Hewlett Packard Enterprise

Deployment Guide

HPE Reference Configuration for Docker Containers as a Service on HPE Synergy Composable Infrastructure

Contents

Executive Summary

HPE Reference Configuration for Docker Containers as a Service on HPE Synergy Composable Infrastructure is a complete solution from Hewlett Packard Enterprise that includes all the hardware, software, professional services, and support you need to deploy a Containers-as-a-Service (CaaS) platform, allowing you to get up and running quickly and efficiently. The solution takes the HPE Synergy infrastructure and combines it with Docker's enterprise-grade container platform, popular open source tools, along with deployment and advisory services from HPE Pointnext.

HPE Enterprise Containers as a Service with Docker EE is ideal for customers migrating legacy applications to containers, transitioning to a container DevOps development model or needing a hybrid environment to support container and non-containerized applications on a common VM platform. This Reference Configuration provides a solution for IT operations, addressing the need for a production-ready environment that is easy to deploy and manage.

This release supports Kubernetes 1.11 via Docker Enterprise Edition (EE) 2.1, which is the only platform that manages and secures applications on Kubernetes in multi-Linux, multi-OS and multi-cloud customer environments. This document describes the best practices for deploying and operating HPE Enterprise Containers as a Service with Docker Enterprise Edition (EE). It shows how to automate the provisioning of the environment using a set of Ansible playbooks. It also outlines a set of manual steps to harden, secure and audit the overall status of the system.

Target Audience: This document is primarily aimed at technical individuals working in the operations side of the software pipeline, such as infrastructure architects, system administrators and infrastructure engineers, but anybody with an interest in automating the provisioning of virtual servers and containers may find this document useful.

Assumptions: The present document assumes a minimum understanding in concepts such as virtualization and containerization and also some knowledge around Linux®, Microsoft Windows® and VMware® technologies.

Solution overview

The HPE Reference Configuration for Docker Containers as a Service on HPE Synergy Composable Infrastructure consists of a set of Ansible playbooks that run on top of a VMware virtualization platform on HPE Synergy and HPE 3PAR storage hardware. The solution allows you to configure a flexible OS environment (with both RHEL and Windows workers) providing built-in high availability (HA), container monitoring and security, and backup and restore functionality. This solution assumes that you have already set up your HPE Synergy hardware, that you have installed your VMware virtualization platform and have configured HPE 3PAR for storage.

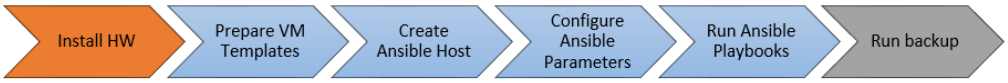

Figure 1. Solution overview

Figure 1 provides an overview of the steps used to deploy the solution. Deploying your hardware and HPE Synergy is specific to your environment and is not covered here. This document shows you how to:

- Prepare the VM templates
- Create the Ansible host
- Configure the Ansible parameters
- Run the Ansible playbooks

Once you are up and running, you should regularly back up the system using the scripts provided as part of this solution.

New in this release

Version 2.1 of the solution provides support for Kubernetes 1.11 via Docker EE 2.1. It is recommended that you set the DTR version to 2.6.4 (released 2019-03-28) to avoid a known issue when restoring DTR after backup. New features in this release include:

- Bare metal deployment for Linux and Windows

Features taken from the most recent release of HPE Express Containers on HPE SimpliVity include:

- **Prometheus/Grafana on Kubernetes:** The playbooks now set up a full monitoring stack for the deployed Kubernetes infrastructure using Prometheus Operator. They install kube-state-metrics and node-exporter components, as well as supporting Kubelet and Apiserver metrics. Sample dashboards for Grafana are installed to help you monitor your Kubernetes infrastructure.
- **Docker UCP metrics for Kubernetes:** A separate, standalone Prometheus/Grafana deployment is provided to support visualization of UCP metrics. This will be integrated into the full stack deployment in a future release.
- **Sysdig for Kubernetes:** The Sysdig deployment has been updated to use Kubernetes 1.11 RBAC and config maps for sensitive data.
- **NFS Provisioner for Kubernetes:** The NFS Provisioner has been updated to use Kubernetes 1.11 RBAC.
- **WordPress and MySQL using NFS Provisioner:** Playbooks are provided to validate the NFS Provisioner, featuring a WordPress and MySQL deployment with persistent storage.
- **kubectl:** A convenience playbook is provided to download and install kubectl.
- **Client bundle:** A convenience playbook is available to download and configure the client bundle from UCP.
- **Helm charts:** Playbooks for downloading, installing and configuring Helm are provided, with the use of sample charts for validation purposes.

For more details on what is new in this release, see the release notes at https://hewlettpackard.github.io/Docker-Synergy/rel-notes/new-features-syn.html.

Solution configuration

The Ansible playbooks are available to download at https://github.com/HewlettPackard/Docker-Synergy. By default, the playbooks are configured as shown in Figure 2 to set up a 3 node environment. This is the minimal starter configuration recommended by HPE and Docker for production.

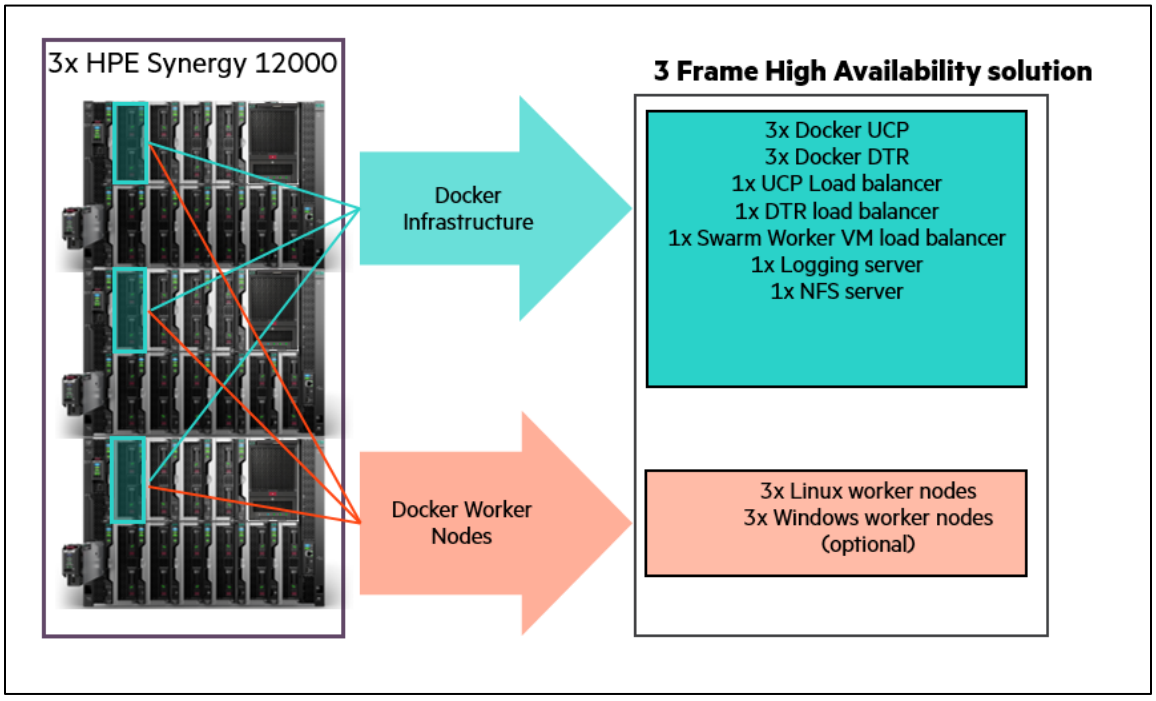

Figure 2. Three node HPE Synergy Configuration

The playbooks can also be used for larger container environments, for example, with a 3 frame, 6 node HPE Synergy system, as shown in Figure 3 with 2 nodes in each frame.

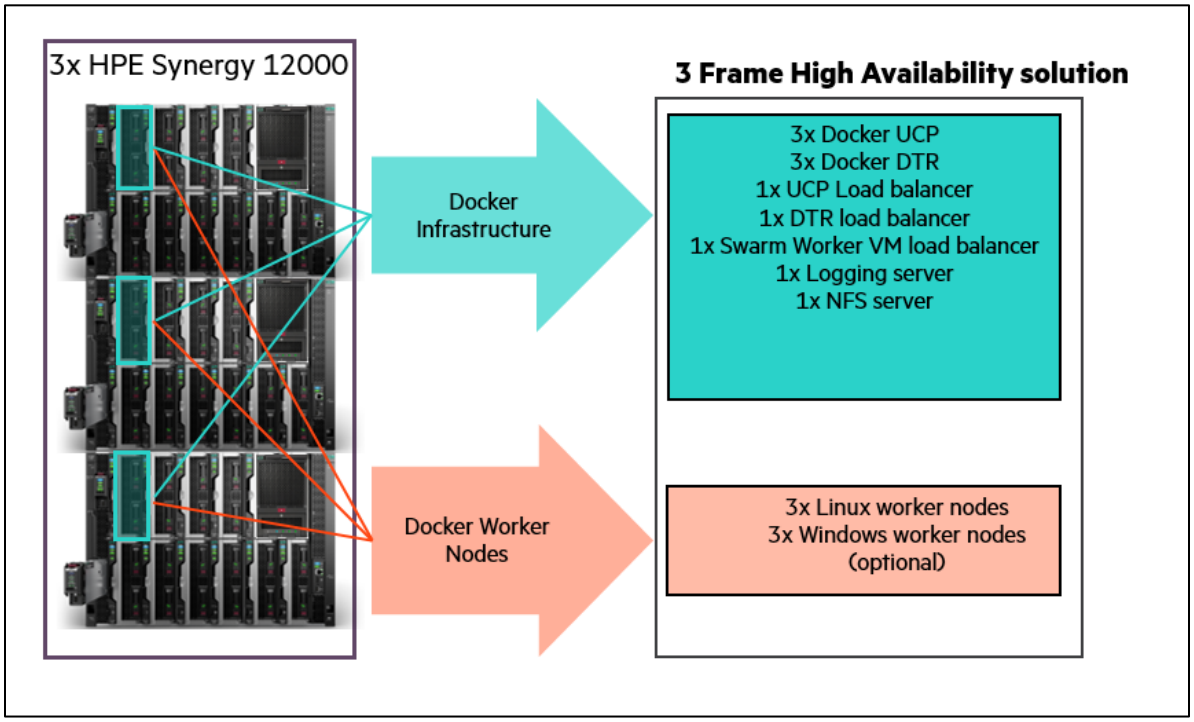

Figure 3. Six node HPE Synergy Configuration

Linux-only VM configuration

- 3 Docker Universal Control Plane (UCP) VM nodes for HA and cluster management
- 3 Docker Trusted Registry (DTR) VM nodes for HA of the container registry

  The Docker UCP and DTR nodes are spread across 3 physical nodes, with one on each physical node. An odd number of manager nodes is recommended to avoid split-brain issues. It is possible to restrict the deployment to 1 UCP and 1 DTR, or to expand to more than 3, but the recommended minimum for an enterprise production deployment is 3 UCPs and 3 DTRs.
- 3 Docker Linux worker VM nodes for container workloads - Kubernetes or Docker swarm or a mix

  The Docker worker nodes will be co-located with the UCP and DTR nodes in a 3 physical node deployment. Where more than 3 physical nodes are available, the worker nodes will typically be separated onto the extra nodes. It is possible to specify that more than one worker node is deployed per physical node but this decision will depend on the requirements of your applications.
- 1 Docker UCP load balancer VM to ensure access to UCP in the event of a node failure
- 1 Docker DTR load balancer VM to ensure access to DTR in the event of a node failure

  By default, two load balancers are deployed to increase availability of UCP and DTR and these are placed on separate physical nodes. Load balancing for applications running on worker nodes can achieved by using the playbooks to deploy additional load balancers, or by manually configuring the existing two to support your applications in addition to supporting UCP and DTR.
- 1 Logging server VM for central logging
- 1 NFS server VM for storage of Docker DTR images

With the addition of the NFS and logging VMs, a total of 13 VMs are created for the default Linux-only deployment. In addition to these VMs, the playbooks also set up the Docker persistent storage plug-in from VMware. The vSphere Docker volume plug-in facilitates the storage of data in a shared datastore that can be accessed from any machine in the cluster.

Hybrid VM configuration (Windows and Linux)

The hybrid deployment will typically add 3 Windows worker nodes to the above configuration, co-located with the Linux workers.

- 3 Docker swarm Windows worker VM nodes for container workloads (optional)

Bare metal (BM) configuration (Windows and Linux)

This solution leverages HPE Synergy OneView 4.10 and HPE Image Streamer 4.10 to provision bare metal servers with an operating system so they can be added to a Docker/Kubernetes cluster as worker nodes. The bare metal worker nodes can be used in conjuction with VM worker nodes or on their own with a virtualized control plane.

Note

Some of the application software supported by this configuration does not currently run on Windows, for example, the Sysdig Software Agent (see the section Monitoring with Sysdig).

High availability

Uptime is paramount for businesses implementing Docker containers in business critical environments. The HPE Enterprise Containers as a Service with Docker EE solution offers various levels of high availability (HA) to support continuous availability. The Docker EE system components run on multiple manager nodes in the cluster. The management plane continues to operate even in the event of a manager node failure. Application containers can be protected through the use of `services` running on top of swarm. The swarm orchestrator works to maintain the number of containers declared as part of the service. The Ansible playbooks can be modified to fit your environment and your high availability (HA) needs.

Load Balancers

This solution also deploys load balancers in the system to help with container traffic management. There are two load balancer VMs – the UCP load balancer and DTR load balancer. The playbooks can be configured to deploy one or more worker load balancers depending on the requirements of your applications. A typical load balancer architecture for applications running on Docker EE is shown in Figure 4. The playbooks now support load balancers based on VRRP, using HAproxy and `keepalived`. The solution can be deployed using these loadbalancers, or external load balancers, or no load balancers or the legacy version of standalone load balancers. For more information on HAproxy, see http://www.haproxy.com/solutions/high-availability/.

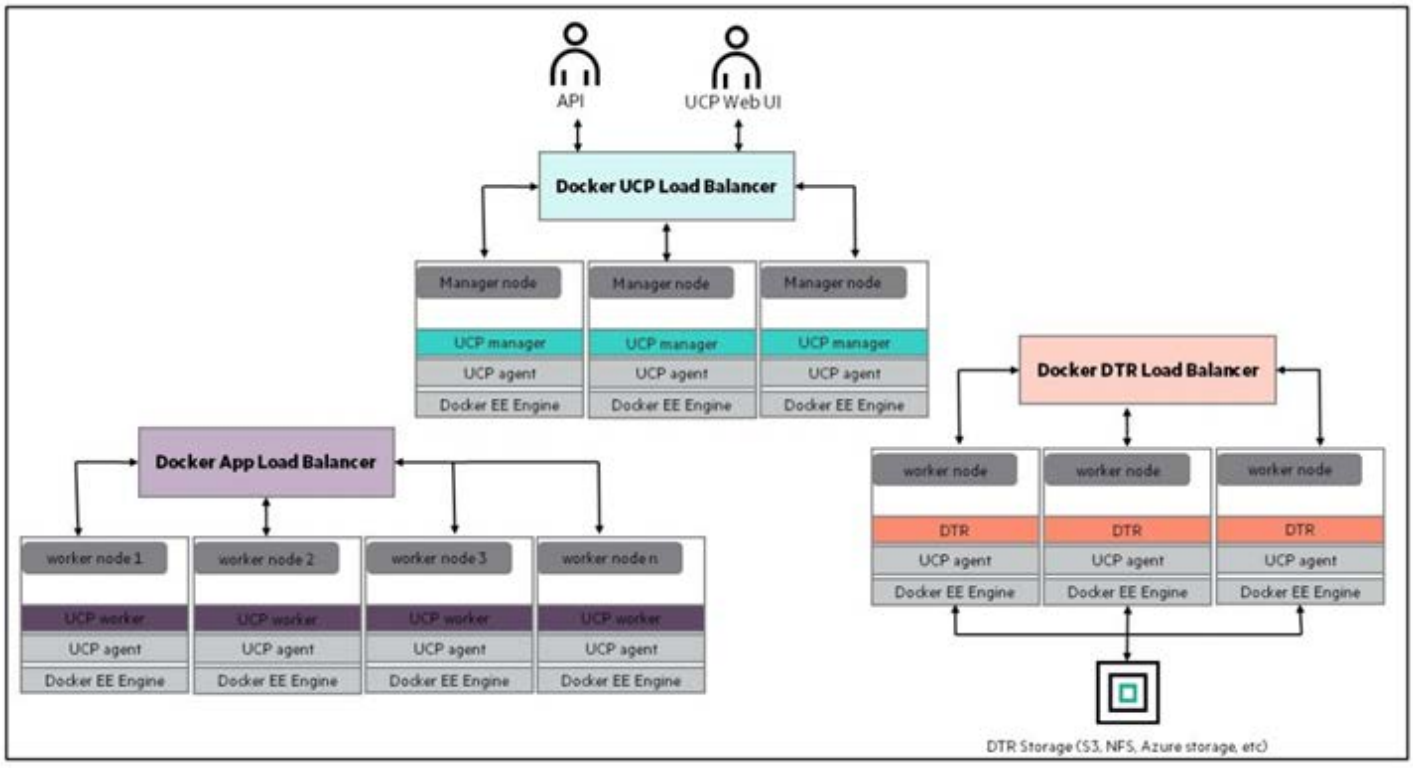

Figure 4. Load balancer architecture

Sizing considerations

A node is a machine in the cluster (virtual or physical) with Docker Engine running on it. There are two types of nodes: managers and workers. UCP will run on the manager nodes. Although DTR runs on a worker node, Docker does not recommend running other application containers on them. To decide what size the node should be in terms of CPU, RAM, and storage resources, consider the following:

1. All nodes should at least fulfil the minimal requirements, for UCP 3.0, 8GB of RAM and 6GB of storage. For production systems, 16GB of RAM and 25-100GB of free disk space is recommended for manager nodes. More detailed requirements are in the Docker EE UCP documentation at https://docs.docker.com/ee/ucp/admin/install/system-requirements/.
2. UCP controller nodes should be provided with more than the minimal requirements, but won't need much more if nothing else runs on them.
3. Ideally, worker node size will vary based on your workloads so it is impossible to define a universal standard size.
4. Other considerations like target density (average number of containers per node), whether one standard node type or several are preferred, and other operational considerations might also influence sizing.

If possible, node size should be determined by experimentation and testing actual workloads; and they should be refined iteratively. A good starting point is to select a standard or default machine type for all nodes in the environment. If your standard machine type provides more resources than the UCP controller nodes need, it makes sense to have a smaller node size for these. Whatever the starting choice, it is important to monitor resource usage and cost to improve the model.

For this solution, the following tables describe sizing configurations, assuming 3 Linux workers and 3 Windows workers. The vCPU allocations are described in Table 1.

Table 1. vCPU

vCPUs	node01	node02	node03
ucp1	4		
ucp2		4	
ucp3			4
dtr1	2		
dtr2		2	
dtr3			2
worker1	4		
worker2		4	
worker3			4
win-worker1	4		
win-worker2		4	
win-worker3			4
lb1	2		
lb2		2	
nfs			2
logger		2	
Total vCPU per node	**16**	**18**	**16**

Note

In the case of one ESX host failure, 2 nodes are enough to accommodate the amount of vCPU required.

Memory allocation for 3 node solution

The memory allocation for this solution (3 Linux workers and 3 Windows workers), is described in Table 2.

Table 2. Memory allocation

RAM (GB)	node01	node02	node03
ucp1	16		
ucp2		16	
ucp3			16
dtr1	16		
dtr2		16	
dtr3			16
worker1	64		
worker2		64	
worker3			64
win-worker1	64		
win-worker2		64	

win-worker3			64
lb1	4		
lb2		4	
nfs			4
logger		4	
Total RAM required (per node)	**164**	**168**	**164**
Available RAM	**384**	**384**	**384**

Note
In the case of one ESX host failure, the two surviving hosts can accommodate the amount of RAM required for all VMs.

Memory allocation for 6 node solution

For a 6 node solution, Table 3 outlines the memory requirements where the control plane is on 3 nodes and the worker nodes are on the other 3 nodes. In this example, it is assumed that there are 2 Linux worker nodes and 1 Windows worker node, but the actual number of worker nodes is not limited to 3 and depends entirely on the workload requirements.

Table 3. Memory allocation for 6 nodes

RAM (GB)	node01	node02	node03	node04	node05	node06
ucp1	16					
ucp2		16				
ucp3			16			
dtr1	16					
dtr2		16				
dtr3			16			
worker1				64		
worker2					64	
win-worker1						64
lb1	4					
lb2		4				
nfs			4			
logger		4				
Total RAM required (per node)	**36**	**44**	**44**	**64**	**64**	**64**
Available RAM	**128**	**128**	**128**	**128**	**128**	**128**

Note
In the case of one ESX host failure, the two surviving hosts can accommodate the amount of RAM required for all VMs.

Disaster Recovery

Recovery Time Objective (RTO) refers to the time that it takes to recover your data and applications while Recovery Point Objective (RPO) refers to the point in time you can recover to, in the event of a disaster. In essence, RPO tells you how often you will need to make new backups.

In order to protect your installation from disasters, you need to take regular backups and transfer the backups to a safe location. This solution provides a range of convenience scripts and Ansible playbooks to help automate the backup of UCP, DTR, your swarm and your Docker volumes. See the section Backup and restore for best practices, procedures and utilities for implementing disaster recovery.

Security

The Docker Reference architecture for Securing Docker EE and Security Best Practices is available at https://success.docker.com/article/Docker_Reference_Architecture-_Securing_Docker_EE_and_Security_Best_Practices

In addition to having all logs centralized in a single place and the image scanning feature enabled for the DTR nodes, there are other guidelines that should be followed in order to keep your Docker environment as secure as possible. The HPE Reference Configuration paper for securing Docker on HPE Hardware places a special emphasis on securing Docker in DevOps environments and covers best practices in terms of Docker security. The document can be found at http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00020437enw.

In addition, the Sysdig product also provides a strong level of container security and monitoring (see the section Monitoring with Sysdig).

Solution components

This section describes the various components that were utilized in this Reference Configuration.

Hardware

Table 4 lists the hardware components that are utilized in this Reference Configuration.

Table 4. Hardware

Component	Purpose
HPE Synergy 12000 Frame	Rack enclosure for compute, storage, and network hardware
HPE Synergy 480 Gen10 Compute Modules	Hosts for running ESX servers that support UCP, DTR, worker and other nodes in the solution
HPE 3PAR StoreServ 8200	Provides the storage for the virtual machines and the Docker backups
HPE StoreOnce	High performance backup system

About 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.

Server requirements

The minimum platform requirement for this configuration, shown in Figure 2, is a three node HPE Synergy 480 Gen10 deployment. There is a single ESXi cluster with both the control plane and the Docker workers spread out on all three nodes. A single node in each Synergy frame has the following suggested requirements:

- 384 GB DDR4-2133 RAM
- 2 Intel® Xeon® CPU Gold 6130 2.10GHz x 16 core

The solution has also been tested on a 6 node HPE Synergy environment, with 2 nodes in each frame. In this setup, there is a single ESXi cluster with the control plane on 3 nodes while the extra 3 nodes are dedicated exclusively to Docker worker nodes. The 6 node deployment is depicted graphically in Figure 3 with the following suggested requirements for each node:

- 128 GB DDR4-2133 RAM
- 2 Intel® Xeon® CPU Gold 6130 2.10GHz x 16 core

Storage requirements

An HPE 3PAR array is required for the ESXi datastore. This solution makes use of an HPE 3PAR StoreServ 8200 populated with:

- 8x 480GB SSD for the vSphere cluster datastore
- 8x 1.8TB HDD for the backup datastore

You should create a large virtual volume on the HPE 3PAR StoreServ to host the virtual machines and another large virtual volume for Docker backups. Create datastores on your vSphere cluster using these virtual volumes. If desired, you can create separate HPE 3PAR StoreServ virtual volumes and attach them to all vSphere cluster hosts for backing up Docker persistent volumes. It is recommended that you configure the volumes that are used for virtual machine deployments on the SSDs. Storage for backups can be configured on the HDDs.

Software

The software components used in this Reference Configuration are listed in Table 5 and Table 6.

Table 5. Third-party software

Component	Version
Ansible	2.7
Docker EE	2.1 with Docker EE Engine 18.09 (tested with UCP 3.1.4 and DTR 2.6.4)
Red Hat Enterprise Linux	7.6
Microsoft Windows	Server 2016
VMware	ESXi 6.5.0 and vCenter 6.5.0

Table 6. HPE Software

Component	Version
HPE Recovery Manager Central	5.0.1
HPE Synergy OneView	4.1
HPE Image Streamer	4.1

About Ansible

Ansible is an open-source automation engine that automates software provisioning, configuration management and application deployment.

As with most configuration management software, Ansible has two types of servers: the controlling machine and the nodes. A single controlling machine orchestrates the nodes by deploying modules to the Linux nodes over SSH. The modules are temporarily stored on the nodes and communicate with the controlling machine through a JSON protocol over the standard output. When Ansible is not managing nodes, it does not consume resources because no daemons or programs are executing for Ansible in the background. Ansible uses one or more inventory files to manage the configuration of the multiple nodes in the system.

When deploying Windows nodes in a hybrid deployment, the Ansible playbooks make use of the Python `pywinrm` module which carries out actions via the Windows remote manager.

More information about Ansible can be found at http://docs.ansible.com.

About Docker Enterprise Edition

Docker Enterprise Edition (EE) is the leading enterprise-ready container platform for IT that manages and secures diverse applications across disparate infrastructure, both on-premises and in the cloud. Docker EE provides integrated container management and security from development to production. Enterprise-ready capabilities like multi-architecture orchestration and secure software supply chain give IT teams the ability to manage and secure containers without breaking the developer experience.

Docker EE provides:

- Integrated management of all application resources from a single web admin UI.
- Frictionless deployment of applications and Compose files to production in a few clicks.
- Multi-tenant system with granular role-based access control (RBAC) and LDAP/AD integration.
- Self-healing application deployment with the ability to apply rolling application updates.
- End-to-end security model with secrets management, image signing and image security scanning.

More information about Docker Enterprise Edition can be found at https://www.docker.com/enterprise-edition.

Application software

A number of different logging and monitoring solutions are supported by this solution:

- Splunk
- Sysdig
- Prometheus and Grafana

The application software components used in this Reference Configuration are listed in Table 7.

Table 7. Application software

Component	Version
Splunk	7.1.2
Sysdig	latest
Prometheus	V2.3.2
Grafana	5.2.3

Monitoring with Splunk and Sysdig

The solution can be configured to use either Splunk or Sysdig or to enable both simultaneously. While there is some overlap in the functionality provided by these tools, they are ultimately complimentary in what they offer. Splunk aggregates logging and tracing for a wide variety of sources and provides a clean, high-level dashboard for all your enterprise systems. Sysdig, on the other hand, has been engineered from the ground up to focus on containerized environments and includes both monitoring and security features, with built-in understanding of the different workloads running on your cloud.

More information on configuring Splunk and running the relevant playbooks can be found in the section Deploying Splunk.

For more information on configuring Sysdig and running the relevant playbooks, see the section Deploying Sysdig monitoring.

Monitoring with Prometheus and Grafana

The solution can be configured to enable the use of Prometheus and Grafana for monitoring. In this setup, there is no need for native installs and all the required monitoring software runs in containers, deployed as either services or stacks.

The solution supports two separate monitoring stacks, with one running on Kubernetes and the other using Docker swarm.

For more information on running Prometheus and Grafana on Kubernetes, see section Monitoring Kubernetes with Prometheus and Grafana.

For more information on running Prometheus and Grafana on Docker swarm, see section Deploying Prometheus and Grafana on Docker swarm.

Preparing the environment

This section describes in detail how to prepare the environment that was outlined in the architecture section. The following high level steps are required:

- Verify prerequisites
- Enable vSphere High Availability (HA)
- Install vSphere Docker Volume Service driver on all ESXi hosts
- Create the Ansible node
- Create the Red Hat Linux Template and configure the `yum` repositories
- Create the Windows Template (optional)
- Finalize the template

Verify prerequisites

Before you start deployment, you must assemble the information required to assign values for each and every variable used by the playbooks. The variables are fully documented in the section Configuring the solution components. A brief overview of the information required is presented in Table 8.

Table 8. Summary of information required

Component	Details
Virtual Infrastructure	The FQDN of your vCenter server and the name of the Datacenter. You will also need administrator credentials in order to create templates and spin up virtual machines.
L3 Network requirements	You will need one IP address for each and every VM and bare metal node configured in the Ansible inventory (see the section Configuring the solution components). The recommended minimal deployment (Linux-only) configures 13 virtual machines so you would need to allocate 13 IP addresses to use this example inventory. If you have a hybrid environment with Windows workers, you will need to increase the allocation. Note that **the Ansible playbooks do not support DHCP** so you need static IP addresses. All the IPs should be in the same subnet. You will also have to specify the size of the subnet (for example /22 or /24) and the L3 gateway for this subnet.
DNS	You will need to know the IP addresses of your DNS server. In addition, all the VMs and bare metal nodes you configure in the inventory must have their names registered in DNS prior to deployment. In addition, you will need to know the domain name to use for configuring the virtual machines (such as `example.com`)
NTP Services	You need time services configured in your environment. The deployed solution uses certificates that are time-sensitive. You will need to specify the IP addresses of your time servers (NTP).
RHEL Subscription	A RHEL subscription is required to pull extra packages that are not on the DVD.
Docker Prerequisites	You will need a URL for the official Docker EE software download and a license file. Refer to the Docker documentation to learn more about this URL and the licensing requirements at: https://docs.docker.com/engine/installation/linux/docker-ee/rhel/ in the section entitled Find your Docker EE repo URL.
Proxy	The playbooks pull the Docker packages from the Internet. If your environment accesses the Internet through a proxy, you will need the details of the proxy including the fully qualified domain name and the port number.

Enable vSphere High Availability (HA)

You must enable vSphere High Availability (HA) to support virtual machine failover during a HA event such as a host failure. Sufficient CPU and memory resources must be reserved across the system so that all VMs on the affected host(s) can fail over to remaining available hosts in the system. You configure an Admission Control Policy (ACP) to specify the percentage CPU and memory to reserve on all the hosts in the cluster to support HA functionality.

Note
You should not use the default Admission Control Policy. Instead, you should calculate the memory and CPU requirements that are specific to your environment.

Install vSphere Docker Volume Service driver on all ESXi hosts

vSphere Docker Volume Service technology enables stateful containers to access the storage volumes. Setting this up is a one-off manual step. In order to be able to use Docker volumes using the vSphere driver, you must first install the latest release of the vSphere Docker Volume Service (vDVS) driver, which is available as a vSphere Installation Bundle (VIB). To perform this operation, log in to each of the ESXi hosts and then download and install the latest release of vDVS driver.

```
# esxcli software vib install -v /tmp/vmware-esx-vmdkops-<version>.vib --no-sig-check
```

More information on how to download and install the driver can be found at http://vmware.github.io/vsphere-storage-for-docker/documentation/install.html.The version of the driver tested in this configuration is 0.21.2.

Create the Ansible node on Fedora

The Docker Synergy playbooks rely on the Ansible Modules for HPE OneView project when deploying bare metal resources. As a result, there is a requirement to run a newer version of Python than is available by default on RHEL. In this release of the Docker Synergy solution, it is required to deploy your Ansible contoller on Fedora, to take advantage of the built-in support for Python 3.

Create Fedora VM

Create a Virtual Machine with the following characteristics:

- **Guest OS:** Red Hat Fedora Server 29 (64-bit)
- **Disk:** 50G (thin provisioning)
- **CPU:** 2
- **RAM:** 4 GB
- **Ethernet Adapter:** VMXNET 3, connected to your Ansible or management network

Install Fedora Server 29 using the appropriate ISO image for the distro (x86 64 bit) and in the Software Selection section, choose:

- **Base Environment:** Fedora Server Edition
- **Add-Ons for Selected Environment:** Guest Agent

Select your language, keyboard, and timezone settings and re-boot when the installation finishes.

Configure your networking and check your connectivity before moving on to the next section. If you are operating behind a proxy, configure DNF by editing `/etc/dnf/dnf.conf`, as outlined here.

Install Ansible and required modules

Login the root account and run the following commands:

```
dnf update -y
dnf install -y git ansible python3-netaddr python3-requests python3-pyvmomi python3-pip python3-winrm

cd /usr/bin
ln -s python3.7 python

# install the python HPE OneView SDK
cd
git clone https://github.com/HewlettPackard/python-hpOneView.git
cd python-hpOneView/
pip3 install .
```

```
# Install the Oneview Ansible Modules
cd
git clone https://github.com/HewlettPackard/oneview-ansible.git

# Configure Ansible
cat <<EOF >>~/.bashrc
export ANSIBLE_LIBRARY=/root/oneview-ansible/library
export ANSIBLE_MODULE_UTILS=/root/oneview-ansible/library/module_utils
EOF

source ~/.bashrc
```

Create the Red Hat Linux template

To create the Red Hat Linux template that you will use as the base for all your VM nodes, you first create a Virtual Machine with the OS installed and then convert the Virtual Machine to a VM Template. The VM Template is created as lean as possible, with any additional software installs and/or system configuration performed subsequently using Ansible.

As the creation of the template is a one-off task, this procedure has not been automated. The steps required to manually create a VM template are outlined below.

Log in to vCenter and create a new Virtual Machine with the following characteristics:

- Guest OS Family: Linux, Guest OS Version: Red Hat Enterprise Linux (64-bit)
- Hard Disk size: 50GB, (Thin provisioning)
- A single network controller connected to the network or VLAN of your choice. All VMs will connect to this same network.
- Optionally you can remove the floppy drive

Install Red Hat Enterprise 7:

1. Select a language which is supported by Docker
2. For the software selection, choose **Infrastructure Server** as the base environment and add the **Guest Agents** from the lists of add-ons available for this environment. The Infrastructure Server environment is selected here versus the Minimal Install because Customization of Linux guest operating systems requires that Perl is installed in the Linux guest operating system.
3. Configure the network settings so that you can later access the VM using SSH. Specify an IP address for the network interface, a default gateway, DNS settings and possibly any HTTP/HTTPS proxies that apply in your environment.
4. Specify a password for the root account and optionally created an admin user.
5. Wait for the installation to finish and for the VM to reboot.

Update packages

Use `yum update` to install the latest packages, configuring a proxy if required.

```
# subscription-manager config --server.proxy_hostname=<proxy IP> --server.proxy_port=<proxy port>
# subscription-manager register --auto-attach

# subscription-manager repos \
--enable=rhel-7-server-rpms \
--enable=rhel-7-server-extras-rpms

# yum -y update
# subscription-manager unregister
```

Finalize the template

Log in to the `root` account on the Ansible box and copy the SSH public key to the VM Template. This will allow your Ansible node to SSH to all the Virtual Machines created from the VM Template without the need for a password.

```
ssh-copy-id root@<IP of your VM_Template>
```

Perform the following steps on the VM Template to finalize its creation:

1. Clean up the template by running the following commands from the **Virtual Machine Console:**

```
# rm /etc/ssh/ssh_host_*
# nmcli con del ens192
# logrotate -f /etc/logrotate.conf
# rm /var/log/*-201?*
# history -c
```

2. Shutdown the VM

```
# shutdown -h now
```

3. Turn the VM into a template by right-clicking on your VM and selecting `Template -> Convert to Template`. This will create a new template visible under VM Templates in Folders, ready for future use.

Note

In both the Ansible node and the VM Template, you might need to configure the network so one node can reach the other. Instructions for this step have been omitted since it is a basic step and could vary depending on the user's environment.

Configuring the solution components

Once you have prepared your environment, you need to download the solution software and edit the configuration variables to match your setup.

Ansible configuration

1. On the Ansible node, retrieve the latest version of the playbooks using Git.

```
# git clone https://github.com/HewlettPackard/Docker-Synergy.git
```

2. Change to the directory that you just cloned:

```
# cd ~/Docker-Synergy
```

Note

All subsequent file names are relative to the `Docker-Synergy` directory. For example `hosts` is located in `~/Docker-Synergy/` and `group_vars/all/vars` corresponds to `~/Docker-Synergy/group_vars/all/vars`.

You now need to prepare the configuration to match your own environment, prior to deploying Docker EE and the rest of the nodes. To do so, you will need to modify a number of files including:

- `site.yml`, the main entry point for the playbooks.
- `hosts`, the inventory file.

You also need to create and populate a number of files:

- `group_vars/all/vars`, the group variables file.

- `group_vars/all/vault`, containing sensitive information that needs to be protected.
- `group_vars/all/backups`, containing backup-related variables.

For the latter group, a set of sample files has been provided to help you get started:

- `group_vars/all/vars.sample`, a sample group variables file.
- `group_vars/all/vault.sample`, a sample vault file.
- `group_vars/all/backups.sample`, a sample backup configuration file.

The file `group_vars/win_worker.yml` supports advanced configuration of Windows remote management and in general should not require modification.

You should work from the `root` account for the configuration steps and also later on when you run the playbooks.

Editing the inventory

The inventory is the file named `hosts` in the `~/Docker-Synergy` directory. You need to edit this file to describe the configuration you want to deploy.

The nodes inside the inventory are organized in groups. The groups are defined by brackets and the group names are static so they must not be changed. Other fields (hostnames, specifications, IP addresses…) are edited to match your setup. The groups are as follows:

Control plane

- `[ucp_main]`: A group containing one single node which will be the main UCP node and swarm leader. Do not add more than one node under this group.
- `[ucp]`: A group containing all the UCP nodes, including the main UCP node. Typically you should have either 3 or 5 nodes under this group.
- `[dtr_main]`: A group containing one single node which will be the first DTR node to be installed. Do not add more than one node under this group.
- `[dtr]`: A group containing all the DTR nodes, including the main DTR node. Typically you should have either 3 or 5 nodes under this group.
- `[nfs]`: A group containing one single node which will be the NFS node. Do not add more than one node under this group.
- `[logger]`: A group containing one single node which will be the logger node. Do not add more than one node under this group.

Load balancers

If you are deploying the new active-active load balancers, using floating IPs managed by `keepalived`:

- `[loadbalancer]`: A group containing the UCP, DTR and any worker load balancers you are deploying.

If you are using the legacy, standalone load balancers:

- `[ucp_lb]`: A group containing one single node which will be the load balancer for the UCP nodes. Do not add more than one node under this group.
- `[dtr_lb]`: A group containing one single node which will be the load balancer for the DTR nodes. Do not add more than one node under this group.
- `[worker_lb]`: A group containing one single node which will be the load balancer for the worker nodes. Do not add more than one node under this group.
- `[lbs]`: A group containing all the load balancers. This group will have 3 nodes, also defined individually in the three groups above.

Note

Even if you are using the new `[loadbalancer]` group, you must still declare the legacy group `[lbs]` and its sub-groups in your inventory. The sub-groups do not need to declare any entries if the new `[loadbalancer]` group is used.

Worker nodes

- [vm_wrk_lnx]: A group containing all the Linux worker nodes on Virtual Machines.
- [bm_wrk_lnx]: A group containing all the bare metal Linux worker nodes.
- [vm_wrk_win]: A group containing all the Windows worker nodes on Virtual Machines.
- [bm_wrk_win]: A group containing all the bare metal Windows worker nodes.

Ansible controller

- [local]: A group containing the local Ansible host. It contains an entry that should not be modified.

Groups of groups

A number of "groups of groups" simplify the handling of sets of nodes:

ctlrplane group

All the nodes that make up the control plane:

```
[ctlrplane:children]
ucp
dtr
lbs
nfs
loadbalancer
logger
```

worker group

All the Docker worker nodes:

```
[worker:children]
vm_wrk_lnx
vm_wrk_win
bm_wrk_lnx
bm_wrk_win
```

bms group

All the bare metal nodes:

```
[bms:children]
bm_wrk_lnx
bm_wrk_win
```

docker group

All the nodes running Docker:

```
[docker:children]
ucp
dtr
worker
```

linux_box group

All the nodes running Linux:

```
[linux_box:children]
ctlrplane
vm_wrk_lnx
bm_wrk_lnx
```

windows_box group
All the nodes running Windows:

```
[windows_box:children]
bm_wrk_win
vm_wrk_win
```

Bare metal variables
When deploying bare metal worker nodes, you must specify the name of the Server Profile Template (SPT), together with the names of the two connections for your Ansible controller. If you have multiple server types in your HPE Synergy setup, you will need to set the name of the server profile template for each individual bare metal node, typically on the node decalration in the inventory file itself, rather than using a common name in the group file.

Bare metal Linux variables
Variables specific to bare metal Linux worker nodes are specified in `group_vars/bm_wrk_lnx.yml`

```
ov_template: 'RedHat760_fcoe_v1.0.2'
ov_ansible_connection_name: 'ansibleA'
ov_ansible_redundant_connection_name: ansibleB

disk2: '/dev/mapper/mpatha'
disk2_part: '/dev/mapper/mpatha1'
orchestrator: kubernetes # or swarm
fcoe_devices: ['ens3f2','ens3f3']
```

Bare metal Windows variables
Variables specific to bare metal Windows worker nodes are specified in `group_vars/bm_wrk_win.yml`

```
ov_template: 'Windows Worker Node (Gen9)'
ov_ansible_connection_name: 'Ansible-A'
ov_ansible_redundant_connection_name: 'Ansible-B'
```

Inventory group variables

Additional configuration files for each group in the inventory are available, including `group_vars/vms.yml`, `group_vars/ucp.yml`, `group_vars/dtr.yml`, `group_vars/worker.yml` and `group_vars/nfs.yml`.

The following files, in the group_vars folder, contain variable definitions for each group.

- **ucp.yml:** Variables defined for all UCP nodes.
- **dtr.yml:** Variables defined for all DTR nodes.
- **nfs.yml:** Variables defined for all NFS nodes.
- **logger.yml:** Variables defined for all logger nodes.
- **loadbalancer.yml:** Variables defined for all nodes in the [loadbalancer] group.
- **lbs.yml:** Variables defined for all nodes in the legacy [lbs] group.
- **vm_wrk_lnx.yml:** Variables defined for all Linux VM worker nodes.
- **vm_wrk_win.yml:** Variables defined for all Windows VM worker nodes.
- **worker.yml:** Variables defined for all worker nodes.
- **windows_box.yml:** Variables defined for all Windows nodes.
- **vms.yml:** Variables defined for all the VMware Virtual Machines deployed by the solution.
- **bms.yml:** Variables defined for all the bare metal machines deployed by the solution.

These group files facilitate more sophisticated settings, such as additional drives and additional network interfaces. For example, here is the `group_vars/nfs.yml` file.

```
networks:
  - name:  '{{ vm_portgroup }}'
    ip:  "{{ ip_addr | ipaddr('address') }}"
    netmask: "{{ ip_addr | ipaddr('netmask') }}"
    gateway: "{{ gateway }}"

disks_specs:
  - size_gb:  '{{ disk1_size }}'
    type: thin
    datastore: "{{ datastores | random }}"
  - size_gb: '{{ disk2_size }}'
    type: thin
    datastore: "{{ datastores | random }}"
  - size_gb: 10
    type: thin
    datastore: "{{ datastores | random }}"
```

In this example, the size of the first two drives is specified using the values of the variables `disk1_size` and `disk2_size` that are declared in the `group_vars/all/vars` file. This maintains compatibility with `hosts` inventories from the previous release of the playbooks. However, it is possible to provide explicit values, depending on your requirements, for the individual UCP, DTR, worker or NFS VMs. For example, you may want to increase the size of the second disk for the NFS VM as this is used to store the DTR images, so the default value of 500GB may not be sufficient to meet your needs.

In this release, support has been added for configuring a third drive that can be used to hold Kubernetes persistent volume data. The default size (10GB) is set low as the use of the NFS VM for storing persistent volume data is only considered suitable for demo purposes and should not be used in a production environment.

In the following example, the `group_vars/nfs.yml` has been modified to configure the NFS VM with a 50GB boot disk, a 500GB drive for DTR images and an 800GB drive for Kubernetes persistent volumes data.

```
networks:
  - name:  '{{ vm_portgroup }}'
    ip:  "{{ ip_addr | ipaddr('address') }}"
    netmask: "{{ ip_addr | ipaddr('netmask') }}"
    gateway: "{{ gateway }}"

disks_specs:
  - size_gb:  50
    type: thin
    datastore: "{{ datastores | random }}"
  - size_gb: 500
    type: thin
    datastore: "{{ datastores | random }}"
  - size_gb: 800
    type: thin
    datastore: "{{ datastores | random }}"
```

Note

The number of drives and the purpose of each drive is determined by the role of the VM and the specific playbooks that use the information. The first disk is always used as the boot disk, irrespective of VM role, while the purpose of the second or third disk is specific to the role.

Overriding group variables

If you wish to configure your nodes with different specifications to the ones defined by the group, it is possible to declare the same variables at the node level, overriding the group value. For instance, you could have one of your workers with higher specifications by setting:

In the file `vm_wrk_lnx.yml`:

```
cpus: '4'
ram: '65536'
disk2_size: '500'
```

In the `hosts` file:

```
[vm_wrk_lnx]
worker01 ip_addr='10.60.59.10/16' esxi_host='esx04.cloudra.local'
worker02 ip_addr='10.60.59.11/16' esxi_host='esx05.cloudra.local'
worker03 ip_addr='10.60.59.12/16' esxi_host='esx06.cloudra.local' cpus='16' ram'131072'
```

In the example above, the worker03 Linux VM node would have 4 times more CPU and double the RAM compared to the rest of the Linux VM worker nodes.

The different variables you can use are described in Table 9 below. They are all mandatory unless otherwise specified.

Table 9. Variables

Variable	Scope	Description
ip_addr	Node	IP address in CIDR format to be given to a node
esxi_host	Node	ESXi host where the node will be deployed. If the cluster is configured with DRS, this option will be overridden
cpus	Node/Group	Number of CPUs to assign to a VM or a group of VMs
ram	Node/Group	Amount of RAM in MB to assign to a VM or a group of VMs
disk2_size	Node/Group	Size of the second disk in GB to attach to a VM or a group of VMs. This variable is only mandatory on Docker nodes (UCP, DTR, worker) and NFS node. It is not required for the logger node or the load balancers.

VMware configuration

All VMware-related variables are mandatory and are described in Table 10.

Table 10. VMware variables

Variable	File	Description
vcenter_hostname	group_vars/all/vars	IP or hostname of the vCenter appliance
vcenter_username	group_vars/all/vars	Username to log in to the vCenter appliance. It might include a domain, for example, '`administrator@vsphere.local`'.
vcenter_password	**group_vars/all/vault**	The password corresponding to the `vcenter_username` user above.
vcenter_validate_certs	group_vars/all/vars	'no'
datacenter	group_vars/all/vars	Name of the datacenter where the environment will be provisioned
vm_username	group_vars/all/vars	Username to log into the VMs. It needs to match the one from the VM Template, so unless you have created a user, you must use '`root`'.
vm_password	**group_vars/all/vault**	The password for the `vm_username` user above.
vm_template	group_vars/all/vars	Name of the RHEL VM Template to be use. Note that this is the name from a vCenter perspective, not the hostname.
folder_name	group_vars/all/vars	vCenter folder to deploy the VMs. If you do not wish to deploy in a particular folder, the value should be /. Note: If you want to deploy in a specific folder, you need to create this folder in the inventory of the selected datacenter before starting the deployment.

datastores	group_vars/all/vars	List of datastores to be used, in list format, i.e. ['Datastore1','Datastore2'...]. The datastores must exist before you run the playbooks. Note that each datastore should be mounted on each of the ESXi hosts.
disk2	group_vars/all/vars	UNIX® name of the second disk for the Docker VMs. Typically `/dev/sdb`
disk2_part	group_vars/all/vars	UNIX name of the partition of the second disk for the Docker VMs. Typically `/dev/sdb1`
vsphere_plugin_version	group_vars/all/vars	Version of the vSphere plugin for Docker. The default is 0.21.2 which is the latest version at the time of writing this document. The version of the plugin should match the version of the vSphere Installation Bundle (VIB) that you installed on the ESXi servers.
vm_portgroup	group_vars/all/vars	Used by the playbook `create_vms.yml`, this variable is used to specify the portgroup connected to the network that connects all the VMs. There is currently only one network. It is recommended that the template which is used as the base for all deployed VMs specifies a network adapter but it is not required. If a network adapter is specified, you should not attach this adapter to a standard switch if the portgroup designated by `vm_portgroup` is connected to a distributed vSwitch. In addition, you should make sure that the adapter specifies `Connect At Power On`.

Networking configuration

All network-related variables are mandatory and are described in Table 11.

Table 11. Network variables

Variable	File	Description
nic_name	group_vars/all/vars	Name of the device, for RHEL this is typically `ens192` and it is recommended to leave it as is.
gateway	group_vars/all/vars	IP address of the gateway to be used
dns	group_vars/all/vars	List of DNS servers to be used, in list format, i.e. ['10.60.59.1','10.60.59.2'...]
domain_name	group_vars/all/vars	Domain name for your Virtual Machines
ntp_servers	group_vars/all/vars	List of NTP servers to be used, in list format, i.e. ['1.2.3.4','0.us.pool.net.org'...]

Environment configuration

All Environment-related variables are described in Table 12 below.

Table 12. Environment variables

Variable	File	Description
env	group_vars/all/vars	Dictionary containing all environment variables. It contains three entries described below. Please leave the proxy related settings empty if not required: `http_proxy`: HTTP proxy URL, such as `'http://15.184.4.2:8080'`. This variable defines the HTTP proxy URL if your environment is behind a proxy. `https_proxy`: HTTPS proxy URL, such as `'http://15.184.4.2:8080'`. This variable defines the HTTPS proxy URL if your environment is behind a proxy. `no_proxy`: List of hostnames or IPs that don't require proxy, such as `'localhost,127.0.0.1,.cloudra.local,10.60.59.'`

Docker configuration

All Docker-related variables are mandatory and are described in Table 13.

Table 13. Docker variables

Variable	File	Description
docker_ee_url	**group_vars/all/vault**	Note: This is a private link to your Docker EE subscription. The value for `docker_ee_url` is the URL documented at the following address: https://docs.docker.com/engine/installation/linux/docker-ee/rhel/.

docker_ee_reponame	group_vars/all/vars	For Docker EE 2.1, this variable must be set to the value `stable-18.09`
docker_ee_version	group_vars/all/vars	Specify an exact version of Docker EE to download from the repo defined by docker_ee_reponame
rhel_version	group_vars/all/vars	For the Docker installation, this sets the version of your RHEL OS, such as `7.6`. The playbooks were tested with RHEL 7.6.
dtr_version	group_vars/all/vars	Version of the Docker DTR you wish to install. You can use a numeric version or `latest` for the most recent one. The playbooks were tested with 2.6.4.
ucp_version	group_vars/all/vars	Version of the Docker UCP you wish to install. You can use a numeric version or `latest` for the most recent one. The playbooks were tested with UCP 3.1.4.
images_folder	group_vars/all/vars	Directory in the NFS server that will be mounted in the DTR nodes and that will host your Docker images.
license_file	group_vars/all/vars	Full path to your Docker EE license file on your Ansible host. The license file is available from the Docker Store
ucp_username	group_vars/all/vars	Username of the administrator user for UCP and DTR, typically `admin`.
ucp_password	**group_vars/all/vault**	The password for the `ucp_username` account.
docker_storage_driver	group_vars/all/vars	Storage driver for Docker nodes. Accepted values are `overlay2` (the default) and `devicemapper`. For RHEL 7.6, only `overlay2` is supported.

To see how to use customer-supplied certificates with UCP and DTR, see Appendix B.

Orchestrator configuration

The variable `orchestrator` in the [`worker`] group is used to specify if a worker node should be assigned to the Kubernetes orchestrator (`orchestrator: 'kubernetes'`) or to the swarm orchestrator (`orchestrator: 'swarm'`). In general, you should only change the orchestrator for worker nodes.

Note

Docker supports a third type, `mixed`, that enables workloads to be scheduled by both Kubernetes and Docker swarm on the same node. Mixing orchestrator types on the same node is not recommended for production deployments because of the likelihood of resource contention. As a result, these playbooks do not support the `mixed` type.

The following example shows how to set Kubernetes as the default orchestrator for worker nodes, and how to override the default to use Docker swarm on one specific node instead.

In the `vm_wrk_lnx.yml` file:

```
cpus: '4'
ram: '65536'
disk2_size: '500'
disk2: '/dev/sdb'
disk2_part: '/dev/sdb1'
orchestrator: kubernetes
```

In the `hosts` file:

```
[vm_wrk_lnx]
hpe-worker01 ip_addr='10.60.59.122/22' esxi_host='esx04.cloudra.local'
hpe-worker02 ip_addr='10.60.59.123/22' esxi_host='esx05.cloudra.local'
hpe-worker03 ip_addr='10.60.59.124/22' esxi_host='esx06.cloudra.local' orchestrator=swarm
```

Note

The playbooks do not change Docker's default orchestrator type which is `swarm`. Instead, the inventory is used to configure worker nodes for Kubernetes workloads or swarm workloads as explained above. If you want to change the default orchestrator type, use the method explained in

the Docker documentation at https://docs.docker.com/ee/ucp/admin/configure/set-orchestrator-type/#set-the-default-orchestrator-type-for-new-nodes.

It is possible to manually change the orchestrator type for a node. When you do this, existing workloads are evicted and they are not migrated automatically to the new orchestrator. If you want the workloads to be scheduled by the new orchestrator, you must migrate them manually. More information is available in the Docker documentation at https://docs.docker.com/ee/ucp/admin/configure/set-orchestrator-type/#what-happens-when-you-change-a-nodes-orchestrator.

Kubernetes configuration

The current playbooks support the deployment of UCP 3.1.* which deploys Kubernetes version 1.11.*. This version of the playbooks will not work with a version of UCP that is lower than 3. If you wish to deploy using UCP 2.*, you will need to download a previous release of the playbooks, which is available on the GitHub site.

The preceding section Orchestrator configuration explains how to assign a worker node to the Kubernetes orchestrator. This section covers specific Kubernetes configuration, including how to set the pod CIDR and how to configure Kubernetes Persistent Volumes.

Pod CIDR

The variable `k8s_pod_cidr` is specified in `group_vars/all/vars` and configures a custom range of IP addresses to be used by pods. The specific range that you use should be dedicated to the cluster.

The default value is `192.168.0.0/16`. To set an alternative value, use the variable as shown in the example:

```
k8s_pod_cidr: 192.168.128.0/17
```

Kubernetes Persistent Volume configuration

Variables related to the configuration of Kubernetes Persistent Volumes are shown in Table 14.

Table 14. Kubernetes Persistent Volume variables

Variable	File	Description
nfs_provisioner_namespace	group_vars/all/vars	The Kubernetes namespace, for example, `nfsstorage`
nfs_provisioner_role	group_vars/all/vars	Name of the role to create, for example, `nfs-provisioner-runner`.
nfs_provisioner_serviceaccount	group_vars/all/vars	The Kubernetes service account name to use for RBAC purposes, for example, `nfs-provisioner`
nfs_provisioner_name	group_vars/all/vars	Name of the provisioner, for example, `hpe.com/nfs`
nfs_provisioner_storage_class_name	group_vars/all/vars	Name of the storage class to create, for example, `nfs`
nfs_provisioner_server_ip	group_vars/all/vars	IP address (or FQDN) of your external NFS server, for example, `hpe2-nfs.cloudra.local`
nfs_provisioner_server_share	group_vars/all/vars	Name of the NFS share where all the persistent volume data will be stored, for example, `/k8s`

Related playbooks

The playbook `playbooks/nfs-provisioner.yml` is used to enable a dynamic NFS provisioner which can be used to automatically create and allocate Kubernetes persistent volumes. This playbook is run from the Ansible box after downloading a UCP client bundle for the `admin` account and sourcing the downloaded `env.sh` file. For more information on using this playbook, see the section Configuring storage.

Protecting sensitive information

A vault file is used to protect any sensitive variables that should not appear in clear text in your `group_vars/all/vars` file. The vault file will be encrypted and will require a password to be entered before it can be read or updated.

A sample vault file is provided named `group_vars/all/vault.sample` that you can use as a model for your vault file. To create a vault, you create a new file called `group_vars/all/vault` and add entries similar to:

```
---
docker_ee_url: 'your_url_here'
```

```
vcenter_password: 'xxxx'
vm_password: 'xxxx'
simplivity_password: 'xxxx'
ucp_password: 'zzzz'
win_password: 'yourpass'
sysdig_access_key: 'enter_sysdig_access_key'
rhn_orgid: "YourOrgId"
rhn_key: "YourActivationKey"
redhat_user: 'YourUserName'
redhat_pass: 'YourPassword'

#password for the splunk universal forwarder. Must meet password complexiy requirement (see splunk
documentation)
splunk_uf_password: 'YourPa$$word12'
oneview_config_password: 'EnterOneViewPa$$word'
#backup_passphrase must be at least 12 characters long
backup_passphrase: 'EnterYourSecretpassphrase123'
```

`rhn_orgid` and `rhn_key` are the credentials needed to subscribe the virtual machines with Red Hat Customer Portal. If these are not supplied, the playbooks will fallback to using the `redhat_user`/`redhat_pass` combination instead. For more information regarding activation keys, see the following URL: https://access.redhat.com/articles/1378093

To encrypt the vault you need to run the following command:

```
# ansible-vault encrypt group_vars/all/vault
```

You will be prompted for a password that will decrypt the vault when required. You can update the values in your vault by running:

```
# ansible-vault edit group_vars/all/vault
```

In order for Ansible to be able to read the vault, you need to specify a file where the password is stored, for instance, in a file called `.vault_pass`. Once the file is created, take the following precautions to avoid illegitimate access to this file:

- Change the permissions so only `root` can read it using `# chmod 600 .vault_pass`
- Add the file to your `.gitignore` file if you are using a Git repository to manage your playbooks.

Overview of the playbooks

The Ansible playbooks are available to download at https://github.com/HewlettPackard/Docker-Synergy. Once you have cloned the repository, change directory to `/root/Docker-Synergy`.

You can use the playbook `site.yml` as the day 0 playbook to deploy the solution. It is simply a wrapper around a number of required and optional playbooks that allow you to configure the deployment to your needs.

To start a deployment, use the following command:

```
# ansible-playbook -i hosts site.yml --vault-password-file .vault_pass
```

The playbooks should run for approximately 35-40 minutes for the default deployment with 3 UCP, 3 DTR and 3 Linux VM worker nodes (depending on your server specifications and the size of your environment).

Core components

The playbooks for deploying the core components are described in the following sections:

- Provisioning RHEL VMs
- Provisioning load balancers for UCP and DTR
- Installing Docker UCP and DTR on RHEL VMs

- Deploying RHEL workers

Optional components

The playbooks for deploying optional components are described in the following sections:

- Playbooks for adding Windows workers
- Playbooks for deploying bare metal workers on Linux and Windows
- Playbooks for installing Sysdig on RHEL
- Playbooks for installing Splunk
- Playbooks for installing Prometheus and Grafana on Kubernetes
- Playbooks for installing Prometheus and Grafana on Docker swarm

Backup and restore playbooks

Best practices and procedures are described in the section Backup and restore. The following playbooks are used to perform backups:

- `playbooks/backup_swarm.yml` is used to back up the swarm data
- `playbooks/backup_ucp.yml` is used to back up UCP
- `playbooks/backup_dtr_meta.yml` is used to back up DTR metadata
- `playbooks/backup_dtr_images.yml` is used to back up DTR images

The following playbooks are used to restore the system:

- `playbooks/restore_dtr_images.yml` is used to restore DTR images
- `playbooks/restore_dtr_metadata.yml` is used to restore DTR metadata
- `playbooks/restore_ucp.yml` is used to restore UCP

Convenience playbooks

- `playbooks/install_kubectl.yml` downloads and installs kubectl on the Ansible controller.
- `playbooks/install_client_bundle.yml` installs and configures the UCP bundle on the Ansible controller.
- `playbooks/install_helm.yml` downloads and installs helm on the Ansible controller.
- `playbooks/clean_all.yml` powers off and deletes all VMs in your inventory.
- `playbooks/distribute_keys.yml` distributes public keys between all nodes, to allow each node to password-less log in to every other node. As this is not essential and can be regarded as a security risk (a worker node probably should not be able to log in to a UCP node, for instance), this playbook is not included in `site.yml` by default.

Convenience scripts

- `backup.sh` can be used to take a backup of the swarm, UCP, DTR metadata and the DTR images in one go.
- `restore_dtr.sh` can be used to restore DTR metadata and DTR images.
- `scale_worker.sh` can be used to scale the worker nodes.

Deploying the core components

At this point, the system is ready to be deployed. Make sure you are logged on as `root` in your Ansible box and that your current directory is `/root/Docker-Synergy`

Note
As well as configuring your `vars` and `vault` files, you must also provide a `backups` configuration file in the `group_vars/all` folder when running `site.yml`. An example file is provided in the repository named `backups.sample`. Rename it to `backups` before running the playbooks. Details on how to configure this file are available in the section Backup and restore.

Provisioning RHEL VMs

The following playbooks are used to provision RHEL VMs:

- `playbooks/provision_nodes.yml` will create all the necessary virtual machines for the environment from the VM Template defined in the `vm_template` variable. All Linux VMs are now created in one go, regardless of the number of drives they have. This playbook also has the potential to configure additional network adapters. Note that this playbook will also provision any Linux or Windows bare metal nodes that are configured in the inventory.
- `playbooks/config_networking.yml` will configure the network settings in all the virtual machines.
- `playbooks/resize_syspart.yml` resizes the logical volume that holds the `/` partition of the Linux VMs to use all the space available on the drive.
- `playbooks/config_subscription.yml` registers and subscribes all virtual machines to the Red Hat Customer Portal.
- `playbooks/config_ntp.yml` configures the **chrony** client package in all virtual machines in order to have a synchronized clock across the environment. It will use the list of servers specified in the `ntp_servers` variable in the file `group_vars/all/vars`.

Provisioning load balancers for UCP and DTR

The playbook `playbooks/loadbalancer.yml` is used to deploy load balancers in an `active-active` configuration to provide highly-available access to UCP and DTR.

At least two nodes are specified in the [`loadbalancer`] group in the inventory, along with group variables defining CPU and RAM requirements. These nodes run `keepalived` and `HAproxy`.

```
[loadbalancer]
hpe-lb1 ip_addr='10.60.59.248/22' esxi_host='esx04.am2.cloudra.local' ucp=true
hpe-lb2 ip_addr='10.60.59.249/22' esxi_host='esx05.am2.cloudra.local' dtr=true

[loadbalancer:vars]
cpus='2'
ram='4096'
```

The virtual IP for UCP will be handled by `hpe-lb1` by default, which will split the traffic across the three UCP VMs. In the case of a failure of `hpe-lb1`, the virtual IP for UCP will automatically move to the second load balancer node `hpe-lb2` which will again distribute the traffic to the UCP VMs.

Similarly, the virtual IP for DTR will be handled by default by the load balancer `hpe-lb2`, splitting the traffic across the three DTR VMs. In the case of a failure of `hpe-lb2`, the virtual IP for DTR will automatically move to the first load balancer node `hpe-lb1` which will again distribute the traffic to the DTR VMs.

To configure the virtual IPs for UCP and DTR, you need to add a `loadbalancers` dictionary to your `group_vars/all/vars` file as shown in the excerpt below:

```
loadbalancers:
  ucp:
    public_interface: 'ens192'
    public_vip: '10.60.59.251'
    public_fqdn: hpe-ucpvip.cloudra.local
    virtual_router_id: 54
  dtr:
```

```
    public_interface: 'ens192'
    public_vip: '10.60.59.252'
    public_fqdn: hpe-dtrvip.cloudra.local
    virtual_router_id: 55
```

Warning
If you re-run `playbooks/loadbalancer.yml` after a configuration change, you may need to subsequently run `playbooks/reconfigure_dtr.yml` as the latter playbook configures the virtual IP address for accessing the UCP Single-Sign-On (SSO) page. If there is no virtual IP or FQDN defined for UCP in the variables file, the playbook will choose the address of the first UCP node in the [`ucp`] group. This scenario introduces a single point of failure and should be avoided.

Note
By default, the playbook supports ports `443` and `6443` for UCP and port `433` for DTR. If you deploy Prometheus and Grafana on Docker Swarm, the Grafana port `3000` will be handled as well.

Note
The playbook `playbooks/loadbalancer.yml` can be used to create one or more load balancers for applications running on your worker nodes. However, it is impossible for the playbooks to know what ports to support, so manual configuration of HAproxy and `keepalived` may be required. By default, the playbooks support ports `80` and `443` for worker nodes.

Legacy stand-alone load balancers

The playbook `playbooks/install_haproxy.yml` is used to deploy three separate load balancers, for the UCP, DTR and worker nodes. It is recommended that you use the HAproxy and `keepalived` solution documented above instead of this option.

Deploying without load balancers

If you do not want to deploy load balancers when running `site.yml`, you should comment out any declarations in the inventory and variables files. This includes any legacy stand-alone load balancers.

Deploying with your own load balancers

If you are using external load balancers for UCP and DTR, you can configure UCP and DTR to use these external load balancers by specifying FQDNs in the loadbalancers dictionary in `group_vars/all/vars`:

```
loadbalancers:
  ucp:
    public_fqdn: external-ucpvip.am2.cloudra.local
  dtr:
    public_fqdn: external-dtrvip.am2.cloudra.local
```

Installing Docker UCP and DTR on RHEL VMs

The following playbooks are used to install Docker UCP and DTR on RHEL VMs.

- `playbooks/config_storage_driver.yml` prepares drives for local Docker volumes and container images. It also configures Docker with the `overlay2` storage driver (the default). This playbook was previously called `playbooks/config_docker_lvs.yml` in earlier releases of the solution.
- `playbooks/install_docker.yml` installs Docker along with all of its dependencies.
- `playbooks/install_rsyslog.yml` installs and configures **rsyslog** in the logger node and in all Docker nodes. The logger node will be configured to receive all `syslogs` on port `514` and the Docker nodes will be configured to send all logs (including container logs) to the logger node.

- `playbooks/docker_post_config.yml` performs a variety of tasks to complete the installation of the Docker environment, including configuration of the HTTP/HTTPS proxies, if any, and installation of the VMware vSphere Storage for Docker volume plugin.
- `playbooks/install_nfs_server.yml` installs and configures an NFS server on the NFS node.

  This playbook has been updated to configure a third drive which is used to hold the data of the persistent volumes created with the NFS provisioner. The default size for this drive is purposefully kept small because using the NFS VM to store persistent volumes is not recommended for production use. However, this can be useful for demo purposes.
- `playbooks/install_nfs_clients.yml` installs the required packages on the DTR nodes to be able to mount an NFS share.
- `playbooks/create_main_ucp.yml` installs and configures the first Docker UCP instance on the target node defined by the group `ucp_main` in the `hosts` inventory.
- `playbooks/scale_ucp.yml` installs and configures additional instances of UCP on the target nodes defined by the group `ucp` in the `hosts` inventory, except for the node defined in the group `ucp_main`.
- `playbooks/create_main_dtr.yml` installs and configures the first Docker DTR instance on the target node defined by the group `dtr_main` in the `hosts` inventory.
- `playbooks/config_scheduler.yml` configures the scheduler to prevent regular users (i.e. non-admin users) scheduling containers on the Docker nodes running instances of UCP and DTR.
- `playbooks/scale_dtr.yml` installs and configures additional instances (or replicas) of DTR on the target nodes defined by the group `dtr` in the `hosts` inventory, with the exception of the node defined in the group `dtr_main`.
- `playbooks/reconfigure_dtr.yml` is used to reconfigure DTR with the FQDN of the UCP Load Balancer for Single Sign On (SSO) purposes and also enables image scanning.

Deploying RHEL workers

By default, `site.yml` will automatically deploy any RHEL (and / or Windows) worker nodes that are declared in the inventory.

If you subsequently want additional RHEL worker nodes, add them to the inventory as appropriate and then run the playbooks for Provisioning RHEL VMs, followed by the specific playbooks for RHEL worker nodes outlined below:

- `playbooks/scale_workers.yml` installs and configures additional Linux workers on the target nodes defined by the group `worker` in the `hosts` inventory.

A utility script `scale_worker.sh` is provided to assist you in adding worker nodes after the initial deployment.

Post deployment

The playbooks in `site.yml` are intended to be used to deploy a new environment. You should only use them for Day 0 deployment purposes.

The Ansible log is stored in the folder `/root/Docker-Synergy`. If the deployment fails, you may find useful hints in this log. To see how to check if your certs have been deployed correctly, see Appendix D: How to check that certs were deployed correctly.

Installing kubectl

A convenience playbook is provided to make it easy to install `kubectl` on the Ansible controller. This playbook uses variables in `group_vars/all/vars` to determine which version to download. The default version specified by the variable `kubectl_version` in the sample variables file is `1.11.5`. Details of the `1.11` release are available at https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG-1.11.md. In particular, the playbook requires a checksum to be present in the variable `kubectl_checksum`. The appropriate value can be found in the details for the specific version of `kubectl` to be downloaded, in this case for version 1.11.5 of `kubernetes-client-linux-amd64.tar.gz`, available at https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG-1.11.md#downloads-for-v1115.

The `vars.sample` file that ships with this release has the following values:

```
kubectl_version: "1.11.5"
kubectl_checksum:
```

```
"sha512:7028d357f65603398c35b7578793a153248e17c2ad631541a587f4ae13ef93f058db130390eea4820c2fd7707509ed
0eb581cb129790b12680e869829a6fc241"
```

To run the playbook:

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts playbooks/install_kubectl.yml
```

Test the installation by running the `kubectl version` command:

```
# kubectl version

Client Version: version.Info{Major:"1", Minor:"11", GitVersion:"v1.11.5",
GitCommit:"753b2dbc622f5cc417845f0ff8a77f539a4213ea", GitTreeState:"clean", BuildDate:"2018-11-
26T14:41:50Z", GoVersion:"go1.10.3", Compiler:"gc", Platform:"linux/amd64"}

The connection to the server localhost:8080 was refused - did you specify the right host or port?
```

The client version is reported correctly. However, `kubectl` cannot connect to the server until you set up a client bundle - this is described in the section titled Installing the client bundle.

Manually installing kubectl

You can find the version number for the current stable version of `kubectl` at https://kubernetes.io/docs/tasks/tools/install-kubectl/. At the time of writing, the stable version is `1.13`.

The following is an example of manually downloading and installing a specific version of `kubectl`.

```
# version=v1.10.4
# wget -O kubectl https://storage.googleapis.com/kubernetes-
release/release/${version}/bin/linux/amd64/kubectl
# chmod +x ./kubectl
# sudo mv ./kubectl /usr/local/bin/kubectl
```

```
# kubectl version
Client Version: version.Info{Major:"1", Minor:"10", GitVersion:"v1.10.4",
GitCommit:"5ca598b4ba5abb89bb773071ce452e33fb66339d", GitTreeState:"clean", BuildDate:"2018-06-
06T08:13:03Z", GoVersion:"go1.9.3", Compiler:"gc", Platform:"linux/amd64"}

Server Version: version.Info{Major:"1", Minor:"8+", GitVersion:"v1.8.11-docker-8d637ae",
GitCommit:"8d637aedf46b9c21dde723e29c645b9f27106fa5", GitTreeState:"clean", BuildDate:"2018-04-
26T16:51:21Z", GoVersion:"go1.8.3", Compiler:"gc", Platform:"linux/amd64"}
```

More details on installing `kubectl` are available at https://kubernetes.io/docs/tasks/tools/install-kubectl/.

Installing the client bundle

A convenience playbook is provided to install and apply the client bundle on the Ansible controller. To run the playbook:

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts playbooks/install_client_bundle.yml --vault-password-file .vault_pass
```

The client bundle is downloaded to `~/certs.<<ucp_instance>>.<<ucp_username>>` where `ucp_instance` will be specific to the cluster you are running against, for example, `hpe2-ucp01` and the `ucp-username` is typically `admin`.

The playbook downloads the client bundle, but does not configure it for use. Change to the download folder and execute `eval "$(<env.sh)"`

```
# cd ~/certs.hpe2-ucp01.admin
# eval "$(<env.sh)"
```

Test the configuration by again running the `kubectl version` command. It should now report the server version as well as the client version:

```
# kubectl version

Client Version: version.Info{Major:"1", Minor:"11", GitVersion:"v1.11.5",
GitCommit:"753b2dbc622f5cc417845f0ff8a77f539a4213ea", GitTreeState:"clean", BuildDate:"2018-11-
26T14:41:50Z", GoVersion:"go1.10.3", Compiler:"gc", Platform:"linux/amd64"}

Server Version: version.Info{Major:"1", Minor:"11+", GitVersion:"v1.11.5-docker-1",
GitCommit:"d512ba512d0de40cd80258f480ff66bf71f2d8a4", GitTreeState:"clean", BuildDate:"2018-12-
03T19:55:14Z", GoVersion:"go1.10.3", Compiler:"gc", Platform:"linux/amd64"}
```

More information on the client bundle is available at https://docs.docker.com/ee/ucp/user-access/cli/#download-client-certificates-by-using-the-rest-api.

Installing Helm

Prerequisites

- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Running the playbook

To run the playbook on your Ansible controller:

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts playbooks/install_helm.yml --vault-password-file .vault_pass
```

The playbook relies on the variable `helm_version` to determine the version of Helm to download. The playbooks have been tested using version `2.12.3`. You must also specify the appropriate checksum for the download in the variable `helm_checksum`. This value can be obtained from the downloads page at https://github.com/helm/helm/releases. The `vars.sample` file that ships with this release contains the following values:

```
helm_version: "2.12.3"
helm_checksum: "sha256:3425a1b37954dabdf2ba37d5d8a0bd24a225bb8454a06f12b115c55907809107"
```

Install sample charts

A number of sample charts are delivered with the solution, for the purposes of demonstration.

Alpine

A simple chart is provided in the `~/Docker-Synergy/test/files/helm/alpine` directory to run a single pod of Alpine Linux.

The `templates/` directory contains a very simple pod resource with a couple of parameters. The `values.yaml` file contains the default values for the `alpine-pod.yaml` template.

```
# cd ~/Docker-Synergy
# helm install test/files/helm/alpine
```

The output shows that a single pod was deployed.

```
NAME:   old-mole
LAST DEPLOYED: Fri Feb  8 17:27:35 2019
NAMESPACE: default
STATUS: DEPLOYED

RESOURCES:
==> v1/Pod
NAME             READY  STATUS   RESTARTS  AGE
old-mole-alpine  1/1    Running  0         0s
```

Nginx

An example chart is provided in the `~/Docker-Synergy/test/files/helm/nginx` directory to install a simple nginx server according to the following pattern:

- A ConfigMap is used to store the files the server will serve. (`templates/configmap.yaml`)
- A Deployment is used to create a Replica Set of nginx pods. (`templates/deployment.yaml`)
- A Service is used to create a gateway to the pods running in the replica set (`templates/service.yaml`)

The `values.yaml` exposes a few of the configuration options in the charts.

```
# cd ~/Docker-Synergy
# helm install test/files/helm/nginx
```

The output shows a service being created with a NodePort at `34567`. This value comes from the `values.yml` file in the folder.

```
NAME:   worn-olm
LAST DEPLOYED: Fri Feb  8 16:23:21 2019
NAMESPACE: default
STATUS: DEPLOYED

RESOURCES:
==> v1/Deployment
NAME            DESIRED  CURRENT  UP-TO-DATE  AVAILABLE  AGE
worn-olm-nginx  1        1        1           1          14s

==> v1/Pod(related)
NAME                             READY  STATUS     RESTARTS  AGE
worn-olm-nginx-7d648f7dfb-gg2jk  1/1    Running    0         14s
worn-olm-nginx-vhwc7             0/1    Completed  0         14s

==> v1/ConfigMap
NAME            DATA  AGE
worn-olm-nginx  2     14s

==> v1/Service
NAME            TYPE      CLUSTER-IP    EXTERNAL-IP  PORT(S)       AGE
worn-olm-nginx  NodePort  10.96.30.222  <none>       80:34567/TCP  14s
```

Helm also allows you to easily delete installed releases. List the installed releases to find the name of the release you wish to delete.

```
# helm list
NAME            REVISION        UPDATED                         STATUS          CHART           APP
VERSION     NAMESPACE
worn-olm        1               Fri Feb  8 16:23:21 2019        DEPLOYED        nginx-0.1.0
default
```

Use the `helm delete` command to remove the named release.

```
# helm delete worn-olm
release "worn-olm" deleted
```

Post-deploy validation

Many sample Kubernetes applications are available at https://kubernetes.io/docs/tutorials/. This section details how to deploy the stateless `guestbook` application with Redis as documented at https://kubernetes.io/docs/tutorials/stateless-application/guestbook/.

When deploying applications, you must be aware that Kubernetes version 1.11 shipped with Docker 2.1. If you are testing examples that are designed to work with a newer (or older) version of Kubernetes, you may have to make changes in some places to the configuration files.

Prerequisites

- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Kubernetes guestbook example with Redis

The playbook for the Kubernetes example guestbook is based on the example taken from the GitHub repo at https://github.com/kubernetes/examples.

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts test/playbooks/k8s-guestbook.yml --vault-password-file .vault_pass
```

You can run the playbook directly, but it can be informative to walk through the individual files to see what is going on under the covers.

Quickstart

```
# cd ~/Docker-Synergy/test/files/k8s-examples/guestbook
# kubectl apply -f redis-master-deployment.yaml
# kubectl apply -f redis-master-service.yaml
# kubectl apply -f redis-slave-deployment.yaml
# kubectl apply -f redis-slave-service.yaml
# kubectl apply -f frontend-deployment.yaml
# kubectl apply -f frontend-service.yaml
# kubectl get svc frontend
```

Details

Change to the directory containing the guestbook YAML files.

```
# cd ~/Docker-Synergy/test/files/k8s-examples/guestbook
```

The manifest file `redis-master-deployment.yaml`, included below, specifies a deployment controller that runs a single replica Redis master pod.

```
# cat redis-master-deployment.yaml
```

```
apiVersion: apps/v1 #  for k8s versions before 1.9.0 use apps/v1beta2  and before 1.8.0 use extensions/v1beta1
kind: Deployment
metadata:
  name: redis-master
spec:
  selector:
    matchLabels:
      app: redis
      role: master
      tier: backend
  replicas: 1
  template:
    metadata:
      labels:
        app: redis
        role: master
        tier: backend
    spec:
      containers:
      - name: master
        image: k8s.gcr.io/redis:e2e  # or just image: redis
        resources:
          requests:
```

```
            cpu: 100m
            memory: 100Mi
        ports:
        - containerPort: 6379
```

Apply the Redis master deployment from the redis-master-deployment.yaml file:

```
# kubectl apply -f redis-master-deployment.yaml
```

Query the list of Pods to verify that the Redis master pod is running.

```
# kubectl get pods | grep redis
redis-master-57657796fc-psvhc     1/1       Running   0          32s
```

Use the `kubectl logs` command to view the logs from the Redis master pod:

```
# kubectl logs -f redis-master-57657796fc-psvhc
```

```
                _.-
            _.-``__ ''-._
       _.-``    `.  `_.  ''-._           Redis 2.8.19 (00000000/0) 64 bit
   .-`` .-```.  ```\/    _.,_ ''-._
  (    '      ,       .-`  | `,    )     Running in stand alone mode
  |`-._`-...-` __...-.``-._|'` _.-'|     Port: 6379
  |    `-._   `._    /     _.-'    |     PID: 1
   `-._    `-._  `-./  _.-'    _.-'
  |`-._`-._    `-.__.-'    _.-'_.-'|
  |    `-._`-._        _.-'_.-'    |           http://redis.io
   `-._    `-._`-.__.-'_.-'    _.-'
  |`-._`-._    `-.__.-'    _.-'_.-'|
  |    `-._`-._        _.-'_.-'    |
   `-._    `-._`-.__.-'_.-'    _.-'
       `-._    `-.__.-'    _.-'
           `-._        _.-'
               `-.__.-'
```

```
[1] 07 Feb 15:04:32.189 # Server started, Redis version 2.8.19
[1] 07 Feb 15:04:32.189 # WARNING you have Transparent Huge Pages (THP) support enabled in your
kernel. This will create latency and memory usage issues with Redis. To fix this issue run the command
'echo never > /sys/kernel/mm/transparent_hugepage/enabled' as root, and add it to your /etc/rc.local
in order to retain the setting after a reboot. Redis must be restarted after THP is disabled.
[1] 07 Feb 15:04:32.189 # WARNING: The TCP backlog setting of 511 cannot be enforced because
/proc/sys/net/core/somaxconn is set to the lower value of 128.
[1] 07 Feb 15:04:32.190 * The server is now ready to accept connections on port 6379
```

The guestbook application needs to communicate with the Redis master to write its data. You need to apply a service to proxy the traffic to the Redis master pod. A service defines a policy to access the pods.

```
# cat redis-master-service.yaml
```

```
apiVersion: v1
kind: Service
metadata:
  name: redis-master
  labels:
    app: redis
```

```
    role: master
    tier: backend
spec:
  ports:
  - port: 6379
    targetPort: 6379
  selector:
    app: redis
    role: master
    tier: backend
```

Apply the Redis master service from the `redis-master-service.yaml` file:

```
#  kubectl apply -f redis-master-service.yaml
service "redis-master" created
```

Query the list of services to verify that the Redis master service is running.

```
# kubectl get svc
NAME           TYPE        CLUSTER-IP      EXTERNAL-IP   PORT(S)          AGE
redis-master   ClusterIP   10.96.240.18    <none>        6379/TCP         1m
```

Although the Redis master is a single pod, you can make it highly available to meet traffic demands by adding replica Redis slaves.

```
# cat redis-slave-deployment.yaml
```

```
apiVersion: apps/v1 #  for k8s versions before 1.9.0 use apps/v1beta2  and before 1.8.0 use
extensions/v1beta1
kind: Deployment
metadata:
  name: redis-slave
spec:
  selector:
    matchLabels:
      app: redis
      role: slave
      tier: backend
  replicas: 2
  template:
    metadata:
      labels:
        app: redis
        role: slave
        tier: backend
    spec:
      containers:
      - name: slave
        image: gcr.io/google_samples/gb-redisslave:v1
        resources:
          requests:
            cpu: 100m
            memory: 100Mi
        env:
        - name: GET_HOSTS_FROM
          value: dns
          # If your cluster config does not include a dns service, then to
```

```
          # instead access an environment variable to find the master
          # service's host, comment out the 'value: dns' line above, and
          # uncomment the line below:
          # value: env
        ports:
        - containerPort: 6379
```

Create the Redis slaves from the `redis-slave-deployment.yaml` file.

```
# kubectl apply -f  redis-slave-deployment.yaml
deployment.apps "redis-slave" created
```

Query the list of Pods to verify that the Redis slave pods are running.

```
# kubectl get pods | grep redis
redis-master-57657796fc-psvhc     1/1       Running   0          7m
redis-slave-5cb5956459-bqqlg      1/1       Running   0          19s
redis-slave-5cb5956459-gql5x      1/1       Running   0          19s
```

The guestbook application needs to communicate to Redis slaves to read data. To make the Redis slaves discoverable, you need to set up a service that provides transparent load balancing to the set of pods.

```
# cat redis-slave-service.yaml
apiVersion: v1
kind: Service
metadata:
  name: redis-slave
  labels:
    app: redis
    role: slave
    tier: backend
spec:
  ports:
  - port: 6379
  selector:
    app: redis
    role: slave
    tier: backend
```

Deploy the Redis slave service from the `redis-slave-service.yaml` file.

```
# kubectl apply -f redis-slave-service.yaml
service "redis-slave" created
```

Query the list of services to verify that the Redis slave service is running.

```
# kubectl get services | grep redis
redis-master   ClusterIP   10.96.240.18    <none>        6379/TCP         4m
redis-slave    ClusterIP   10.96.200.85    <none>        6379/TCP         22s
```

The guestbook application has a web frontend written in PHP serving the HTTP requests. It is configured to connect to the `redis-master` service for write requests and the `redis-slave` service for read requests.

```
# cat frontend-deployment.yaml
apiVersion: apps/v1 #  for k8s versions before 1.9.0 use apps/v1beta2  and before 1.8.0 use
extensions/v1beta1
```

```
kind: Deployment
metadata:
  name: frontend
spec:
  selector:
    matchLabels:
      app: guestbook
      tier: frontend
  replicas: 3
  template:
    metadata:
      labels:
        app: guestbook
        tier: frontend
    spec:
      containers:
      - name: php-redis
        image: gcr.io/google-samples/gb-frontend:v4
        resources:
          requests:
            cpu: 100m
            memory: 100Mi
        env:
        - name: GET_HOSTS_FROM
          value: dns
          # If your cluster config does not include a dns service, then to
          # instead access environment variables to find service host
          # info, comment out the 'value: dns' line above, and uncomment the
          # line below:
          # value: env
        ports:
        - containerPort: 80
```

Create the frontend deployment using the `frontend-deployment.yaml` file.

```
# kubectl apply -f frontend-deployment.yaml
deployment.apps "frontend" created
```

Query the list of pods to verify that the three frontend replicas are running.

```
# kubectl get pods -l app=guestbook -l tier=frontend
NAME                        READY     STATUS    RESTARTS   AGE
frontend-7f5cd767dc-28j6b   1/1       Running   0          23s
frontend-7f5cd767dc-mqcbv   1/1       Running   0          23s
frontend-7f5cd767dc-v6lwc   1/1       Running   0          23s
```

If you want guests to be able to access your guestbook, you must configure the frontend service to be externally visible, so a client can request the service from outside the container cluster.

```
# cat frontend-service.yaml
apiVersion: v1
kind: Service
metadata:
  name: frontend
```

```
  labels:
    app: guestbook
    tier: frontend
spec:
  # comment or delete the following line if you want to use a LoadBalancer
  type: NodePort
  # if your cluster supports it, uncomment the following to automatically create
  # an external load-balanced IP for the frontend service.
  # type: LoadBalancer
  ports:
  - port: 80
  selector:
    app: guestbook
    tier: frontend
```

Deploy the frontend service using the `frontend-service.yaml` file.

```
# kubectl apply -f frontend-service.yaml
service "frontend" created
```

Query the list of services to verify that the frontend service is running.

```
#  kubectl get services | grep frontend
frontend       NodePort    10.96.16.200    <none>        80:33444/TCP     25s
```

Access the UI using the identified port on any node in your cluster, for example, `http://hpe2-ucp01.am2.cloudra.local:33444/` as shown in Figure 5.

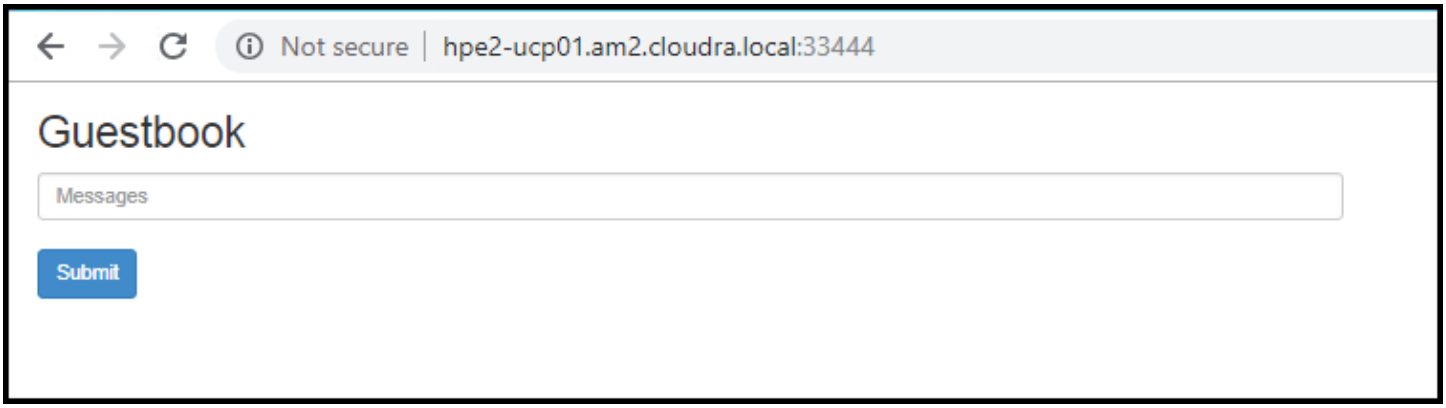

Figure 5. Guestbook UI

Teardown

A playbook is provided to remove the deployed `guestbook` artifacts.

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts test/playbooks/k8s-guestbook-teardown.yml --vault-password-file
.vault_pass
```

UCP metrics in Prometheus

Docker EE 2.1 uses a built-in deployment of Prometheus to power the performance graphs in the web UI for UCP. The metrics that UCP generates can be routed to a separate Prometheus, if required. A convenience playbook has been provided to configure a minimal Prometheus and Grafana deployment that can help vizualize all of the metrics that UCP generates.

For more information on UCP cluster metrics, see the article at https://docs.docker.com/ee/ucp/admin/configure/collect-cluster-metrics/.

Prerequisites

- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Deploy Prometheus and Grafana

The playbook `playbooks/ucp-metrics-prometheus.yml` deploys pods for Prometheus and Grafana and configures them to use the client bundle to access the UCP metrics. To run the playbook:

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts playbooks/ucp-metrics-prometheus.yml --vault-password-file .vault_pass
```

Prometheus UI

The playbook exposes a port to access the user interface for Prometheus - to find the port, get the details of the `prometheus` service:

```
# kubectl get svc Prometheus
NAME         TYPE       CLUSTER-IP      EXTERNAL-IP   PORT(S)          AGE
prometheus   NodePort   10.96.216.220   <none>        9090:34713/TCP   6d
```

The Prometheus UI can be accessed on any node in your cluster, using the port returned by `kubectl get svc`. In this instance, it is accessed at `http://hpe2-ucp01.am2.cloudra.local:34713` as shown in Figure 6.

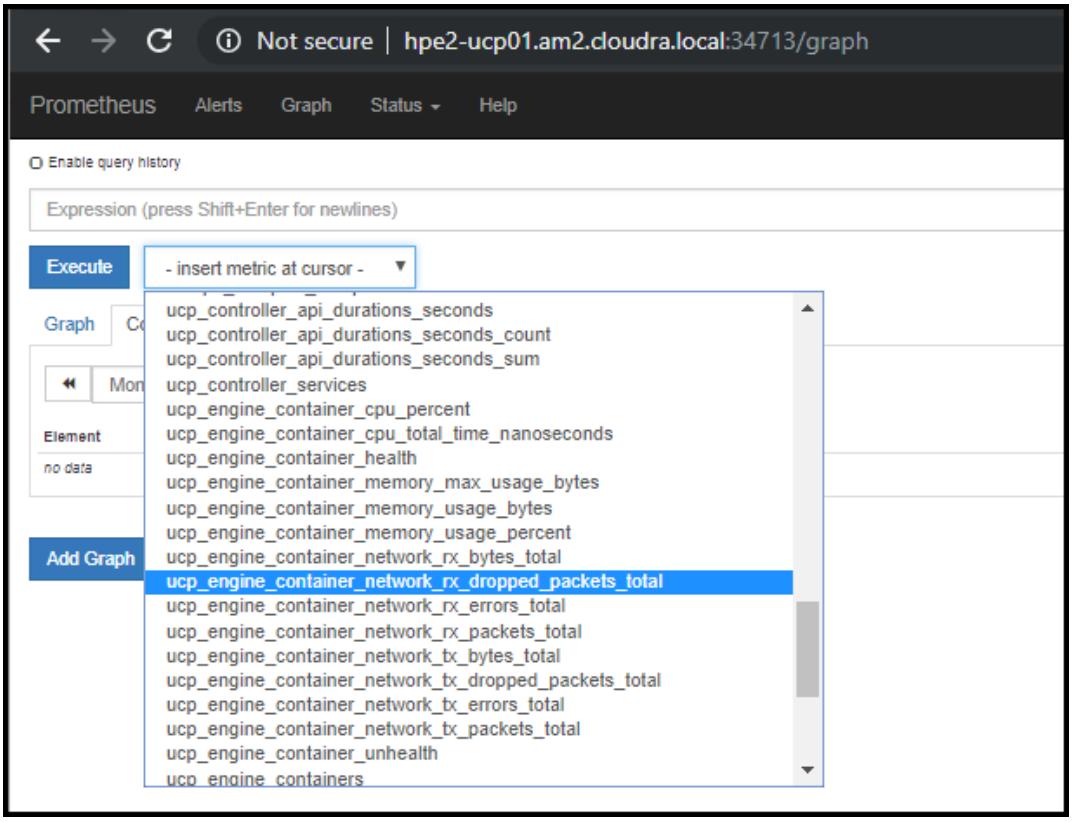

Figure 6. UCP metrics in Prometheus

Using Grafana to vizualize UCP metrics

The playbook also exposes a port to access the Grafana UI - to find the port, get the details of the grafana service:

```
# kubectl get svc grafana
NAME      TYPE       CLUSTER-IP      EXTERNAL-IP   PORT(S)          AGE
grafana   NodePort   10.96.177.108   <none>        3000:33118/TCP   6d
```

The Grafana UI can be accessed on any node in your cluster, using the port returned by `kubectl get svc`. In this instance, it is accessed at `http://hpe2-ucp01.am2.cloudra.local:33118`. The example UCP Dashboard shown in Figure 7 is taken from https://grafana.com/dashboards/9309.

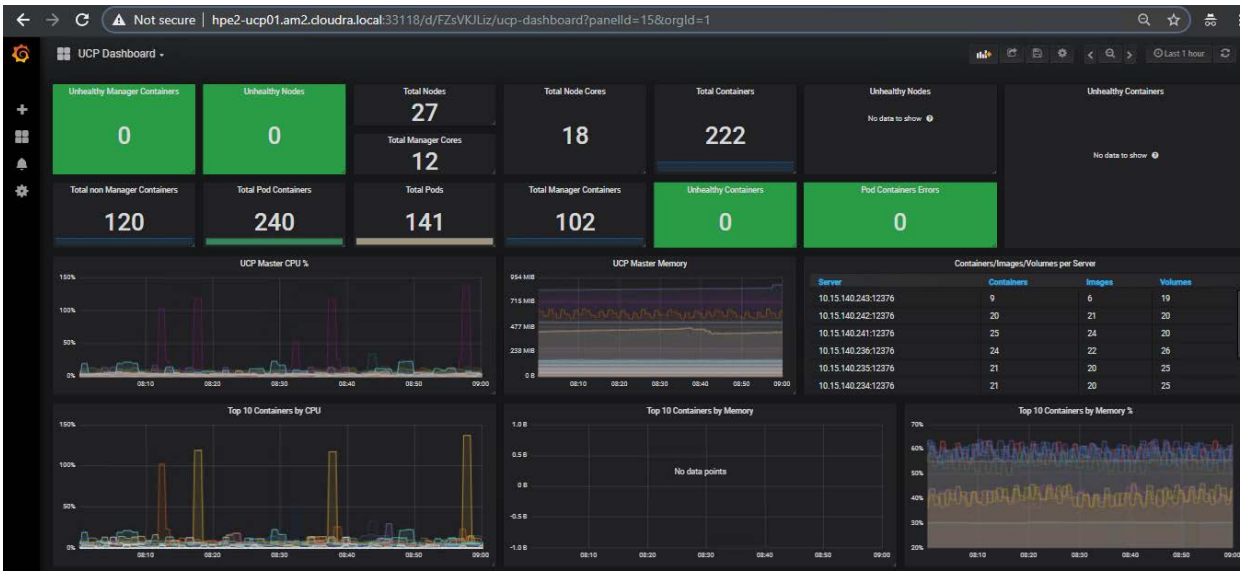

Figure 7. UCP Dashboard in Grafana

Configuring storage

Using HPE 3PAR when deploying NFS provisioner for Kubernetes

Prerequisites

- Configure the variables described in the section Kubernetes Persistent Volume configuration
- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Setting up HPE 3PAR

The following section outlines the steps you need to follow in order to configure a Virtual File Server and a share for use by the Kubernetes NFS provisioner.

Log in to the HPE 3PAR StoreServ Management console and perform the following tasks.

Create a virtual file server (VFS):

1. In the General section, specify a name, in this instance `hpe_vfs3par`

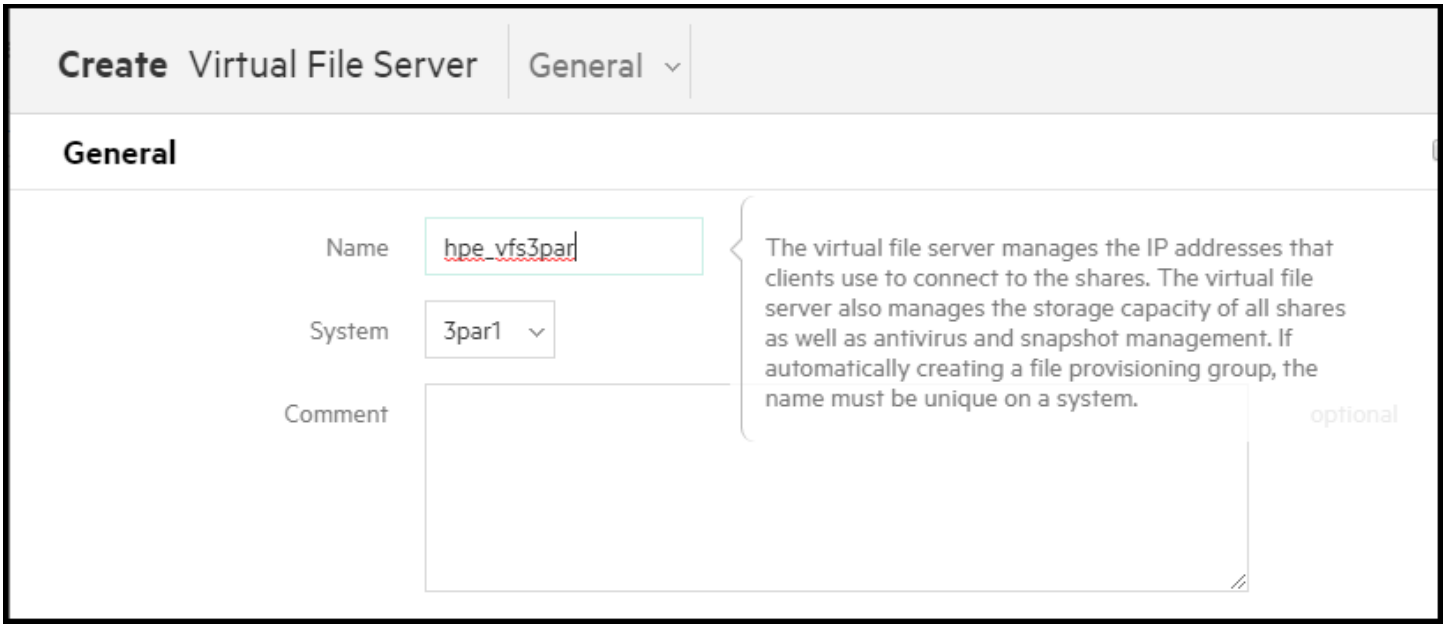

Figure 8. Create Virtual File Server - General

2. In the Storage Allocation Settings section, set the Provisioning to Thin Provisioned, select an appropriate CPG, in this instance FC_r1, and set the size, for example, 1 terabyte.

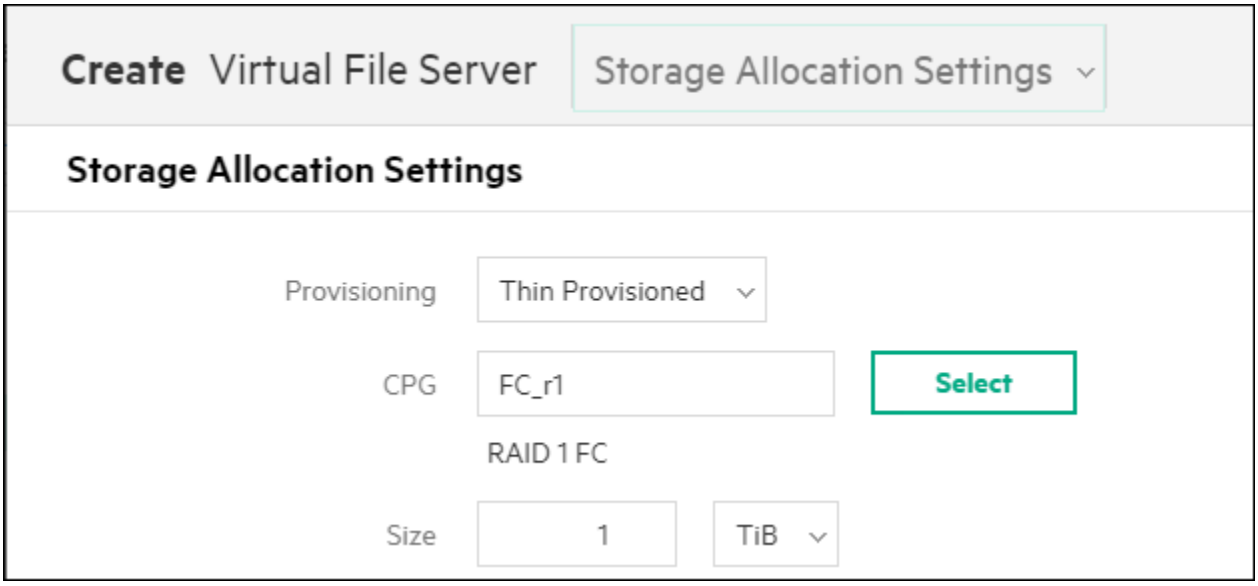

Figure 9. Create Virtual File Server - Storage Allocation Settings

3. Add a virtual IP address

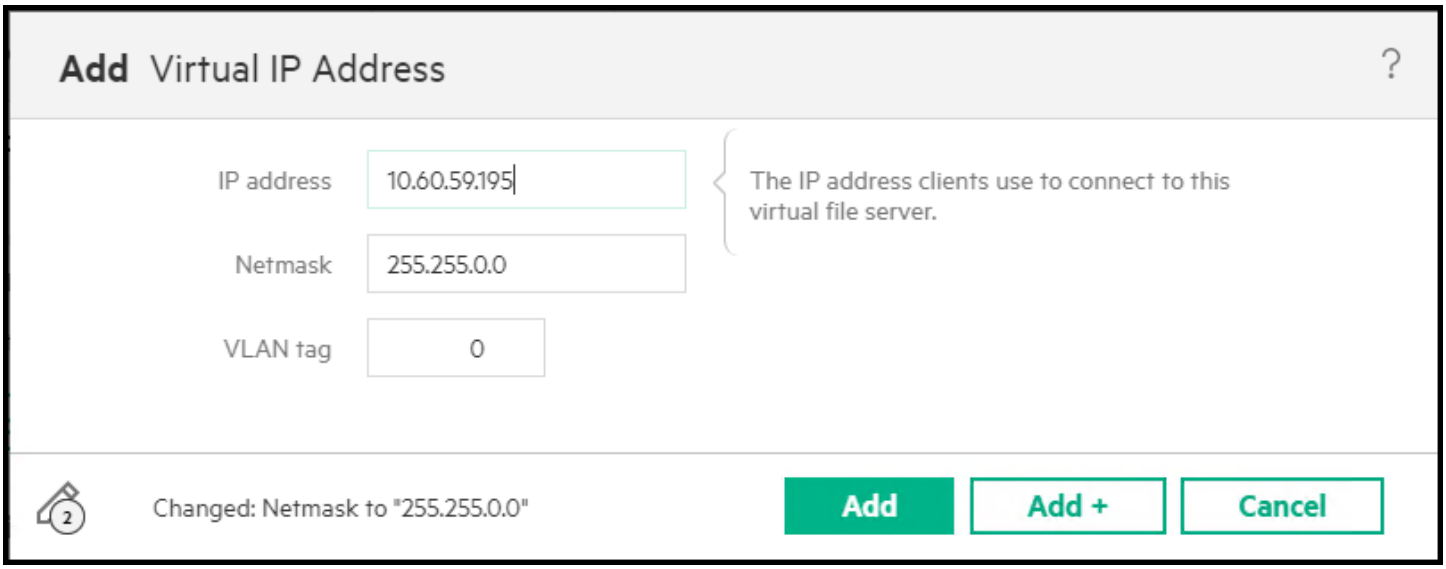

Figure 10. Create Virtual File Server - Add Virtual IP Address

These steps result in a Virtual File Server:

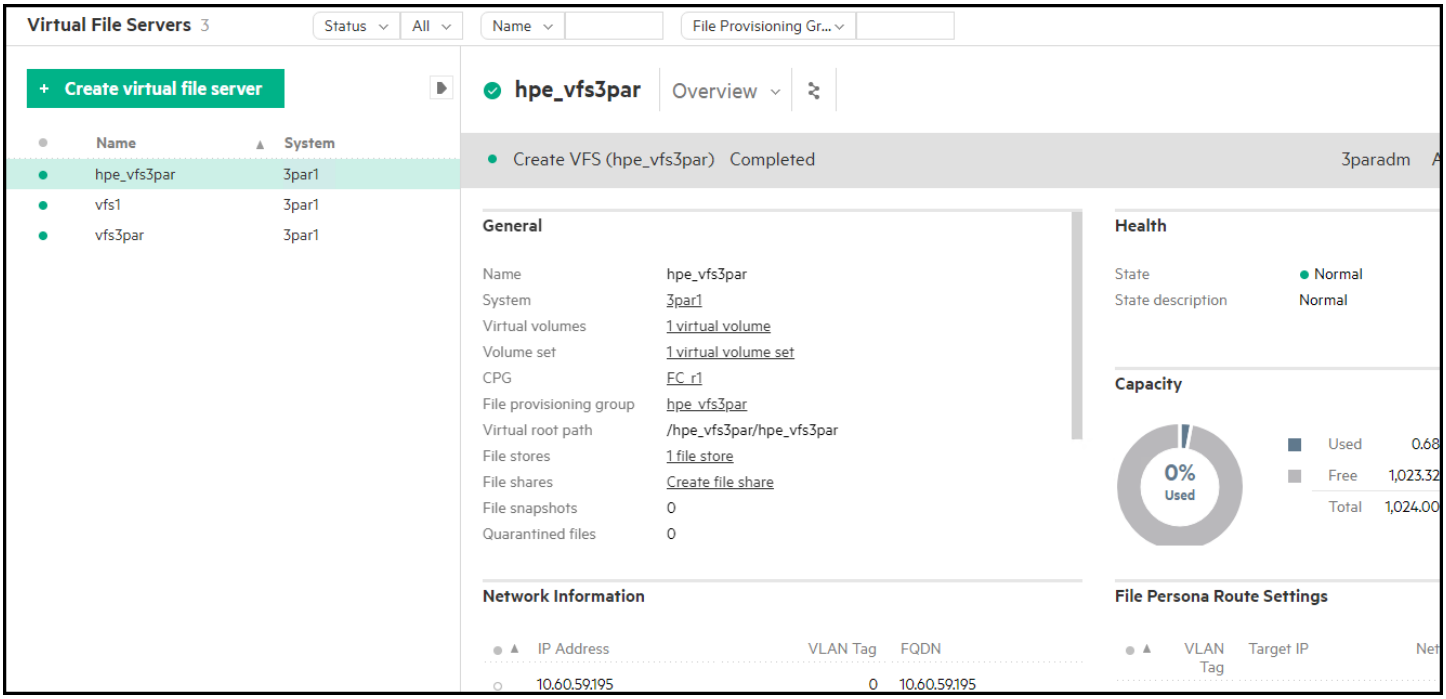

Figure 11. Virtual File Server

Create a File Store:

1. In the General section, specify a name, in this instance HPE_filestore3par, and select the Virtual File Server that you just created.

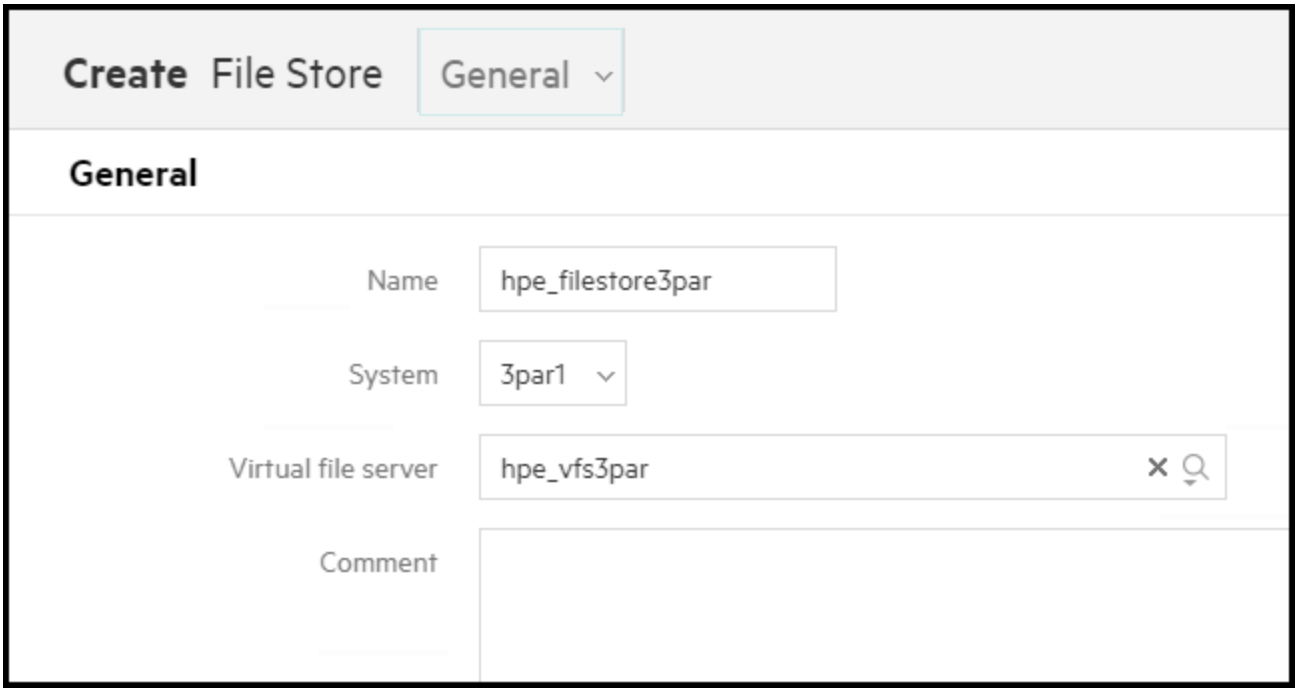

Figure 12. Create File Store – General

2. Use the default Security settings:

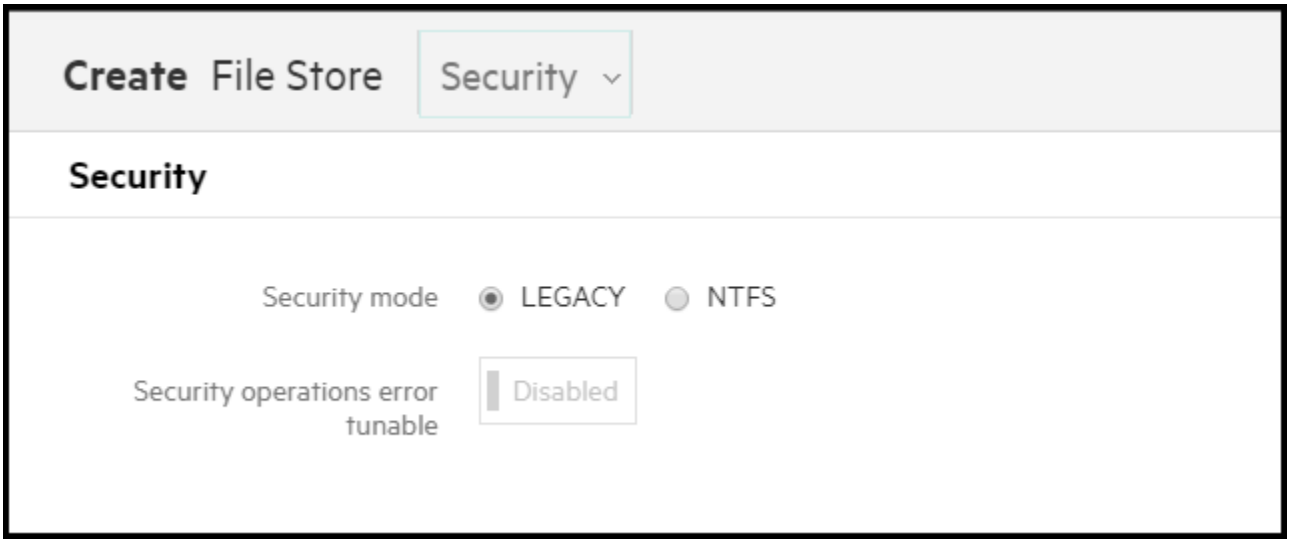

Figure 13. Create File Store – Security

These steps result in the File Store shown below:

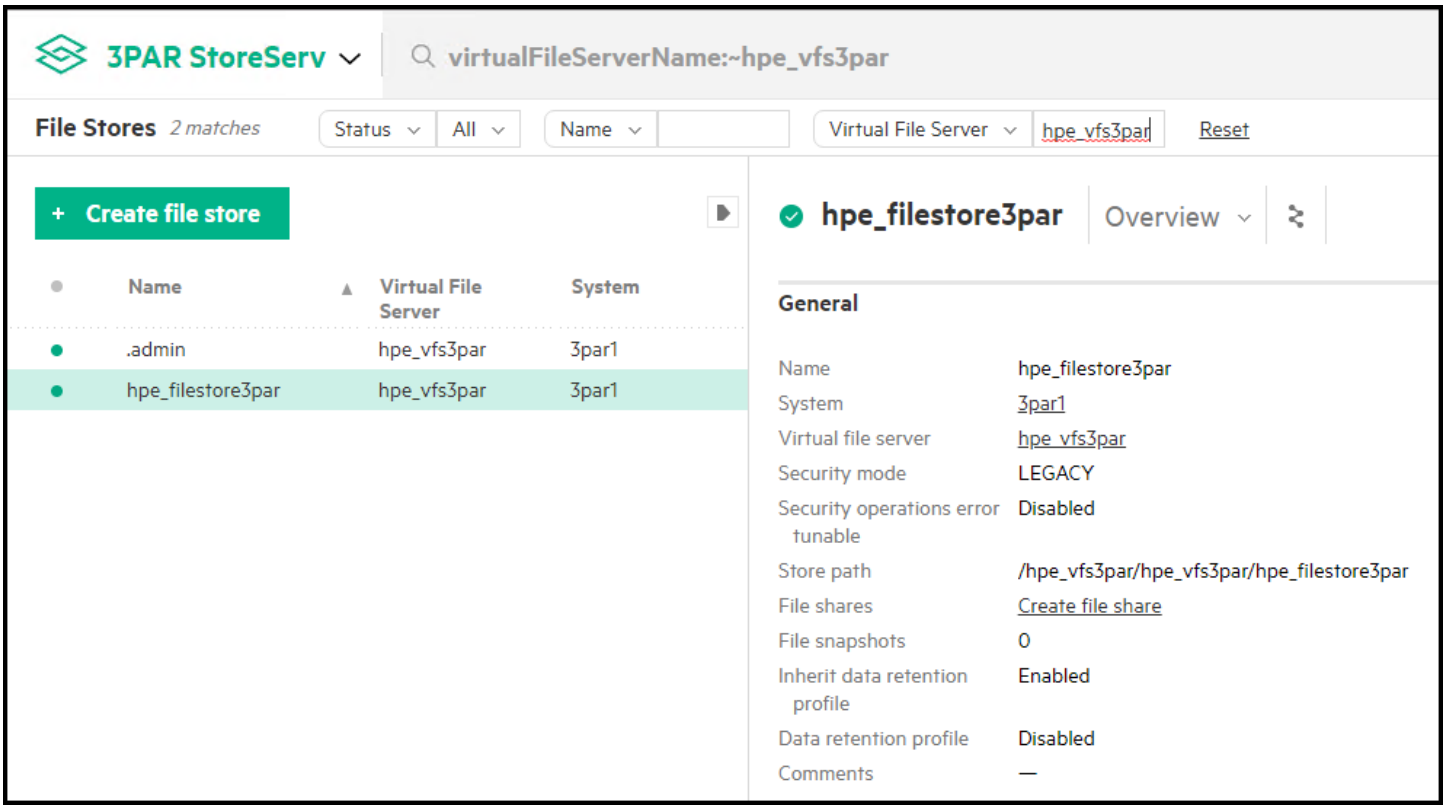

Figure 14. File Store

Create a File Share:

1. In the General section of the Create File Share dialog, set the share type to NFS Share and set a share name, for example, hpe_fileshare3par.

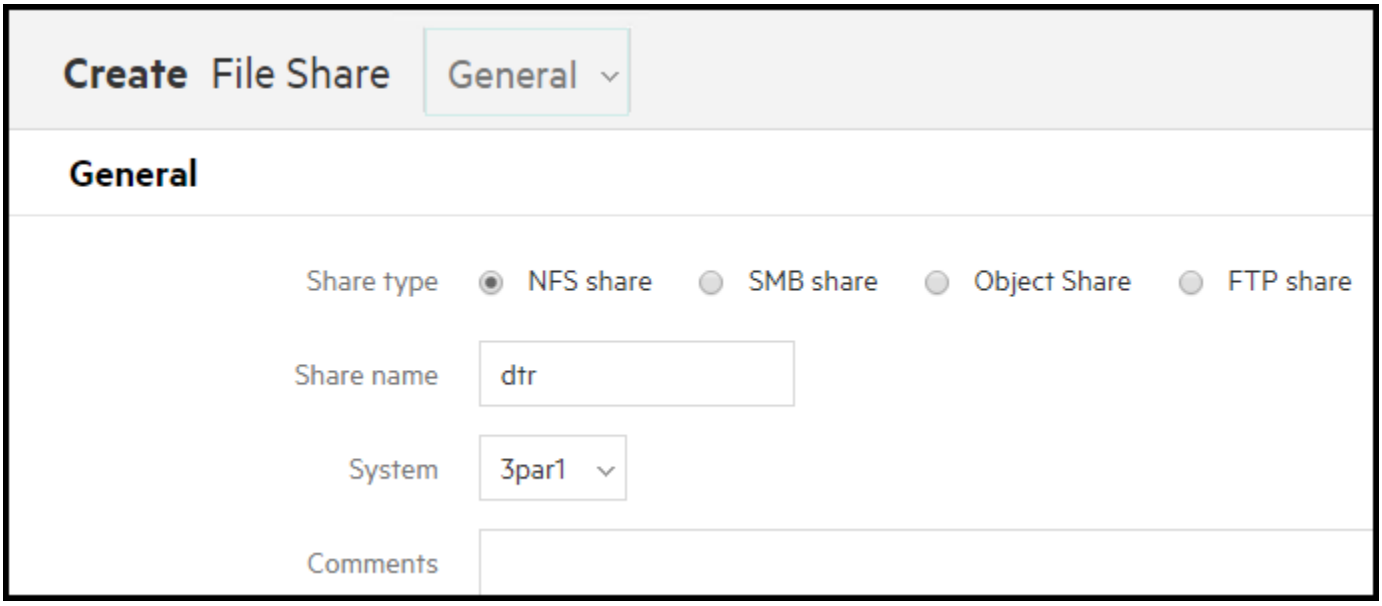

Figure 15. Create File Share – General

2. In the Share Path section, select the virtual file server and file store that you created earlier and set the sub-directory to k8s.

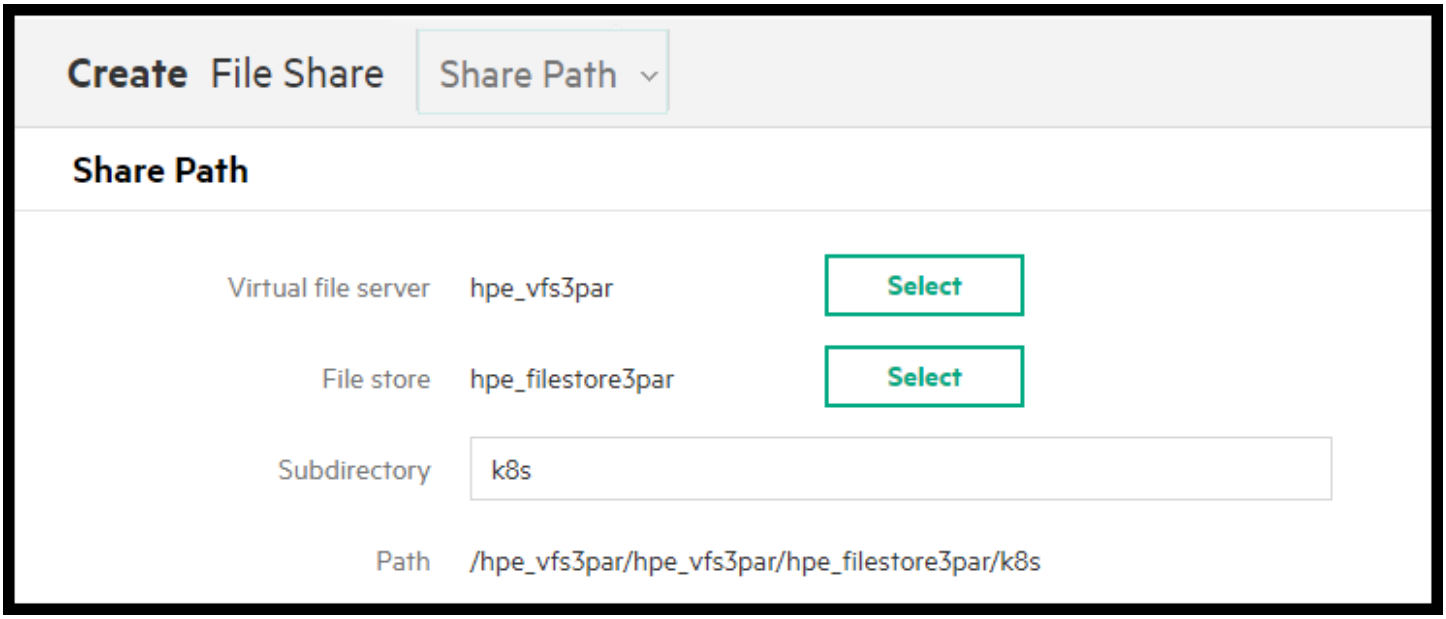

Figure 16. Create File Share - Share Path

3. In the Additional Settings section, set the Permission to Read/Write allowed and the Privilege to root squashing is off (no root squash):

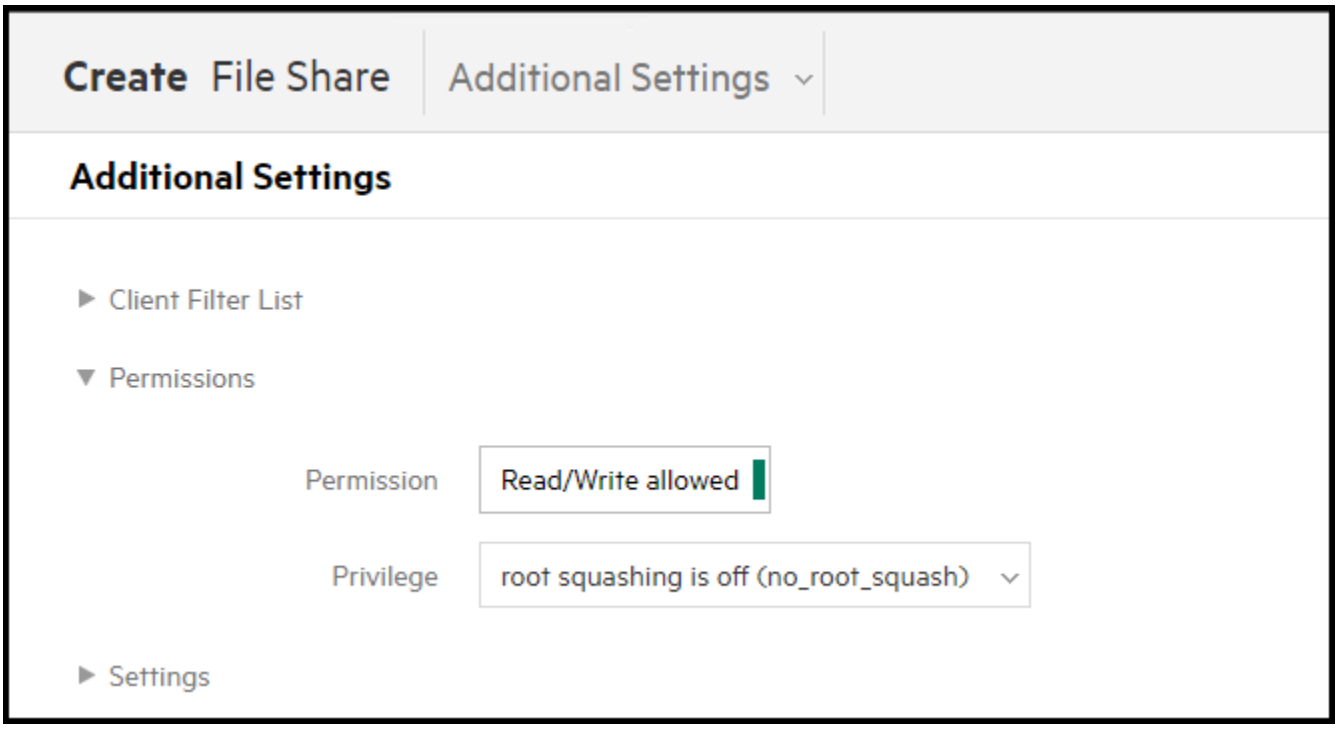

Figure 17. Create File Share - Additional Settings

The overview for the created File Share is shown below and contains the information you need to specify the configuration variables.

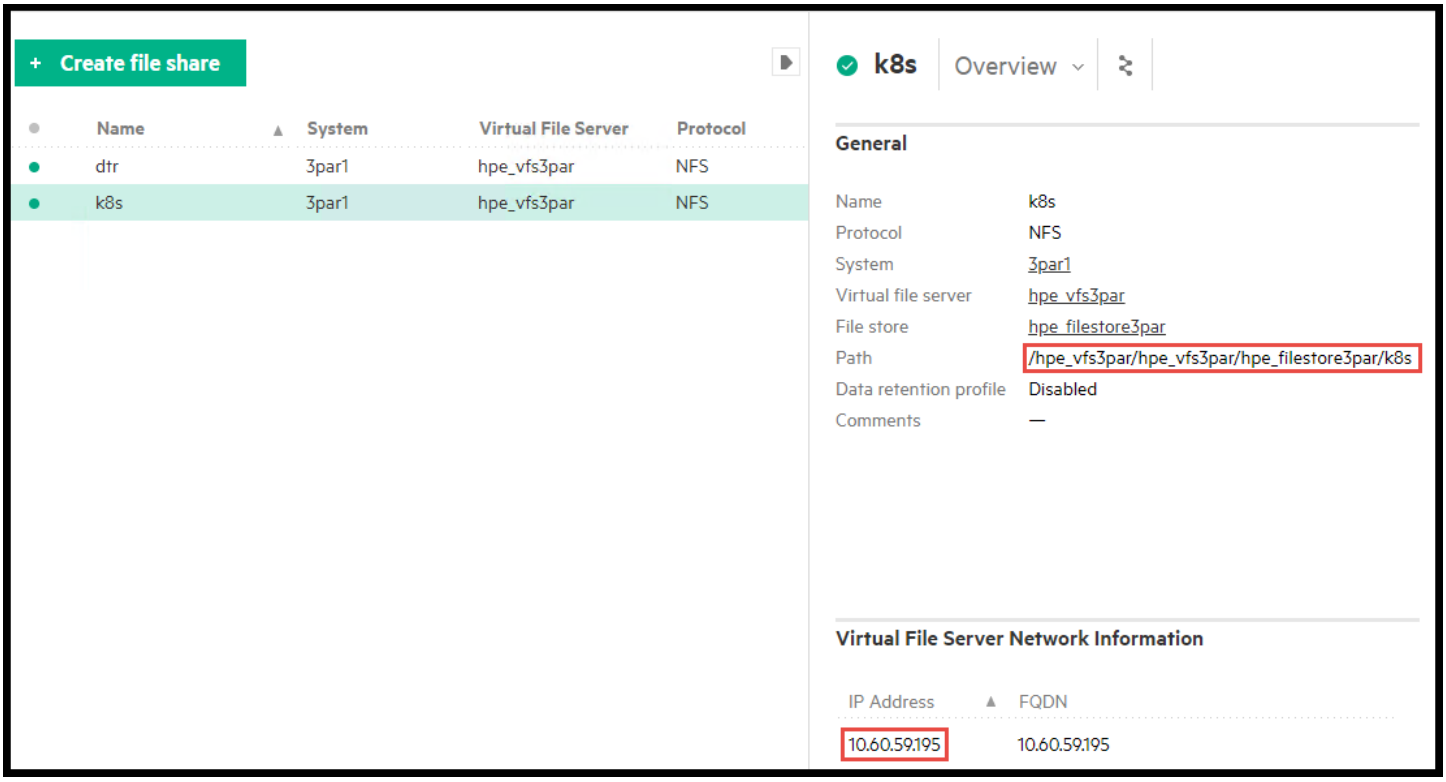

Figure 18. File Share

Configurating NFS on HPE 3PAR for post-deployment verification

In this example, it is assumed that the relevant variables are configured as follows:

Table 15. NFS on HPE 3PAR configuration

Variable	Value
nfs_provisioner_namespace	`nfsstorage`
nfs_provisioner_role	`nfs-provisioner-runner-3par`
nfs_provisioner_serviceaccount	`nfs-provisioner`
nfs_provisioner_name	`hpe.com/nfs-3par`
nfs_provisioner_storage_class_name	`nfs-3par`
nfs_provisioner_server_ip	`hpe_vfs3par.cloudra.local`
nfs_provisioner_server_share	`/hpe_vfs3par/hpe_vfs3par/hpe_filestore3par/k8s`
nfs_mount_options	`'rw,sync,actimeo=0'`

Running the playbook

Once the appropriate configuration has been establised, run the playbook:

```
# cd ~/Docker-Synergy
# ansible-playbook -i hosts playbooks/nfs-provisioner.yml --vault-password-file .vault_pass
```

Running the command `kubectl get sc` will show the storage class named `nfs-3par`:

```
# kubectl get sc

NAME       PROVISIONER        AGE
nfs-3par   hpe.com/nfs-3par   5m
```

The playbook has a built-in test to validate the provisioining. A pod is deployed to write some test content:

`templates/nfs-provisioner/nfs-provisioner-writer-pod.yml.j2`

```
kind: Pod
apiVersion: v1
metadata:
  name: writer-pod
spec:
  containers:
  - name: writer-pod
    image: gcr.io/google_containers/busybox:1.24
    command:
      - "/bin/sh"
    args:
      - "-c"
      - "echo '{{ TestMessage }}' >/mnt/bar.txt && while [ 1 ] ; do sleep 2 ; done "
    volumeMounts:
      - name: nfs-pvc
        mountPath: "/mnt"
  restartPolicy: "Never"
  volumes:
    - name: nfs-pvc
      persistentVolumeClaim:
        claimName: test-claim
```

This pod is then deleted, and a new pod is deployed to check that the test content has been persisted after the writer pod went away.

`templates/nfs-provisioner/nfs-provisioner-reader-pod.yml.j2`

```
kind: Pod
apiVersion: v1
metadata:
  name: reader-pod
spec:
  containers:
  - name: reader-pod
    image: gcr.io/google_containers/busybox:1.24
    command:
      - "/bin/sh"
    args:
      - "-c"
      - "cat /mnt/bar.txt  && while [ 1 ] ; do sleep 1 ; done "
    volumeMounts:
      - name: nfs-pvc
        mountPath: "/mnt"
  restartPolicy: "Never"
  volumes:
    - name: nfs-pvc
```

```
      persistentVolumeClaim:
        claimName: test-claim
```

You should see the following output if the provisioning succeeds:

```
ok: [localhost] => {
    "msg": "Successfully tested NFS persistent storage"
}
```

Using NFS VM when deploying NFS provisioner for Kubernetes

NFS can be provisioned using the NFS VM for proof of concept or demo systems.

Prerequisites

- Configure the variables described in the section Kubernetes Persistent Volume configuration
- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Using NFS VM for post-deployment verification

In this example, it is assumed that the relevant variables are configured as shown in Table 16.

Table 16. NFS provisioner configuration values

Variable	Value
nfs_provisioner_namespace	`nfsstorage`
nfs_provisioner_role	`nfs-provisioner-runner-vm`
nfs_provisioner_serviceaccount	`nfs-provisioner`
nfs_provisioner_name	`hpe.com/nfs-vm`
nfs_provisioner_storage_class_name	`nfs-vm`
nfs_provisioner_server_ip	`hpe-nfs.cloudra.local`
nfs_provisioner_server_share	`/k8s`
nfs_mount_options	`'rw,sync,actimeo=0'`

In this instance, the variable `nfs_external_server` is commented out, resulting in the NFS VM being used, rather than any external server.

Note

When using an external NFS server such as the one hosted by 3PAR, you need to create the file shares manually as shown in the previous section. If you are using the NFS VM, the file share is created automatically when running `site.yml` by the playbook `playbooks/install_nfs_server.yml`. If you wish to change the file share after initial deployment, you must update the variable `nfs_provisioner_server_share` and then re-run the playbook `playbooks/install_nfs_server.yml`.

Running the playbook

Once the prerequisites are satisfied, run the appropriate playbook on your Ansible node.

```
# cd Docker-Synergy
# ansible-playbook -i hosts playbooks/nfs-provisioner.yml --vault-password-file .vault_pass
```

Running the command kubectl get sc will show the storage class named `nfs-vm`:

```
# kubectl get sc

NAME      PROVISIONER   AGE
nfs-vm     hpe.com/nfs-vm     5m
```

The playbook has a built-in test to validate the provisioining. A pod is deployed to write some test content:

The playbook has a built-in test to validate the provisioining. A pod is deployed to write some test content:

`templates/nfs-provisioner/nfs-provisioner-writer-pod.yml.j2`

```
kind: Pod
apiVersion: v1
metadata:
  name: writer-pod
spec:
  containers:
  - name: writer-pod
    image: gcr.io/google_containers/busybox:1.24
    command:
      - "/bin/sh"
    args:
      - "-c"
      - "echo '{{ TestMessage }}' >/mnt/bar.txt && while [ 1 ] ; do sleep 2 ; done "
    volumeMounts:
      - name: nfs-pvc
        mountPath: "/mnt"
  restartPolicy: "Never"
  volumes:
    - name: nfs-pvc
      persistentVolumeClaim:
        claimName: test-claim
```

This pod is then deleted, and a new pod is deployed to check that the test content has been persisted after the writer pod went away.

`templates/nfs-provisioner/nfs-provisioner-reader-pod.yml.j2`

```
kind: Pod
apiVersion: v1
metadata:
  name: reader-pod
spec:
  containers:
  - name: reader-pod
    image: gcr.io/google_containers/busybox:1.24
    command:
      - "/bin/sh"
    args:
      - "-c"
      - "cat /mnt/bar.txt  && while [ 1 ] ; do sleep 1 ; done "
    volumeMounts:
      - name: nfs-pvc
        mountPath: "/mnt"
  restartPolicy: "Never"
```

```
  volumes:
    - name: nfs-pvc
      persistentVolumeClaim:
        claimName: test-claim
```

You should see the following output if the provisioning succeeds:

```
ok: [localhost] => {
    "msg": "Successfully tested NFS persistent storage"
}
```

Validating the NFS provisioner using WordPress and MySQL

A sample playbook has been provided to show how to use the NFS provioner for perstent storage for a WordPress and MySQL deployment.

Prerequisites

- Install the `kubectl` binary on your Ansible box
- Install the UCP Client bundle for the admin user
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`

Deploy the NFS provisioner as outlined in the preceeding section. The article assumes that the NFS configuration is the same as used in that section, as shown in Table 17:

Table 17. NFS provisioner configuration values

Variable	Value
nfs_provisioner_namespace	nfsstorage
nfs_provisioner_role	nfs-provisioner-runner
nfs_provisioner_serviceaccount	nfs-provisioner
nfs_provisioner_name	hpe.com/nfs
nfs_provisioner_storage_class_name	nfs
nfs_provisioner_server_ip	hpe2-nfs.cloudra.local
nfs_provisioner_server_share	/k8s

Running the playbook

The playbook `test/playbooks/wordpress-mysql-nfs.yml` creates Persistent Volume Claims for both Wordpress and MySQL, deploys both applications and makes the WordPress UI available via a NodePort.

```
# cd ~/Docker-Synergy
#  ansible-playbook -i hosts ./test/playbooks/wordpress-mysql-nfs.yml --vault-password-file
.vault_pass
```

The output shows the components created along with the NodePort for the `wordpress` service.

```
ok: [localhost] => {
    "ps.stdout_lines": [
        "Cluster \"ucp_hpe2-ucp01.am2.cloudra.local:6443_admin\" set.",
        "User \"ucp_hpe2-ucp01.am2.cloudra.local:6443_admin\" set.",
        "Context \"ucp_hpe2-ucp01.am2.cloudra.local:6443_admin\" modified.",
        "namespace/wordpress-mysql created",
        "secret/mysql-pass created",
        "persistentvolumeclaim/mysql-pv-claim created",
        "persistentvolumeclaim/wp-pv-claim created",
```

```
    "deployment.apps/wordpress-mysql created",
    "deployment.apps/wordpress created",
    "service/wordpress-mysql created",
    "service/wordpress created",
    "NAME              TYPE        CLUSTER-IP      EXTERNAL-IP   PORT(S)        AGE",
    "wordpress         NodePort    10.96.216.103   <none>        80:33790/TCP   0s",
    "wordpress-mysql   ClusterIP   None            <none>        3306/TCP       0s"
]
```

Browse to the specified port on any node in your cluster.

```
http://hpe2-ucp01.am2.cloudra.local:33790
```

Configuring WordPress

You need to configure the language and password before WordPress is ready to use, as shown in Figure 19.

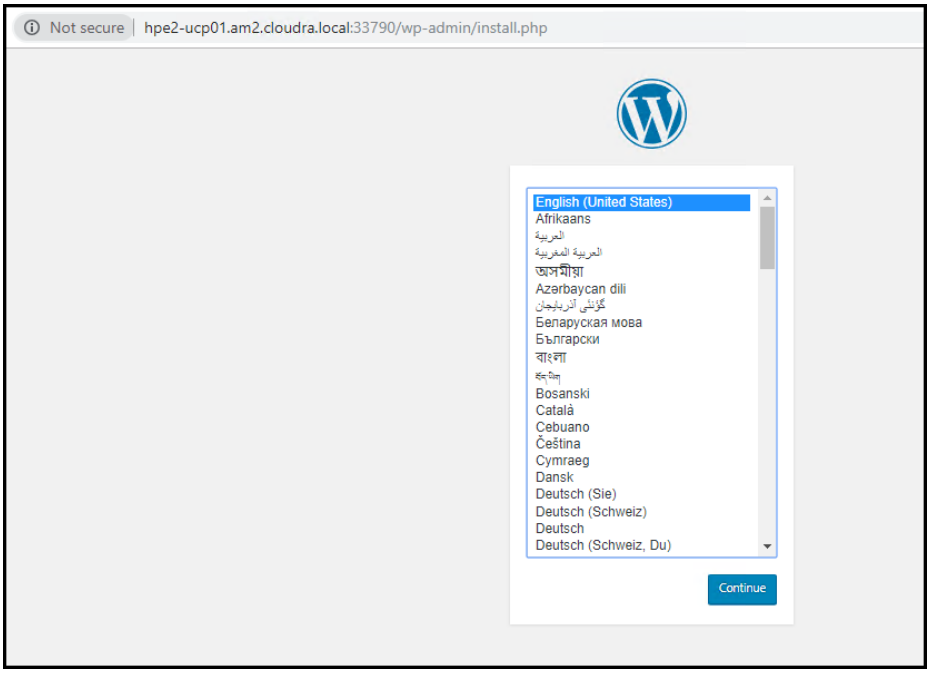

Figure 19. Configure WordPress language

Add a username, password and other configuration details, as shown in Figure 20.

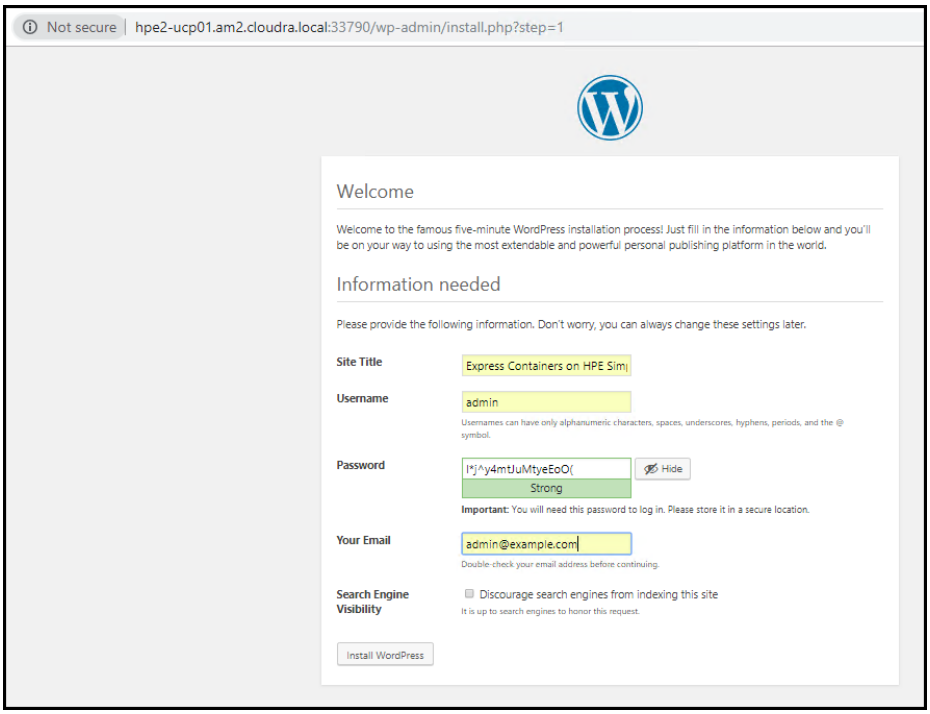

Figure 20. Configure WordPress password

Log in to WordPress, as shown in Figure 21, with the user name and password you have just set up.

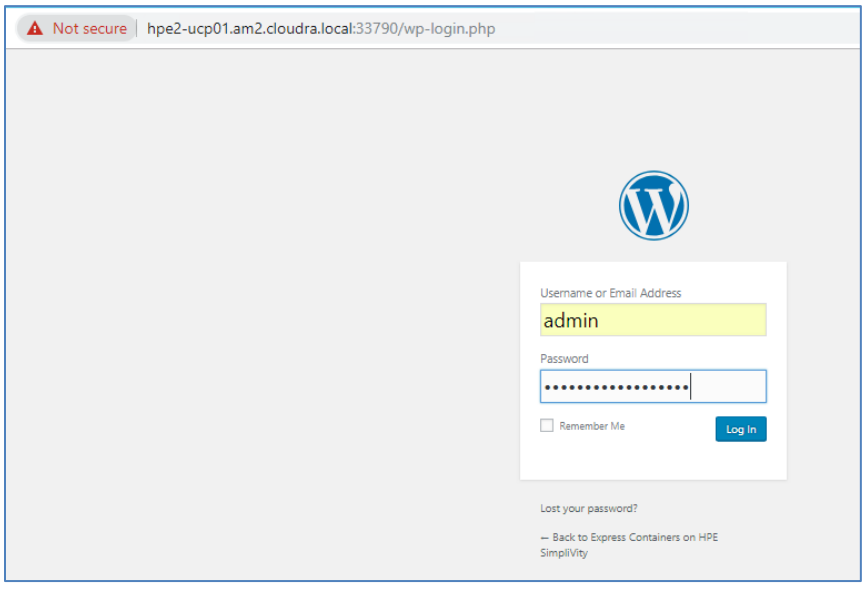

Figure 21. WordPress login

The welcome page is displayed, as shown in Figure 22.

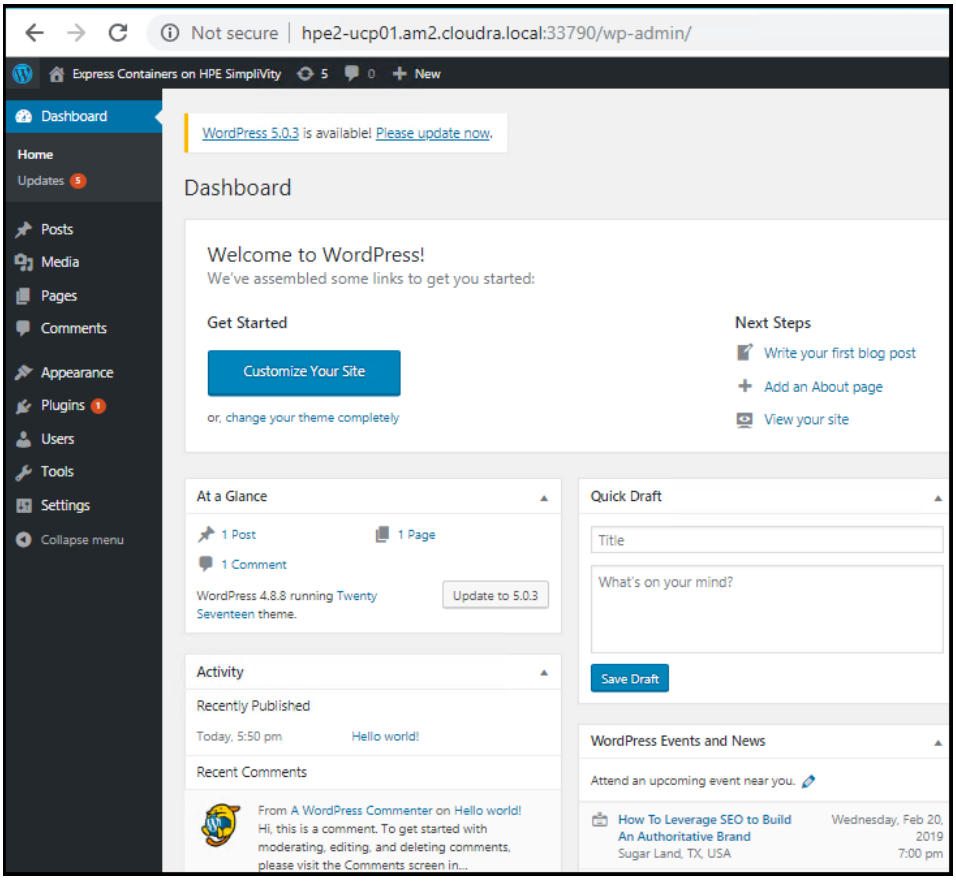

Figure 22. WordPress welcome

Create your first post

Click on `Write your first blog post` and start creating some content. Add a blog title and then click `Add Media` to upload an image to the Media Library and then `Insert` into post. In this example, as shown in Figure 23, the image is a file named `380 with OmniStack.jpg`.

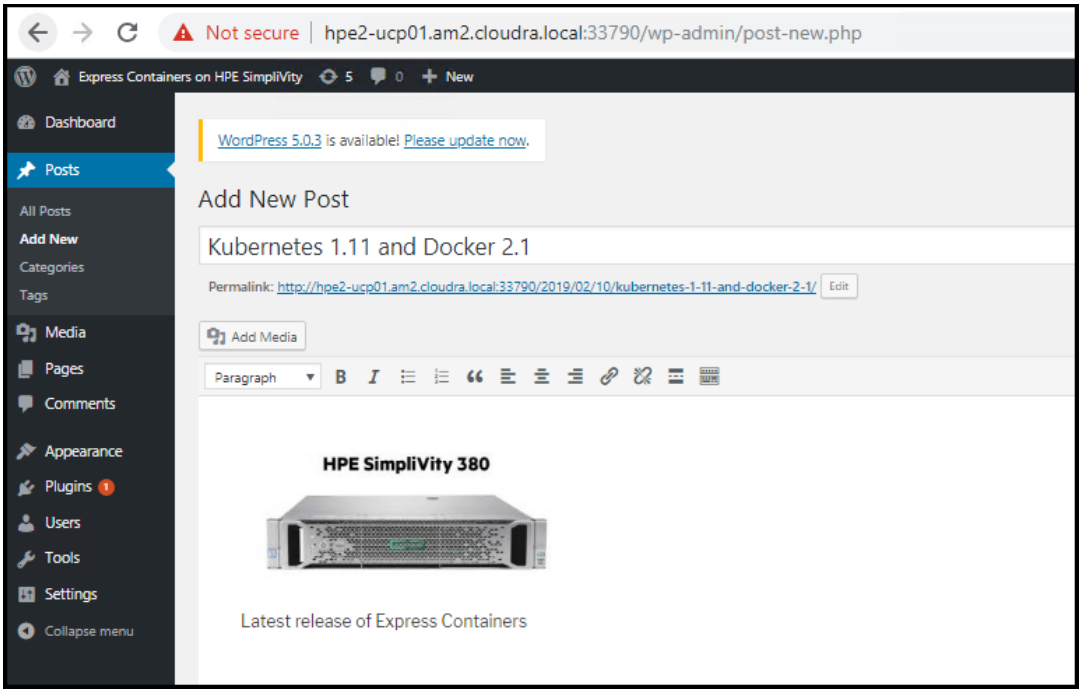

Figure 23. Create your first WordPress blog post

Click `Publish` and then `View post` to see your new blog post, as shown in Figure 24.

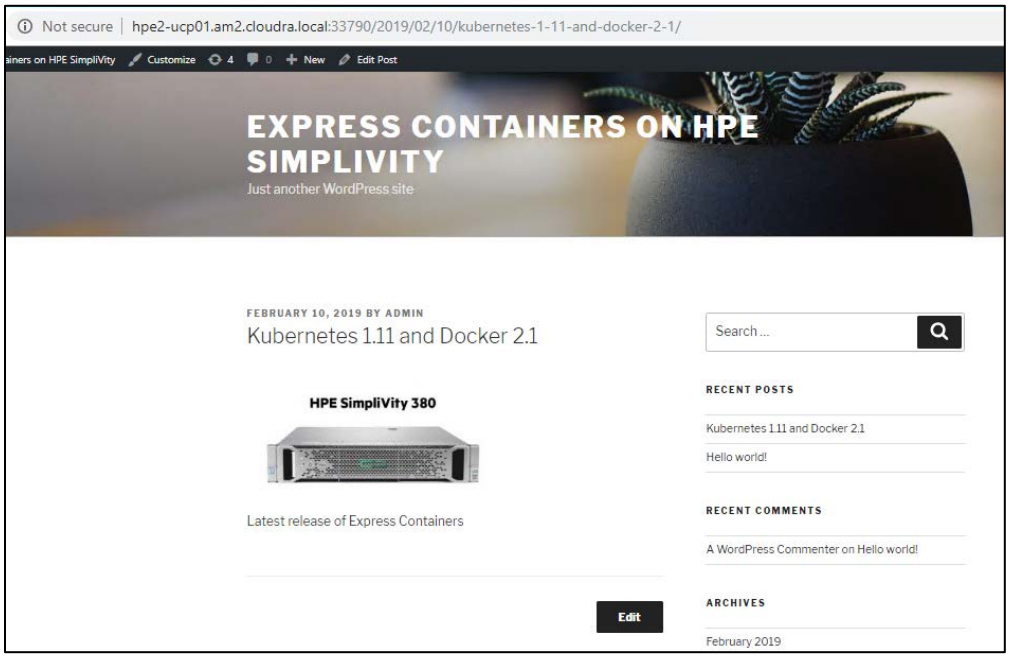

Figure 24. View your first post

Test persistence for WordPress

Find your WordPress Persistent Volume Claim (PVC).

```
# kubectl -n wordpress-mysql get pvc
NAME             STATUS    VOLUME                                     CAPACITY   ACCESS   STORAGECLASS
AGE
mysql-pv-claim   Bound     pvc-d48880e3-2d58-11e9-adb2-0242ac110003   1Gi        RWO            nfs
1h
wp-pv-claim      Bound     pvc-d4bc101f-2d58-11e9-adb2-0242ac110003   20Gi       RWO            nfs
1h
```

Connect to the NFS VM and browse the `/k8s` folder to find the volume for the WordPress claim `wp-pv-claim`.

```
# ssh hpe2-nfs ls /k8s
wordpress-mysql-mysql-pv-claim-pvc-d48880e3-2d58-11e9-adb2-0242ac110003
wordpress-mysql-wp-pv-claim-pvc-d4bc101f-2d58-11e9-adb2-0242ac110003
```

Locate the `wp-content` folder.

```
# ssh hpe2-nfs ls /k8s/wordpress-mysql-wp-pv-claim-pvc-d4bc101f-2d58-11e9-adb2-0242ac110003
index.php
license.txt
readme.html
wp-activate.php
wp-admin
wp-blog-header.php
wp-comments-post.php
wp-config.php
wp-config-sample.php
wp-content
wp-cron.php
wp-includes
wp-links-opml.php
wp-load.php
wp-login.php
wp-mail.php
wp-settings.php
wp-signup.php
wp-trackback.php
xmlrpc.php
```

Now find the image used in the blog post.

```
# ssh hpe2-nfs ls /k8s/wordpress-mysql-wp-pv-claim-pvc-d4bc101f-2d58-11e9-adb2-0242ac110003/wp-
content/uploads/2019/02
380-with-OmniStack-100x100.jpg
380-with-OmniStack-150x150.jpg
380-with-OmniStack-300x150.jpg
380-with-OmniStack-768x384.jpg
380-with-OmniStack.jpg
```

Note that WordPress has created a number of variations of the original image, for different screen sizes. Shutdown wordpress (leave MySQL running for now).

```
# kubectl -n wordpress-mysql delete -f /tmp/wordpress-mysql-nfs/wordpress-deployment.yml
deployment.apps "wordpress" deleted
```

Refresh the page in the browser to confirm that WordPress is indeed inaccessible.

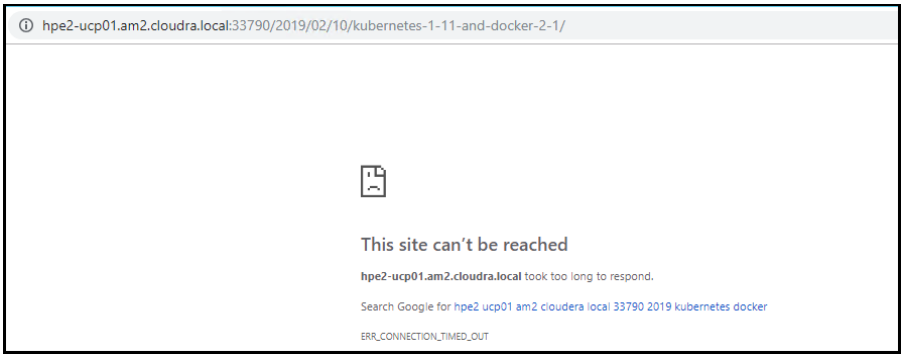

Figure 25. Cannot connect to WordPress

Now redeploy Wordpress

```
# kubectl -n wordpress-mysql apply -f /tmp/wordpress-mysql-nfs/wordpress-deployment.yml
deployment.apps/wordpress created
```

Refresh the page in the browser to confirm that WordPress is now accessible and that the image in the blog post has survived the shutdown, as shown in Figure 26.

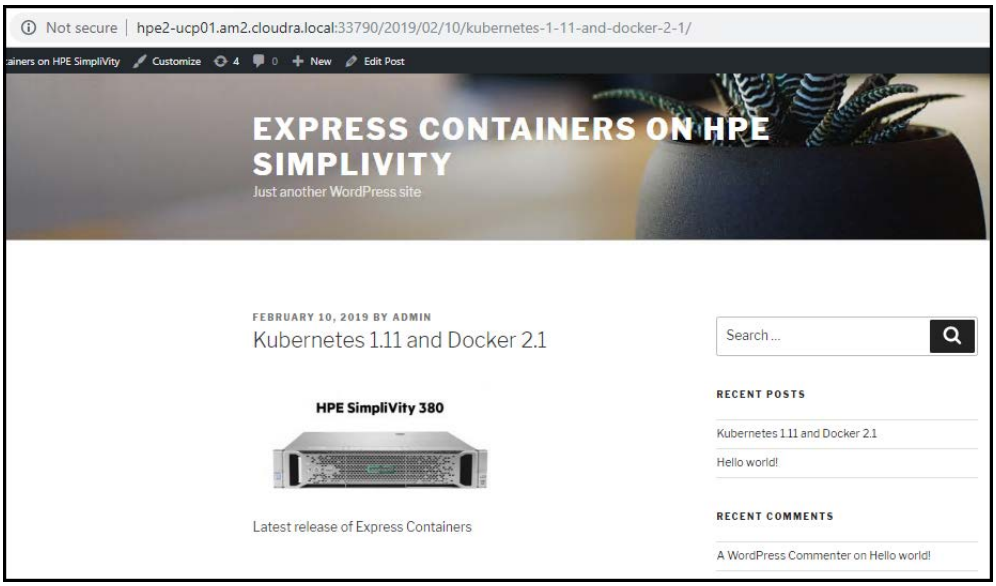

Figure 26. View restored post

Test persistence in MySQL

A similar procedure can be performed for MySQL. While assets such as images, CSS files, etc are stored in the WordPress volume, information about users, posts, comments, tags, etc are stored in the MySQL database. It is possible to browse the tables in the database and identify the rows related to the blog post you created.

Shut down MySQL as follows:

```
# kubectl -n wordpress-mysql delete -f /tmp/wordpress-mysql-nfs/mysql-deployment.yml

deployment.apps "wordpress-mysql" deleted
```

Refresh the page for your blog post, and you will see that WordPress can no longer connect to the database, as shown in Figure 27.

Figure 27. Cannot connect to MySQL

Restore the MySQL deployment:

```
# kubectl -n wordpress-mysql apply -f /tmp/wordpress-mysql-nfs/mysql-deployment.yml
deployment.apps/wordpress-mysql created
```

Refresh the page in the browser, as shown in Figure 28, to confirm that WordPress can now access the database and that the blog post has survived the database shutdown.

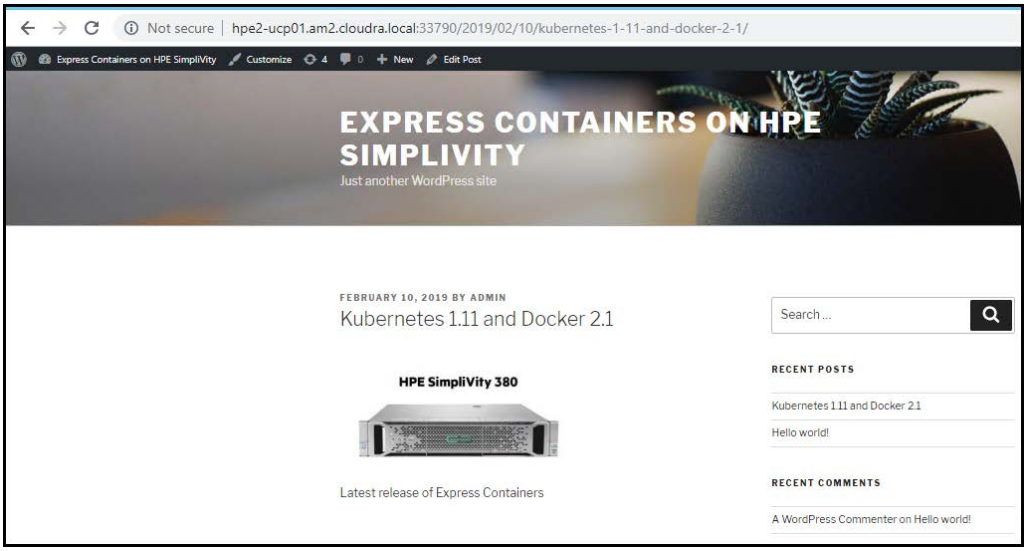

Figure 28. Check after MySQL restored

Deploying Windows workers

The `site.yml` playbook will automatically deploy any Windows workers declared in the inventory. The playbooks should run for approximately 70-80 minutes with 3 Windows workers added to the default deployment (depending on your server specifications and the size of your environment). The increase in running time is primarily due to the need to update Windows after creating the VMs.

This section describes the functionality and configuration of the Windows-specific playbooks. It also details how to create the initial Windows template and how to manage deploying Windows worker nodes behind a proxy.

Create the Windows Template

To create the Windows Template that you will use as the base for all your Windows VM worker nodes, you will first create a Virtual Machine with the OS installed and then convert the Virtual Machine to a VM Template. The VM Template is created as lean as possible, with any additional software installs and/or system configuration performed subsequently using Ansible.

As the creation of the template is a one-off task, this procedure has not been automated. The steps to create a VM template manually are outlined below.

Log in to vCenter and create a new Virtual Machine with the following characteristics:

- Guest OS Family: Windows, Guest OS Version: Microsoft Windows Server 2016 (64-bit)
- Hard Disk size: 100GB (Thin provisioning), 1 vCPU and 4 GB of RAM. Both vCPU and memory can be altered later after you deploy from this template.
- A single network controller connected to the network or VLAN of your choice. All VMs will connect to this same network.
- Change the network type to `VMXNET3`, and attach the Windows Server 2016 ISO image from a datastore ensuring you connect the CD/DVD drive on boot.
- Click on the `VM Options` tab, and in the `Boot Options` section, select `Force BIOS setup[*]` to ensure that the machine enters the BIOS setup screen on next boot of this VM. This will allow you to adjust the boot order, placing the virtual CDROM in front of your hard drive.
- Optionally you can remove the floppy drive.

Install Windows Server 2016:

1. Power on the selected VM and then Open Console. Once connected to the console, you will be placed in the BIOS setup screen.
1. Select the Boot tab, click on CD-ROM Drive and move up the CDROM drive above the hard drive. This allows your Windows Server 2016 ISO image to be loaded first on boot. F10 Save and exit is next step.
2. Enter your choices for Language, Time/Currency Format, Keyboard and then Install Now.
3. Select the OS you want to install, and then select Custom: Install Windows Only.
4. Select drive 0, the 100 GB drive you specified earlier, as the location for installing Windows.
5. Add a password for the Administrator user.
6. Install VMware Tools and reboot.
7. Once the VM has re-booted, add a temporary network IP address.
8. Use the `sconfig` utility from (MS-DOS) command line to install Windows updates and enable remote desktop.
9. Perform any other customizations you require at this point.
10. Prior to converting the VM to Template, run `Sysprep: C:\Windows\System32\Sysprep\Sysprep.exe`
11. Ensure 'System Out-of-Box Experience (OOBE)' is selected.
12. Select the 'Generalize' option.
13. Select 'Shutdown' from the Shutdown Options.
14. Shutdown VM, and untick `Connect CD/DVD` so that the Windows Server 2016 ISO is no longer mounted.
15. Boot the Windows VM one final time and enter regional settings applicable to your location and keyboard mapping, then enter a password and Shutdown VM.

Note
The `vmware_guest` module used by the playbooks will generate a new SID.

Turn the VM into a template by right-clicking on your VM and selecting `Template -> Convert to Template`. This will create a new template visible under VM Templates in Folders, ready for future use.

Playbooks for adding Windows workers

- `playbooks/provision_nodes.yml` will create all the necessary Windows 2016 VMs for the environment based on the Windows VM Template defined in the `win_vm_template` variable. Windows VM workers nodes are defined in the group `vm_wrk_win` in the `hosts` inventory.

- `playbooks/install_docker.yml` installs Docker along with all its dependencies on your Windows VMs
- `playbooks/scale_workers.yml` installs and configures additional Windows workers on the target nodes defined by the group `vm_wrk_win` in the `hosts` inventory.
- `playbooks/splunk_uf_win.yml` installs and configures the Splunk Universal Forwarder on each Windows machine in the inventory.

Windows configuration

Window-related variables are shown in Table 18. Variables for all Windows nodes (VM and bare metal) are in the file `group_vars/windows_box.yml`. Windows VM-specific variables are in `group_vars/vm_wrk_win.yml` while Windows bare metal variables are in `group_vars/bm_wrk_win.yml`

Table 18. Windows variables

Variable	File	Description
win_username	**group_vars/windows_box.yml**	Windows user name. The default is `Administrator`
win_password	**group_vars/all/vault**	The password for the Windows account.
docker_ee_version_windows	**group_vars/windows_box.yml**	It is important that the version of the Docker engine running on your Windows worker nodes is the same as that running on RHEL in the rest of your cluster. You should use this variable to explicitly match up the versions. For Docker 2.1, the recommended value is '`18.09`'. If you do not explicitly set this value, you may end up with an incompatible newer version running on your Windows workers.
windows_update	**group_vars/windows_box.yml**	Variable used to determine if Windows updates are automatically downloaded when installing Docker on Windows worker nodes (in the `playbooks/install_docker.yml`). Defaults to `true`.
windows_docker_drive	**group_vars/windows_box.yml**	'D'
windows_docker_directory	**group_vars/windows_box.yml**	'D:\\DockerData'
windows_docker_volume_label	**group_vars/windows_box.yml**	'DockerVol'
windows_tz	**group_vars/windows_box.yml**	'Pacific Standard Time' This is different from the `windows_timezone` variable. It is important that this value matches the timezone used by UCP servers for certificate validation. See https://msdn.microsoft.com/en-us/library/ms912391.aspx.
windows_winrm_script	**group_vars/windows_box.yml**	Variable used to determine where the `winrm` Powershell script will be downloaded from. See the following section for more information.

Configuring the winrm remoting script

The playbooks for deploying Windows workers rely on a Powershell script for remote access from the Ansible machine. The script `ConfigureRemotingForAnsible.ps1` is available online on GitHub at https://raw.githubusercontent.com/ansible/ansible/devel/examples/scripts/ConfigureRemotingForAnsible.ps1.

You need to make this script available locally on the Fedora 29 Ansible controller:

1. Download the script:

```
wget
https://raw.githubusercontent.com/ansible/ansible/devel/examples/scripts/ConfigureRemotingForAnsible.ps1
```

2. Deploy a local HTTP server, enabling port 80, for example:

```
firewall-cmd --permanent --add-port 80/tcp --zone=public
firewall-cmd --permanent --change-interface=ens192 --zone=public
firewall-cmd -reload

dnf install httpd
systemctl enable httpd
systemctl start httpd
```

3. Copy the downloaded script to the web server:

```
cp ConfigureRemotingForAnsible.ps1 /var/www/html
```

4. Configure the variable to point at the local web server, using the name or the IP address of the Ansible controller, for example:

```
windows_winrm_script: 'http://10.60.59.230/ConfigureRemotingForAnsible.ps1'
```

Windows VM variables

The following table shows the variables specific to Windows VMs.

Table 19. Windows VM variables

Variable	File	Description
win_vm_template	**group_vars/vm_wrk_win.yml**	Name of the Windows 2016 VM Template to use. Note that this is the name from a vCenter perspective, not the hostname.
windows_vdvs_ps	**group_vars/vm_wrk_win.yml**	Variable used to download the PowerShell script that is used to install vDVS for Windows. For example, https://raw.githubusercontent.com/vmware/vsphere-storage-for-docker/master/install-vdvs.ps1
windows_vdvs_path	**group_vars/vm_wrk_win.yml**	Variable used to download vSphere Docker Volume Service software. This variable is combined with `windows_vdvs_version` (below) to generate a URL of the form <windows_vdvs_path>_<windows_vdvs_version>.zip to download the software. For example, to download version 0.21, set `windows_vdvs_path` equal to `https://vmware.bintray.com/vDVS/vsphere-storage-for-docker_windows` and `windows_vdvs_version` equal to `0.21`
windows_vdvs_version	**group_vars/vm_wrk_win.yml**	Combined with `windows_vdvs_path`, this variable is used to generate the URL for downloading the software.
windows_vdvs_directory	**group_vars/vm_wrk_win.yml**	Variable used to determine where vDVS software will be unzipped and installed from. The default is `C:\Users\Administrator\Downloads`
windows_timezone	**group_vars/vm_wrk_win.yml**	Defaults to 15. Valid values are available at https://msdn.microsoft.com/en-us/library/ms912391.aspx

group_vars/win_worker.yml

In general, it should not be necessary to modify the following advanced variables, but they are documented in Table 20 for the sake of completeness.

Table 20. Advanced windows variables

Variable	File	Description
ansible_user	**group_vars/windows_box.yml**	Defaults to the Windows user account `win_username` as specified in `group_vars/all/vars`
ansible_password	**group_vars/windows_box.yml**	Defaults to the Windows user password `win_password` as specified in `group_vars/all/vault`
ansible_port	**group_vars/windows_box.yml**	5986

ansible_connection	**group_vars/windows_box .yml**	winrm
ansible_winrm_server_cert_validation	**group_vars/windows_box .yml**	Defaults to `ignore`
ansible_winrm_operation_timeout_sec	**group_vars/windows_box .yml**	Defaults to `250`
ansible_winrm_read_timeout_sec	**group_vars/windows_box .yml**	Defaults to `300`
windows_timezone	**group_vars/windows_box .yml**	Defaults to `15`. Valid values are available at https://msdn.microsoft.com/en-us/library/ms912391.aspx

Windows operating system and Docker EE

Docker Enterprise Edition for Windows Server (Docker EE) enables native Docker containers on Windows Server. This solution has been tested with Windows worker nodes running Windows Server 2016 and with Docker EE 18.09. More recent versions of Windows Server may work but have not been tested.

Note

Docker Universal Control Plane is not currently supported on Windows Server 1709 due to image incompatibility issues. For more information, see the Docker documentation Install Docker Enterprise Edition for Windows Server.

This solution recommends that you only run Windows Server 2016 on your Windows worker nodes and that you install any required updates to your Windows nodes in a timely manner.

For information on how to update Docker EE on Windows Server 2016, see the Docker documentation Update Docker EE.

Deploying bare metal workers

Introduction to bare metal workers

This solution leverages HPE Synergy OneView 4.10 and HPE Image Streamer 4.10 to provision bare metal servers with an operating system so they can be added to a Docker/Kubernetes cluster as worker nodes. Before you can provision servers using the playbooks, you need to create one or more Image Streamer Operating System Deployment Plans (OSDP) and one or more OneView Server Profile Templates (SPT).

HPE OneView Server Profile Templates are used to create the OneView Server Profiles (SP) that are applied to the Synergy compute modules, also known as bare metal servers. Each bare metal server listed in the Ansible inventory maps to exactly one OneView Server Profile Template. Depending on the environment, you may need to create one or more SPTs depending on the type of servers available in your Synergy environment. In the simplest case, where all servers are of the same Server Hardware Type and there is a single enclosure group, a single SPT can be used. If, on the other hand, the pool of compute modules consists of different server types (for example Gen9 and Gen10 compute modules), then a separate SPT must be created for each Server Hardware Type. When creating the SPT, an OSDP is specified. In most cases, the same OSDP can be used for all compute modules running the same operating system. If you want to deploy both Windows and Linux worker nodes in the same cluster, you need to create a minimum of two SPTs and two OSDPs. One SPT will specify an OSDP that deploys Linux, while a separate SPT will specify a different OSDP that deploys Windows.

Image Streamer Operating System Deployment Plans leverage Operating System Build Plans (OSBP), each of which contains one or more Plan Scripts that are used to configure the deployed Operating System. Each Plan Script may expose one or more OS custom attributes. Custom attributes are parameters that can either be hard-coded to specific values or exposed to the deployment plan and configured by the SPT using the deployment plan. Custom attributes can hold various data types such as IP addresses, host names, product keys etc. The OSDP also specifies a golden image, which will be used when deploying the OS on the server.

When it comes to the provisioning of bare metal servers, the Ansible playbooks create Server Profiles (SP) based on specified SPT and assign the server profiles to physical compute modules in the Synergy enclosures. The provisioning of the operating system is done when the server profile is applied using the Image Streamer OSDP specified in the SPT. Once the servers are provisioned, they are powered on by the playbooks.

Playbooks and configuration

The following table shows the basic variables needed for OneView configuration.

Table 21. HPE OneView variables

Variable	File	Description
oneview_config_hostname	**group_vars/all/vars**	The server hosting HPE OneView
oneview_config_username	**group_vars/all/vars**	HPE OneView user name. Defaults to `Administrator`
oneview_config_password	**group_vars/all/vault**	HPE OneView password.
oneview_config_api_version	**group_vars/all/vars**	HPE OneView API version. Defaults to 600

When it comes to the provisioning of bare-metal servers, the Ansible playbooks create Server Profiles (SP) based on specified Server Profile Templates (SPT) and assign the server profiles to physical compute modules in the Synergy enclosures. The provisioning of the operating system is done when the server profile is applied using the Image Streamer OSDP specified in the SPT. Once the servers are provisioned, they are powered on by the playbooks.

The playbook responsible for the provisioning of the bare metal servers uses the following information stored in Ansible variables for each worker node:

- **ov_template:** The name of the SPT to use when creating the SP for this compute module
- **ov_ansible_connection_name** and **ov_ansible_redundant_connection_name:** The names of the network connections in the server profile template that maps to the network where the Ansible controller node resides. Currently redundant connections are supported so you must specify two connections on the Ansible network/VLAN
- **enclosure** and **bay:** The target compute module to provision, specified by the name of the Synergy enclosure where the compute module resides and the bay number of the compute module

Below is an excerpt of a sample inventory file. The enclosure and bay number is specified for each bare-metal server. Because this particular HPE Synergy environment contains compute modules of different hardware types, each worker node entry also specifies the HPE OneView Server Profile Template to use when deploying the OS.

In this example, both Gen9 and Gen10 compute modules are used and Linux and Windows worker nodes are being deployed.

```
[bm_wrk_lnx]
clh-worker04 ip_addr='10.60.59.25/16' enclosure='Rack1-Mid-CN759000BZ' bay=8
ov_template='RedHat760_fcoe_gen9_4_v1.0.3'
clh-worker06 ip_addr='10.60.59.27/16' enclosure='Rack1-Top-CN7515048P' bay=5
ov_template='RedHat760_fcoe_gen9_3_v1.0.3'

[bm_wrk_win]
clh-worker05 ip_addr='10.60.59.26/16'  enclosure='Rack1-Top-CN7515048P' bay=2
ov_template='Win2016_fcoe_gen10_3_v1.0.3'
```

Note the difference in the Linux template names for the separate Server Hardware Types of **Gen 9 4** and **Gen 9 3**. This can be seen in the following figure taken from HPE OneView:

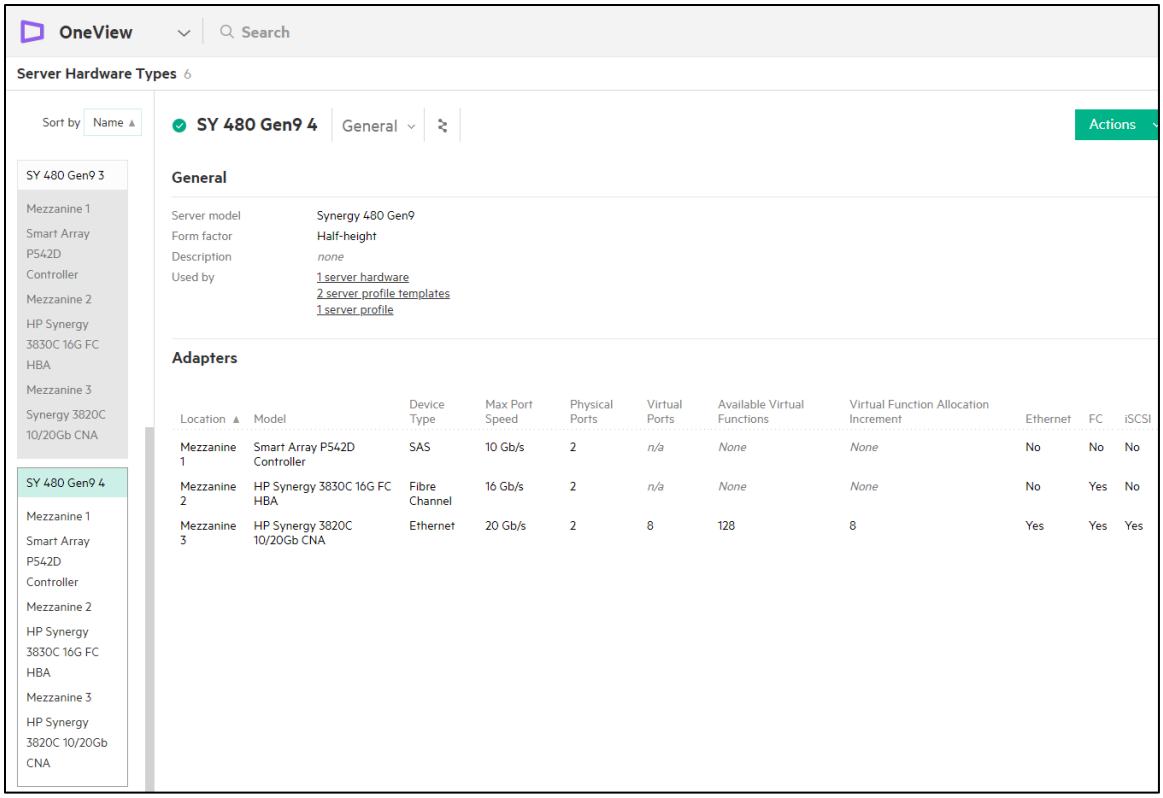

Figure 29. HPE OneView Server Hardware Types

Common variables for all Windows nodes (VM and bare metal) are specified in the file `group_vars/windows_box.yml`. Windows VM-specific variables are in `group_vars/vm_wrk_win.yml` while Windows bare metal variables are in `group_vars/bm_wrk_win.yml`

OS Deployment Plan Custom Attributes

RHEL OS Deployment Plan

Currently, the code responsible for the provisioning of server profiles expects the OS Deployment Plans to expose two and only two custom attributes named 'NIC1' and 'NIC2'. This means the server profiles templates using the OSBP will only see the NIC1 and NIC2 attributes as shown in the figure below, which illustrates the OS Deployment section of the `RedHat760_fcoe_gen9_4_v1.0.3` server profile template.

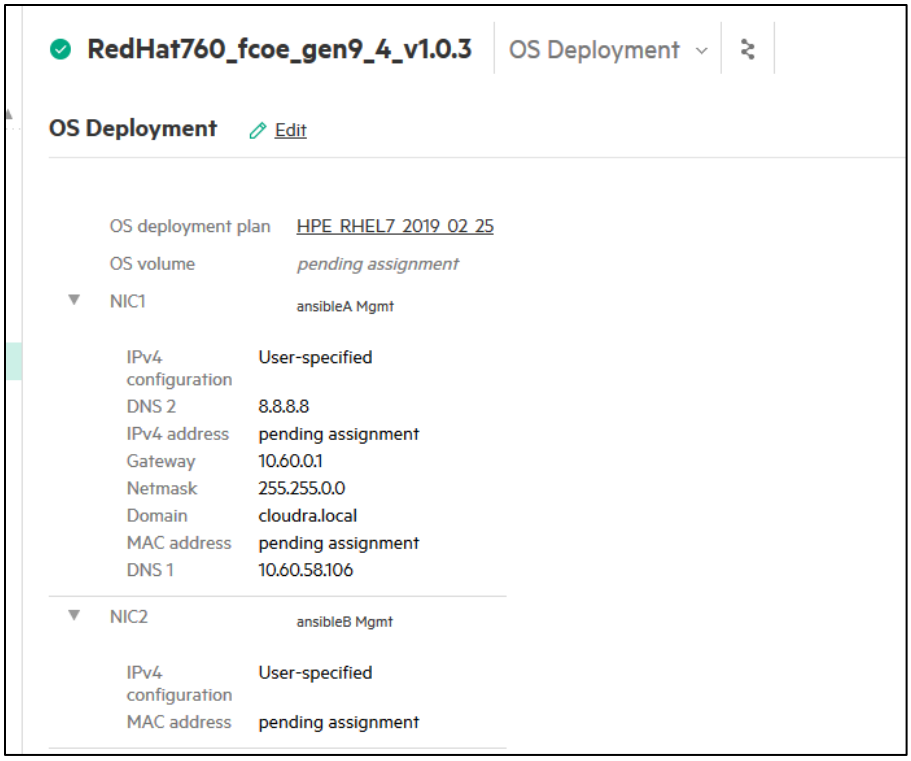

Figure 30. Server Profile Template - OS Deployment

The IPV4 configuration should be configured using "User-specified" because the playbooks will assign the IP addresses from the data in the `hosts` inventory file. All other attributes are populated automatically.

It is possible to specify additional custom attributes in the OS Deployment Plan and the underlying OS Build Plan but these attributes should be hard-coded to the desired values and should not be made visible on deployment.

For example, the Red Hat OS Deployment Plan includes four custom attributes used by the underlying OS Build Plan that are not exposed by the OS Deployment Plan.

- **NewRootPassword:** This attribute is used to configure the password for the root account.
- **NewUser** and **NewUserPassword:** These two custom attributes are used to configure an additional user.
- **ssh:** The underlying OS Build Plan specifies that SSH is enabled since this is required for Ansible to work.

Again, these non-visible custom attributes are all hard-coded to specific values in the OS Build Plan, which effectively means any compute module deployed using this OS Deployment Plan will have these custom attributes set to these hard-coded values.

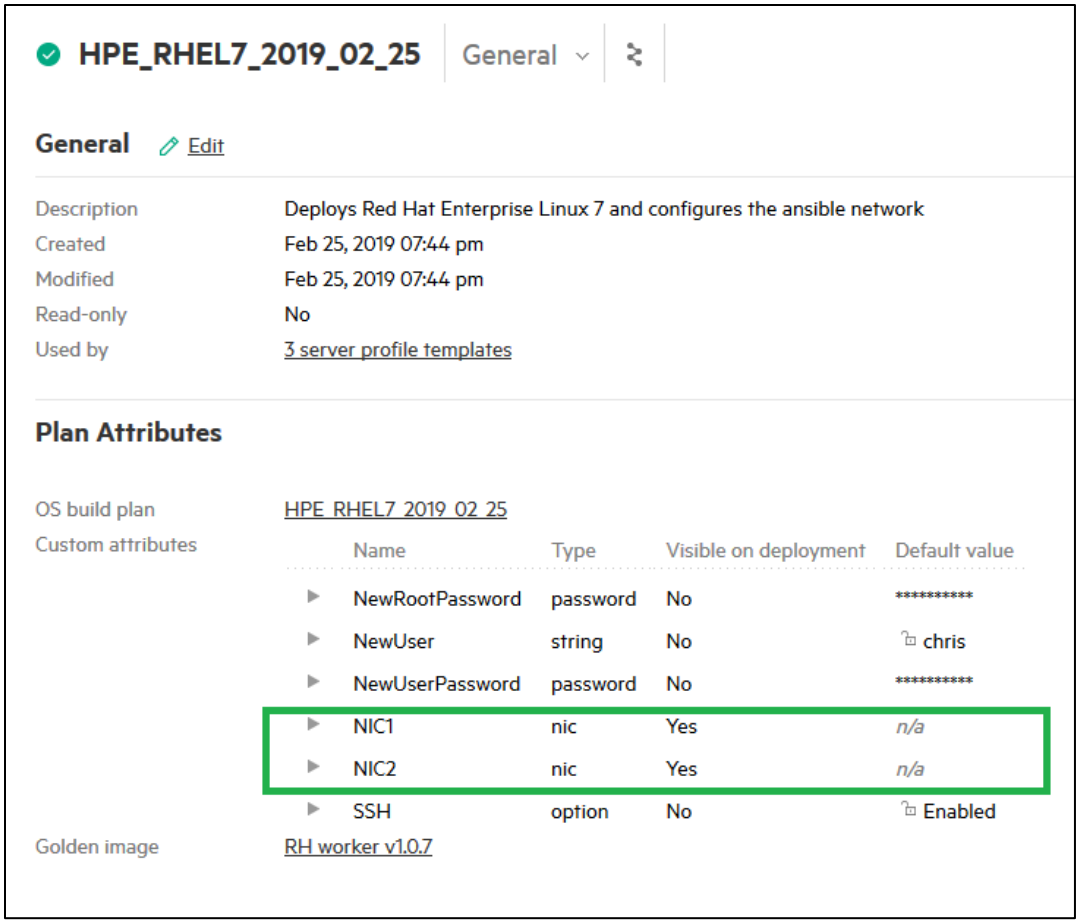

Figure 31. Deployment Plan attributes

Windows 2016 OS Deployment Plan

The following figure shows the Windows 2016 OS Deployment Plan shipping with this solution where only the NIC1 and NIC2 attributes are exposed but additional custom attributes are present and used to configure the Windows OS during deployment. Among other things, the password for the administrative user, the desired Power Plan, Remote Desktop settings, and the Windows Product Key are specified using custom attributes.

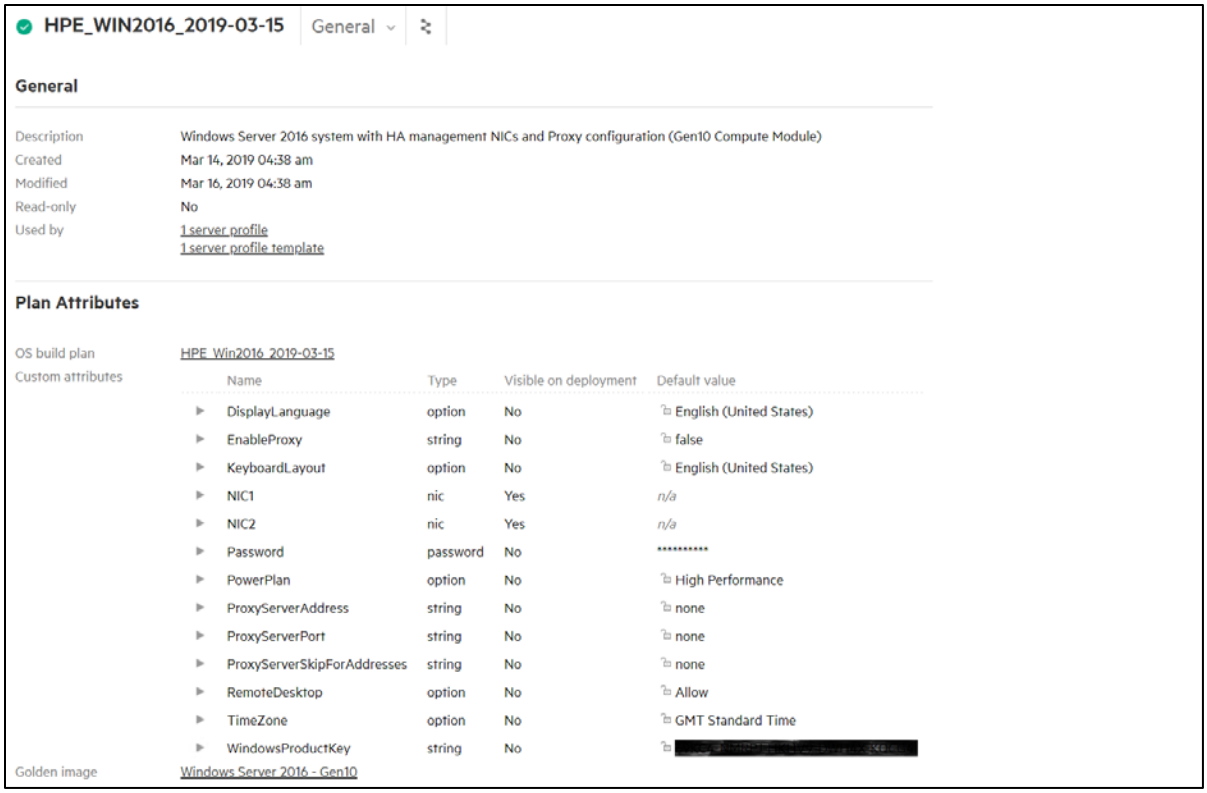

Figure 32. Windows 2016 Deployment Plan attributes

Windows Proxy Server Configuration

This Deployment Plan includes the ability to configure a Proxy server if needed. There are four custom attributes related to proxy server configuration:

Table 22. Custom attributes for proxy configuration

Custom attribute name	Purpose	Default value
EnableProxy	Controls whether the remaining proxy-related custom attributes are applied to the server during OS deployment	false
ProxyServerAddress	The hostname or IP address of the proxy server	none
ProxyServerPort	The numeric port number used by the proxy server	none
ProxyServerSkipForAddresses	Hostnames or IP addresses that are excluded from the proxy server	none

By default the EnableProxy custom attribute is set to "false" which causes the other three proxy-related custom attributes to be ignored. However, all of these custom attributes require a string value be configured (i.e. they cannot be left blank), which is why the remaining proxy attributes are set to "none". In environments where a proxy server is required to reach the internet, the `EnableProxy` attribute must be set to "true" and the `ProxyServerAddress`, `ProxyServerPort`, and `ProxyServerSkipForAddresses` attributes should be configured with their appropriate values.

For more information about custom attributes and the type of attributes available, see the HPE Synergy Image Streamer 4.1 User's Guide.

RHEL Golden Images

OS installation and configuration with HPE Synergy Image Streamer

The bare metal RHEL worker nodes will be deployed and customized using HPE Synergy Image Streamer. This section outlines the steps required to install the host. At a high level, these steps can be described as:

1. Download the artifacts for HPE Image Streamer from the HPE GitHub site.
2. Add the artifact bundles to HPE Image Streamer.
3. Prepare a compute module for the installation of the Operating System.
4. Create a Server Profile.
5. Install and customize the Operating System.
6. Capture a Golden Image from the compute module.
7. Deploy the hosts.

Download the artifacts for HPE Synergy Image Streamer

Red Hat Enterprise Linux bundles for HPE Image Streamer may be downloaded from https://github.com/HewlettPackard/image-streamerrhel/tree/V4.1/artifact-bundles/.

Sample foundation artifact bundles should be downloaded from https://github.com/HewlettPackard/imagestreamer-tools/tree/v4.0/foundation/artifact-bundles.

Add the artifact bundles to HPE Image Streamer

The following steps show how to add the artifact bundles to Image Streamer:

1. From within the HPE Image Streamer interface navigate to the Artifact Bundles page.
2. From the Actions menu, Add the downloaded RHEL artifact bundle. If not already present, add the sample foundation bundle.
3. From the Actions menu, select Extract to extract the artifacts from each downloaded bundle.

Prepare the compute module for the installation of the Operating System

1. Attach a Red Hat Enterprise Linux 7.* Server ISO to the iLO of a worker node host by selecting the Action menu and then Launch Console.
2. When the console launches, select Virtual Drives and then Image File CD-ROM/DVD. Browse to the location where your ISO resides and select it.

Create a Server Profile

1. Use a Server Profile Template to deploy a new Server Profile to the worker node host you attached the ISO to in the prior step.
2. Select the new Server Profile and choose Edit.
3. Under the OS Deployment section, choose HPE-Create Empty Volume and enter a Volume Size of 30720 MB.
4. Validate the network and SAN connections exist on the host.
5. Ensure that, under boot settings, Boot mode is set to UEFI optimized and that the Primary boot device is Hard disk.
6. Click OK. It will take some time for the profile to create.
7. While waiting on profile creation to complete, select the Actions menu and then click Launch Console. Click Allow to launch the console.

Install and customize the Operating System

1. After profile creation is completed, power on the server. From the console window, select Power Switch and then Momentary Press.
2. When the screen shown in the following figure appears, select Install Red Hat Enterprise Linux 7.* and then hit the letter ‘e’ on the keyboard.

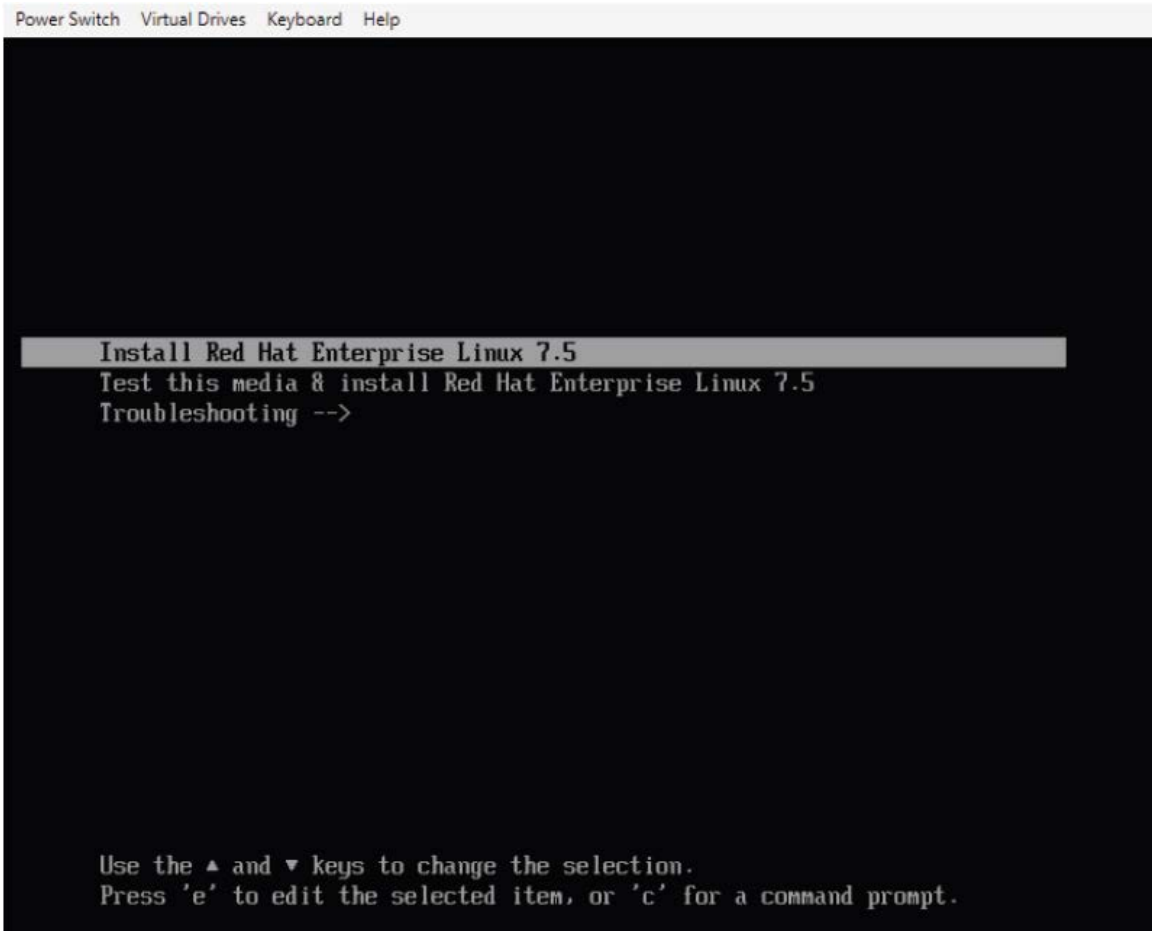

Figure 33. Selecting OS to install

3. Append the following to the install kernel boot parameter: `rd.iscsi.ibft=1`
4. Type `Ctrl-x` to continue the boot process.
5. When the installer screen appears, insure you select your local language, set the date and time, keyboard layout and language support. When done, click `Installation Destination`.
6. At the Installation Destination screen, select `Add a disk` and then choose the 30 GiB volume from HPE Image Streamer. Select `Done` once you have chosen this disk.
7. Under `Other Storage Options`, select the radio button for `I will configure partitioning` and then click `Done`.
8. At the `Manual Partitioning` screen, select `Click here to create them automatically`. This will display a new Manual Partitioning screen.
9. Highlight the `/boot` partition and on the right side of the page select `ext4` as the File System. Click the `Update Settings` button.
10. Highlight the `/` partition and on the right side of the screen, reduce the `Desired Capacity` to `8 GiB` and then choose `ext4` as the File System. Click the `Update Settings` button.
11. Highlight the `swap` partition and on the right side of the screen, change `Desired Capacity` from `3000 MiB` to `4092 MiB.` Click the `Update Settings` button.
12. Click the **"+"** button below the list of partitions. For `Mount Point`, select `/var` from the dropdown and leave the `Desired Capacity` blank. This will allow the `/var` partition to use all remaining space.

13. At the `Manual Partitioning` screen, highlight the `/var` partition and choose `/ext4` for the File System. Click `Update Settings.`
14. The screen should appear as shown in the following figure.

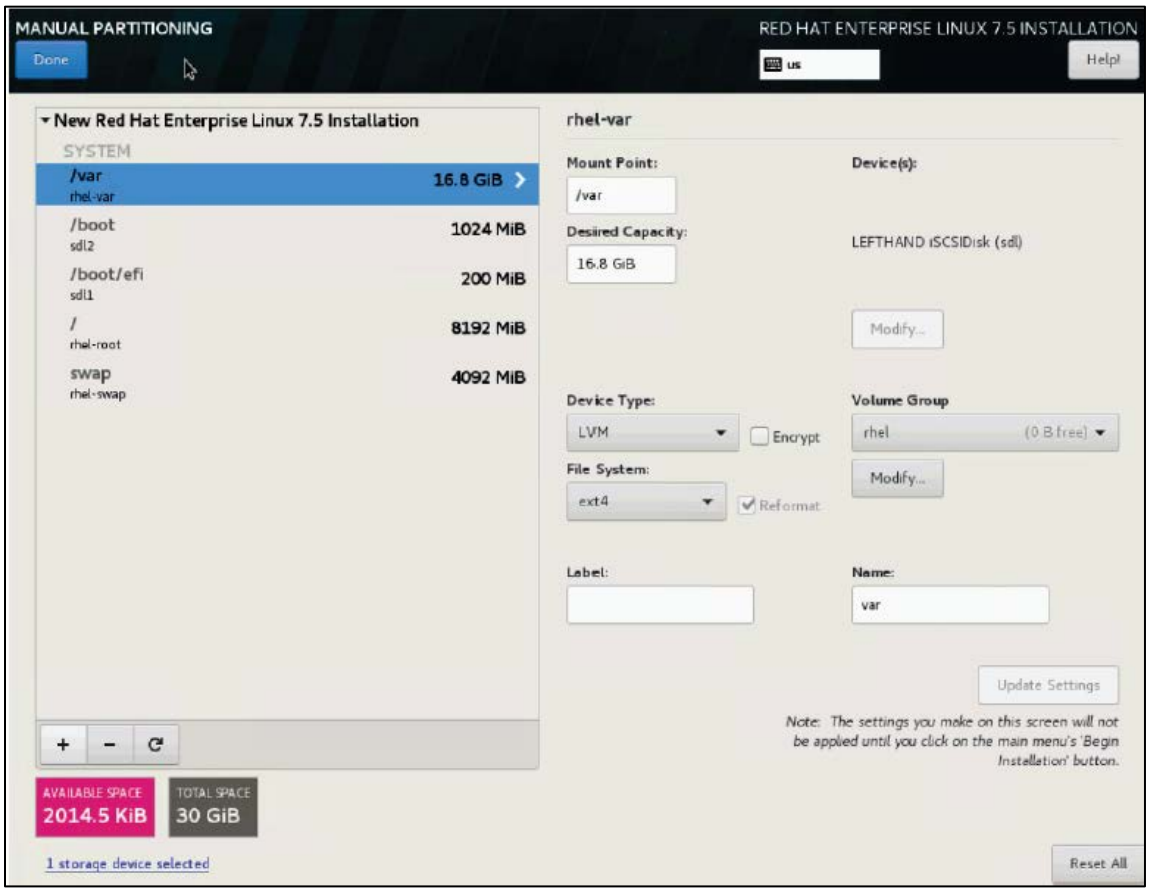

Figure 34. Manual partitioning

15. Once you have validated the file systems and partition sizes are correct, click Done.
16. When prompted, click Accept Changes.
17. Click the Network & Hostname link. At the resulting screen, highlight Ethernet (ens3f4) and set it to 'ON' in the descriptor screen as in the following figure. Click Done.

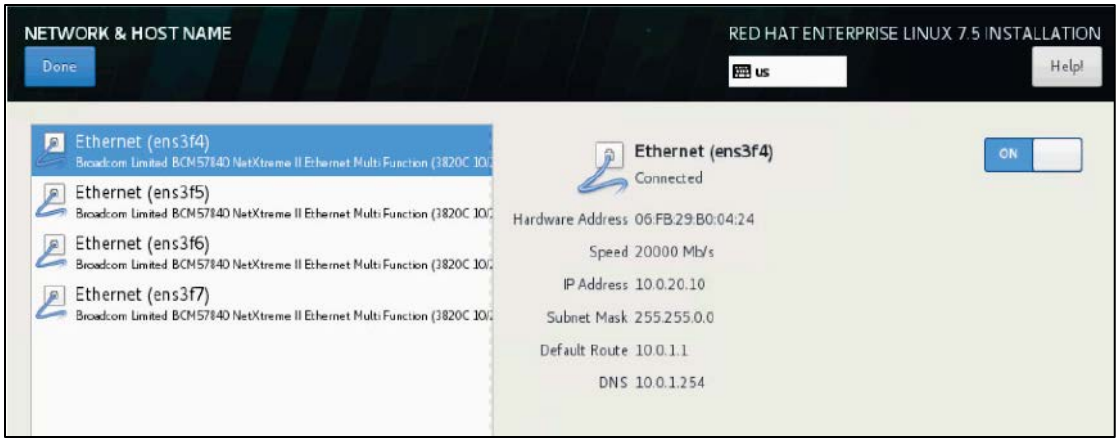

Figure 35. Network & Host Name

18. Click Begin Installation. Set a root password for the host. Do not configure extra users. Click Done.
19. Once the OS installation is complete you can reboot the host. Log on at the iLO console when the host becomes active again.
20. Configure a temporary hostname for the system
21. Configure your networking and ensure you have connectivity.
22. Register the host with Red Hat by running the following command. Enter the username and password when prompted.

```
# subscription-manager register
```

23. Use Red Hat subscription manager to register your system to give you access to the official Red Hat repositories. Use the subscription-manager register command as follows.

```
# subscription-manager auto-attach
```

24. Enable the required repos:

```
# subscription-manager repos \
--enable=rhel-7-server-rpms \
--enable=rhel-7-server-extras-rpms
```

25. Update the host by running the following command.

```
# yum update
```

26. Copy the SSH public key from your Ansible box. This will allow your Ansible node to SSH without the need for a password to all the bare metal REHL nodes.

```
# ssh-copy-id root@<IP of your bare metal node>
```

27. Gracefully shut down the host.

```
# yum update
```

Windows Golden Images

Prepare Image Streamer with Windows Artifact Bundle

- Download the "HPE - Windows - 2018-10-26.zip" artifact bundle from the GitHub repository at https://github.com/HewlettPackard/image-streamer-windows. The file is available in the artifact-bundles directory. The artifacts are supported on HPE Image Streamer 4.1 and higher for Windows 2016, while version 4.2 and higher are required for Windows 2019. This solution has been tested using Windows 2016 on HPE Image Streamer 4.1.
- Upload the Artifact bundle to the Image Streamer appliance
- Extract the Artifact Bundle on the Image Streamer appliance

Create Windows Golden Image

The procedure for creating a Windows Server 2016 golden image are documented in the Image Streamer GitHub repostiory at https://github.com/HewlettPackard/image-streamer-windows. See the appropriate file in the `docs` directory here.

The instructions are repeated here for convenience, but you should rely on the Image Streamer repository for the definitive version of the documentation.

1. Ensure that you have access to Windows 2016 or 2019 ISO file.
2. Create a server profile with “HPE - Foundation 1.0 - create empty OS Volume” as OS Deployment plan and a server hardware of desired hardware type (see section on Golden Image Compatibility below). Set an appropriate value for volume size in MiB units, say 40000 MiB. The HPE Synergy Server will be configured for access to this empty OS Volume.
3. Launch iLO Integrated Remote Console of this server and set the Windows 2016 or 2019 ISO file as virtual CD-ROM/DVD image file. Power on the server.
4. Windows should present an option of installing from CD/DVD. Continue with this option.
5. Install Windows 2016 or 2019.
6. (Optional) To take a backup of this installation at this stage:
   a. Shutdown the server
   b. Perform an as-is capture using "HPE - Windows - Capture - As-Is" build plan to create the "as-is" golden image of the OS.
   c. Deploy another server with the golden image captured in previous step and boot the server.
7. Install any additional software or roles if required.

Note

The next six steps can be automated using the “PrepareForImageStreamerOSVolumeCapture.ps1” script in “scripts” directory on the GitHub repository where Windows artifact bundles are available for download.

---

8. Create a FAT32 partition which will be used by the artifacts for personalization: FAT 32 partition can be created either from UI using Disk Management utility (8.1) or using CMD Diskpart commands (8.2).

   8.1 FAT32 partition creation from UI

   a. Open "Computer Management" > "Disk Management"
   b. Select C: partition
   c. Shrink volume
   d. Change amount of space to shrink to 100 MB
   e. Select Shrink
   f. Select new Unallocated space
   g. Select New Simple Volume

h. Leave size
i. Assign drive letter, (Choose S)
j. Format as FAT32 file system type (this requires changing from the default)
k. Give Volume label as "ISDEPLOY"
l. Finish
m. “ISDEPLOY (S:)” should be shown

8.2 FAT32 partition creation using CMD commands
Use list volume command to get volume number for C: partition. Here C: partition resides in Volume 0.

```
C:\Users\Administrator>diskpart
DISKPART>list volume
DISKPART >select volume 0
DISKPART >shrink desired=100
DISKPART >create partition primary size=100
DISKPART >format fs=fat32 quick label=ISDEPLOY
DISKPART >assign letter=S
```

9. Backup drive-letters

```
reg export HKLM\System\MountedDevices C:\driveletters.reg
```

10. Generalize Windows using Sysprep

WARNING: This operation is destructive and will remove all configuration. To take backup of the system at this stage, capture an as-is golden image.

Open Command Prompt window and run the following

```
cd Windows\System32\Sysprep
Sysprep /generalize /oobe /quit
```

This will take a few minutes to complete and will generalize the system. All settings will be lost. This does not remove any additional user accounts that are created. Any user accounts not required in the captured golden image must be manually deleted.

11. Restore drive-letters

```
reg import C:\driveletters.reg
```

12. Set Unattend.xml location to the FAT32 partition

```
reg add HKLM\System\Setup /v UnattendFile /t REG_SZ /d ”S:\ISdeploy\Unattend.xml”
```

13. Set SetupComplete.cmd location to the FAT32 partition

```
mkdir C:\Windows\Setup\Scripts
echo S:\ISdeploy\SetupComplete.cmd > C:\Windows\Setup\Scripts\SetupComplete.cmd
```

14. Shutdown the server.

15. Capture a golden image using the "HPE - Windows - Capture - As-Is" build plan as described in the following section.

Capture the Golden Image

- Determine the OS Volume that was created for the Server Profile created earlier:

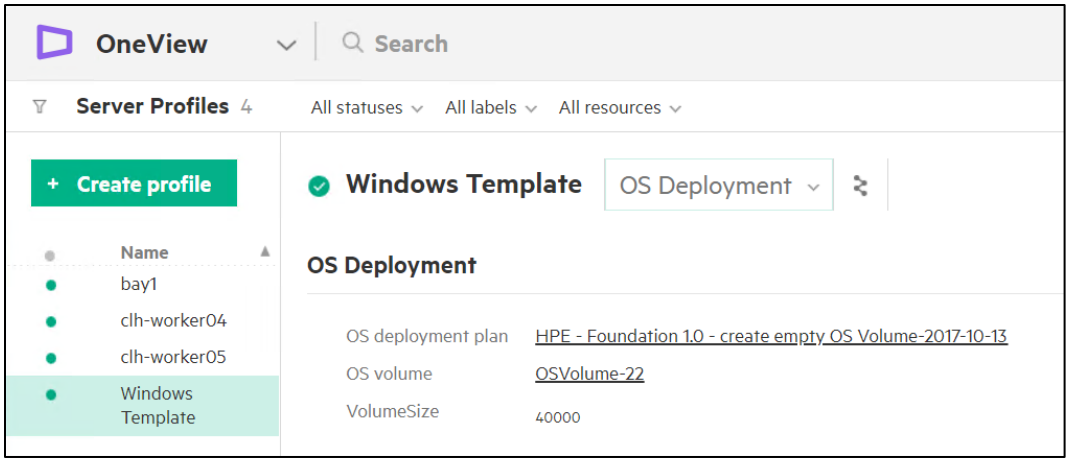

Figure 36. Server profile

- Navigate to the Image Streamer Golden Images page
- Select "Create golden image" specifying the OS Volume and the "HPE - Windows - Capture - As-Is" build plan:

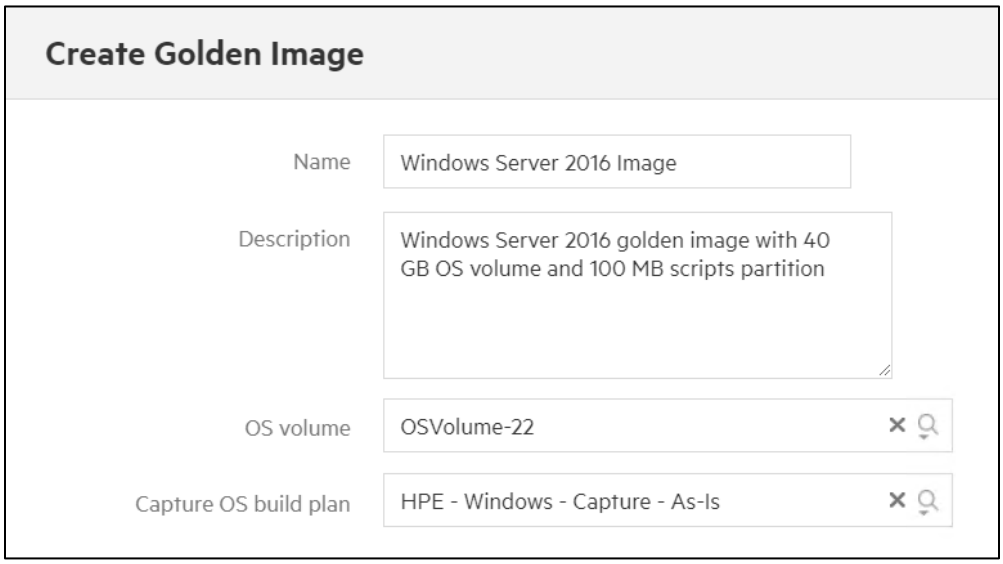

Figure 37. Create Golden Image

- Select "Create"
- Delete the Server Profile "Windows Template" used to create the golden image

Golden Image Compatibility

The golden image created using the above method will work only when the image is deployed on server hardware of the same model. Specifically, if the number of processors on server where the image is deployed is different from the server where the image was captured, server boot after deployment will fail. Also, if the boot controller is moved from one Mezzanine slot on the server to another, Windows will not boot correctly.

OS Deployment Plans

The solution delivers two artifact bundles, one for Windows Server 2016 systems and one for Red Hat Linux 7 systems. Each artifact bundle contains one Deployment Plan, one OS Build Plan and all dependent Plan Scripts.

The artifact bundles are included in the Docker-Synergy repository:

```
# cd ~/Docker-Synergy
# ls ./files/ImageStreamer

HPE_RHEL7_2019_02_25.zip
HPE_WIN2016_2019-03-15.zip
```

In the Image Streamer UI, use the `Add Artifact bundle` button in the `Artifact Bundles` screen to upload the two files. When the upload is finished, select the Artifact bundles corresponding to the files (without the .zip extension) and use the `Actions` button to extract artifacts from the bundles, as shown in the following figure.

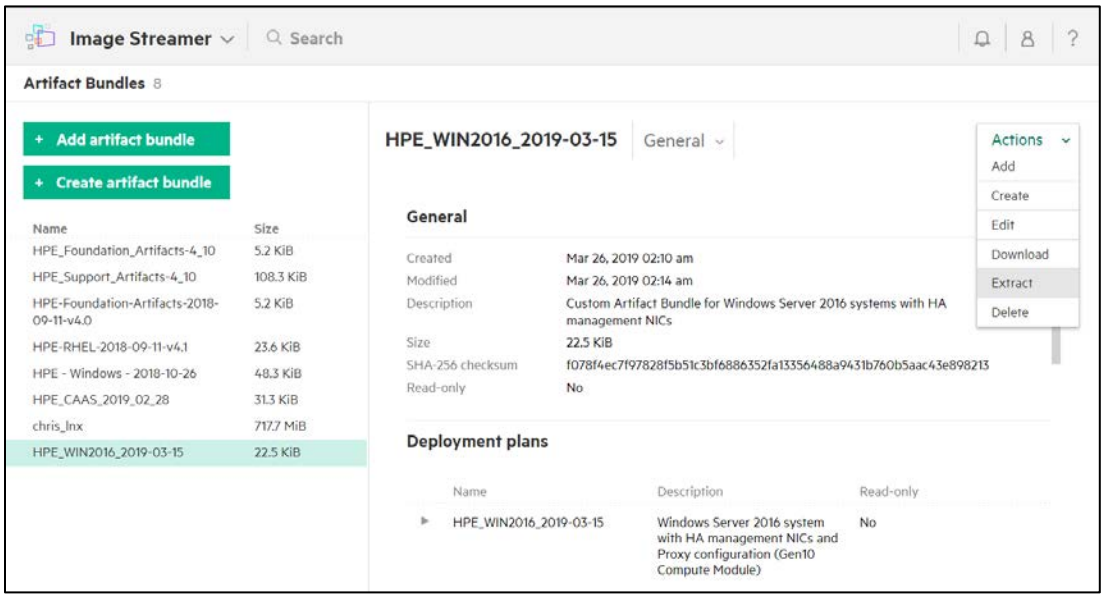

Figure 38. Extract bundle

After the extraction completes, you should find two new deployment plans in your Image Streamer appliance named:

- HPE_RHEL7_2019_02_25 which is the Red Hat Enterprise Linux 7 OS Deployment Plan
- HPE_WIN2016_2019-03-15 which is the Microsoft Windows Server 2016 OS Deployment Plan

The deployment plans are shipped without a golden image. Golden images for each OS must be created as outlined in the previous sections.

Update the Red Hat OS Deployment plan

1. Select the OS Deployment Plan named `HPE_RHEL7_2019_02_25` on the `Deployment Plans` menu in the Image Streamer UI.
2. Click the `Action` button, then `Edit` to edit the deployment plan.
3. In the Edit screen, locate the Golden Image drop-down widget and select the golden image created with Red Hat Linux 7.
4. Ensure that the visibility of the custom attributes is configured as explained earlier (i.e. only NIC1 and NIC2 should be visible).
5. Save your changes.

Update the Windows Server 2016 Deployment plan

1. Select the OS Deployment Plan named `HPE_WIN2016_2019-03-15` on the `Deployment Plans` menu in the Image Streamer UI.
2. Click the `Action` button, then `Edit` to edit the deployment plan.
3. In the Edit screen, locate the Golden Image drop-down widget and select the golden image created with Microsoft Windows Server 2016.
4. Make sure the visibility of the custom attributes is configured as explained earlier (ie only NIC1 and NIC2 should be visible).
5. Save your changes

OneView Server Profile Templates

The server profile template must meet the following criteria:

- The template must specify an Image Streamer Deployment Plan and the deployment plan must match the constraints explained in the section OS Deployment Plan Custom Attributes.
- There must be at least one data drive in addition to the boot device provided by the Image Streamer. The playbooks supports local drives as well as drives configured from a Synergy D3940 storage module or LUNs from an HPE 3PAR array.
- There must be two Ethernet connections mapped to the Ethernet network used by the Ansible controller node.

Deploying Sysdig monitoring

By default, the playbooks for deploying Sysdig are commented out in `site.yml` and must be explicitly enabled in that file if you want it included in the initial deployment. Alternatively, you can run the specific playbooks detailed in this section in a stand-alone manner, subsequent to the initial deployment.

Note

By default, you must have outgoing port `6666` open in your firewall, to allow data to flow to `collector.sysdigcloud.com`. You can configure the agent to use a different port by using the variable `sysdig_collector_port` in `group_vars/all/vars.`

If you are using a proxy, it must be configured to be "fully-transparent". Non-transparent proxies will not allow the agent to connect.

Monitoring with Sysdig

Sysdig's approach to Docker monitoring uses transparent instrumentation to see inside containers from the outside, with no need for agents in each container. Metrics from Docker containers, and from your applications running inside them, are aggregated in real-time across each service to provide meaningful monitoring dashboards and alerts for your application. Figure 39 provides an overview of the Sysdig architecture.

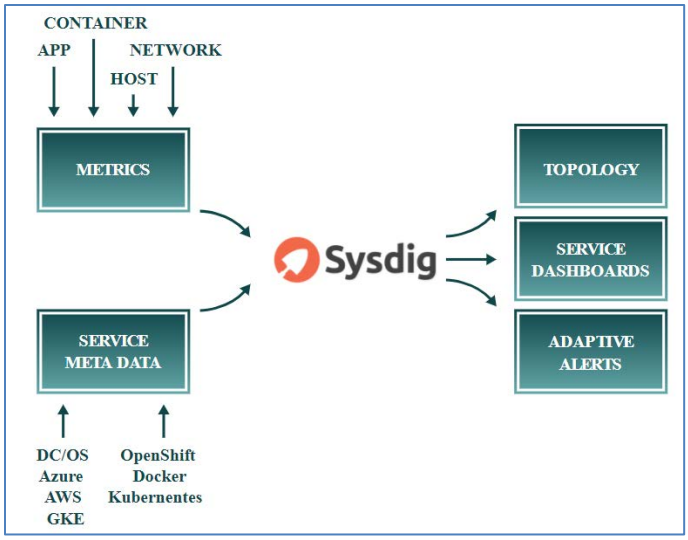

Figure 39. Sysdig architecture

Sysdig Monitor allows you to analyze response times, application performance metrics, container and server utilization metrics, and network metrics. You can build dashboards across applications, micro-services, container and networks, and explore metadata from Docker, Kubernetes, Mesos and AWS. For more information, see the Sysdig Container Monitoring video overview and the Sysdig Monitor 101 training course.

Sysdig Secure provides security at the orchestrator as well as the container level. You create service-aware policies that allow you to take actions (like killing a container) or send alerts (to Slack, Splunk, etc) whenever a policy violation occurs. All commands are audited to help you identify anomalous actions, along with taking snapshots of all activities pre-and-post a policy violation. For more information, see the Sysdig Secure video overview and the Sysdig Secure 101 training course.

The implementation in this solution uses the Software as a Service (SaaS) version of Sysdig. The playbooks deploy Sysdig Agent software on each UCP, DTR and Linux worker node, as well as the NFS, logger and load balancer VMs 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 - see the section on Sysdig configuration 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://hewlettpackard.github.io/Docker-Synergy/sysdig/sysdig-trial.html.

Playbooks for installing Sysdig on RHEL

The following playbooks are used when deploying Sysdig:

- `playbooks/sysdig-k8s-rbac.yml` is used to configure Sysdig for Kubernetes.
- `playbooks/install_sysdig.yml` is used to configure Sysdig for Docker swarm. It opens the required port in the firewall, and installs the latest version of the Sysdig agent image on the nodes. By default, this playbook is commented out in `site.yml`, so if you want to use the solution to automatically configure Sysdig for Docker swarm, you must uncomment this line.

Sysdig configuration

Separate playbooks are used to install Sysdig for Docker swarm and Sysdig for Kubernetes.

Sysdig configuration for Docker swarm

The playbook `playbooks/install_sysdig.yml` is used to automate the configuration of the SaaS setup for Docker swarm. By default, this playbook is commented out in `site.yml` and must be explicitly enabled. The variables used to configure Sysdig for Docker swarm are detailed in Table 23.

Table 23. Sysdig variables for Docker swarm

Variable	File	Description
sysdig_access_key	**group_vars/all/vault**	After the activation of your account on the Sysdig portal, you will be provided with your access key. This is used by the playbooks to install the agent on each UCP, DTR and Linux worker node, as well as the NFS, logger and load balancer VMs.
sysdig_agent	group_vars/all/vars	Specifies the URL to the Sysdig Linux native install agent, for example, `https://s3.amazonaws.com/download.draios.com/stable/install-agent`
sysdig_tags	group_vars/all/vars	Tagging your hosts is highly recommended. Tags allow you to sort the nodes of your infrastructure into custom groups in Sysdig Monitor. Specify location, role, and owner in the format: `'location:City,role:Express Containers,owner:Customer Name'`

Sysdig configuration for Kubernetes

The playbook `playbooks/sysdig-k8s-rbac.yml` is used to automate the configuration of the SaaS setup for Kubernetes. The variables used to configure Sysdig for Kubernetes are detailed in Table 24.

Table 24. Sysdig variables for Kubernetes

Variable	File	Description
sysdig_access_key	**group_vars/all/vault**	After the activation of your account on the Sysdig portal, you will be provided with your access key. This is used by the playbooks to install the agent on each UCP, DTR and Linux Kubernetes worker nodes.
sysdig_collector	group_vars/all/vars	The URL for the Sysdig SaaS, by default, '`collector.sysdigcloud.com`'
sysdig_collector_port	group_vars/all/vars	The port used by the agent, by default, `'6666'`
sysdig_tags	group_vars/all/vars	Tagging your hosts is highly recommended. Tags allow you to sort the nodes of your infrastructure into custom groups in Sysdig Monitor. Specify location, role, and owner in the format: `'location:City,role:Express Containers,owner:Customer Name'`

k8s_cluster	group_vars/all/vars	This should match the cluster name displayed when you source the environment setup script, for example:
		
		`# source env.sh`
		`Cluster "ucp_hpe-ucp.cloudra.local:6443_admin" set.`
		`User "ucp_hpe-ucp.cloudra.local:6443_admin" set.`
		
		For more information, see the section on installing the UCP client bundle in the section Deploying Sysdig monitoring on Kubernetes.

Registering for Sysdig trial

Hewlett Packard Enterprise has teamed up with Sysdig to offer a fully featured 90-day trial version of Sysdig Monitor and Secure as part of the HPE Enterprise Containers as a Service with Docker EE solution. For more details on how to sign up, see the GitHub repository at https://github.com/HewlettPackard/Docker-Synergy.

After registering for the trial, you will be presented with options for setting up your environment, as shown in Figure 40.

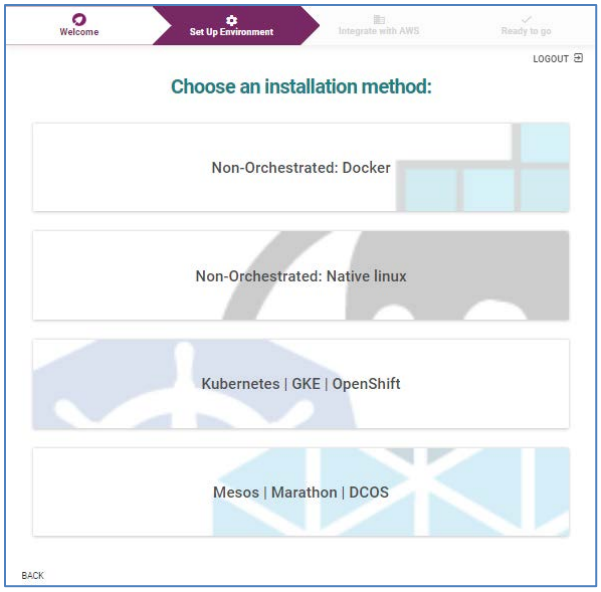

Figure 40. Sysdig Monitor set up environment

Sysdig Monitoring for Kubernetes

If you are deploying Sysdig monitoring on Kubernetes, select the `Kubernetes | GKE | OpenShift` option. You will be presented with an access code, as shown in Figure 41.

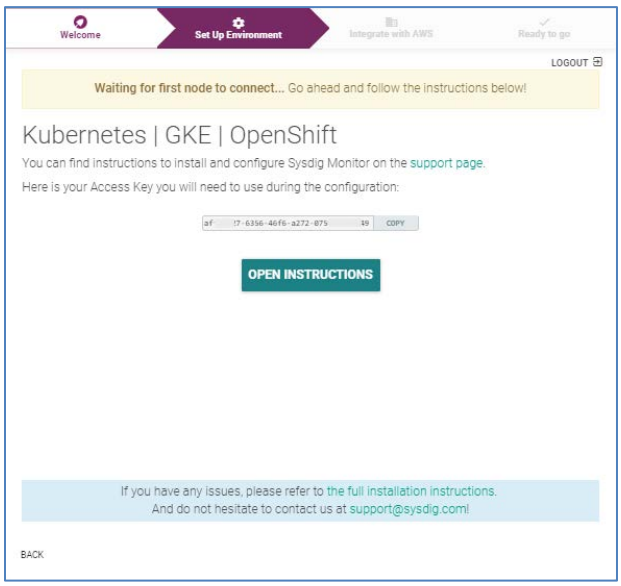

Figure 41. Sysdig Monitor access code for Kubernetes

Use the `sysdig_access_key` field in your `group_vars/all/vault`, as described in the section Sysdig configuration for Kubernetes. Once you deploy your environment and your Kubernetes nodes connect to the Sysdig SaaS platform, Sysdig will automatically display information regarding your setup, as shown in Figure 42.

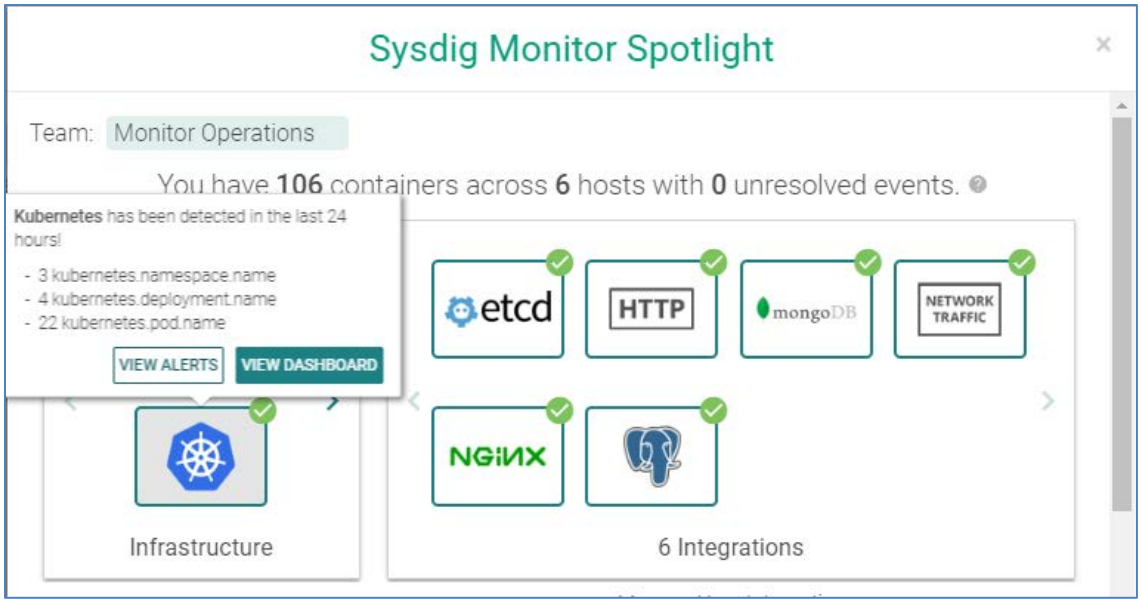

Figure 42. Sysdig Monitor Spotlight for Kubernetes

Select `View Dashboard` for an entry point to accessing all your monitoring data. Alternatively, you can browse to https://app.sysdigcloud.com at any time to access your dashboards.

Sysdig Monitor for Docker swarm

If you are deploying Sysdig monitoring on Docker swarm, select the `Non-Orchestrated: Native Linux` option. You will be presented with a screen containing details for the URL to download the Sysdig agent, along with your access code embedded in the command, as shown in Figure 43.

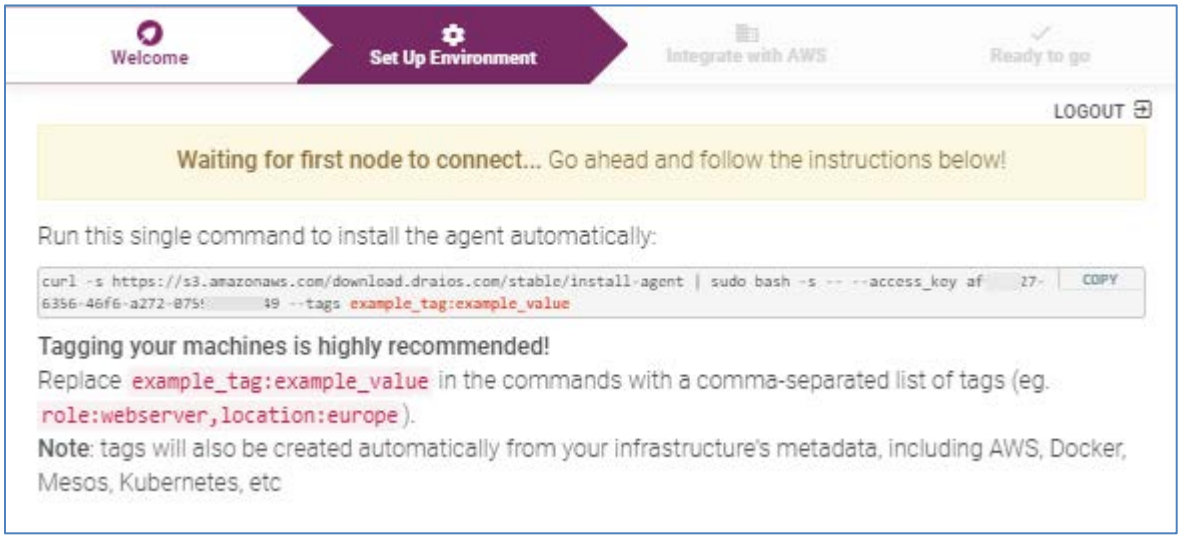

Figure 43. Sysdig Monitor download location and access code for Docker

The download URL is used in the `sysdig_agent` field in `group_vars/all/vars`, while the access code is stored in the `sysdig_access_key` field in your `group_vars/all/vault`, as described in the section Sysdig configuration for Docker swarm.

Once you deploy your environment and your Docker swarm nodes connect to the Sysdig SaaS platform, Sysdig will automatically display information regarding your setup. Alternatively, you can browse to https://app.sysdigcloud.com at any time to access your dashboards.

Deploying Sysdig monitoring on Kubernetes

The latest version of Sysdig supports monitoring of Kubernetes logs and metrics.

Prerequisites

- Install the `kubectl` binary on your Ansible box.
- Install the UCP Client bundle for the admin user.
- Confirm that you can connect to the cluster by running a test command, for example, `kubectl get nodes`
- Ensure that you have configured the required variables, as described in the section Sysdig configuration for Kubernetes

For example, you add the relevant variables in the `group_vars/all/vars` file.

```
sysdig_collector: 'collector.sysdigcloud.com'
sysdig_collector_port: '6666'
sysdig_tags: 'location:Enter city,role:Enter role,owner:Customer name'
k8s_cluster: 'ucp_hpe2-ucp.cloudra.local'
```

You should add the access key to the encrypted `group_vars/all/vault` using the command `ansible-vault edit group_vars/all/vault`.

```
sysdig_access_key: '10****97-9160-****-9061-84bfd0f****0'
```

Running the playbook

The playbook `playbooks/k8s-install-sysdig.yml` is used to automate the configuration of the SaaS setup for Docker swarm.

```
# cd Docker-Synergy
# ansible-playbook -i hosts playbooks/sysdig-k8s-rbac.yml --vault-password-file .vault_pass
```

Using the Sysdig software as a solution (SaaS) website https://app.sysdigcloud.com, you are able to view, analyze and inspect various different dashboards. Initially, you will just see the monitoring information for the infrastructure itself. Deploy a sample application, as detailed in the section Kubernetes guestbook example with Redis, and use the Sysdig solution to analyze the different facets of the deployed application.

Deploying Sysdig monitoring on Docker Swarm

The playbook `playbooks/install_sysdig.yml` is used to automate the configuration of the SaaS setup for Docker swarm. By default, this playbook is commented out in `site.yml` and must be explicitly enabled. An access key variable must be set in the `group_vars/all/vault` file as detailed in Table 23.

```
# cd Docker-Synergy
# ansible-playbook -i hosts playbooks/install_sysdig.yml --vault-password-file .vault_pass
```

Using the Sysdig software as a solution (SaaS) website https://app.sysdigcloud.com, you are able to view, analyze and inspect various different dashboards.

Deploying Splunk

This section provides an overview of Splunk, outlines how to configure and run the relevant playbooks and shows how to access the UI to see the resultant Docker and Kubernetes dashboards.

Monitoring with Splunk

Splunk Enterprise allows you to collect and index any data from any source, and to monitor systems and infrastructure in real time to preempt issues before they happen. It allows you to analyze your data to understand trends, patterns of activity and behavior, giving you valuable intelligence across your entire organization. The solution architecture for Splunk is shown in Figure 44.

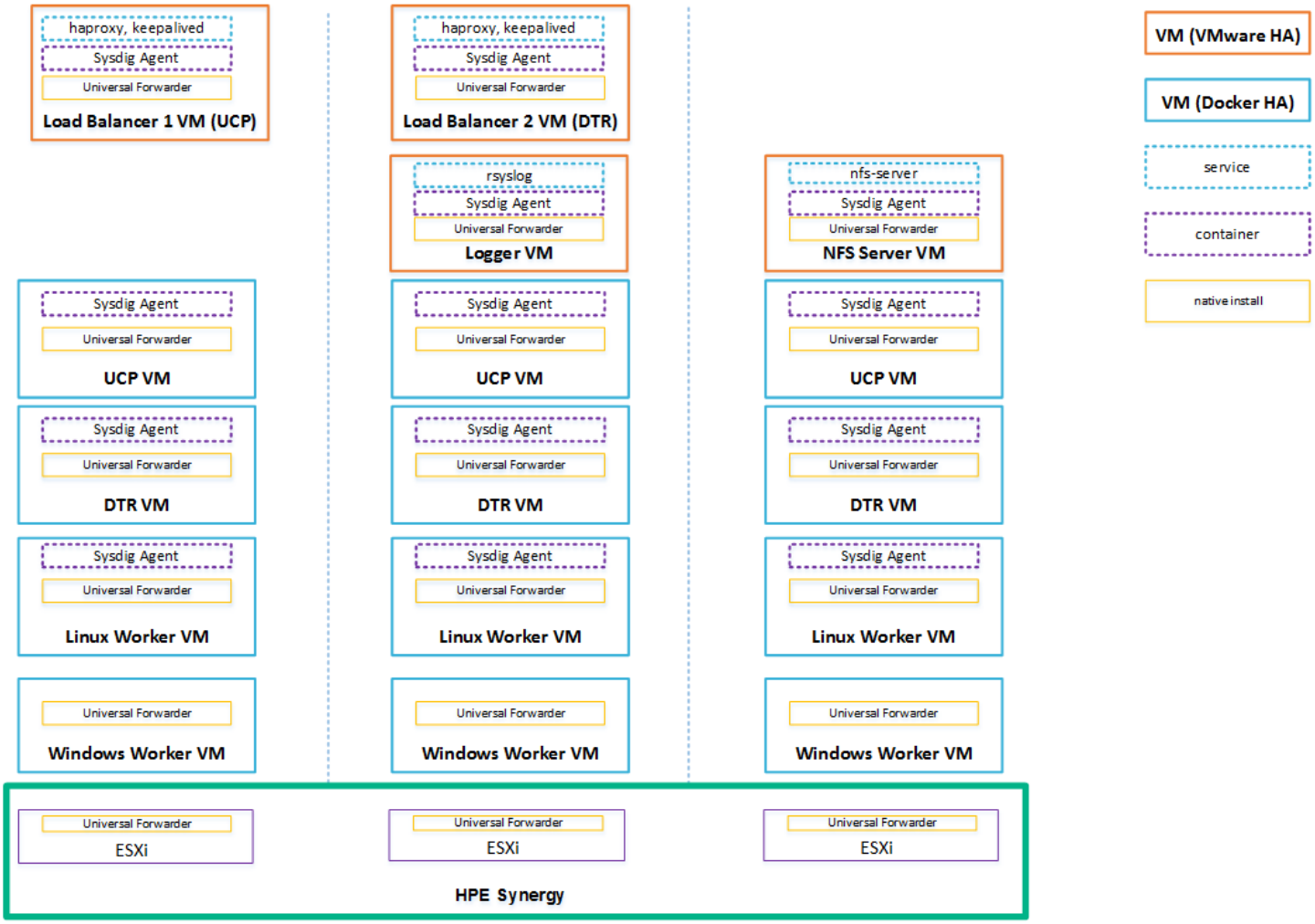

Figure 44. Solution architecture: Hybrid Linux and Windows workers with Splunk and Sysdig

This solution allows you to integrate your CaaS deployment with an existing Splunk Enterprise installation or to deploy a stand-alone Splunk Enterprise demo environment as a Docker stack in your cloud. In both instances, Universal Forwarders are used to collect data from your applications running on your Linux and Windows worker nodes in your cloud, as well as log data from the Docker platform itself and from the infrastructure VMs and servers. Figure 45 shows the Splunk architecture.

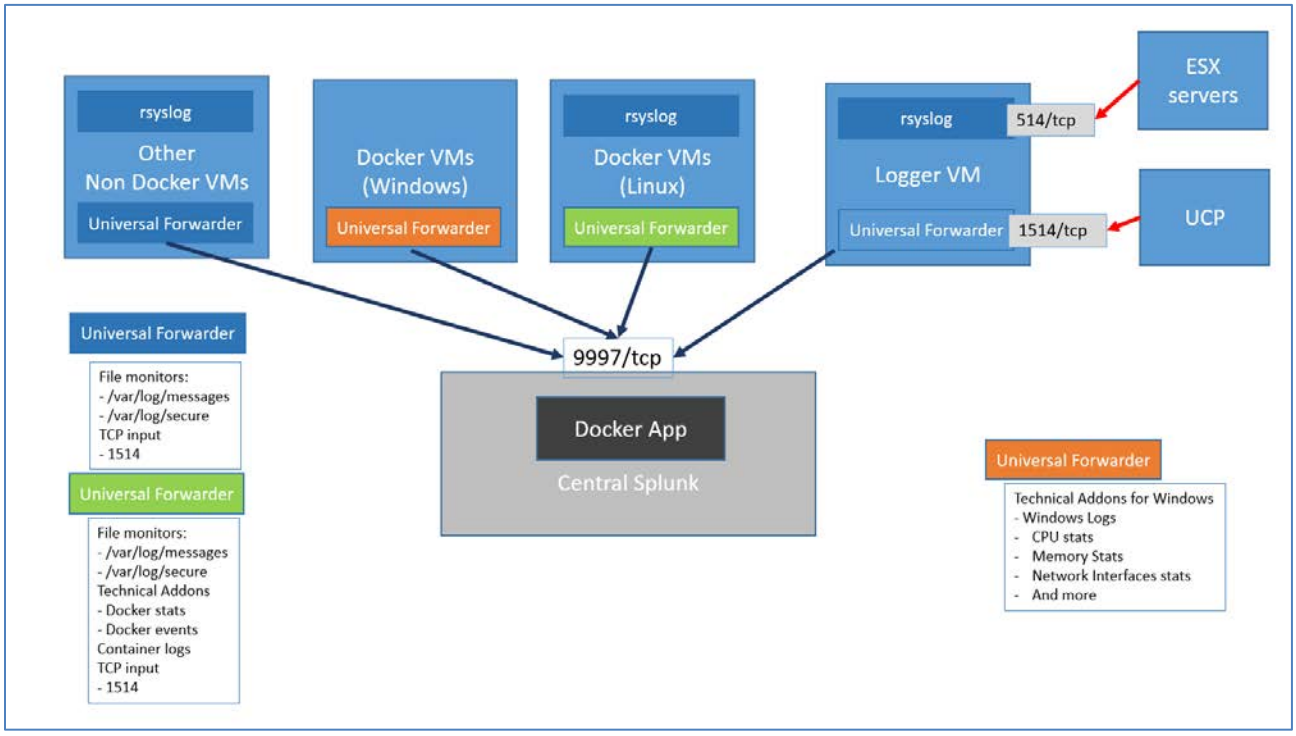

Figure 45. Splunk architecture

All the Universal Forwarders run natively on the operating system to allow greater flexibility in terms of configuration options. Each forwarder sends the data it collects to one or more indexers in the central Splunk.

Linux worker nodes: The Universal Forwarders on the Linux worker nodes collect log and metrics data. The log data includes:

- `/var/log/messages` from the Docker host (including the daemon engine logs)
- `/var/log/secure` from the Docker hosts
- container logs via a Splunk technical add-on

The metrics data is collected via a technical add-on and includes:

- `docker stats`
- `docker top`
- `docker events`
- `docker service stats`

Windows worker nodes: The Universal Forwarders running on the Windows worker nodes collect the following data:

- Windows logs
- CPU stats
- Memory stats
- Network Interface stats
- and more

For more information on configuring standalone Splunk for Linux and Windows worker nodes, see the section on Splunk prerequisites.

UCP and ESXi: UCP operational logs and ESXi logs are forwarded to the logger VM via TCP ports 1514 and 514 respectively. Port 1514 is assigned a special `sourcetype` of `ucp` which is then used by the Splunk Docker APP to interpret UCP logs. The Universal Forwarder runs the rsyslog daemon which will record the log messages coming from the ESX machines into the `/var/log/messages` file on the VM.

Non-Docker VMs: Other VMs, for example, NFS, use a Splunk `monitor` to collect and forward data from the following files:

- /var/log/messages
- /var/log/secure (Red Hat)

Note
You can configure the list of files monitored by the Universal Forwarder.

Other syslog senders can be configured to send their data to the logger VM or directly to central Splunk.

Playbooks for installing Splunk

The following playbooks are used to install Splunk.

- `playbooks/splunk_demo.yml` installs a demo of Splunk Enterprise in the cluster (if the `splunk_demo` deployment option is selected. A value of `splunk` is used to configure an external production Splunk deployment.)
- `playbooks/splunk_uf.yml` installs and configures the Splunk Universal Forwarder on each Linux and Windows node in the inventory

Splunk configuration

This solution supports two types of Splunk deployments. Firstly, there is a built-in deployment useful for demos and for getting up to speed with Splunk. Alternatively, the solution can be configured to interact with a standalone, production Splunk deployment that you set up independently. In this case, you must explicitly configure the universal forwarders with external "forward servers" (Splunk indexers), whereas this happens automatically with the built-in option.

In the standalone deployment, you can enable SSL authentication between the universal forwarders and the indexers, by setting the `splunk_ssl` variable to `yes` in the file `group_vars/all/vars`. The built-in demo deployment does not support SSL and so, in this instance, the value of the `splunk_ssl` variable is ignored. For more information on enabling SSL, see Appendix C.

After the installation is complete, the Splunk UI can be reached at `http://<fqdn>:8000`, where `<fqdn>` is the FQDN of one of your Linux Docker nodes. Mesh routing does not currently work on Windows so you must use a Linux node to access the UI.

Splunk prerequisites

You should select the Splunk deployment type that you require by setting the variable `monitoring_stack` in the `group_vars/all/vars` file to either **splunk**, to use a standalone Splunk deployment, or **splunk_demo** for the built-in version. If you omit this variable, or if it has an invalid value, no Splunk deployment will be configured.

For both types of deployment, you need to download the Splunk universal forwarder images/packages from https://www.splunk.com/en_us/download/universal-forwarder.html. Packages are available for 64-bit Linux and 64-bit Windows 8.1/Windows 10. Download the RPM package for Linux 64-bit (2.6+ kernel Linux distributions) to `./files/splunk/linux`. If you are deploying Windows

nodes, download the MSI package for Windows 64 bit to `./files/splunk/windows`. For a dual Linux/Windows deployment, the images and packages must have same name and version, along with the appropriate extensions, for example:

- files/splunk/windows/splunkforwarder-7.1.2.msi
- files/splunk/linux/splunkforwarder-7.1.2.rpm

You need to set the variable `splunk_architecture_universal_forwarder_package` to the name you selected for the package(s), not including the file extension. Depending on the Splunk deployment you have chosen, edit the file `templates/monitoring/`**`splunk`**`/vars.yml` or the file `templates/monitoring/`**`splunk_demo`**`/vars.yml` and set the variable, for example:

```
splunk_architecture_universal_forwarder_package: 'splunkforwarder-7.1.2'
```

As of Splunk version 7.1, the Splunk universal forwarder must be deployed with a password. This password is specified using the variable `splunk_uf_password which is configured in group_vars/all/vault.`

If you are using a standalone Splunk deployment, you must specify the list of indexers using the variable `splunk_architecture_forward_servers` in `group_vars/all/vars`, for example:

```
splunk_architecture_forward_servers:
- splunk-indexer1.cloudra.local:9997
- splunk-indexer2.cloudra.local:9997
```

By default, the indexers are configured in a single load balancing group. This can be changed by editing the file `outputs.conf.j2` in the folder `template/monitoring/splunk/`. For more information on forwarding using Universal Forwarder, see the Splunk documentation at http://docs.splunk.com/Documentation/Forwarder/7.0.2/Forwarder/Configureforwardingwithoutputs.conf.

On your standalone Splunk installation, you need to install the following add-ons and apps.

To monitor **Linux worker nodes**, the **Docker app** should be installed on central Splunk. More info on this Docker app can be found at https://github.com/splunk/docker-itmonitoring and at https://hub.docker.com/r/splunk/universalforwarder/.

To monitor the **Windows worker nodes**, install the **Splunk App for Windows Infrastructure** on central Splunk and its dependencies:

- Splunk App for Windows Infrastructure. The Splunk App for Windows Infrastructure is not compatible with The Splunk Add-on for Windows 5.0 at this time. See https://splunkbase.splunk.com/app/1680/
- Splunk Add-on for Microsoft Windows version 4.8.4 - see https://splunkbase.splunk.com/app/742/
- Splunk Add-On for Microsoft Active Directory version 1.0.0 - see https://splunkbase.splunk.com/app/3207/
- Splunk Add-on for Microsoft Windows DNS version 1.0.1 (if this is not installed on central Splunk, you will see yellow icons on some dashboards with the message `eventtype wineventlog-dns does not exist or is disabled`) - see https://splunkbase.splunk.com/app/3208/
- Splunk Supporting Add-on for Active Directory version 2.1.7 (if this is not installed on central Splunk, you will see yellow icons on some dashboards with the message `eventtype wineventlog-ds does not exist or is disabled`) - see https://splunkbase.splunk.com/app/1151/

If you want to use your own certificates in your standalone Splunk deployment to secure the communications between the indexers and the universal forwarders, see Appendix D.

You can specify advanced Splunk configuration in the following files:

- files/splunk/linux/SPLUNK_HOME
- files/splunk/linux/DOCKER_TAS
- files/splunk/windows/SPLUNK_HOME

These files will be copied as-is to the systems running the universal forwarder.

Configuring syslog in UCP

In order to see some data in the UCP operational dashboard, you need to have UCP send its logs to the VM configured in the [logger] group. For example, for the following `vm_host` file:

```
[logger]
hpe-logger ip_addr='10.60.59.24/16' esxi_host='esx02.cloudra.local'
```

This will configure UCP to send its logs to `hpe-logger.cloudra.local:1514`. You need to select the TCP protocol as shown in Figure 46.

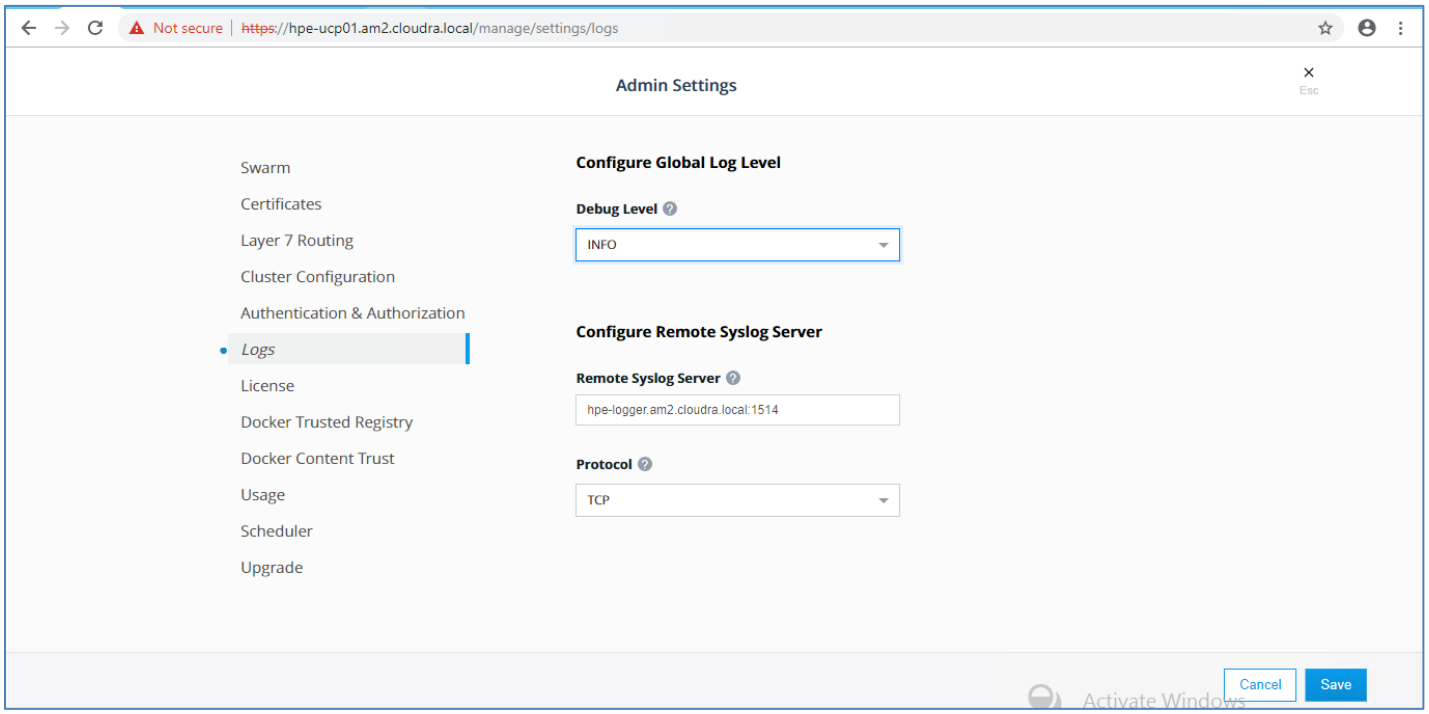

Figure 46. Configure Remote Syslog Server in UCP

Configuring syslog in ESX

This configuration must be done manually for each ESX server. The syslog server should be the server configured in the [`logger`] group in your `hosts` inventory. The protocol should be `tcp` and the port `514` as shown in Figure 47.

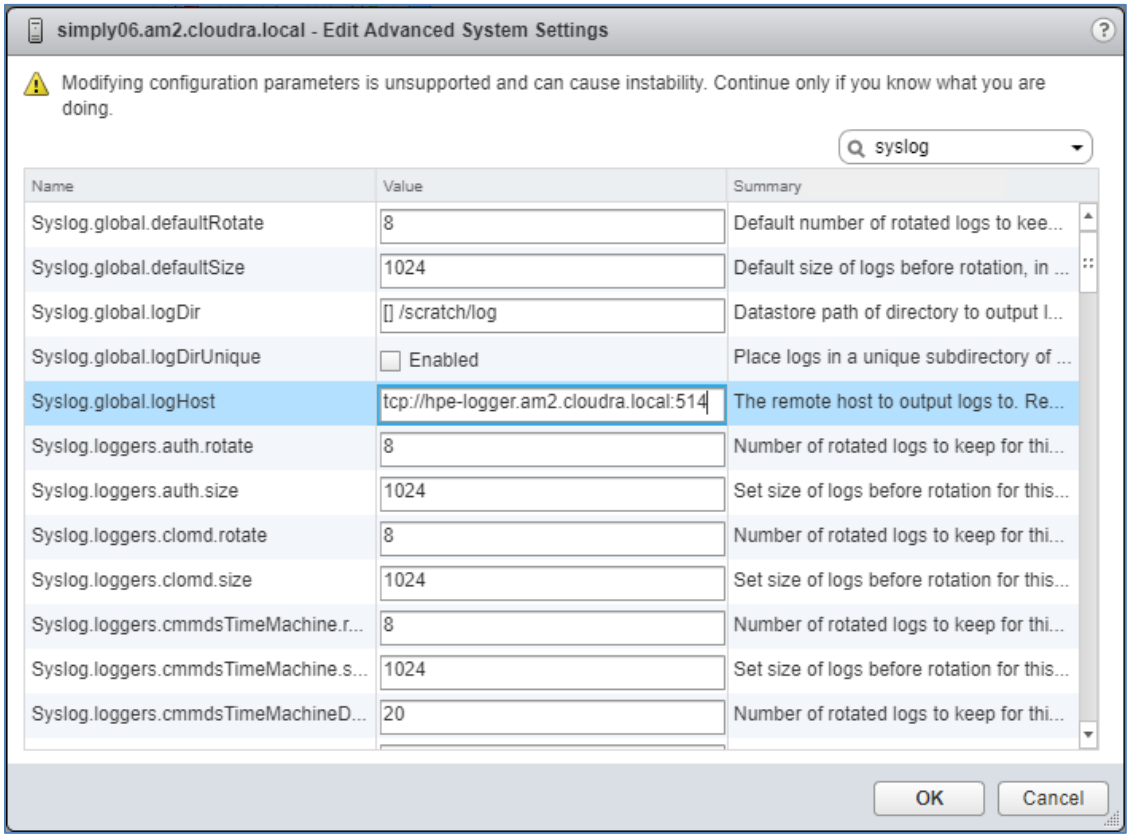

Figure 47**.** Configure Syslog on ESXi Hosts

For more information, see the VMware documentation at https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.security.doc/GUID-9F67DB52-F469-451F-B6C8-DAE8D95976E7.html.

Limitations

- The Dockerized Splunk App has a number of open issues
  - https://github.com/splunk/docker-itmonitoring/issues/19
  - https://github.com/splunk/docker-itmonitoring/issues/20
- The Docker events tab is not working

Accessing Splunk UI

After the installation is complete, the Splunk UI can be reached at http://<fqdn>:8000, where <fqdn> is the FQDN of one of your Linux Docker nodes. Mesh routing does not currently work on Windows so you must use a Linux node to access the UI. For example:

```
http://hpe-ucp01.am2.cloudra.local:8000/
```

The default username and password for Splunk is `admin` / `changeme`.

Use the Docker App to view the Docker overview as shown in Figure 48 and the Docker stats as shown in Figure 49.

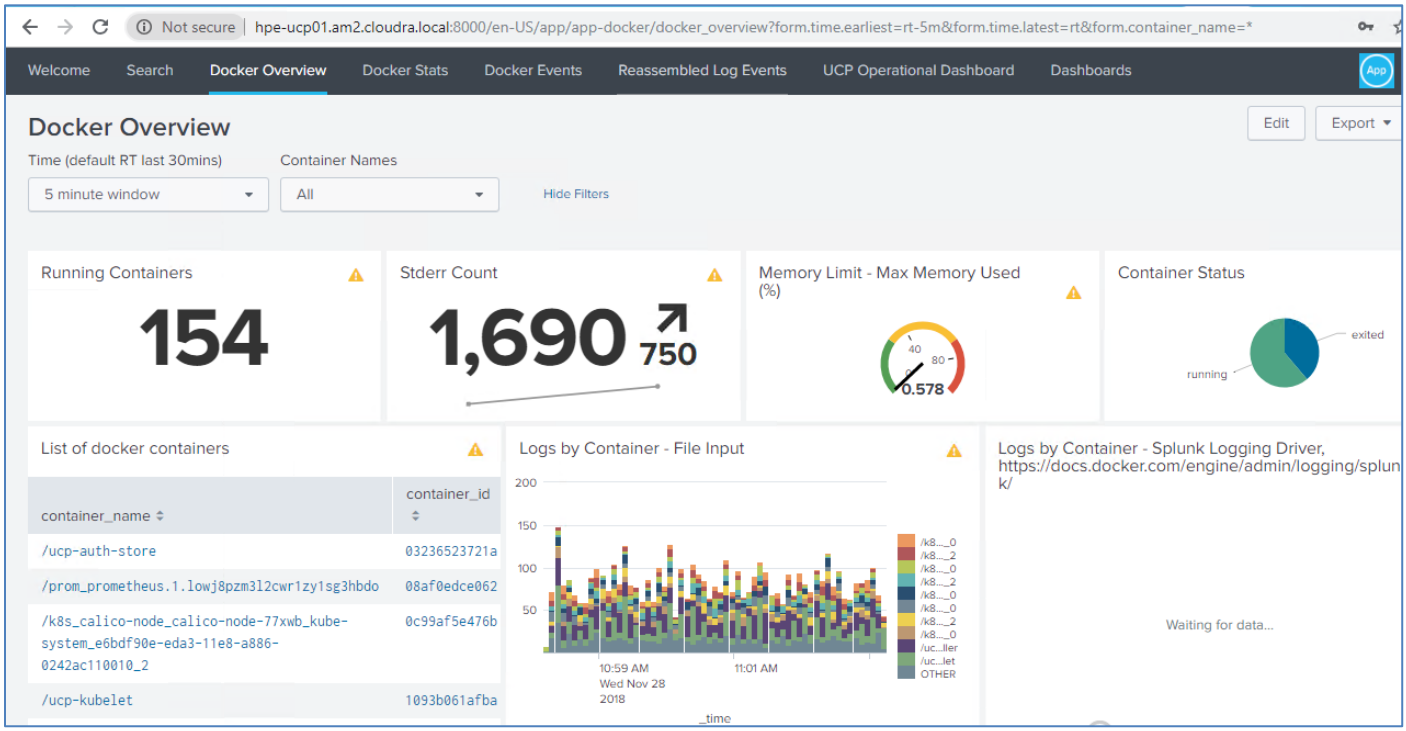

Figure 48. Docker overview

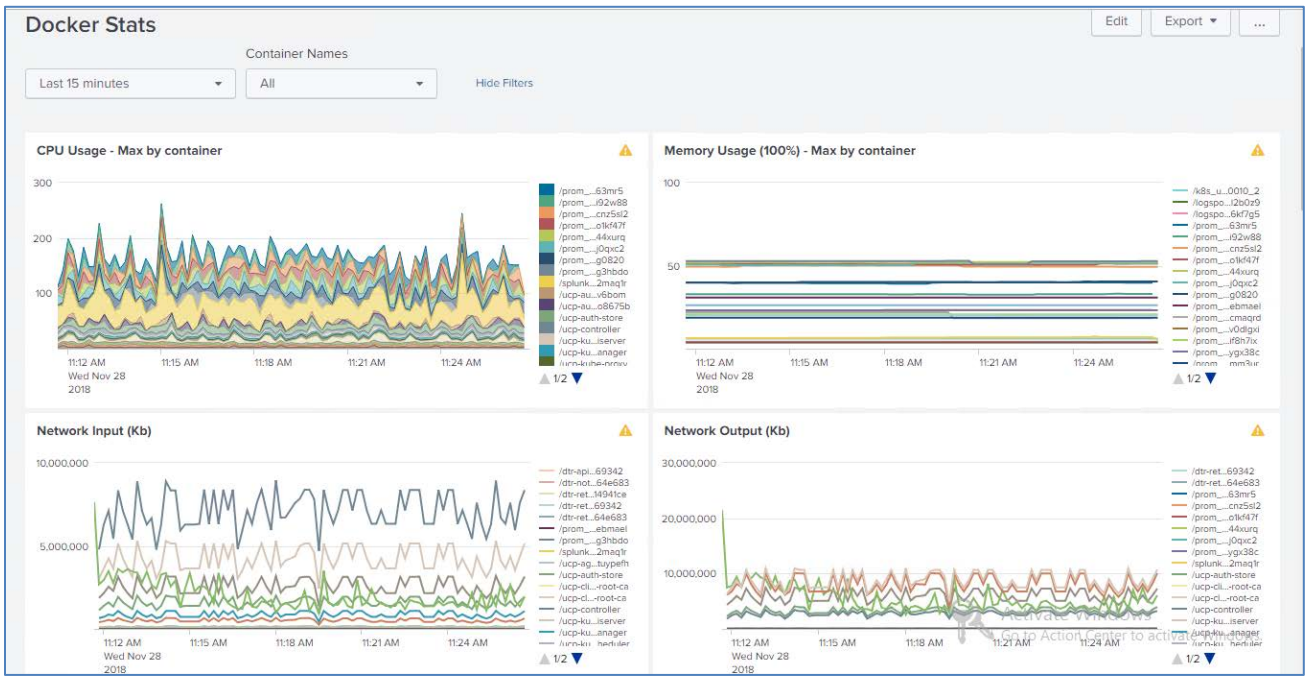

Figure 49. Docker stats

Use the `k8s App` to see the Kubernetes overview as shown in Figure 50 and then access the details for deployments, daemon sets, replica sets, services, etc.

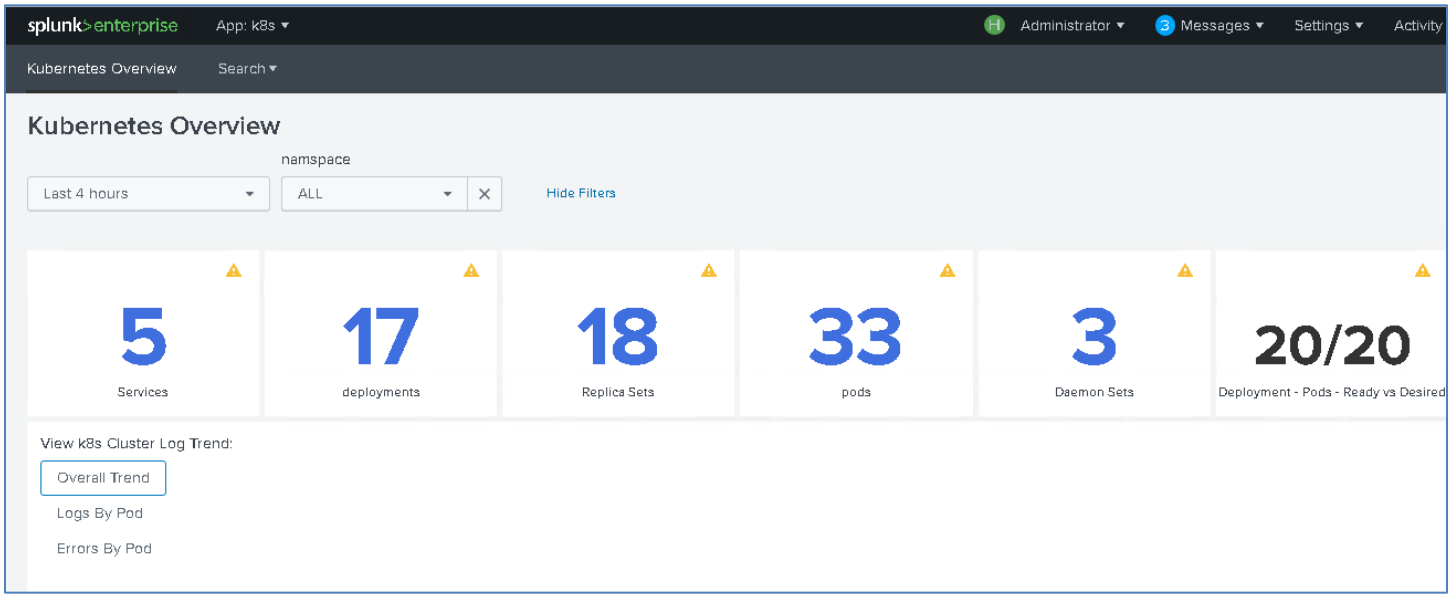

Figure 50. Kubernetes overview

Redeploying Splunk demo

The Splunk demo deployment, whilst fully featured, is severely restricted in the amount of data it can process. Once this limit has been reached, often after running for just one or two days, it is necessary to redeploy the application if you want to continue experimenting with the demo.

Before you redeploy, it is necessary to remove the corresponding Docker stack and delete the associated volumes.

```
# ssh hpe-ucp02

# docker stack rm splunk_demo
Removing service splunk_demo_splunkenterprise
Removing network splunk_demo_default

# docker volume ls | grep splunk
vsphere:latest      splunk_demo_vsplunk-opt-splunk-etc@Docker_HPE
vsphere:latest      splunk_demo_vsplunk-opt-splunk-var@Docker_HPE

# docker volume rm splunk_demo_vsplunk-opt-splunk-etc@Docker_HPE
splunk_demo_vsplunk-opt-splunk-etc@Docker_HPE

# docker volume rm splunk_demo_vsplunk-opt-splunk-var@Docker_HPE
splunk_demo_vsplunk-opt-splunk-var@Docker_HPE
```

Then re-run the playbook on your Ansible node.

```
ansible-playbook -i hosts playbooks/splunk_demo.yml --vault-password-file .vault_pass
```

Deploying Prometheus and Grafana on Kubernetes

Monitoring Kubernetes with Prometheus and Grafana

Monitoring a Kubernetes cluster with Prometheus is a natural choice as Kubernetes components themselves are instrumented with Prometheus metrics, therefore those components simply have to be discovered by Prometheus and most of the cluster is monitored.

The solution uses the Prometheus Operator to deploy Prometheus and Grafana. The playbooks install `kube-state-metrics` and `node-exporter` components, as well as supporting `kubelet` and `apiserver` metrics. Sample dashboards for Grafana are installed to help you monitor your Kubernetes infrastructure.

The Prometheus Operator, shown in Figure 51, makes running Prometheus on top of Kubernetes as easy as possible, while preserving Kubernetes-native configuration options. It introduces additional resources in Kubernetes to declare the desired state and configuration of Prometheus. The `Prometheus` resource declaratively describes the desired state of a Prometheus deployment, while a `ServiceMonitor` describes the set of targets to be monitored by Prometheus.

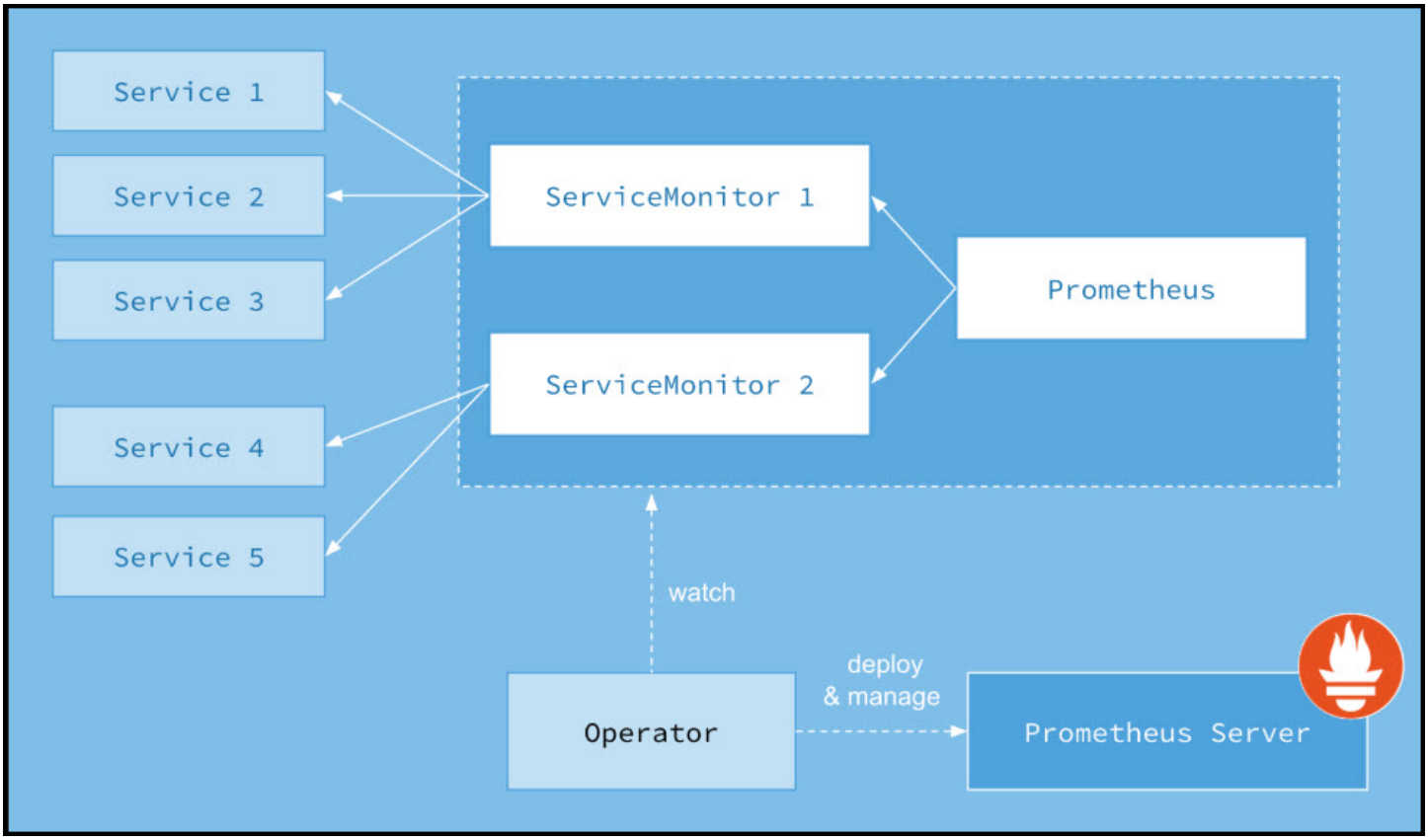

Figure 51. Prometheus Operator

Playbooks for installing Prometheus and Grafana on Kubernetes

Prerequisites

Before you run the playbook to install Prometheus and Grafana on Kubernetes, you need to ensure that you have already downloaded and installed kubectl and set up your client bundle. Two convenience playbooks have been provided to make this process easier.

The playbook `playbooks/install-kubectl.yml` installs a specific version of `kubectl` based on the settings in your `group_vars/all/vars` file.

The playbook `playbooks/kube-prometheus.yml` is used to deploy the Prometheus/Grafana stack on Kubernetes. It is a wrapper for a number of separate playbooks outlined below.

- `playbooks/kube-prometheus/operator.yml`
- `playbooks/kube-prometheus/kube-state-metrics.yml`
- `playbooks/kube-prometheus/node-exporter.yml`
- `playbooks/kube-prometheus/monitors.yml`
- `playbooks/kube-prometheus/prometheus.yml`
- `playbooks/kube-prometheus/grafana.yml`

You can choose not to install certain components, such as `node-exporter` or `kube-state-metrics`, by commenting out the appropriate line in the wrapper playbook.

Prometheus Operator
The Prometheus Operator makes running Prometheus on top of Kubernetes as easy as possible, while preserving Kubernetes-native configuration options. For more information on Prometheus Operator, see https://coreos.com/operators/prometheus/docs/latest/user-guides/getting-started.html.

The playbook `playbooks/kube-prometheus/operator.yml` installs the operator itself.

Kube state metrics
`kube-state-metrics` is a simple service that listens to the Kubernetes API server and generates metrics about the state of the objects. It is not focused on the health of the individual Kubernetes components, but rather on the health of the various objects inside, such as deployments, nodes and pods. For more information on kube-state-metrics, see https://github.com/kubernetes/kube-state-metrics.

The playbook `playbooks/kube-prometheus/kube-state-metrics.yml` installs kube-state-metrics on all UCP, DTR and Kubernetes worker nodes.

Node exporter
The node-exporter provides an overview of cluster node resources including CPU, memory and disk utilization and more. For more information on node-exporter, see https://github.com/prometheus/node_exporter.

The playbook `playbooks/kube-prometheus/node-exporter.yml` installs `node-exporter` as a set of Docker containers on all UCP, DTR and Kubernetes worker nodes. Port `9100` is opened in the firewall on each node where it is installed.

Monitors
While all the other Kubernetes components run on top of Kubernetes itself, `kubelet` and `apiserver` do not, and so they just need service monitors to access these metrics.

The playbook `playbooks/kube-prometheus/monitors.yml` installs Service Monitors for `kubelet` and `apiserver`.

cAdvisor
Support for cAdvisor is built-in to Kubernetes, so cAdvisor metrics will automatically be available within Prometheus, without any other configuration required.

Note
Because Docker EE provides a hosted version of Kubernetes, it is not possible to access metrics for `kube-scheduler` and `kube-controller-manager`.

Prometheus
For convenience, the playbook sets up a NodePort so that the Prometheus UI can be accessed on port `33090`, as shown in the following code extract:

```
# kubectl -n monitoring patch svc prometheus-k8s --type='json' -p
'[{"op":"replace","path":"/spec/type","value":"NodePort"}]'

# kubectl -n monitoring patch svc prometheus-k8s --type='json' -p '[{"op": "add",
"path":"/spec/ports/0/nodePort", "value":33090}]'
```

On a production system, it is likely that you will want to remove this NodePort. The following code segment shows how you can use the `patch` command to remove the NodePort.

```
# kubectl -n monitoring patch svc prometheus-k8s --type='json' -p '[{"op": "remove",
"path":"/spec/ports/0/nodePort"}]'

# kubectl -n monitoring patch svc prometheus-k8s --type='json' -p '[{"op": "remove",
"path":"/spec/type"}]'
```

Grafana

For convenience, the playbook sets up a NodePort so that the Grafana UI can be access on the port `33030`, as shown in the following code extract:

```
# kubectl -n monitoring patch svc grafana --type='json' -p '[{"op":"replace","path":"/spec/type",
"value":"NodePort"}]'

# kubectl -n monitoring patch svc grafana --type='json' -p '[{"op": "add",
"path":"/spec/ports/0/nodePort", "value":33030}]'
```

On a production system, it is likely that you will want to remove this NodePort. The following code segment shows how you can use the patch command to remove the NodePort.

```
# kubectl -n monitoring patch svc grafana --type='json' -p '[{"op": "remove",
"path":"/spec/ports/0/nodePort"}]'

# kubectl -n monitoring patch svc grafana --type='json' -p '[{"op": "remove", "path":"/spec/type"}]'
```

Teardown

The playbook `playbooks/kube-prometheus-teardown.yml` removes the installed Prometheus\Grafana stack.

Prometheus UI

The Prometheus UI is available via your UCP, DTR or Kubernetes worker nodes, using HTTP on port `33090`, for example,

```
http://hpe-ucp01.am2.cloudra.local:33090
```

To see what services are being monitored, access the service discovery page, via `Status -> Service Discovery`, or using the `/service-discovery` endpoint:

```
http://hpe2-ucp01.am2.cloudra.local:33090/service-discovery
```

The monitored services are listed as shown in Figure 52.

Figure 52. Prometheus service discovery

To see the status for the monitored services, access the targets page via `Status -> Targets` or using the endpoint `/targets`.

`http://hpe2-ucp01.am2.cloudra.local:33090/targets`

The status of the various monitors are displayed, as shown in Figure 53.

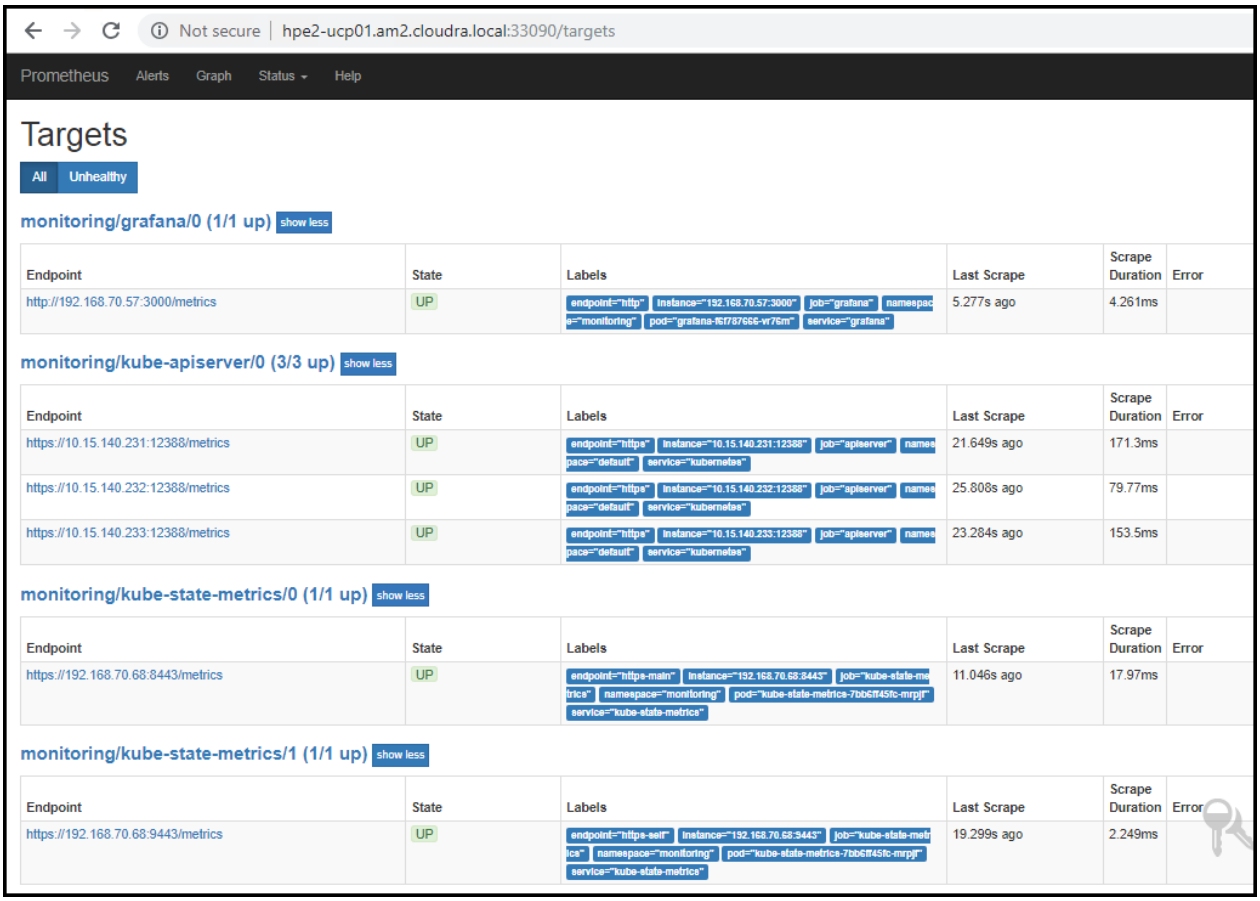

Figure 53. Prometheus targets

To see all the metrics available, click on `Graph` or use the endpoint `/graph`:

`http://hpe2-ucp01.am2.cloudra.local:33090/graph`

Click on the drop-down titled "`- insert metric at cursor -`" to see all the metrics that are available to Prometheus as shown in Figure 54.

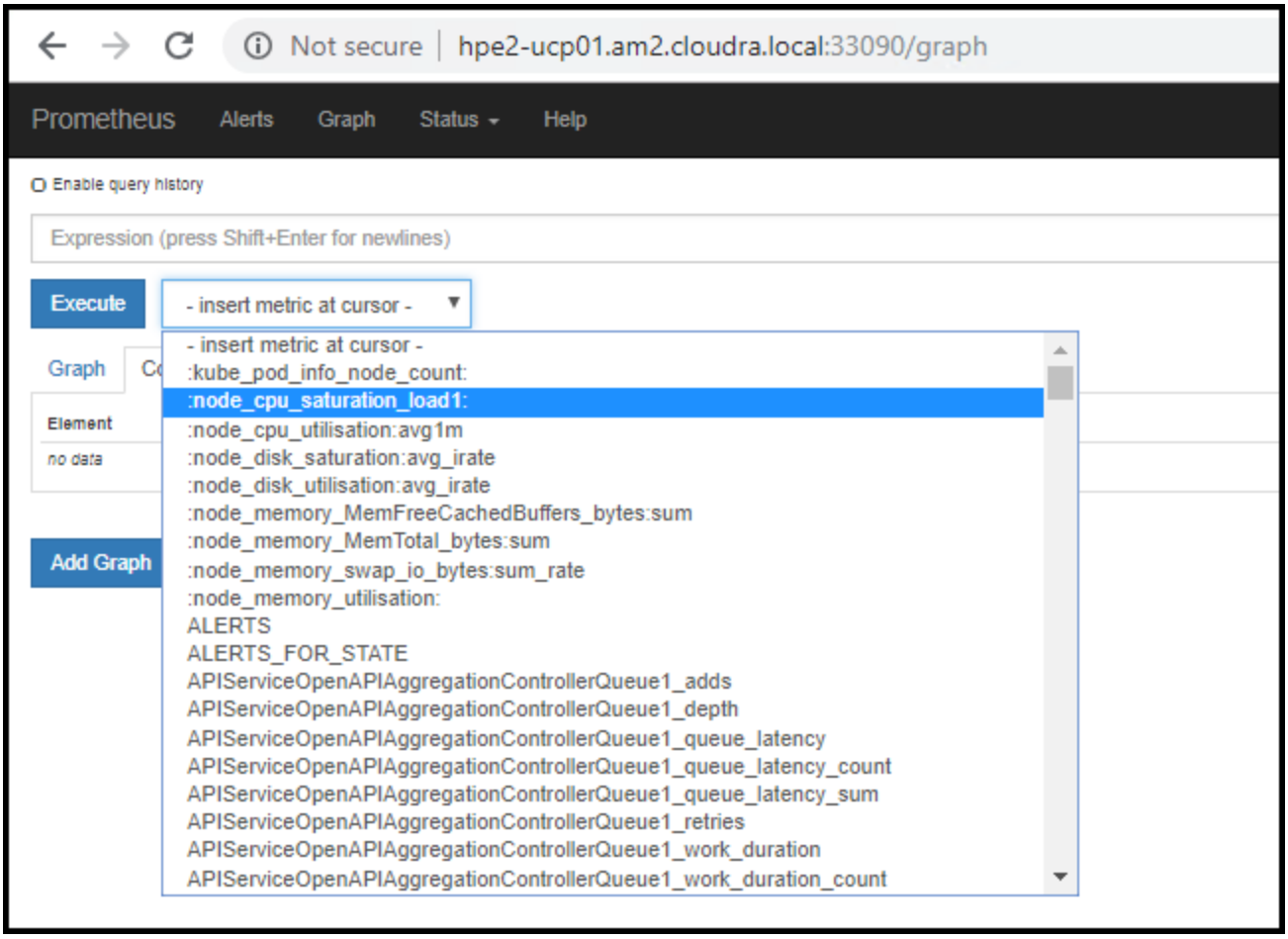

Figure 54. Prometheus metrics

Node Exporter

Metrics specific to the Node Exporter are prefixed with `node_` and include metrics like `node_cpu_seconds_total` and `node_exporter_build_info`. Table 25 below lists some example expressions.

Table 25. Sample Node Exporter metrics

Metric	Meaning
rate(node_cpu_seconds_total{mode="system"}[1m])	The average amount of CPU time spent in system mode, per second, over the last minute (in seconds)
node_filesystem_avail_bytes	The filesystem space available to non-root users (in bytes)
rate(node_network_receive_bytes_total[1m])	The average network traffic received, per second, over the last minute (in bytes)

More information on the use of `node-exporter` metrics is available at https://github.com/prometheus/node_exporter.

cAdvisor

cAdvisor is an open source container resource usage and performance analysis agent. It is purpose-built for containers and supports Docker containers natively. In Kubernetes, cAdvisor is integrated into the Kubelet binary. cAdvisor auto-discovers all containers in the machine and collects CPU, memory, filesystem, and network usage statistics. cAdvisor also provides the overall machine usage by analyzing the 'root' container on the machine.

Kubelet exposes a simple cAdvisor UI for containers on a machine, via the default port `4194`. However, this feature has been marked deprecated in v1.10 and completely removed in v1.12. For more inforation on how upcoming releases will reduce the set of metrics exposed by the `kubelet`, see the relevant issue page at https://github.com/kubernetes/kubernetes/issues/68522.

The Kubelet also starts an internal HTTP server on port 10255 and exposes endpoints including `/metrics` and `/metrics/cadvisor`. As this release of Express Containers uses Kubernetes 1.11, it is able to use this feature. In future releases, it will be necessary to deploy cAdvisor as a DaemonSet for access to the cAdvisor UI.

Table 26 lists some example cAdvisor expressions.

Table 26. Sample cAdvisor metrics

Expression	Description	For
rate(container_cpu_usage_seconds_total{name="redis"}[1m])	The cgroup's CPU usage in the last minute (split up by core)	The `redis` container
container_memory_usage_bytes{name="redis"}	The cgroup's total memory usage (in bytes)	The `redis` container
rate(container_network_transmit_bytes_total[1m])	Bytes transmitted over the network by the container per second in the last minute	All containers
rate(container_network_receive_bytes_total[1m])	Bytes received over the network by the container per second in the last minute	All containers

A full listing of cAdvisor-gathered container metrics exposed to Prometheus can be found in the cAdvisor documentation at https://github.com/google/cadvisor/blob/master/docs/storage/prometheus.md.

Grafana UI

The Grafana UI is available via your UCP, DTR or Kubernetes worker nodes, using HTTP on port `33030`, for example,

```
http://hpe-ucp01.am2.cloudra.local:33030
```

The default username and password for Grafana is `admin`/`admin`. The first time you login, you will be asked to reset the default `admin` password.

A number of dashboards are installed by default. The following figues illustrate some of the dashboard provided.

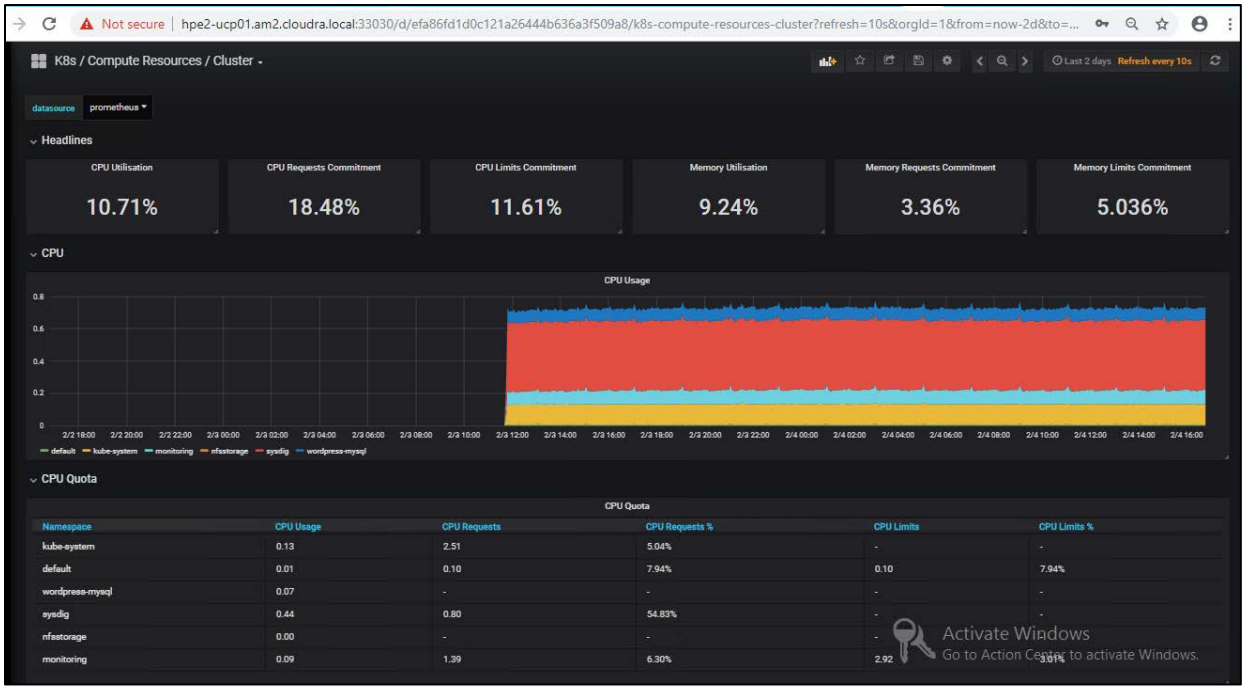

Figure 55. Compute resources dashboard

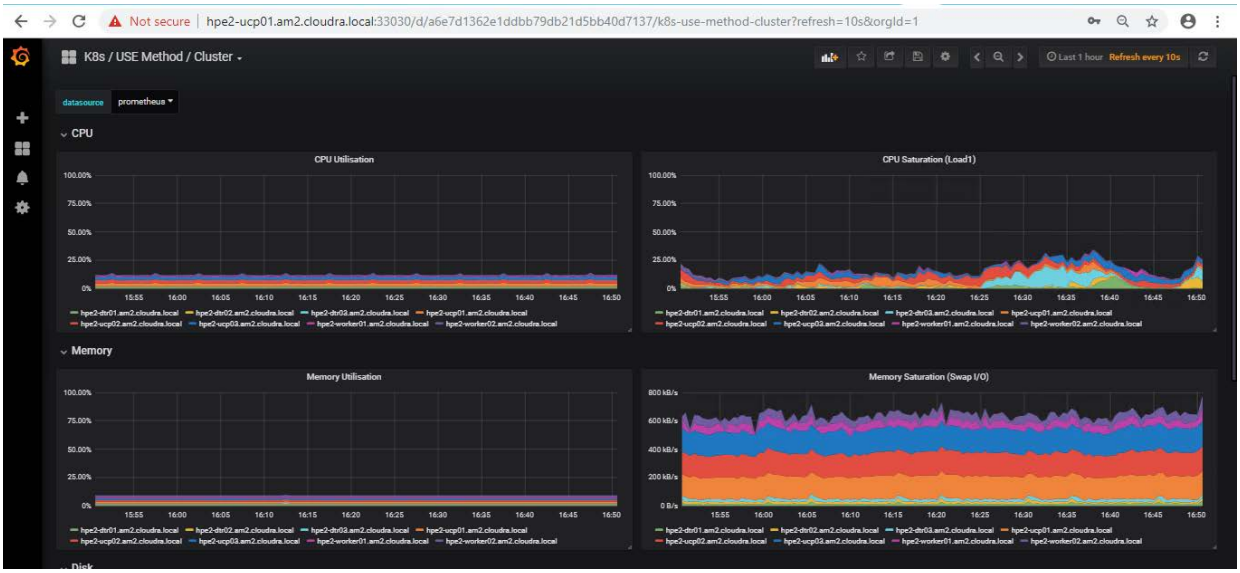

Figure 56. USE method cluster dashboard

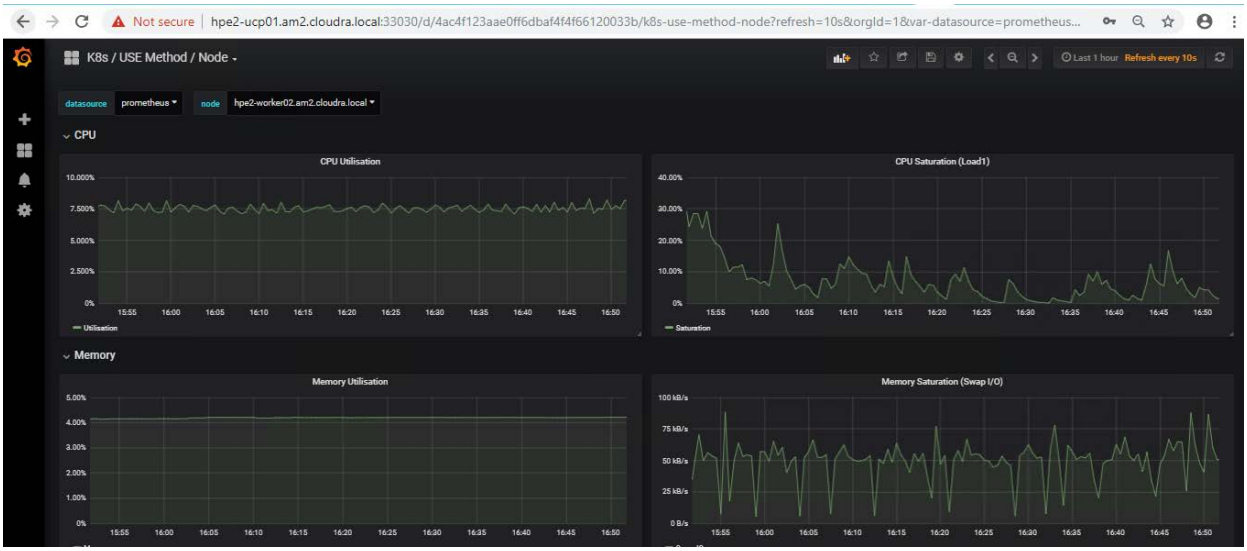

Figure 57. USE method node dashboard

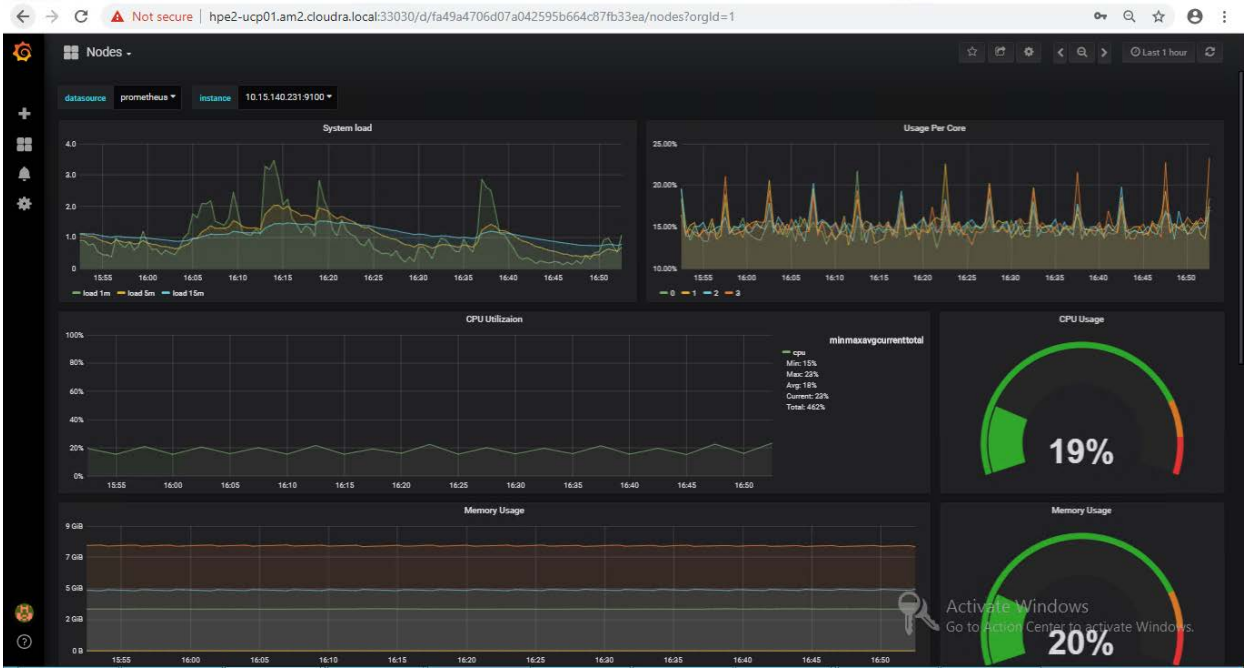

Figure 58. Nodes dashboard

Deploying Prometheus and Grafana on Docker swarm

Monitoring with Prometheus and Grafana

The solution can be configured to enable the use of Prometheus and Grafana for monitoring. In this setup, there is no need for native installs and all the required monitoring software runs in containers, deployed as either services or stacks. The load among the three hosts will be shared as per Figure 59.

Figure 59. Solution architecture: Linux workers with Prometheus and Grafana

The Prometheus and Grafana services are declared in a Docker stack as replicated services with one replica each, so that if they fail, Docker EE will ensure that they are restarted on one of the UCP VMs. `cAdvisor` and `node-exporter` are declared in the same stack as global services, so Docker EE will ensure that there is always one copy of each running on every machine in the cluster.

Note

Prometheus and Grafana functionality is not turned on by default in this solution - see the section on Prometheus and Grafana configuration for more information on how to enable these tools. Additionally, this functionality will not work for the Windows worker nodes in your environment at present.

Playbooks for installing Prometheus and Grafana on Docker swarm

The following playbooks are used to deploy Prometheus and Grafana on Docker RHEL nodes.

- `playbooks/install_logspout.yml` installs and configures **Logspout** on all Docker nodes. Logspout is responsible for sending logs produced by containers running on the Docker nodes to the central logger VM. By default, this playbook is commented out in `site.yml`.
- `playbooks/config_monitoring.yml` configures a monitoring system for the Docker environment based on Grafana, Prometheus, cAdvisor and node-exporter Docker containers. By default, this playbook is commented out in `site.yml`, so if you want to use the solution to automatically deploy a Prometheus/Grafana monitoring system, you must explicitly uncomment both this and the `playbooks/install_logspout.yml` playbook.

Prometheus and Grafana configuration

All monitoring-related variables for Prometheus and Grafana are described in Table 27. The variables determine the versions of various software tools that are used and it is recommended that the values given below are used.

Table 27. Monitoring variables

Variable	Description
cadvisor_version	`v0.28.3`
node_exporter_version	`v1.15.0`
prometheus_version	`V2.3.2`
grafana_version	`5.2.3`
logspout_version	`v3.2.4`
prom_persistent_vol_name	The name of the volume which will be used to store the monitoring data. The volume is created using the vSphere Docker Volume plugin.
prom_persistent_vol_size	The size of the volume which will hold the monitoring data. The exact syntax is dictated by the vSphere Docker Volume plugin. The default value is 10GB.

Accessing Grafana UI

The Grafana UI is available at the UCP VIP, using HTTP on port 3000, for example,

`http://hpe-ucpvip.am2.cloudra.local:3000`

The default username and password for Grafana is `admin/admin`. The first time you login, you will be asked to reset the default `admin` password. Select the Docker Swarm Monitor dashboard that has already been loaded by the playbooks, as shown in Figure 60 and Figure 61.

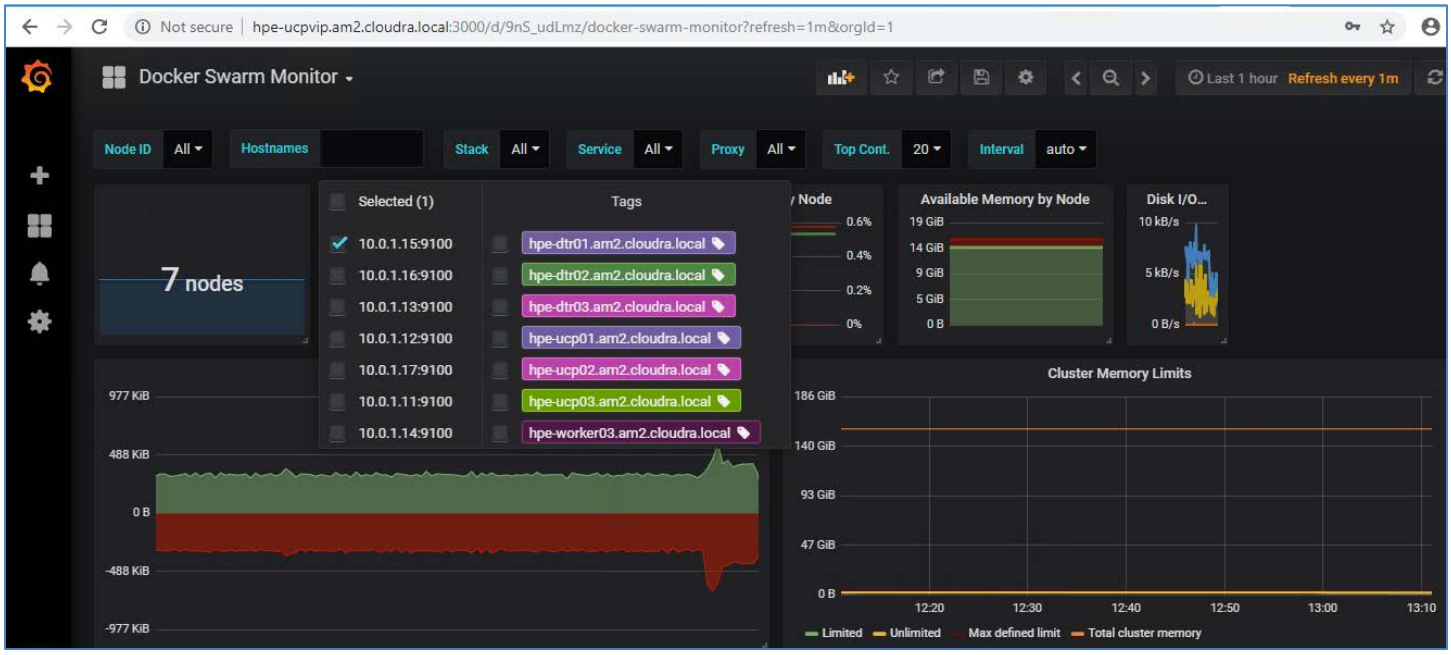

Figure 60. Docker Swarm Monitor dashboard

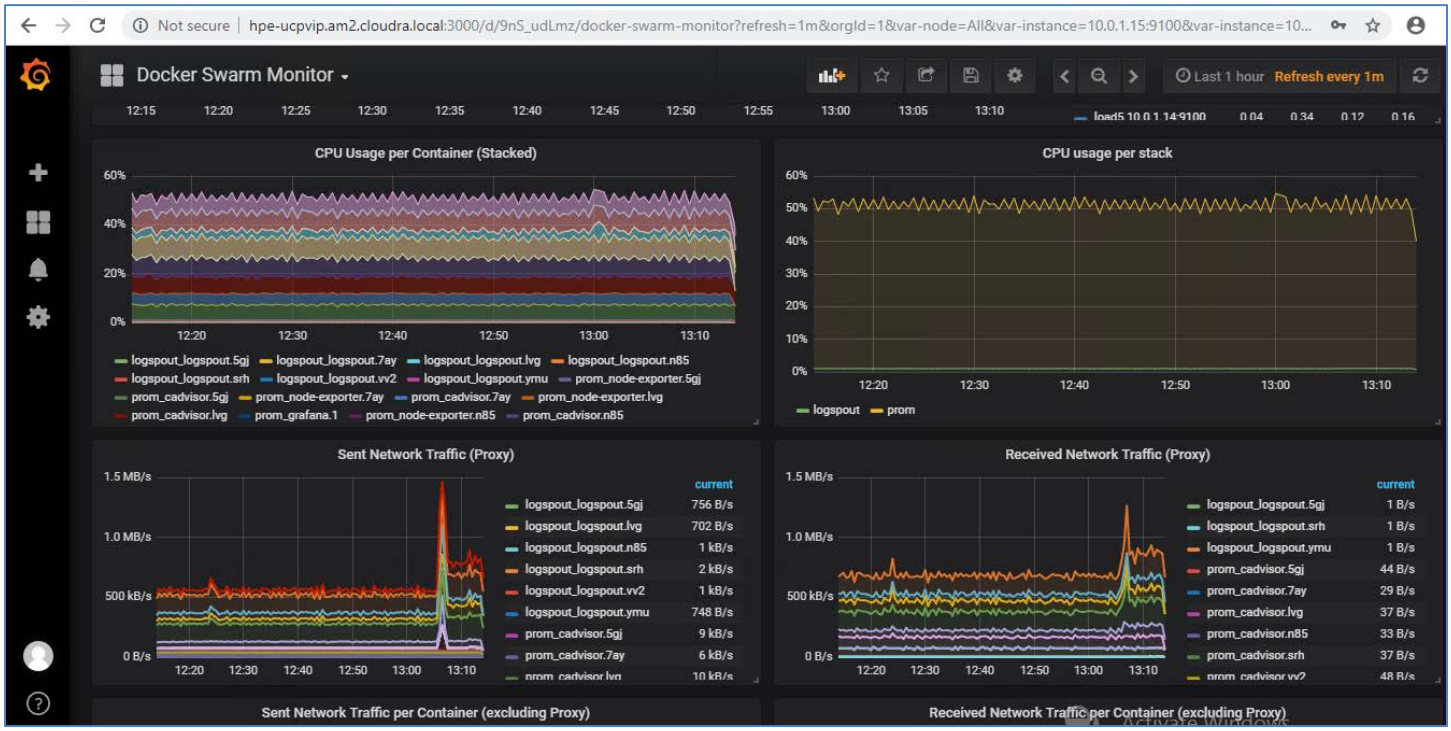

Figure 61. Docker Swarm Monitor dashboard

Backup and restore

This Reference Configuration provides playbooks and scripts to help you back up and restore:

- Docker UCP and DTR
- Docker volumes

Backup and restore UCP and DTR

The playbooks provided in this solution implement the backup and restore procedures as they are described in the Docker documentation at https://docs.docker.com/enterprise/backup/. The solution follows the recommendations in the Docker best practices document at https://success.docker.com/article/backup-restore-best-practices.

Note

It is important that you make copies of the backed up data and that you store the copies in a separate physical location. You must also recognize that the backed up data contains sensitive information such as private keys and so it is important to restrict access to the generated files. However, the playbooks do not backup the sensitive information in your `group_vars/all/vault` file so you should make sure to keep track of the credentials for the UCP Administrator.

Warning

The restore procedures do not restore swarm data. You should follow infrastructure as code (IaC) guidelines and maintain your service, stack and network definitions using source code or configuration management tools. You must also ensure that you safely manage the credentials of your administration accounts, as existing UCP Client bundles will not work when you restore UCP on a new swarm.

Backup UCP and DTR

The playbooks support backing up the swarm, UCP, DTR metadata and DTR images.

Backup configuration variables

Table 28 shows the variables related to backing up UCP and DTR. All these variables are defined in the file **group_vars/all/backups**. All the data that is backed up is streamed over an SSH connection to the backup server. Currently, the playbooks only support the use of the Ansible box as the backup server.

Table 28. Backup variables

Variable	File	Description
backup_server	**group_vars/all/backups**	Currently, the playbooks only support the use of the Ansible box as the backup server.
backup_dest	**group_vars/all/backups**	This variable should point to an existing folder on your Ansible box where the `root` user has write access. All the backups will be stored in this folder. For example, `/root/backups`
backup_passphrase	**group_vars/all/vault**	This variable is used to encrypt the tar file with a passphrase that must be at least 12 characters long.
swarm_offline_backup	**group_vars/all/backups**	This variable is commented out by default. More information on this variable is provided below.

Backing up the swarm

When you back up the swarm, your services and stack definitions are backed up together with the networks definitions. However, Docker volumes or their contents will not be backed up. (If Docker volumes are defined in stacks, they will be re-created when you restore the stacks, but their content will be lost). You can back up the swarm using the playbook named `backup_swarm.yml` which is located in the `playbooks` folder on your Ansible server. The playbook is invoked as follows:

```
# ansible-playbook -i hosts playbooks/backup_swarm.yml
```

This playbook creates two archives in the folder specified by the variable `backup_dest` in `group_vars/all/backups`. By default, the file is named using the following pattern:

```
<backup_dest>/backup_swarm_<vmname>_<timestamp>.tgz
<backup_dest>/backup_swarm_<vmname>_<timestamp>.vars.tgz
```

`<vmname>` is the name of the host (in the inventory) that was used to take the backup, and `<timestamp>` is the time at which the backup was taken. The file with the extension `.vars.tgz` contains information regarding the system that was backed up.

You can override the generated file name by defining the variable **backup_name** on the command line when running the playbook. In the example below:

```
# ansible-playbook -i hosts playbooks/backup_swarm.yml -e backup_name=my_swarm_backup
```

The generated files won't have `<vmname>` or `<timestamp>` appended:

```
<backup_dest>/my_swarm_backup.tgz
<backup_dest>/my_swarm_backup.vars.tgz
```

Warning

Online versus offline backups: By default, the playbook performs online backups. You can take offline backups by setting the variable `swarm_backup_offline` to **"true"**. The playbook will then stop the Docker daemon on the machine used to take the backup (a manager or UCP node). Before it does so, the playbook will verify that enough managers are running in the cluster to maintain the quorum. If this is not the case, the playbook will exit with an error. For more information, see the Docker documentation at https://docs.docker.com/engine/swarm/admin_guide/#recover-from-disaster

Backing up the Universal Control Plane (UCP)

When you backup UCP, you save the data/metadata outlined in Table 29.

Table 29. UCP data backed up

Data	Description
Configurations	The UCP cluster configurations, as shown by `docker config ls`, including Docker EE license and swarm and client CAs
Access control	Permissions for team access to swarm resources, including collections, grants, and roles
Certificates and keys	The certificates, public keys, and private keys that are used for authentication and mutual TLS communication
Metrics data	Monitoring data gathered by UCP
Organizations	Your users, teams, and orgs
Volumes	All UCP named volumes, which include all UCP component certs and data

To make a backup of UCP, use `playbook/backup_ucp.yml` as follows:

```
# ansible-playbook -i vm_host playbooks/backup_ucp.yml
```

This playbook creates two archives in the folder specified by the variable `backup_dest` in `group_vars/all/backups`. By default, the files are named using the following pattern:

```
<backup_dest>/backup_ucp_<ucpid>_<vmname>_<timestamp>.tgz
<backup_dest>/backup_ucp_<ucpid>_<vmname>_<timestamp>.vars.tgz
```

`<ucpid>` is the ID of the UCP instance, `<vmname>` is the name of the host (in the inventory) that was used to take the backup, and `<timestamp>` is the time at which the backup was taken. The file with the extension `.vars.tgz` contains information regarding the system which was backed up.

You can override the generated file name by defining the variable **backup_name** on the command line when running the playbook. In the example below:

```
# ansible-playbook -i hosts playbooks/backup_ucp.yml -e backup_name=my_ucp_backup
```

The generated files won't have `<vmname>` or `<timestamp>` appended:

```
<backup_dest>/my_ucp_backup.tgz
<backup_dest>/my_ucp_backup.vars.tgz
```

Warning

To create a consistent backup, the backup command **temporarily stops the UCP containers running on the node where the backup is being performed**. User resources, such as services, containers, and stacks are not affected by this operation and will continue to operate as expected. Any long-lasting `docker exec`, `docker logs`, `docker events`, or `docker attach` operations on the affected manager node will be disconnected.

For more information on UCP backup, see the Docker documentation at https://docs.docker.com/datacenter/ucp/3.0/guides/admin/backups-and-disaster-recovery/

Backing up the Docker Trusted Registry (DTR)

When you backup DTR, you save the data/metadata outlined in Table 30.

Table 30. DTR data backed up

Data	Backed up?	Description
Configurations	yes	DTR settings
Repository metadata	yes	Metadata like image architecture and size
Access control to repos and images	yes	Data about who has access to which images
Notary data	yes	Signatures and digests for images that are signed
Scan results	yes	Information about vulnerabilities in your images
Certificates and keys	yes	TLS certificates and keys used by DTR
Image content	no	Needs to be backed up separately, depends on DTR configuration
Users, orgs, teams	no	Create a UCP backup to backup this data
Vulnerability database	no	Can be re-downloaded after a restore

To make a backup of DTR metadata, use `playbook/backup_dtr_metadata.yml`

```
# ansible-playbook -i vm_host playbooks/backup_dtr_metadata.yml
```

This playbook creates two archives in the folder specified by the variable `backup_dest` in `group_vars/all/backups`. By default, the file is named using the following pattern:

```
<backup_dest>/backup_dtr_meta_<replica_id>_<vmname>_<timestamp>.tgz
<backup_dest>/backup_dtr_meta_<replica_id>_<vmname>_<timestamp>.vars.tgz
```

`<replica_id>` is the ID of the DTR replica that was backed up, `<vmname>` is the name of the host (in the inventory) that was used to take the backup, and `<timestamp>` is the time at which the backup was taken. The file with the extension `.vars.tgz` contains information regarding the system that was backed up.

You can override the generated file name by defining the variable **backup_name** on the command line when running the playbook. In the example below:

```
# ansible-playbook -i hosts playbooks/backup_dtr_metadata.yml -e backup_name=my_dtr_metadata_backup
```

The generated files won't have `<vmname>` or `<timestamp>` appended:

```
<backup_dest>/my_dtr_metadata_backup.tgz
<backup_dest>/my_dtr_metadata_backup.vars.tgz
```

For more information on DTR backups, see the Docker documentation at https://docs.docker.com/datacenter/dtr/2.5/guides/admin/backups-and-disaster-recovery/

Backing up DTR data (images)

To make a backup of the images that are inventoried in DTR and stored on the NFS server, use `playbooks/backup_dtr_images.yml`

```
# ansible-playbook -i vm_host playbooks/backup_dtr_images.yml
```

This playbook creates two archives in the folder specified by the variable `backup_dest` in `group_vars/all/backups`. By default, the files are named using the following pattern:

```
<backup_dest>/backup_dtr_data_<replica_id>_<vmname>_<timestamp>.tgz
<backup_dest>/backup_dtr_data_<replica_id>_<vmname>_<timestamp>.vars.tgz
```

<replica_id> is the ID of the DTR replica that was backed up, <vmname> is the name of the host (in the inventory) that was used to take the backup, and <timestamp> is the time at which the backup was taken.

You can override the generated file names by defining the variable **backup_name** on the command line when running the playbook, as shown in the example below:

```
# ansible-playbook -i hosts playbooks/backup_dtr_images.yml -e backup_name=my_dtr_data_backup
```

The generated files won't have <vmname> or <timestamp> appended:

```
<backup_dest>/my_dtr_data_backup.tgz
<backup_dest>/my_dtr_data_backup.vars.tgz
```

For more information on DTR backups, see the Docker documentation at https://docs.docker.com/datacenter/dtr/2.5/guides/admin/backups-and-disaster-recovery/

Backing up other metadata, including passwords

The backup playbooks do not backup the sensitive data in your group_vars/all/vault file. The information stored in the .vars.tgz files includes backups of the following files:

- **hosts**, a copy of the hosts file at the time the backup was taken
- **vars**, a copy of your group_vars/all/vars file at the time the backup was taken
- **meta.yml**, a generated file containing information pertaining to the backup

The **meta.yml** file contains the following information:

```
backup_node="<node that took the backup>"
replica_id="<ID of DTR replica if DTR backup>"
backup_source=""
ucp_version="<UCP version if UCP backup>"
dtr_version="<DTR version of DTR backup>"
```

Backup Utility

The script backup.sh can be used to take a backup of the swarm, UCP, DTR metadata and the DTR images in one go. You can pass this script an argument (tag) that will be used to prefix the backup filenames, thereby overriding the default naming. Table 31 shows the file names produced by backup.sh based on the argument passed in the command line.

Table 31. Backup utility

Example	Command line	Generated filenames
Default	./backup.sh	backup_swarm_<vmname>_<timestamp>.tgz, backup_ucp_<ucpid>_<vmname>_<timestamp>.tgz, backup_dtr_meta_<replica_id>_<vmname>_<timestamp>.tgz, backup_dtr_data_<replica_id>_<vmname>_<timestamp>.tgz and the corresponding .vars.tgz files
Custom	./backup.sh my_backup	my_backup_swarm.tgz, my_backup_ucp.tgz, my_backup_dtr_meta.tgz, my_backup_dtr_data.tgz, and the corresponding .vars.tgz files
Date	./backup.sh $(date '+%Y_%m_%d_%H%M%S')	<date>_swarm.tgz, <date>_ucp.tgz, <date>_dtr_meta.tgz, <date>_dtr_data.tgz, and the corresponding .vars.tgz files

In addition, the backup.sh script accepts an optional switch that will let you specify the location of the password file that will be passed to the ansible-playbook commands in the script. This is required if you have encrypted the group_vars/all/vault file. The general syntax for this script is as follows:

```
./backup.sh [ -v <Vault Password File> ] [ tag ]
```

Related playbooks

- `playbooks/backup_swarm.yml` is used to back up the swarm data
- `playbooks/backup_ucp.yml` is used to back up UCP
- `playbooks/backup_dtr_meta.yml` is used to back up DTR metadata
- `playbooks/backup_dtr_images.yml` is used to back up DTR images

Restoring your cluster after a disaster

The playbooks address a disaster recovery scenario where you have lost your entire cluster and all the VMs. Other scenarios and how to handle them are described in the Docker documentation including the following scenarios:

- You have lost one UCP instance but your cluster still has the quorum. The easiest way is to recreate the missing UCP instance from scratch.
- You have lost the quorum in your UCP cluster but there is still one UCP instance running.
- You have lost one instance of DTR but still have a quorum of replicas. The easiest way is to recreate the missing DTR instance from scratch.
- You have lost the quorum of your DTR cluster but still have one DTR instance running.

Before you restore

Step 1. Retrieve the backup files using your chosen backup solution and save them to a folder on your Ansible server. If you have used timestamps in the naming of your backup files, you can use them to determine the chronological order. If you used the `backup.sh` script specifying a date prefix, you can use that to identify the matching set of backup files. You should choose the files in the following reverse chronological order, from the most recent to the oldest file. Make sure you restore both the `*.tgz` and the `*.vars.tgz` files.

1. DTR images backup
2. DTR metadata backup
3. UCP backup
4. Swarm backup

In this example, we will assume a set of backup files stored in `/root/restore` that were created specifying a date prefix. These will have names like `2018_04_17_151734_swarm.tgz`, `2018_04_17_151734_ucp.tgz`, etc and the corresponding `.vars.tgz` files.

Step 2: Retrieve the DTR replica ID, the DTR version and the UCP version

To retrieve the ID of the replica that was backed up, as well as the version of DTR, you need to extract the data from the `.vars.tgz` file associated with the archive of the DTR metadata. You can retrieve this as follows:

```
# tar -Oxf /root/restore/2018_04_17_151734_dtr_meta.vars.tgz meta.yml
backup_node="hpe-dtr01"
replica_id="ad5204e8a4d0"
backup_source=""
ucp_version=""
dtr_version="2.4.3"


# tar -Oxf /root/restore/2018_04_17_151734_ucp.vars.tgz meta.yml
backup_node="hpe-ucp01"
replica_id=""
backup_source=""
ucp_version="3.0.4"
dtr_version=""
```

Take note of the replica ID (`ad5204e8a4d0`), the version of DTR (`2.5.3`) and the version of UCP (`3.0.4`).

Step 3: Populate the `group_vars/all/backups` file

```
backup_swarm: "/root/restore/2018_04_17_151734_swarm.tgz"
backup_ucp: "/root/restore/2018_04_17_151734_ucp.tgz"
backup_dtr_meta: "/root/restore/2018_04_17_151734_dtr_meta.tgz"
backup_dtr_data: "/root/restore/2018_04_17_151734_dtr_data.tgz"
backup_dtr_id: "ad5204e8a4d0"
backup_dest: "/root/backups"
backup_server: <IP of your ansible box>
```

You should populate your `group_vars/all/backups` file as above, with the `backup_dtr_id` variable containing the value you retrieved in the preceding step as `replica_id=`**`"ad5204e8a4d0"`**.

Step 4: Verify that your `group_vars/all/vars` file specifies the correct versions of DTR and UCP.

The playbooks use the versions of UCP and DTR as specified in your `group_vars/all/vars` file to restore your backups. You must ensure that the versions specified in your current `group_vars/all/vars` file correspond to the versions in the backups as determined above.

```
# cat group_vars/all/vars | grep dtr_version
dtr_version: '2.5.3'


# cat group_vars/all/vars | grep ucp_version
ucp_version: '3.0.4'
```

Step 5: Restore UCP admin credentials if required

You must ensure that the UCP admin credentials in your current `group_vars/all/vars` file are those that were in effect when you generated the backup files. If they have changed since then, you must restore the original credentials for the duration of the restore procedure.

Step 6: Restore your inventory (`hosts`)

Your inventory must reflect the environment that was present when the backup files were created. You can find a copy of the inventory as it was when the backup was taken in the `*.vars.tgz` files.

Restore UCP and DTR

Warning

This procedure is aimed at restoring a cluster after a disaster. It assumes you have lost all the VMs in your cluster and want to redeploy using data that you backed up earlier. The solution follows Docker best practice, which means the swarm artifacts are not restored. You will need to restore your Docker volumes and your applications (stacks and services) when this procedure is complete.

---

1. Ensure that you have completed all the preliminary steps as outlined in the section Before you restore.
2. Run the restore playbook

   ```
   ansible-playbook -i hosts restore.yml
   ```

3. Reload your Docker licence, using the Docker UCP UI under `Admin Settings -> Licence` or directly by using the route `/manage/settings/license`.
4. If you are using the image scanning functionality in DTR, you will need to re-download the vulnerability database. For more information, see the Docker documentation here.

You are now ready to restore your Docker volumes and your applications.

Restore DTR metadata and DTR images

Note

This procedure restores DTR metadata and images and assumes you have lost all the DTR VMs in your cluster. It will redeploy using the DTR data that you backed up earlier and will also restore the images if the folder exported by the NFS VM is empty.

1. Ensure that you have completed all the preliminary steps as outlined in the section Before you restore. In this scenario, you need the archives for the DTR metadata and the DTR images.
2. Ensure that all the DTR VMs listed in your inventory are destroyed, using the vSphere Web Client to delete them if required. If you want to restore the DTR images you should also delete the NFS VM.
3. Remove the DTR nodes from the swarm by running the `docker node rm <DTR node>` command on a UCP node for each DTR node in your cluster. The following example shows the sequence of commands to use to remove the DTR nodes:

```
# docker node ls
ID       HOSTNAME                     STATUS              AVAILABILITY
aiz… *   hpe-ucp02.cloudra.local      Ready               Active
gvf…     hpe-dtr01.cloudra.local      Down                Active
ir4…     hpe-ucp03.cloudra.local      Ready               Active
mwf…     hpe-dtr02.cloudra.local      Down                Active
oqy…     hpe-ucp01.cloudra.local      Ready               Active
xqe…     hpe-worker01.cloudra.local   Ready               Active
zdu…     hpe-dtr03.cloudra.local      Down                Active


# docker node rm hpe-dtr01.cloudra.local
hpe-dtr01.cloudra.local
# docker node rm hpe-dtr02.cloudra.local
hpe-dtr02.cloudra.local
# docker node rm hpe-dtr03.cloudra.local
hpe-dtr03.cloudra.local


# docker node ls
ID       HOSTNAME                     STATUS              AVAILABILITY
aiz…     hpe-ucp02.cloudra.local      Ready               Active
ir4…     hpe-ucp03.cloudra.local      Ready               Active
oqy… *   hpe-ucp01.cloudra.local      Ready               Active
xqe…     hpe-worker01.cloudra.local   Ready               Active
```

4. Run the restore script:

```
./restore_dtr.sh
```

5. If you are using the image scanning functionality in DTR, you will need to re-download the vulnerability database. For more information, see the Docker documentation here.

Related playbooks

- `playbooks/restore_swarm.yml` is used to restore the swarm data
- `playbooks/restore_dtr_meta.yml` is used to restore DTR metadata
- `playbooks/restore_dtr_images.yml` is used to restore DTR images

Backup and restore Docker persistent volumes

There are a number of prerequisites that must be fulfilled before you backup and restore your Docker persistent volumes.

- VSphere clusters should have access to a datastore specifically for backups. This is a separate Virtual Volume created on the HPE 3PAR StoreServ and presented to all the hosts in the vSphere cluster.
- Backup software must be available. HPE Recovery Manager Central and HPE 3PAR StoreServ is recommended but other customer backup and restore solutions are acceptable.

A number of restrictions also apply:

- Volumes may not be in use when a volume is cloned. Any container that has the volume attached must be paused prior to creating the clone. The container can be resumed once the clone is complete.
- When Docker volumes need to be restored from backup, the backup datastore needs to be detached from all vSphere cluster servers prior to restoration.

Persistent storage backup solution

Creating the volume

Docker persistent volumes can be created from a worker node using the following command:

```
docker volume create --driver=vsphere --name=MyVolume@MyDatastore -o size=10gb
```

Cloning the volume

Note

Prior to creating a clone of a volume, any containers accessing the volume should be paused or stopped.

Docker volumes can be cloned to a new datastore:

```
docker volume create --driver=vsphere --name=CloneVolumme@DockerBackup -o clone-
from=MyVolume@MyDatastore -o access=read-only
```

Snapshot and back up HPE 3PAR Virtual Volumes with HPE Recovery Manager Central and HPE StoreOnce

HPE Recovery Manager Central (RMC) software integrates HPE 3PAR StoreServ All-Flash arrays with HPE StoreOnce Systems to leverage the performance of snapshots with the protection of backups. RMC uses a direct backup model to orchestrate data protection between the array and the backup system without a backup application. When the first full backup is complete, each subsequent backup is incremental, making it significantly faster than traditional backup methods, particularly for higher volumes of data. Backups to HPE StoreOnce are block-level copies of volumes, de-duplicated to save space. Because RMC snapshots are self-contained, fully independent volumes, they can be restored to any HPE 3PAR array in the event of a disaster. See Figure 62 for an overview of the architecture.

HPE Recovery Manager Central enables you to replicate data from the source storage system (HPE 3PAR StoreServ) to the destination storage system (HPE StoreOnce). The replication is based on point-in-time snapshots.

HPE Recovery Manager Central is installed as a VM on VMware vSphere ESXi. It can be installed on the HPE Synergy platform on a separate (from the Docker Solution) vSphere cluster or external to the Synergy environment as long as the external server has connectivity to the HPE 3PAR StoreServ and HPE StoreOnce. HPE RMC can be installed directly on an ESXi host or can be deployed to a VMware vCenter managed environment. For this solution, the standalone "RMC only" is installed. If HPE RMC is installed in the HPE Synergy environment, iSCSI connection to the HPE 3PAR StoreServ is required.

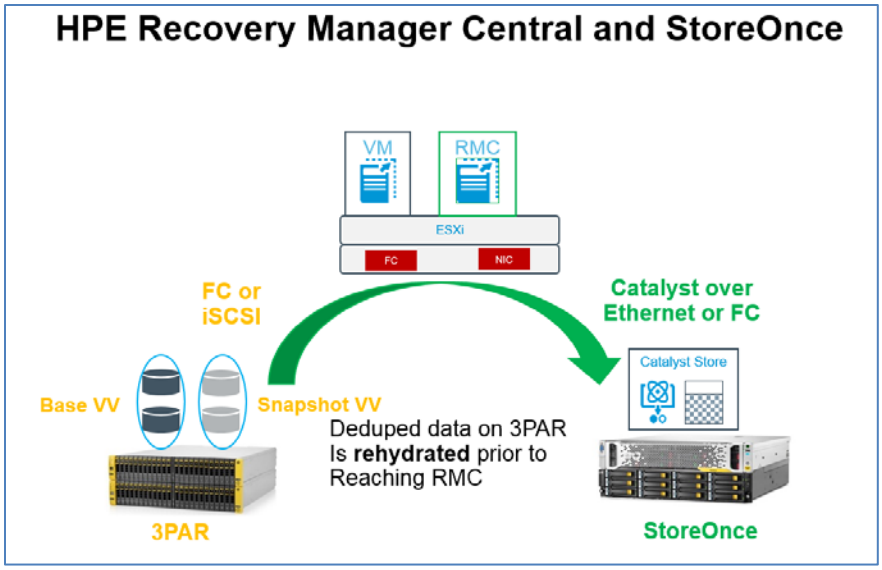

Figure 62. HPE Recovery Manger Central and HPE StoreOnce

- The connectivity between HPE 3PAR StoreServ and HPE RMC for data traffic is over iSCSI.
- The connectivity between HPE StoreOnce and HPE RMC is over CoEthernet (Catalyst OverEthernet)
- The connectivity between HPE RMC, HPE 3PAR StoreServ, and HPE StoreOnce for management traffic is over IP.

Figure 63 illustrates the connectivity between various components.

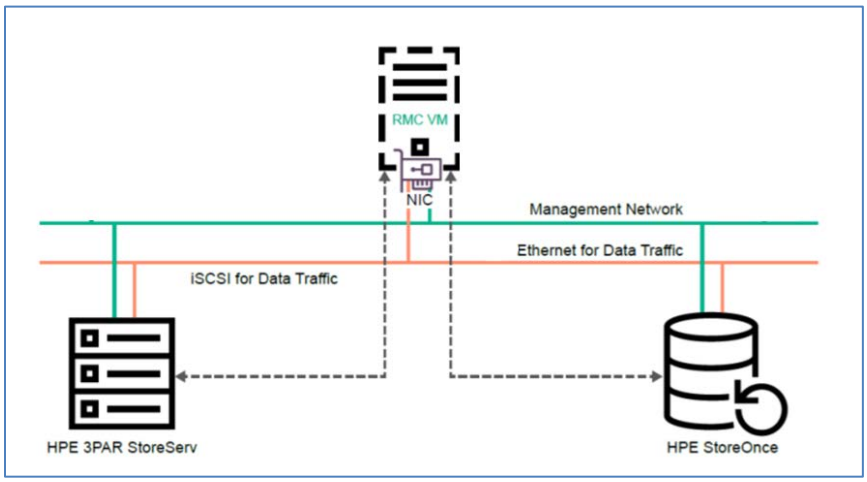

Figure 63. Connectivity

Refer to HPE RMC User guide for detailed instructions on setup and configuration of HPE RMC and HPE StoreOnce. When RMC is installed, it can be configured with the Backup Appliance Persona. The Backup persona allows the RMC to manage snapshots and Express Protect Backups. During installation, RMC configuration should specify Data Protection of RMC Core. The initial configuration of backups can be set up using the Protection Wizard. The Protection Wizard assists with creation of a Recovery Set.

Create a Recovery Set as shown in Figure 64 and select to protect your `DockerBackup` volume. Once you have created your Recovery Set, the next step is to create Protection Jobs. The Auto Protection Job simplifies the initial configuration of policies. The Auto Protection Job will automatically configure the storage, define default backup policies and protection policies and will schedule snapshots or express protect jobs with the created policies.

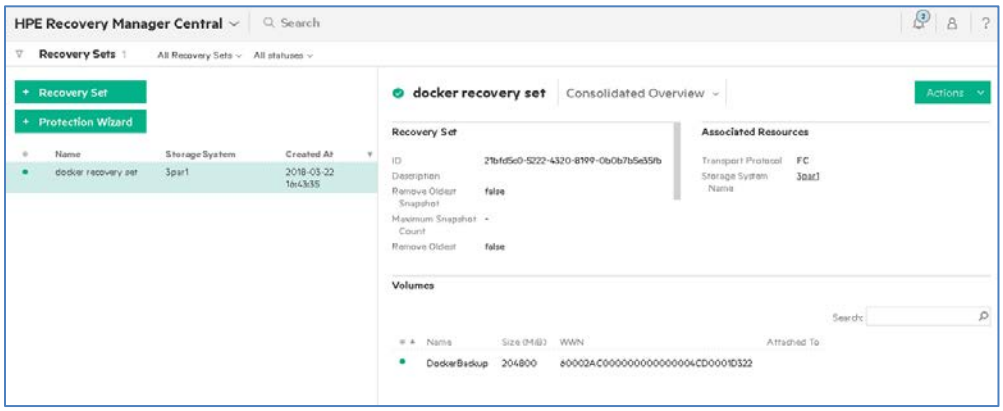

Figure 64. Recovery Set Overview

RMC uses the Express Protect feature, as shown in Figure 65, to enable the backup of the snapshot data from the HPE 3PAR array to the HPE StoreOnce system for deduplication and long-term retention.

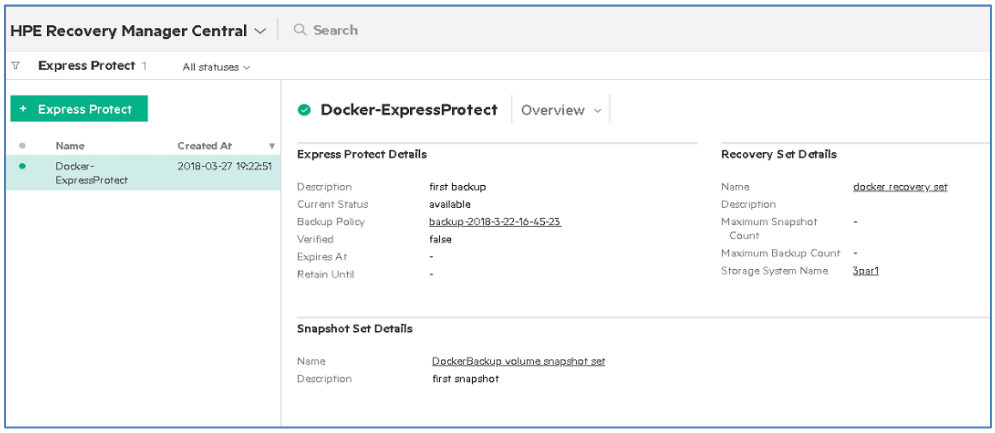

Figure 65. Express Protect

The Express Restore feature restores either snapshots or base volumes.

RMC leverages HPE 3PAR StoreServ SnapDiff technology to create an application-consistent snapshot. Only changed blocks are sent to the HPE StoreOnce system, which minimizes network traffic and saves disk space on the backup system.

Restoring the volume

If a Docker persistent storage volume needs to be restored from backup, the HPE 3PAR volume can be restored either from a snapshot saved on the HPE 3PAR or from a backup on HPE StoreOnce. Stop any applications using the Docker volume. Use the vSphere Web UI to unmount the datastore from the vSphere cluster. Use RMC to detach the HPE 3PAR virtual volumes prior to restoring the backup. The volume can be restored from a Recovery Set restore point as shown in Figure 66. The Express Protect restore point will restore the volume from the HPE StoreOnce system. A Snapshot Set restore point will restore an HPE 3PAR StoreServ snapshot.

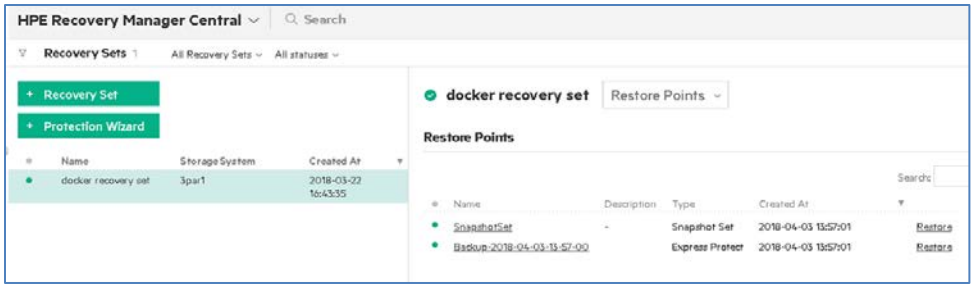

Figure 66. Restore points

Once the HPE 3PAR virtual volume is restored, the volume must be reattached to the vSphere cluster from RMC. After the volume is reattached, the datastore must be mounted. Applications can then access the restored docker volume.

Integrate UCP and DTR backup with HPE RMC and HPE StoreOnce

You can take advantage of HPE Recovery Manager Central and HPE StoreOnce to provide scheduled snapshots and backup protection for the data generated by the backup procedure for Docker UCP and DTR.

1. Create a datastore from the Backup virtual volume you created and present it to all hosts in the vSphere cluster. This backup datastore is used for storing copies of Docker persistent volumes as well as backups of DTR and UCP.
2. The Ansible server is used to create backup and restore files for DTR and UCP on the local hard drive. The backup files should be copied to the `DockerBackup` datastore which can be automatically configured for snapshots and offsite backup.
3. Edit the Ansible server configuration from vCenter. Add a new hard disk and specify the location as the Docker Backup datastore as shown in Figure 67.

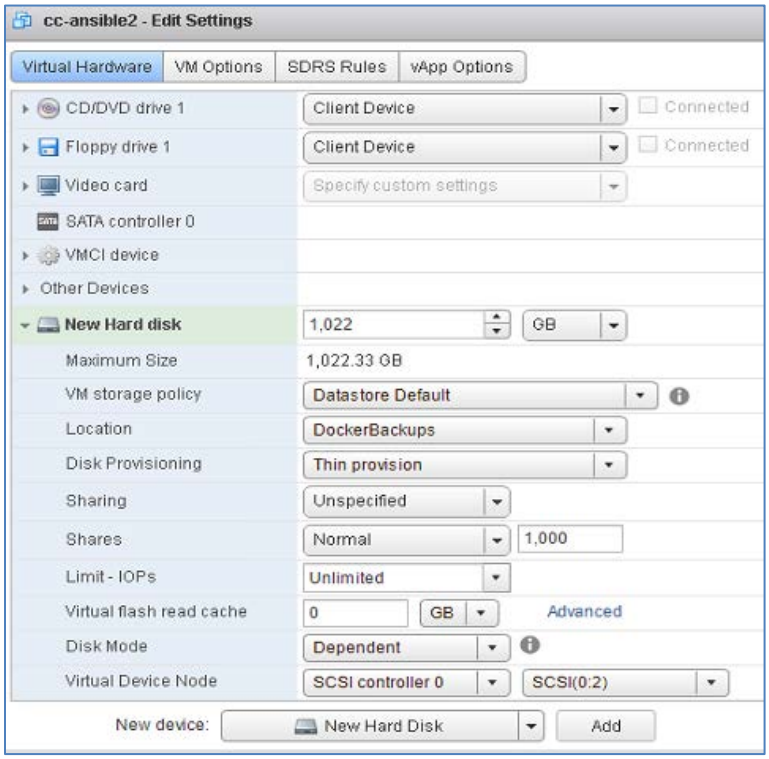

Figure 67. Add new hard disk

4. After the hard disk is added, it is visible from the Linux operating system. From the Ansible server:

   ```
   # ls /dev/sd*
   ```

5. The newly added storage should appear as `/dev/sdb`. Now, make a filesystem, ignoring any warnings:

   ```
   # mkfs -t ext4 /dev/sdb
   ```

6. Create a mount point for the new disk:

   ```
   # mkdir /docker_backup
   ```

7. Edit the `/etc/fstab` file and add the following line:

   ```
   /dev/sdb /docker_backup ext4 defaults 0  0
   ```

8. After saving the change, mount the new volume using:

   ```
   #mount -a
   ```

Each time you backup Docker UCP and DTR using the `backup.sh` script, you should copy the generated files from the `/root/backups` folder to `/docker_backup`. You may wish to add a command to the backup script to automate this process.

The virtual volume used to host the `Docker_Backup` datastore can be scheduled for snapshot and backup protection with HPE Recovery Manager Central and HPE StoreOnce as described in the section Backup and restore Docker persistent volumes. Data backed up to HPE StoreOnce can be restored to the HPE 3PAR StoreServ and attached to the Ansible host for recovery.

Solution lifecycle management

Lifecycle management with respect to this solution refers to the maintenance and management of software and hardware of various components that make up the solution stack. Lifecycle management is required to keep the solution up-to-date and ensure that the latest versions of the software are running to provide optimal performance, security and to fix any existing defects within the product.

In this section, we will cover life cycle management of the different components that are used in this solution. The lifecycle of the following stacks need to be maintained and managed:

- Monitoring Tools (Splunk or Prometheus and Grafana)
- Docker Enterprise Edition Environment
- Virtual Machine Operating systems
- HPE Synergy environment

The general practice and recommendation is to follow a bottom-up approach for updating all components of the environment and making sure the dependencies are met. In our solution, we would start with HPE Synergy and end with the monitoring environment. If all components are not being updated at the same time, the same approach can be followed – updating only the components that require updates while adhering to the interdependencies of each component that is being updated.

HPE Synergy

HPE Synergy Composer powered by HPE OneView provides fast, reliable, and simplified firmware and driver management across many HPE Synergy components. HPE OneView manages firmware to reduce manual interactions and errors, in addition to minimizing downtime. Firmware updates of management appliances and shared infrastructure are non-disruptive to the production workload.

More information is available in the Best Practices for HPE Synergy Firmware and Driver Updates guide at https://support.hpe.com/hpsc/doc/public/display?docId=c05212310.

vSphere Docker Volume Service Plug-in

vSphere Docker Volume service plug-in is part of an open source project by VMware that enables running stateful containers by providing persistent Docker volumes leveraging existing storage technology from VMware. There are two parts to the plug-in, namely, client software and server software (see Table 32). Every version of the plug-in that is released includes both pieces of software and it is imperative that the version number installed on the client side and server side are the same.

When updating the Docker Volume service plug-in, ensure the ESXi version you are running is supported and that the client software is compatible with the operating system.

Table 32. vSphere Docker Volume service components

Order	Component	Dependency (compatibility)	Download/Documentation
1.	Server Software	1. VMware ESXi 2. Docker EE	vSphere Docker Volume Service on GitHub
2.	Client Software	1. VM Operating System 2. Docker EE	

Red Hat Enterprise Linux operating system

This solution is built using Red Hat Enterprise Linux (see Table 33) as the base operating system. When upgrading the operating system on the VMs, first verify that the OS version is compatible with Docker EE by looking at the Docker OS compatibility matrix.

Table 33. Operating system

Order	Component	Dependency (compatibility)	Download/Documentation
1.	Red Hat Enterprise Linux	1. Docker EE 2. vDVS client software plugin	RHEL

Docker EE Environment

Each release of Docker Enterprise Edition contains three technology components – UCP, DTR and the Docker Daemon or Engine. It is imperative that the components belonging to the same version are deployed or upgraded together – see Table 34.

A banner will be displayed on the UI, as shown in Figure 68, when an update is available for UCP or DTR. You can start the upgrade process by clicking on the banner.

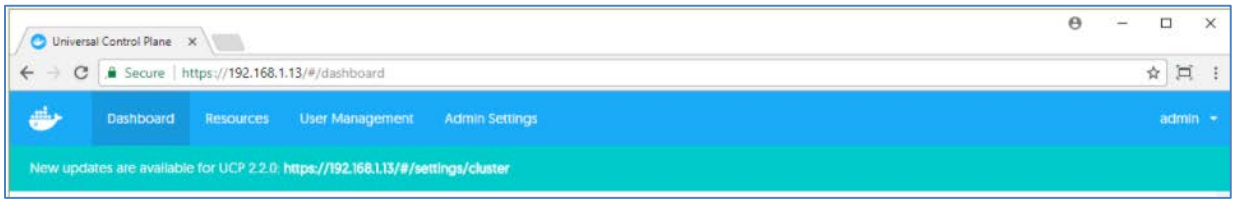

Figure 68. Docker update notification

Table 34. Docker EE components

Order	Component	Dependency (compatibility)	Download/Documentation
1.	Docker Daemon/Engine	1. VM Operating System 2. vDVS plugin 3. Prometheus and Grafana	Docker Lifecycle Maintenance Docker Compatibility Matrix
2.	Universal Control Plane		
3.	Docker Trusted Registry		

Monitoring Tools

To learn more about upgrading Splunk, see the relevant documentation at How to upgrade Splunk Enterprise.

The Sysdig agent runs as a container and the latest version is pulled from the Docker hub on first installation. Re-run the `install_sysdig.yml` playbook to update to the newest version if required.

Prometheus and Grafana monitoring tools (see Table 35) run as containers within the Docker environment. Newer versions of these tools can be deployed by pulling the Docker images from Docker Hub. Verify that the version of Prometheus that is being used is compatible with the version of Docker EE.

Table 35. Monitoring tools: Prometheus and Grafana

Order	Component	Dependency (compatibility)	Download/Documentation
1.	Prometheus	1. Grafana 2. Docker EE	1. Prometheus Images on Docker Hub 2. Upgrading Grafana
2.	Grafana	1. Prometheus 2. Docker EE	

Summary

This document has described how to architect and deploy an HPE Enterprise Containers as a Service with Docker EE solution, using Ansible playbooks to quickly install and deploy a production-ready container environment. This deployment includes a highly available container cluster with backup services and persistent data support. This solution is ideal for customers looking to run containers on VMs to take advantage of the resource efficient usage of virtual machines for Docker containers, and having the ability to run legacy and new container applications side-by-side. Customers deploying Docker containers on a large scale, on Linux and Microsoft Windows, should consider HPE Synergy as the deployment infrastructure.

Appendix A: Software Licenses

Licenses are required for the following software components:

- VMware
- Red Hat Linux
- Microsoft Windows Server
- Docker EE
- Splunk (optional software)
- Sysdig (optional software)

Appendix B: Using customer supplied certificates for UCP and DTR

Table 36 lists the variables used when configuring customer supplied certificates for UCP and DTR.

Table 36. Customer certs variables

Variable	File	Description
ucp_certs_dir	group_vars/all/vars	If **ucp_certs_dir** is not defined, UCP is installed with self-signed certificates and DTR is installed with the `--ucp-insecure-tls` switch If **ucp_certs_dir** is defined, this is a folder on the Ansible machine that must contain 3 files: `ca.pem`, the root CA certificate in PEM format `cert.pem`, the server certificate optionally followed by intermediate CAs `key.pem`, the private key that comes with the `cert.pem` certificates
dtr_certs_dir	group_vars/all/vars	If **dtr_certs_dir** is not defined, DTR is installed with self-signed certificates If **dtr_certs_dir** is defined, this is a folder on the Ansible machine that must contain 3 files: `ca.pem`, the root CA certificate in PEM format `cert.pem`, the server certificate optionally followed by intermediate CAs `key.pem`, the private key that comes with the `cert.pem` certificates

Note

The installation will fail if the `ca.pem`, `cert.pem` and `key.pem` files cannot be found in the folders designated by `dtr_certs_dir` and `ucp_certs_dir` or if they don't constitute valid certificates.

The certificates should specify the names of the FQDNs of the load balancer and the FQDNs of the VMs themselves. This applies to both the UCP server certificate and the DTR server certificate.

Generating and testing certificates

In the example described here we have a root CA named `Example root CA` and an intermediate CA named `Intermediate CA valid 3 years`. The intermediate CA signs the server certificates for UCP and DTR.

Below is the start of the output displayed by running the `openssl x509` utility against the `ca.pem` file (the root CA certificate).

```
[root@ansible ucp_certs]# openssl x509 -text -noout -in ca.pem|head -14
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
```

```
            0d:07:ca:ea:00:37:77:6e:25:e0:18:3e:0e:db:80:0f:11:cb:1b:3f
    Signature Algorithm: sha256WithRSAEncryption
        Issuer: CN=Example Root CA
        Validity
            Not Before: Apr 24 20:12:01 2018 GMT
            Not After : Apr 21 20:12:30 2028 GMT
        Subject: CN=Example Root CA
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (4096 bit)
```

Here is an excerpt from the example `ca.pem` file:

```
-----BEGIN CERTIFICATE-----
MIIFJTCCAw2gAwIBAgIUDQfK6gA3d24l4Bg+DtuADxHLGz8wDQYJKoZIhvcNAQEL
BQAwGjEYMBYGA1UEAxMPRXhhbXBsZSBSb290IENBMB4XDTE4MDQyNDIwMTIwMVoX
...
...
uXzYbCtU6Jt9B3fayAeWWswQv+jQSzuuA3reOM1x838iIZWDx93f4yLJWLJl7xsY
btvKBmqKDCsAqsQLFLnNj/JyYq4e9a6Xxcyn9FXNpzuEsfjfNGHn+csY+w3f987T
MNviy376xZbyAc1CV5kgmnZzjU5bDkgT8Q==
-----END CERTIFICATE-----
```

The `cert.pem` file should contain the server certificate itself, followed by your intermediate CA's certificate. The following example shows how to extract the intermediate CA certificate from `cert.pem` and to save it to a file named `intca.pem`. Using the `openssl x509` utility, you can display the content of the `intca.pem` file in human readable form. This certificate was signed by the example CA above (`Issuer = 'Example Root CA'`).

```
[root@ansible ucp_certs]# openssl x509 -text -noout -in intca.pem|head -14
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            6b:1e:0c:86:20:cf:f0:88:d2:52:0d:5d:b9:56:fa:91:87:a0:49:18
    Signature Algorithm: sha256WithRSAEncryption
        Issuer: CN=Example Root CA
        Validity
            Not Before: Apr 24 20:12:09 2018 GMT
            Not After : Apr 23 20:12:39 2021 GMT
        Subject: CN=Intermediate CA valid 3 years
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (4096 bit)
```

Here is an excerpt from the `incta.pem` file showing the example Intermediate CA certificate:

```
-----BEGIN CERTIFICATE-----
MIIFcjCCA1qgAwIBAgIUax4MhiDP8IjSUg1duVb6kYegSRgwDQYJKoZIhvcNAQEL
BQAwGjEYMBYGA1UEAxMPRXhhbXBsZSBSb290IENBMB4XDTE4MDQyNDIwMTIwOVoX
...
...
o2tL5nwR7ROiAr/kk9MIRzWrLNbc4cYth7jEjspU9dBqsXgsTozzWlwqI9ybZwvL
Ni1JnZandVlyQdoOaB2M/1DNFfKvwW3JeArKvDA9j95n/BWFTjoZ+YOz9pYit6T7
1GCGu3be
-----END CERTIFICATE-----
```

The `openssl x509` utility will only decrypt the first certificate found in `cert.pem`, so you don't need to extract the server certificate from `cert.pem`. In this example, the server certificate is signed by the intermediate CA above. Note the `Subject Alternate Names: hpe-ucp.cloudra.local` is the FQDN of the UCP load balancer, and the other names are those of the UCP instances (`hpe-ucp01.cloudra.local`, `hpe-ucp02.clodura.local`, `hpe-ucp03.cloudra.local`).

```
[root@ansible ucp_certs]# openssl x509 -text -noout -in server.pem
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            25:d9:f8:1d:9b:1d:23:f1:21:56:54:f2:43:cc:4f:0e:73:22:be:ec
    Signature Algorithm: sha256WithRSAEncryption
        Issuer: CN=Intermediate CA valid 3 years
        Validity
            Not Before: Apr 24 20:17:30 2018 GMT
            Not After : Apr 24 20:18:00 2019 GMT
        Subject: O=HPE, OU=CloudRA Team, CN=hpe-ucp.cloudra.local
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                CA Issuers - URI:http://localhost:8200/v1/intca
      { portions removed }

            X509v3 Subject Alternative Name:
                DNS:hpe-ucp.cloudra.local, DNS:hpe-ucp01.cloudra.local, DNS:hpe-
ucp02.cloudra.local, DNS:hpe-ucp03.cloudra.local
```

The following excerpts from `cert.pem` show the first certificate which is the server certificate itself and the second certificate which is the intermediate CA's certificate.

```
-----BEGIN CERTIFICATE-----
MIIFGTCCAwGgAwIBAgIUJdn4HZsdI/EhVlTyQ8xPDnMivuwwDQYJKoZIhvcNAQEL
BQAwKDEmMCQGA1UEAxMdSW50ZXJtZWRpYXRlIENBIHZhbGlkIDMgeWVhcnMwHhcN
...
...
sOR4I3Qnc5OoNISng5l7wW1d4RMMwmXQhG1H5QKAUjHfJXH4bNtIzKxw/zGTVr4Z
llYKbEwJcgAvvfkn+w==
-----END CERTIFICATE-----
```

```
-----BEGIN CERTIFICATE-----
MIIFcjCCA1qgAwIBAgIUax4MhiDP8IjSUg1duVb6kYegSRgwDQYJKoZIhvcNAQEL
BQAwGjEYMBYGA1UEAxMPRXhhbXBsZSBSb29IIENBMB4XDTE4MDQyNDIwMTIwOVoX
...
...
Ni1JnZandVlyQdoOaB2M/1DNFfKvwW3JeArKvDA9j95n/BWFTjoZ+YOz9pYit6T7
1GCGu3be
-----END CERTIFICATE-----
```

Finally, here is an excerpt from key.pem, the private key that goes with the server certificate.

```
-----BEGIN RSA PRIVATE KEY-----
MIIEpQIBAAKCAQEA5rmmb52ufE8Oa3cXhY2HSRZNazb7/fipXY1rZ+U5+rJv9BN5
d/X3NTroSE8/PvoS/maGkHCnURGNqbu/G2umKN/tm/eSpDY861YnGWxj+bcOgtiU
...
...
```

```
AOSGidSMk3hFX1Iaftgx4EUGbrzZO7I8M5RO64U1aMFNFyj4XghJ2mZTdNelwNBw
pr/fYulyi5lYPalQHYH3OyvNqQQ3arEbTbZp8hEyYOgxtZRXmmaoqOY=
-----END RSA PRIVATE KEY-----
```

Verify your certificates

The playbooks do not verify the validity of the certificate files you supply so you should verify them manually before you start your deployment.

Verify that the private and the server cert match

On the Ansible box, run the following commands:

```
ckcert=$(openssl x509 -noout -modulus -in cert.pem | openssl md5)
ckkey=$(openssl rsa -noout -modulus -in key.pem| openssl md5)
if [ "$ckkey" == "$ckcert" ] ; then echo "Private key and Certificate match" ; else echo "STOP! Privat
e Key and Certificate don't match" ; fi
```

Verify that the server certificate was signed by the CA

Extract all but the first certificate from `cert.pem` (i.e. extract the certs for the intermediate CA authorities) into the file `int.pem`

```
sed -e '1,/-----END CERTIFICATE-----/d' cert.pem >intca.pem
```

Combine `intca.pem` and `ca.pem` to form `cachain.pem`:

```
cat intca.pem ca.pem > cachain.pem
```

Finally, verify that `cert.pem` was signed by the CA or by an intermediate CA:

```
openssl verify -verbose -CAfile cachain.pem  cert.pem
```

A successful check will generate output similar to:

```
[root@ansible ucp_certs]# cat intca.pem ca.pem > cachain.pem
[root@ansible ucp_certs]# openssl verify -verbose -CAfile cachain.pem  cert.pem
cert.pem: OK
```

An unsuccessful check will generate output similar to:

```
[root@ansible ucp_certs]# openssl verify -verbose -CAfile cachain.pem  certsignedbyanotherca.pem
certsignedbyanotherca.pem: O = HPE, OU = CloudRA Team, CN = hpe-ucp.cloudra.local
error 20 at 0 depth lookup:unable to get local issuer certificate
```

Appendix C: Enabling SSL between the universal forwarders and the Splunk indexers using your certificates

The procedure for enabling SSL between the universal forwarders and the Splunk indexers using your certificates is described below. In summary, the following steps are required:

1. Set the variable `splunk_ssl` to `yes` in `group_vars/all/vars`
2. Put your root CA certificate and your server certificate files in `/root/Docker-Synergy/files/splunk/linux/SPLUNK_HOME/etc/mycerts`
3. Uncomment the `[sslConfig]` stanza in the file `/files/splunk/linux/SPLUNK_HOME/etc/system/local/server.conf`

Limitations

SSL only works with Linux worker nodes. The Universal Forwarders verify that the indexers they connect to have a certificate signed by the configured root CA and that the Common Name in the certificate presented by the indexer matches its FQDN as listed by the variable `splunk_architecture_forward_servers`.

Prerequisites

Configure your indexers to use SSL on port 9998. The following is an example `inputs.conf` file located in `$SPLUNK_HOME/etc/system/local` that enables SSL on port 9998 and configures the certificate file for use by the indexer itself, in this instance `/opt/splunk/etc/mycerts/indexer.pem`.

```
[splunktcp-ssl://9998]
disabled=0
connection_host = ip

[SSL]
serverCert=/opt/splunk/etc/mycerts/indexer.pem
#requireClientCert = true
#sslAltNameToCheck = forwarder,forwarder.cloudca.local

[tcp://1514]
connection_host = dns
sourcetype = ucp
```

For more information, see the documentation at https://docs.splunk.com/Documentation/Splunk/7.1.2/Security/ConfigureSplunkforwardingtousesignedcertificates. In addition, you can see how to create your own certificates and the content of the file designated with `serverCert` at http://docs.splunk.com/Documentation/Splunk/7.1.2/Security/Howtoself-signcertificates.

In this instance, the folder `mycerts` was created under `/opt/splunk/etc` and the file `indexer.pem` was copied to this folder.

Indexers are configured with the Root CA cert used to sign all certificates. This can be achieved by editing the file `server.conf` in `$SPLUNK_HOME/etc/system/local` on your indexer(s). The following code block shows the relevant portion of this file where `sssRootCaPath` is pointing to the root CA certificate.

```
[sslConfig]
sslRootCAPath = /opt/splunk/etc/mycerts/ca.pem
```

Note

In order to be able to download and install additional applications, you may want to append the file `$SPLUNK_HOME/auth/appsCA.pem` to your `ca.pem` file. If you don't do this, the Splunk UI will make this suggestion when you attempt to `Find more apps`.

Splunk should be restarted on the indexers if you had to make these changes (see the Splunk documentation for more information).

Before you deploy

Generate the forwarder certificate and name it `forwarder.pem`. Make sure that you copy the root CA certificate to `ca.pem`

1. Copy both the `ca.pem` and the `forwarder.pem` files to `files/splunk/linux/SPLUNK_HOME/etc/mycerts/` (overwriting any existing files).
2. Edit the file `server.conf` in the folder `files/splunk/linux/SPLUNK_HOME/etc/system/local` and uncomment the last two lines as suggested in the file itself. Your file should look like this:

```
#
# uncomment the section below if you want to enable SSL
#
[sslConfig]
sslRootCAPath = /opt/splunkforwarder/etc/mycerts/ca.pem
```

3. Set splunk_ssl to yes in the file group_vars/all/vars, uncommenting the line if required. Make sure that the splunk_architecture_forward_servers list specifies all your indexers together with the port that was configured to accept SSL:

```
monitoring_stack: splunk
splunk_ssl: yes
splunk_architecture_forward_servers:
- indexer1.cloudra.local:9998
- indexer2.cloudra.local:9998
```

Hybrid environment Linux / Windows

Currently, you cannot deploy your own certificates for use by the Universal Forwarders deployed on Windows machines. If you want to have your Linux machines in a hybrid deployment to use SSL, proceed as follows.

1. Comment out the `splunk_architecture_forward_servers` variable (and its values) from `group_vars/all/vars`

```
monitoring_stack: splunk
splunk_ssl: yes
#splunk_architecture_forward_servers:
#  - hpe2-ansible.cloudra.local:9998
```

2. Create a file named `vms.yml` in the folder `group_vars` and specify the list of forward servers to use by the Linux servers. This list is typically the same as the one used for Windows servers but specifies a TCP port that enables SSL.

```
splunk_architecture_forward_servers:
- hpe2-ansible.cloudra.local:9998
```

3. Edit the `group_vars/win_worker.yml` file and specify the list of forward servers to be used by the Windows servers. This list is typically the same as the one used for Linux servers but specifies a TCP port that does not enable SSL.

```
splunk_architecture_forward_servers:
- hpe2-ansible.cloudra.local:9997
```

Appendix D: How to check that certs were deployed correctly

The following commands should return the CA certificates used by UCP / DTR. This certificates is the same as the one pointed to by the `--cacert` switch.

```
# curl --cacert <ucp_certs_dir>/ca.pem https://<your ucp fqdn>/ca
# curl --cacert <dtr_certs_dir>/ca.pem https://<your dtr fqdn>/ca
```

Output 1: certificates successfully deployed (content will depend on your own CA certificate)

```
-----BEGIN CERTIFICATE-----
MIIDyTCCArGgAwIBAgIUUeo+H6xGSB7/9gqq9T2SUwJPLggwDQYJKoZIhvcNAQEL
BQAwbDELMAkGA1UEBhMCRlIxFTATBgNVBAcTDFRoZSBJbnRlcm5ldDETMBEGA1UE
ChMKQ2hyaXNOb3BoZTEUMBIGA1UECxMLQOEgU2VydmljZXMxGzAZBgNVBAMTEkNo
...
XkJ8WcsHocJO8J9J3RaWsM2BQc7wRntJcOkA7ooTH13OtQTP1jFcQp5xNdI4J3Mz
j9BAYERjkGqu7v9tfOem99oVGUal2Opu4r73eWUm1mL948xuw6PgiRSLZrXhn/RS
uvFVnS/vPYJozOXIZA==
-----END CERTIFICATE-----
```

If the deployment was not successful, `curl` will output something like **Output 2.**

Output 2: certificates were not successfully deployed

```
curl: (60) Peer's Certificate issuer is not recognized.
More details here: http://curl.haxx.se/docs/sslcerts.html
...
```

Enable certs for browser (Windows 2016 example)

Choose `Manage computer certificates` in the control panel as shown in Figure 69.

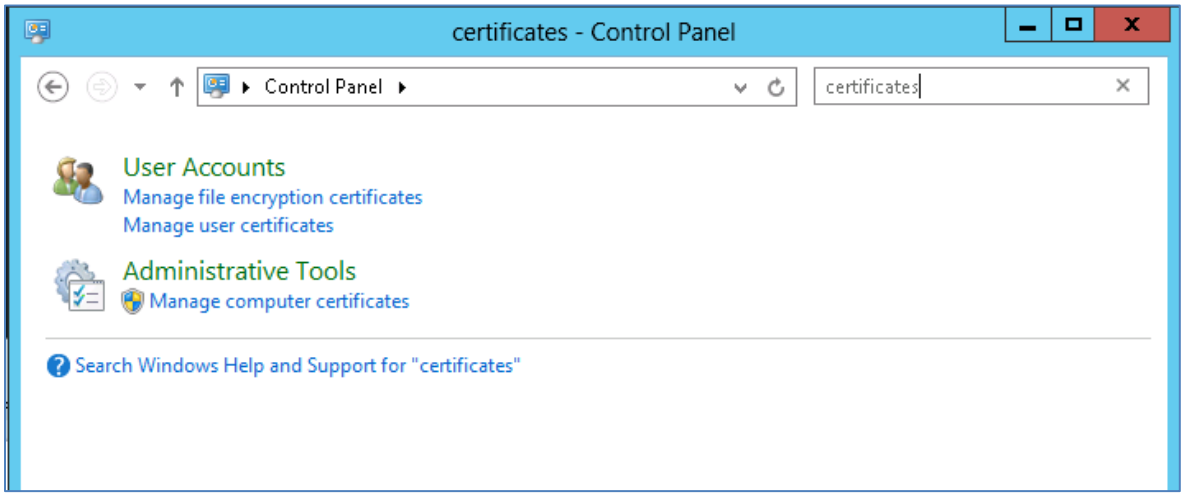

Figure 69. Manage computer certificates

Import the `ca.pem` for UCP into the Trusted Root Certification Authorities, as shown in Figure 70.

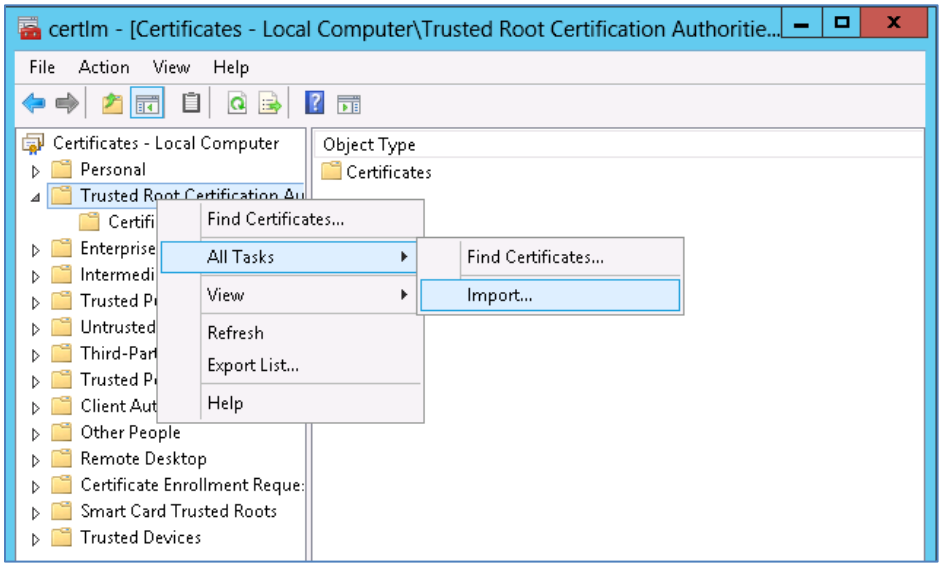

Figure 70. Import the ca.pem

It should now show up in the list of certificates. You may need to restart your browser to see the green, secure lock symbol as shown in Figure 71.

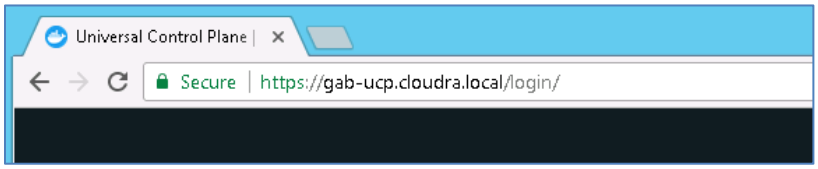

Figure 71. Secure HTTPS

Resources and additional links

HPE Reference Architectures, hpe.com/info/ra

HPE Synergy, hpe.com/synergy

HPE Servers, hpe.com/servers

HPE Storage, hpe.com/storage

HPE Networking, hpe.com/networking

HPE Technology Consulting Services, hpe.com/us/en/services/consulting.html

Docker Reference Architectures, https://success.docker.com/architectures

Splunk Validate Architectures, https://www.splunk.com/pdfs/white-papers/splunk-validated-architectures.pdf

Sysdig Resources, https://sysdig.com/resources/

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

Sign up for updates

a000xxxxxenw, April 2019