPE Synergy Compute Modules used in this solution.

HPE Nimble Storage AF40

HPE Nimble Storage All Flash Arrays combine a flash-efficient architecture with HPE InfoSight predictive analytics to achieve fast, reliable access to data and 99.9999% guaranteed availability.[1] Radically simple to deploy and use, the arrays are cloud-ready, providing data mobility to the cloud through HPE Cloud Volumes. Your storage investment made today will support you well into the future, thanks to the technology and business-model innovations. HPE Nimble Storage All Flash Arrays include all-inclusive licensing, easy upgrades, and flexible payment options – while also being future-proofed for new technologies, such as NVMe and SCM.

HPE Nimble Storage abstracts a rich feature set to Red Hat OpenShift Container Platform to enable a breadth of modern use cases. Features are exposed as parameters to a Kubernetes StorageClass that users may leverage through Persistent Volume Claims.

Features include:

- Advanced lifecycle controls to enable ephemeral clones and volume annotation through custom metadata
- Application optimized performance policies and quality of service (QoS) controls
- Volume placement directives, per cluster or StorageClass, to enforce performance caps, storage capacity or media (Flash or Hybrid)
- Protection templates to fulfill data protection SLA/SLO defined by the business, including replication to HPE Cloud Volumes
- Provision data at rest (DAR) encrypted volumes to meet compliance and regulations
- Satisfy a wide range of stateful application by provisioning volumes up to 127 TB – either thick or thin-provisioned
- Toggle data reduction features such as variable block size, deduplication, and adaptive compression per workload
- Enable users to clone and restore Persistent Volume Claims using native tools and APIs
- Import existing Nimble volumes into a container from traditional virtualized environments when modernizing applications

Each of the features, or combinations of features enable popular use cases in DevOps centric and cloud-native environments. They include:

- CI/CD – Applications are being developed and productized faster than ever. Automation reduces development cycles to minutes to deploy a new feature in order to stay competitive. HPE Nimble Storage lifecycle controls, cloning, and data offloading allow developers to integrate production data into CI/CD systems to perform tests on an entire microservice architected application, stateless and stateful.
- Lift & Shift – Organizations looking to modernize existing applications are often stuck with data in virtual machines or on bare-metal servers. If the data could be migrated to an HPE Nimble Storage array in its original environment, the container integration would provide rich functionality to leverage existing volumes in iterative migration workflows using snapshots and clones.
- IT Operations – Many IT shops today have discovered the benefits of standardizing IT applications in containers which include radically simplified security, lifecycle management, and reliability. The majority of IT applications are either stateful themselves or require parts of it to be on persistent storage, such as a transactional databases or fixed content store. With HPE Nimble Storage, it is very easy to define a set of defaults for the stateful workloads and enjoy completely hassle-free operations for containers.
- CaaS – Large organizations are providing clusters of Red Hat OpenShift Container Platform to their organizations for different purposes. It could either be for subsidiaries or the simple segregation of production, test, stage, and QA. With HPE Nimble Storage placement controls, QoS and capacity limits, operations teams may partition resources from a Nimble array group for different purposes and co-host application workloads without additional investments.
- Big Data – AI and machine learning workloads require vast amounts of data, low-latency storage, and the ability to iterate computations on a given dataset. Data may reside in multiple sources and it may be challenging to run queries on an ever-changing dataset. With HPE Nimble Storage, the user may use clones of a larger dataset that is easy to revert to and create reproducible results. Once a query has been fine-tuned, the whole set may be either refreshed from its source or destroyed depending on the nature of the use case.

[1] HPE Get Six Nines Guarantee: http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00026086enw

Solution software

Red Hat Enterprise Linux

Red Hat Enterprise Linux Server powers the applications that run your business with the control, confidence, and freedom that comes from a consistent foundation across hybrid cloud deployments. As the premier platform provider for enterprise workloads, Red Hat works side by side with engineers from major hardware vendors and cloud providers to make sure that the operating system takes full advantage of the latest innovations. This leadership with partners, as well as Red Hat's influence and contributions to upstream communities, provides a stable, secure, and performance driven foundation for the applications that run the business of today and tomorrow. Red Hat Enterprise Linux is at the core of this solution, each of the Red Hat OpenShift Container Platform control plane nodes running as virtual machines are running Red Hat Enterprise Linux. This solution is designed to use dedicated physical servers for the Kubernetes application nodes. Each Kubernetes application node is a dedicated SY480 Compute Module running Red Hat Enterprise Linux 7.6 for maximum application and pod scalability.

VMware vSphere Virtualization

The Red Hat OpenShift Container Platform control plane, including the master, infrastructure, and load balancer roles, are deployed on virtual machines that are distributed across a three-node vSphere cluster. When iSCSI storage connectivity is utilized it is possible to create worker nodes as virtual machines.

Note

Hewlett Packard Enterprise has documented the use of Red Hat Enterprise Virtualization as a hypervisor as part of this solution. Instructions are available at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere.

Red Hat OpenShift Container Platform

Red Hat OpenShift Container Platform unites developers and IT operations on a single platform to build, deploy, and manage applications consistently across hybrid cloud and multi-cloud infrastructures. Red Hat OpenShift helps businesses achieve greater value by delivering modern and traditional applications with shorter development cycles and lower operating costs. Red Hat OpenShift is built on open source innovation and industry standards, including Kubernetes and Red Hat Enterprise Linux, the world's leading enterprise Linux distribution.[2]

Sysdig Agent

The Sysdig Agent is installed as part of the Sysdig Cloud-Native Intelligence Platform, Sysdig Secure, and Sysdig Monitor, in order to offer unified container security, monitoring, and forensics for container and Kubernetes based environments. Sysdig Agent moves the data collected to the Sysdig SaaS Platform where the security and monitoring services will be available to the user based on the subscription model they have selected.

[2] redhat.com/cms/managed-files/cl-openshift-container-platform-datasheet-f9695kc-201711-en.pdf

Red Hat OpenShift Container Platform layout

Figure 8 highlights how the individual Red Hat OpenShift Container Platform pieces are laid out within the solution. Three (3) application/worker nodes are shown for clarity, but six (6) nodes were tested in the solution. Virtual worker nodes can also be deployed as part of the solution.

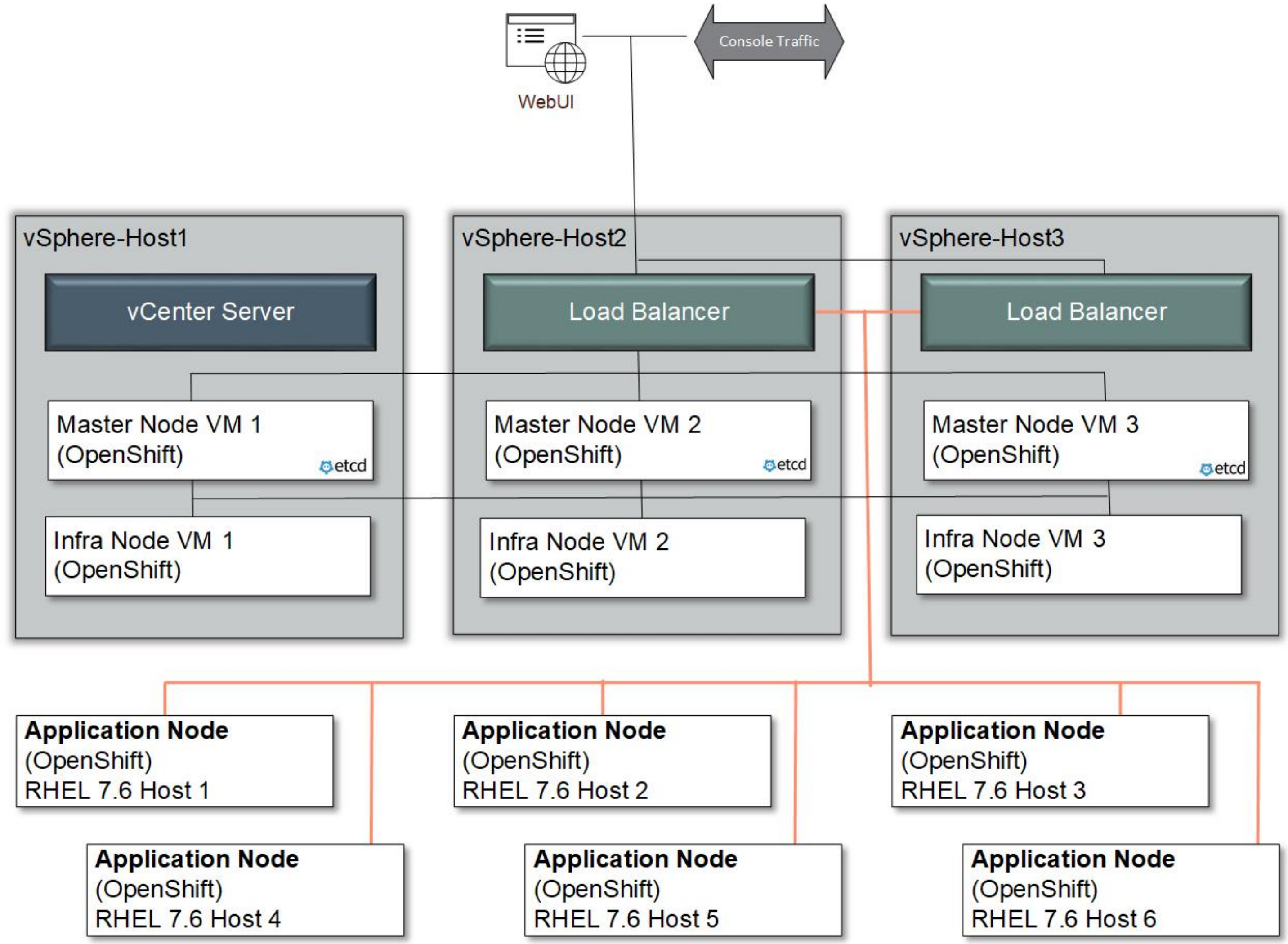

WebUI
Console Traffic
vSphere-Host1
vCenter Server
Master Node VM 1 (OpenShift)
etcd
Infra Node VM 1 (OpenShift)
vSphere-Host2
Load Balancer
Master Node VM 2 (OpenShift)
etcd
Infra Node VM 2 (OpenShift)
vSphere-Host3
Load Balancer
Master Node VM 3 (OpenShift)
etcd
Infra Node VM 3 (OpenShift)
Application Node (OpenShift) RHEL 7.6 Host 1
Application Node (OpenShift) RHEL 7.6 Host 2
Application Node (OpenShift) RHEL 7.6 Host 3
Application Node (OpenShift) RHEL 7.6 Host 4
Application Node (OpenShift) RHEL 7.6 Host 5
Application Node (OpenShift) RHEL 7.6 Host 6

Figure 8. Red Hat OpenShift Container Platform solution component layout

Best practices and configuration guidance for the solution

This section discusses the high-level cabling and configuration of the solution hardware and software. For a detailed explanation of how to build and deploy the entire solution stack, refer the deployment guide and accompanying Ansible deployment playbooks at, https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere.

Solution cabling

Figure 9 describes the cabling configuration of the three (3) Synergy 12000 Frames as well as the HPE FlexFabric 5945 switches and Intelligent Resilient Fabric (IRF) within the context of this solution. These cables carry frame management, inter-frame, and interconnect traffic between frames.

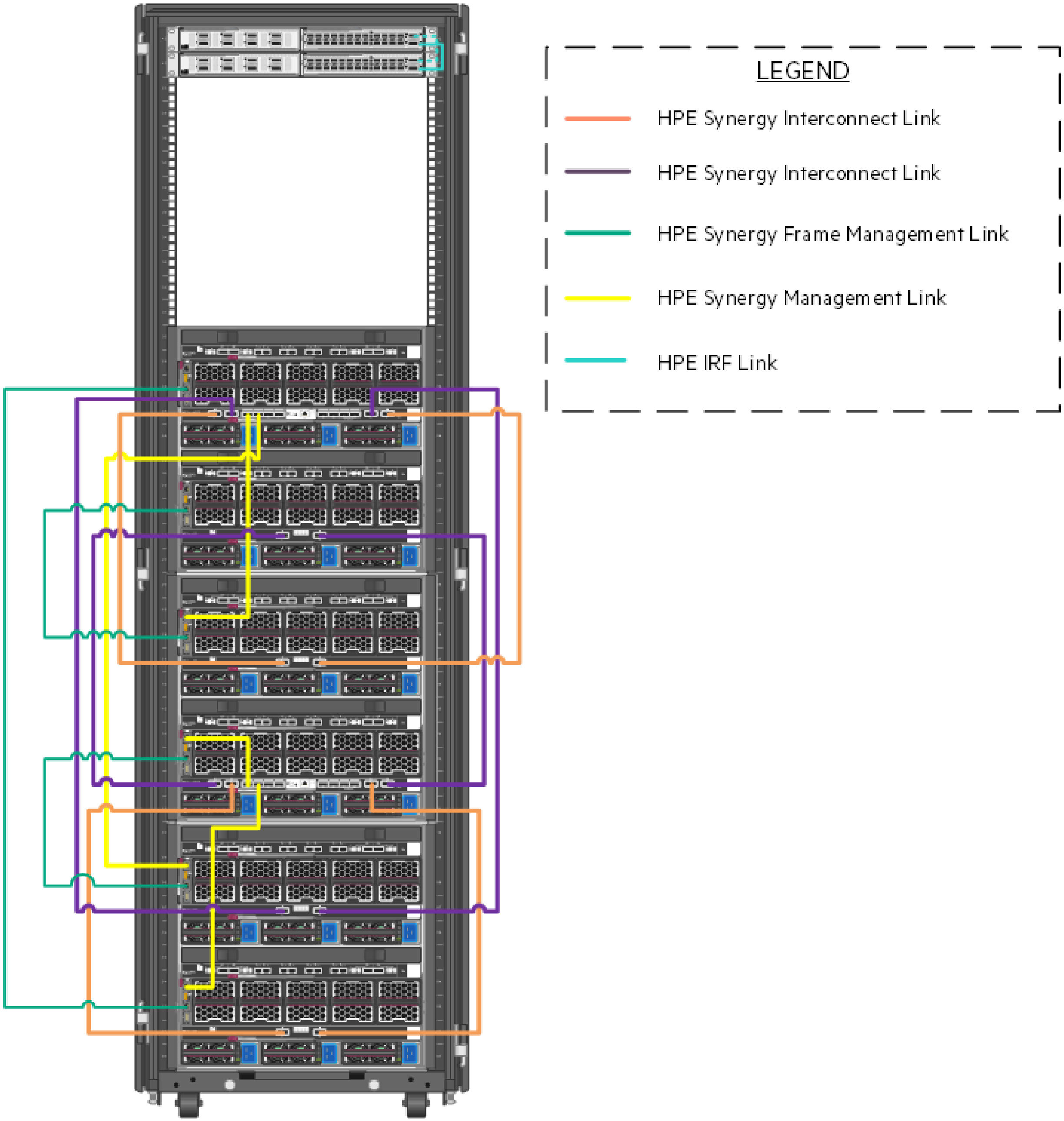

LEGEND
HPE Synergy Interconnect Link
HPE Synergy Interconnect Link
HPE Synergy Frame Management Link
HPE Synergy Management Link
HPE IRF Link

Figure 9. Frame and switch cabling within the solution

Networking

Figure 10 documents the cabling of the solution from the HPE Virtual Connect SE 40 Gb modules to the switches as well as the traffic carried on each connection. In Synergy environments, this first layer of switching may be end of row rather than top of rack switching. Top of rack was used for solution development to facilitate the connectivity to the HPE Nimble Storage iSCSI NICs in an adjacent rack. Switching is flexible based on the installation environment.

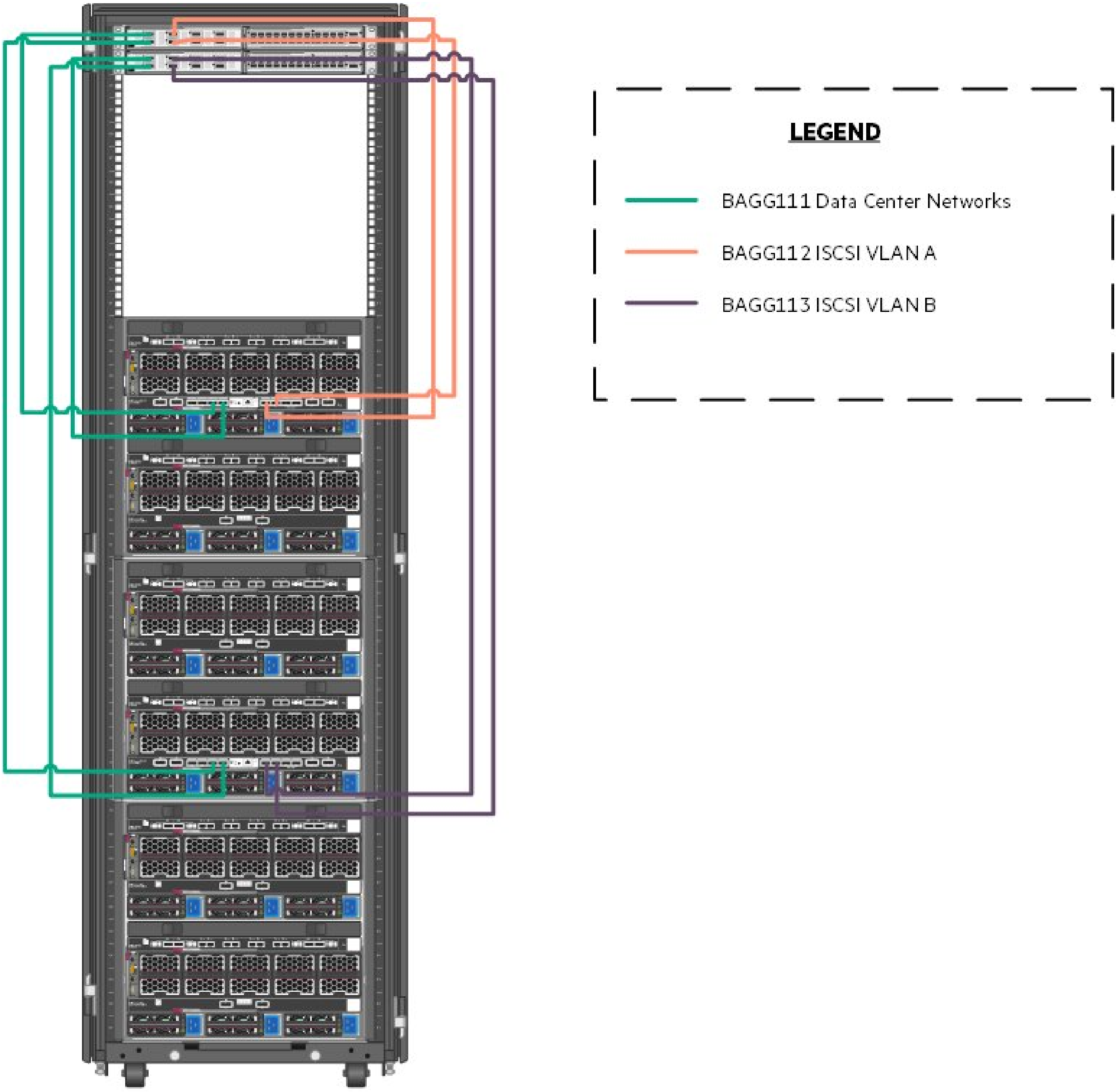


Figure 10. Network cabling from the HPE Synergy 12000 Frames to the switches

Table 3 describes the configuration of the networks as defined within HPE OneView for HPE Synergy and the bandwidth associated with each network.

Table 3. Networks defined within HPE Synergy Composer for this solution

Network Name	Type	VLAN Number	Purpose	Requested Bandwidth (Gb)	Maximum Bandwidth (Gb)
Management	Ethernet	1193	Solution management	5	20
Synergy Management	Ethernet	193	Image Streamer, OneView for Synergy	1	20
Data_Center	Ethernet	2193	Production Application, authentication and other user networks	5	20
ISCSI_VLAN_A	Ethernet	3193	ISCSI VLAN A	8	20
ISCSI_VLAN_B	Ethernet	3194	ISCSI VLAN B	8	20
Deployment	Ethernet	500	Deployment Network	2	20

Table 4 describes the cabling of the HPE Virtual Connect SE 40 Gb F8 Modules for Synergy to the switches and highlights what networks are carried on the connections. All Ethernet networks as described in Table 3 are carried on the Uplink Set labeled “Network”.

Table 4. Networks used in this solution

Uplink Set	Synergy Source	Switch Destination
Network	Enclosure 1 Port Q3	FortyGigE1/1/1
	Enclosure 1 Port Q4	FortyGigE2/1/1
	Enclosure 2 Port Q3	FortyGigE1/1/2
	Enclosure 2 Port Q4	FortyGigE2/1/2
iSCSI_SAN_A	Enclosure 1 Port Q5	FortyGigE1/1/5
	Enclosure 1 Port Q6	FortyGigE1/1/6
iSCSI_SAN_B	Enclosure 2 Port Q5	FortyGigE2/1/5
	Enclosure 2 Port Q6	FortyGigE2/1/6

Table 5 describes the bridge aggregation groups and the networks they contain.

Table 5. VLAN definitions by bridge aggregation group

Network Function	VLAN Number	Bridge Aggregation Group
Deployment	500	111
Solution_Management	1193	111
Data_Center	2193	111
iSCSI_A	3193	112
iSCSI_B	3194	113

By utilizing HPE Synergy, the non-storage networks within the solution are able to traverse the Synergy infrastructure in an east-west fashion across high speed, low latency links, both within and between HPE Virtual Connect Modules. In particular, communication between the core OpenShift management pieces remains within the HPE Synergy Frames. Figure 11 places the network configuration in the context of the HPE Synergy Compute Modules utilized for management functions. In this part of the solution, the compute modules are virtualized and as such, Ethernet traffic traverses a virtual switch prior to the Synergy Link Aggregation Group (LAG).

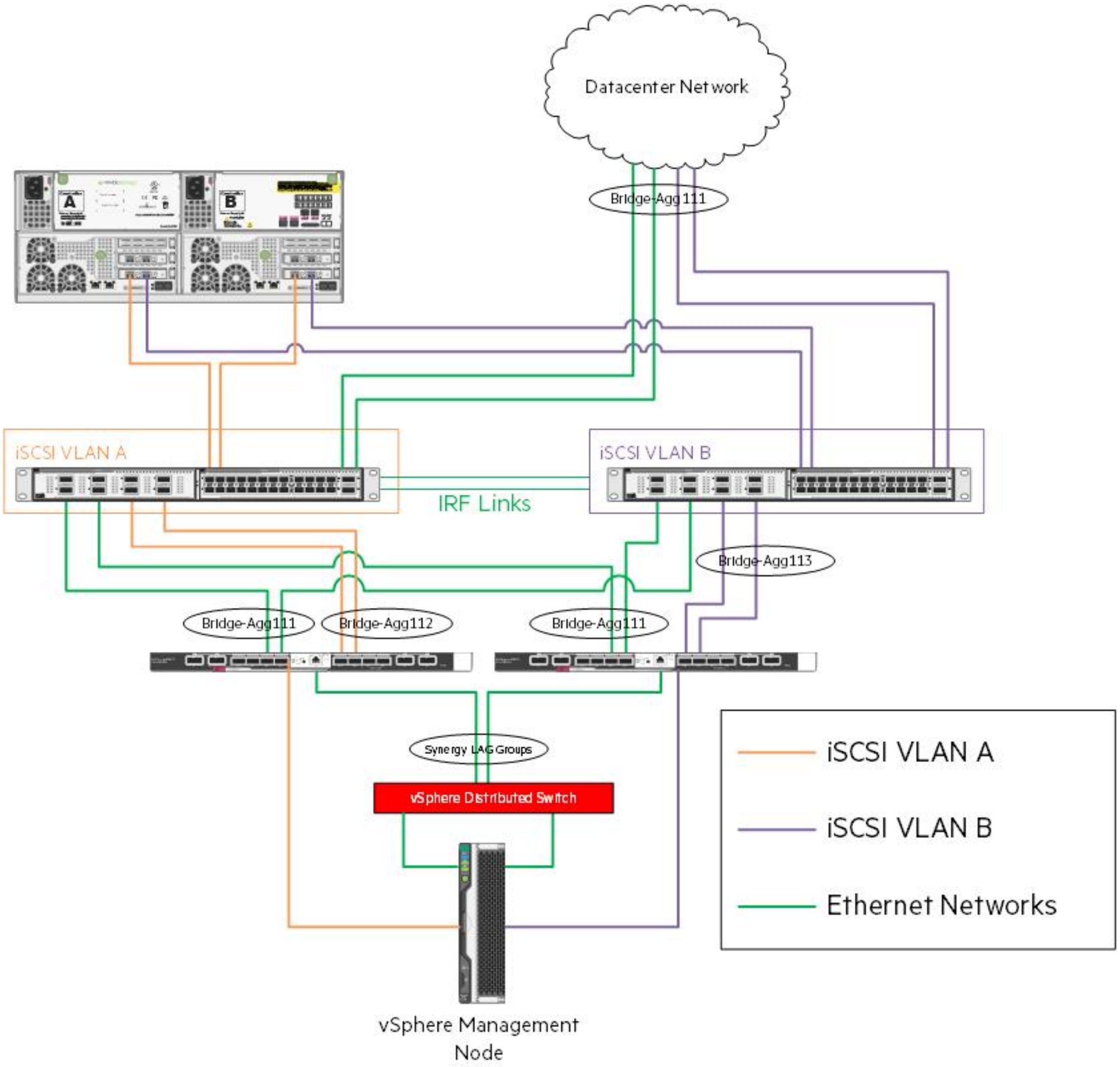


Figure 11. Network configuration for virtualized compute modules

Figure 12 presents the same diagram but in the context of the OpenShift worker nodes. These compute modules run Red Hat Enterprise Linux and do not have a virtual switch. iSCSI traffic in both cases does not pass through a LAG, but rather is untagged from source to final destination.

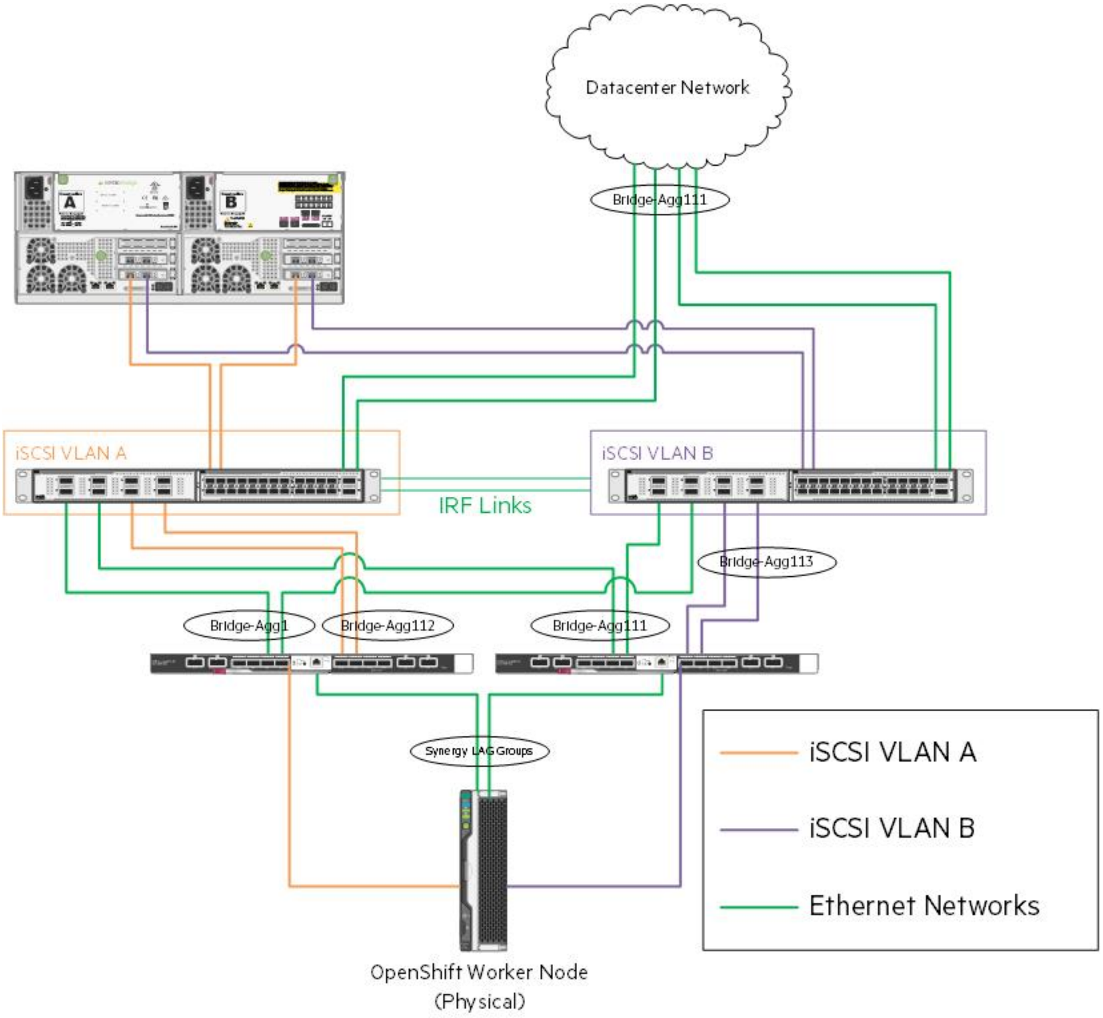


Figure 12. Network configuration for OpenShift worker nodes

Storage

This section describes the configuration of the HPE Nimble Storage AF40 in the context of this solution. The HPE Nimble Storage AF40 provides the shared storage for:

- Docker storage
- Virtual machines including storage for VMware vCenter Server appliance
- OpenShift Container Registry
- Persistent volumes

Figure 13 describes the logical storage layout used in the solution. Local storage is used for the operating system installation the virtualized hosts while HPE Image Streamer provides the OS volumes to the bare metal worker nodes. The HPE Nimble Storage AF40 provides dedicated and shared volumes as outlined in the figure. Storage redundancy is covered in the deployment guide at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere.

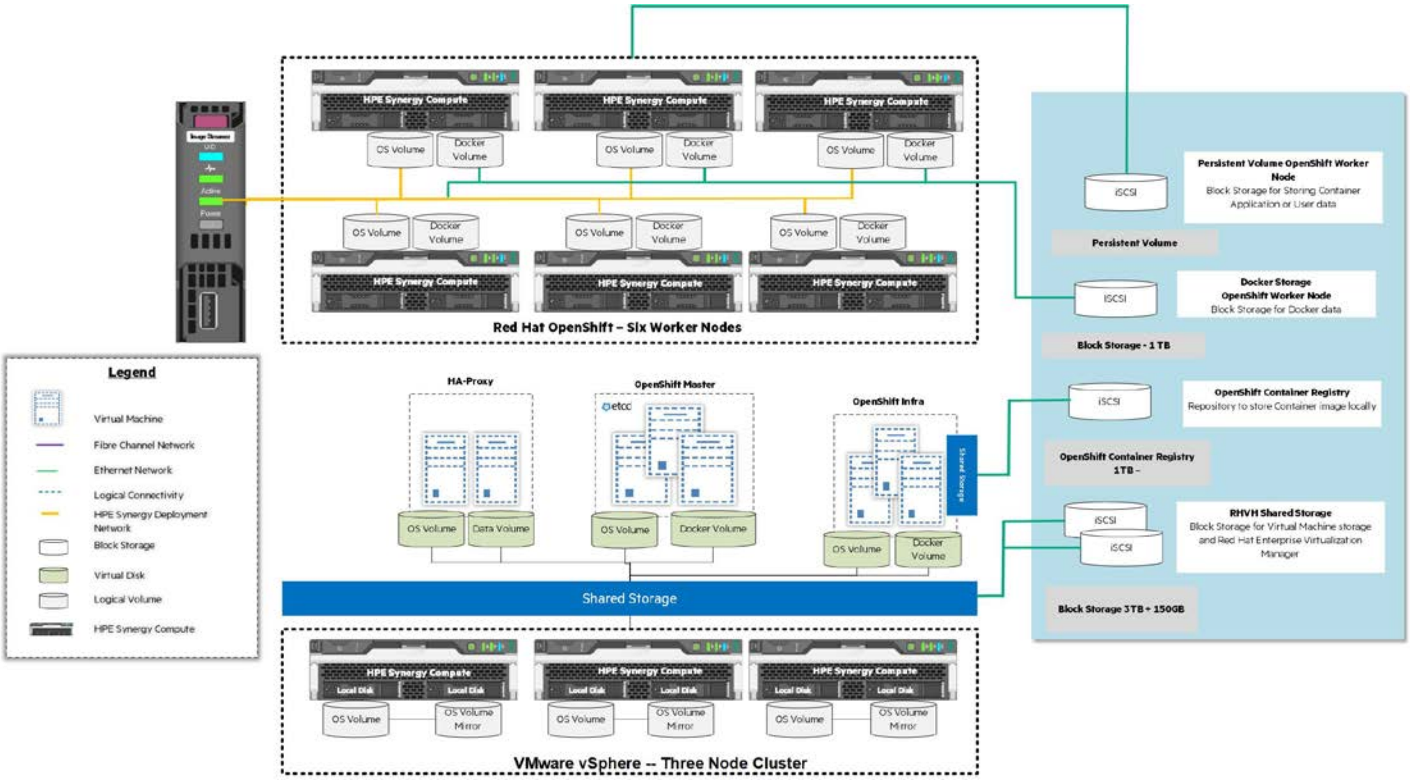

HPE Synergy Compute
OS Volume
Docker Volume
Red Hat OpenShift – Six Worker Nodes
Legend
Virtual Machine
Fibre Channel Network
Ethernet Network
Logical Connectivity
HPE Synergy Deployment Network
Block Storage
Virtual Disk
Logical Volume
HPE Synergy Compute
HA-Proxy
Data Volume
OpenShift Master
etcd
OpenShift Infra
Shared Storage
Local Disk
OS Volume Mirror
VMware vSphere -- Three Node Cluster
iSCSI
Persistent Volume OpenShift Worker Node
Block Storage for Storing Container Application or User data
Persistent Volume
Docker Storage
OpenShift Worker Node
Block Storage for Docker data
Block Storage - 1 TB
OpenShift Container Registry
Repository to store Container image locally
OpenShift Container Registry
1TB -
RHVH Shared Storage
Block Storage for Virtual Machine storage and Red Hat Enterprise Virtualization Manager
Block Storage 3TB + 150GB

Figure 13. Logical storage layout within the solution

Table 6 lists all volumes used within the solution and highlights what storage provides the capacity for each function.

Table 6. Volumes used in this solution

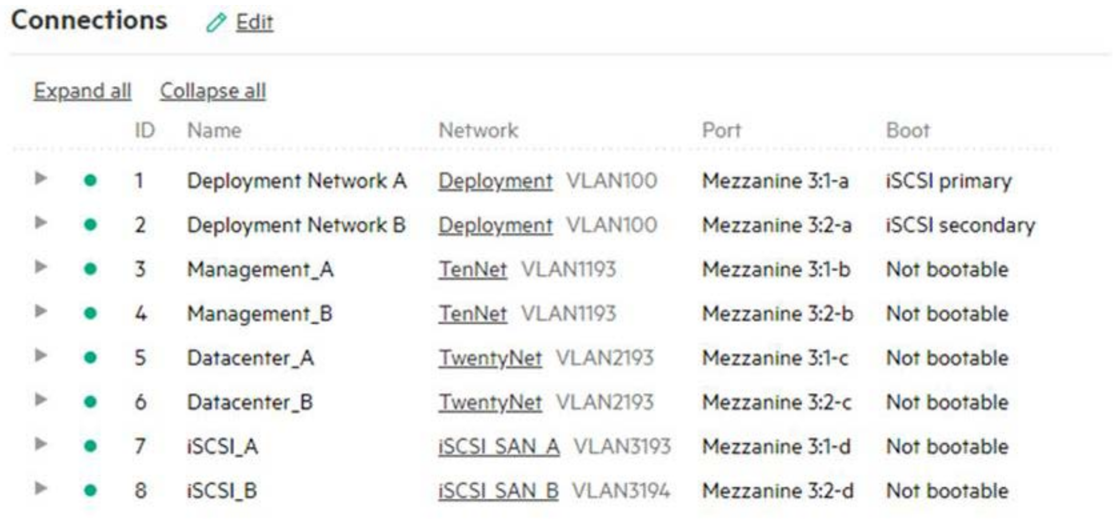
Volume/Disk Function	Qty	Size	Source	Hosts	Shared/Dedicated
Hypervisor	3	40GB	HPE Image Streamer	vSphere hosts	Dedicated
Operating System	6	40GB	HPE Image Streamer	OpenShift worker nodes	Both
Virtual Machine Hosting	1	3TB	HPE Nimble	vSphere hosts	Shared
Persistent Application Data	N	App Specific	HPE Nimble	Worker nodes	Dedicated
Docker Local Storage	6	1TB	HPE Nimble	Worker nodes	Dedicated
VMware vCenter Server	1	150GB	HPE Nimble	vSphere hosts or external	Shared
OpenShift Container Registry	1	1TB	HPE Nimble	Infrastructure node	Shared

The array used in the creation of this solution was built to suit configuration testing. Customer requirements around application performance and capacity should be taken into account when selecting an HPE Nimble Storage array for production Red Hat OpenShift Container Platform environments.

Server profiles

HPE Synergy Composable Infrastructure using HPE Virtual Connect provides the construct of a server profile. A server profile allows a suite of configuration parameters, including network and SAN connectivity, BIOS tuning, boot order configuration, local storage configuration, and more to be templatized and applied programmatically to compute resources. These templates are the key to delivering the "infrastructure as code" capabilities of the HPE Synergy platform. For the purpose of this solution, a single template was created that was applied to all compute modules.

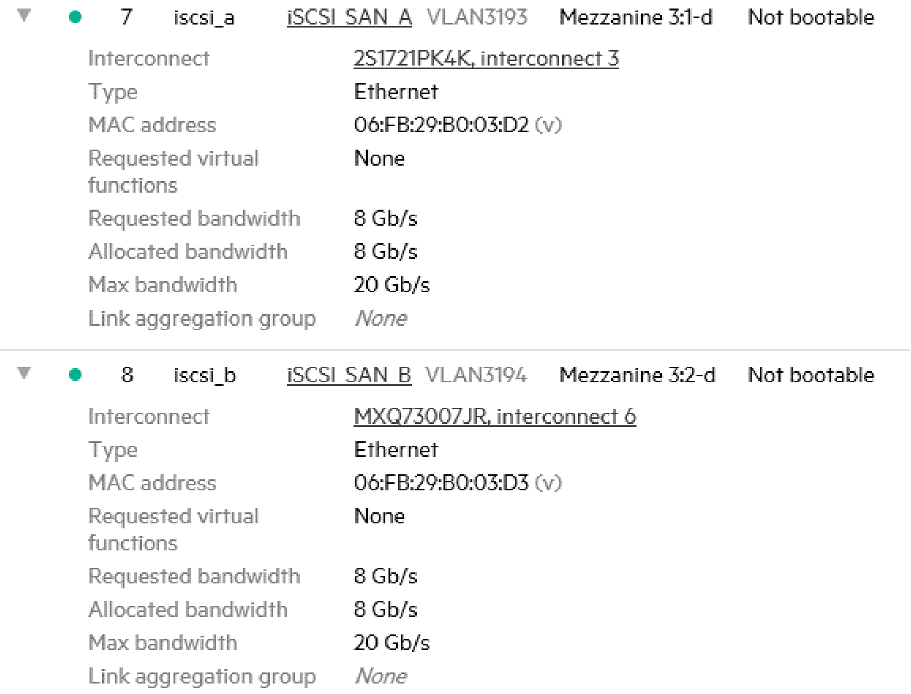
The critical items configured as part of a template supporting Red Hat OpenShift Container Platform are the network connections and storage. Figure 14 describes the configuration of the network interfaces as part of the profile template for the virtualized management as well as the physical worker nodes. There are 8 redundant Ethernet networks that are defined. Deployment networks are used by HPE Synergy Image Streamer to deploy operating systems using Golden images.



		ID	Name	Network	Port	Boot
▶	●	1	Deployment Network A	Deployment VLAN100	Mezzanine 3:1-a	iSCSI primary
▶	●	2	Deployment Network B	Deployment VLAN100	Mezzanine 3:2-a	iSCSI secondary
▶	●	3	Management_A	TenNet VLAN1193	Mezzanine 3:1-b	Not bootable
▶	●	4	Management_B	TenNet VLAN1193	Mezzanine 3:2-b	Not bootable
▶	●	5	Datacenter_A	TwentyNet VLAN2193	Mezzanine 3:1-c	Not bootable
▶	●	6	Datacenter_B	TwentyNet VLAN2193	Mezzanine 3:2-c	Not bootable
▶	●	7	iSCSI_A	iSCSI SAN A VLAN3193	Mezzanine 3:1-d	Not bootable
▶	●	8	iSCSI_B	iSCSI SAN B VLAN3194	Mezzanine 3:2-d	Not bootable

Figure 14. Network interfaces as defined by the server profile template

Figure 15 describes the iSCSI connections of an individual profile regardless of whether it is virtual or physical. It is possible to define the type for this connection as either ISCSI or Ethernet. This solution uses Ethernet which uses software iSCSI initiators rather than hardware-based offload. The solution calls for bandwidth of 8 Gb per adapter but is flexible based on customer need.



▼ ● 7 iscsi_a iSCSI_SAN_A VLAN3193 Mezzanine 3:1-d Not bootable

Interconnect	2S1721PK4K, interconnect 3
Type	Ethernet
MAC address	06:FB:29:B0:03:D2 (v)
Requested virtual functions	None
Requested bandwidth	8 Gb/s
Allocated bandwidth	8 Gb/s
Max bandwidth	20 Gb/s
Link aggregation group	*None*

▼ ● 8 iscsi_b iSCSI_SAN_B VLAN3194 Mezzanine 3:2-d Not bootable

Interconnect	MXQ73007JR, interconnect 6
Type	Ethernet
MAC address	06:FB:29:B0:03:D3 (v)
Requested virtual functions	None
Requested bandwidth	8 Gb/s
Allocated bandwidth	8 Gb/s
Max bandwidth	20 Gb/s
Link aggregation group	*None*

Figure 15. iSCSI configuration parameters as defined within the server profile template

Red Hat OpenShift virtual machines

Red Hat OpenShift Container Platform requires a number of redundant functions. These functions may be hosted on either physical compute modules or on virtual machines, both of which run Red Hat Enterprise Linux 7.6. For this solution, Hewlett Packard Enterprise chose to implement these functions as virtual machines. This approach reduces the amount of infrastructure required while introducing enhanced options for management and high availability. Three (3) HPE Synergy 480 Gen10 Compute Modules host the virtual machines shown in Figure 16. This figure also shows that the worker nodes run on bare metal. While the solution was tested with six (6) worker nodes, it is scalable to include many more. While worker nodes are depicted as physical resources in Figure 16, the hosts can be deployed as hypervisors and the workers can be deployed as virtual machines in variable quantities according to the container requirements.

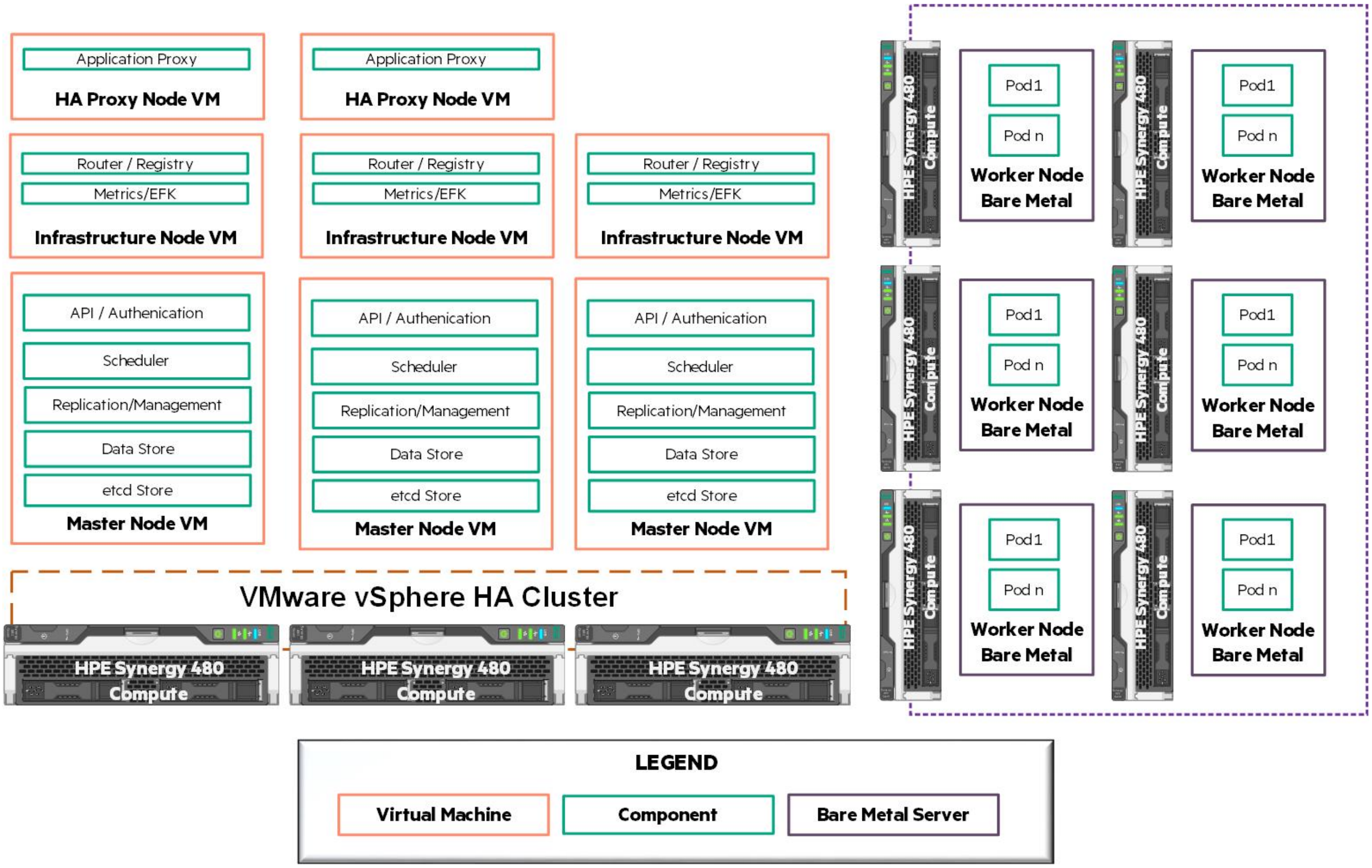


Figure 16. Red Hat OpenShift implementation with virtual machines for core management functions

The virtual environment is managed by VMware vCenter Server which runs as a virtual machine within the environment. This virtual machine appliance is deployed to a dedicated volume hosted on HPE Nimble Storage.

Software

This section describes the software versions utilized in the solution as well as noting any special installation or configuration requirements. For detailed descriptions of how to install and configure the software, refer the deployment guide and accompanying scripts found at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere. Table 7 lists the versions of required software used in the creation of this solution.

Table 7. Software versions used in the solution

Component	Version
Red Hat Enterprise Linux Server	7.6
vCenter, vSphere	6.7
Red Hat OpenShift Container Platform	3.11
Python	2.7.9 or above
Ansible version	2.6
HPE Nimble Storage Linux Toolkit	2.4.1.13
Nimble Kube Storage Controller	2.4.1
Sysdig Agent	0.90.2

Capacity and sizing

Sizing for a Red Hat OpenShift Container Platform environment varies depending on the requirements of the specific organization and type of deployment. This section discusses sizing considerations for Red Hat OpenShift Container Platform virtual machines, host requirements, and cluster sizing.

Red Hat OpenShift Container Platform role sizing

- Master – The minimum size for a physical or virtual machine running the master node is 4 vCPU and 16 GB RAM with a 40 GB disk space for /var, 1 GB disk space for /usr/local/bin, and 1 GB disk space for the system's temporary directory. Master nodes should be configured with an additional 1 CPU core and 1.5 GB RAM for each additional 1,000 pods. Three (3) master node virtual machines were deployed for this solution.
- Infrastructure nodes – Application and infrastructure nodes require a minimum of 1 vCPU and 8 GB RAM. The node requires at least 15 GB of disk space for /var/, 1 GB of disk space for /usr/local/bin, and 1 GB of disk space for the system's temporary directory. A total of three (3) infrastructure nodes were deployed for this solution.
- HAProxy – A total of two (2) HAProxy load balancer VMs were deployed for this solution. Default values for CPU and memory were utilized. In a production environment, the expectation is that a dedicated software or hardware load balancing solution will be implemented.

Red Hat OpenShift Container Platform cluster sizing

The number of application nodes in an OpenShift Cluster depends on the number of pods that an organization is planning to deploy. Red Hat OpenShift Container Platform can support the following:

- Maximum of 2000 nodes per cluster
- Maximum of 150,000 pods per cluster
- Maximum of 250 pods per node
- Maximum of 10 pods per CPU core

To determine the number of nodes required in a cluster, estimate the number of pods the organization is planning to deploy and divide by the maximum number of pods per node. For example, if the organization expects to deploy 5000 pods, then the organization should expect to deploy 20 application nodes with 250 pods per node (5000 / 250 =20). In this environment with a default configuration of six physical application nodes, the Red Hat OpenShift Cluster should be expected to support 1500 pods (250 pods x 6 nodes = 1500 pods).

For more information about Red Hat OpenShift Container Platform sizing, refer to the Red Hat OpenShift Container Platform Scaling and Performance Guide which can be found at, https://docs.openshift.com/container-platform/3.11/scaling_performance/index.html.

Solution deployment overview

In depth details of solution deployment are documented in the deployment guide and accompanying Ansible automation scripts at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/nimble-vsphere. This document also outlines about configuring OpenShift in a disconnected environment, Configuring Sysdig security, and Configuring Data protection for Red Hat OpenShift Cluster. Figure 17 shows the overall deployment flow. For more detailed flow instructions, refer to the deployment guide.

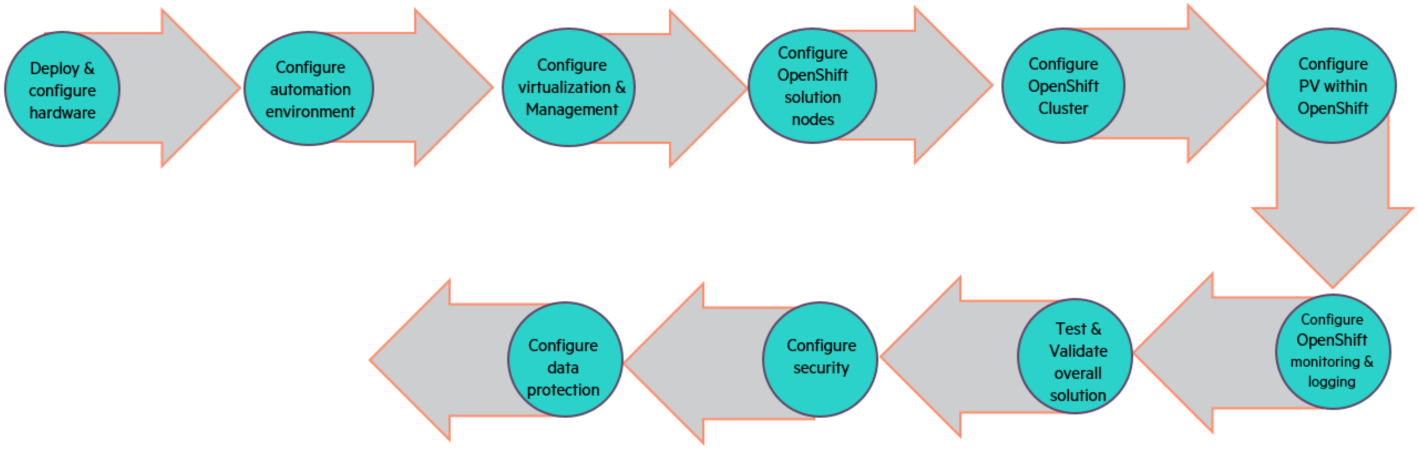

Deploy & configure hardware
Configure automation environment
Configure virtualization & Management
Configure OpenShift solution nodes
Configure OpenShift Cluster
Configure PV within OpenShift
Configure OpenShift monitoring & logging
Test & Validate overall solution
Configure security
Configure data protection

Figure 17. Red Hat OpenShift Solution Setup Overview

Summary

Red Hat OpenShift Container Platform on HPE Synergy provides an end-to-end fully integrated container solution that, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale. Red Hat OpenShift Container Platform provides organizations with a reliable platform for deploying and scaling container-based applications. HPE Synergy provides the flexible infrastructure you need to run the container platform to dynamically provision and scale applications, whether they run in VMs or containers, or are hosted on-premises, in the cloud, or in a hybrid environment.

This Reference Configuration demonstrates the following benefits of utilizing HPE Synergy for Red Hat OpenShift Container Platform:

- Automated initial installation and configuration of highly available vSphere hosts and the management virtual machine for the Red Hat OpenShift Container Platform, is reduced from more than 3 hours to approximately one hour, and complexity of the manual operation is reduced from needing to perform more than 500 steps, to running two Ansible play books.
- Automated deployment and configuration of physical worker nodes is reduced from 8 hours to under 20 minutes, and complexity of manual operation is reduced from needing to perform close to 300 steps, to running two Ansible play books.
- Automated host preparation of the Red Hat OpenShift Container Platform Nodes is reduced from up to 8 hours to as little as 20 minutes, and the complexity of the manual operation is reduced from performing more than three hundred steps, to one Ansible playbook.
- Deploying the management and infrastructure nodes on VMs to optimize resource usage while keeping the worker nodes on bare metal to optimize for performance.
- Using an enterprise grade storage solution such as HPE Nimble Storage for persistent storage with containers enables speed, portability, and agility for traditional enterprise applications and data.
- The HPE Synergy Composable Infrastructure solution provides a layered view of security controls. The objective of choosing this layered security view is to ensure that customers become aware of the depth of security risk that an infrastructure can have and also make them aware of the depth of defense that is built into the HPE design.
- Container platform security provided by Sysdig Cloud-Native Intelligence Platform, Sysdig Secure, and Sysdig Monitor, offer unified container security, monitoring, and forensics for container and Kubernetes based environments.
- Taking advantage of HPE Nimble storage array's data replication for protection and disaster recovery helps in faster and affordable recovery.
- Provides deployment capabilities for Red Hat OpenShift Container Platform in an offline fashion.

Appendix A: Bill of materials

The following bill of materials contains the core components utilized in the creation of this solution. Services, support, and software are not included in the BOM and should be customized based on customer needs.

Note

Part numbers are accurate at time of testing and are subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, consult with your HPE Reseller or HPE Sales Representative for more details. Refer to hpe.com/us/en/services/consulting.html.

Table A1. Bill of materials

Qty	Part number	Description
		Rack and Network Infrastructure
2	P9K10A	HPE 42U 600mmx1200mm G2 Kitted Advanced Shock Rack with Side Panels and Baying
2	P9K10A 001	HPE Factory Express Base Racking Service
2	H6J85A	HPE Rack Hardware Kit

Qty	Part number	Description
2	BW932A	HPE 600mm Rack Stabilizer Kit
2	BW932A B01	HPE 600mm Rack include with Complete System Stabilizer Kit
6	AF533A	HPE Intelligent Modular 3Ph 14.4kVA/CS8365C 40A/208V Outlets (6) C19/Horizontal NA/JP PDU
		HPE Synergy Composable Infrastructure
3	797740-B21	HPE Synergy 12000 Configure-to-order Frame with 1x Frame Link Module 10x Fans
4	779218-B21	HPE Synergy 20Gb Interconnect Link Module
2	794502-B23	HPE Virtual Connect SE 40Gb F8 Module for Synergy
2	804937-B21	HPE Synergy Image Streamer
3	798096-B21	HPE 6x 2650W Performance Hot Plug Titanium Plus FIO Power Supply Kit
2	804353-B21	HPE Synergy Composer
3	804938-B21	HPE Synergy Frame Rack Rail Kit
3	804942-B21	HPE Synergy Frame Link Module
1	804943-B21	HPE Synergy Frame 4x Lift Handles
1	859493-B21	Synergy Multi Frame Master1 FIO
1	859494-B22	Synergy Multi Frame Master2 FIO
8	804101-B21	HPE Synergy Interconnect Link 3m Active Optical Cable
2	720199-B21	HPE BladeSystem c-Class 40G QSFP+ to QSFP+ 3m Direct Attach Copper Cable
2	861412-B21	HPE Synergy Frame Link Module CAT6A 1.2m Cable
1	861413-B21	HPE Synergy Frame Link Module CAT6A 3m Cable
		Virtualized Hosts
3	871940-B21	HPE Synergy 480 Gen10 Configure-to-order Compute Module
3	873381-L21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) FIO Processor Kit
3	873381-B21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) Processor Kit
54	815097-B21	HPE 8GB (1x8GB) Single Rank x8 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
18	815098-B21	HPE 16GB (1x16GB) Single Rank x4 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
6	875478-B21	HPE 1.92TB SATA 6G Mixed Use SFF (2.5in) SC 3yr Wty Digitally Signed Firmware SSD
3	P01367-B1	HPE 96W Smart Storage Battery (up to 20 Devices) with 260mm Cable Kit
3	804424-B21	HPE Smart Array P204i-c SR Gen10 (4 Internal Lanes/1GB Cache) 12G SAS Modular Controller
3	777430-B21	HPE Synergy 3820C 10/20Gb Converged Network Adapter
		Worker Nodes
6	871943-B21	HPE Synergy 480 Gen10 6130 2P 64GB-R P204i-c SAS Performance Compute Module
6	873381-L21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) FIO Processor Kit
6	873381-B21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) Processor Kit
108	815097-B21	HPE 8GB (1x8GB) Single Rank x8 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
36	815098-B21	HPE 16GB (1x16GB) Single Rank x4 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
6	P01367-B1	HPE 96W Smart Storage Battery (up to 20 Devices) with 260mm Cable Kit
6	804424-B21	HPE Smart Array P204i-c SR Gen10 (4 Internal Lanes/1GB Cache) 12G SAS Modular Controller
6	777430-B21	HPE Synergy 3820C 10/20Gb Converged Network Adapter

Qty	Part number	Description
		Primary HPE Nimble Storage
1	Q8H41A	HPE Nimble Storage AF40 All Flash Dual Controller 10GBASE-T 2-port Configure-to-order Base Array
1	Q8B88B	HPE Nimble Storage 2x10GbE 2-port FIO Adapter Kit
1	Q8G27B	HPE Nimble Storage NOS Default FIO Software
1	Q8H47A	HPE Nimble Storage AF40 All Flash Array R2 11.52TB (24x480GB) FIO Flash Bundle
2	R0P84A	HPE Nimble Storage NEMA IEC 60320 C14 to C19 250V 15 Amp 1.8m FIO Power Cord
1	Q8F56A	HPE Nimble Storage 10GbE 2-port Spare Adapter
2	P9Q66A	HPE G2 IEC C20 Input/(8) C13 Expansion Outlets/PDU Extension Bar Kit
		Replication target HPE Nimble Storage
1	Q8H41A	HPE Nimble Storage AF40 All Flash Dual Controller 10GBASE-T 2-port Configure-to-order Base Array
1	Q8B88B	HPE Nimble Storage 2x10GbE 2-port FIO Adapter Kit
1	Q8G27B	HPE Nimble Storage NOS Default FIO Software
1	Q8H47A	HPE Nimble Storage AF40 All Flash Array R2 11.52TB (24x480GB) FIO Flash Bundle
2	R0P84A	HPE Nimble Storage NEMA IEC 60320 C14 to C19 250V 15 Amp 1.8m FIO Power Cord
1	Q8F56A	HPE Nimble Storage 10GbE 2-port Spare Adapter
2	P9Q66A	HPE G2 IEC C20 Input/(8) C13 Expansion Outlets/PDU Extension Bar Kit
		HPE 5945 FlexFabric Switching
2	JQ075A	HPE FF 5945 2-Slot Switch
2	JH180A	HPE 5930 24p SFP+ and 2p QSFP+ Module
2	JH183A	HPE 5930 8-port QSFP+ Module
4	JH389A	HPE X712 Back (Power Side) to Front (Port Side) Airflow High Volume Fan Tray
4	JC680A	HPE 58x0AF 650W AC Power Supply
4	JC680A B2B	INCLUDED: Jumper Cable - NA/JP/TW
2	JG326A	HPE X240 40G QSFP+ QSFP+ 1m DAC Cable
4	JG327A	HPE X240 40G QSFP+ QSFP+ 3m DAC Cable
		Red Hat OpenShift Container Platform
1	R1Z92AAE	Red Hat OpenShift Container Platform for HPE Synergy 1-32 Cores 1yr Subscription 24x7
		Red Hat Enterprise Linux Server
6	J8J36AAE	Red Hat Enterprise Linux Server 2 Sockets 1 Guest 1 Year Subscription 24x7 Support

Version history

Project. HPE Reference Configuration for Red Hat OpenShift Container Platform on HPE Synergy and HPE Nimble Storage

Status. Final

Document version	Date	Description of change
1	September 2018	Phase 1 Initial Publication.

Document version	Date	Description of change
2	January 2019	Phase 2 Updates to CA750 Recipe, host automation, install to OCP 3.10, Ansible Tower, Nimble automation, BURA Nimble, BURA 3PAR
3	June 2019	Phase 3 Description and objective of the Business Challenges.
4	December 2019	Phase 4 This update integrates vSphere into the core solution stack, adds GPUs into the solution, consolidates the OCP control plane, adds enhanced security around API authentication build processes, and the registry.

Resources and additional links

Red Hat, redhat.com

Red Hat OpenShift Container Platform 3.11 documentation, https://docs.openshift.com/container-platform/3.11/welcome/index.html

HPE Synergy, hpe.com/info/synergy

HPE Nimble Storage, hpe.com/storage/nimble

HPE FlexFabric 5945 switching, https://buy.hpe.com/us/en/networking/networking-switches/hpe-flexfabric-5945-switch-series/p/1010907030

HPE OpenShift Solutions on GitHub, https://github.com/hewlettpackard/hpe-solutions-openshift

Sysdig Platform Documentation https://sysdigdocs.atlassian.net/wiki/spaces/Platform/overview

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Share now ✉

a00056101enw, Version 4, December 2019