Infrastructure

HCI

NetApp
November 04, 2020

Table of Contents

Infrastructure

NVA-1148: NetApp HCI with Red Hat Virtualization

Alan Cowles, Nikhil M Kulkarni, NetApp

NetApp HCI with Red Hat Virtualization is a verified, best-practice architecture for the deployment of an on-premises virtual datacenter environment in a reliable and dependable manner.

This architecture reference document serves as both a design guide and a deployment validation of the Red Hat Virtualization solution on NetApp HCI. The architecture described in this document has been validated by subject matter experts at NetApp and Red Hat to provide a best-practice implementation for an enterprise virtual datacenter deployment using Red Hat Virtualization on NetApp HCI within your own enterprise datacenter environment.

Use Cases

The NetApp HCI for Red Hat OpenShift on Red Hat Virtualization solution is architected to deliver exceptional value for customers with the following use cases:

1. Infrastructure to scale on demand with NetApp HCI
2. Enterprise virtualized workloads in Red Hat Virtualization

Value Proposition and Differentiation of NetApp HCI with Red Hat Virtualization

NetApp HCI provides the following advantages with this virtual infrastructure solution:

- A disaggregated architecture that allows for independent scaling of compute and storage.
- The elimination of virtualization licensing costs and a performance tax on independent NetApp HCI storage nodes.
- NetApp Element storage provides quality of service (QoS) per storage volume and allows for guaranteed storage performance for workloads on NetApp HCI, preventing adjacent workloads from negatively affecting performance.
- The data fabric powered by NetApp allows data to be replicated from an on-premise to on- premise location or replicated to the cloud to move the data closer to where the application needs the data.
- Support through NetApp Support or Red Hat Support.

NetApp HCI Design

NetApp HCI, is the industry’s first and leading disaggregated hybrid cloud infrastructure, providing the widely recognized benefits of hyperconverged solutions. Benefits include lower TCO and ease of acquisition, deployment, and management for virtualized workloads, while also allowing enterprise customers to independently scale compute and storage resources as needed. NetApp HCI with Red Hat

Virtualization provides an open source, enterprise virtualization environment based on Red Hat Enterprise Linux.

By providing an agile turnkey infrastructure platform, NetApp HCI enables you to run enterprise-class virtualized and containerized workloads in an accelerated manner. At its core, NetApp HCI is designed to provide predictable performance, linear scalability of both compute and storage resources, and a simple deployment and management experience.

Predictable

One of the biggest challenges in a multitenant environment is delivering consistent, predictable performance for all your workloads. Running multiple enterprise-grade workloads can result in resource contention, where one workload interferes with the performance of another. NetApp HCI alleviates this concern with storage quality-of-service (QoS) limits that are available natively with NetApp Element software. Element enables the granular control of every application and volume, helps to eliminate noisy neighbors, and satisfies enterprise performance SLAs. NetApp HCI multitenancy capabilities can help eliminate many traditional performance-related problems.

Flexible

Previous generations of hyperconverged infrastructure typically required fixed resource ratios, limiting deployments to four-node and eight-node configurations. NetApp HCI is a disaggregated hyper-converged infrastructure that can scale compute and storage resources independently. Independent scaling prevents costly and inefficient overprovisioning, eliminates the 10% to 30% HCI tax from controller virtual machine (VM) overhead, and simplifies capacity and performance planning. NetApp HCI is available in mix-and-match, small, medium, and large storage and compute configurations.

The architectural design choices offered enable you to confidently scale on your terms, making HCI viable for core Tier-1 data center applications and platforms. NetApp HCI is architected in building blocks at either the chassis or the node level. Each chassis can hold four nodes in a mixed configuration of storage or compute nodes.

Simple

A driving imperative within the IT community is to simplify deployment and automate routine tasks, eliminating the risk of user error while freeing up resources to focus on more interesting, higher-value projects. NetApp HCI can help your IT department become more agile and responsive by both simplifying deployment and ongoing management.

Business Value

Enterprises that perform virtualization in an open-source data center with Red Hat products can realize the value of this solution by following the recommended design, deployment, and best practices described in this document. The detailed setup of RHV on NetApp HCI provides several benefits when deployed as part of an enterprise virtualization solution:

- High availability at all layers of the stack
- Thoroughly documented deployment procedures
- Nondisruptive operations and upgrades to hypervisors and the manager VM
- API-driven, programmable infrastructure to facilitate management
- Multitenancy with performance guarantees
- The ability to run virtualized workloads based on KVM with enterprise-grade features and support
- The ability to scale infrastructure independently based on workload demands

NetApp HCI with Red Hat Virtualization acknowledges these challenges and helps address each concern by implementing a verified architecture for solution deployment.

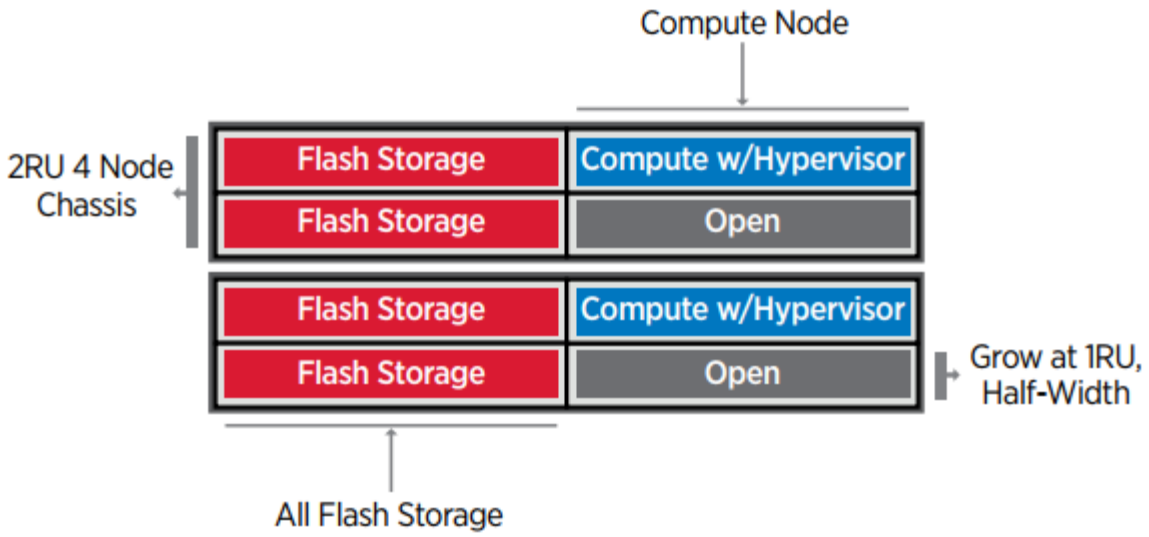
Technology Overview

With NetApp HCI for Red Hat Virtualization, you can deploy a fully integrated, production-grade virtual data center that allows you to take advantage of the following features:

- NetApp HCI compute and storage nodes
  - Enterprise-grade hyperconverged infrastructure designed for hybrid cloud workloads
  - NetApp Element storage software
  - Intel- based server compute nodes, including options for NVIDIA GPUs
- Red Hat Virtualization
  - Enterprise hypervisor solution for deployment and management of virtual infrastructures

NetApp HCI

NetApp HCI is an enterprise-scale disaggregated hybrid cloud infrastructure (HCI) solution that delivers compute and storage resources in an agile, scalable, and easy-to-manage two-rack unit (2RU) four-node building block. It can also be configured with 1RU compute and server nodes. The minimum deployment consists of four NetApp HCI storage nodes and two NetApp HCI compute nodes. The compute nodes are installed as RHV-H hypervisors in an HA cluster. This minimum deployment can be easily scaled to fit customer enterprise workload demands by adding additional NetApp HCI storage or compute nodes to expand available resources.



The design for NetApp HCI for Red Hat Virtualization consists of the following components in a minimum starting configuration:

- NetApp H-Series all-flash storage nodes running NetApp Element software
- NetApp H-Series compute nodes running the Red Hat Virtualization RHV-H hypervisor

For more information about compute and storage nodes in NetApp HCI, see the NetApp HCI Datasheet.

NetApp Element Software

NetApp Element software provides modular, scalable performance, with each storage node delivering guaranteed capacity and throughput to the environment. You can also specify per-volume storage QoS policies to support dedicated performance levels for even the most demanding workloads.

iSCSI Login Redirection and Self-Healing Capabilities

NetApp Element software uses the iSCSI storage protocol, a standard way to encapsulate SCSI commands on a traditional TCP/IP network. When SCSI standards change or when Ethernet network performance improves, the iSCSI storage protocol benefits without the need for any changes.

Although all storage nodes have a management IP and a storage IP, NetApp Element software advertises a single storage virtual IP address (SVIP address) for all storage traffic in the cluster. As a part of the iSCSI login process, storage can respond that the target volume has been moved to a different address, and therefore it cannot proceed with the negotiation process. The host then reissues the login request to the new address in a process that requires no host-side reconfiguration. This process is known as iSCSI login redirection.

iSCSI login redirection is a key part of the NetApp Element software cluster. When a host login request is received, the node decides which member of the cluster should handle the traffic based on IOPS and the capacity requirements for the volume. Volumes are distributed across the NetApp Element software cluster and are redistributed if a single node is handling too much traffic for its volumes or if

a new node is added. Multiple copies of a given volume are allocated across the array. In this manner, if a node failure is followed by volume redistribution, there is no effect on host connectivity beyond a logout and login with redirection to the new location. With iSCSI login redirection, a NetApp Element software cluster is a self-healing, scale-out architecture that is capable of non- disruptive upgrades and operations.

NetApp Element Software Cluster QoS

A NetApp Element software cluster allows QoS to be dynamically configured on a per-volume basis. You can use per-volume QoS settings to control storage performance based on SLAs that you define. The following three configurable parameters define the QoS:

- **Minimum IOPS.** The minimum number of sustained IOPS that the NetApp Element software cluster provides to a volume. The minimum IOPS configured for a volume is the guaranteed level of performance for a volume. Per-volume performance does not drop below this level.
- **Maximum IOPS.** The maximum number of sustained IOPS that the NetApp Element software cluster provides to a specific volume.
- **Burst IOPS.** The maximum number of IOPS allowed in a short burst scenario. The burst duration setting is configurable, with a default of 1 minute. If a volume has been running below the maximum IOPS level, burst credits are accumulated. When performance levels become very high and are pushed, short bursts of IOPS beyond the maximum IOPS are allowed on the volume.

Multitenancy

Secure multitenancy is achieved with the following features:

- **Secure authentication.** The Challenge-Handshake Authentication Protocol (CHAP) is used for secure volume access. The Lightweight Directory Access Protocol (LDAP) is used for secure access to the cluster for management and reporting.
- **Volume access groups (VAGs).** Optionally, VAGs can be used in lieu of authentication, mapping any number of iSCSI initiator-specific iSCSI Qualified Names (IQNs) to one or more volumes. To access a volume in a VAG, the initiator's IQN must be in the allowed IQN list for the group of volumes.
- **Tenant virtual LANs (VLANs).** At the network level, end-to-end network security between iSCSI initiators and the NetApp Element software cluster is facilitated by using VLANs. For any VLAN that is created to isolate a workload or a tenant, Element software creates a separate iSCSI target SVIP address that is accessible only through the specific VLAN.
- **VPN routing/forwarding (VRF)-enabled VLANs.** To further support security and scalability in the data center, Element software allows you to enable any tenant VLAN for VRF-like functionality. This feature adds these two key capabilities:
  - **L3 routing to a tenant SVIP address.** This feature allows you to situate iSCSI initiators on a separate network or VLAN from that of the NetApp Element software cluster.
  - **Overlapping or duplicate IP subnets.** This feature enables you to add a template to tenant

environments, allowing each respective tenant VLAN to be assigned IP addresses from the same IP subnet. This capability can be useful for service provider environments where scale and preservation of IP- space are important.

Enterprise Storage Efficiencies

The NetApp Element software cluster increases overall storage efficiency and performance. The following features are performed inline, are always on, and require no manual configuration by the user:

- **Deduplication.** The system only stores unique 4K blocks. Any duplicate 4K blocks are automatically associated with an already stored version of the data. Data is on block drives and is mirrored with Element Helix data protection. This system significantly reduces capacity consumption and write operations within the system.
- **Compression.** Compression is performed inline before data is written to NVRAM. Data is compressed, stored in 4K blocks, and remains compressed in the system. This compression significantly reduces capacity consumption, write operations, and bandwidth consumption across the cluster.
- **Thin provisioning.** This capability provides the right amount of storage at the time that you need it, eliminating capacity consumption that caused by overprovisioned volumes or underutilized volumes.
- **Helix.** The metadata for an individual volume is stored on a metadata drive and is replicated to a secondary metadata drive for redundancy.

Element was designed for automation. All the storage features mentioned above can be managed with APIs. These APIs are the only method that the UI uses to control the system and can be incorporated into user workflows to ease the management of the solution.

Red Hat Virtualization

Red Hat Virtualization (RHV) is an enterprise virtual data center platform that runs on Red Hat Enterprise Linux using the KVM hypervisor.

For more information about Red Hat Virtualization, see the website located here.

RHV provides the following features:

- **Centralized management of VMs and hosts.** The RHV manager runs as a physical or VM in the deployment and provides a web-based GUI for the management of the solution from a central interface.
- **Self-Hosted Engine.** To minimize the hardware requirements, RHV allows RHV Manager to be deployed as a VM on the same hosts that run guest VMs.
- **High Availability.** To avoid disruption from host failures, RHV allows VMs to be configured for

high availability. The highly available VMs are controlled at the cluster level using resiliency policies.

- **High Scalability.** A single RHV cluster can have up to 200 hypervisor hosts, enabling it to support the requirements of massive VMs to hold resource-greedy enterprise-class workloads.
- **Enhanced security.** Inherited from RHEL, Secure Virtualization (sVirt) and Security Enhanced Linux (SELinux) technologies are employed by RHV for the purposes of elevated security and hardening for the hosts and VMs. The key advantage from these features is logical isolation of a VM and its associated resources.

Red Hat Virtualization Manager

Red Hat Virtualization Manager (RHV-M) provides centralized enterprise-grade management for the physical and logical resources within the RHV virtualized environment. A web-based GUI with different role- based portals is provided to access RHV-M features.

RHV-M exposes configuration and management of RHV resources with open-source, community-driven RESTful APIs. It also supports full-fledged integration with Red Hat CloudForms and Red Hat Ansible for automation and orchestration.

Red Hat Virtualization Hosts

Hosts (also called hypervisors) are the physical servers that provide hardware resources for the VMs to run on. A kernel-based virtual machine (KVM) provides full virtualization support, and Virtual Desktop Server Manager (VDSM) is the host agent that is responsible for host communication with the RHV-M.

The two types of hosts supported in Red Hat Virtualization are Red Hat Virtualization Hosts (RHV-H) and Red Hat Enterprise Linux hosts (RHEL).

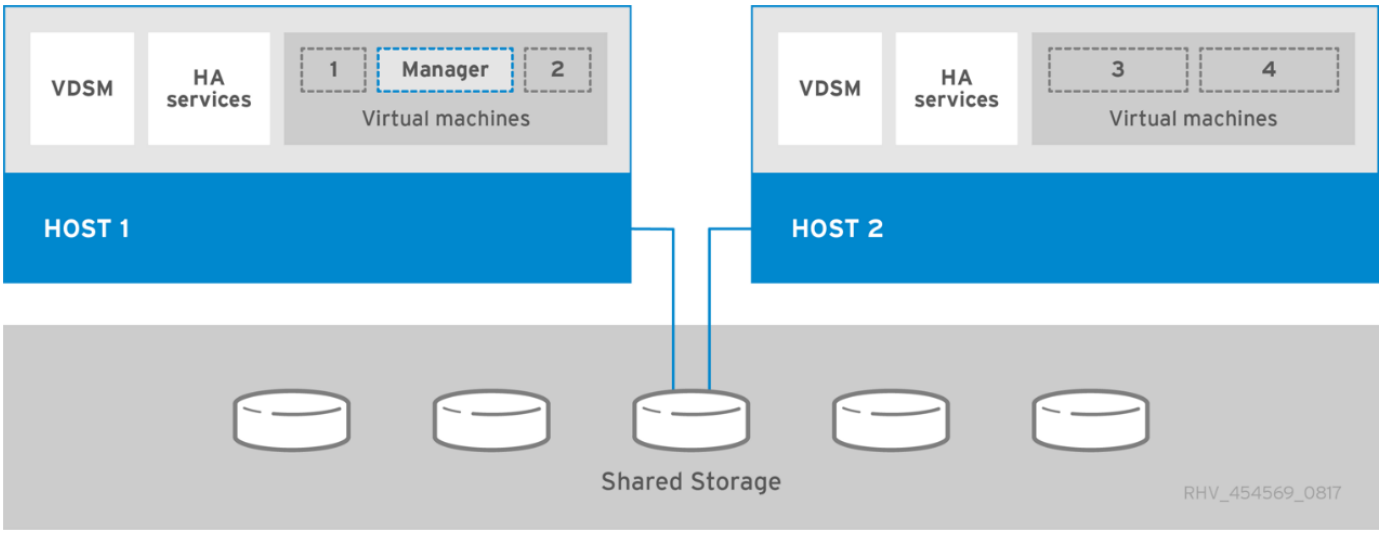
RHV-H is a minimal, light-weight operating system based on Red Hat Enterprise Linux that is optimized for the ease of setting up physical servers as RHV hypervisors.

RHEL hosts are servers that run the standard Red Hat Enterprise Linux operating system. They can then be configured with the required subscriptions to install the packages required to permit the physical servers to be used as RHV hosts.

Red Hat Virtualization Architecture

Red Hat Virtualization can be deployed in two different architectures, with the RHV-M as a physical server in the infrastructure or with the RHV-M configured as a self-hosted engine. NetApp recommends using the self-hosted engine deployment, in which the RHV-M is a VM hosted in the same environment as other VMs, as we do in this guide.

A minimum of two self-hosted nodes are required for high availability of guest VMs and RHV-M. To provide high availability for the manager VM, HA services are enabled and run on all the self-hosted engine nodes.



Architecture Overview: NetApp HCI with RHV

Hardware Requirements

The following table lists the minimum number of hardware components that are required to implement the solution. The hardware components that are used in specific implementations of the solution might vary based on customer requirements.

Hardware	Model	Quantity
NetApp HCI compute nodes	NetApp H410C	2
NetApp HCI storage nodes	NetApp H410S	4
Data switches	Mellanox SN2010	2
Management switches	Cisco Nexus 3048	2

Software Requirements

The following table lists the software components that are required to implement the solution. The software components that are used in any implementation of the solution might vary based on customer requirements.

Software	Purpose	Version
NetApp HCI	Infrastructure (compute/storage)	1.8
NetApp Element	Storage	12.0
Red Hat Virtualization	Virtualization	4.3.9

Design Considerations: NetApp HCI with RHV

Review the following design considerations when developing your deployment

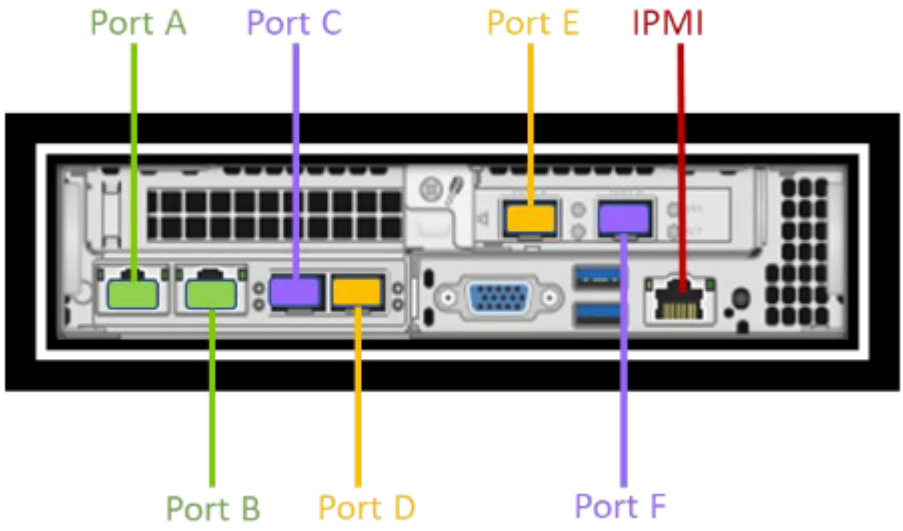
strategy.

Networking Requirements

This section describes the networking requirements for the deployment of Red Hat Virtualization on NetApp HCI as a validated solution. It provides physical diagrams of the network ports on both the NetApp HCI compute nodes and the switches deployed in the solution. This section also describes the arrangement and purpose of each virtual network segment used in the solution.

Port Identification

NetApp HCI consists of NetApp H-Series nodes dedicated to either compute or storage. Both node configurations are available with two 1GbE ports (ports A and B) and two 10/25GbE ports (ports C and D) on board. The compute nodes have additional 10/25GbE ports (ports E and F) available in the first mezzanine slot. Each node also has an additional out-of-band management port that supports Intelligent Platform Management Interface (IPMI) functionality. Each of these ports on the rear of an H410C node can be seen in the following figure.



Network Design

The NetApp HCI with Red Hat Virtualization solution uses two data switches to provide primary data connectivity at 25Gbps. It also uses two additional management switches that provide connectivity at 1Gbps for in-band management for the storage nodes and out-of-band management for IPMI functionality.

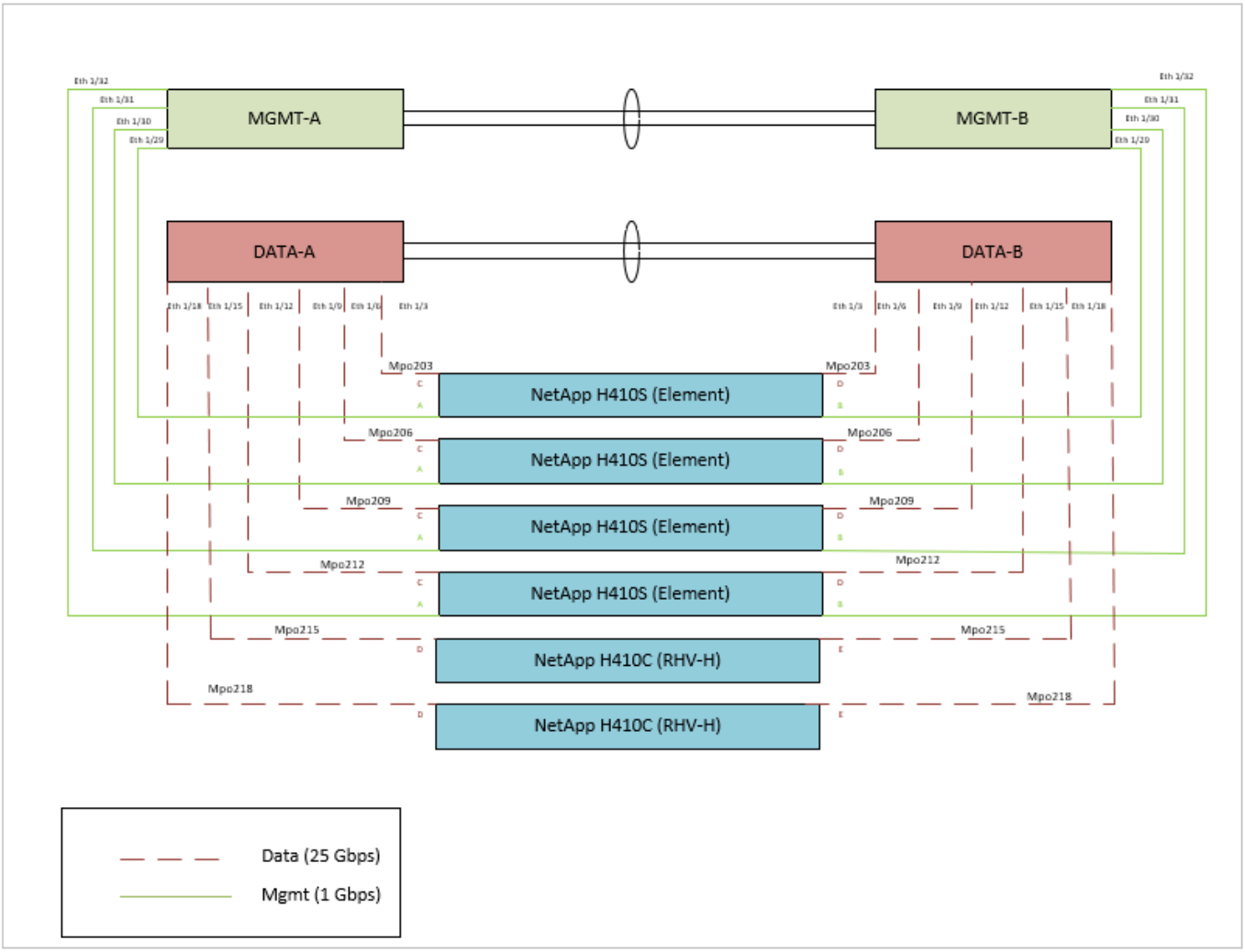
Cabling Storage Nodes

The management ports A and B must be active on each storage node to configure the NetApp HCI cluster, and provide management accessibility to Element after the solution is deployed. The two 25Gbps ports (C and D) should be connected, one to each data switch, to provide physical fault tolerance. The switch ports should be configured for multi-chassis link aggregation (MLAG) and the data ports on the node should be configured for LACP with jumbo-frames support enabled. The IPMI

ports on each node can be used to remotely manage the node after it is installed in a data center. With IPMI, the node can be accessed with a web-browser-based console to run the initial installation, run diagnostics, and reboot or shut down the node if necessary.

Cabling Compute Nodes

The two 25Gbps ports (C and E) should be connected, one to each data switch, to provide physical fault tolerance. The switch ports should be configured for multi-chassis link aggregation (MLAG), and the data ports on the node should be configured for LACP with jumbo-frames support enabled. The IPMI ports can also be used to remotely manage the node after it is installed in a data center. With IPMI, the node can be accessed with a web-browser- based console to run the initial installation, run diagnostics, and reboot or shut down the node if necessary.



VLAN Requirements

The solution is designed to logically separate network traffic for different purposes by using Virtual Local Area Networks (VLANs). NetApp HCI requires a minimum of three network segments. However, this configuration can be scaled to meet customer demands or to provide further isolation for specific network services. The following table lists the VLANs that are required to implement the solution, as

well as the specific VLAN IDs that are used later in the validated architecture deployment.

VLANs	Purpose	VLAN Used
Out-of-band management network	Management for HCI nodes / IPMI	16
In-band management network	Management for HCI nodes / ovirtmgmt	1172
Storage network	Storage network for NetApp Element.	3343
Migration network	Network for virtual guest migration.	3345
VM network	Network for virtual guests.	3346

Network Infrastructure Support Resources

The following infrastructure should be in place prior to the deployment of the Red Hat Virtualization on NetApp HCI solution:

- At least one DNS server providing full host-name resolution that is accessible from the in-band management network and the VM network.
- At least one NTP server that is accessible from the in-band management network and the VM network.
- Outbound internet connectivity is recommended, but not required, for both the in-band management network and the VM network.

Deployment Procedures NetApp HCI with RHV

Deployment Summary: NetApp HCI with RHV

The detailed steps provided in this section provide a validation for the minimum hardware and software configuration required to deploy and validate the NetApp HCI with Red Hat Virtualization solution.

Deploying Red Hat Virtualization for NetApp HCI involves the following high-level tasks:

1. Configure Management Switches
2. Configure Data Switches
3. Deploy Element Storage System on HCI Storage Nodes
4. Install RHV-H to HCI Compute Nodes
5. Deploy RHV Manager as a Self-hosted Engine

6. Deploy Test VMs
7. Test HA Functionality

1. Configure Management Switches: NetApp HCI with RHV

Cisco Nexus 3048 switches are used in this deployment procedure to provide 1Gbps connectivity for in and out-of-band management of the compute and storage nodes. These steps begin after the switches have been racked, powered, and put through the initial setup process. To configure the switches to provide management connectivity to the infrastructure, complete the following steps:

Enable Advanced Features for Cisco Nexus

Run the following commands on each Cisco Nexus 3048 switch to configure advanced features:

1. Enter configuration mode.

```
Switch-01# configure terminal
```

2. Enable VLAN functionality.

```
Switch-01(config)# feature interface-vlan
```

3. Enable LACP.

```
Switch-01(config)# feature lacp
```

4. Enable virtual port channels (vPCs).

```
Switch-01(config)# feature vpc
```

5. Set the global port-channel load-balancing configuration.

```
Switch-01(config)# port-channel load-balance src-dst ip-l4port
```

6. Perform global spanning-tree configuration.

```
Switch-01(config)# spanning-tree port type network default
Switch-01(config)# spanning-tree port type edge bpduguard default
```

Configure Ports on the Switch for In-Band Management

1. Run the following commands to create VLANs for management purposes:

```
Switch-01(config)# vlan 2
Switch-01(config-vlan)# Name Native_VLAN
Switch-01(config-vlan)# vlan 16
Switch-01(config-vlan)# Name OOB_Network
Switch-01(config-vlan)# vlan 1172
Switch-01(config-vlan)# Name MGMT_Network
Switch-01(config-vlan)# exit
```

2. Configure the ports ETH1/29-32 as VLAN trunk ports that connect to management interfaces on each HCI storage node.

```
Switch-01(config)# int eth 1/29
Switch-01(config-if)# description HCI-STG-01 PortA
Switch-01(config-if)# switchport mode trunk
Switch-01(config-if)# switchport trunk native vlan 2
Switch-01(config-if)# switchport trunk allowed vlan 1172
Switch-01(config-if)# spanning tree port type edge trunk
Switch-01(config-if)# int eth 1/30
Switch-01(config-if)# description HCI-STG-02 PortA
Switch-01(config-if)# switchport mode trunk
Switch-01(config-if)# switchport trunk native vlan 2
Switch-01(config-if)# switchport trunk allowed vlan 1172
Switch-01(config-if)# spanning tree port type edge trunk
Switch-01(config-if)# int eth 1/31
Switch-01(config-if)# description HCI-STG-03 PortA
Switch-01(config-if)# switchport mode trunk
Switch-01(config-if)# switchport trunk native vlan 2
Switch-01(config-if)# switchport trunk allowed vlan 1172
Switch-01(config-if)# spanning tree port type edge trunk
Switch-01(config-if)# int eth 1/32
Switch-01(config-if)# description HCI-STG-04 PortA
Switch-01(config-if)# switchport mode trunk
Switch-01(config-if)# switchport trunk native vlan 2
Switch-01(config-if)# switchport trunk allowed vlan 1172
Switch-01(config-if)# spanning tree port type edge trunk
Switch-01(config-if)# exit
```

Configure Ports on the Switch for Out-of-Band Management

Run the following commands to configure the ports for cabling the IPMI interfaces on each HCI node.

```
Switch-01(config)# int eth 1/13
Switch-01(config-if)# description HCI-CMP-01 IPMI
Switch-01(config-if)# switchport mode access
Switch-01(config-if)# switchport access vlan 16
Switch-01(config-if)# spanning-tree port type edge
Switch-01(config-if)# int eth 1/14
Switch-01(config-if)# description HCI-STG-01 IPMI
Switch-01(config-if)# switchport mode access
Switch-01(config-if)# switchport access vlan 16
Switch-01(config-if)# spanning-tree port type edge
Switch-01(config-if)# int eth 1/15
Switch-01(config-if)# description HCI-STG-03 IPMI
Switch-01(config-if)# switchport mode access
Switch-01(config-if)# switchport access vlan 16
Switch-01(config-if)# spanning-tree port type edge
Switch-01(config-if)# exit
```

In the validated configuration, we cabled odd-node IPMI interfaces to Switch-01 and even-node IPMI interfaces to Switch-02.

Create a vPC Domain to Ensure Fault Tolerance

1. Activate the ports used for the vPC peer-link between the two switches.

   ```
   Switch-01(config)# int eth 1/1
   Switch-01(config-if)# description vPC peer-link Switch-02 1/1
   Switch-01(config-if)# int eth 1/2
   Switch-01(config-if)# description vPC peer-link Switch-02 1/2
   Switch-01(config-if)# exit
   ```

2. Perform the vPC global configuration.

```
Switch-01(config)# vpc domain 1
Switch-01(config-vpc-domain)# role priority 10
Switch-01(config-vpc-domain)# peer-keepalive destination <switch-02_mgmt_address> source
<switch-01_mgmt_address> vrf managment
Switch-01(config-vpc-domain)# peer-gateway
Switch-01(config-vpc-domain)# auto recovery
Switch-01(config-vpc-domain)# ip arp synchronize
Switch-01(config-vpc-domain)# int eth 1/1-2
Switch-01(config-vpc-domain)# channel-group 10 mode active
Switch-01(config-vpc-domain)# int Po10
Switch-01(config-if)# description vPC peer-link
Switch-01(config-if)# switchport mode trunk
Switch-01(config-if)# switchport trunk native vlan 2
Switch-01(config-if)# switchport trunk allowed vlan 16, 1172
Switch-01(config-if)# spanning-tree port type network
Switch-01(config-if)# vpc peer-link
Switch-01(config-if)# exit
```

2. Configure Data Switches: NetApp HCI with RHV

Mellanox SN2010 switches are used in this deployment procedure to provide 25Gbps connectivity for the data plane of the compute and storage nodes. These steps begin after the switches have been racked, cabled, and put through the initial setup process. To configure the switches to provide data connectivity to the infrastructure, complete the following steps:

Create MLAG Cluster to Provide Fault Tolerance

1. Run the following commands on each Mellanox SN210 switch for general configuration:
   a. Enter configuration mode.

   ```
   Switch-01 enable
   Switch-01 configure terminal
   ```

   b. Enable the LACP required for the Inter-Peer Link (IPL).

   ```
   Switch-01 (config) # lacp
   ```

   c. Enable the Link Layer Discovery Protocol (LLDP).

```
Switch-01 (config) # lldp
```

d. Enable IP routing.

```
Switch-01 (config) # ip routing
```

e. Enable the MLAG protocol.

```
Switch-01 (config) # protocol mlag
```

f. Enable global QoS.

```
Switch-01 (config) # dcb priority-flow-control enable force
```

2. For MLAG to function, the switches must be made peers to each other through an IPL. This should consist of two or more physical links for redundancy. The MTU for the IPL is set for jumbo frames (9216), and all VLANs are enabled by default. Run the following commands on each switch in the domain:

a. Create port channel 10 for the IPL.

```
Switch-01 (config) # interface port-channel 10
Switch-01 (config interface port-channel 10) # description IPL
Switch-01 (config interface port-channel 10) # exit
```

b. Add interfaces ETH 1/20 and 1/22 to the port channel.

```
Switch-01 (config) # interface ethernet 1/20 channel-group 10 mode active
Switch-01 (config) # interface ethernet 1/20 description ISL-SWB_01
Switch-01 (config) # interface ethernet 1/22 channel-group 10 mode active
Switch-01 (config) # interface ethernet 1/22 description ISL-SWB_02
```

c. Create a VLAN outside of the standard range dedicated to IPL traffic.

```
Switch-01 (config) # vlan 4000
Switch-01 (config vlan 4000) # name IPL VLAN
Switch-01 (config vlan 4000) # exit
```

d. Define the port channel as the IPL.

```
Switch-01 (config) # interface port-channel 10 ipl 1
Switch-01 (config) # interface port-channel 10 dcb priority-flow-control mode on
force
```

e. Set an IP for each IPL member (non-routable; it is not advertised outside of the switch).

```
Switch-01 (config) # interface vlan 4000
Switch-01 (config vlan 4000) # ip address 10.0.0.1 255.255.255.0
Switch-01 (config vlan 4000) # ipl 1 peer-address 10.0.0.2
Switch-01 (config vlan 4000) # exit
```

3. Create a unique MLAG domain name for the two switches and assign a MLAG virtual IP (VIP). This IP is used for keep-alive heartbeat messages between the two switches. Run these commands on each switch in the domain:

   a. Create the MLAG domain and set the IP address and subnet.

   ```
   Switch-01 (config) # mlag-vip MLAG-VIP-DOM ip a.b.c.d /24 force
   ```

   b. Create a virtual MAC address for the system MLAG.

   ```
   Switch-01 (config) # mlag system-mac AA:BB:CC:DD:EE:FF
   ```

   c. Configure the MLAG domain so that it is active globally.

   ```
   Switch-01 (config) # no mlag shutdown
   ```

The IP used for the MLAG VIP must be in the same subnet as the switch management network (mgmt0). Also, The MAC address used can be any unicast MAC address and must be set to the same value on both switches in the MLAG domain.

Configure Ports to Connect to Storage and Compute Hosts

1. Create each of the VLANs needed to support the services for NetApp HCI. Run these commands on each switch in the domain:

   a. Create the VLANs.

```
Switch-01 (config) # vlan 1172
Switch-01 (config vlan 1172) exit
Switch-01 (config) # vlan 3343
Switch-01 (config vlan 3343) exit
Switch-01 (config) # vlan 3344
Switch-01 (config vlan 3345) exit
Switch-01 (config) # vlan 3345
Switch-01 (config vlan 3346) exit
```

b. Create names for each VLAN for easier accounting.

```
Switch-01 (config) # vlan 1172 name “MGMT_Network”
Switch-01 (config) # vlan 3343 name “Storage_Network”
Switch-01 (config) # vlan 3345 name “Migration_Network”
Switch-01 (config) # vlan 3346 name “VM_Network”
```

2. Create MLAG interfaces and hybrid VLANs on ports identified so that you can distribute connectivity between the switches and tag the appropriate VLANs for the NetApp HCI compute nodes.

a. Select the ports you want to work with.

```
Switch-01 (config) # interface ethernet 1/15
```

b. Set the MTU for each port.

```
Switch-01 (config interface ethernet 1/15) # mtu 9216 force
```

c. Modify spanning- tree settings for each port.

```
Switch-01 (config interface ethernet 1/15) # spanning-tree bpdufilter enable
Switch-01 (config interface ethernet 1/15) # spanning-tree port type edge
Switch-01 (config interface ethernet 1/15) # spanning-tree bpduguard enable
```

d. Set the switchport mode to hybrid.

```
Switch-01 (config interface ethernet 1/15) # switchport mode hybrid
Switch-01 (config interface ethernet 1/15) # exit
```

e. Create descriptions for each port being modified.

```
Switch-01 (config) # interface ethernet 1/15 description HCI-CMP-01 PortD
```

f. Create and configure the MLAG port channels.

```
Switch-01 (config) # interface mlag-port-channel 215
Switch-01 (config interface mlag-port-channel 215) # exit
Switch-01 (config) # interface mlag-port-channel 215 no shutdown
Switch-01 (config) # interface mlag-port-channel 215 mtu 9216 force
Switch-01 (config) # interface ethernet 1/15 lacp port-priority 10
Switch-01 (config) # interface ethernet 1/15 lacp rate fast
Switch-01 (config) # interface ethernet 1/15 mlag-channel-group 215 mode active
```

g. Tag the appropriate VLANs for the NetApp HCI environment.

```
Switch-01 (config) # interface mlag-port-channel 215 switchport hybrid
Switch-01 (config) # interface mlag-port-channel 215 switchport hybrid allowed-vlan
add 1172
Switch-01 (config) # interface mlag-port-channel 215 switchport hybrid allowed-vlan
add 3343
Switch-01 (config) # interface mlag-port-channel 215 switchport hybrid allowed-vlan
add 3345
Switch-01 (config) # interface mlag-port-channel 215 switchport hybrid allowed-vlan
add 3346
```

3. Create MLAG interfaces and hybrid VLAN ports identified so that you can distribute connectivity between the switches and tag the appropriate VLANs for the NetApp HCI storage nodes.

a. Select the ports that you want to work with.

```
Switch-01 (config) # interface ethernet 1/3
```

b. Set the MTU for each port.

```
Switch-01 (config interface ethernet 1/3) # mtu 9216 force
```

c. Modify spanning tree settings for each port.

```
Switch-01 (config interface ethernet 1/3) # spanning-tree bpdufilter enable
Switch-01 (config interface ethernet 1/3) # spanning-tree port type edge
Switch-01 (config interface ethernet 1/3) # spanning-tree bpduguard enable
```

d. Set the switchport mode to hybrid.

```
Switch-01 (config interface ethernet 1/3) # switchport mode hybrid
Switch-01 (config interface ethernet 1/3) # exit
```

e. Create descriptions for each port being modified.

```
Switch-01 (config) # interface ethernet 1/3 description HCI-STG-01 PortD
```

f. Create and configure the MLAG port channels.

```
Switch-01 (config) # interface mlag-port-channel 203
Switch-01 (config interface mlag-port-channel 203) # exit
Switch-01 (config) # interface mlag-port-channel 203 no shutdown
Switch-01 (config) # interface mlag-port-channel 203 mtu 9216 force
Switch-01 (config) # interface mlag-port-channel 203 lacp-individual enable force
Switch-01 (config) # interface ethernet 203 lacp port-priority 10
Switch-01 (config) # interface ethernet 203 lacp rate fast
Switch-01 (config) # interface ethernet 1/3 mlag-channel-group 203 mode active
```

g. Tag the appropriate VLANs for the storage environment.

```
Switch-01 (config) # interface mlag-port-channel 203 switchport mode hybrid
Switch-01 (config) # interface mlag-port-channel 203 switchport hybrid allowed-vlan
add 1172
Switch-01 (config) # interface mlag-port-channel 203 switchport hybrid allowed-vlan
add 3343
```

The configurations in this section show the configuration for a single port as example. They must also be run for each additional port connected in the solution, as well as on the associated port of the second switch in the MLAG domain. NetApp recommends that the descriptions for each port are updated to reflect the device ports that are being cabled and configured on the other switch.

Create Uplink Ports for the Switches

1. Create an MLAG interface to provide uplinks to both Mellanox SN2010 switches from the core network.

```
Switch-01 (config) # interface mlag port-channel 201
Switch-01 (config interface mlag port-channel) # description Uplink CORE-SWITCH port
PORT
Switch-01 (config interface mlag port-channel) # exit
```

2. Configure the MLAG members.

```
Switch-01 (config) # interface ethernet 1/1 description Uplink to CORE-SWITCH port
PORT
Switch-01 (config) # interface ethernet 1/1 speed 10000 force
Switch-01 (config) # interface mlag-port-channel 201 mtu 9216 force
Switch-01 (config) # interface ethernet 1/1 mlag-channel-group 201 mode active
```

3. Set the switchport mode to hybrid and allow all VLANs from the core uplink switches.

```
Switch-01 (config) # interface mlag-port-channel switchport mode hybrid
Switch-01 (config) # interface mlag-port-channel switchport hybrid allowed-vlan all
```

4. Verify that the MLAG interface is up.

```
Switch-01 (config) # interface mlag-port-channel 201 no shutdown
Switch-01 (config) # exit
```

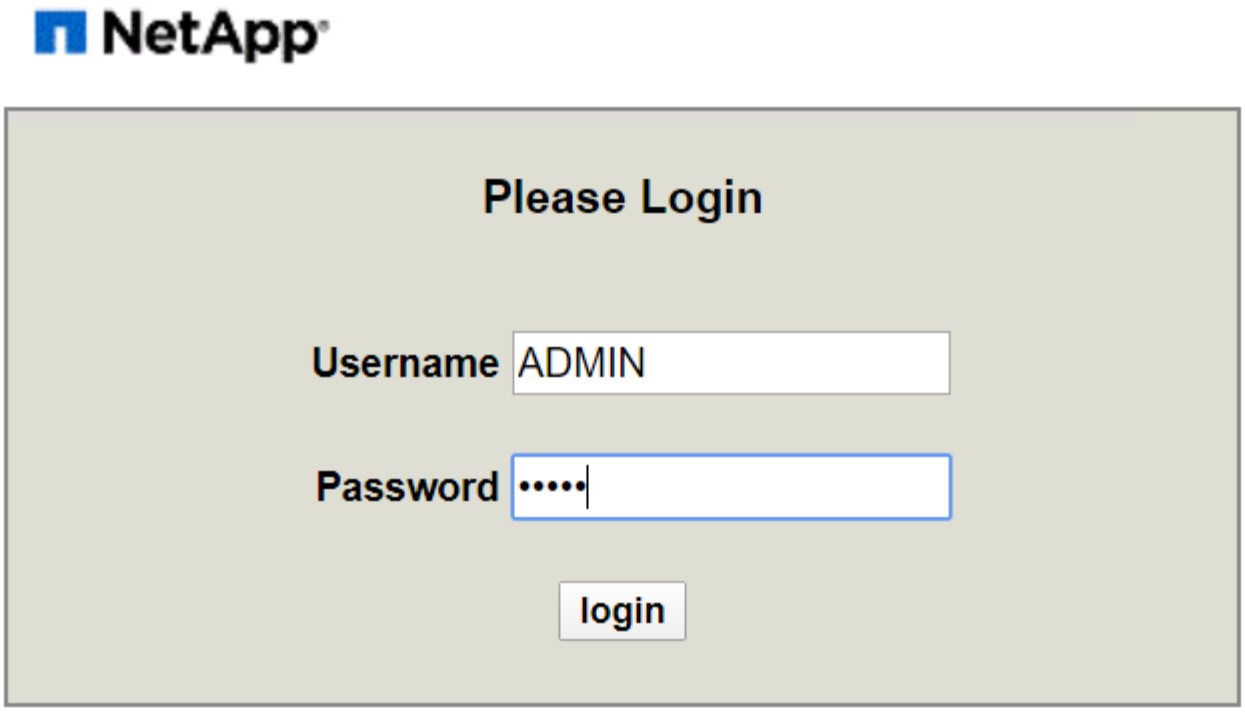
The configurations in this section must also be run on the second switch in the MLAG domain. NetApp recommends that the descriptions for each port are updated to reflect the device ports that are being cabled and configured on the other switch.

3. Deploy the Element Storage System on the HCI Storage Nodes: NetApp HCI with RHV

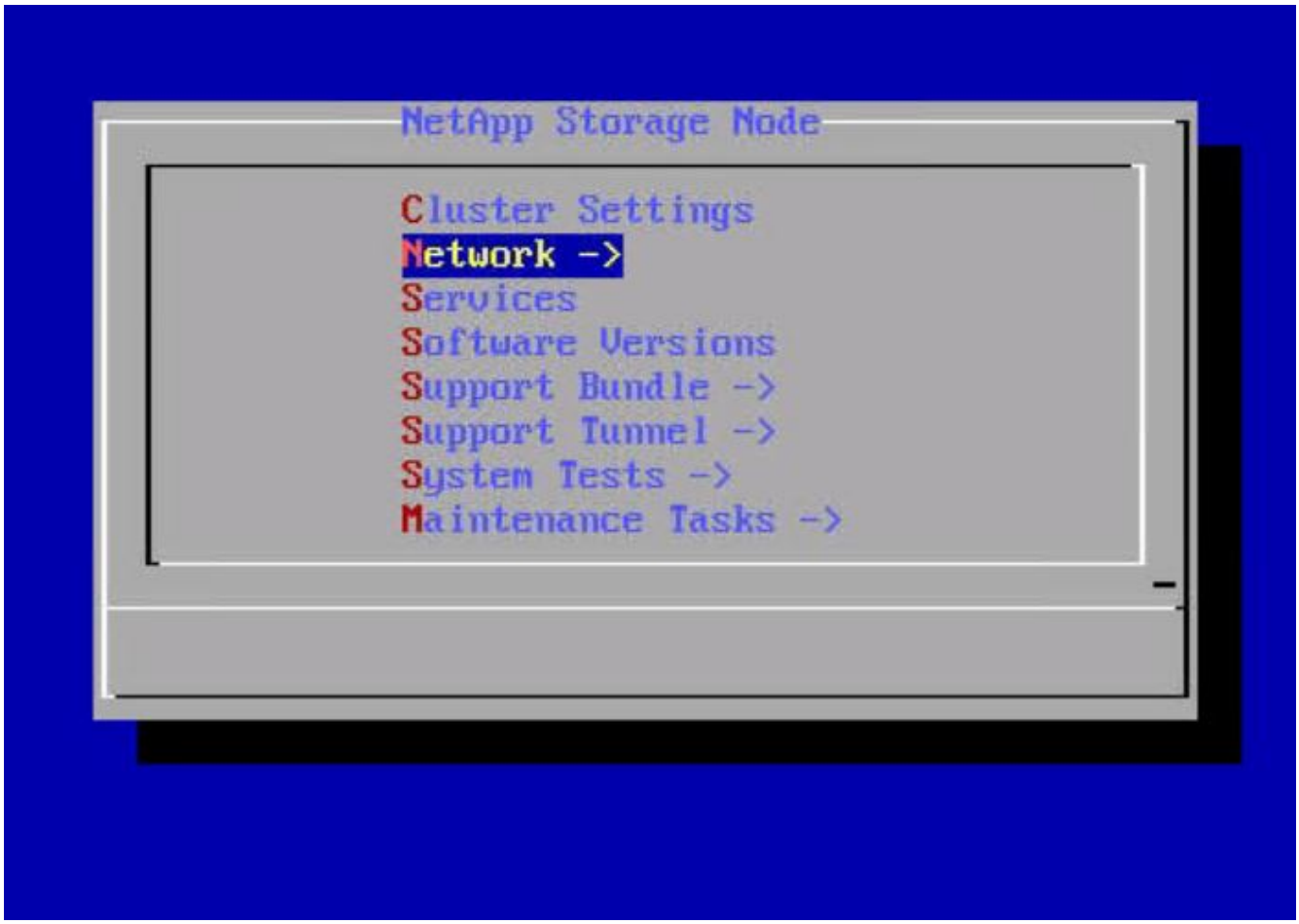
Basic NetApp Element Storage Setup

NetApp Element cluster setup is performed in a manner similar to a standalone NetApp SolidFire storage setup. These steps begin after the nodes have been racked, and cabled, and the IPMI port has been configured on each node using the console. To setup a storage cluster, complete the following steps:

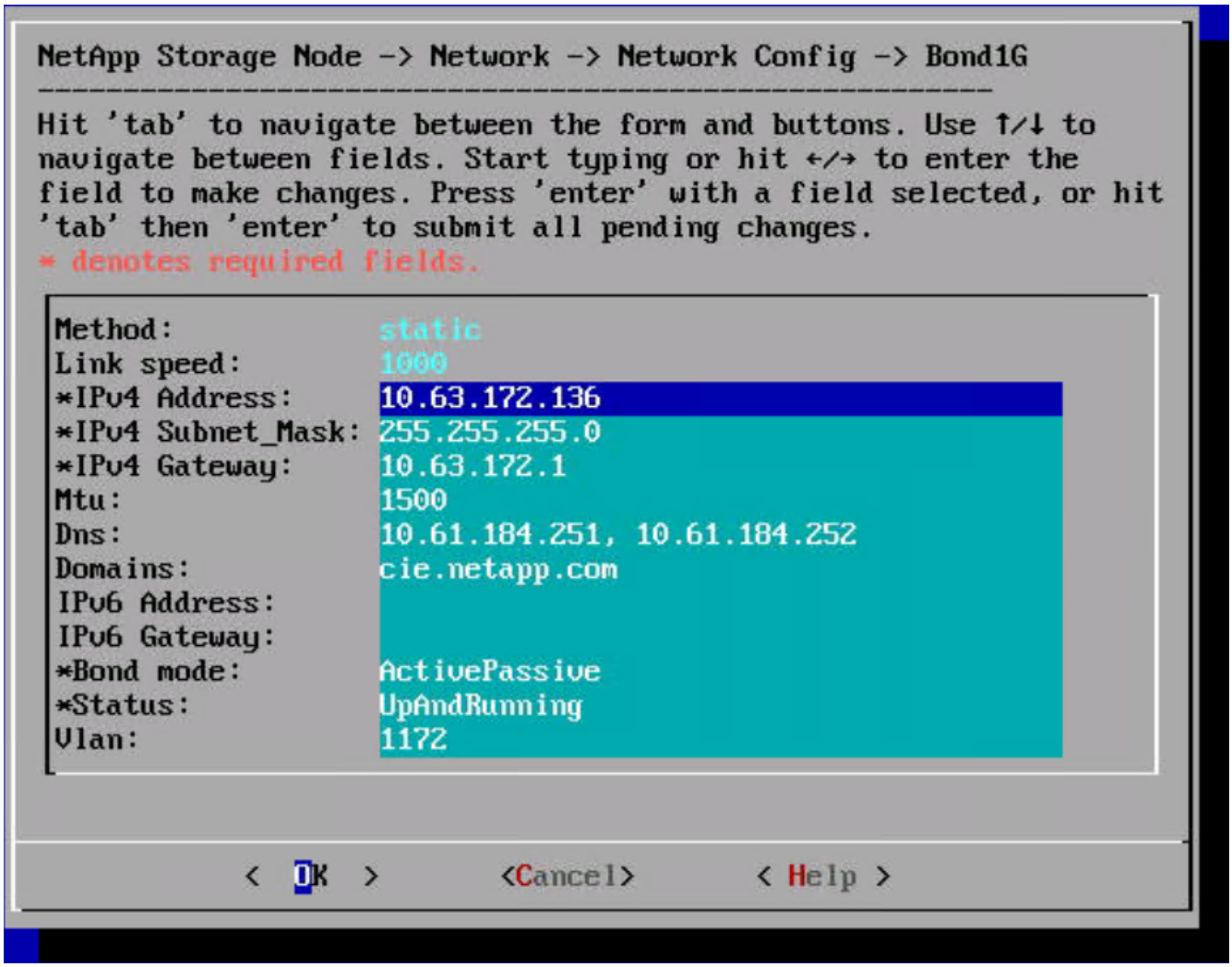
1. Access the out-of-band management console for the storage nodes in the cluster and log in with the default credentials ADMIN/ADMIN.



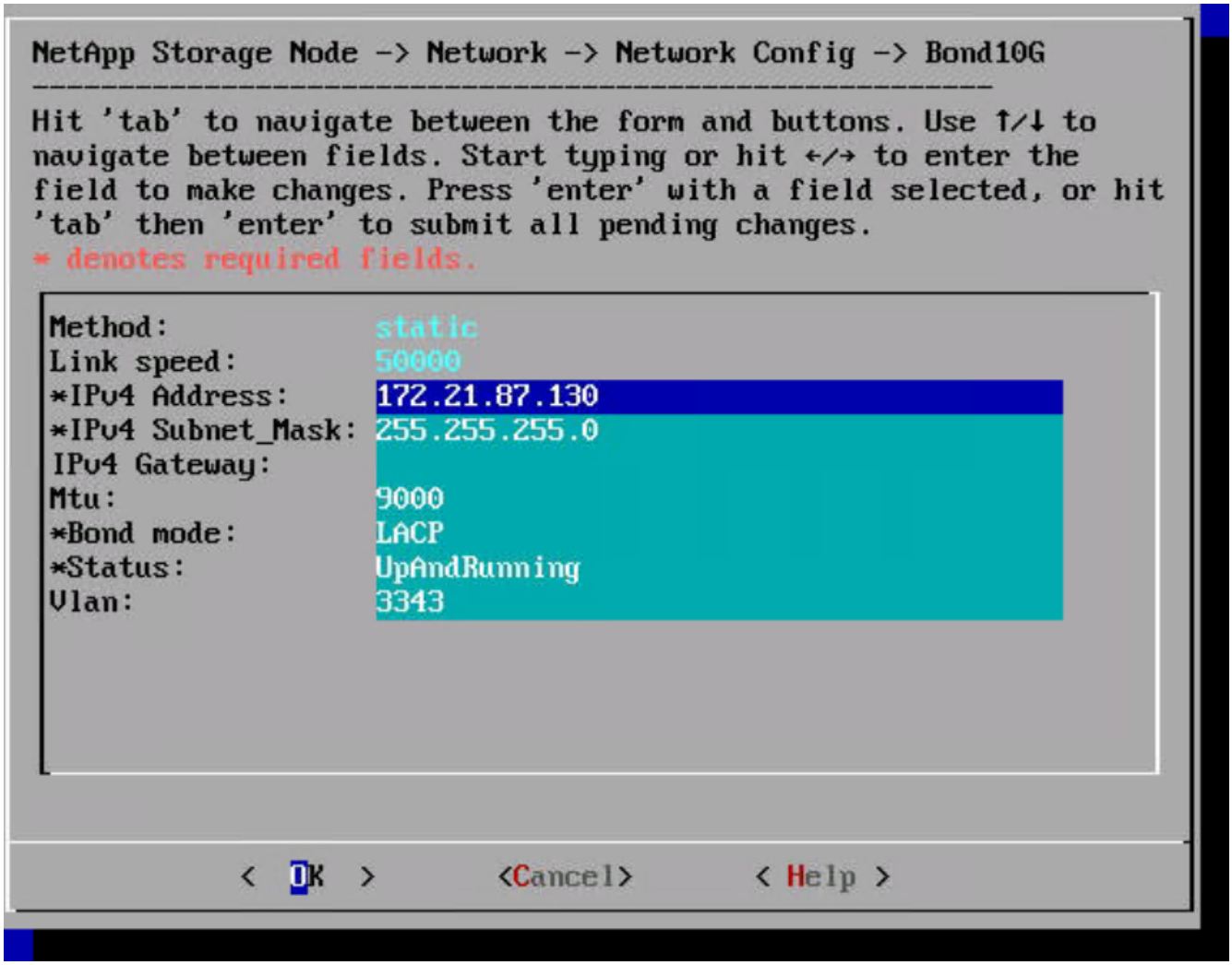
2. Click the Remote Console Preview image in the center of the screen to download a JNLP file launched by Java Web Start, which launches an interactive console to the system.



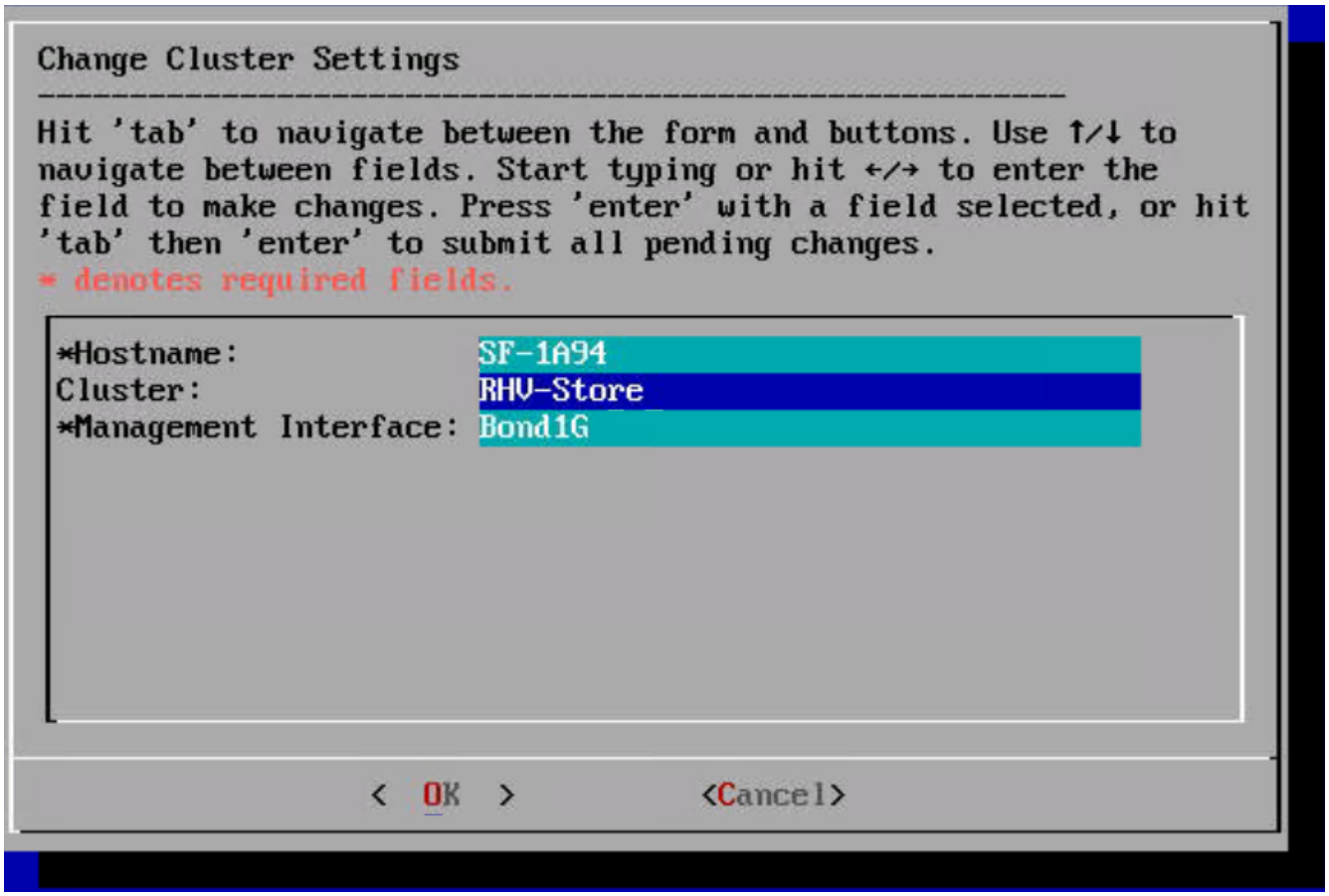
3. Navigate to Network > Network Config > Bond1G (Management) and configure the Bond1G interface. The Bond1G interface should be in ActivePassive bond mode and must have an IP, a netmask, and a gateway set statically. Its VLAN must correspond to IB Management network and DNS servers defined for the environment. Then click OK.



4. Select Bond10G (Storage) and configure the Bond10G interface. The Bond 10G interface must be in LACP bonding mode and have the MTU set to 9000 to enable jumbo frames. It must be assigned an IP address and netmask that are available on the defined storage VLAN. Click OK after entering the details.

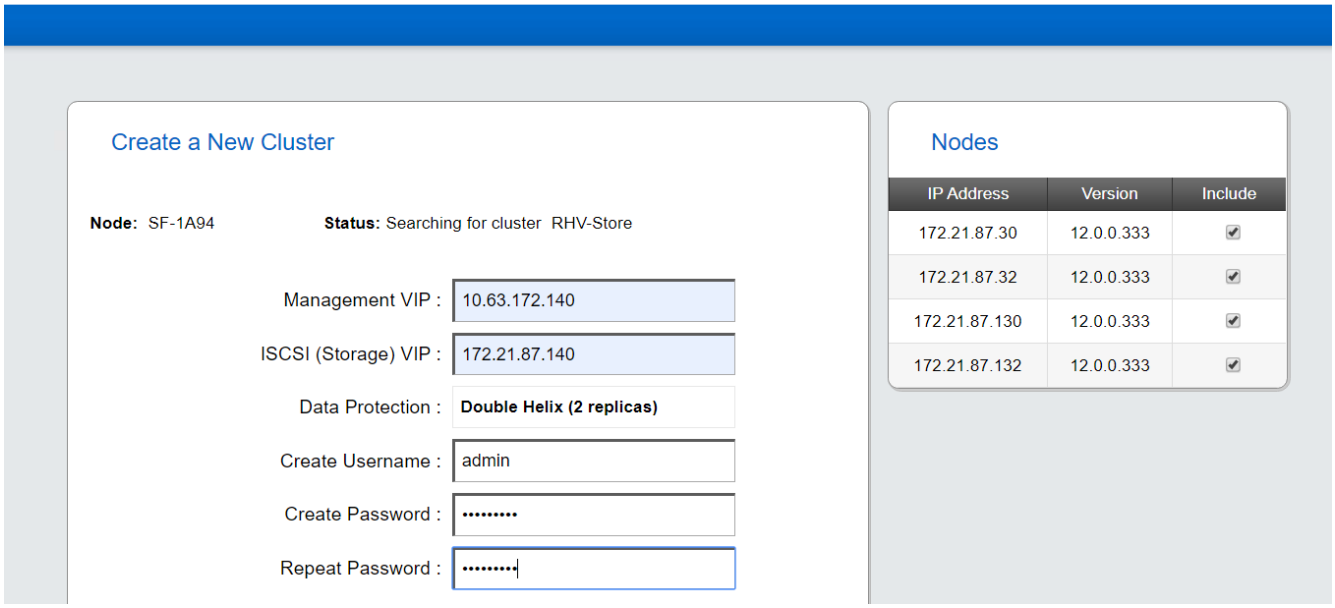


5. Go back to the initial screen, navigate to Cluster Settings, and click Change Settings. Enter the Cluster Name of your choice and click OK.

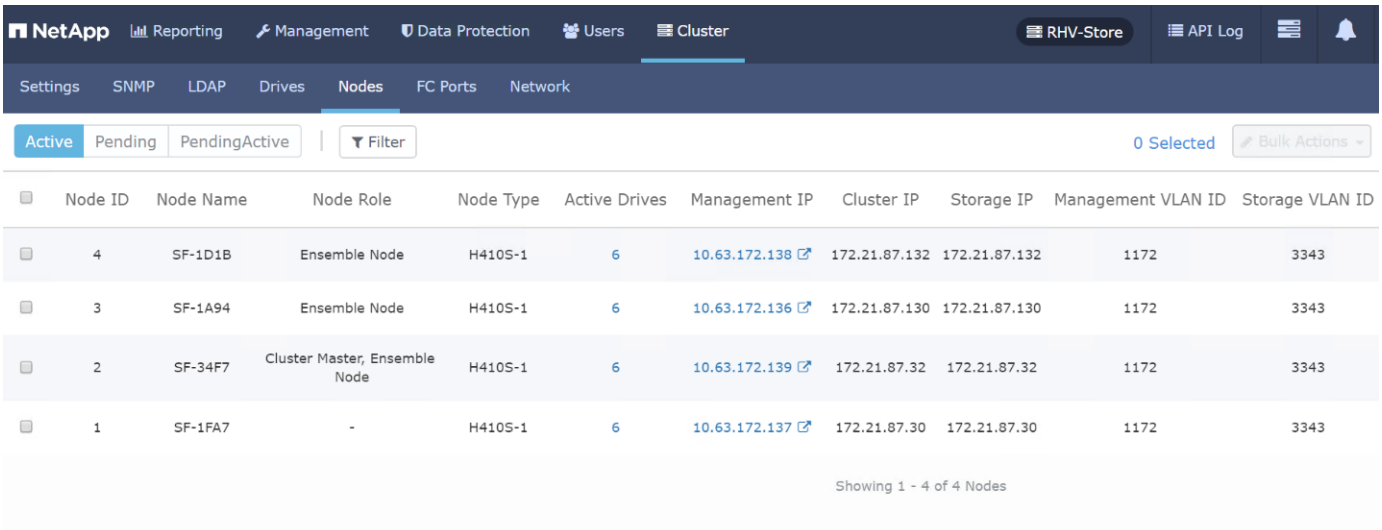


6. Repeat steps 1 to 5 for all HCI storage nodes.
7. After all the storage nodes are configured, use a web browser to log into the IB Management IP of one of the storage nodes. This presents the setup page with the Create a New Cluster dialog. Management VIP, storage VIP, and other details of the Element cluster are configured on this page. The storage nodes that were configured in the previous step are automatically detected. Make sure that any nodes that you do not want in the cluster are unchecked before proceeding. Accept the End User License Agreement and click Create New Cluster to begin the cluster creation process. It takes a few minutes to get the cluster up.

In some cases, visiting the IB management address automatically connects on port 442 and launches the NDE setup wizard. If this happens, delete the port specification from the URL and reconnect to the page.



8. After the cluster is created, it redirects to the Element cluster management interface available at the assigned MVIP address. Log in with the credentials provided in the previous step.
9. After you log in, the cluster automatically detects the number of available drives and requests for confirmation to add all drives. Click Add Drives to add all drives at once.
10. The Element cluster is ready to use. Navigate to Cluster > Nodes, and all four nodes should be in a healthy state with active drives.



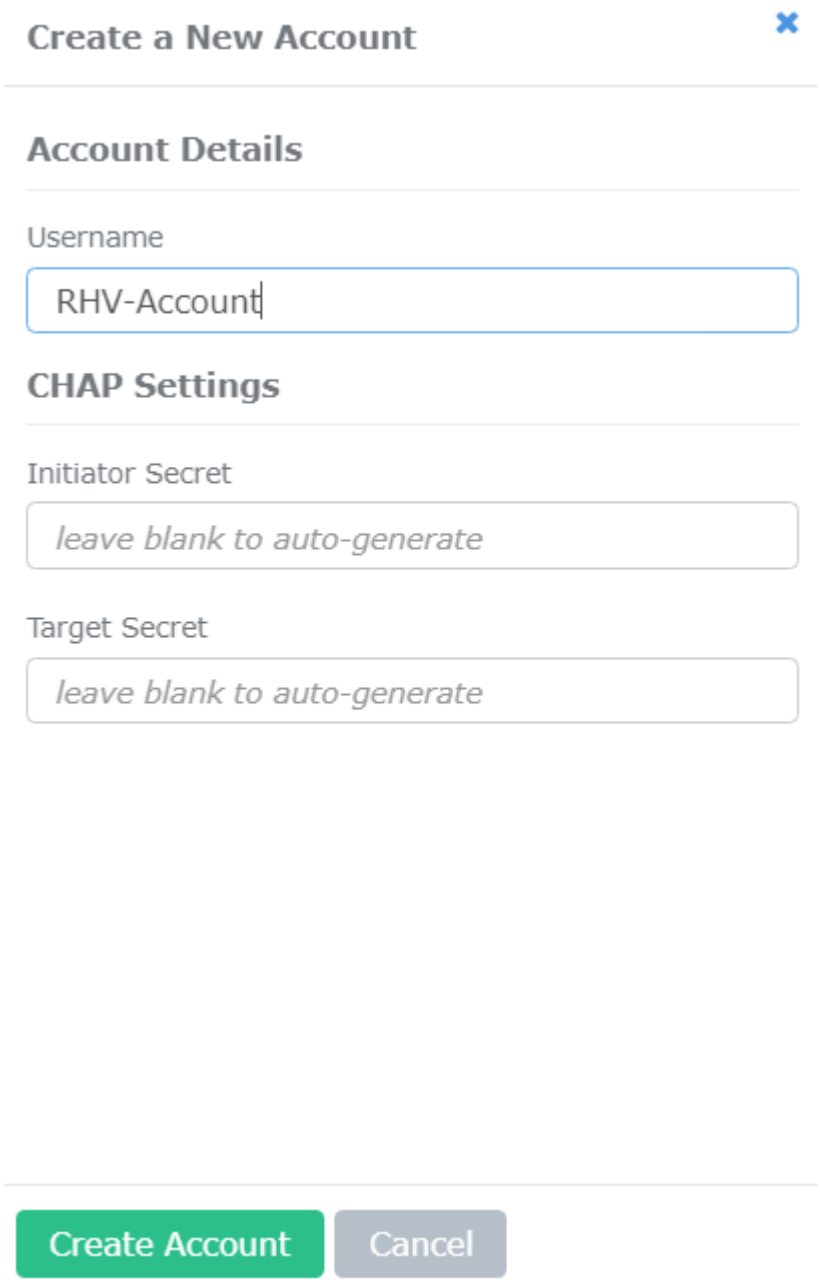
Element Storage Configuration to Support RHV Deployment

In our NetApp HCI for Red Hat Virtualization solution, we use a NetApp Element storage system to provide the backend storage support for RHV's requirement of shared storage domains. The self-hosted engine architecture of RHV deployment requires two storage domains at a minimum—one for the hosted engine storage domain and one for the guest VM data domain.

For this part of deployment, you must configure an account, two volumes of appropriate size, and the associated initiators. Then map these components to an access group that allows the RHV hosts to map the block volumes for use. Each of these actions can be performed through the web user interface or through the native API for the Element system. For this deployment guide, we go through the steps with the GUI.

Log in to the NetApp Element cluster GUI at its MVIP address using a web browser. Navigate to the Management tab and complete the following steps:

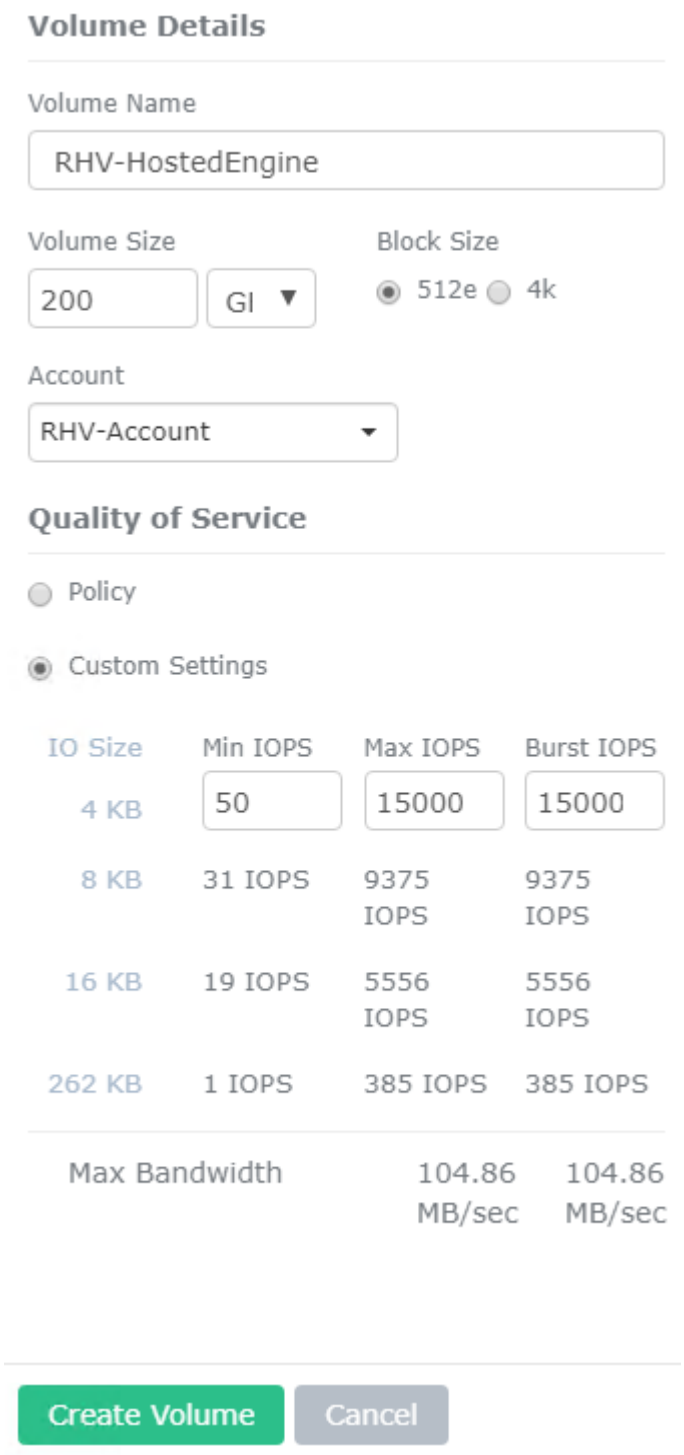
1. To create accounts, go to the Accounts sub-tab and click Create Account. Enter the name of your choice and click Create Account.



2. To create volumes, complete the following steps:

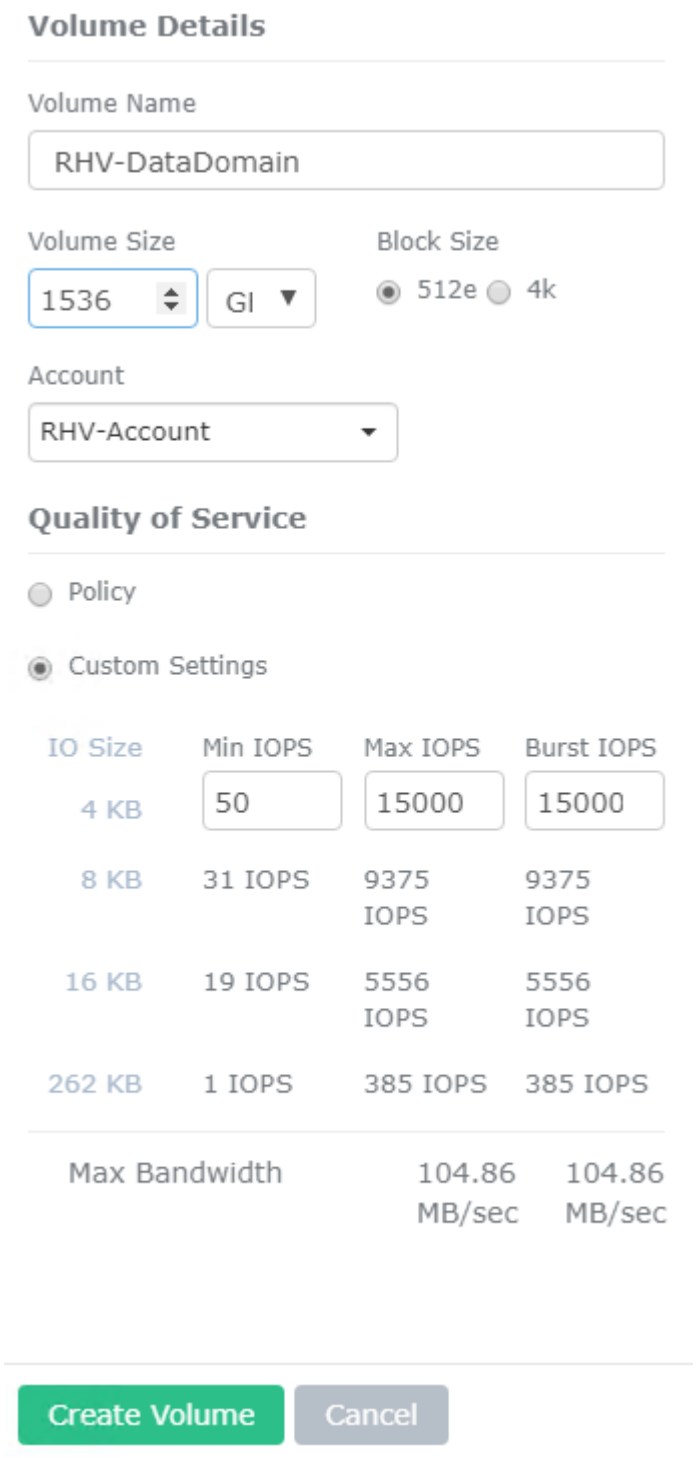
a. Navigate to the Volumes sub-tab and click Create Volume.

b. To create the volume for the self-hosted engine storage domain, enter the name of your choice, select the account you created in the last step, enter the size of the volume for the self-hosted engine storage domain, configure the QoS setting, and click Create Volume.



The minimum size for the hosted engine volume is 75GB. In our design, we added additional space to allow for future extents to be added to the RHV-M VM if necessary.

c. To create the volume for the guest VMs data storage domain, enter the name of your choice, select the account you created in the last step, enter the size of the volume for the data storage domain, configure the QoS setting and click Create Volume.

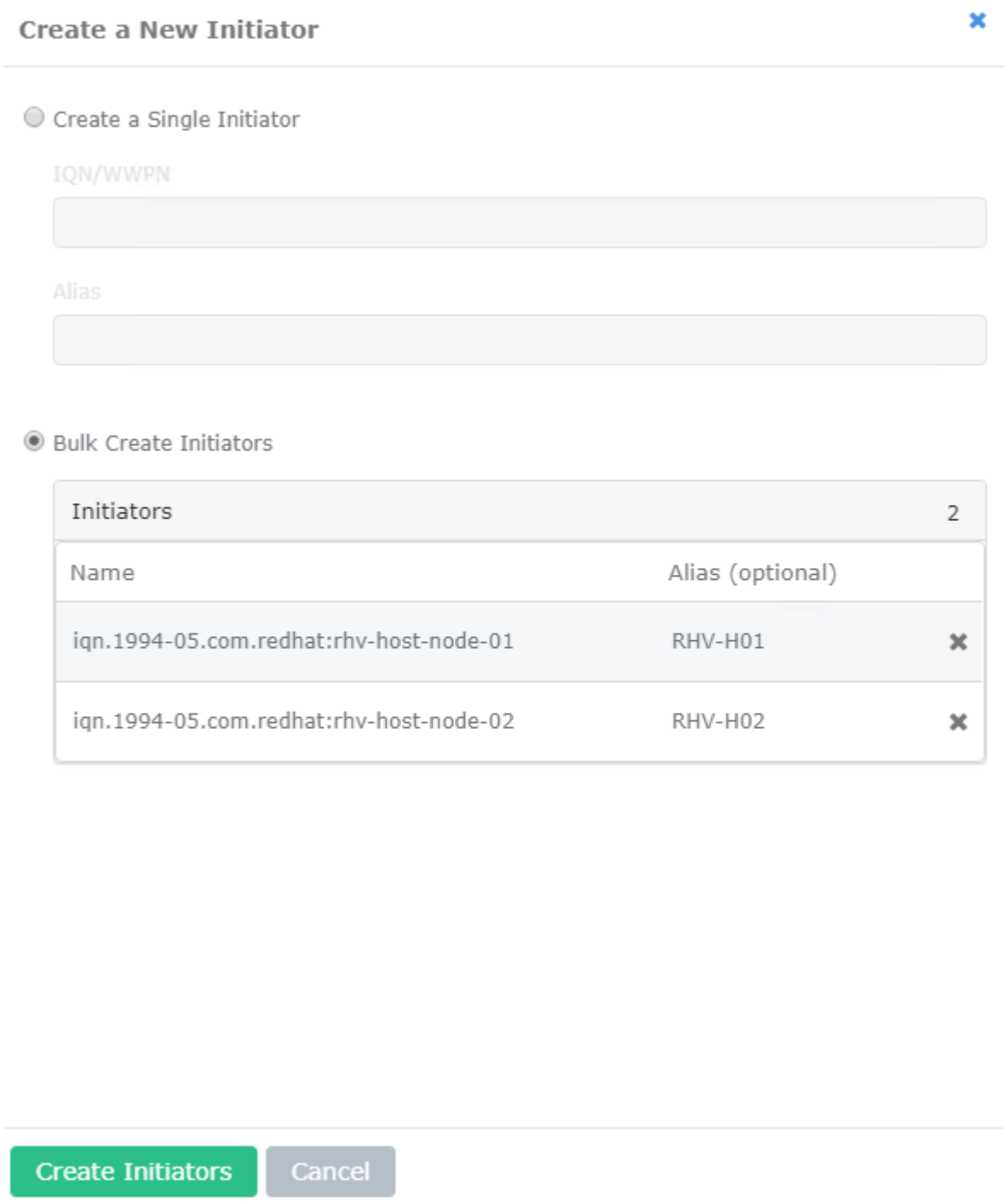


The size of the data domain depends on the kind of VMs run in the environment and the space required to support them. Adjust the size of this volume to meet the needs of your environment.

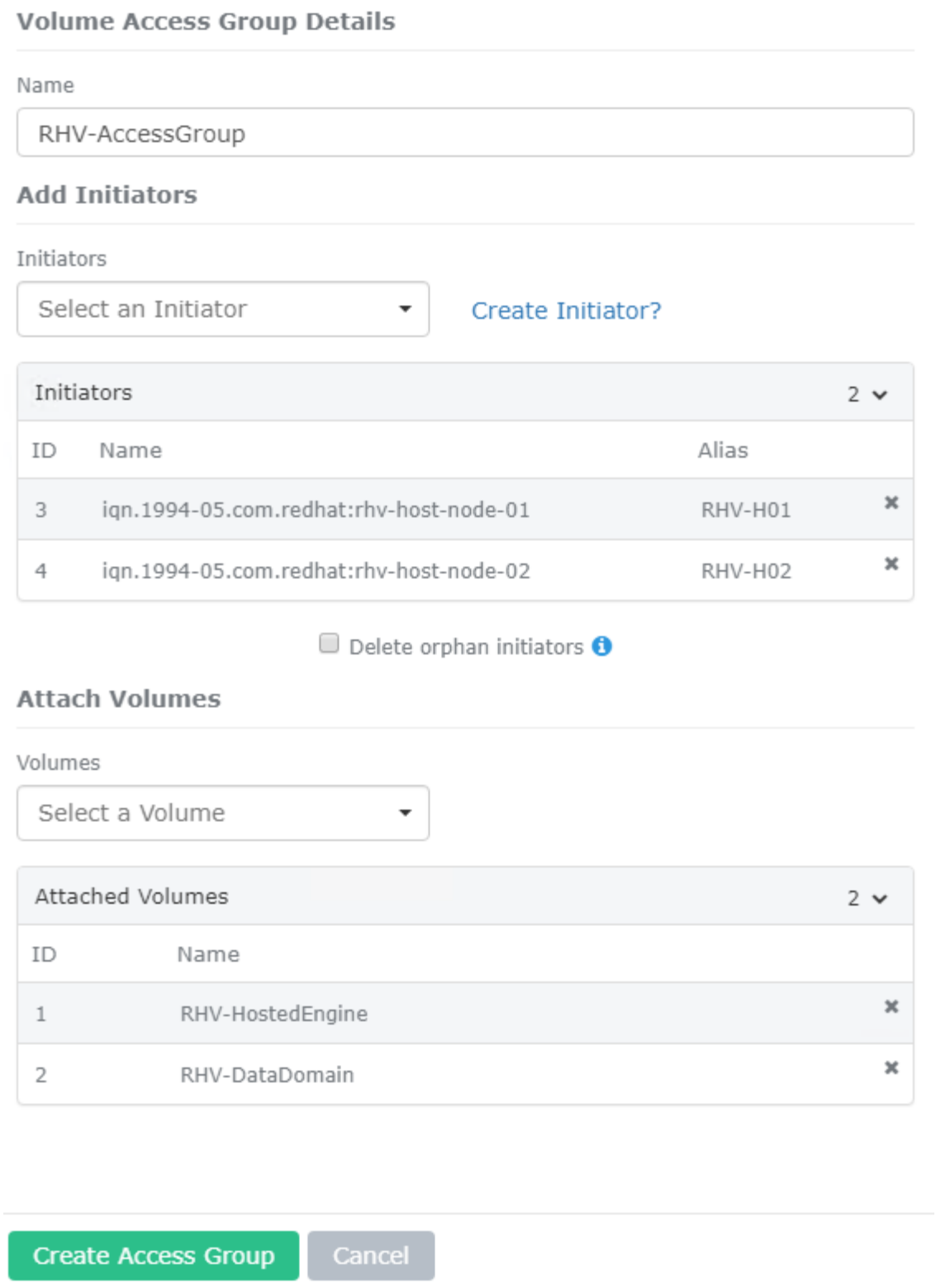
3. To create initiators, complete the following steps:

   a. Go to the Initiators sub-tab and click Create Initiator.

b. Select the Bulk Create Initiators radio button and enter the initiators' details of both the RHV-H nodes with comma separated values. Then click Add Initiators, enter the aliases for the initiators, and click the tick button. Verify the details and click Create Initiators.



4. To create access groups, complete the following steps:
   a. Go to the Access Groups sub-tab and click Create Access Groups.
   b. Enter the name of your choice, select the initiators for both RHV-H nodes that were created in the previous step, select the volumes, and click Create Access Group.

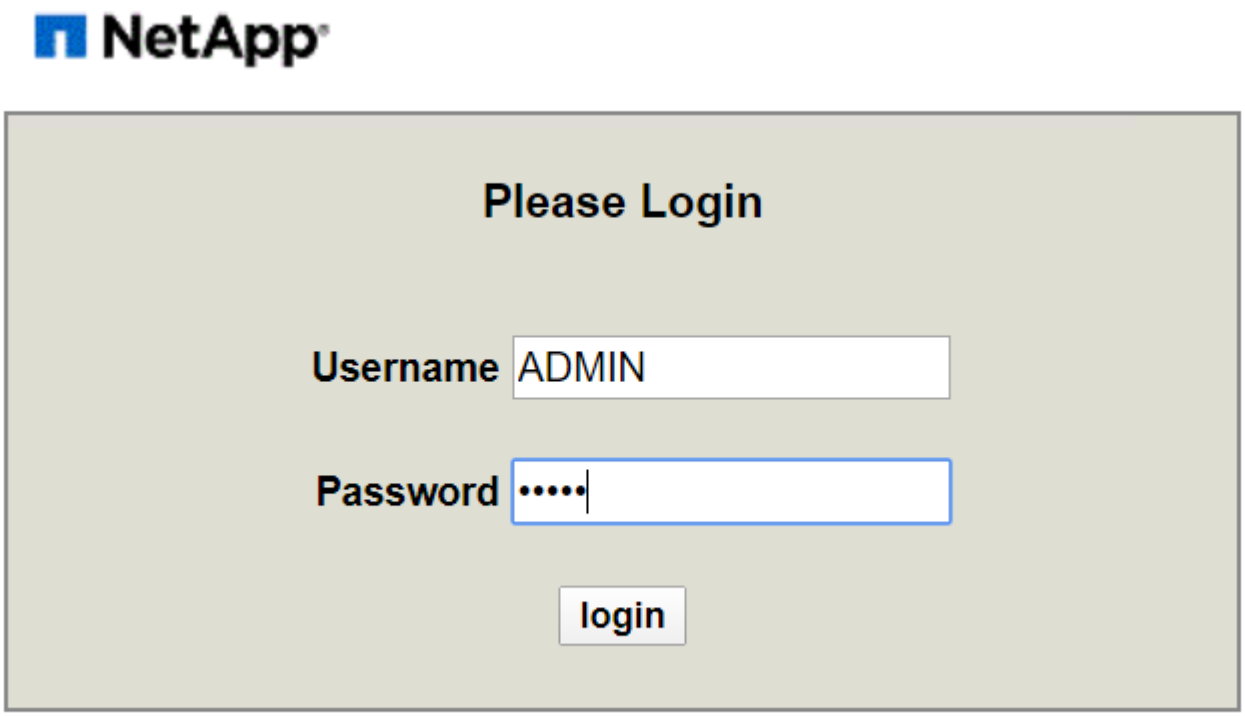


4. Deploy the RHV-H Hypervisor on the HCI Compute Nodes: NetApp HCI with RHV

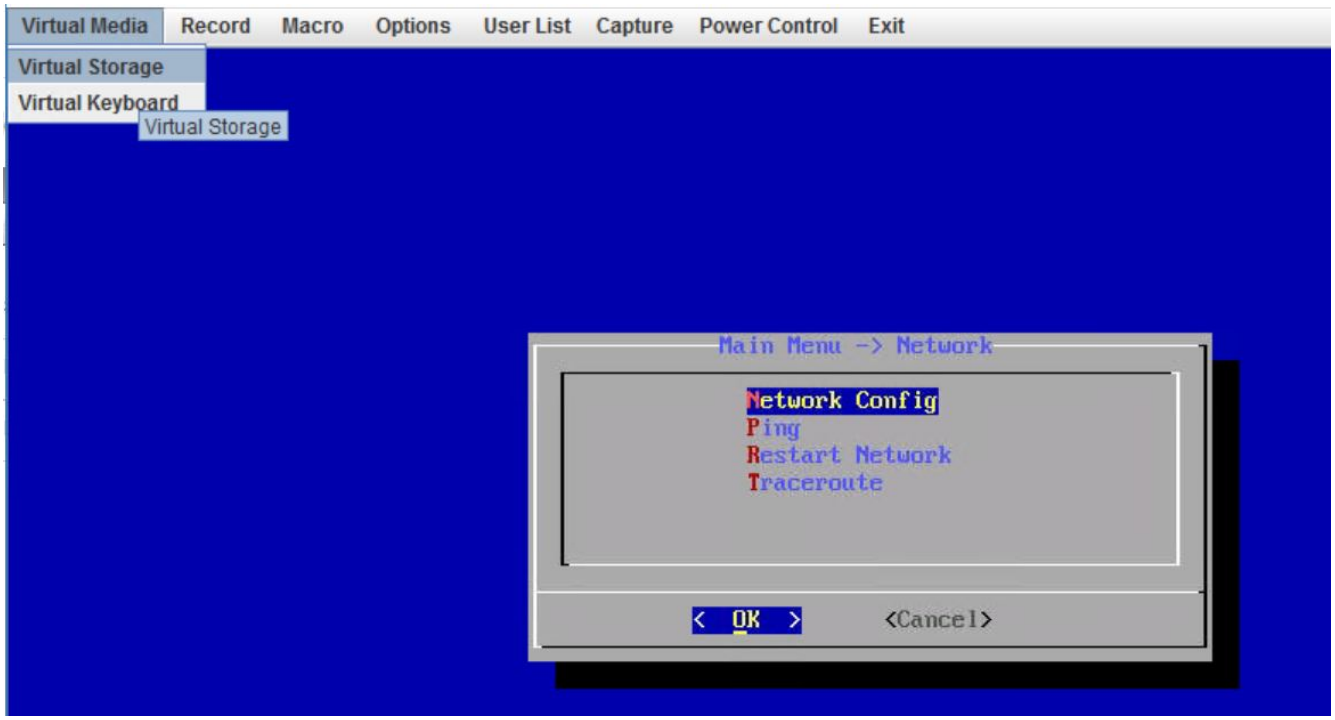
This solution employs the recommended self-hosted engine architecture of RHV deployment with the minimum setup (two self-hosted engine nodes). These steps begin after the nodes have been racked and cabled and the IPMI port has been

configured on each node for using the console. To deploy the RHV-H hypervisor on HCI compute nodes, complete the following steps:

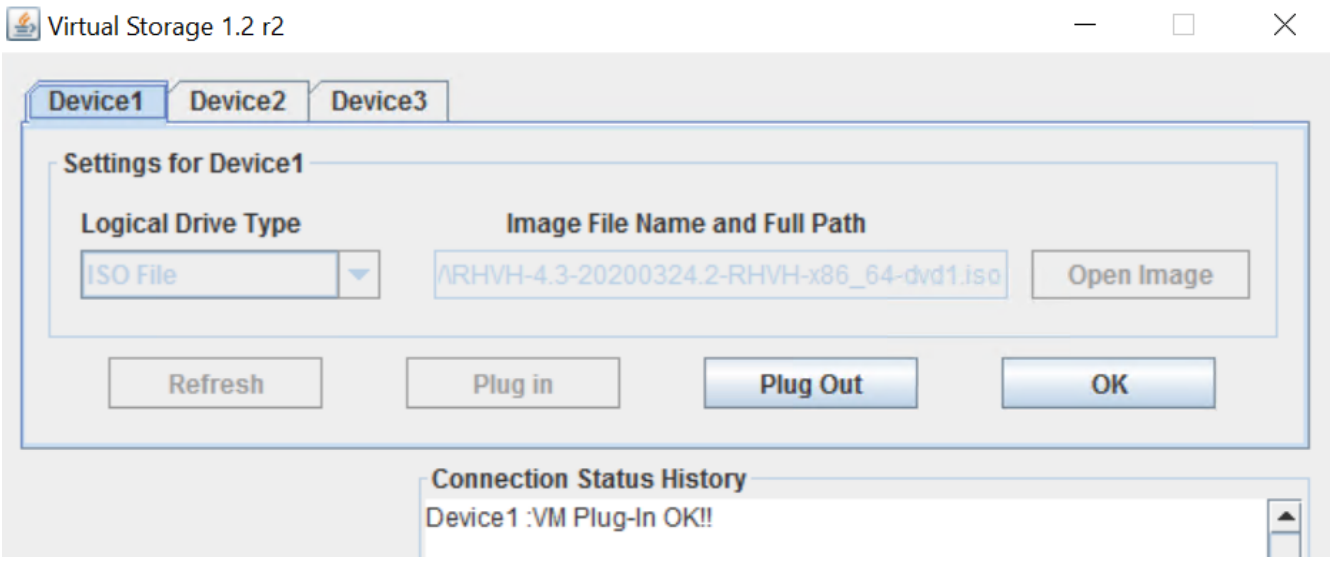
1. Access the out-of-band management console for the compute nodes in the cluster and log in with the default credentials ADMIN/ADMIN.



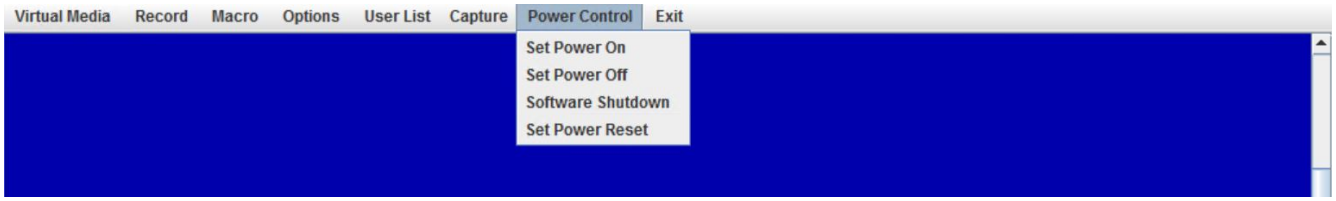
2. Click the Remote Console Preview image in the center of the screen to download a JNLP file launched by Java Web Start, which launches an interactive console to the system.
3. After the virtual console launches, attach the RHV-H 4.3.9 ISO by navigating to and clicking Virtual Media > Virtual Storage.



4. For Logical Drive Type, select ISO File from the drop down. Provide the full path and full name of the RHV-H 4. 3.9 ISO file or attach it by clicking the Open Image button. Then click Plug In.

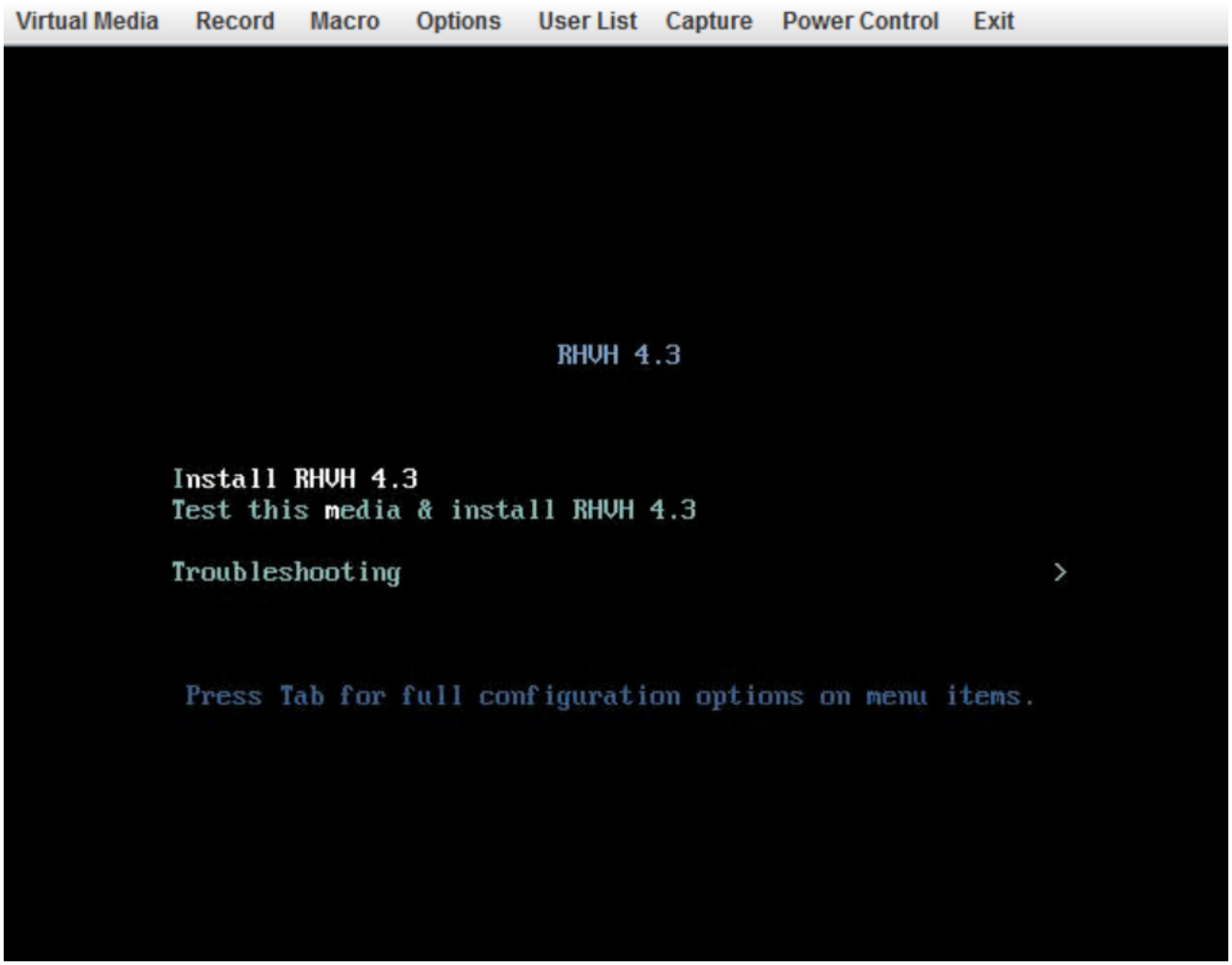


5. Reboot the server so that it boots using RHV-H 4.3.9 ISO by navigating and clicking Power Control > Set Power Reset.

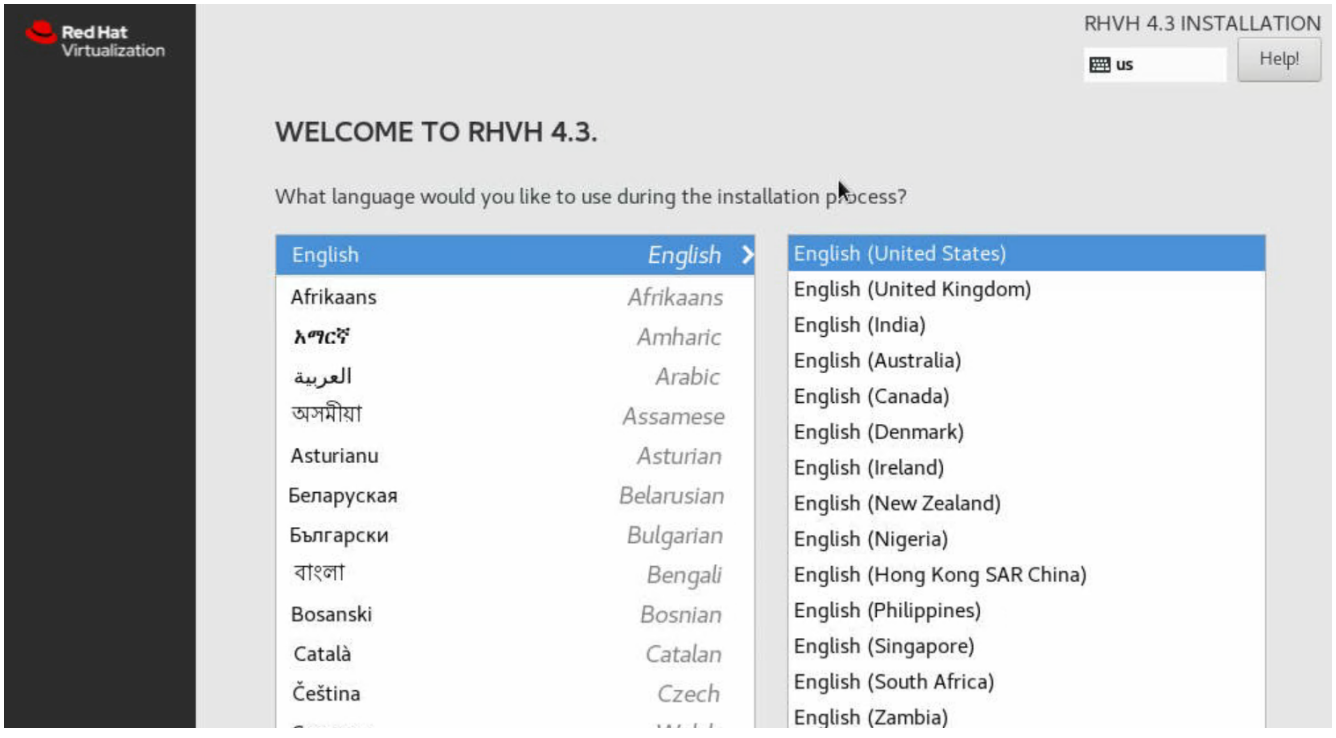


6. When the node reboots and the initial screen appears, press F11 to enter the boot menu. From the boot menu, navigate to and click ATEN Virtual CDROM YSOJ.

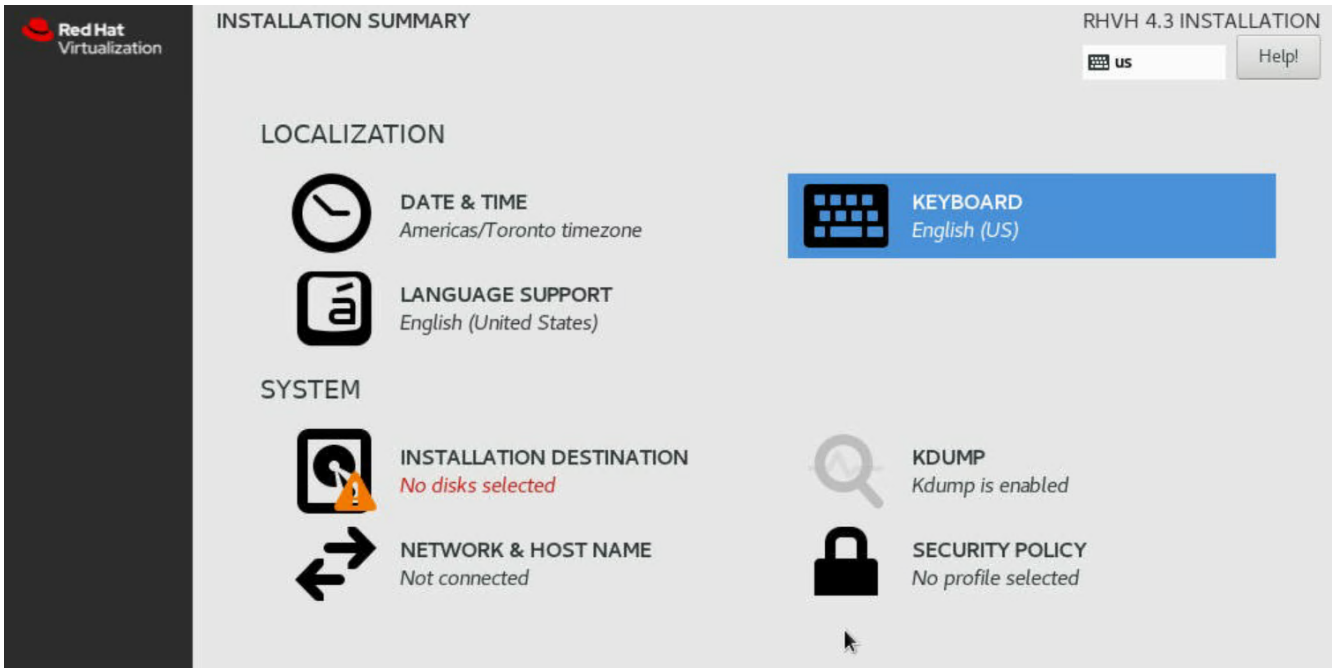
7. On the next screen, navigate to and click Install RHV 4.3. This loads the image, runs the pre-installation scripts, and starts Anaconda, the Red Hat Enterprise Linux system installer.



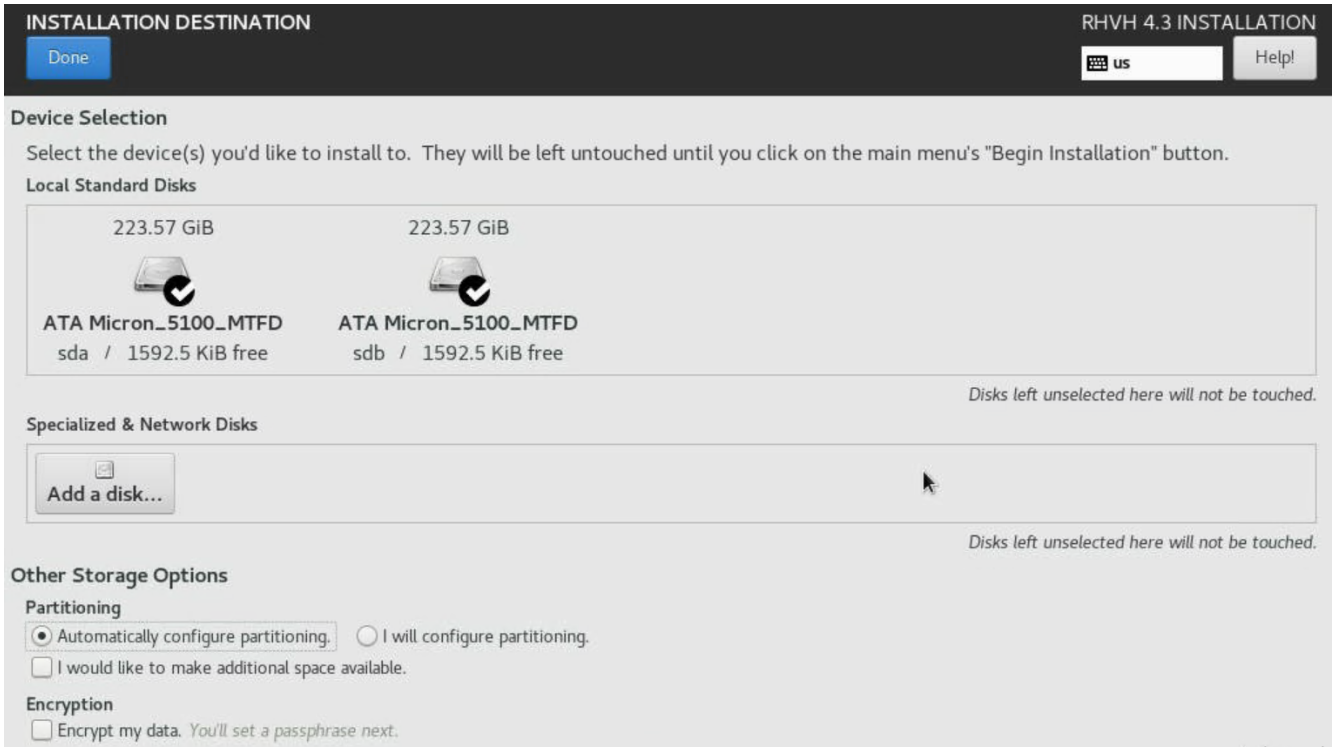
8. The installation welcome screen appears. Select the preferred language and click Next.



9. In the next screen, select your time zone under Date & Time. The default is UTC. However, NetApp recommends that you configure NTP servers for your environment on this screen. Then select the keyboard language and click Done.



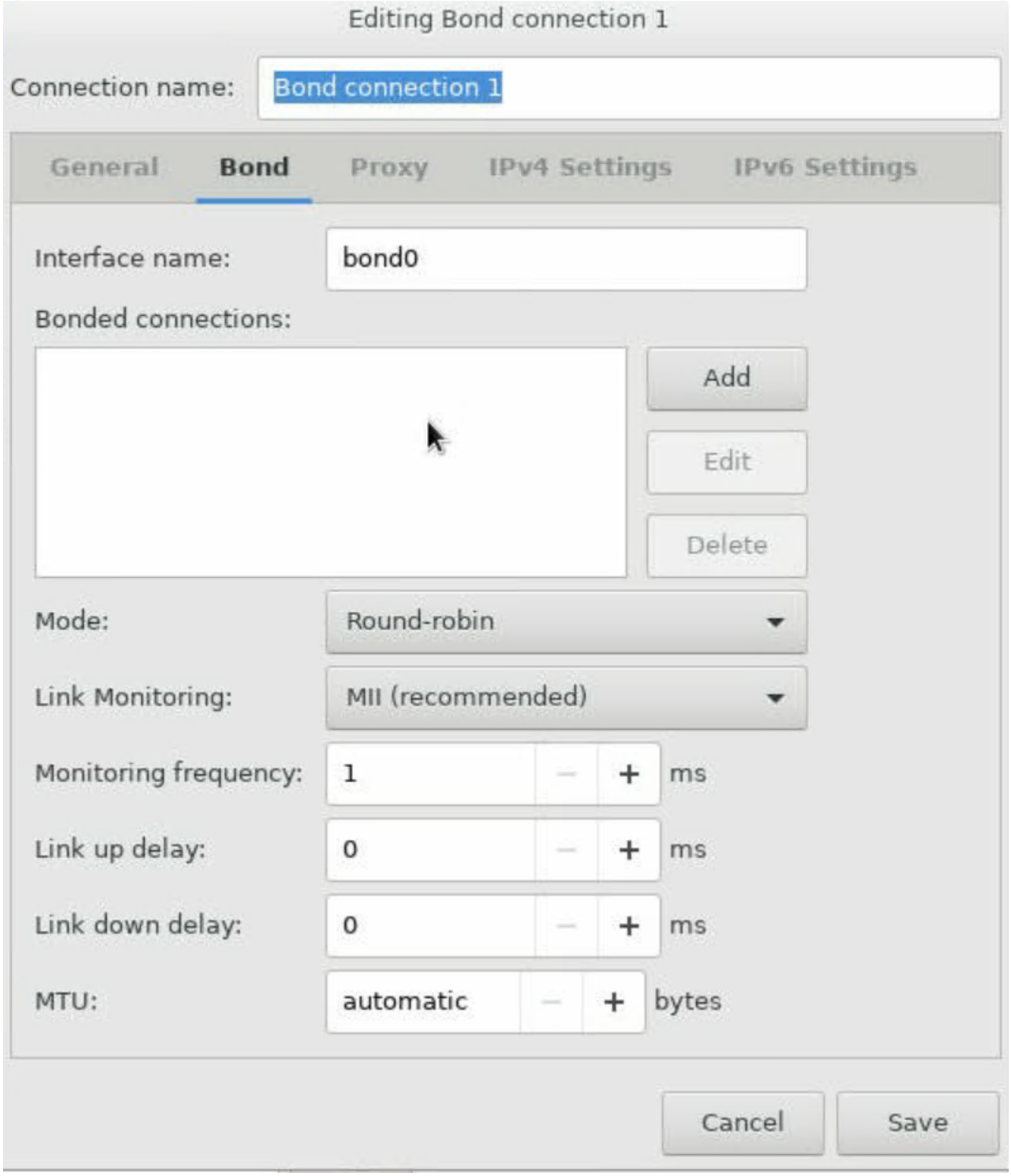
10. Next, click Installation Destination. In the Installation Destination screen, select the drives on which you want to install RHV-H. Verify that Automatically Configure Partitioning is selected in the Partitioning section. Optionally, you can enable encryption by checking the box next to Encrypt My Data. Click Done to confirm the settings.



11. Click Network & Host Name. Provide the desired host name at the bottom of the screen. Then click

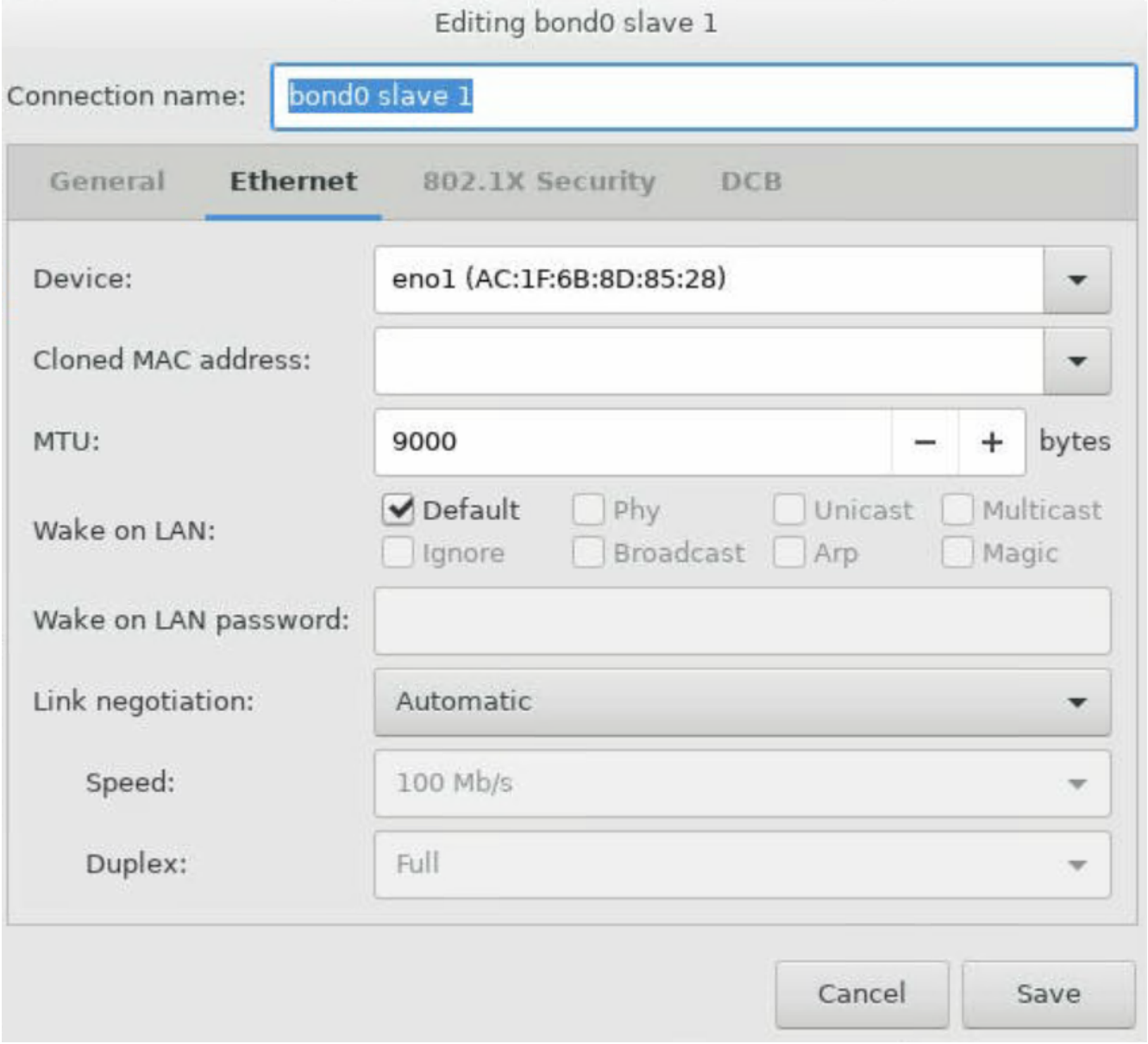
the (+) button at the bottom. Select the Bond from the drop down and click Add.

12. Next, in the bond configuration screen, click Add to add the member interfaces to the bond interface.

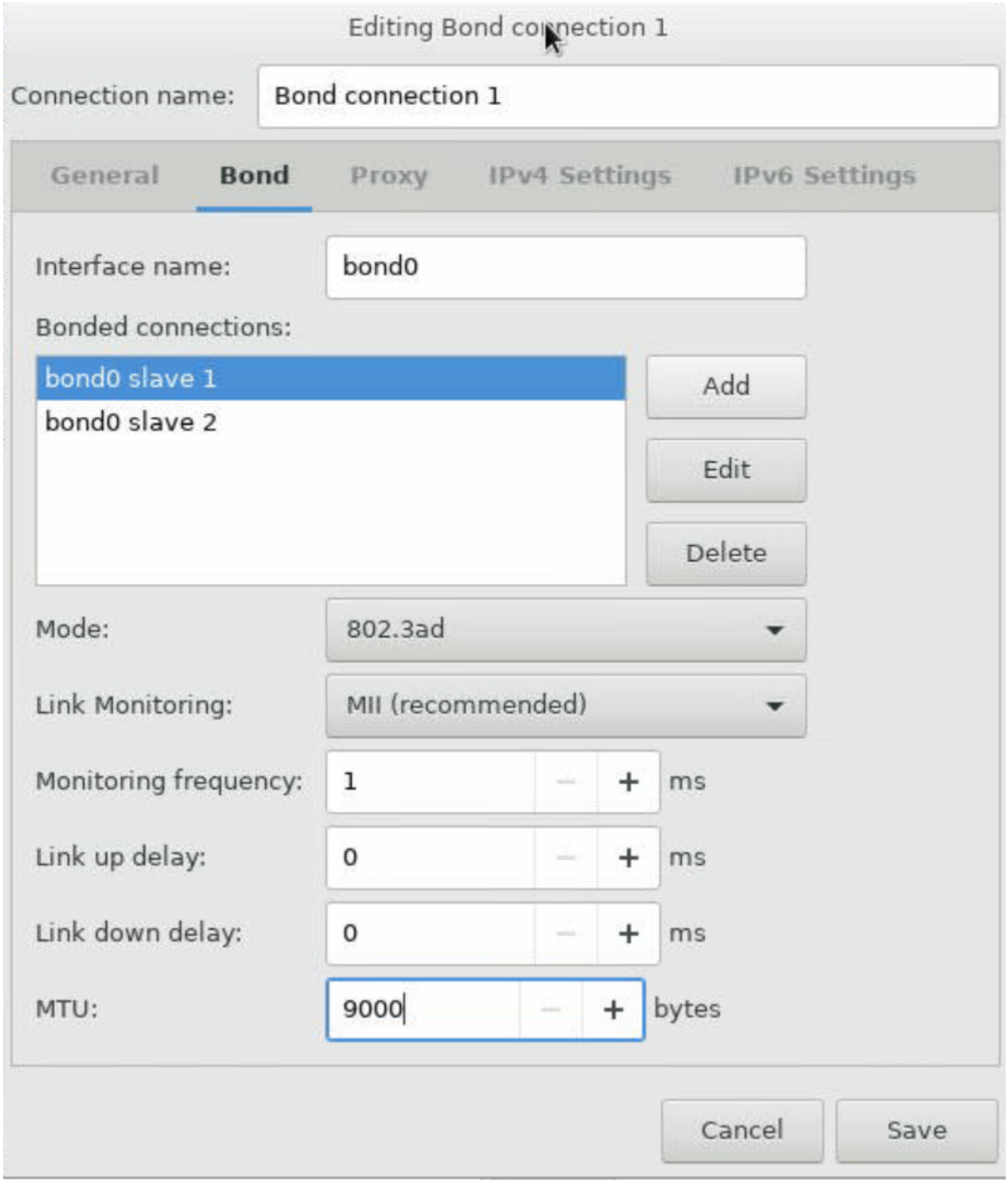


13. Select Ethernet from the drop down, indicating that the Ethernet interface is added as a member to the bond interface. Click Create.

14. From the Device dropdown in the slave 1 configuration screen, select the Ethernet interface. Verify that the MTU is set to 9000. Click Save.

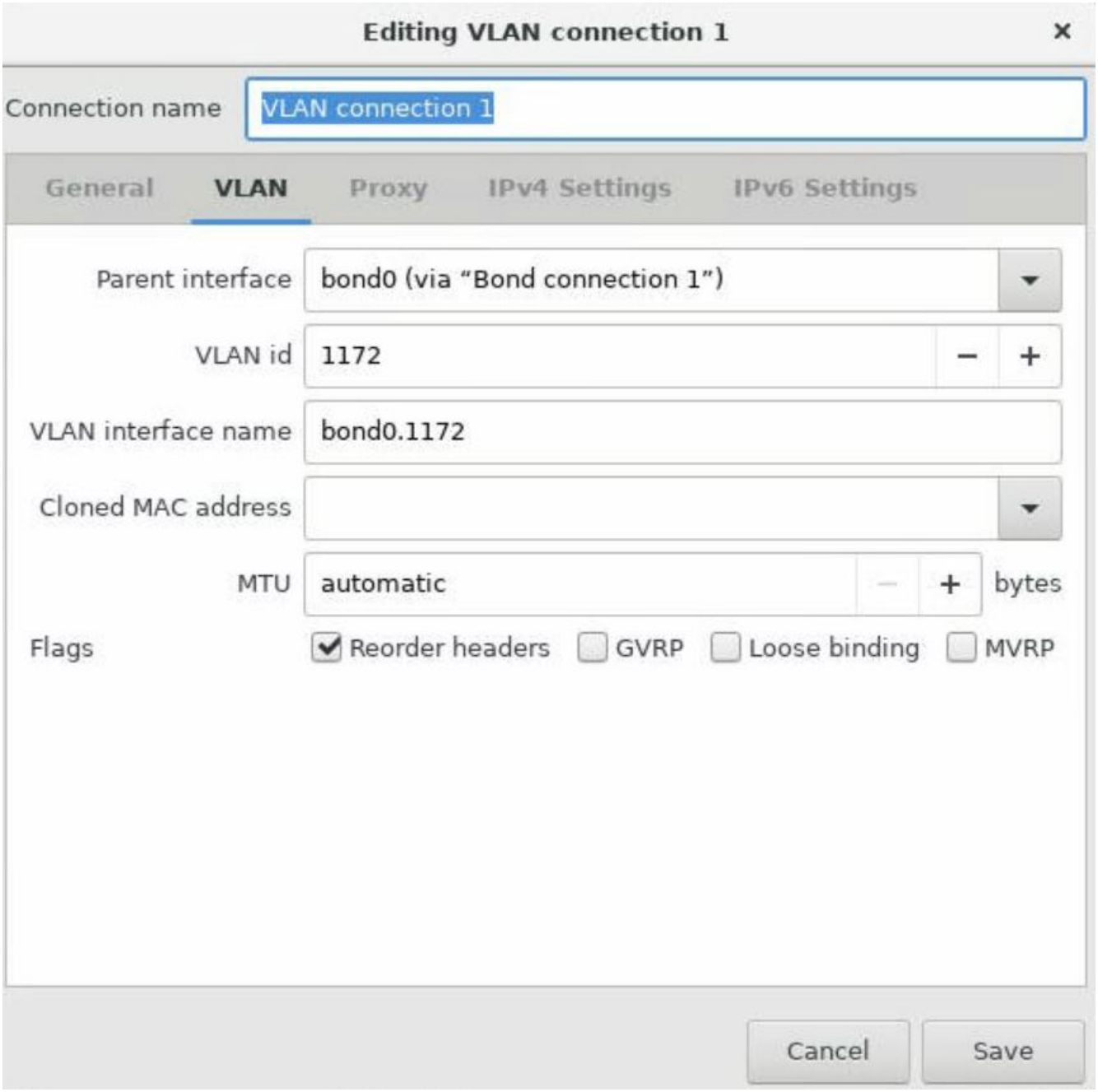


15. Repeat steps 12, 13, and 14 to add the other Ethernet port to the bond0 interface.
16. From the Mode dropdown in the bond configuration screen, select 802.3ad for LACP. Verify that the MTU is set to 9000. Then click Save.

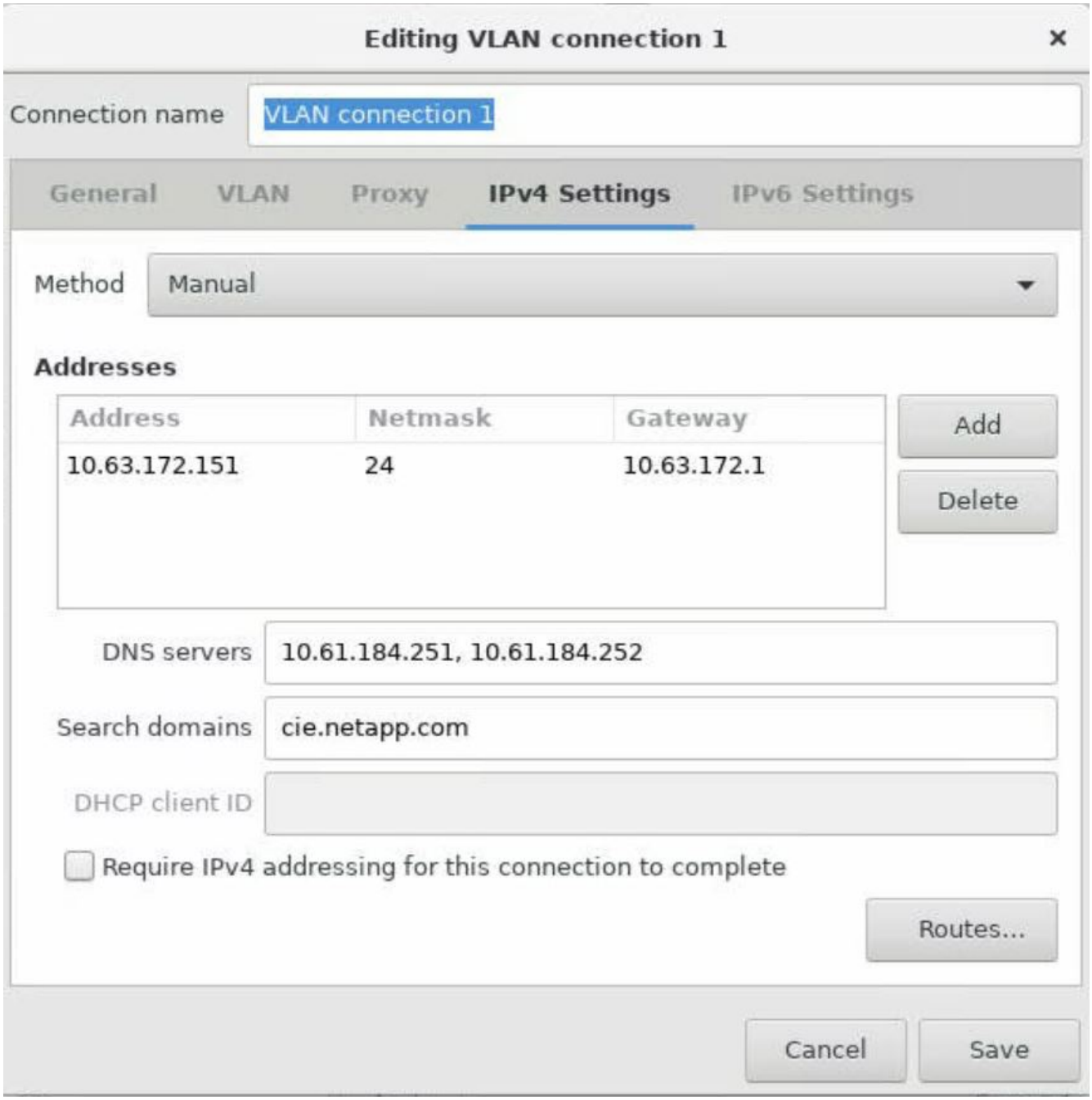


17. Create the VLAN interface for the in-band management network. Click the (+) button again, select VLAN from the dropdown and click Create.

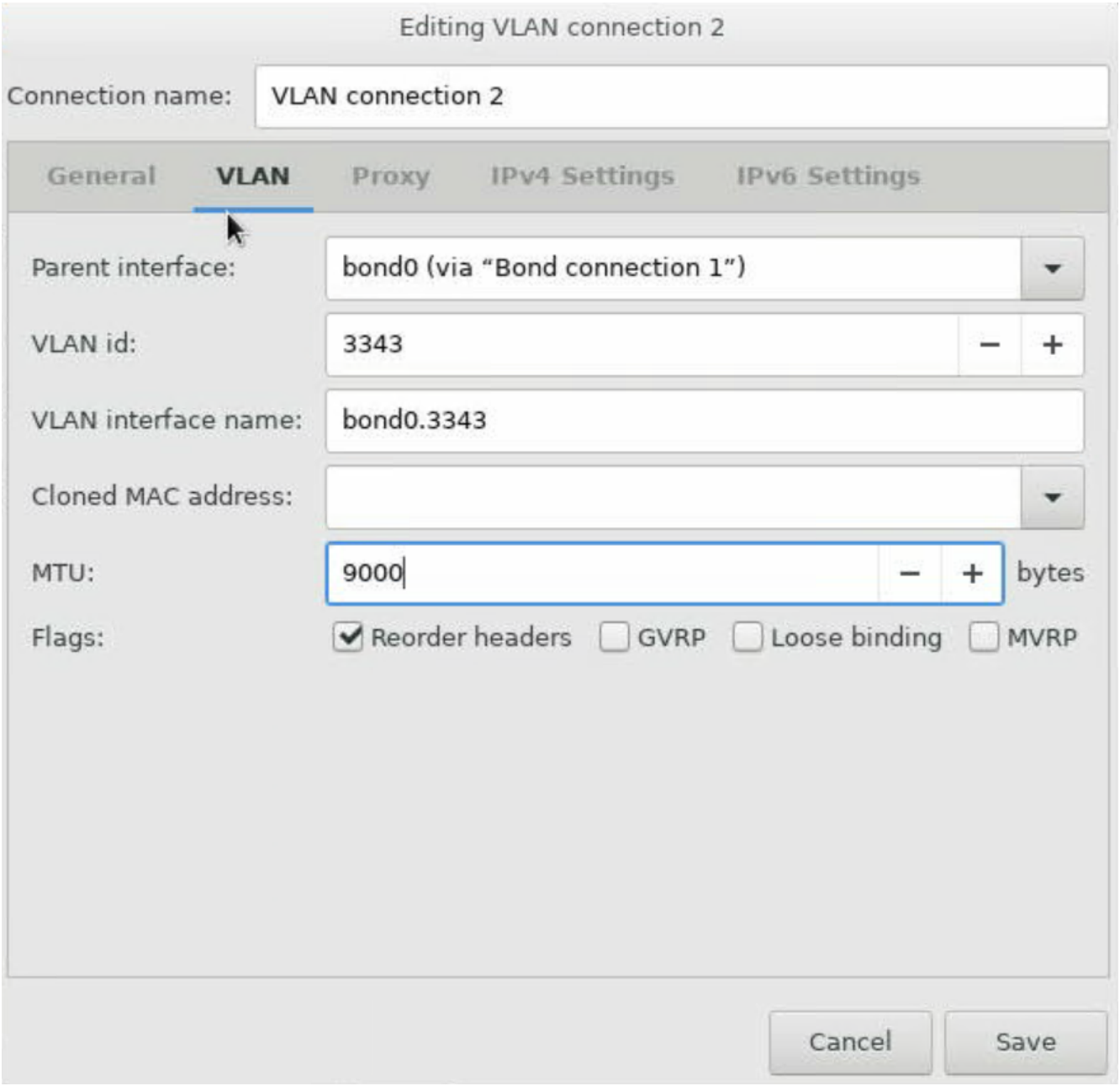
18. In the Editing VLAN connection screen, select bond0 in the Parent Interface dropdown, enter the VLAN ID of the in-band management network. Provide the name of the VLAN interface in `bond 0.< vlan_id >` format.



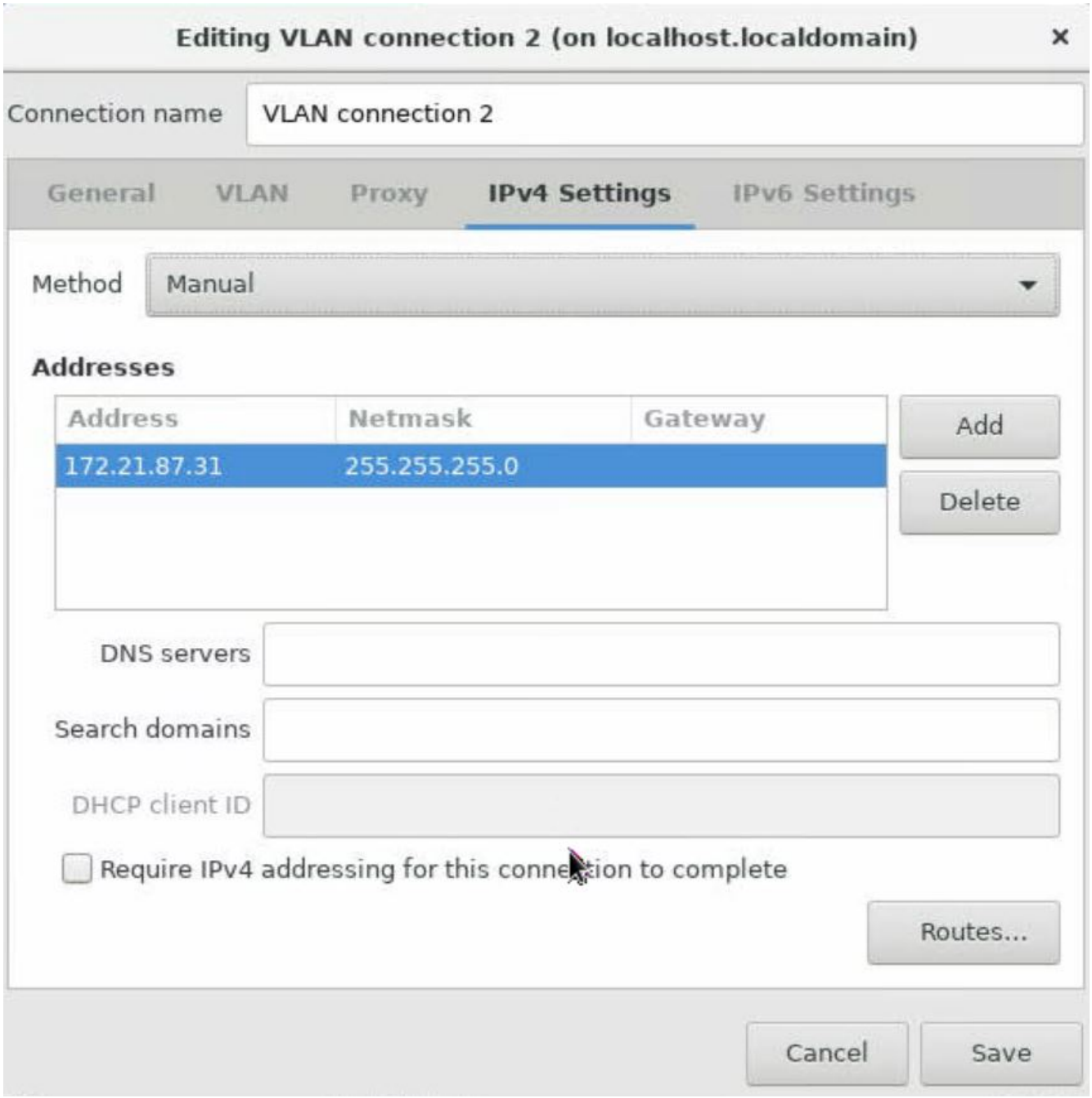
19. In the Editing VLAN connection screen, click the IPv4 Settings sub-tab. In the IPv4 Settings sub-tab, configure the network address, netmask, gateway, and DNS servers corresponding to the in-band management network. Click Save to confirm the settings.



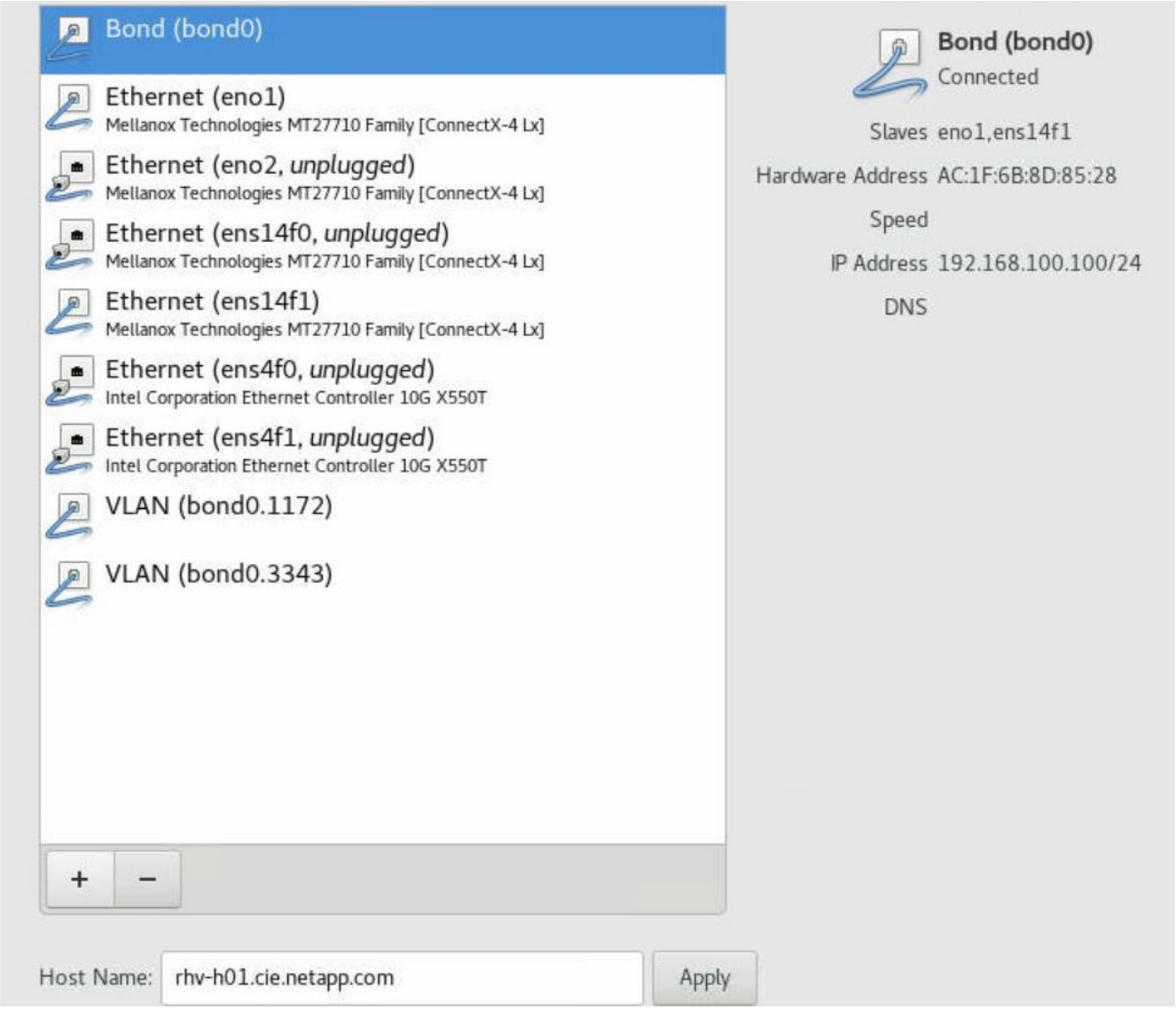
20. Create the VLAN interface for the storage network. Click the (+) button again, select VLAN from the dropdown, and click Create. In the Editing VLAN Connection screen, select bond0 in the Parent Interface dropdown, enter the VLAN ID of the storage network, provide the name of the VLAN interface in the `bond 0.< vlan_id >` format. Adjust the MTU to 9000 to allow jumbo frame support. Click Save.



21. In the Editing VLAN Connection screen, click the IPv4 Settings sub-tab. In the IPv4 Settings sub-tab, configure the network address and the netmask corresponding to the storage network. Click Save to confirm the settings.



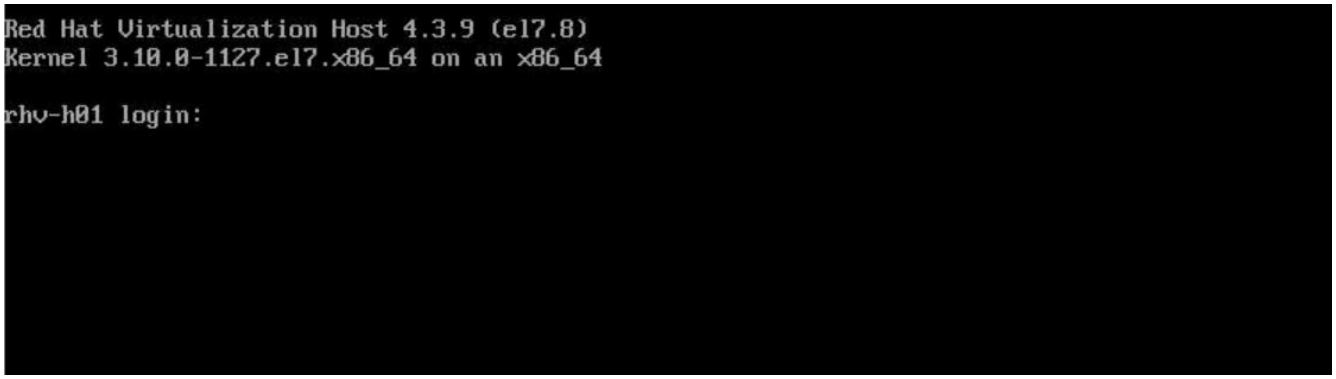
22. Confirm that the network interfaces are up and click Done.



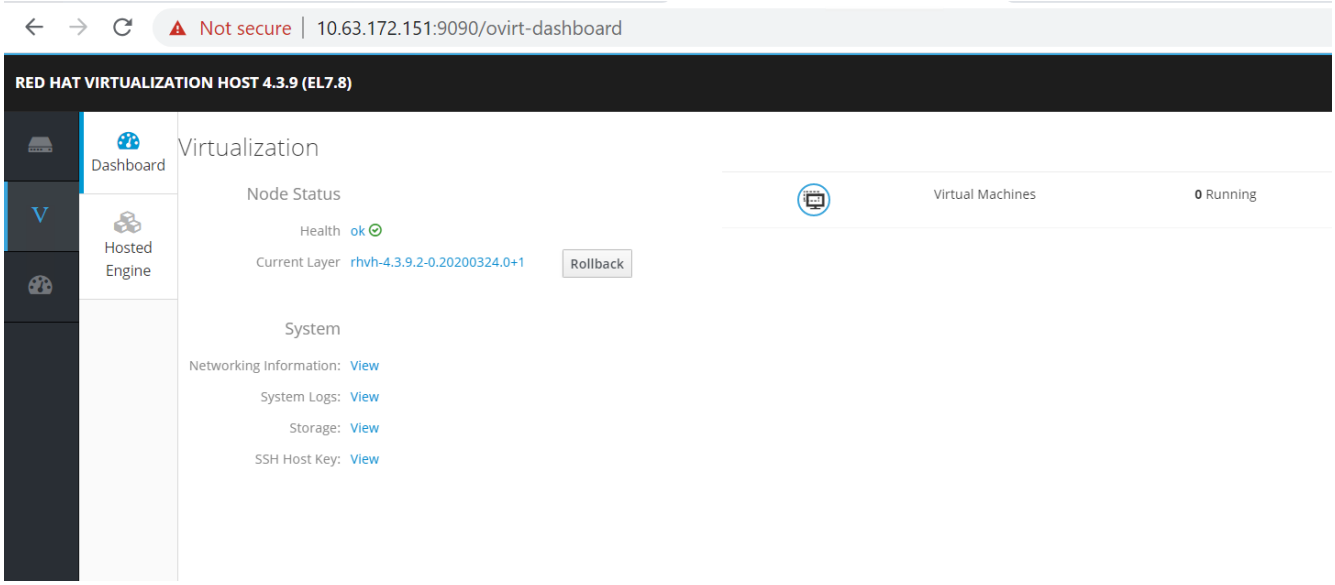
23. After the wizard navigates back to the configuration page, click Begin Installation. The next screen prompts you to configure the root password and optionally to create another user for logging into RHV-H.

24. After the installation completes, unmount the ISO file by navigating to Virtual media > Virtual Storage in the virtual console and click Plug Out. Then click Reboot on the Anaconda GUI to complete the installation process. The node then reboots.

After the node comes up, it displays the login screen.

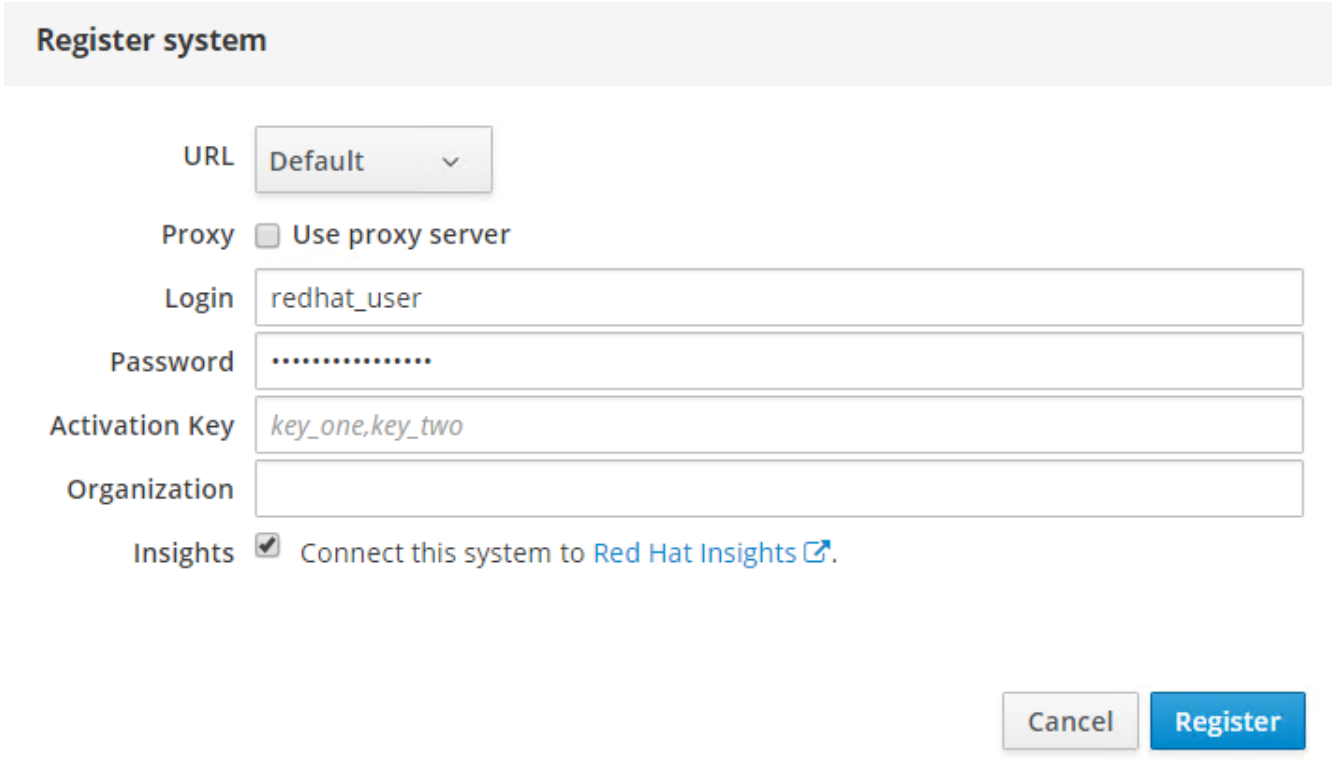


25. Now that the installation is complete, you must then register RHV-H and enable the required repositories. Open a browser and log in to the Cockpit user interface at `https://<HostFQDN/IP>:9090` using the root credentials provided during the installation.



26. Navigate to localhost > Subscriptions and click Register. Enter your Red Hat Portal username and password, click the check box Connect this System to Red Hat Insights, and click Register. The system automatically subscribes to the Red Hat Virtualization Host entitlement.

    Red Hat Insights provide continuous analysis of registered systems to proactively recognize threats to availability, security, performance, and stability across physical, virtual, and cloud environments.



27. Navigate to localhost > Terminal to display the CLI. Optionally you can use any SSH client to log in to the RHV- H CLI. Confirm that the required subscription is attached, and then enable the Red Hat Virtualization Host 7 repository to allow further updates and make sure that all other repositories

are disabled.

```
# subscription-manager list
+----------------------------------------+
    Installed Product Status
+----------------------------------------+
Product Name:   Red Hat Virtualization Host
Product ID:     328
Version:        4.3
Arch:           x86_64
Status:         Subscribed
# subscription-manager repos --disable=*
Repository 'rhel-7-server- rhvh-4-source-rpms' is disabled for this system.
Repository 'rhvh-4-build-beta-for-rhel-8-x86_64-source-rpms' is disabled for this
system.
Repository 'rhel-7-server- rhvh-4-beta-debug-rpms' is disabled for this system.
Repository 'rhvh-4-beta-for-rhel-8-x86_64-debug-rpms' is disabled for this system.
Repository 'jb-eap-textonly-1-for-middleware-rpms' is disabled for this system.
Repository 'rhvh-4-build-beta-for-rhel-8-x86_64-rpms' is disabled for this system.
Repository 'rhvh-4-beta-for-rhel-8-x86_64-source-rpms' is disabled for this system.
Repository 'rhel-7-server- rhvh-4-debug-rpms' is disabled for this system.
Repository 'rhvh-4-build-beta-for-rhel-8-x86_64-debug-rpms' is disabled for this
system.
Repository 'rhel-7-server- rhvh-4-beta-source-rpms' is disabled for this system.
Repository 'rhel-7-server- rhvh-4-rpms' is disabled for this system.
Repository 'jb-coreservices-textonly-1-for-middleware-rpms' is disabled for this
system.
Repository 'rhvh-4-beta-for-rhel-8-x86_64-rpms' is disabled for this system.
Repository 'rhel-7-server- rhvh-4-beta-rpms' is disabled for this system.
# subscription-manager repos --enable=rhel-7-server- rhvh-4-rpms
Repository 'rhel-7-server- rhvh-4-rpms' is enabled for this system.
```

28. From the console, modify the iSCSI initiator ID to match the one you set in the Element access group previously by running the following command.

```
rhv-h01 # echo InitiatorName=iqn.1994-05.com.redhat:rhv-host-node- 01 >
/etc/iscsi/initiatorname.iscsi
```

29. Enable and restart the iscsid service.

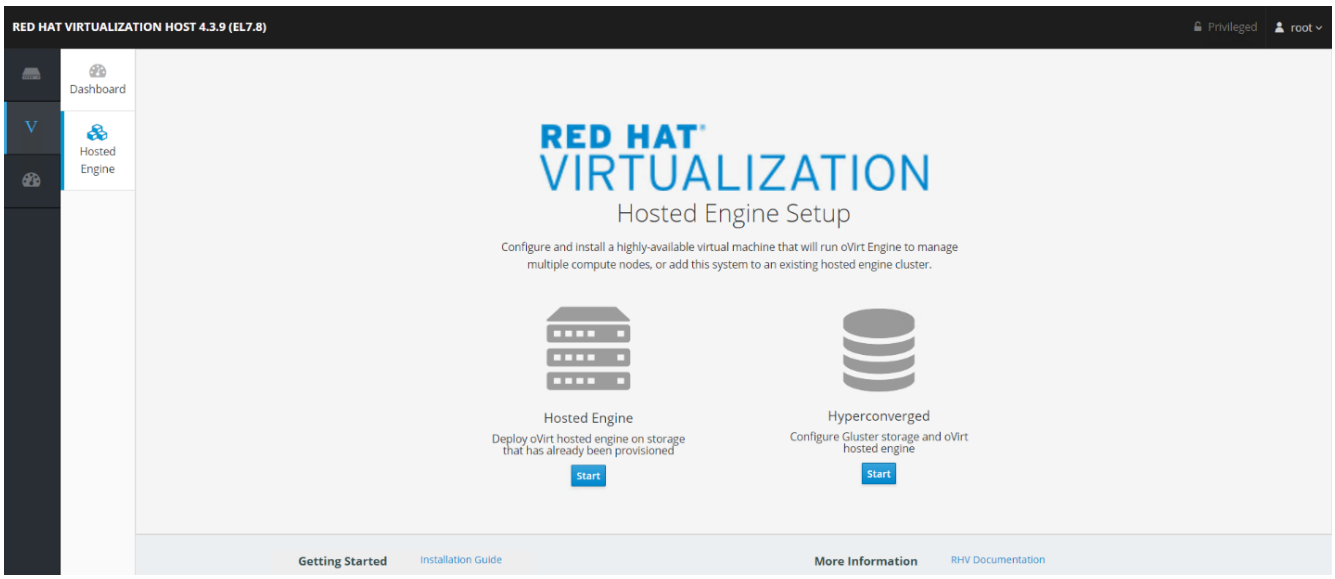
```
 # systemctl enable iscsid
Created symlink from /etc/systemd/system/multi-user.target.wants/iscsid.service to
/usr/lib/systemd/system/iscsid.service
 # systemctl start iscsid
 # systemctl status iscsid
● iscsid.service - Open-iSCSI
   Loaded: loaded (/usr/lib/systemd/system/iscsid.service; enabled; vendor preset:
disabled)
   Active: active (running) since Thu 2020-05-14 16:08:52 EDT; 3 days ago
     Docs: man:iscsid(8)
           man:iscsiuio(8)
           man:iscsiadm(8)
 Main PID: 5422 (iscsid)
   Status: "Syncing existing session(s)"
   CGroup: /system.slice/iscsid.service
           ├─5422 /sbin/iscsid -f
           └─5423 /sbin/iscsid -f
```

30. Install and prepare the other RHV host by repeating the steps 1 to 29.

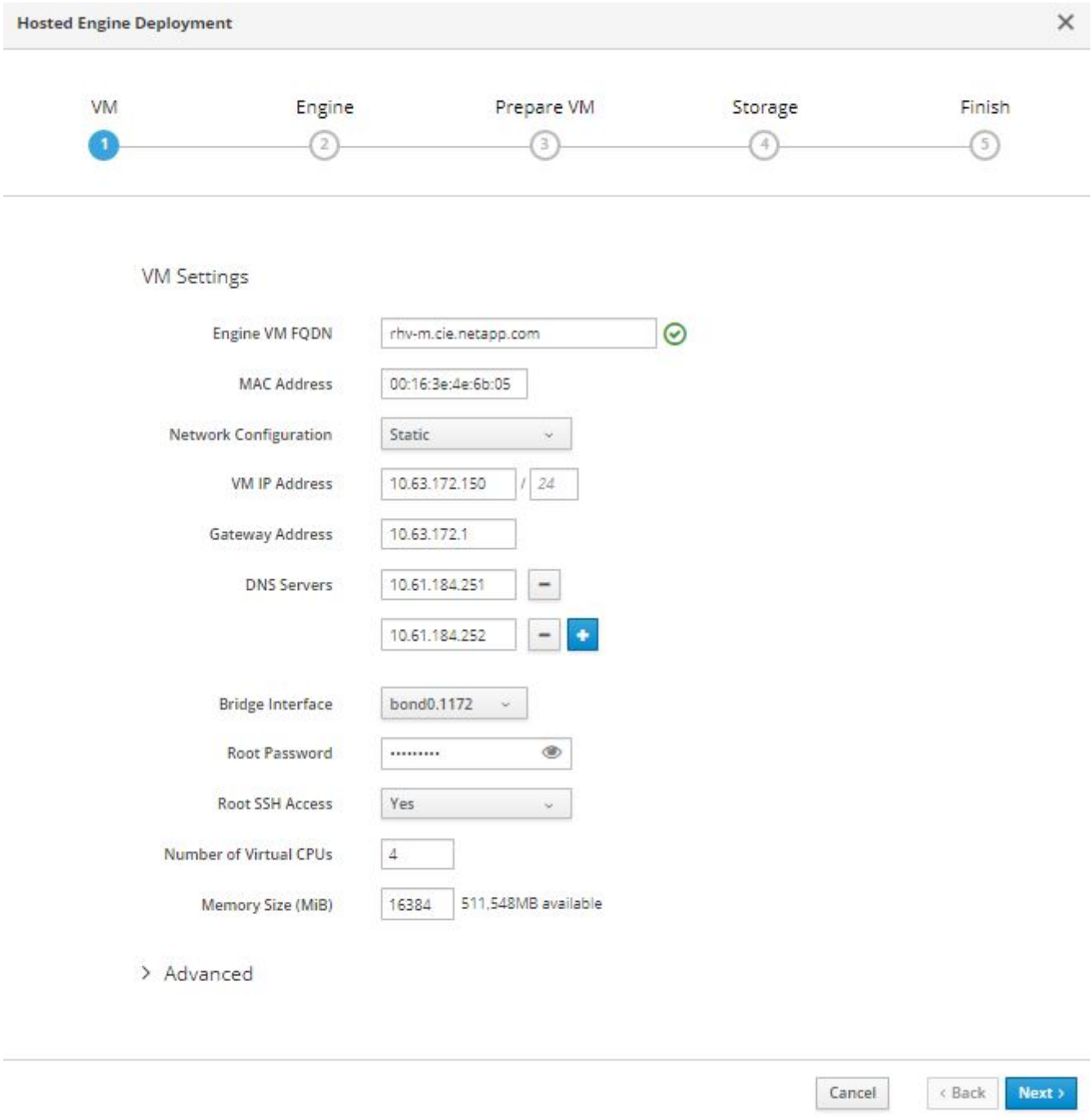
5. Deploy the RHV Manager as a Self-Hosted Engine: NetApp HCI with RHV

This section describes the detailed steps for installing the Red Hat Virtualization Manager as a self-hosted engine. These steps begin after the RHV hosts are registered and the Cockpit GUI is accessible.

1. Log in to the Cockpit GUI of one of the RHV hosts at `https://<HostFQDN/IP>:9090` using the root credentials. Navigate to the Virtualization sub-tab and click Hosted Engine. Then click the Start button below the Hosted Engine content to initiate the engine deployment.

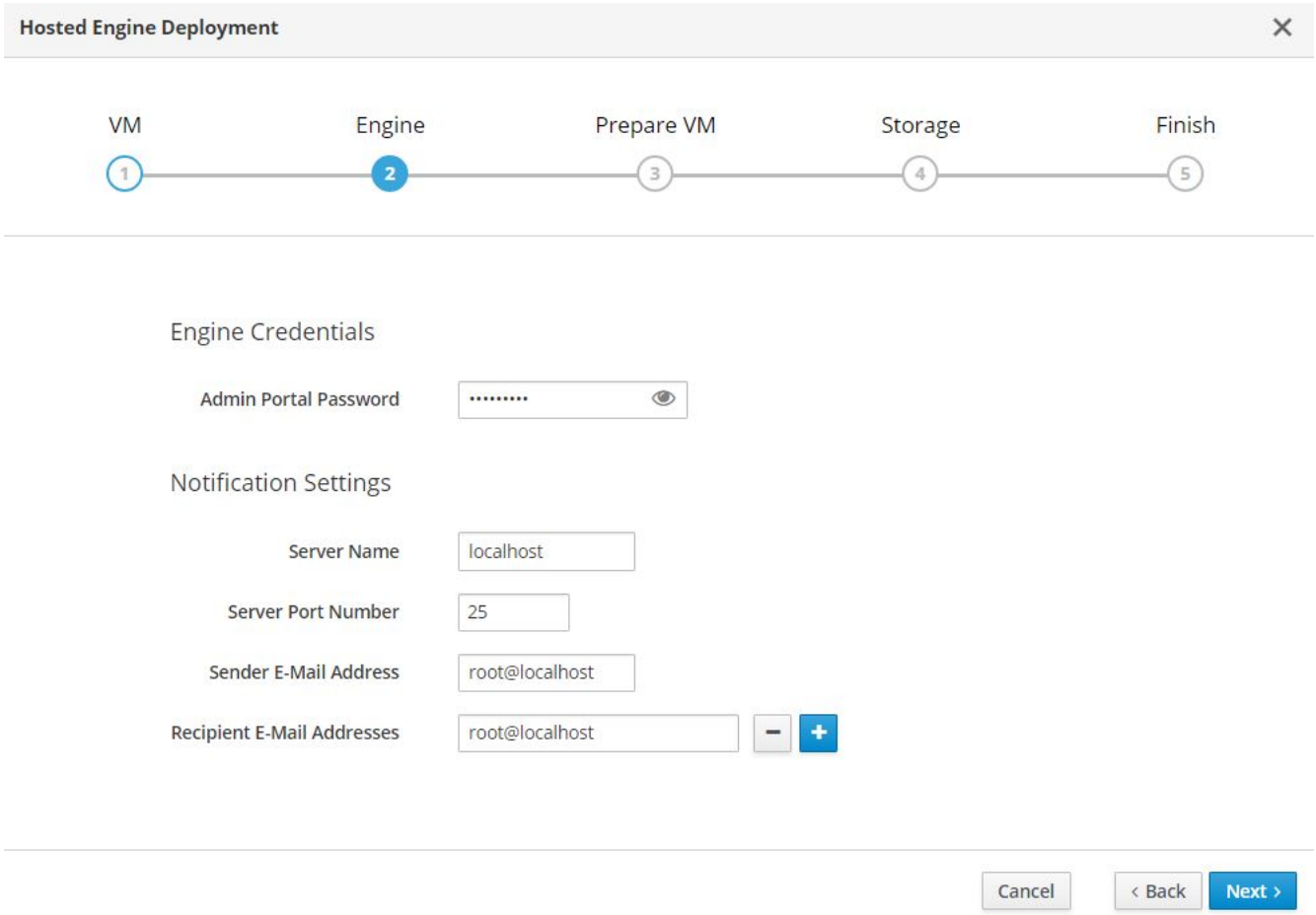


2. In the first screen of engine deployment, configure the RHV-M FQDN, network related configuration, root password, and resources for the engine VM (at least 4 CPUs and 16GB memory). Confirm the other configuration settings as required and click Next.

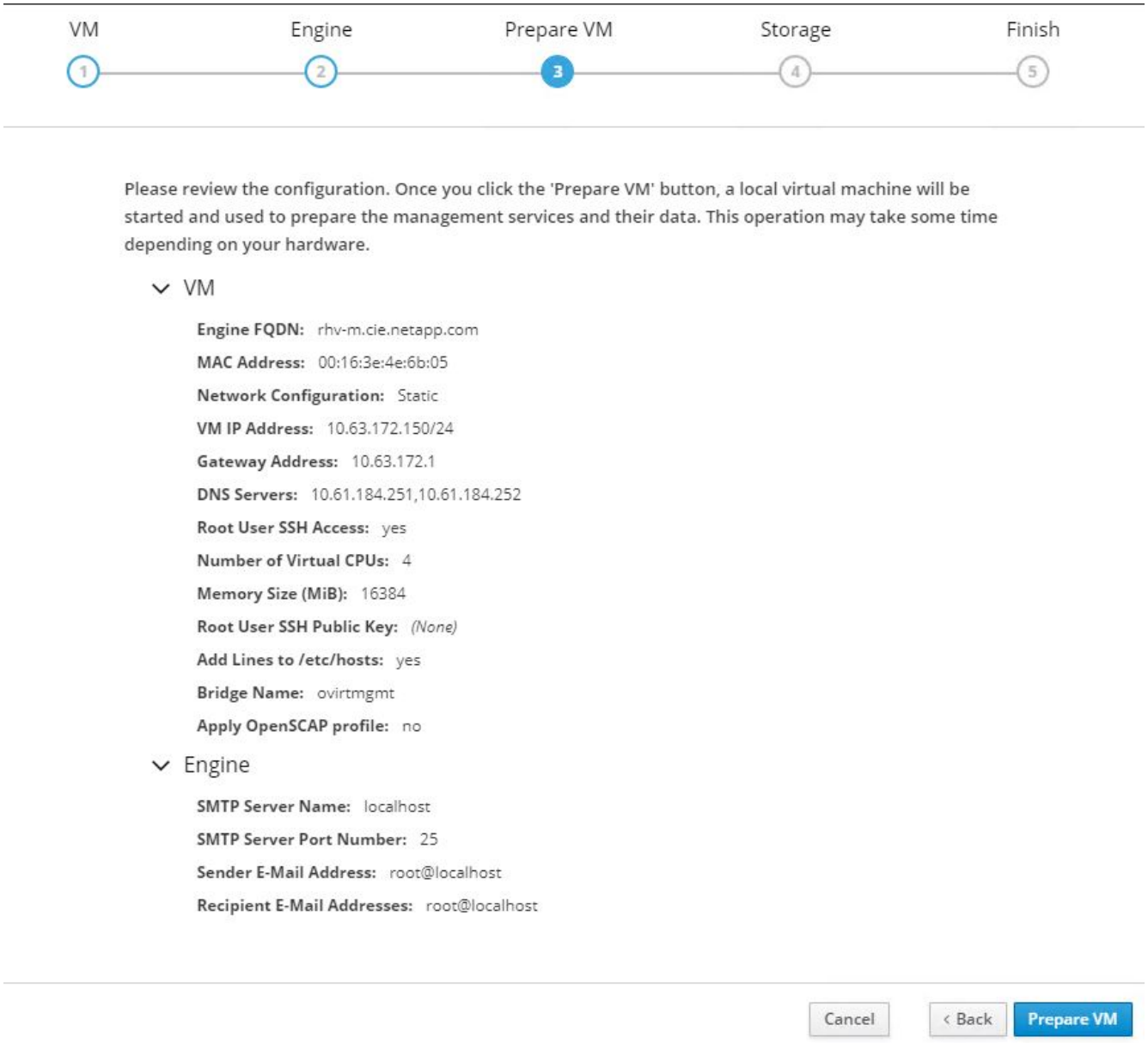


Make sure that the engine VM FQDN is resolvable by the specified DNS servers.

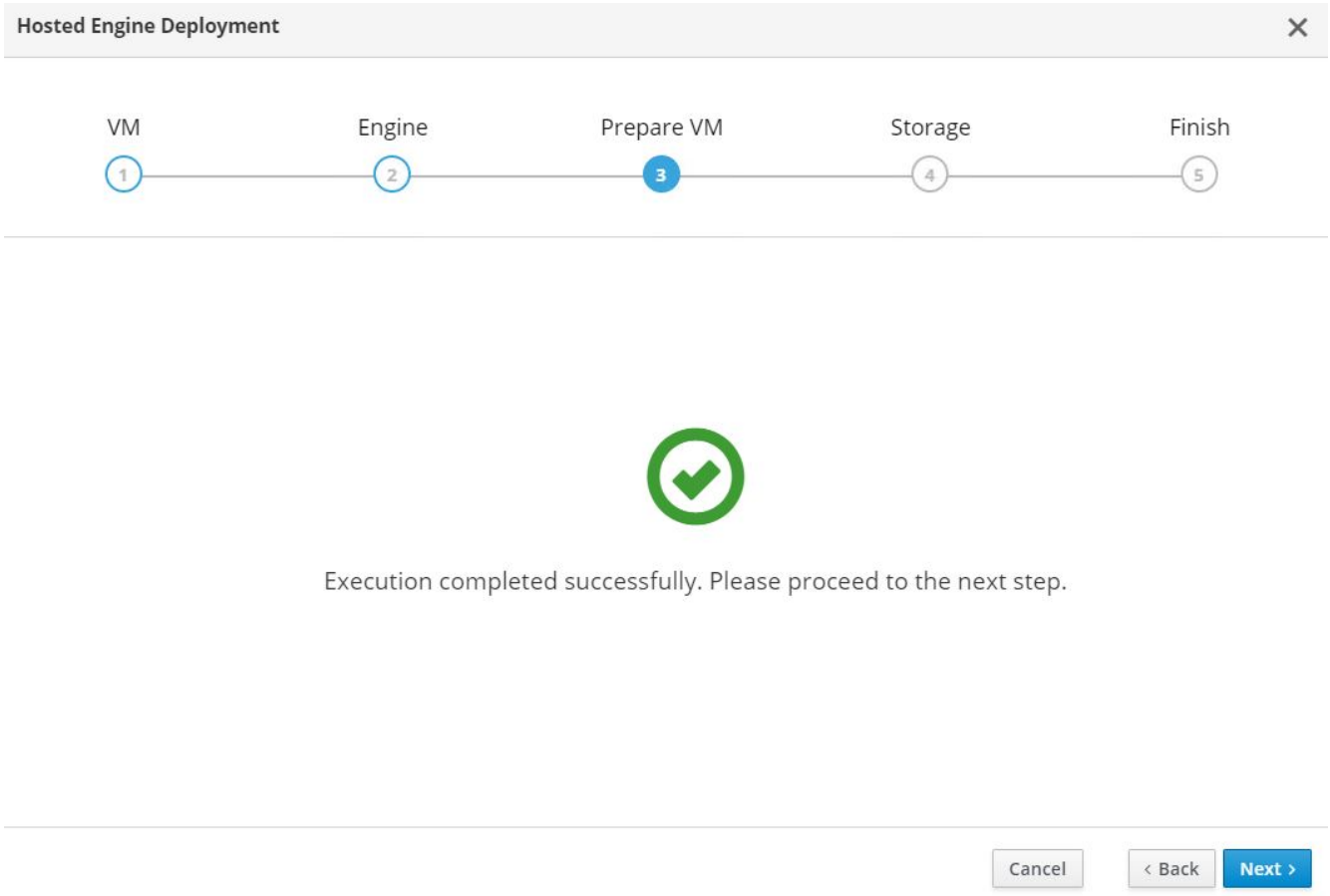
3. In the next screen, enter the admin portal password. Optionally, enter the notification settings for alerts to be sent by email. Then click Next.



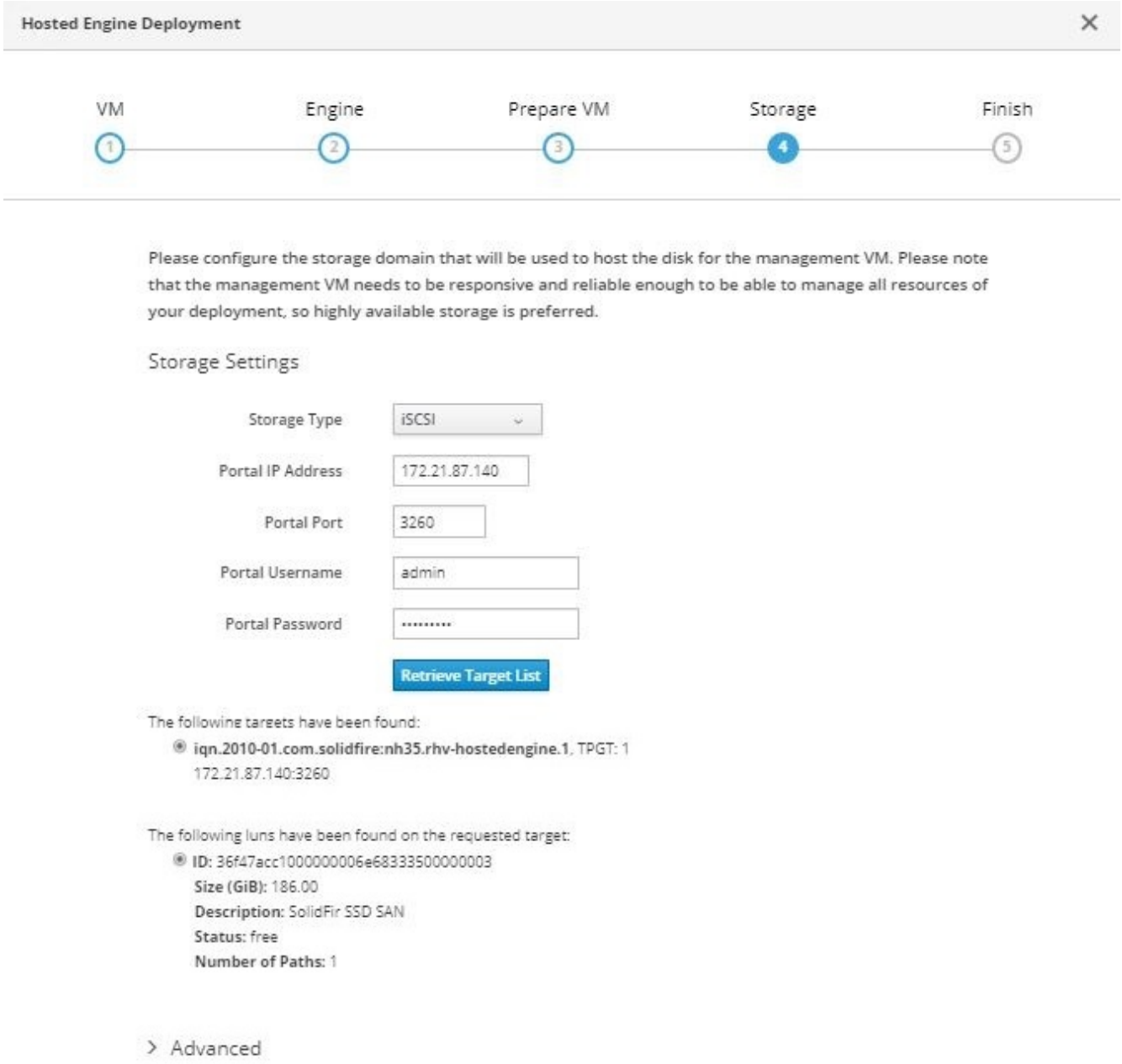
4. In the next screen, review the configuration for the engine VM. If any changes are desired, go back at this point and make them. If the information is correct, click Prepare the VM.



5. The VM installation begins and can take some time to complete as it downloads a machine image and stages the VM locally. After it has completed, it displays the Execution Completed Successfully message. Click Next.

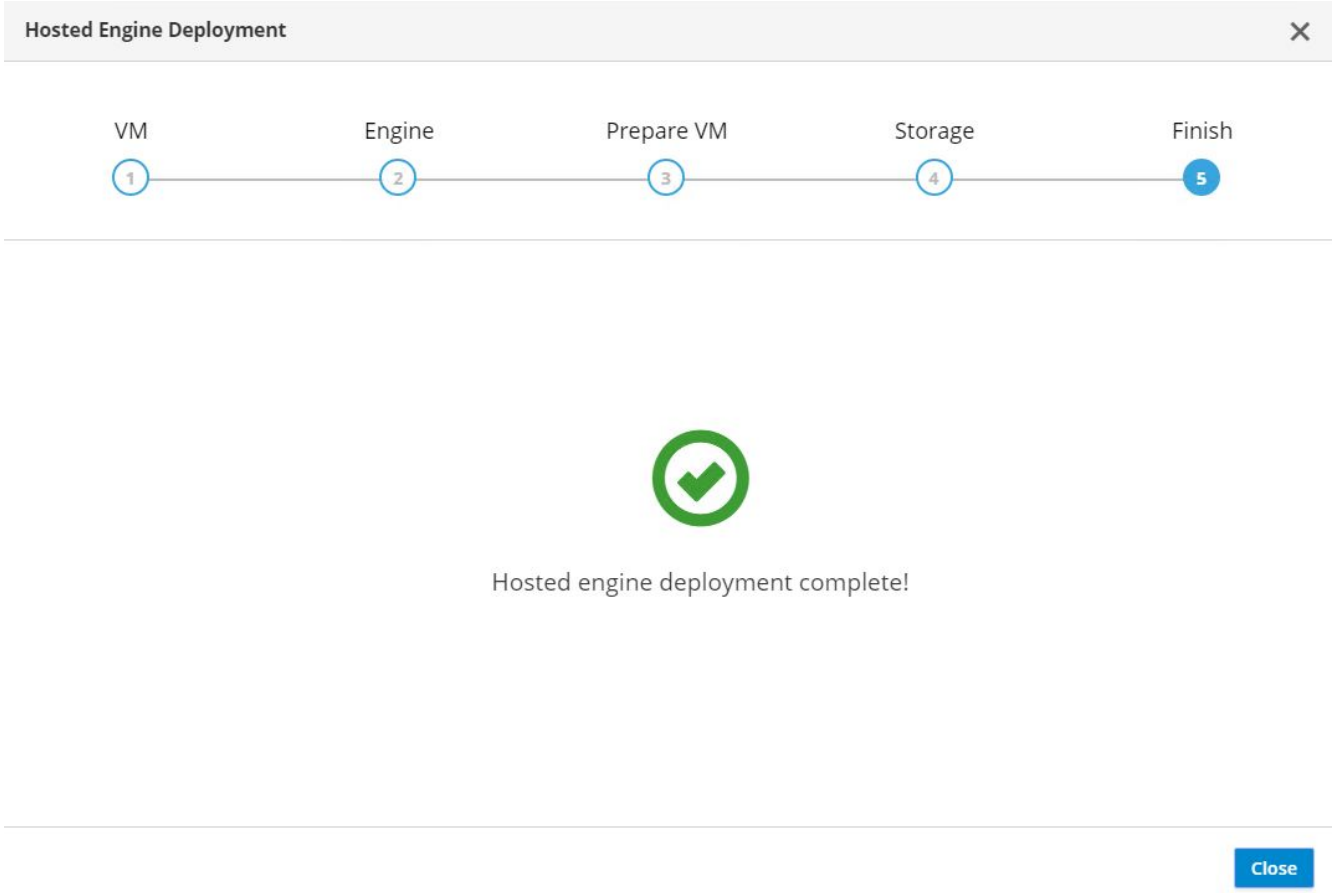


6. After RHV-M is installed, enter the details of the hosted engine storage domain where it copies the VM from local storage to the shared storage domain to facilitate a high availability engine quorum.
7. Enter the Storage Type as iSCSI, provide the iSCSI portal details, click Retrieve Target List, which fetches the iSCSI target list corresponding to the portal, and select the volume and LUN to be mapped to the hosted engine storage domain. Click Next.



If the Hosted Engine setup is unable to discover the storage, open an interactive SSH session to the node and verify that you can reach the SVIP IP address through your node's storage interface. If the network is reachable, you might need to manually discover or log in to the iSCSI LUN intended for the Hosted Engine install.

8. On the next screen, review the storage configuration and, if any changes are desired, go back and make them. If the information is correct, click Finish Deployment. It takes some time as the VM is copied to the storage domain. After deployment is complete, click Close.



9. The next step is to register and enable the Red Hat Virtualization Manager repositories. Log in to the RHV-M VM with SSH to register it with Subscription Manager.

```
# subscription-manager register
Registering to: subscription.rhsm.redhat.com:443/subscription
Username: redhat_user
Password:  redhat_password
The system has been registered with ID: 99d06fcb-a3fd74-41230f-bad583-0ae61264f9a3
The registered system name is: rhv-m.cie.netapp.com
```

10. After registration, list the available subscriptions and record the pool ID for RHV-M.

```
# subscription-manager list --available
<snip>
Subscription Name:   Red Hat Virtualization Manager
Provides:            Red Hat Beta
                     Red Hat Enterprise Linux Server
                     Red Hat CodeReady Linux Builder for x86_64
                     Red Hat Enterprise Linux for x86_64
                     Red Hat Virtualization Manager
                     Red Hat OpenShift Container Platform
                     Red Hat Ansible Engine
                     Red Hat Enterprise Linux Fast Datapath
                     Red Hat JBoss Core Services
                     JBoss Enterprise Application Platform
SKU:                 RV00045
Contract:
Pool ID:             8a85f9937a1a2a57c0171a366b5682540112a313 ß Pool ID
Provides Management: No
Available:           6
Suggested:           0
Service Type:        L1-L3
Roles:
Service Level:       Layered
Usage:
Add-ons:
Subscription Type:   Stackable
Starts:              04/22/2020
Ends:                04/21/2021
Entitlement Type:    Physical
<snip>
```

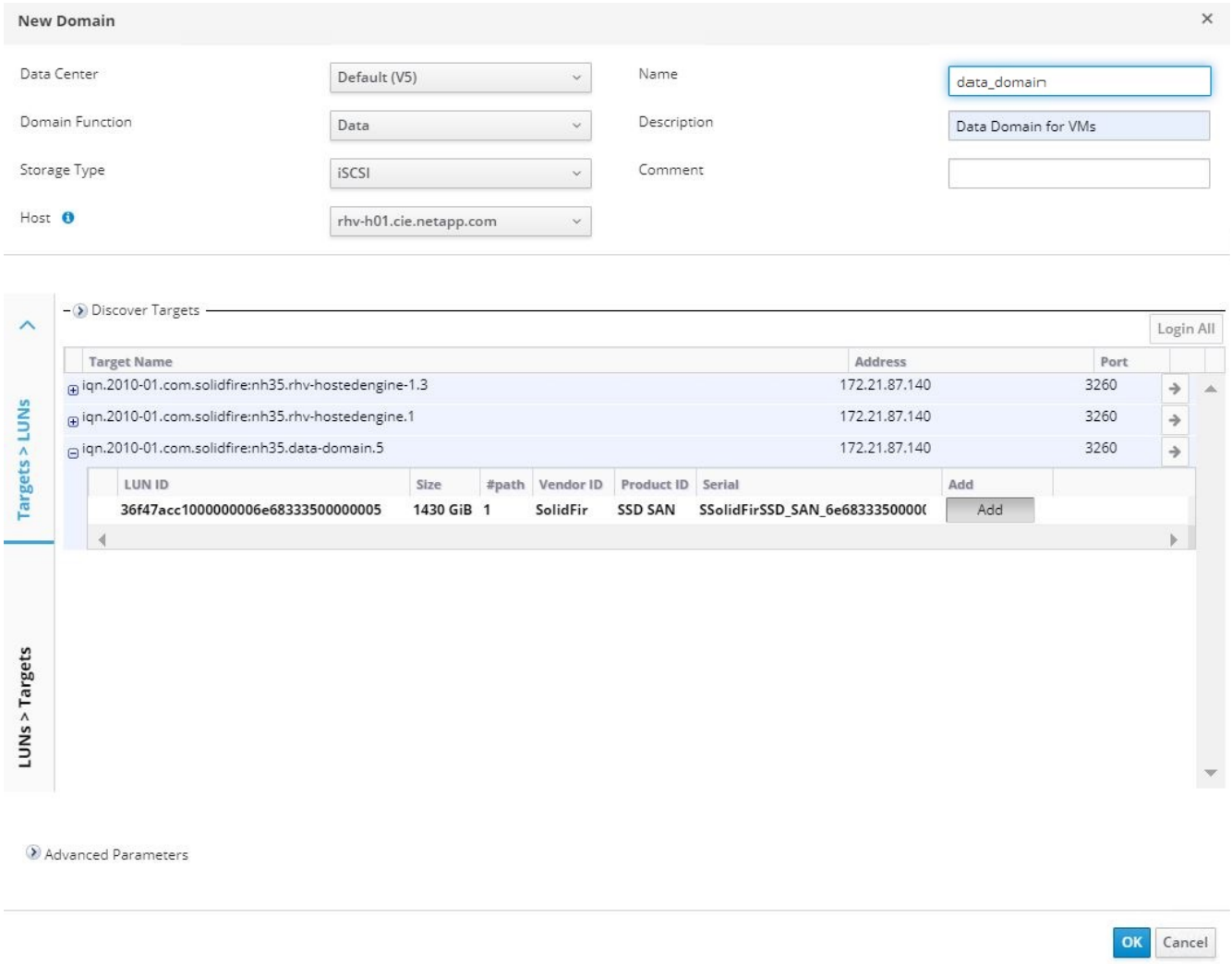
11. Attach the RHV-M subscription using the recorded pool ID.

```
# subscription-manager attach --pool=8a85f9937a1a2a57c0171a366b5682540112a313
Successfully attached a subscription for: Red Hat Virtualization Manager
```

12. Enable the required RHV-M repositories.

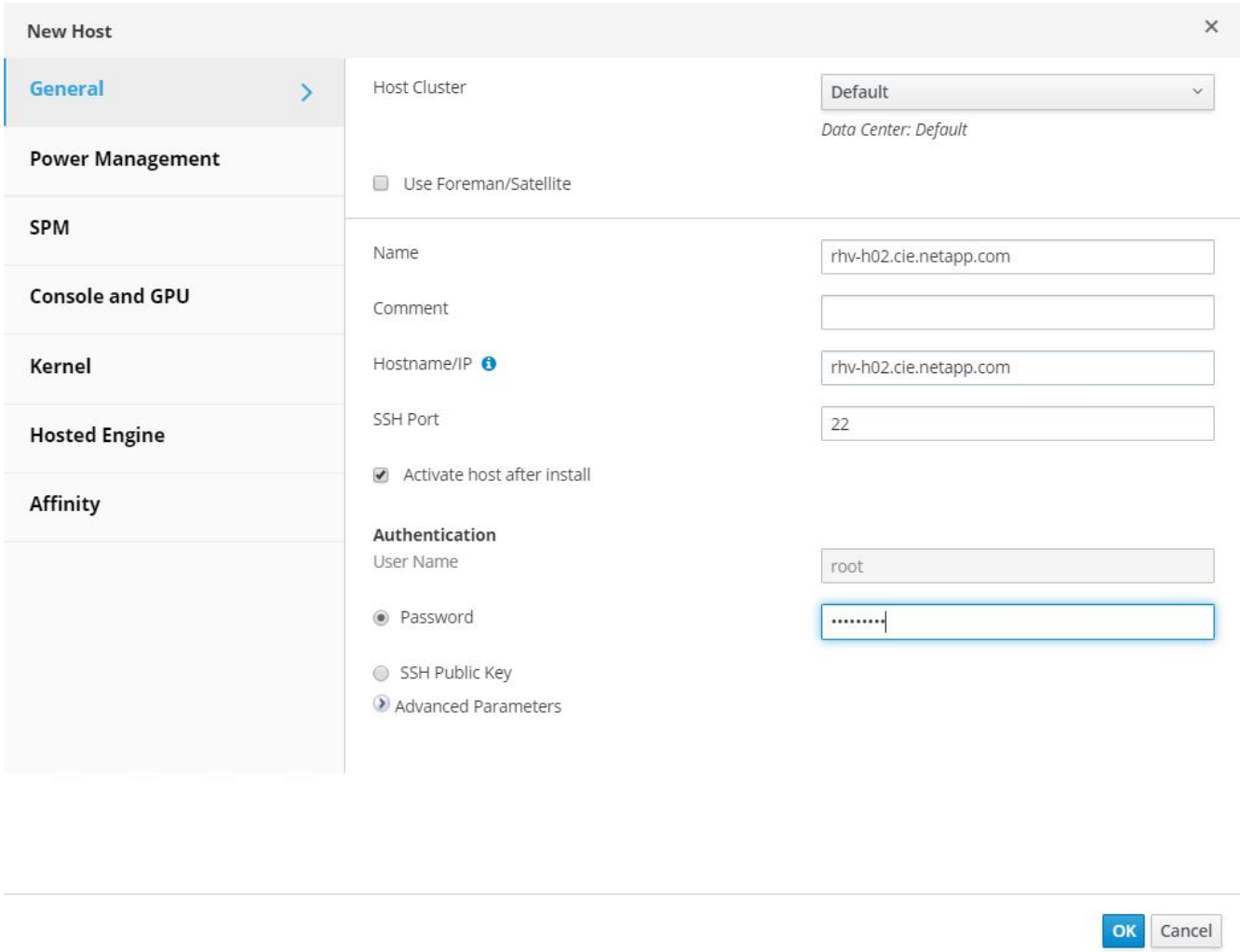
```
# subscription-manager repos \
    --disable='*' \
    --enable=rhel-7-server-rpms \
    --enable=rhel-7-server-supplementary-rpms \
    --enable=rhel-7-server-rhv-4.3-manager-rpms \
    --enable=rhel-7-server-rhv-4-manager-tools-rpms \
    --enable=rhel-7-server-ansible-2-rpms \
    --enable=jb-eap-7.2-for-rhel-7-server-rpms
Repository 'rhel-7-server-ansible-2-rpms' is enabled for this system.
Repository 'rhel-7-server-rhv-4-manager-tools-rpms' is enabled for this system.
Repository 'rhel-7-server-rhv-4.3-manager-rpms' is enabled for this system.
Repository 'rhel-7-server-rpms' is enabled for this system.
Repository 'jb-eap-7.2-for-rhel-7-server-rpms' is enabled for this system.
Repository 'rhel-7-server-supplementary-rpms' is enabled for this system.
```

13. Next, create a storage domain to hold the VM disks or OVF files for all VMs in the same datacenter as that of the hosts.
14. To log into the RHV-M Administrative portal using a browser, log into `https://<ManagerFQDN>/ovirt-engine`, select Administrative Portal, and log in as the `admin@internal` user.
15. Navigate to Storage > Storage Domains and click New Domain.
16. From the dropdown menu, select Data for the Domain Function, select iSCSI for the Storage Type, select the host to map the volume, enter a name of your choice, confirm that the data center is correct, and then expand the data domain iSCSI target and add the LUN. Click OK to create the domain.

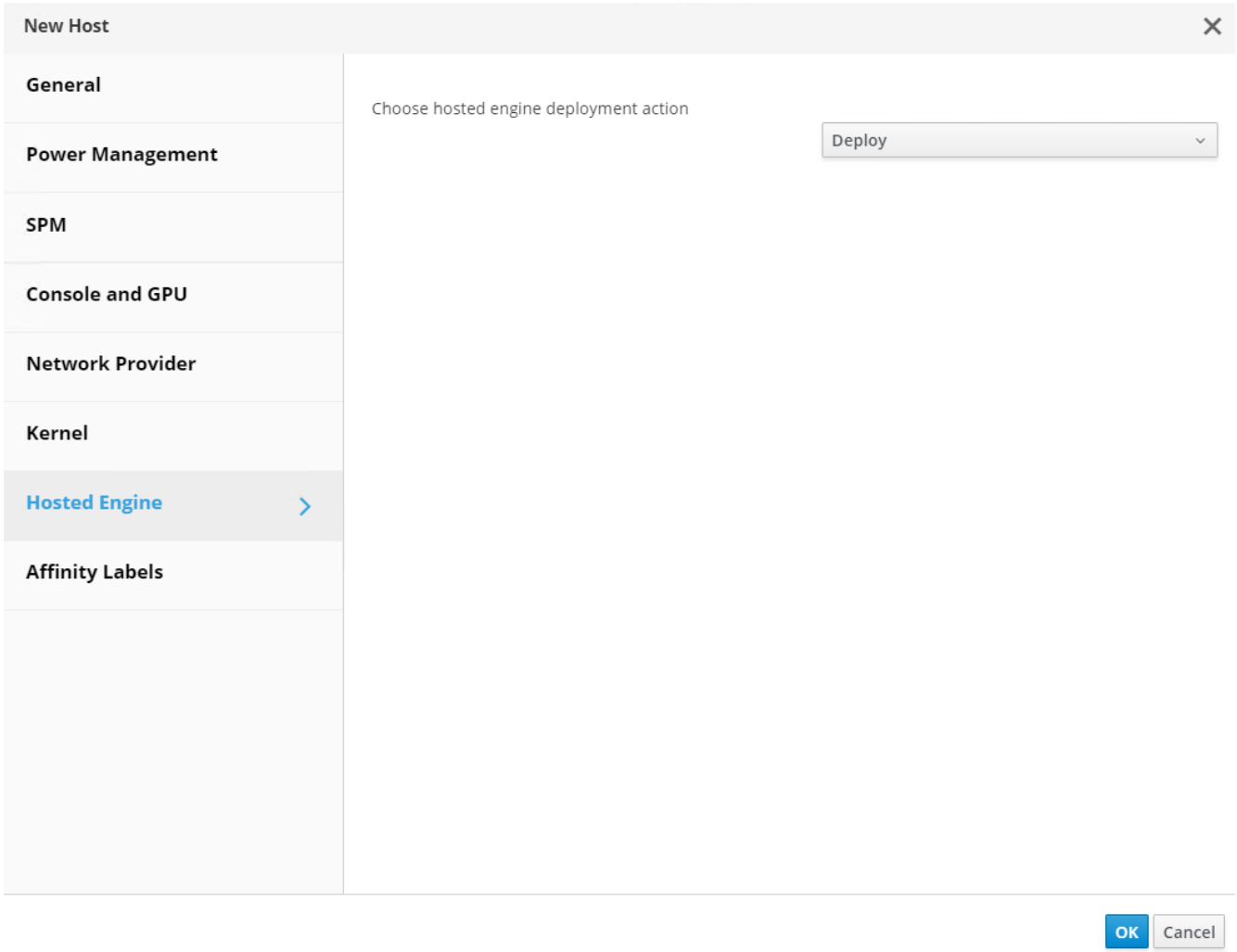


If the Hosted Engine setup is unable to discover the storage, you might need to manually discover or log in to the iSCSI LUN intended for the data domain.

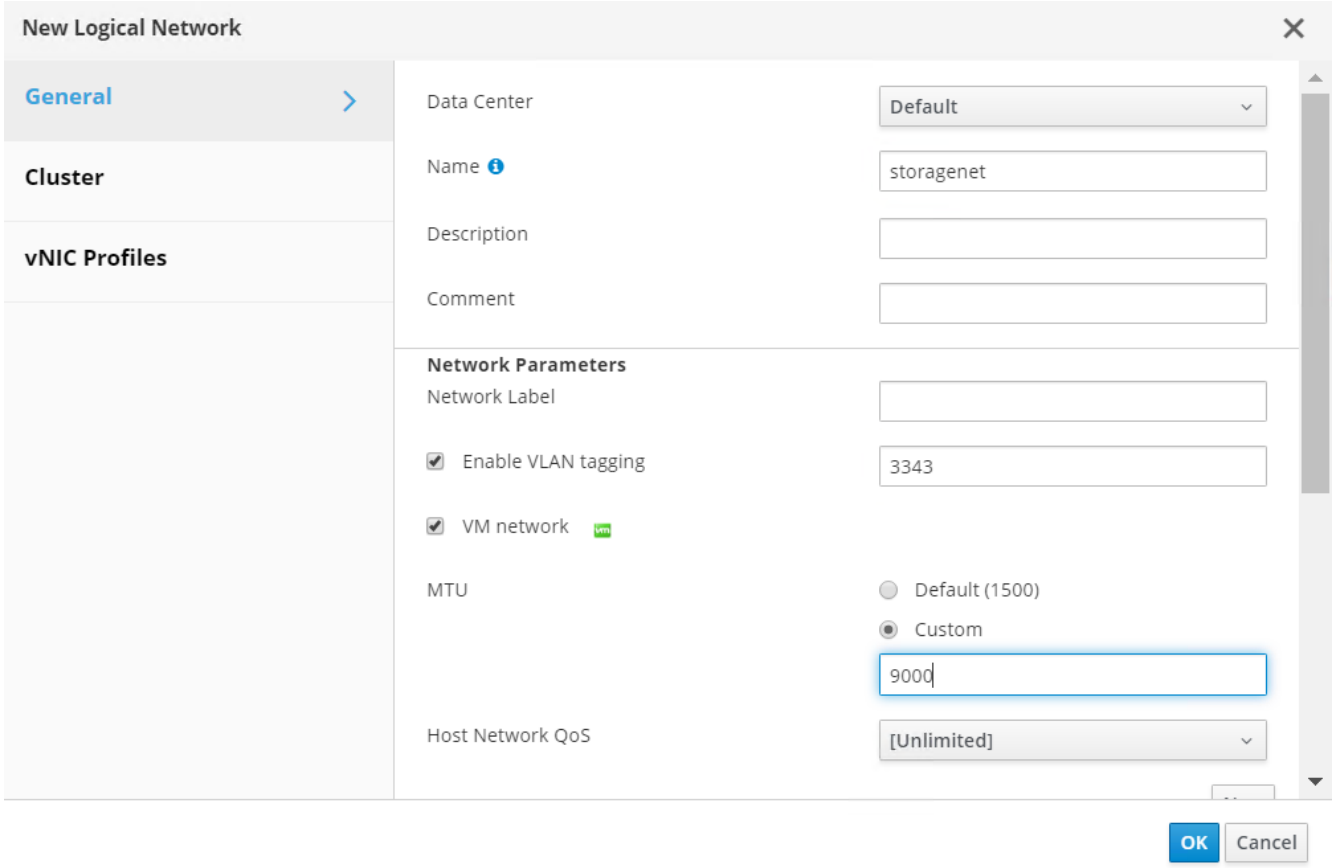
17. Add the second host to the hosted engine quorum. Navigate to Compute > Hosts and click New. In the New Host pane, select the appropriate cluster, provide the details of the second host, and check the Activate Host After Install checkbox.



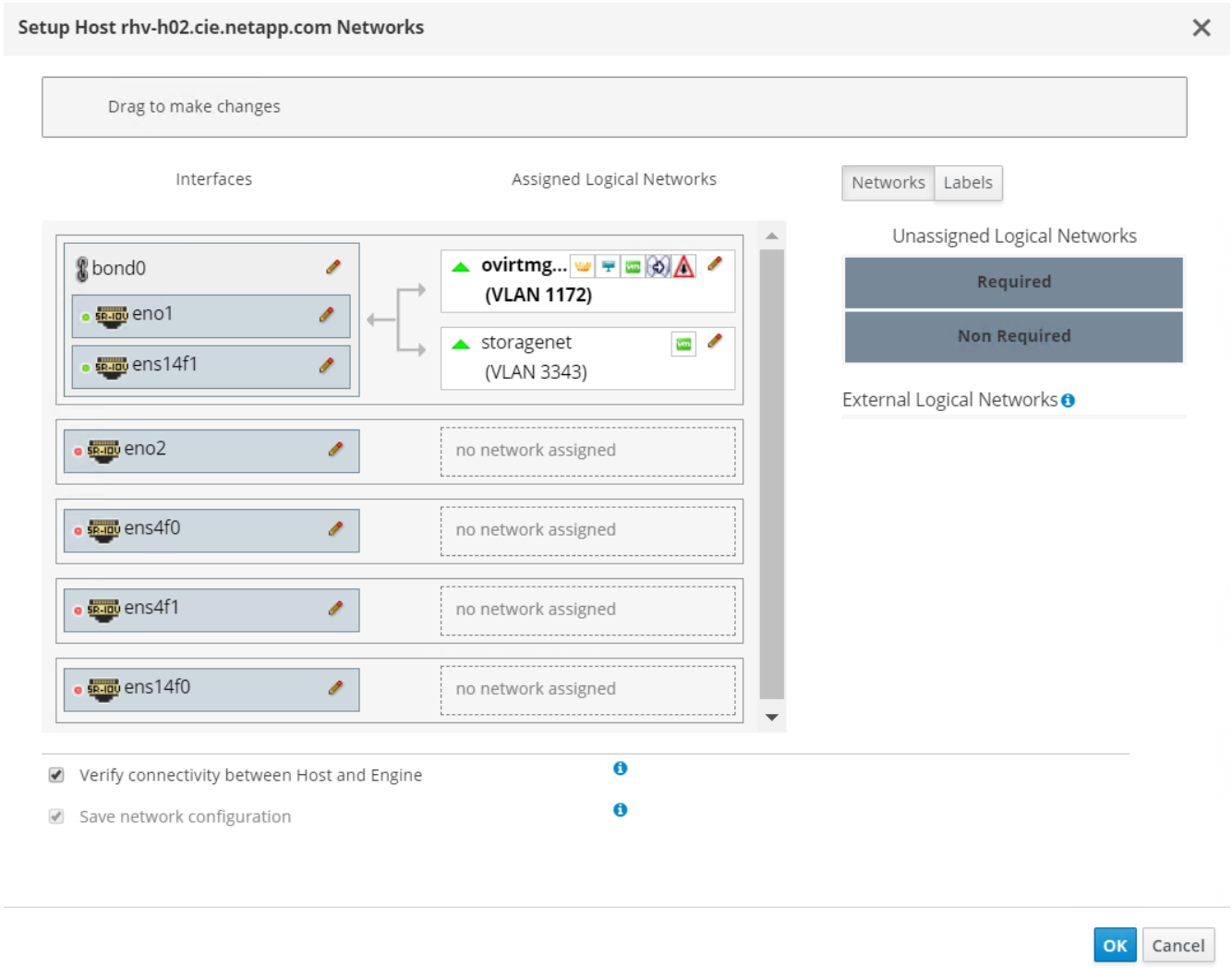
18. Click the Hosted Engine sub-tab in the New Host pane dropdown and select Deploy from the hosted engine deployment action. Click OK to add the host to the quorum. This begins the installation of the necessary packages to support the hosted engine and activate the host. This process might take a while.



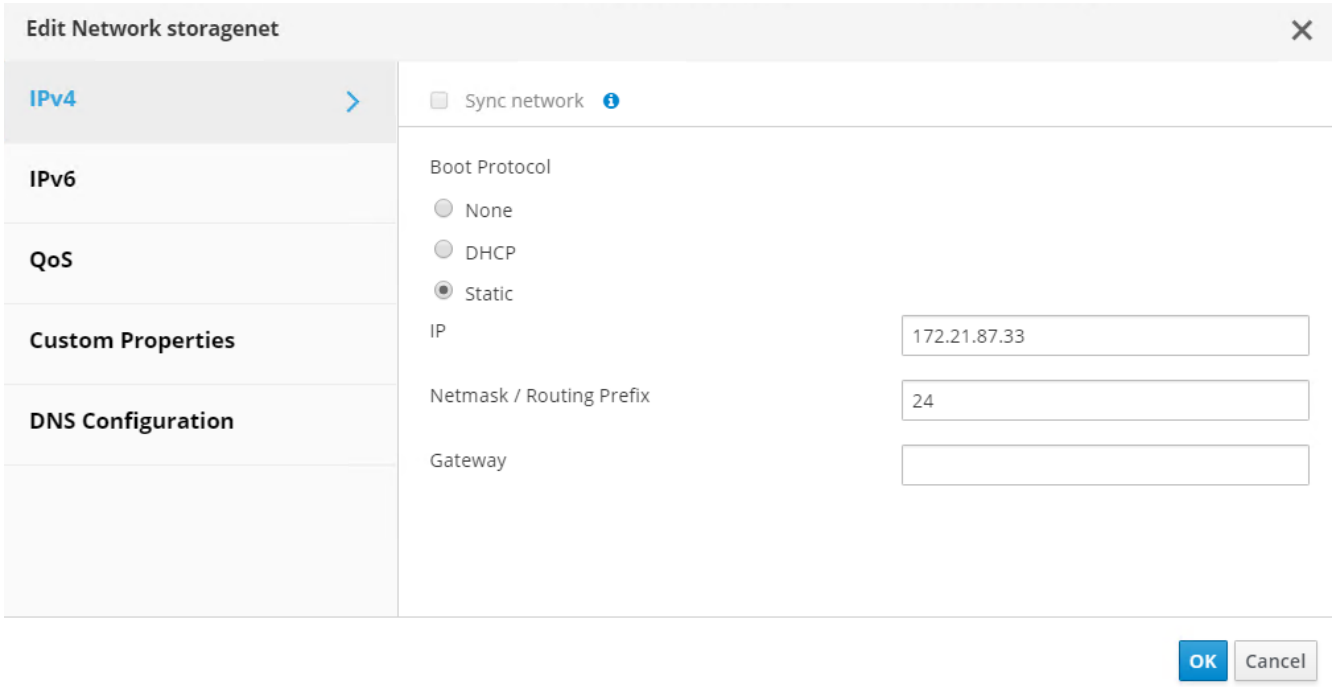
19. Next, create a storage virtual network for hosts. Navigate to Network > Networks and click New. Enter the name of your choice, enable VLAN tagging, and enter the VLAN ID for the Storage network. Confirm that the VM Network checkbox is checked and that the MTU is set to 9000. Go to the Cluster sub-tab and make sure that Attach and Require are checked. Then click OK to create the storage network.



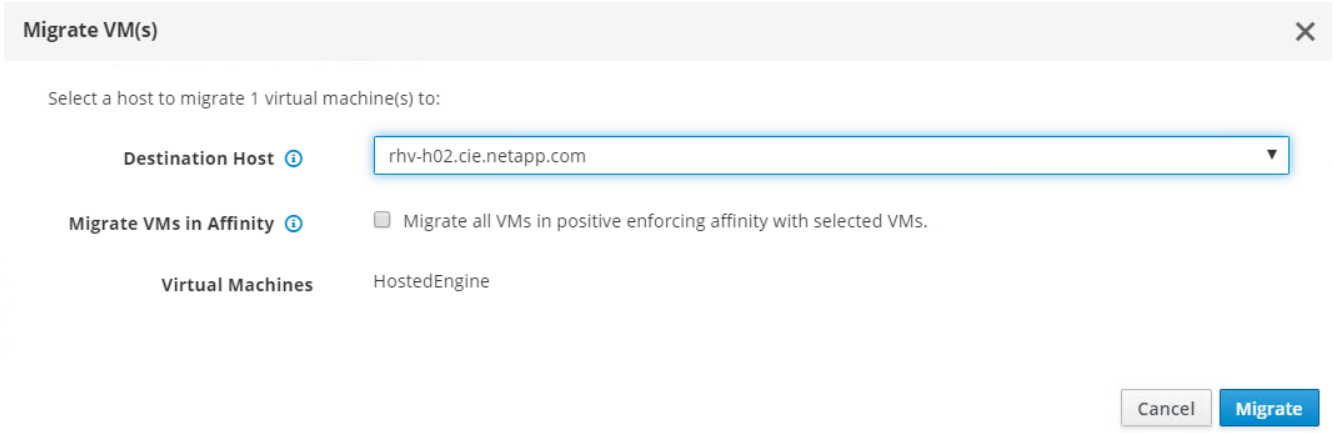
20. Assign the storage logical network to the second host in the cluster or to whichever host is not currently hosting the hosted engine VM.
21. Navigate to Compute > Hosts, and click the host that has silver crown in the second column. Then navigate to the Network Interfaces sub-tab, click Setup Host Networks, and drag and drop the storage logical network into the Assigned Logical Networks column to the right of bond0.



22. Click the pen symbol on the storage network interface under bond0. Configure the IP address and the netmask, and then click OK. Click OK again in the Setup Host Networks pane.

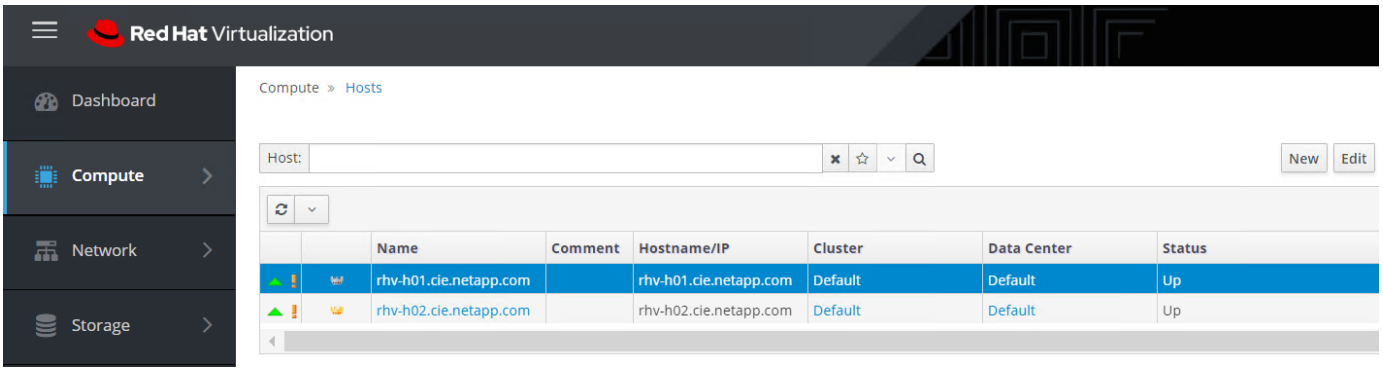


23. Migrate the hosted engine VM to the host that was just configured so that the storage logical network can be configured on the second host. Navigate to Compute > Virtual Machines, click HostedEngine and then click Migrate. Select the second host from the dropdown menu Destination Host and click Migrate.



After the migration is successful and the hosted engine VM is migrated to the second host, repeat steps 21 and 22 for the host that currently possesses the silver crown.

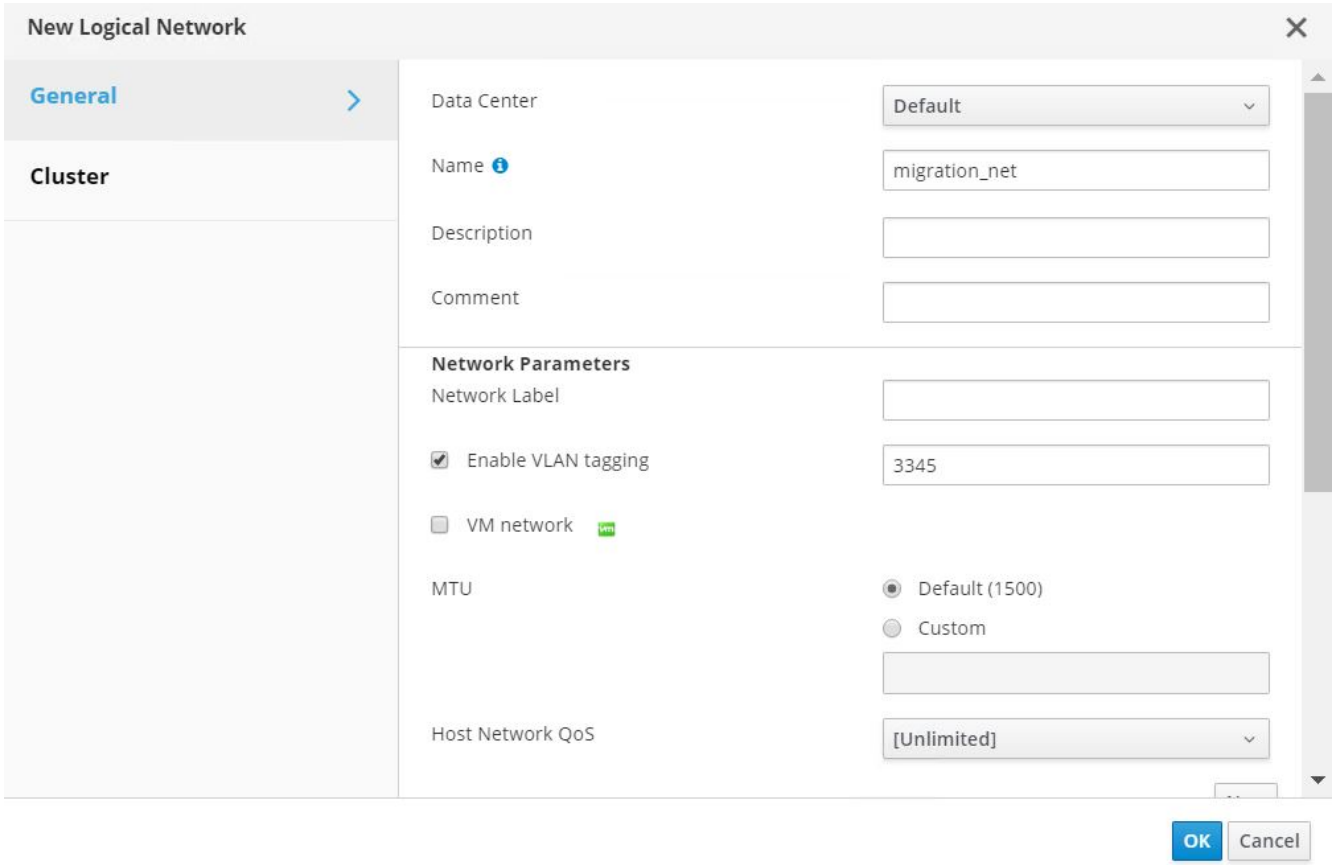
24. After you have completed this process, you should see that both the hosts are up. One of the hosts has a golden crown, indicating that it is hosting the hosted engine VM, and the other host has a silver crown indicating that it is capable of hosting the hosted engine VM.



6. Configure RHV-M Infrastructure: NetApp HCI with RHV

To configure the RHV-M infrastructure, complete the following steps:

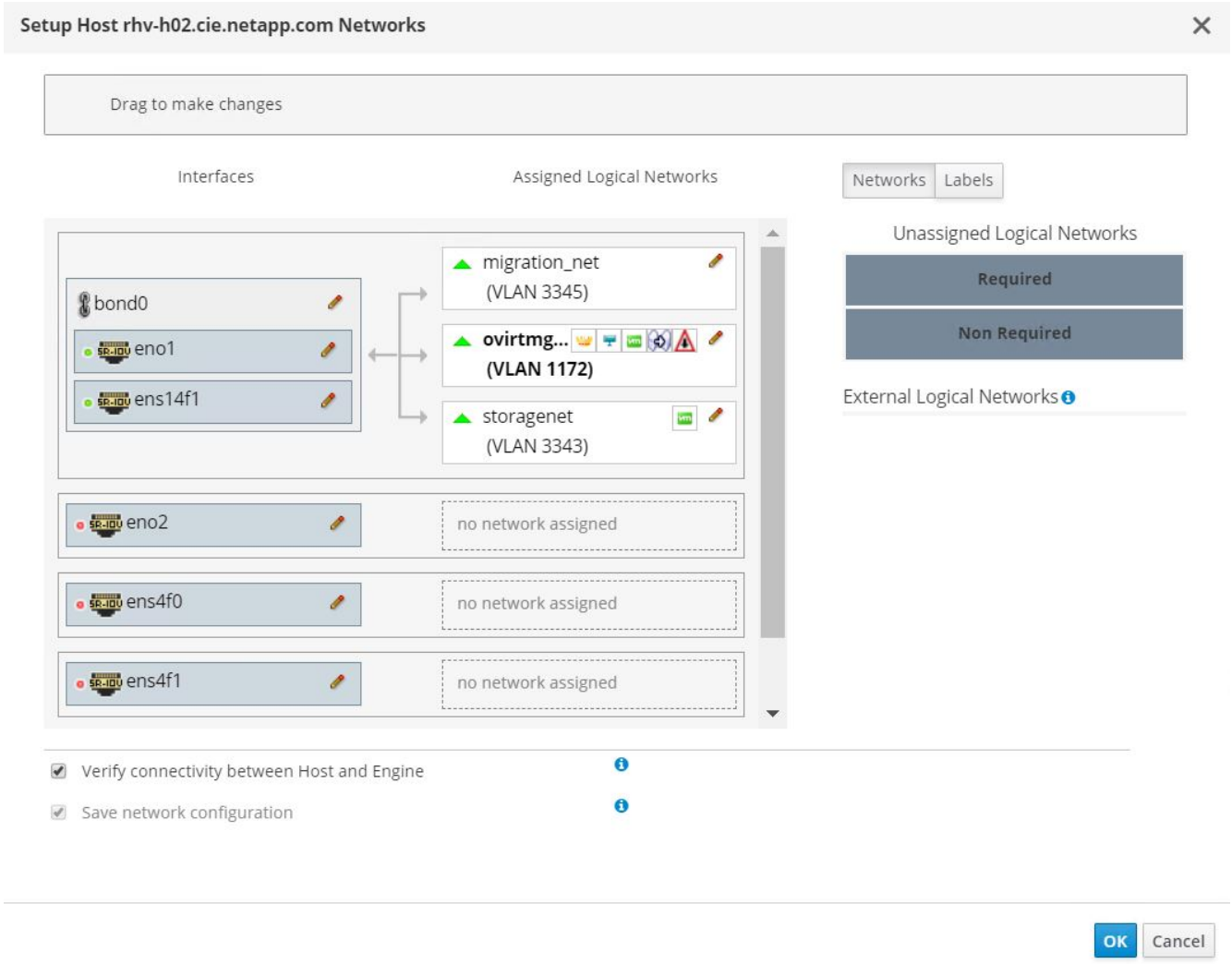
1. By default, the ovirtmgmt network is used for all purposes, including the migration of VMs and virtual guest data.
2. It is a best practice to specify different networks for these purposes. To configure the migration network, navigate to Network > Networks and click New. Enter the name of your choice, enable VLAN tagging, and enter the VLAN ID for the migration network.
3. Make sure that the VM Network checkbox is unchecked. Go to the Cluster sub-tab and make sure that Attach and Require are checked. Then click OK to create the network.



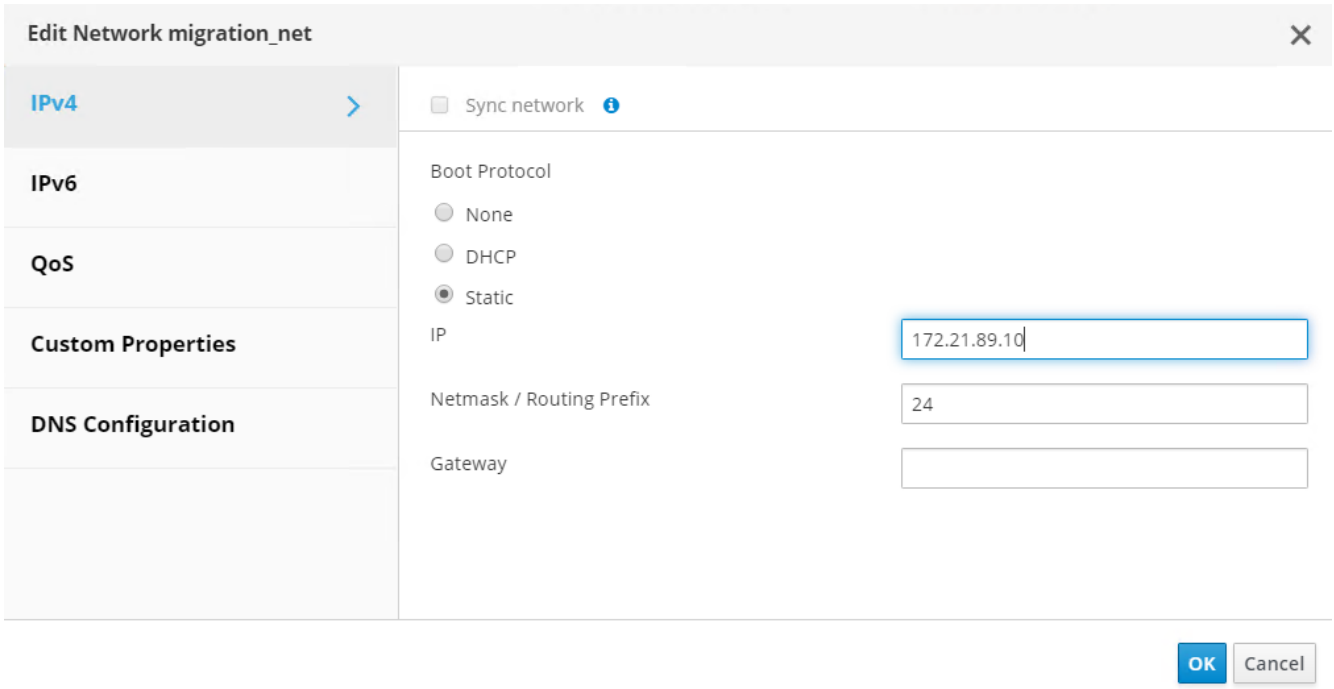
4. To assign the migration logical network to both the hosts, navigate to Compute > Hosts, click the

hosts, and navigate to the Network Interfaces sub-tab.

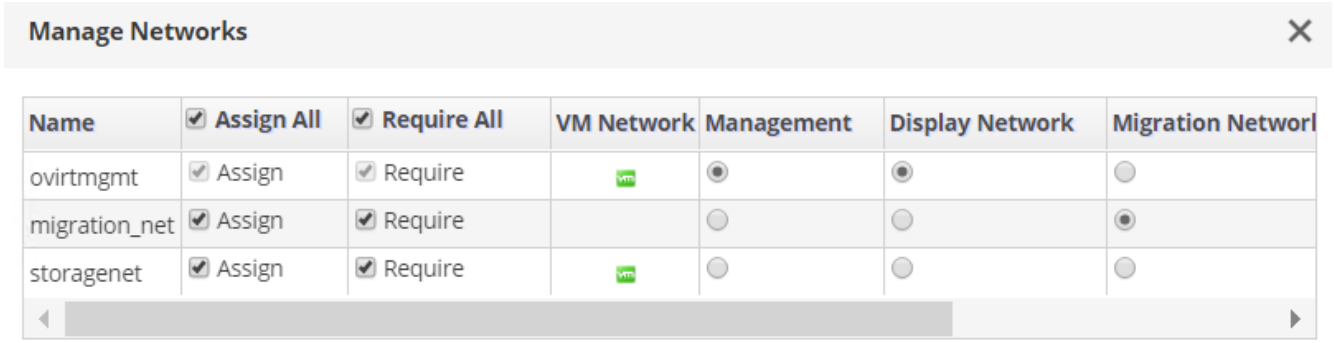
5. Then click Setup Host Networks and drag and drop the migration logical network into the Assigned Logical Networks column to the right of bond0.



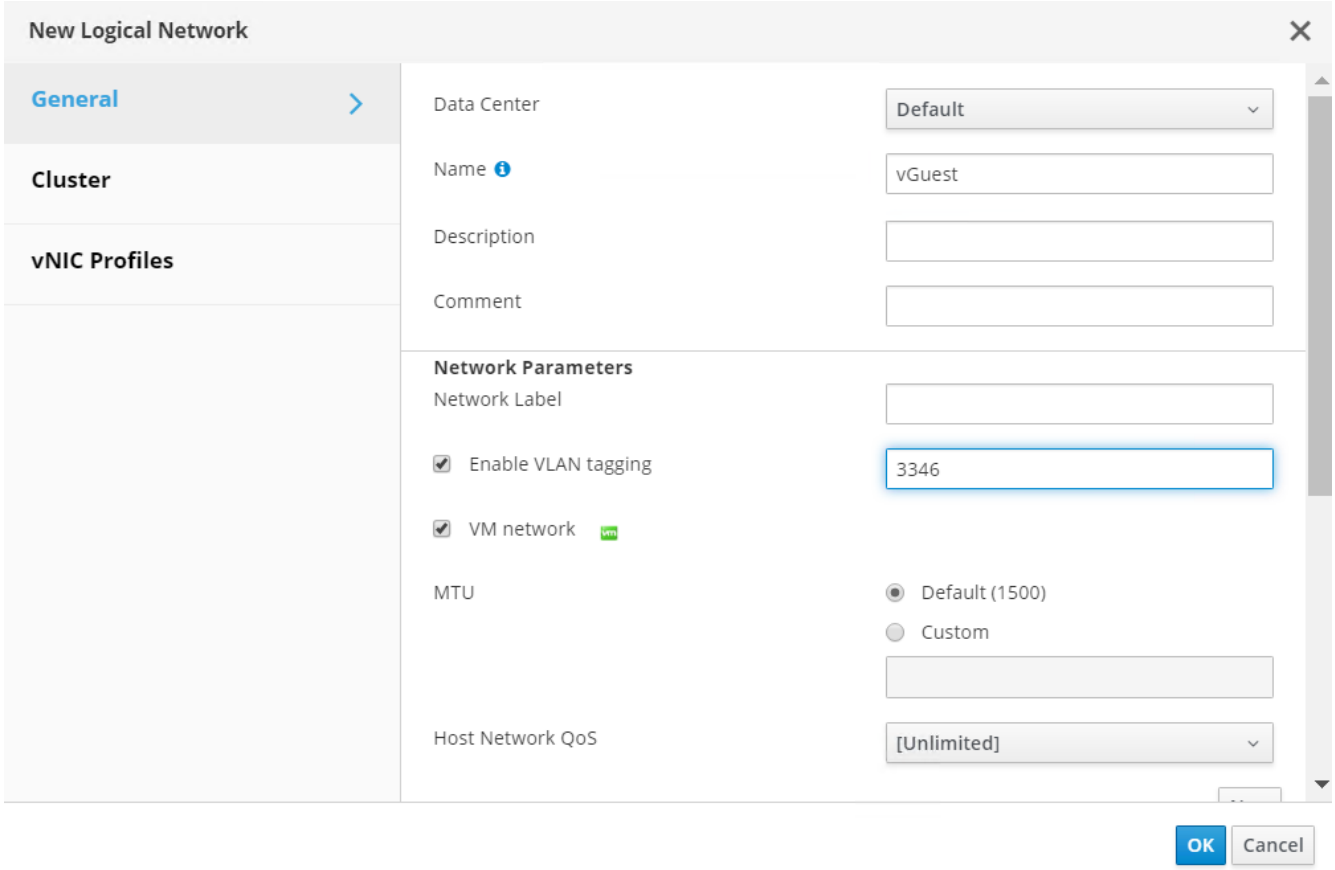
6. Click the pen symbol on the migration network interface under bond0. Configure the IP address details and click OK. Then click OK again in the Setup Host Networks pane.



7. Repeat steps 4 through 6 for the other host as well.
8. The newly created network must be assigned the role of the migration network. Navigate to Compute > Clusters and click the cluster that the RHV hosts belong to, click the Logical Networks sub-tab, and click Manage Networks. For the migration network, enable the checkbox under Migration Network column. Click OK.



9. Next, as a best practice, create a separate VM network rather than using the ovirtmgmt network for VMs.
10. Navigate to Network > Networks and click New. Enter the name of your choice, enable VLAN tagging, and enter the VLAN ID for the VM guest network. Make sure that the checkbox VM Network is checked. Go to the Cluster's sub-tab and make sure that Attach and Require are checked. Then click OK to create the VM guest network.



11. Assign the VM guest logical network to both the hosts. Navigate to Compute > Hosts, click the host names and navigate to the Network Interfaces sub-tab. Then click Setup Host Networks and drag and drop the VM guest logical network into the Assigned Logical Networks column to the right of bond0. There is no need to assign an IP to this logical network, because it provides passthrough networking for the VMs.

    The VM guest network should be able to reach the internet to allow guests to register with Red Hat Subscription Manager.

7. Deploy the NetApp mNode: NetApp HCI with RHV

The management node (mNode) is a VM that runs in parallel with one or more Element software-based storage clusters. It is used for the following purposes:

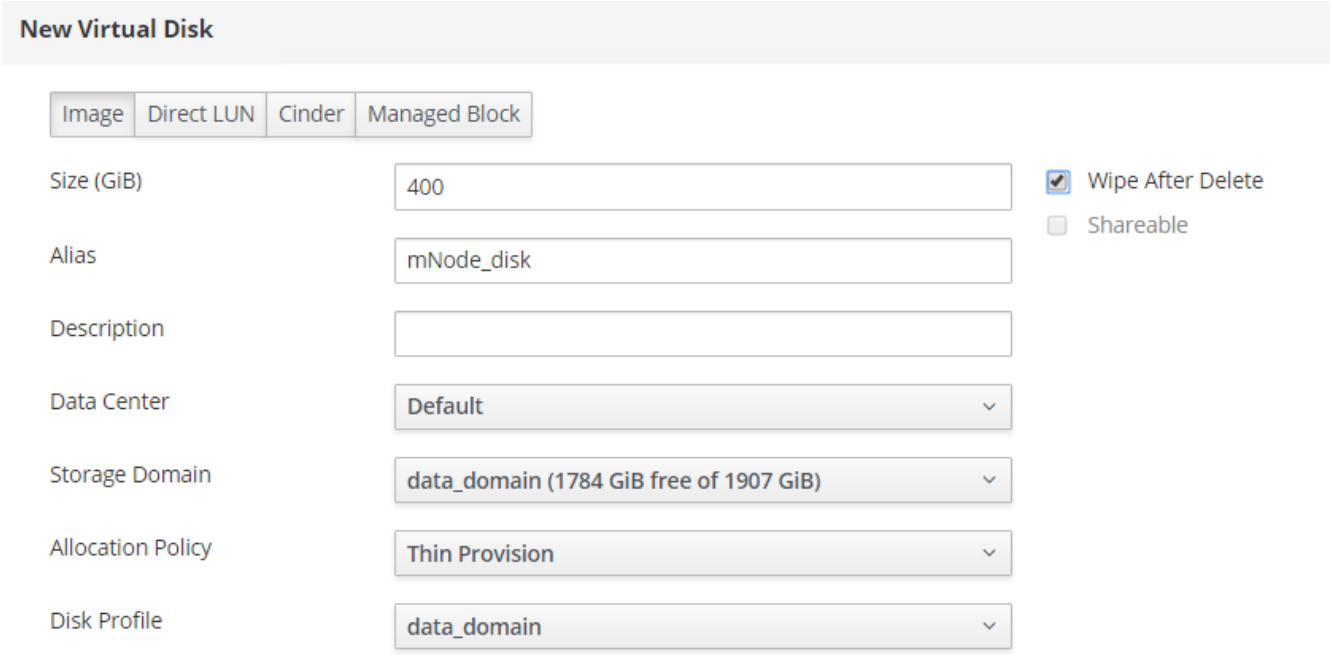
- Providing system services including monitoring and telemetry
- Managing cluster assets and settings
- Running system diagnostic tests and utilities
- Enabling callhome for NetApp ActiveIQ for additional support

To install the NetApp mNode on Red Hat Virtualization, complete the following steps:

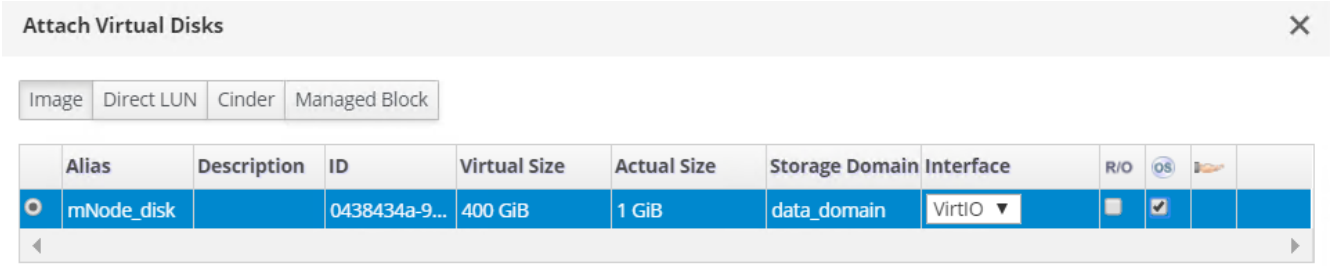
1. Upload the mNode ISO as a disk to the storage domain. Navigate to Storage > Disks > Upload and

click Start. Then click Upload Image and select the downloaded mNode ISO image. Verify the storage domain, the host to perform the upload, and additional details. Then click OK to upload the image to the domain. A progress bar indicates when the upload is complete and the ISO is usable.

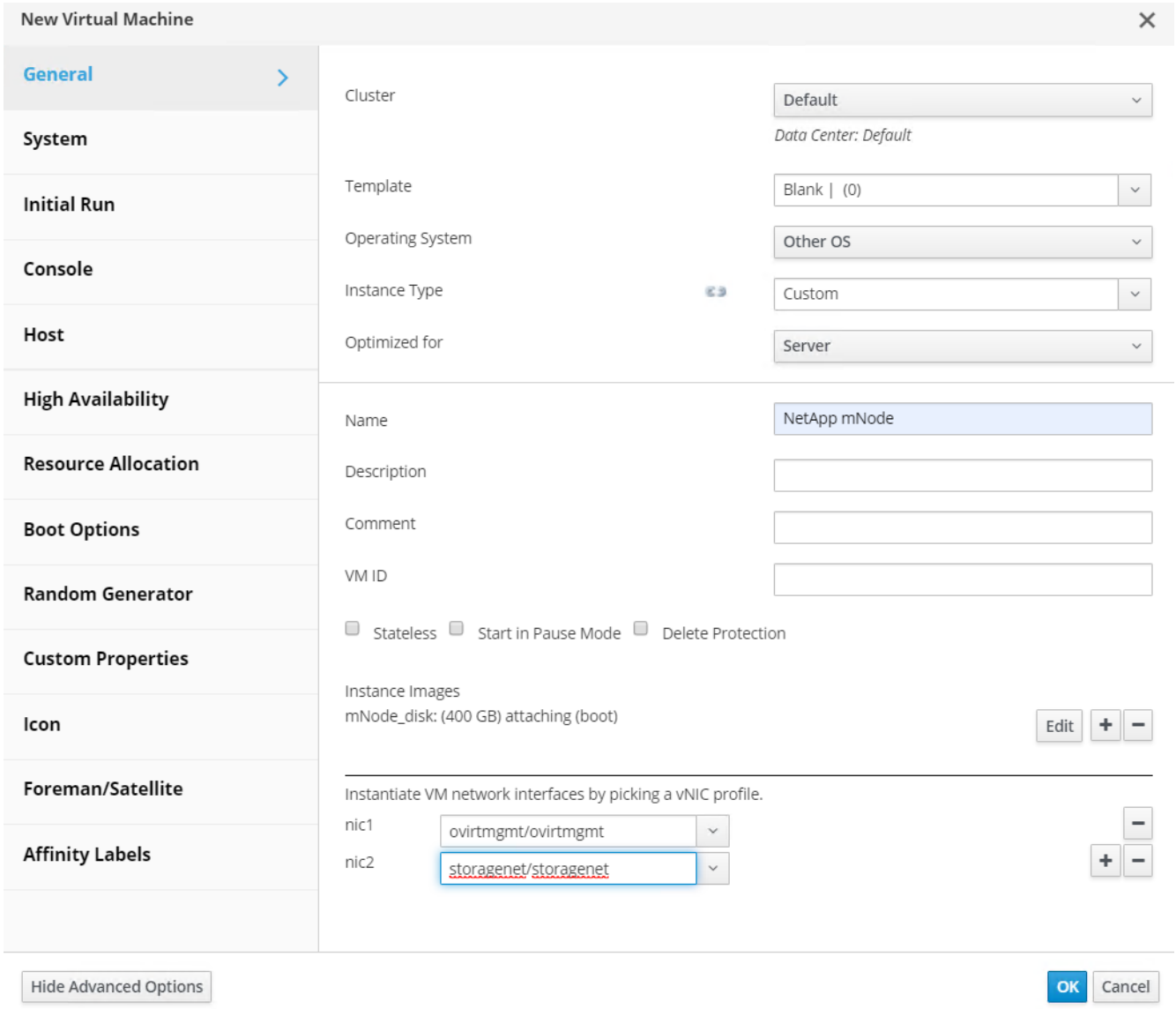
2. Create a VM disk by navigating to Storage > Disks and click New. The mNode disk must be at least 400 GB in size but can be thin-provisioned. In the wizard, enter the name of your choice, select the proper data center, make sure that the proper storage domain is selected, select Thin Provisioning for the allocation policy, and check the Wipe After Delete checkbox. Click OK.



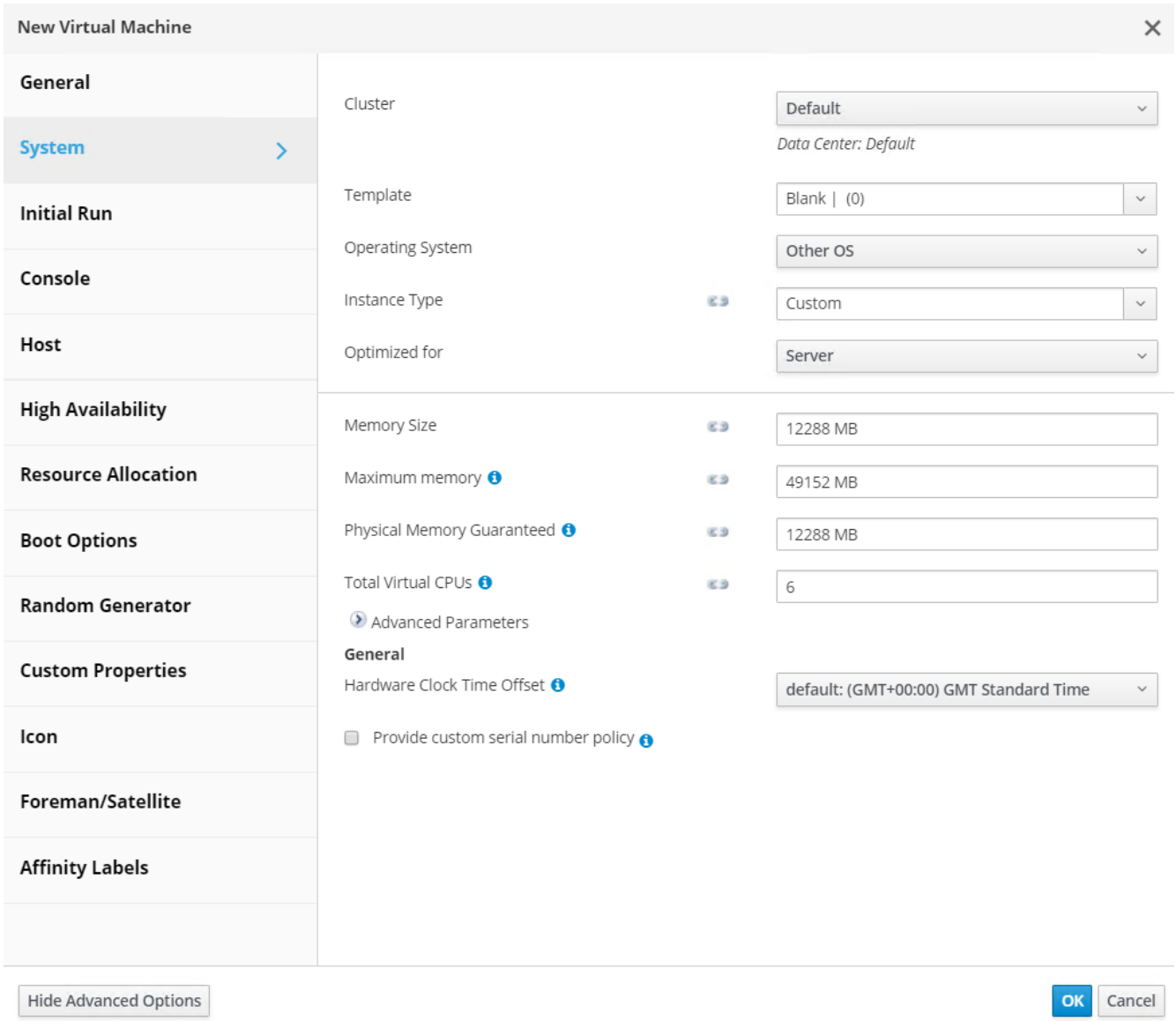
3. Next, navigate to Compute > Virtual Machines and click New. In the General sub-tab, select the appropriate cluster, enter the name of your choice, click attach, and select the disk created in the previous step. Check the box below OS to emphasize that it is a bootable drive. Click OK.



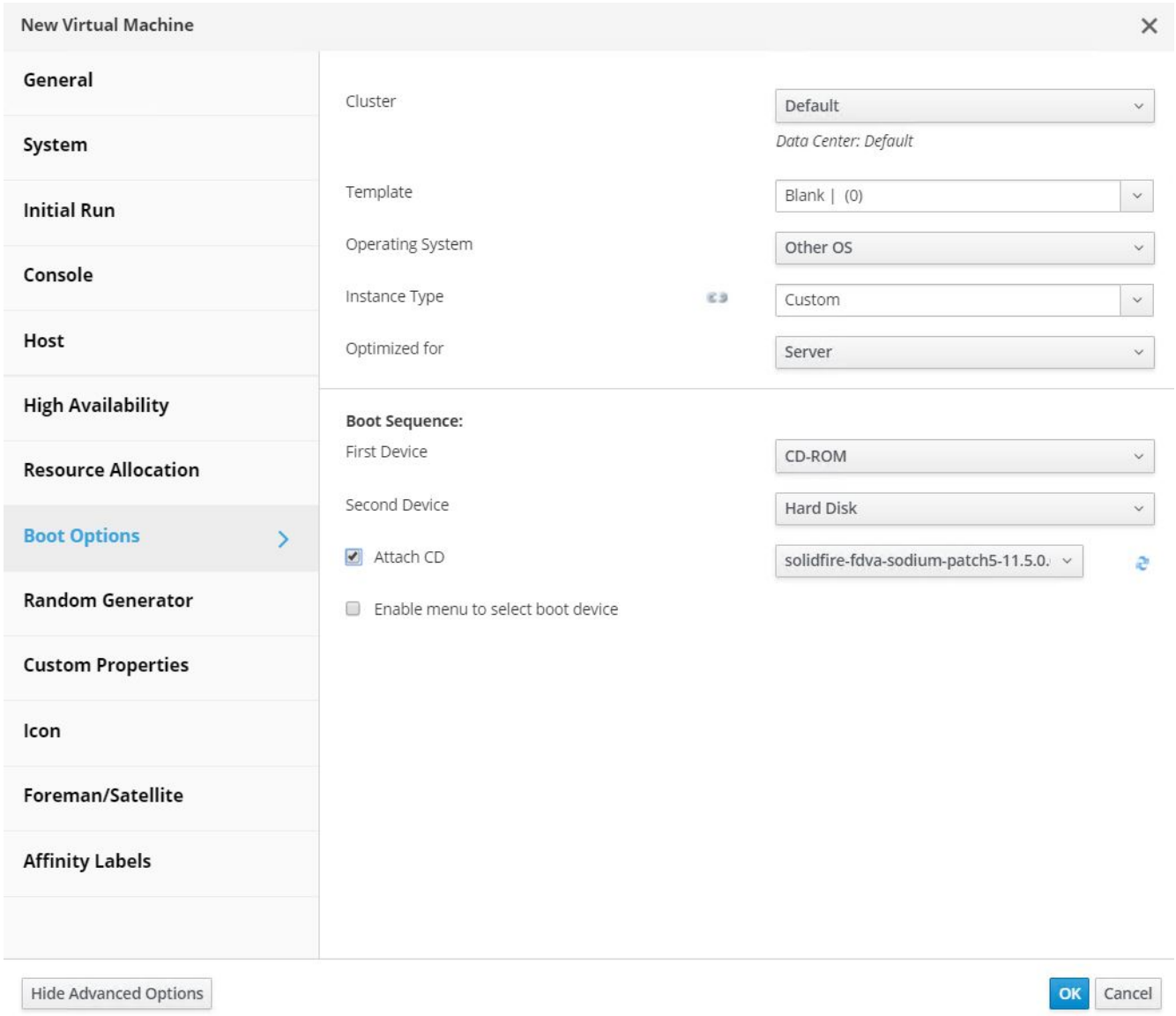
4. Select ovirtmgmt from the dropdown for nic1. Click the (+) sign and select the storage network interface from the dropdown list for nic2.



5. Click the System sub-tab and make sure that it has at least 12GB of memory and 6 virtual CPUs as recommended.



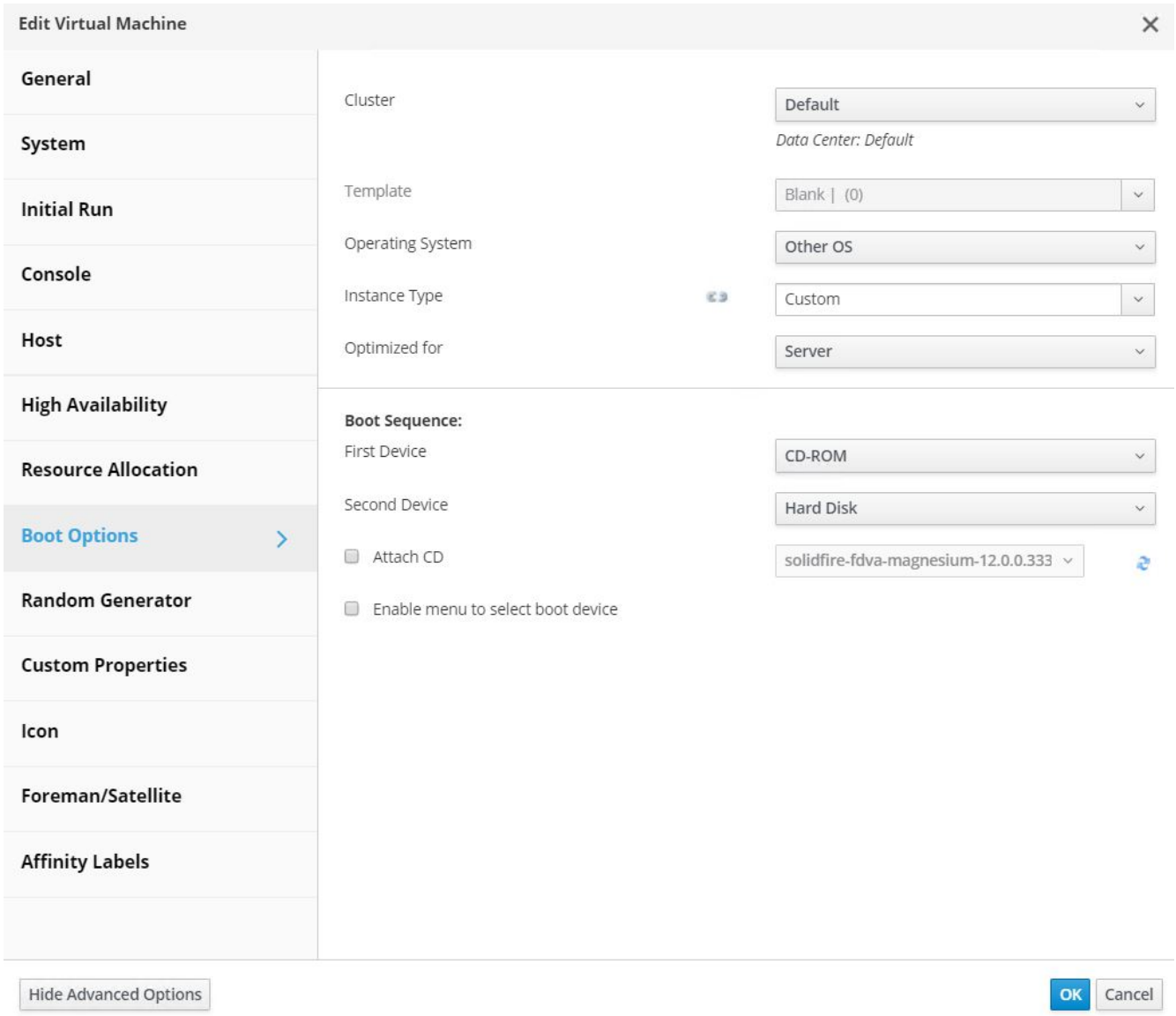
6. Click the Boot Options sub-tab, select CD-ROM as the first device in the boot sequence, select Hard Drive as the second device. Enable Attach CD and attach the mNode ISO. Then click OK.



The VM is created.

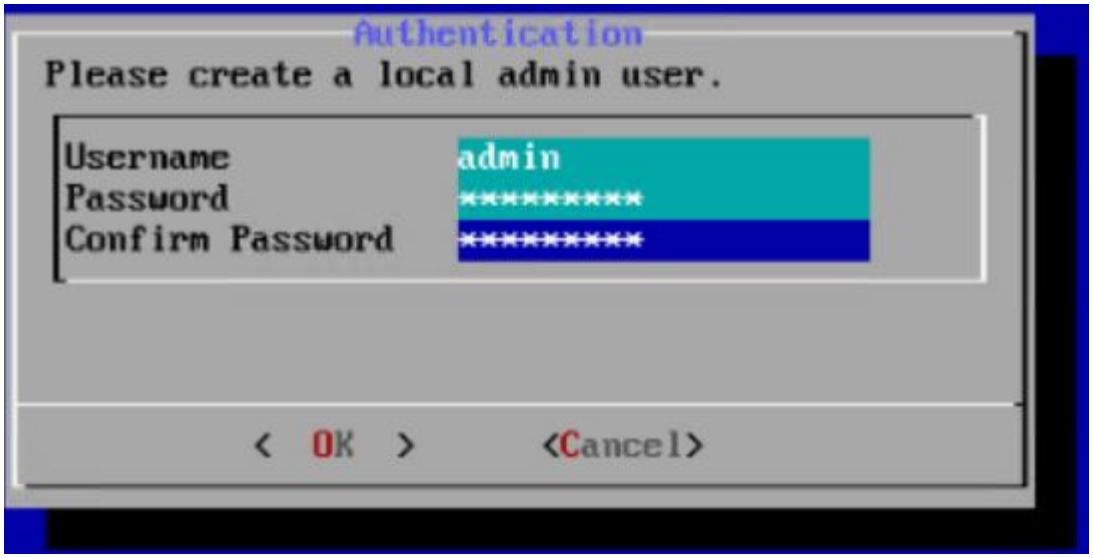
7. After the VM becomes available, power it on, and open a console to it. It begins to load the NetApp Solidfire mNode installer. When the installer is loaded, you are prompted to start the RTFI magnesium installation; type `yes` and press Enter. The installation process begins, and after it is complete, it automatically powers off the VM.

8. Next, click the mNode VM and click Edit. In the Boot Options sub-tab, uncheck the Attach CD checkbox and click the OK button.

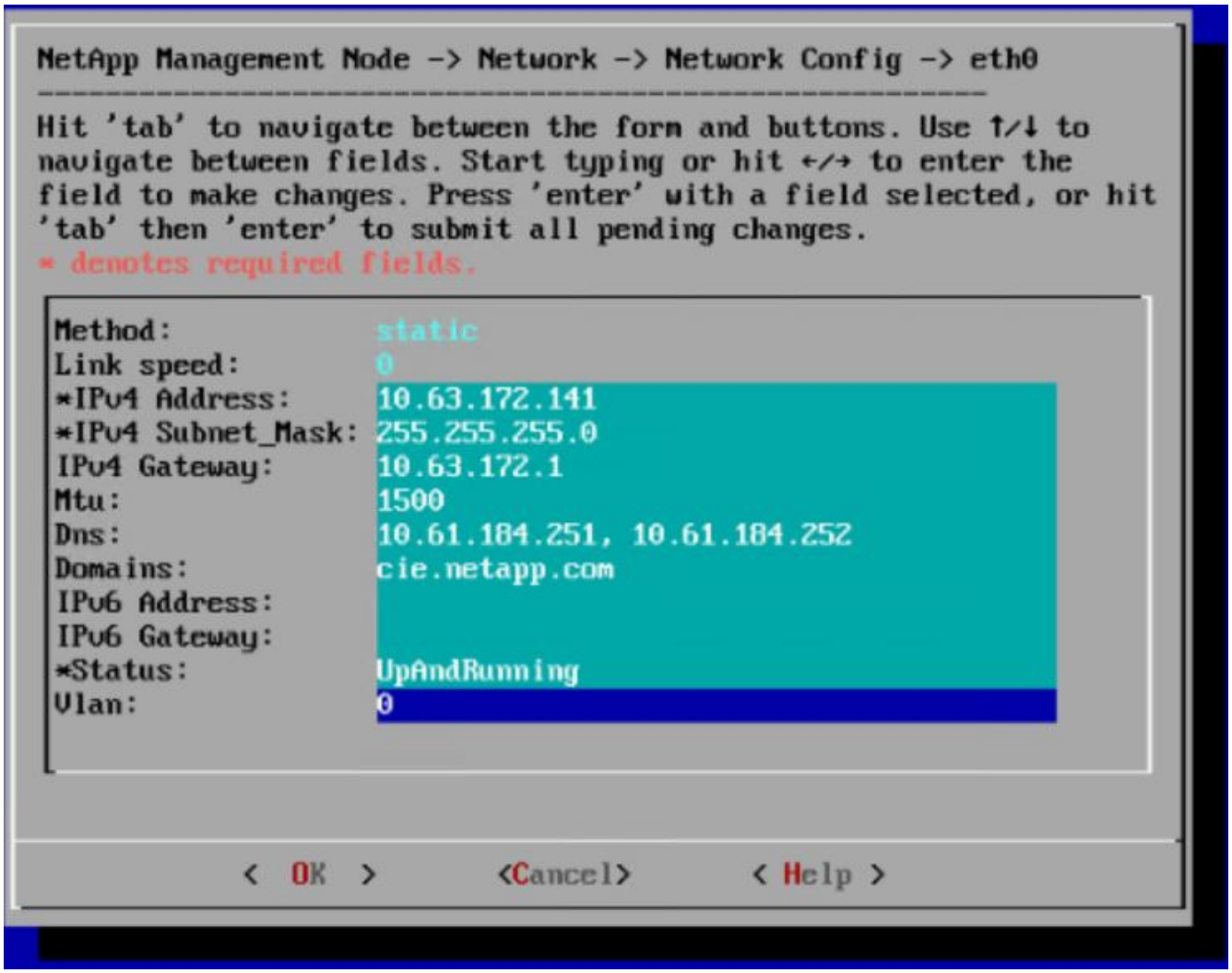


9. Power on the mNode VM. Using the terminal user interface (TUI), create a management node admin user.

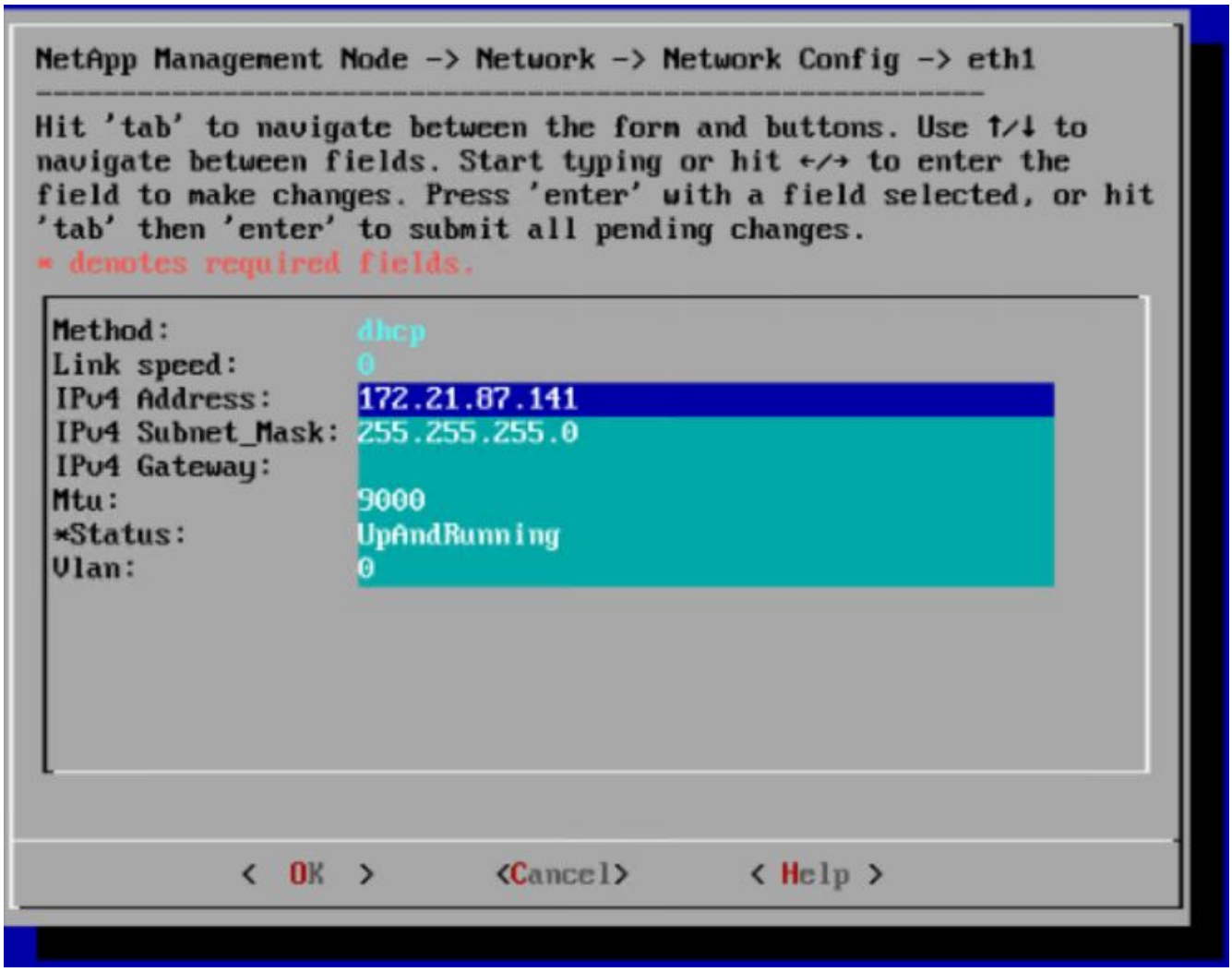
To move through the menu options, press the Up or Down arrow keys. To move through the buttons, press Tab. To move from the buttons to the fields, press Tab. To navigate between fields, press the Up or Down arrow keys.



10. After the user is created, you are returned to a login screen. Log in with the credentials that were just created.
11. To configure the network interfaces starting with the management interface, navigate to Network > Network Config > eth0 and enter the IP address, netmask, gateway, DNS servers, and search domain for your environment. Click OK.



12. Next, configure eth1 to access the storage network. Navigate to Network > Network Config > eth1 and enter the IP address and netmask. Verify that the MTU is 9000. Then click OK.



You can now close the TUI interface.

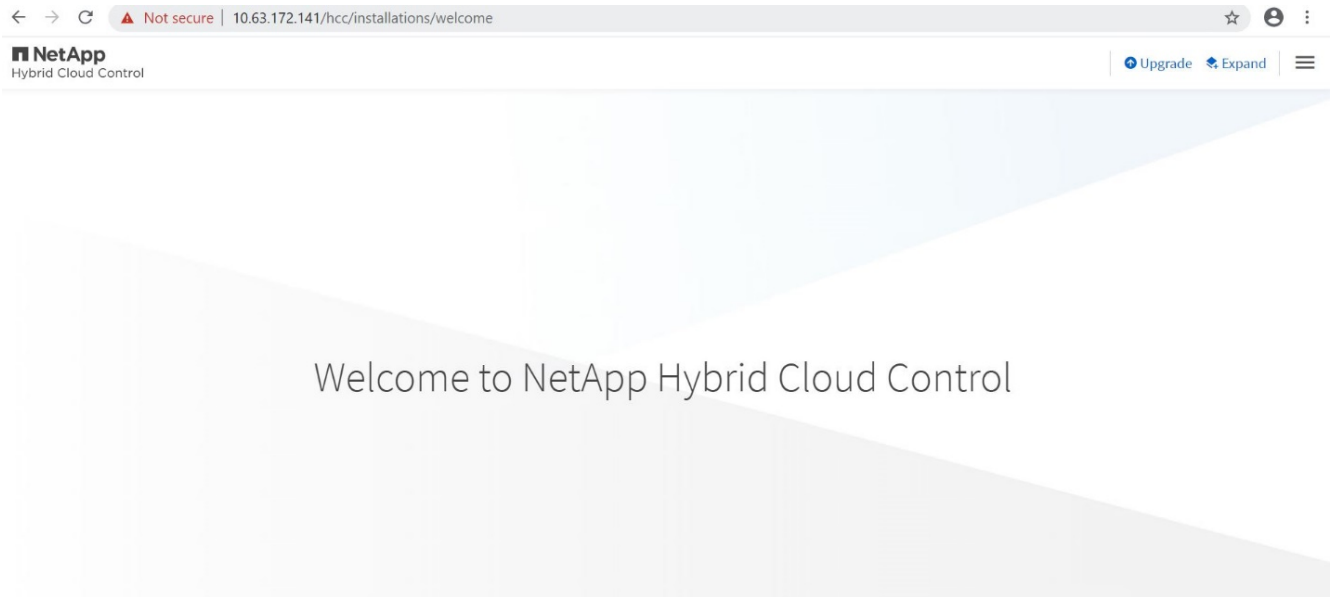
13. SSH into the management node using the management IP, escalate to root and register the mNode with the HCI storage cluster.

```
admin@SF-3D1C ~ $ sudo su

SF-3D1C /home/admin # /sf/packages/mnode/setup-mnode --mnode_admin_user admin
--storage_mvip 10.63.172.140 --storage_username admin --telemetry_active true

Enter the password for storage user admin:
Enter password for mNode user admin:
[2020-05-21T17:19:53.281657Z]:[setup_mnode:296] INFO:Starting mNode deployment
[2020-05-21T17:19:53.286153Z]:[config_util:1313] INFO:No previously running mNode.
Continuing with deployment.
[2020-05-21T17:19:53.286687Z]:[config_util:1320] INFO:Validating credentials for mNode
host.
[2020-05-21T17:19:53.316270Z]:[config_util:1232] INFO:Checking Cluster information.
[2020-05-21T17:19:53.380168Z]:[config_util:112] INFO:Cluster credentials verification
successful.
```

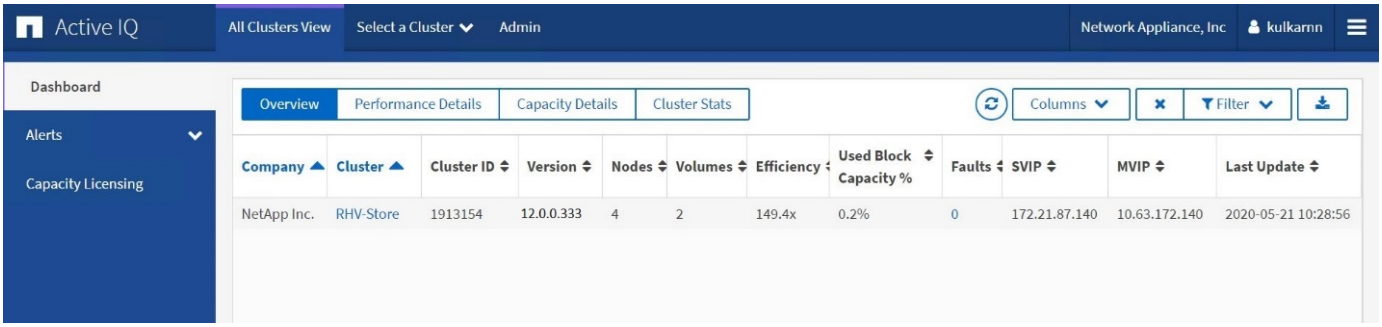
```
[2020-05-21T17:19:53.380665Z]:[config_util:1252] INFO:Cluster version check
successful.
[2020-05-21T17:19:53.458271Z]:[config_util:112] INFO:Successfully queried system
configuration
[2020-05-21T17:19:53.463611Z]:[config_util:497] INFO:CIDR range 172.16.0.0/22 open.
Using for docker ingress.
[2020-05-21T17:19:53.464179Z]:[mnodecfg:141] INFO:Configuring mNode
[2020-05-21T17:19:53.464687Z]:[config_util:194] INFO:Wait for ping of 127.0.0.1 to
succeed
[2020-05-21T17:19:53.475619Z]:[mnodecfg:145] INFO:Validating the supplied MNode
network configuration
[2020-05-21T17:19:53.476119Z]:[mnodecfg:155] INFO:Testing the MNode network
configuration
[2020-05-21T17:19:53.476687Z]:[config_util:353] INFO:Testing network connection to
storage MVIP: 10.63.172.140
[2020-05-21T17:19:53.477165Z]:[config_util:194] INFO:Wait for ping of 10.63.172.140 to
succeed
[2020-05-21T17:19:53.488045Z]:[config_util:356] INFO:Successfully reached storage
MVIP: 10.63.172.140
[2020-05-21T17:19:53.488569Z]:[mnodecfg:158] INFO:Configuring MNode storage (this can
take several minutes)
[2020-05-21T17:19:57.057435Z]:[config_util:536] INFO:Configuring MNode storage
succeeded.
[2020-05-21T17:19:57.057938Z]:[config_util:445] INFO:Replacing default ingress
network.
[2020-05-21T17:19:57.078685Z]:[mnodecfg:163] INFO:Extracting services tar (this can
take several minutes)
[2020-05-21T17:20:36.066185Z]:[config_util:1282] INFO:Extracting services tar
succeeded
[2020-05-21T17:20:36.066808Z]:[mnodecfg:166] INFO:Configuring MNode authentication
[2020-05-21T17:20:36.067950Z]:[config_util:1485] INFO:Updating element-auth
configuration
[2020-05-21T17:20:41.581716Z]:[mnodecfg:169] INFO:Deploying MNode services (this can
take several minutes)
[2020-05-21T17:20:41.810264Z]:[config_util:557] INFO:Deploying MNode services
succeeded
[2020-05-21T17:20:41.810768Z]:[mnodecfg:172] INFO:Deploying MNode Assets
[2020-05-21T17:20:42.162081Z]:[config_util:122] INFO:Retrying 1/45 time...
[2020-05-21T17:20:42.162640Z]:[config_util:125] INFO:Waiting 10 seconds before next
attempt.
[2020-05-21T17:20:52.199224Z]:[config_util:112] INFO:Mnode is up!
[2020-05-21T17:20:52.280329Z]:[config_util:112] INFO:Root asset created.
[2020-05-21T17:20:52.280859Z]:[config_util:122] INFO:Retrying 1/5 time...
[2020-05-21T17:20:52.281280Z]:[config_util:125] INFO:Waiting 10 seconds before next
attempt.
[2020-05-21T17:21:02.299565Z]:[config_util:112] INFO:Successfully queried storage
assets
[2020-05-21T17:21:02.696930Z]:[config_util:112] INFO:Storage asset created.
```

```
[2020-05-21T17:21:03.238455Z]:[config_util:112] INFO:Storage asset registered.
[2020-05-21T17:21:03.241966Z]:[mnodecfg:175] INFO:Attempting to set up VCP-SIOC
credentials
[2020-05-21T17:21:03.242659Z]:[config_util:953] INFO:No VCP-SIOC credential given from
NDE. Using default credentials for VCP-SIOC service.
[2020-05-21T17:21:03.243117Z]:[mnodecfg:185] INFO:Configuration Successfully Completed
```

14. Using a browser, log into the management node GUI using `https://<mNodeIP >`. mNode or Hybrid Cloud Control facilitates expansion, monitoring, and upgrading the Element cluster.



15. Click the three parallel lines on the top right and click View Active IQ. Search for the HCI storage cluster by filtering the cluster name and make sure that it is logging the most recent updates.



Best Practices for Production Deployments

Updating RHV Manager and RHV-H Hosts: NetApp HCI with RHV

It is a recommended best practice to make sure that both the RHV Manager and the RHV-H hosts have the latest security and stability updates applied to make sure that the environment is protected and continues to run as expected. To apply the updates to the hosts in the deployment, they must first be subscribed to either the

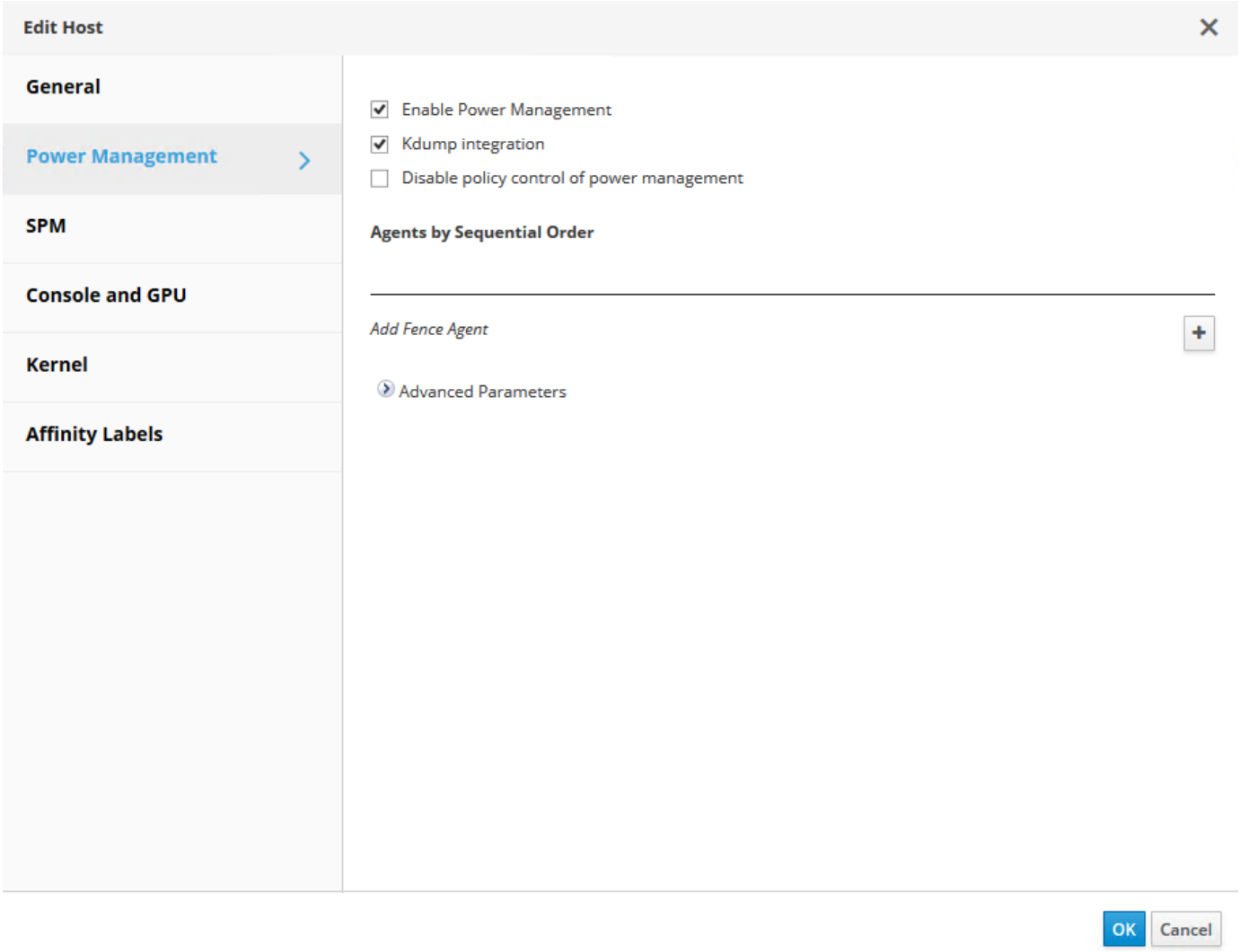
Red Hat Content Delivery Network or a local Red Hat Satellite repository. The tasks involved in updating the platform include updating the manager VM and afterward updating each physical host non-disruptively after ensuring virtual guests are migrated to another node in the cluster.

Official documentation to support the upgrade of RHV 4.3 between minor releases can be found here.

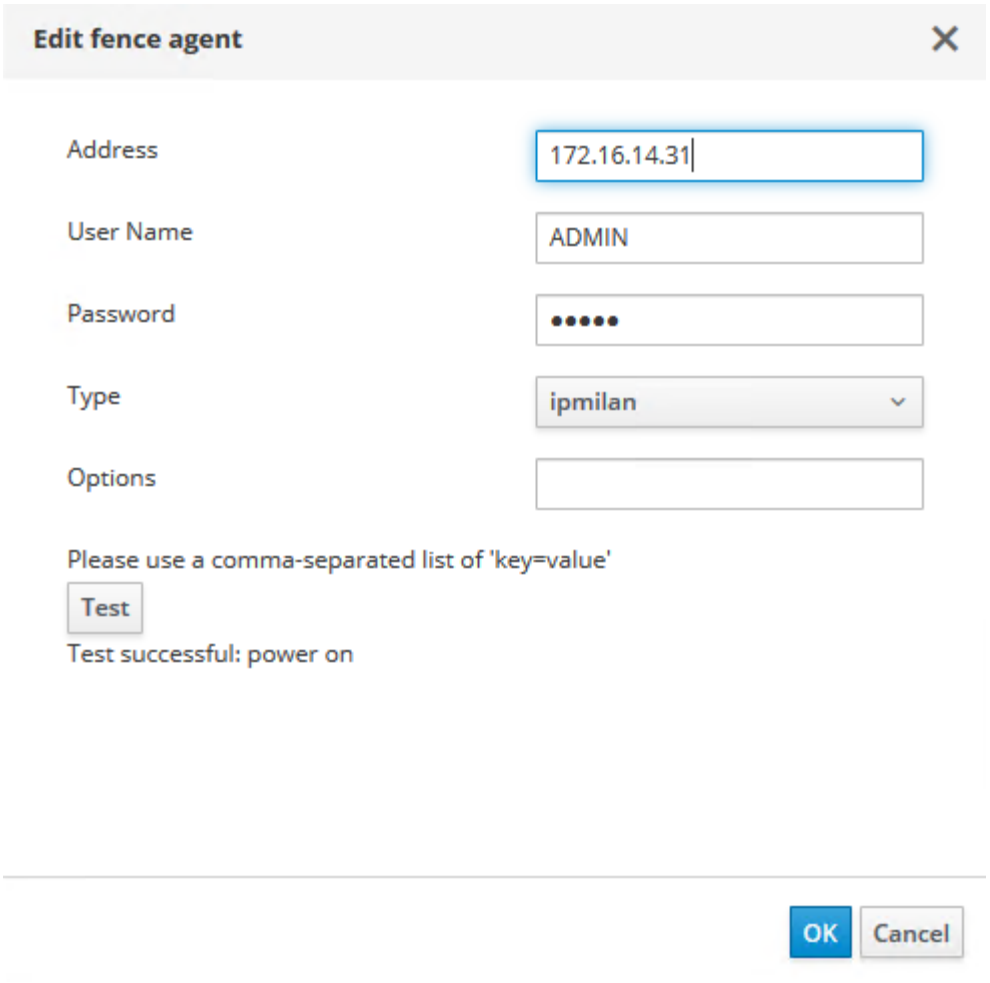
Enabling Fencing for RHV-H Hosts: NetApp HCI with RHV

Fencing is a process by which the RHV Manager can provide high availability of the VMs in the environment by automatically shutting down a non-responsive hypervisor host. It does this by sending commands to a fencing agent, which in the case of NetApp HCI is available through the IPMI out-of-band management interface on the compute nodes and rebooting the host. This action releases the locks that the non-responsive hypervisor node has on VM disks and allows for those virtual guests to be restarted on another node in the cluster without risking data corruption. After the host completes its boot process, it automatically attempts to rejoin the cluster it was a part of prior to the shutdown. If it is successful, it is once again allowed to host VMs.

To enable fencing, each host must have power management enabled; this can be found by highlighting the host and clicking the Edit button in the upper right-hand corner or by right-clicking on the host and selecting Edit.



After power management is enabled, the next step involves configuring a fencing agent. Click on the plus sign (+) near the Add Fence Agent, and a new window pops up that must be filled out with the information for the IPMI connection on the NetApp HCI compute nodes. The type of connection is IPMILAN, and the agent needs the IP address, username, and password for the console login. After you have provided this information, you can click test to validate the configuration. If properly configured, it should report the current power status of the node.



With fencing enabled, the RHV environment is configured to support a highly available deployment should one of the hypervisor nodes become nonresponsive.

Optimizing Memory for Red Hat Virtualization: NetApp HCI with RHV

One of the primary benefits for deploying a virtual infrastructure is to enable the more efficient use of physical resources in the environment. In a case in which the guest VMs underutilize the memory allotted, you can use memory overcommitment to optimize memory usage. With this feature, the sum of the memory allocated to guest VMs on a host is allowed to exceed the amount of physical memory on that host.

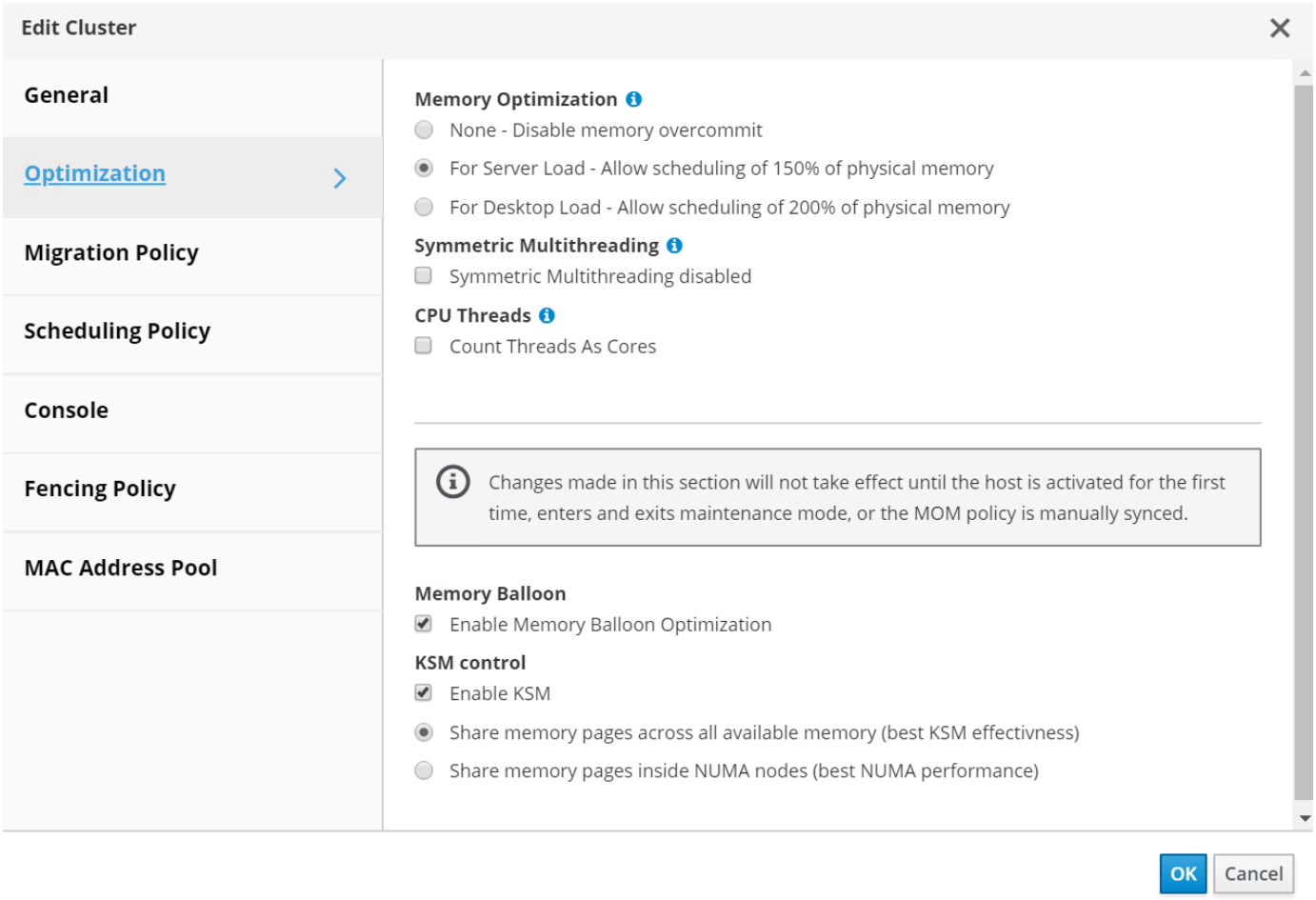
The concept behind memory overcommitment is similar to thin provisioning of storage resources. At any given moment, every VM on the host does not use the total amount of memory allocated to it. When one VM has excess memory, its unused memory is available for other VMs to use. Therefore, an end user can deploy more VMs that the physical infrastructure would not normally allow. Memory overcommitment on the hosts in the cluster is handled by Memory Overcommit Manager (MoM). Techniques like memory ballooning and Kernel Same-page Merging (KSM) can improve memory overcommitment depending on the kind of workload.

Memory ballooning is a memory management technique which allows a host to artificially expand its memory by reclaiming unused memory that was previously allocated to various VMs, with a limitation of the guaranteed memory size of every VM. For memory ballooning to work, each VM by default has a balloon device with the necessary drivers. Ballooning essentially is a cooperative operation between the VM driver and the host. Depending on the memory needs of the host, it instructs the guest OS to inflate (provide memory to host) or deflate (regain the memory) the balloon which is controlled by the balloon device.

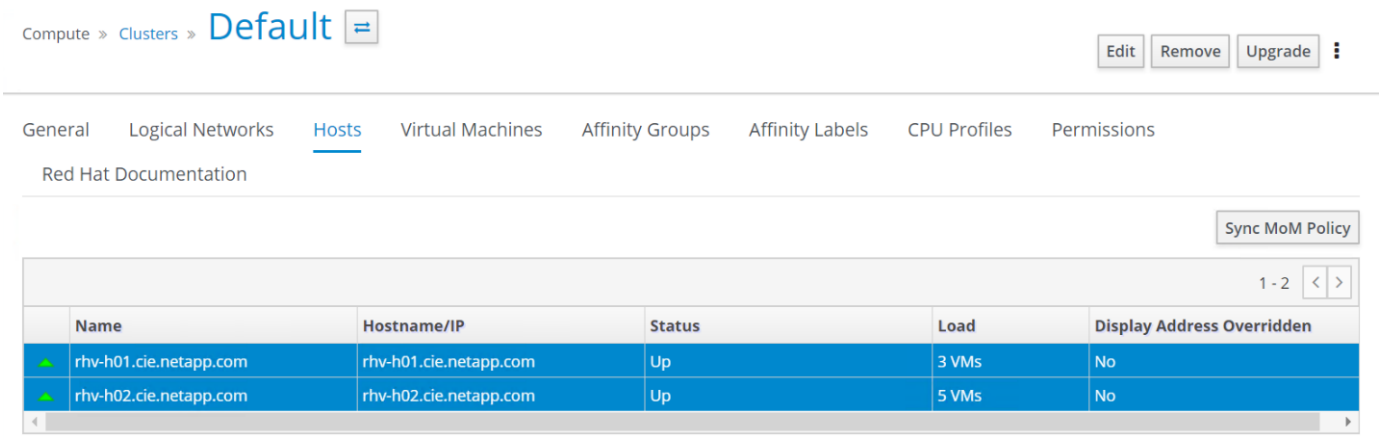
Kernel Same-page Merging (KSM) allows the host kernel to examine two or more running VMs and compare their image and memory. If any memory regions or pages are identical, KSM reduces multiple identical memory pages to a single page. This page is then marked ‘copy on write’ and a new page is created for that guest VM if the contents of the page are modified by a guest VM.

Both features can be enabled at a cluster level to apply to all hosts in that cluster. To enable these features, navigate to Compute > Clusters, select the desired cluster and click Edit. Then click the Optimization sub-tab and perform the following steps based on your requirements:

1. Depending on the use-case and workload, enable Memory Optimization to allow overcommitment of memory to either 150% or 200% of the available physical memory.
2. To enable memory ballooning, check the Enable Memory Balloon Optimization checkbox.
3. To enable KSM, check the Enable KSM checkbox.
4. Click Ok to confirm the changes.



Be aware that after these changes have been applied, they do not take effect until you manually sync the MoM policy. To sync the MoM policy, navigate to Compute > Clusters and click the cluster for which you made the optimization changes. Navigate to the Hosts sub-tab, select all the hosts, and then click Sync MoM Policy.



KSM and ballooning can free up some memory on the host and facilitate overcommitment, but, if the amount of shareable memory decreases and the use of physical memory increases, it might cause an out-of-memory condition. Therefore, the administrator should be sure to reserve enough memory to

avoid out-of-memory conditions if the shareable memory decreases.

In some scenarios, memory ballooning may collide with KSM. In such situations, MoM tries to adjust the balloon size to minimize collisions. Also, there can be scenarios for which ballooning might cause sub-optimal performance. Therefore, depending on the workload requirements, you can consider enabling either or both the techniques.

Where to Find Additional Information: NetApp HCI with RHV

To learn more about the information described in this document, review the following documents and/or websites:

- NetApp HCI Documentation https://www.netapp.com/us/documentation/hci.aspx
- Red Hat Virtualization Documentation https://access.redhat.com/documentation/en-us/red_hat_virtualization/4.3/

TR-4857: NetApp HCI with Cisco ACI

Abhinav Singh, Nikhil M Kulkarni, NetApp

Cisco Application Centric Infrastructure (Cisco ACI) is an industry-leading, secure, open, and comprehensive Software-Defined Networking (SDN) solution. Cisco ACI radically simplifies, optimizes, and accelerates infrastructure deployment and governance, and it expedites the application deployment lifecycle. Cisco ACI deployed in data centers is proven to work with NetApp HCI with full interoperability. You can manage Ethernet networks for compute, storage, and access with Cisco ACI. You can establish and manage secure network segments for server-to-server and virtual machine (VM)-to-VM communications as well as secure storage-network access through iSCSI from server-to-NetApp HCI storage. This level of endpoint-to-endpoint network security allows customers to architect and operate NetApp HCI in a more secure fashion.

Next: Use Cases

Use Cases

The NetApp HCI with Cisco ACI solution delivers exceptional value for customers with the following use cases:

- On-premises software-defined compute, storage, and networking infrastructure
- Large enterprise and service-provider environments
- Private cloud (VMware and Red Hat)
- End User Computing and Virtual Desktop Infrastructure
- Mixed-workload and mixed-storage environments

Next: Architecture

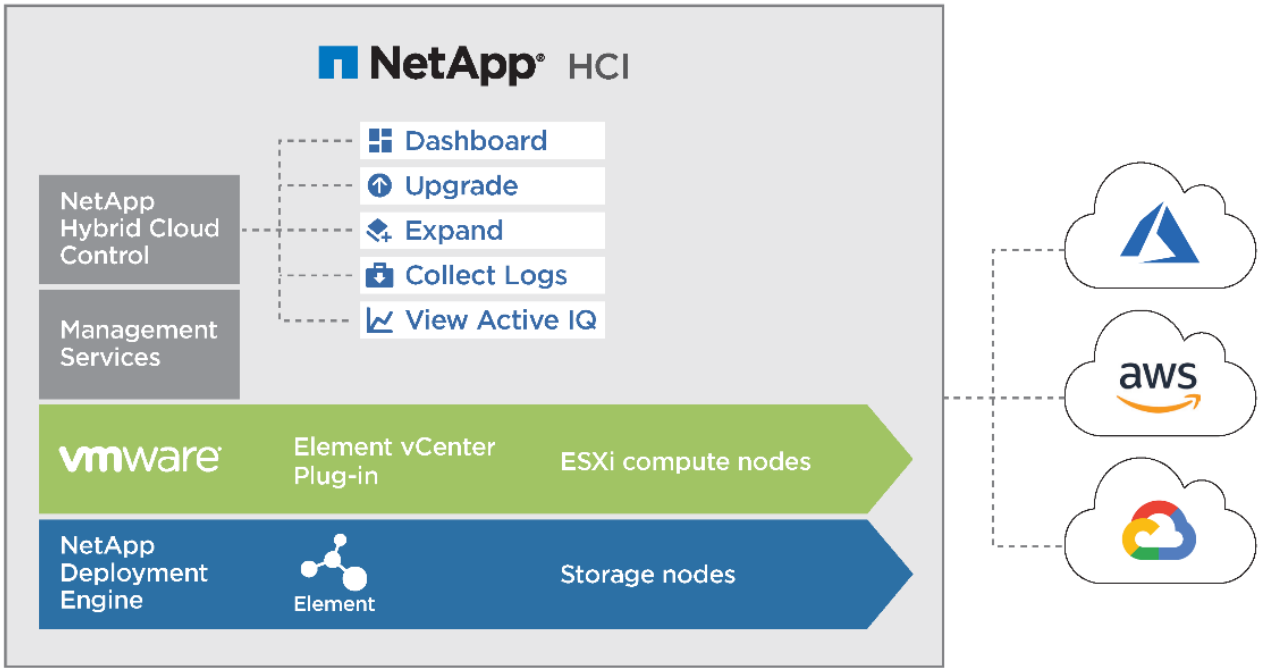
Architecture

Solution Technology

This document outlines the best practices to follow for a fully featured on-premises data center or private cloud while interoperating NetApp HCI with Cisco ACI. To demonstrate workload independence, networking best practices are extended to virtualization solutions, including VMware vSphere and Red Hat Virtualization when deployed over NetApp HCI, and to other storage solutions like NetApp ONTAP and StorageGRID. It also emphasizes the interoperability of Cisco ACI switches with different virtual switches, for example, VMware Distributed Switch (VDS), Cisco ACI Virtual Edge (AVE), Linux Bridge, or Open vSwitch.

NetApp HCI

NetApp HCI is an enterprise-scale, hyper-converged infrastructure solution that delivers compute and storage resources in an agile, scalable, easy-to-manage architecture. Running multiple enterprise-grade workloads can result in resource contention, where one workload interferes with the performance of another. NetApp HCI alleviates this concern with storage quality-of-service (QoS) limits that are available natively within NetApp Element software. Element enables the granular control of every application and volume, helps to eliminate noisy neighbors, and satisfies enterprise performance SLAs. NetApp HCI multitenancy capabilities can help eliminate many traditional performance related problems. See the following graphic for an overview of NetApp HCI.



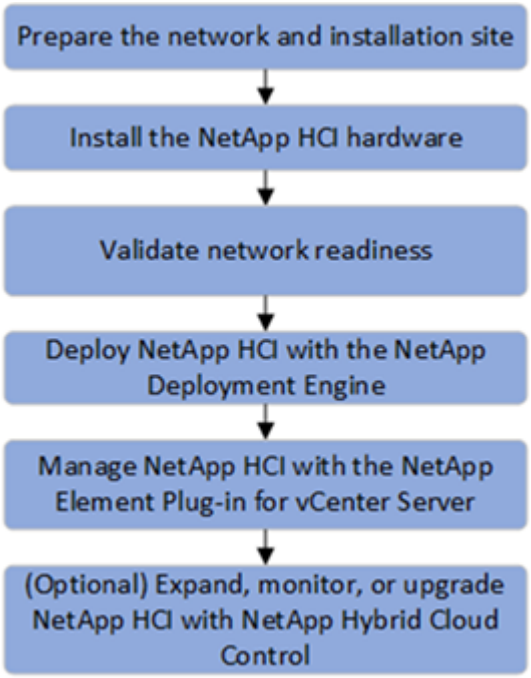
NetApp HCI streamlines installation through the NetApp Deployment Engine (NDE), an intuitive

deployment engine that automates more than 400 inputs to fewer than 30 to get your setup running in about 45 minutes. In addition, a robust suite of APIs enables seamless integration into higher-level management, orchestration, backup, and disaster recovery tools. With the NetApp Hybrid Cloud Control management suite, you can manage, monitor, and upgrade your entire infrastructure throughout its lifecycle through a single pane of glass.

Software-Defined Architecture

NetApp HCI provides a software-defined approach for deploying and managing data and storage resources. NetApp HCI uses NetApp Element software to provide an easy-to-use GUI-based portal and REST-based API for storage automation, configuration, and management. NetApp Element software provides modular and scalable performance, with each storage node delivering guaranteed capacity and throughput to the environment.

NetApp HCI uses the NetApp Deployment Engine (NDE) to automate the configuration and deployment of physical infrastructure, including the installation and configuration of the VMware vSphere environment and the integration of the NetApp Element Plug-in for vCenter Server. The following figure depicts an overview of the process for deploying NetApp HCI.



Performance Guarantee

A common challenge is delivering predictable performance when multiple applications are sharing the same infrastructure. An application interfering with other applications creates performance degradation. Mainstream applications have unique I/O patterns that can affect each other's performance when deployed in a shared environment. To address these issues, the NetApp HCI Quality of Service (QoS) feature allows fine-grained control of performance for every application, thereby eliminating noisy neighbors and satisfying performance SLAs. In NetApp HCI, each volume is configured with minimum, maximum, and burst IOPS values. The minimum IOPS setting guarantees

performance, independent of what other applications on the system are doing. The maximum and burst values control allocation, enabling the system to deliver consistent performance to all workloads.

NetApp Element software uses the iSCSI storage protocol, a standard way to encapsulate SCSI commands on a traditional TCP/IP network. Element uses a technique called iSCSI login redirection for better performance. iSCSI login redirection is a key part of the NetApp Element software cluster. When a host login request is received, the node decides which member of the cluster should handle the traffic based on IOPS and the capacity requirements for the volume. Volumes are distributed across the NetApp Element software cluster and are redistributed if a single node is handling too much traffic for its volumes or if a new node is added. Multiple copies of a given volume are allocated across the array. In this manner, if a node failure is followed by volume redistribution, there is no effect on host connectivity beyond a logout and login with redirection to the new location. With iSCSI login redirection, a NetApp Element software cluster is a self-healing, scale-out architecture that is capable of nondisruptive upgrades and operations.

Interoperability

Previous generations of hyperconverged infrastructure typically required fixed resource ratios, limiting deployments to four-node and eight-node configurations. NetApp HCI is a disaggregated hyper-converged infrastructure that can scale compute and storage resources independently. Independent scaling prevents costly and inefficient overprovisioning and simplifies capacity and performance planning.

The architectural design choices offered enables you to confidently scale on your terms, making HCI viable for core Tier-1 data center applications and platforms. It is architected in building blocks at either the chassis or the node level. Each chassis can hold four nodes in a mixed configuration of storage or compute nodes. NetApp HCI is available in mix-and-match, small, medium, and large storage and compute configurations.

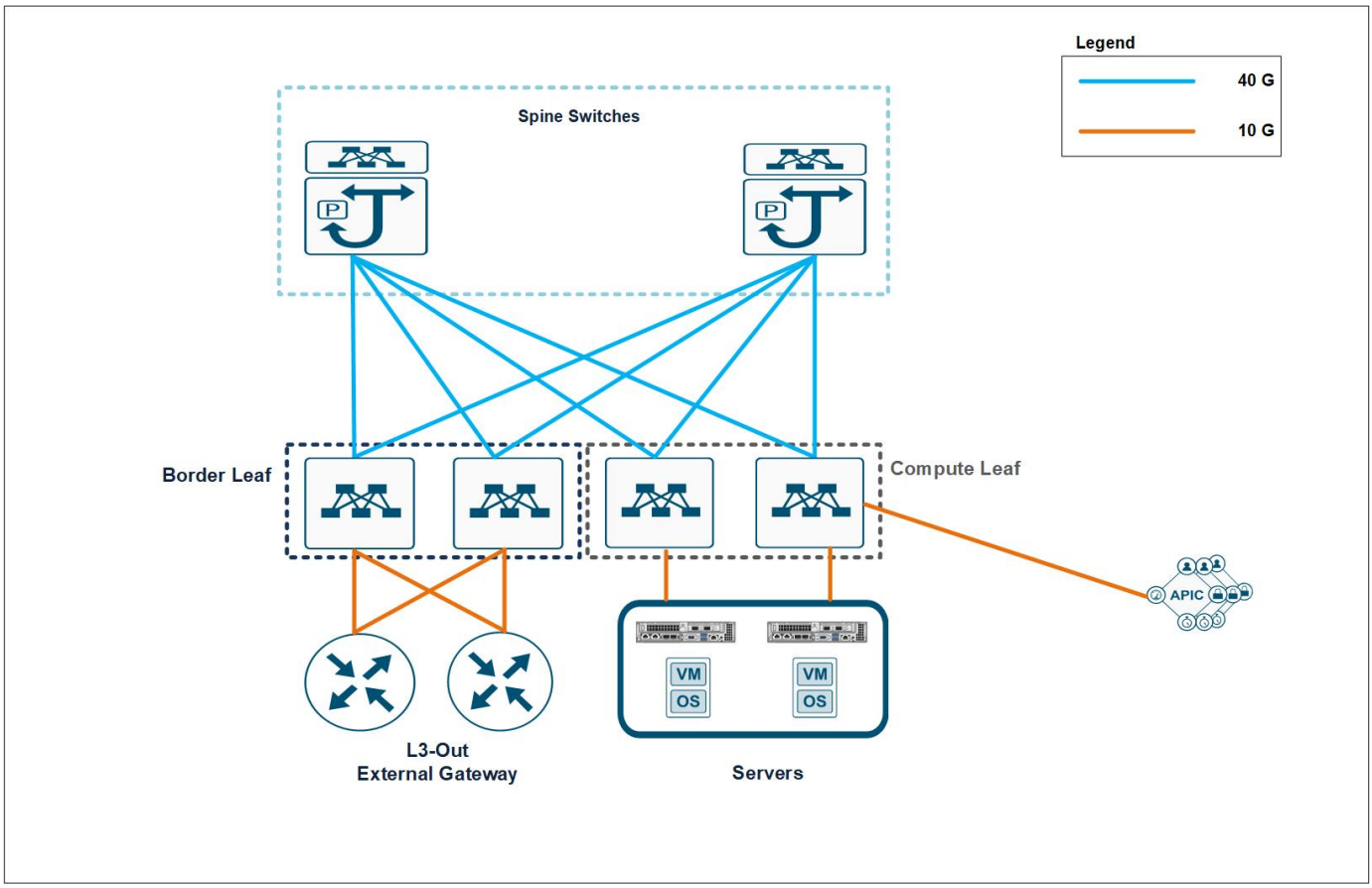
NetApp HCI provides proven multiprotocol and hybrid- cloud support with enterprise grade features. It also offers easy interoperability with multiple different host virtualization technologies and storage solutions. Deploying ONTAP Select and StorageGRID as appliances expands NetApp HCI storage capabilities to include file, block, and object storage services. NetApp HCI provides an agile infrastructure platform for virtual data centers of different flavors. VMware vSphere, Red Hat Virtualization, KVM, Citrix Hypervisor, and so on are supported platforms that can use the NetApp HCI infrastructure to provide a scalable, enterprise-grade on-premises virtual environment.

For more details, see the NetApp HCI documentation.

Cisco ACI

Cisco ACI is an industry leading software-defined networking solution that facilitates application agility and data center automation. Cisco ACI has a holistic architecture with a centralized policy-driven management. It implements a programmable data center Virtual Extensible LAN (VXLAN) fabric that delivers distributed networking and security for any workload, regardless of its nature (virtual, physical, container, and so on).

Cisco pioneered the introduction of intent-based networking with Cisco ACI in the data center. It combines the high- performance hardware and robust software integrated with two important SDN features—overlays and centralized control. The ACI fabric consists of Cisco Nexus 9000 series switches running in ACI mode and a cluster of at least three centrally managed Application Policy Infrastructure Controllers (APIC) servers. The following figure provides an overview of Cisco ACI.

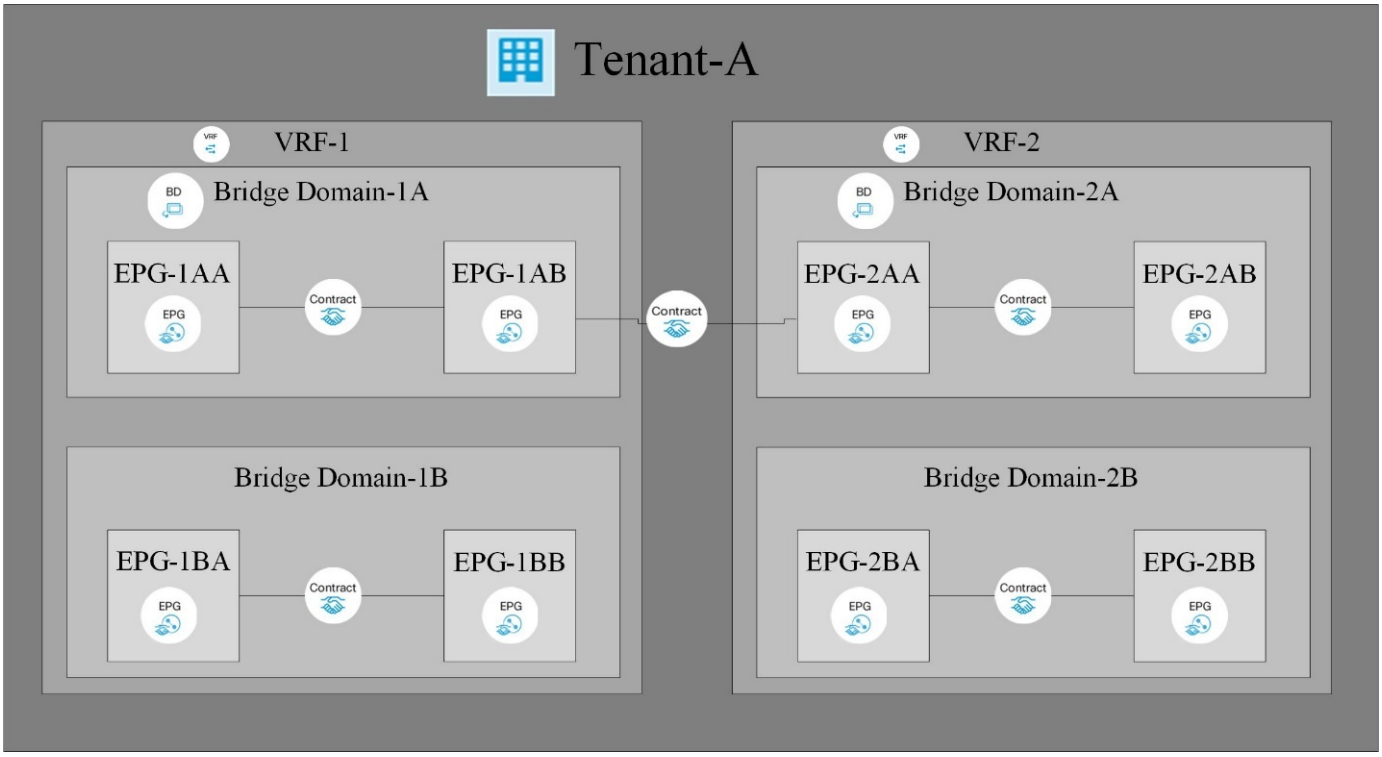


Policy-Driven Networking

Cisco ACI, with its policy driven model, makes network hardware stateless. The Application Policy Infrastructure Controller (APIC) acts as the central controller managing and configuring all the switches in the ACI fabric. The Cisco ACI fabric consists of Cisco Nexus 9000 series switches which are centrally configured and managed by the cluster of APICs using the declarative policy model.

Cisco ACI uses logical constructs to form a layered policy architecture to define and manage the different functions of the entire fabric, including infrastructure, authentication, security, services, applications, and diagnostics.

The following figure depicts the categorization and relation between different logical constructs in Cisco ACI.



Tenants are logical containers with administrative boundaries that exercise domain-based access control. It is a logical policy isolation and does not equate to a real network construct.

Within the tenant, a context is a unique layer-3 forwarding policy domain. A context can be directly mapped to the Virtual Routing and Forwarding (VRF) concept of traditional networks. In fact, a context is also called VRF. Because each context is a separate layer- 3 domain, two different contexts can have overlapping IP spaces.

Within a context, a bridge domain (BD) represents a unique layer-2 forwarding construct. The bridge domain defines the unique layer-2 MAC address space and can be equated to a layer-2 flood domain or to a layer-3 gateway. A bridge domain can have zero subnets, but it must have at least one subnet if it is to perform routing for the hosts residing in the BD.

In ACI, an endpoint is anything that communicates on the network, be it a compute host, a storage device, a network entity that is not part of the ACI fabric, a VM, and so on. A group of endpoints that have the same policy requirements are categorized into an Endpoint Group (EPG). An EPG is used to configure and manage multiple endpoints together. An EPG is a member of a bridge domain. One EPG cannot be a member of multiple bridge domains, but multiple EPGs can be members of a single bridge domain.

All the endpoints that belong to the same EPG can communicate with each other. However, endpoints in different EPGs cannot communicate by default, but they can communicate if a contract exists between the two EPGs allowing that communication. Contracts can be equated to ACLs in traditional networking. However, it differs from an ACL in the way that it doesn't involve specifying specific IP addresses as source and destination and that contracts are applied to an EPG as a whole.

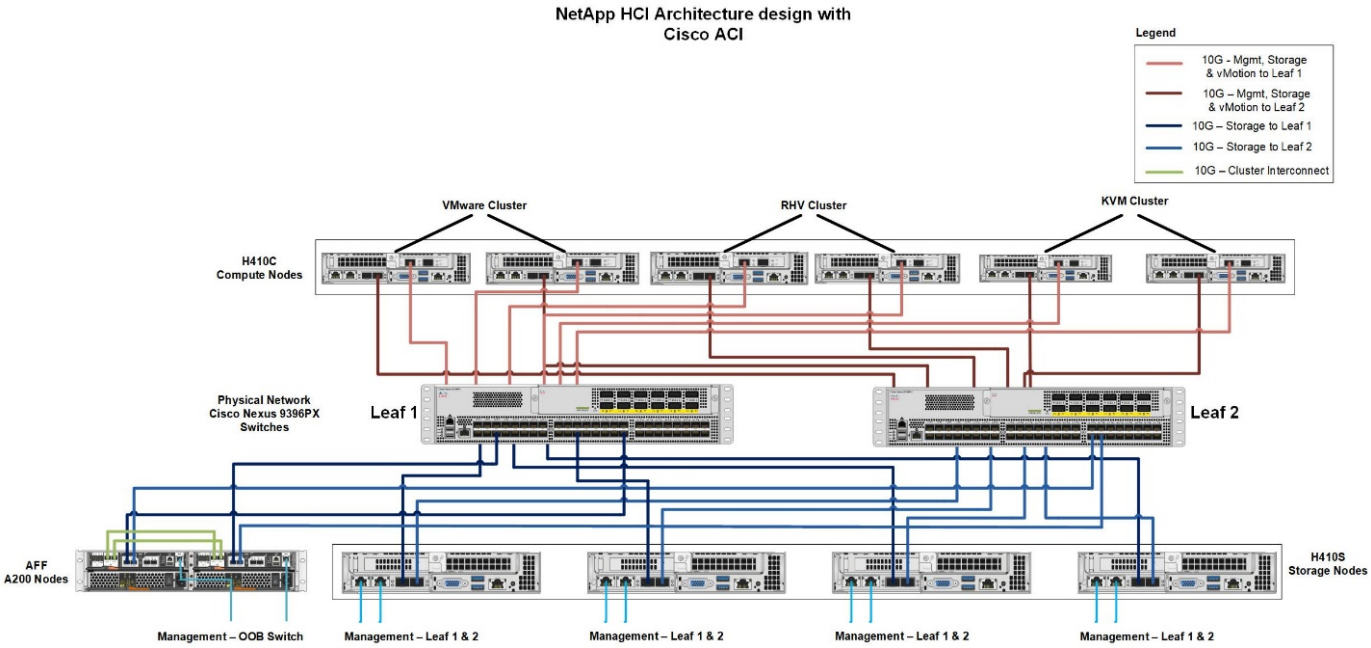
See the Cisco ACI documentation for more information.

Networking Advantages

Cisco ACI provides many advantages over traditional networking. Programmability and automation are critical features of a scalable data center virtualization infrastructure and the policy driven mechanism of Cisco ACI opens a lot of opportunities for providing optimal physical and virtual networking.

- **Virtual Machine Manager (VMM) Integration.** With the Cisco ACI open REST API features, integration with virtualized environments is easy. Cisco ACI supports VMM integration with multiple hypervisors and provides automated access and control over the hypervisor virtual switches to the networking constructs in ACI. VMM integration in ACI seamlessly extends the ACI policy framework to virtual workloads. In other words, VMM integration allows Cisco ACI to control the virtual switches running on virtualization hosts and to extend the ACI fabric access policies to virtual workloads. The integration also automates the hypervisor's virtual switch deployment and configuration tasks. Cisco ACI VMM integration provides the following benefits:
  - Single point of policy management for physical and virtual environments through APIC
  - Faster application deployment, with transparent instantiation of applications in virtual environments
  - Full integrated visibility into the health of the application through holistic aggregation of information across physical and virtual environments
  - Simplified networking configuration for virtual workloads because the port-group or VM NIC profiles required to attach to the VMs are created automatically. For more information on Cisco ACI VMM integration, see the Cisco documentation. In addition, see the Cisco ACI virtualization compatibility matrix for version compatibility details.
- **Micro-segmentation.** Micro-segmentation in Cisco ACI allows you to classify the endpoints in existing application EPGs into microsegment (uSeg) EPGs using network-based or VM-based attributes. This helps for filtering the endpoints more granularly and apply specific dynamic policies on those endpoints. Micro-segmentation can be applied to any endpoints within the tenant. Cisco supports micro-segmentation on a variety of virtual switches - Cisco ACI Virtual Edge, VMware VDS and Microsoft vSwitch. uSeg EPGs can be configured with multiple attributes but an endpoint can be assigned to only one EPG. For more details, see the Cisco ACI Virtualization guide for the specific version.
- **Intra-EPG Isolation.** By default, all endpoints belonging to the same EPG can communicate with each other. Intra-EPG Isolation in Cisco ACI is a feature to prevent endpoints in the same EPG communicate with each other. It achieves isolation by using different VLANs for traffic from ACI leaf to hypervisor hosts and from hypervisor hosts to ACI leaf. Intra-EPG isolation can be enforced on both application EPGs and microsegment EPGs. See the specific version of the Cisco ACI virtualization guide for more information.

Architectural Diagram



This diagram represents the physical architecture of NetApp HCI with Cisco ACI that was designed for this solution. Two leaf switches connected via spines and managed by a cluster of three APICs forms the ACI fabric. The leaf switches are connected to upstream routers for external connectivity. Three pairs of NetApp HCI compute nodes (each pair dedicated for a hypervisor) are configured with a two-cable option. Four storage nodes were configured with four-cable option to form the Element cluster. A pair of AFF A200 nodes are used to provide the ONTAP capabilities to the system.

Hardware and Software Requirements

Compute

The following tables list the hardware and software compute resources utilized in the solution. The components that are used in any implementation of the solution might vary based on customer requirements.

Hardware	Model	Quantity
NetApp HCI compute nodes	NetApp H410C	6

Software	Purpose	Version
VMware ESXi	Virtualization	6.7
VMware vCenter Server Appliance	Virtualization management	6.7
Red Hat Enterprise Linux	Operating system	7.7
KVM	Virtualization	1.5.3-167
Red Hat Virtualization	Virtualization	4.3.9

Storage

The following tables list the hardware and software storage resources used in this solution. The components that are used in any particular implementation of the solution might vary based on customer requirements.

Hardware	Model	Quantity
NetApp HCI storage nodes	NetApp H410S	4
AFF	A200	2

Software	Purpose	Version
NetApp HCI	Infrastructure	1.8
NetApp Element	Storage	12.0
ONTAP	Storage	9.7P6
ONTAP Select	Storage	9.7
Storage Grid	Storage	11.3

Networking

The following tables list the hardware and software network resources used in this solution. The components that are used in any particular implementation of the solution might vary based on customer requirements.

Hardware	Model	Quantity
Cisco UCS server	UCS C-220 M3	3
Cisco Nexus	N9K-C9336-PQ	2
Cisco Nexus	N9K-C9396-PX	2

Software	Purpose	Version
Cisco APIC	Network Management	3.2(9h)
Cisco Nexus ACI-mode Switch	Network	13.2(9h)
Cisco AVE	Network	1.2.9
Open vSwitch (OVS)	Network	2.9.2
VMware Virtual Distributed Switch	Network	6.6

Next: Design Considerations

Design Considerations

Network Design

The minimum configuration of a Cisco ACI fabric consists of two leaf switches and two spine switches with a cluster at least three APICs managing and controlling the whole fabric. All the workloads connect to leaf switches. Spine switches are the backbone of the network and are responsible for interconnecting all leaf switches. No two leaf switches can be interconnected. Each leaf switch is connected to each of the spine switches in a full-mesh topology.

With this two-tier spine-and-leaf architecture, no matter which leaf switch the server is connected to, it's traffic always crosses the same number of devices to get to another server attached to the fabric (unless the other server is located on the same leaf). This approach keeps latency at a predictable level.

Compute Design

The minimum number of compute nodes required for a highly available infrastructure using NetApp HCI is two. NetApp HCI provides two options for cabling: two-cable and six-cable. NetApp HCI H410C compute nodes are available with two 1GbE ports (ports A and B) and four 10/25GbE ports (ports C, D, E, and F) on board. For a two-cable option, ports D and E are used for connectivity to uplink switches, and, for a six-cable option, all ports from A to F are used. Each node also has an additional out-of-band management port that supports Intelligent Platform Management Interface (IPMI) functionality. This solution utilizes the two-cable option for compute nodes.

For VMware deployments, NetApp HCI comes with an automated deployment tool called the NetApp Deployment Engine (NDE). For non-VMware deployments, manual installation of hypervisors or operating systems is required on the compute nodes.

Storage Design

NetApp HCI uses four-cable option for storage nodes. NetApp HCI H410S storage nodes are available with two 1GbE ports (ports A and B) and two 10/25GbE ports (ports C and D) on board. The two 1GbE ports are bundled as Bond1G (active/passive mode) used for management traffic and the two 10/25GbE ports are bundled as Bond10G (LACP active mode) used for storage data traffic.

For non-VMware deployments, the minimum configuration of NetApp HCI storage cluster is four nodes. For NetApp HCI versions earlier than 1.8 with VMware deployments, the minimum configuration is four storage nodes. However, for HCI version 1.8 with VMware deployments, the minimum configuration for NetApp HCI storage cluster is two nodes. For more information on NetApp HCI two-node storage cluster, see the documentation here.

Next: VMware vSphere: NetApp HCI with Cisco ACI

Deploying NetApp HCI with Cisco ACI

VMware vSphere: NetApp HCI with Cisco ACI

VMware vSphere is an industry-leading virtualization platform that provides a way to build a resilient and reliable virtual infrastructure. vSphere contains virtualization, management, and interface layers. The two core components of VMware vSphere are ESXi server and the vCenter Server. VMware ESXi is hypervisor software installed on a physical machine that facilitates hosting of VMs and virtual appliances. vCenter Server is the service through which you manage multiple ESXi hosts connected in a network and pool host resources. For more information on VMware vSphere, see the documentation here.

Workflow

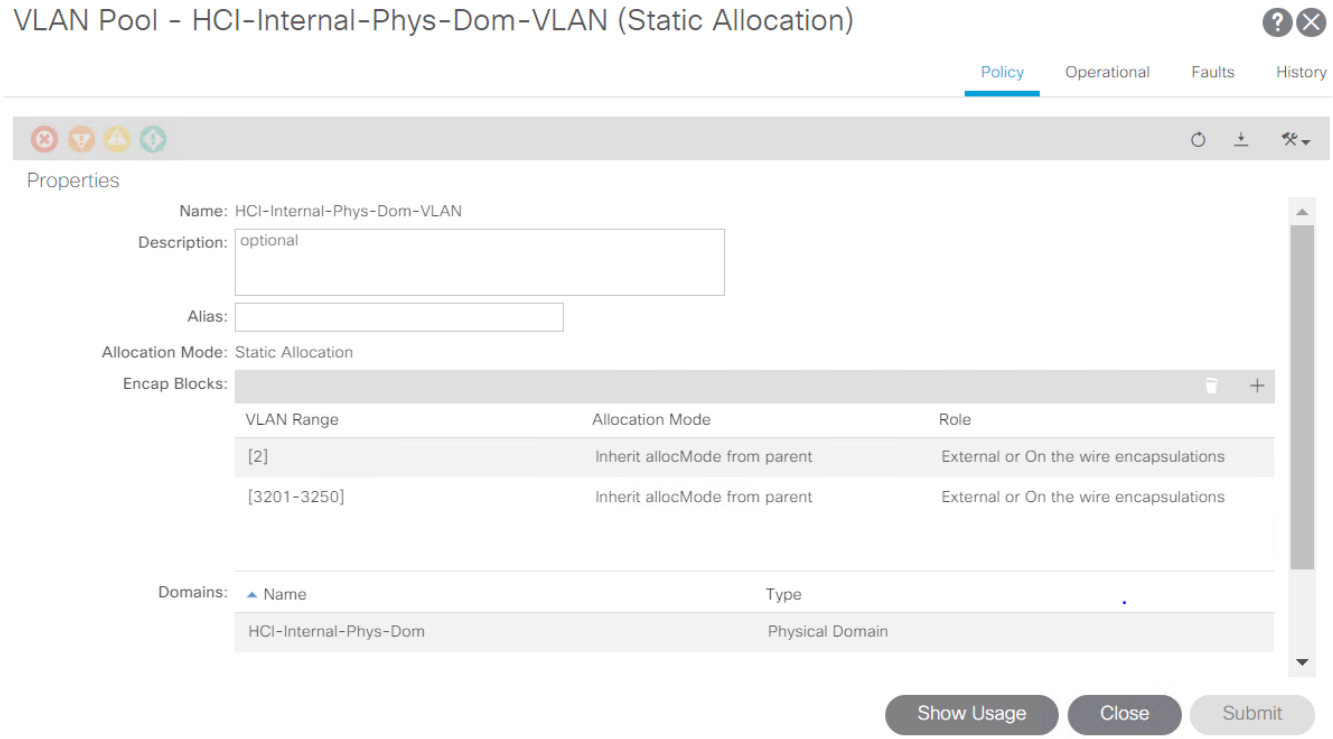
The following workflow was used to up the virtual environment. Each of these steps might involve several individual tasks.

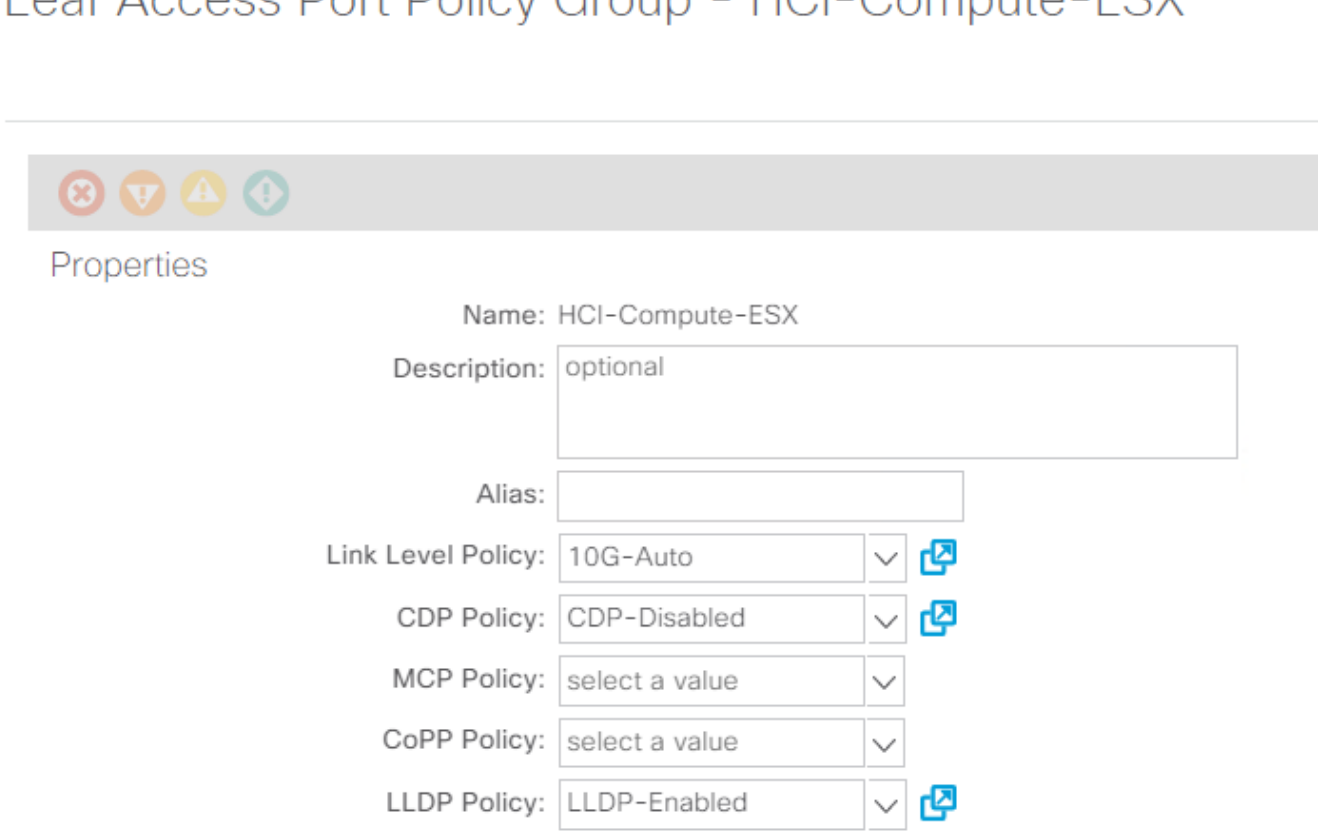
1. Install and configure Nexus 9000 switches in ACI mode and APIC software on the UCS C-series server. See the Install and Upgrade documentation for detailed steps.
2. Configure and setup ACI fabric by referring to the documentation.
3. Configure the tenants, application profiles, bridge domains, and EPGs required for NetApp HCI nodes. NetApp recommends using one BD to one EPG framework, except for iSCSI. See the documentation here for more details. The minimum set of EPGs required are in-band management, iSCSI, iSCSI-A, iSCSI-B, VM motion, VM-data network, and native.

iSCSI multipathing requires two iSCSI EPGs: iSCSI-A and iSCSI-B, each with one active uplink.

NetApp mNode requires an iSCSI EPG with both uplinks active.

4. Create the VLAN pool, physical domain, and AEP based on the requirements. Create the switch and interface profiles for individual ports. Then attach the physical domain and configure the static paths to the EPGs. See the configuration guide for more details.



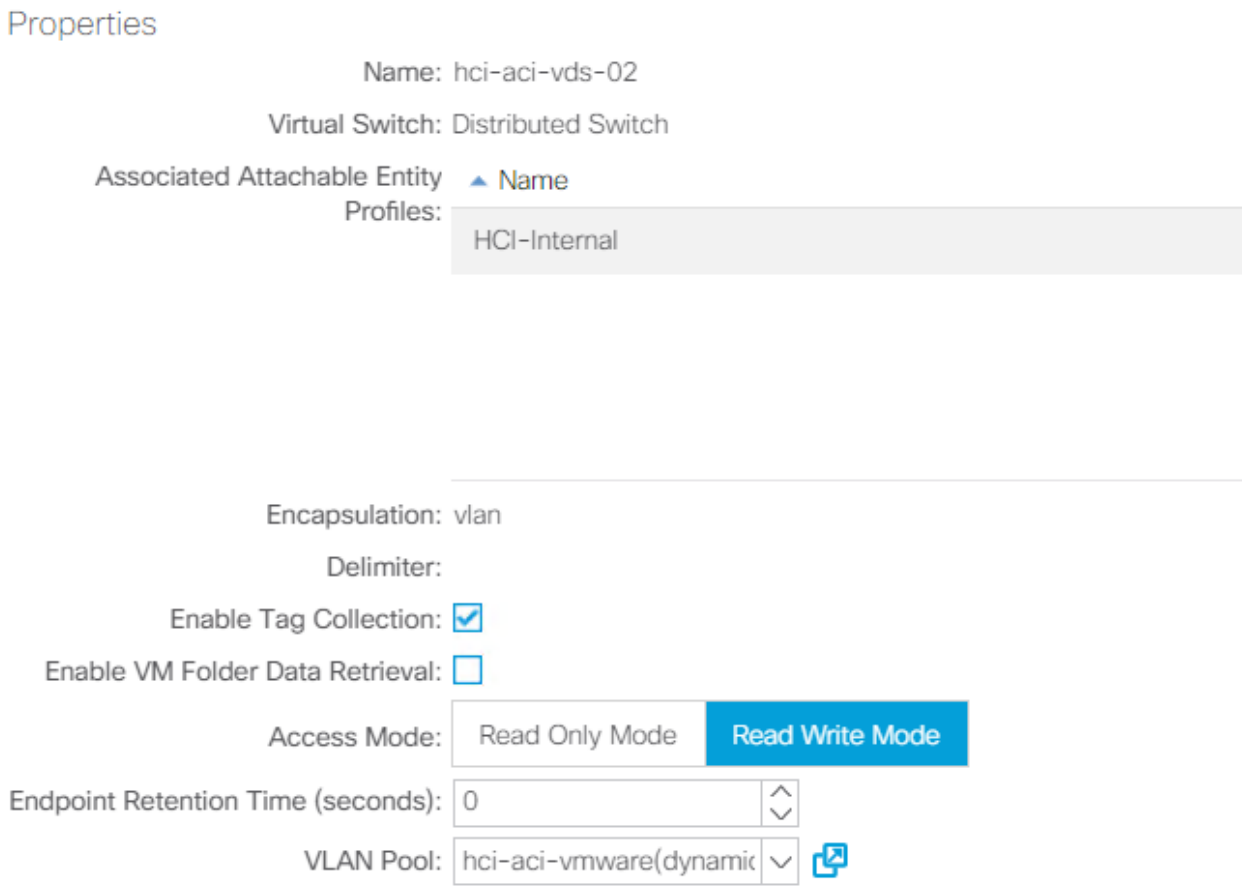


Use an access port policy group for interfaces connecting to NetApp HCI compute nodes, and use vPC policy group for interfaces to NetApp HCI storage nodes.

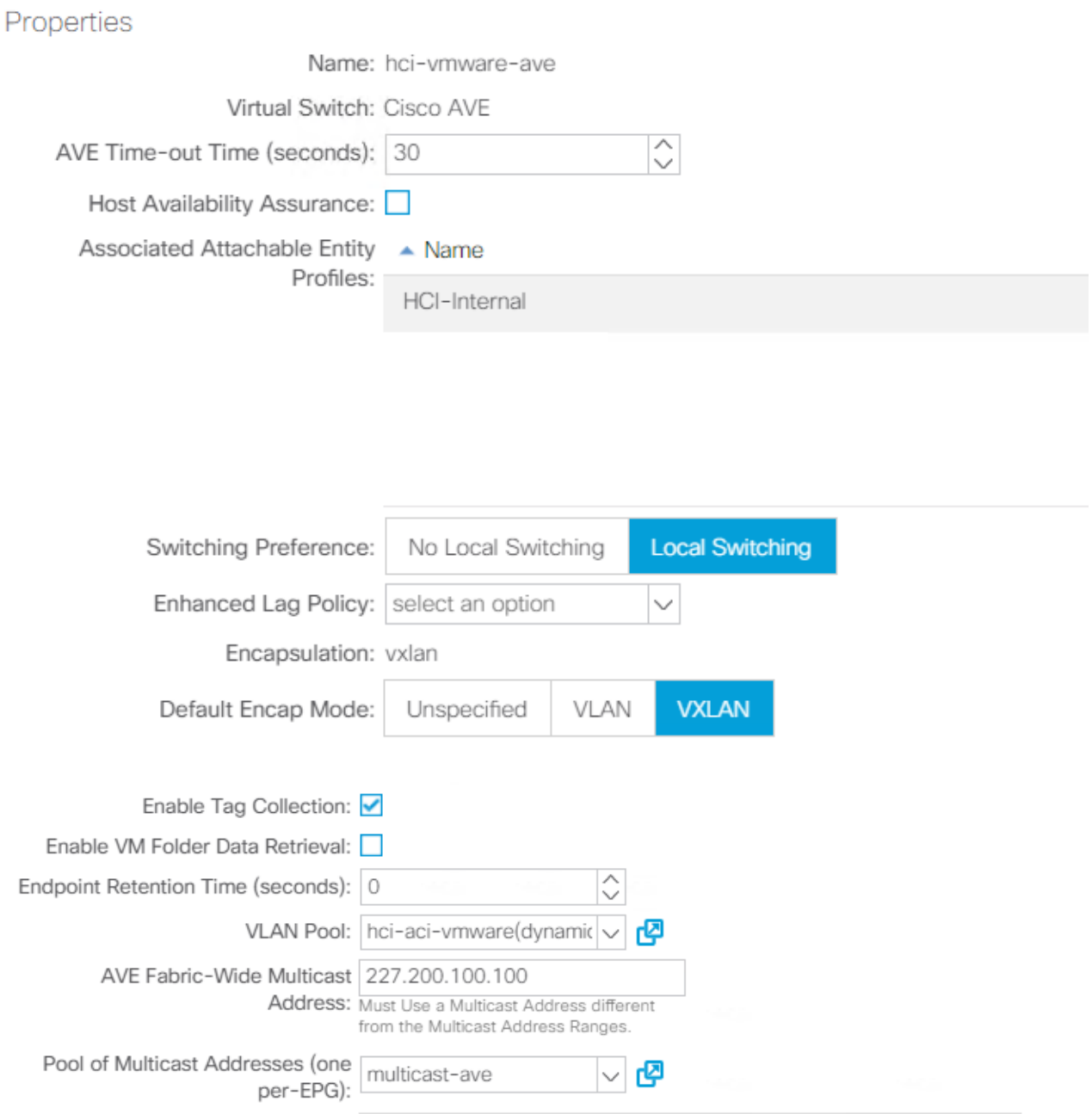
5. Create and assign contracts for tightly-controlled access between workloads. For more information

on configuring the contracts, see the guide here.

6. Install and configure NetApp HCI using NDE. NDE configures all the required parameters, including VDS port groups for networking, and also installs the mNode VM. See the deployment guide for more information.
7. Though VMM integration of Cisco ACI with VMware VDS is optional, using the VMM integration feature is a best practice. When not using VMM integration, an NDE-installed VDS can be used for networking with physical domain attachment on Cisco ACI.
8. If you are using VMM integration, NDE-installed VDS cannot be fully managed by ACI and can be added as read-only VMM domain. To avoid that scenario and make efficient use of Cisco ACI's VMM networking feature, create a new VMware VMM domain in ACI with an explicit dynamic VLAN pool. The VMM domain created can integrate with any supported virtual switch.
    a. **Integrate with VDS.** If you wish to integrate ACI with VDS, select the virtual switch type to be VMware Distributed Switch. Consider the configuration best practices noted in the following table. See the configuration guide for more details.

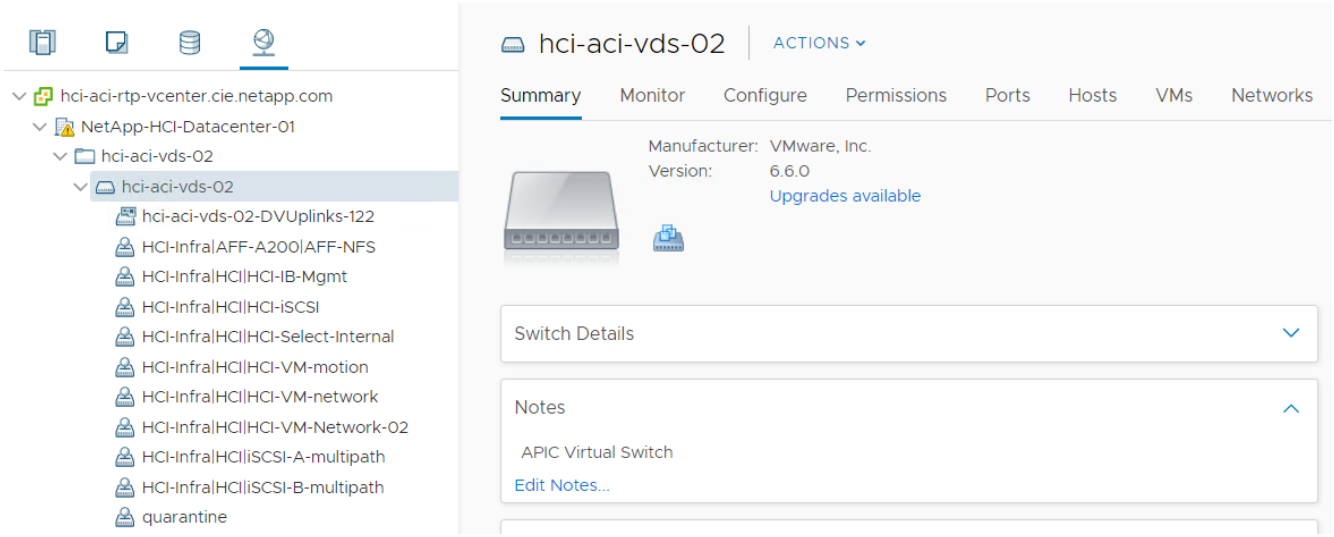


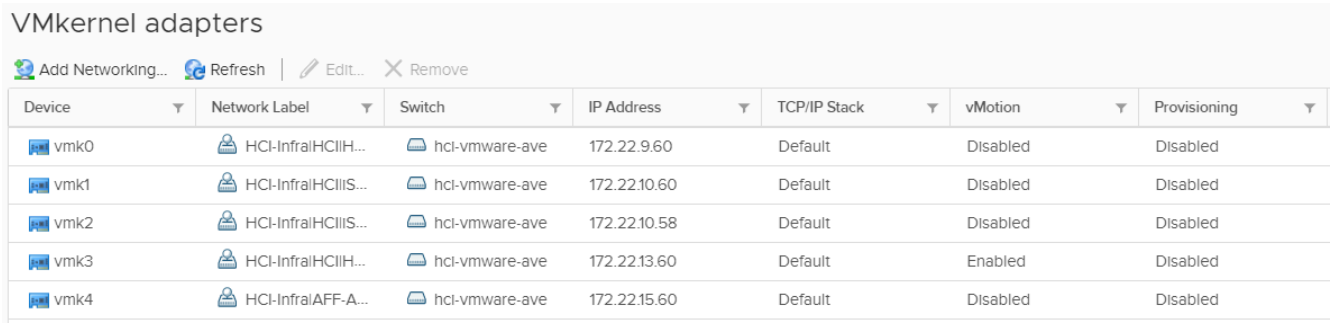
    b. **Integrate with Cisco AVE.** If you are integrating Cisco AVE with Cisco ACI, select the virtual switch type to be Cisco AVE. Cisco AVE requires a unique VLAN pool of type Internal for communicating between internal and external port groups. Follow the configuration best practices noted in this table. See the installation guide to install and configure Cisco AVE.



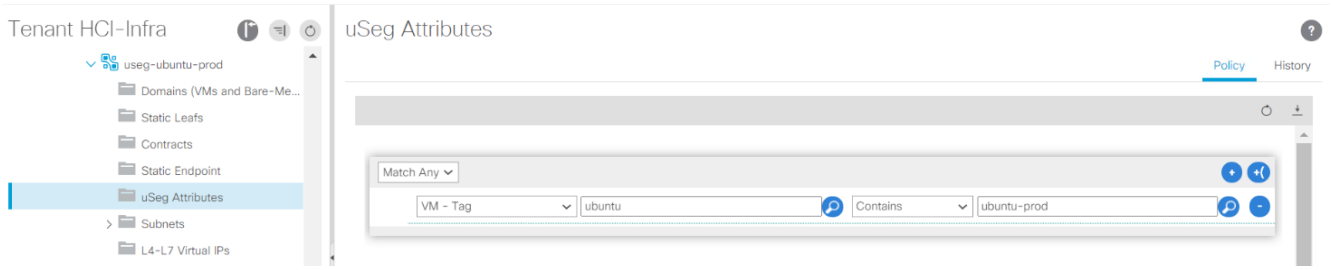
9. Attach the VMM domain to the EPGs using Pre-Provision Resolution Immediacy. Then migrate all the VMNICs, VMkernel ports, and VNICs from the NDE-created VDS to ACI-created VDS or AVE and so on. Configure the uplink failover and teaming policy for iSCSI-A and iSCSI-B to have one active uplink each. VMs can now attach their VMNICs to ACI-created port groups to access network resources. The port groups on VDS that are managed by Cisco ACI are in the format of `<tenant-name>|<application-profile-name>|<epg-name>`.

Pre-Provision Resolution Immediacy is required to ensure the port policies are downloaded to the leaf switch even before the VMM controller is attached to the virtual switch.





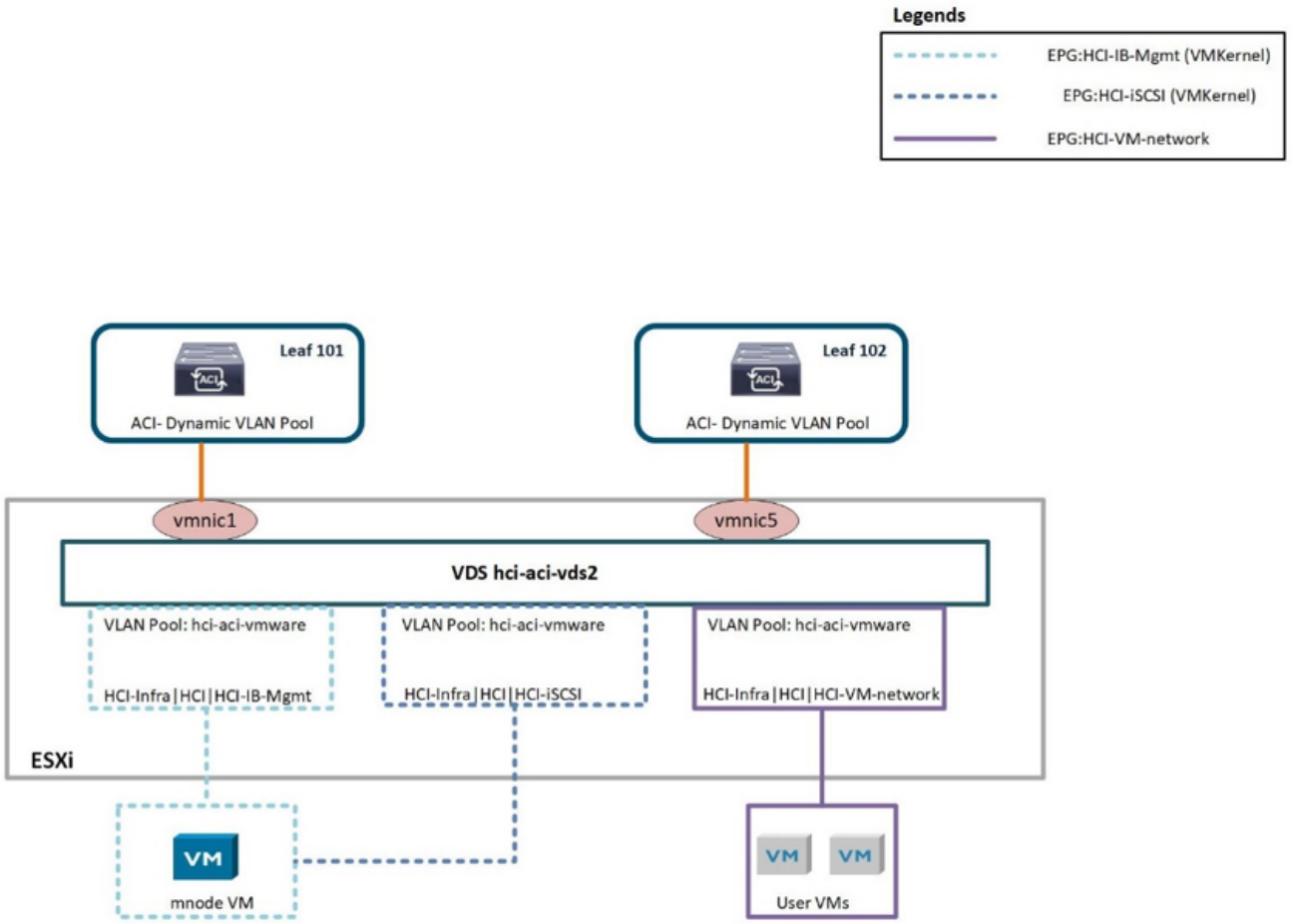
10. If you intend to use micro-segmentation, then create micro-segment (uSeg) EPGs attaching to the right BD. Create attributes in VMware vSphere and attach them to the required VMs. Ensure the VMM domain has Enable Tag Collection enabled. Configure the uSeg EPGs with the corresponding attribute and attach the VMM domain to it. This provides more granular control of communication on the endpoint VMs.



The networking functionality for VMware vSphere on NetApp HCI in this solution is provided either using VMware VDS or Cisco AVE.

VMware VDS

VMware vSphere Distributed Switch (VDS) is a virtual switch that connects to multiple ESXi hosts in the cluster or set of clusters allowing virtual machines to maintain consistent network configuration as they migrate across multiple hosts. VDS also provides for centralized management of network configurations in a vSphere environment. For more details, see the VDS documentation.



The following table outlines the necessary parameters and best practices for configuring and integrating Cisco ACI with VMware VDS.

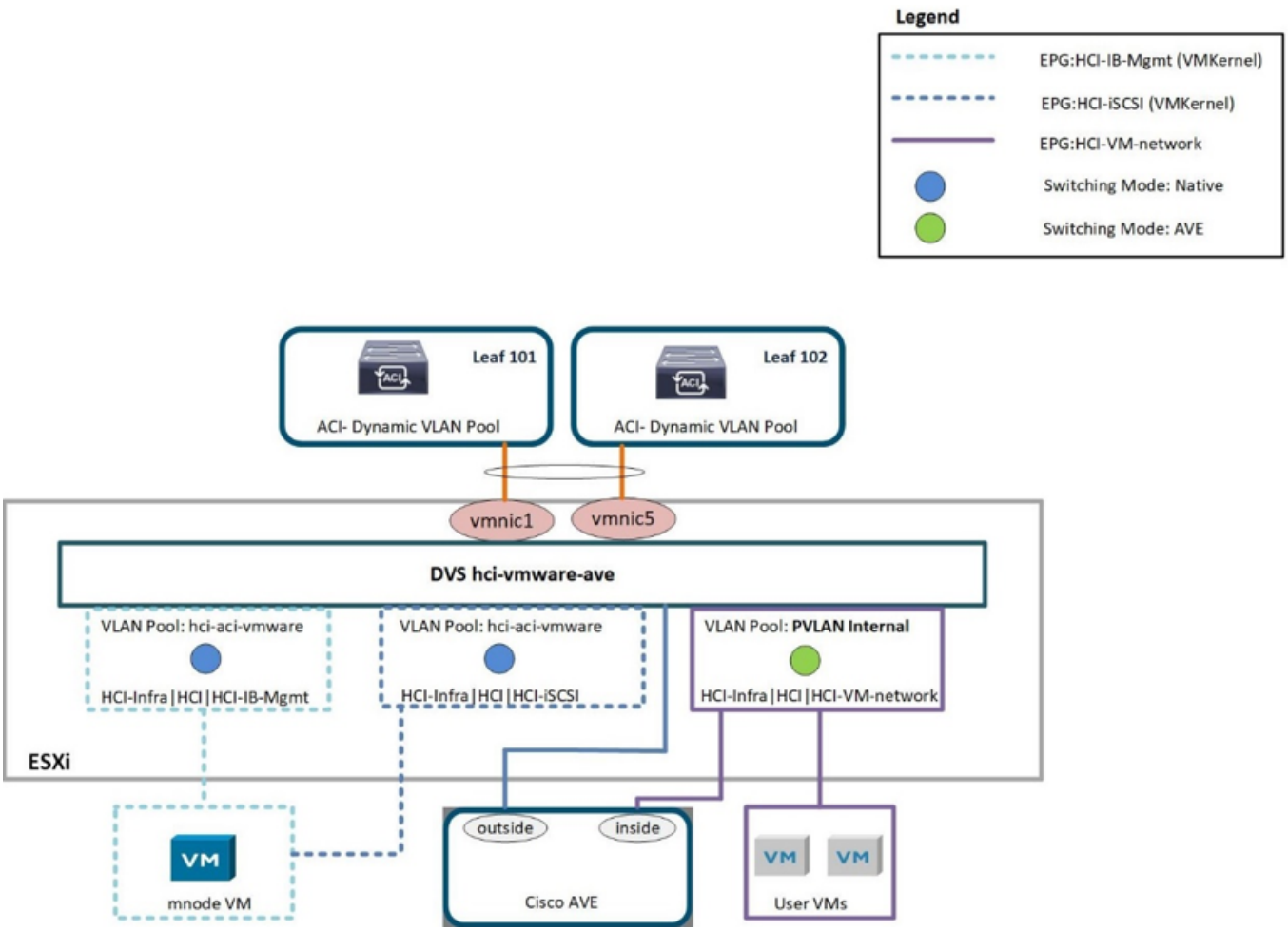
Resource	Configuration Considerations	Best Practices
Endpoint groups	• Separate EPG for native VLANs • Static binding of interfaces to HCI storage and compute nodes in native VLAN EPG uses 802.1P mode. This is required for node discovery to run NDE. • Separate EPGs for iSCSI, iSCSI-A, and iSCSI-B with a common BD • iSCSI-A and iSCSI-B are for iSCSI multipathing and are used for VMkernel ports on ESXi hosts • Physical domain to be attached to iSCSI EPG before running NDE • VMM domain to be attached to iSCSI, iSCSI-A, and iSCSI-B EPGs	• Contracts between EPGs to be well defined. Allow only required ports for communication. • Use unique native VLAN for NDE node discovery • For EPGs corresponding to port-groups being attached to VMkernel ports, VMM domain to be attached with Pre-Provision for Resolution Immediacy
Interface policy	• A common leaf access port policy group for all ESXi hosts • One vPC policy group per NetApp HCI storage node • LLDP enabled, CDP disabled	• Separate VLAN pool for VMM domain with dynamic allocation turned on • Recommended to use vPC with LACP Active port-channel policy for interfaces towards NetApp HCI Storage Nodes • Recommended to use individual interfaces for Compute Nodes, No LACP.

Resource	Configuration Considerations	Best Practices
VMM Integration	• Local switching preference • Access mode is Read Write.	• MAC-Pinning-Physical-NIC-Load for vSwitch policy • LLDP for discovery policy • Enable Tag collection if micro-segmentation is used
VDS	• Both uplinks active for iSCSI port-group • One uplink each for iSCSI-A and iSCSI-B	• Load balancing method for all port-groups to be 'Route based on physical NIC load' • iSCSI VMkernel port migration to be done one at a time from NDE deployed VDS to ACI integrated VDS

For traffic load-balancing, port channels with vPCs can be used on Cisco ACI along with LAGs on VDS with LACP in active mode. However, using LACP can affect storage performance when compared to iSCSI multipathing.

Cisco AVE

Cisco ACI Virtual Edge (AVE) is a virtual switch offering by Cisco that extends the Cisco ACI policy model to virtual infrastructure. It is a hypervisor- independent distributed network service that sits on top of the native virtual switch of the hypervisor. It leverages the underlying virtual switch using a VM-based solution to provide network visibility into the virtual environments. For more details on Cisco AVE, see the documentation. The following figure depicts the internal networking of Cisco AVE on an ESXi host (as tested).



The following table lists the necessary parameters and best practices for configuring and integrating Cisco ACI with Cisco AVE on VMware ESXi. Cisco AVE is currently only supported with VMware vSphere.

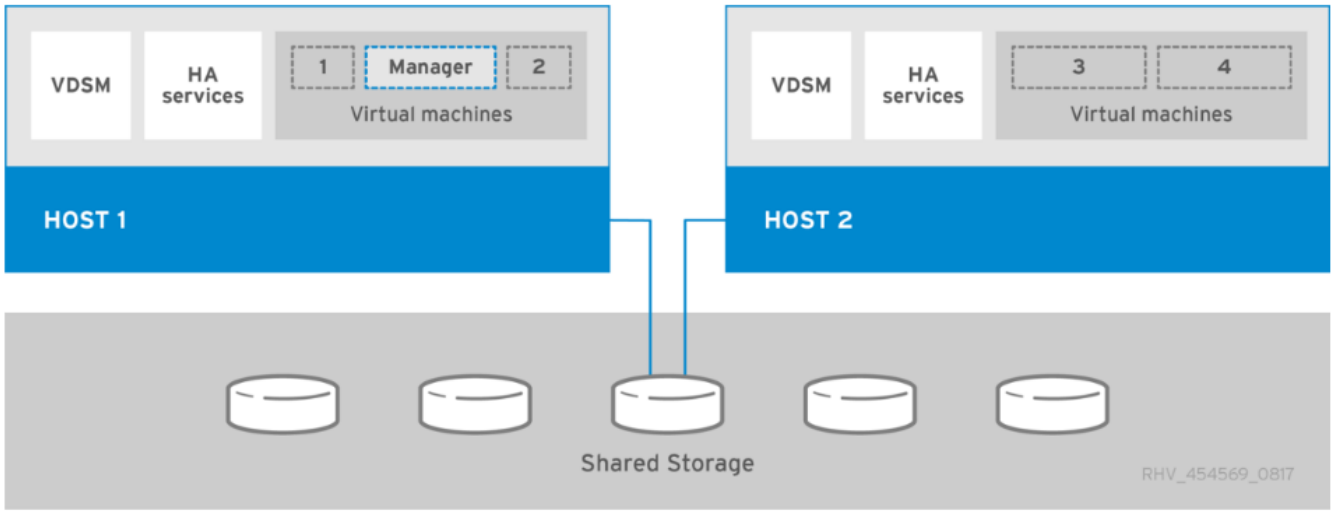
Resource	Configuration Considerations	Best Practices
Endpoint Groups	Separate EPG for native VLANs Static binding of interfaces towards HCI storage and compute nodes in native VLAN EPG uses 802.1P mode. This is required for node discovery to run NDE. Separate EPGs for iSCSI, iSCSI-A and iSCSI-B with a common BD iSCSI-A and iSCSI-B are for iSCSI multipathing and are used for VMkernel ports on ESXi hosts Physical domain to be attached to iSCSI EPG before running NDE VMM domain is attached to iSCSI, iSCSI-A, and iSCSI-B EPGs	Separate VLAN pool for VMM domain with dynamic allocation turned on Contracts between EPGs to be well defined. Allow only required ports for communication. Use unique native VLAN for NDE node discovery Use native switching mode in VMM domain for EPGs that correspond to port groups being attached to host's VMkernel adapters Use AVE switching mode in VMM domain for EPGs corresponding to port groups carrying user VM traffic For EPGs corresponding to port-groups being attached to VMkernel ports, VMM domain is attached with Pre-Provision for Resolution Immediacy
Interface Policy	• One vPC policy group per ESXi host • One vPC policy group per NetApp HCI storage node • LLDP enabled, CDP disabled	• NetApp recommends using vPCs to ESXi hosts • Use static mode on port-channel policy for vPCs to ESXi • Use Layer-4 SRC port load balancing hashing method for port-channel policy • NetApp recommends using vPC with LACP active port-channel policy for interfaces to NetApp HCI storage nodes

Resource	Configuration Considerations	Best Practices
VMM Integration	• Create a new VLAN range [or Encap Block] with role Internal and Dynamic allocation' attached to the VLAN pool intended for VMM domain • Create a pool of multicast addresses (one address per EPG) • Reserve another multicast address different from the pool of multicast addresses intended for AVE fabric-wide multicast address • Local switching preference • Access mode to be Read Write mode	• Static mode on for vSwitch policy • Ensure that vSwitch port-channel policy and interface policy group's port-channel policy are using the same mode • LLDP for discovery policy • Enable Tag collection if using micro-segmentation • Recommended option for Default Encap mode is VXLAN
VDS	• - Both uplinks active for iSCSI port-group • - One uplink each for iSCSI-A and iSCSI-B	• iSCSI VMkernel port migration is done one at a time from NDE deployed VDS to ACI integrated VDS • Load balancing method for all port-groups to be Route based on IP hash

For traffic load balancing, port channel with vPCs can be used on Cisco ACI along with LAGs on ESXi hosts with LACP in active mode. However, using LACP can affect storage performance when compared to iSCSI multipathing.

Red Hat Virtualization: NetApp HCI with Cisco ACI

Red Hat Virtualization (RHV) is an enterprise virtual data center platform that runs on Red Hat Enterprise Linux using the KVM hypervisor. The key components of RHV include Red Hat Virtualization Hosts (RHV- H) and the Red Hat Virtualization Manager (RHV- M). RHV-M provides centralized, enterprise-grade management for the physical and logical resources within the virtualized RHV environment. RHV-H is a minimal, light-weight operating system based on Red Hat Enterprise Linux that is optimized for the ease of setting up physical servers as RHV hypervisors. For more information on RHV, see the documentation here. The following figure provides an overview of RHV.

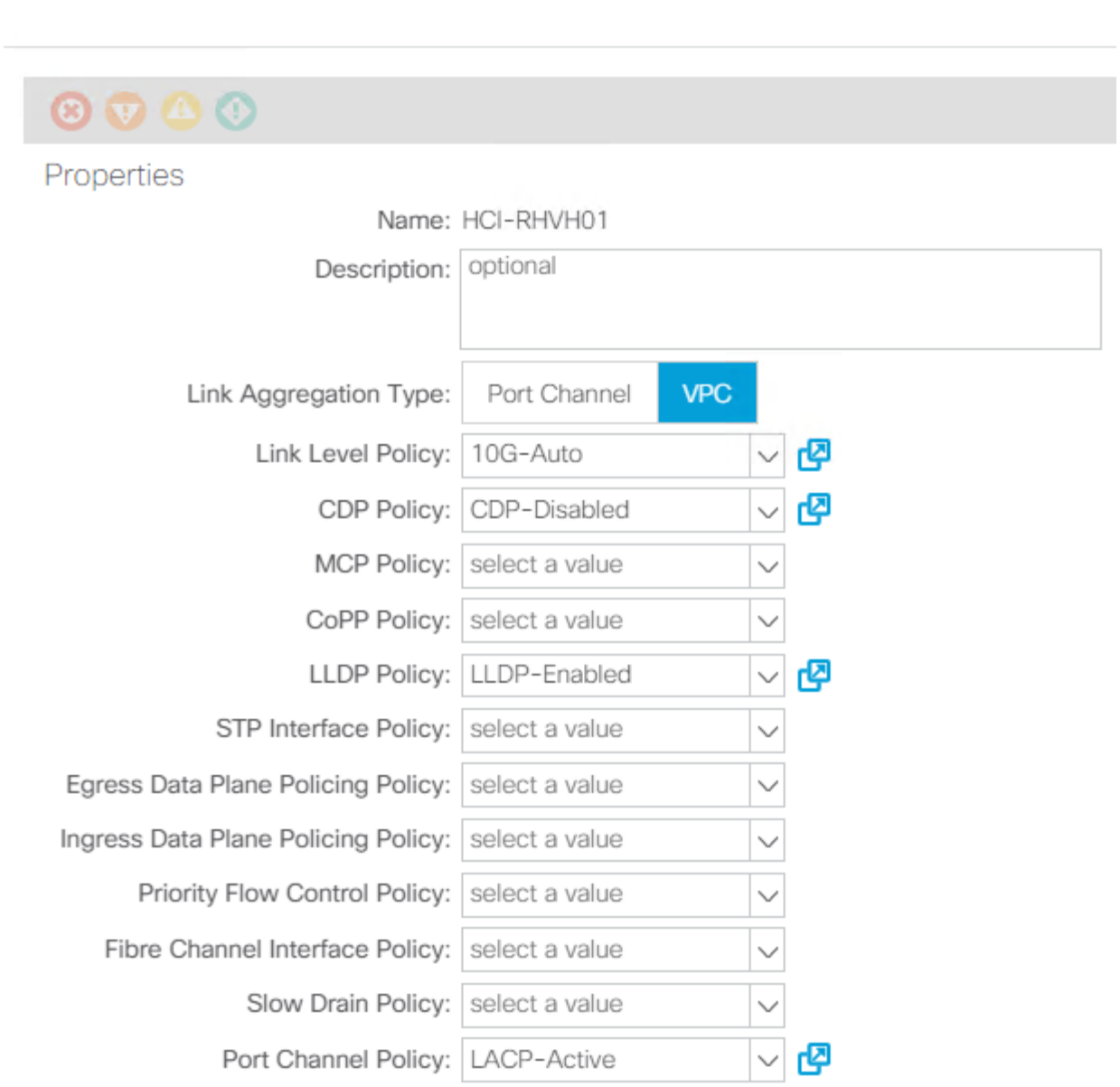


Starting with Cisco APIC release 3.1, Cisco ACI supports VMM integration with Red Hat Virtualization environments. The RHV VMM domain in Cisco APIC is connected to RHV-M and directly associated with a data center object. All the RHV-H clusters under this data center are considered part of the VMM domain. Cisco ACI automatically creates logical networks in RHV- M when the EPGs are attached to the RHV VMM domain in ACI. RHV hosts that are part of a Red Hat VMM domain can use Linux bridge or Open vSwitch as its virtual switch. This integration simplifies and automates networking configuration on RHV-M, saving a lot of manual work for system and network administrators.

Workflow

The following workflow is used to set up the virtual environment. Each of these steps might involve several individual tasks.

1. Install and configure Nexus 9000 switches in ACI mode and APIC software on the UCS C-series server. Refer to the Install and Upgrade documentation for detailed steps.
2. Configure and setup the ACI fabric by referring to the documentation.
3. Configure tenants, application profiles, bridge domains, and EPGs required for NetApp HCI nodes. NetApp recommends using one BD to one EPG framework, except for iSCSI. See the documentation here for more details. The minimum set of EPGs required are in-band management, iSCSI, VM motion, VM-data network, and native.
4. Create the VLAN pool, physical domain, and AEP based on the requirements. Create the switch and interface profiles and policies for vPCs and individual ports. Then attach the physical domain and configure the static paths to the EPGs. see the configuration guide for more details. This table lists best practices for integrating ACI with Linux bridge on RHV.

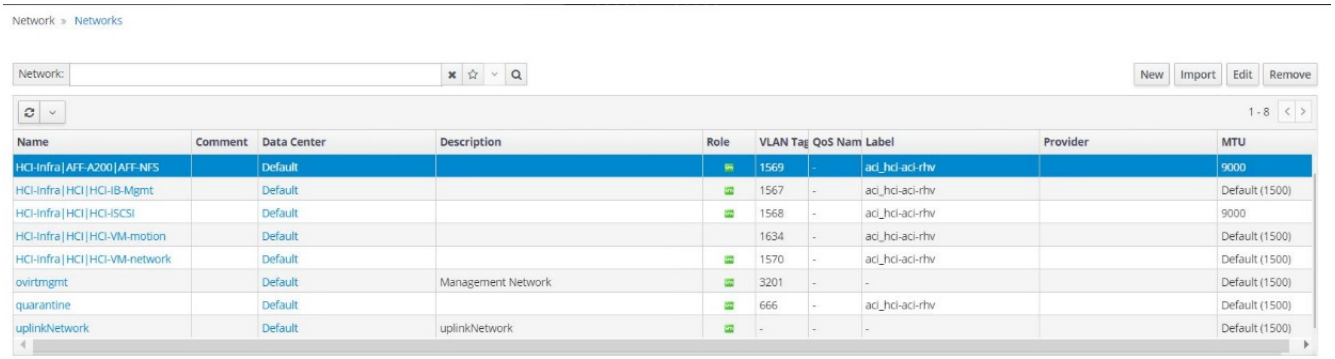


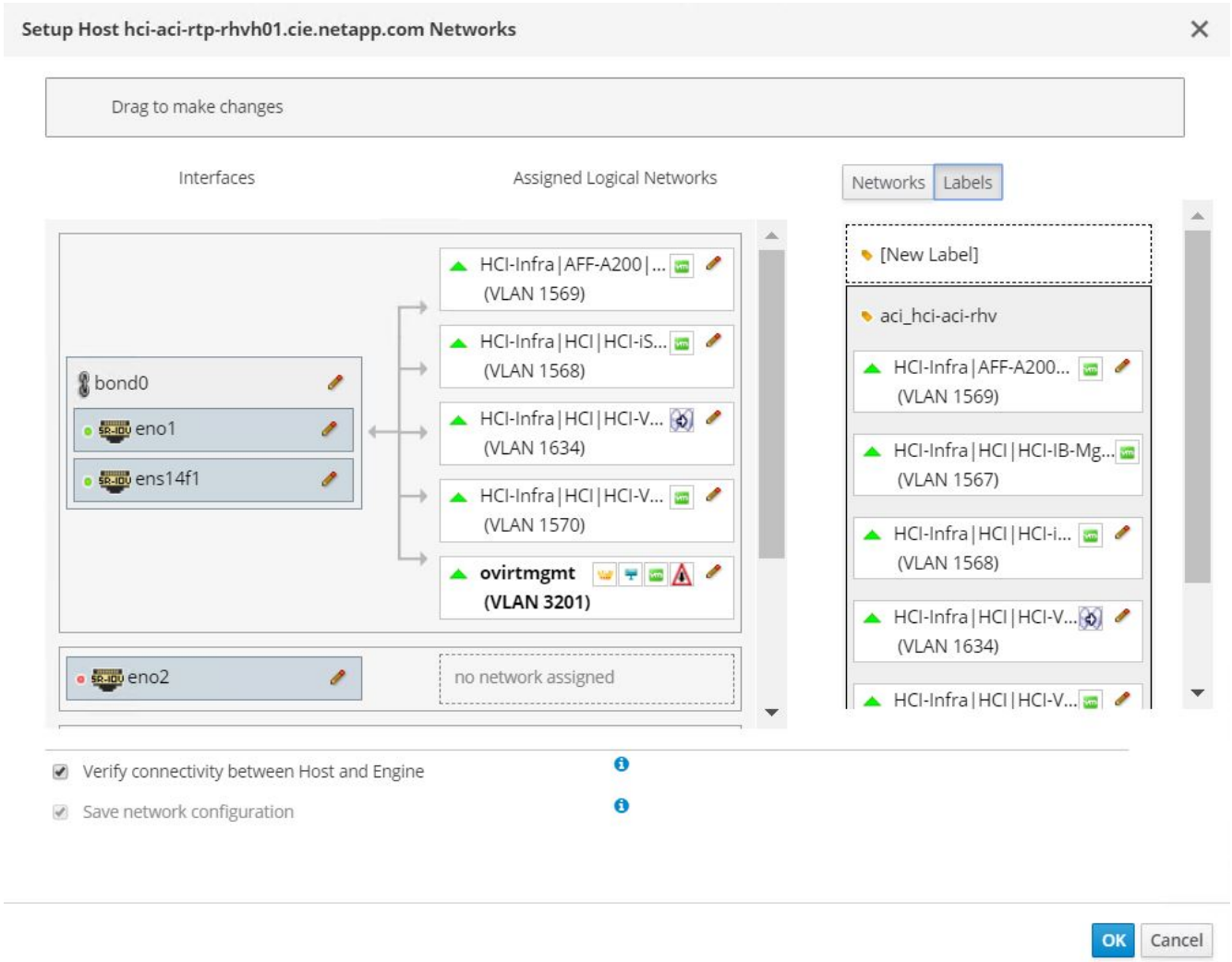
Use a vPC policy group for interfaces connecting to NetApp HCI storage and compute nodes.

5. Create and assign contracts for tightly controlled access between workloads. For more information on configuring the contracts, see the guide here.
6. Install and configure the NetApp HCI Element cluster. Do not use NDE for this install; rather, install a standalone Element cluster on the HCI storage nodes. Then configure the required volumes for installation of RHV. Install RHV on NetApp HCI. Refer to RHV on NetApp HCI NVA for more details.
7. RHV installation creates a default management network called ovirtmgmt. Though VMM integration of Cisco ACI with RHV is optional, leveraging VMM integration is preferred. Do not create other logical networks manually. To use Cisco ACI VMM integration, create a Red Hat VMM

domain and attach the VMM domain to all the required EPGs, using Pre- Provision Resolution Immediacy. This process automatically creates corresponding logical networks and vNIC profiles. The vNIC profiles can be directly used to attach to hosts and VMs for their communication. The networks that are managed by Cisco ACI are in the format `<tenant-name>|<application-profile-name>|<epg-name>` tagged with a label of format `aci_<rhv-vmm-domain-name>`. See Cisco's whitepaper for creating and configuring a VMM domain for RHV. Also, see this table for best practices when integrating RHV on NetApp HCI with Cisco ACI.

Except for ovirtmgmt, all other logical networks can be managed by Cisco ACI.

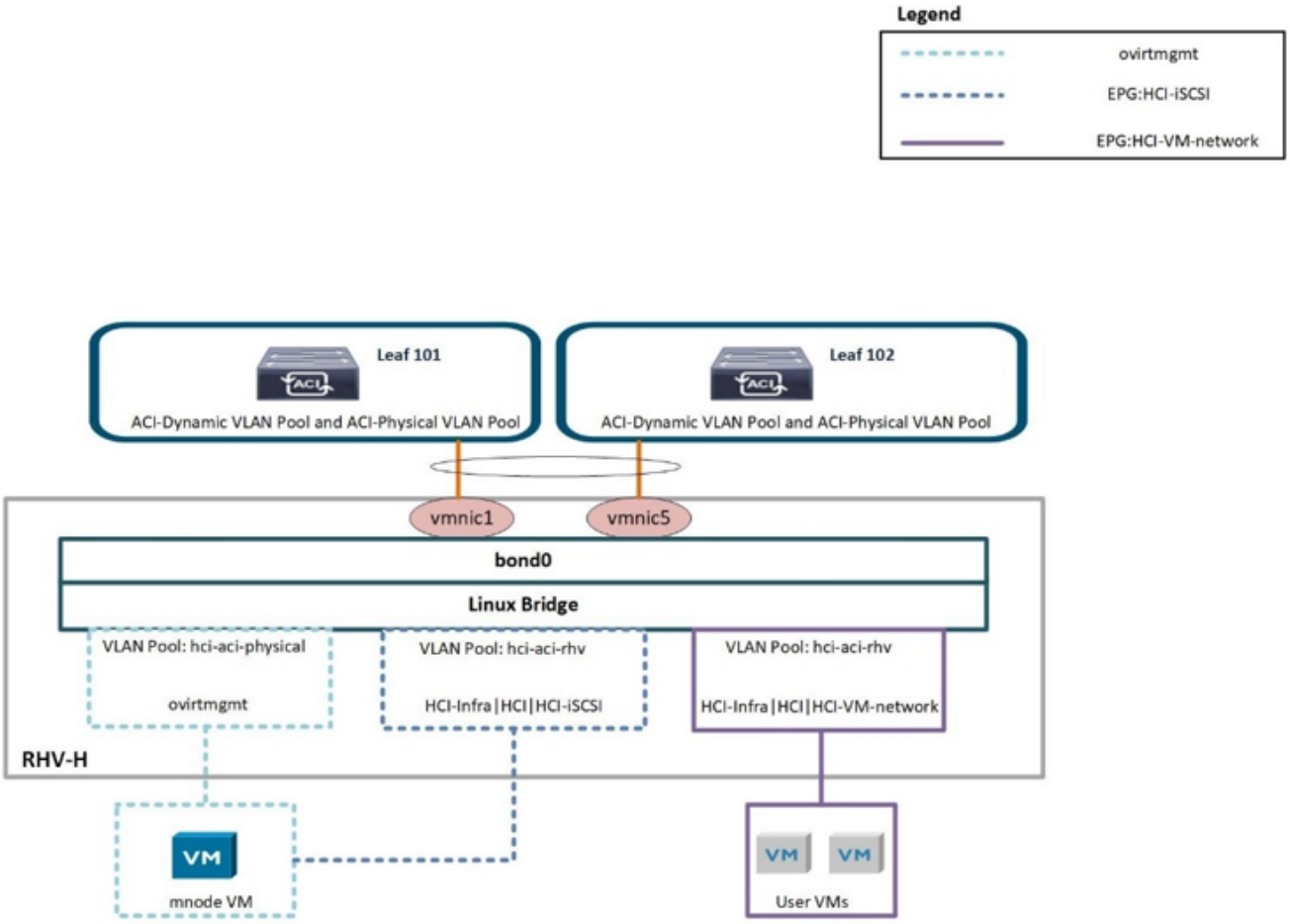




The networking functionality for RHVH hosts in this solution is provided by Linux bridge.

Linux Bridge

Linux Bridge is a default virtual switch on all Linux distributions that is usually used with KVM/QEMU-based hypervisors. It is designated to forward traffic between networks based on MAC addresses and thus is regarded as a layer-2 virtual switch. For more information, see the documentation here. The following figure depicts the internal networking of Linux Bridge on RHV-H (as tested).



The following table outlines the necessary parameters and best practices for configuring and integrating Cisco ACI with Linux Bridge on RHV hosts.

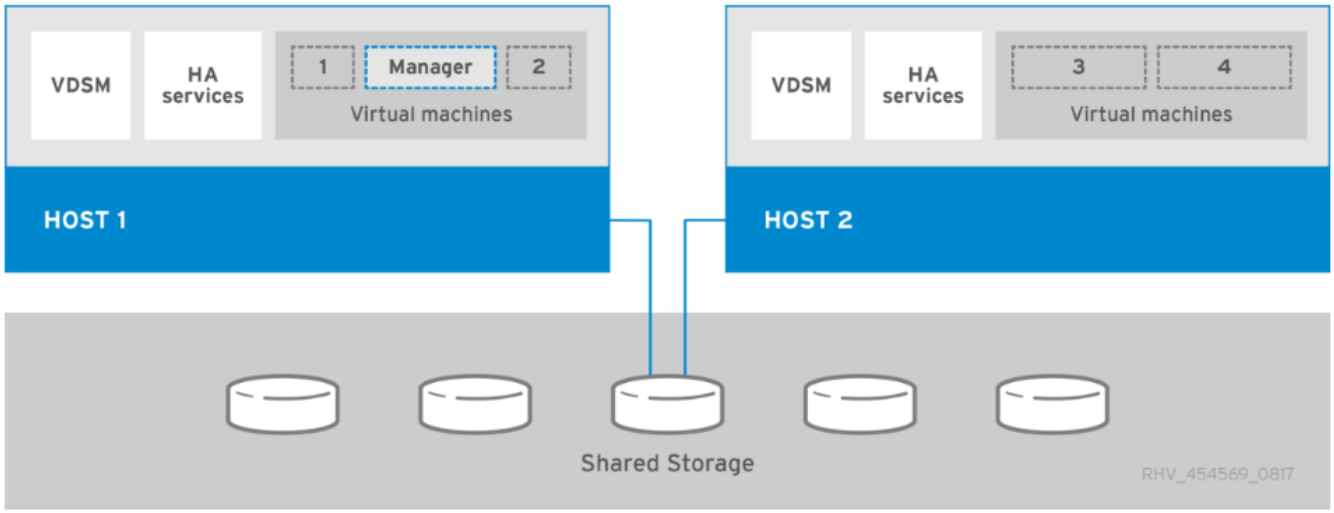
Resource	Configuration considerations	Best Practices
Endpoint groups	• Separate EPG for native VLAN • Static binding of interfaces towards HCI storage and compute nodes in native VLAN EPG to be on 802.1P mode • Static binding of vPCs required on In-band management EPG and iSCSI EPG before RHV installation	• Separate VLAN pool for VMM domain with dynamic allocation turned on • Contracts between EPGs to be well defined. Allow only required ports for communication. • Use unique native VLAN for discovery during Element cluster formation • For EPGs corresponding to port-groups being attached to VMkernel ports, VMM domain to be attached with 'Pre-Provision' for Resolution Immediacy
Interface policy	• One vPC policy group per RHV-H host • One vPC policy group per NetApp HCI storage node • LLDP enabled, CDP disabled	• Recommended to use vPC towards RHV-H hosts • Use 'LACP Active' for the port-channel policy • Use only 'Graceful Convergence' and 'Symmetric Hashing' control bits for port-channel policy Use 'Layer4 Src-port' load balancing hashing method for port-channel policy Recommended to use vPC with LACP Active port-channel policy for interfaces towards NetApp HCI storage nodes
VMM Integration	Do not migrate host management logical interfaces from ovirtmgmt to any other logical network	iSCSI host logical interface to be migrated to iSCSI logical network managed by ACI VMM integration

Except for the ovirtmgmt logical network, it is possible to create all other infrastructure logical networks on Cisco APIC and map them to the VMM domain. 'ovirtmgmt' logical network uses the static path binding on the In-band management EPG attached with the physical domain.

Next: KVM on RHEL: NetApp HCI with Cisco ACI

Red Hat Virtualization: NetApp HCI with Cisco ACI

Red Hat Virtualization (RHV) is an enterprise virtual data center platform that runs on Red Hat Enterprise Linux using the KVM hypervisor. The key components of RHV include Red Hat Virtualization Hosts (RHV-H) and the Red Hat Virtualization Manager (RHV- M). RHV-M provides centralized, enterprise-grade management for the physical and logical resources within the virtualized RHV environment. RHV-H is a minimal, light-weight operating system based on Red Hat Enterprise Linux that is optimized for the ease of setting up physical servers as RHV hypervisors. For more information on RHV, see the documentation here. The following figure provides an overview of RHV.



Starting with Cisco APIC release 3.1, Cisco ACI supports VMM integration with Red Hat Virtualization environments. The RHV VMM domain in Cisco APIC is connected to RHV-M and directly associated with a data center object. All the RHV-H clusters under this data center are considered part of the VMM domain. Cisco ACI automatically creates logical networks in RHV- M when the EPGs are attached to the RHV VMM domain in ACI. RHV hosts that are part of a Red Hat VMM domain can use Linux bridge or Open vSwitch as its virtual switch. This integration simplifies and automates networking configuration on RHV-M, saving a lot of manual work for system and network administrators.

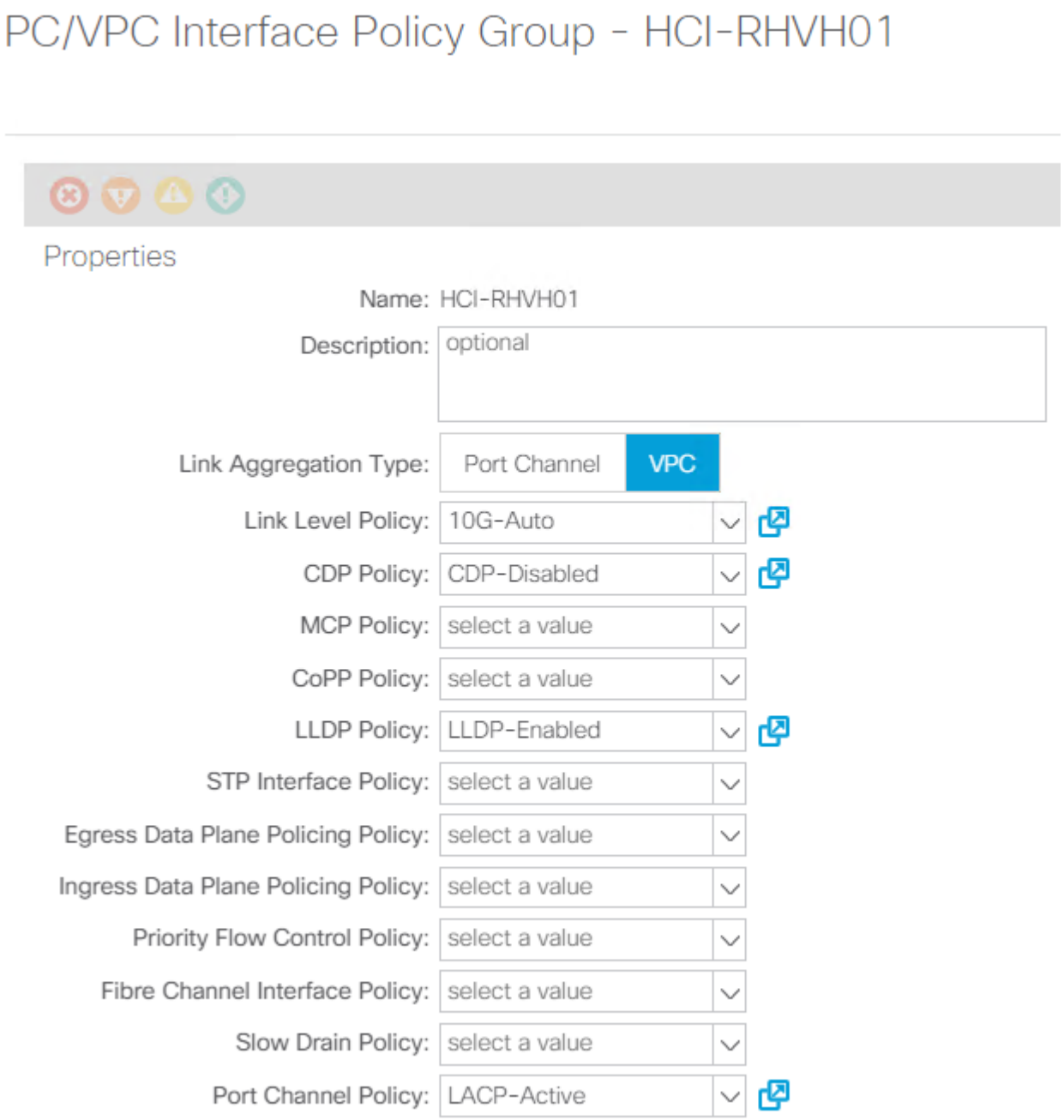
Workflow

The following workflow is used to set up the virtual environment. Each of these steps might involve several individual tasks.

1. Install and configure Nexus 9000 switches in ACI mode and APIC software on the UCS C-series

server. Refer to the Install and Upgrade documentation for detailed steps.

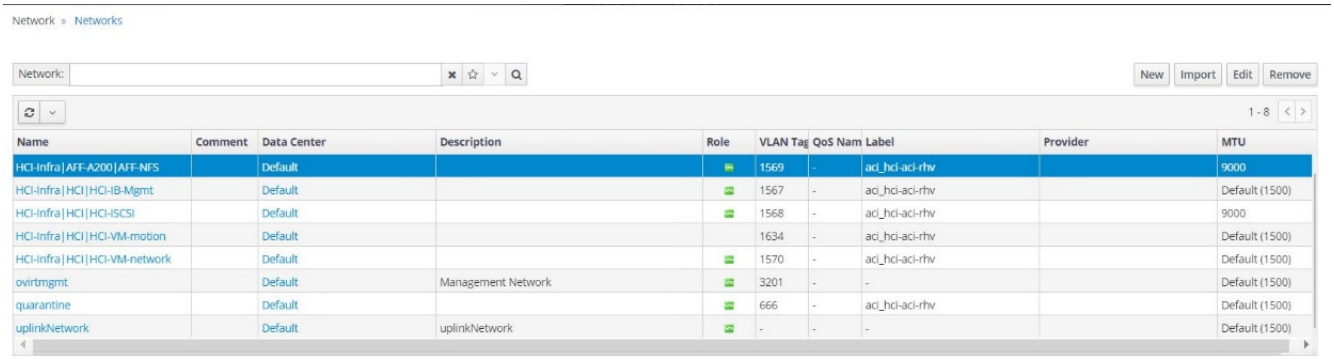
2. Configure and setup the ACI fabric by referring to the documentation.
3. Configure tenants, application profiles, bridge domains, and EPGs required for NetApp HCI nodes. NetApp recommends using one BD to one EPG framework, except for iSCSI. See the documentation here for more details. The minimum set of EPGs required are in-band management, iSCSI, VM motion, VM-data network, and native.
4. Create the VLAN pool, physical domain, and AEP based on the requirements. Create the switch and interface profiles and policies for vPCs and individual ports. Then attach the physical domain and configure the static paths to the EPGs. see the configuration guide for more details. This table lists best practices for integrating ACI with Linux bridge on RHV.

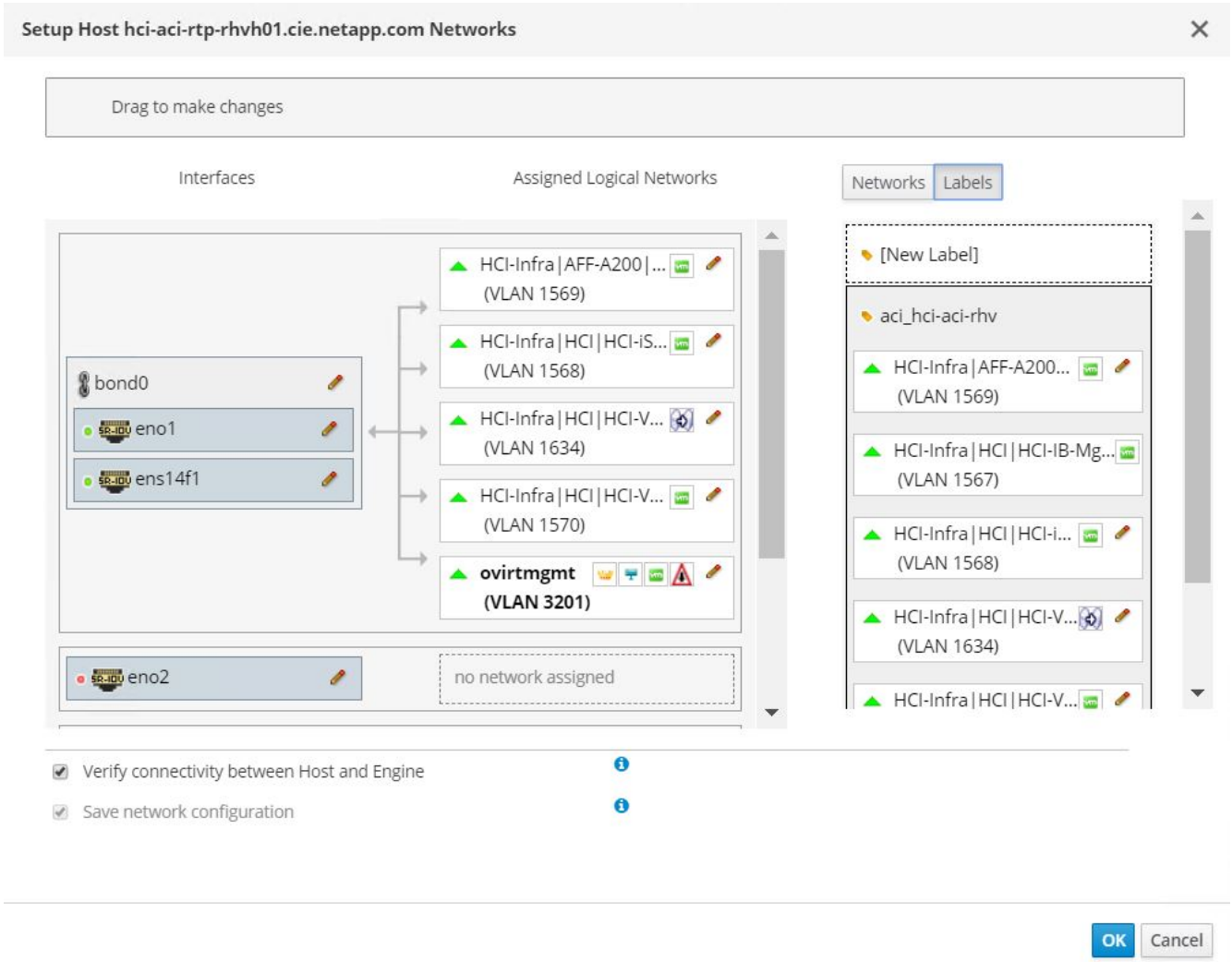


Use a vPC policy group for interfaces connecting to NetApp HCI storage and compute nodes.

5. Create and assign contracts for tightly controlled access between workloads. For more information on configuring the contracts, see the guide here.
6. Install and configure the NetApp HCI Element cluster. Do not use NDE for this install; rather, install a standalone Element cluster on the HCI storage nodes. Then configure the required volumes for installation of RHV. Install RHV on NetApp HCI. Refer to RHV on NetApp HCI NVA for more details.
7. RHV installation creates a default management network called ovirtmgmt. Though VMM integration of Cisco ACI with RHV is optional, leveraging VMM integration is preferred. Do not create other logical networks manually. To use Cisco ACI VMM integration, create a Red Hat VMM domain and attach the VMM domain to all the required EPGs, using Pre- Provision Resolution Immediacy. This process automatically creates corresponding logical networks and vNIC profiles. The vNIC profiles can be directly used to attach to hosts and VMs for their communication. The networks that are managed by Cisco ACI are in the format `<tenant-name>|<application-profile-name>|<epg-name>` tagged with a label of format `aci_<rhv-vmm-domain-name>`. See Cisco's whitepaper for creating and configuring a VMM domain for RHV. Also, see this table for best practices when integrating RHV on NetApp HCI with Cisco ACI.

Except for ovirtmgmt, all other logical networks can be managed by Cisco ACI.

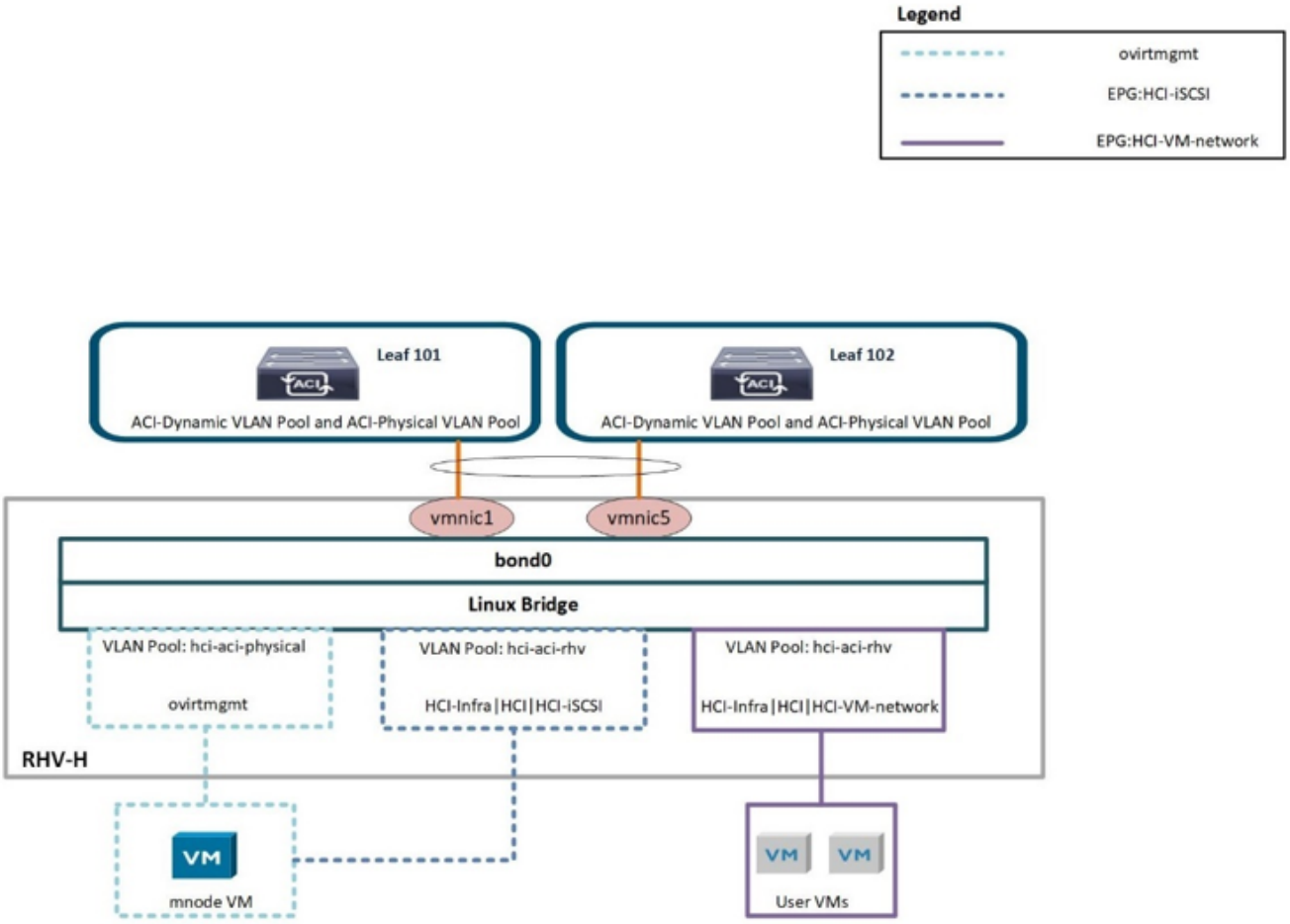




The networking functionality for RHVH hosts in this solution is provided by Linux bridge.

Linux Bridge

Linux Bridge is a default virtual switch on all Linux distributions that is usually used with KVM/QEMU-based hypervisors. It is designated to forward traffic between networks based on MAC addresses and thus is regarded as a layer-2 virtual switch. For more information, see the documentation here. The following figure depicts the internal networking of Linux Bridge on RHV-H (as tested).



The following table outlines the necessary parameters and best practices for configuring and integrating Cisco ACI with Linux Bridge on RHV hosts.

Resource	Configuration considerations	Best Practices
Endpoint groups	• Separate EPG for native VLAN • Static binding of interfaces towards HCI storage and compute nodes in native VLAN EPG to be on 802.1P mode • Static binding of vPCs required on In-band management EPG and iSCSI EPG before RHV installation	• Separate VLAN pool for VMM domain with dynamic allocation turned on • Contracts between EPGs to be well defined. Allow only required ports for communication. • Use unique native VLAN for discovery during Element cluster formation • For EPGs corresponding to port-groups being attached to VMkernel ports, VMM domain to be attached with 'Pre-Provision' for Resolution Immediacy
Interface policy	• One vPC policy group per RHV-H host • One vPC policy group per NetApp HCI storage node • LLDP enabled, CDP disabled	• Recommended to use vPC towards RHV-H hosts • Use 'LACP Active' for the port-channel policy • Use only 'Graceful Convergence' and 'Symmetric Hashing' control bits for port-channel policy Use 'Layer4 Src-port' load balancing hashing method for port-channel policy Recommended to use vPC with LACP Active port-channel policy for interfaces towards NetApp HCI storage nodes
VMM Integration	Do not migrate host management logical interfaces from ovirtmgmt to any other logical network	iSCSI host logical interface to be migrated to iSCSI logical network managed by ACI VMM integration

Except for the ovirtmgmt logical network, it is possible to create all other infrastructure logical networks on Cisco APIC and map them to the VMM domain. 'ovirtmgmt' logical network uses the static path binding on the In-band management EPG attached with the physical domain.

Next: KVM on RHEL: NetApp HCI with Cisco ACI

KVM on RHEL: NetApp HCI with Cisco ACI

KVM (for Kernel-based Virtual Machine) is an open-source full virtualization solution for Linux on x86 hardware such as Intel VT or AMD-V. In other words, KVM lets you turn a Linux machine into a hypervisor that allows the host to run multiple, isolated VMs.

KVM converts any Linux machine into a type-1 (bare-metal) hypervisor. KVM can be implemented on any Linux distribution, but implementing KVM on a supported Linux distribution—like Red Hat Enterprise Linux—expands KVM's capabilities. You can swap resources among guests, share common libraries, and optimize system performance.

Workflow

The following high-level workflow was used to set up the virtual environment. Each of these steps might involve several individual tasks.

1. Install and configure Nexus 9000 switches in ACI mode, and install and configure APIC software on a UCS C-series server. See the Install and Upgrade documentation for detailed steps.
2. Configure and set up the ACI fabric by referring to the documentation.
3. Configure the tenants, application profiles, bridge domains, and EPGs required for NetApp HCI nodes. NetApp recommends using a one-BD-to-one-EPG framework except for iSCSI. See the documentation here for more details. The minimum set of EPGs required are in-band management, iSCSI, VM Motion, VM-data network, and native.
4. Create the VLAN pool, physical domain, and AEP based on the requirements. Create the switch and interface profiles and policies for vPCs and individual ports. Then attach the physical domain and configure the static paths to the EPGs. See the configuration guide for more details. Also see this table <link> for best practices for integrating ACI with Open vSwitch on the RHEL–KVM hypervisor.

Use a vPC policy group for interfaces connecting to NetApp HCI storage and compute nodes.

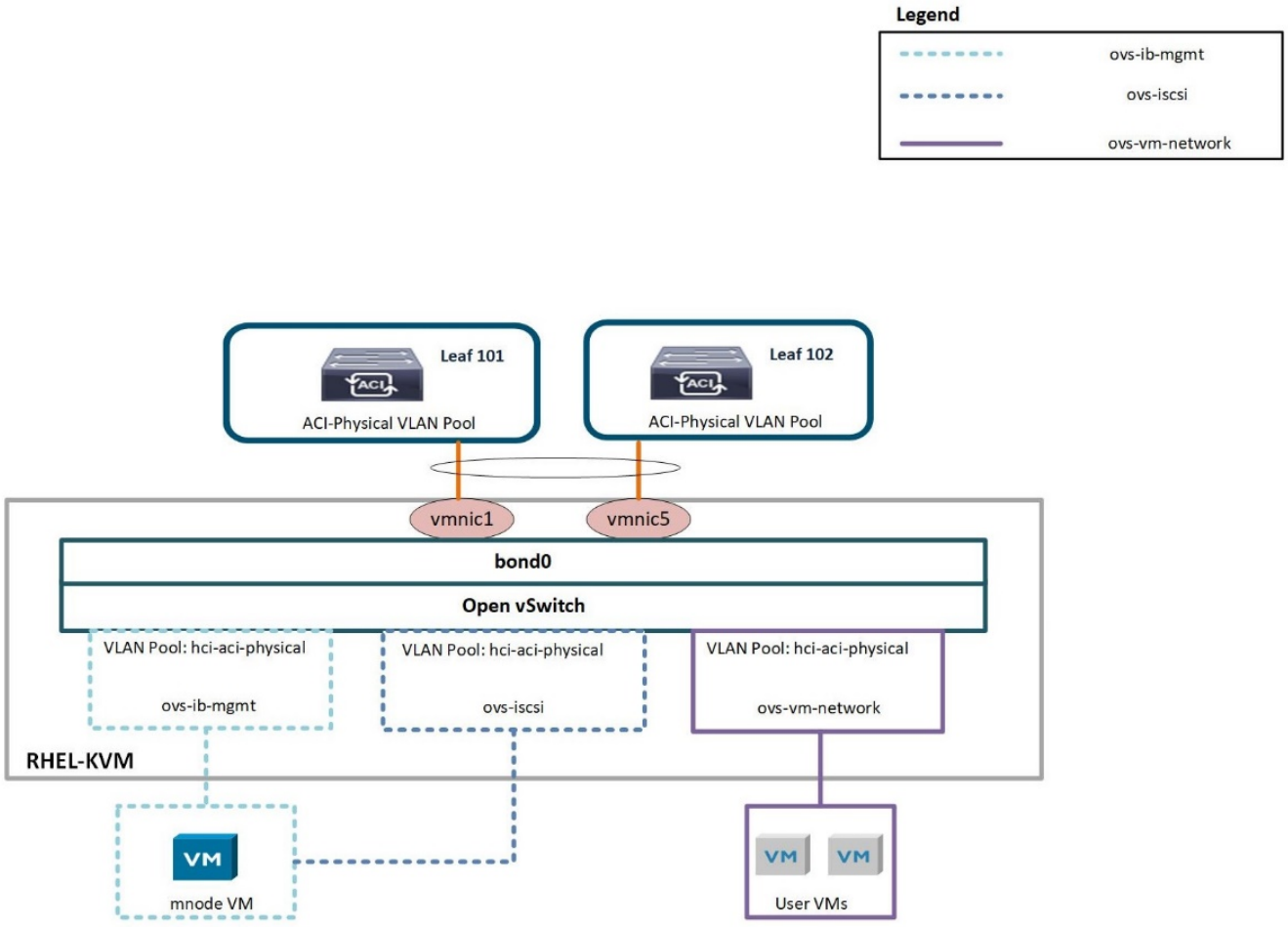
5. Create and assign contracts for tightly-controlled access between workloads. For more details on configuring the contracts, see the guide here.
6. Install and configure a NetApp HCI Element cluster. Do not use NDE for this installation; rather, install a standalone Element cluster on HCI storage nodes. Then configure the required volumes for

the installation of RHEL. Install RHEL, KVM, and Open vSwitch on the NetApp HCI compute nodes. Configure storage pools on the hypervisor using Element volumes for a shared storage service for hosts and VMs. For more details on installation and configuration of KVM on RHEL, see the Red Hat documentation. See the OVS documentation for details on configuring Open vSwitch.

7. RHEL KVM hypervisor's Open vSwitch cannot be VMM integrated with Cisco ACI. Physical domain and static paths must be configured on all required EPGs to allow the required VLANs on the interfaces connecting the ACI leaf switches and RHEL hosts. Also configure the corresponding OVS bridges on RHEL hosts and configure VMs to use those bridges. The networking functionality for the RHEL KVM hosts in this solution is achieved using Open vSwitch virtual switch.

Open vSwitch

Open vSwitch is an open-source, enterprise-grade virtual switch platform. It uses virtual network bridges and flow rules to forward packets between hosts. Programming flow rules work differently in OVS than in the standard Linux Bridge. The OVS plugin does not use VLANs to tag traffic. Instead, it programs flow rules on the virtual switches that dictate how traffic should be manipulated before forwarded to the exit interface. Flow rules determine how inbound and outbound traffic should be treated. The following figure depicts the internal networking of Open vSwitch on an RHEL-based KVM host.



The following table outlines the necessary parameters and best practices for configuring Cisco ACI and Open vSwitch on RHEL based KVM hosts.

Resource	Configuration Considerations	Best Practices
Endpoint groups	• Separate EPG for native VLAN • Static binding of interfaces towards HCI storage and compute nodes in native VLAN EPG to be on 802.1P mode • Static binding of vPCs required on in-band management EPG and iSCSI EPG before KVM installation	• Separate VLAN pool for physical domain with static allocation turned on • Contracts between EPGs to be well defined. Allow only required ports for communication. • Use unique native VLAN for discovery during Element cluster formation
Interface Policy	- One vPC policy group per RHEL host - One vPC policy group per NetApp HCI storage node - LLDP enabled, CDP disabled	• NetApp recommends using vPC towards RHV-H hosts • Use LACP Active for the port-channel policy • Use only Graceful Convergence and Symmetric Hashing control bits for port-channel policy • Use Layer4 Src-Port load-balancing hashing method for port-channel policy • NetApp recommends using vPC with LACP Active port-channel policy for interfaces towards NetApp HCI storage nodes

Next: ONTAP on AFF: NetApp HCI and Cisco ACI

ONTAP on AFF: NetApp HCI and Cisco ACI

NetApp AFF is a robust storage platform that provides low-latency performance, integrated data protection, multiprotocol support, and nondisruptive operations. Powered by NetApp ONTAP data management software, NetApp AFF ensures

nondisruptive operations, from maintenance to upgrades to complete replacement of your storage system.

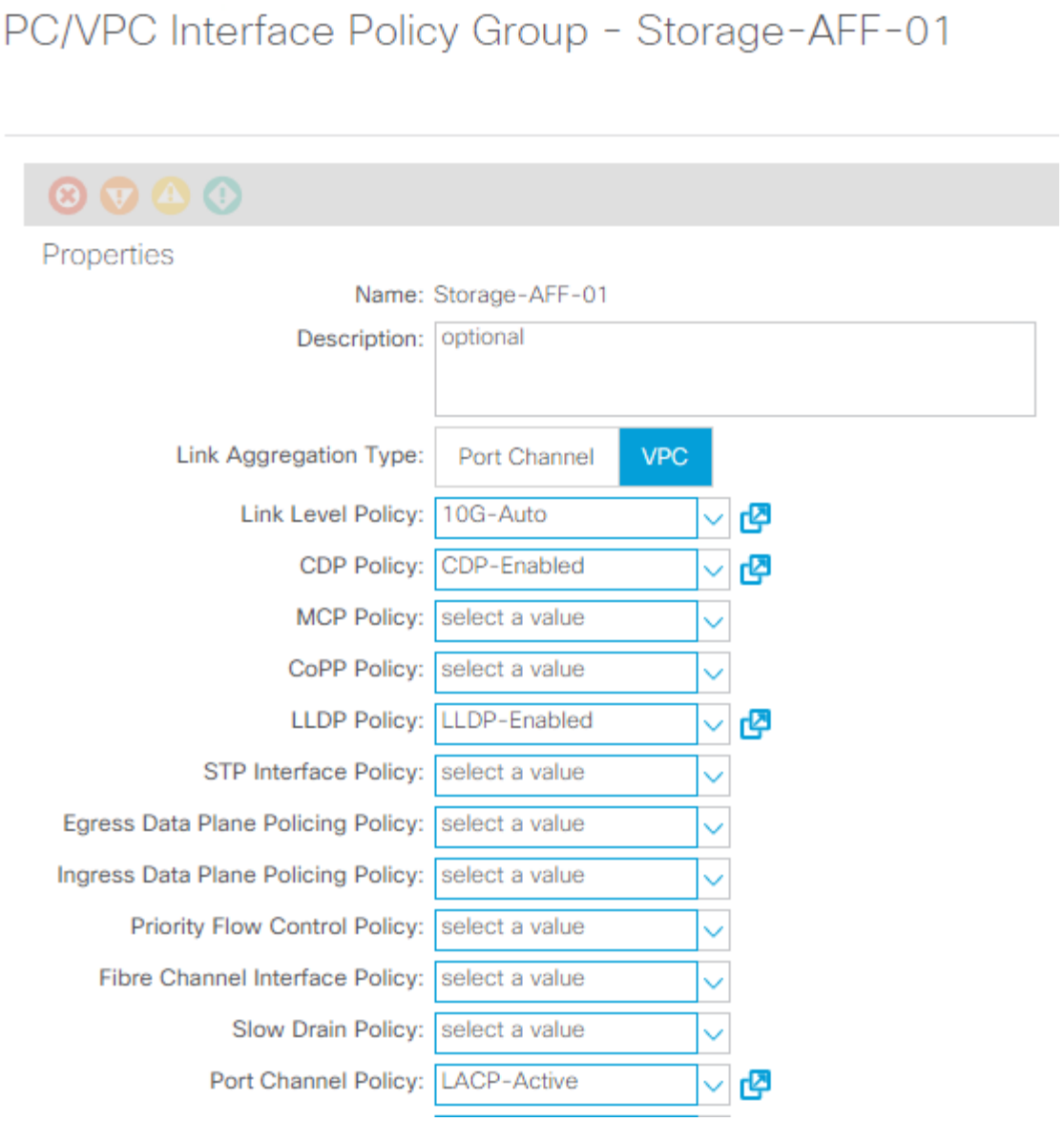
NetApp ONTAP is a powerful storage operating system with capabilities like inline compression, nondisruptive hardware upgrades, and cross-storage import. A NetApp ONTAP cluster provides a unified storage system with simultaneous data access and management of Network File System (NFS), Common Internet File System (CIFS), iSCSI, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), and NVMe/FC protocols. ONTAP provides robust data protection capabilities, such as NetApp MetroCluster, SnapLock, Snapshot copies, SnapVault, SnapMirror, SyncMirror technologies and more. For more information, see the ONTAP documentation.

To extend the capabilities of storage to file services and add many more data protection abilities, ONTAP can be used in conjunction with NetApp HCI. If NetApp ONTAP already exists in your environment, you can easily integrate it with NetApp HCI and Cisco ACI.

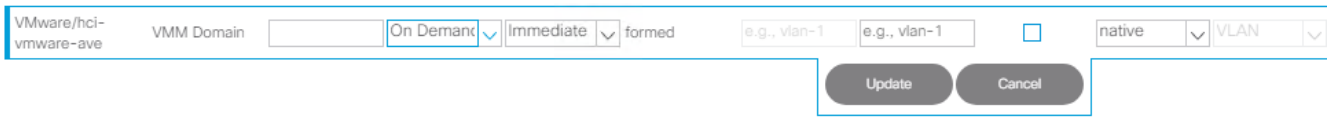
Workflow

The following high-level workflow was used to set up the environment. Each of these steps might involve several individual tasks.

1. Create a separate bridge domain and EPG on ACI for NFS and/or other protocols with the corresponding subnets. You can use the same HCI-related iSCSI EPGs.
2. Make sure you have proper contracts in place to allow inter-EPG communication for only the required ports.
3. Configure the interface policy group and selector for interfaces towards AFF controllers. Create a vPC policy group with the LACP Active mode for port-channel policy.



4. Attach both a physical and VMM domain to the EPGs created. Attach the vPC policy as static paths and, in the case of theCisco AVE virtual switch, use Native switching mode when you attach the VMM domain.



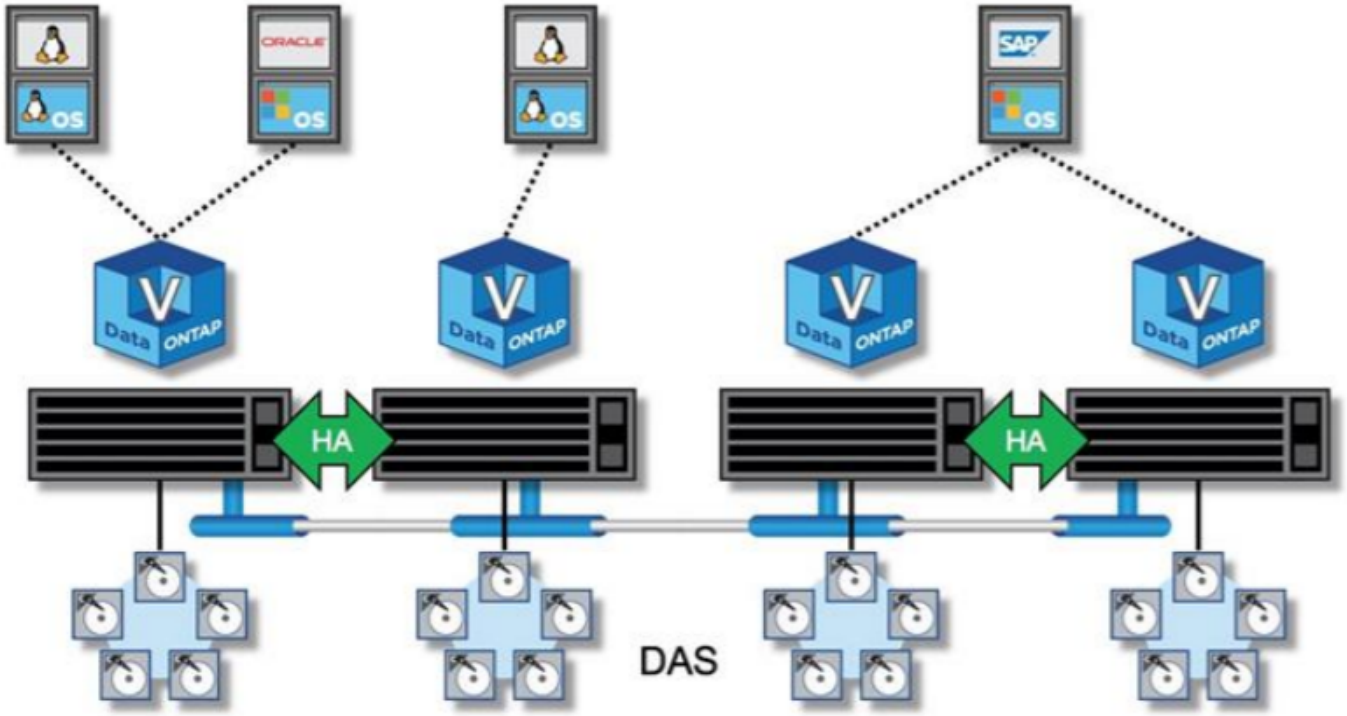
5. Install and configure an ONTAP cluster on the AFF controllers. Then create and configure NFS and/or iSCSI volumes/LUNs. See the AFF and ONTAP documentation for more information.
6. Create a VMkernel adapter (in the case of VMware ESXi) or a logical interface (in the case of RHV-H and RHEL-KVM hosts) attaching the NFS (or other protocols) port group or logical network.
7. Create additional datastores, storage domains, or storage pools on hypervisors (VMware, RHV, or KVM) using AFF storage.

Next: ONTAP Select with VMware vSphere: NetApp HCI and Cisco ACI

ONTAP Select with VMware vSphere: NetApp HCI and Cisco ACI

NetApp ONTAP Select is the NetApp solution for software-defined storage (SDS), bringing enterprise-class storage management features to the software-defined data center. ONTAP Select extends ONTAP functionality to extreme edge use cases including IoT and tactical servers as a software-defined storage appliance that acts as a full storage system. It can run as a simple VM on top of a virtual environment to provide a flexible and scalable storage solution.

Running ONTAP as software on top of another software application allows you to leverage much of the qualification work done by the hypervisor. This capability is critical for helping us to rapidly expand our list of supported platforms. Also, positioning ONTAP as a virtual machine (VM) allows customers to plug into existing management and orchestration frameworks, which allows rapid provisioning and end-to-end automation from deployment to sunsetting. The following figure provides an overview of a four-node ONTAP Select instance.

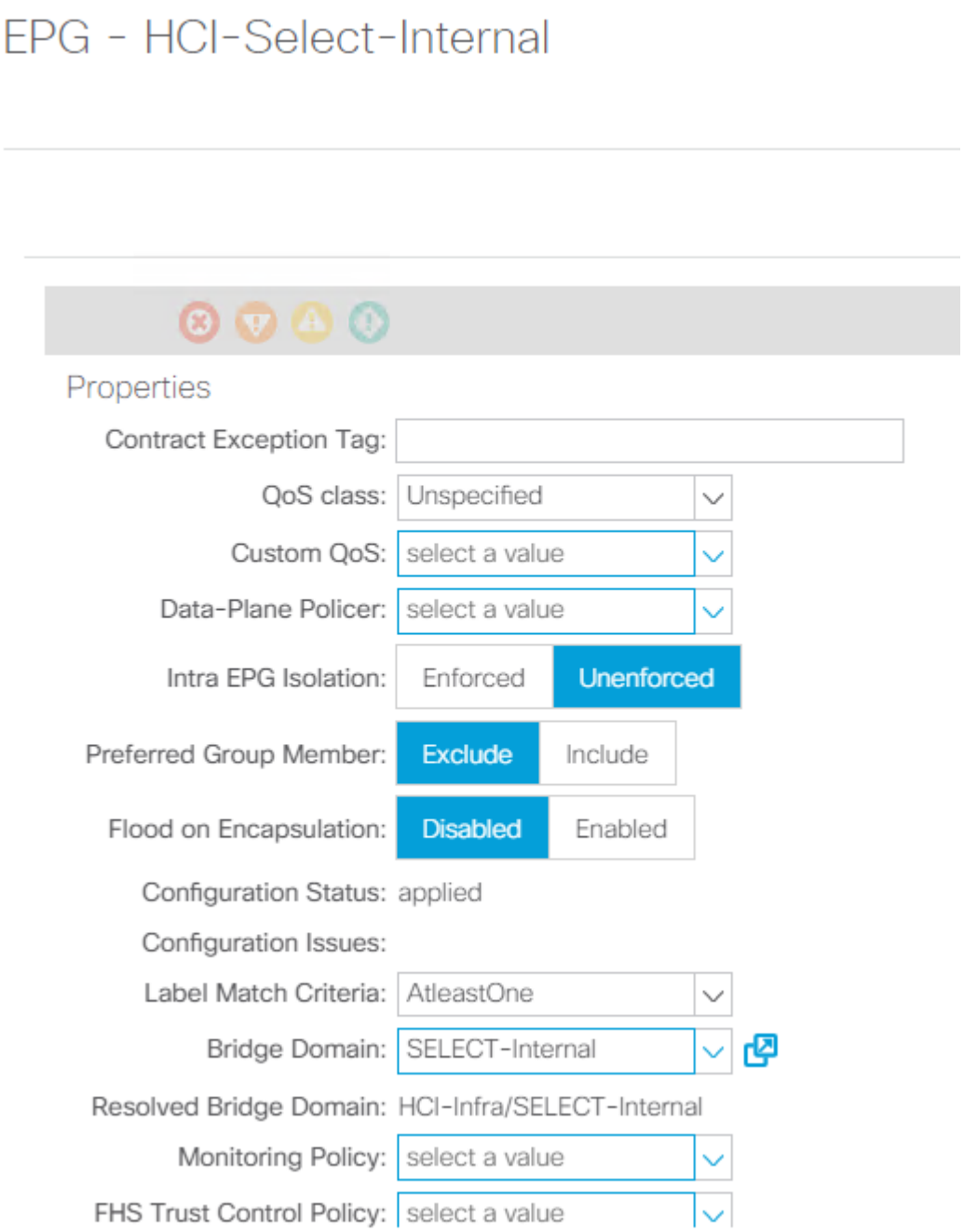


Deploying ONTAP Select in the environment to use the storage offered by NetApp HCI extends the capabilities of NetApp Element.

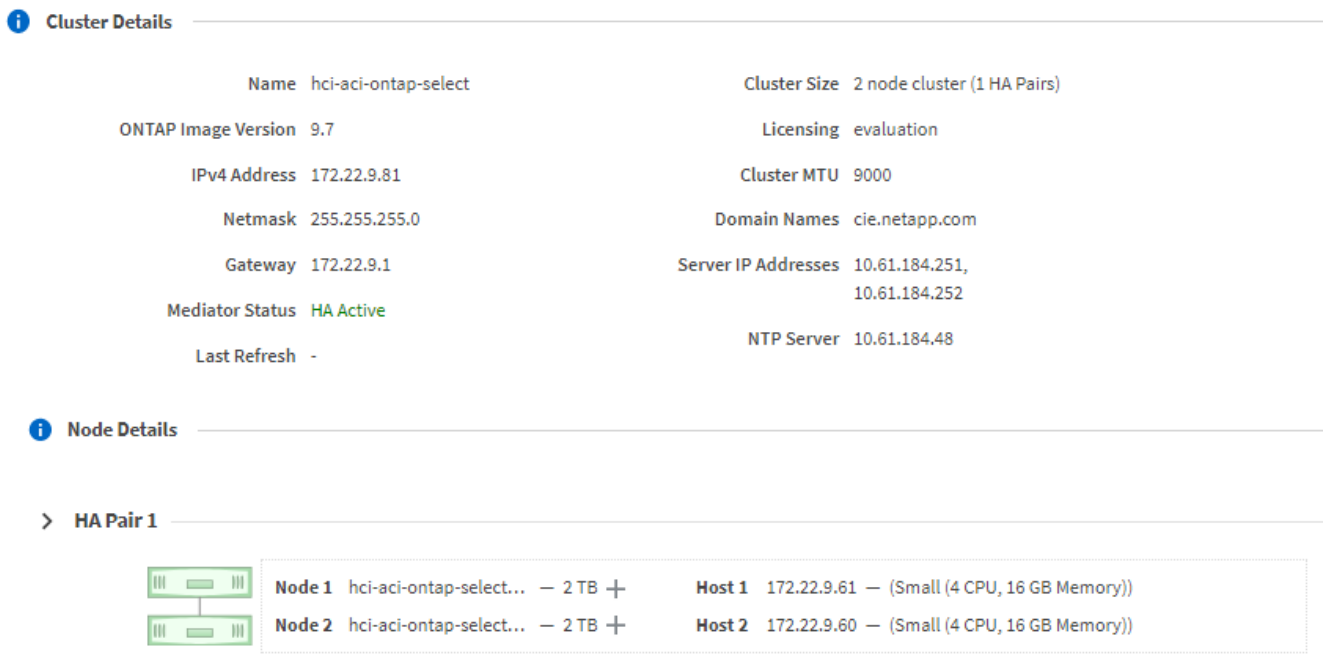
Workflow

The following workflow was used to set up the environment. In this solution, we deployed a two-node ONTAP Select cluster. Each of these steps might involve several individual tasks.

1. Create an L2 BD and EPG for the OTS cluster's internal communication and attach the VMM domain to the EPG in the Native switching mode (in case of a Cisco AVE virtual switch) with Pre-Provision Resolution Immediacy.



2. Verify that you have a VMware vSphere license.
3. Create a datastore that hosts OTS.
4. Deploy and configure ONTAP Select according to the ONTAP Select documentation.

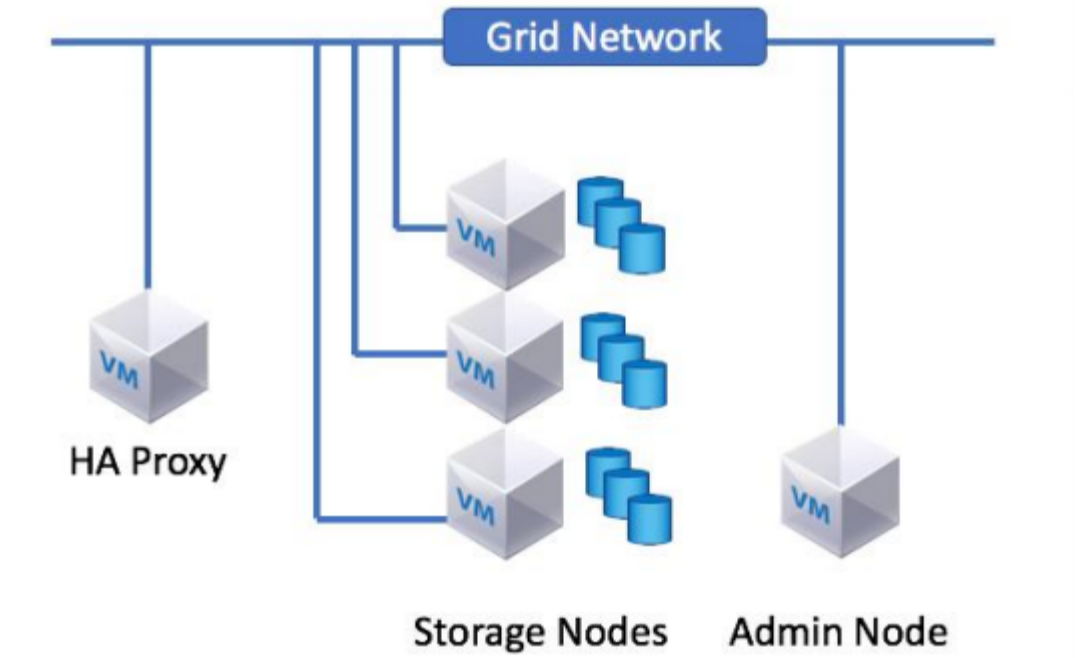


5. Create additional datastores using ONTAP Select to make use of additional capabilities.

Next: StorageGRID with VMware vSphere: NetApp HCI and Cisco ACI

StorageGRID with VMware vSphere: NetApp HCI and Cisco ACI

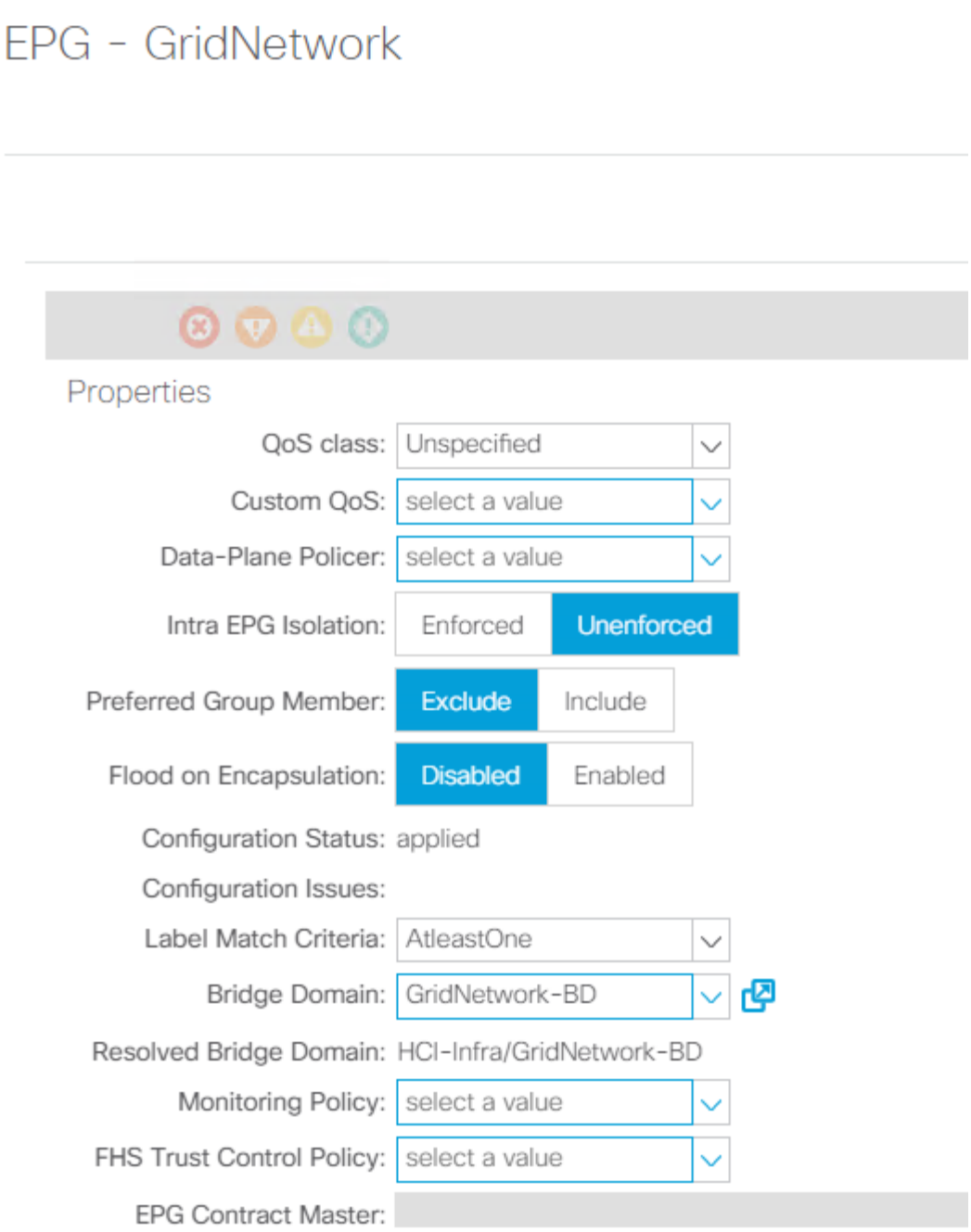
StorageGRID is a robust software-defined, object-based storage platform that stores and manages unstructured data with a tiered approach along with intelligent policy-driven management. It allows you to manage data while optimizing durability, protection, and performance. StorageGRID can also be deployed as hardware or as an appliance on top of a virtual environment that decouples storage management software from the underlying hardware. StorageGRID opens a new realm of supported storage platforms, increasing flexibility and scalability. StorageGRID platform services are also the foundation for realizing the promise of the hybrid cloud, letting you tier and replicate data to public or other S3-compatible clouds. See the StorageGRID documentation for more details. The following figure provides an overview of StorageGRID nodes.



Workflow

The following workflow was used to set up the environment. Each of these steps might involve several individual tasks.

1. Create an L2 BD and EPG for the grid network used for internal communication between the nodes in the StorageGRID system. However, if your network design for StorageGRID consists of multiple grid networks, then create an L3 BD instead of an L2 BD. Attach the VMM domain to the EPG with the Native switching mode (in the case of a Cisco AVE virtual switch) and with Pre-Provision Resolution Immediacy. The corresponding port group is used for the grid network on StorageGRID nodes.



2. Create a datastore to host the StorageGRID nodes.
3. Deploy and configure StorageGRID. For more details on installation and configuration, see the StorageGRID documentation. If the environment already has ONTAP or ONTAP Select, then you can use the NetApp Fabric Pool feature. Fabric Pool is an automated storage tiering feature in which active data resides on local high-performance solid-state drives (SSDs) and inactive data is tiered to low-cost object storage. It was first made available in NetApp ONTAP 9.2. For more information on Fabric Pool, see the documentation here.

Next: Validation Results

Validation Results

We used the iPerf tool for testing network throughput, and the baseline expectation was that the test systems should achieve throughput within 10% of the maximum line rate. Test results for different virtual switches is indicated in the following table.

For storage IOPS subsystem measurement, we used the IOmeter tool. The baseline expectation was that the test systems should achieve read/write throughput within 10% of the maximum. Test results for different hypervisors is indicated in the following table.

We considered the following scenarios for the network line rate and storage IOPS testing:

VMware

- VMs on a NetApp HCI datastore (with and without micro-segmentation)
- VMs on a NetApp ONTAP datastore
- VMs on a NetApp ONTAP Select datastore

Red Hat Virtualization

- VMs on a NetApp HCI datastore
- VMs on a NetApp ONTAP datastore

KVM (RHEL)

- VMs on a NetApp HCI datastore

Miscellaneous

- One VM on RHV with a NetApp HCI datastore and one VM on VMware vSphere with a NetApp ONTAP datastore.

Hypervisor	Virtual Switch	iPerf	IOmeter	Micro-segmentation
VMware	VDS	Pass	Pass	Pass
RHV	Linux Bridge	Pass	Pass	N/A
RHEL-KVM	Open vSwitch	Pass	Pass	N/A

Next: Where to Find Additional Information

Where to Find Additional Information

To learn more about the information that is described in this document, review the following documents and/or websites:

- NetApp HCI Documentation

  https://www.netapp.com/us/documentation/hci.aspx

- Cisco ACI Documentation

  https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/index.html

- Cisco Nexus 9000 Series Switches

  http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

- NetApp AFF A-series

  http://www.netapp.com/us/products/storage-systems/all-flash-array/aff-a-series.aspx

- ONTAP Documentation

  https://docs.netapp.com/ontap-9/index.jsp

- ONTAP Select Documentation

  https://docs.netapp.com/us-en/ontap-select/

- StorageGRID Documentation

  https://docs.netapp.com/sgws-113/index.jsp

- Red Hat Virtualization

  https://access.redhat.com/documentation/en-us/red_hat_virtualization/4.3/

- VMware vSphere

  https://docs.vmware.com/en/VMware-vSphere/index.html

- VMware vCenter Server

  http://www.vmware.com/products/vcenter-server/overview.html

- NetApp Interoperability Matrix Tool

  http://now.netapp.com/matrix

- Cisco ACI Virtualization Compatibility Matrix

  https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html

- VMware Compatibility Guide

  http://www.vmware.com/resources/compatibility