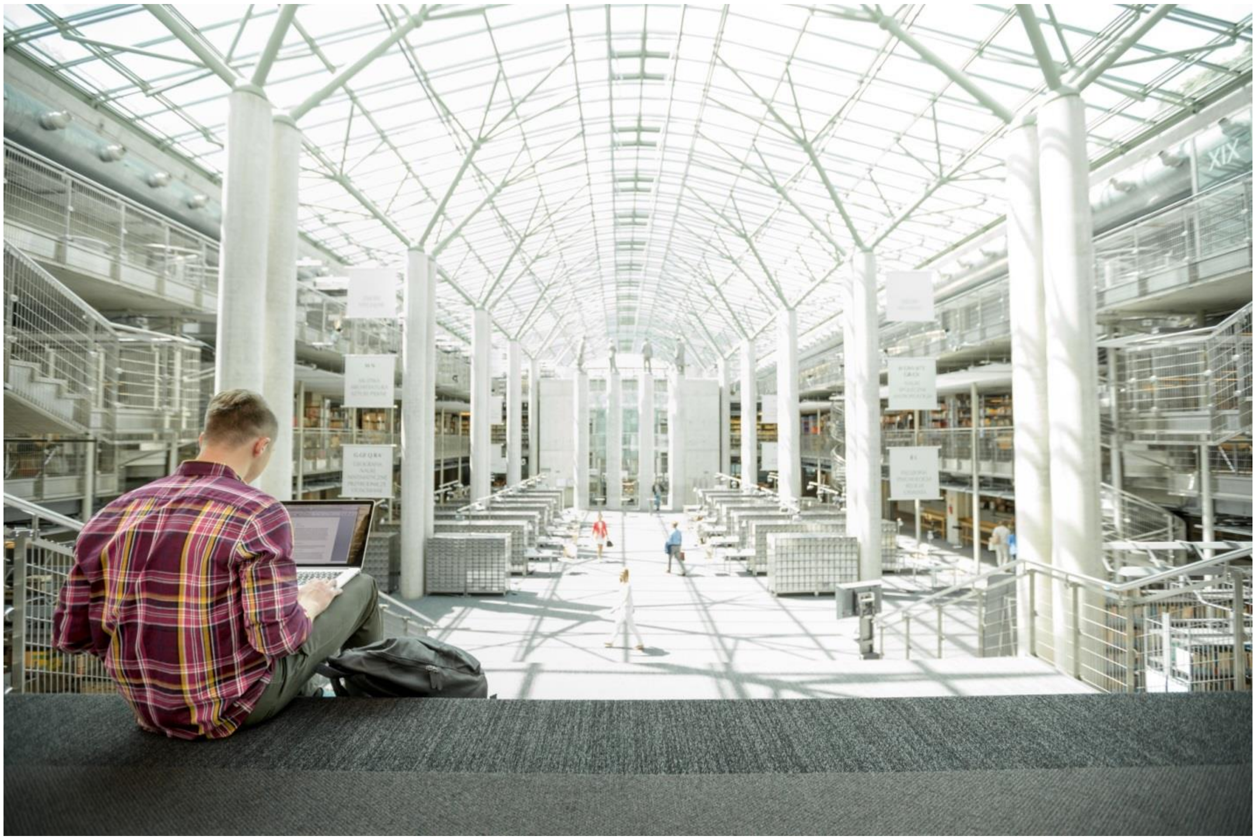


# Cisco UCS Integrated Infrastructure with Red Hat OpenStack Platform 8 and Red Hat Ceph Storage

## Deployment Guide

**Last Updated:** November 13, 2016

# About Cisco Validated Designs

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit

http://www.cisco.com/go/designzone.

ALL DESIGNS, SPECIFICATIONS, STATEMENTS, INFORMATION, AND RECOMMENDATIONS (COLLECTIVELY, "DESIGNS") IN THIS MANUAL ARE PRESENTED "AS IS," WITH ALL FAULTS. CISCO AND ITS SUPPLIERS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE. IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THE DESIGNS, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

THE DESIGNS ARE SUBJECT TO CHANGE WITHOUT NOTICE. USERS ARE SOLELY RESPONSIBLE FOR THEIR APPLICATION OF THE DESIGNS. THE DESIGNS DO NOT CONSTITUTE THE TECHNICAL OR OTHER PROFESSIONAL ADVICE OF CISCO, ITS SUPPLIERS OR PARTNERS. USERS SHOULD CONSULT THEIR OWN TECHNICAL ADVISORS BEFORE IMPLEMENTING THE DESIGNS. RESULTS MAY VARY DEPENDING ON FACTORS NOT TESTED BY CISCO.

# Table of Contents

# Executive Summary

The Cisco Validated Design program consists of systems and solutions that are designed, tested, and documented to facilitate and improve customer deployments. These designs incorporate a wide range of technologies and products into a portfolio of solutions that have been developed to address the business needs of our customers.

The reference architecture described in this document is a realistic use case for deploying Red Hat OpenStack Platform 8 on Cisco UCS Blade and Rack-Mount servers. This document provides step-by-step instructions for setting up Cisco UCS hardware, installing Red Hat OpenStack Platform director, issues and workarounds evolved during installation, integration of Cisco Plugins with OpenStack, requirements for leveraging High Availability from both hardware and software, and lessons learned while validating the solution and including a few troubleshooting steps.

Cisco UCS Integrated Infrastructure for Red Hat OpenStack Platform is an all-in-one solution for deploying OpenStack based private cloud using Cisco Infrastructure and Red Hat OpenStack Platform. This solution is validated and supported by Cisco and Red Hat, for rapid infrastructure deployment and reduce the risk of scaling from proof-of-concept to enterprise production environment completely.

# Solution Overview

## Introduction

Automation, virtualization, cost, and ease of deployment are the key criteria to meet the growing IT challenges. Virtualization is a key and critical strategic deployment model for reducing the Total Cost of Ownership (TCO) and achieving better utilization of the platform components like hardware, software, network and storage. The platform should be flexible, reliable and cost effective for enterprise applications.

The Cisco UCS solution implementing Red Hat OpenStack Platform provides a very simplistic yet fully integrated and validated infrastructure to deploy virtual machines (VMs) in various sizes to suit your application needs. Cisco Unified Computing System (UCS) is a next-generation data center platform that unifies computing, network, storage access, and virtualization into a single interconnected system, which makes Cisco UCS an ideal platform for OpenStack architecture. The combined architecture of Cisco UCS platform, Red Hat OpenStack Platform and Red Hat Ceph Storage can accelerate your IT transformation by enabling faster deployments, greater flexibility of choice, efficiency, and lower risks. Furthermore, Cisco Nexus series of switches provide the network foundation for the next-generation data center.

This deployment guide provides you with step-by-step instructions to install Red Hat OpenStack Platform director and Red Hat Ceph Storage on Cisco UCS Blades and Rack-Mount servers. The traditional complexities of installing OpenStack are simplified by Red Hat Openstack Platform director while Cisco UCS Manager capabilities bring an integrated, scalable, multi-chassis platform in which all resources participate in a unified management domain. The solution included in this CVD is an effort by Cisco Systems, Inc. in partnership with Red Hat, Inc., and Intel Corporation.

## Audience

The audience for this document includes, but is not limited to, sales engineers, field consultants, professional services, IT managers, partner engineers, IT architects, and customers who want to take advantage of an infrastructure that is built to deliver IT efficiency and enable IT innovation. The reader of this document is expected to have the necessary training and background to install and configure Red Hat Enterprise Linux, Cisco Unified Computing System (UCS) and Cisco Nexus Switches as well as a high level understanding of OpenStack components. External references are provided where applicable and it is recommended that the reader be familiar with these documents.

Readers are also expected to be familiar with the infrastructure, network and security policies of the customer installation.

## Purpose of the Document

This document details the installation steps for Red Hat OpenStack Platform 8 and Red Hat Ceph Storage 1.3 architecture on the Cisco UCS platform. It also describes the daily operational challenges in running OpenStack and steps to mitigate them, High Availability use cases, Live Migration, common troubleshooting aspects of OpenStack along with Operational best practices.

## Solution Summary

This solution is focused on Red Hat OpenStack Platform 8 (based on the upstream OpenStack Liberty release) and Red Hat Ceph Storage 1.3 on Cisco Unified Computing System. The advantages of Cisco UCS and Red Hat OpenStack Platform combine to deliver an OpenStack Infrastructure as a Service (IaaS) deployment that is quick and easy to setup. The solution can scale up for greater performance and capacity or scale out for environments that require consistent, multiple deployments. Converged infrastructure of Compute, Networking, and Storage components from Cisco UCS is a validated enterprise-class IT platform, rapid deployment for business critical applications, reduces costs, minimizes risks, and increase flexibility and business agility Scales up for future growth.

Red Hat OpenStack Platform 8 on Cisco UCS helps IT organizations accelerate cloud deployments while retaining control and choice over their environments with open and interoperable cloud solutions. It also offers redundant architecture on compute, network, and storage perspective. The solution comprises the following key components:

- Cisco Unified Computing System (UCS)
  - Cisco UCS 6200 Series Fabric Interconnects
  - Cisco VIC 1340
  - Cisco VIC 1227
  - Cisco 2204XP IO Module or Cisco UCS Fabric Extenders
  - Cisco B200 M4 Servers
  - Cisco C240 M4 Servers
- Cisco Nexus 9300 Series Switches
- Cisco Nexus Plugin for Nexus Switches
- Cisco UCS Manager Plugin for Cisco UCS
- Red Hat Enterprise Linux 7.2
- Red Hat OpenStack Platform director
- Red Hat OpenStack Platform 8
- Red Hat Ceph Storage 1.3

The scope is limited to the infrastructure pieces of the solution. It does not address the vast area of the OpenStack components and multiple configuration choices available in OpenStack.

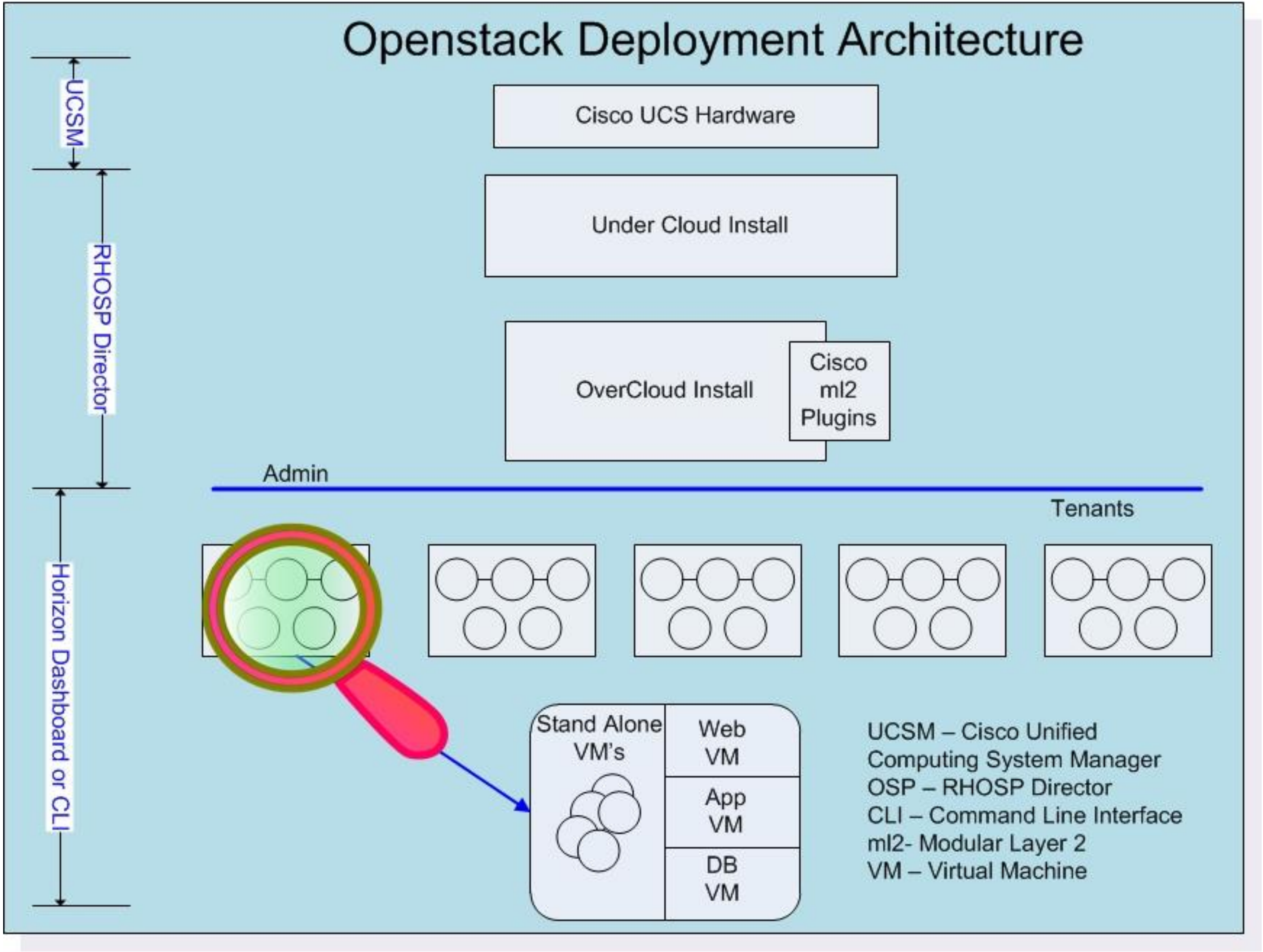
# Deployment Architecture

This architecture is based on Red Hat OpenStack Platform built on Cisco UCS hardware is an integrated foundation to create, deploy, and scale OpenStack cloud based on Liberty OpenStack community release. This deployment is further to Cisco's validated design on Red Hat OpenStack Platform 7 released earlier. The earlier deployment guide can be referenced at: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/ucs_openstack_osp7.html

The reference architecture use case provides a comprehensive, end-to-end example of deploying RHOSP 8 cloud on bare metal using Red Hat OpenStack Platform director and services through heat templates.

The first section in this CVD covers setting up of Cisco hardware - the blade and rack servers, chassis and Fabric Interconnects and the peripherals like Cisco Nexus 9000 switches. The second section explains how to install cloud through Red Hat OpenStack Platform director. The final section includes the functional and High Availability tests on the configuration, and the best practices evolved while validating the solution.

**Figure 1** OpenStack Deployment Architecture

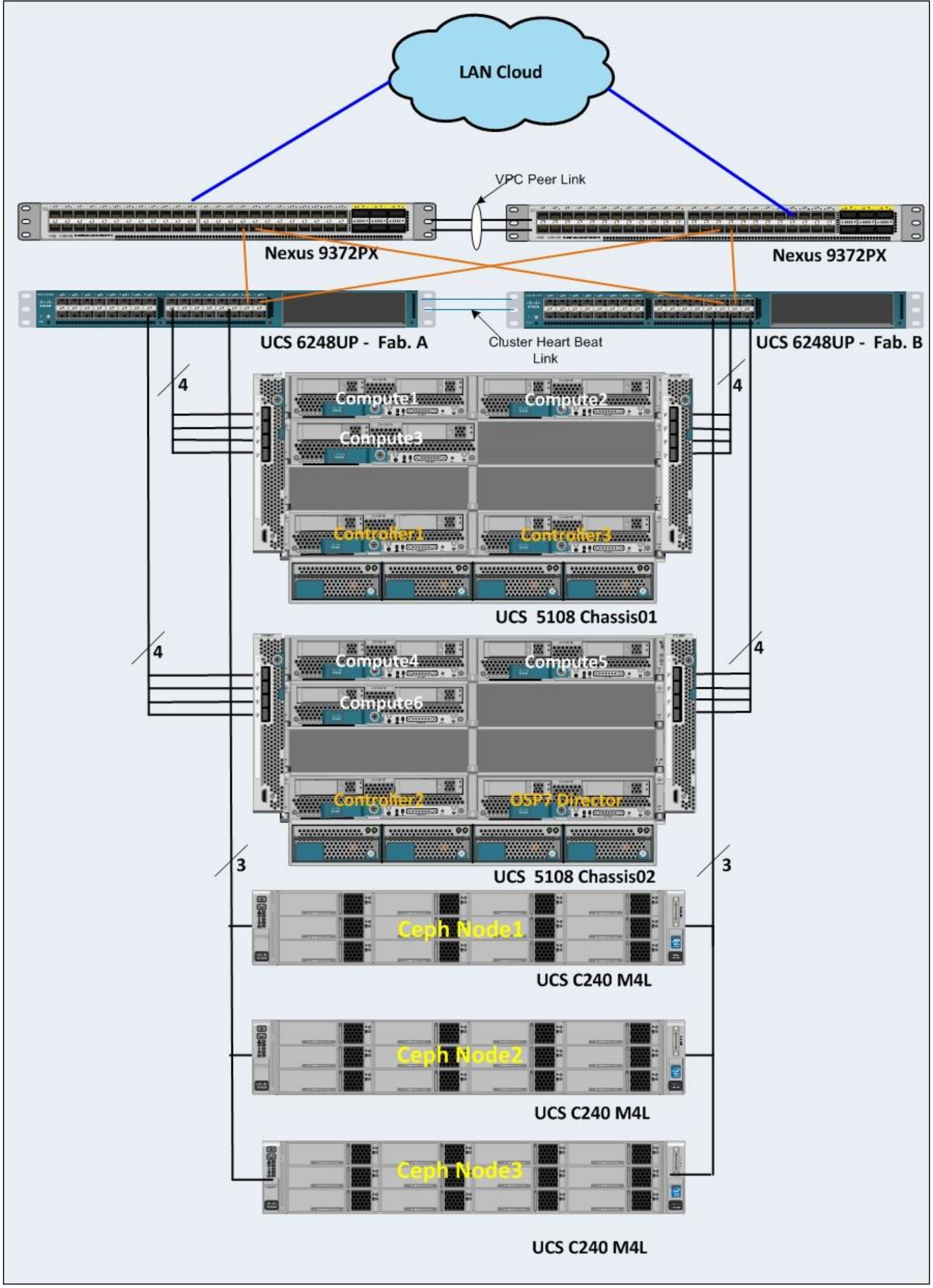


# Solution Design

## Physical Topology

Figure 2 illustrates the physical topology of this solution.

**Figure 2** Physical Topology



The configuration is comprised of three controller nodes, six compute nodes, three storage nodes, and a pair of Cisco UCS Fabric Interconnect and Cisco Nexus Switches, where most of the tests were conducted. This architecture is scalable horizontally and vertically within the chassis.

- More Compute Nodes and Chassis can be added as desired.
- More Ceph Nodes for storage can be added. The Ceph nodes can be Cisco UCS C240M4L or C240M4S.
- If there is a higher bandwidth requirement then Cisco UCS Fabric Extender 2208XP can be used instead of 2204XP model, which is used in this configuration.
- Both Cisco UCS Fabric Interconnects and Cisco Nexus Switches can be 96 port switches instead of 48 ports as shown in Figure 2.

## Solution Overview

The Cisco Unified Computing System is an integrated, scalable, multi-chassis platform in which all resources participate in a unified management domain. The Cisco Unified Computing System accelerates the delivery of new services simply, reliably, and securely through end-to-end provisioning and migration support for both virtualized and non-virtualized systems. Cisco UCS Manager using single connect technology manages servers and chassis and performs auto-discovery to detect inventory, manage, and provision system components that are added or changed.

The Red Hat OpenStack Platform IaaS cloud on Cisco UCS servers is implemented as a collection of interacting services that control compute, storage, and networking resources.

OpenStack Networking handles creation and management of a virtual networking infrastructure in the OpenStack cloud. Infrastructure elements include networks, subnets, and routers. Because OpenStack Networking is software-defined, it can react in real-time to changing network needs, such as creation and assignment of new IP addresses.

Compute serves as the core of the OpenStack cloud by providing virtual machines on demand. Compute servers support the libvirt driver that uses KVM as the hypervisor. The hypervisor creates virtual machines and enables live migration from node to node. OpenStack also provides storage services to meet the storage requirements for the above mentioned virtual machines.

The solution also includes OpenStack Networking ML2 Core components.

The Cisco Nexus driver for OpenStack Neutron allows customers to easily build their infrastructure-as-a-service (IaaS) networks using the industry’s leading networking platform, delivering performance, scalability, and stability with the familiar manageability and control you expect from Cisco® technology.

Cisco UCS Manager Plugin configures compute blades with necessary VLANs. The Cisco UCS Manager Plugin talks to the Cisco UCS Manager application running on the Fabric Interconnect.

# System Hardware and Software Specifications

Table 1 and Table 2 list the Hardware and Software releases used for solution verification.

**Table 1 Required Hardware Components**

| | **Hardware** | **Quantity** | **Firmware Details** |
|---|---|---|---|
| OSP director | Cisco UCS B200M4 blade | 1 | 2.2(5) |
| Controller | Cisco UCS B200M4 blade | 3 | 2.2(5) |
| Compute | Cisco UCS B200M4 blade | 6 | 2.2(5) |
| Storage | Cisco UCS C240M4L or C240M4S Rack server | 3 | 2.2(5) |
| Fabrics Interconnects | Cisco UCS 6248UP FIs | 2 | 2.2(5) |
| Nexus Switches | Cisco Nexus 9372 NX-OS | 2 | 7.0(3)I1(3) |

**Table 2 Software Specifications**

| | **Software** | **Version** |
|---|---|---|
| Operating System | Red Hat Enterprise Linux | 7.2 |
| OpenStack Platform | Red Hat OpenStack Platform | RHOSP 8 |
| | Red Hat OpenStack Platform director | RHOSP 8 |
| | Red Hat Ceph Storage | 1.3 |
| Plugins | Cisco Nexus Plugin | RHOSP 8 |
| | Cisco UCS Manager Plugin | RHOSP 8 |

# Bill of Materials

This section contains the Bill of Materials used in the configuration.

**Table 3 Bill of Materials**

| Component | Model | Quantity | Comments |
|---|---|---|---|
| OpenStack Platform director Node | Cisco UCS B200M4 blade | 1 | CPU – 2 x E5-2630 V3<br>Memory – 8 x 16GB 2133 MHz DIMM – total of 128G<br>Local Disks – 2 x 300 GB SAS disks for Boot<br>Network Card – 1x1340 VIC<br>Raid Controller – Cisco MRAID 12 G SAS Controller |
| Controller Nodes | Cisco UCS B200M4 blades | 3 | CPU – 2 x E5-2630 V3<br>Memory – 8 x 16GB 2133 MHz DIMM – total of 128G<br>Local Disks – 2 x 300 GB SAS disks for Boot<br>Network Card – 1x1340 VIC<br>Raid Controller – Cisco MRAID 12 G SAS Controller |
| Compute Nodes | Cisco UCS B200M4 blades | 6 | CPU – 2 x E5-2660 – V3<br>Memory – 16 x 16GB 2133 MHz DIMM – total of 256G<br>Local Disks – 2 x 300 GB SAS disks for Boot<br>Network Card – 1x1340 VIC<br>Raid Controller – Cisco MRAID 12 G SAS Controller |
| Storage Nodes<br>(only one of LFF/SFF) | Cisco UCS C240M4L Rack Servers | 3 | CPU – 2 x E5-2630 – V3<br>Memory – 8 x 16GB 2133 MHz DIMM – total of 128G<br>Internal HDD – None<br>Ceph OSD's – 8 x 6TB SAS Disks<br>Ceph Journals – 2 x 400GB SSD's<br>OS Boot – 2 x 1TB SAS Disks<br>Network Cards – 1 x VIC 1227<br>Raid Controller – Cisco MRAID 12 G SAS Controller |

| Component | Model | Quantity | Comments |
|---|---|---|---|
| | Cisco UCS C240M4S Rack Servers | 3 | CPU – 2  x E5-2630 – V3<br>Memory – 8 x 16GB 2133 MHz DIMM – total of 128G<br>Internal HDD – None<br>Ceph OSD's – 18 x 1.2 TB SAS Disks<br>Ceph Journals – 4 x 400GB SSD's<br>OS Boot – 2 x 1TB SAS Disks<br>Network Cards – 1 x VIC 1227<br>Raid Controller – Cisco MRAID 12 G SAS Controller |
| Chassis | Cisco UCS 5108 Chassis | 2 | |
| IO Modules | Cisco UCS 2204XP Fabric Extenders | 4 | |
| Fabric Interconnects | Cisco UCS 6248UP Fabric Interconnects | 2 | |
| Switches | Cisco Nexus 9372PX Switches | 2 | |

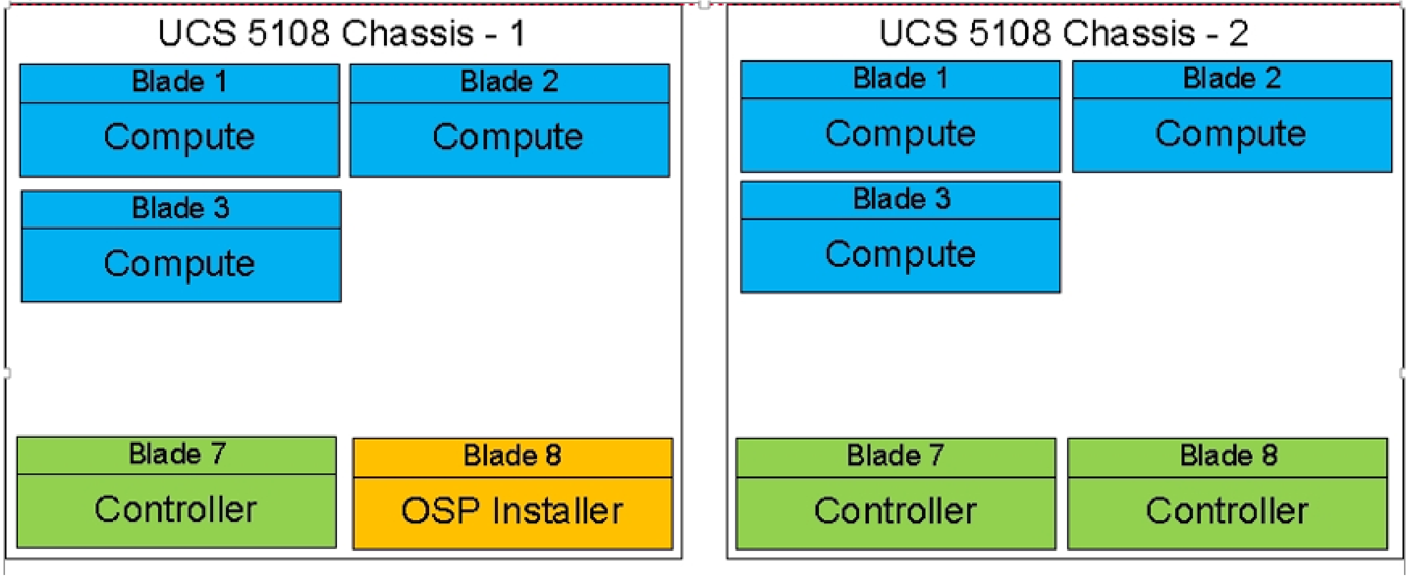
# Solution Components

This section provides an overview of the components used in this solution.

## Cisco UCS Blades Distribution in Chassis

Figure 3 lists the server distribution in the Cisco UCS Chassis.

**Figure 3** Servers Distribution in Cisco UCS Chassis



The controller and compute nodes are distributed across the chassis. This gives High Availability to the stack though a failure of Chassis per se does not happen. There is only one Installer node in the system and can be added in any one of the Chassis as above. In case of larger deployments having 3 or more chassis, it is recommended to distribute one controller in each chassis.

In larger deployments where the chassis are fully loaded with blades, a better approach would be to distribute the tenant and storage traffic across the Fabric Interconnects. This method ensures that the tenant traffic is distributed evenly across both the fabrics.

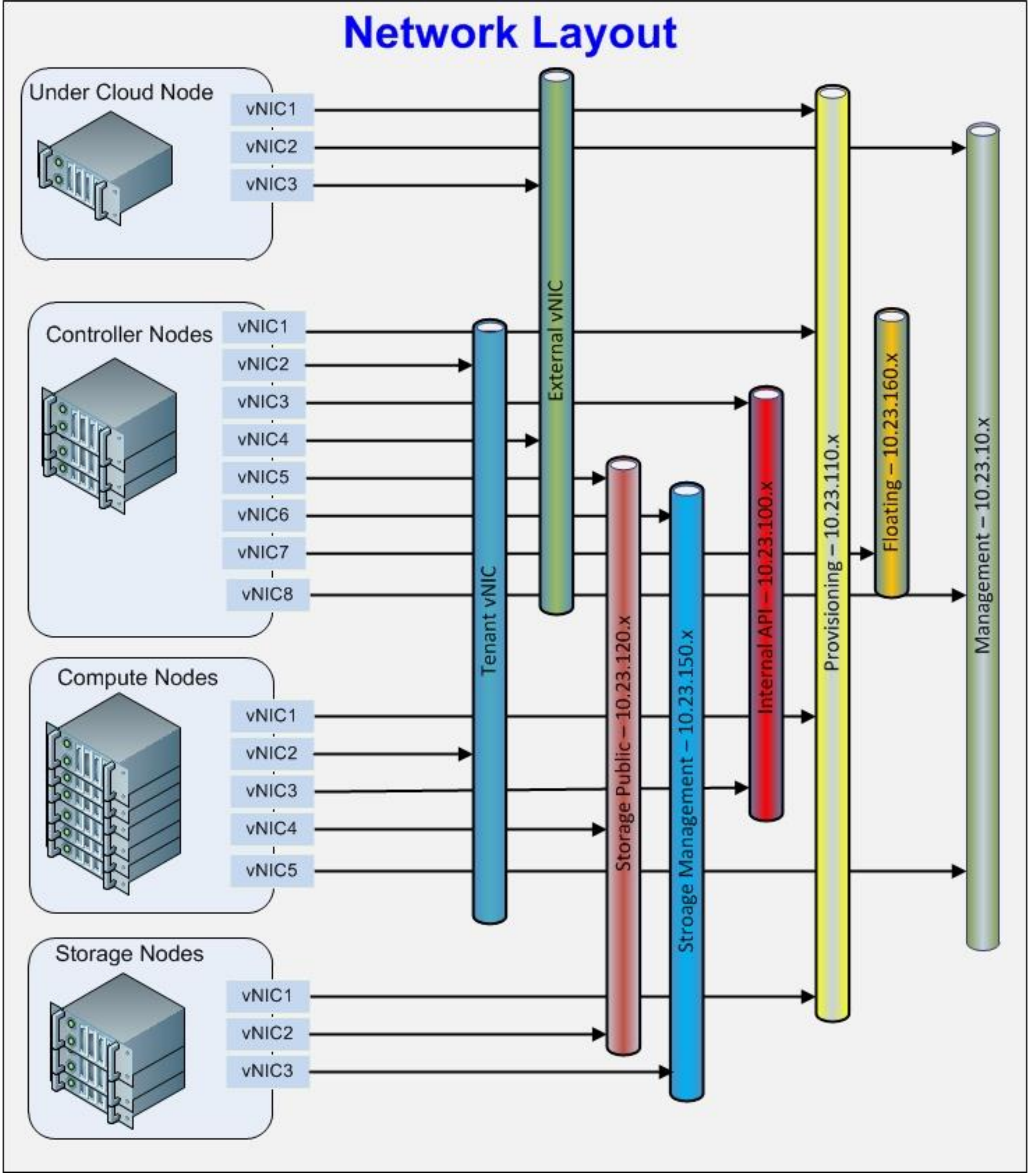
## Service Profiles

Service profiles will be created from the Service Profile Templates. However once successfully created, they will be unbound from the templates.

## Cisco UCS vNIC Configuration

Figure 4 illustrates the network layout.

**Figure 4** Network Layout



A Floating or Provider network is not necessary. VMs can be accessed either through floating point network or through external network. This is determined how the external bridge is configured, covered later in this document.

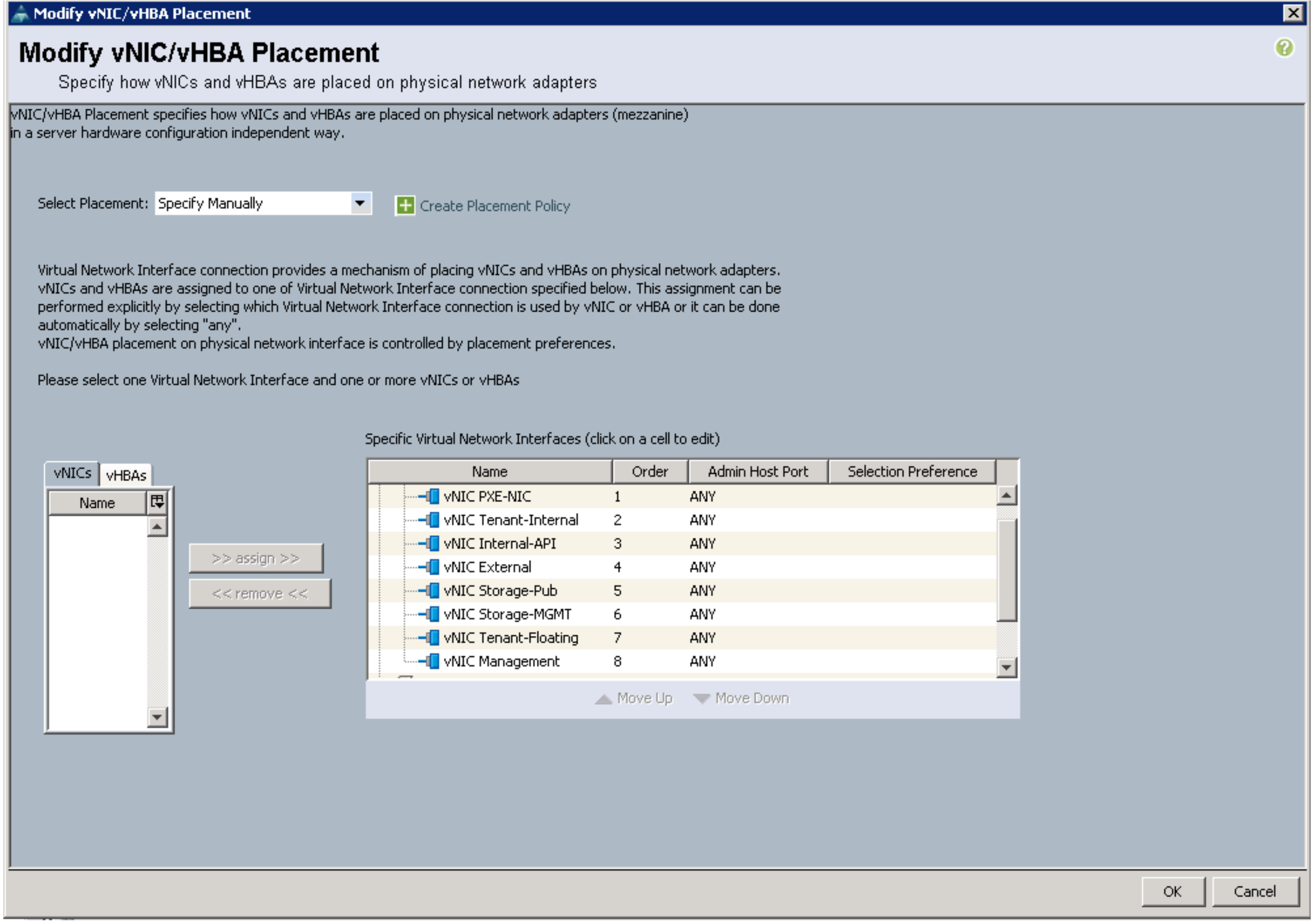
The family of vNICs are placed in the same Fabric Interconnect to avoid an extra hop to the upstream Nexus switches.

The following categories of vNICs are used in the setup:

- Provisioning Interfaces pxe vNICs are pinned to Fabric A
- Tenant vNICs are pinned to Fabric A
- Internal API vNICs are pinned to Fabric B
- External Interfaces vNICs are pinned to Fabric A
- Storage Public Interfaces are pinned to Fabric A
- Storage Management Interfaces are pinned to Fabric B
- Management vNICs are pinned to Fabric A

While configuring vNICs in templates and with failover option enabled in Fabrics, the vNICs order has to be specified manually as shown below.

**Figure 5** vNIC Placement



Modify vNIC/vHBA Placement

Modify vNIC/vHBA Placement

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement: Specify Manually

Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vNIC PXE-NIC | 1 | ANY | |
| vNIC Tenant-Internal | 2 | ANY | |
| vNIC Internal-API | 3 | ANY | |
| vNIC External | 4 | ANY | |
| vNIC Storage-Pub | 5 | ANY | |
| vNIC Storage-MGMT | 6 | ANY | |
| vNIC Tenant-Floating | 7 | ANY | |
| vNIC Management | 8 | ANY | |

The order of vNICs has to be pinned as above for consistent PCI device naming options. The above is an example of controller blade. The same has to be done for all the other servers, the compute and storage nodes. This order should match the Overcloud heat templates NIC1, NIC2, NIC3, NIC4 etc.

## Red Hat OpenStack Platform 8 director

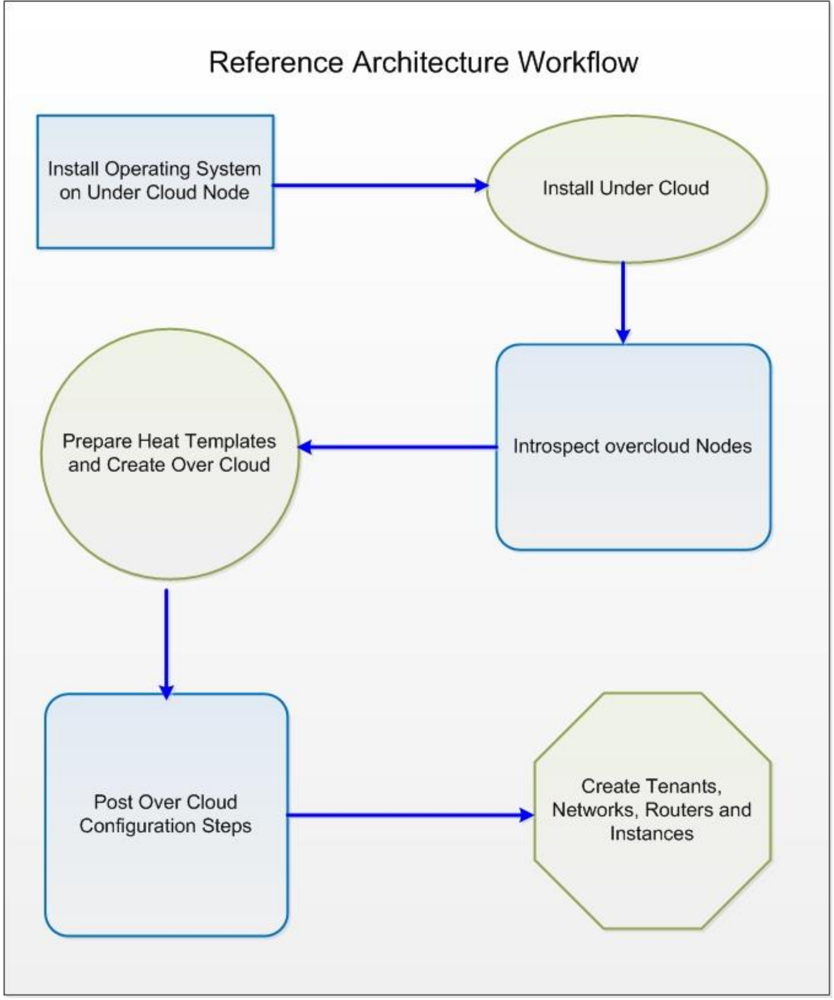
Red Hat OpenStack Platform 8 (RHOSP 8) delivers an integrated foundation to create, deploy, and scale a more secure and reliable public or private OpenStack cloud. RHOSP 8 starts with the proven foundation of Red Hat Enterprise Linux and integrates Red Hat's OpenStack Platform technology to provide production ready cloud platform. RHOSP 8 director is based on community based Liberty OpenStack release. Red Hat OpenStack Platform 8 introduces a cloud installation and lifecycle management tool chain. It provides the following:

- Simplified deployment through ready-state provisioning of bare metal resources
- Flexible network definitions
- High Availability with Red Hat Enterprise Linux Server High Availability
- Integrated setup and Installation of Red Hat Ceph Storage 1.3

## Reference Architecture Workflow

Figure 6 illustrates the reference architecture workflow.

**Figure 6** Reference Architecture Workflow


Reference Architecture Workflow
Install Operating System on Under Cloud Node
Install Under Cloud
Introspect overcloud Nodes
Prepare Heat Templates and Create Over Cloud
Post Over Cloud Configuration Steps
Create Tenants, Networks, Routers and Instances

Red Hat OpenStack Platform director is a new set of tool chain introduced with Kilo that automates the creation of Undercloud and Overcloud nodes as above. It performs the following:

- Install Operating System on Undercloud Node
- Install Undercloud Node
- Perform Hardware Introspection
- Prepare Heat templates and Install Overcloud

Undercloud Node is the deployment environment while Overcloud nodes are referred to nodes actually rendering the cloud services to the tenants.

The Undercloud is the TripleO (OOO – OpenStack on OpenStack) control plane. It uses native OpenStack APIs and services to deploy, configure, and manage the production OpenStack deployment. The Undercloud defines the Overcloud with Heat templates and then deploys it through the Ironic bare metal provisioning service. Red Hat OpenStack Platform director includes predefined Heat templates for the basic server roles that comprise the Overcloud. Customizable templates allow director to deploy, redeploy, and scale complex Overclouds in a repeatable fashion.

Ironic gathers information about bare metal servers through a discovery mechanism known as introspection. Ironic pairs servers with bootable images and installs them through PXE and remote power management.

Red Hat Openstack Platform director deploys all servers with the same generic image by injecting Puppet modules into the image to tailor it for specific server roles. It then applies host-specific customizations through Puppet including network and storage configurations. While the Undercloud is primarily used to deploy OpenStack, the Overcloud is a functional cloud available to run virtual machines and workloads.

The following subsections detail the roles that comprise the Overcloud.

### Control

This role provides endpoints for REST-based API queries to the majority of the OpenStack services. These include Compute, Image, Identity, Block, Network, and Data processing. The controller nodes also provide the supporting facilities for the APIs, database, load balancing, messaging, and distributed memory objects. They also provide external access to virtual machines. The controller can run as a standalone server or as a High Availability (HA) cluster. The current configuration was configured with HA.

### Compute

This role provides the processing, memory, storage, and networking resources to run virtual machine instances. It runs the KVM hypervisor by default. New instances are spawned across compute nodes in a round-robin fashion based on resource availability by default. The default filters can be altered if needed; for more information, see OpenStack documentation.

### Ceph-Storage

Ceph is a distributed block, object store and file system. This role deploys Object Storage Daemon (OSD) nodes for Ceph clusters. It also installs the Ceph Monitor service on the controller. The instance distribution is influenced by the currently set filters.

## Network Isolation

OpenStack requires multiple network functions. While it is possible to collapse all network functions onto a single network interface, isolating communication streams in their own physical or virtual networks provides better performance and scalability. Each OpenStack service is bound to an IP on a particular network. In a cluster a service virtual IP is shared by the HA controllers.

### Provisioning

The Control plane installs Overcloud through this network. All nodes must have a physical interface attached to the provisioning network. This network carries DHCP/PXE and TFTP traffic. It must be provided on a dedicated interface or native VLAN to the boot interface. The provisioning interface can also act as a default gateway for the Overcloud; the compute and storage nodes use this provisioning gateway interface on the Undercloud node.

### External

The External network is used for the Horizon dashboard and the Public APIs, as well as hosting the floating IPs that are assigned to VMs. The Neutron L3 routers which perform NAT are attached to this interface. The range of IPs that are assigned to floating IPs should not include the IPs used for hosts and VIPs on this network.

### Internal API

This network is used for connections to the API servers, as well as RPC messages using RabbitMQ and connections to the database. The Glance Registry API uses this network, as does the Cinder API. This network is typically only reachable from inside the OpenStack Overcloud environment, so API calls from outside the cloud will use the Public APIs via the external Network.

## Management

Red Hat OpenStack Platform 8 introduces a new network called as Management Network that provides access for system administration functions such as SSH access, DNS, NTP traffic etc. In the current validated design, this network was used to communicate with Cisco Nexus switches and UCS Manager.

## Tenant

Virtual machines communicate over the tenant network. It supports three modes of operation: VXLAN, GRE, and VLAN. VXLAN and GRE tenant traffic is delivered through software tunnels on a single VLAN. Individual VLANs correspond to tenant networks in cases where the VLAN tenant networks are used.

## Storage

This network carries storage communication including Ceph, Cinder, and Swift traffic. The virtual machine instances communicate with the storage servers through this network. Data-intensive OpenStack deployments should isolate storage traffic on a dedicated high bandwidth interface, that is, 10 GB interface. The Glance API, Swift proxy, and Ceph Public interface services are all delivered through this network.

## Storage Management

Storage management communication can generate large amounts of network traffic. This network is shared between the front and back end storage nodes. Storage controllers use this network to access data storage nodes. This network is also used for storage clustering and replication traffic.

Network traffic types are assigned to network interfaces through Heat template customizations prior to deploying the Overcloud. Red Hat OpenStack Platform director supports several network interface types including physical interfaces, bonded interfaces and either tagged or native 802.1Q VLANs.

## Network Types by Server Role

Server role was discussed in the previous section. Each server role requires access to specific types of network traffic. The network isolation feature allows Red Hat OpenStack Platform director to segment network traffic by particular network types. When using network isolation, each server role must have access to its required network traffic types.

By default, Red Hat OpenStack Platform director collapses all network traffic to the provisioning interface. This configuration is suitable for evaluation, proof of concept, and development environments. It is not recommended for production environments where scaling and performance are a primary concern.

## Tenant Network Types

Red Hat OpenStack Platform 8 supports tenant network communication through the OpenStack Networking (Neutron) service. OpenStack Networking supports overlapping IP address ranges across tenants through the Linux kernel’s network namespace capability. It also supports three default networking types:

### VLAN segmentation mode

Each tenant is assigned a network subnet mapped to an 802.1q VLAN on the physical network. This tenant networking type requires VLAN-assignment to the appropriate switch ports on the physical network.

### VXLAN segmentation mode

In the VXLAN mechanism driver encapsulates each layer 2 Ethernet frame sent by the VMs in a layer 3 UDP packet. The UDP packet includes an 8-byte field, within which a 24-bit value is used for the VXLAN Segment ID. The VXLAN Segment ID is used to designate the individual VXLAN overlay network on which the communicating VMs are housed. This provides segmentation for each Tenant network.

### GRE segmentation mode

In the GRE mechanism driver encapsulates each layer 2 Ethernet frame sent by the VMs in a special IP packet using the GRE protocol (IP type 47). The GRE header contains a 32-bit *key* which is used to identify a flow or virtual network in a tunnel. This provides segmentation for each Tenant network.

Cisco Nexus Plugin is bundled in OpenStack Platform 8, Liberty release. While it can support both VLAN and VXLAN configurations, only VLAN mode is validated as part of this design. VXLAN will be considered in future releases when the current Cisco VIC 1340 adapter will be certified on VXLAN and Red Hat operating system.

## Cluster Manager and Proxy Server

Two components drive HA for all core and non-core OpenStack services:

- Cluster Manager
- Proxy Server

The cluster manager is responsible for the startup and recovery of an inter-related services across a set of physical machines. It tracks the cluster's internal state across multiple machines. State changes trigger appropriate responses from the cluster manager to ensure service availability and data integrity.

This section describes the steps to configure network for Overcloud. The network setup used in the configuration is shown in Figure 4.

The configuration is done using Heat Templates on the Undercloud prior to deploying the Overcloud. These steps need to be followed after the Undercloud installation. In order to use network isolation, we have to define the Overcloud networks. Each will have an IP subnet, a range of IP addresses to use on the subnet and a VLAN ID. These parameters will be defined in the network environment file. In addition to the global settings there is a template for each of the nodes like controller, compute and Ceph that determines the NIC configuration for each role. These have to be customized to match the actual hardware configuration.

Heat communicates with Neutron API running on the Undercloud node to create isolated networks and to assign neutron ports on these networks. Neutron will assign a static IP to each port and Heat will use these static IPs to configure networking on the Overcloud nodes. A utility called os-net-config runs on each node at the time of provisioning to configure host level networking.

Table 4 lists the VLANs that are created on the configuration.

**Table 4 VLANs**

| VLAN Name | VLAN Purpose | VLAN ID or VLAN Range Used in This Design for Reference |
|---|---|---|
| Management | Management Network to UCSM and Nexus Switches | 10 |
| PXE | Provisioning Network VLAN | 110 |
| Internal-API | Internal API Network | 100 |
| External | External Network | 215 |
| Storage Public | Storage Public Network | 120 |
| Storage Management | Storage Cluster or Management Network | 150 |
| Floating | Floating Network | 160 |

# High Availability

Red Hat OpenStack Platform director's approach is to leverage Red Hat's distributed cluster system.

## Cluster Manager and Proxy Server

The Cluster Manager is responsible for the startup and recovery of an inter-related services across a set of physical machines. It tracks the cluster's internal state across multiple machines. State changes trigger appropriate responses from the cluster manager to ensure service availability and data integrity.

In the HA model Clients do not directly connect to service endpoints. Connection requests are routed to service endpoints by a proxy server.

The Cluster Manager provides state awareness of other machines to coordinate service startup and recovery, shared quorum to determine majority set of surviving cluster nodes after failure, data integrity through fencing and automated recovery of failed instances.

The Proxy servers help with load balancing connections across service end points. The nodes can be added or removed without interrupting service.

Red Hat OpenStack Platform director uses HAproxy and Pacemaker to manage HA services and load balance connection requests. With the exception of RabbitMQ and Galera, HAproxy distributes connection requests to active nodes in a round-robin fashion. Galera and RabbitMQ use persistent options to ensure requests go only to active and/or synchronized nodes. Pacemaker checks service health at every one second interval. Timeout settings vary by service.

The combination of Pacemaker and HAProxy:

- Detects and recovers machine and application failures
- Starts and stops OpenStack services in the correct order
- Responds to cluster failures with appropriate actions including resource failover and machine restart and fencing

RabbitMQ, memcached, and mongodb do not use HAProxy server. These services have their own failover and HA mechanisms.

# Cisco ML2 Plugins

OpenStack Modular Layer 2 (ML2) allows the separation of network segment types and the device specific implementation of segment types. ML2 architecture consists of multiple 'type drivers' and 'mechanism drivers'. Type drivers manage the common aspects of a specific type of network while the mechanism driver manages specific device to implement network types.

Type drivers:

- VLAN
- GRE
- VXLAN

Mechanism drivers:

- Cisco UCS Manager
- Cisco Nexus
- Openvswitch, Linuxbridge

The Cisco Nexus driver for OpenStack Neutron allows customers to easily build their Infrastructure-as-a-Service (IaaS) networks using the industry's leading networking platform, delivering performance, scalability, and stability with the familiar manageability and control you expect from Cisco® technology. ML2 Nexus drivers dynamically provision OpenStack managed VLANs on Cisco Nexus switches. They configure the trunk ports with dynamically created VLANs solving the logical port count issue on the Nexus switches. They provide better manageability of the network infrastructure.

ML2 Cisco UCS Manager drivers dynamically provision OpenStack managed VLANs on Fabric Interconnects. They configure VLANs on Controller and Compute node VNICs. The Cisco UCS Manager Plugin talks to the Cisco UCS Manager application running on Fabric Interconnect and is part of an ecosystem for Cisco UCS Servers that consists of Fabric Interconnects and IO modules. The ML2 Cisco UCS Manager driver does not support configuration of Cisco UCS Servers, whose service profiles are attached to Service Templates. This is to prevent that same VLAN configuration to be pushed to all the service profiles based on that template. The plugin can be used after the Service Profile has been unbound from the template.

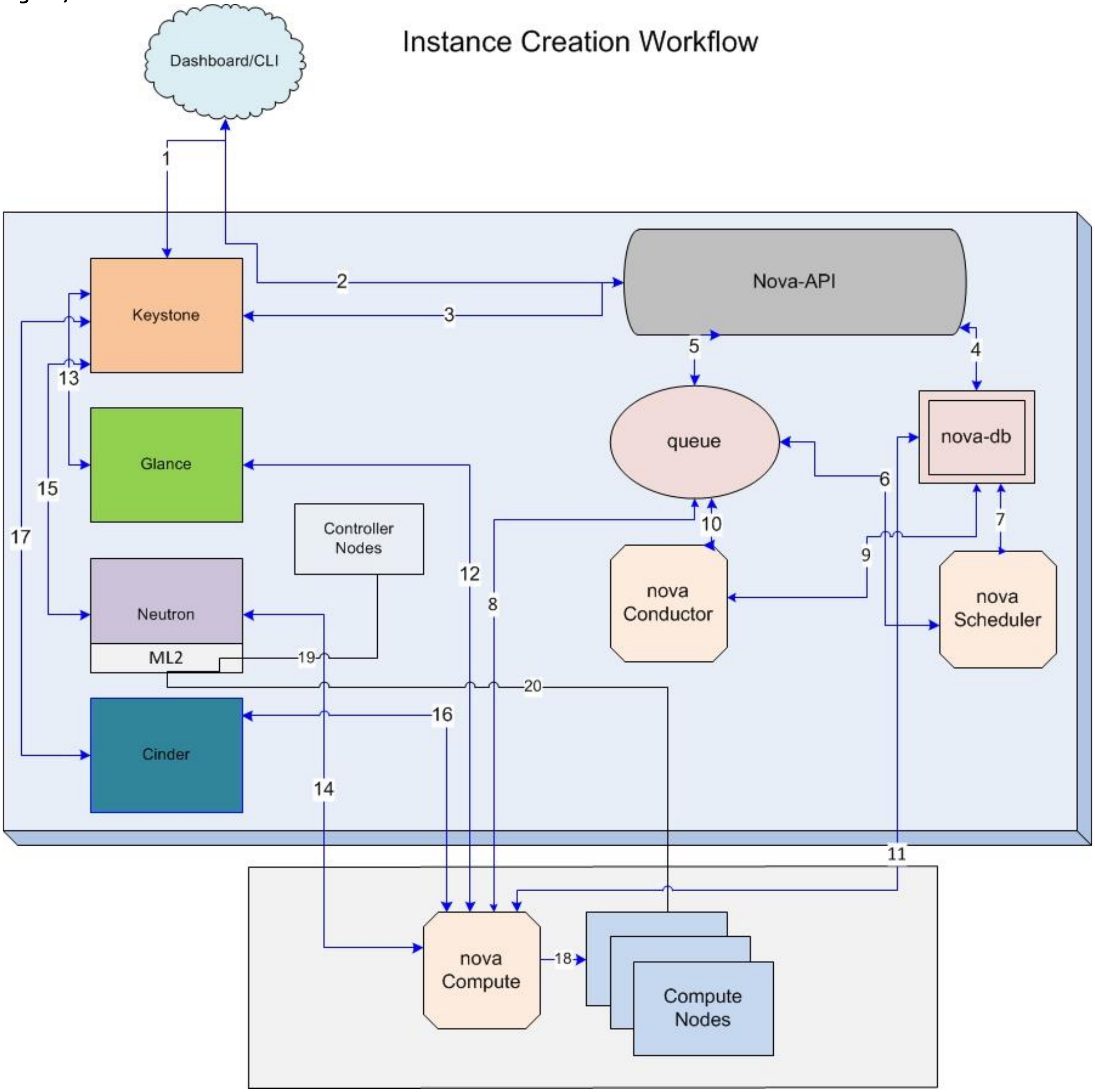
## Instance creation work flow

To create a virtual machine, complete the following steps:

1. Dashboard/CLI authenticates with Keystone.
2. Dashboard/CLI sends nova-boot to nova-api.
3. nova-api validates the token with keystone.
4. nova-api checks for conflicts, if not creates a new entry in database.
5. nova-api sends rpc.call to nova-scheduler and gets updated host-entry with host-id.
6. nova-scheduler picks up the request from the queue.
7. nova-scheduler sends the rpc.cast request to nova-compute for launching an instance on the appropriate host after applying filters.
8. nova-compute picks up the request from the queue.
9. nova-compute sends the rpc.call request to nova-conductor to fetch the instance information such as host ID and flavor (RAM, CPU, and Disk).
10. nova-conductor picks up the request from the queue.
11. nova-conductor interacts with nova-database and picks up instance information from queue.
12. nova-compute performs the REST with auth-token to glance-api. Then, nova-compute retrieves the Image URI from the Image Service, and loads the image from the image storage.
13. glance-api validates the auth-token with keystone and nova-compute gets the image data.
14. nova-compute performs the REST call to network API to allocated and configure the network
15. neutron server validates the token and creates network info.
16. Nova-compute performs REST to volume API to attach volume to the instance.
17. Cinder-api validates the token and provides block storage info to nova-compute.
18. Nova compute generates data for the hypervisor driver.

19. DHCP and/or Router port bindings by neutron on controller nodes triggers Cisco ML2 plugins:
    a. Cisco UCS Manager driver creates VLAN and trunks the eth1 vNICs for the controller node's service-profile
    b. Nexus driver creates VLAN and trunks the switch port(s) mapped to the controller node
20. Virtual Machine Instance's Port bindings to a Compute Node triggers again ML2:
    a. Cisco UCS Manager driver creates VLAN and trunks the eth1 vNICs for the compute node's service-profile
    b. Nexus driver creates VLAN and trunks the switches' port(s) mapped to the compute node

**Figure 7** Instance Creation Workflow



# Deployment Hardware

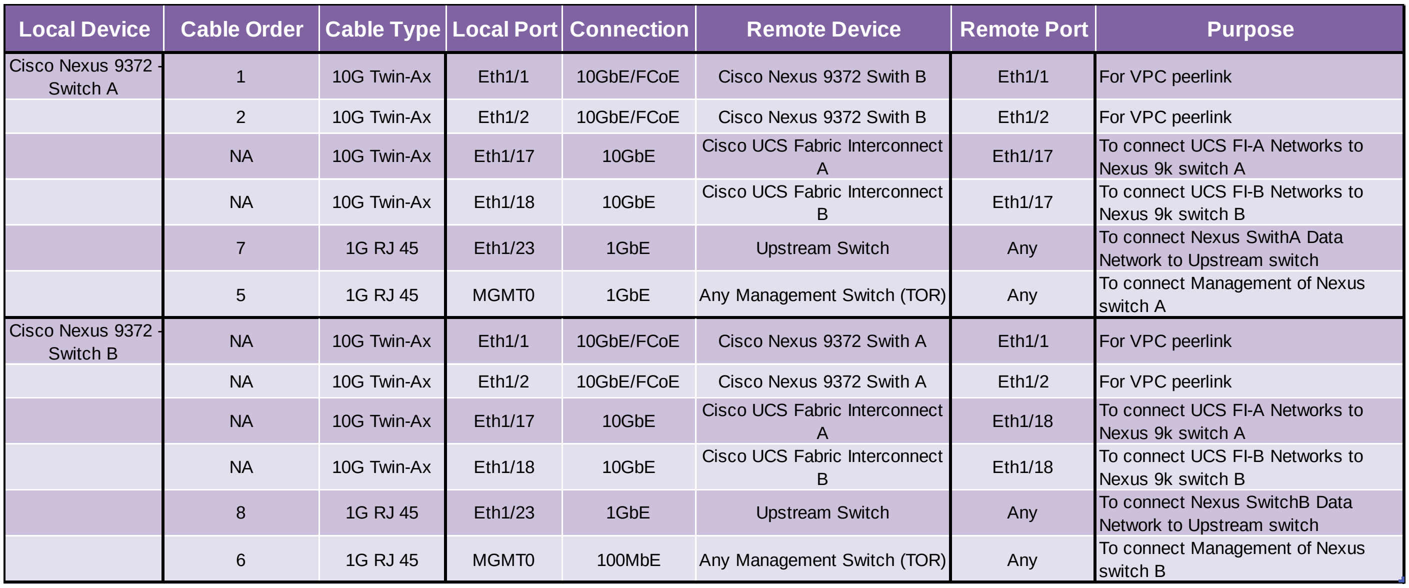
This section details the deployment hardware used in this solution.

## Cabling Details

Table 5 lists the cabling information.

Table 5 Cabling Details

| Local Device | Cable Order | Cable Type | Local Port | Connection | Remote Device | Remote Port | Purpose |
|---|---|---|---|---|---|---|---|
| Cisco UCS Fabric Interconnect A | 1 | 10G Twin-Ax | Eth1/1 | 10GbE/FCoE | Chassis 1 FEX A (left) | port 1 | To connect UCS chassis1 to UCS Fabric InterconnectA |
| | 2 | 10G Twin-Ax | Eth1/2 | 10GbE/FCoE | Chassis 1 FEX A (left) | port 2 | To connect UCS chassis1 to UCS Fabric InterconnectA |
| | 3 | 10G Twin-Ax | Eth1/3 | 10GbE/FCoE | Chassis 1 FEX A (left) | port 3 | To connect UCS chassis1 to UCS Fabric InterconnectA |
| | 4 | 10G Twin-Ax | Eth1/4 | 10GbE/FCoE | Chassis 1 FEX A (left) | port 4 | To connect UCS chassis1 to UCS Fabric InterconnectA |
| | 5 | 10G Twin-Ax | Eth1/5 | 10GbE/FCoE | Chassis 2 FEX A (left) | port 1 | To connect UCS chassis2 to UCS Fabric InterconnectA |
| | 6 | 10G Twin-Ax | Eth1/6 | 10GbE/FCoE | Chassis 2 FEX A (left) | port 2 | To connect UCS chassis2 to UCS Fabric InterconnectA |
| | 7 | 10G Twin-Ax | Eth1/7 | 10GbE/FCoE | Chassis 2 FEX A (left) | port 3 | To connect UCS chassis2 to UCS Fabric InterconnectA |
| | 8 | 10G Twin-Ax | Eth1/8 | 10GbE/FCoE | Chassis 2 FEX A (left) | port 4 | To connect UCS chassis2 to UCS Fabric InterconnectA |
| | 9 | 10G Twin-Ax | Eth1/9 | 10GbE/FCoE | C240 M4 - Server1 - VIC1227 | Port 1 | To connect UCS C240 Srv1 to UCS Fabric InterconnectA |
| | 10 | 10G Twin-Ax | Eth1/10 | 10GbE/FCoE | C240 M4 - Server2 - VIC1227 | Port 1 | To connect UCS C240 Srv2 to UCS Fabric InterconnectA |
| | 11 | 10G Twin-Ax | Eth1/11 | 10GbE/FCoE | C240 M4 - Server3 - VIC1227 | Port 1 | To connect UCS C240 Srv3 to UCS Fabric InterconnectA |
| | 3 | 10G Twin-Ax | Eth1/17 | 10GbE/FCoE | Nexus 9372 Switch A | Eth 1/17 | To connect UCS FI-A Networks to Nexus 9k switch A |
| | 4 | 10G Twin-Ax | Eth1/18 | 10GbE/FCoE | Nexus 9372 Switch B | Eth 1/17 | To connect UCS FI-B Networks to Nexus 9k switch B |
| | 1 | 1G RJ 45 | MGMT0 | 1GbE | Any Management Switch (TOR) | Any | To Connect Management of UCS Fabric Interconnect |
| | 2 | 1G RJ 45 | L1 | 1GbE | UCS Fabric Interconnect B | L1 | Cluster connection between UCS FIs. |
| | 3 | 1G RJ 45 | L2 | 1GbE | UCS Fabric Interconnect B | L2 | Cluster connection between UCS FIs. |
| Cisco UCS Fabric Interconnect B | 12 | 10G Twin-Ax | Eth1/1 | 10GbE/FCoE | Chassis 1 FEX B (Right) | port 1 | To connect UCS chassis1 to UCS Fabric InterconnectB |
| | 13 | 10G Twin-Ax | Eth1/2 | 10GbE/FCoE | Chassis 1 FEX B (Right) | port 2 | To connect UCS chassis1 to UCS Fabric InterconnectB |
| | 14 | 10G Twin-Ax | Eth1/3 | 10GbE/FCoE | Chassis 1 FEX B (Right) | port 3 | To connect UCS chassis1 to UCS Fabric InterconnectB |
| | 15 | 10G Twin-Ax | Eth1/4 | 10GbE/FCoE | Chassis 1 FEX B (Right) | port 4 | To connect UCS chassis1 to UCS Fabric InterconnectB |
| | 16 | 10G Twin-Ax | Eth1/5 | 10GbE/FCoE | Chassis 2 FEX B (Right) | port 1 | To connect UCS chassis2 to UCS Fabric InterconnectB |
| | 17 | 10G Twin-Ax | Eth1/6 | 10GbE/FCoE | Chassis 2 FEX B (Right) | port 2 | To connect UCS chassis2 to UCS Fabric InterconnectB |
| | 18 | 10G Twin-Ax | Eth1/7 | 10GbE/FCoE | Chassis 2 FEX B (Right) | port 3 | To connect UCS chassis2 to UCS Fabric InterconnectB |
| | 19 | 10G Twin-Ax | Eth1/8 | 10GbE/FCoE | Chassis 2 FEX B (Right) | port 4 | To connect UCS chassis2 to UCS Fabric InterconnectB |
| | 20 | 10G Twin-Ax | Eth1/9 | 10GbE/FCoE | C240 M4 - Server1 - VIC1227 | Port 2 | To connect UCS C240 Srv1 to UCS Fabric InterconnectB |
| | 21 | 10G Twin-Ax | Eth1/10 | 10GbE/FCoE | C240 M4 - Server2 - VIC1227 | Port 2 | To connect UCS C240 Srv2 to UCS Fabric InterconnectB |
| | 22 | 10G Twin-Ax | Eth1/11 | 10GbE/FCoE | C240 M4 - Server3 - VIC1227 | Port 2 | To connect UCS C240 Srv3 to UCS Fabric InterconnectB |
| | 5 | 10G Twin-Ax | Eth1/17 | 10GbE/FCoE | Nexus 9372 Switch A | Eth 1/18 | To connect UCS FI-A Networks to Nexus 9k switch A |
| | 6 | 10G Twin-Ax | Eth1/18 | 10GbE/FCoE | Nexus 9372 Switch B | Eth 1/18 | To connect UCS FI-B Networks to Nexus 9k switch B |
| | 4 | 1G RJ 45 | MGMT0 | 1GbE | Any Management Switch (TOR) | Any | To Connect Management of UCS Fabric Interconnect |
| | NA | 1G RJ 45 | L1 | 1GbE | UCS Fabric Interconnect A | L1 | Cluster connection between UCS FIs. |
| | NA | 1G RJ 45 | L2 | 1GbE | UCS Fabric Interconnect A | L2 | Cluster connection between UCS FIs. |

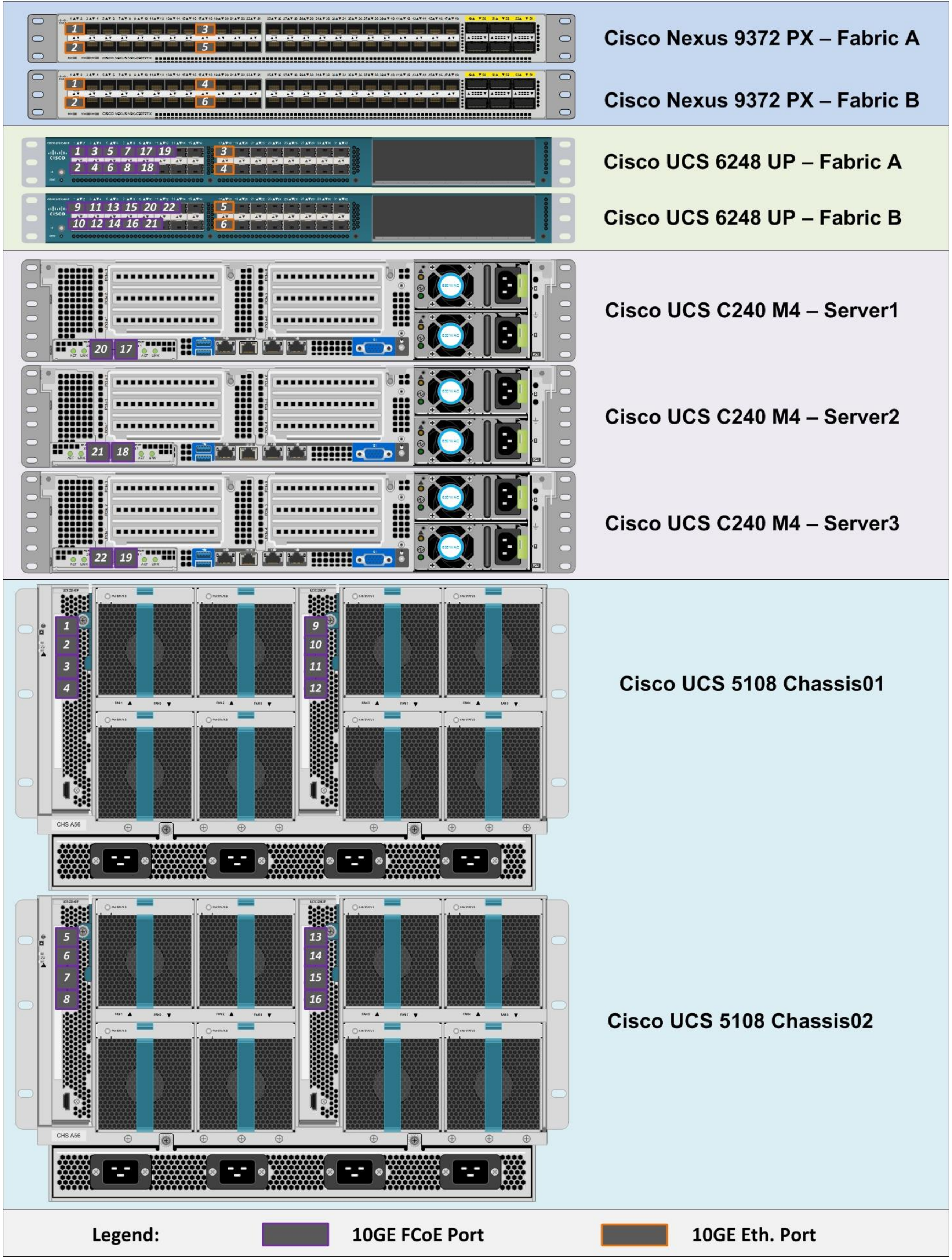


| Local Device | Cable Order | Cable Type | Local Port | Connection | Remote Device | Remote Port | Purpose |
|---|---|---|---|---|---|---|---|
| Cisco Nexus 9372 - Switch A | 1 | 10G Twin-Ax | Eth1/1 | 10GbE/FCoE | Cisco Nexus 9372 Swith B | Eth1/1 | For VPC peerlink |
| | 2 | 10G Twin-Ax | Eth1/2 | 10GbE/FCoE | Cisco Nexus 9372 Swith B | Eth1/2 | For VPC peerlink |
| | NA | 10G Twin-Ax | Eth1/17 | 10GbE | Cisco UCS Fabric Interconnect A | Eth1/17 | To connect UCS FI-A Networks to Nexus 9k switch A |
| | NA | 10G Twin-Ax | Eth1/18 | 10GbE | Cisco UCS Fabric Interconnect B | Eth1/17 | To connect UCS FI-B Networks to Nexus 9k switch B |
| | 7 | 1G RJ 45 | Eth1/23 | 1GbE | Upstream Switch | Any | To connect Nexus SwithA Data Network to Upstream switch |
| | 5 | 1G RJ 45 | MGMT0 | 1GbE | Any Management Switch (TOR) | Any | To connect Management of Nexus switch A |
| Cisco Nexus 9372 - Switch B | NA | 10G Twin-Ax | Eth1/1 | 10GbE/FCoE | Cisco Nexus 9372 Swith A | Eth1/1 | For VPC peerlink |
| | NA | 10G Twin-Ax | Eth1/2 | 10GbE/FCoE | Cisco Nexus 9372 Swith A | Eth1/2 | For VPC peerlink |
| | NA | 10G Twin-Ax | Eth1/17 | 10GbE | Cisco UCS Fabric Interconnect A | Eth1/18 | To connect UCS FI-A Networks to Nexus 9k switch A |
| | NA | 10G Twin-Ax | Eth1/18 | 10GbE | Cisco UCS Fabric Interconnect B | Eth1/18 | To connect UCS FI-B Networks to Nexus 9k switch B |
| | 8 | 1G RJ 45 | Eth1/23 | 1GbE | Upstream Switch | Any | To connect Nexus SwitchB Data Network to Upstream switch |
| | 6 | 1G RJ 45 | MGMT0 | 100MbE | Any Management Switch (TOR) | Any | To connect Management of Nexus switch B |

## Physical Cabling

Figure 8 illustrates the physical cabling used in this solution.

**Figure 8** Physical Cabling


Cisco Nexus 9372 PX – Fabric A
1
2
3
5
Cisco Nexus 9372 PX – Fabric B
1
2
4
6
Cisco UCS 6248 UP – Fabric A
1 3 5 7 17 19
2 4 6 8 18
3
4
Cisco UCS 6248 UP – Fabric B
9 11 13 15 20 22
10 12 14 16 21
5
6
Cisco UCS C240 M4 – Server1
20 17
Cisco UCS C240 M4 – Server2
21 18
Cisco UCS C240 M4 – Server3
22 19
Cisco UCS 5108 Chassis01
1
2
3
4
9
10
11
12
Cisco UCS 5108 Chassis02
5
6
7
8
13
14
15
16
Legend:
10GE FCoE Port
10GE Eth. Port

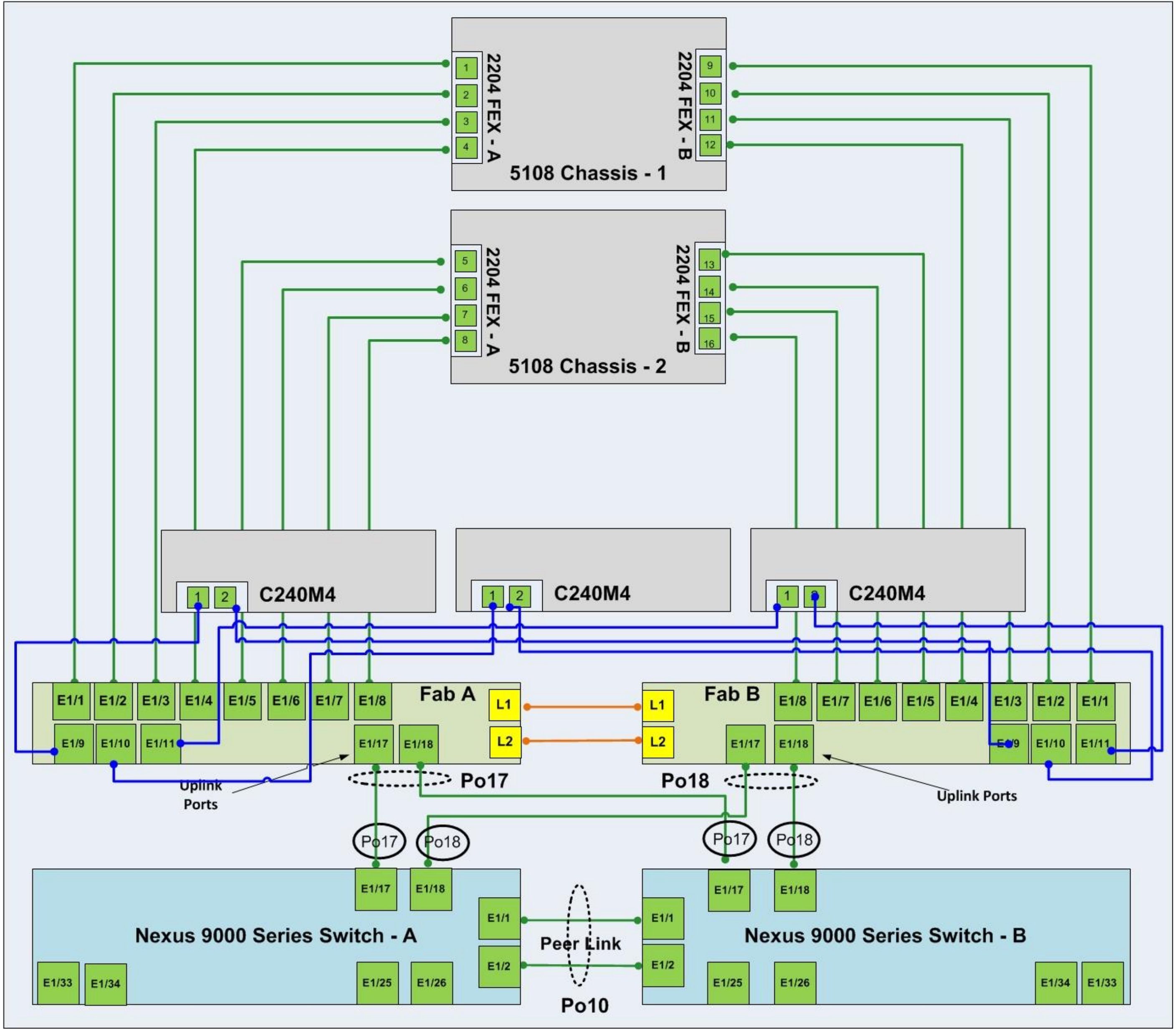
Please note the port numbers on VIC1227 card. As shown in the figure Port 1 is on the right and Port 2 is on the left.

http://www.cisco.com/c/dam/en/us/products/interfaces-modules/ucs-virtual-interface-card-1227/kO71144-large.jpg

## Cabling Logic

Figure 9 illustrates the cabling logic used in this solution.

**Figure 9** Cabling Logic



## Cisco UCS Configuration

### Configure Cisco UCS Fabric Interconnects

Configure the Fabric Interconnects after the cabling is complete. To hook up the console port on the Fabrics, complete the following steps:

Please replace the appropriate addresses for your setup.

#### Cisco UCS 6248UP Switch A

Connect the console port to the UCS 6248 Fabric Interconnect switch designated for Fabric A:

```
Enter the configuration method: console
Enter the setup mode; setup newly or restore from backup.(setup/restore)? setup
```

```
You have chosen to setup a new fabric interconnect? Continue? (y/n): y
Enforce strong passwords? (y/n) [y]: y
Enter the password for "admin": <password>
Enter the same password for "admin": <password>
Is this fabric interconnect part of a cluster (select 'no' for standalone)?
(yes/no) [n]:y
Which switch fabric (A|B): A
Enter the system name: UCS-6248-FAB
Physical switch Mgmt0 IPv4 address: 10.23.10.6
Physical switch Mgmt0 IPv4 netmask: 255.255.255.0
IPv4 address of the default gateway: 10.23.10.1
Cluster IPv4 address: 10.23.10.5
Configure DNS Server IPv4 address? (yes/no) [no]: y
DNS IPv4 address: <<var_nameserver_ip>>
Configure the default domain name? y
Default domain name: <<var_dns_domain_name>>
Join centralized management environment (UCS Central)? (yes/no) [n]: Press Enter
You will be prompted to review the settings.
If they are correct, answer yes to apply and save the configuration. Wait for the
login prompt to make sure that the configuration has been saved.
```

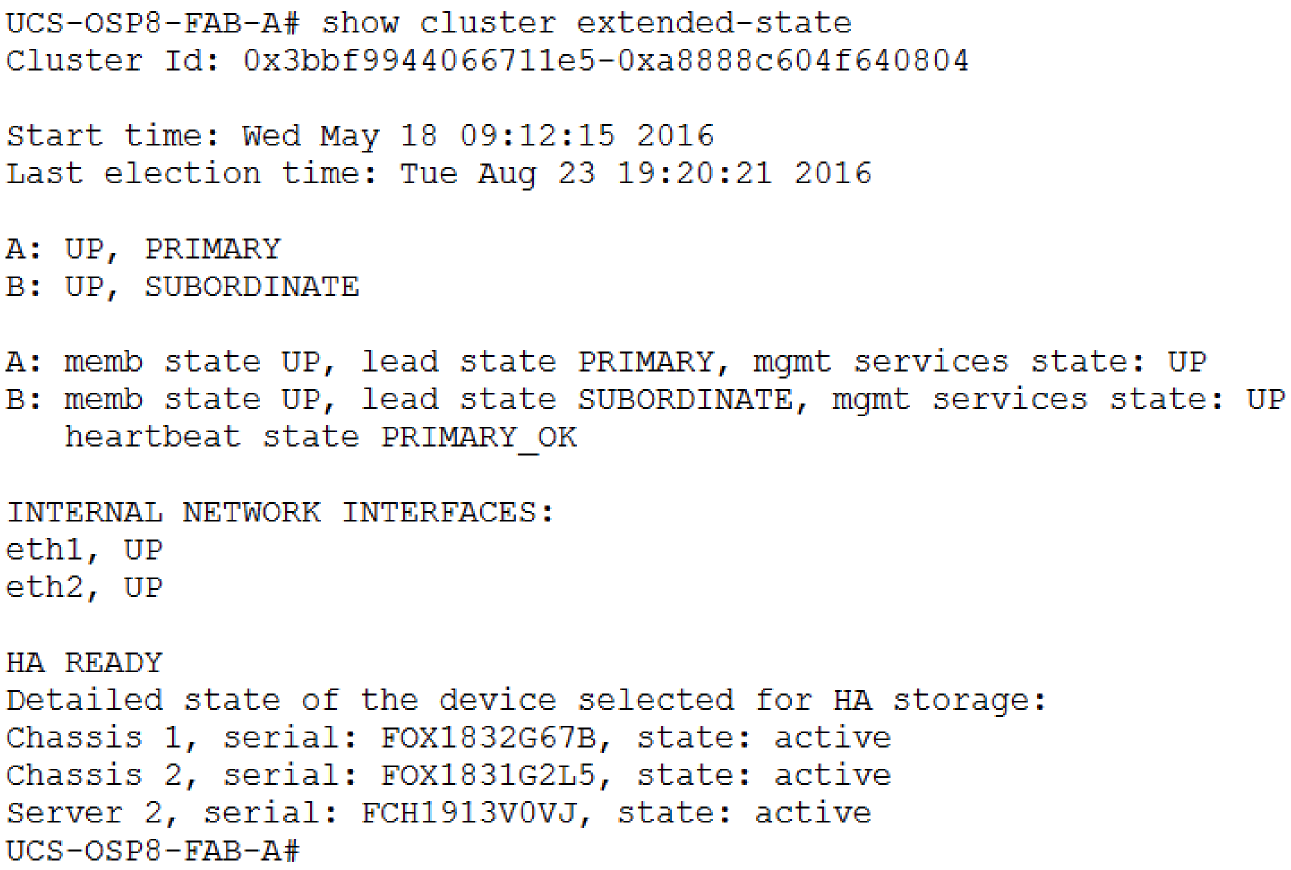
### Cisco UCS 6248UP Switch B

Connect the console port to Peer UCS 6248 Fabric Interconnect switch designated for Fabric B:

```
Enter the configuration method: console
Installer has detected the presence of a peer Fabric interconnect. This Fabric
interconnect will be added to the cluster. Do you want to continue {y|n}? y
Enter the admin password for the peer fabric interconnect: <password>
Physical switch Mgmt0 IPv4 address: 10.23.10.7
Apply and save the configuration (select “no” if you want to re-enter)? (yes/no):
yes
```

Verify the connectivity:

After completing the FI configuration, verify the connectivity as detailed below by logging to one of the Fabrics or the VIP address and checking the cluster state or extended state as shown below:



```
UCS-OSP8-FAB-A# show cluster extended-state
Cluster Id: 0x3bbf9944066711e5-0xa8888c604f640804

Start time: Wed May 18 09:12:15 2016
Last election time: Tue Aug 23 19:20:21 2016

A: UP, PRIMARY
B: UP, SUBORDINATE

A: memb state UP, lead state PRIMARY, mgmt services state: UP
B: memb state UP, lead state SUBORDINATE, mgmt services state: UP
   heartbeat state PRIMARY_OK

INTERNAL NETWORK INTERFACES:
eth1, UP
eth2, UP

HA READY
Detailed state of the device selected for HA storage:
Chassis 1, serial: FOX1832G67B, state: active
Chassis 2, serial: FOX1831G2L5, state: active
Server 2, serial: FCH1913V0VJ, state: active
UCS-OSP8-FAB-A#
```

## Configure the Cisco UCS Global Policies

To configure the Global policies, log into UCS Manager GUI, and complete the following steps:

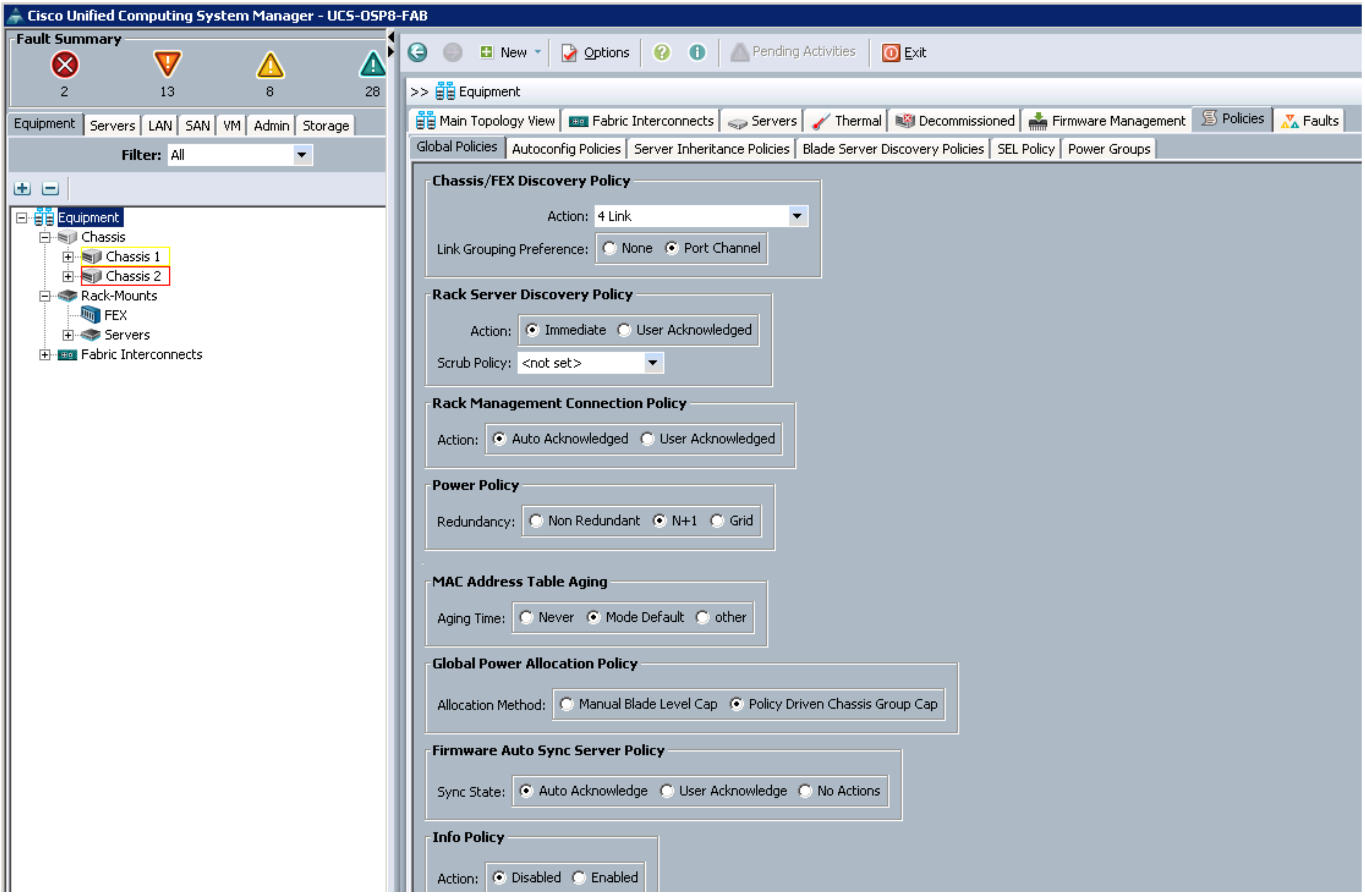
1. Under Equipment → Global Policies;

    a. Set the Chassis/FEX Discovery Policy to match the number of uplink ports that are cabled between the chassis or fabric extenders and to the fabric interconnects.

    b. Set the Power policy based on the input power supply to the UCS chassis. In general, UCS chassis with 5 or more blades recommends minimum of 3 power supplies with N+1 configuration. With 4 power supplies, 2 on each PDUs the recommended power policy is Grid.

    c. Set the Global Power allocation Policy as Policy driven Chassis Group cap.

    d. Click Save changes to save the configuration.



**Chassis/FEX Discovery Policy**

Action: 4 Link

Link Grouping Preference: ○ None ● Port Channel

**Rack Server Discovery Policy**

Action: ● Immediate ○ User Acknowledged

Scrub Policy: <not set>

**Rack Management Connection Policy**

Action: ● Auto Acknowledged ○ User Acknowledged

**Power Policy**

Redundancy: ○ Non Redundant ● N+1 ○ Grid

**MAC Address Table Aging**

Aging Time: ○ Never ● Mode Default ○ other

**Global Power Allocation Policy**

Allocation Method: ○ Manual Blade Level Cap ● Policy Driven Chassis Group Cap

**Firmware Auto Sync Server Policy**

Sync State: ● Auto Acknowledge ○ User Acknowledge ○ No Actions

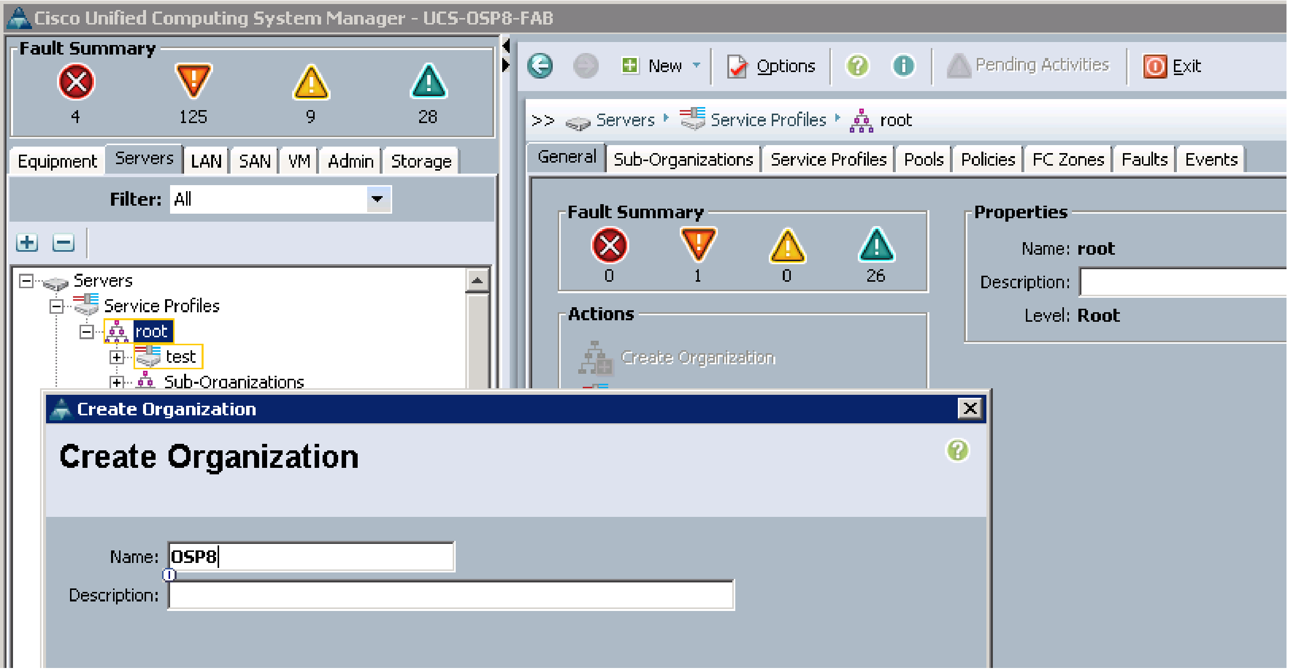
**Info Policy**

Action: ● Disabled ○ Enabled

## Configure Sub-Orgs

In case you wish to have sub-orgs in UCS, create sub-orgs as below:

Under Servers Tab -> Service Profiles -> root -> Sub-Organizations -> Create Organization
Enter the Organization Name of your choice as below and click OK to continue.



Properties

Name: root

Description:

Level: Root

Create Organization

Name: OSP8

Description:

## Configure Server Ports for Blade Discovery and Rack Discovery

Navigate to each fabric interconnect and configure the server ports on the fabric interconnects. Complete the following steps:

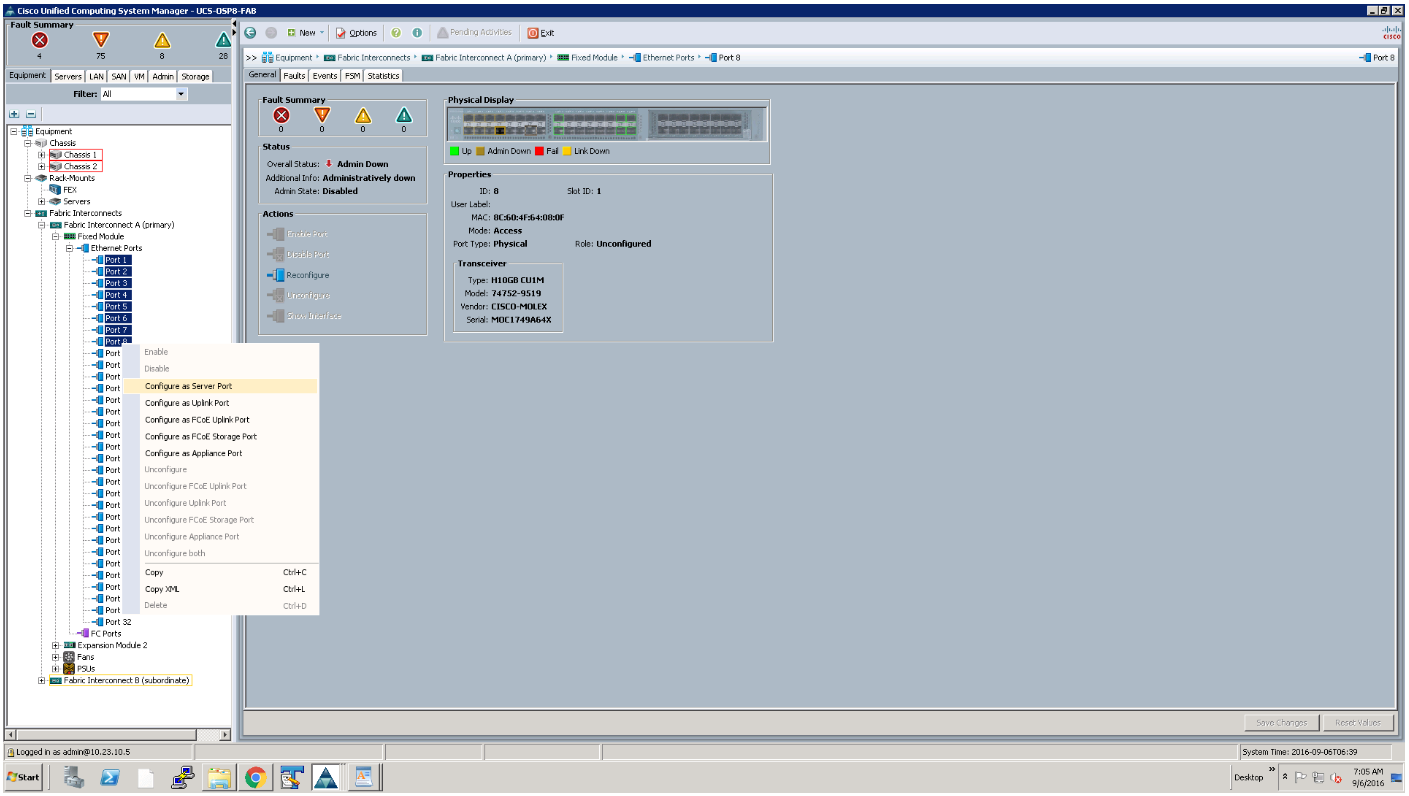
Under Equipment → Fabric Interconnects → Fabric Interconnect A → Fixed Module → Ethernet Ports:

Select the ports (Port 1 to 8) that are connected to the left side of each UCS chassis FEX 2204, right-click them and select Configure as Server Port.

Select the ports (Port 9 to 11) that are connected to the 10G MLOM (VIC1227) port1 of each UCS C240 M4, right-click them, and select Configure as Server Port.

Click Save Changes to save the configuration.

Repeat steps 1 and 2 on Fabric Interconnect B and save the configuration.



Status

Overall Status: Admin Down
Additional Info: Administratively down
Admin State: Disabled

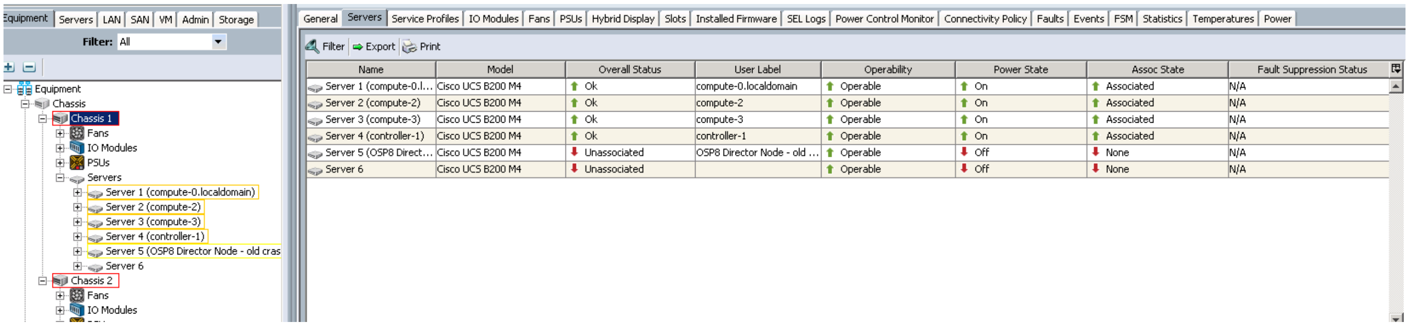
Properties

ID: 8 Slot ID: 1
User Label:
MAC: 8C:60:4F:64:08:0F
Mode: Access
Port Type: Physical Role: Unconfigured

Transceiver

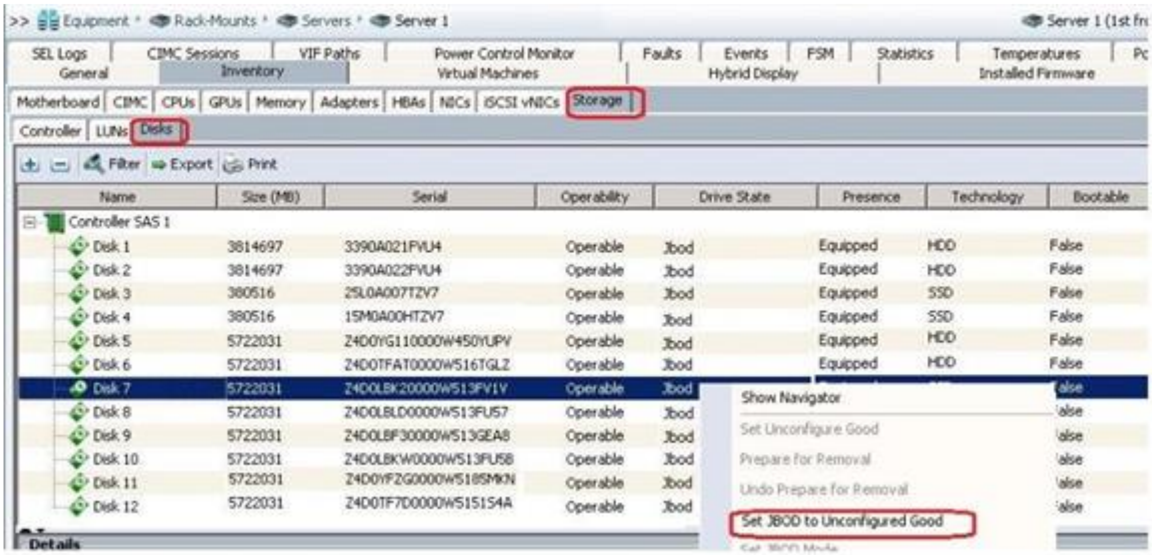
Type: H10GB CU1M
Model: 74752-9519
Vendor: CISCO-MOLEX
Serial: MOC1749A64X

The blades and rack servers are discovered as shown below:



| Name | Model | Overall Status | User Label | Operability | Power State | Assoc State | Fault Suppression Status |
|---|---|---|---|---|---|---|---|
| Server 1 (compute-0.l... | Cisco UCS B200 M4 | Ok | compute-0.localdomain | Operable | On | Associated | N/A |
| Server 2 (compute-2) | Cisco UCS B200 M4 | Ok | compute-2 | Operable | On | Associated | N/A |
| Server 3 (compute-3) | Cisco UCS B200 M4 | Ok | compute-3 | Operable | On | Associated | N/A |
| Server 4 (controller-1) | Cisco UCS B200 M4 | Ok | controller-1 | Operable | On | Associated | N/A |
| Server 5 (OSP8 Direct... | Cisco UCS B200 M4 | Unassociated | OSP8 Director Node - old ... | Operable | Off | None | N/A |
| Server 6 | Cisco UCS B200 M4 | Unassociated | | Operable | Off | None | N/A |

Navigate to each blade and rack servers to make sure that the disks are in Unconfigured Good state; if not, convert JBOD to Unconfigured as shown in the screen shot below. The screen shot illustrates how to convert a disk to the Unconfigured Good state.


>> Equipment ▸ Rack-Mounts ▸ Servers ▸ Server 1
Server 1 (1st fr
SEL Logs | CIMC Sessions | VIF Paths | Power Control Monitor | Faults | Events | FSM | Statistics | Temperatures | Pc
General | Inventory | Virtual Machines | Hybrid Display | Installed Firmware
Motherboard | CIMC | CPUs | GPUs | Memory | Adapters | HBAs | NICs | iSCSI vNICs | Storage
Controller | LUNs | Disks
Filter | Export | Print
Show Navigator
Set Unconfigure Good
Prepare for Removal
Undo Prepare for Removal
Set JBOD to Unconfigured Good
Set JBOD Mode
Details

| Name | Size (MB) | Serial | Operability | Drive State | Presence | Technology | Bootable |
|---|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | | |
| Disk 1 | 3814697 | 3390A021FVU4 | Operable | Jbod | Equipped | HDD | False |
| Disk 2 | 3814697 | 3390A022FVU4 | Operable | Jbod | Equipped | HDD | False |
| Disk 3 | 380516 | 25L0A007TZV7 | Operable | Jbod | Equipped | SSD | False |
| Disk 4 | 380516 | 15M0A00HTZV7 | Operable | Jbod | Equipped | SSD | False |
| Disk 5 | 5722031 | Z4D0YG110000W450YUPV | Operable | Jbod | Equipped | HDD | False |
| Disk 6 | 5722031 | Z4D0TFAT0000W516TGLZ | Operable | Jbod | Equipped | HDD | False |
| Disk 7 | 5722031 | Z4D0LBK20000W513FV1V | Operable | Jbod | | | alse |
| Disk 8 | 5722031 | Z4D0LBLD0000W513FU57 | Operable | Jbod | | | alse |
| Disk 9 | 5722031 | Z4D0LBF30000W513GEA8 | Operable | Jbod | | | alse |
| Disk 10 | 5722031 | Z4D0LBKW0000W513FU58 | Operable | Jbod | | | alse |
| Disk 11 | 5722031 | Z4D0YFZG0000W518SMKN | Operable | Jbod | | | alse |
| Disk 12 | 5722031 | Z4D0TF7D0000W51515 4A | Operable | Jbod | | | alse |

## Configure Network Uplinks

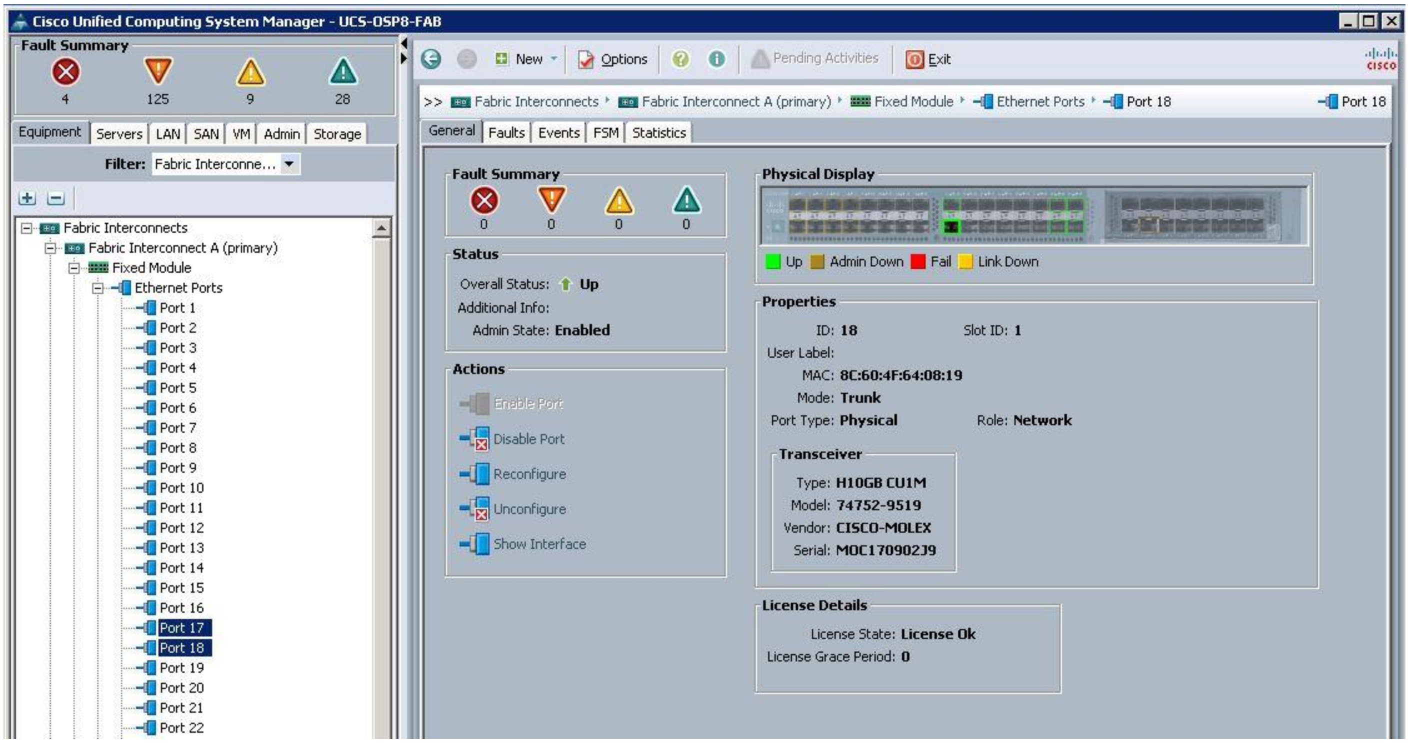
Navigate to each Fabric Interconnects and configure the Network Uplink ports on Fabric Interconnects. Complete the following steps:

Under Equipment → Fabric Interconnects → Fabric Interconnect A → Fixed Module → Ethernet Ports:

Select the port 17 and Port18 that are connected to Nexus 9k switches, right-click them and select Configure as Uplink Port.

Click Save Changes to save the configuration.

Repeat the steps 1 and 2 on Fabric Interconnect B.



Status

Overall Status: Up

Additional Info:

Admin State: **Enabled**

Properties

ID: **18** Slot ID: **1**

User Label:

MAC: **8C:60:4F:64:08:19**

Mode: **Trunk**

Port Type: **Physical** Role: **Network**

Transceiver

Type: **H10GB CU1M**

Model: **74752-9519**

Vendor: **CISCO-MOLEX**

Serial: **MOC17090239**

License Details

License State: **License Ok**

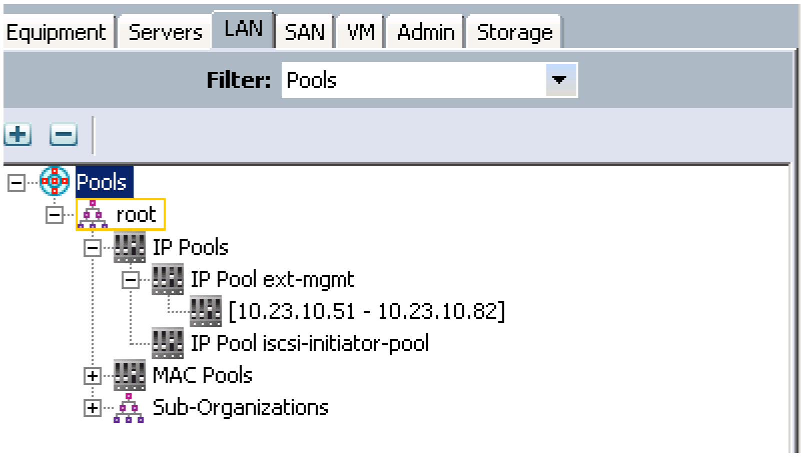
License Grace Period: **0**

## Create KVM IP Pools

To access the KVM console of each UCS Server, create the KVM IP pools from the UCS Manager GUI, and complete the following steps:

Under LAN → Pools → root → IP Pools → IP Pool ext-mgmt → right-click and select Create Block of IPV4 addresses.

Specify the Starting IP address, subnet mask and gateway and size.



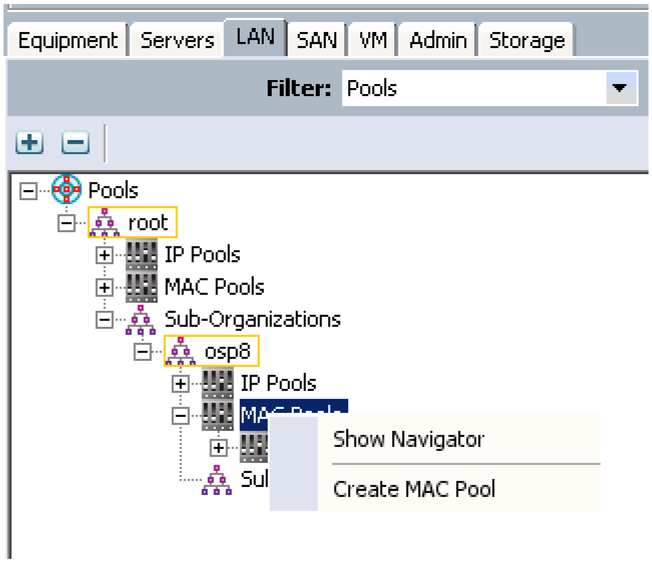
Filter: Pools

## Create MAC Pools

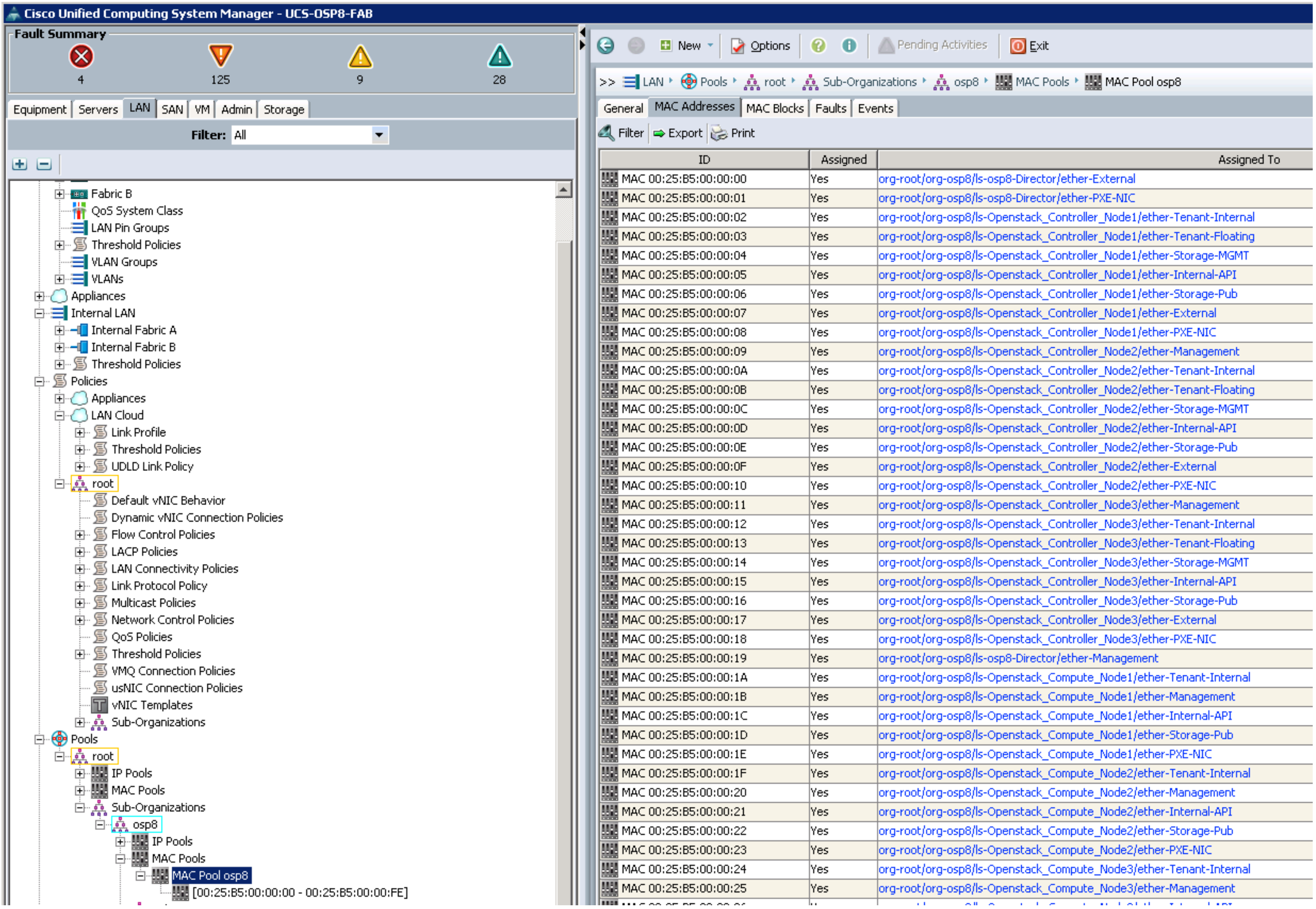
To configure a MAC address for each Cisco UCS Server VNIC interface, create the MAC pools from the Cisco UCS Manager GUI, and complete the following steps:

Under LAN → Pools → root → Sub-Organizations -> osp8 -> MAC Pools → right-click and select Create MAC Pool.

Specify the name and description for the MAC pool.



Filter: Pools



| ID | Assigned | Assigned To |
|---|---|---|
| MAC 00:25:B5:00:00:00 | Yes | org-root/org-osp8/ls-osp8-Director/ether-External |
| MAC 00:25:B5:00:00:01 | Yes | org-root/org-osp8/ls-osp8-Director/ether-PXE-NIC |
| MAC 00:25:B5:00:00:02 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:03 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-Tenant-Floating |
| MAC 00:25:B5:00:00:04 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-Storage-MGMT |
| MAC 00:25:B5:00:00:05 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-Internal-API |
| MAC 00:25:B5:00:00:06 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-Storage-Pub |
| MAC 00:25:B5:00:00:07 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-External |
| MAC 00:25:B5:00:00:08 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node1/ether-PXE-NIC |
| MAC 00:25:B5:00:00:09 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Management |
| MAC 00:25:B5:00:00:0A | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:0B | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Tenant-Floating |
| MAC 00:25:B5:00:00:0C | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Storage-MGMT |
| MAC 00:25:B5:00:00:0D | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Internal-API |
| MAC 00:25:B5:00:00:0E | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-Storage-Pub |
| MAC 00:25:B5:00:00:0F | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-External |
| MAC 00:25:B5:00:00:10 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node2/ether-PXE-NIC |
| MAC 00:25:B5:00:00:11 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Management |
| MAC 00:25:B5:00:00:12 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:13 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Tenant-Floating |
| MAC 00:25:B5:00:00:14 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Storage-MGMT |
| MAC 00:25:B5:00:00:15 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Internal-API |
| MAC 00:25:B5:00:00:16 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-Storage-Pub |
| MAC 00:25:B5:00:00:17 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-External |
| MAC 00:25:B5:00:00:18 | Yes | org-root/org-osp8/ls-Openstack_Controller_Node3/ether-PXE-NIC |
| MAC 00:25:B5:00:00:19 | Yes | org-root/org-osp8/ls-osp8-Director/ether-Management |
| MAC 00:25:B5:00:00:1A | Yes | org-root/org-osp8/ls-Openstack_Compute_Node1/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:1B | Yes | org-root/org-osp8/ls-Openstack_Compute_Node1/ether-Management |
| MAC 00:25:B5:00:00:1C | Yes | org-root/org-osp8/ls-Openstack_Compute_Node1/ether-Internal-API |
| MAC 00:25:B5:00:00:1D | Yes | org-root/org-osp8/ls-Openstack_Compute_Node1/ether-Storage-Pub |
| MAC 00:25:B5:00:00:1E | Yes | org-root/org-osp8/ls-Openstack_Compute_Node1/ether-PXE-NIC |
| MAC 00:25:B5:00:00:1F | Yes | org-root/org-osp8/ls-Openstack_Compute_Node2/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:20 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node2/ether-Management |
| MAC 00:25:B5:00:00:21 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node2/ether-Internal-API |
| MAC 00:25:B5:00:00:22 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node2/ether-Storage-Pub |
| MAC 00:25:B5:00:00:23 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node2/ether-PXE-NIC |
| MAC 00:25:B5:00:00:24 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node3/ether-Tenant-Internal |
| MAC 00:25:B5:00:00:25 | Yes | org-root/org-osp8/ls-Openstack_Compute_Node3/ether-Management |

## Create UUID Pools

To configure the UUID pools for each UCS Server, create the UUID pools from the Cisco UCS Manager GUI, complete the following steps:

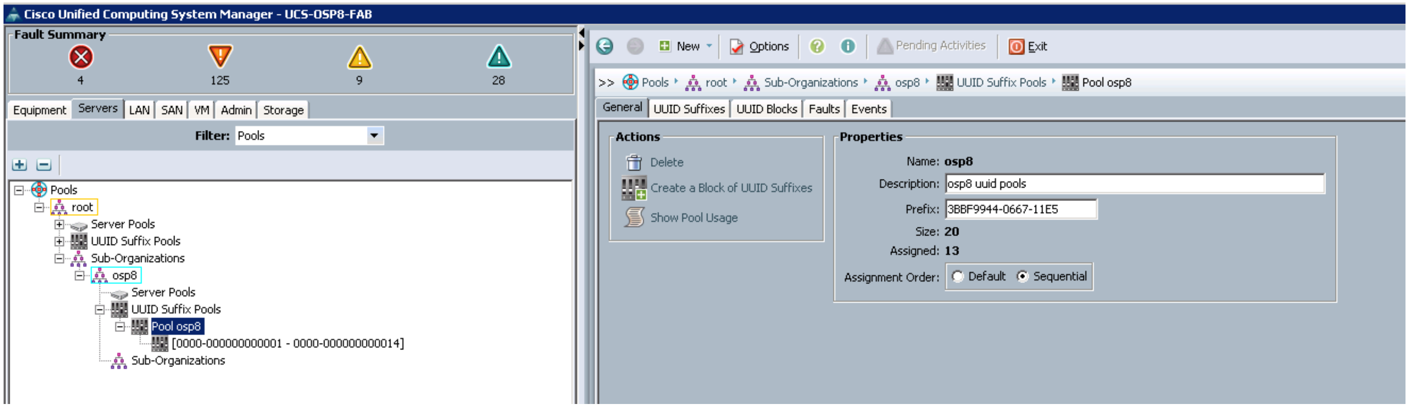
Under Servers → Pools → root → Sub-Organizations -> osp8 -> UUID Suffix Pools → right-click and select Create UUID Suffix Pool.

Specify the name and description for the UUID pool.

Click Add.

Specify the UUID Suffixes and size for the UUID pool.

Click Finish to complete the UUID pool creation.

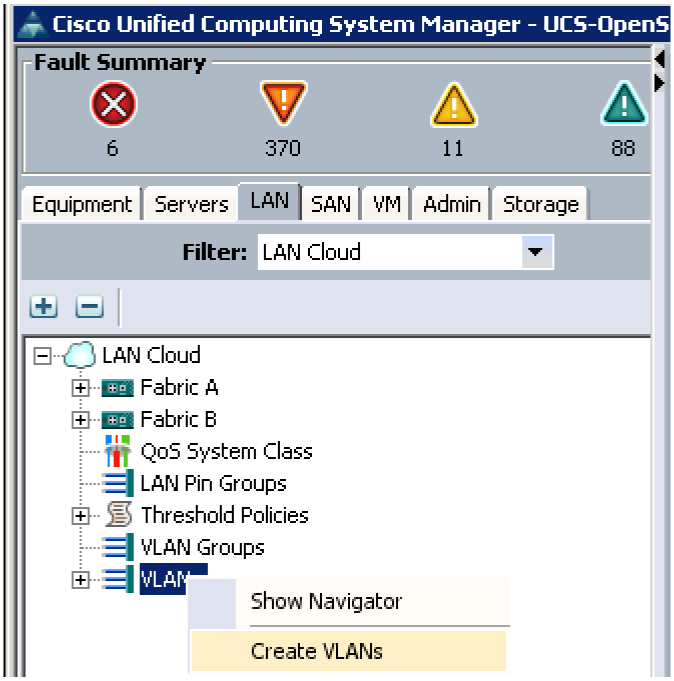


Name: **osp8**
Description: osp8 uuid pools
Prefix: 3BBF9944-0667-11E5
Size: **20**
Assigned: **13**
Assignment Order: ○ Default ● Sequential

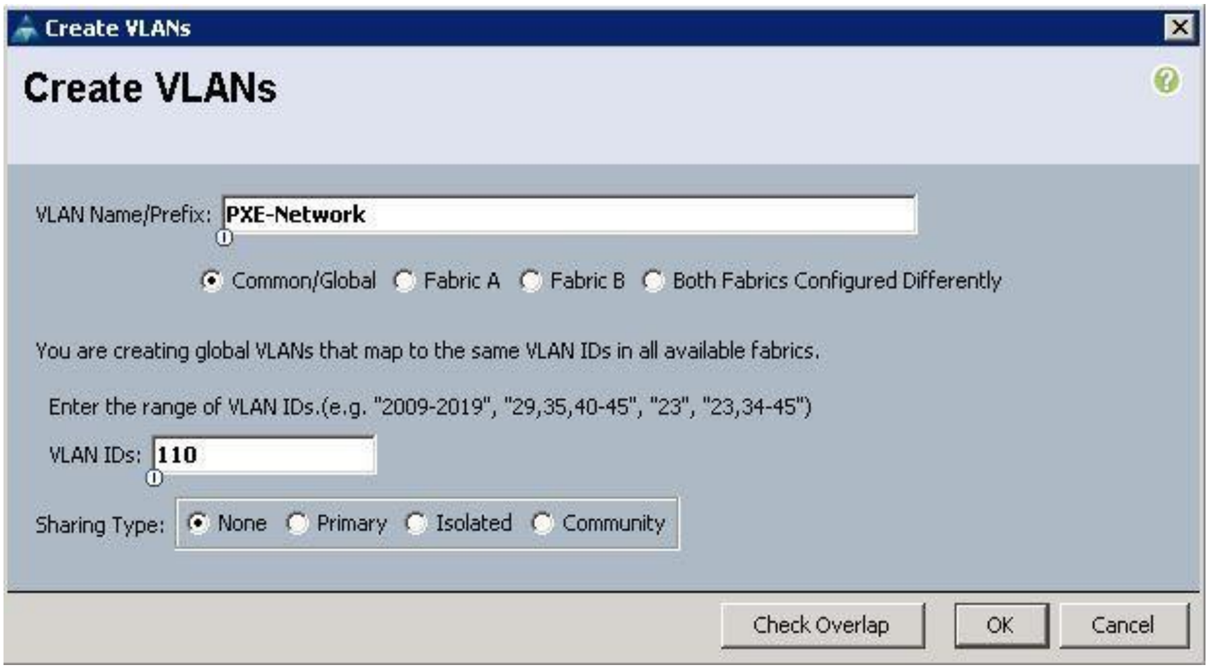
## Create VLANs

To create VLANs for all OpenStack networks for Controller, Compute and Ceph Storage Servers, from the UCS Manager GUI, complete the following steps:

Under LAN → Cloud → VLANs → right-click and select Create VLANs.



Specify the VLAN name as PXE-Network for Provisioning and specify the VLAN ID as 110 and click OK.



Create VLANs

VLAN Name/Prefix: PXE-Network

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

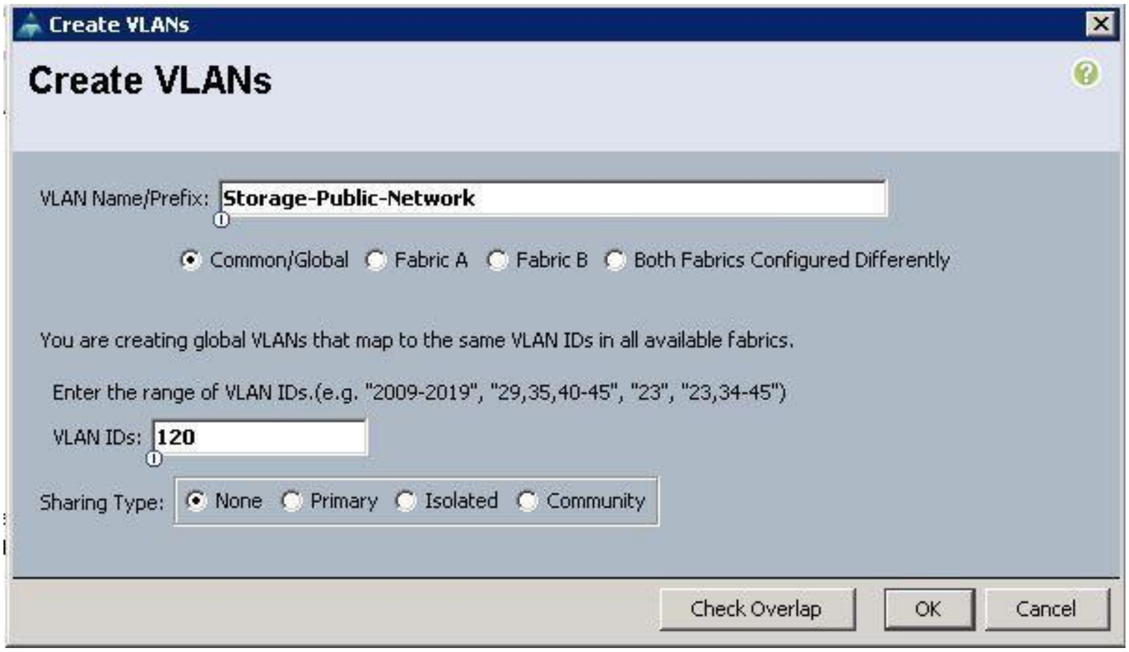
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 110

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

Specify the VLAN name as Storage-Public for accessing Ceph Storage Public Network and specify the VLAN ID as 120 and click OK.



Create VLANs

VLAN Name/Prefix: Storage-Public-Network

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

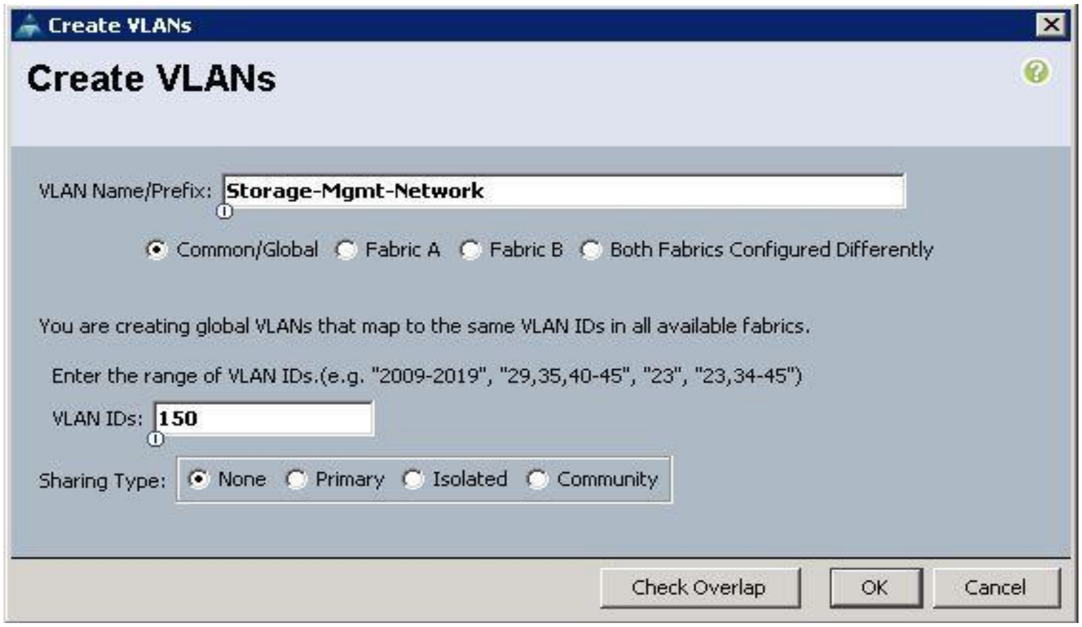
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 120

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

Specify the VLAN name as Storage-Mgmt-Network for Managing Ceph Storage Cluster and specify the VLAN ID as 150 and click OK.



Create VLANs

VLAN Name/Prefix: Storage-Mgmt-Network

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

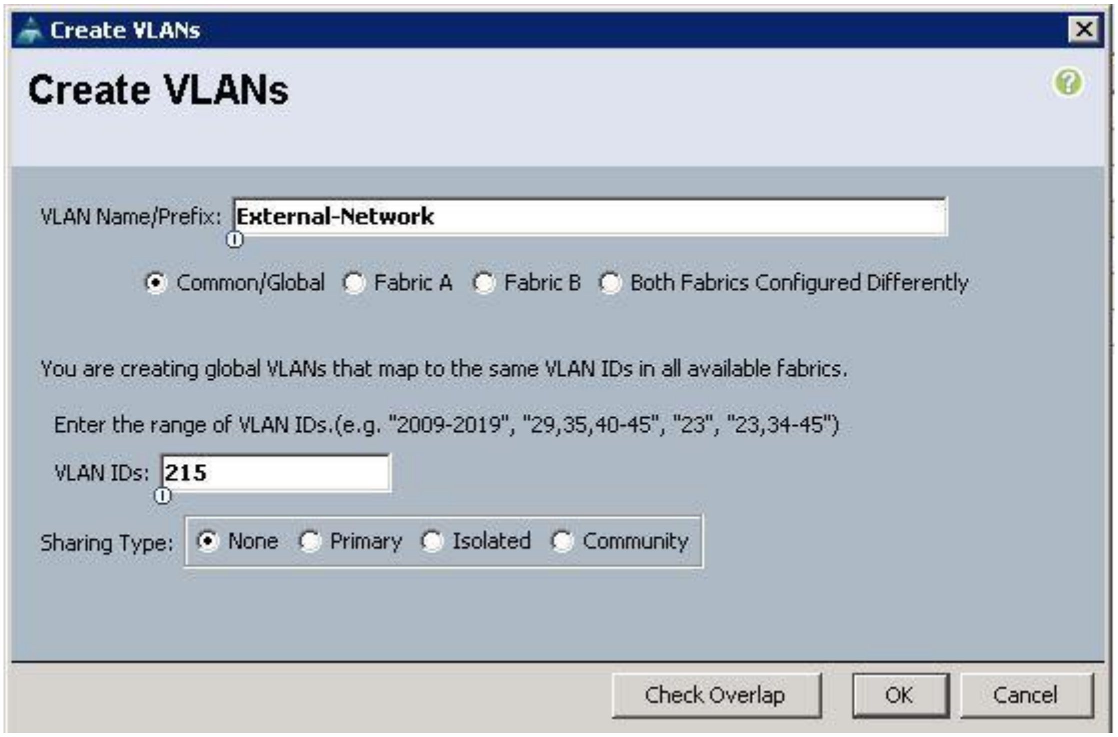
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 150

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

Specify the VLAN name as External-Network and specify the VLAN ID as 215 and click OK.



Create VLANs

VLAN Name/Prefix: External-Network

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

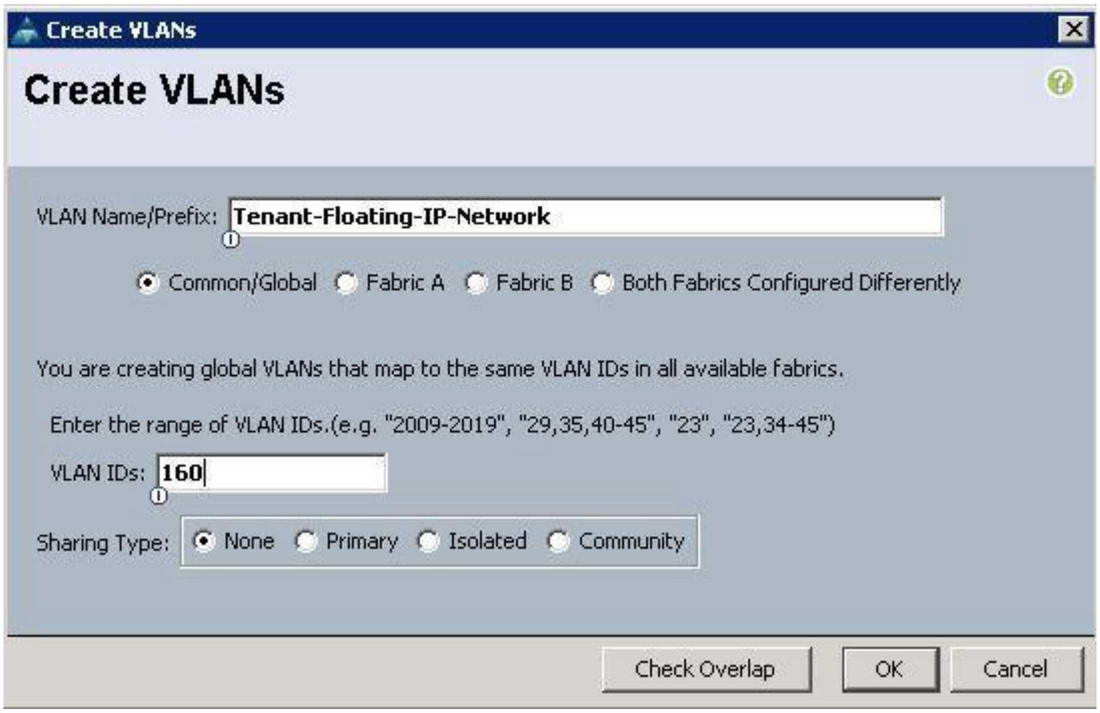
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 215

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

Specify the VLAN name as Tenant-Floating-Network for accessing Tenant instances externally and specify the VLAN ID as 160 and click OK.



Create VLANs

VLAN Name/Prefix: Tenant-Floating-IP-Network

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

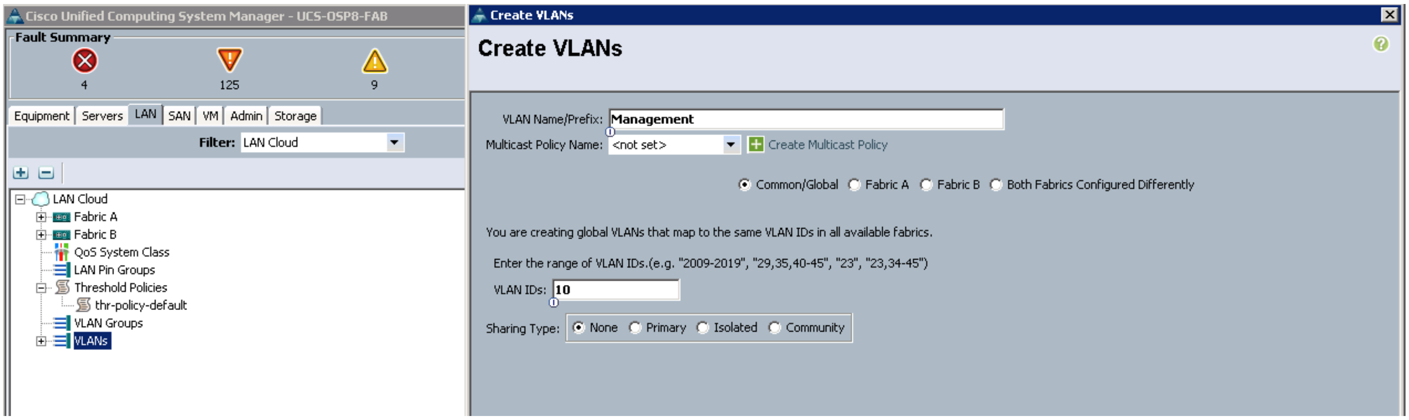
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 160

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

Specify the VLAN name as Management for accessing UCSM and Nexus 9k Networks and specify the VLAN ID as 10 and click OK.



Create VLANs

VLAN Name/Prefix: Management

Multicast Policy Name: <not set>

● Common/Global ○ Fabric A ○ Fabric B ○ Both Fabrics Configured Differently

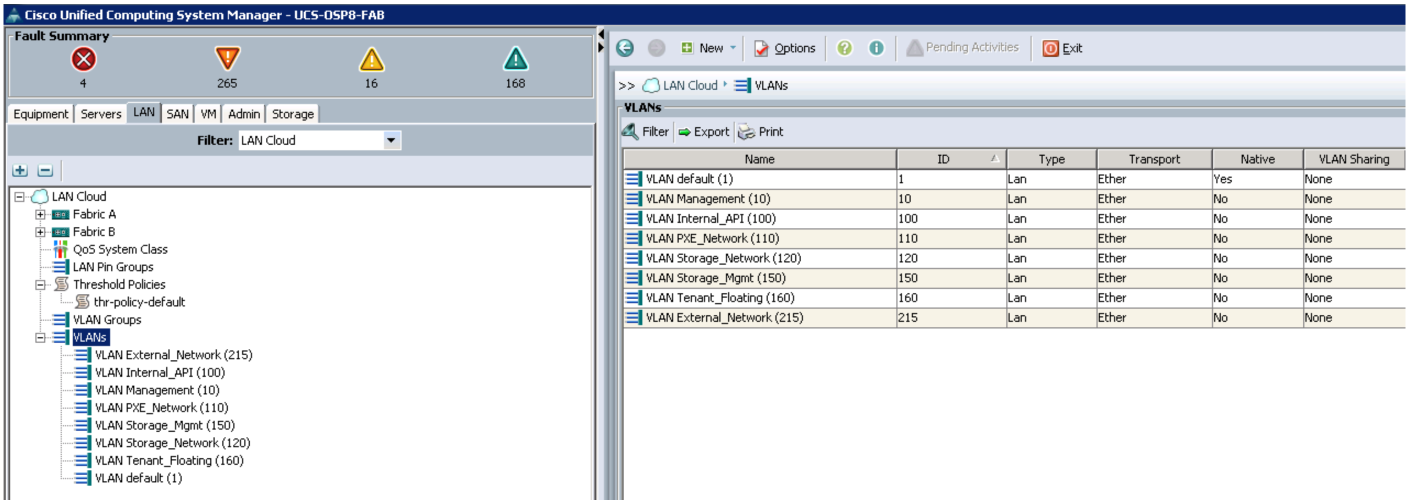
You are creating global VLANs that map to the same VLAN IDs in all available fabrics.

Enter the range of VLAN IDs.(e.g. "2009-2019", "29,35,40-45", "23", "23,34-45")

VLAN IDs: 10

Sharing Type: ● None ○ Primary ○ Isolated ○ Community

The screenshot below shows the output of VLANs for all the OpenStack Networks created above:



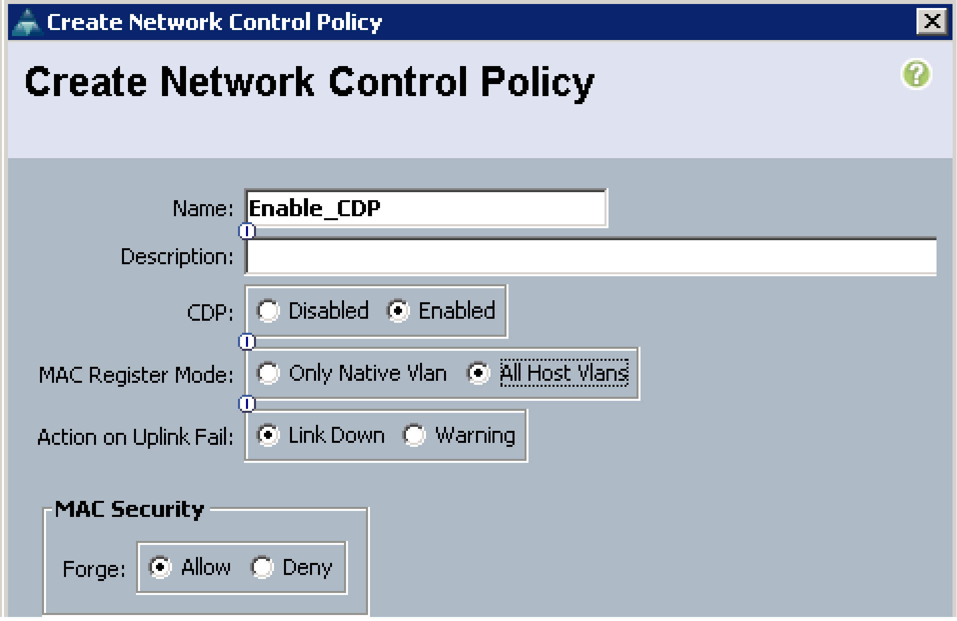
| Name | ID | Type | Transport | Native | VLAN Sharing |
|---|---|---|---|---|---|
| VLAN default (1) | 1 | Lan | Ether | Yes | None |
| VLAN Management (10) | 10 | Lan | Ether | No | None |
| VLAN Internal_API (100) | 100 | Lan | Ether | No | None |
| VLAN PXE_Network (110) | 110 | Lan | Ether | No | None |
| VLAN Storage_Network (120) | 120 | Lan | Ether | No | None |
| VLAN Storage_Mgmt (150) | 150 | Lan | Ether | No | None |
| VLAN Tenant_Floating (160) | 160 | Lan | Ether | No | None |
| VLAN External_Network (215) | 215 | Lan | Ether | No | None |

## Create a Network Control Policy

To configure the Network Control policy from the UCS Manager, complete the following steps:

Under LAN → Policies → root → Sub-Organizations -> Network Control Policies → right-click and select Create Network Control Policy.

Specify the name and choose CDP as Enabled. Select the MAC register mode as "All hosts VLANs" and Action on Uplink fail as "Link Down" and click OK.



Create Network Control Policy

Name: Enable_CDP

Description:

CDP: ○ Disabled ● Enabled

MAC Register Mode: ○ Only Native Vlan ● All Host Vlans

Action on Uplink Fail: ● Link Down ○ Warning

MAC Security

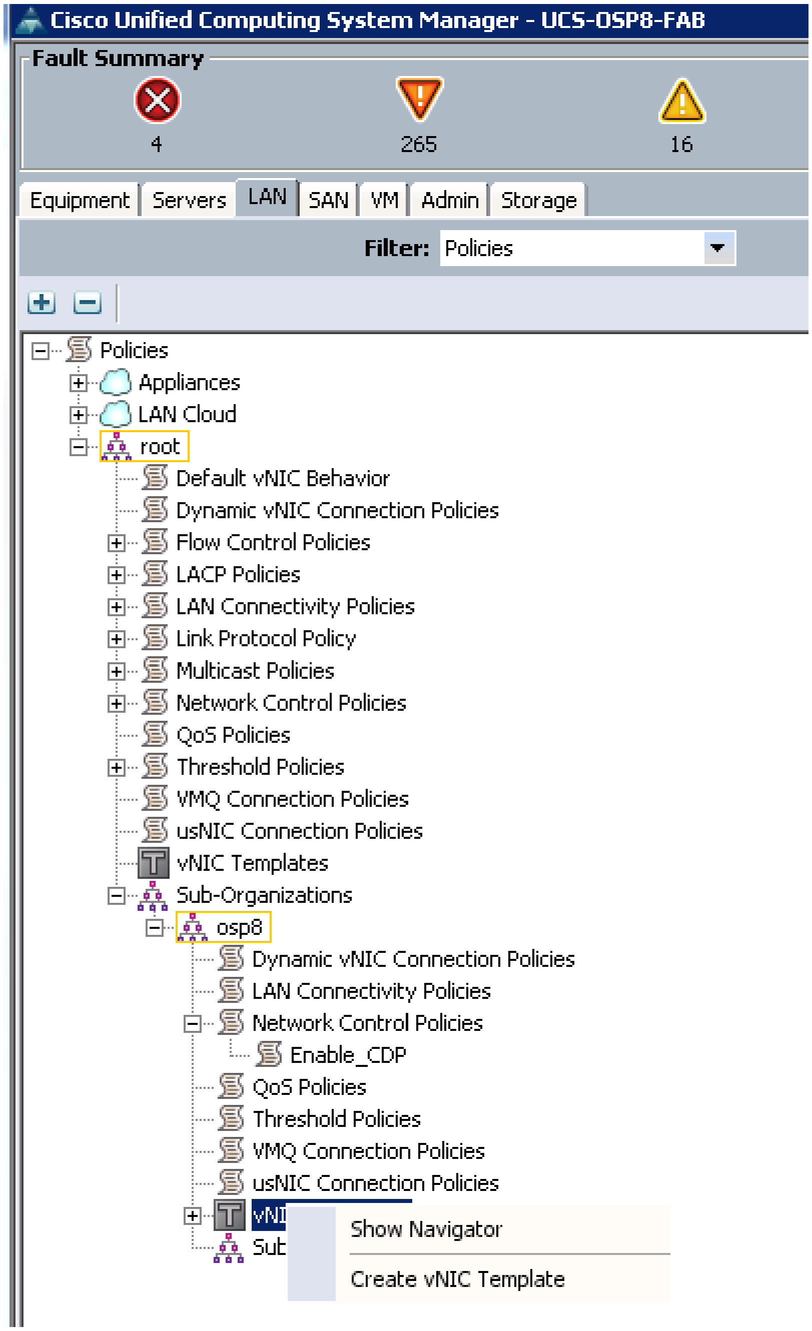
Forge: ● Allow ○ Deny

## Create vNIC Templates

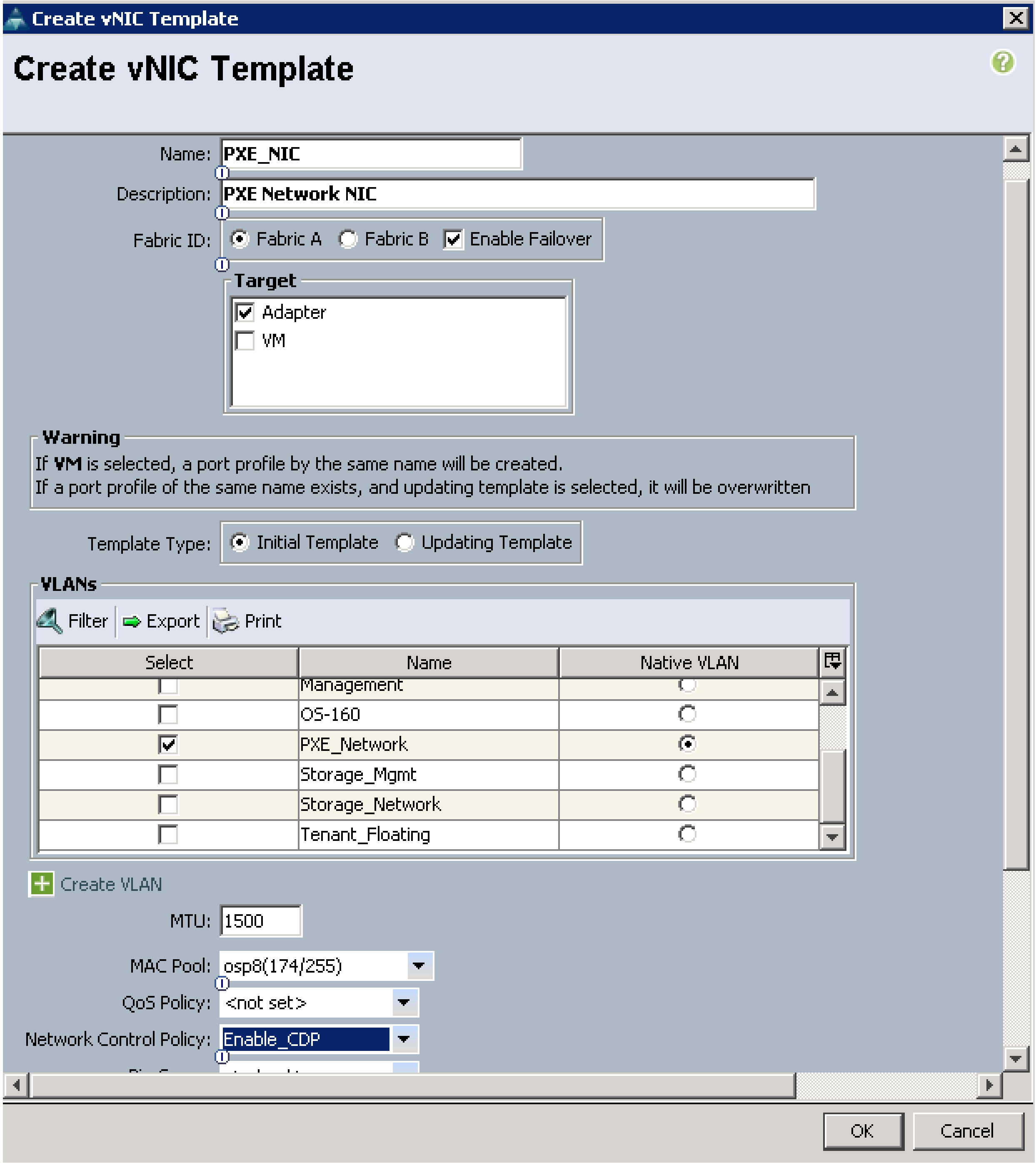
To Configure VNIC templates for each UCS Server VNIC interfaces, create VNIC templates from the Cisco UCS Manager GUI, complete the following steps:

The storage networks are configured with 9000 MTU.

Under LAN → Policies → root → Sub-Organizations ->osp8 -> VNIC Templates → right-click and select Create VNIC Template.



Create VNIC template for PXE or Provisioning network. Specify the name, description, Fabric ID, VLAN ID and choose MAC pools from the drop-down list.



Create vNIC Template

**Create vNIC Template**

Name: PXE_NIC

Description: PXE Network NIC

Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover

**Target**

☑ Adapter
□ VM

**Warning**
If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

**VLANs**

| Select | Name | Native VLAN |
|---|---|---|
| □ | Management | ○ |
| □ | OS-160 | ○ |
| ☑ | PXE_Network | ● |
| □ | Storage_Mgmt | ○ |
| □ | Storage_Network | ○ |
| □ | Tenant_Floating | ○ |

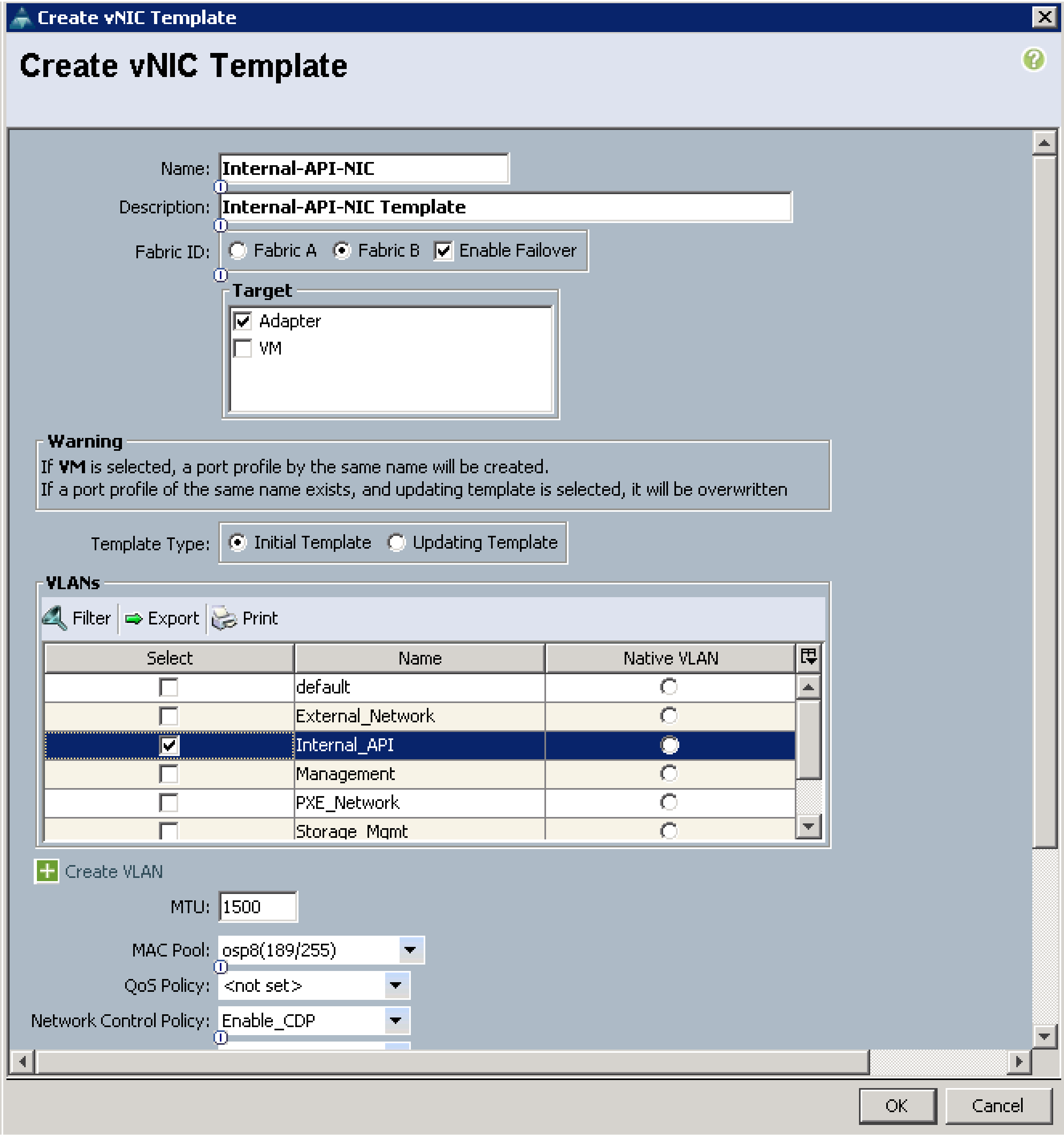
MTU: 1500

MAC Pool: osp8(174/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Create VNIC template for Internal-API network. Specify the name, description, Fabric ID, VLAN ID and choose MAC pools from the drop-down list.



Create vNIC Template

Create vNIC Template

Name: Internal-API-NIC

Description: Internal-API-NIC Template

Fabric ID: ○ Fabric A ● Fabric B ☑ Enable Failover

Target

☑ Adapter
□ VM

Warning

If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| □ | External_Network | ○ |
| ☑ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |

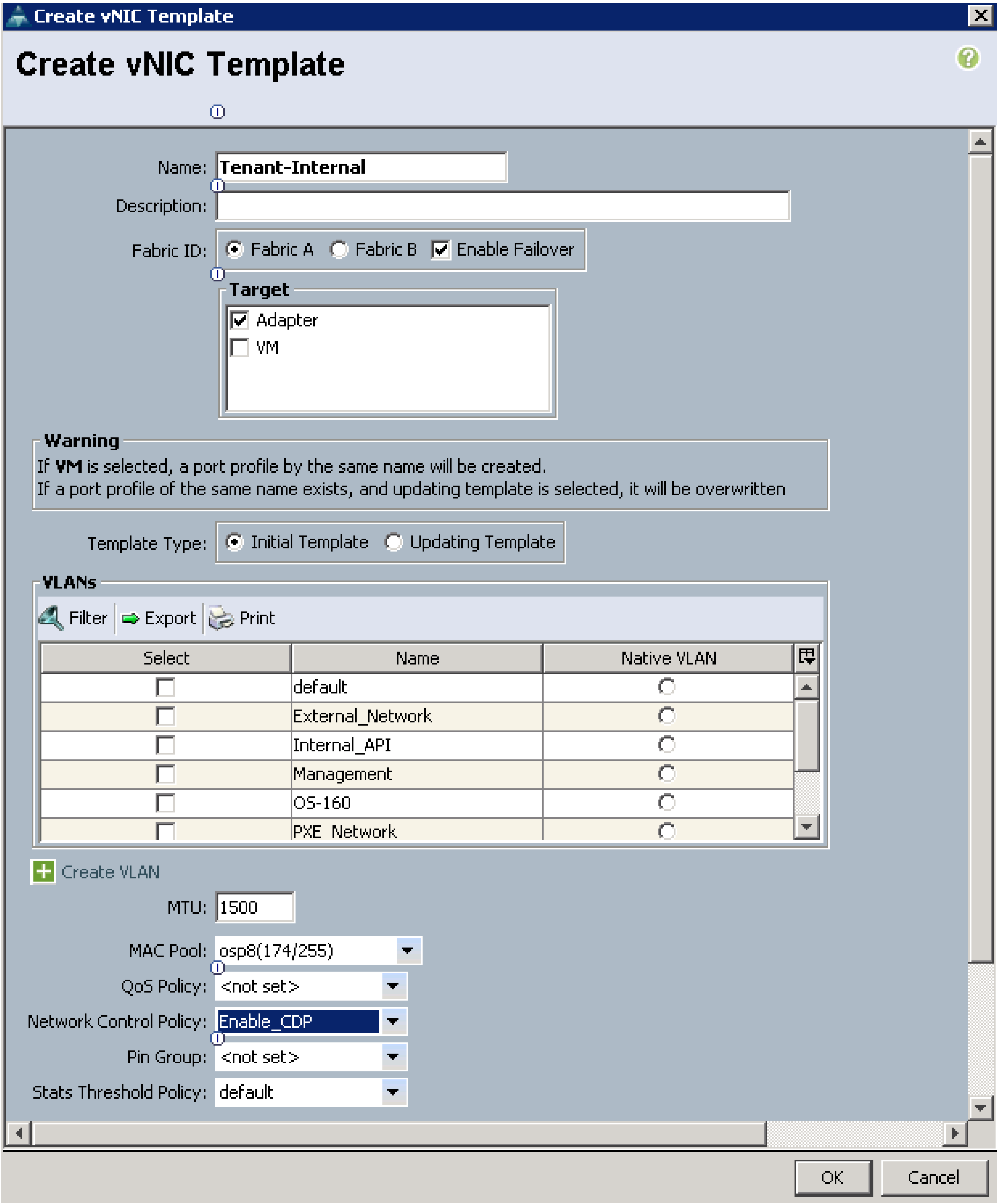
MTU: 1500

MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Create VNIC template for Tenant-Internal Network. You need not associate any VLAN's for Tenant-Internal network. The vlans for a tenant will be created globally and also on each compute blade's Tenant interface by UCSM plugin.



Create vNIC Template

Name: Tenant-Internal

Description:

Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover

Target

☑ Adapter

□ VM

Warning

If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| □ | External_Network | ○ |
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | OS-160 | ○ |
| □ | PXE_Network | ○ |

MTU: 1500

MAC Pool: osp8(174/255)

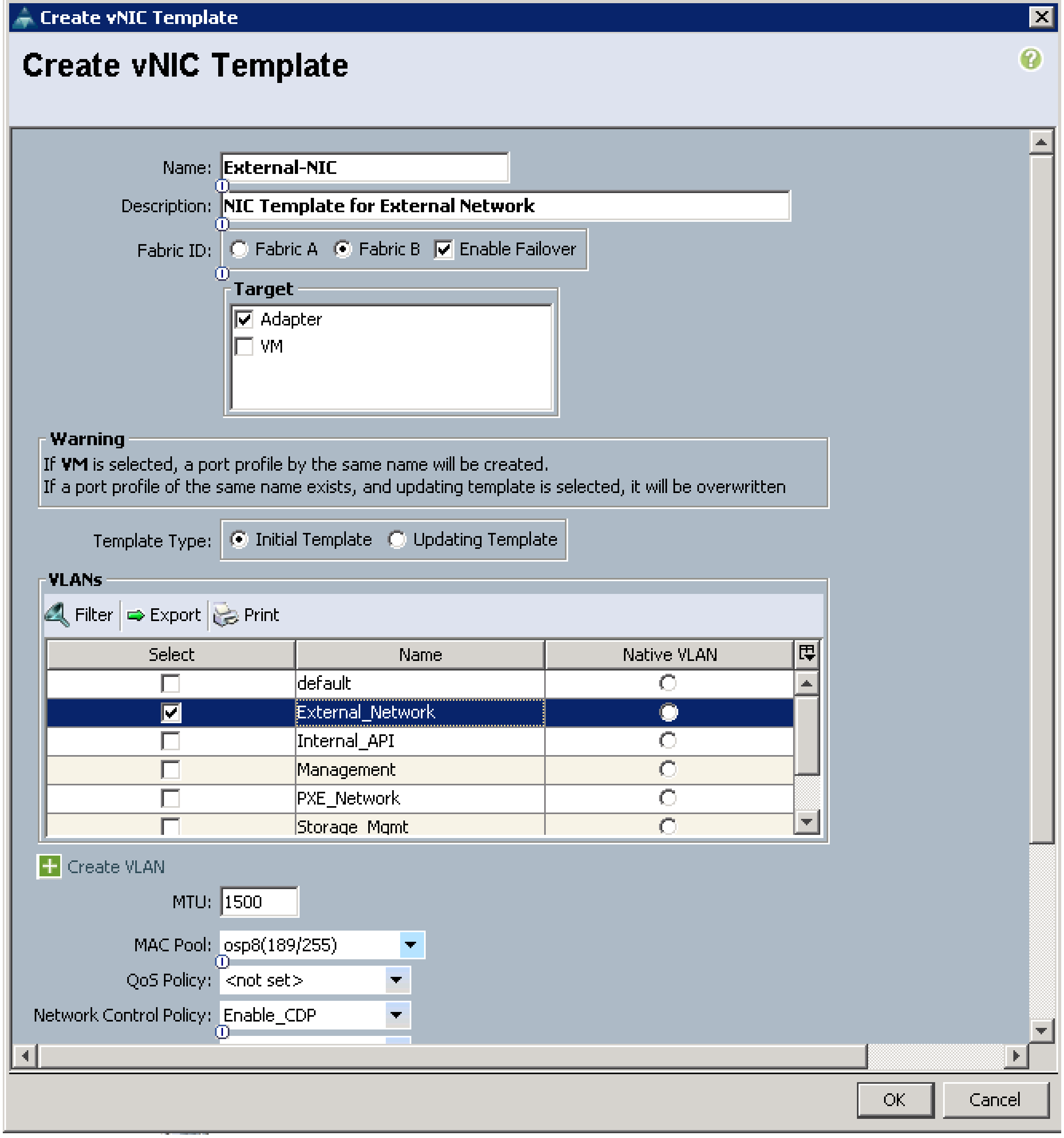
QoS Policy: <not set>

Network Control Policy: Enable_CDP

Pin Group: <not set>

Stats Threshold Policy: default

Create VNIC template for External Network.



Create vNIC Template

Name: External-NIC

Description: NIC Template for External Network

Fabric ID: ○ Fabric A ● Fabric B ☑ Enable Failover

Target

☑ Adapter

□ VM

Warning

If **VM** is selected, a port profile by the same name will be created.

If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| ☑ | External_Network | ● |
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |

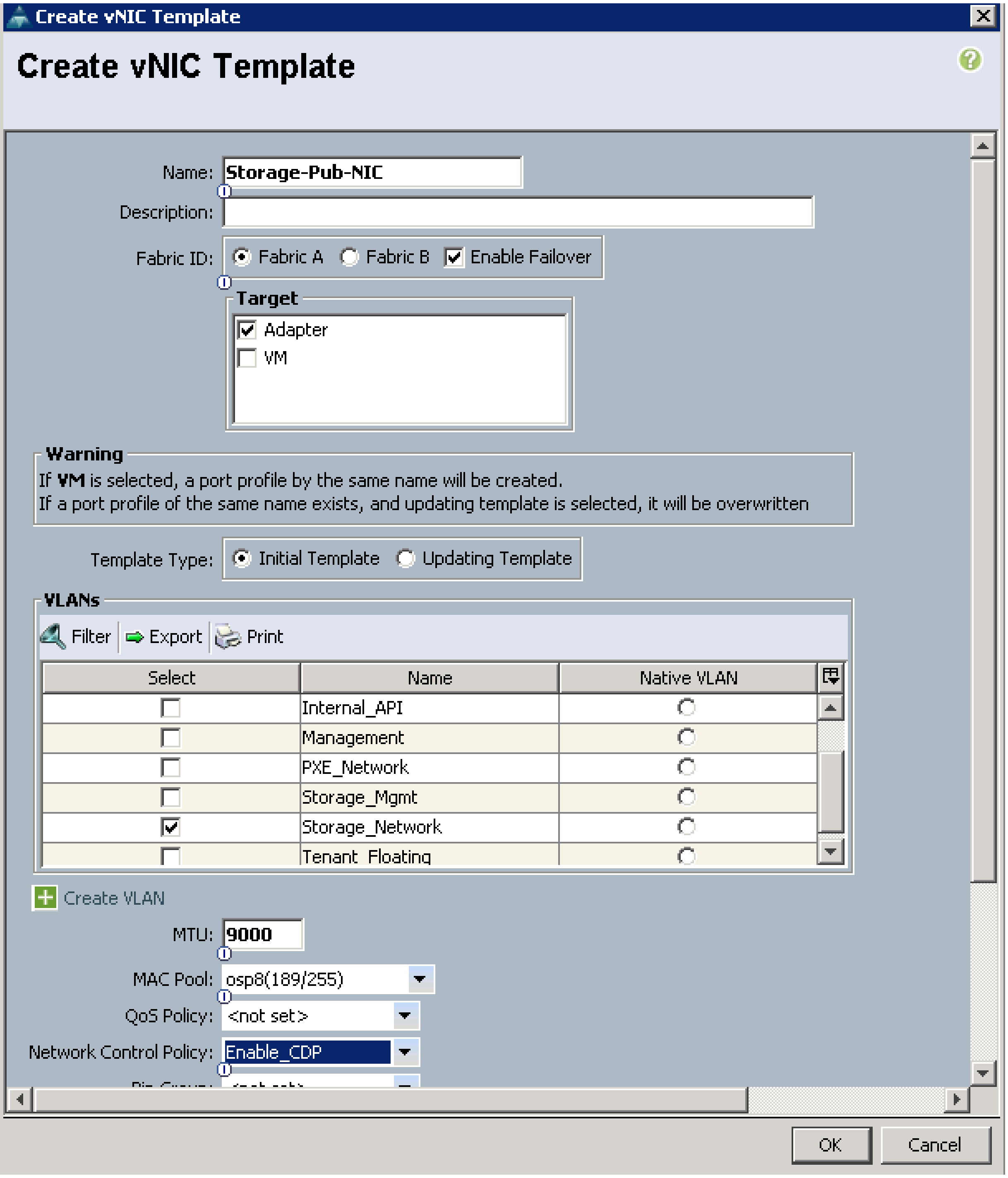
MTU: 1500

MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Create the VNIC template for Storage Public Network.



Create vNIC Template

Name: Storage-Pub-NIC

Description:

Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover

Target
☑ Adapter
□ VM

Warning
If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |
| ☑ | Storage_Network | ○ |
| □ | Tenant_Floating | ○ |

MTU: 9000

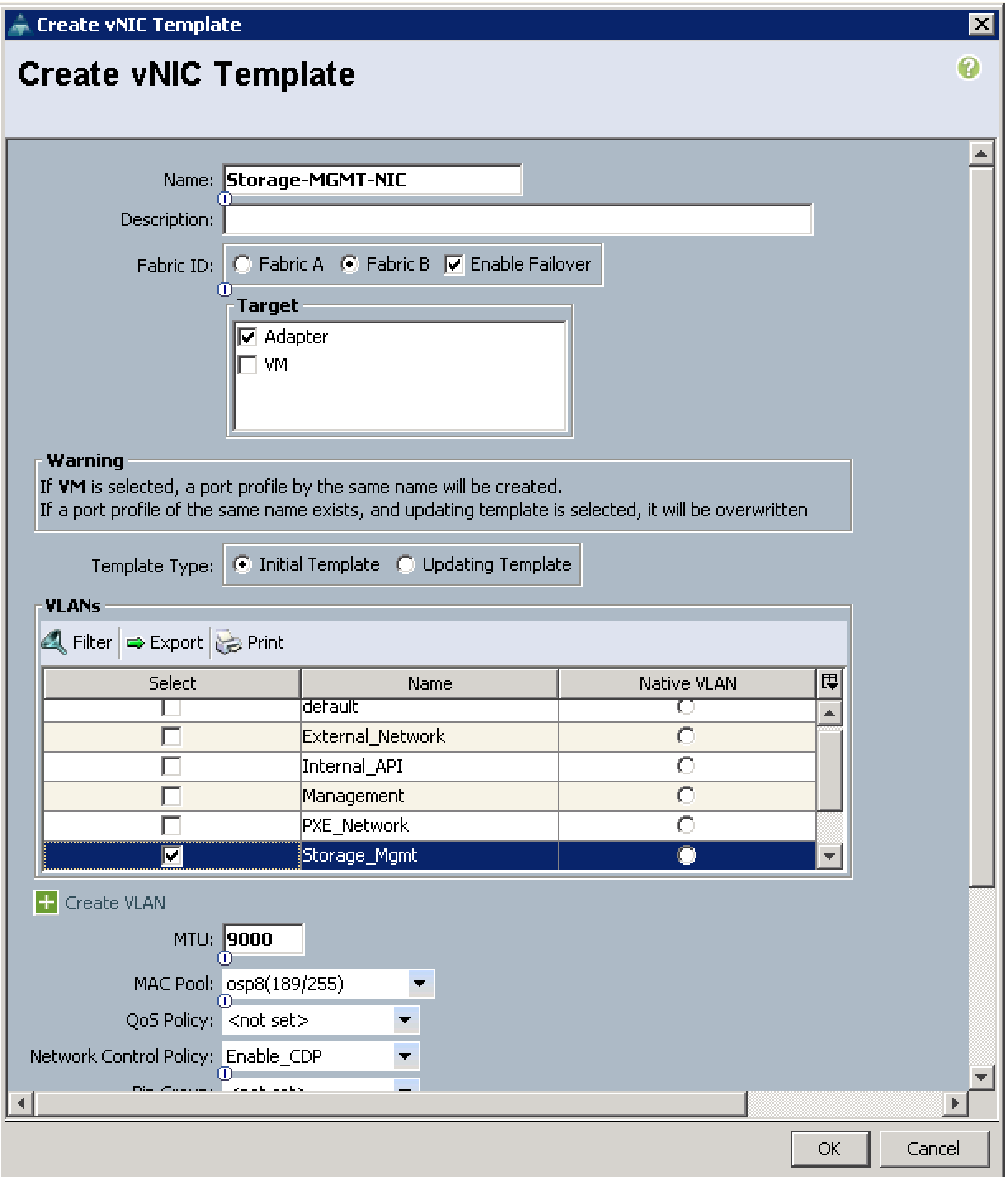
MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Pin Group: <not set>

Create VNIC template for Storage Mgmt Cluster network.



Create vNIC Template

Name: Storage-MGMT-NIC

Description:

Fabric ID: ○ Fabric A ● Fabric B ☑ Enable Failover

Target

☑ Adapter

□ VM

Warning

If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| □ | External_Network | ○ |
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| ☑ | Storage_Mgmt | ● |

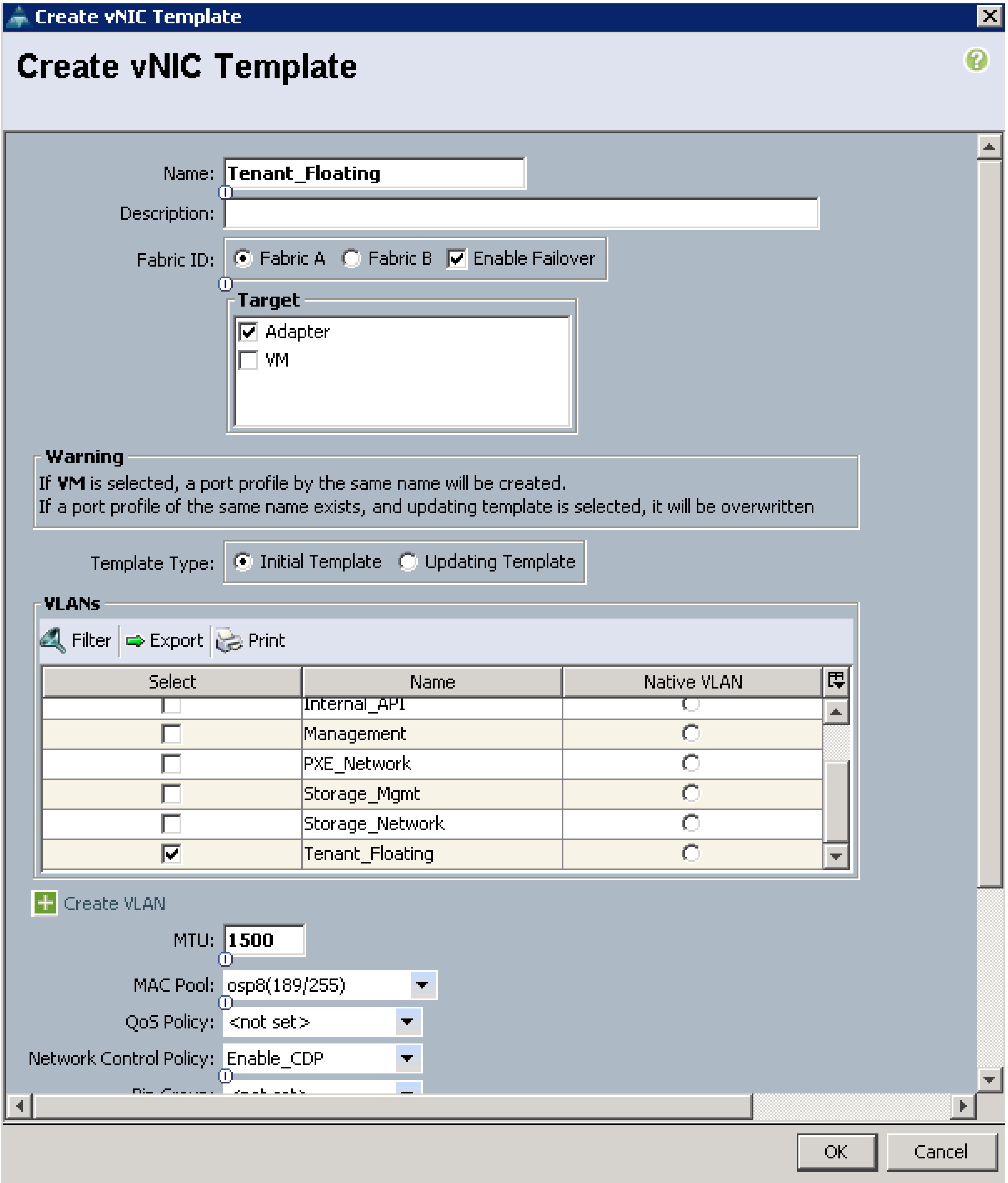
MTU: 9000

MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Create VNIC template for Tenant Floating Network.



Create vNIC Template

Name: Tenant_Floating

Description:

Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover

Target

☑ Adapter

□ VM

Warning

If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |
| □ | Storage_Network | ○ |
| ☑ | Tenant_Floating | ○ |

MTU: 1500

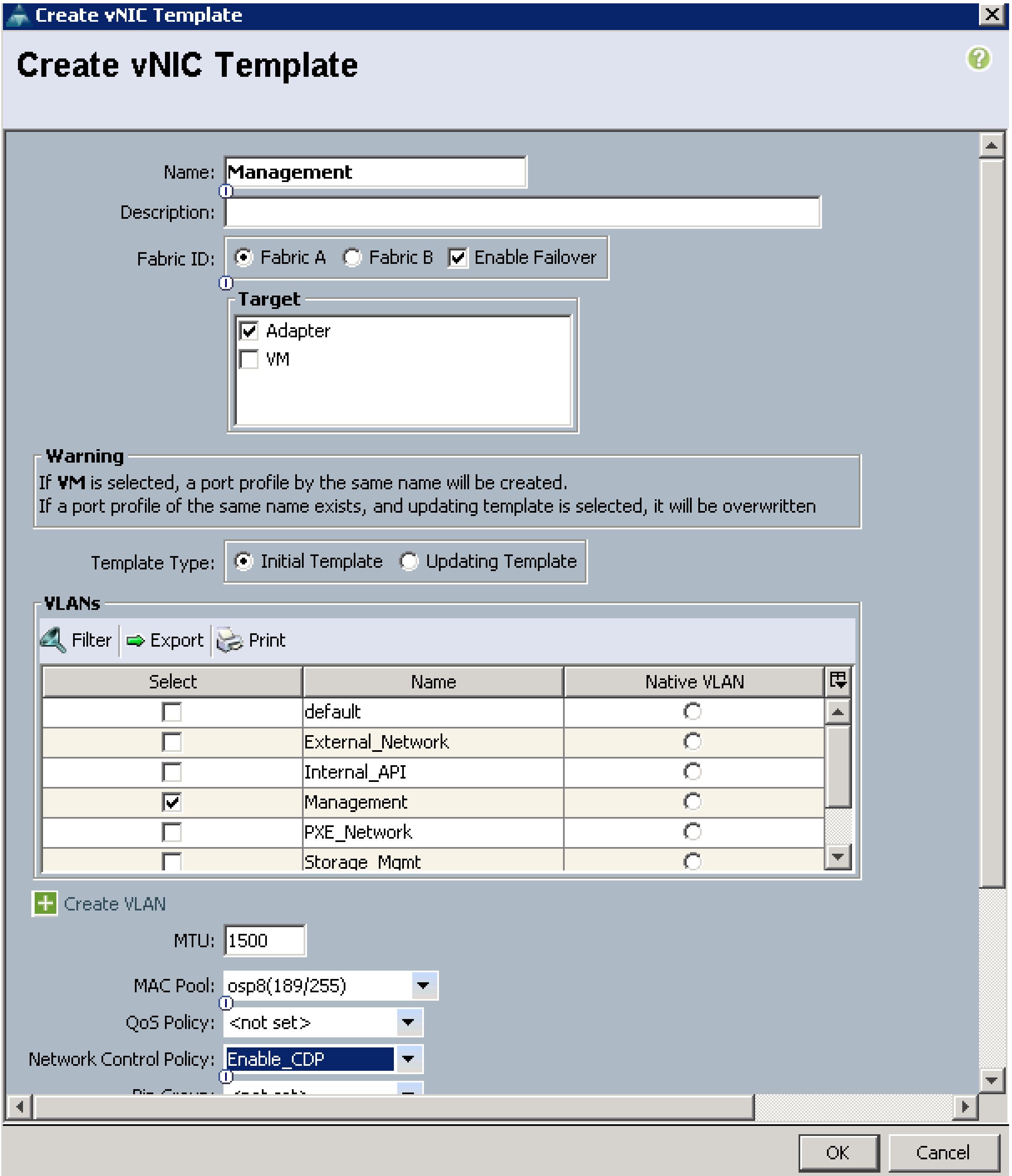
MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

Pin Group: <not set>

Create VNIC template for Management Network.



Create vNIC Template

Name: Management

Description:

Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover

Target

☑ Adapter
□ VM

Warning
If **VM** is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ● Initial Template ○ Updating Template

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| □ | External_Network | ○ |
| □ | Internal_API | ○ |
| ☑ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |

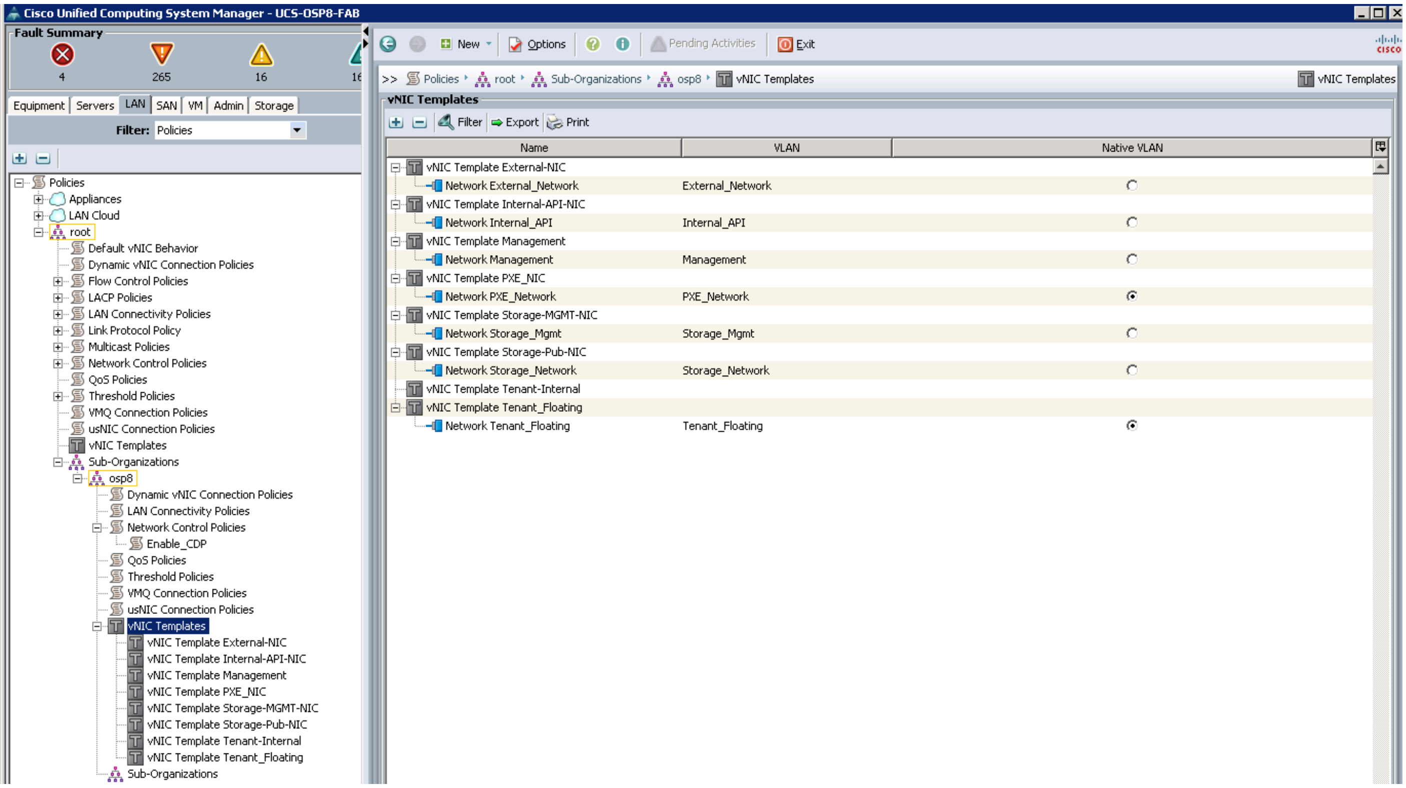
MTU: 1500

MAC Pool: osp8(189/255)

QoS Policy: <not set>

Network Control Policy: Enable_CDP

After completion, you can see the VNIC templates for each traffic.



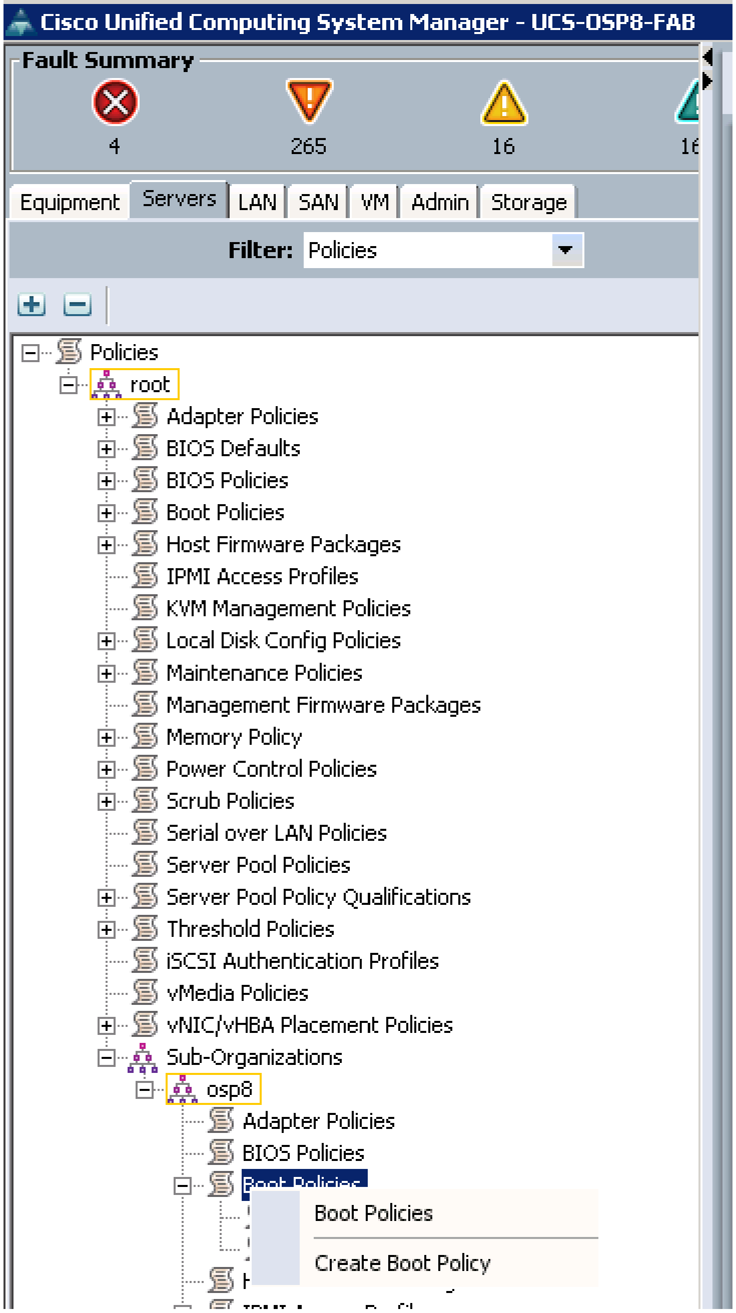
| Name | VLAN | Native VLAN |
|---|---|---|
| vNIC Template External-NIC | | |
| Network External_Network | External_Network | ○ |
| vNIC Template Internal-API-NIC | | |
| Network Internal_API | Internal_API | ○ |
| vNIC Template Management | | |
| Network Management | Management | ○ |
| vNIC Template PXE_NIC | | |
| Network PXE_Network | PXE_Network | ● |
| vNIC Template Storage-MGMT-NIC | | |
| Network Storage_Mgmt | Storage_Mgmt | ○ |
| vNIC Template Storage-Pub-NIC | | |
| Network Storage_Network | Storage_Network | ○ |
| vNIC Template Tenant-Internal | | |
| vNIC Template Tenant_Floating | | |
| Network Tenant_Floating | Tenant_Floating | ● |

For storage interfaces, a MTU value of 9000 has been added.

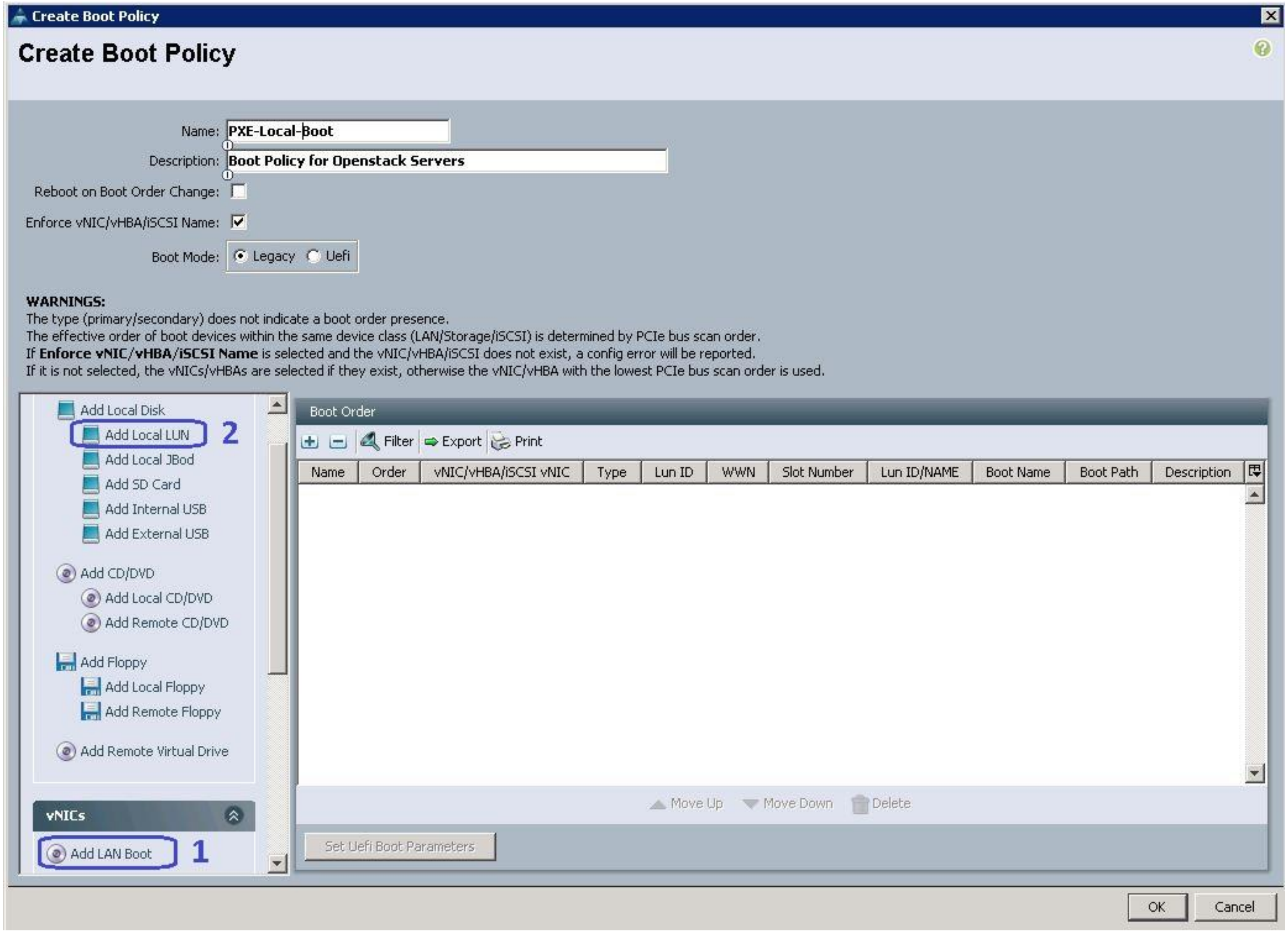
## Create Boot Policy

To configure the Boot policy for the Cisco UCS Servers, create a Boot Policy from the Cisco UCS Manager GUI and complete the following steps:

Under Server → Policies → root → Sub-Organizations -> Boot Policies → right-click and select Create Boot Policy.



Specify the name and description. Select the First boot order as LAN boot and specify the actual VNIC name of the PXE network (PXE-NIC). Then select the second boot order and click Add Local LUN.



Create Boot Policy

Name: PXE-Local-Boot

Description: Boot Policy for Openstack Servers

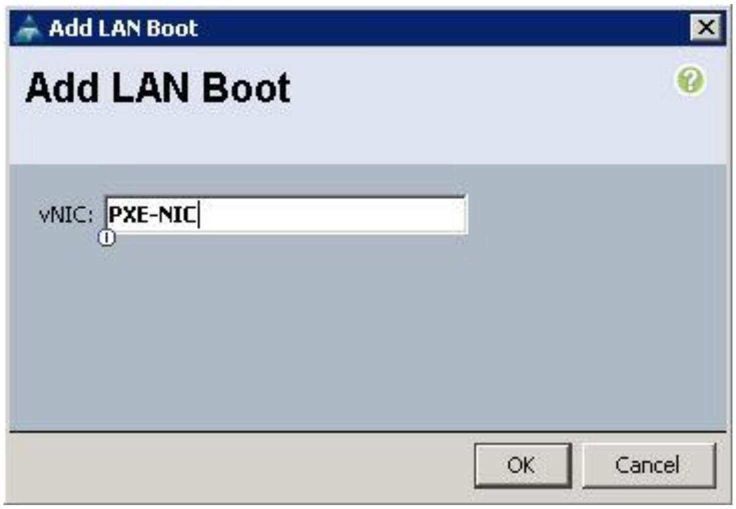
Reboot on Boot Order Change: ☐

Enforce vNIC/vHBA/iSCSI Name: ☑

Boot Mode: ● Legacy ○ Uefi

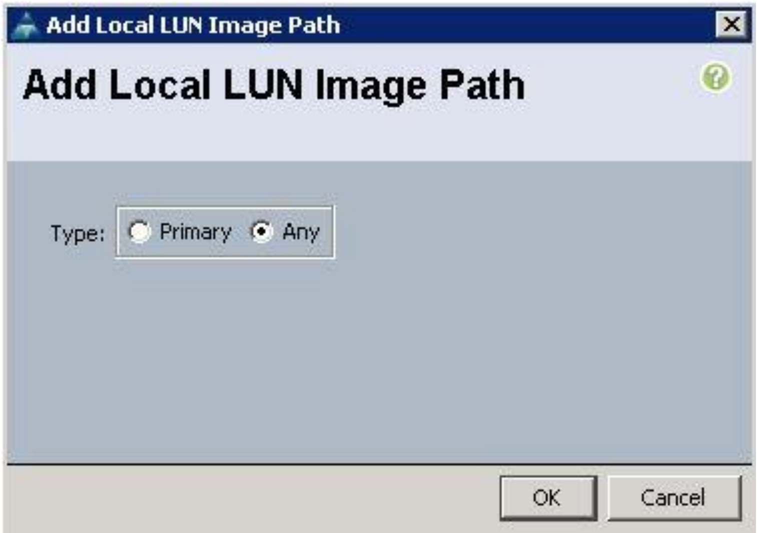
**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Specify the VNIC Name as PXE-NIC.



Add LAN Boot

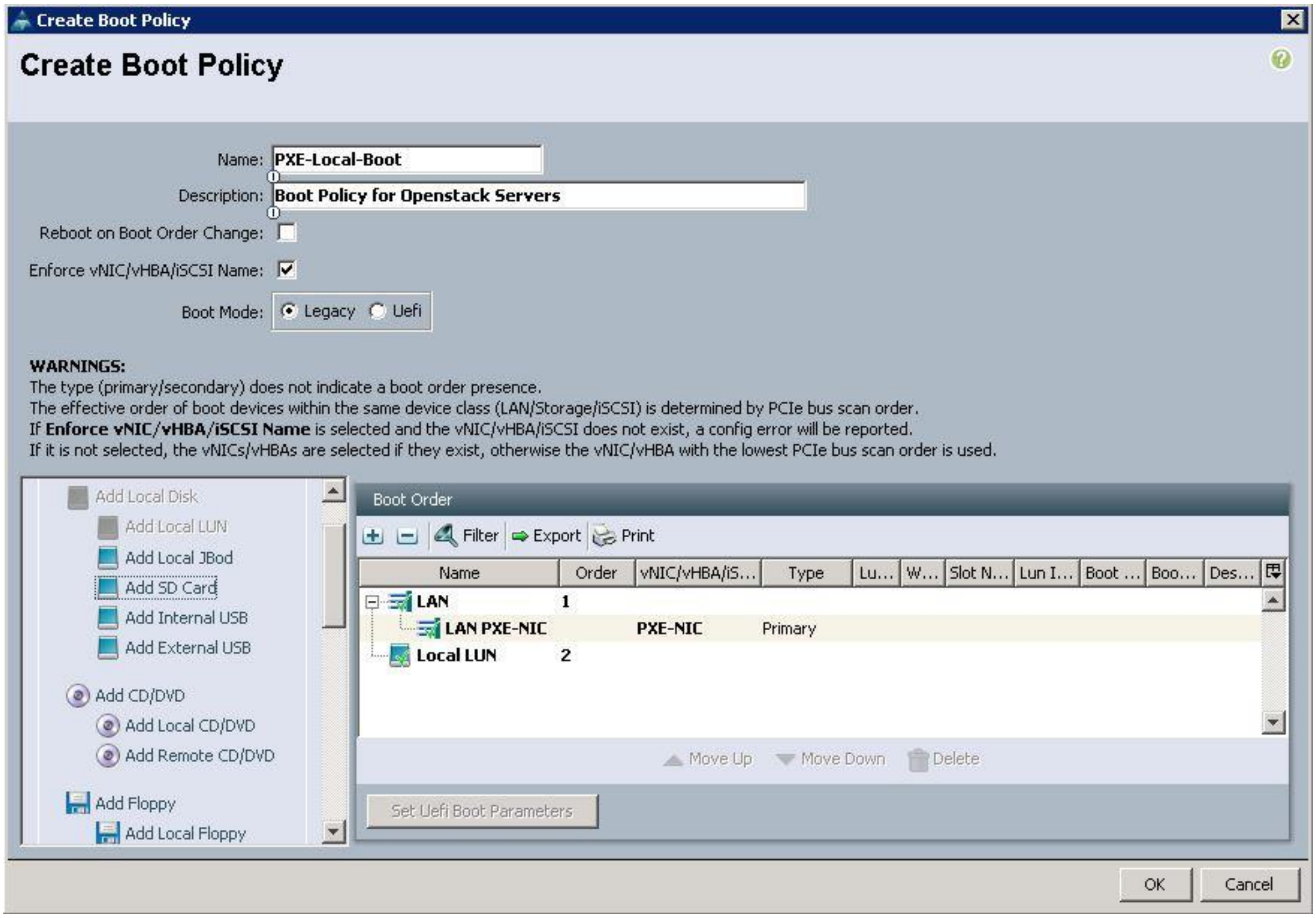
vNIC: PXE-NIC



Add Local LUN Image Path

Type: ○ Primary ● Any

Make sure the First boot order is PXE NIC and second boot order is Local LUN and click OK.



Create Boot Policy

Name: PXE-Local-Boot

Description: Boot Policy for Openstack Servers

Reboot on Boot Order Change: □

Enforce vNIC/vHBA/iSCSI Name: ☑

Boot Mode: ● Legacy ○ Uefi

**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Boot Order

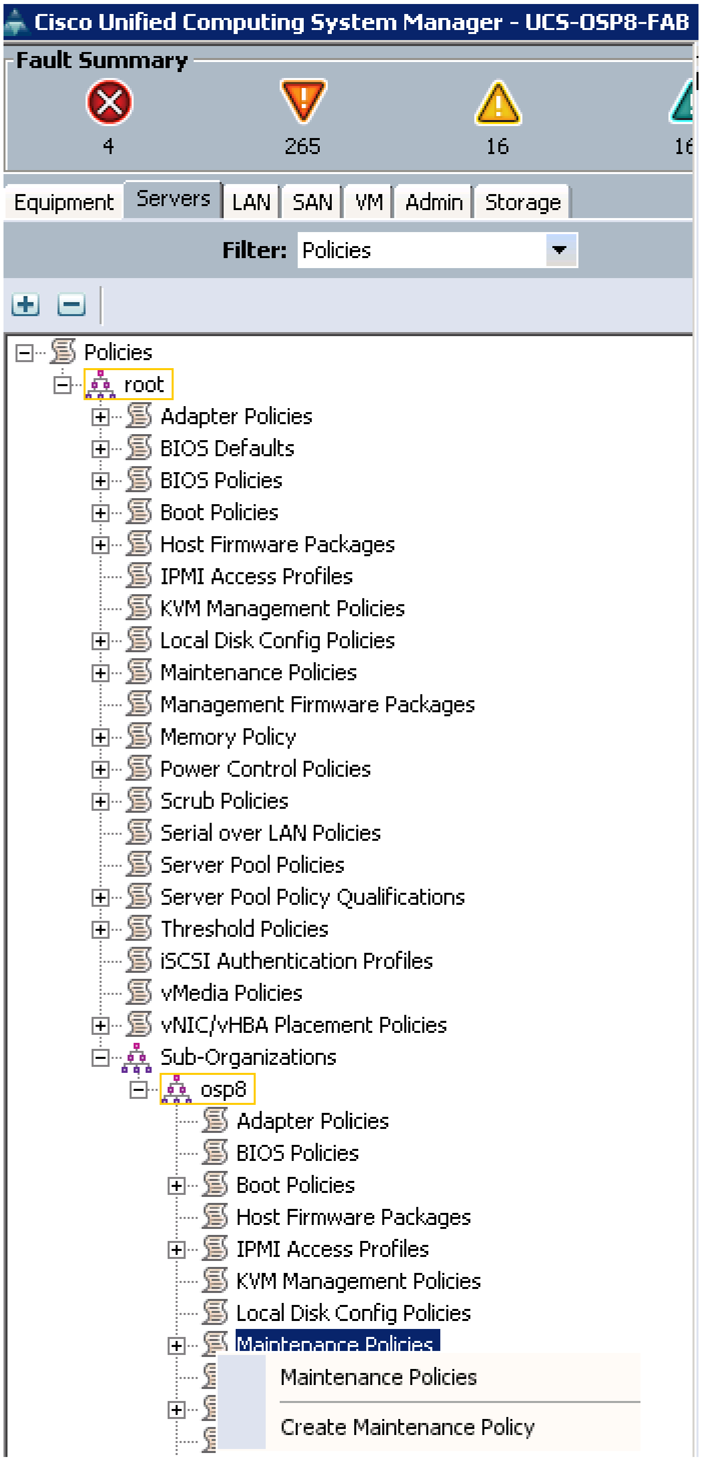
| Name | Order | vNIC/vHBA/iS... | Type | Lu... | W... | Slot N... | Lun I... | Boot ... | Boo... | Des... |
|---|---|---|---|---|---|---|---|---|---|---|
| LAN | 1 | | | | | | | | | |
| LAN PXE-NIC | | PXE-NIC | Primary | | | | | | | |
| Local LUN | 2 | | | | | | | | | |

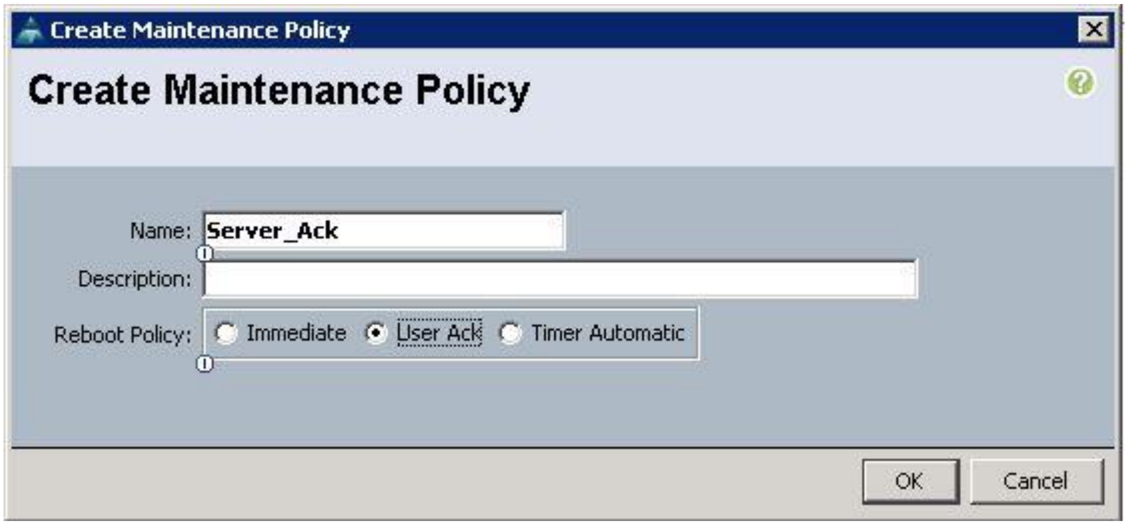
## Create a Maintenance Policy

A maintenance policy determines a pre-defined action to take when there is a disruptive change made to the service profile associated with a server. When creating a maintenance policy you have to select a reboot policy which defines when the server can reboot once the changes are applied.

To configure the Maintenance policy from the Cisco UCS Manager, complete the following steps:

Under Server → Policies → root → Sub-Organizations -> Maintenance Policies → right-click and select Create Maintenance Policy.





Create Maintenance Policy

Name: Server_Ack

Description:

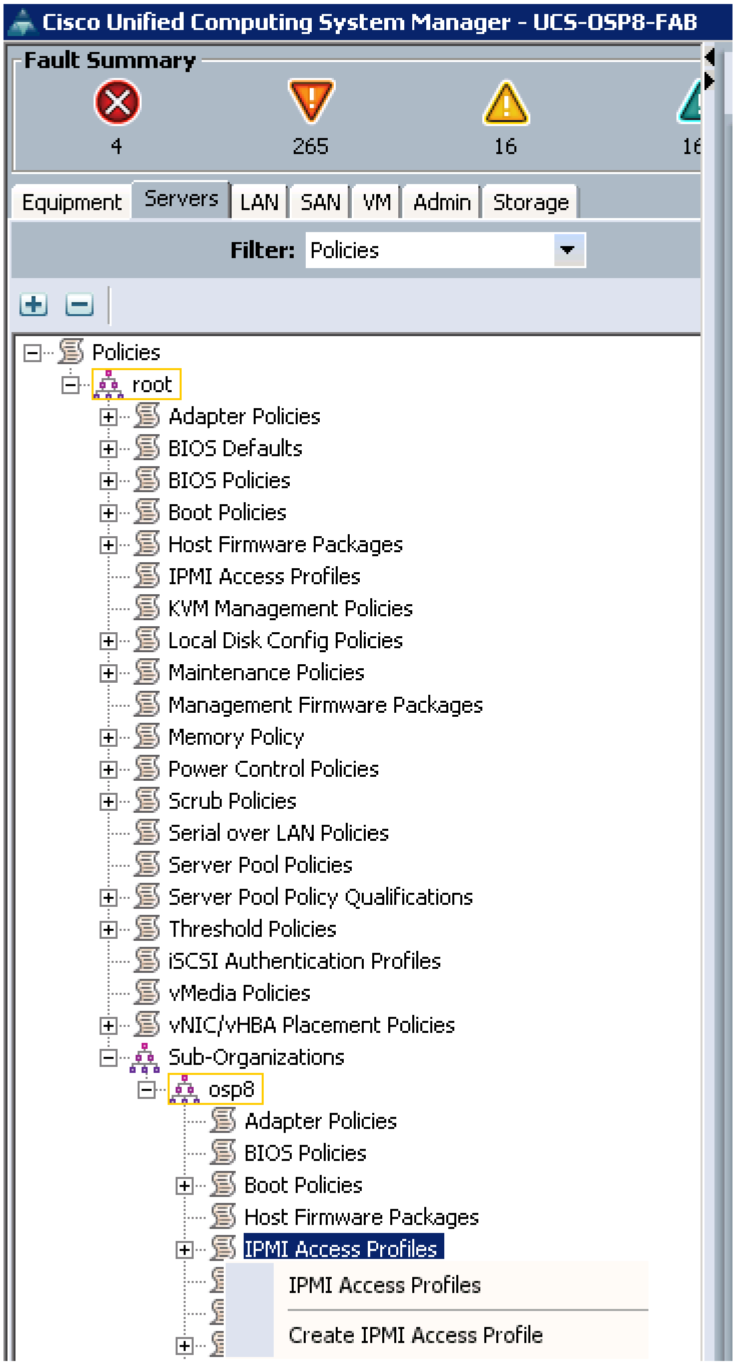
Reboot Policy: ○ Immediate ● User Ack ○ Timer Automatic

## Create an IPMI Access Policy

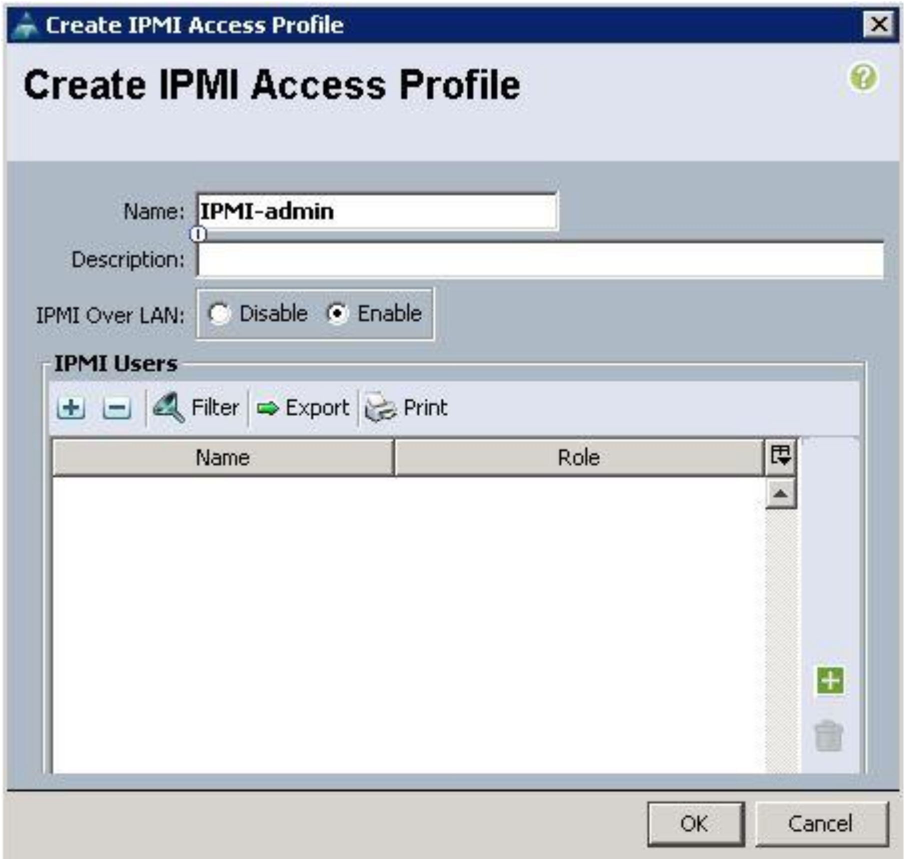
This policy allows you to determine whether IPMI commands can be sent directly to the server, using the IP address (KVM IP address).

To configure the IPMI Access profiles from the Cisco UCS Manager, complete the following steps:

Under Server → Policies → root → Sub-Organizations -> IPMI Access profiles → right-click and select Create IPMI Access Profile.



Specify the name and click IPMI over LAN as Enabled and click "+".



Create IPMI Access Profile

Name: IPMI-admin

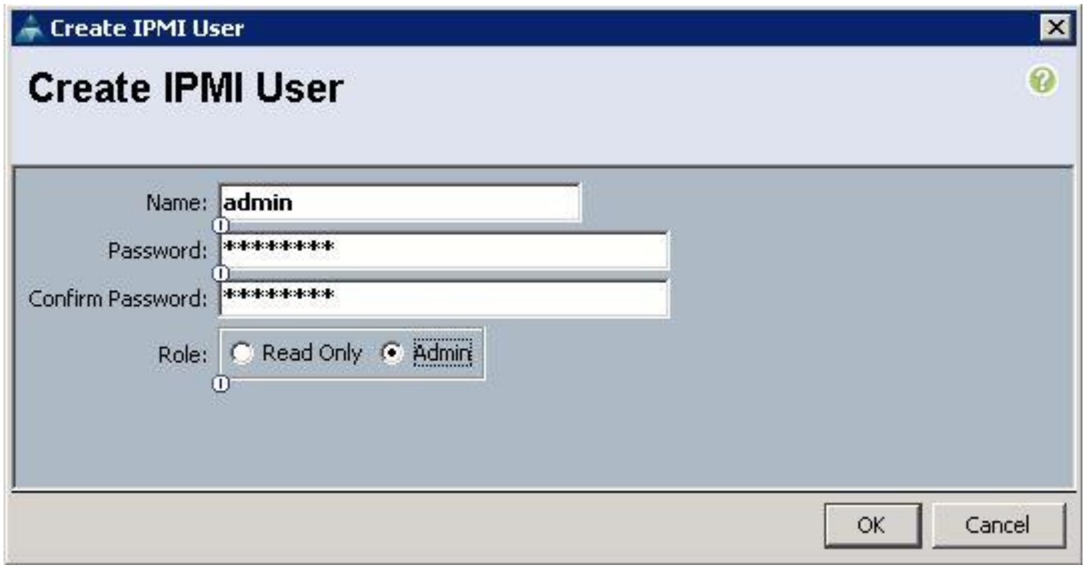
Description:

IPMI Over LAN: ○ Disable ● Enable

IPMI Users

| Name | Role |
|---|---|

Specify the username and password. Choose Admin for the Role and click OK.



Create IPMI User

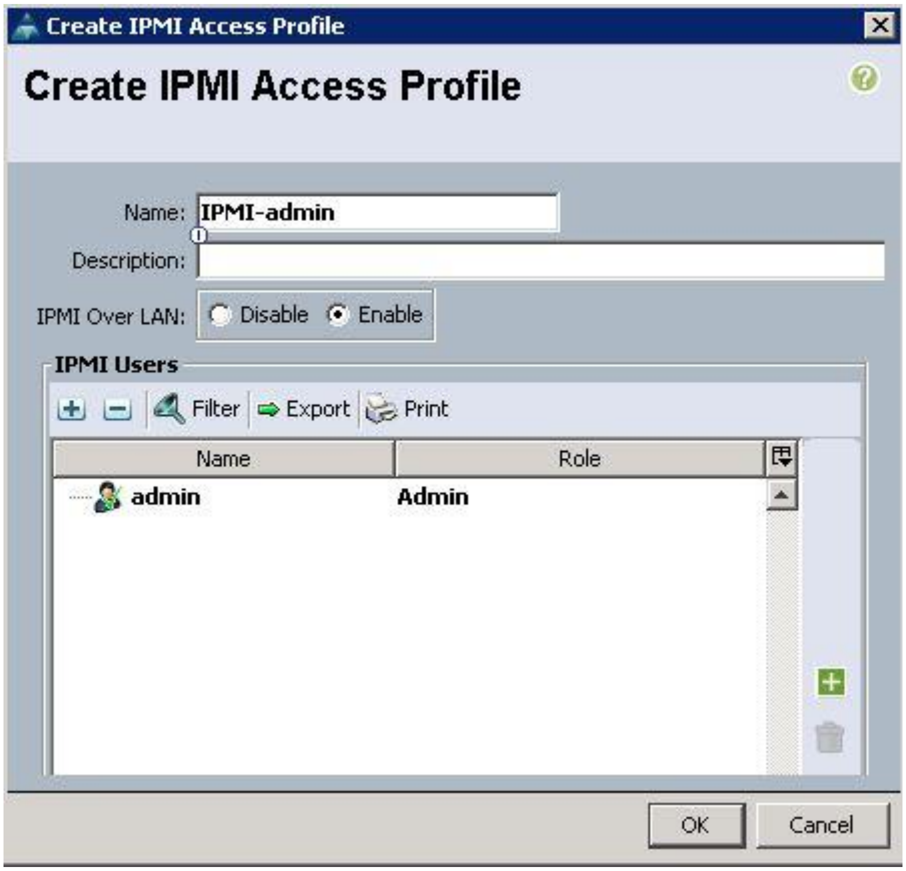
Name: admin

Password: *********

Confirm Password: *********

Role: ○ Read Only ● Admin

Click OK to create the IPMI access profile.



Create IPMI Access Profile

Name: IPMI-admin

Description:

IPMI Over LAN: ○ Disable ● Enable

IPMI Users

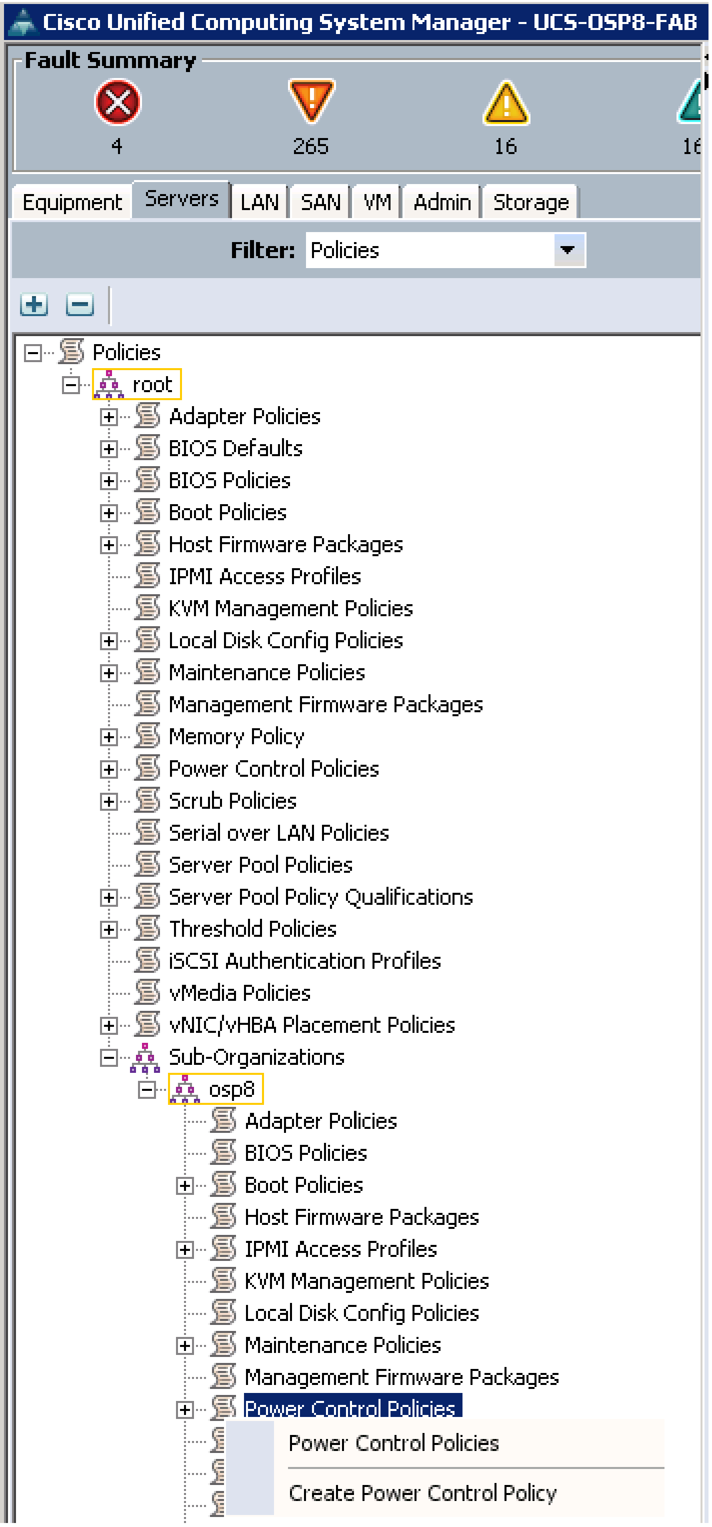
| Name | Role |
|---|---|
| admin | Admin |

## Create a Power Policy

Cisco UCS uses the priority set in the power control policy, along with the blade type and configuration, to calculate the initial power allocation for each blade within a chassis. During normal operation, the active blades within a chassis can borrow power from idle blades within the same chassis. If all blades are active and reach the power cap, service profiles with higher priority power control policies take precedence over service profiles with lower priority power control policies.

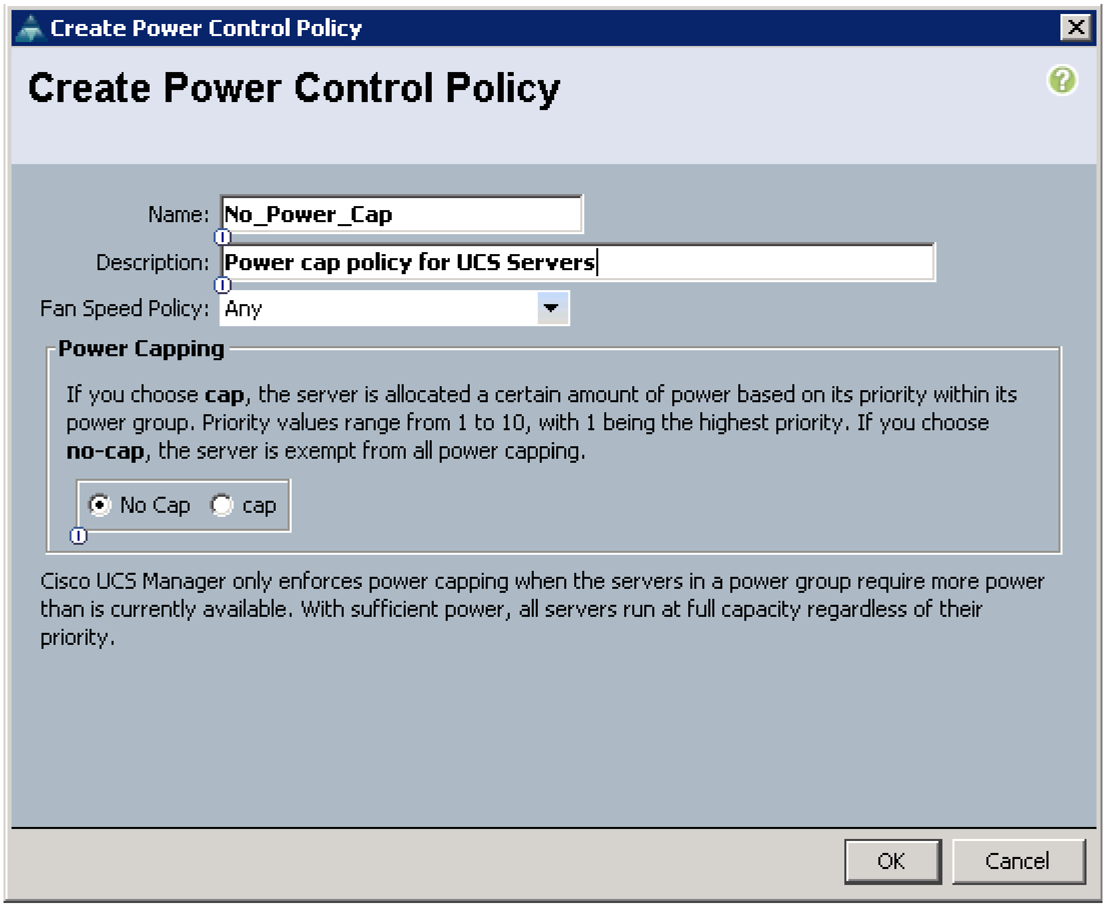
To configure the Power Control policy from the UCS Manager, complete the following steps:

> Under Server → Policies → root → Sub-Organizations -> Power Control Policies → right-click and select Create Power Control Policy.



Specify the name and description. Choose Power Capping as No Cap.

No Cap keeps the server runs at full capacity regardless of the power requirements of the other servers in its power group. Setting the priority to no-cap prevents Cisco UCS from leveraging unused power from that particular blade server. The server is allocated the maximum amount of power that that blade can reach.



## Create Power Control Policy

Name: No_Power_Cap

Description: Power cap policy for UCS Servers

Fan Speed Policy: Any

**Power Capping**

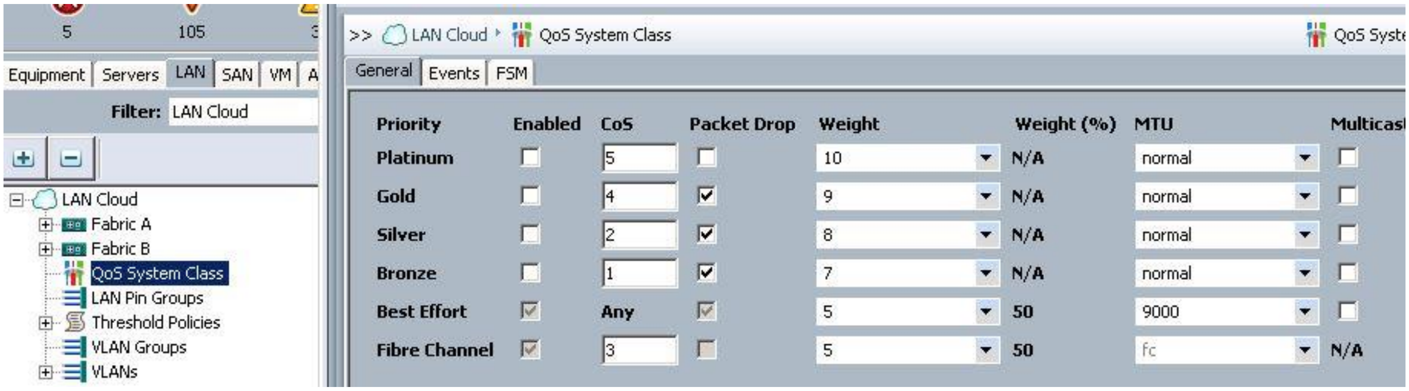
If you choose **cap**, the server is allocated a certain amount of power based on its priority within its power group. Priority values range from 1 to 10, with 1 being the highest priority. If you choose **no-cap**, the server is exempt from all power capping.

● No Cap ○ cap

Cisco UCS Manager only enforces power capping when the servers in a power group require more power than is currently available. With sufficient power, all servers run at full capacity regardless of their priority.

## Create a QOS system class

Create a QOS system class as shown below:



| Priority | Enabled | CoS | Packet Drop | Weight | Weight (%) | MTU | Multicast |
|---|---|---|---|---|---|---|---|
| Platinum | □ | 5 | □ | 10 | N/A | normal | □ |
| Gold | □ | 4 | ☑ | 9 | N/A | normal | □ |
| Silver | □ | 2 | ☑ | 8 | N/A | normal | □ |
| Bronze | □ | 1 | ☑ | 7 | N/A | normal | □ |
| Best Effort | ☑ | Any | ☑ | 5 | 50 | 9000 | □ |
| Fibre Channel | ☑ | 3 | □ | 5 | 50 | fc | N/A |

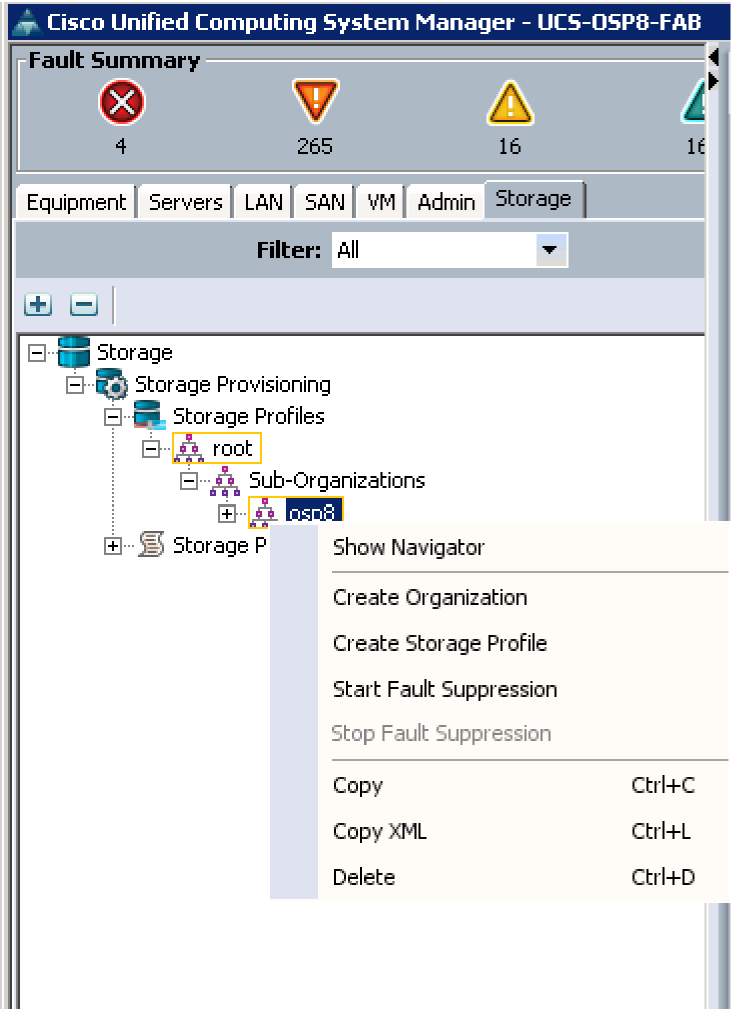
Select the Best Effort class as MTU 9000, which will be leveraged in vNIC templates for storage public and storage management vNICs.

## Create Storage Profiles for the Controller and Compute Blades

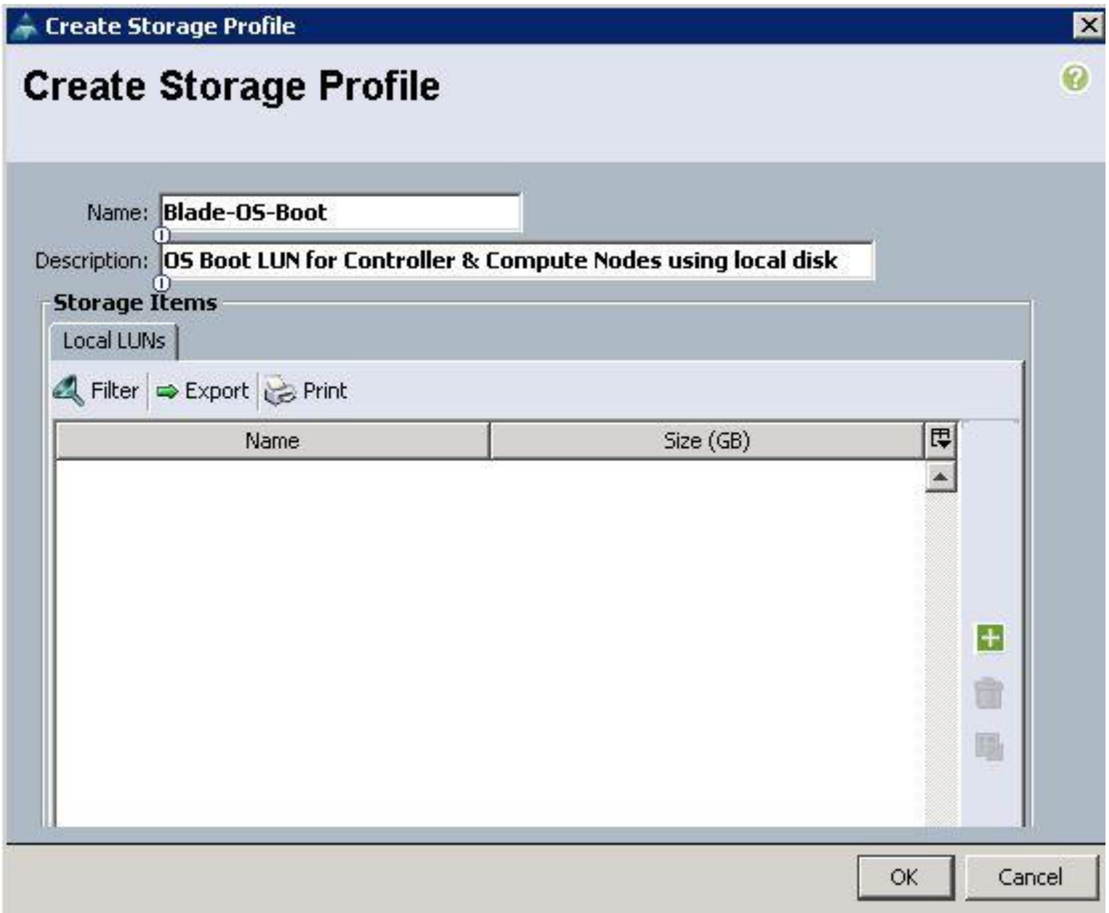
To allow flexibility in defining the number of storage disks, roles and usage of these disks, and other storage parameters, you can create and use storage profiles. LUNs configured in a storage profile can be used as boot LUNs or data LUNs, and can be dedicated to a specific server. You can also specify a local LUN as a boot device. However, LUN resizing is not supported.

To configure Storage profiles from the Cisco UCS Manager, complete the following steps:

Under Storage → Storage Provisioning → Storage Profiles → root → Sub-Organizations -> right-click and select Create Storage Profile.



Specify the name and click "+".

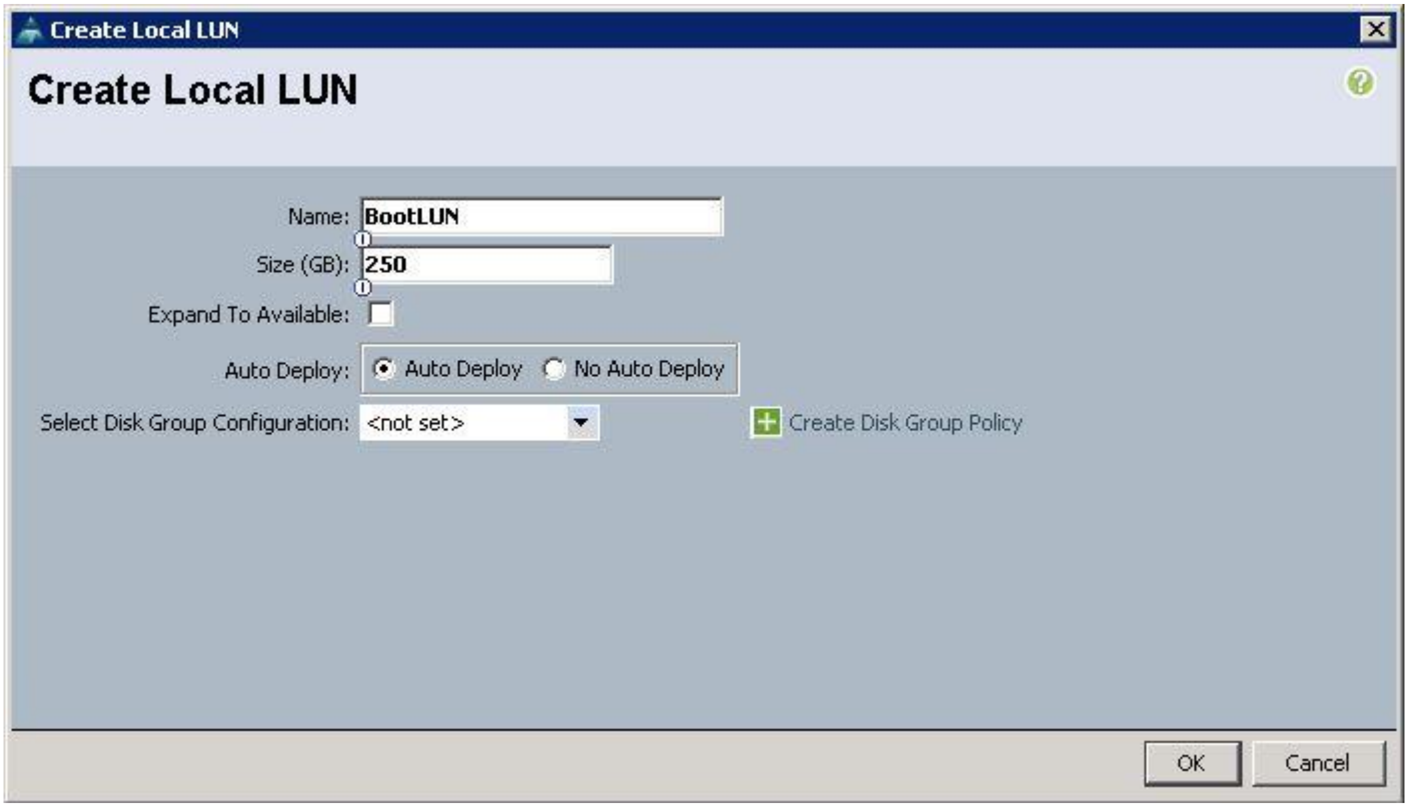


Name: Blade-OS-Boot

Description: OS Boot LUN for Controller & Compute Nodes using local disk

Specify the Local LUN name and size as 250 in GB and click Auto Deploy.

To configure RAID levels and configure the number of disks for the disk group, select Create Disk Group Policy.



Name: BootLUN

Size (GB): 250

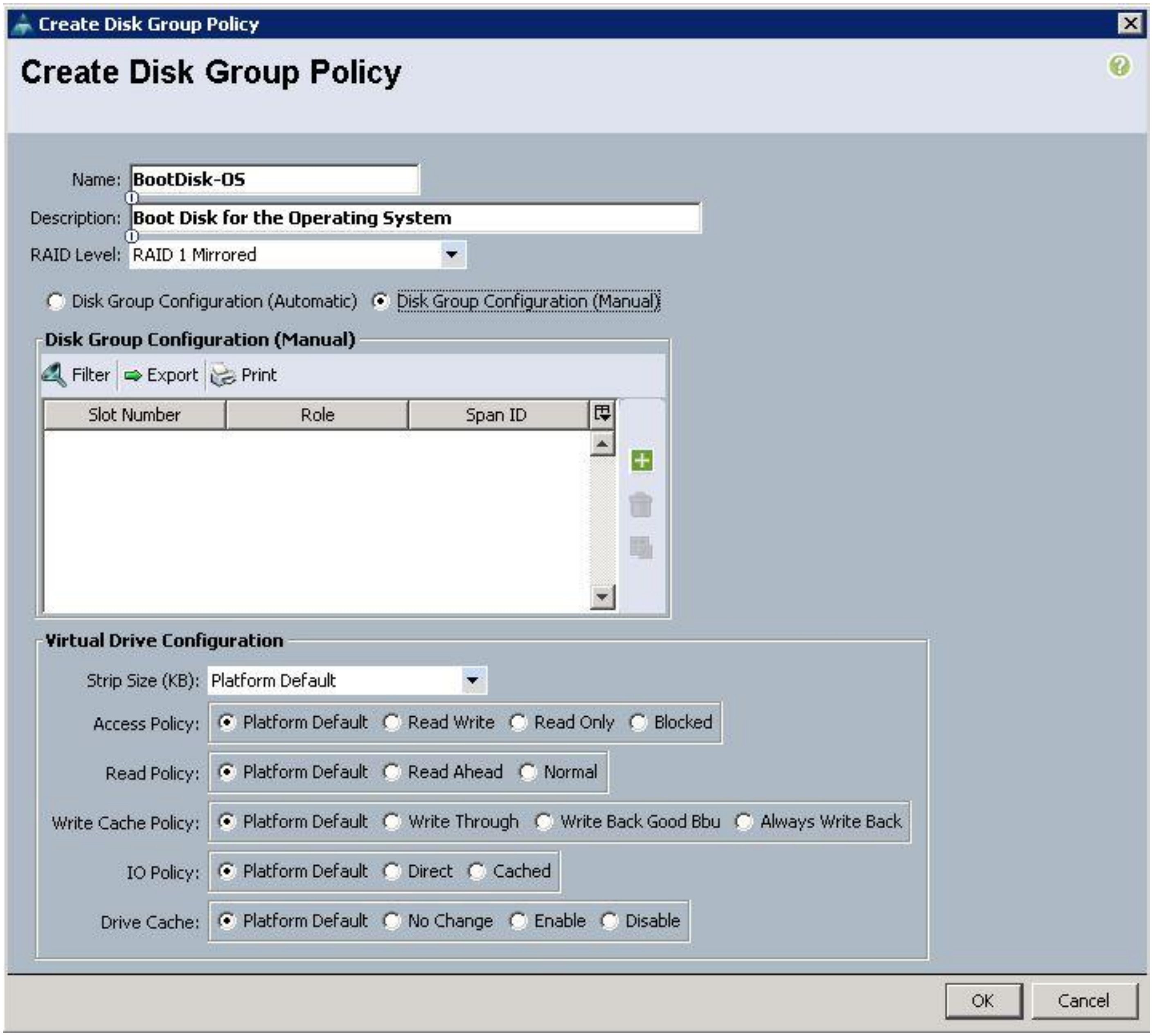
Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: <not set>

Specify the name and choose RAID level as RAID 1 Mirrored. RAID1 is recommended for the Local boot LUNs.

Select Disk group Configuration (Manual) and click “+”. Keep the Virtual Drive configuration with the default values.



Create Disk Group Policy

Name: BootDisk-OS

Description: Boot Disk for the Operating System

RAID Level: RAID 1 Mirrored

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

Disk Group Configuration (Manual)

| Slot Number | Role | Span ID |
|---|---|---|

Virtual Drive Configuration

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

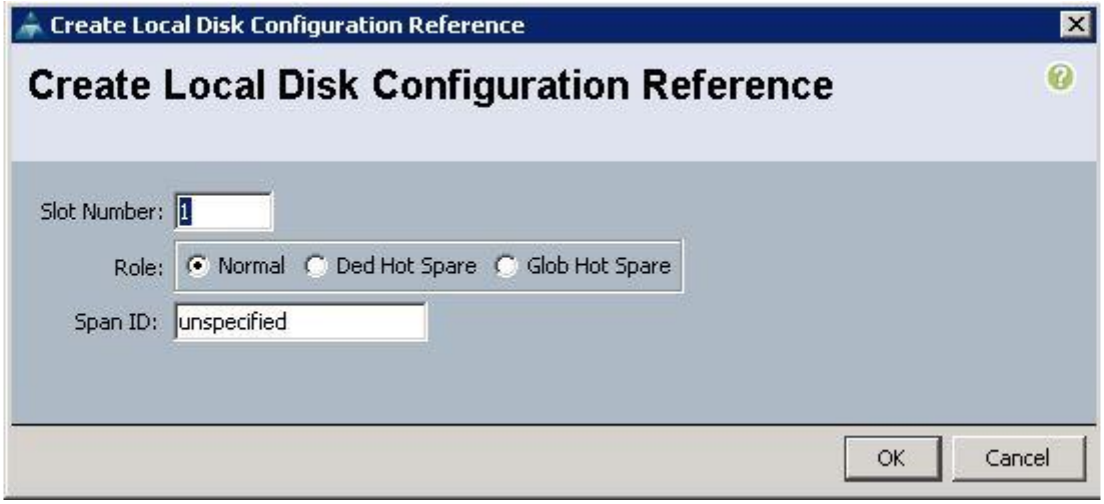
Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

Specify Disk Slot Number as 1 and Role as Normal.


Create Local Disk Configuration Reference
Create Local Disk Configuration Reference

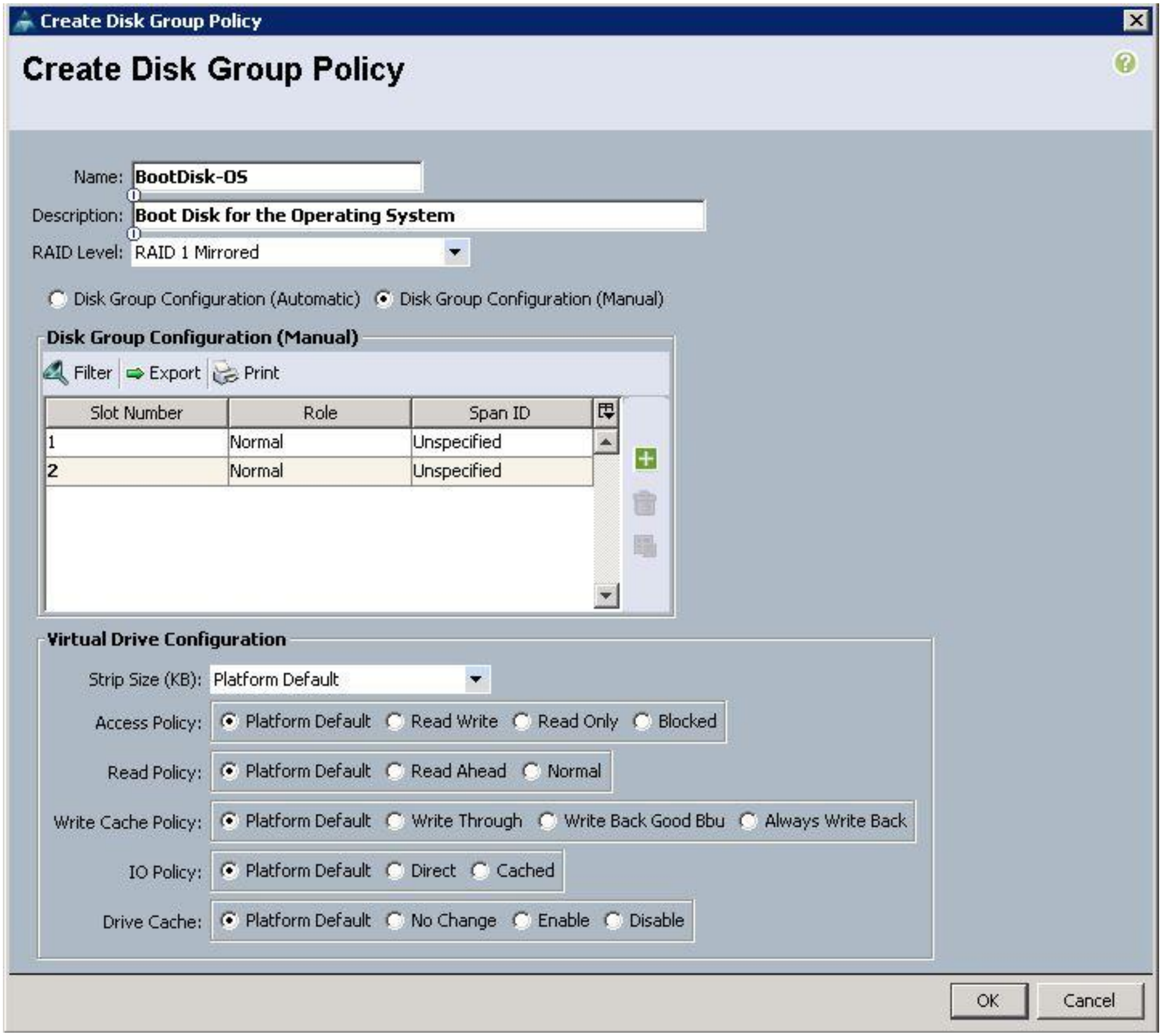
Slot Number: 1

Role: ● Normal ○ Ded Hot Spare ○ Glob Hot Spare

Span ID: unspecified

Create another Local Disk configuration with the Slot number as 2 and click OK.

In this solution, we used Local Disk 1 and Disk 2 as the boot LUNs with RAID 1 mirror configuration.



Create Disk Group Policy

Name: BootDisk-OS

Description: Boot Disk for the Operating System

RAID Level: RAID 1 Mirrored

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

Disk Group Configuration (Manual)

| Slot Number | Role | Span ID |
|---|---|---|
| 1 | Normal | Unspecified |
| 2 | Normal | Unspecified |

Virtual Drive Configuration

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

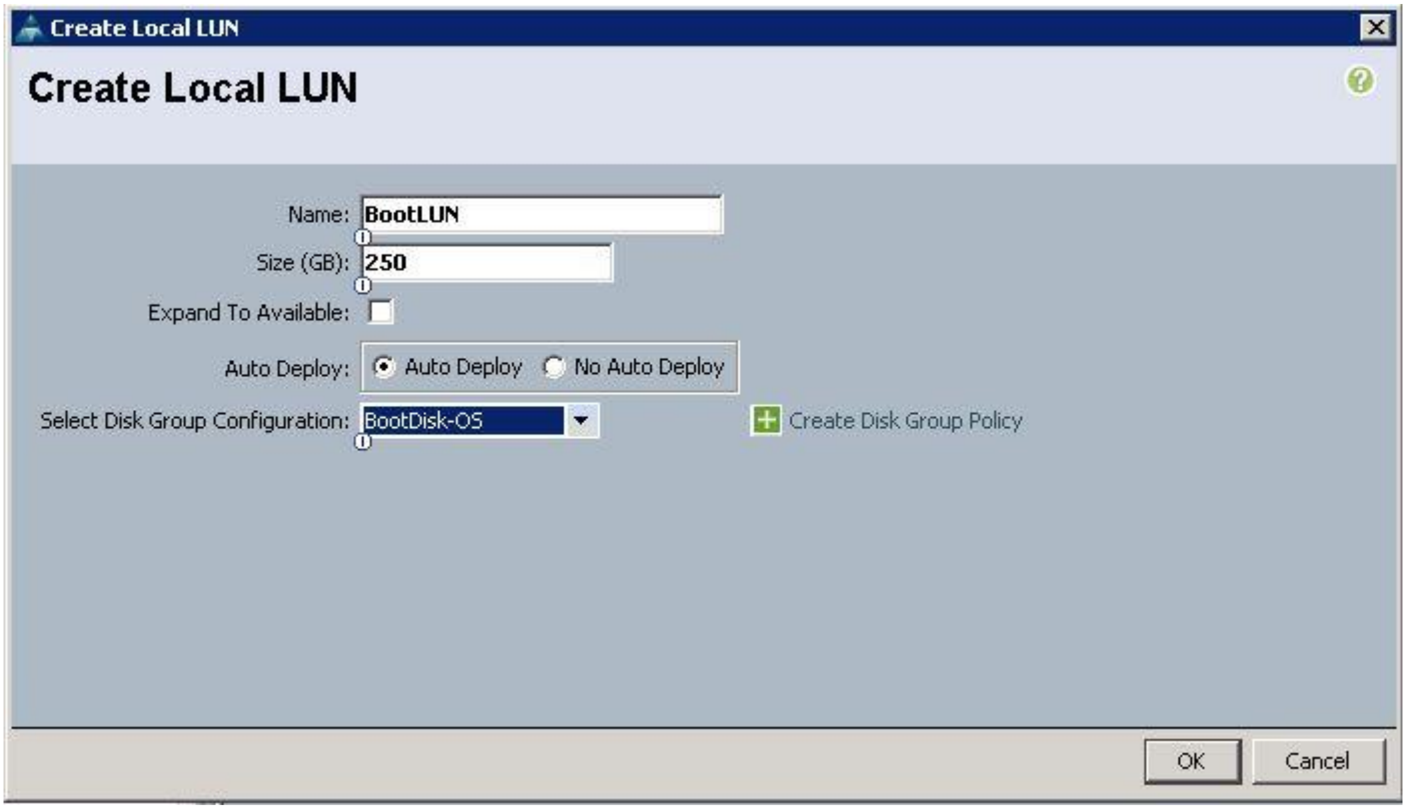
Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

Choose the Disk group policy Boot Disk-OS for the Local Boot LUN.



Create Local LUN

Name: BootLUN

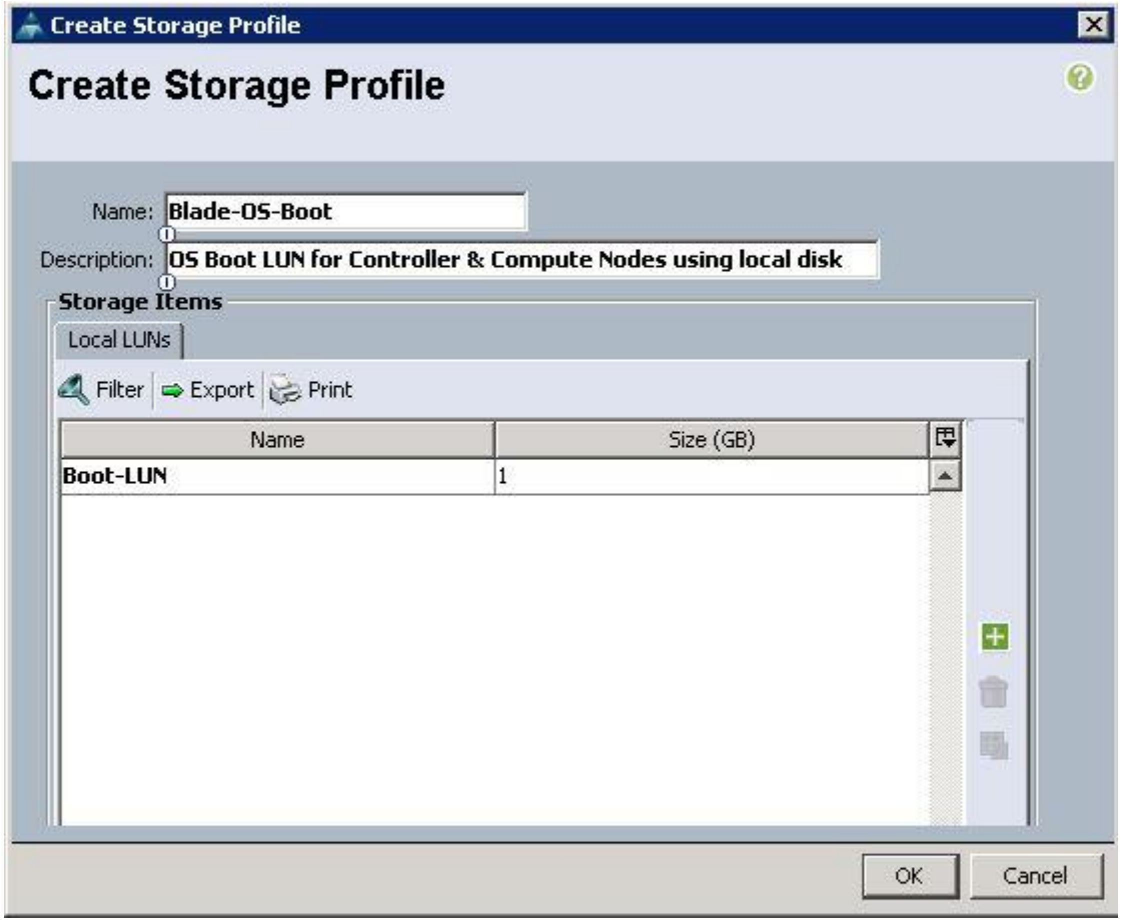
Size (GB): 250

Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: BootDisk-OS

Click OK to confirm the Storage profile creation.



Create Storage Profile

Name: Blade-OS-Boot

Description: OS Boot LUN for Controller & Compute Nodes using local disk

Storage Items

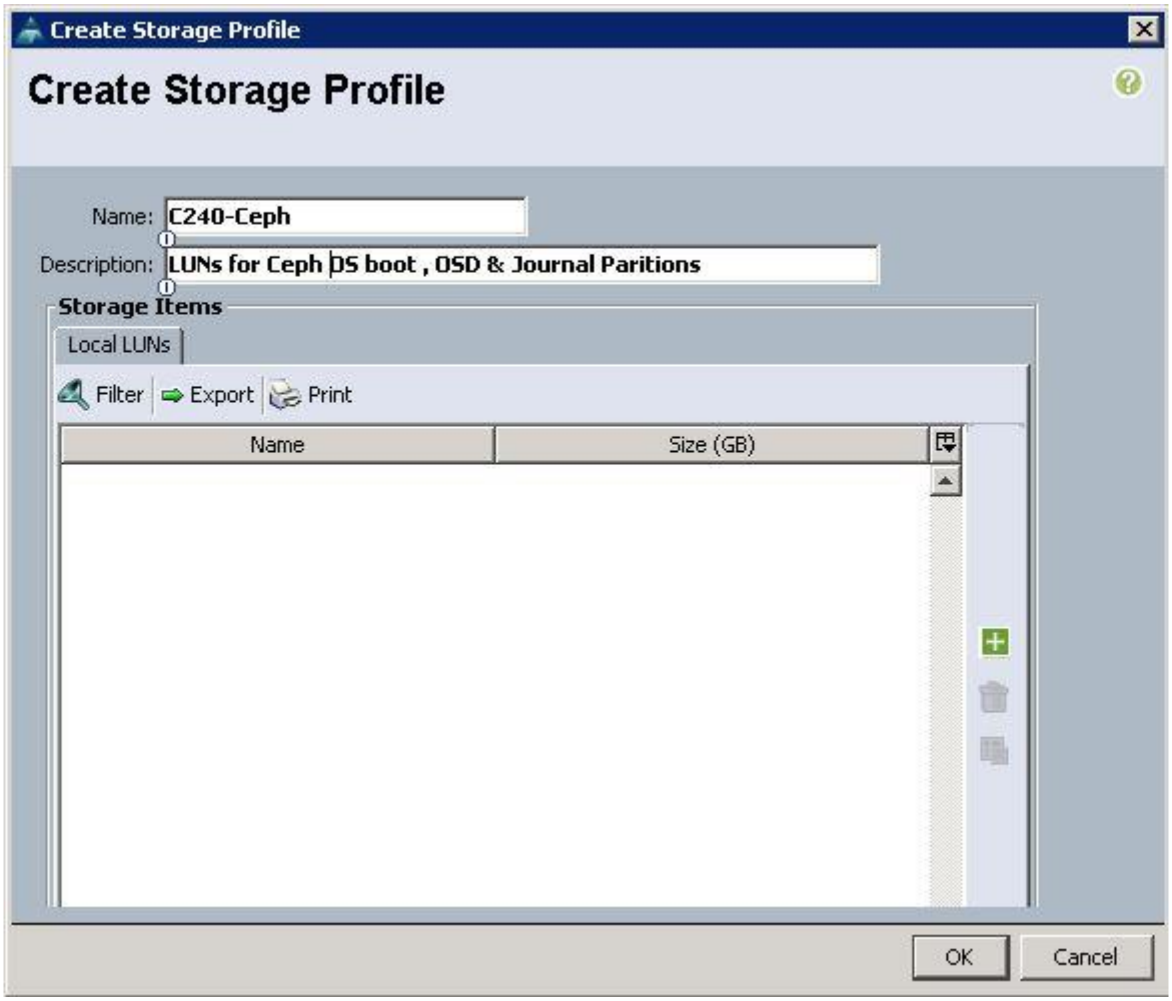
Local LUNs

| Name | Size (GB) |
|---|---|
| Boot-LUN | 1 |

## Create Storage Profiles for Cisco UCS C240 M4 Blade Server

To configure the Storage profiles from the Cisco UCS Manager, complete the following steps:

Specify the Storage profile name as C240-Ceph for the Ceph Storage Servers. Click "+".

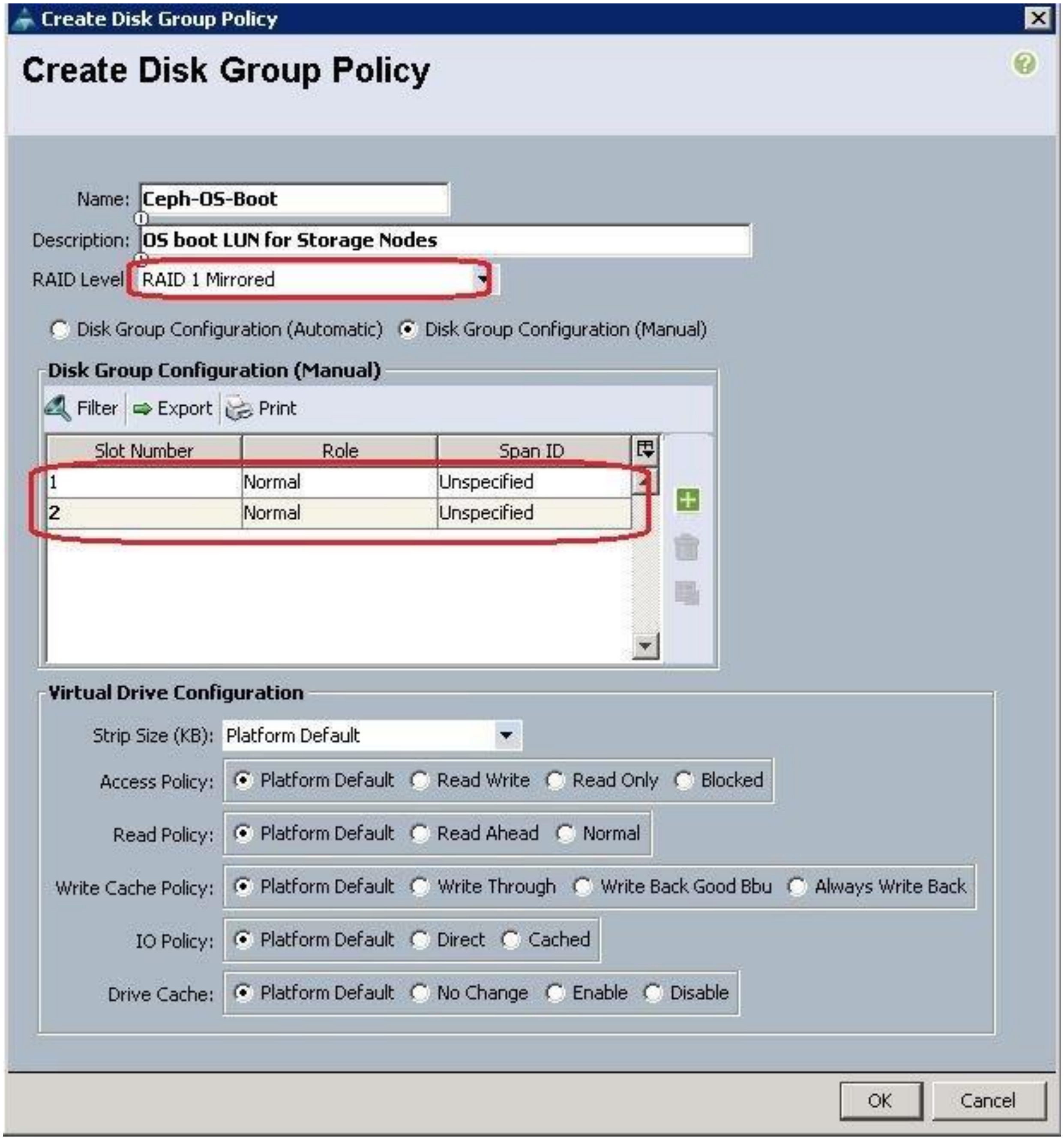


Create Storage Profile

Name: C240-Ceph

Description: LUNs for Ceph OS boot , OSD & Journal Paritions

Specify the LUN name and size in GB. For the Disk group policy creation, select Disk Group Configuration for Ceph nodes as Ceph-OS-Boot similar to "BootDisk-OS" disk group policy as above.



Create Disk Group Policy

Name: Ceph-OS-Boot

Description: OS boot LUN for Storage Nodes

RAID Level: RAID 1 Mirrored

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

Disk Group Configuration (Manual)

| Slot Number | Role | Span ID |
|---|---|---|
| 1 | Normal | Unspecified |
| 2 | Normal | Unspecified |

Virtual Drive Configuration

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

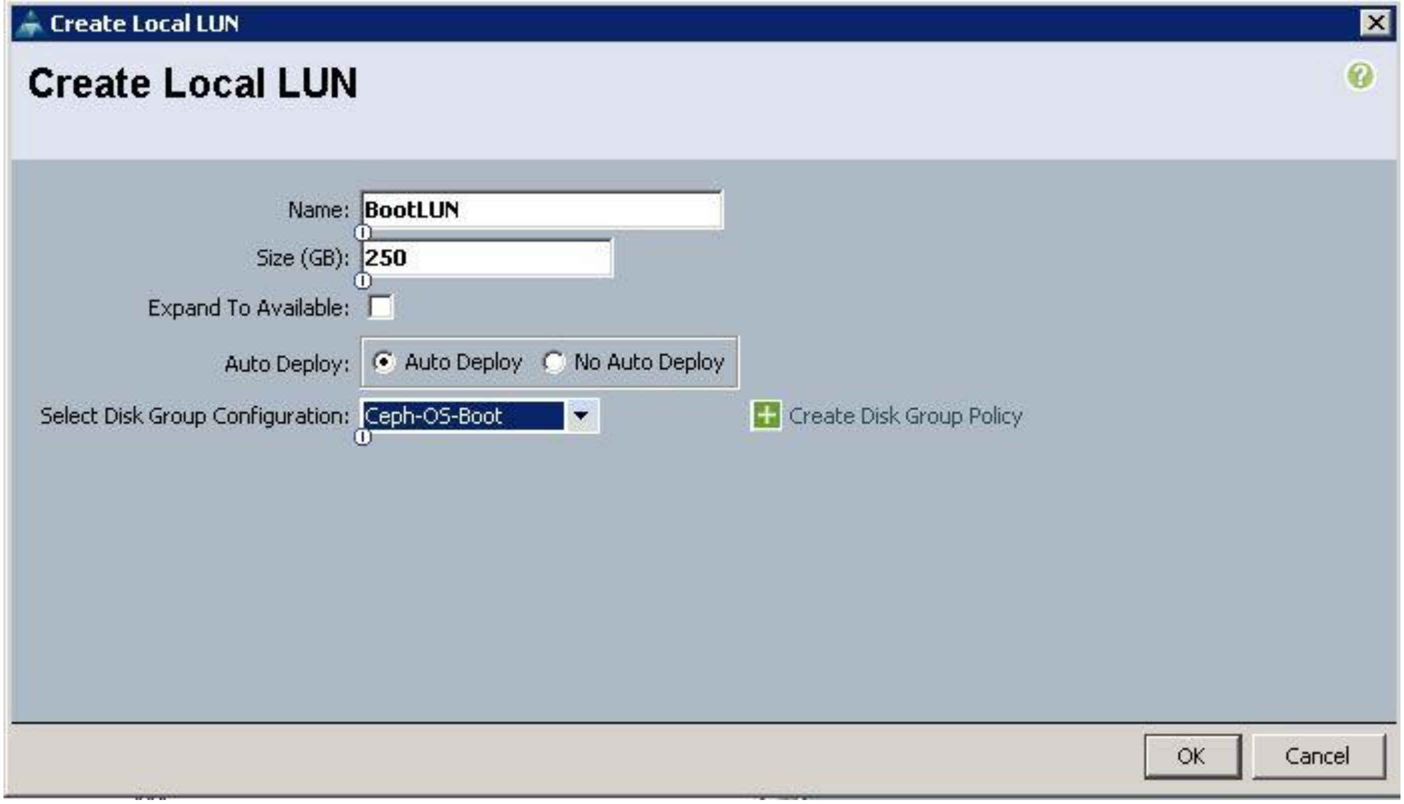
Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

After successful creation of Disk Group Policy, choose Disk Group Configuration as Ceph-OS-Boot and click OK.



Create Local LUN

Name: BootLUN

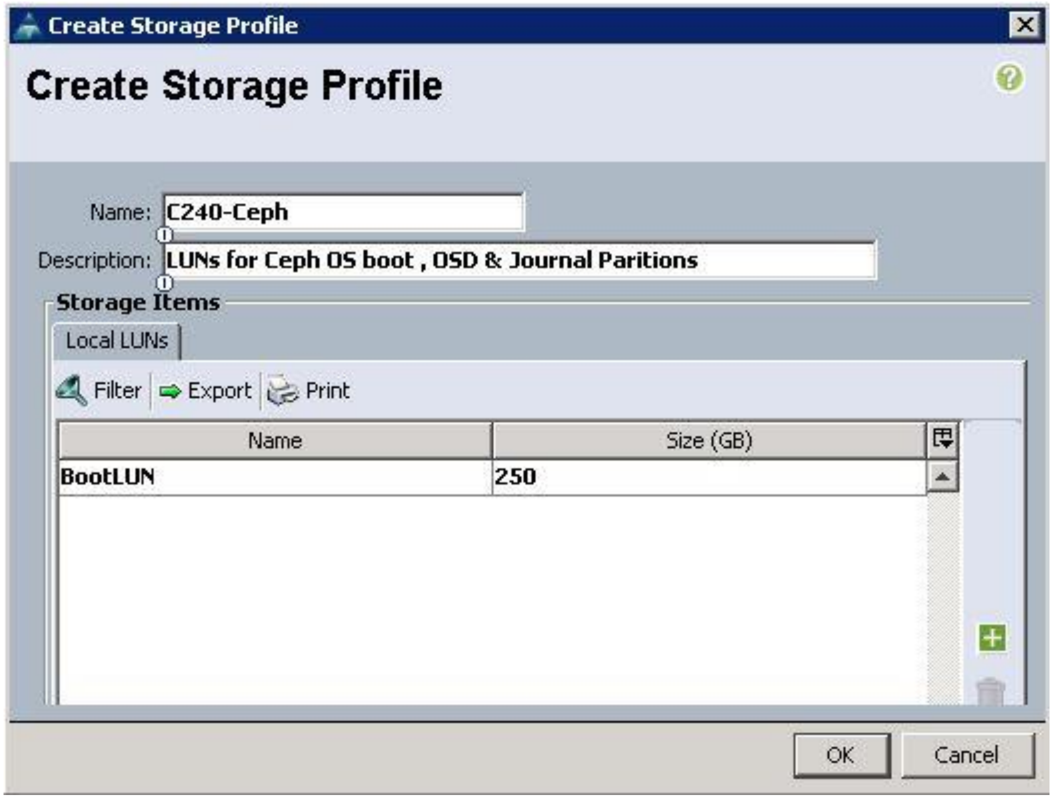
Size (GB): 250

Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: Ceph-OS-Boot

Click OK to complete the Storage Profile creation for the Ceph Nodes.



Create Storage Profile

Name: C240-Ceph

Description: LUNs for Ceph OS boot , OSD & Journal Paritions

Storage Items

Local LUNs

| Name | Size (GB) |
|---|---|
| BootLUN | 250 |

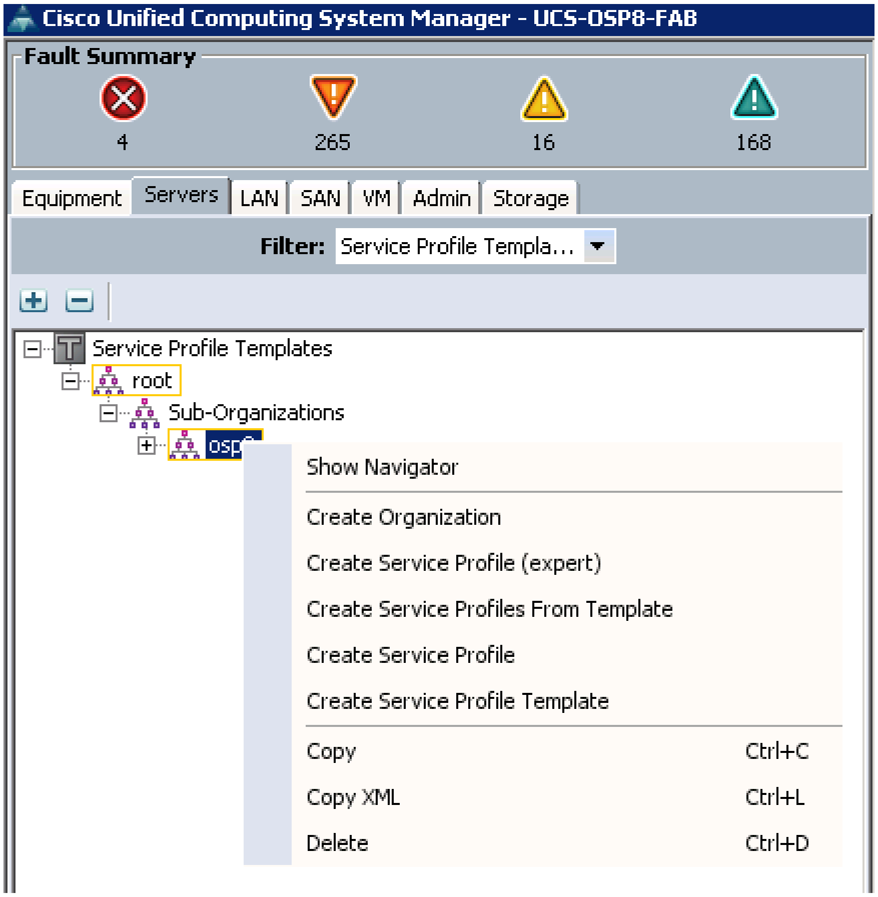
For the Cisco UCS C240 M4 servers, the LUN creation for Ceph OSD disks (6TB SAS) and Ceph Journal disks (400GB SSDs) still remains on the Ceph Storage profile. Due to the Cisco UCS Manager limitations, we have to create OSD

LUNs and Journal LUNs after the Cisco UCS C240 M4 server has been successfully associated with the Ceph Storage Service profiles.

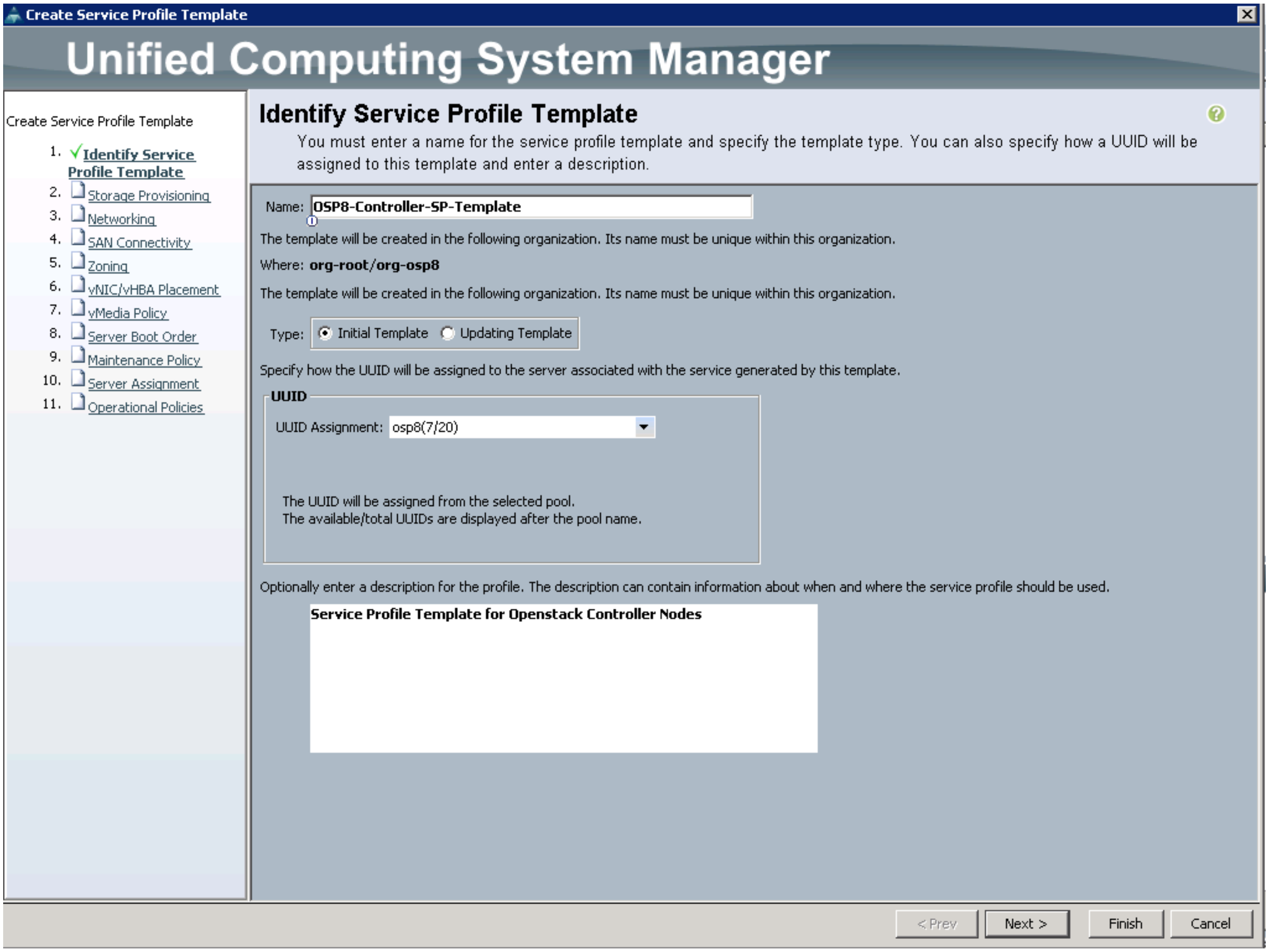
## Create Service Profile Templates for Controller Nodes

To configure the Service Profile Templates for the Controller Nodes, complete the following steps:

Under Servers → Service Profile Templates → root → Sub-Organizations -> right-click and select Create Service Profile Template.


Cisco Unified Computing System Manager - UCS-OSP8-FAB
Fault Summary
4
265
16
168
Equipment
Servers
LAN
SAN
VM
Admin
Storage
Filter: Service Profile Templa...
Service Profile Templates
root
Sub-Organizations
osp8
Show Navigator
Create Organization
Create Service Profile (expert)
Create Service Profiles From Template
Create Service Profile
Create Service Profile Template
Copy Ctrl+C
Copy XML Ctrl+L
Delete Ctrl+D

Specify the Service profile template name for the Controller node as OSP8-Controller-SP-Template. Choose the UUID pools previously created from the drop-down list and click Next.



**Identify Service Profile Template**

You must enter a name for the service profile template and specify the template type. You can also specify how a UUID will be assigned to this template and enter a description.

Name: OSP8-Controller-SP-Template

The template will be created in the following organization. Its name must be unique within this organization.

Where: **org-root/org-osp8**

The template will be created in the following organization. Its name must be unique within this organization.

Type: ● Initial Template ○ Updating Template

Specify how the UUID will be assigned to the server associated with the service generated by this template.

**UUID**

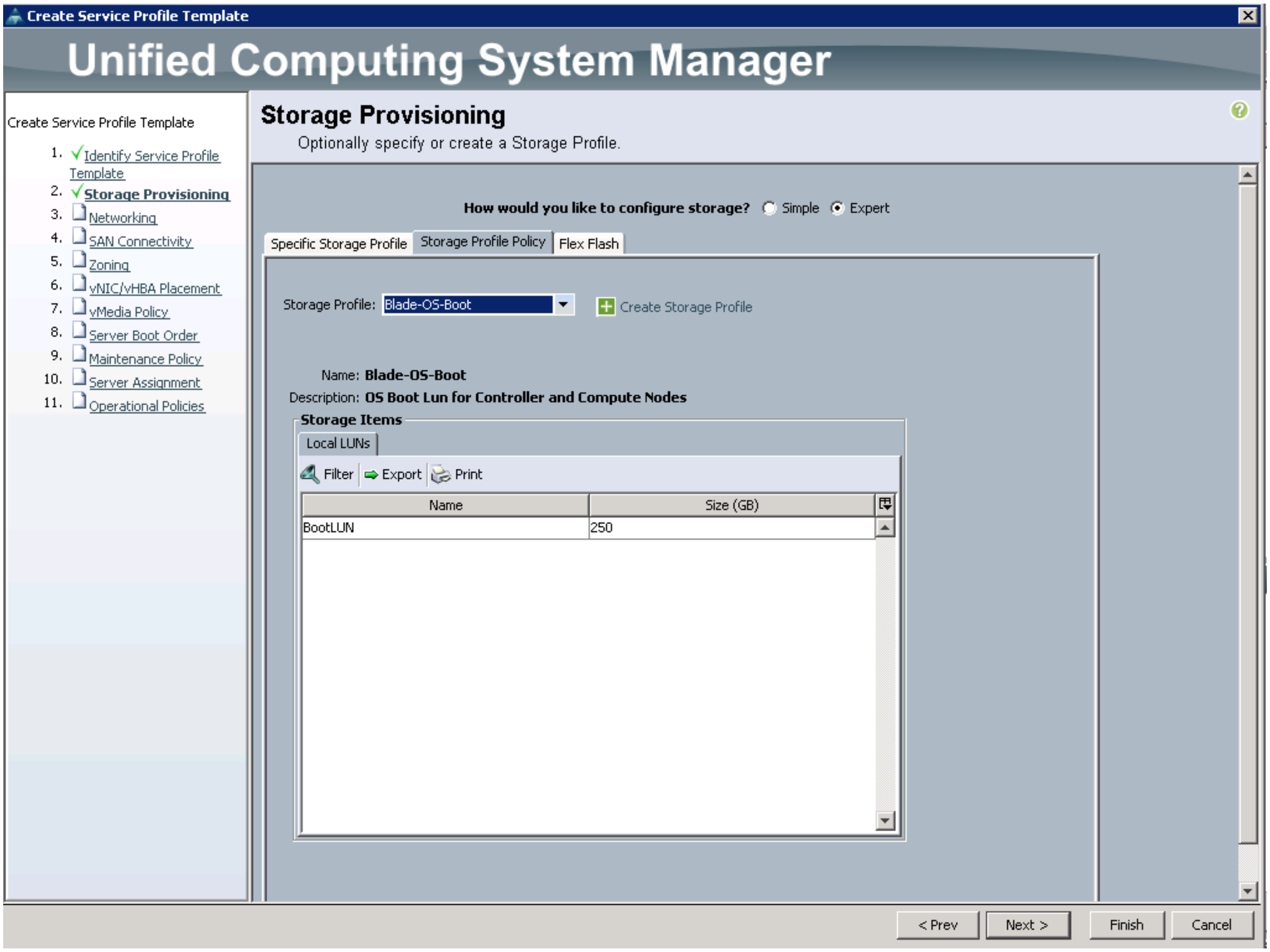
UUID Assignment: osp8(7/20)

The UUID will be assigned from the selected pool.
The available/total UUIDs are displayed after the pool name.

Optionally enter a description for the profile. The description can contain information about when and where the service profile should be used.

Service Profile Template for Openstack Controller Nodes

For Storage Provisioning, choose Expert and click Storage profile Policy and choose the Storage profile Blade-OS-boot previously created from the drop-down list and click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. Networking
4. SAN Connectivity
5. Zoning
6. vNIC/vHBA Placement
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

Storage Provisioning

Optionally specify or create a Storage Profile.

How would you like to configure storage? ○ Simple ● Expert

Storage Profile: Blade-OS-Boot

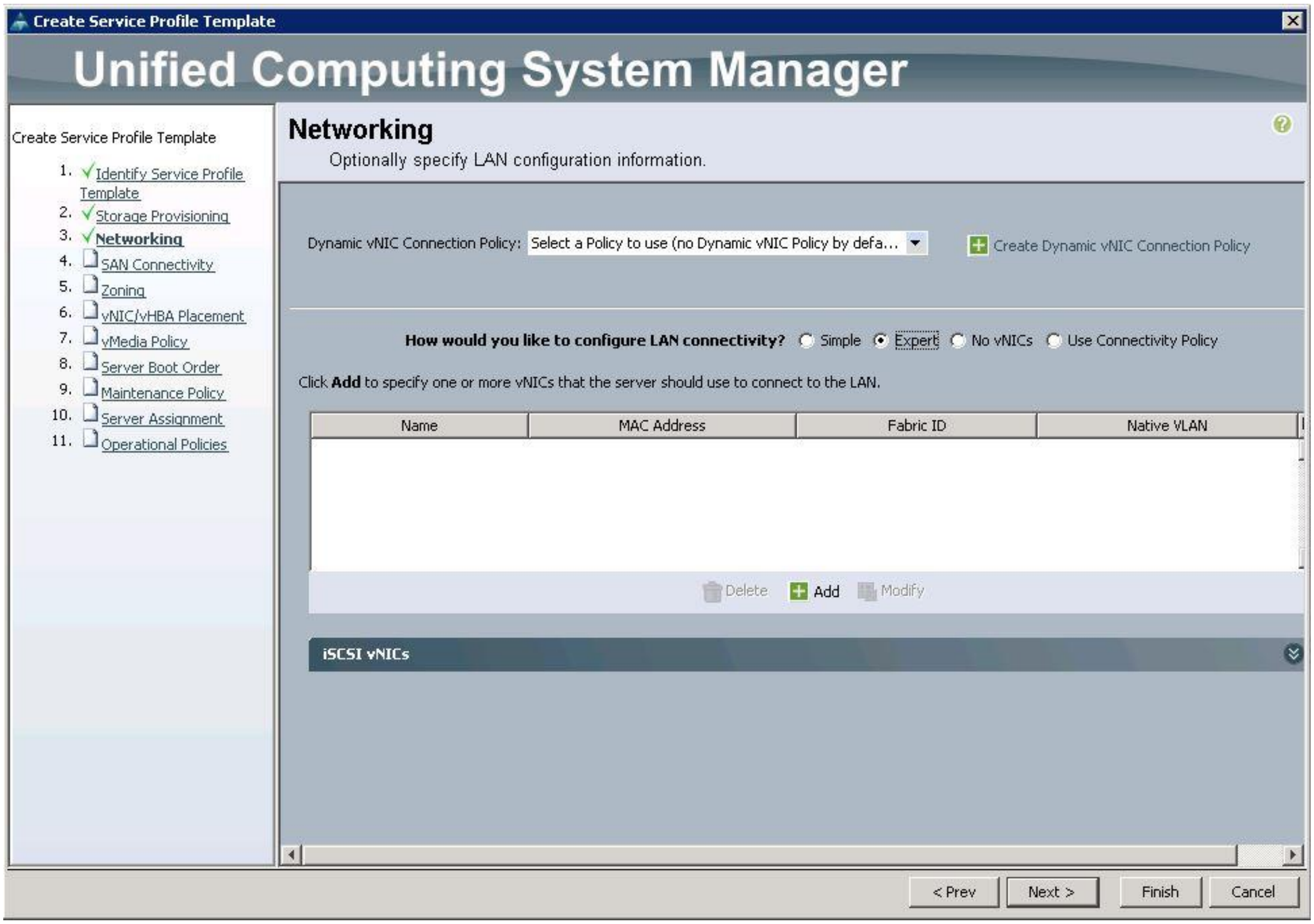
Name: **Blade-OS-Boot**

Description: **OS Boot Lun for Controller and Compute Nodes**

**Storage Items**

| Name | Size (GB) |
|---|---|
| BootLUN | 250 |

For Networking, choose Expert and click “+”.



Networking

Optionally specify LAN configuration information.

Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

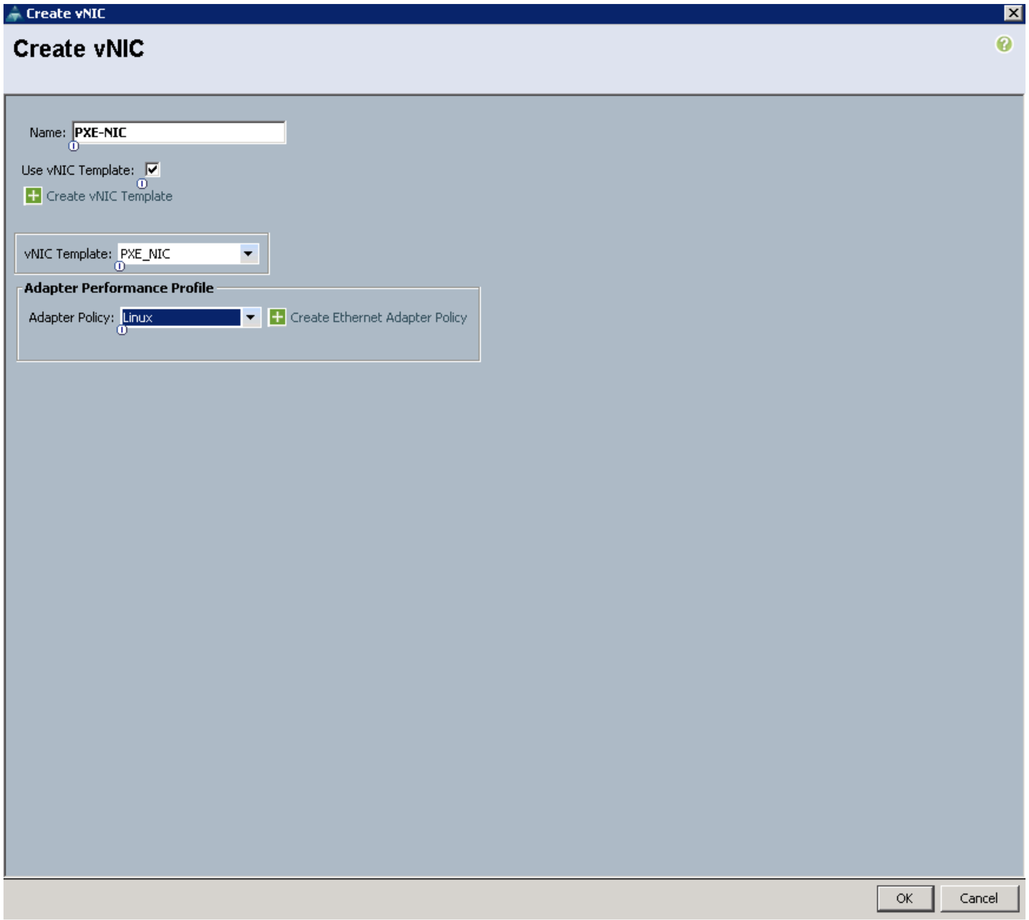
**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|

Create the VNIC interface for PXE or Provisioning network as PXE-NIC and click the check box Use VNIC template.

Under vNIC template, choose the PXE-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Name: PXE-NIC

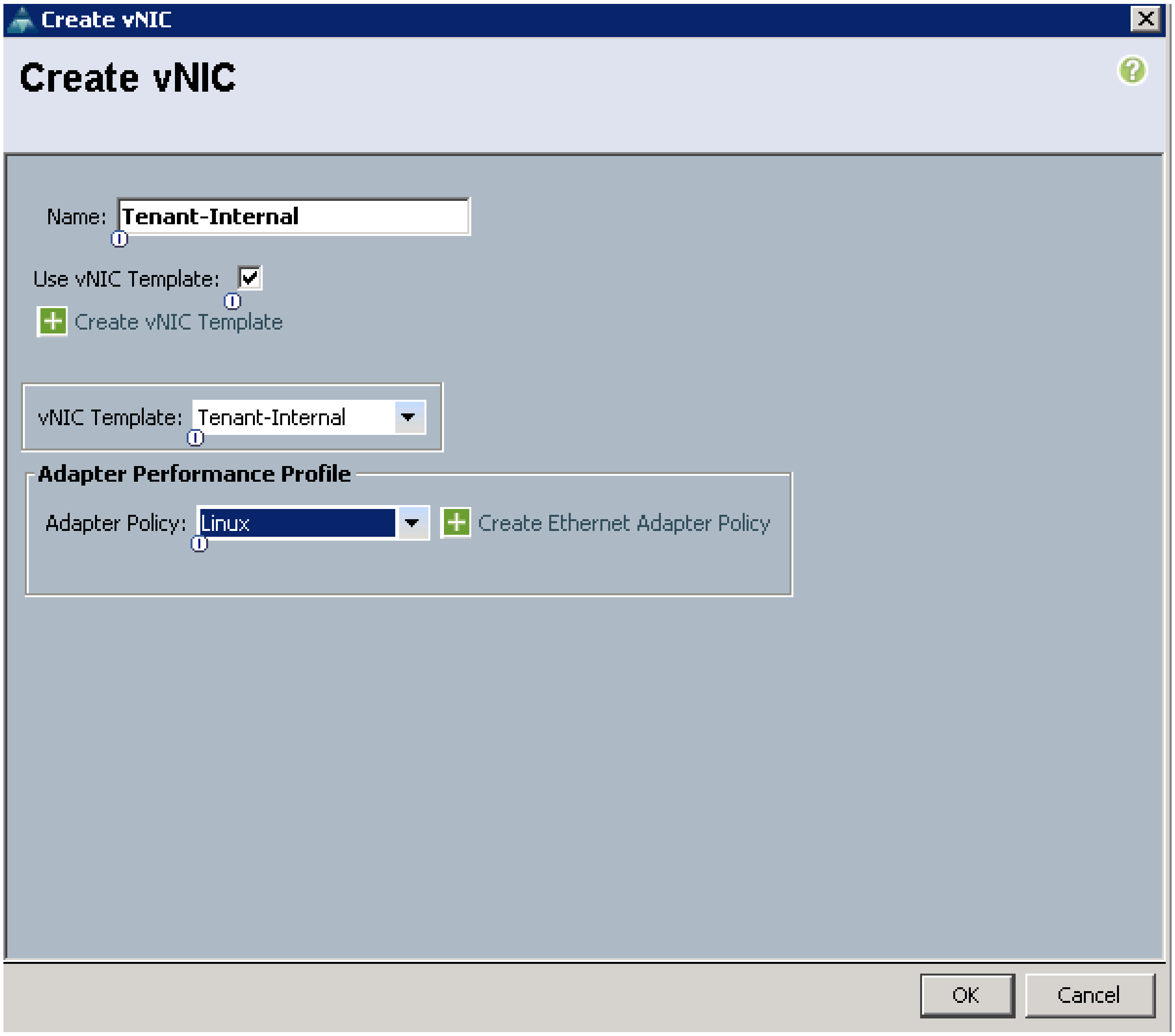
Use vNIC Template: ☑

vNIC Template: PXE_NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Tenant Internal Network as Tenant-Internal and then under vNIC template, choose the "Tenant-Internal" template we created before from the drop-down list and choose Adapter Policy as "Linux".



Create vNIC

Name: Tenant-Internal

Use vNIC Template: ☑

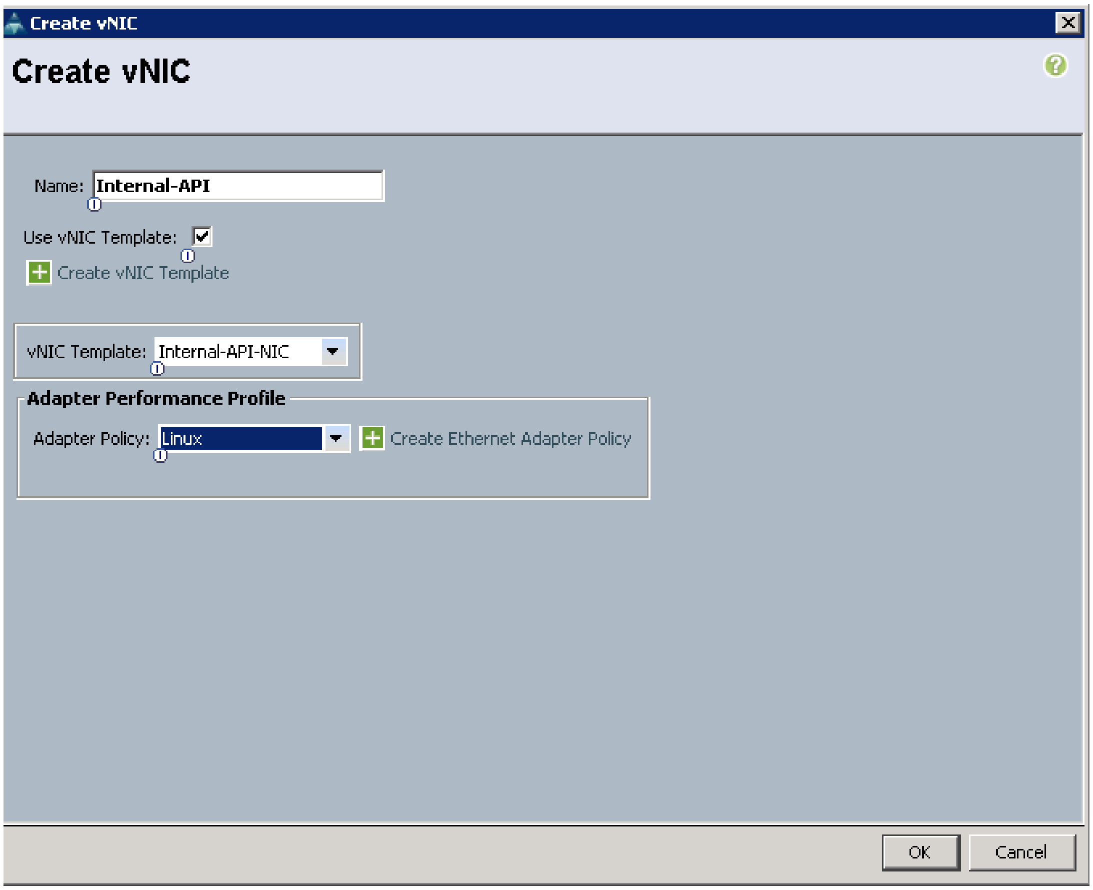
vNIC Template: Tenant-Internal

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Internal API network as Internal-API and click the check box for Use VNIC template.

Under vNIC template, choose the Internal-API-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: Internal-API

Use vNIC Template: ☑

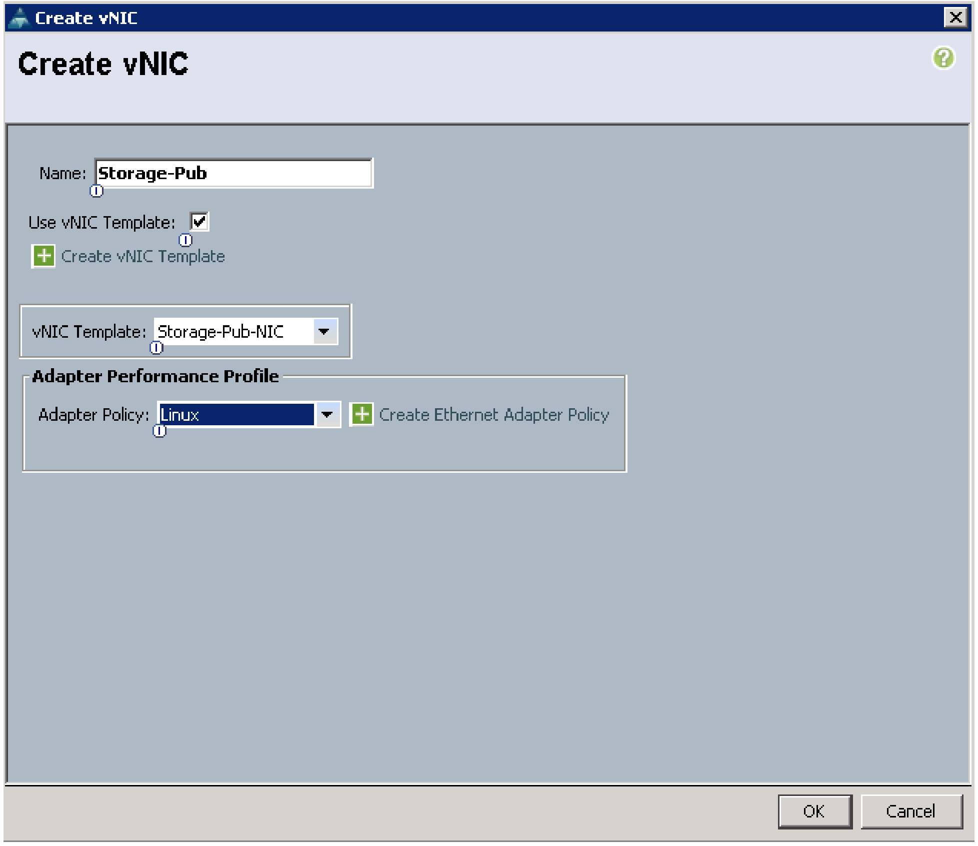
vNIC Template: Internal-API-NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Storage Public Network as Storage-Pub and click the check box for Use VNIC template.

Under vNIC template, choose the Storage-Pub-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Name: Storage-Pub

Use vNIC Template: ☑

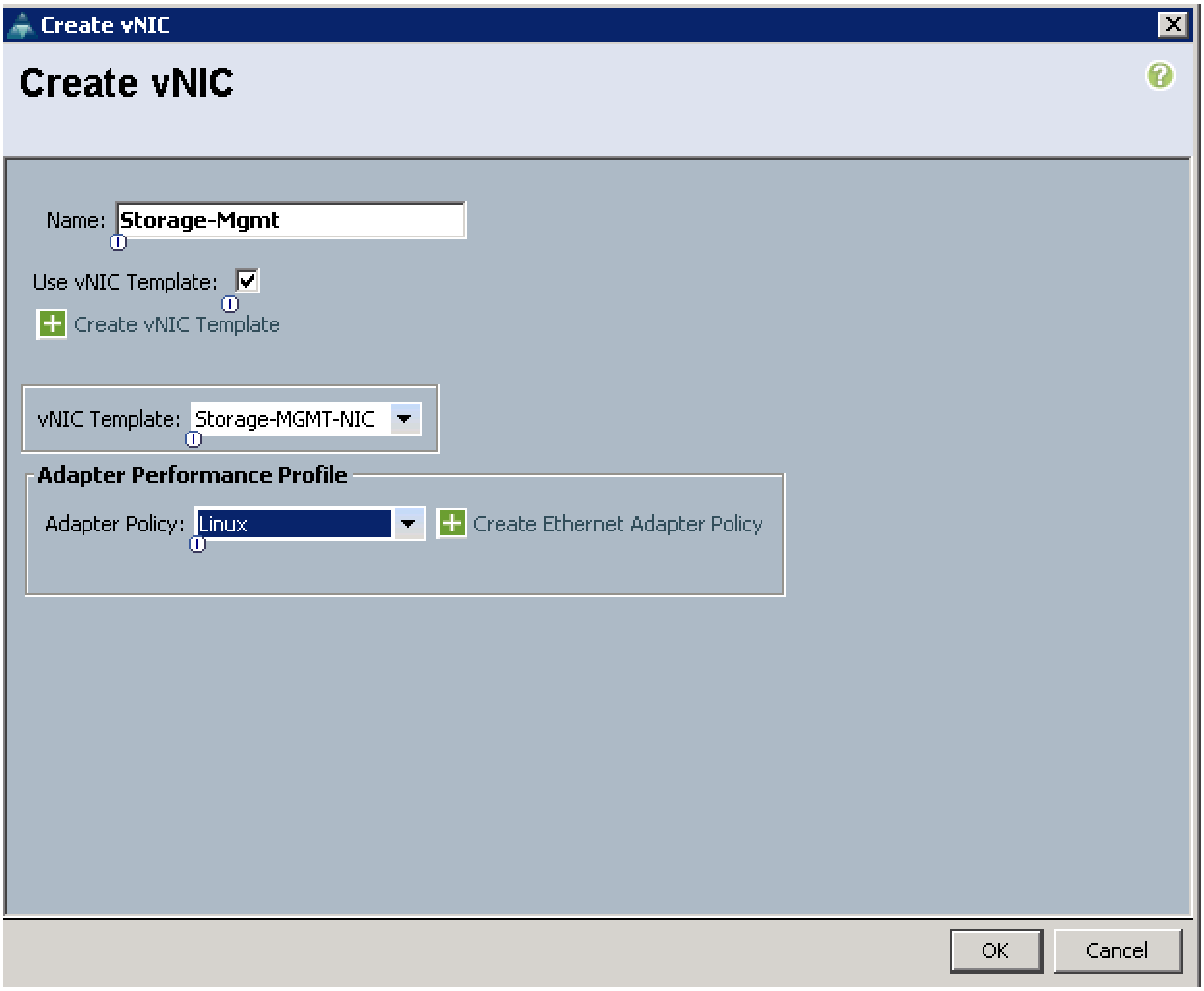
vNIC Template: Storage-Pub-NIC

**Adapter Performance Profile**

Adapter Policy: Linux

Create the VNIC interface for Storage Mgmt Cluster Network as Storage-Mgmt and click the check box for Use VNIC template.

Under vNIC template, choose the Storage-Mgmt-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: Storage-Mgmt

Use vNIC Template: ☑

Create vNIC Template

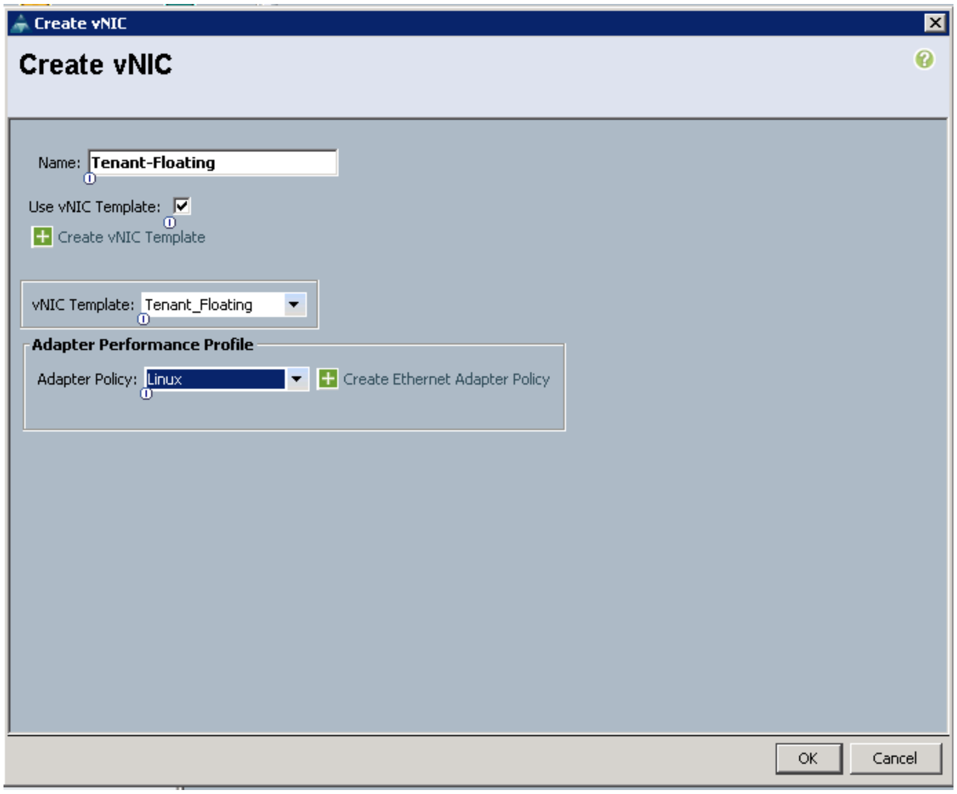
vNIC Template: Storage-MGMT-NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Floating Network as Tenant-Floating and click the check box the Use VNIC template.

Under the vNIC template, choose the Tenant-Floating template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: Tenant-Floating

Use vNIC Template: ☑

Create vNIC Template

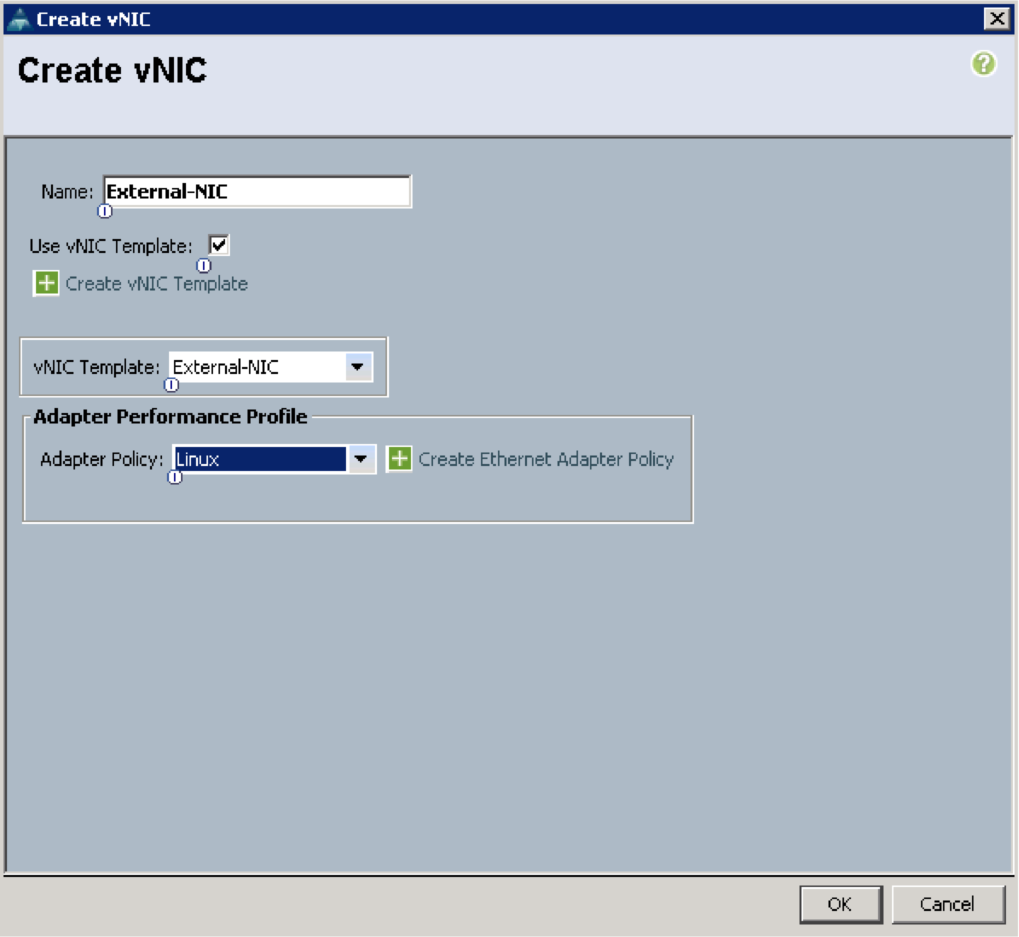
vNIC Template: Tenant_Floating

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for External Network as External-NIC and click the check box the Use VNIC template.

Under the vNIC template, choose the External-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: External-NIC

Use vNIC Template: ☑

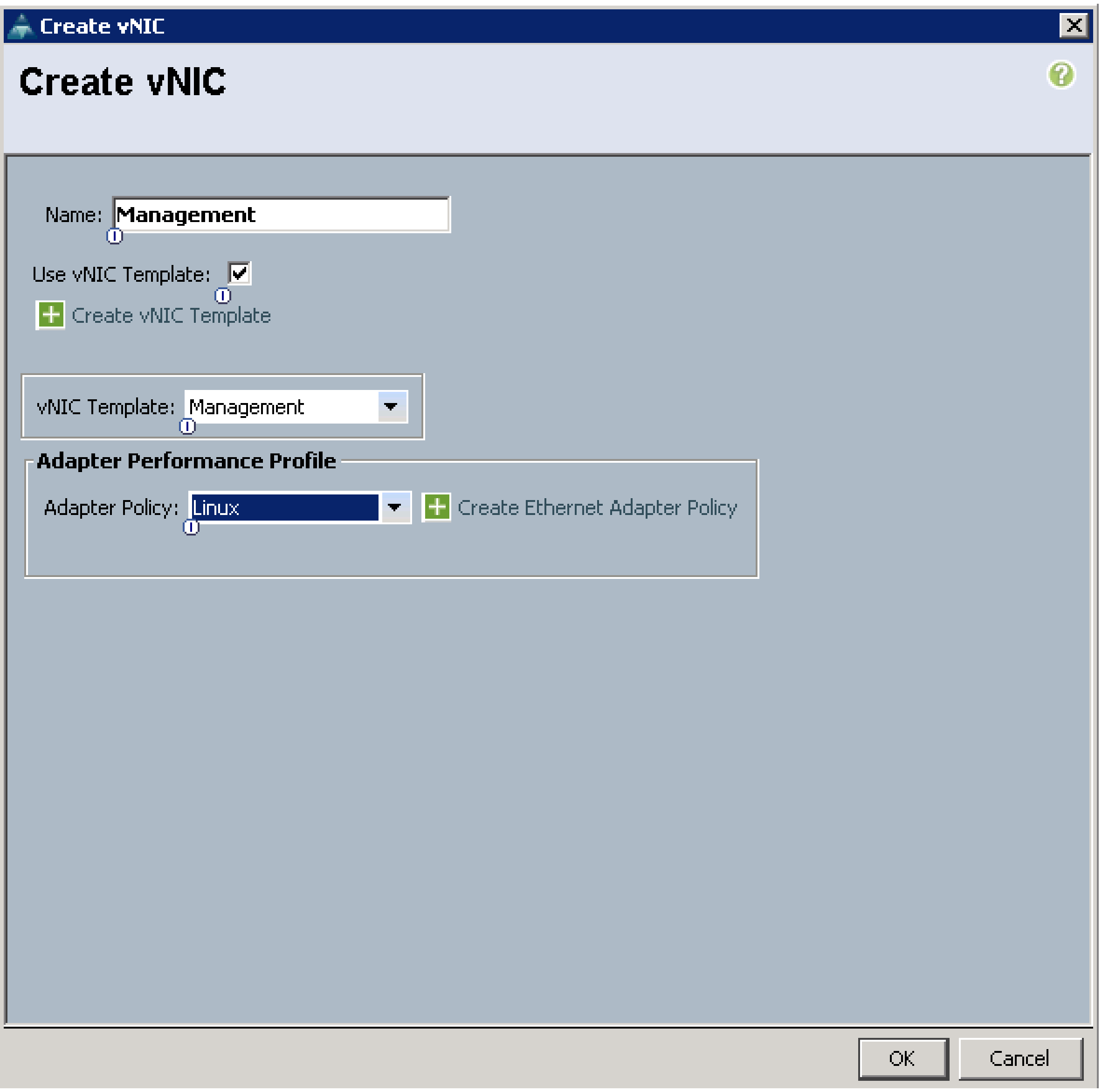
vNIC Template: External-NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Management Network as Management-NIC and click the check box the Use VNIC template.

Under the vNIC template, choose the Management-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: Management

Use vNIC Template: ☑

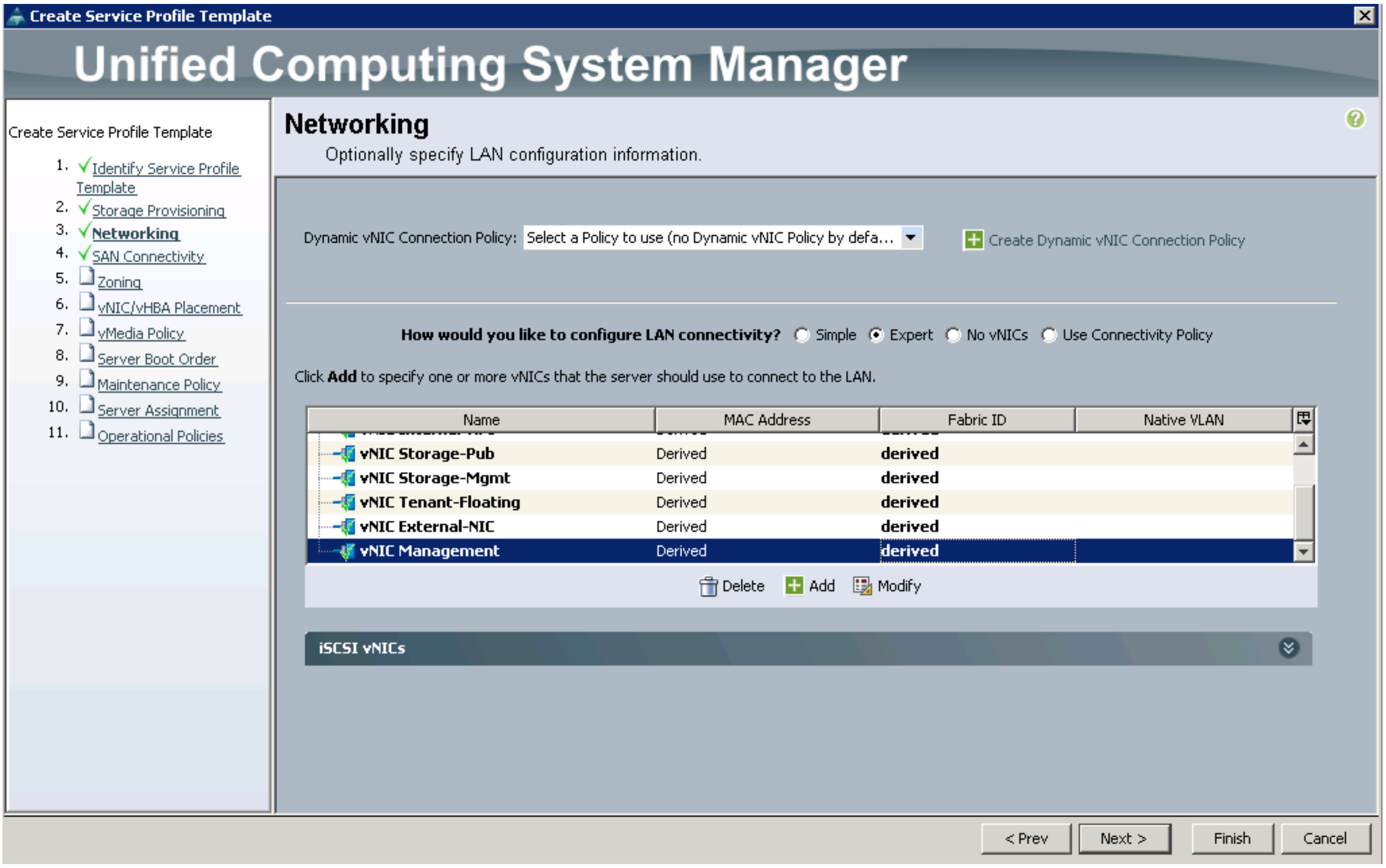
Create vNIC Template

vNIC Template: Management

Adapter Performance Profile

Adapter Policy: Linux

After a successful VNIC creation, click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ **Networking**
4. ✓ SAN Connectivity
5. □ Zoning
6. □ vNIC/vHBA Placement
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

**Networking**

Optionally specify LAN configuration information.

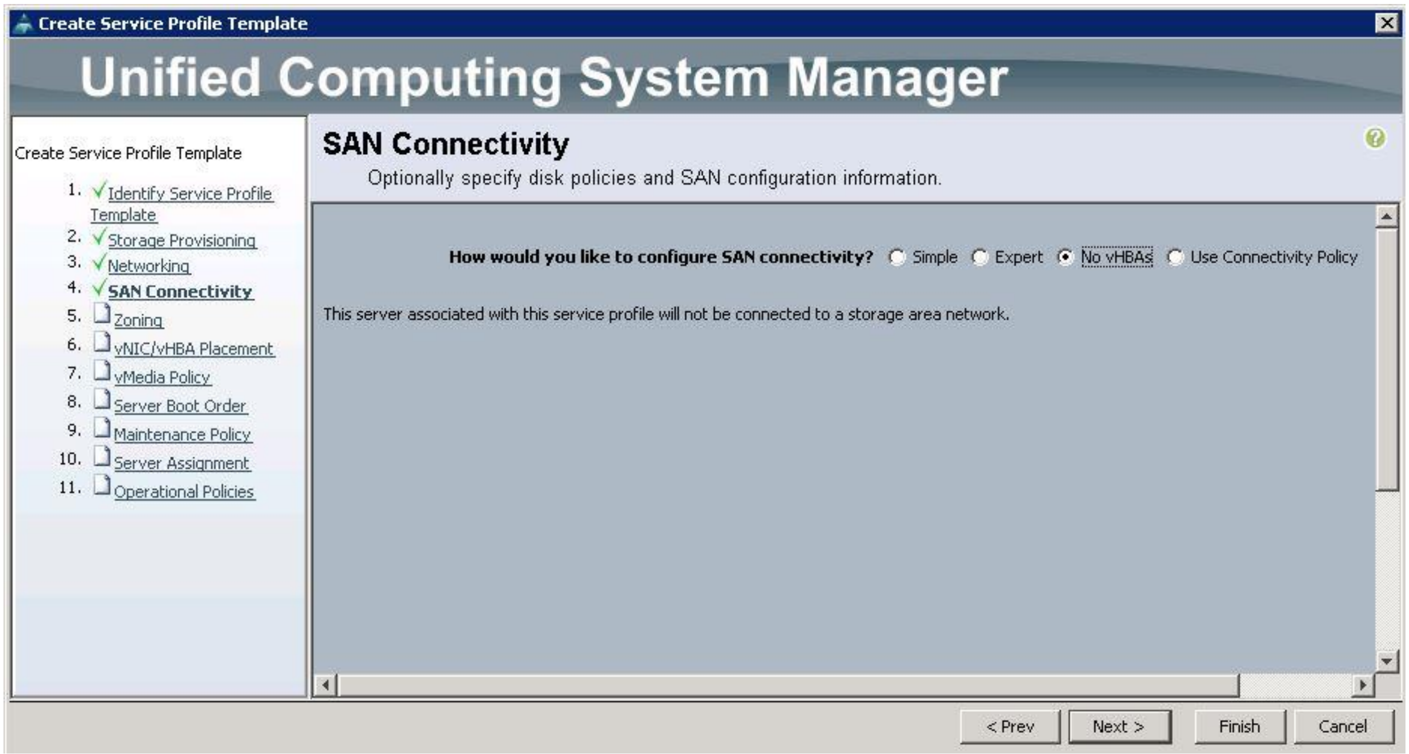
Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa... ▾   Create Dynamic vNIC Connection Policy

**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|
| vNIC Storage-Pub | Derived | derived | |
| vNIC Storage-Mgmt | Derived | derived | |
| vNIC Tenant-Floating | Derived | derived | |
| vNIC External-NIC | Derived | derived | |
| vNIC Management | Derived | derived | |

Under the SAN connectivity, choose No VHBAs and click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ **SAN Connectivity**
5. □ Zoning
6. □ vNIC/vHBA Placement
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

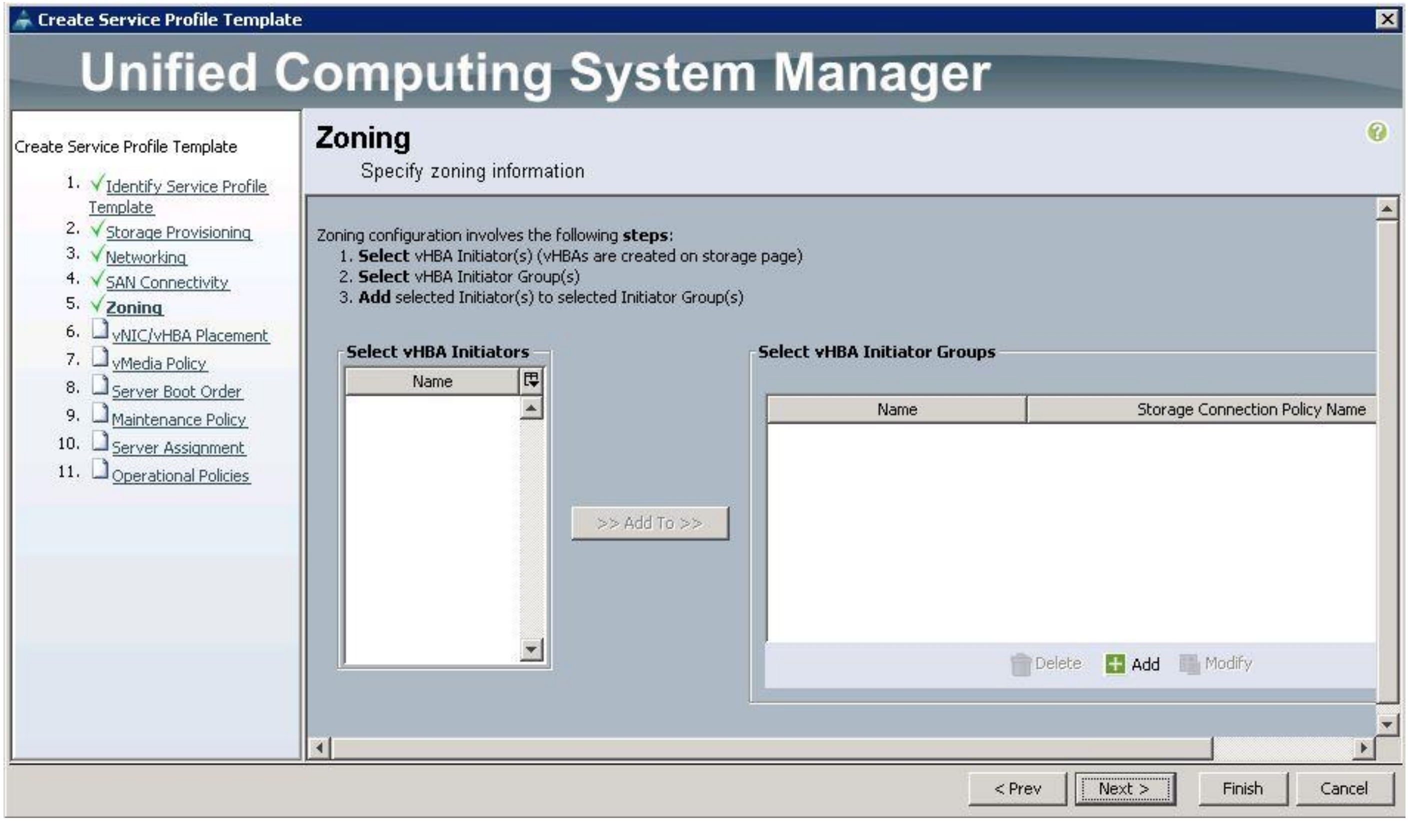
**SAN Connectivity**

Optionally specify disk policies and SAN configuration information.

**How would you like to configure SAN connectivity?** ○ Simple ○ Expert ● No vHBAs ○ Use Connectivity Policy

This server associated with this service profile will not be connected to a storage area network.

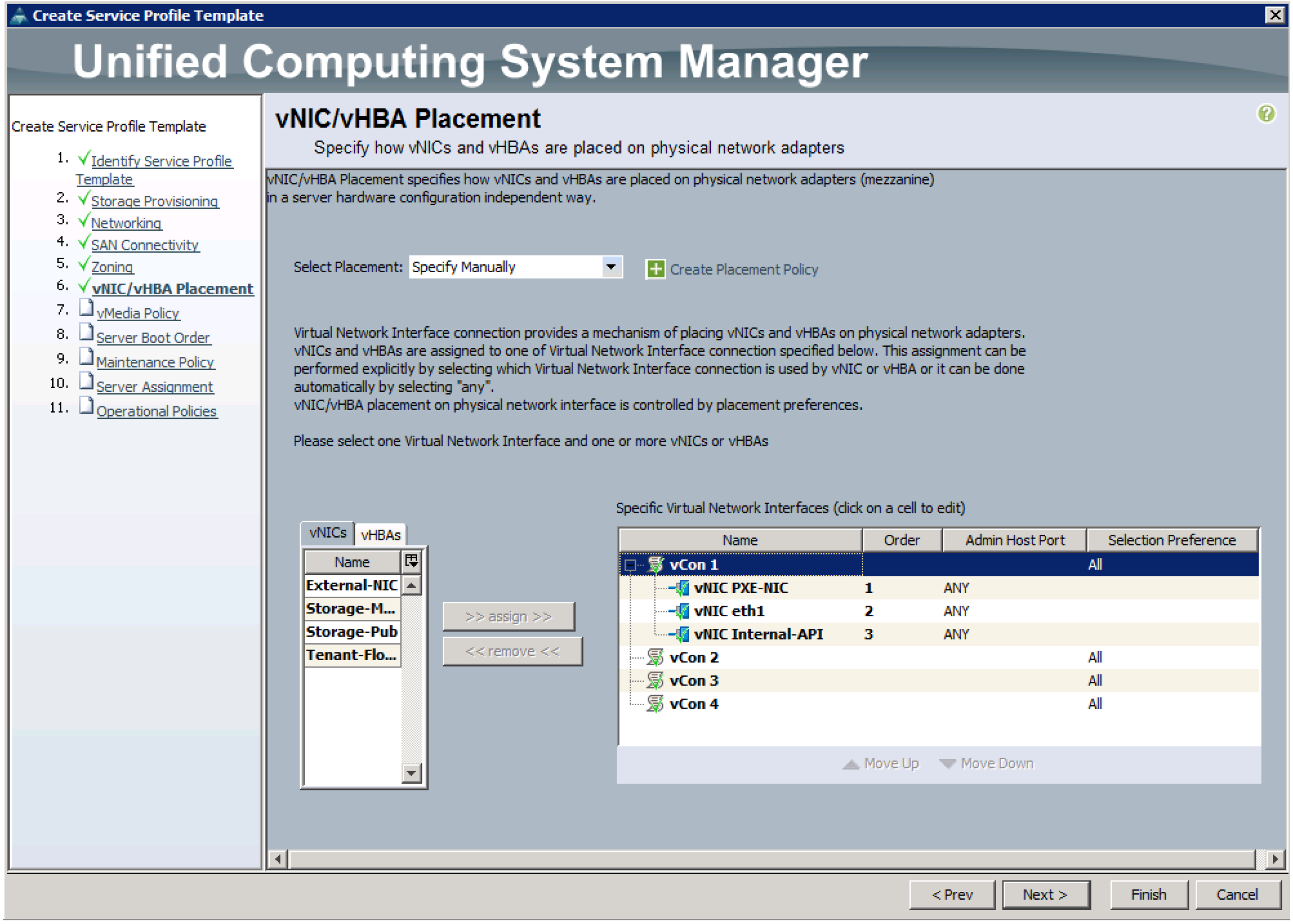
Under Zoning, click Next.



Zoning configuration involves the following **steps**:

1. **Select** vHBA Initiator(s) (vHBAs are created on storage page)
2. **Select** vHBA Initiator Group(s)
3. **Add** selected Initiator(s) to selected Initiator Group(s)

Under VNIC/VHBA Placement, choose the vNICs PCI order as shown below and click Next.



Create Service Profile Template
Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ **vNIC/vHBA Placement**
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**vNIC/vHBA Placement**

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement: Specify Manually

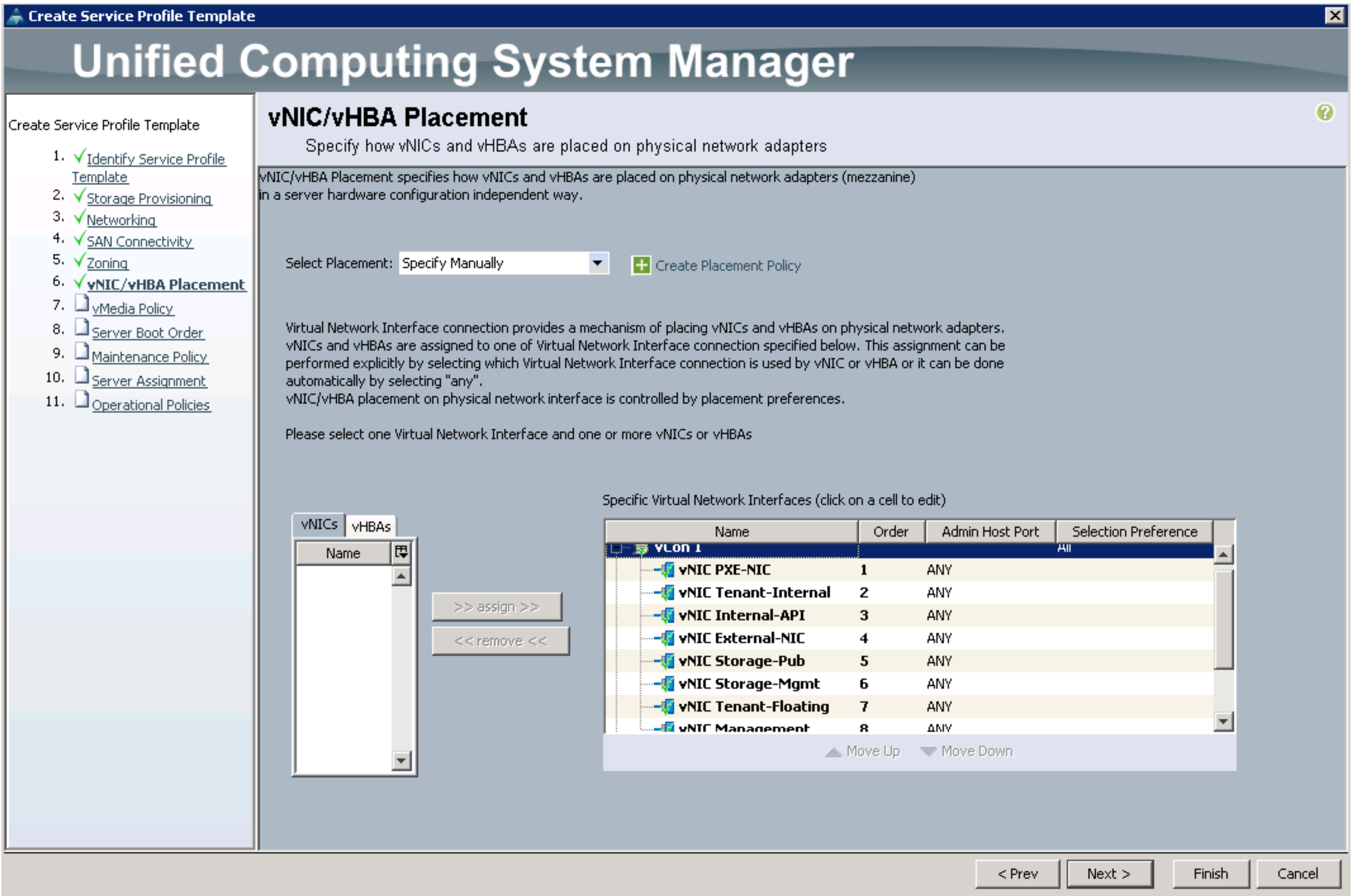
Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

| Name |
|---|
| External-NIC |
| Storage-M... |
| Storage-Pub |
| Tenant-Flo... |

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vCon 1 | | | All |
| vNIC PXE-NIC | 1 | ANY | |
| vNIC eth1 | 2 | ANY | |
| vNIC Internal-API | 3 | ANY | |
| vCon 2 | | | All |
| vCon 3 | | | All |
| vCon 4 | | | All |



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ **vNIC/vHBA Placement**
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

**vNIC/vHBA Placement**

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement: Specify Manually

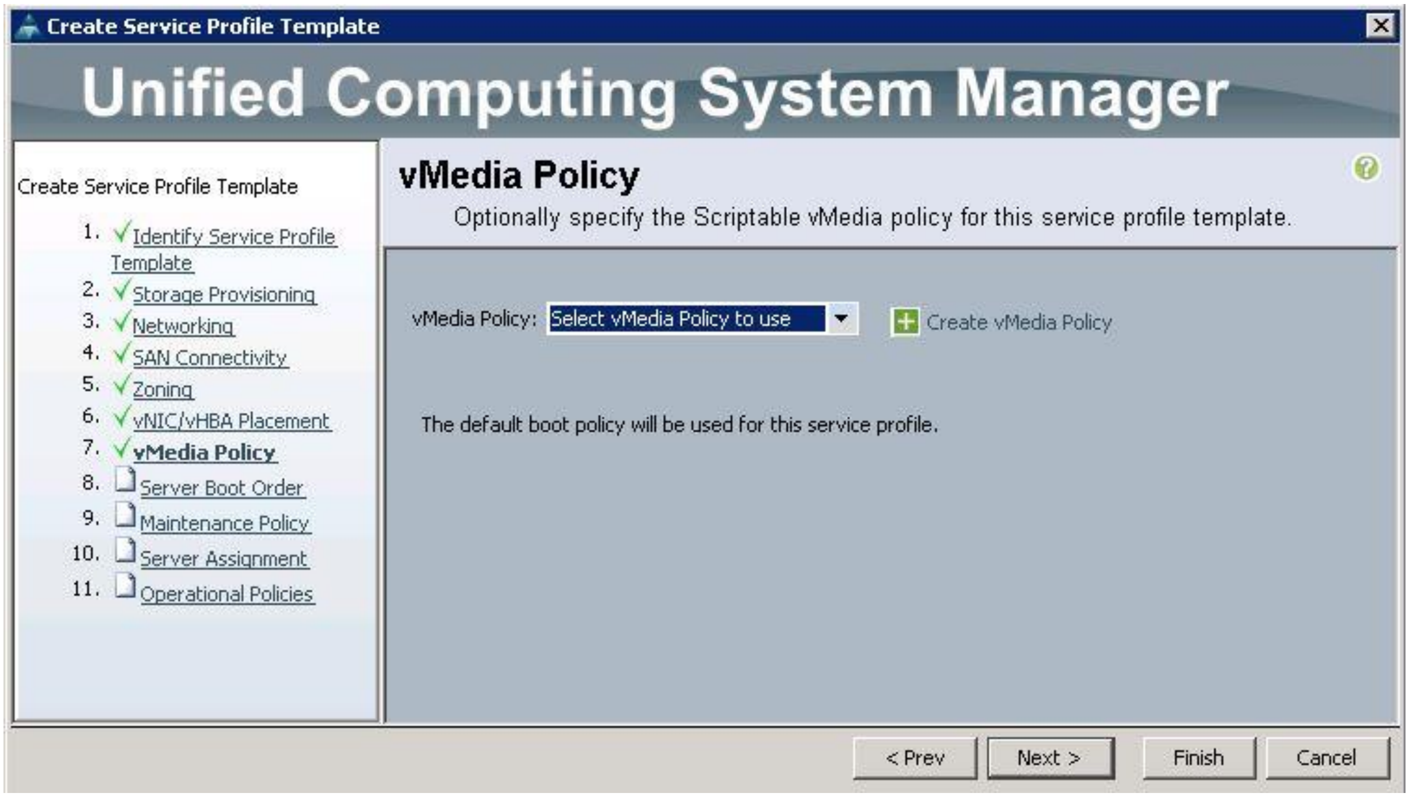
Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vCon 1 | | | All |
| vNIC PXE-NIC | 1 | ANY | |
| vNIC Tenant-Internal | 2 | ANY | |
| vNIC Internal-API | 3 | ANY | |
| vNIC External-NIC | 4 | ANY | |
| vNIC Storage-Pub | 5 | ANY | |
| vNIC Storage-Mgmt | 6 | ANY | |
| vNIC Tenant-Floating | 7 | ANY | |
| vNIC Management | 8 | ANY | |

Under vMedia Policy, click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. ✓ **vMedia Policy**
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

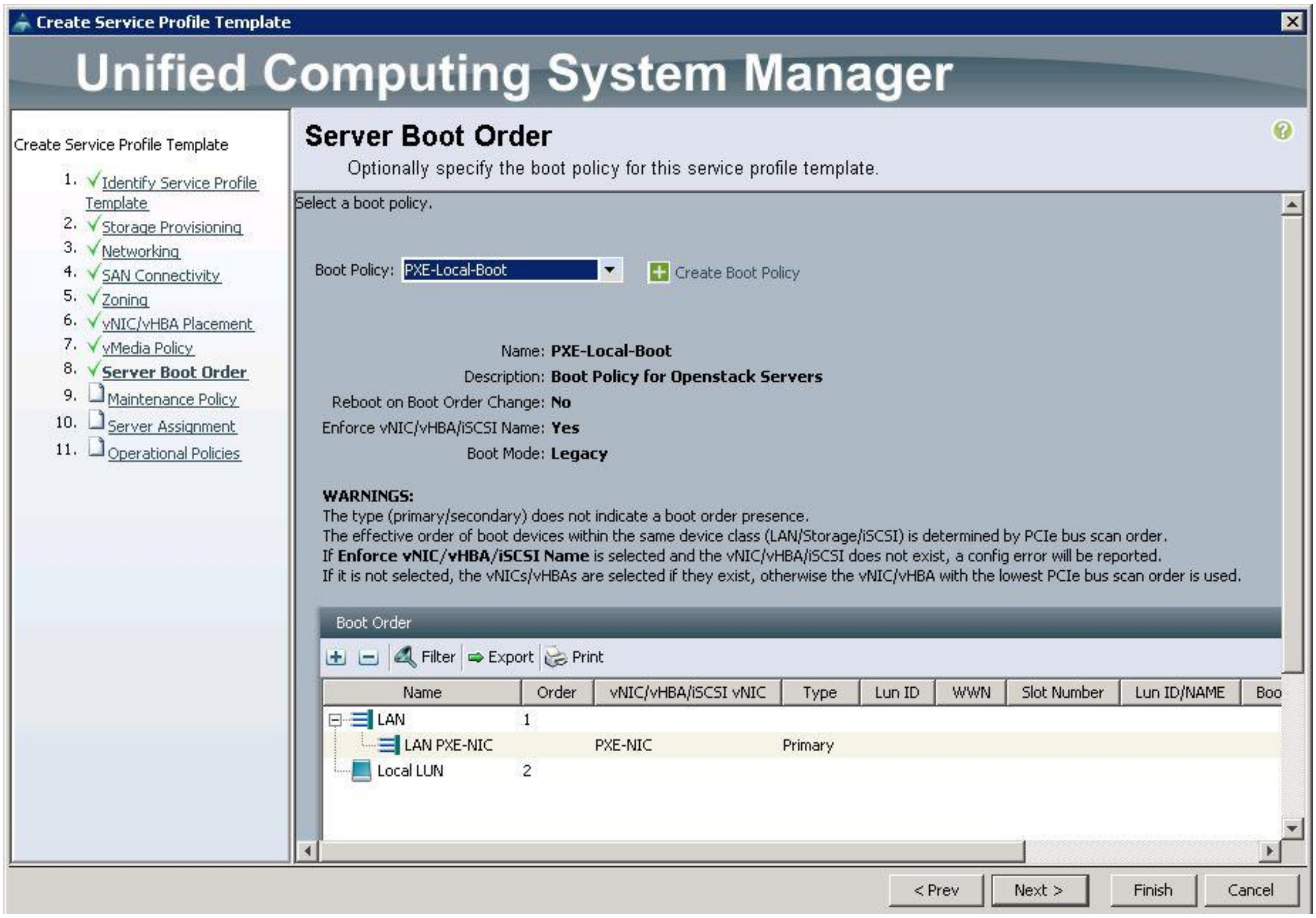
**vMedia Policy**

Optionally specify the Scriptable vMedia policy for this service profile template.

vMedia Policy: Select vMedia Policy to use

The default boot policy will be used for this service profile.

Under Server Boot Order, choose the boot policy as PXE-LocalBoot previously created, from the drop-down list and click Next.



Create Service Profile Template
Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. ✓ vMedia Policy
8. ✓ **Server Boot Order**
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**Server Boot Order**

Optionally specify the boot policy for this service profile template.

Select a boot policy.

Boot Policy: PXE-Local-Boot

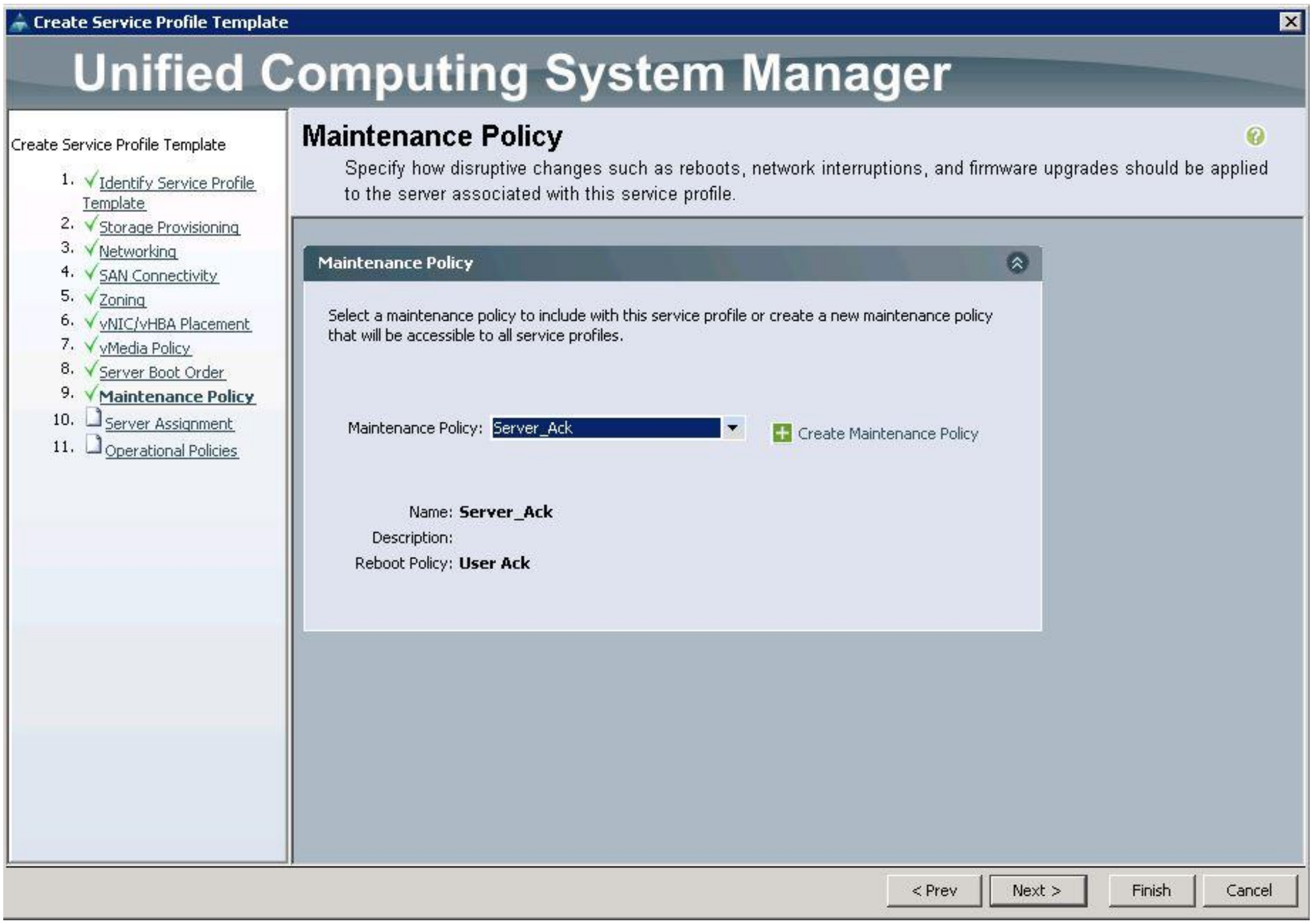
Name: **PXE-Local-Boot**
Description: **Boot Policy for Openstack Servers**
Reboot on Boot Order Change: **No**
Enforce vNIC/vHBA/iSCSI Name: **Yes**
Boot Mode: **Legacy**

**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Boot Order

| Name | Order | vNIC/vHBA/iSCSI vNIC | Type | Lun ID | WWN | Slot Number | Lun ID/NAME | Boo |
|---|---|---|---|---|---|---|---|---|
| LAN | 1 | | | | | | | |
| LAN PXE-NIC | | PXE-NIC | Primary | | | | | |
| Local LUN | 2 | | | | | | | |

Under Maintenance Policy, choose Server_Ack previously created, from the drop-down list and click Next.



Maintenance Policy

Specify how disruptive changes such as reboots, network interruptions, and firmware upgrades should be applied to the server associated with this service profile.

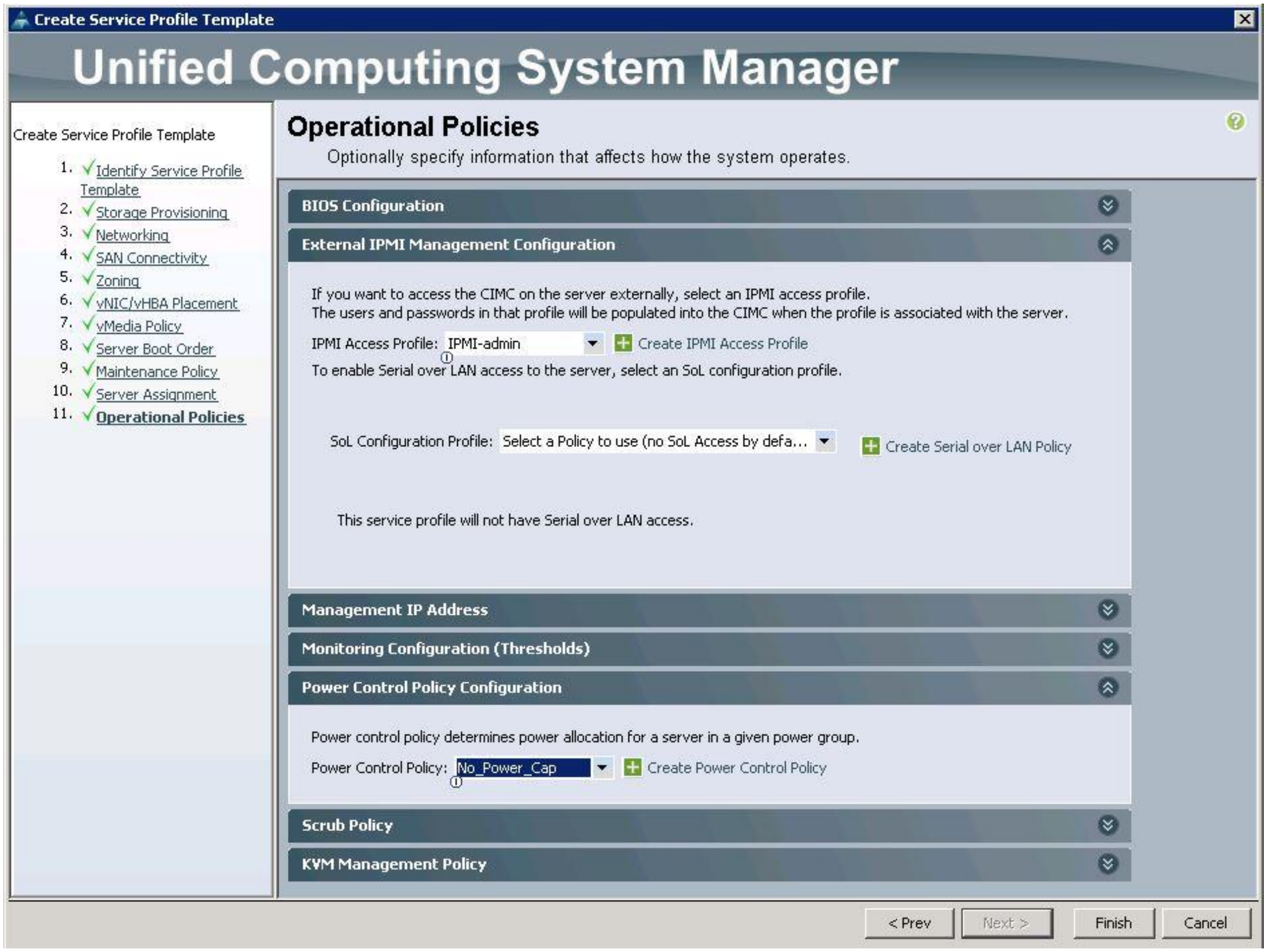
Maintenance Policy

Select a maintenance policy to include with this service profile or create a new maintenance policy that will be accessible to all service profiles.

Maintenance Policy: Server_Ack

Name: **Server_Ack**
Description:
Reboot Policy: **User Ack**

Under Operational Policies, choose the IPMI Access Profile as IPMI_admin previously created, from the drop-down list and choose the Power Control Policy as No_Power_Cap and click Finish.



Operational Policies

Optionally specify information that affects how the system operates.

If you want to access the CIMC on the server externally, select an IPMI access profile.
The users and passwords in that profile will be populated into the CIMC when the profile is associated with the server.

IPMI Access Profile: IPMI-admin ▼ Create IPMI Access Profile

To enable Serial over LAN access to the server, select an SoL configuration profile.

SoL Configuration Profile: Select a Policy to use (no SoL Access by defa... ▼ Create Serial over LAN Policy

This service profile will not have Serial over LAN access.

Power control policy determines power allocation for a server in a given power group.

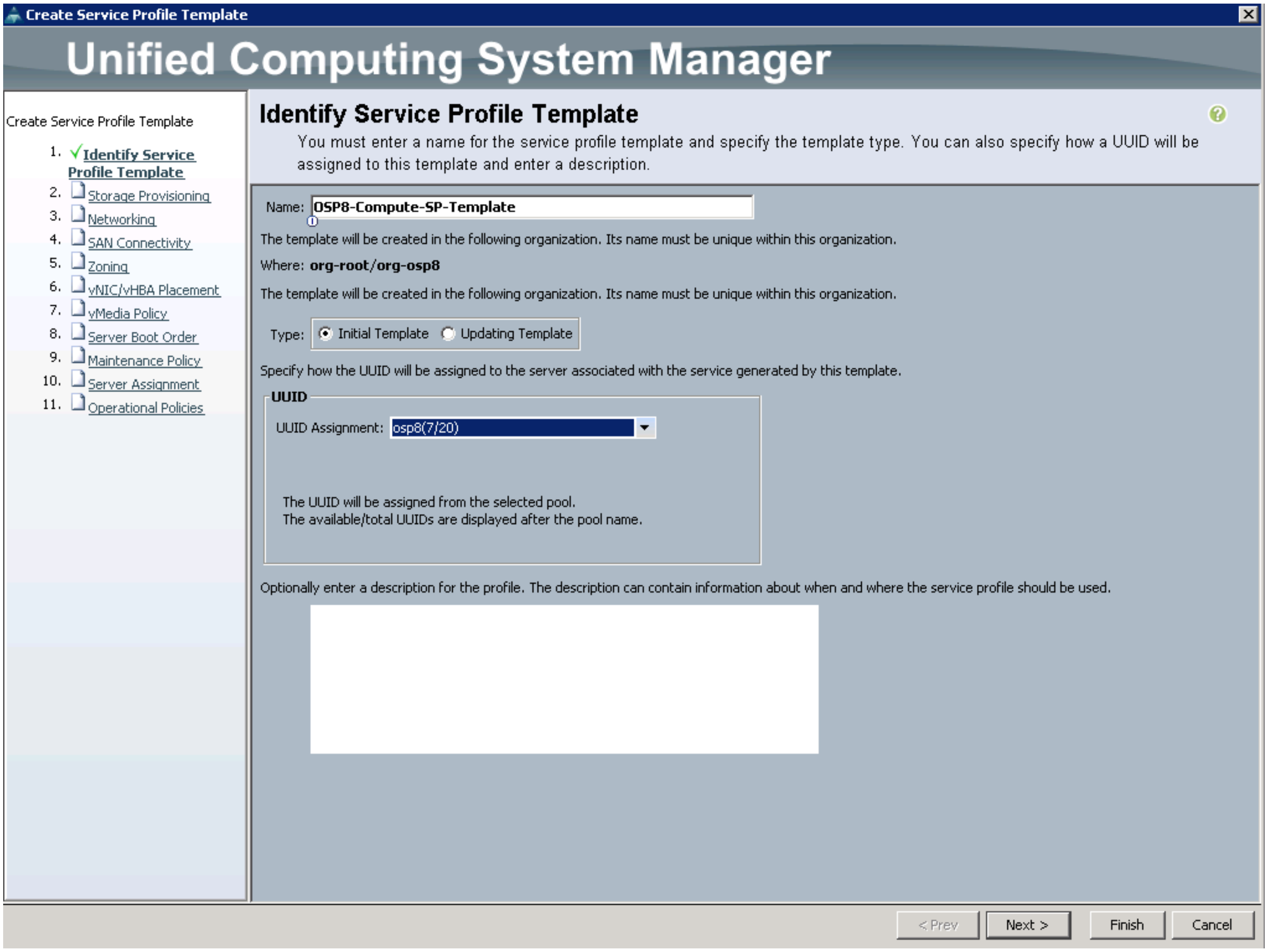
Power Control Policy: No_Power_Cap ▼ Create Power Control Policy

## Create Service Profile Templates for Compute Nodes

To create the Service Profile templates for the Compute nodes, complete the following steps:

Specify the Service profile template name for the Controller node as OSP8-Compute-SP-Template.

Choose the UUID pools previously created from the drop-down list and click Next.



Identify Service Profile Template

You must enter a name for the service profile template and specify the template type. You can also specify how a UUID will be assigned to this template and enter a description.

Name: OSP8-Compute-SP-Template

The template will be created in the following organization. Its name must be unique within this organization.

Where: **org-root/org-osp8**

The template will be created in the following organization. Its name must be unique within this organization.

Type: ● Initial Template ○ Updating Template

Specify how the UUID will be assigned to the server associated with the service generated by this template.

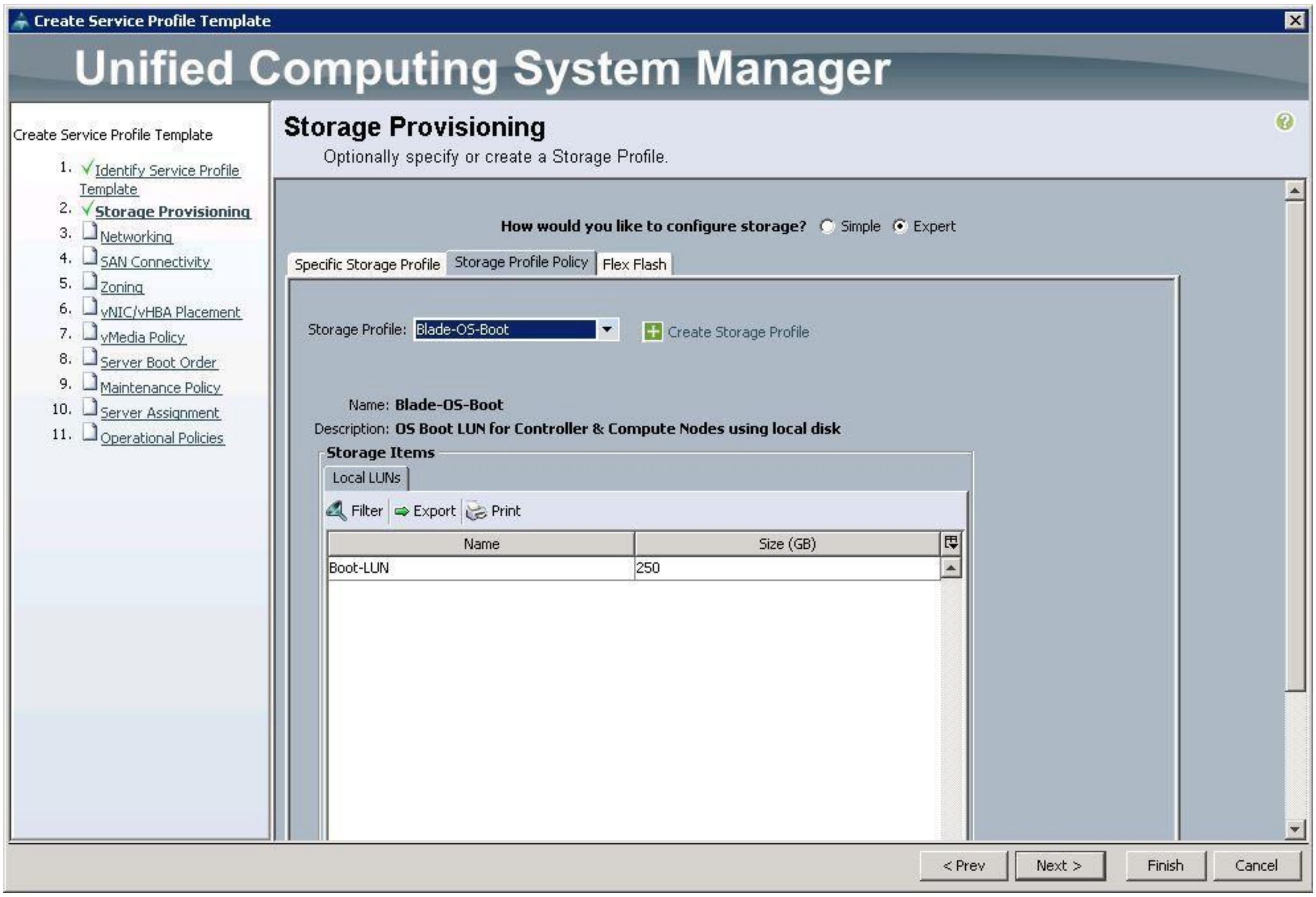
UUID

UUID Assignment: osp8(7/20)

The UUID will be assigned from the selected pool.
The available/total UUIDs are displayed after the pool name.

Optionally enter a description for the profile. The description can contain information about when and where the service profile should be used.

For Storage Provisioning, choose Expert and click Storage Profile Policy and choose the Storage profile Blade-OS-boot previously created, from the drop-down list and click Next.



Storage Provisioning

Optionally specify or create a Storage Profile.

How would you like to configure storage? ○ Simple ● Expert

Storage Profile: Blade-OS-Boot

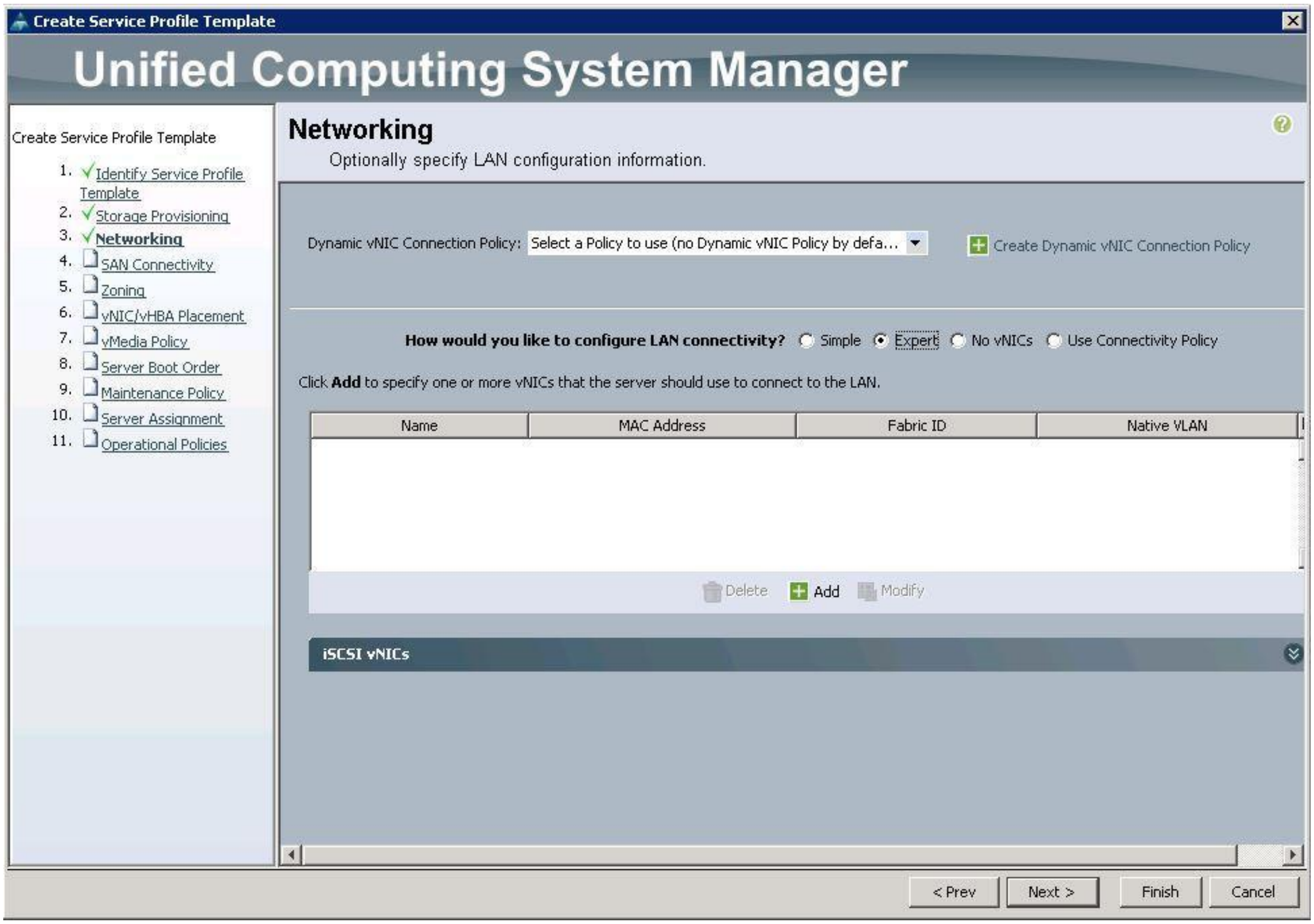
Name: Blade-OS-Boot

Description: OS Boot LUN for Controller & Compute Nodes using local disk

Storage Items

| Name | Size (GB) |
|---|---|
| Boot-LUN | 250 |

For Networking, choose Expert and click “+".



Create Service Profile Template

Networking

Optionally specify LAN configuration information.

Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

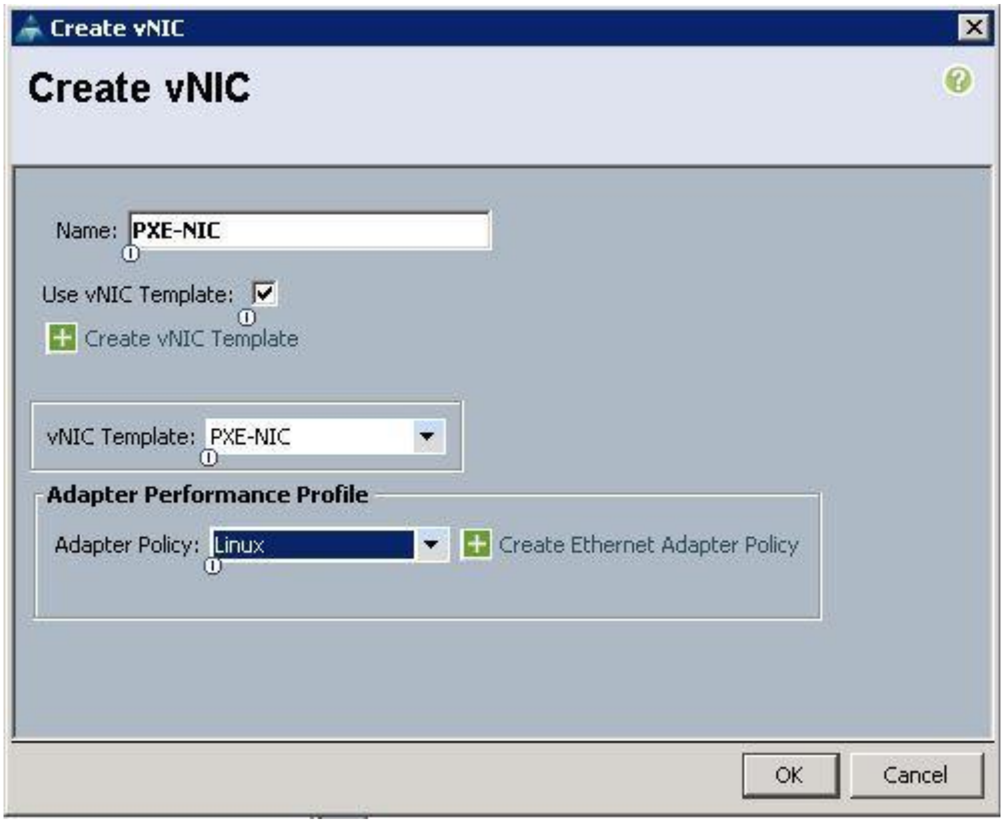
**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|

Create the VNIC interface for PXE or Provisioning network as PXE-NIC and click the check box for Use VNIC template.

Under the vNIC template, choose the PXE-NIC template previously created, from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: PXE-NIC

Use vNIC Template: ☑

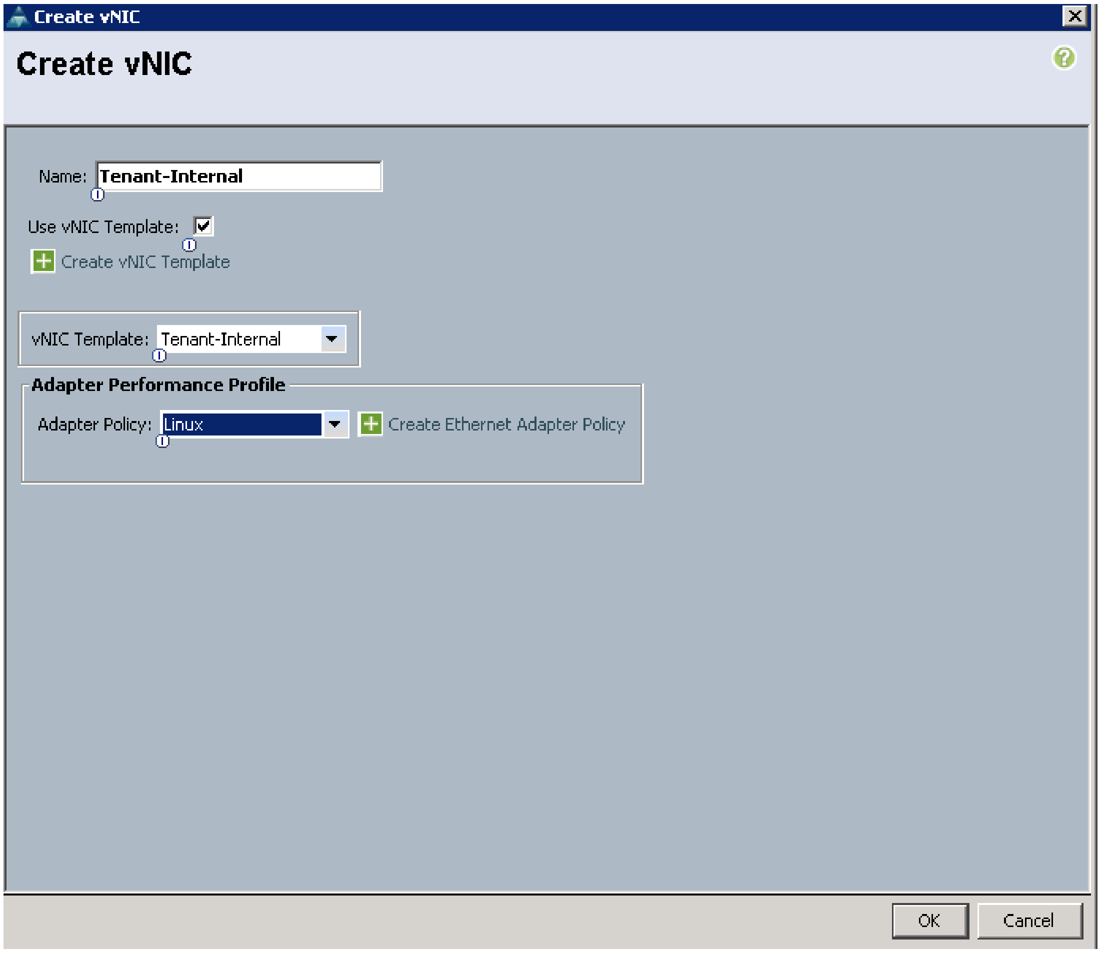
vNIC Template: PXE-NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for Tenant Internal Network as Tenant-Internal and then under vNIC template, choose the "Tenant-Internal" template we created before from the drop-down list and choose Adapter Policy as "Linux".

Due to the Cisco UCS Manager Plugin limitations, we have created eth1 as VNIC for Tenant Internal Network.



Create vNIC

Name: Tenant-Internal

Use vNIC Template: ☑

Create vNIC Template

vNIC Template: Tenant-Internal

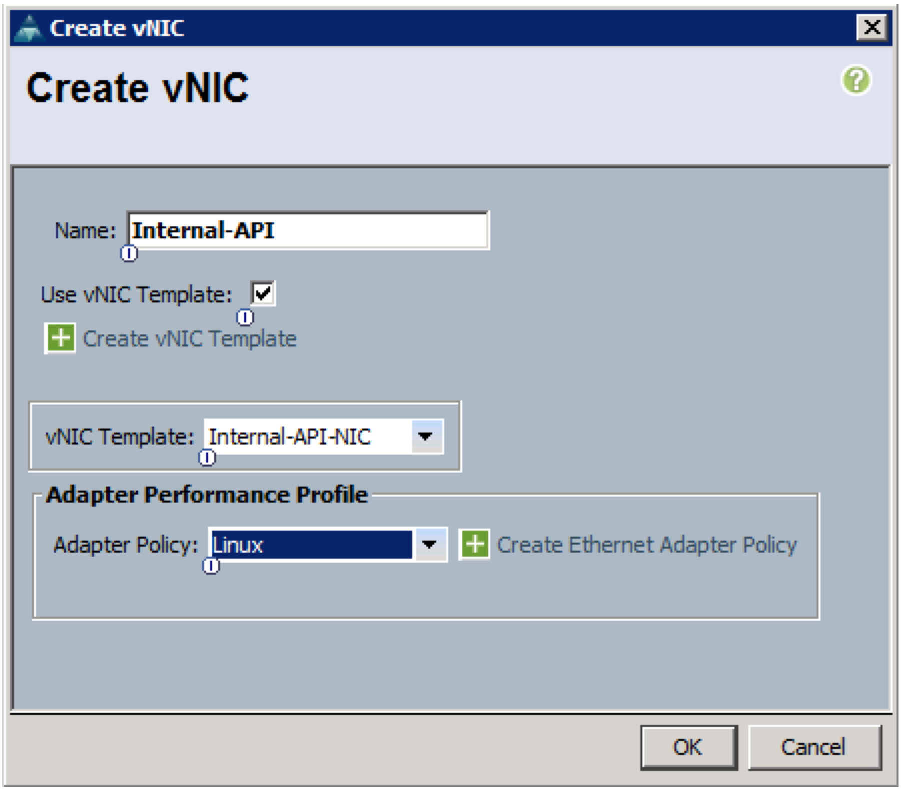
Adapter Performance Profile

Adapter Policy: Linux

Create Ethernet Adapter Policy

Create the VNIC interface for Internal API network as Internal-API and click the check box for VNIC template.

Under the vNIC template, choose the Internal-API template previously created, from the drop-down list and choose Linux for the Adapter Policy.



Name: Internal-API

Use vNIC Template: ☑

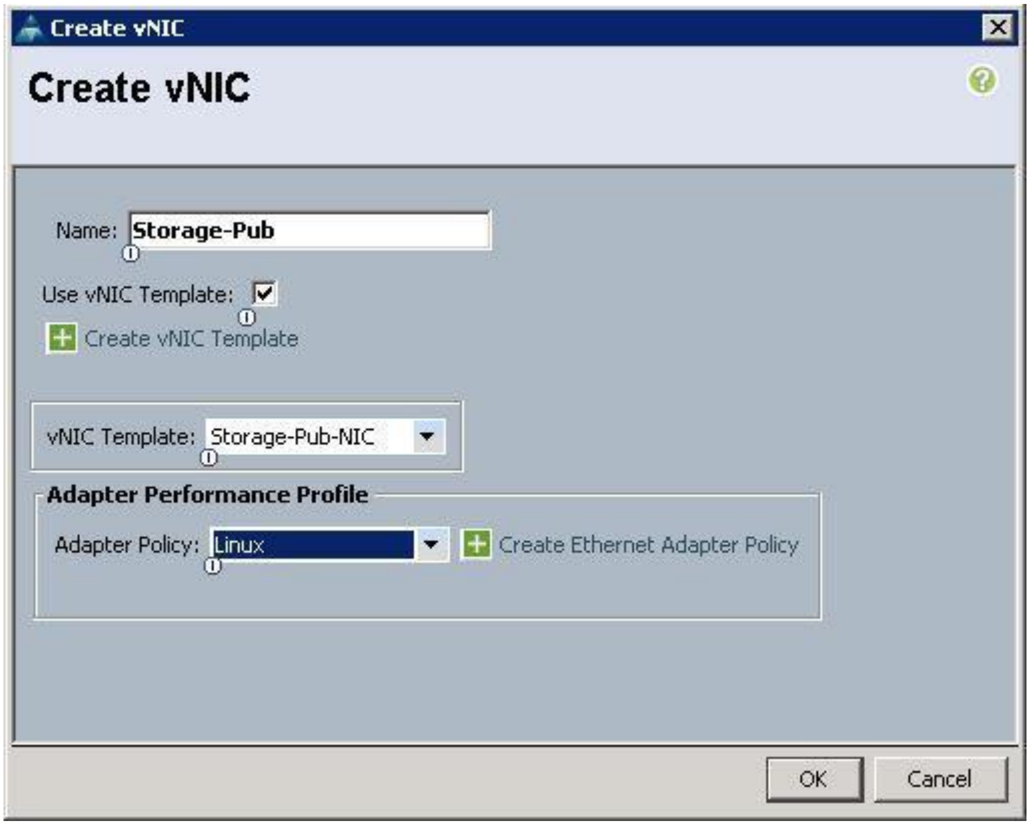
vNIC Template: Internal-API-NIC

**Adapter Performance Profile**

Adapter Policy: Linux

Create the VNIC interface for Storage Public Network as Storage-Pub and click the check box for Use VNIC template.

Under the vNIC template, choose the Storage-Pub-NIC template previously created, from the drop-down list and choose Linux for the Adapter Policy.



Name: Storage-Pub

Use vNIC Template: ☑

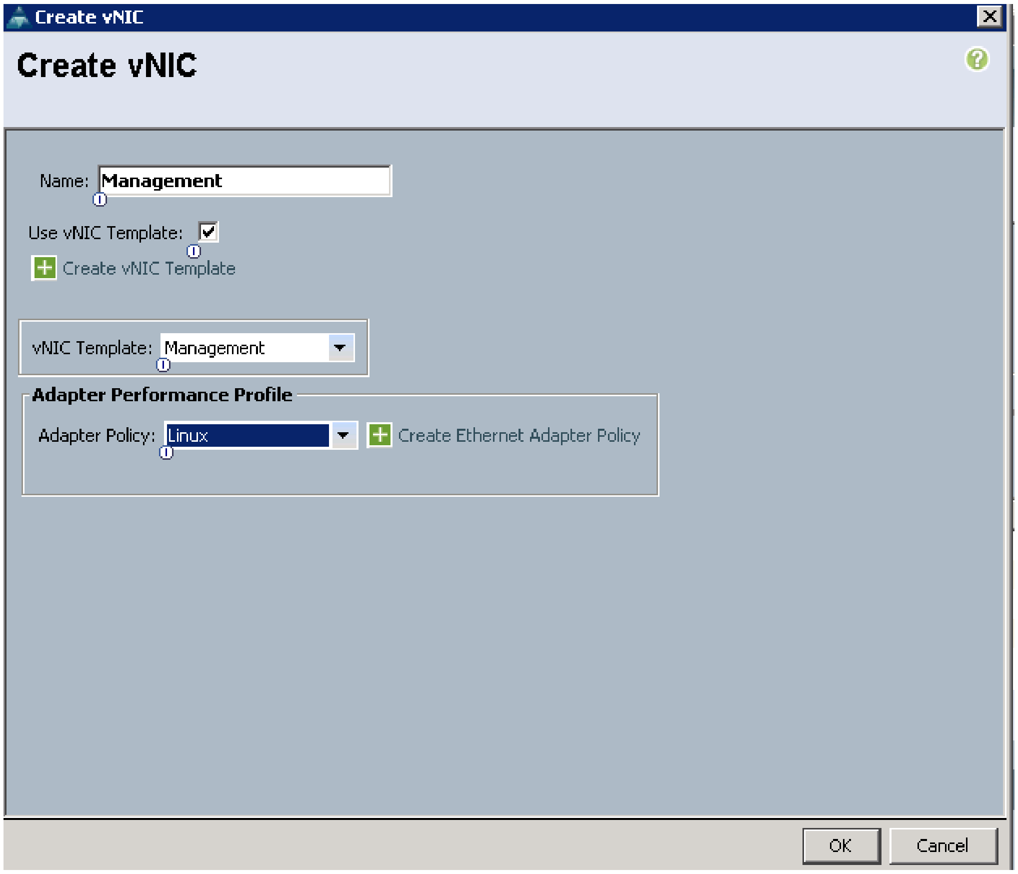
vNIC Template: Storage-Pub-NIC

**Adapter Performance Profile**

Adapter Policy: Linux

Create the VNIC interface for Management Network as Management and click the check box for Use VNIC template.

Under the vNIC template, choose the Management template previously created, from the drop-down list and choose Linux for the Adapter Policy.



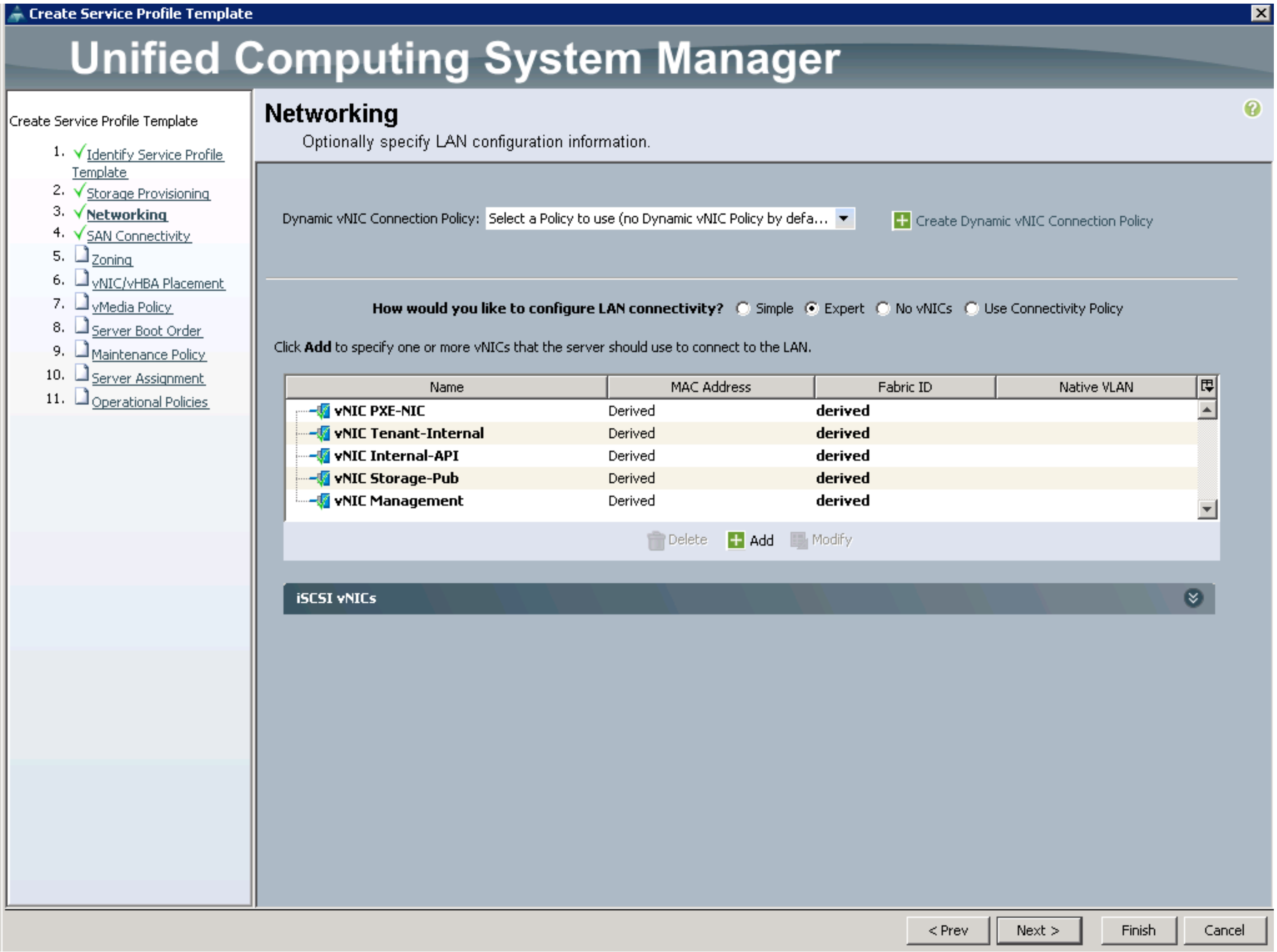
Name: Management

Use vNIC Template: ☑

vNIC Template: Management

Adapter Policy: Linux

After a successful VNIC creation, click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ **Networking**
4. ✓ SAN Connectivity
5. Zoning
6. vNIC/vHBA Placement
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**Networking**

Optionally specify LAN configuration information.

Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

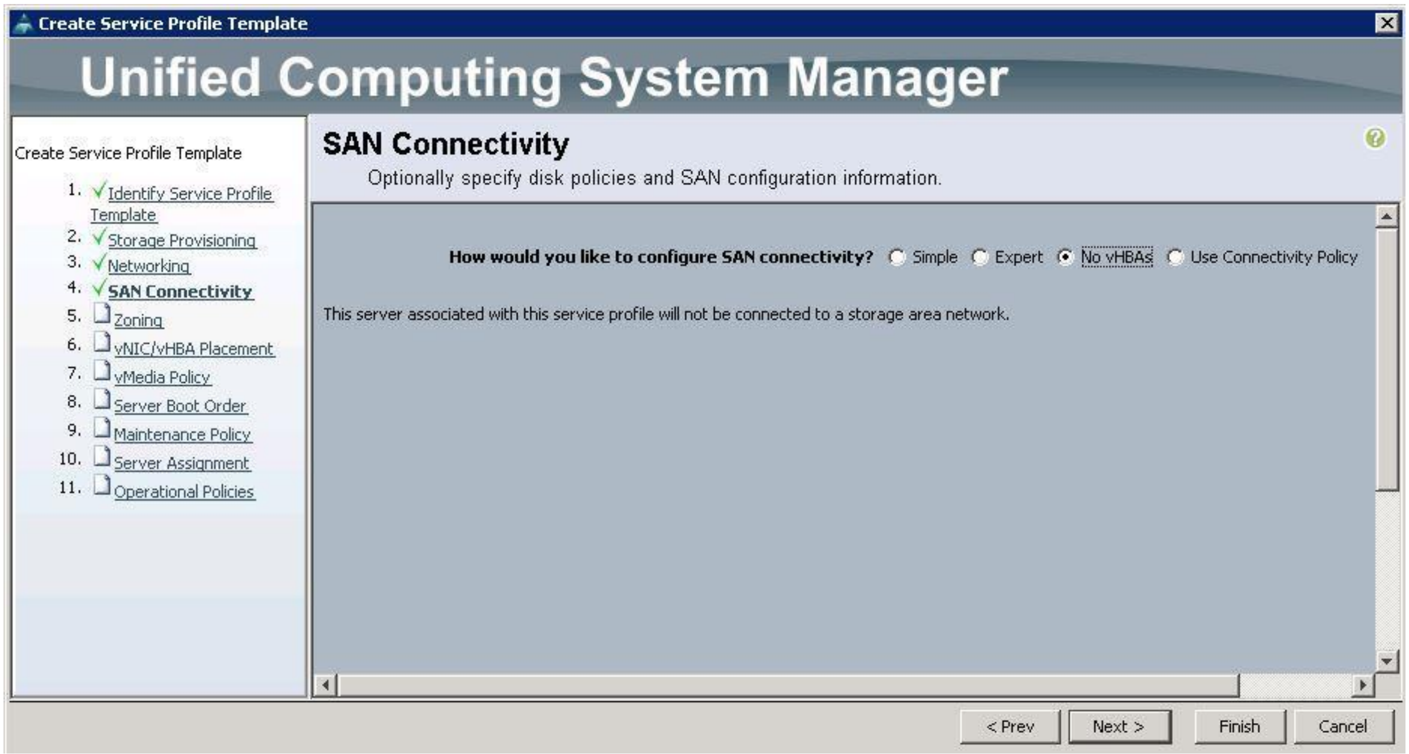
Create Dynamic vNIC Connection Policy

**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|
| **vNIC PXE-NIC** | Derived | **derived** | |
| **vNIC Tenant-Internal** | Derived | **derived** | |
| **vNIC Internal-API** | Derived | **derived** | |
| **vNIC Storage-Pub** | Derived | **derived** | |
| **vNIC Management** | Derived | **derived** | |

Under SAN connectivity, choose No VHBAs and click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ **SAN Connectivity**
5. □ Zoning
6. □ vNIC/vHBA Placement
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

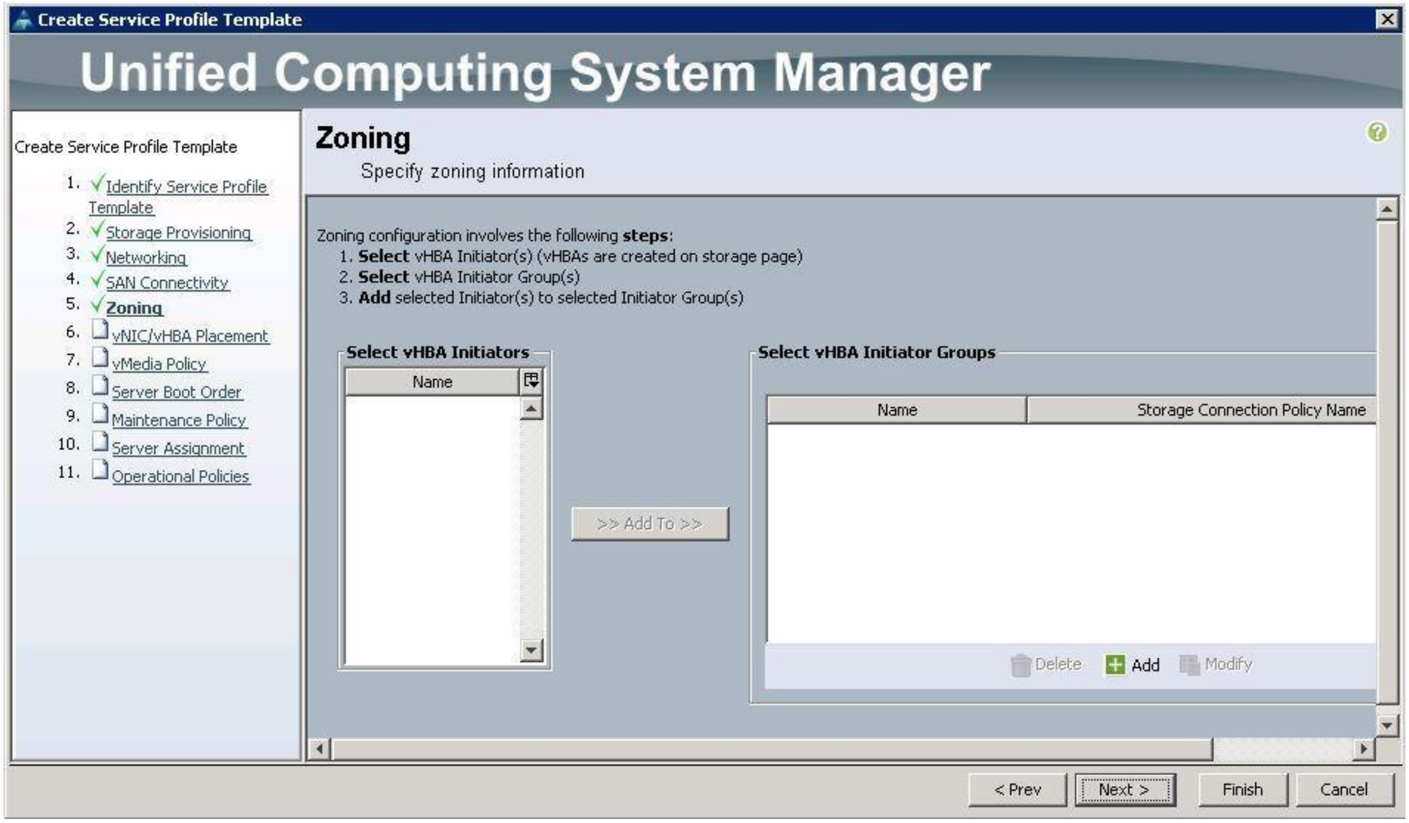
**SAN Connectivity**

Optionally specify disk policies and SAN configuration information.

**How would you like to configure SAN connectivity?** ○ Simple ○ Expert ● No vHBAs ○ Use Connectivity Policy

This server associated with this service profile will not be connected to a storage area network.

Under Zoning, click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ **Zoning**
6. □ vNIC/vHBA Placement
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

**Zoning**

Specify zoning information

Zoning configuration involves the following **steps**:

1. **Select** vHBA Initiator(s) (vHBAs are created on storage page)
2. **Select** vHBA Initiator Group(s)
3. **Add** selected Initiator(s) to selected Initiator Group(s)

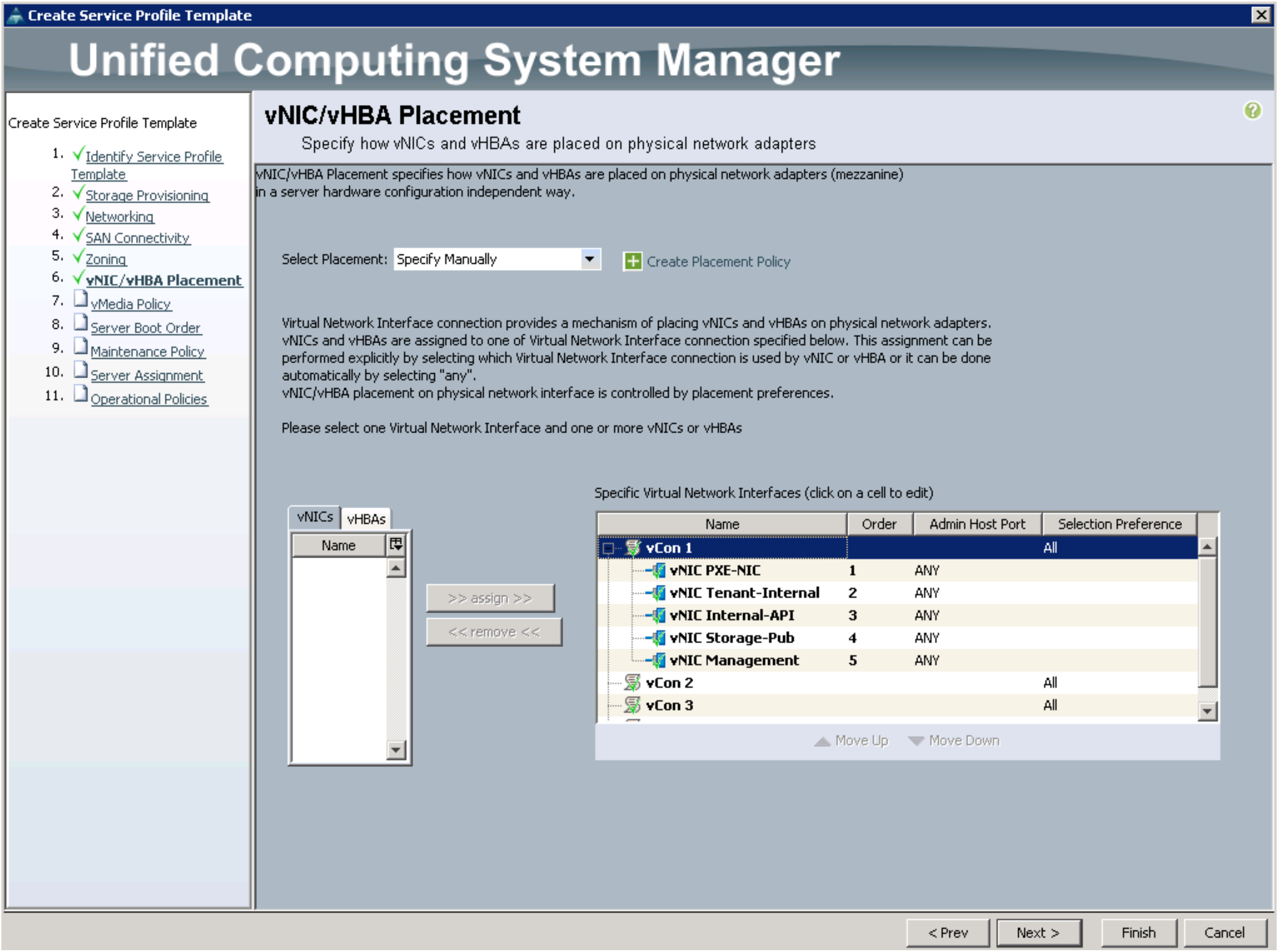
**Select vHBA Initiators**

| Name |
|---|

**Select vHBA Initiator Groups**

| Name | Storage Connection Policy Name |
|---|---|

Under VNIC/VHBA Placement, choose the vNICs PCI order as shown below and click Next.



Create Service Profile Template
Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ **vNIC/vHBA Placement**
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

**vNIC/vHBA Placement**

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement: Specify Manually

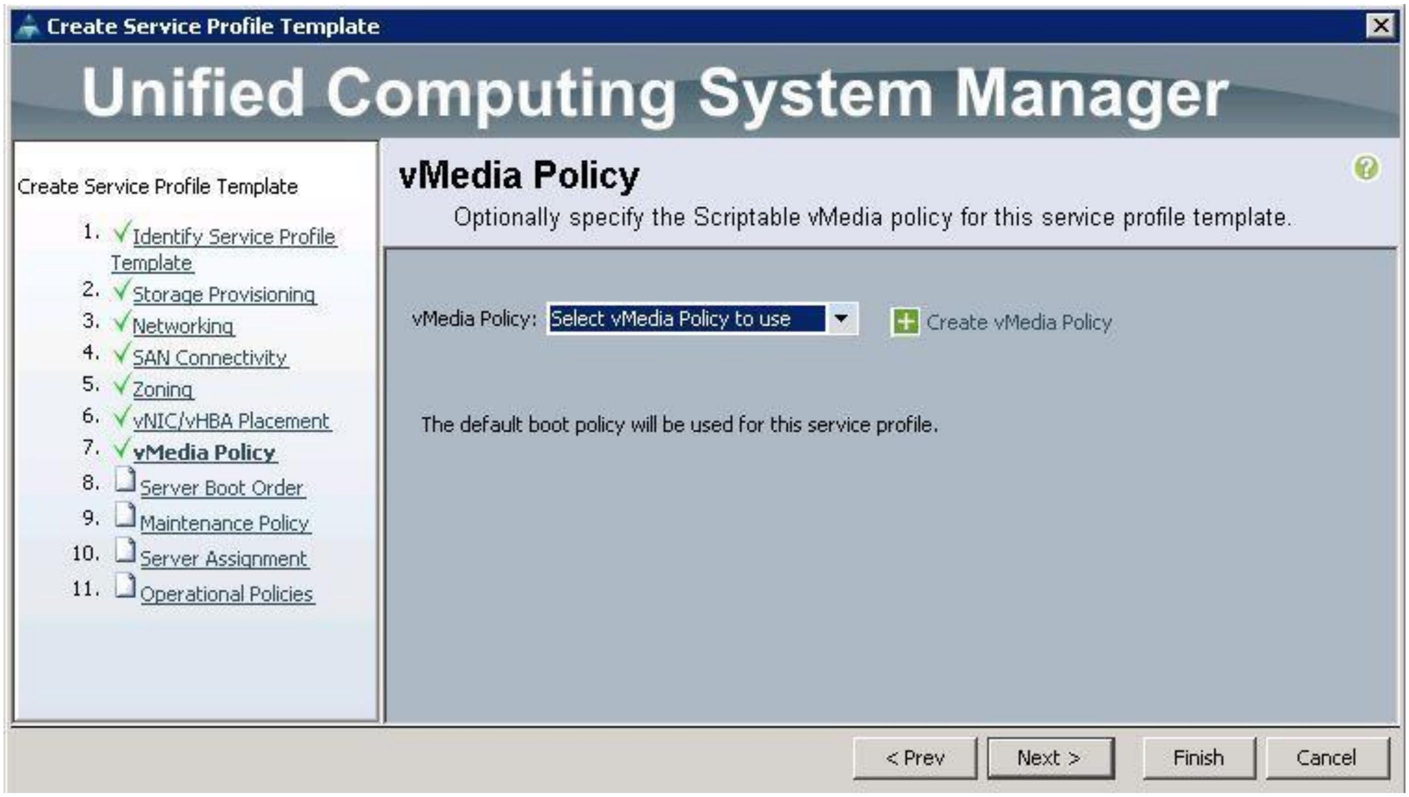
Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vCon 1 | | | All |
| vNIC PXE-NIC | 1 | ANY | |
| vNIC Tenant-Internal | 2 | ANY | |
| vNIC Internal-API | 3 | ANY | |
| vNIC Storage-Pub | 4 | ANY | |
| vNIC Management | 5 | ANY | |
| vCon 2 | | | All |
| vCon 3 | | | All |

Under vMedia Policy, click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. ✓ **vMedia Policy**
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

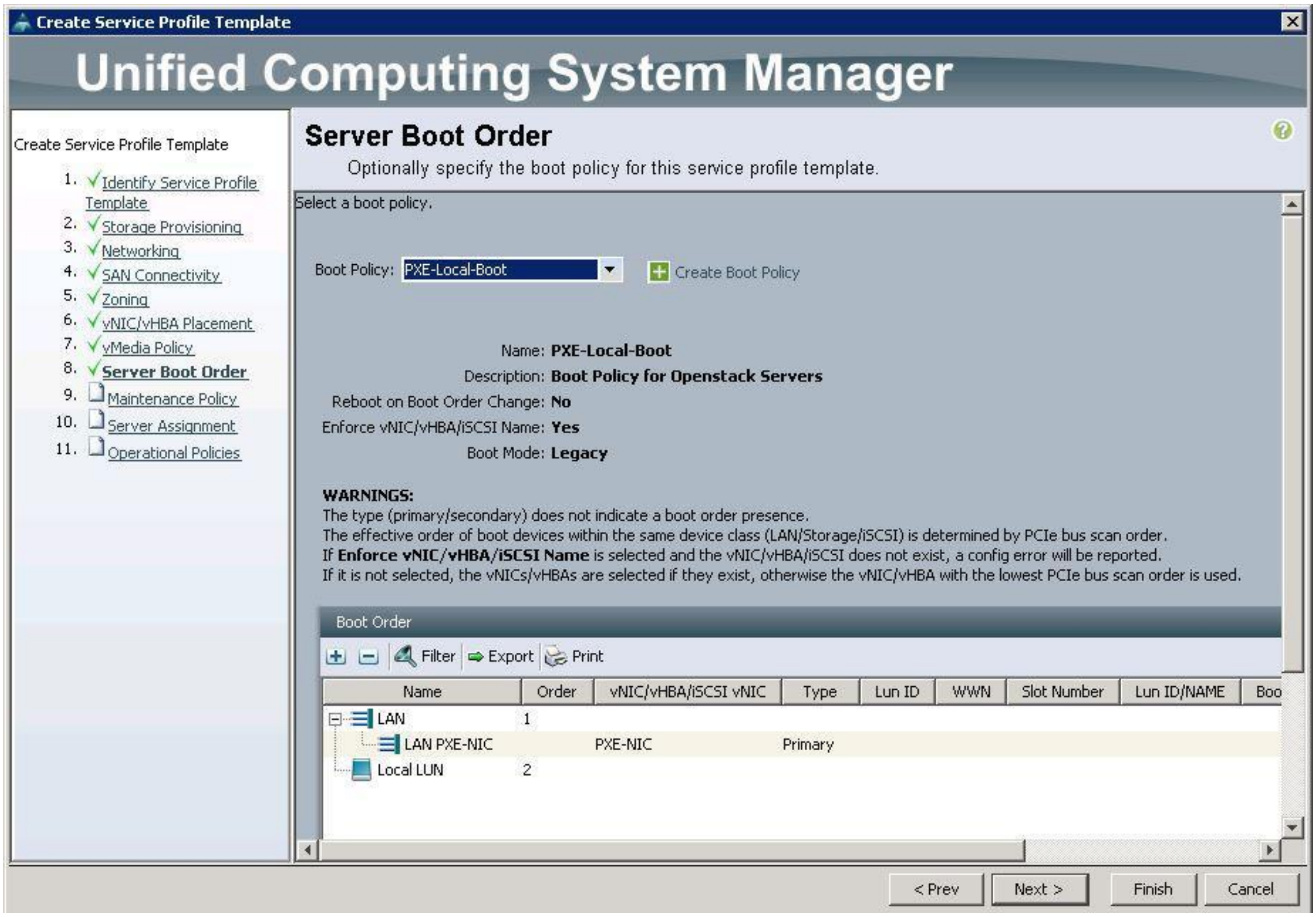
**vMedia Policy**

Optionally specify the Scriptable vMedia policy for this service profile template.

vMedia Policy: Select vMedia Policy to use

The default boot policy will be used for this service profile.

Under Server Boot Order, choose boot policy as PXE-LocalBoot we created from the drop-down list and click Next.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. ✓ vMedia Policy
8. ✓ **Server Boot Order**
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

**Server Boot Order**

Optionally specify the boot policy for this service profile template.

Select a boot policy.

Boot Policy: PXE-Local-Boot

Create Boot Policy

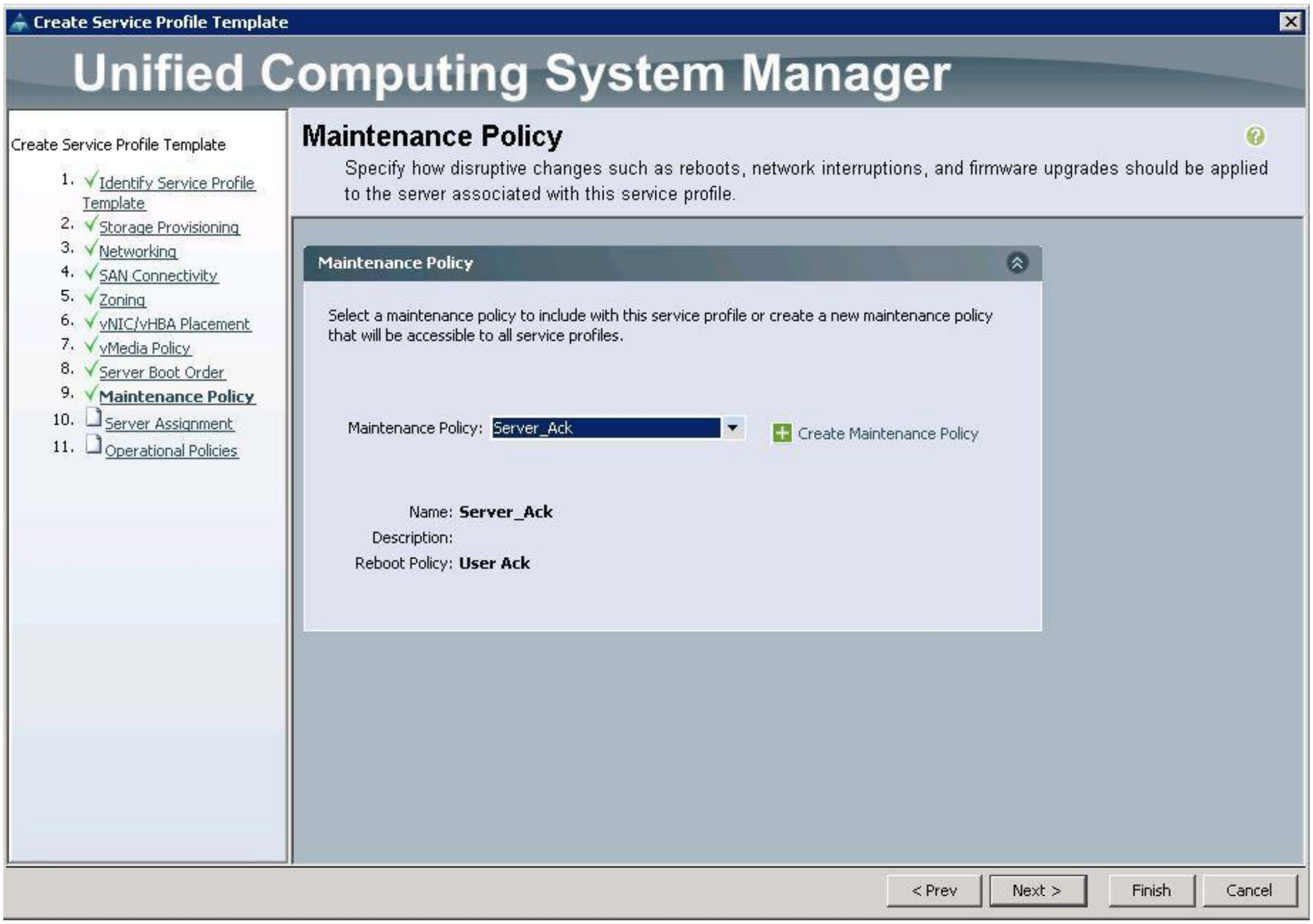
Name: **PXE-Local-Boot**
Description: **Boot Policy for Openstack Servers**
Reboot on Boot Order Change: **No**
Enforce vNIC/vHBA/iSCSI Name: **Yes**
Boot Mode: **Legacy**

**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Boot Order

| Name | Order | vNIC/vHBA/iSCSI vNIC | Type | Lun ID | WWN | Slot Number | Lun ID/NAME | Boo |
|---|---|---|---|---|---|---|---|---|
| LAN | 1 | | | | | | | |
| LAN PXE-NIC | | PXE-NIC | Primary | | | | | |
| Local LUN | 2 | | | | | | | |

Under Maintenance Policy, choose Server_Ack previously created, from the drop-down list and click Next.



Maintenance Policy

Specify how disruptive changes such as reboots, network interruptions, and firmware upgrades should be applied to the server associated with this service profile.

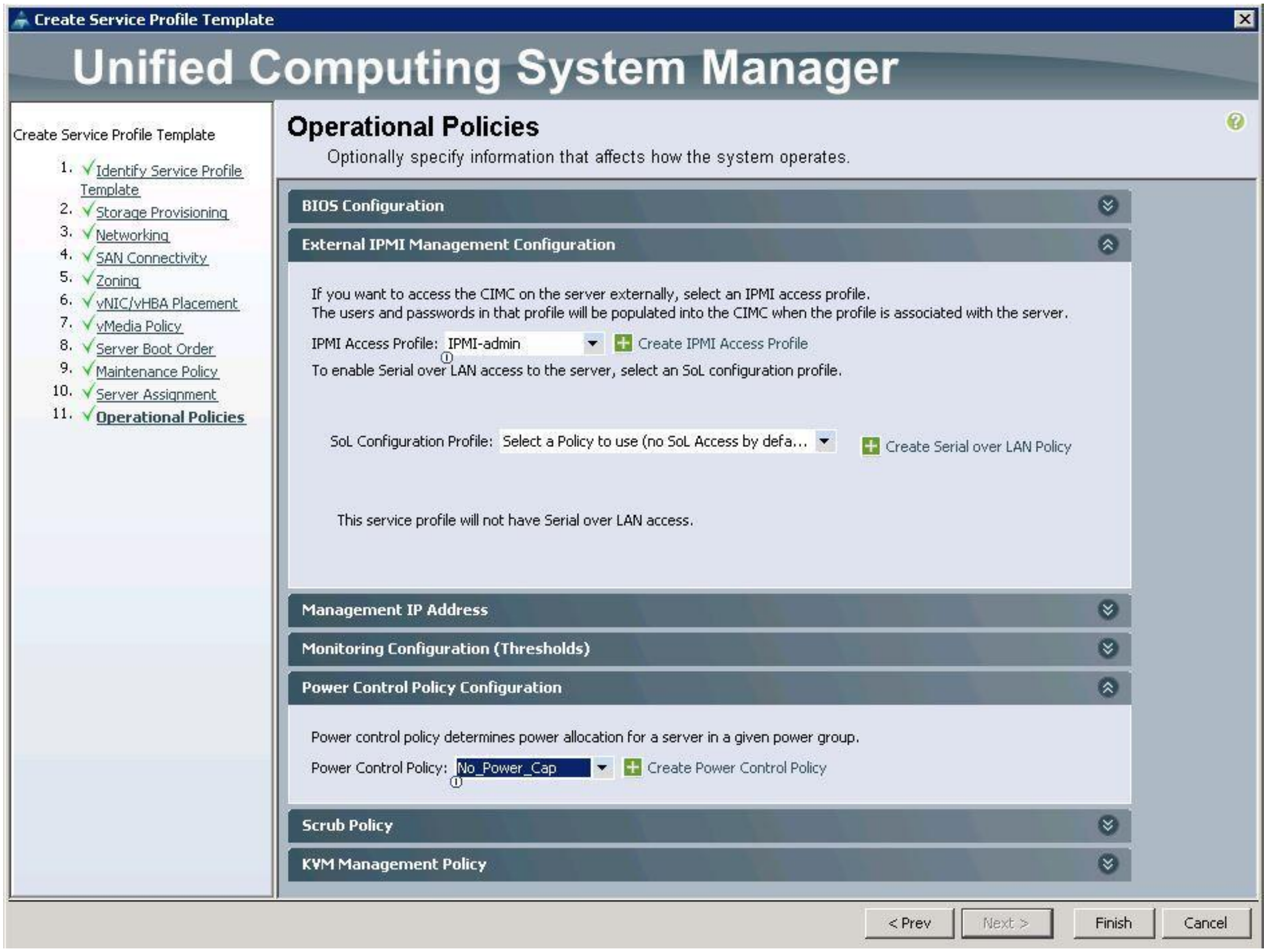
Maintenance Policy

Select a maintenance policy to include with this service profile or create a new maintenance policy that will be accessible to all service profiles.

Maintenance Policy: Server_Ack

Name: **Server_Ack**
Description:
Reboot Policy: **User Ack**

Under Operational Policies, choose the IPMI Access Profile as IPMI_admin previously created, from the drop-down list and choose the Power Control Policy as No_Power_Cap and click Finish.



Operational Policies

Optionally specify information that affects how the system operates.

External IPMI Management Configuration

If you want to access the CIMC on the server externally, select an IPMI access profile.
The users and passwords in that profile will be populated into the CIMC when the profile is associated with the server.

IPMI Access Profile: IPMI-admin   Create IPMI Access Profile

To enable Serial over LAN access to the server, select an SoL configuration profile.

SoL Configuration Profile: Select a Policy to use (no SoL Access by defa...   Create Serial over LAN Policy

This service profile will not have Serial over LAN access.

Power Control Policy Configuration

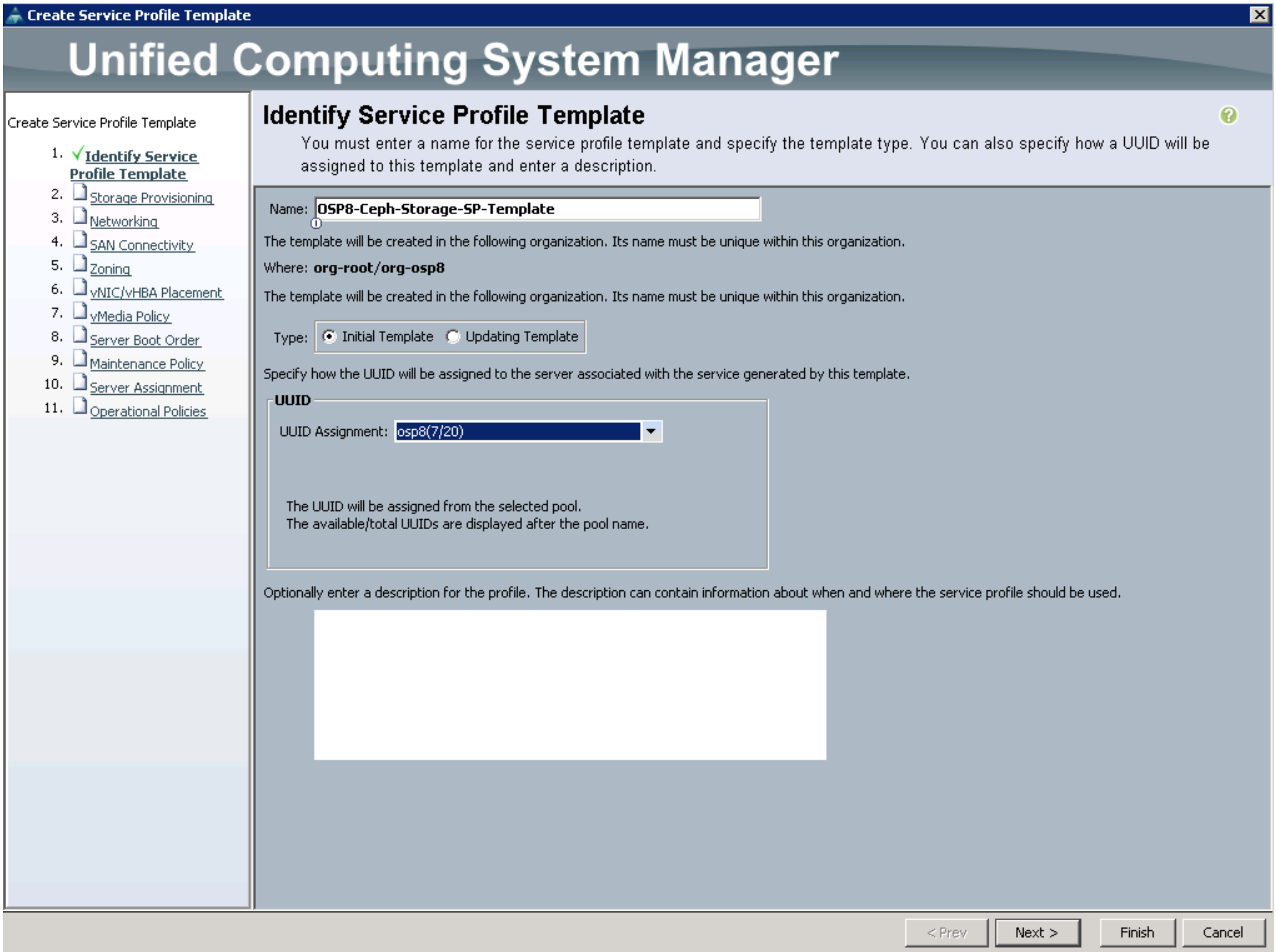
Power control policy determines power allocation for a server in a given power group.

Power Control Policy: No_Power_Cap   Create Power Control Policy

## Create Service Profile Templates for Ceph Storage Nodes

To create the Service Profile templates for the Ceph Storage nodes, complete the following steps:

Specify the Service profile template name for the Ceph storage node as OSP8-Ceph-Storage-SP-Template. Choose the UUID pools previously created, from the drop-down list and click Next.



Identify Service Profile Template

You must enter a name for the service profile template and specify the template type. You can also specify how a UUID will be assigned to this template and enter a description.

Name: OSP8-Ceph-Storage-SP-Template

The template will be created in the following organization. Its name must be unique within this organization.

Where: **org-root/org-osp8**

The template will be created in the following organization. Its name must be unique within this organization.

Type: ● Initial Template ○ Updating Template

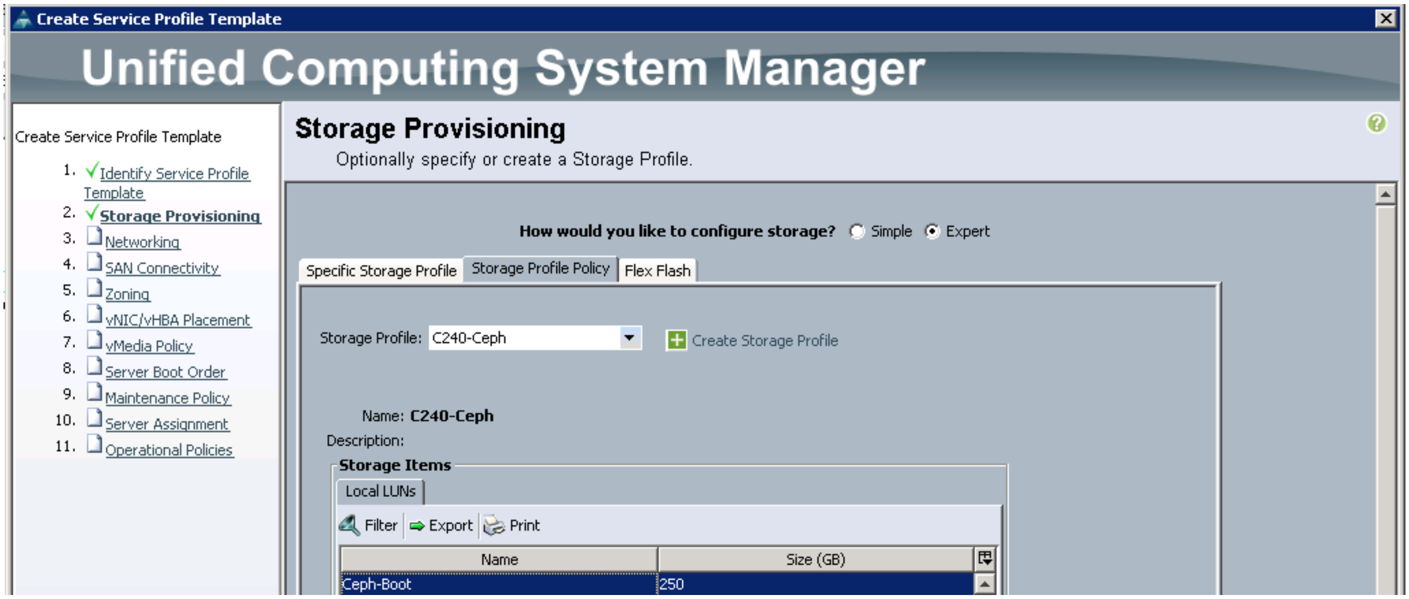
Specify how the UUID will be assigned to the server associated with the service generated by this template.

**UUID**

UUID Assignment: osp8(7/20)

The UUID will be assigned from the selected pool.
The available/total UUIDs are displayed after the pool name.

Optionally enter a description for the profile. The description can contain information about when and where the service profile should be used.



Storage Provisioning

Optionally specify or create a Storage Profile.

**How would you like to configure storage?** ○ Simple ● Expert

Storage Profile: C240-Ceph

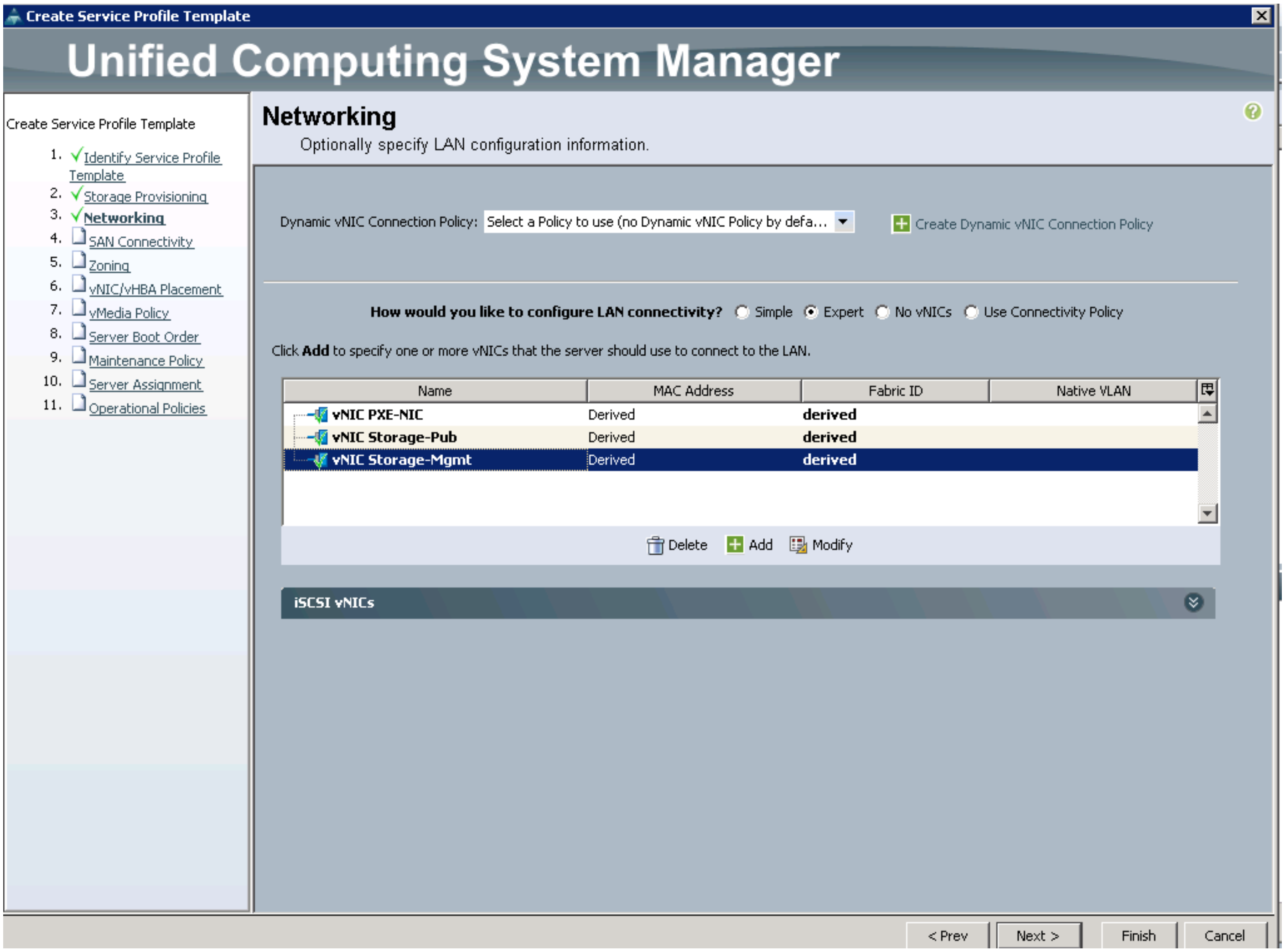
Name: **C240-Ceph**

Description:

**Storage Items**

| Name | Size (GB) |
|---|---|
| Ceph-Boot | 250 |

Create vNICs for PXE, Stroage-Pub and Storage-Mgmt by following the steps detailed here. You do not need Management Network on Ceph Nodes since this network was created only for UCSM and Nexus Switches.



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ **Networking**
4. □ SAN Connectivity
5. □ Zoning
6. □ vNIC/vHBA Placement
7. □ vMedia Policy
8. □ Server Boot Order
9. □ Maintenance Policy
10. □ Server Assignment
11. □ Operational Policies

## Networking

Optionally specify LAN configuration information.

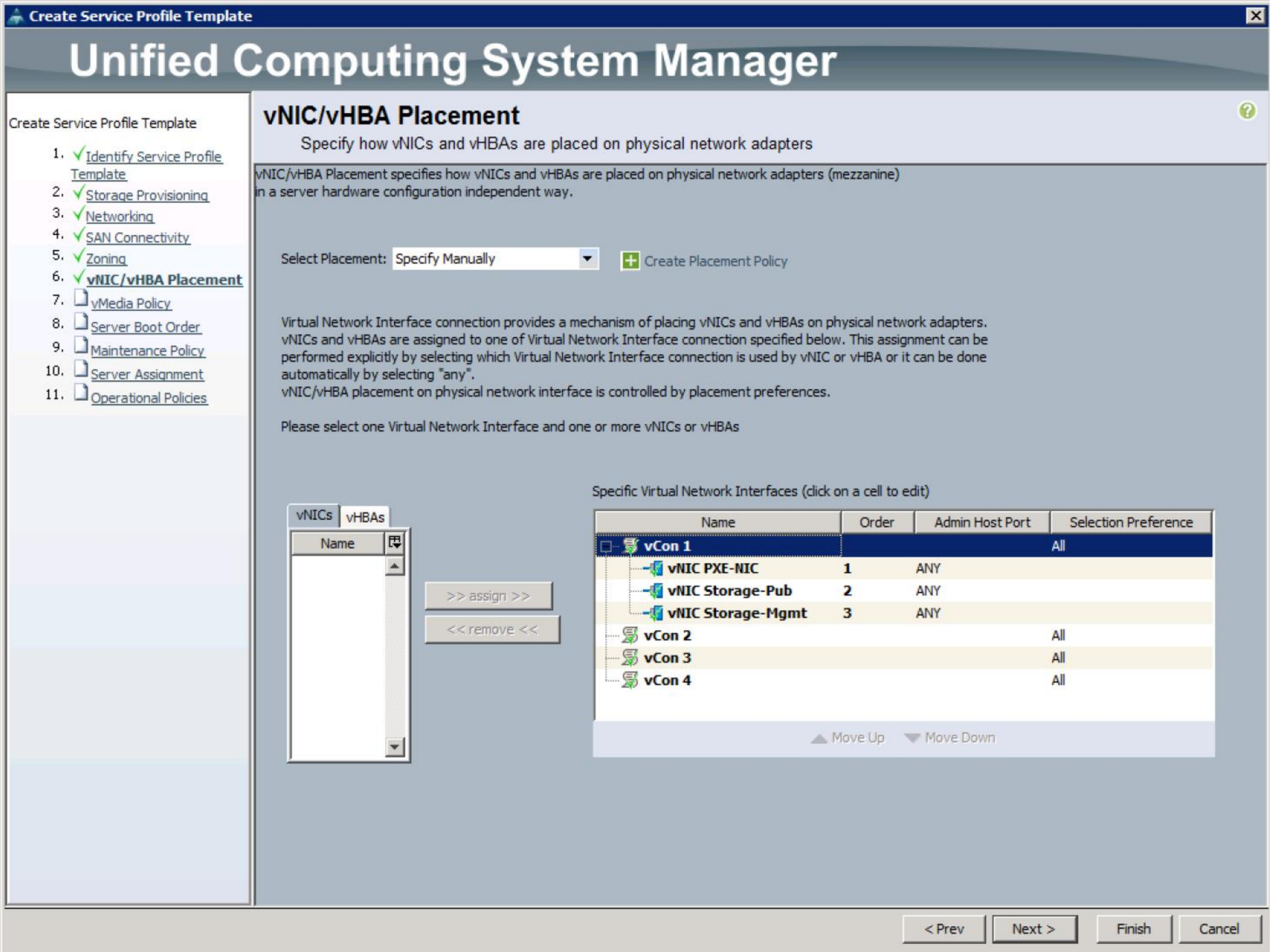
Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

Create Dynamic vNIC Connection Policy

**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|
| **vNIC PXE-NIC** | Derived | **derived** | |
| **vNIC Storage-Pub** | Derived | **derived** | |
| **vNIC Storage-Mgmt** | Derived | **derived** | |



Create Service Profile Template

Unified Computing System Manager

Create Service Profile Template

1. ✓ Identify Service Profile Template
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ **vNIC/vHBA Placement**
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

## vNIC/vHBA Placement

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine)
in a server hardware configuration independent way.

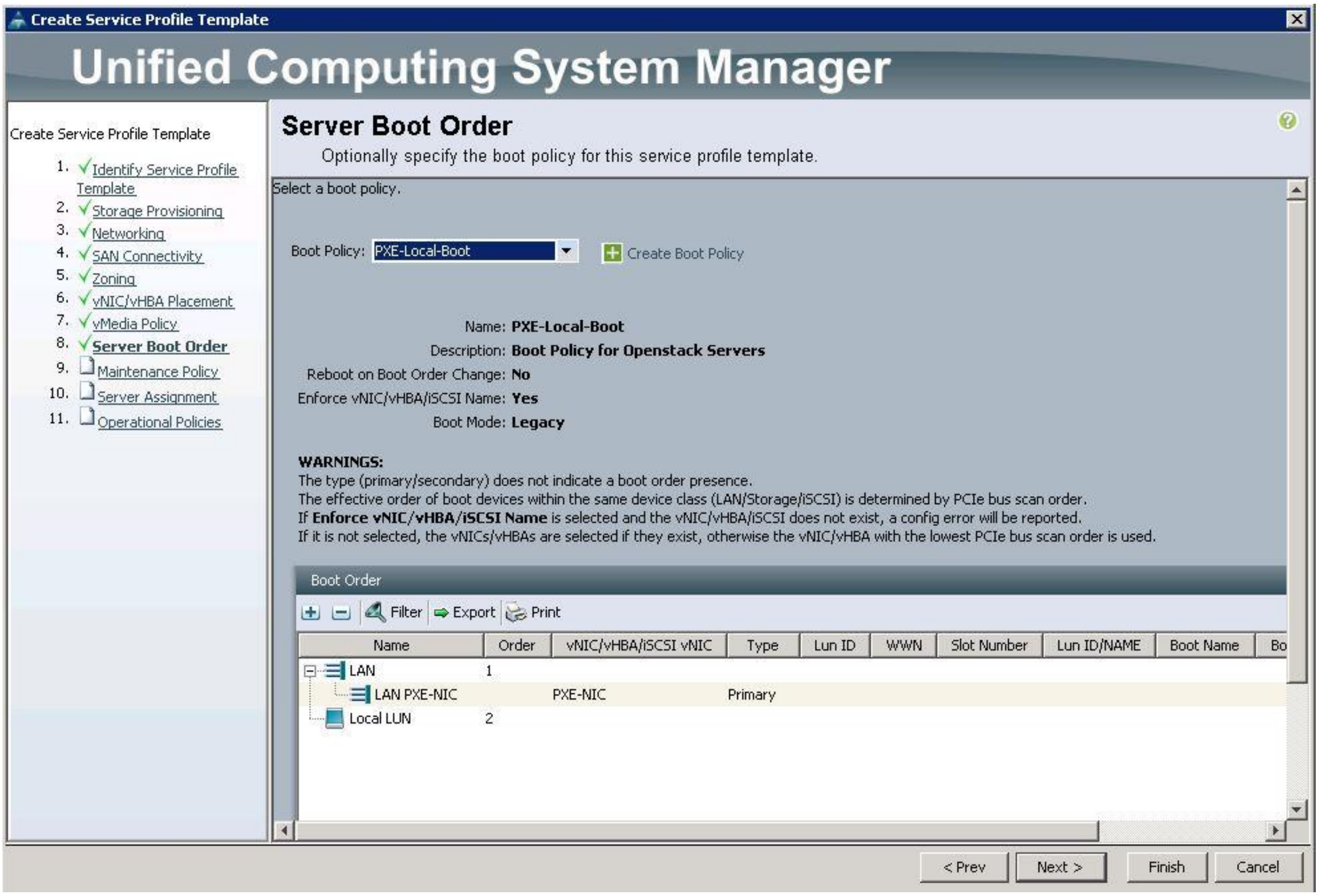
Select Placement: Specify Manually

Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vCon 1 | | | All |
| vNIC PXE-NIC | 1 | ANY | |
| vNIC Storage-Pub | 2 | ANY | |
| vNIC Storage-Mgmt | 3 | ANY | |
| vCon 2 | | | All |
| vCon 3 | | | All |
| vCon 4 | | | All |



Create Service Profile Template

Unified Computing System Manager

Server Boot Order

Optionally specify the boot policy for this service profile template.

Select a boot policy.

Boot Policy: PXE-Local-Boot

Name: **PXE-Local-Boot**
Description: **Boot Policy for Openstack Servers**
Reboot on Boot Order Change: **No**
Enforce vNIC/vHBA/iSCSI Name: **Yes**
Boot Mode: **Legacy**

**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determined by PCIe bus scan order.
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a config error will be reported.
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the lowest PCIe bus scan order is used.

Boot Order

| Name | Order | vNIC/vHBA/iSCSI vNIC | Type | Lun ID | WWN | Slot Number | Lun ID/NAME | Boot Name | Bo |
|---|---|---|---|---|---|---|---|---|---|
| LAN | 1 | | | | | | | | |
| LAN PXE-NIC | | PXE-NIC | Primary | | | | | | |
| Local LUN | 2 | | | | | | | | |

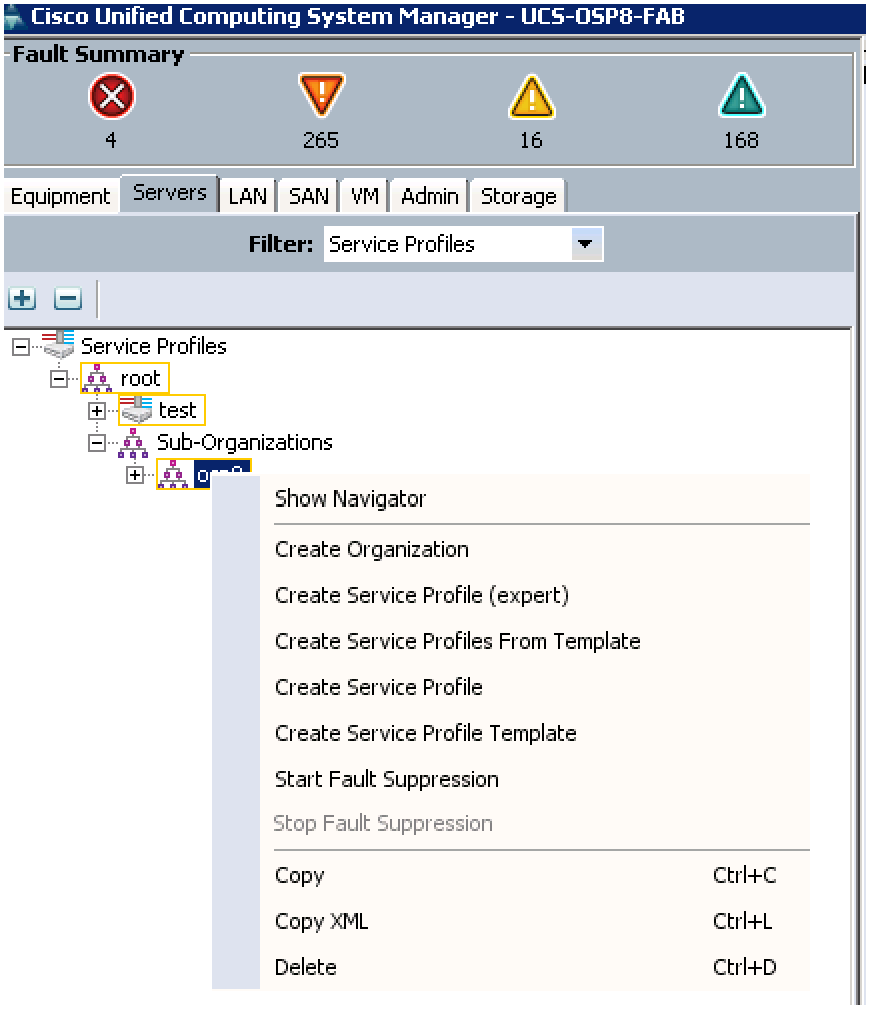
Click Next and then Choose “Server_Ack” under Maintenance Policy. Then select “No-power-cap” under power control policy. Click on Finish to complete the Service profile template creation for Ceph nodes.

## Create Service Profile for Undercloud ( OSP8 director ) Node

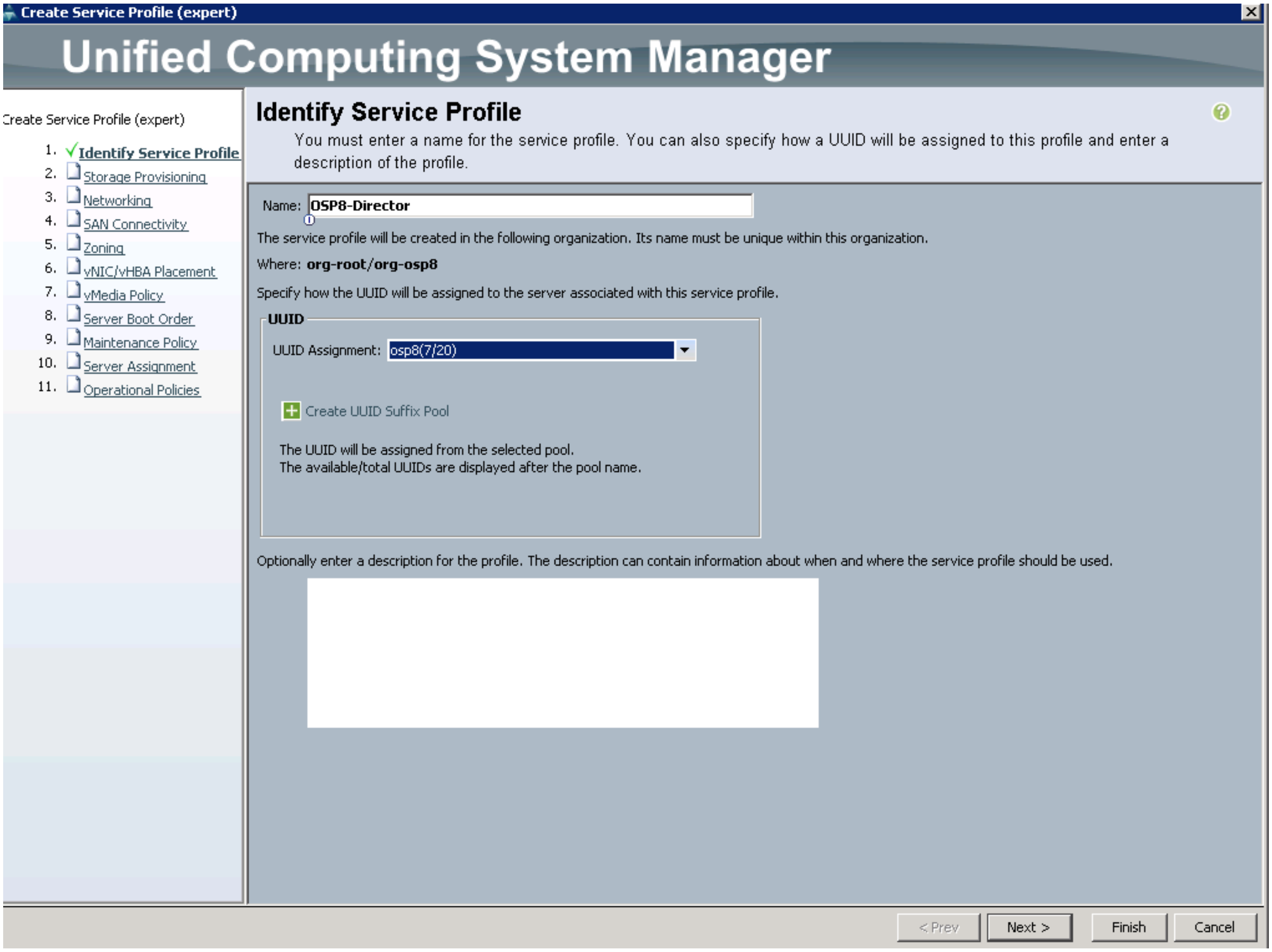
To configure the Service Profile for Undercloud (OSP8 director) Node, complete the following steps:

As there is only one node for Undercloud, a single Service Profile is created. There are no Service Profile Templates for the Undercloud node.

Under Servers > Service Profiles > root > Sub-Organizations -> osp8 -> right-click and select “Create Service Profile (expert)”



Specify the Service profile name for Undercloud node as OSP8-director. Choose the UUID pools previously created from the drop-down list and click Next.



**Identify Service Profile**

You must enter a name for the service profile. You can also specify how a UUID will be assigned to this profile and enter a description of the profile.

Name: OSP8-Director

The service profile will be created in the following organization. Its name must be unique within this organization.

Where: **org-root/org-osp8**

Specify how the UUID will be assigned to the server associated with this service profile.

**UUID**

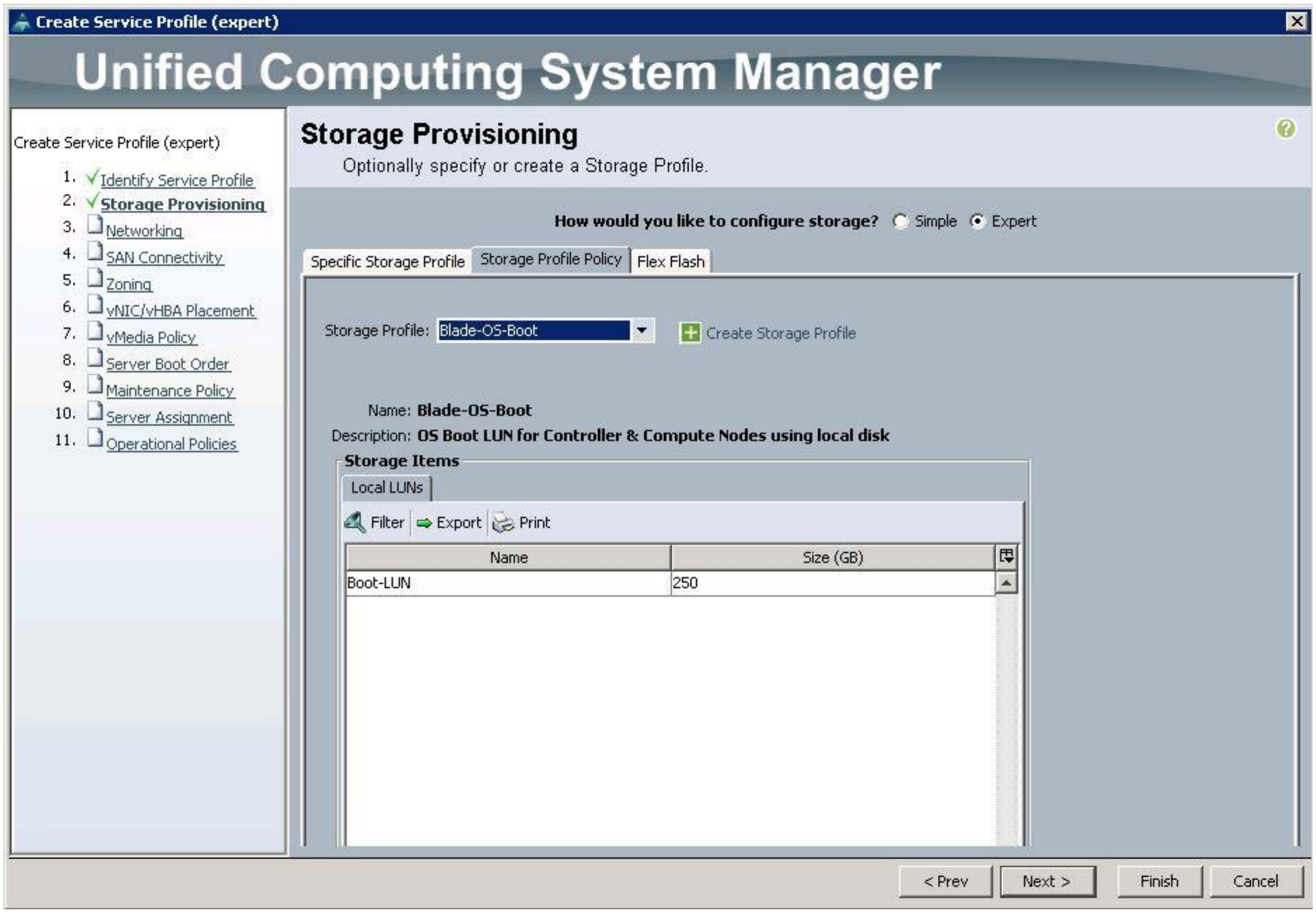
UUID Assignment: osp8(7/20)

Create UUID Suffix Pool

The UUID will be assigned from the selected pool.
The available/total UUIDs are displayed after the pool name.

Optionally enter a description for the profile. The description can contain information about when and where the service profile should be used.

For Storage Provisioning, choose Expert and click Storage profile Policy and choose the Storage profile Blade-OS-boot previously created from the drop-down list and click Next.



How would you like to configure storage? ○ Simple ● Expert

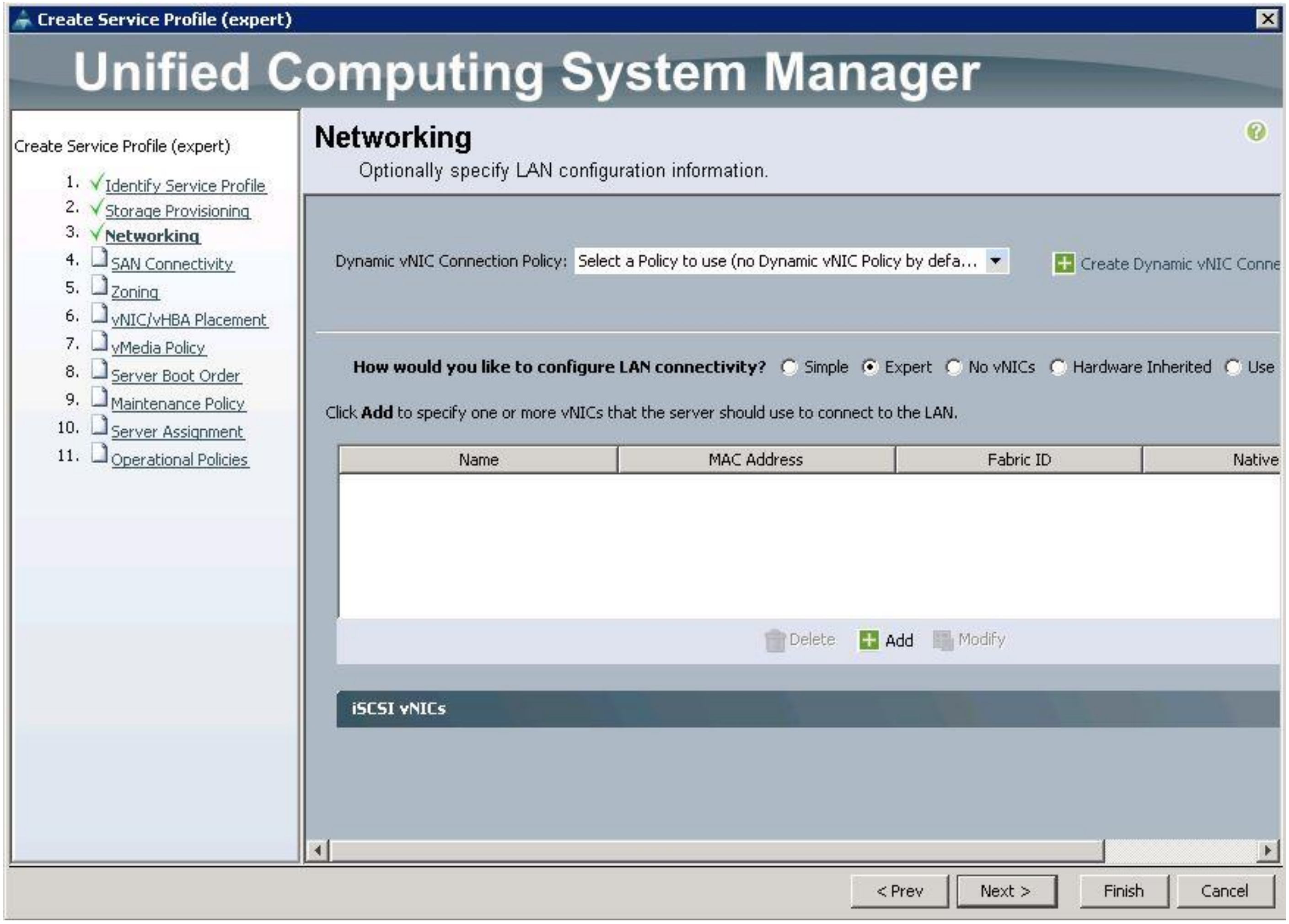
Storage Profile: Blade-OS-Boot

Name: Blade-OS-Boot

Description: OS Boot LUN for Controller & Compute Nodes using local disk

| Name | Size (GB) |
|---|---|
| Boot-LUN | 250 |

For Networking, choose Expert and click "+".



**Networking**

Optionally specify LAN configuration information.

Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

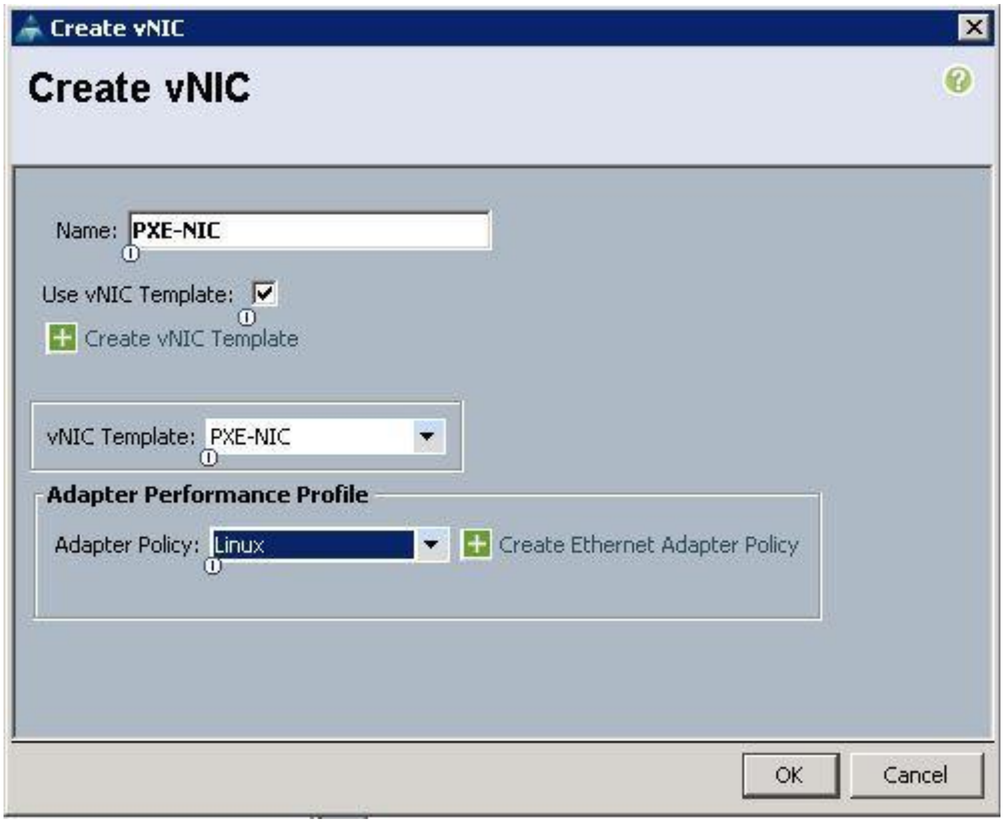
**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Hardware Inherited ○ Use

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

| Name | MAC Address | Fabric ID | Native |
|---|---|---|---|

Create the VNIC interface for PXE or Provisioning network as PXE-NIC and click the check box Use VNIC template.

Under vNIC template, choose the PXE-NIC template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: PXE-NIC

Use vNIC Template: ☑

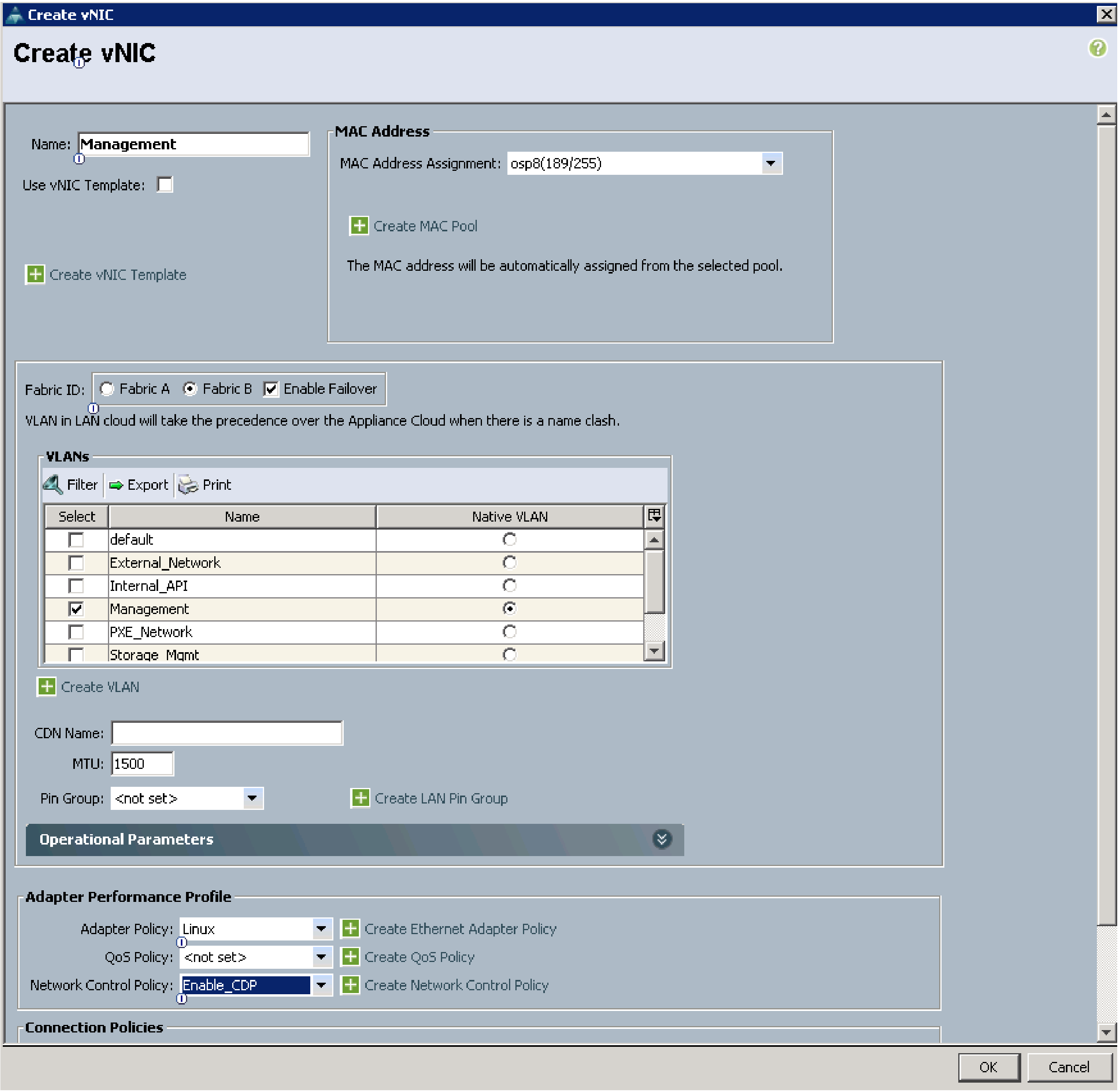
vNIC Template: PXE-NIC

Adapter Performance Profile

Adapter Policy: Linux

Create the VNIC interface for  Management network as Management and from the drop-down list choose MAC pools created before. Click Fabric B and check Enable Failover.

Under VLANs, Select “Management” as Native VLAN, then  choose Adapter Policy as “Linux” and Network Controller Policy as “Enable_CDP”



Create vNIC

## Create vNIC

Name: Management

Use vNIC Template: □

Create vNIC Template

**MAC Address**

MAC Address Assignment: osp8(189/255)

Create MAC Pool

The MAC address will be automatically assigned from the selected pool.

Fabric ID: ○ Fabric A ● Fabric B ☑ Enable Failover

VLAN in LAN cloud will take the precedence over the Appliance Cloud when there is a name clash.

**VLANs**

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| □ | External_Network | ○ |
| □ | Internal_API | ○ |
| ☑ | Management | ● |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |

Create VLAN

CDN Name:

MTU: 1500

Pin Group: <not set>

Create LAN Pin Group

**Operational Parameters**

**Adapter Performance Profile**

Adapter Policy: Linux

Create Ethernet Adapter Policy

QoS Policy: <not set>

Create QoS Policy

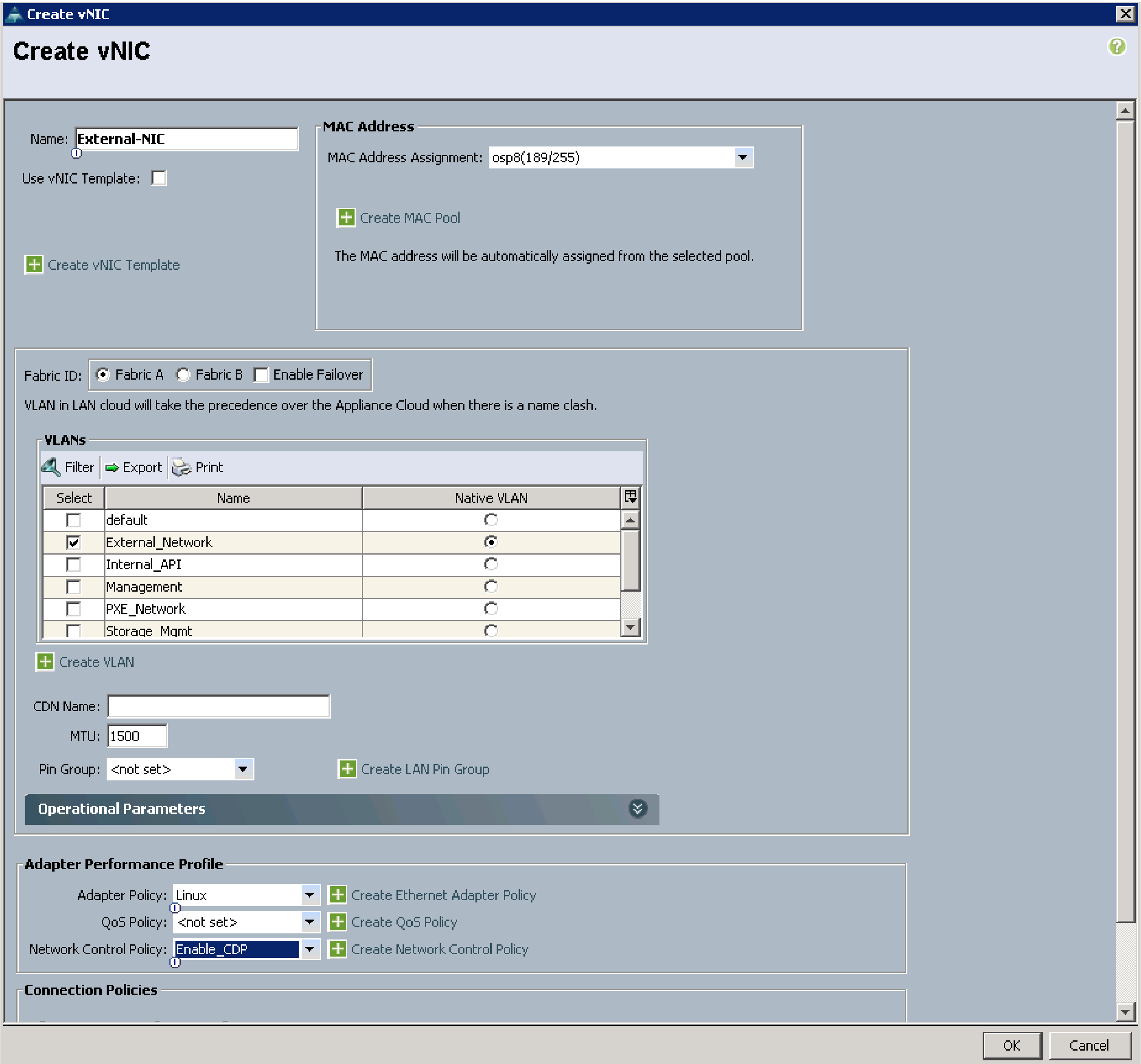
Network Control Policy: Enable_CDP

Create Network Control Policy

**Connection Policies**

Create the VNIC interface for External network as External-NIC and from the drop-down list choose MAC pools created before. Click Fabric B and check Enable Failover.

Under VLANs, Select “External” as Native VLAN, then choose Adapter Policy as “Linux” and Network Controller Policy as “Enable_CDP”



Create vNIC

Name: External-NIC

Use vNIC Template: □

MAC Address

MAC Address Assignment: osp8(189/255)

The MAC address will be automatically assigned from the selected pool.

Fabric ID: ● Fabric A ○ Fabric B □ Enable Failover

VLAN in LAN cloud will take the precedence over the Appliance Cloud when there is a name clash.

VLANs

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| ☑ | External_Network | ● |
| □ | Internal_API | ○ |
| □ | Management | ○ |
| □ | PXE_Network | ○ |
| □ | Storage_Mgmt | ○ |

CDN Name:

MTU: 1500

Pin Group: <not set>

Adapter Performance Profile

Adapter Policy: Linux

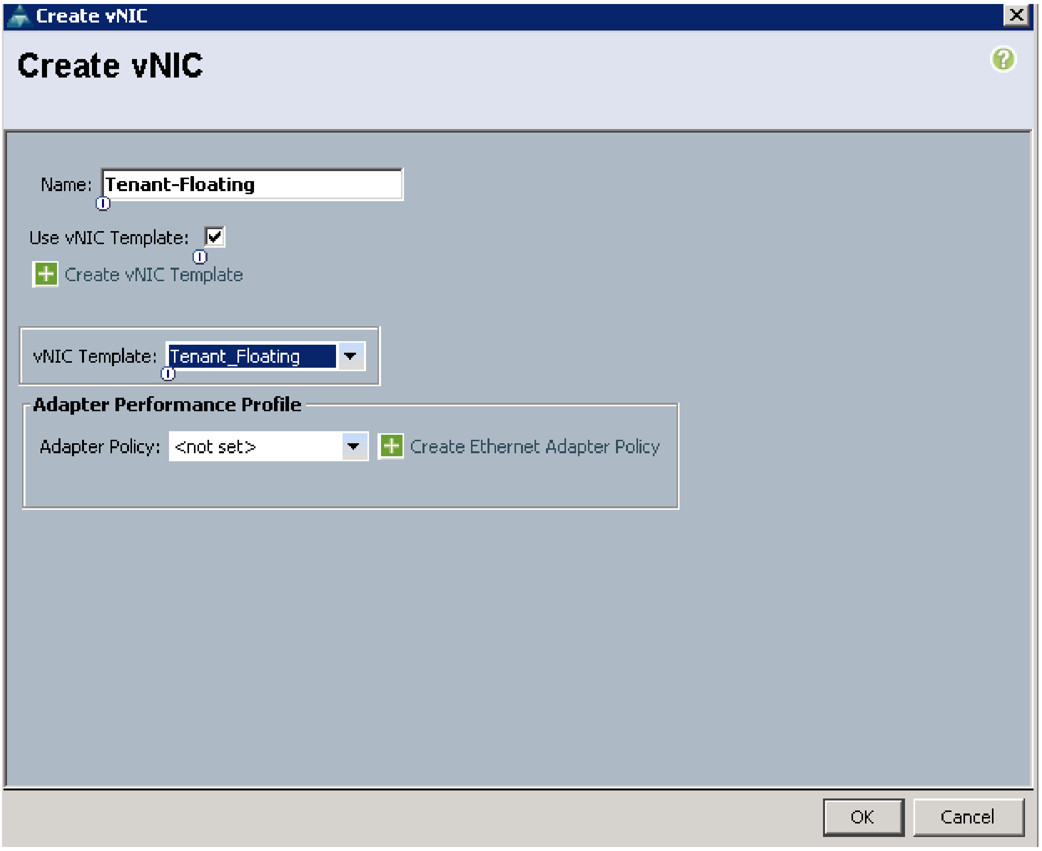
QoS Policy: <not set>

Network Control Policy: Enable_CDP

Connection Policies

Create the VNIC interface for Floating network as Tenant-Floating and click the check box Use VNIC template.

Under vNIC template, choose the Floating Network (optional) template previously created from the drop-down list and choose Linux for the Adapter Policy.



Create vNIC

Name: Tenant-Floating

Use vNIC Template: ☑

Create vNIC Template

vNIC Template: Tenant_Floating

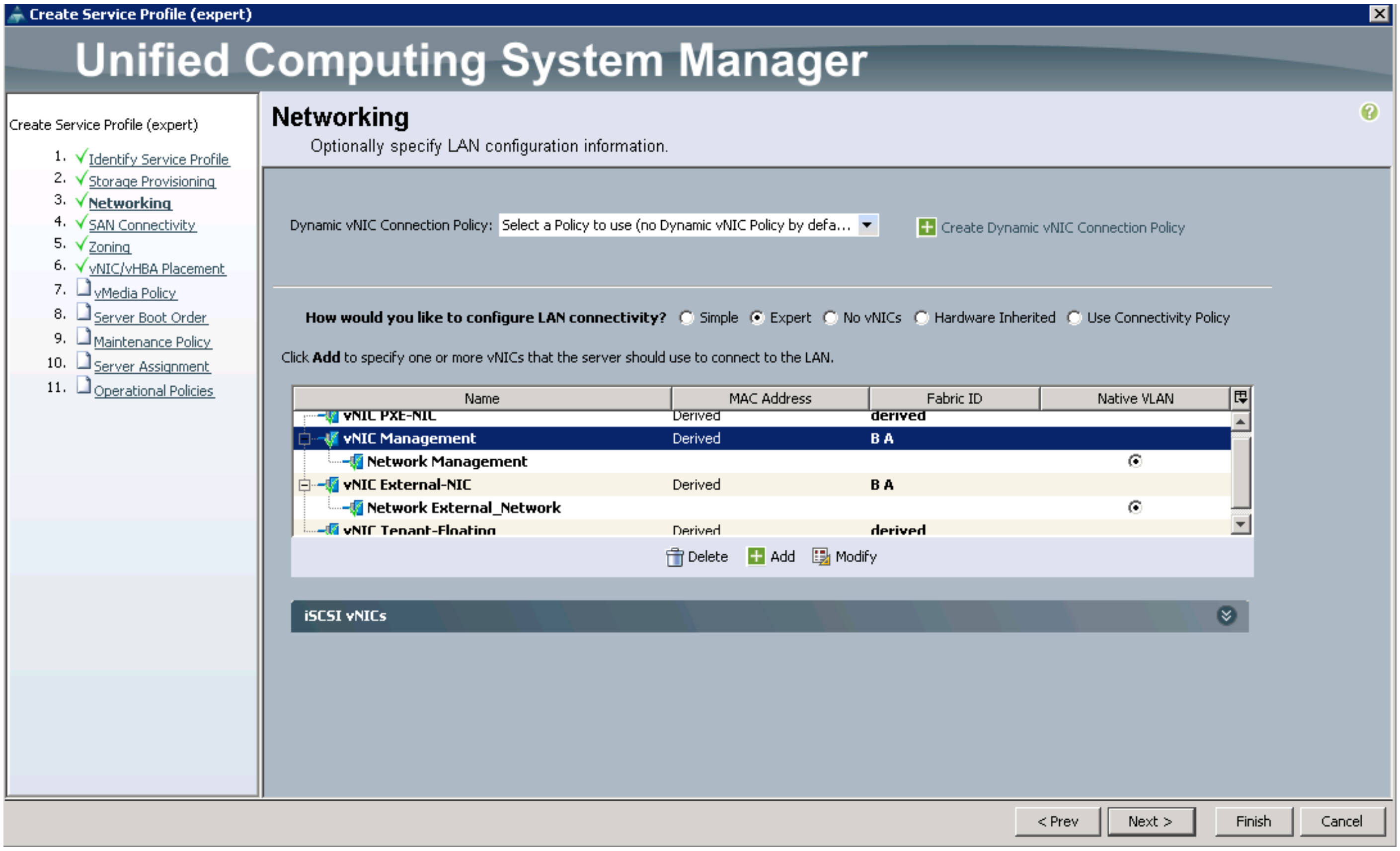
Adapter Performance Profile

Adapter Policy: <not set>

Create Ethernet Adapter Policy

Tenant Floating NIC has been created on the system just to verify that access to the VMs from director node works. If you would like to access the VMs externally from another node outside of this setup and this floating network is routable, this step isn't necessary.

After a successful VNIC creation, click Next.



Create Service Profile (expert)

Unified Computing System Manager

Create Service Profile (expert)

1. ✓ Identify Service Profile
2. ✓ Storage Provisioning
3. ✓ **Networking**
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**Networking**

Optionally specify LAN configuration information.

Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by defa...

Create Dynamic vNIC Connection Policy

**How would you like to configure LAN connectivity?** ○ Simple ● Expert ○ No vNICs ○ Hardware Inherited ○ Use Connectivity Policy

Click **Add** to specify one or more vNICs that the server should use to connect to the LAN.

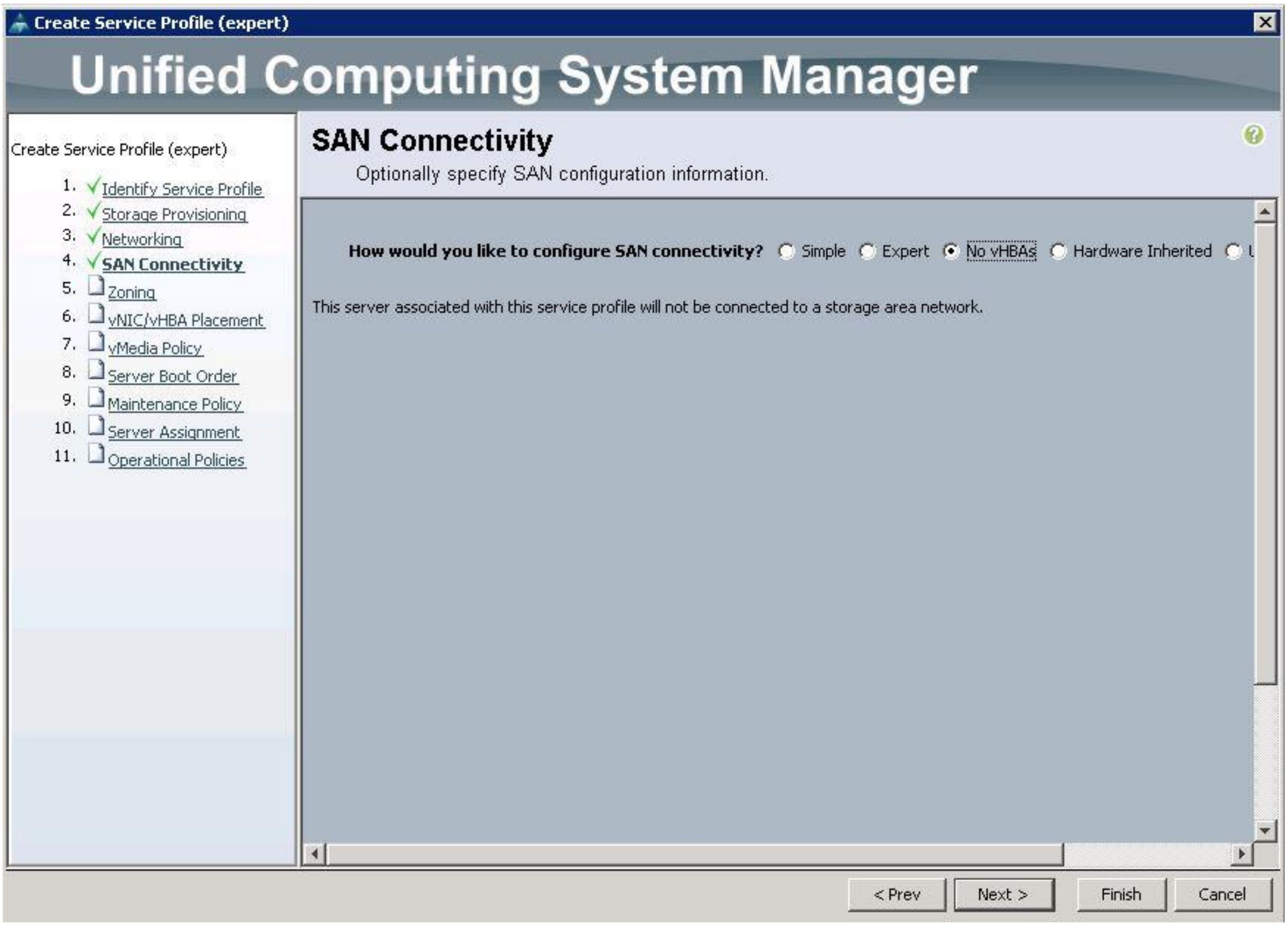
| Name | MAC Address | Fabric ID | Native VLAN |
|---|---|---|---|
| vNIC PXE-NIC | Derived | derived | |
| vNIC Management | Derived | B A | |
| Network Management | | | ● |
| vNIC External-NIC | Derived | B A | |
| Network External_Network | | | ● |
| vNIC Tenant-Floating | Derived | derived | |

Delete Add Modify

iSCSI vNICs

< Prev Next > Finish Cancel

Under the SAN connectivity, choose No VHBAs and click Next.



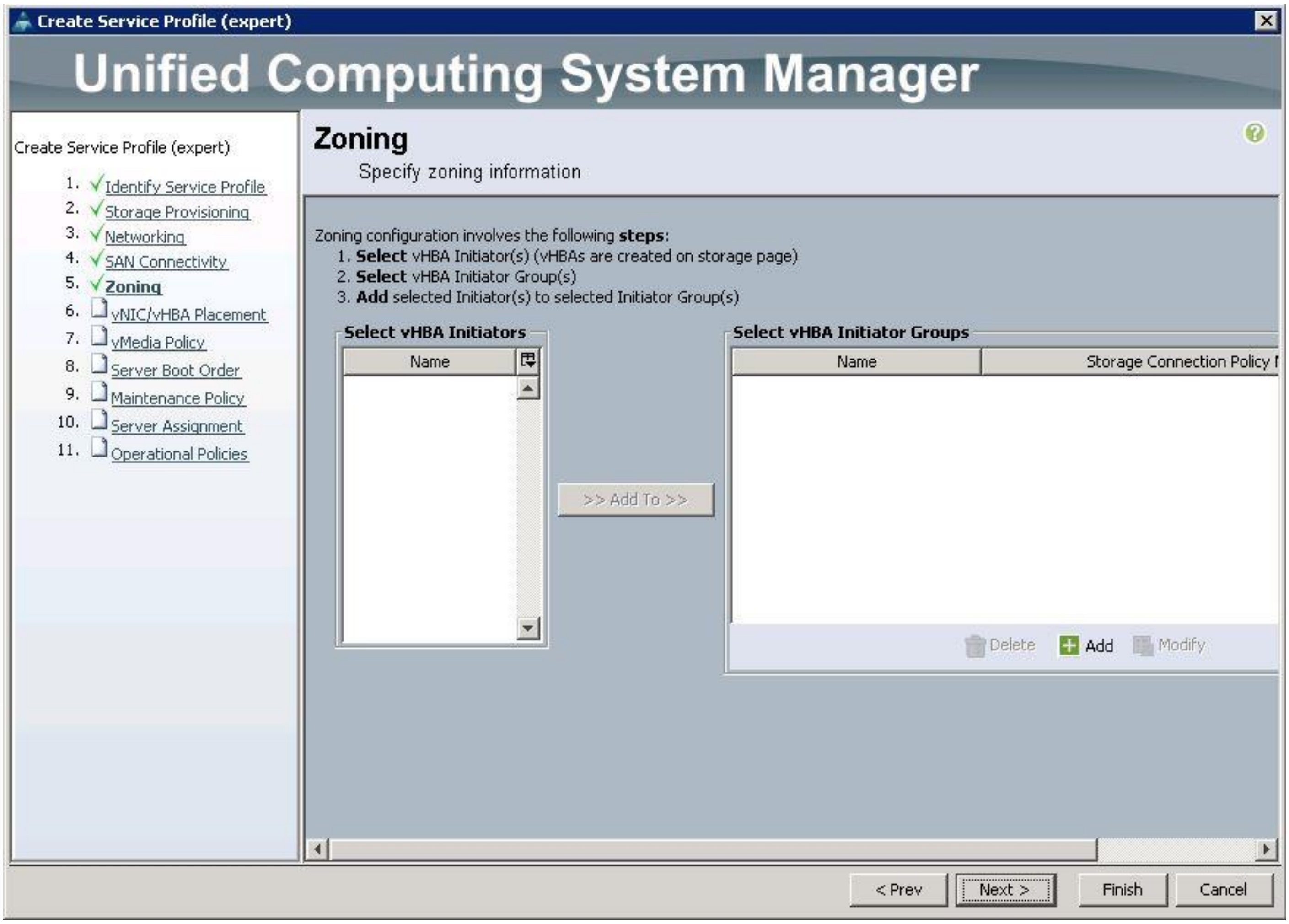
SAN Connectivity

Optionally specify SAN configuration information.

How would you like to configure SAN connectivity? ○ Simple ○ Expert ● No vHBAs ○ Hardware Inherited ○ U

This server associated with this service profile will not be connected to a storage area network.

Under Zoning, click Next.



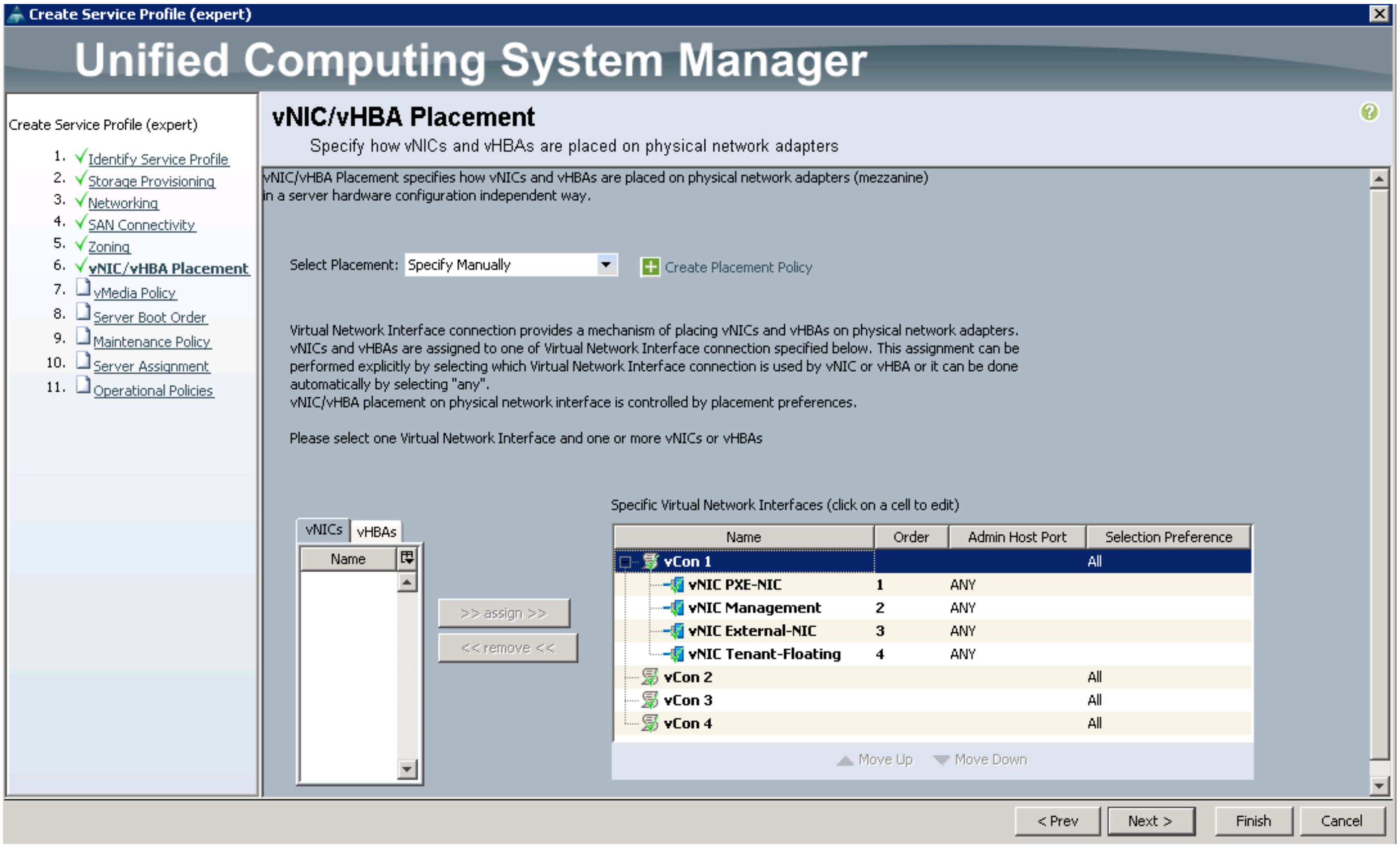
Zoning

Specify zoning information

Zoning configuration involves the following **steps**:

1. **Select** vHBA Initiator(s) (vHBAs are created on storage page)
2. **Select** vHBA Initiator Group(s)
3. **Add** selected Initiator(s) to selected Initiator Group(s)

Under VNIC/VHBA Placement, choose the vNICs PCI order as shown below and click Next.



Create Service Profile (expert)

Unified Computing System Manager

Create Service Profile (expert)

1. ✓ Identify Service Profile
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ **vNIC/vHBA Placement**
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**vNIC/vHBA Placement**

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement: Specify Manually

Create Placement Policy

Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any".

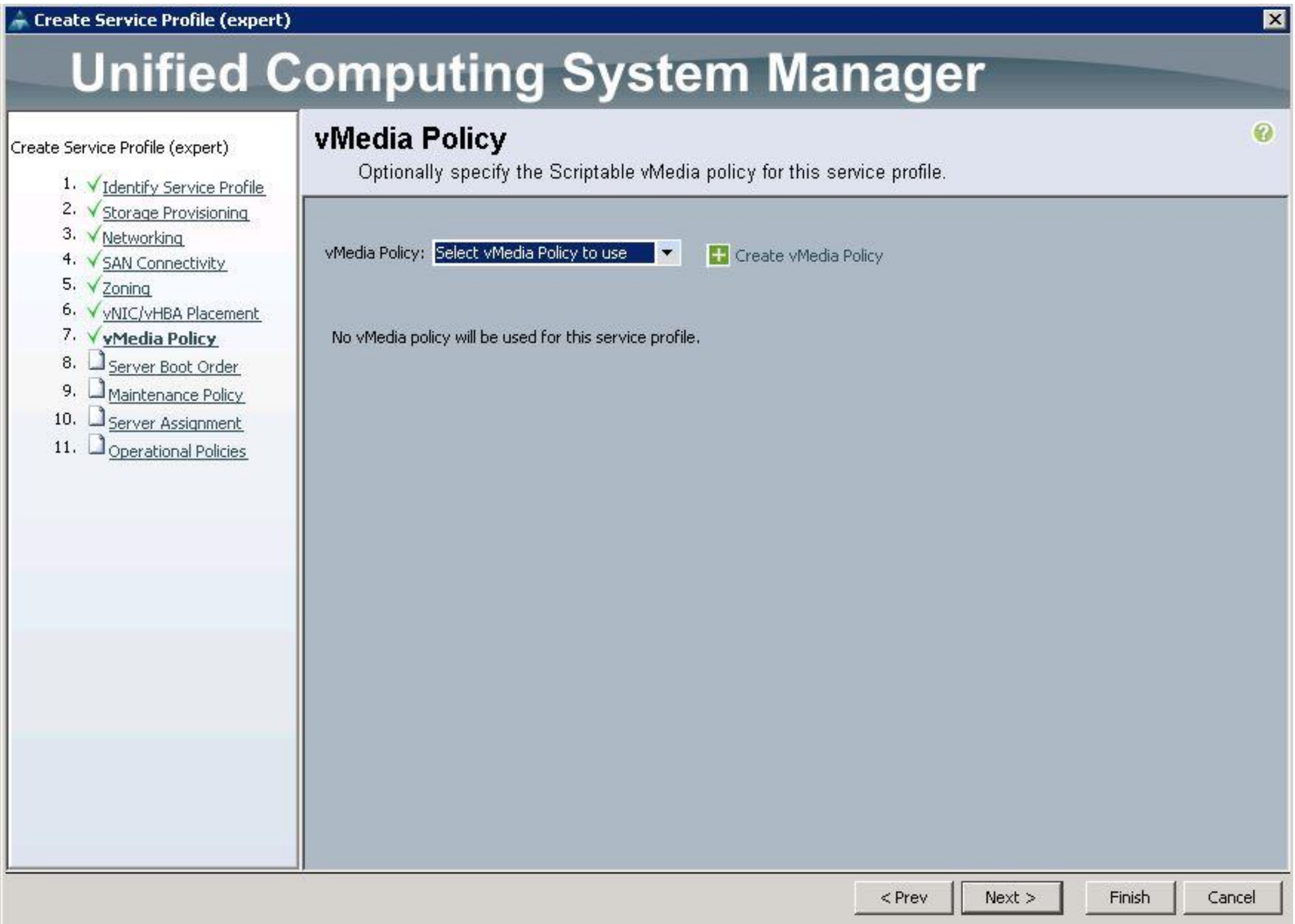
vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Specific Virtual Network Interfaces (click on a cell to edit)

| Name | Order | Admin Host Port | Selection Preference |
|---|---|---|---|
| vCon 1 | | | All |
| vNIC PXE-NIC | 1 | ANY | |
| vNIC Management | 2 | ANY | |
| vNIC External-NIC | 3 | ANY | |
| vNIC Tenant-Floating | 4 | ANY | |
| vCon 2 | | | All |
| vCon 3 | | | All |
| vCon 4 | | | All |

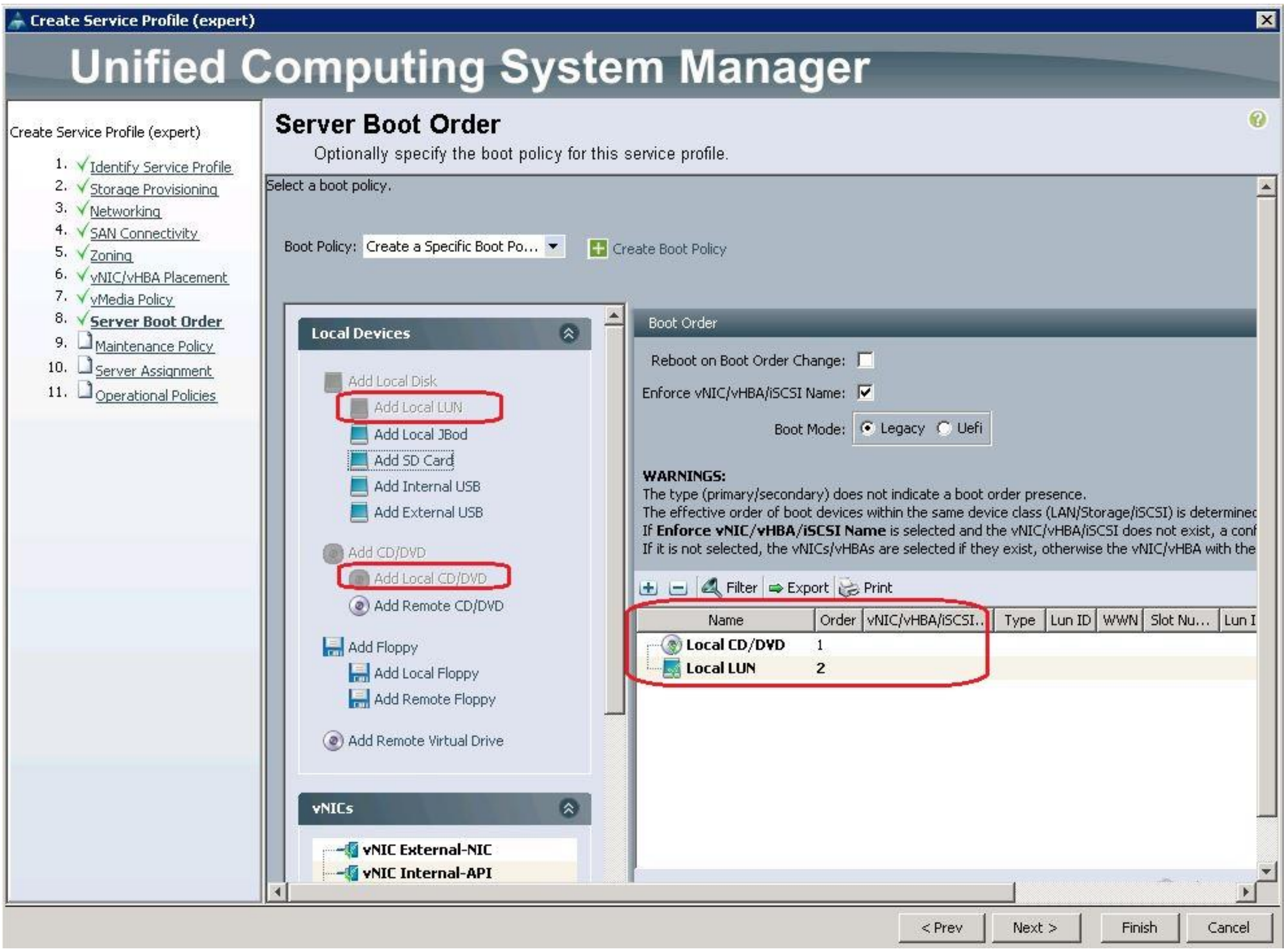
Under vMedia Policy, click Next.



vMedia Policy: Select vMedia Policy to use

No vMedia policy will be used for this service profile.

Under Server Boot Order, choose the boot policy as "Create a Specific Boot Policy", from the drop-down list and click Next. Make sure you select " local CD/DVD" as first boot order and " local LUN" as second boot order and click Next.



Create Service Profile (expert)

Unified Computing System Manager

Create Service Profile (expert)

1. ✓ Identify Service Profile
2. ✓ Storage Provisioning
3. ✓ Networking
4. ✓ SAN Connectivity
5. ✓ Zoning
6. ✓ vNIC/vHBA Placement
7. ✓ vMedia Policy
8. ✓ **Server Boot Order**
9. Maintenance Policy
10. Server Assignment
11. Operational Policies

**Server Boot Order**

Optionally specify the boot policy for this service profile.

Select a boot policy.

Boot Policy: Create a Specific Boot Po...

Reboot on Boot Order Change: ☐

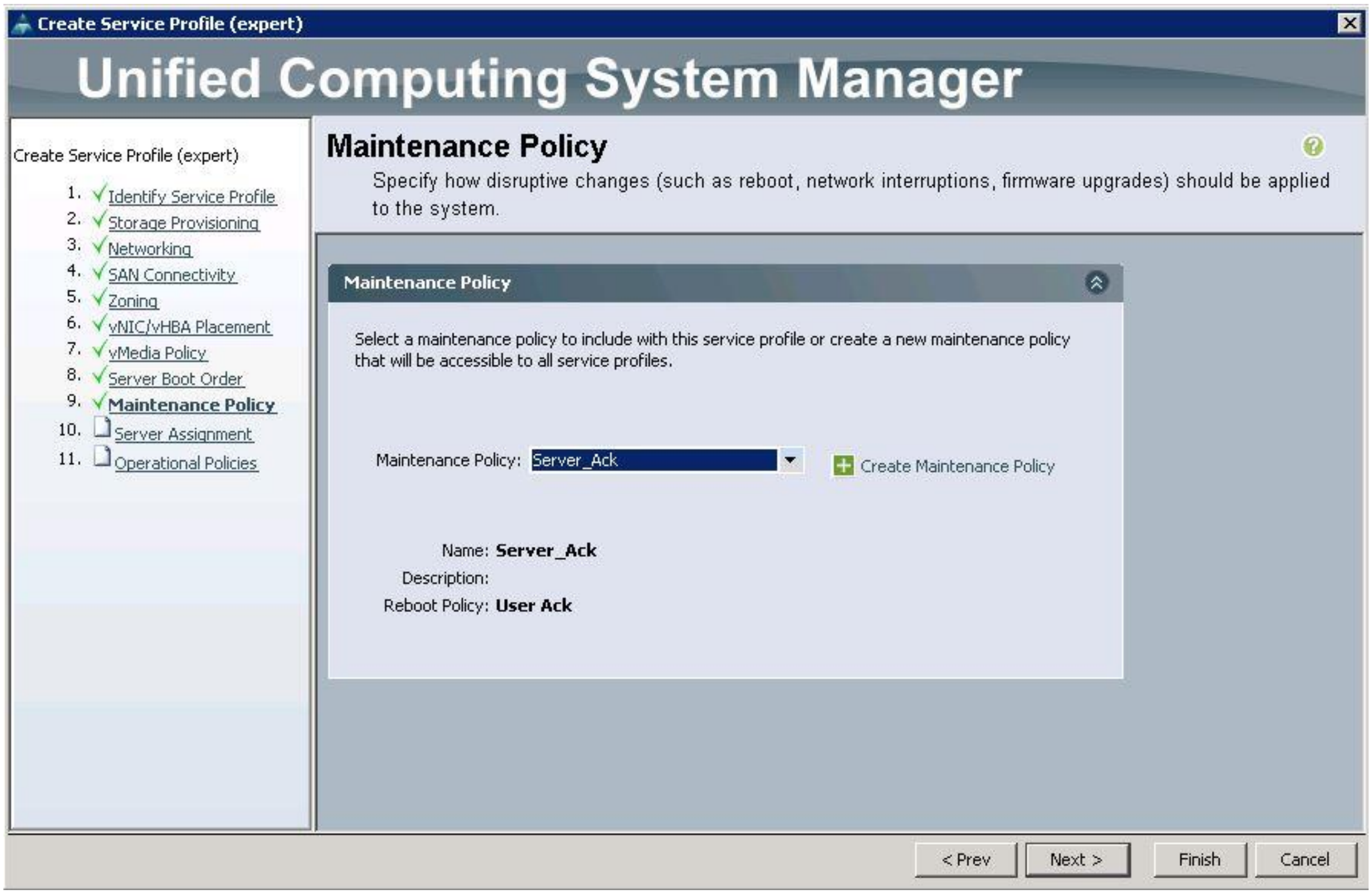
Enforce vNIC/vHBA/iSCSI Name: ☑

Boot Mode: ● Legacy ○ Uefi

**WARNINGS:**
The type (primary/secondary) does not indicate a boot order presence.
The effective order of boot devices within the same device class (LAN/Storage/iSCSI) is determinec
If **Enforce vNIC/vHBA/iSCSI Name** is selected and the vNIC/vHBA/iSCSI does not exist, a conf
If it is not selected, the vNICs/vHBAs are selected if they exist, otherwise the vNIC/vHBA with the

| Name | Order | vNIC/vHBA/iSCSI... | Type | Lun ID | WWN | Slot Nu... | Lun I |
|---|---|---|---|---|---|---|---|
| Local CD/DVD | 1 | | | | | | |
| Local LUN | 2 | | | | | | |

Under Maintenance Policy, choose Server_Ack previously created, from the drop-down list and click Next.



Maintenance Policy

Specify how disruptive changes (such as reboot, network interruptions, firmware upgrades) should be applied to the system.

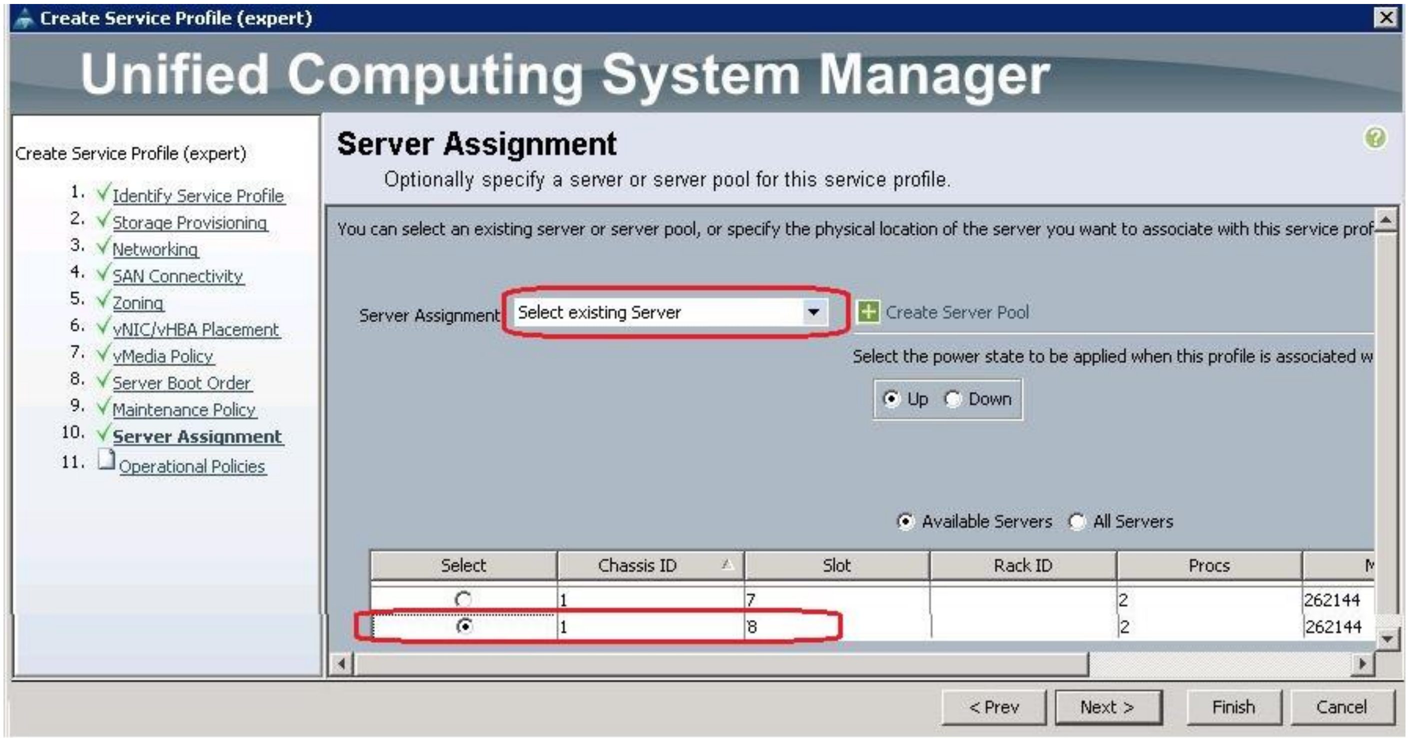
Maintenance Policy

Select a maintenance policy to include with this service profile or create a new maintenance policy that will be accessible to all service profiles.

Maintenance Policy: Server_Ack

Name: **Server_Ack**
Description:
Reboot Policy: **User Ack**

Under Server Assignment, choose  “Select existing server” and select the respective blade assigned for director node and click Next.



Server Assignment

Optionally specify a server or server pool for this service profile.

You can select an existing server or server pool, or specify the physical location of the server you want to associate with this service prof

Server Assignment Select existing Server

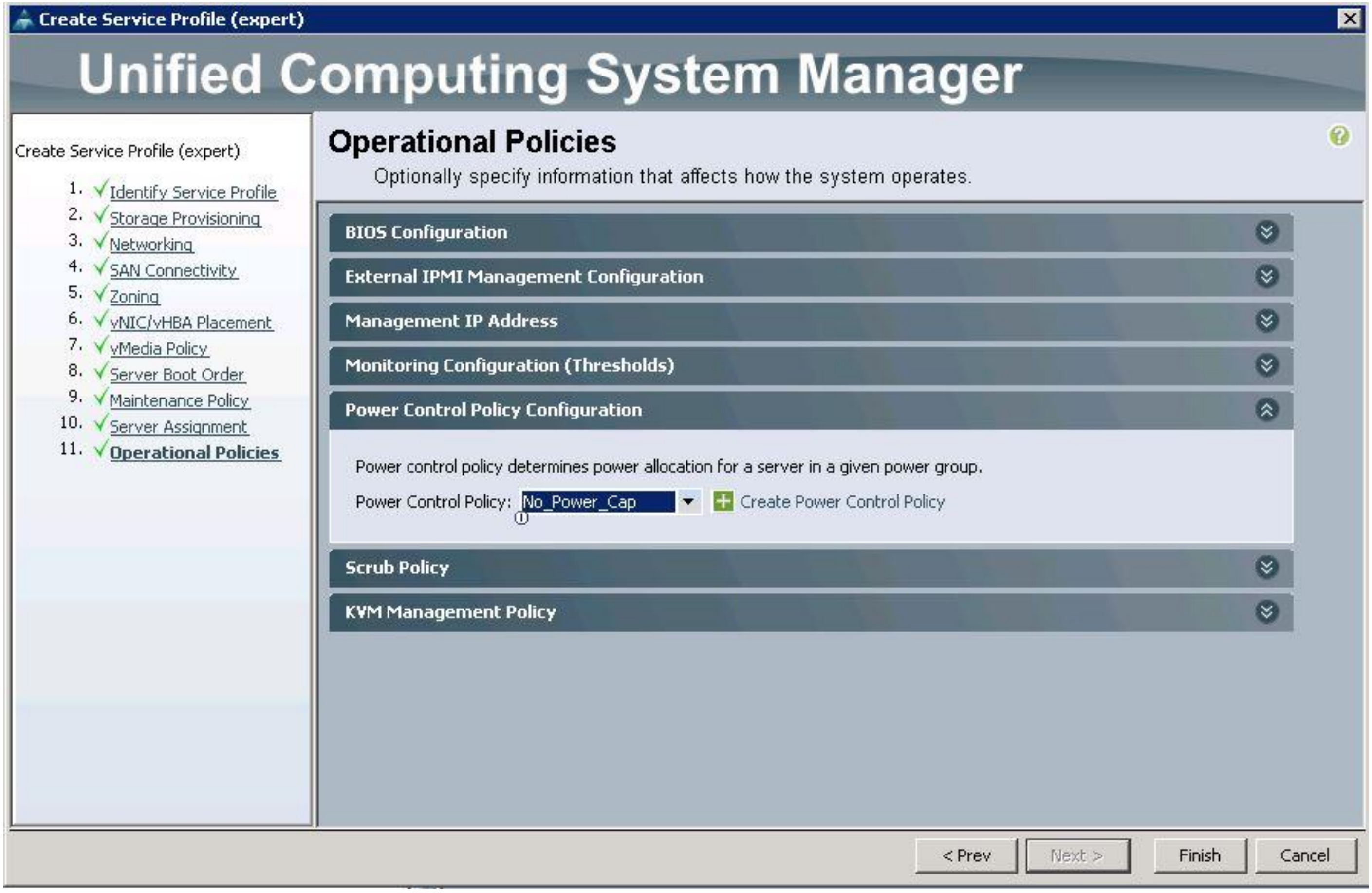
Select the power state to be applied when this profile is associated w

● Up ○ Down

● Available Servers ○ All Servers

| Select | Chassis ID | Slot | Rack ID | Procs | M |
|---|---|---|---|---|---|
| ○ | 1 | 7 | | 2 | 262144 |
| ● | 1 | 8 | | 2 | 262144 |

Under Operational Policies, choose the Power Control Policy as "No_Power_Cap" and click Finish.


Create Service Profile (expert)
Unified Computing System Manager
Create Service Profile (expert)
1. Identify Service Profile
2. Storage Provisioning
3. Networking
4. SAN Connectivity
5. Zoning
6. vNIC/vHBA Placement
7. vMedia Policy
8. Server Boot Order
9. Maintenance Policy
10. Server Assignment
11. Operational Policies
Operational Policies
Optionally specify information that affects how the system operates.
BIOS Configuration
External IPMI Management Configuration
Management IP Address
Monitoring Configuration (Thresholds)
Power Control Policy Configuration
Scrub Policy
KVM Management Policy
Create Power Control Policy
< Prev
Next >
Finish
Cancel

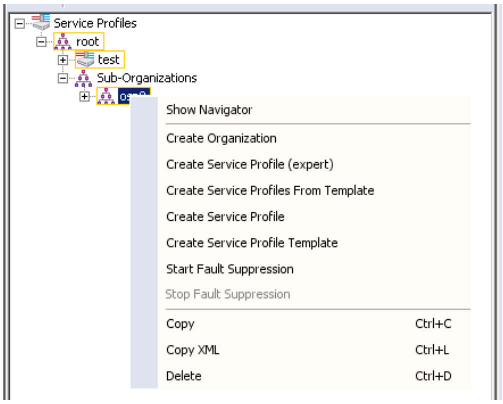
Power control policy determines power allocation for a server in a given power group.

Power Control Policy: No_Power_Cap

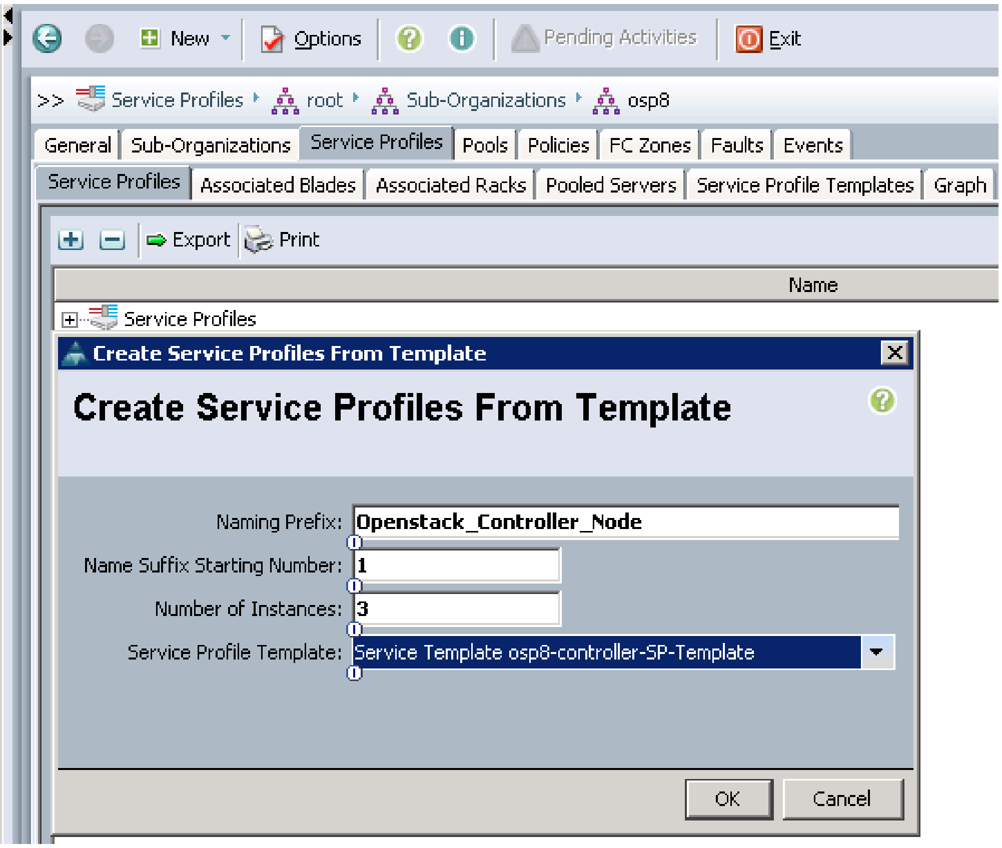
## Create Service Profiles for Controller Nodes

To create Service profiles for Controller nodes, complete the following steps:

Under Servers → Service Profile Templates → root → Sub-Organizations -> select the Controller Service profile template and click Create Service Profiles from Templates.



Specify the Service profile name and the number of instances as 3 for the Controller nodes.


New
Options
Pending Activities
Exit
>> Service Profiles ▸ root ▸ Sub-Organizations ▸ osp8
General | Sub-Organizations | Service Profiles | Pools | Policies | FC Zones | Faults | Events
Service Profiles | Associated Blades | Associated Racks | Pooled Servers | Service Profile Templates | Graph
Export
Print
Name
Service Profiles
Create Service Profiles From Template
OK
Cancel

Create Service Profiles From Template

Naming Prefix: Openstack_Controller_Node

Name Suffix Starting Number: 1

Number of Instances: 3

Service Profile Template: Service Template osp8-controller-SP-Template

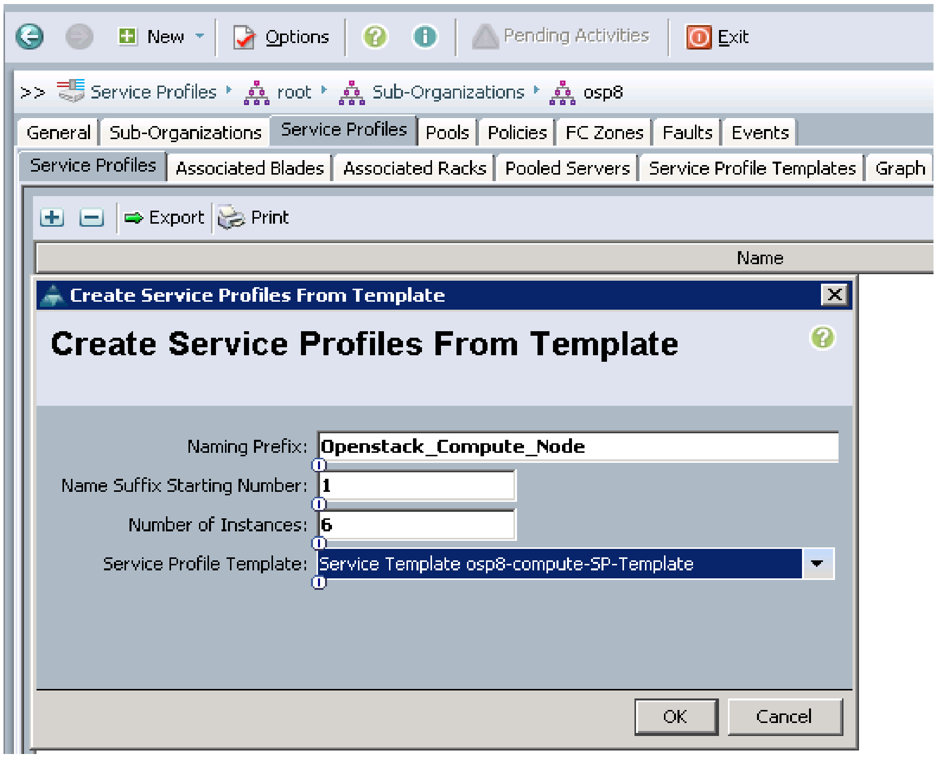
Make sure the Service profiles for the Controller nodes have been created. Under Servers → Service profiles → root -> Sub-Organizations (osp8).

## Create Service Profiles for Compute Nodes

To create Service profiles for Compute nodes, complete the following steps:

Under Servers → Service Profile Templates → root → Sub-Organizations and select the Compute Service profile template and click Create Service Profiles from Templates, similar to Controller above.

Specify the profile name and set the number of instances to 6 for compute nodes.



Create Service Profiles From Template

Create Service Profiles From Template

Naming Prefix: Openstack_Compute_Node

Name Suffix Starting Number: 1

Number of Instances: 6

Service Profile Template: Service Template osp8-compute-SP-Template

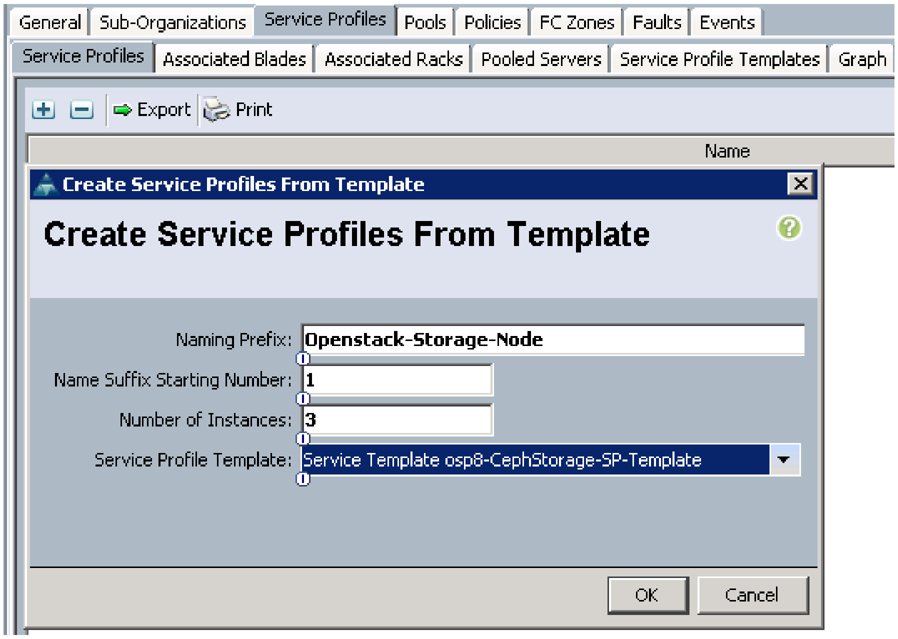
Make sure the Service profiles for the Compute nodes have been created under the Sub Orgs.

## Create Service Profiles for Ceph Storage Nodes

To create Service profiles for Ceph Storage nodes, complete the following steps:

Under Servers → Service Profile Templates → root → Sub-Organizations and select the Ceph Storage Service profile template and click Create Service Profiles from Templates.

Specify the Service profile name and set the number of instances to 3 for the Ceph Storage nodes.



Create Service Profiles From Template

Create Service Profiles From Template

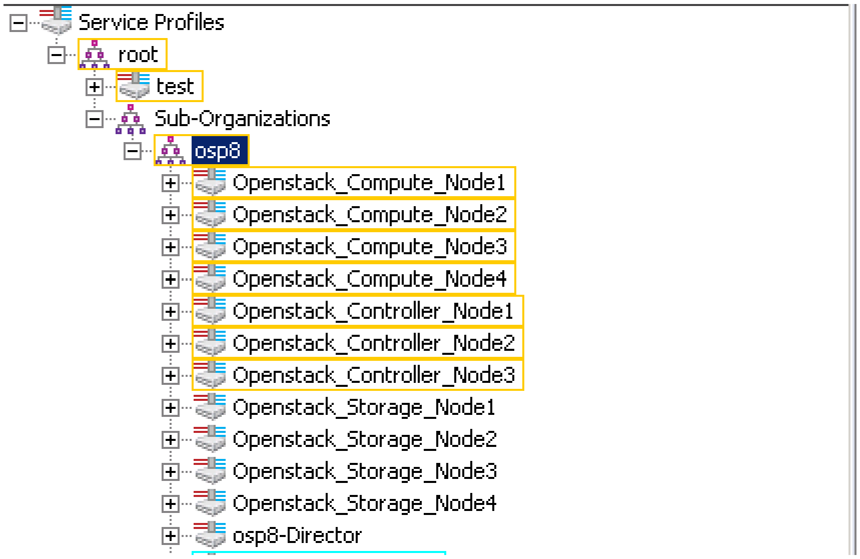
Naming Prefix: Openstack-Storage-Node

Name Suffix Starting Number: 1

Number of Instances: 3

Service Profile Template: Service Template osp8-CephStorage-SP-Template

Make sure the Service profiles for the Ceph Storage nodes have been created under Sub Orgs.



Verify the Service profile association with the respective UCS Servers.

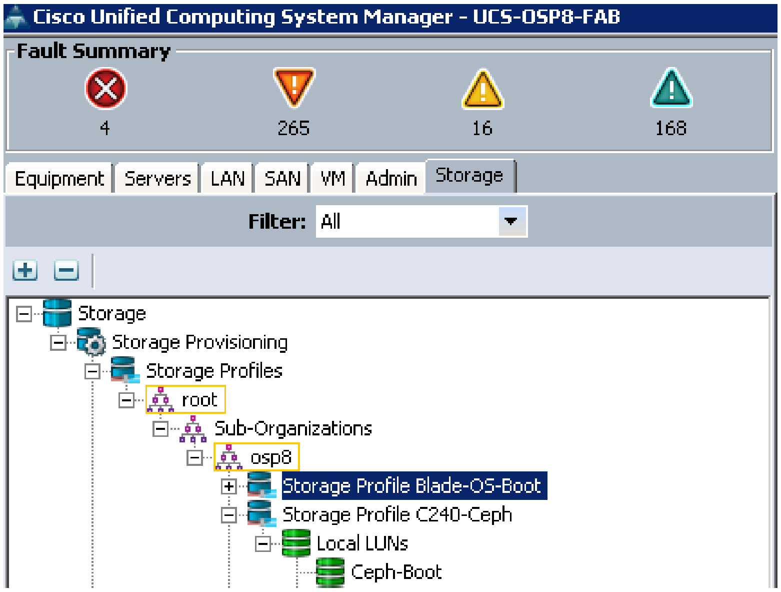
## Create LUNs for the Ceph OSD and Journal Disks

After a successful CephStorage Server association, create the remaining LUNs for the Ceph OSD disks and Journal disks.

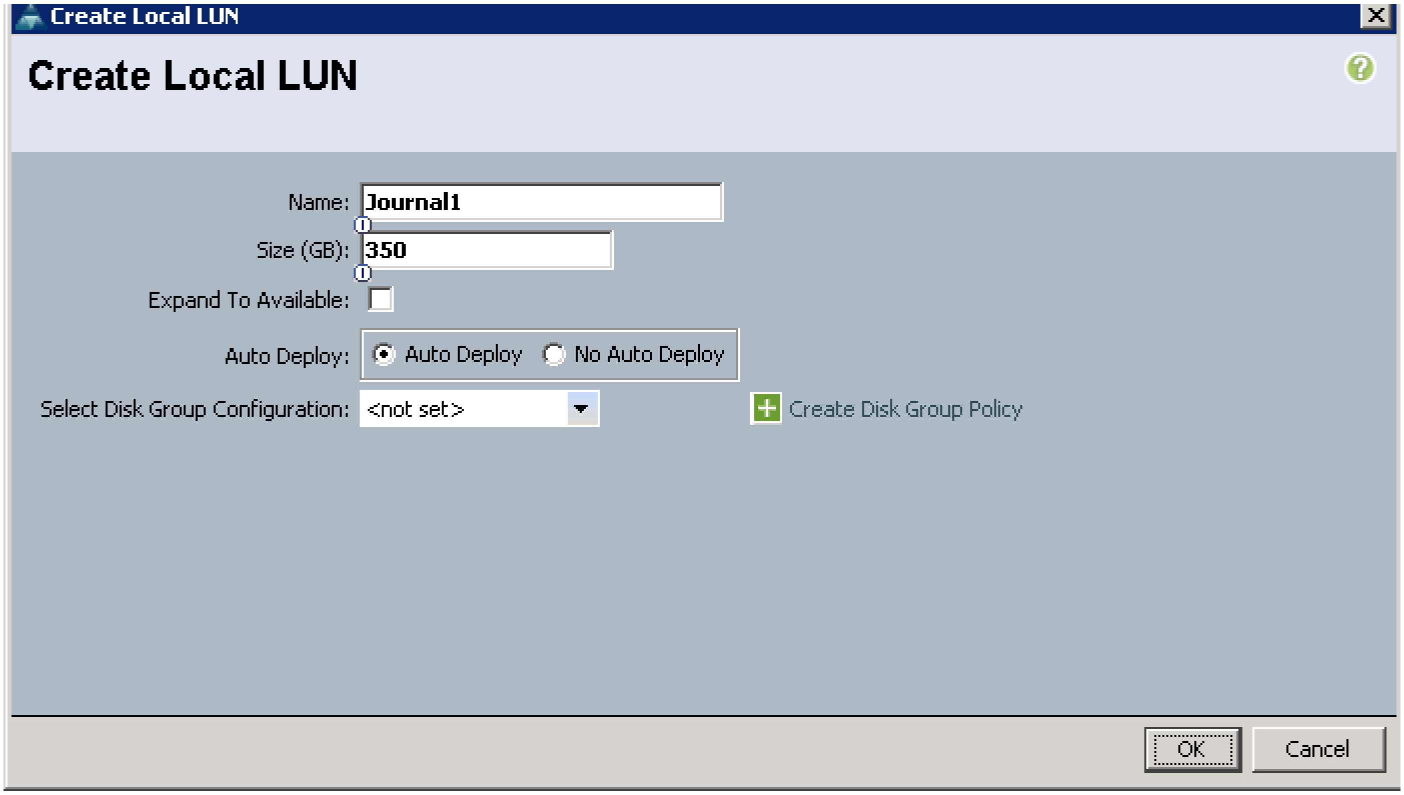
### Create the Ceph Journal LUN

To create the Ceph Journal LUNs, complete the following steps:

Under Storage →Storage Provisioning → root → Sub-Organizations -> select the previously created Ceph Sotrage profile C240-Ceph → click Local LUNs → click Create Local LUNs.



Specify the name as Journal1 and set the size in GB to 350 for the 400GB SSD disks and click Create Disk Group Policy.



Create Local LUN

Name: Journal1

Size (GB): 350

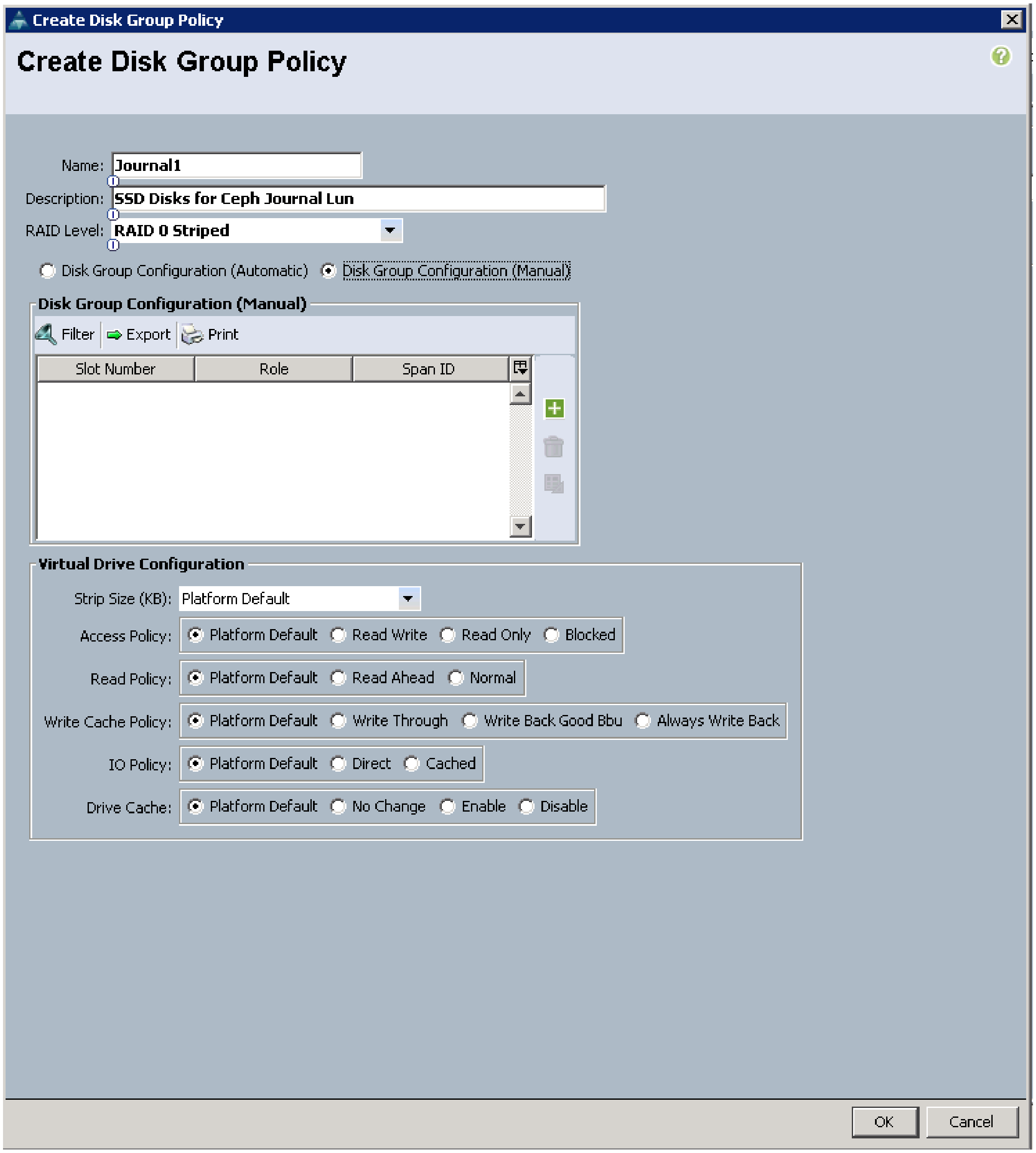
Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: <not set>

Create Disk Group Policy

Specify the Disk group policy name and choose the RAID level as RAID o and select Disk Group Configuration (Manual).



## Create Disk Group Policy

Name: Journal1

Description: SSD Disks for Ceph Journal Lun

RAID Level: RAID 0 Striped

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

**Disk Group Configuration (Manual)**

| Slot Number | Role | Span ID |
|---|---|---|

**Virtual Drive Configuration**

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

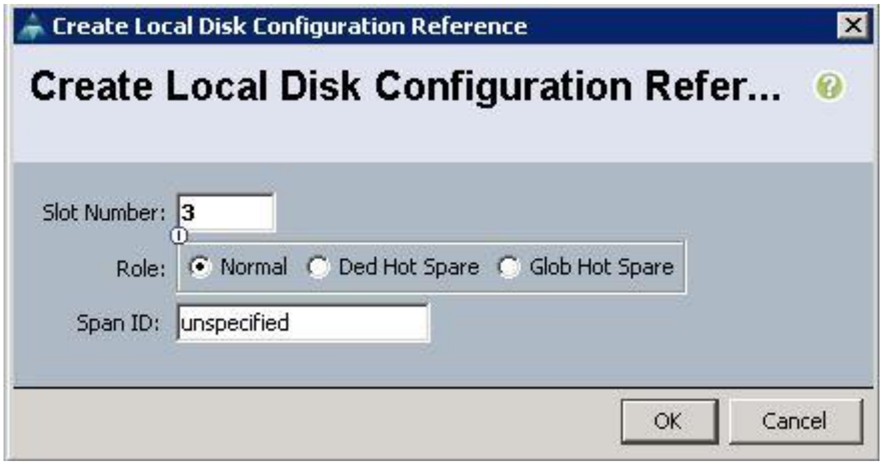
Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

Specify the Slot ID as 3, which is the physical disk slot number for 400GB SSDs for the Journal LUN1 and click OK.



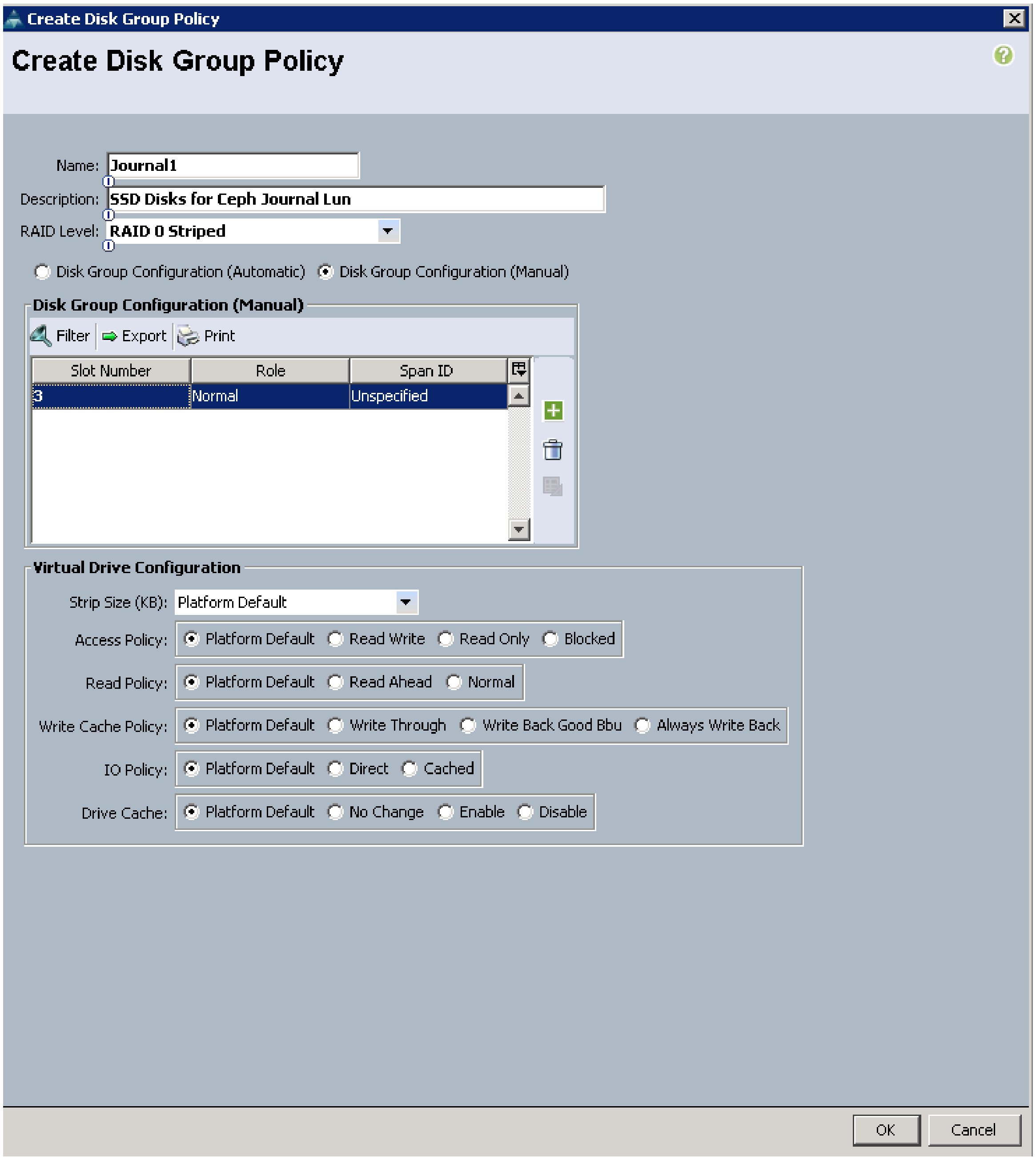
Create Local Disk Configuration Reference
Create Local Disk Configuration Refer...

Slot Number: 3

Role: ● Normal ○ Ded Hot Spare ○ Glob Hot Spare

Span ID: unspecified

Click OK to confirm the Disk group policy creation.



Create Disk Group Policy

## Create Disk Group Policy

Name: Journal1

Description: SSD Disks for Ceph Journal Lun

RAID Level: RAID 0 Striped

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

**Disk Group Configuration (Manual)**

| Slot Number | Role | Span ID |
|---|---|---|
| 3 | Normal | Unspecified |

**Virtual Drive Configuration**

Strip Size (KB): Platform Default

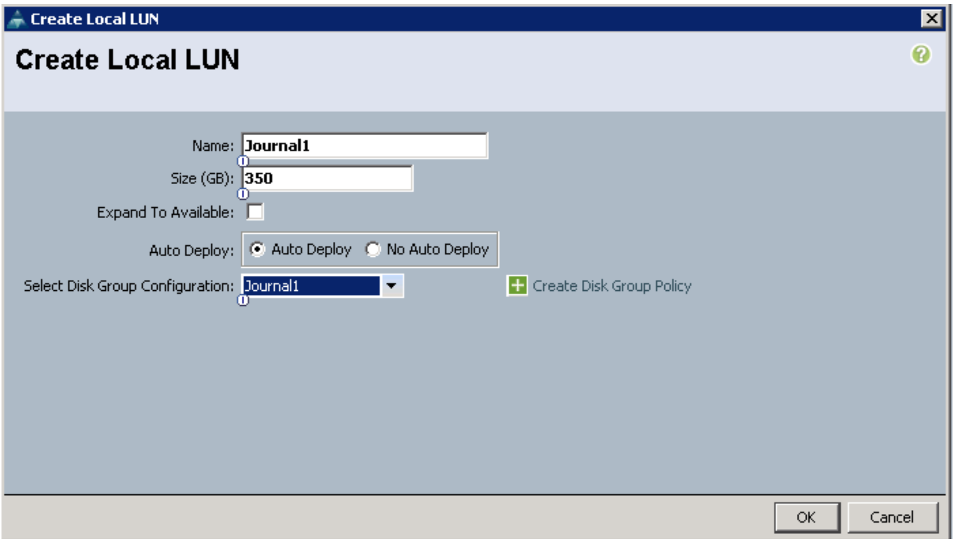
Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable



Create Local LUN

Name: Journal1

Size (GB): 350

Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: Journal1

Create Disk Group Policy

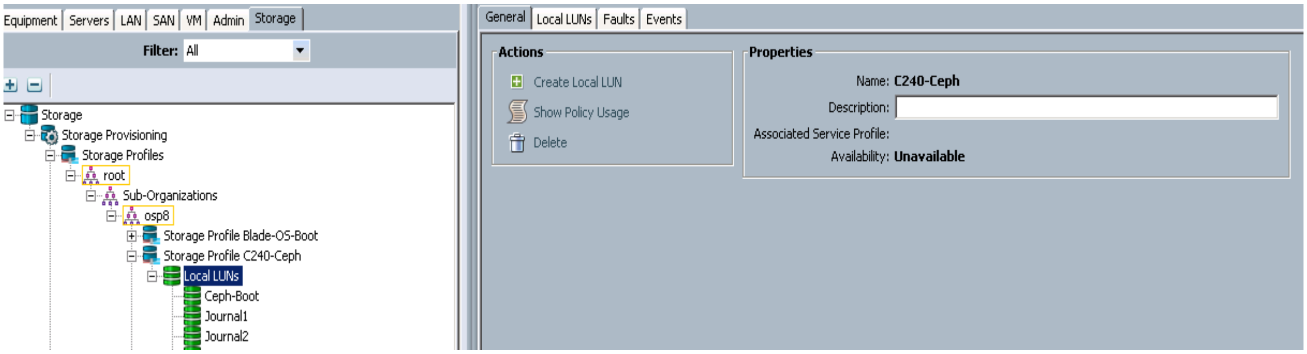
From the drop-down list, choose the Disk group policy for the Journal LUN as Journal1.

Similar to the above, create the Local LUN as Journal2 with Disk group policy as Journal2 using 400GB SSD on Disk Slot4.

### Create the Ceph OSD LUN

To create the Ceph OSD LUN, complete the following steps:

Under Storage → Provisioning → root → Sub-Organizations and select the previously created Ceph Storage profile C240-Ceph → click Local LUNs → click Create Local LUN.



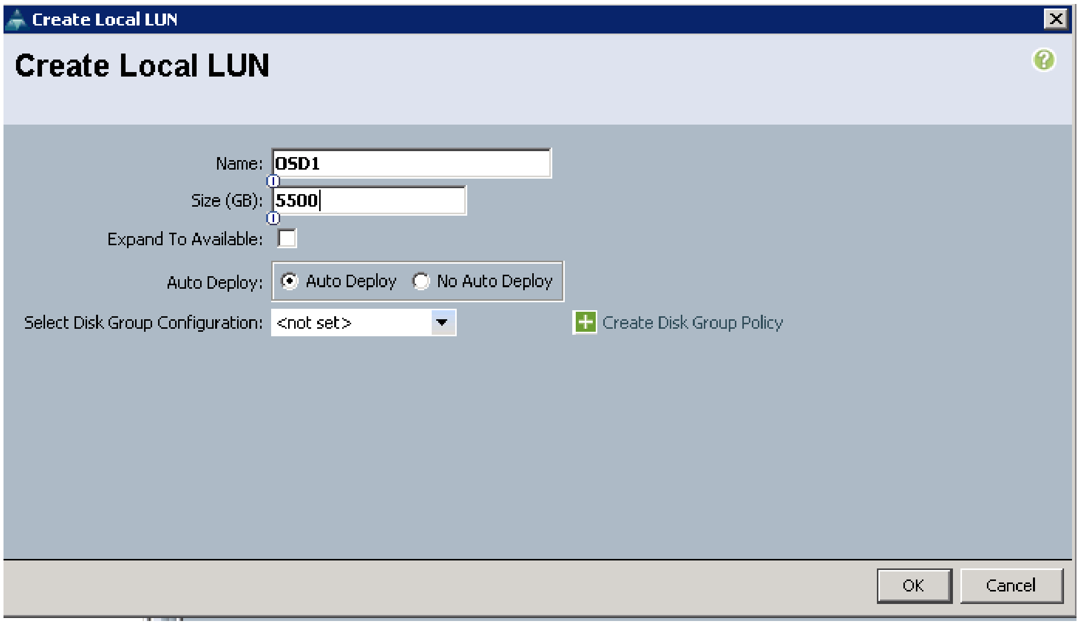
Name: C240-Ceph

Description:

Associated Service Profile:

Availability: Unavailable

Specify the name as OSD1 and the size in GB as 5500 for the 6TB SAS disks and click Create Disk Group Policy.



Create Local LUN

Name: OSD1

Size (GB): 5500

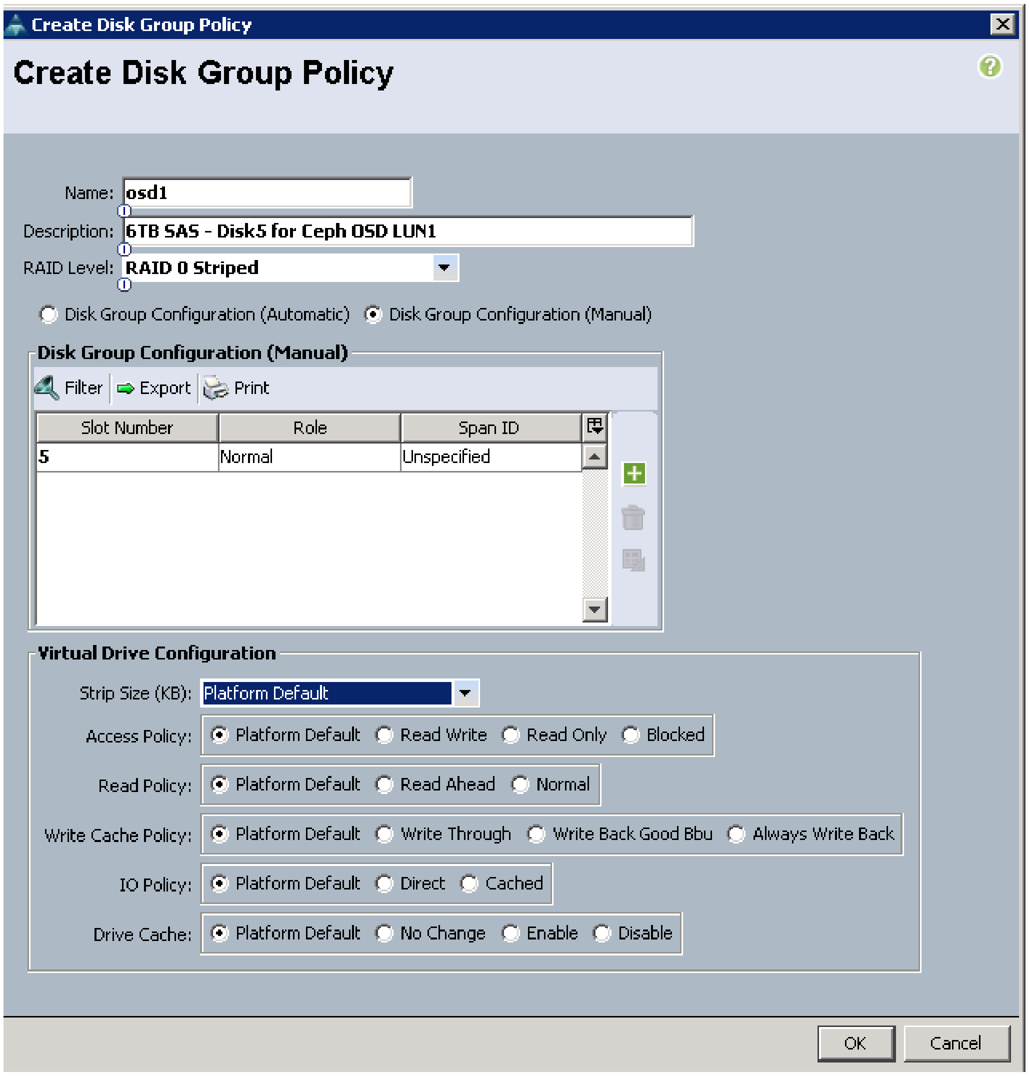
Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: <not set>

Specify Disk group policy name and Choose RAID level as RAID 0 and select Disk Group Configuration(Manual)

Click "+" and Specify Slot ID as 5, which is physical disk slot number for 6TB SAS disks for Ceph OSD LUN1 and click OK.



Create Disk Group Policy

Name: osd1

Description: 6TB SAS - Disk5 for Ceph OSD LUN1

RAID Level: RAID 0 Striped

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

Disk Group Configuration (Manual)

| Slot Number | Role | Span ID |
|---|---|---|
| 5 | Normal | Unspecified |

Virtual Drive Configuration

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

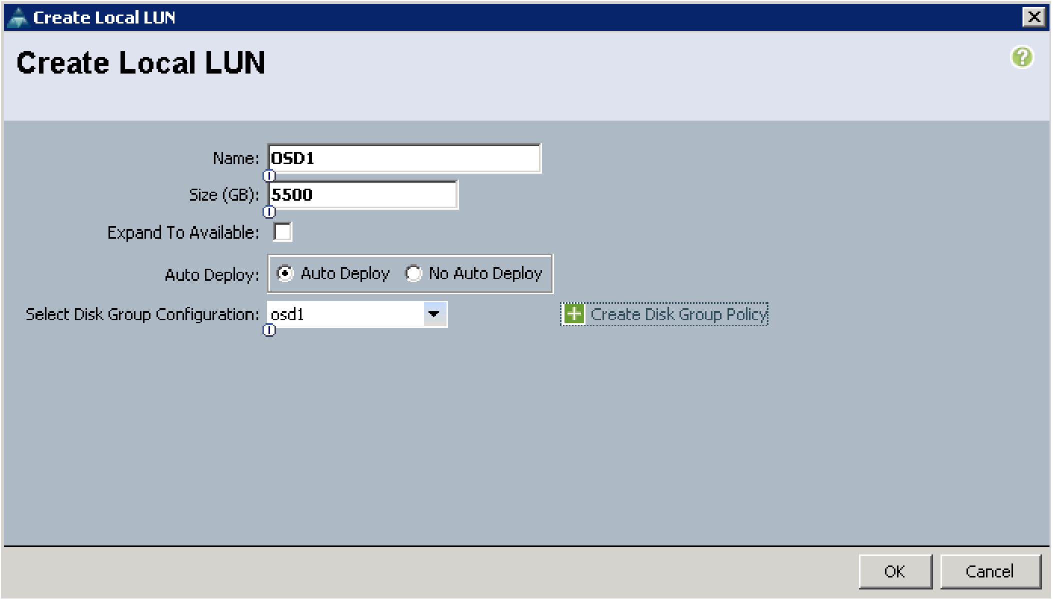
Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

From the drop-down list, choose the Disk group policy for OSD1 as OSD1.



Create Local LUN

Name: OSD1

Size (GB): 5500

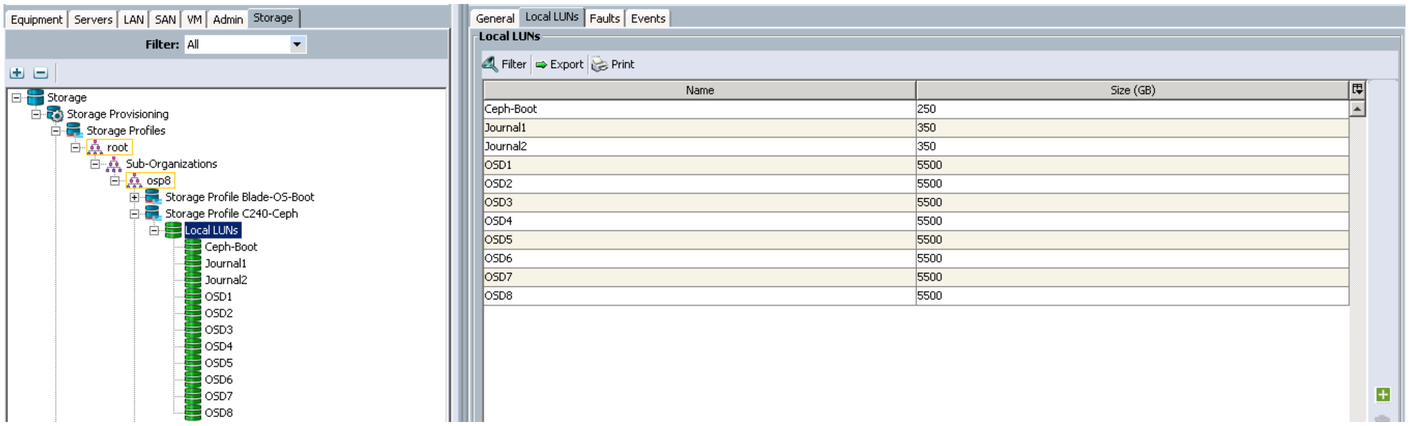
Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: osd1

Create the remaining OSD LUNs 3, 4, 5, 6, 7, 8 with the Disk group policy using 6TB SAS disks 6, 7, 8, 9, 10, 11, and 12.

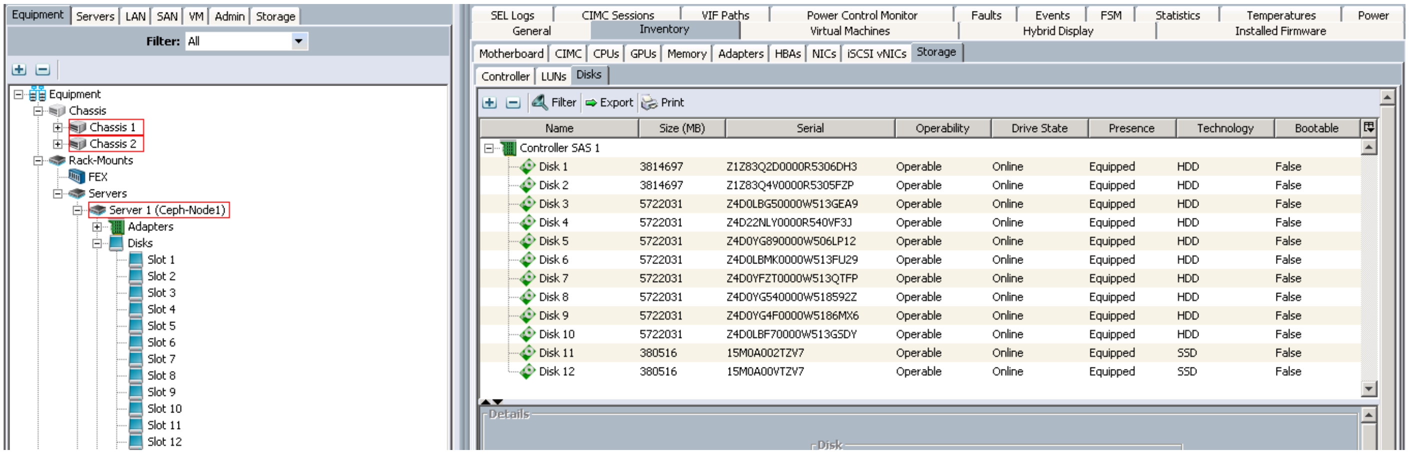
Make sure the LUNs for Journals and OSDs are created as shown below.



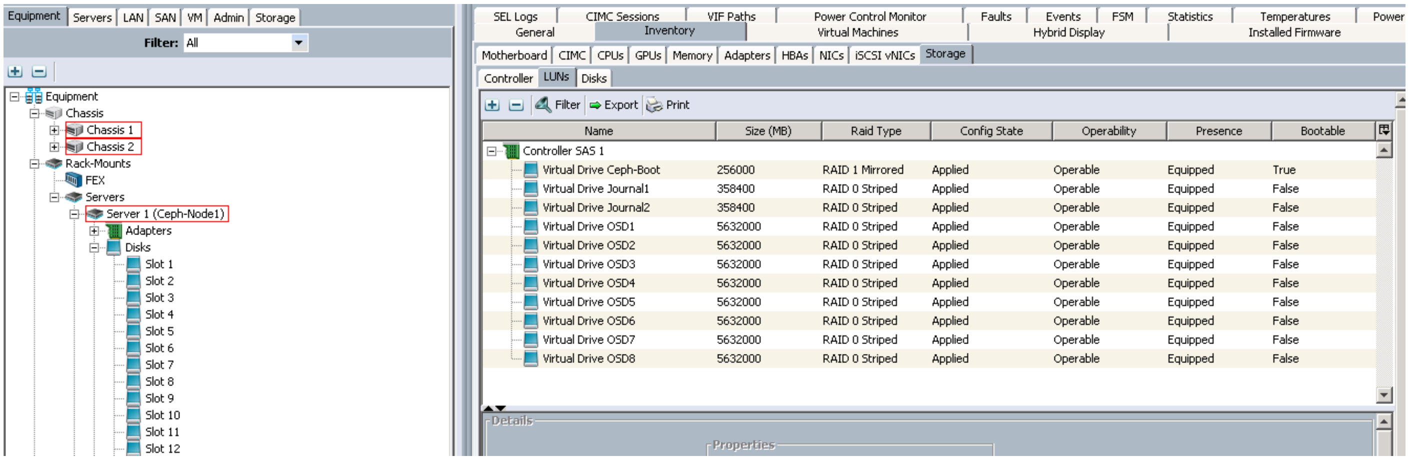
| Name | Size (GB) |
|---|---|
| Ceph-Boot | 250 |
| Journal1 | 350 |
| Journal2 | 350 |
| OSD1 | 5500 |
| OSD2 | 5500 |
| OSD3 | 5500 |
| OSD4 | 5500 |
| OSD5 | 5500 |
| OSD6 | 5500 |
| OSD7 | 5500 |
| OSD8 | 5500 |

Make sure all the Ceph Storage Servers have the identical LUN ID and Device ID for all the LUNs (OS-boot, Journal and OSD) as shown in the table below:

| Physical Disk Slot | Disk Type | Disk Size | RAID Level | LUN Size | LUN ID | Device ID |
|---|---|---|---|---|---|---|
| **Disk 1** | SAS | 300 GB | RAID 1 | 250 GB | 1000 | 0 |
| **Disk 2** | SAS | 300 GB | | | | |
| **Disk 3** | SSD | 400 GB | RAID 0 | 350 GB | 1001 | 1 |
| **Disk 4** | SSD | 400 GB | RAID 0 | 350 GB | 1002 | 2 |
| **Disk 5** | SAS | 6 TB | RAID 0 | 5500 GB | 1003 | 3 |
| **Disk 6** | SAS | 6 TB | RAID 0 | 5500 GB | 1004 | 4 |
| **Disk 7** | SAS | 6 TB | RAID 0 | 5500 GB | 1005 | 5 |
| **Disk 8** | SAS | 6 TB | RAID 0 | 5500 GB | 1006 | 6 |
| **Disk 9** | SAS | 6 TB | RAID 0 | 5500 GB | 1007 | 7 |
| **Disk 10** | SAS | 6 TB | RAID 0 | 5500 GB | 1008 | 8 |
| **Disk 11** | SAS | 6 TB | RAID 0 | 5500 GB | 1009 | 9 |
| **Disk 12** | SAS | 6 TB | RAID 0 | 5500 GB | 1010 | 10 |



| Name | Size (MB) | Serial | Operability | Drive State | Presence | Technology | Bootable |
|---|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | | |
| Disk 1 | 3814697 | Z1Z83Q2D0000R5306DH3 | Operable | Online | Equipped | HDD | False |
| Disk 2 | 3814697 | Z1Z83Q4V0000R5305FZP | Operable | Online | Equipped | HDD | False |
| Disk 3 | 5722031 | Z4D0LBG50000W513GEA9 | Operable | Online | Equipped | HDD | False |
| Disk 4 | 5722031 | Z4D22NLY0000R540VF3J | Operable | Online | Equipped | HDD | False |
| Disk 5 | 5722031 | Z4D0YG890000W506LP12 | Operable | Online | Equipped | HDD | False |
| Disk 6 | 5722031 | Z4D0LBMK0000W513FU29 | Operable | Online | Equipped | HDD | False |
| Disk 7 | 5722031 | Z4D0YFZT0000W513QTFP | Operable | Online | Equipped | HDD | False |
| Disk 8 | 5722031 | Z4D0YG540000W518592Z | Operable | Online | Equipped | HDD | False |
| Disk 9 | 5722031 | Z4D0YG4F0000W5186MX6 | Operable | Online | Equipped | HDD | False |
| Disk 10 | 5722031 | Z4D0LBF70000W513GSDY | Operable | Online | Equipped | HDD | False |
| Disk 11 | 380516 | 15M0A002TZV7 | Operable | Online | Equipped | SSD | False |
| Disk 12 | 380516 | 15M0A00VTZV7 | Operable | Online | Equipped | SSD | False |



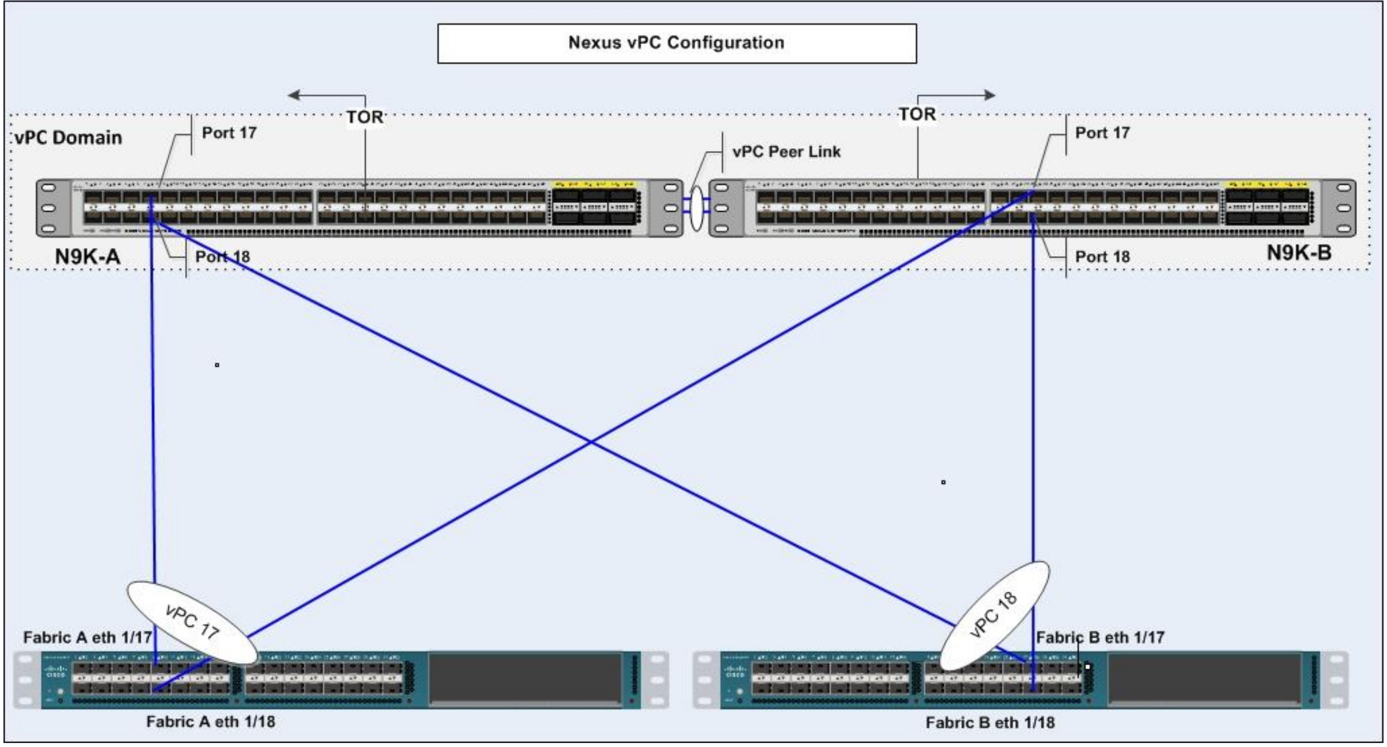
| Name | Size (MB) | Raid Type | Config State | Operability | Presence | Bootable |
|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | |
| Virtual Drive Ceph-Boot | 256000 | RAID 1 Mirrored | Applied | Operable | Equipped | True |
| Virtual Drive Journal1 | 358400 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive Journal2 | 358400 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD1 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD2 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD3 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD4 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD5 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD6 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD7 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD8 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |

All the disks have to be in Equipped state; LUNs have to be in Applied and Operable state as shown above.

## Create Port Channels for Cisco UCS Fabrics

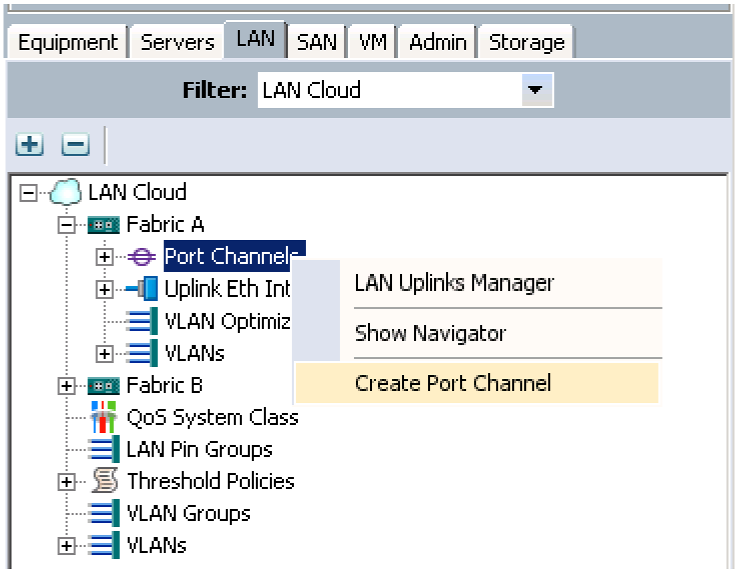
Figure 10 illustrates the virtual Port Channel configuration.

**Figure 10** Nexus vPC Configuration

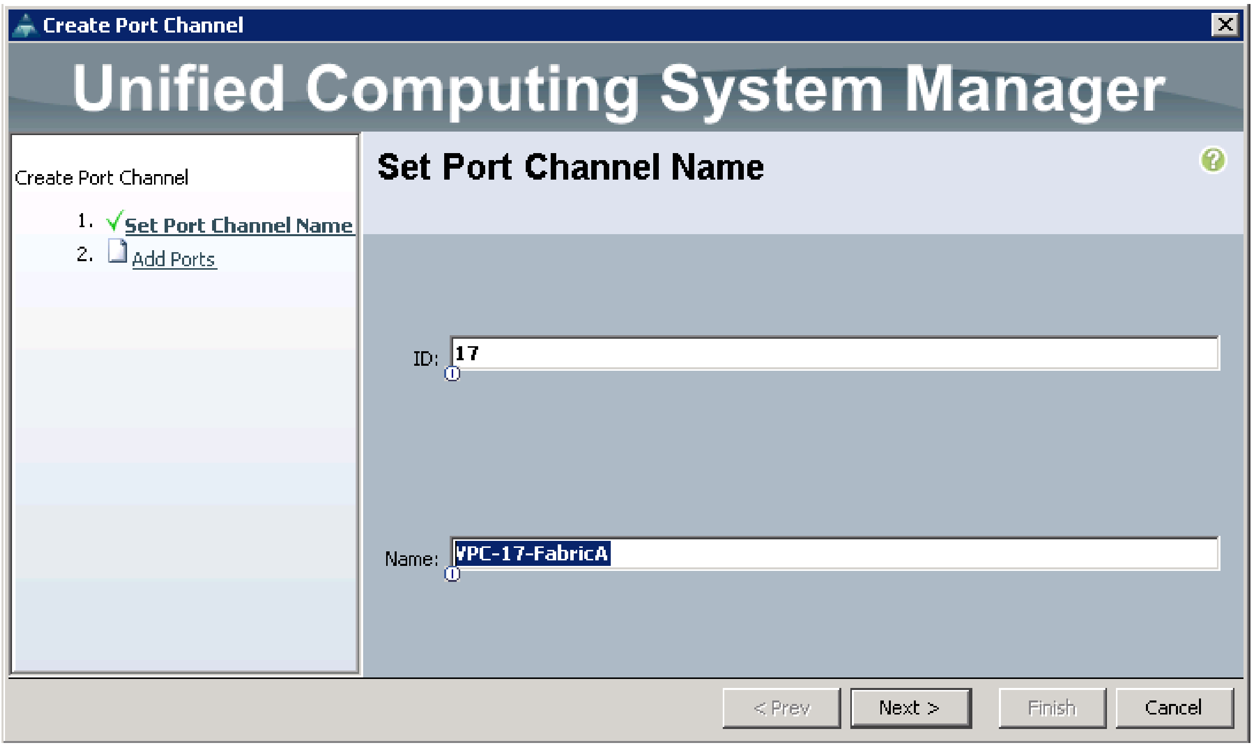


To create Port Channels from the UCS Manager GUI, complete the following steps:

Under LAN → Cloud → Fabric A → Port Channels → right-click and select Create Port Channel.



Specify the ID and name for the port channel and click Next.



Create Port Channel

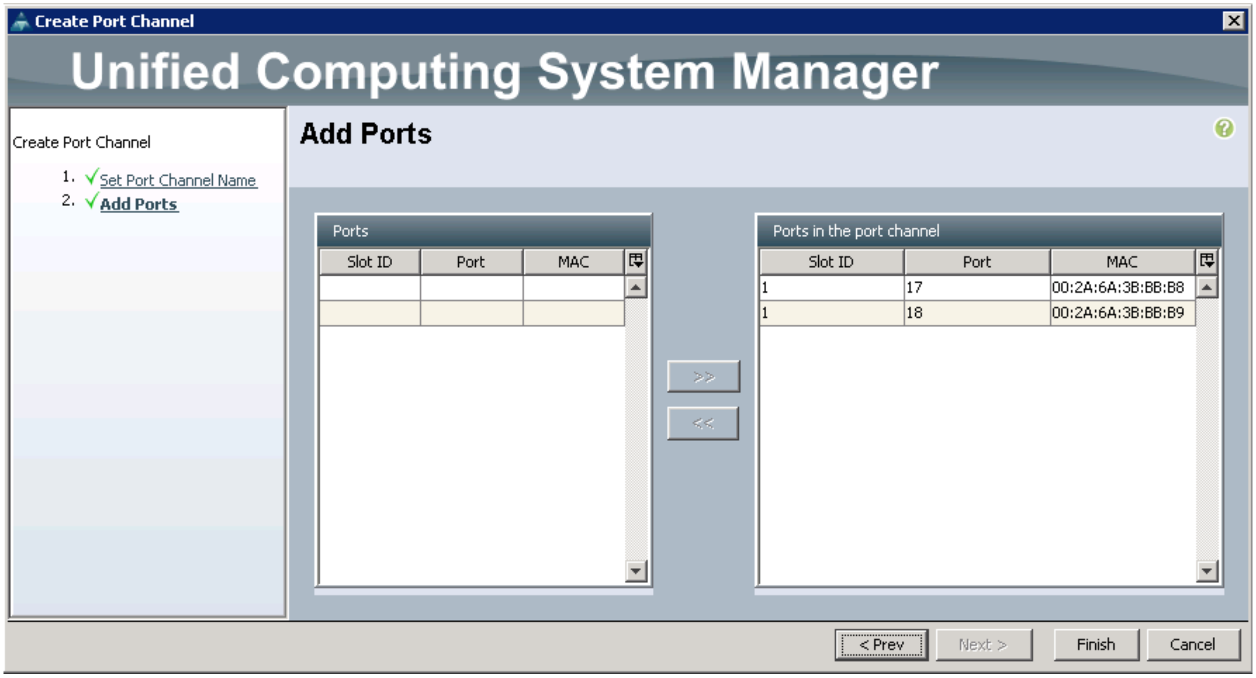
1. ✓ Set Port Channel Name
2. Add Ports

Set Port Channel Name

ID: 17

Name: VPC-17-FabricA

Select the ports 17 and 18 from left pane and move to the right pane into Ports in the Port Channel and click Finish.



Create Port Channel

1. ✓ Set Port Channel Name
2. ✓ Add Ports

Add Ports

Ports

| Slot ID | Port | MAC |
|---|---|---|

Ports in the port channel

| Slot ID | Port | MAC |
|---|---|---|
| 1 | 17 | 00:2A:6A:3B:BB:B8 |
| 1 | 18 | 00:2A:6A:3B:BB:B9 |

Repeat the steps shown above on Fabric B with Port-Channel as18.

## Cisco Nexus Configuration

### Configure the Cisco Nexus 9372 PX Switch A

To configure the Cisco Nexus 9372 PX Switch A, complete the following step:

Connect the console port to the Nexus 9372 PX switch designated for Fabric A:

```
---- Basic System Configuration Dialog VDC: 1 ----
This setup utility will guide you through the basic configuration of the system.
Setup configures only enough connectivity for management of the system.
*Note: setup is mainly used for configuring the system initially, when no
configuration is present. So setup always assumes system defaults and not the
current system configuration values.
Press Enter at anytime to skip a dialog. Use ctrl-c at anytime to skip the
remaining dialogs.
Would you like to enter the basic configuration dialog (yes/no): yes
Do you want to enforce secure password standard (yes/no) [y]:
Create another login account (yes/no) [n]:
Configure read-only SNMP community string (yes/no) [n]:
Configure read-write SNMP community string (yes/no) [n]:
Enter the switch name : OSP8-N9K-FAB-A
Continue with Out-of-band (mgmt0) management configuration? (yes/no) [y]:
             Mgmt0 IPv4 address : 10.23.10.3
 Mgmt0 IPv4 netmask : 255.255.255.0
            Configure the default gateway? (yes/no) [y]:
            IPv4 address of the default gateway : 10.23.10.1
            Configure advanced IP options? (yes/no) [n]:
            Enable the telnet service? (yes/no) [n]:
            Enable the ssh service? (yes/no) [y]:
            Type of ssh key you would like to generate (dsa/rsa) [rsa]:
            Number of rsa key bits <1024-2048> [2048]:
            Configure the ntp server? (yes/no) [n]: y
            NTP server IPv4 address : <<ntp_server_ip>>

            Configure CoPP system profile (strict/moderate/lenient/dense/skip)
[strict]:
The following configuration will be applied:
   password strength-check
   switchname OSP8-N9k-FAB-A
   vrf context management
   ip route 0.0.0.0/0 10.23.10.1
   exit
   no feature telnet
   ssh key rsa 2048 force
    feature ssh
    ntp server <<var_global_ntp_server_ip>>
    copp profile strict
     interface mgmt0
     ip address 10.23.10.3 255.255.255.0
     no shutdown
       Would you like to edit the configuration? (yes/no) [n]: Enter
       Use this configuration and save it? (yes/no) [y]: Enter
       [########################################] 100%
       Copy complete.
```

## Configure the Cisco Nexus 9372 PX Switch B

To configure the Cisco Nexus 9372 PX Switch B, complete the following step:

Connect the console port to the Nexus 9372 PX switch designated for Fabric B:

```
---- Basic System Configuration Dialog VDC: 1 ----
This setup utility will guide you through the basic configuration of the system.
Setup configures only enough connectivity for management of the system.
```

```
*Note: setup is mainly used for configuring the system initially, when no
configuration is present. So setup always assumes system defaults and not the
current system configuration values.
Press Enter at anytime to skip a dialog. Use ctrl-c at anytime to skip the
remaining dialogs.
Would you like to enter the basic configuration dialog (yes/no): yes
Do you want to enforce secure password standard (yes/no) [y]:
Create another login account (yes/no) [n]:
Configure read-only SNMP community string (yes/no) [n]:
Configure read-write SNMP community string (yes/no) [n]:
Enter the switch name : OSP8-N9k-FAB-B
Continue with Out-of-band (mgmt0) management configuration? (yes/no) [y]:
             Mgmt0 IPv4 address : 10.23.10.4
 Mgmt0 IPv4 netmask : 255.255.255.0
            Configure the default gateway? (yes/no) [y]:
            IPv4 address of the default gateway : 10.23.10.1
            Configure advanced IP options? (yes/no) [n]:
            Enable the telnet service? (yes/no) [n]:
            Enable the ssh service? (yes/no) [y]:
            Type of ssh key you would like to generate (dsa/rsa) [rsa]:
            Number of rsa key bits <1024-2048> [2048]:
            Configure the ntp server? (yes/no) [n]: y
            NTP server IPv4 address : <<ntp_server_ip>>

            Configure CoPP system profile (strict/moderate/lenient/dense/skip)
[strict]:
The following configuration will be applied:
   password strength-check
   switchname OSP8-N9k-FAB-B
   vrf context management
   ip route 0.0.0.0/0 10.23.10.1
   exit
   no feature telnet
   ssh key rsa 2048 force
    feature ssh
    ntp server <<var_global_ntp_server_ip>>
    copp profile strict
     interface mgmt0
     ip address 10.23.10.4 255.255.255.0
     no shutdown
       Would you like to edit the configuration? (yes/no) [n]: Enter
       Use this configuration and save it? (yes/no) [y]: Enter
       [########################################] 100%
       Copy complete.
```

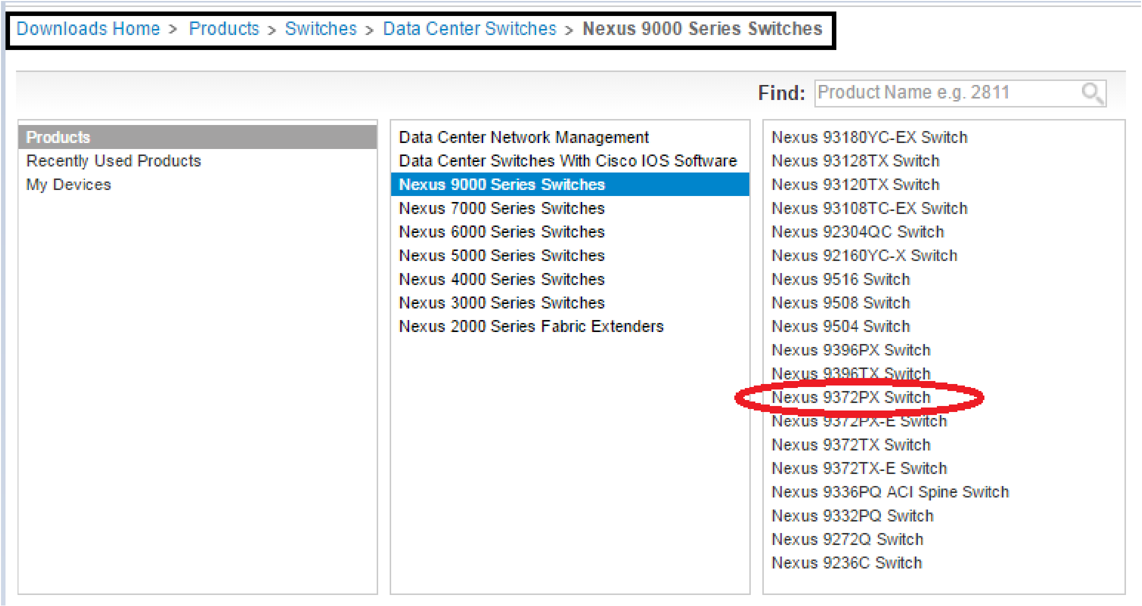
## Check Nexus OS compatibility

Login into each of the nexus switches and check for the NXOS version as below.

```
OSP8-N9K-FAB-A# show version
Cisco Nexus Operating System (NX-OS) Software
.........
.........
Software
  BIOS: version 07.34
  NXOS: version 7.0(3)I1(3)
  BIOS compile time:  08/11/2015
  NXOS image file is: bootflash:///n9000-dk9.7.0.3.I1.3.bin
  NXOS compile time:  8/21/2015 3:00:00 [08/21/2015 10:27:18]
```

Make sure that the software version of Nexus OS is 7.0 (3) I1(3) as this is the version of Nexus OS that was validated. Either Upgrade or downgrade the switch to this version as below:

- Go to https://software.cisco.com/download/navigator.html
- On the products tab, select switches and then Data Center Switches and then Nexus 9000 series switch
- Select the model say 9372 as below



Downloads Home > Products > Switches > Data Center Switches > Nexus 9000 Series Switches

Find: Product Name e.g. 2811

Products
Recently Used Products
My Devices

Data Center Network Management
Data Center Switches With Cisco IOS Software
Nexus 9000 Series Switches
Nexus 7000 Series Switches
Nexus 6000 Series Switches
Nexus 5000 Series Switches
Nexus 4000 Series Switches
Nexus 3000 Series Switches
Nexus 2000 Series Fabric Extenders

Nexus 93180YC-EX Switch
Nexus 93128TX Switch
Nexus 93120TX Switch
Nexus 93108TC-EX Switch
Nexus 92304QC Switch
Nexus 92160YC-X Switch
Nexus 9516 Switch
Nexus 9508 Switch
Nexus 9504 Switch
Nexus 9396PX Switch
Nexus 9396TX Switch
Nexus 9372PX Switch
Nexus 9372PX-E Switch
Nexus 9372TX Switch
Nexus 9372TX-E Switch
Nexus 9336PQ ACI Spine Switch
Nexus 9332PQ Switch
Nexus 9272Q Switch
Nexus 9236C Switch

-
- Select NX-OS system software, expand All Releases and download the version 7.0(3)I1.3
- Upgrade or downgrade the software by following instructions from Nexus 9000 Guide. The Upgrade/downgrade can also be referred from here.

## Enable Features on the Switch

To enable the features on the switch, enter the following:

```
OSP8-N9K-FAB-A# config terminal
OSP8-N9k-FAB-A(config)# feature udld
OSP8-N9K-FAB-A(config)# feature interface-vlan
OSP8-N9K-FAB-A(config)# feature hsrp
OSP8-N9K-FAB-A(config)# feature lacp
OSP8-N9K-FAB-A(config)# feature vpc
OSP8-N9K-FAB-A(config)# exit
```

Repeat the same steps on Nexus 9372 Switch B.

## Enable Jumbo MTU

To enable the Jumbo MTU, enter the following:

```
OSP8-N9K-FAB-A# config terminal
OSP8-N9K-FAB-A(config)# system jumbomtu 9216
```

```
OSP8-N9K-FAB-A(config)# exit
```

Repeat the same steps on Nexus 9372 Switch B.

## Create VLANs

To create VLANs, enter the following:

```
OSP8-N9K-FAB-A# config terminal
OSP8-N9K-FAB-A(config)# vlan 10
OSP8-N9K-FAB-A(config-vlan)# name Management
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 100
OSP8-N9K-FAB-A(config-vlan)# name Internal-API
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 110
OSP8-N9K-FAB-A(config-vlan)# name PXE-Network
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 120
OSP8-N9K-FAB-A(config-vlan)# name Storage-Public-Network
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 150
OSP8-N9K-FAB-A(config-vlan)# name Storage-Mgmt-Network
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 160
OSP8-N9K-FAB-A(config-vlan)# name Tenant-Floating-IP-Network
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)# vlan 215
OSP8-N9K-FAB-A(config-vlan)# name External-Network
OSP8-N9K-FAB-A(config-vlan)# no shut
OSP8-N9K-FAB-A(config-vlan)# exit
OSP8-N9K-FAB-A(config)#
```

Repeat the same steps on Nexus 9372 Switch B.

## Configure the Interface VLAN (SVI) on the Cisco Nexus 9K Switch A

To configure the Interface VLAN on the Cisco Nexus 9K Switch A, enter the following:

```
OSP8-N9K-FAB-A(config)#
OSP8-N9K-FAB-A(config)# interface Vlan10
OSP8-N9K-FAB-A(config-if)# description Management
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 10.23.10.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 10
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
```

```
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.10.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit


OSP8-N9K-FAB-A(config)# interface Vlan100
OSP8-N9K-FAB-A(config-if)# description Internal-API
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 10.23.100.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 100
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.100.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit

OSP8-N9K-FAB-A(config)# interface Vlan110
OSP8-N9K-FAB-A(config-if)# description PXE_Network
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 10.23.110.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 110
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.110.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit

OSP8-N9K-FAB-A(config)# interface Vlan120
OSP8-N9K-FAB-A(config-if)# description Storage_Public_Network
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 10.23.120.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 120
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.120.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit

OSP8-N9K-FAB-A(config)# interface Vlan150
OSP8-N9K-FAB-A(config-if)# description Storage_ClusterMgmt_Network
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 10.23.150.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 150
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.150.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit

OSP8-N9K-FAB-A(config)# interface Vlan160
OSP8-N9K-FAB-A(config-if)# description Tenanat_Floating_Network
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
```

```
OSP8-N9K-FAB-A(config-if)# ip address 10.23.160.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)#  hsrp version 2
OSP8-N9K-FAB-A(config-if-hsrp)# hsrp 160
OSP8-N9K-FAB-A(config-if-hsrp)# preempt
OSP8-N9K-FAB-A(config-if-hsrp)# priority 110
OSP8-N9K-FAB-A(config-if-hsrp)# ip 10.23.160.1
OSP8-N9K-FAB-A(config-if-hsrp)#exit

OSP8-N9K-FAB-A(config)# interface Vlan215
OSP8-N9K-FAB-A(config-if)# description External_Network
OSP8-N9K-FAB-A(config-if)# no shutdown
OSP8-N9K-FAB-A(config-if)# no ip redirects
OSP8-N9K-FAB-A(config-if)# ip address 172.22.215.253/24
OSP8-N9K-FAB-A(config-if)# no ipv6 redirects
OSP8-N9K-FAB-A(config-if)# exit
OSP8-N9K-FAB-A(config)# Copy running-config Startup-config
```

## Configure the Interface VLAN (SVI) on the Cisco Nexus 9K Switch B

To configure the Interface VLAN on the Cisco Nexus 9K Switch B, enter the following:

```
OSP8-N9k-FAB-B(config)#
OSP8-N9k-FAB-B(config)# interface Vlan10
OSP8-N9k-FAB-B(config-if)# description Management
OSP8-N9k-FAB-B(config-if)#  no shutdown
OSP8-N9k-FAB-B(config-if)#  no ip redirects
OSP8-N9k-FAB-B(config-if)#  ip address 10.23.10.254/24
OSP8-N9k-FAB-B(config-if)#  no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 100
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.100.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan100
OSP8-N9k-FAB-B(config-if)# description Internal-API
OSP8-N9k-FAB-B(config-if)#  no shutdown
OSP8-N9k-FAB-B(config-if)#  no ip redirects
OSP8-N9k-FAB-B(config-if)#  ip address 10.23.100.254/24
OSP8-N9k-FAB-B(config-if)#  no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 100
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.100.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan110
OSP8-N9k-FAB-B(config-if)# description PXE_Network
OSP8-N9k-FAB-B(config-if)#  no shutdown
OSP8-N9k-FAB-B(config-if)#  no ip redirects
OSP8-N9k-FAB-B(config-if)#  ip address 10.23.110.254/24
OSP8-N9k-FAB-B(config-if)#  no ipv6 redirects
OSP8-N9k-FAB-B(config-if)#  hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 110
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
```

```
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.110.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan120
OSP8-N9k-FAB-B(config-if)# description Storage_Public_Network
OSP8-N9k-FAB-B(config-if)# no shutdown
OSP8-N9k-FAB-B(config-if)# no ip redirects
OSP8-N9k-FAB-B(config-if)# ip address 10.23.120.254/24
OSP8-N9k-FAB-B(config-if)# no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 120
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.120.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan150
OSP8-N9k-FAB-B(config-if)# description Storage_ClusterMgmt_Network
OSP8-N9k-FAB-B(config-if)# no shutdown
OSP8-N9k-FAB-B(config-if)# no ip redirects
OSP8-N9k-FAB-B(config-if)# ip address 10.23.150.254/24
OSP8-N9k-FAB-B(config-if)# no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 150
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.150.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan160
OSP8-N9k-FAB-B(config-if)# description Tenanat_Floating_Network
OSP8-N9k-FAB-B(config-if)# no shutdown
OSP8-N9k-FAB-B(config-if)# no ip redirects
OSP8-N9k-FAB-B(config-if)# ip address 10.23.160.254/24
OSP8-N9k-FAB-B(config-if)# no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# hsrp version 2
OSP8-N9k-FAB-B(config-if-hsrp)# hsrp 160
OSP8-N9k-FAB-B(config-if-hsrp)# preempt
OSP8-N9k-FAB-B(config-if-hsrp)# priority 100
OSP8-N9k-FAB-B(config-if-hsrp)# ip 10.23.160.1
OSP8-N9k-FAB-B(config-if-hsrp)# exit

OSP8-N9k-FAB-B(config)# interface Vlan215
OSP8-N9k-FAB-B(config-if)# description External_Network
OSP8-N9k-FAB-B(config-if)# no shutdown
OSP8-N9k-FAB-B(config-if)# no ip redirects
OSP8-N9k-FAB-B(config-if)# ip address 172.22.215.254/24
OSP8-N9k-FAB-B(config-if)# no ipv6 redirects
OSP8-N9k-FAB-B(config-if)# exit
OSP8-N9K-FAB-B(config)# Copy running-config Startup-config
```

## Configure the VPC and Port Channels on Switch A

To configure the VPC and Port Channels on Switch A, enter the following:

```
OSP8-N9K-FAB-A(config)# vpc domain 1
OSP8-N9K-FAB-A(config-vpc-domain)# role priority 10
OSP8-N9K-FAB-A(config-vpc-domain)# peer-keepalive destination 10.23.100.4
OSP8-N9K-FAB-A(config-vpc-domain)# peer-gateway
```

```
OSP8-N9K-FAB-A(config-vpc-domain)# exit


OSP8-N9K-FAB-A(config)# interface port-channel1
OSP8-N9K-FAB-A(config-if)# description VPC peerlink for Nexus 9k Switch A & B
OSP8-N9K-FAB-A(config-if)# switchport mode trunk
OSP8-N9K-FAB-A(config-if)# spanning-tree port type network
OSP8-N9K-FAB-A(config-if)# speed 10000
OSP8-N9K-FAB-A(config-if)# vpc peer-link
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface Ethernet1/1
OSP8-N9K-FAB-A(config-if)# description connected to Peer Nexus 9k-B port1/1
OSP8-N9K-FAB-A(config-if)# switchport mode trunk
OSP8-N9K-FAB-A(config-if)# speed 10000
OSP8-N9K-FAB-A(config-if)# channel-group 1 mode active
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface Ethernet1/2
OSP8-N9K-FAB-A(config-if)# description connected to Peer Nexus 9k-B port1/2
OSP8-N9K-FAB-A(config-if)# switchport mode trunk
OSP8-N9K-FAB-A(config-if)# speed 10000
OSP8-N9K-FAB-A(config-if)# channel-group 1 mode active
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface port-channel17
OSP8-N9K-FAB-A(config-if)# description Port-channel for UCS_Fabric_A port_17 &
port_18
OSP8-N9K-FAB-A(config-if)# vpc 17
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface port-channel18
OSP8-N9K-FAB-A(config-if)# description Port-channel for UCS_Fabric_B port_17 &
port_18
OSP8-N9K-FAB-A(config-if)# vpc 18
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface Ethernet1/17
OSP8-N9K-FAB-A(config-if)# description Uplink from UCS_Fabric_A_Port_17
OSP8-N9K-FAB-A(config-if)# channel-group 17 mode active
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface Ethernet1/18
OSP8-N9K-FAB-A(config-if)# description Uplink from UCS_Fabric_B_Port_17
OSP8-N9K-FAB-A(config-if)# channel-group 18 mode active
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface port-channel17
OSP8-N9K-FAB-A(config-if)# switchport mode trunk
OSP8-N9K-FAB-A(config-if)# switchport trunk allowed vlan 10,100,110,120,
150,160,215
OSP8-N9K-FAB-A(config-if)# spanning-tree port type edge trunk
OSP8-N9K-FAB-A(config-if)# mtu 9216
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# interface port-channel18
OSP8-N9K-FAB-A(config-if)# switchport mode trunk
OSP8-N9K-FAB-A(config-if)# switchport trunk allowed vlan 10,100,110,120,
150,160,215
OSP8-N9K-FAB-A(config-if)# spanning-tree port type edge trunk
```

```
OSP8-N9K-FAB-A(config-if)# mtu 9216
OSP8-N9K-FAB-A(config-if)# exit

OSP8-N9K-FAB-A(config)# copy running-config startup-config
```

## Configure the VPC and Port Channels on the Cisco Nexus 9K Switch B

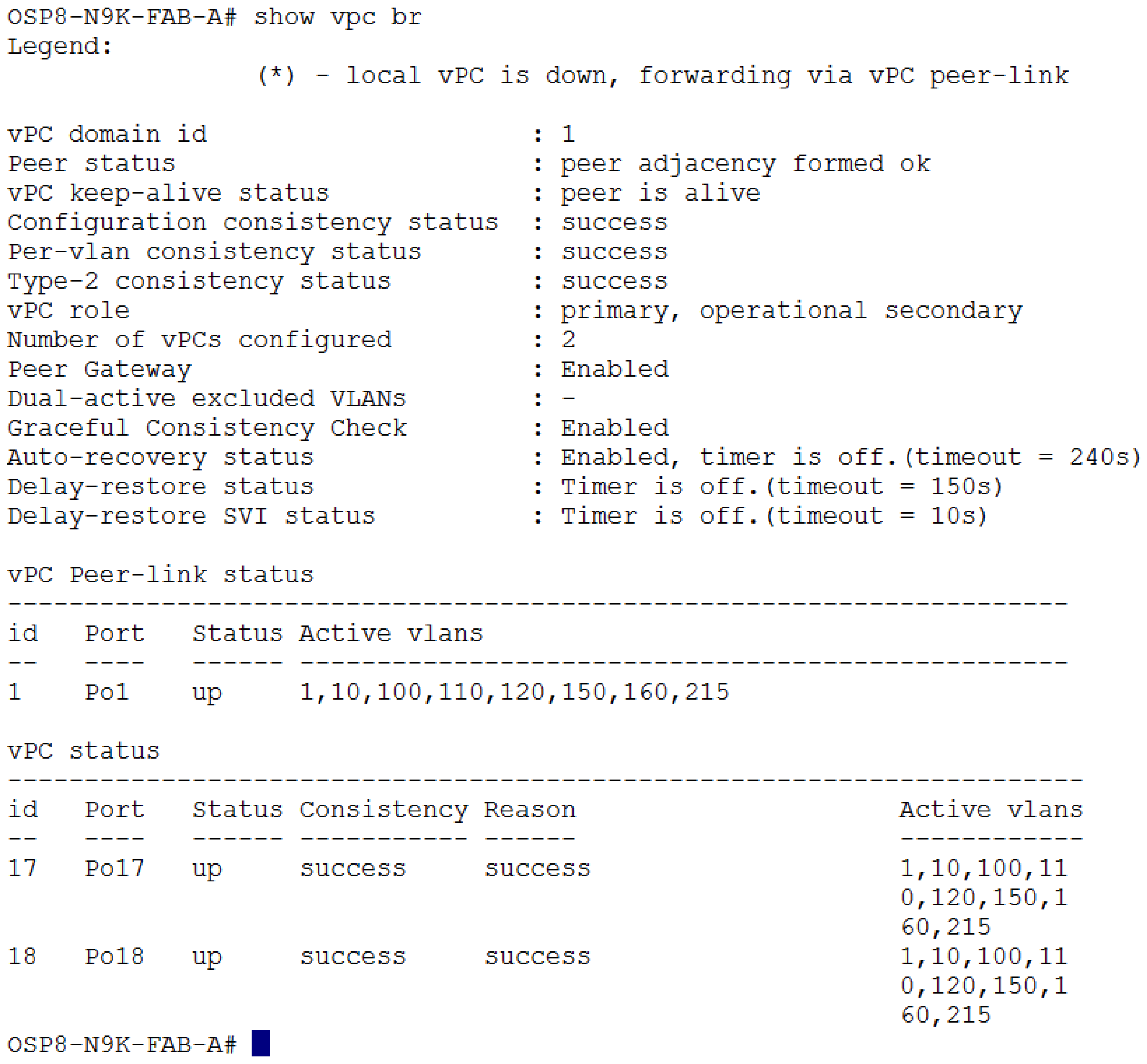
To configure the VPC and Port Channels on the Cisco Nexus 9K Switch B, enter the following:

```
OSP8-N9k-FAB-B(config)# vpc domain 1
OSP8-N9K-FAB-B(config-vpc-domain)# role priority 10
OSP8-N9k-FAB-B(config-vpc-domain)# peer-keepalive destination 10.23.100.3
OSP8-N9k-FAB-B(config-vpc-domain)# peer-gateway
OSP8-N9k-FAB-B(config-vpc-domain)# exit

OSP8-N9k-FAB-B(config)# interface port-channel1
OSP8-N9k-FAB-B(config-if)# description VPC peerlink for Nexus 9k Switch A & B
OSP8-N9k-FAB-B(config-if)# switchport mode trunk
OSP8-N9k-FAB-B(config-if)# spanning-tree port type network
OSP8-N9k-FAB-B(config-if)# speed 10000
OSP8-N9k-FAB-B(config-if)# vpc peer-link
OSP8-N9k-FAB-B(config-if)# exit

OSP8-N9k-FAB-B(config)# interface Ethernet1/1
OSP8-N9k-FAB-B(config-if)# description connected to Peer Nexus 9k-A port1/1
OSP8-N9k-FAB-B(config-if)# switchport mode trunk
OSP8-N9k-FAB-B(config-if)# speed 10000
OSP8-N9k-FAB-B(config-if)# channel-group 1 mode active
OSP8-N9k-FAB-B(config-if)# exit

OSP8-N9k-FAB-B(config)# interface Ethernet1/2
OSP8-N9k-FAB-B(config-if)# description connected to Peer Nexus 9k-A port1/2
OSP8-N9k-FAB-B(config-if)# switchport mode trunk
OSP8-N9k-FAB-B(config-if)# speed 10000
OSP8-N9k-FAB-B(config-if)# channel-group 1 mode active
OSP8-N9k-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# interface port-channel17
OSP8-N9K-FAB-B(config-if)# description Port-channel for UCS_Fabric_A port_17 &
port_18
OSP8-N9K-FAB-B(config-if)# vpc 17
OSP8-N9K-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# interface port-channel18
OSP8-N9K-FAB-B(config-if)# description Port-channel for UCS_Fabric_B port_17 &
port_18
OSP8-N9K-FAB-B(config-if)# vpc 18
OSP8-N9K-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# interface Ethernet1/17
OSP8-N9K-FAB-B(config-if)# description Uplink from UCS_Fabric_A_Port_18
OSP8-N9K-FAB-B(config-if)# channel-group 17 mode active
OSP8-N9K-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# interface Ethernet1/18
OSP8-N9K-FAB-B(config-if)# description Uplink from UCS_Fabric_B_Port_18
OSP8-N9K-FAB-B(config-if)# channel-group 18 mode active
OSP8-N9K-FAB-B(config-if)# exit
```

```
OSP8-N9K-FAB-B(config)# interface port-channel17
OSP8-N9K-FAB-B(config-if)# switchport mode trunk
OSP8-N9K-FAB-B(config-if)# switchport trunk allowed vlan 10,100,110,120,
150,160,215
OSP8-N9K-FAB-B(config-if)# spanning-tree port type edge trunk
OSP8-N9K-FAB-B(config-if)# mtu 9216
OSP8-N9K-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# interface port-channel18
OSP8-N9K-FAB-B(config-if)# switchport mode trunk
OSP8-N9K-FAB-B(config-if)# switchport trunk allowed vlan 10,100,110,120,
150,160,215
OSP8-N9K-FAB-B(config-if)# spanning-tree port type edge trunk
OSP8-N9K-FAB-B(config-if)# mtu 9216
OSP8-N9K-FAB-B(config-if)# exit

OSP8-N9K-FAB-B(config)# copy running-config startup-config
```

## Verify the Port Channel Status on the Cisco Nexus Switches

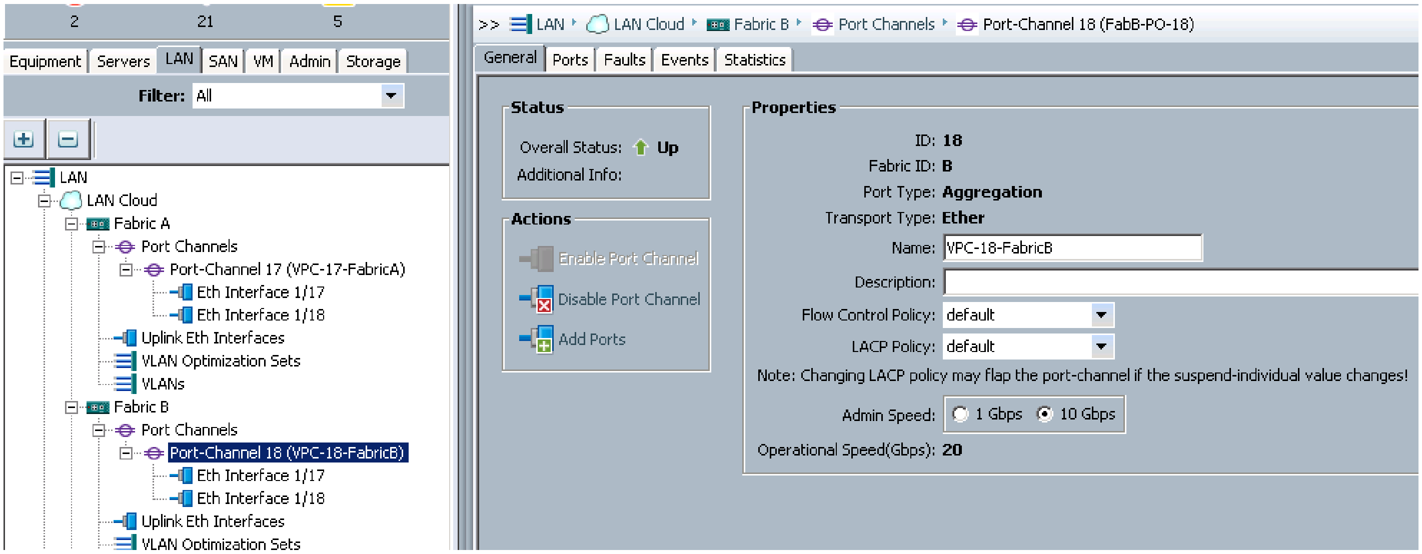
After successfully creating a Virtual Port Channel on both Nexus switches, verify the Port Channel status on the Nexus 9K Switch. To verify the status, enter the following:



```
OSP8-N9K-FAB-A# show vpc br
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary, operational secondary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Enabled, timer is off.(timeout = 240s)
Delay-restore status              : Timer is off.(timeout = 150s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)

vPC Peer-link status
---------------------------------------------------------------------
id    Port   Status Active vlans
--    ----   ------ -------------------------------------------------
1     Po1    up     1,10,100,110,120,150,160,215

vPC status
----------------------------------------------------------------------
id    Port   Status Consistency Reason                     Active vlans
--    ----   ------ ----------- ------                     ------------
17    Po17   up     success     success                    1,10,100,11
                                                           0,120,150,1
                                                           60,215
18    Po18   up     success     success                    1,10,100,11
                                                           0,120,150,1
                                                           60,215

OSP8-N9K-FAB-A# █
```

## Verify the Port Channels Status on the Fabrics

To verify the status on the Fabrics, complete the following steps as shown in the screenshots below:



Status

Overall Status: Up

Additional Info:

Properties

ID: 18

Fabric ID: B

Port Type: Aggregation

Transport Type: Ether

Name: VPC-18-FabricB

Description:

Flow Control Policy: default

LACP Policy: default

Note: Changing LACP policy may flap the port-channel if the suspend-individual value changes!

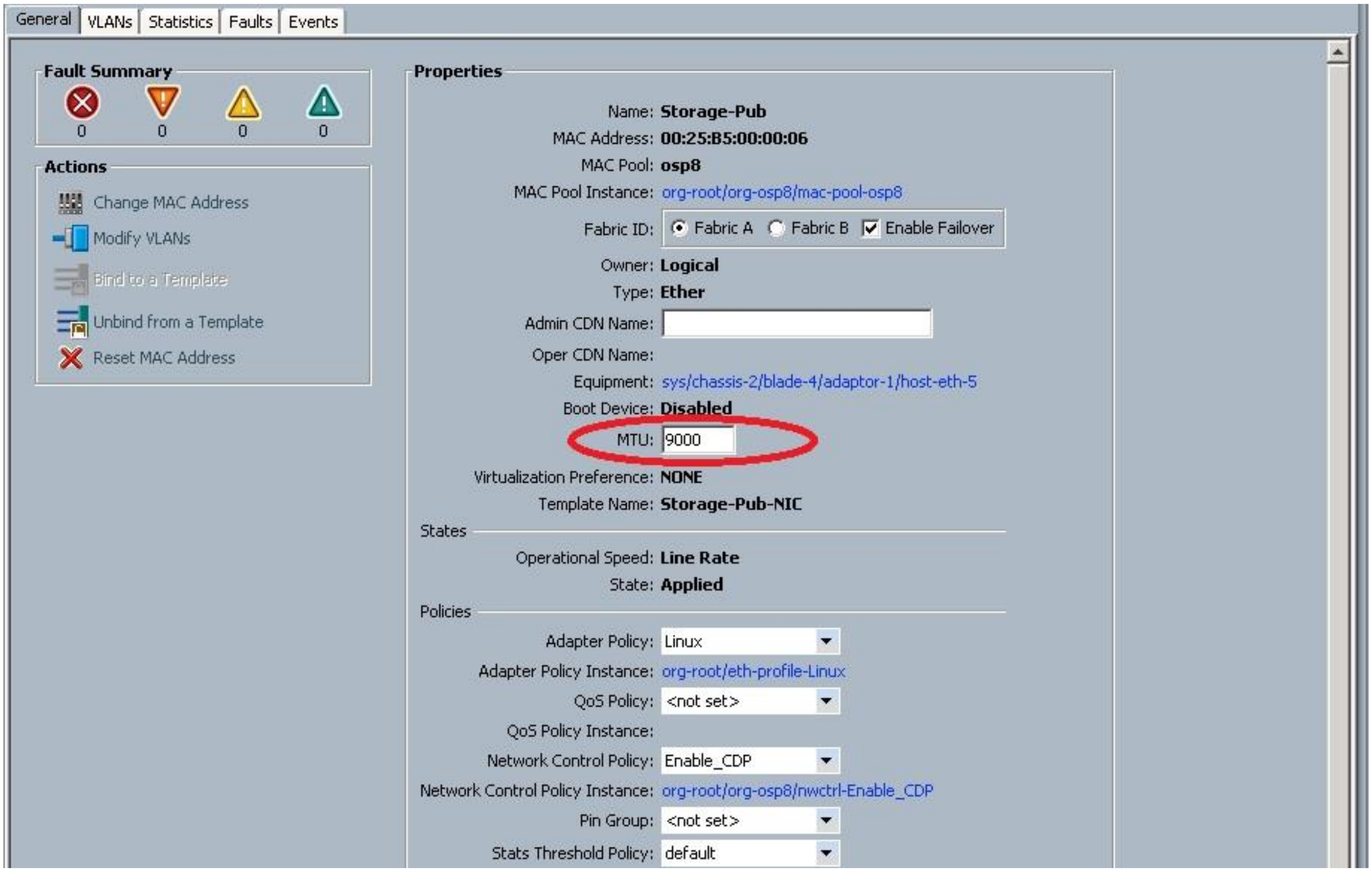
Admin Speed: ○ 1 Gbps ● 10 Gbps

Operational Speed(Gbps): 20

# Cisco UCS Validation Checks

Prior to starting the Operating System installation on the Undercloud Node, you must complete the pre-validation checks. To complete the validation checks, complete the following steps:

If you are planning to use Jumbo frames for the storage network, make sure to enter the following information in the templates as shown in the screenshot below.



Properties

Name: Storage-Pub
MAC Address: 00:25:B5:00:00:06
MAC Pool: osp8
MAC Pool Instance: org-root/org-osp8/mac-pool-osp8
Fabric ID: ● Fabric A ○ Fabric B ☑ Enable Failover
Owner: Logical
Type: Ether
Admin CDN Name:
Oper CDN Name:
Equipment: sys/chassis-2/blade-4/adaptor-1/host-eth-5
Boot Device: Disabled
MTU: 9000
Virtualization Preference: NONE
Template Name: Storage-Pub-NIC

States

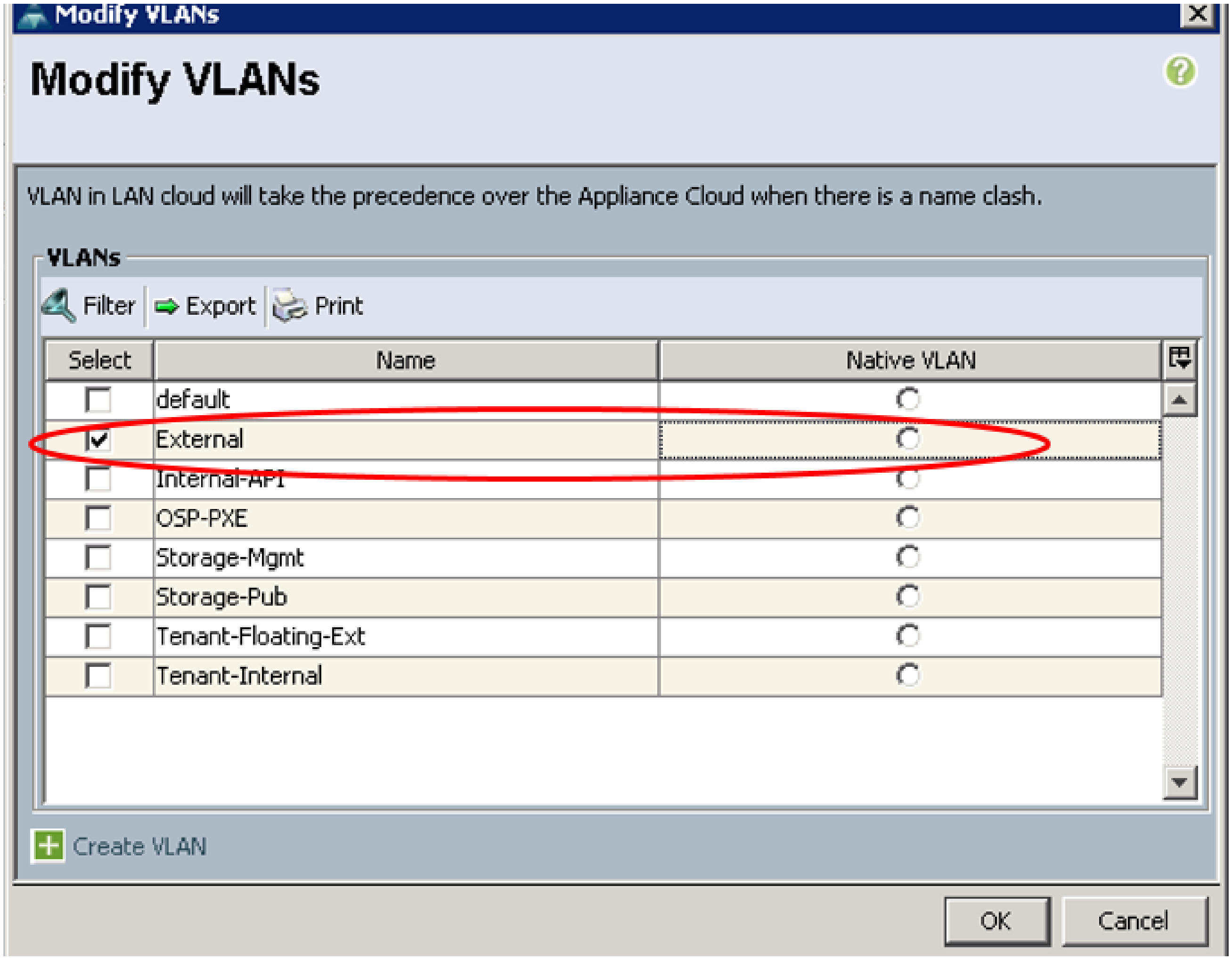
Operational Speed: Line Rate
State: Applied

Policies

Adapter Policy: Linux
Adapter Policy Instance: org-root/eth-profile-Linux
QoS Policy: <not set>
QoS Policy Instance:
Network Control Policy: Enable_CDP
Network Control Policy Instance: org-root/org-osp8/nwctrl-Enable_CDP
Pin Group: <not set>
Stats Threshold Policy: default

When the service profiles are created from the template, unbind from the templates in case they have been created as updating templates. This is to accommodate the Cisco UCS Manager Plugin. Keeping the compute host's service profiles bound to the template does not allow the plugin to individually configure each compute host with tenant based VLANs. The service profiles for each compute host need to be unbound from the template. Please check the current limitations outlined in the UCSM Liberty plugin web page.

VLAN ID is included in the OpenStack configuration. Do not tag the native VLAN for your external interface on the overcloud service profiles.



VLAN in LAN cloud will take the precedence over the Appliance Cloud when there is a name clash.

| Select | Name | Native VLAN |
|---|---|---|
| □ | default | ○ |
| ☑ | External | ○ |
| □ | Internal-API | ○ |
| □ | OSP-PXE | ○ |
| □ | Storage-Mgmt | ○ |
| □ | Storage-Pub | ○ |
| □ | Tenant-Floating-Ext | ○ |
| □ | Tenant-Internal | ○ |

The provisioning interfaces should be Native for both Undercloud and Overcloud setups.

While planning your networks, make sure all the networks defined are not overlapping with any of your data-center networks.

The disks should be in the same order across all storage nodes.

# Install the Operating System on the Undercloud Node

It is highly recommended to install the Operating System with versionlock as outlined in the steps below. Versionlock restricts yum to install or upgrade a package to a fixed specific version than specified using the versionlock plugin of yum.

The steps outlined in this document, including a few of the configurations, are bound to the installed packages. Installing the same set of packages, as in this Cisco Validated Design, ensures the accuracy of the solution with minimal deviations. While installing Red Hat OpenStack Platform 8 on Cisco blade and rack servers without version, lock should still work; it needs to be noted that there could be changes in the configurations and install steps needed that may not exist in this document.

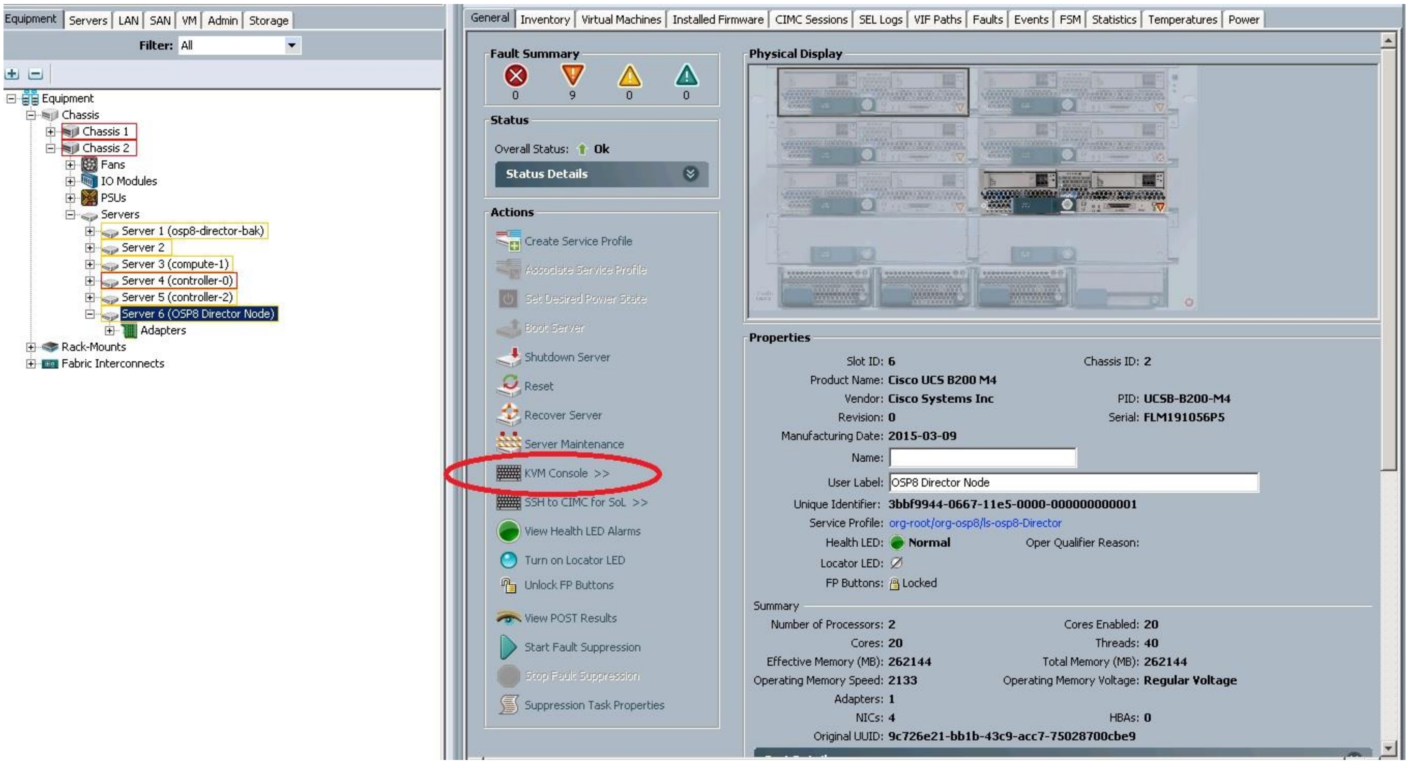
Any updates to the Undercloud stack later through yum install may conflict with the version lock packages. You may have to relax the lock files for such updates, when it is required. It is strongly recommended to complete the install with version lock first followed by Overcloud install before attempting any such updates.

Download the versionlock file from Cisco Systems https://communities.cisco.com/docs/DOC-70256

To install the Operating System on the Undercloud Node, complete the following steps:

Download Red Hat Enterprise Linux 7.2 from http://access.redhat.com.

Launch the KVM Console; UCS Manager > Equipment Tab > General > KVM Console.



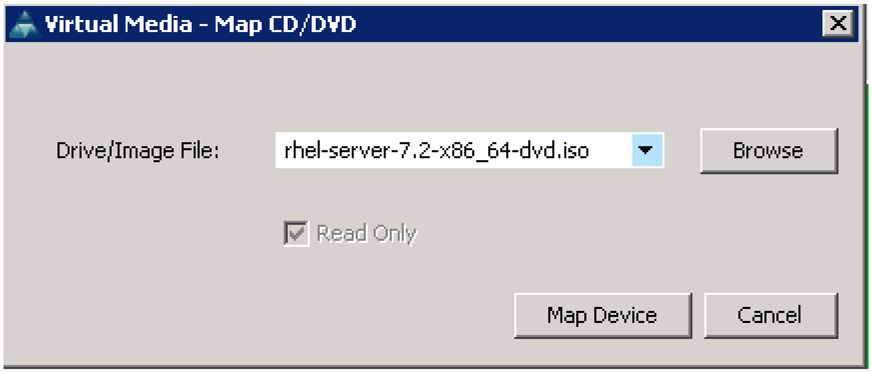
Properties

Slot ID: 6 Chassis ID: 2
Product Name: Cisco UCS B200 M4
Vendor: Cisco Systems Inc PID: UCSB-B200-M4
Revision: 0 Serial: FLM191056P5
Manufacturing Date: 2015-03-09
Name:
User Label: OSP8 Director Node
Unique Identifier: 3bbf9944-0667-11e5-0000-000000000001
Service Profile: org-root/org-osp8/ls-osp8-Director
Health LED: Normal Oper Qualifier Reason:
Locator LED:
FP Buttons: Locked

Summary

Number of Processors: 2 Cores Enabled: 20
Cores: 20 Threads: 40
Effective Memory (MB): 262144 Total Memory (MB): 262144
Operating Memory Speed: 2133 Operating Memory Voltage: Regular Voltage
Adapters: 1
NICs: 4 HBAs: 0
Original UUID: 9c726e21-bb1b-43c9-acc7-75028700cbe9

In the KVM Console Menu, Activate Virtual Devices under Virtual Media and then click Map CD/DVD, attach the downloaded ISO as shown below and then reboot the server.

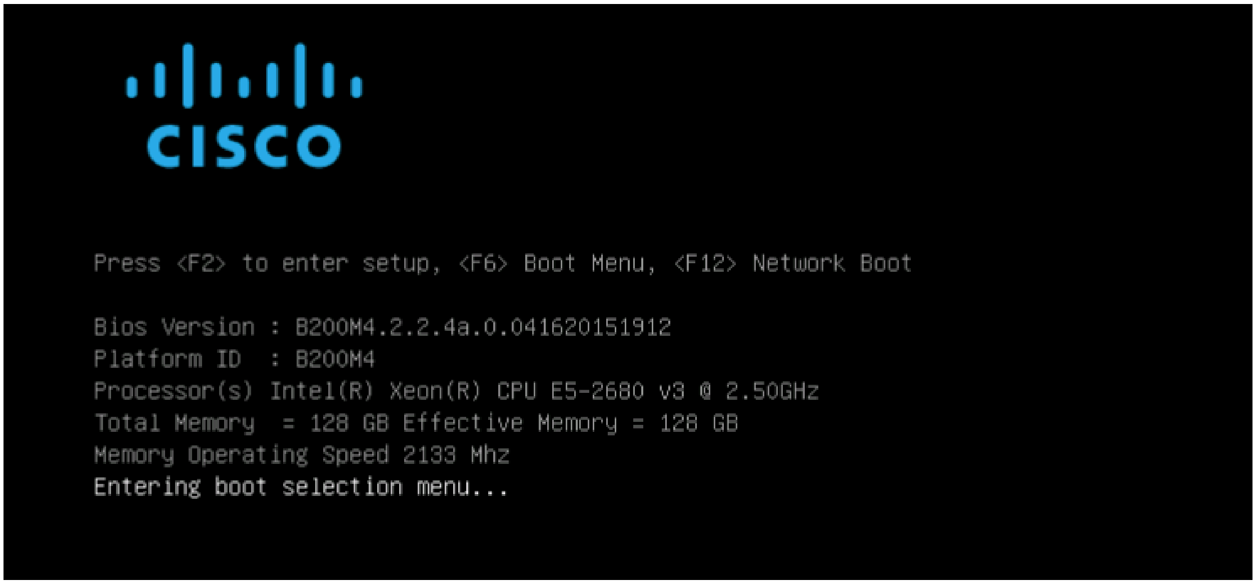


Virtual Media - Map CD/DVD

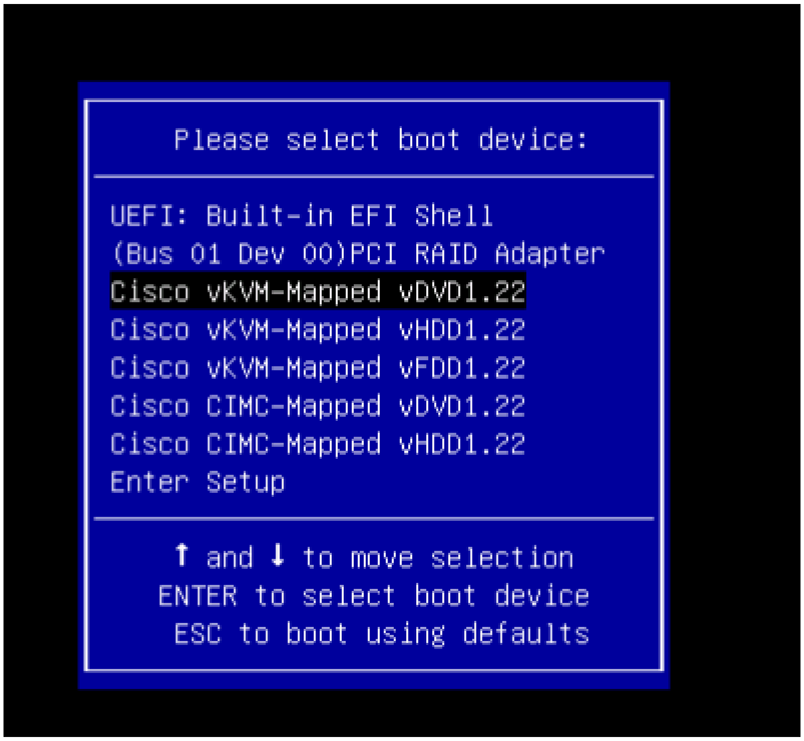
Drive/Image File: rhel-server-7.2-x86_64-dvd.iso

☑ Read Only

When the system boots up, press F6 for the boot menu.

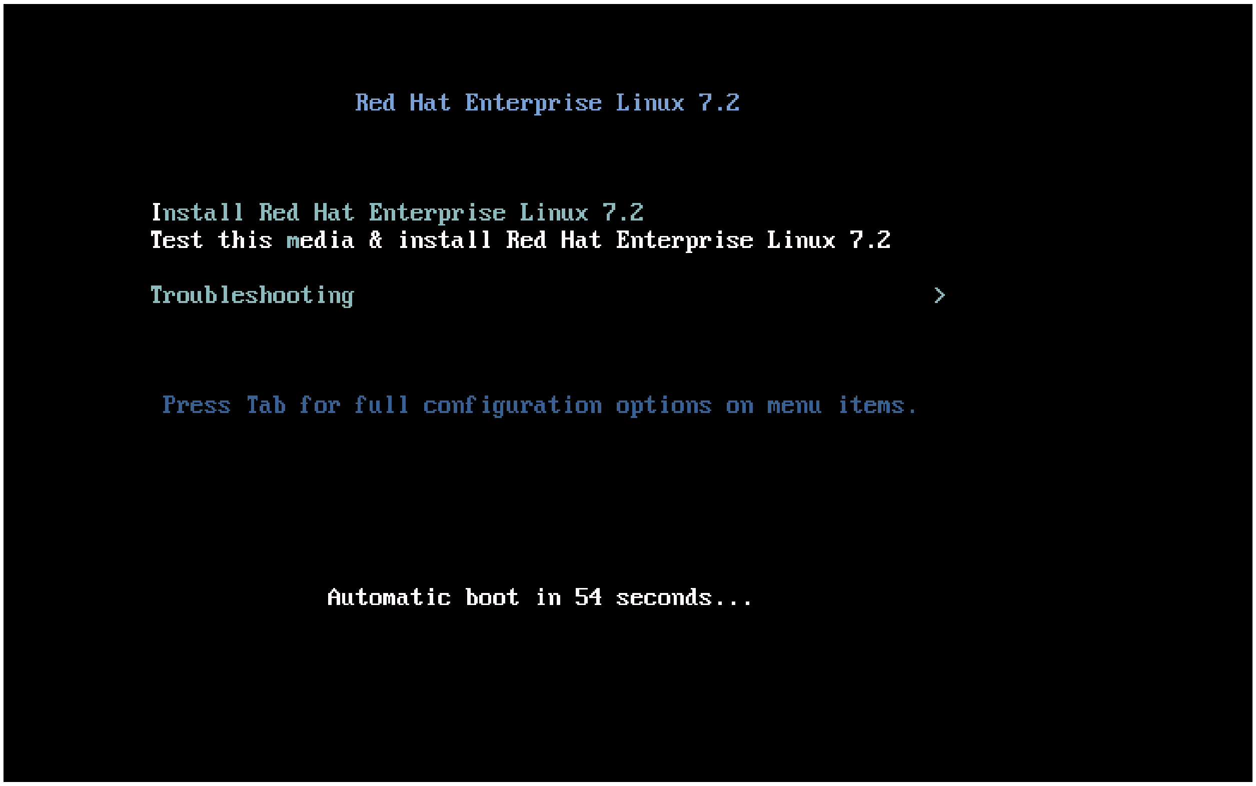


```
Press <F2> to enter setup, <F6> Boot Menu, <F12> Network Boot

Bios Version : B200M4.2.2.4a.0.041620151912
Platform ID  : B200M4
Processor(s) Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz
Total Memory  = 128 GB Effective Memory = 128 GB
Memory Operating Speed 2133 Mhz
Entering boot selection menu...
```



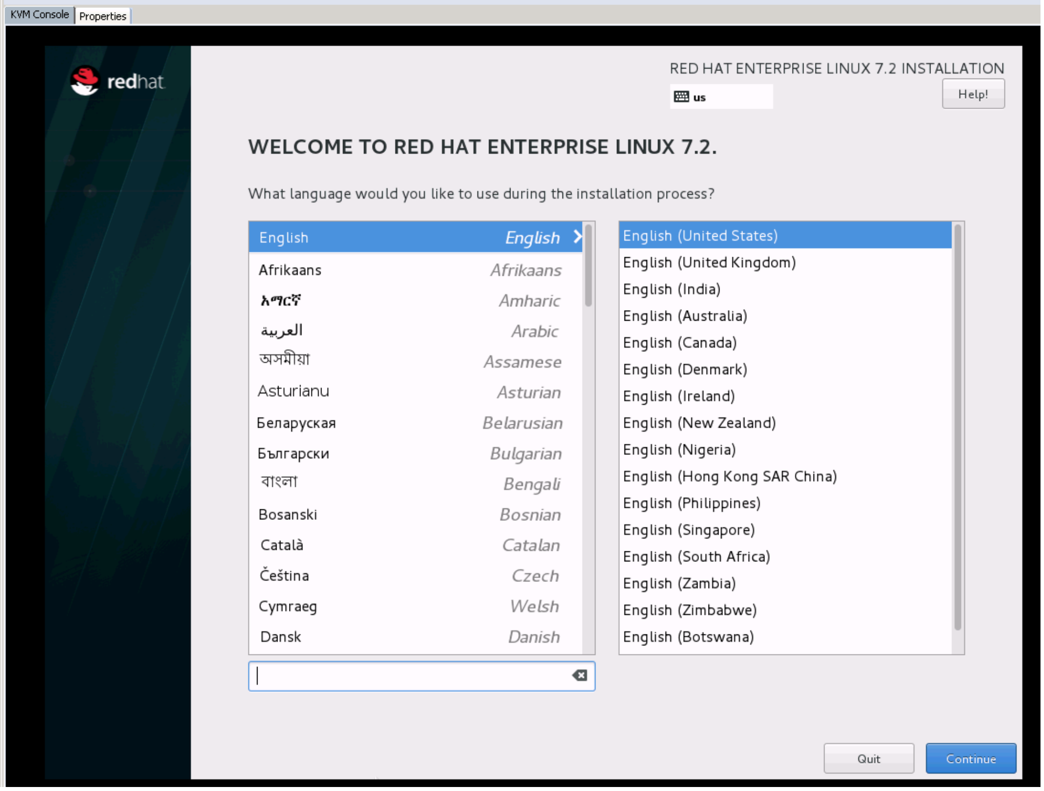
```
     Please select boot device:

 UEFI: Built-in EFI Shell
 (Bus 01 Dev 00)PCI RAID Adapter
 Cisco vKVM-Mapped vDVD1.22
 Cisco vKVM-Mapped vHDD1.22
 Cisco vKVM-Mapped vFDD1.22
 Cisco CIMC-Mapped vDVD1.22
 Cisco CIMC-Mapped vHDD1.22
 Enter Setup

     ↑ and ↓ to move selection
   ENTER to select boot device
     ESC to boot using defaults
```

The ISO image takes you to the below screen. Select Install Red Hat Enterprise Linux 7.2.



```
                Red Hat Enterprise Linux 7.2


Install Red Hat Enterprise Linux 7.2
Test this media & install Red Hat Enterprise Linux 7.2

Troubleshooting                                          >


 Press Tab for full configuration options on menu items.






             Automatic boot in 54 seconds...
```

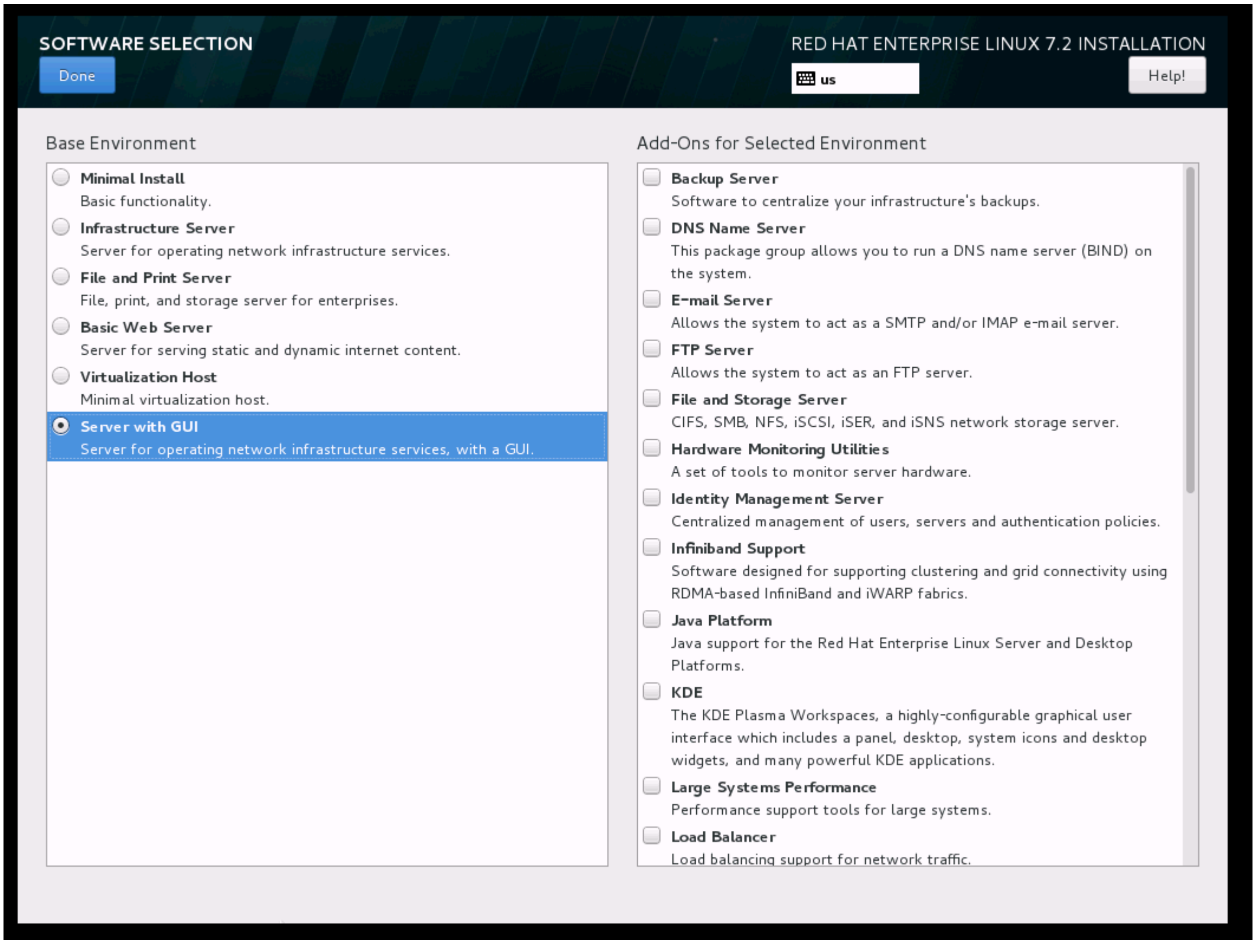
Select the default language and time zone.



WELCOME TO RED HAT ENTERPRISE LINUX 7.2.

What language would you like to use during the installation process?

In the software selection, screen select server with GUI.



Base Environment

○ **Minimal Install**
Basic functionality.

○ **Infrastructure Server**
Server for operating network infrastructure services.

○ **File and Print Server**
File, print, and storage server for enterprises.

○ **Basic Web Server**
Server for serving static and dynamic internet content.

○ **Virtualization Host**
Minimal virtualization host.

● **Server with GUI**
Server for operating network infrastructure services, with a GUI.

Add-Ons for Selected Environment

□ **Backup Server**
Software to centralize your infrastructure's backups.

□ **DNS Name Server**
This package group allows you to run a DNS name server (BIND) on the system.

□ **E-mail Server**
Allows the system to act as a SMTP and/or IMAP e-mail server.

□ **FTP Server**
Allows the system to act as an FTP server.

□ **File and Storage Server**
CIFS, SMB, NFS, iSCSI, iSER, and iSNS network storage server.

□ **Hardware Monitoring Utilities**
A set of tools to monitor server hardware.

□ **Identity Management Server**
Centralized management of users, servers and authentication policies.

□ **Infiniband Support**
Software designed for supporting clustering and grid connectivity using RDMA-based InfiniBand and iWARP fabrics.

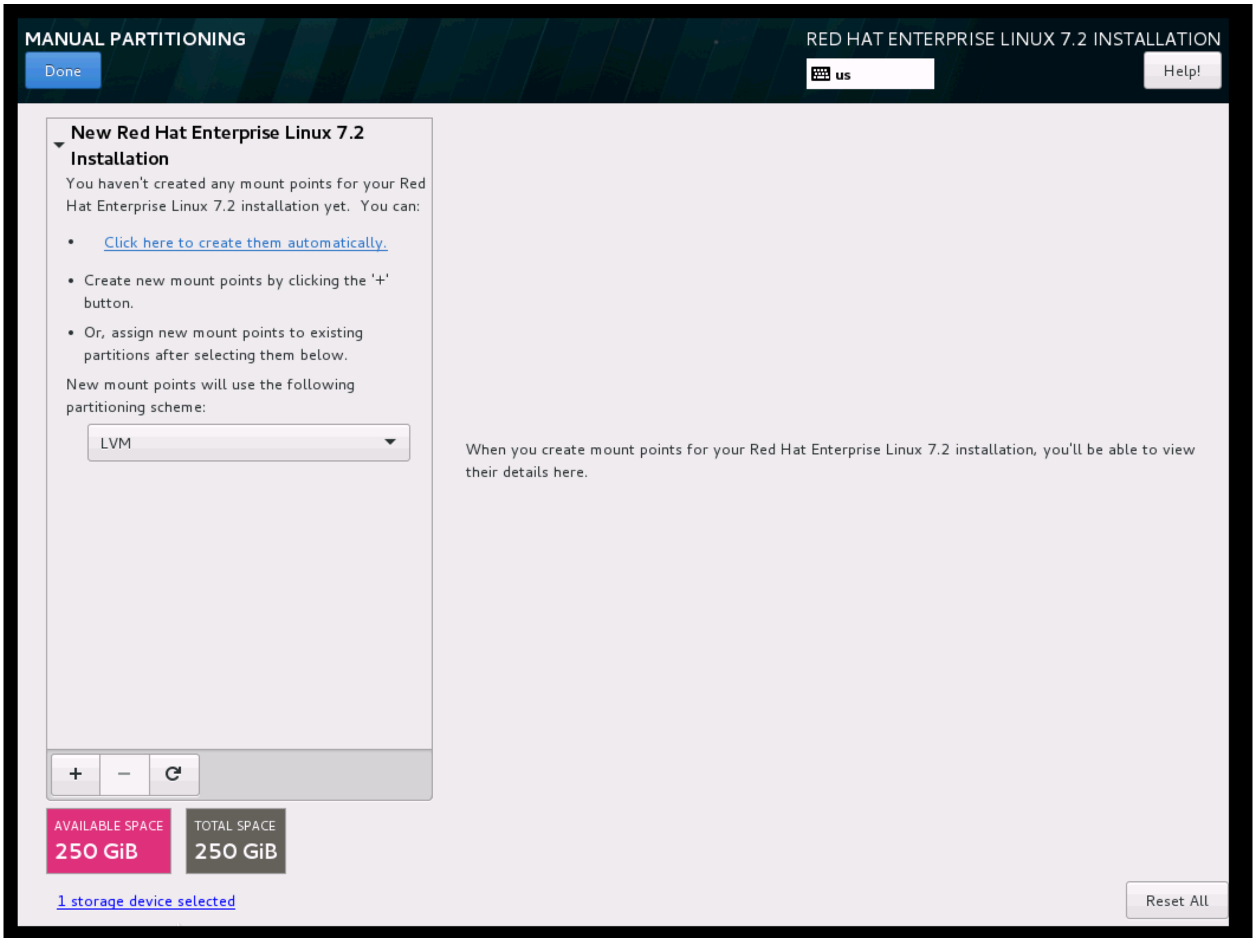
□ **Java Platform**
Java support for the Red Hat Enterprise Linux Server and Desktop Platforms.

□ **KDE**
The KDE Plasma Workspaces, a highly-configurable graphical user interface which includes a panel, desktop, system icons and desktop widgets, and many powerful KDE applications.

□ **Large Systems Performance**
Performance support tools for large systems.

□ **Load Balancer**
Load balancing support for network traffic.

Select manual partitioning as shown below.



MANUAL PARTITIONING

RED HAT ENTERPRISE LINUX 7.2 INSTALLATION

New Red Hat Enterprise Linux 7.2 Installation

You haven't created any mount points for your Red Hat Enterprise Linux 7.2 installation yet. You can:

- Click here to create them automatically.
- Create new mount points by clicking the '+' button.
- Or, assign new mount points to existing partitions after selecting them below.

New mount points will use the following partitioning scheme:

LVM

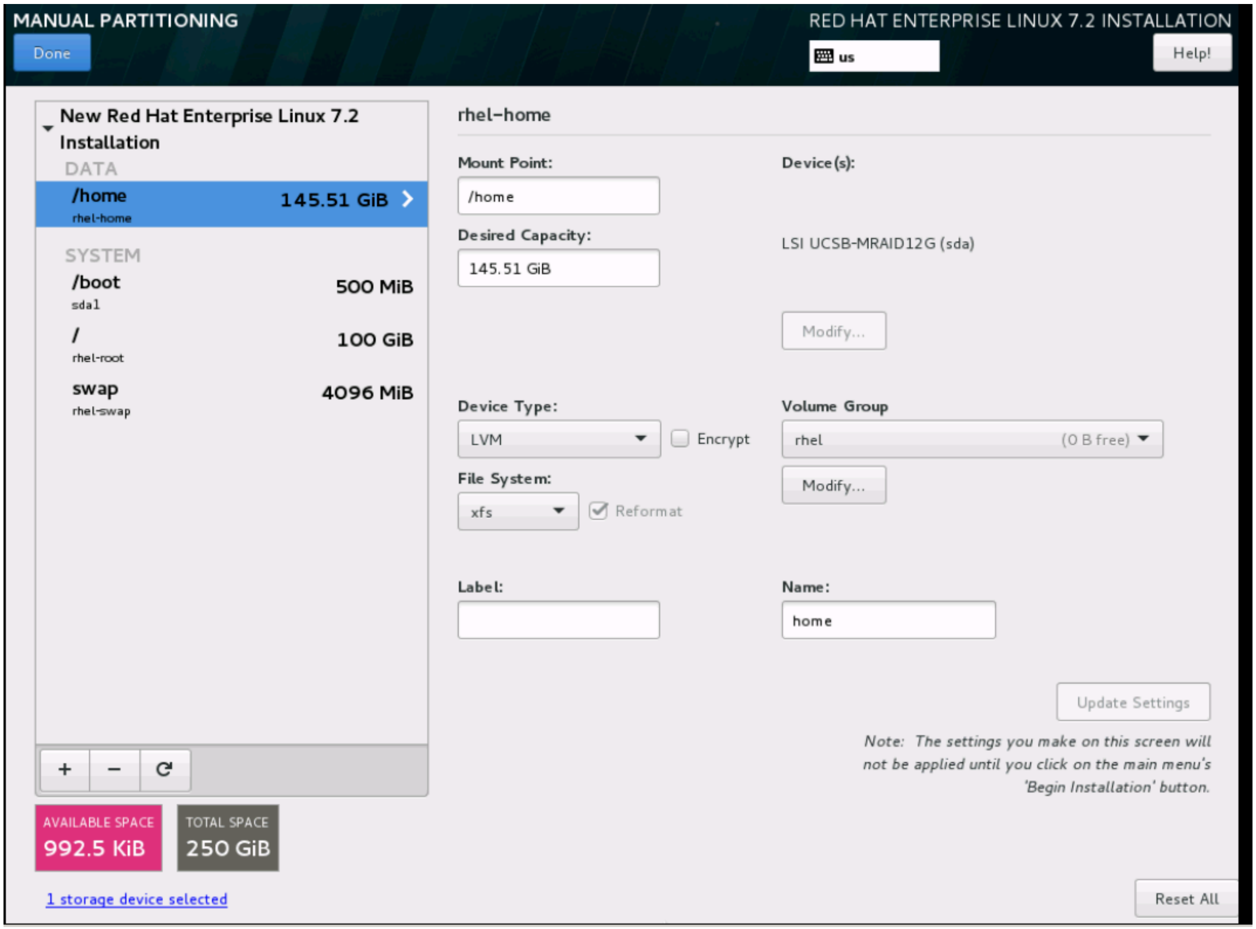
When you create mount points for your Red Hat Enterprise Linux 7.2 installation, you'll be able to view their details here.

AVAILABLE SPACE 250 GiB

TOTAL SPACE 250 GiB

1 storage device selected

Update the values and allocate around 100GB to root partition.



MANUAL PARTITIONING

Done

RED HAT ENTERPRISE LINUX 7.2 INSTALLATION

us

Help!

New Red Hat Enterprise Linux 7.2 Installation

DATA

/home rhel-home 145.51 GiB

SYSTEM

/boot sda1 500 MiB

/ rhel-root 100 GiB

swap rhel-swap 4096 MiB

AVAILABLE SPACE 992.5 KiB

TOTAL SPACE 250 GiB

1 storage device selected

rhel-home

Mount Point: /home

Desired Capacity: 145.51 GiB

Device(s): LSI UCSB-MRAID12G (sda)

Modify...

Device Type: LVM

□ Encrypt

File System: xfs

☑ Reformat

Volume Group: rhel (0 B free)

Modify...

Label:

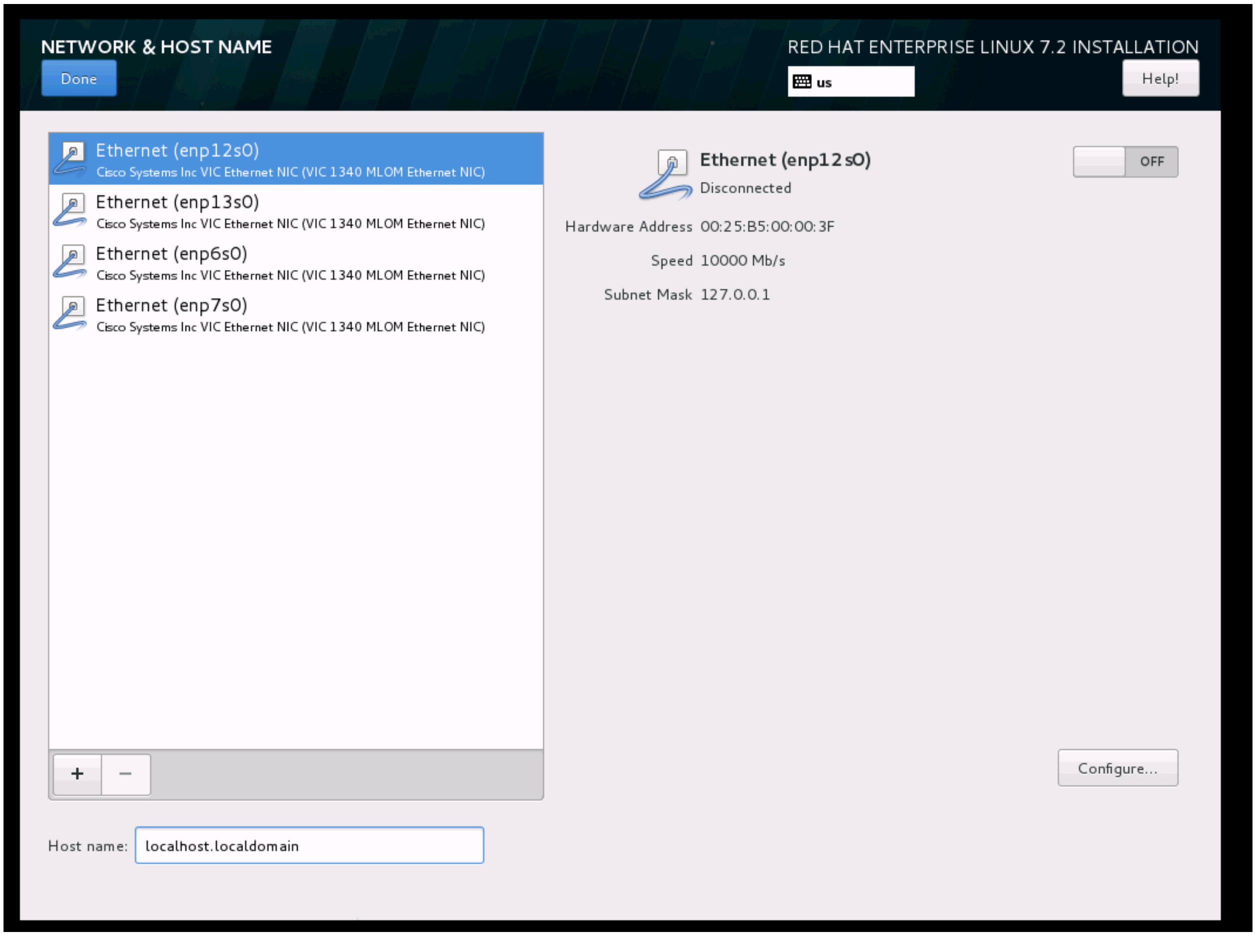
Name: home

Update Settings

*Note: The settings you make on this screen will not be applied until you click on the main menu's 'Begin Installation' button.*

Reset All

Select network tab and configure the external and Management NICs as shown below. Compare the MAC address in UCSM with MAC address displayed by the installer to identify the correct vNIC configuration.



Ethernet (enp12s0)
Cisco Systems Inc VIC Ethernet NIC (VIC 1340 MLOM Ethernet NIC)

Ethernet (enp13s0)
Cisco Systems Inc VIC Ethernet NIC (VIC 1340 MLOM Ethernet NIC)

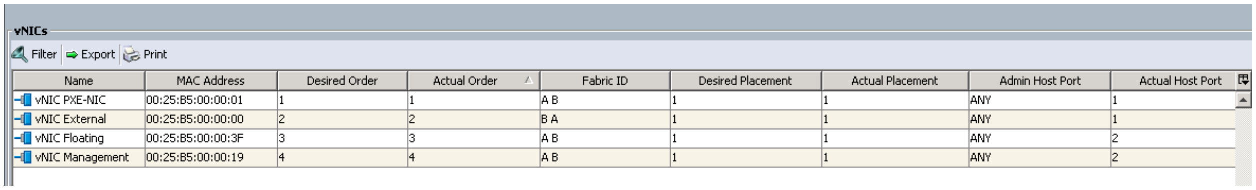
Ethernet (enp6s0)
Cisco Systems Inc VIC Ethernet NIC (VIC 1340 MLOM Ethernet NIC)

Ethernet (enp7s0)
Cisco Systems Inc VIC Ethernet NIC (VIC 1340 MLOM Ethernet NIC)

Ethernet (enp12s0)
Disconnected

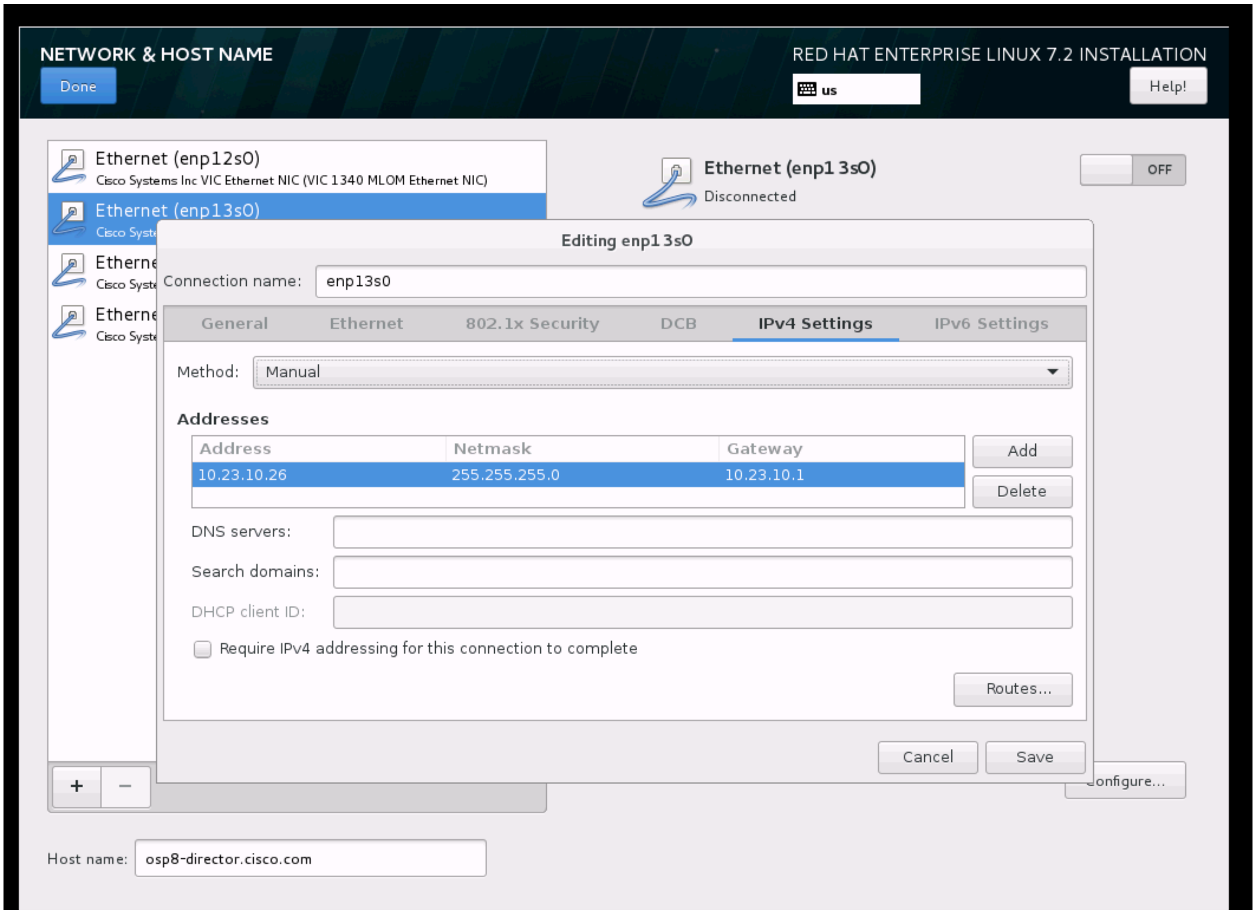
Hardware Address 00:25:B5:00:00:3F
Speed 10000 Mb/s
Subnet Mask 127.0.0.1

Host name: localhost.localdomain



| Name | MAC Address | Desired Order | Actual Order | Fabric ID | Desired Placement | Actual Placement | Admin Host Port | Actual Host Port |
|---|---|---|---|---|---|---|---|---|
| vNIC PXE-NIC | 00:25:B5:00:00:01 | 1 | 1 | A B | 1 | 1 | ANY | 1 |
| vNIC External | 00:25:B5:00:00:00 | 2 | 2 | B A | 1 | 1 | ANY | 1 |
| vNIC Floating | 00:25:B5:00:00:3F | 3 | 3 | A B | 1 | 1 | ANY | 2 |
| vNIC Management | 00:25:B5:00:00:19 | 4 | 4 | A B | 1 | 1 | ANY | 2 |

The Floating network has been added on the test bed for logging into the VMs from director node. This step is optional.



Connection name: enp13s0

Method: Manual

Addresses

| Address | Netmask | Gateway |
|---|---|---|
| 10.23.10.26 | 255.255.255.0 | 10.23.10.1 |

DNS servers:

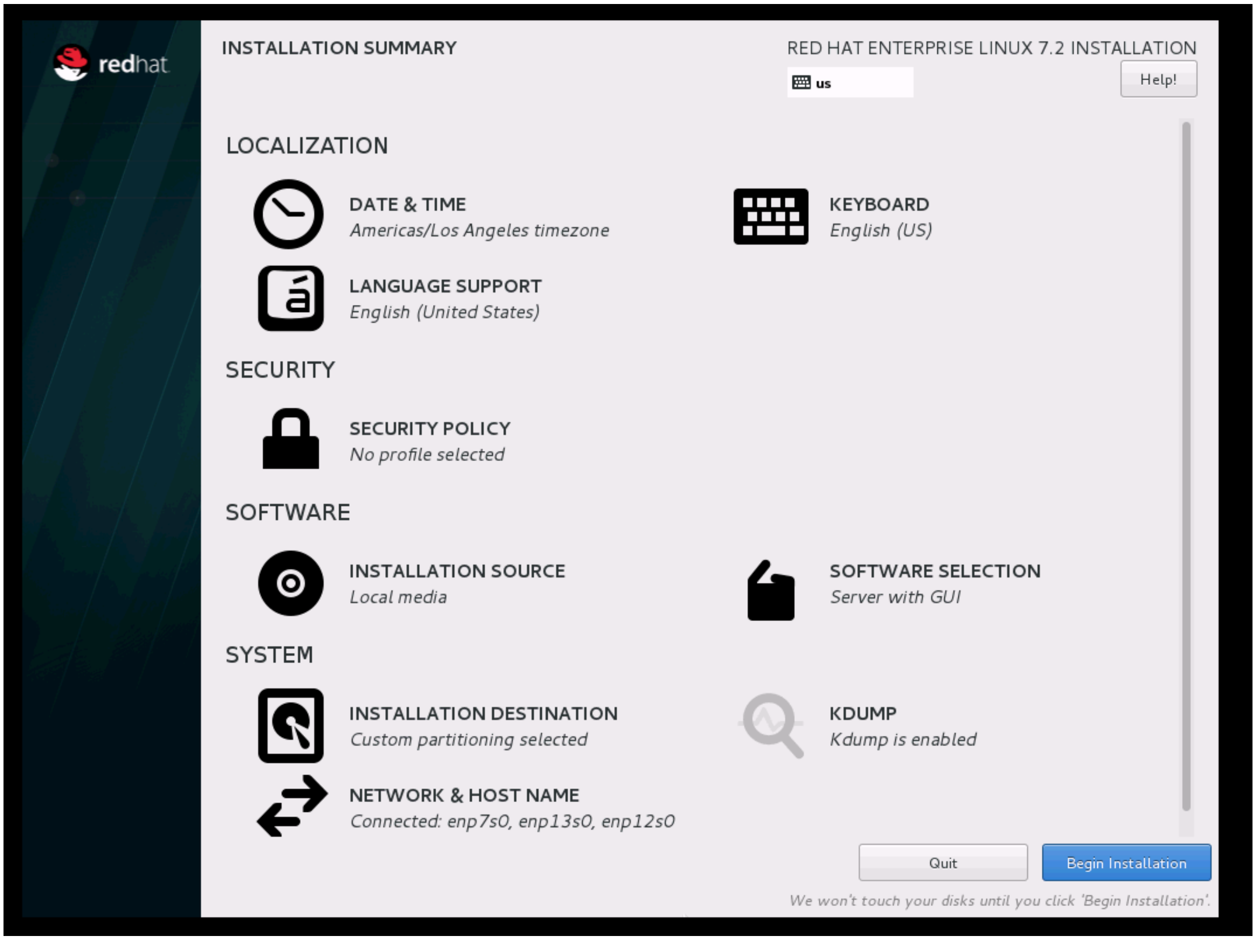
Search domains:

DHCP client ID:

□ Require IPv4 addressing for this connection to complete

Host name: osp8-director.cisco.com

As shown above, configure the External and optional Floating NICs here.



INSTALLATION SUMMARY

RED HAT ENTERPRISE LINUX 7.2 INSTALLATION

us

LOCALIZATION

DATE & TIME
*Americas/Los Angeles timezone*

KEYBOARD
*English (US)*

LANGUAGE SUPPORT
*English (United States)*

SECURITY

SECURITY POLICY
*No profile selected*

SOFTWARE

INSTALLATION SOURCE
*Local media*

SOFTWARE SELECTION
*Server with GUI*

SYSTEM

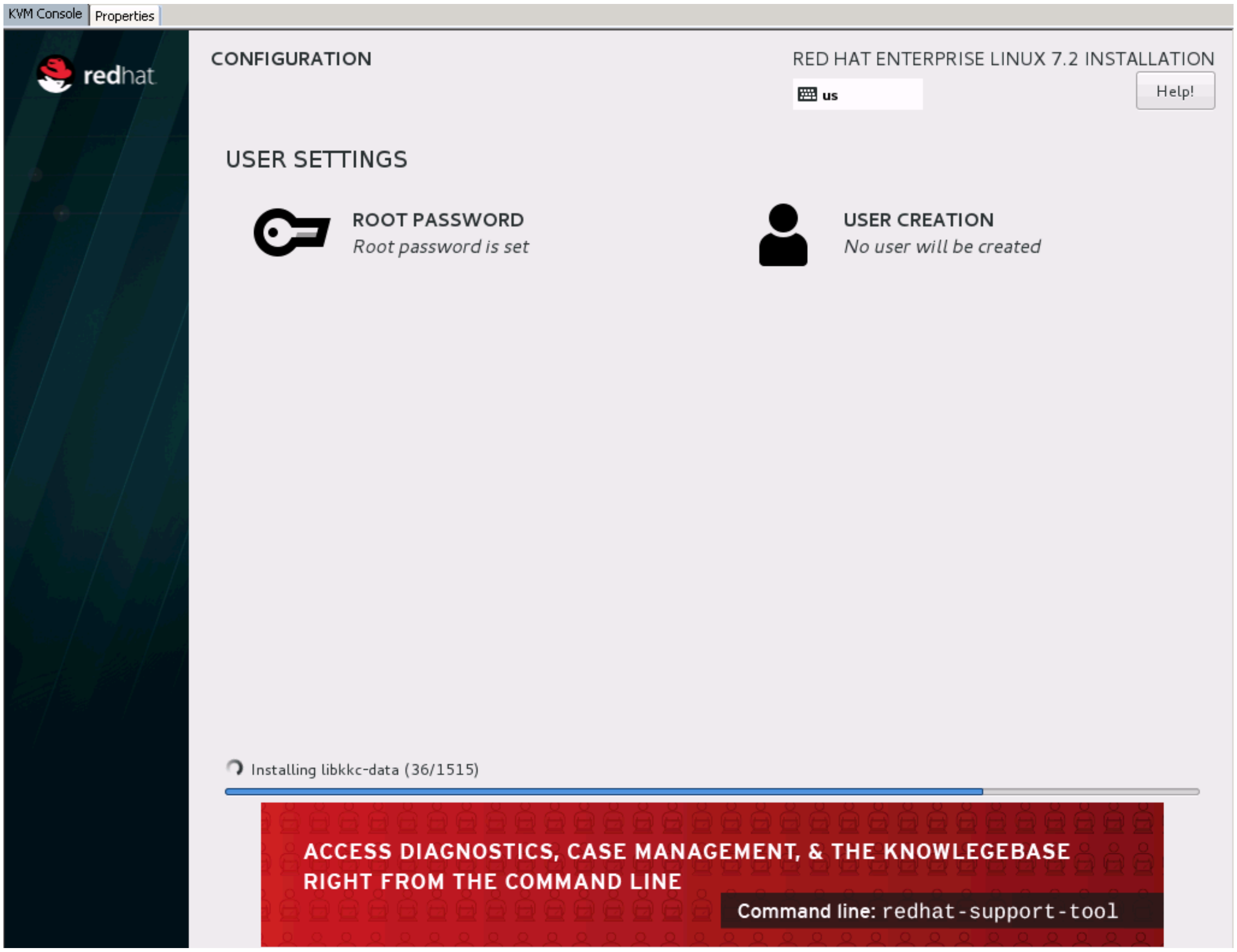
INSTALLATION DESTINATION
*Custom partitioning selected*

KDUMP
*Kdump is enabled*

NETWORK & HOST NAME
*Connected: enp7s0, enp13s0, enp12s0*

*We won't touch your disks until you click 'Begin Installation'.*

Add External-Network for public network. This is the interface that Undercloud will pull the necessary files from the Red Hat website during the installation. Management Interface is on the test bed to log into Fabric Interconnects and/or Nexus switches. This is how IPMI connectivity happens during introspection and Overcloud installations. If you already have a routable network for UCSM and Nexus, you do not need this interface. Leave the pxe interface NIC as unconfigured. It will be configured later through the Undercloud installation. Floating interface on the director node is not mandatory either. It has been added on the test system to log into VMs from director node.


KVM Console
Properties
redhat
CONFIGURATION
RED HAT ENTERPRISE LINUX 7.2 INSTALLATION
us
Help!
ACCESS DIAGNOSTICS, CASE MANAGEMENT, & THE KNOWLEGEBASE RIGHT FROM THE COMMAND LINE
Command line: redhat-support-tool

USER SETTINGS

ROOT PASSWORD
Root password is set

USER CREATION
No user will be created

Installing libkkc-data (36/1515)

Enter the root password and optionally create the stack user and reboot the server when prompted and accept the license agreement.

Run Post Install Steps before proceeding:

Register the director node with subscription Management and with release set as 7.2. Attach the pool with OpenStack entitlements.

subscription-manager **--release=7.2** register,  and then attach to openstack entitlements

Please do not run yum update after registration until you installed the version lock outlined below.

Yum Install the version lock package

```
yumdownloader yum-plugin-versionlock-1.1.31-34.el7.noarch
```

You may ignore any Public key messages for now. Just check for the existence of the downloaded file.

Install this rpm.

```
yum localinstall '/whatever-dir/yum-plugin-versionlock-1.1.31-
34.el7.noarch.rpm'

[root@osp8-director ~]# rpm -qa | grep yum-plugin-versionlock
yum-plugin-versionlock-1.1.31-34.el7.noarch
```

Check for existence of versionlock.conf and versionlock.list in /etc/yum/pluginconf.d/versionlock.list.

Update versionlock.conf as shown below:

Uncomment the line follow_obsoletes

follow_obsoletes = 1

Download the versionlock.list file from https://communities.cisco.com/docs/DOC-70256

Download and Extract the zip file cisco-osp8-cvd.zip from the above web page.

The version lock files are in versionlock directory in the zip file.

Copy the downloaded list file to /etc/yum/pluginconf.d/

yum versionlock list command should reveal the contents for /etc/yum/pluginconf.d/versionlock.list.

Run ifconfig to check the health of the configured interfaces. The pxe interface should not have been configured at this stage.

Check name resolution and external connectivity. This is needed for yum updates and registration.

Validate by running wget www.cisco.com or wget subscription.rhn.redhat.com

Install ntp server and synchronize the clock in director node

yum install ntp –y

```
Update /etc/ntp.conf with appropriate ntp server address and restart ntpd
[root@osp8-director ~]# service ntpd restart
Redirecting to /bin/systemctl restart  ntpd.service
Check the time sync, else restart ntpd to force sync the time:
```

```
[root@osp8-director ~]# ntpdate -dv 171.68.38.66
13 Sep 19:32:28 ntpdate[26862]: ntpdate 4.2.6p5@1.2349-o Tue May  3 14:57:04
UTC 2016 (1)
Looking for host 171.68.38.66 and service ntp
host found : mtv5-ai27-dcm10n-ntp2.cisco.com
transmit(171.68.38.66)
receive(171.68.38.66)
............
13 Sep 19:32:35 ntpdate[26862]: adjust time server 171.68.38.66 offset
0.000115 sec
delay 0.02707, dispersion 0.00000
offset 0.000085
```

The clock is synchronized to 115 micro seconds now. Usually a clock sync of less than 20 milli seconds is recommended. In case ntp servers are not accessible from overcloud nodes or director node, you may setup director node as your ntp server. Please refer Linux/Red Hat documentation for making director node as your ntp server.

Refer bug 1178497. This bug is not in the main stream, at the time of writing this document. Please follow the workaround steps in the bug and reboot the kernel.

Take a backup of /boot/initramfs<kernel> to revert back in case something goes wrong:

edit the /usr/lib/dracut/modules.d/99shutdown/module-setup.sh and files
/usr/lib/dracut/modules.d/99shutdown/shutdown.sh after taking a backup of these files.

**module-setup.sh**

Change

```
inst_multiple umount poweroff reboot halt losetup
```

to

```
        inst_multiple umount poweroff reboot halt losetup stat
```

**shutdown.sh**

insert a block of code

after ./lib/dracut-lib.sh

**add:**

```
if [ "$(stat -c '%T' -f /)" = "tmpfs" ]; then
   mount -o remount,rw /
fi
```

**Recreate initramfs** :

```
dracut --force
```

**Unmask the shutdown :**

```
systemctl unmask dracut-shutdown.service
```

**Reboot the node**

This completes the OS Installation on the director node.

# Undercloud Setup

## Undercloud Installation

To install Undercloud, complete the following steps:

Create Stack User:

If the Stack user was not created as part of the install earlier, it has to be created for the Undercloud now.

```
useradd stack
passwd stack
echo "stack ALL=(root) NOPASSWD:ALL" | tee -a /etc/sudoers.d/stack
chmod 0440 /etc/sudoers.d/stack
```

Become the Stack user and create the following:

```
su - stack
mkdir -p ~/images
mkdir -p ~/templates
sudo hostnamectl set-hostname <FQDN of the director node > as an example
sudo hostnamectl set-hostname osp8-director.cisco.com
sudo hostnamectl set-hostname --transient osp8-director.cisco.com
```

Update /etc/hosts:

```
sudo vi /etc/hosts as below
#External Interface
172.22.215.24    osp8-director.cisco.com osp8-director
# local
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
```

Update resolv.conf if needed:

```
sudo vi /etc/resolv.conf as needed. As an example
search cisco.com
nameserver 8.8.8.8
```

It is recommended to use your organization DNS server. name server 8.8.8.8 is used here for reference purpose only.

In case you have not registered the Undercloud node, please register the system to Red Hat Network and get the appropriate pool id for Open stack entitlements and attach the pool.

Disable and enable only the required repositories:

```
sudo subscription-manager repos --disable=*
sudo subscription-manager repos --enable=rhel-7-server-rpms \
--enable=rhel-7-server-extras-rpms \
--enable=rhel-7-server-openstack-8-rpms \
--enable=rhel-7-server-openstack-8-director-rpms \
--enable rhel-7-server-rh-common-rpms
```

Install Undercloud packages:

Make sure that Versionlock is in place by running `yum versionlock list`.

```
sudo yum install -y python-tripleoclient
sudo yum update -y
```

Create undercloud.conf file:

```
cp /usr/share/instack-undercloud/undercloud.conf.sample ~/undercloud.conf
```

The following are the values used in the configuration. 10.23.110 is the pxe network:

```
      image_path = /home/stack/images
      local_ip = 10.23.110.26/24
      network_gateway = 10.23.110.26
      undercloud_public_vip = 10.23.110.27
      undercloud_admin_vip = 10.23.110.28
      local_interface = enp6s0
      network_cidr = 10.23.110.0/24
      masquerade_network = 10.23.110.0/24
      dhcp_start = 10.23.110.51
      dhcp_end = 10.23.110.80
      inspection_interface = br-ctlplane
      inspection_iprange = 10.23.110.81,10.23.110.110
      undercloud_debug = true
```

By using the provisioning interface on the director node and the local_ip and network_gateway, it configures the system to act as the gateway for all the nodes.

Update enic driver:

Download the enic driver Cisco appropriate for the UCSM version. The version used in the configuration with UCSM 2.2(5) and Red Hat Enterprise Linux 7.2 was 2.1.1.93: go to http://software.cisco.com/download/navigator.html

In the download page, select servers-Unified computing under products. On the right menu select your class of servers say Cisco UCS B-series Blade server software and then select Unified Computing System (UCS) Drivers in the following page.

Select your firmware version under All Releases, as an example 2.2(5d) and download the ISO image of UCS-related drivers for your matching firmware, for example **ucs-bxxx-drivers.2.2.5d.iso.**

Download the iso file to your undercloud machine and mount the iso:

```
root@osp8-director ~]# mount ucs-bxxx-drivers.2.2.5d.iso /mnt
mount: /dev/loop0 is write-protected, mounting read-only
cd /mnt/Linux/Network/Cisco/VIC/RHEL/RHEL7.2
cp kmod-enic-2.1.1.93-rhel7u2.el7.x86_64.rpm /tmp

umount /mnt
```

Install the appropriate enic driver on the director machine.

rpm –ivh /tmp/kmod-enic-2.1.1.93-rhel7u2.el7.x86_64.rpm

Validate by running modinfo:

```
[root@osp8-director ~]# modinfo enic
```

```
filename:       /lib/modules/3.10.0-327.28.3.el7.x86_64/weak-
updates/enic/enic.ko
version:        2.1.1.93
license:        GPL v2
author:         Scott Feldman <scofeldm@cisco.com>
description:    Cisco VIC Ethernet NIC Driver
rhelversion:    7.2
srcversion:     D272F11F27065C9714656F4
alias:          pci:v00001137d00000071sv*sd*bc*sc*i*
alias:          pci:v00001137d00000044sv*sd*bc*sc*i*
alias:          pci:v00001137d00000043sv*sd*bc*sc*i*
depends:
vermagic:       3.10.0-327.el7.x86_64 SMP mod_unload modversions
parm:           rxcopybreak:Maximum size of packet that is copied to a new
buffer on receive (uint)
```

Copy the enic file to your ~/images directory created above.

```
cp /lib/modules/3.10.0-327.28.3.el7.x86_64/weak-updates/enic/enic.ko ~/images
[root@osp8-director images]# ls -l enic.ko
-rw-r--r--. 1 stack stack 3982019 Aug 29 05:55 enic.ko
```

This enic.ko file will be used to customize Overcloud image file later.

Download libguestfs tool needed to customize Overcloud image file:

```
sudo yum install libguestfs-tools -y
```

The system is ready to run the Undercloud installation.

Run the following as stack user:

```
cd /home/stack
openstack undercloud install
```

This might take around 10-15 minutes.

To debug any Undercloud install failures, check files in /home/stack/.instack/*

## Post Undercloud Installation Checks

To perform the Undercloud installation checks, complete the following steps:

Check /etc/resolv.conf

[root@osp8-director ~]# cat /etc/resolv.conf

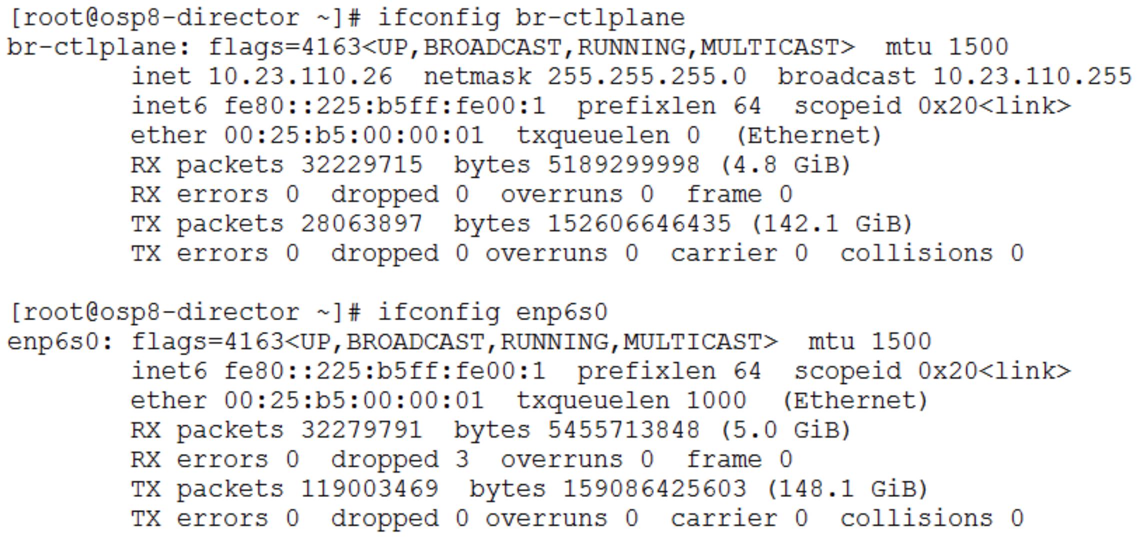
# Generated by NetworkManager

search cisco.com

nameserver 8.8.8.8

Check control plane bridge:

A new bridge br-ctlplane should have been created as part of the Undercloud install on the pxe interface as shown below. Validate MAC and IP’s.



```
[root@osp8-director ~]# ifconfig br-ctlplane
br-ctlplane: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 10.23.110.26  netmask 255.255.255.0  broadcast 10.23.110.255
        inet6 fe80::225:b5ff:fe00:1  prefixlen 64  scopeid 0x20<link>
        ether 00:25:b5:00:00:01  txqueuelen 0  (Ethernet)
        RX packets 32229715  bytes 5189299998 (4.8 GiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 28063897  bytes 152606646435 (142.1 GiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

[root@osp8-director ~]# ifconfig enp6s0
enp6s0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet6 fe80::225:b5ff:fe00:1  prefixlen 64  scopeid 0x20<link>
        ether 00:25:b5:00:00:01  txqueuelen 1000  (Ethernet)
        RX packets 32279791  bytes 5455713848 (5.0 GiB)
        RX errors 0  dropped 3  overruns 0  frame 0
        TX packets 119003469  bytes 159086425603 (148.1 GiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

Check /var/log/ironic/* files, to understand and fix any issues.

# Introspection

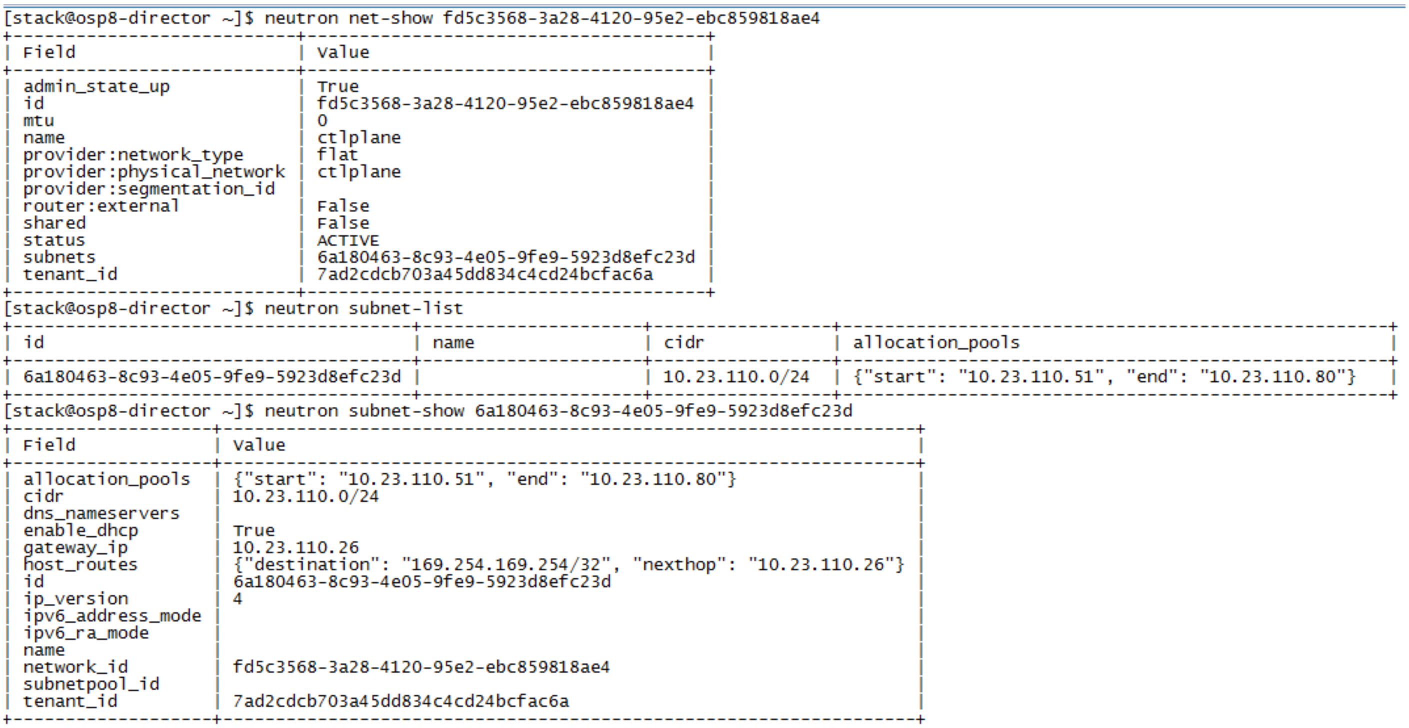
## Pre-Installation Checks for Introspection

To do the pre-installation check, complete the following steps:

Log into the director node as the stack user and source stackrc file.

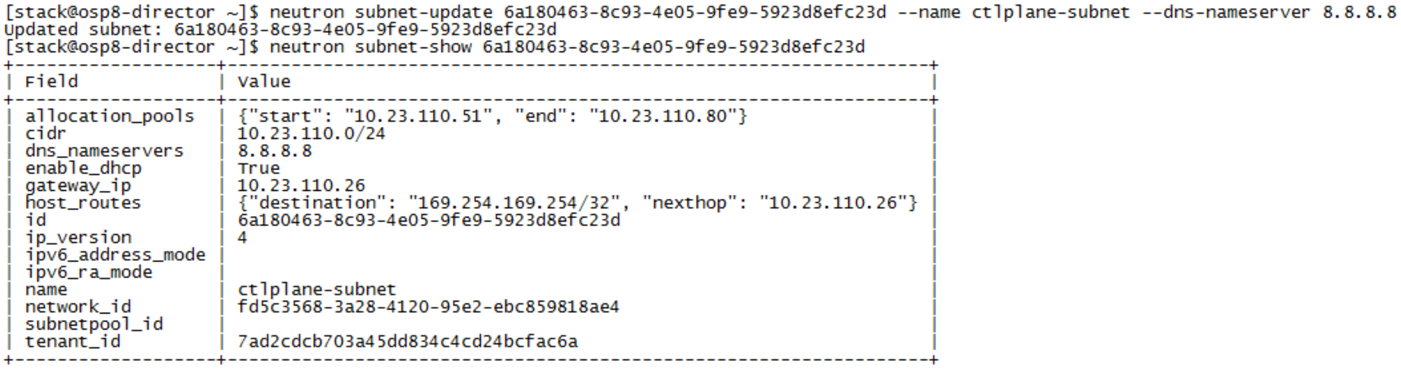
Check neutron and subnet lists:

Source stackrc as stack user and run neutron net-list, neutron net-show, neutron subnet-list and neutron subnet-show for br-ctlplane:



```
[stack@osp8-director ~]$ neutron net-show fd5c3568-3a28-4120-95e2-ebc859818ae4
+---------------------------+--------------------------------------+
| Field                     | Value                                |
+---------------------------+--------------------------------------+
| admin_state_up            | True                                 |
| id                        | fd5c3568-3a28-4120-95e2-ebc859818ae4 |
| mtu                       | 0                                    |
| name                      | ctlplane                             |
| provider:network_type     | flat                                 |
| provider:physical_network | ctlplane                             |
| provider:segmentation_id  |                                      |
| router:external           | False                                |
| shared                    | False                                |
| status                    | ACTIVE                               |
| subnets                   | 6a180463-8c93-4e05-9fe9-5923d8efc23d |
| tenant_id                 | 7ad2cdcb703a45dd834c4cd24bcfac6a     |
+---------------------------+--------------------------------------+
[stack@osp8-director ~]$ neutron subnet-list
+--------------------------------------+----------------------+-----------------+--------------------------------------------------+
| id                                   | name                 | cidr            | allocation_pools                                 |
+--------------------------------------+----------------------+-----------------+--------------------------------------------------+
| 6a180463-8c93-4e05-9fe9-5923d8efc23d |                      | 10.23.110.0/24  | {"start": "10.23.110.51", "end": "10.23.110.80"} |
+--------------------------------------+----------------------+-----------------+--------------------------------------------------+
[stack@osp8-director ~]$ neutron subnet-show 6a180463-8c93-4e05-9fe9-5923d8efc23d
+-------------------+-------------------------------------------------------------------+
| Field             | Value                                                             |
+-------------------+-------------------------------------------------------------------+
| allocation_pools  | {"start": "10.23.110.51", "end": "10.23.110.80"}                  |
| cidr              | 10.23.110.0/24                                                    |
| dns_nameservers   |                                                                   |
| enable_dhcp       | True                                                              |
| gateway_ip        | 10.23.110.26                                                      |
| host_routes       | {"destination": "169.254.169.254/32", "nexthop": "10.23.110.26"} |
| id                | 6a180463-8c93-4e05-9fe9-5923d8efc23d                              |
| ip_version        | 4                                                                 |
| ipv6_address_mode |                                                                   |
| ipv6_ra_mode      |                                                                   |
| name              |                                                                   |
| network_id        | fd5c3568-3a28-4120-95e2-ebc859818ae4                              |
| subnetpool_id     |                                                                   |
| tenant_id         | 7ad2cdcb703a45dd834c4cd24bcfac6a                                  |
+-------------------+-------------------------------------------------------------------+
```

The allocation_pools, dns_nameservers, cidr should match whatever specified earlier in undercloud.conf file. If not, update with neutron subnet-update.



```
[stack@osp8-director ~]$ neutron subnet-update 6a180463-8c93-4e05-9fe9-5923d8efc23d --name ctlplane-subnet --dns-nameserver 8.8.8.8
Updated subnet: 6a180463-8c93-4e05-9fe9-5923d8efc23d
[stack@osp8-director ~]$ neutron subnet-show 6a180463-8c93-4e05-9fe9-5923d8efc23d
+-------------------+-------------------------------------------------------------------+
| Field             | Value                                                             |
+-------------------+-------------------------------------------------------------------+
| allocation_pools  | {"start": "10.23.110.51", "end": "10.23.110.80"}                  |
| cidr              | 10.23.110.0/24                                                    |
| dns_nameservers   | 8.8.8.8                                                           |
| enable_dhcp       | True                                                              |
| gateway_ip        | 10.23.110.26                                                      |
| host_routes       | {"destination": "169.254.169.254/32", "nexthop": "10.23.110.26"} |
| id                | 6a180463-8c93-4e05-9fe9-5923d8efc23d                              |
| ip_version        | 4                                                                 |
| ipv6_address_mode |                                                                   |
| ipv6_ra_mode      |                                                                   |
| name              | ctlplane-subnet                                                   |
| network_id        | fd5c3568-3a28-4120-95e2-ebc859818ae4                              |
| subnetpool_id     |                                                                   |
| tenant_id         | 7ad2cdcb703a45dd834c4cd24bcfac6a                                  |
+-------------------+-------------------------------------------------------------------+
```

Check /etc/ironic-inspector/* files:

```
vi /etc/ironic-inspector/inspector.conf /etc/ironic-inspector/dnsmasq.conf
```

The dnsmasq.conf dhcp_range should match the undercloud.conf file range. This will help you spot any errors that might have gone while running Undercloud install earlier. The default pxe timeout is 60 minutes. This means if you have more servers to be introspected and it takes longer than 60 minutes, introspection is bound to fail.

Update /etc/ironic-inspector/inspector.conf with timeout variable under discovered section:

```
timeout=0
```

Restart ironic in case these files are updated.

```
[root@osp8-director ~]# systemctl restart openstack-ironic-conductor.service
openstack-ironic-inspector-dnsmasq.service openstack-ironic-api.service
openstack-ironic-inspector.service
```

This may be necessary only in larger deployments and depends on the network to download the ramdisk files, CPU speed, etc.

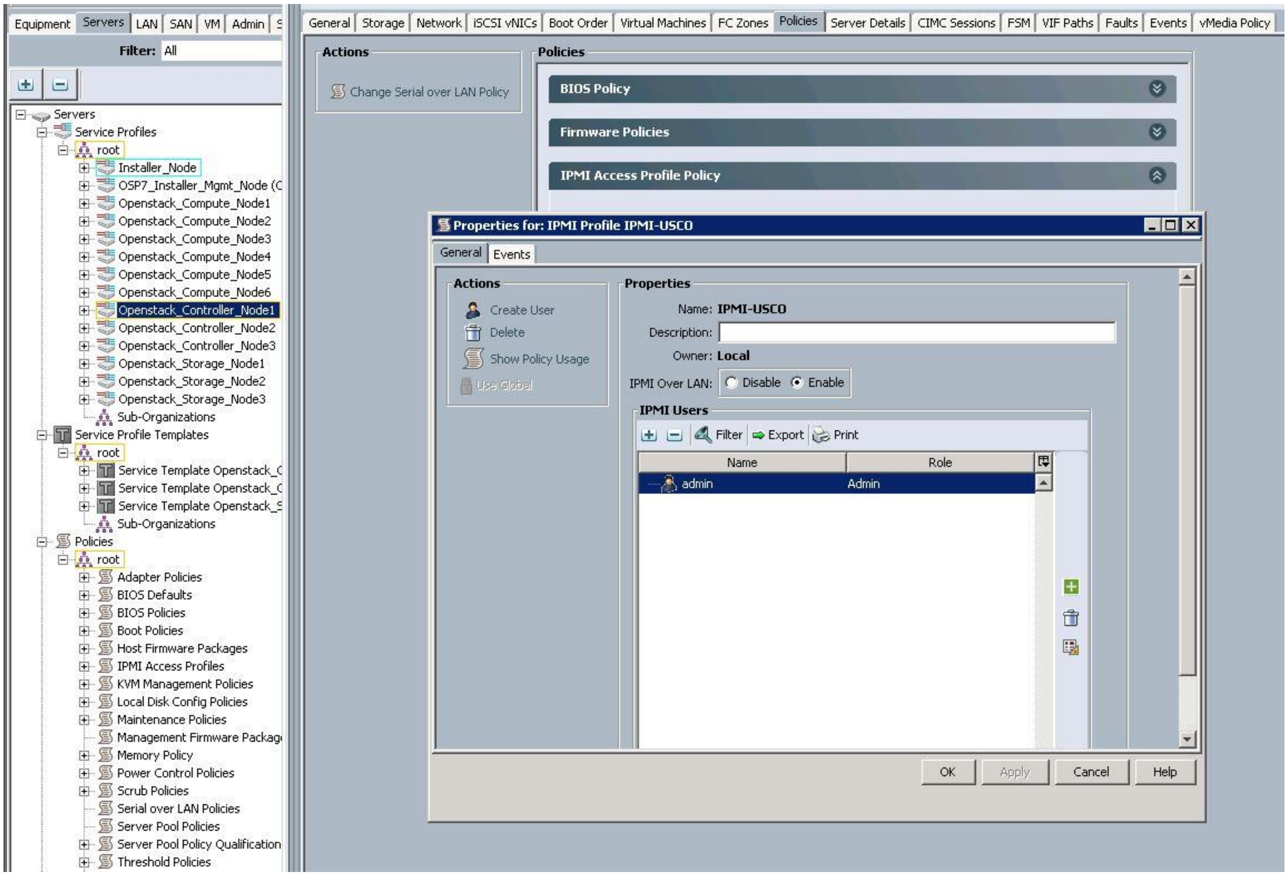
Prepare instack.json file.

This file should contain all the nodes, controllers, computes and storage nodes that need to be introspected.

A sample instackenv.json file is provided in Appendix. Below is an explanation of how to build this file for a node.

```
  [stack@osp8-director ~]$ cat instackenv.json
{
  "nodes": [
    {
      "pm_user": "<ipmi admin user>",
      "pm_password": "<password>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.57",
      "mac": [
        "00:25:b5:00:00:08"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "32"
    },
……………………
…………………… .
```

pm_user and pm_password is the ipmi user and password configured earlier for this node’s service profile or templates.



Properties for: IPMI Profile IPMI-USCO

Properties

Name: IPMI-USCO

Description:

Owner: Local

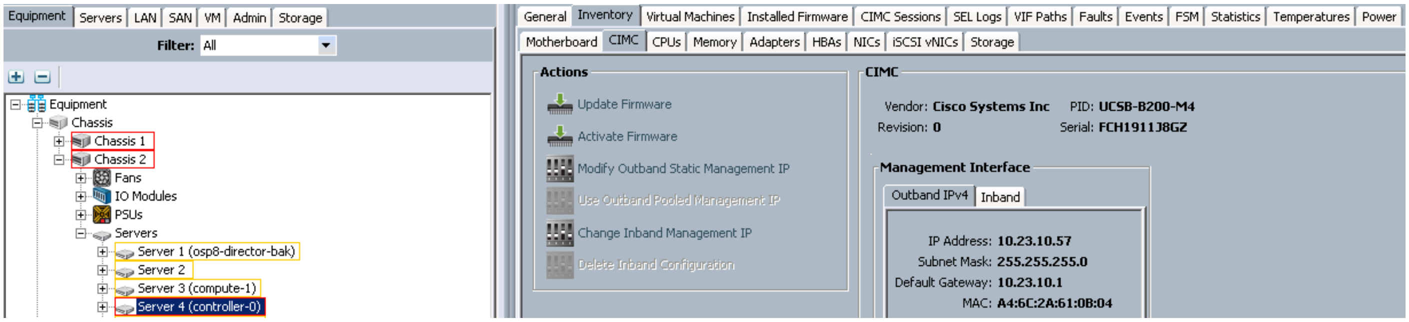
IPMI Over LAN: ○ Disable ● Enable

IPMI Users

| Name | Role |
|---|---|
| admin | Admin |

```
pm_type=”pxe_ipmitool” Leave this, as is
```

pm_addr is the IPMI address allocated to that node. This can be obtained from the CIMC tab in equipment.



CIMC

Vendor: Cisco Systems Inc PID: UCSB-B200-M4

Revision: 0 Serial: FCH1911J8GZ

Management Interface

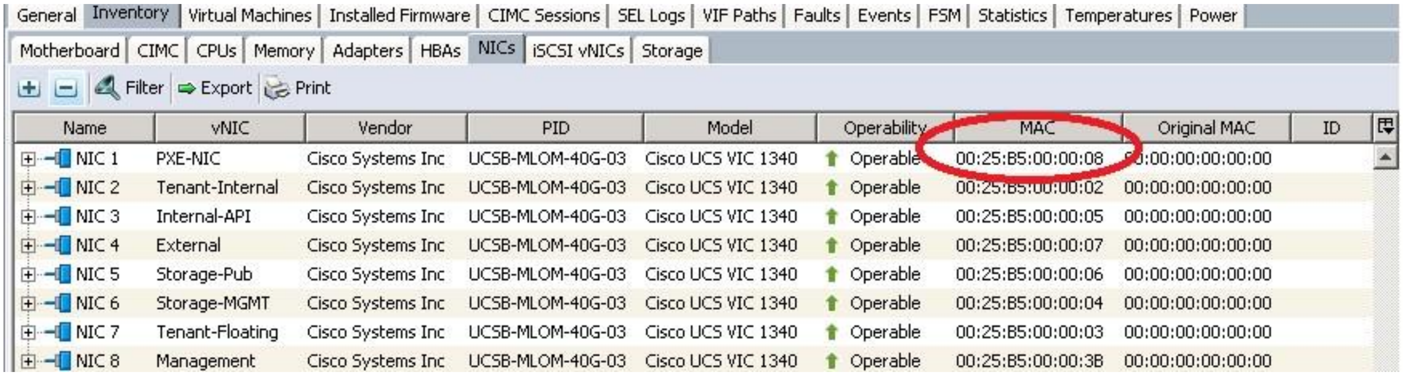
IP Address: 10.23.10.57

Subnet Mask: 255.255.255.0

Default Gateway: 10.23.10.1

MAC: A4:6C:2A:61:0B:04

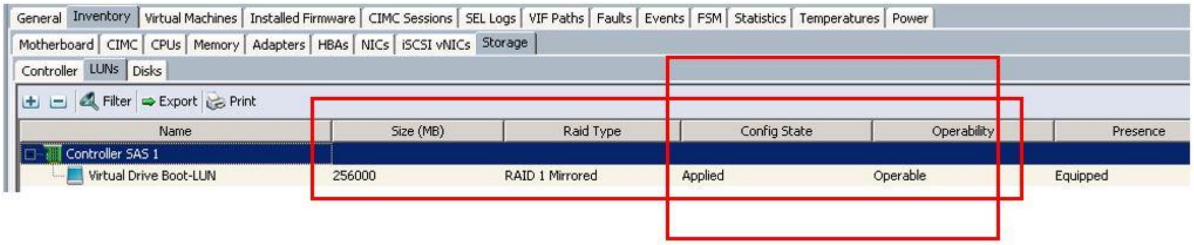
The MAC address is the discovery NIC or pxe interface for that node.



| Name | vNIC | Vendor | PID | Model | Operability | MAC | Original MAC | ID |
|---|---|---|---|---|---|---|---|---|
| NIC 1 | PXE-NIC | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:08 | 00:00:00:00:00:00 | |
| NIC 2 | Tenant-Internal | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:02 | 00:00:00:00:00:00 | |
| NIC 3 | Internal-API | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:05 | 00:00:00:00:00:00 | |
| NIC 4 | External | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:07 | 00:00:00:00:00:00 | |
| NIC 5 | Storage-Pub | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:06 | 00:00:00:00:00:00 | |
| NIC 6 | Storage-MGMT | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:04 | 00:00:00:00:00:00 | |
| NIC 7 | Tenant-Floating | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:03 | 00:00:00:00:00:00 | |
| NIC 8 | Management | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:3B | 00:00:00:00:00:00 | |

The memory, disk, and CPU can be obtained under the same Inventory tab for that node.

Make sure that the storage lun is applied and in operable state after applying the storage policy.



| Name | Size (MB) | Raid Type | Config State | Operability | Presence |
|---|---|---|---|---|---|
| Controller SAS 1 | | | | | |
| Virtual Drive Boot-LUN | 256000 | RAID 1 Mirrored | Applied | Operable | Equipped |

Build the instackenv.json file for all the hosts that have to be introspected as above.

Make sure to maintain consistent indentations with white spaces or tabs.

Check the ipmi connectivity works for all the hosts:

You can run a quick check to validate this from instackenv.json file:

```
[stack@osp8-director ~]$ for i in `grep pm_addr instackenv.json | cut -d "\"" -
f4`
do
ipmitool -I lanplus -H $i -U <ipmi admin user> -P <replace with your ipmi
password> chassis power status
done
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
Chassis Power is off
```

The chassis power status should be either On or Off depending on whether the server is up or down in UCS. However any errors like the example shown below need investigation:

```
Error: Unable to establish IPMI v2 / RMCP+ session
Unable to get Chassis Power Status
```

Make lvm changes for bug 1323024

```
Update /etc/lvm/lvm.conf activation function
# Configuration section activation.
activation {
......
......
#Add the below line where rhel/home comes from vg/lv display or /etc/fstab
......
auto_activation_volume_list = ["rhel/home"]
}
```

Reboot the node after making the above change.

Download the Image files needed for introspection and Overcloud:

```
sudo yum install rhosp-director-images rhosp-director-images-ipa
cp /usr/share/rhosp-director-images/overcloud-full-latest-8.0.tar ~/images/.
cp /usr/share/rhosp-director-images/ironic-python-agent-latest-8.0.tar
~/images/.
cd ~/images
for tarfile in *.tar; do tar -xf $tarfile; done

Download the KVM Guest Image to the directory from access.redhat.com

[stack@osp8-director images]$ /bin/ls -1
enic.ko
ironic-python-agent-latest-8.0.tar
ironic-python-agent.initramfs
ironic-python-agent.kernel
overcloud-full-latest-8.0.tar
overcloud-full.initrd
overcloud-full.qcow2
overcloud-full.vmlinuz
```

You may remove the tar files if desired.

```
rhel-guest-image-7.2-20151102.0.x86_64.qcow2
```

Customize the Overcloud image with enic drivers and fencing packages.

Run the following as root user. Navigate to your download directory and issue the following as root:

```
cd /home/stack/images
export LIBGUESTFS_BACKEND=direct
```

Update fencing packages.

Refer bug 1298430.

```
Download the fencing packages from Red Hat web site.
   sudo yumdownloader fence-agents-cisco-ucs-4.0.11-27.el7_2.9 fence-agents-
   common-4.0.11-27.el7_2.9 fence-agents-scsi-4.0.11-27.el7_2.9

Upload the downloaded files to overcloud image.
   for i in *.rpm
   do
   virt-customize -a overcloud-full.qcow2 --upload  $i:/root
   done

Validate that the packages do exist in /root

   virt-ls -a overcloud-full.qcow2 /root | grep rpm

         fence-agents-cisco-ucs-4.0.11-27.el7_2.9.x86_64.rpm
         fence-agents-common-4.0.11-27.el7_2.9.x86_64.rpm
         fence-agents-scsi-4.0.11-27.el7_2.9.x86_64.rpm
```

Install these packages in the overcloud image file

```
[root@osp8-director images]# virt-customize -a overcloud-full.qcow2 --run-
command 'yum localinstall -y /root/fence-agents-common-4.0.11-
27.el7_2.9.x86_64.rpm /root/fence-agents-cisco-ucs-4.0.11-
27.el7_2.9.x86_64.rpm /root/fence-agents-scsi-4.0.11-27.el7_2.9.x86_64.rpm'

      [   0.0] Examining the guest ...
      [   4.0] Setting a random seed
      [   4.0] Running: yum localinstall -y /root/fence-agents-common-
      4.0.11-27.el7_2.9.x86_64.rpm /root/fence-agents-cisco-ucs-4.0.11-
      27.el7_2.9.x86_64.rpm /root/fence-agents-scsi-4.0.11-
      27.el7_2.9.x86_64.rpm
      [  12.0] Finishing off
```

Update Grub file;

```
virt-copy-out -a overcloud-full.qcow2 /etc/default/grub /home/stack/images/

vi grub file and change the following line
GRUB_CMDLINE_LINUX="console=tty0 console=ttyS0,115200n8 crashkernel=auto rhgb
quiet net.ifnames=0 biosdevname=0"
---(you are appending net.ifnames=0 and biosdevname=0)
virt-copy-in –a overcloud-full.qcow2 ./grub /etc/default/
```

After this proceed, with the remaining customizations.

Update enic drivers;

```
[root@osp8-director images]# virt-ls -R -a overcloud-full.qcow2 /lib/modules |
grep enic # Get the directory where enic exists
/3.10.0-327.18.2.el7.x86_64/kernel/drivers/net/ethernet/cisco/enic
/3.10.0-327.18.2.el7.x86_64/kernel/drivers/net/ethernet/cisco/enic/enic.ko

enic driver exists in /lib/modules/3.10.0-
327.18.2.el7.x86_64/kernel/drivers/net/ethernet/cisco
```

Copy the enic driver to the above location

```
virt-copy-in -a overcloud-full.qcow2 ./enic.ko /lib/modules/3.10.0-
327.18.2.el7.x86_64/kernel/drivers/net/ethernet/cisco/enic/
```

The location of this enic driver is dependent on the kernel packaged in the Overcloud image file. Should be changed if needed.

Update root password;

```
virt-customize -a overcloud-full.qcow2 --root-password password:<password>
```

Change the permissions back to stack user:

```
chown stack:stack /home/stack/images/*
```

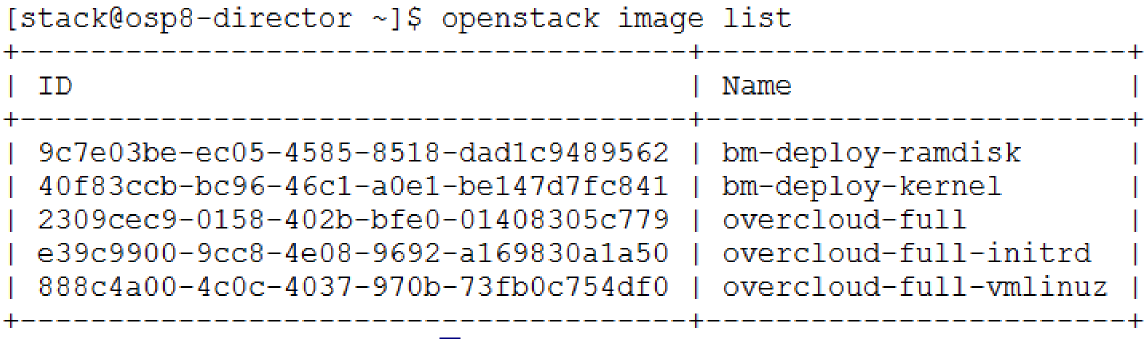
This is how the overcloud-full.qcow2 may look after the update:

```
[root@osp8-director images]# ls -l overcloud-full.qcow2
-rw-r--r--. 1 stack stack 1096679424 Oct 19 14:57 overcloud-full.qcow2
```

While updating, the image with root password is not required; it becomes useful to login through KVM console in case of Overcloud installation failures and debug the issues.
The enic.ko was extracted earlier on the director node after installing the enic rpm. This helps ensure that both director and the Overcloud images will be with same enic driver.
The grub has been modified to have interface names like eth[0], eth[1] …
The fence_cisco_ucs package has been modified to take care of the HA bug 1298430.

Upload the images to openstack. As stack user run the following:

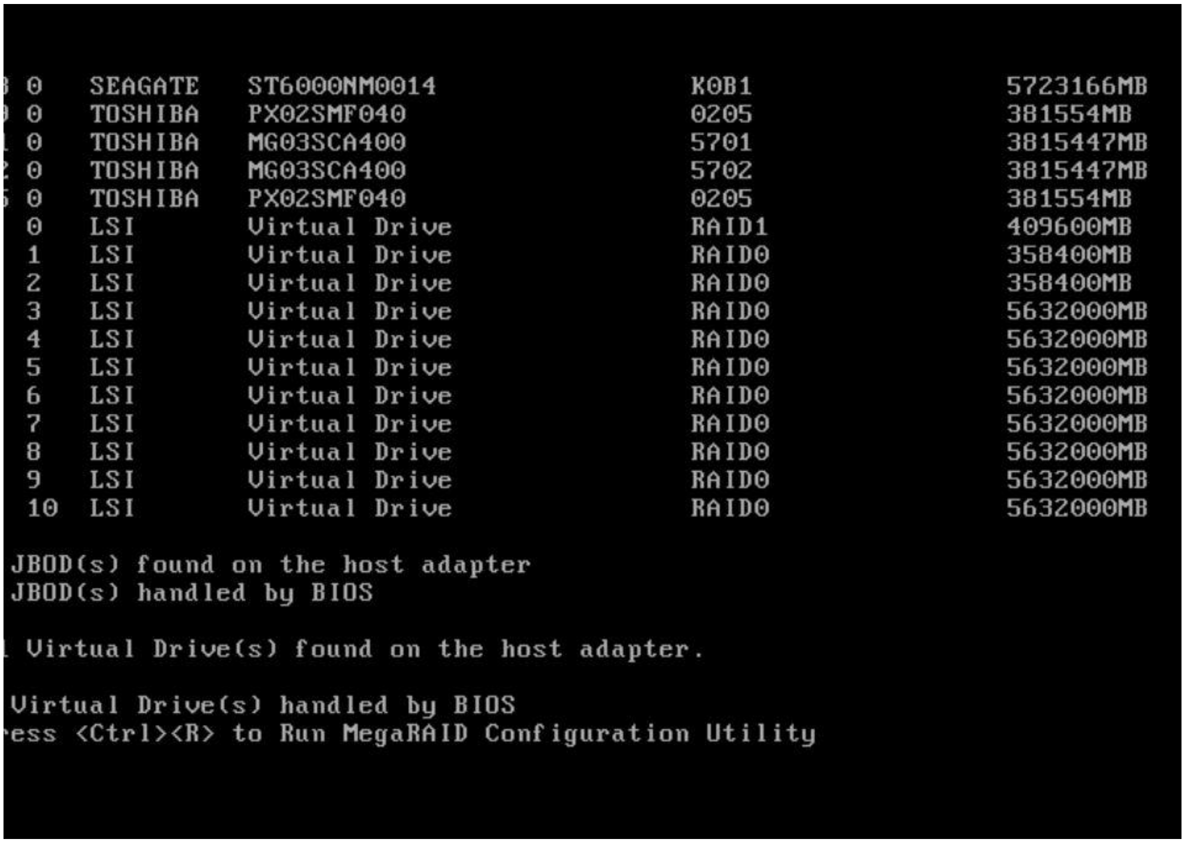
```
su - stack
source stackrc
cd ~/images
openstack overcloud image upload --image-path /home/stack/images/
openstack image list
```



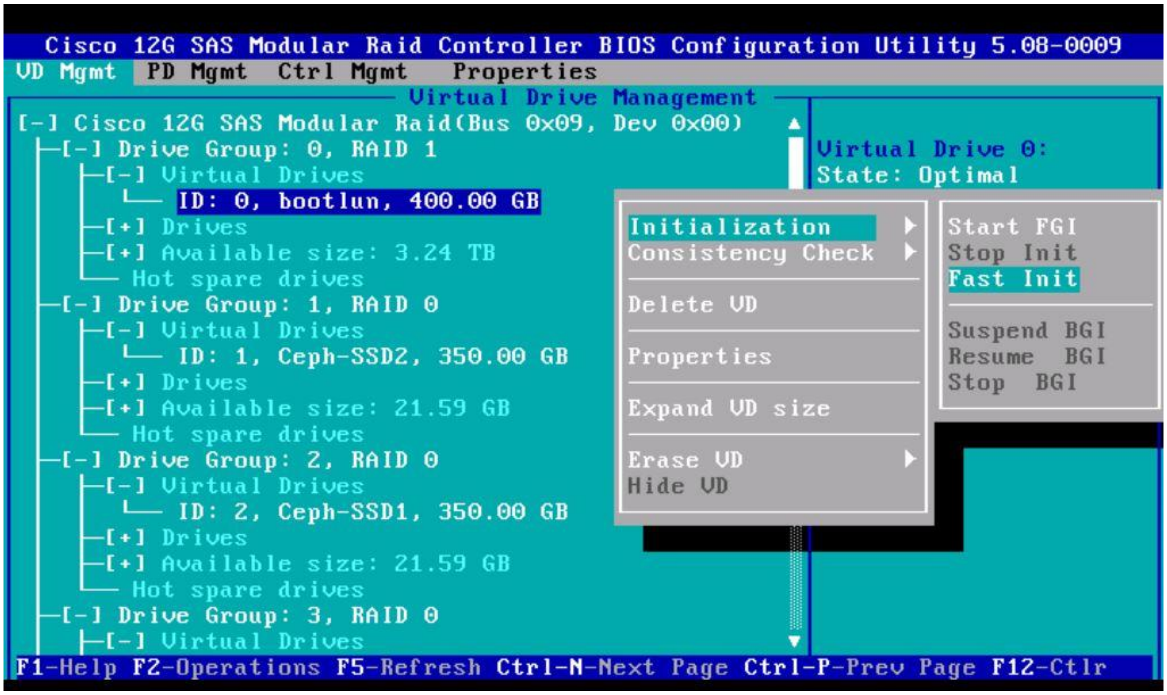
```
[stack@osp8-director ~]$ openstack image list
+--------------------------------------+------------------------+
| ID                                   | Name                   |
+--------------------------------------+------------------------+
| 9c7e03be-ec05-4585-8518-dad1c9489562 | bm-deploy-ramdisk      |
| 40f83ccb-bc96-46c1-a0e1-be147d7fc841 | bm-deploy-kernel       |
| 2309cec9-0158-402b-bfe0-01408305c779 | overcloud-full         |
| e39c9900-9cc8-4e08-9692-a169830a1a50 | overcloud-full-initrd  |
| 888c4a00-4c0c-4037-970b-73fb0c754df0 | overcloud-full-vmlinuz |
+--------------------------------------+------------------------+
```

Before running Introspection and Overcloud installation, it is recommended to initialize the boot LUNs. This is required in case you are repeating or using old disks.

Boot the server in UCS, press CTRL-R, then F2 and re-initialize the boot LUNs as shown below and then power off the servers.



```
 0   SEAGATE   ST6000NM0014                  K0B1          5723166MB
 0   TOSHIBA   PX02SMF040                    0205          381554MB
 0   TOSHIBA   MG03SCA400                    5701          3815447MB
 0   TOSHIBA   MG03SCA400                    5702          3815447MB
 0   TOSHIBA   PX02SMF040                    0205          381554MB
 0   LSI       Virtual Drive                 RAID1         409600MB
 1   LSI       Virtual Drive                 RAID0         358400MB
 2   LSI       Virtual Drive                 RAID0         358400MB
 3   LSI       Virtual Drive                 RAID0         5632000MB
 4   LSI       Virtual Drive                 RAID0         5632000MB
 5   LSI       Virtual Drive                 RAID0         5632000MB
 6   LSI       Virtual Drive                 RAID0         5632000MB
 7   LSI       Virtual Drive                 RAID0         5632000MB
 8   LSI       Virtual Drive                 RAID0         5632000MB
 9   LSI       Virtual Drive                 RAID0         5632000MB
 10  LSI       Virtual Drive                 RAID0         5632000MB

JBOD(s) found on the host adapter
JBOD(s) handled by BIOS

 Virtual Drive(s) found on the host adapter.

Virtual Drive(s) handled by BIOS
ess <Ctrl><R> to Run MegaRAID Configuration Utility
```



```
 Cisco 12G SAS Modular Raid Controller BIOS Configuration Utility 5.08-0009
VD Mgmt  PD Mgmt  Ctrl Mgmt   Properties
                         Virtual Drive Management
[-] Cisco 12G SAS Modular Raid(Bus 0x09, Dev 0x00)
  ├─[-] Drive Group: 0, RAID 1                                Virtual Drive 0:
  │   ├─[-] Virtual Drives                                    State: Optimal
  │   │   └── ID: 0, bootlun, 400.00 GB
  │   ├─[+] Drives                         Initialization     ►  Start FGI
  │   ├─[+] Available size: 3.24 TB        Consistency Check  ►  Stop Init
  │   └── Hot spare drives                                       Fast Init
  ├─[-] Drive Group: 1, RAID 0             Delete VD
  │   ├─[-] Virtual Drives                                       Suspend BGI
  │   │   └── ID: 1, Ceph-SSD2, 350.00 GB  Properties            Resume  BGI
  │   ├─[+] Drives                                               Stop  BGI
  │   ├─[+] Available size: 21.59 GB       Expand VD size
  │   └── Hot spare drives
  ├─[-] Drive Group: 2, RAID 0             Erase VD           ►
  │   ├─[-] Virtual Drives                 Hide VD
  │   │   └── ID: 2, Ceph-SSD1, 350.00 GB
  │   ├─[+] Drives
  │   ├─[+] Available size: 21.59 GB
  │   └── Hot spare drives
  ├─[-] Drive Group: 3, RAID 0
  │   ├─[-] Virtual Drives
F1-Help F2-Operations F5-Refresh Ctrl-N-Next Page Ctrl-P-Prev Page F12-Ctlr
```

Make sure that all the servers are powered off before introspection.

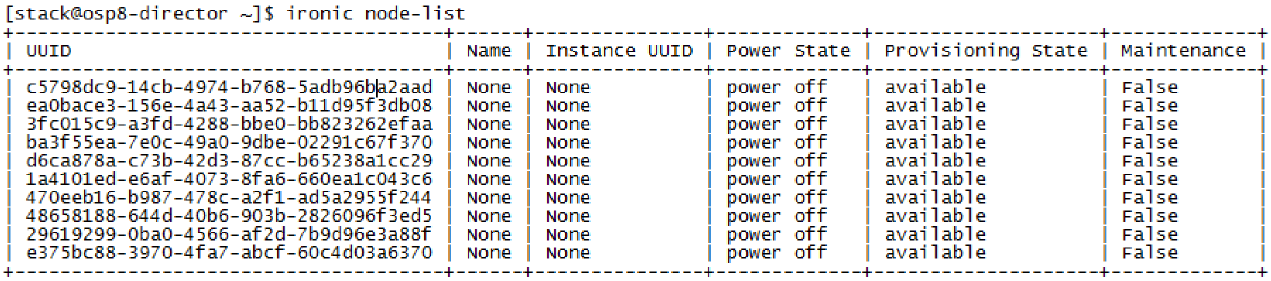
Reboot the Undercloud node and start the introspection.

## Run Introspection

To run Introspection, complete the following steps:

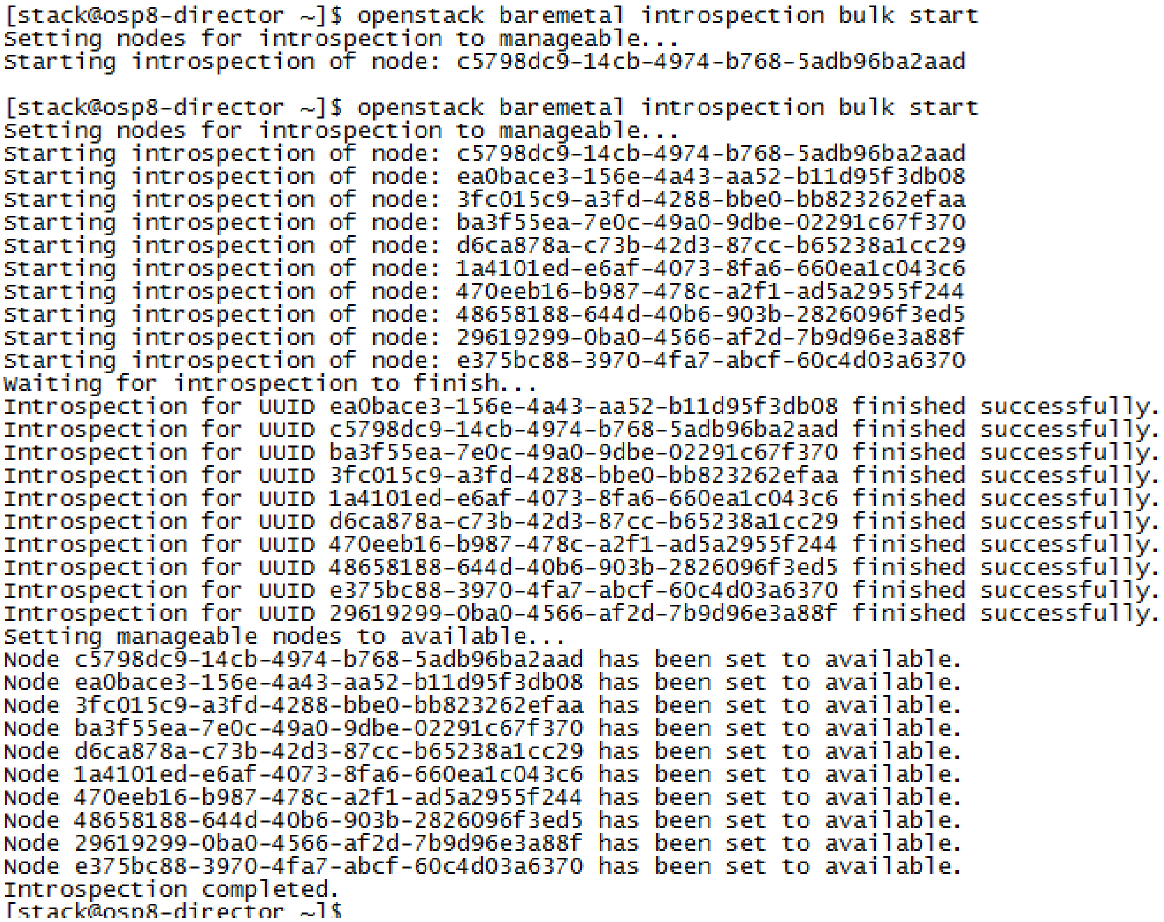
As stack user:

```
source ~/stackrc
openstack baremetal import --json ~/instackenv.json
openstack baremetal configure boot
ironic node-list
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| c5798dc9-14cb-4974-b768-5adb96ba2aad | None | None          | power off   | available          | False       |
| ea0bace3-156e-4a43-aa52-b11d95f3db08 | None | None          | power off   | available          | False       |
| 3fc015c9-a3fd-4288-bbe0-bb823262efaa | None | None          | power off   | available          | False       |
| ba3f55ea-7e0c-49a0-9dbe-02291c67f370 | None | None          | power off   | available          | False       |
| d6ca878a-c73b-42d3-87cc-b65238a1cc29 | None | None          | power off   | available          | False       |
| 1a4101ed-e6af-4073-8fa6-660ea1c043c6 | None | None          | power off   | available          | False       |
| 470eeb16-b987-478c-a2f1-ad5a2955f244 | None | None          | power off   | available          | False       |
| 48658188-644d-40b6-903b-2826096f3ed5 | None | None          | power off   | available          | False       |
| 29619299-0ba0-4566-af2d-7b9d96e3a88f | None | None          | power off   | available          | False       |
| e375bc88-3970-4fa7-abcf-60c4d03a6370 | None | None          | power off   | available          | False       |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
```

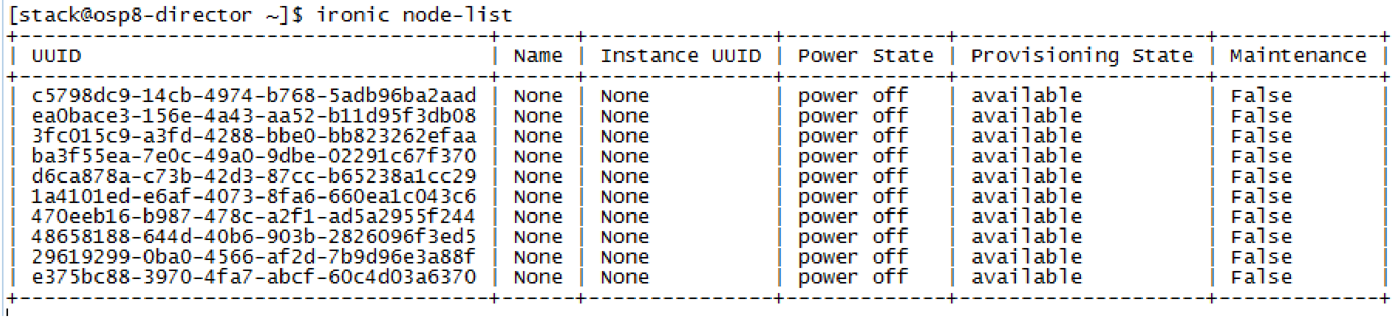
```
openstack baremetal introspection bulk start
```



```
[stack@osp8-director ~]$ openstack baremetal introspection bulk start
Setting nodes for introspection to manageable...
Starting introspection of node: c5798dc9-14cb-4974-b768-5adb96ba2aad

[stack@osp8-director ~]$ openstack baremetal introspection bulk start
Setting nodes for introspection to manageable...
Starting introspection of node: c5798dc9-14cb-4974-b768-5adb96ba2aad
Starting introspection of node: ea0bace3-156e-4a43-aa52-b11d95f3db08
Starting introspection of node: 3fc015c9-a3fd-4288-bbe0-bb823262efaa
Starting introspection of node: ba3f55ea-7e0c-49a0-9dbe-02291c67f370
Starting introspection of node: d6ca878a-c73b-42d3-87cc-b65238a1cc29
Starting introspection of node: 1a4101ed-e6af-4073-8fa6-660ea1c043c6
Starting introspection of node: 470eeb16-b987-478c-a2f1-ad5a2955f244
Starting introspection of node: 48658188-644d-40b6-903b-2826096f3ed5
Starting introspection of node: 29619299-0ba0-4566-af2d-7b9d96e3a88f
Starting introspection of node: e375bc88-3970-4fa7-abcf-60c4d03a6370
Waiting for introspection to finish...
Introspection for UUID ea0bace3-156e-4a43-aa52-b11d95f3db08 finished successfully.
Introspection for UUID c5798dc9-14cb-4974-b768-5adb96ba2aad finished successfully.
Introspection for UUID ba3f55ea-7e0c-49a0-9dbe-02291c67f370 finished successfully.
Introspection for UUID 3fc015c9-a3fd-4288-bbe0-bb823262efaa finished successfully.
Introspection for UUID 1a4101ed-e6af-4073-8fa6-660ea1c043c6 finished successfully.
Introspection for UUID d6ca878a-c73b-42d3-87cc-b65238a1cc29 finished successfully.
Introspection for UUID 470eeb16-b987-478c-a2f1-ad5a2955f244 finished successfully.
Introspection for UUID 48658188-644d-40b6-903b-2826096f3ed5 finished successfully.
Introspection for UUID e375bc88-3970-4fa7-abcf-60c4d03a6370 finished successfully.
Introspection for UUID 29619299-0ba0-4566-af2d-7b9d96e3a88f finished successfully.
Setting manageable nodes to available...
Node c5798dc9-14cb-4974-b768-5adb96ba2aad has been set to available.
Node ea0bace3-156e-4a43-aa52-b11d95f3db08 has been set to available.
Node 3fc015c9-a3fd-4288-bbe0-bb823262efaa has been set to available.
Node ba3f55ea-7e0c-49a0-9dbe-02291c67f370 has been set to available.
Node d6ca878a-c73b-42d3-87cc-b65238a1cc29 has been set to available.
Node 1a4101ed-e6af-4073-8fa6-660ea1c043c6 has been set to available.
Node 470eeb16-b987-478c-a2f1-ad5a2955f244 has been set to available.
Node 48658188-644d-40b6-903b-2826096f3ed5 has been set to available.
Node 29619299-0ba0-4566-af2d-7b9d96e3a88f has been set to available.
Node e375bc88-3970-4fa7-abcf-60c4d03a6370 has been set to available.
Introspection completed.
[stack@osp8-director ~]$
```

Check the status of Introspection:

```
    openstack baremetal introspection bulk status
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| c5798dc9-14cb-4974-b768-5adb96ba2aad | None | None          | power off   | available          | False       |
| ea0bace3-156e-4a43-aa52-b11d95f3db08 | None | None          | power off   | available          | False       |
| 3fc015c9-a3fd-4288-bbe0-bb823262efaa | None | None          | power off   | available          | False       |
| ba3f55ea-7e0c-49a0-9dbe-02291c67f370 | None | None          | power off   | available          | False       |
| d6ca878a-c73b-42d3-87cc-b65238a1cc29 | None | None          | power off   | available          | False       |
| 1a4101ed-e6af-4073-8fa6-660ea1c043c6 | None | None          | power off   | available          | False       |
| 470eeb16-b987-478c-a2f1-ad5a2955f244 | None | None          | power off   | available          | False       |
| 48658188-644d-40b6-903b-2826096f3ed5 | None | None          | power off   | available          | False       |
| 29619299-0ba0-4566-af2d-7b9d96e3a88f | None | None          | power off   | available          | False       |
| e375bc88-3970-4fa7-abcf-60c4d03a6370 | None | None          | power off   | available          | False       |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
```

Refer to the Troubleshooting section for any failures around introspection and how to resolve them.

## Set Flavors

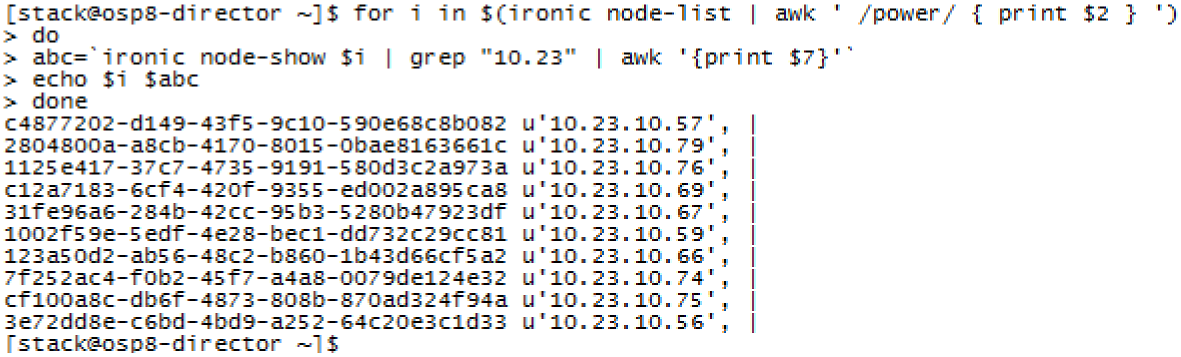
Red Hat OpenStack Platform 8 comes with pre-created flavors that can be queried as follows:

```
[stack@osp8-director ~]$ openstack flavor list
+--------------------------------------+---------------+------+------+-----------+-------+-----------+
| ID                                   | Name          |  RAM | Disk | Ephemeral | VCPUs | Is Public |
+--------------------------------------+---------------+------+------+-----------+-------+-----------+
| 41b37759-0377-4853-af78-42a725d82980 | baremetal     | 4096 |   40 |         0 |     1 | True      |
| 4667ea01-fc7c-4011-9fcb-17be56f0458e | compute       | 4096 |   40 |         0 |     1 | True      |
| 9d6b80f1-04e9-416d-97f5-cf5e6f816e5f | control       | 4096 |   40 |         0 |     1 | True      |
| bf615b53-4c32-4604-8641-a0aaf117014d | ceph-storage  | 4096 |   40 |         0 |     1 | True      |
| c9b7892d-5d77-478a-a67d-6a9aae250349 | swift-storage | 4096 |   40 |         0 |     1 | True      |
| e06bd9e0-6c91-4e13-93be-a3611b845d63 | block-storage | 4096 |   40 |         0 |     1 | True      |
+--------------------------------------+---------------+------+------+-----------+-------+-----------+
```

```
openstack flavor set --property "cpu_arch"="x86_64" --property "capabili-
ties:boot_option"="local" baremetal
openstack flavor set --property "cpu_arch"="x86_64" --property "capabili-
ties:boot_option"="local" --property "capabilities:profile"="compute" compute
openstack flavor set --property "cpu_arch"="x86_64" --property "capabili-
ties:boot_option"="local" --property "capabilities:profile"="control" control
openstack flavor set --property "cpu_arch"="x86_64" --property "capabili-
ties:boot_option"="local" --property "capabilities:profile"="ceph-storage" ceph-
storage
```

The Flavors have to be set to every category of servers. Identify the servers based on IPMI address created earlier in instackenv.json file:

```
[stack@osp8-director ~]$ for i in $(ironic node-list | awk ' /power/ { print $2 }
')
do
abc=`ironic node-show $i | grep "10.23" | awk '{print $7}'`
echo $i $abc
done
```



```
[stack@osp8-director ~]$ for i in $(ironic node-list | awk ' /power/ { print $2 } ')
> do
> abc=`ironic node-show $i | grep "10.23" | awk '{print $7}'`
> echo $i $abc
> done
c4877202-d149-43f5-9c10-590e68c8b082 u'10.23.10.57', |
2804800a-a8cb-4170-8015-0bae8163661c u'10.23.10.79', |
1125e417-37c7-4735-9191-580d3c2a973a u'10.23.10.76', |
c12a7183-6cf4-420f-9355-ed002a895ca8 u'10.23.10.69', |
31fe96a6-284b-42cc-95b3-5280b47923df u'10.23.10.67', |
1002f59e-5edf-4e28-bec1-dd732c29cc81 u'10.23.10.59', |
123a50d2-ab56-48c2-b860-1b43d66cf5a2 u'10.23.10.66', |
7f252ac4-f0b2-45f7-a4a8-0079de124e32 u'10.23.10.74', |
cf100a8c-db6f-4873-808b-870ad324f94a u'10.23.10.75', |
3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 u'10.23.10.56', |
[stack@osp8-director ~]$
```

```
[stack@osp8-director ~]$ for i in b7dde876-354a-4688-8550-aec8f64c582c e4563ca5-
2f12-4e08-9905-f770f740ad2b \
> 285965a9-9713-4301-8ad5-7aa3ef5dd1c2
> do
> ironic node-update $i add
properties/capabilities='profile:control,boot_option:local'
> done
```

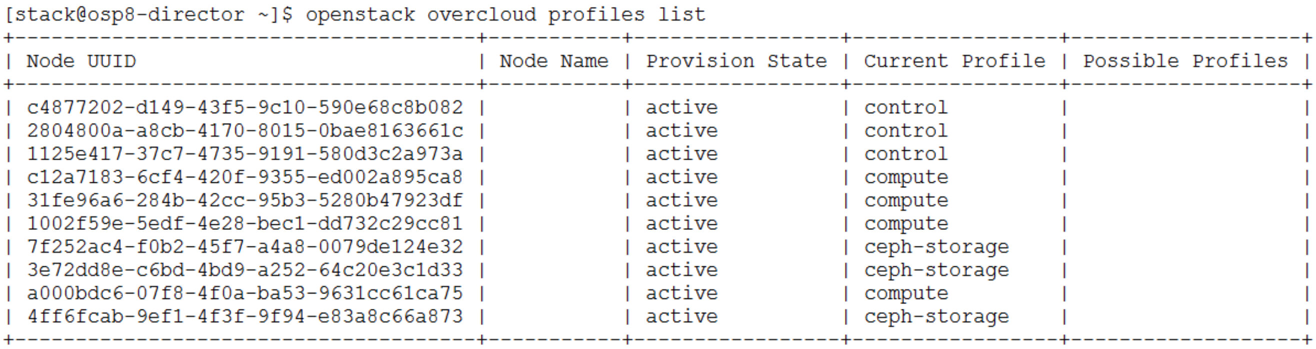
```
[stack@osp8-director ~]$ for i in b4dc04ac-0c69-4000-9c4d-2d82d141905f 036cae70-
bdee-427c-987c-a6a2d8a32292 \
> 8570c96e-f9cd-44ff-a1d8-0252bc405c24 af46cd81-c78e-47c5-94e3-44d9d669410c
19260dbb-29a9-4810-b39d-85cc6e1d886f \
> d4dae332-4595-43be-9b63-5a64331ea33b
> do
> ironic node-update $i add
properties/capabilities='profile:compute,boot_option:local'
> done
```

```
[stack@osp8-director ~]$ for i in 179befe6-2510-4311-ad9f-4880454fdaff \
> ff0dadfe-e2f3-408f-b69d-01398bb9699d b59f57e3-d5e1-499a-80c1-aac0c78c9534
> do
> ironic node-update $i add properties/capabilities='profile:ceph-
storage,boot_option:local'
> done
```

The added profiles can be queried for validation:

```
[stack@osp8-director ~]$ instack-ironic-deployment --show-profile
Preparing for deployment...
  Querying assigned profiles ...
    b7dde876-354a-4688-8550-aec8f64c582c
      "profile:control,boot_option:local"
    e4563ca5-2f12-4e08-9905-f770f740ad2b
      "profile:control,boot_option:local"
    285965a9-9713-4301-8ad5-7aa3ef5dd1c2
      "profile:control,boot_option:local"
    b4dc04ac-0c69-4000-9c4d-2d82d141905f
      "profile:compute,boot_option:local"
    036cae70-bdee-427c-987c-a6a2d8a32292
      "profile:compute,boot_option:local"
    8570c96e-f9cd-44ff-a1d8-0252bc405c24
      "profile:compute,boot_option:local"
    af46cd81-c78e-47c5-94e3-44d9d669410c
      "profile:compute,boot_option:local"
    19260dbb-29a9-4810-b39d-85cc6e1d886f
      "profile:compute,boot_option:local"
    d4dae332-4595-43be-9b63-5a64331ea33b
      "profile:compute,boot_option:local"
    179befe6-2510-4311-ad9f-4880454fdaff
      "profile:ceph-storage,boot_option:local"
    ff0dadfe-e2f3-408f-b69d-01398bb9699d
      "profile:ceph-storage,boot_option:local"
    b59f57e3-d5e1-499a-80c1-aac0c78c9534
      "profile:ceph-storage,boot_option:local"
  DONE.
Prepared.
```

You can also query the servers and associated profiles as openstack overcloud profiles list



```
[stack@osp8-director ~]$ openstack overcloud profiles list
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| Node UUID                            | Node Name | Provision State | Current Profile | Possible Profiles |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| c4877202-d149-43f5-9c10-590e68c8b082 |           | active          | control         |                   |
| 2804800a-a8cb-4170-8015-0bae8163661c |           | active          | control         |                   |
| 1125e417-37c7-4735-9191-580d3c2a973a |           | active          | control         |                   |
| c12a7183-6cf4-420f-9355-ed002a895ca8 |           | active          | compute         |                   |
| 31fe96a6-284b-42cc-95b3-5280b47923df |           | active          | compute         |                   |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 |           | active          | compute         |                   |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 |           | active          | ceph-storage    |                   |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 |           | active          | ceph-storage    |                   |
| a000bdc6-07f8-4f0a-ba53-9631cc61ca75 |           | active          | compute         |                   |
| 4ff6fcab-9ef1-4f3f-9f94-e83a8c66a873 |           | active          | ceph-storage    |                   |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
```

You can validate the ipmi, mac_address and server profiles as shown below:

```
for i in $(ironic node-list | awk '/None/ {print $2}' );
do
ipmi_addr=`ironic node-show $i | grep "10.23" | awk '{print $7}'`
mac_addr=`ironic node-port-list  $i | awk '/00:25/ {print $4}'`
profile=`ironic node-show $i | grep -io "u'profile:.*:local"`
echo $i $ipmi_addr $mac_addr $profile
done
```

# Overcloud Setup

Before delving into the Overcloud installation, it is necessary to understand and change the templates for your configuration. Red Hat OpenSstack Platform Directordirector provides lot of flexibility in configuring Overcloud. At the same time, understanding the parameters and providing the right inputs to heat through these templates is paramount.

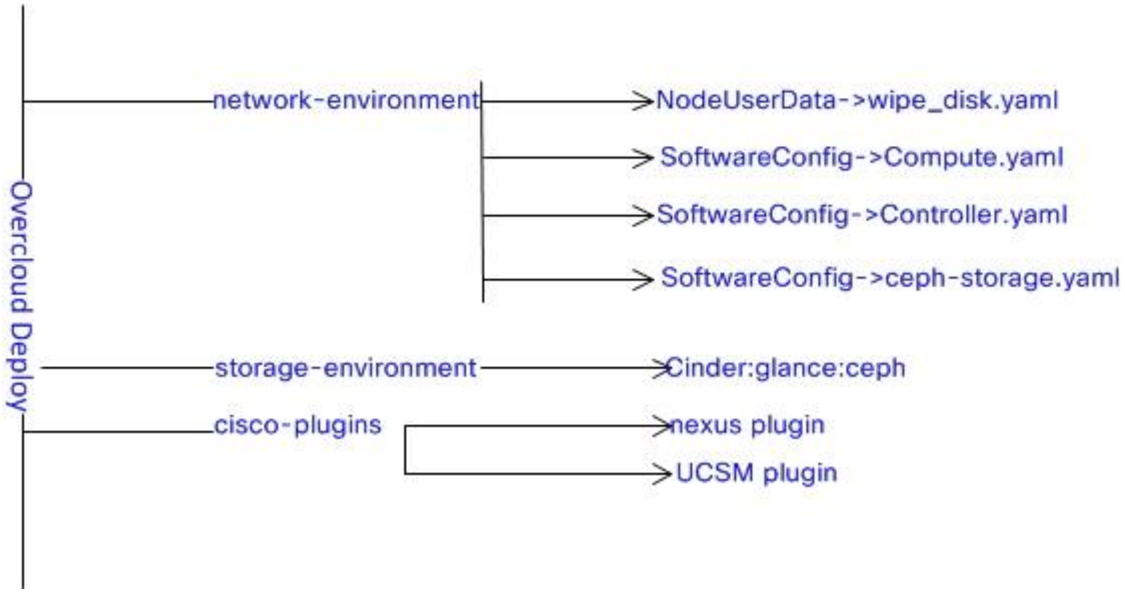
## Customize Heat Templates

Before attempting the Overcloud install, it is necessary to understand and setup the Overcloud heat templates. For complete details of the templates, please refer to the Red Hat online documentation on OpenStack.

Overcloud is installed through command line interface with the following command. A top down approach of the yaml and configuration files is provided here.

The files are sensitive to whitespaces and tabs.

Refer to Appendix, for run.sh, the command used to deploy Overcloud.


Overcloud Deploy
network-environment
NodeUserData->wipe_disk.yaml
SoftwareConfig->Compute.yaml
SoftwareConfig->Controller.yaml
SoftwareConfig->ceph-storage.yaml
storage-environment
Cinder:glance:ceph
cisco-plugins
nexus plugin
UCSM plugin

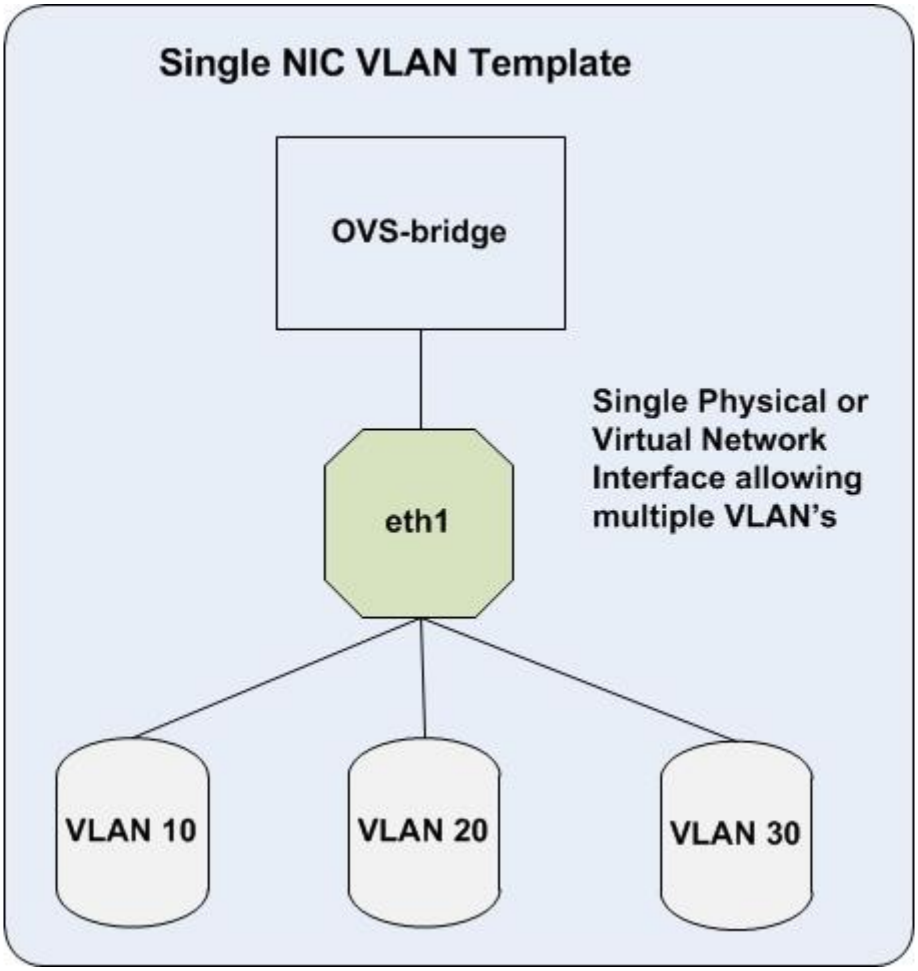
The heat templates have to be customized depending on the network layout and NIC interface configurations in the setup. The templates are standard heat templates in YAML format. They are included in the Appendix section.

The network configuration included in the director are of two categories and are included in /usr/share/openstack-tripleo-heat-templates/network/config.

More details are available at this link.

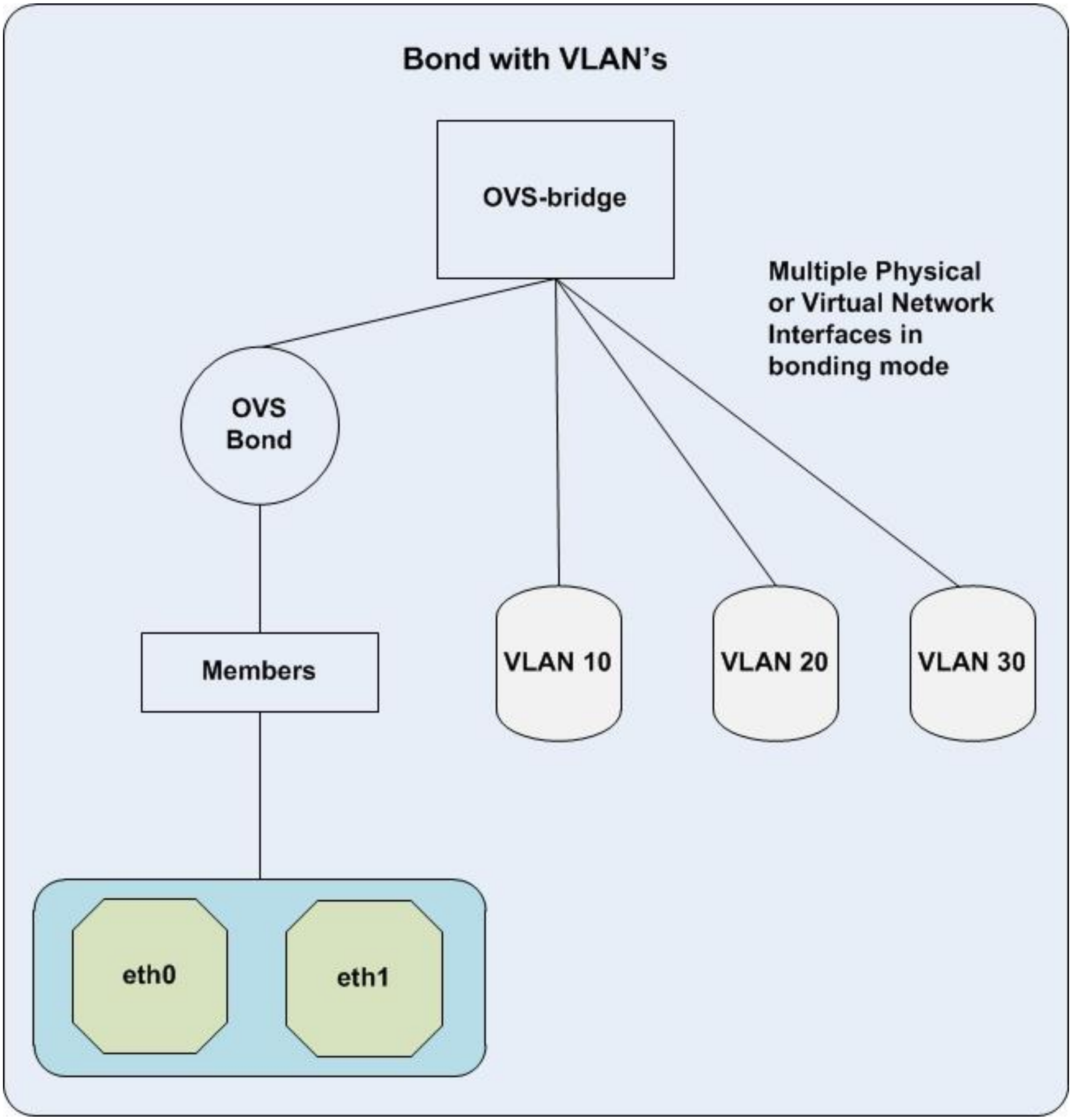
```
Single nic VLANs
Bond with VLANs
```

## Single NIC VLAN Templates


Single NIC VLAN Template
OVS-bridge
eth1
Single Physical or Virtual Network Interface allowing multiple VLAN's
VLAN 10
VLAN 20
VLAN 30

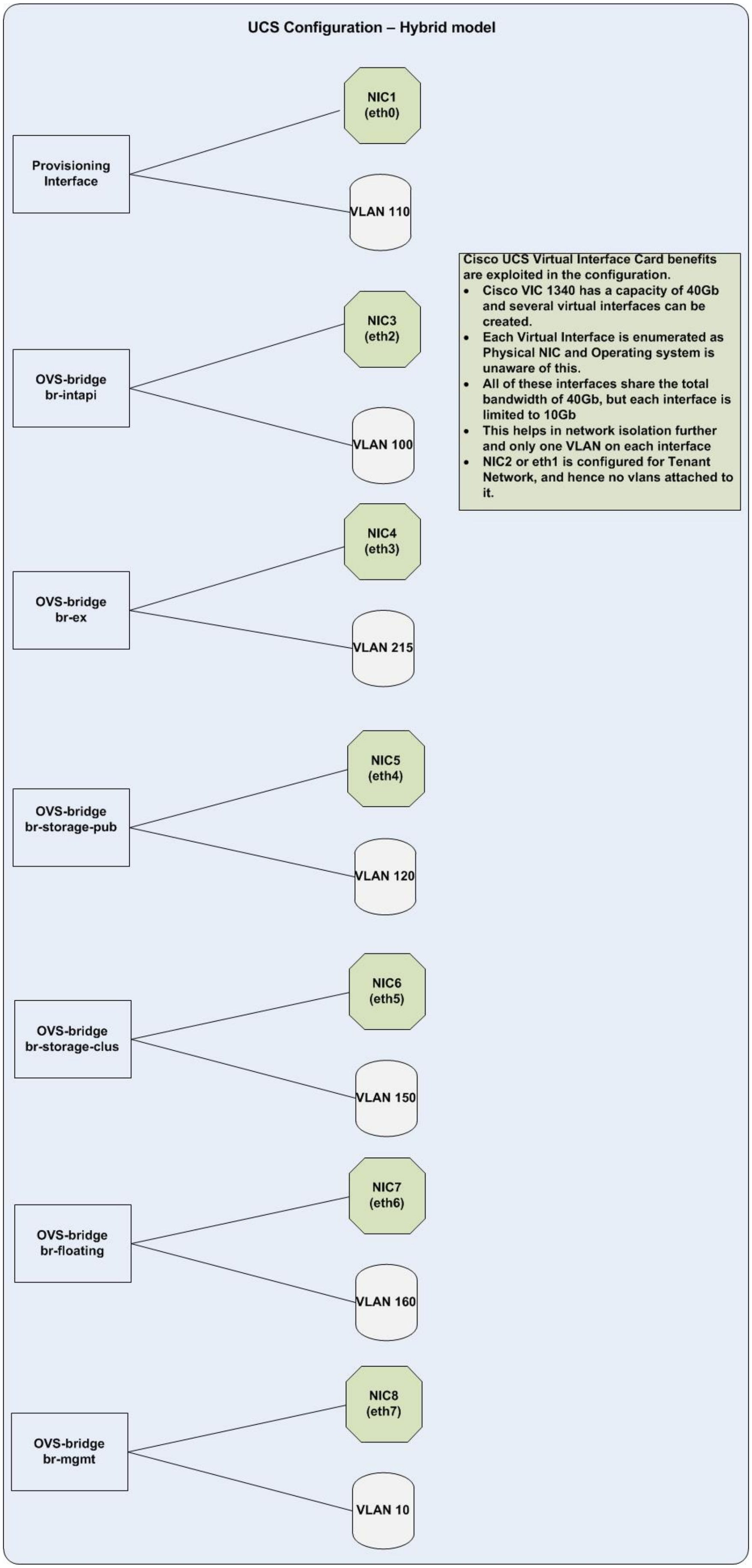
This model assumes that you have a single interface allowing all the VLANs configured in the system.

## Bond with VLAN Templates



## Cisco UCS Configuration

In the Cisco UCS configuration a hybrid model was adopted. This was done for simplicity and also to have a separate VLAN dedicated on each interface for every network. This gives a fine grain control of policies like QOS etc, if needed, but were not adopted for simplicity. NIC2 or eth1 was used as tenant interface.



UCS Configuration – Hybrid model

Cisco UCS Virtual Interface Card benefits are exploited in the configuration.

- Cisco VIC 1340 has a capacity of 40Gb and several virtual interfaces can be created.
- Each Virtual Interface is enumerated as Physical NIC and Operating system is unaware of this.
- All of these interfaces share the total bandwidth of 40Gb, but each interface is limited to 10Gb
- This helps in network isolation further and only one VLAN on each interface
- NIC2 or eth1 is configured for Tenant Network, and hence no vlans attached to it.

```
As stack user mkdir –p /home/stack/templates/nic-configs
```

Copy the template files from /usr/share/openstack-tripleo-heat-templates. Refer to Red Hat online documentation.

Create network-environment.yaml per above documentation or use Appendix for reference. Sample template files can also be downloaded from https://communities.cisco.com/docs/DOC-70256

Download the zip file and extract the templates directory into /home/stack/templates. Make changes as needed to these templates.

```
[stack@osp8-director templates]$ ls *.yaml
ceph.yaml management.yaml network-management.yaml timezone.yaml cisco-
plugins.yaml network-environment.yaml storage-environment.yaml wipe-disk.yaml


[stack@osp8-director nic-configs]$ ls *.yaml
ceph-storage.yaml  compute.yaml  controller.yaml
```

Some of the above files may have to be created. These files are referenced in Overcloud deploy command either directly or through another file. Ceph.yaml has to be modified directly in /usr/share/openstack-tripleo-heat-templates.

## Yaml Configuration Files Overview

### network-environment.yaml

The first section is for resource_registry. The section for parameter defaults have to be customized. The following are a few important points to be noted in network-environment.yaml file:

- Enter the Network Cidr values in the parameter section.
- The Control Plane Default Route is the Gateway Router for the provisioning network or the Undercloud IP. This matches with your network_gateway and masquerade_network in your undercloud.conf file.
- EC2Metadata IP is the Undercloud IP.
- Neturon External Network Bridge should be set to "''". An empty string to allow multiple external networks or VLANs. In case you are using the same external network for VMs instead of floating IP's relace the string " ' ' " with br-ex.
- No bonding used in the configuration. This will be addressed in our future releases.

### controller.yaml

This parameter section overrides the ones mentioned in the networking-environment file. The get_param calls for the defined parameters. The following are important points to be considered for Controller.yaml file:

- The PXE interface NIC1 should have dhcp as false to configure static ips, with next hop going to Undercloud node.
- The external bridge is configured to the External Interface Default Route on the External Network VlanID.
- The MTU value of 9000 to be added as needed. Both the storage networks are configured on mtu 9000.

### compute.yaml

The same rules for the Controller apply:

- The PXE interface NIC1 is configured with dhcp as false. There are no external IP's available for Controller and Storage. Hence NATing is done through Undercloud node. For this purpose, the Control Plane Default Route is the, network gateway defined in undercloud.conf file which is also the Undercloud local_ip.

- Only the Storage Public network is defined along with Tenant networks on Compute nodes.

### ceph-storage.yaml

- Same as Compute.yaml mentioned above.
- Only Storage Public and Storage Cluster are defined in this file.

### ceph.yaml

Configuring Ceph.yaml is tricky and needs to be done carefully. This is because we are configuring the partitions even before installing operating system on it. Also depending on the configuration whether you are using C240M4 LFF or C240M4 SFF the configuration changes.

An overview of the current limitations from the Red Hat OpenStack Platform director and Cisco UCS and the workarounds are provided for reference.

The way disk ordering is done is inconsistent. However for Ceph to work you need a consistent way of disk ordering. Post boot you can setup the disk labels by by-uuid or by-partuuid.

This is also a challenge to use JBODs in Ceph, the conventional way. Using RAID-0 Luns in place of JBOD's is equally challenging. The Lun ID's have to be consistent every time a server reboots. The order that is deployed in UCS is also unpredictable. The following workarounds have evolved with the configuration to meet these requirements. The internal SSD drives in both C240 LFF and SFF models will not be used as they are not visible to the RAID controller in the current version of UCSM and will pose challenges to Red Hat OpenStack Platform director (they are visible to BIOS, Luns cannot be carved out as RAID controller does not see them and they appear as JBODs to the kernel thus breaking the LUN and JBOD id's).

**Figure 11** Cisco UCS C240 M4 – Large Form Factor with 12 Slots

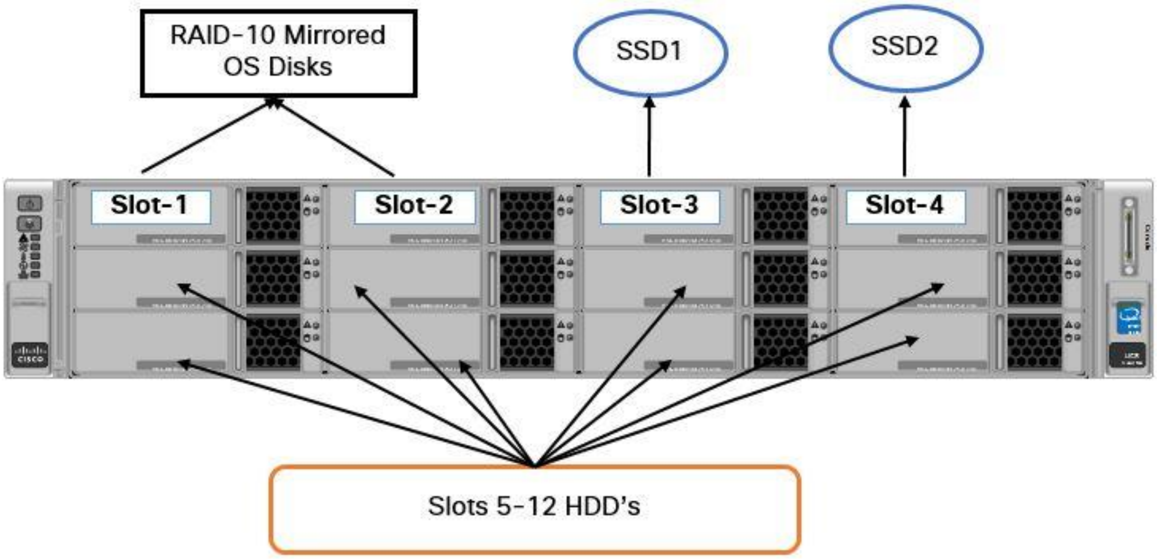
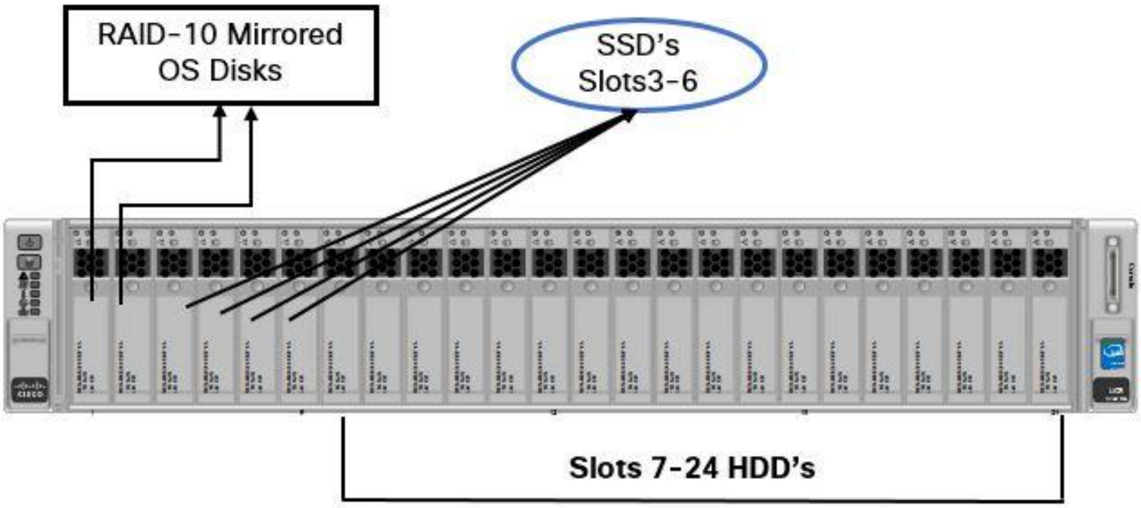


**Figure 12** Cisco UCS C240 M4 – Small Form Factor with 24 Slots



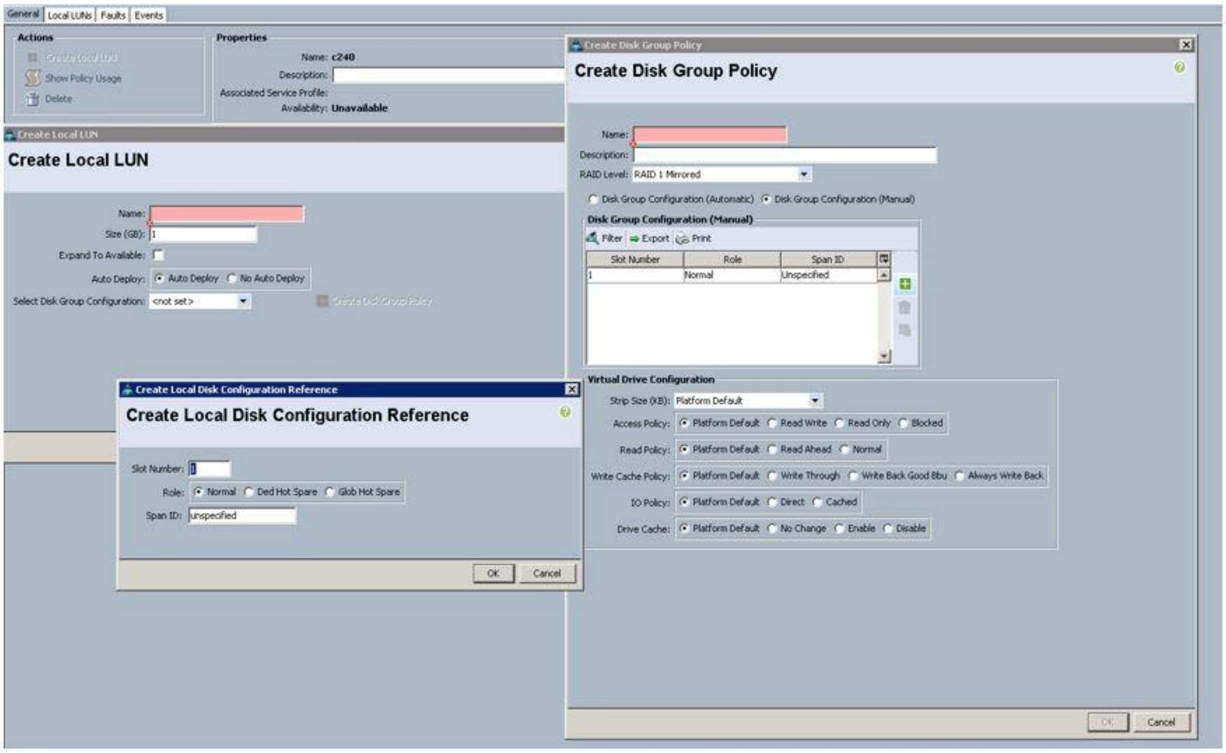
## Cisco UCS Side Fixes to Mitigate the Issue

As mentioned earlier, storage profiles will be used from UCS side on these servers:

Make sure that you do not have local disk configuration policy in UCS for these servers. It should read as no-disk policy.

Create storage profile, disk group policy as below under the template. There will be one Disk group policy for each slot. One policy for RAID-10 for the OS luns and one policy each of RAID-0 for the remaining.

Navigate to Create Storage Profile -> Create Local Lun -> Create Disk Group Policy (Manual) -> Create Local disk configuration. This will help in binding the disk slot to each lun created.



Properties

Name: c240

Description:

Associated Service Profile:

Availability: Unavailable

Create Local LUN

Name:

Size (GB): 1

Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: <not set>

Create Local Disk Configuration Reference

Slot Number: 1

Role: ● Normal ○ Ded Hot Spare ○ Glob Hot Spare

Span ID: unspecified

Create Disk Group Policy

Name:

Description:

RAID Level: RAID 1 Mirrored

○ Disk Group Configuration (Automatic) ● Disk Group Configuration (Manual)

Disk Group Configuration (Manual)

| Slot Number | Role | Span ID |
|---|---|---|
| 1 | Normal | Unspecified |

Virtual Drive Configuration

Strip Size (KB): Platform Default

Access Policy: ● Platform Default ○ Read Write ○ Read Only ○ Blocked

Read Policy: ● Platform Default ○ Read Ahead ○ Normal

Write Cache Policy: ● Platform Default ○ Write Through ○ Write Back Good Bbu ○ Always Write Back

IO Policy: ● Platform Default ○ Direct ○ Cached

Drive Cache: ● Platform Default ○ No Change ○ Enable ○ Disable

Create first the boot LUN from the first 2 slots and then apply. This will give LUN-0 to boot luns.

Create the second and third LUNs from the SSD slots (as in C240M4 LFF ). This would create RAID-0 luns, LUN-1 and LUN-2 on the SSD disks.

The rest of the LUNs can be created and applied in any order.

With the above procedure, we are assured that LUN-0 is for Operating system, LUN-1 and LUN-2 for SSDs and the rest for HDDs. This in turn decodes to /dev/sda for boot lun, /dev/sdb for SSD1 and /dev/sdc for SSD2 and the rest for HDD's.

Do not apply all the luns at the same time in the storage profile. First apply the boot lun, which should become LUN-0, followed by the SSD luns and then the rest of the HDD luns. Failure to comply with the above, will cause lun assignment in random order and heat will deploy on whatever the first boot lun presented to it.

Follow a similar procedure for C240 SFF servers too. A minimum of 4 SSD journals recommended for C240M4 SFF. The first two SSD luns with 5 partitions and the rest two with 4 partitions each.

### OpenStack Side Fixes to Mitigate the Issue

Implementing Red Hat OpenStack Platform director to successfully deploy Ceph on these disks need gpt label pre-created. This can be achieved by including wipe-disk.yaml file which creates these labels with sgdisk utility. Please refer to Appendix for details about wipe-disk.yaml.

In the current version there is only one ceph.yaml file on all the servers. This mapping has to be uniform across the storage servers.

While the contents of ceph.yaml in the Appendix are self-explanatory, the following is how the mappings between SSDs and HDDs need to be done:

```
ceph::profile::params::osds:
    '/dev/sdd':
        journal: '/dev/sdb1'
    '/dev/sde':
        journal: '/dev/sdb2'
    '/dev/sdf':
        journal: '/dev/sdb3'
    '/dev/sdg':
        journal: '/dev/sdb4'
    '/dev/sdh':
        journal: '/dev/sdc1'
    '/dev/sdi':
        journal: '/dev/sdc2'
    '/dev/sdj':
        journal: '/dev/sdc3'
    '/dev/sdk':
        journal: '/dev/sdc4'
```

The above is an example for C240M4 LFF server. Based on the LUN ids created above /dev/sdb and /dev/sdc are journal entries. Four entries for each of these journal directs Red Hat OpenStack Platform to create 4 partitions on each SSD disk. The entries on the left are for HDD disks. Please do not append the partition number to left side HDD partitions.

A similar approach can be followed for SFF servers.

The ceph.yaml was copied to /usr/share/openstack-tripleo-heat-templates/puppet/hieradata/

### cisco-plugins.yaml

The parameters section specifies the parameters.

### Cisco UCS Manager

NetworkUCSMIp: UCS Manager IP

NetworkUCSMHostList: Mapping between tenant mac address derived from UCS with Service profile name, comma separated. This list has to be built for all the compute and controller nodes.

### Nexus

This will list both the Nexus switches details, their IPs and passwords.

> Servers: The list should specify the interface MAC of each controller and compute and the port-channel numbers created on the Nexus switch.
>
> NetworkNexusManagedPhysicalNetwork physnet-tenant, the parameter you pass in the Overcloud deploy command
>
> NetworkNexusVlanNamePrefix: 'q-' These are the vlans that will be created on the switches
>
> NetworkNexusVxlanGlobalConfig: false. Vxlan is not used and is not validated as part of this CVD
>
> NeutronServicePlugins: Leave the default string as is. Any typos may create successfully Overcloud but will fail to create VMs later.
>
> NeutronTypeDrivers: vlan. The only drivers validated in this CVD.
>
> NeutronCorePlugin: 'ml2'
>
> NeutronNetworkVLANRanges: 'physnet-tenant:250:700,floating:160:160' The range you are passing to Overcloud deploy.

Leave the controllerExtraConfig parameters to default as in templates, refer Appendix.

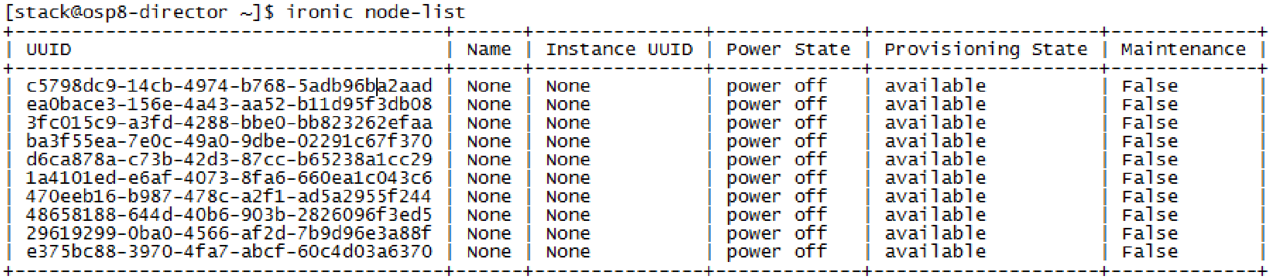
wipe-disks.yaml is configured as part of firstboot to create gpt lables on Storage node disks.

## Pre-Installation Checks Prior to Deploying Overcloud

To perform the pre-installation checks, complete the following steps:

Check for the existence of all the templates in templates and nic-configs directory as mentioned earlier.

Run ironic node-list to check that all the servers are available, powered off and not in maintenance.



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
| c5798dc9-14cb-4974-b768-5adb96ba2aad | None | None          | power off   | available          | False       |
| ea0bace3-156e-4a43-aa52-b11d95f3db08 | None | None          | power off   | available          | False       |
| 3fc015c9-a3fd-4288-bbe0-bb823262efaa | None | None          | power off   | available          | False       |
| ba3f55ea-7e0c-49a0-9dbe-02291c67f370 | None | None          | power off   | available          | False       |
| d6ca878a-c73b-42d3-87cc-b65238a1cc29 | None | None          | power off   | available          | False       |
| 1a4101ed-e6af-4073-8fa6-660ea1c043c6 | None | None          | power off   | available          | False       |
| 470eeb16-b987-478c-a2f1-ad5a2955f244 | None | None          | power off   | available          | False       |
| 48658188-644d-40b6-903b-2826096f3ed5 | None | None          | power off   | available          | False       |
| 29619299-0ba0-4566-af2d-7b9d96e3a88f | None | None          | power off   | available          | False       |
| e375bc88-3970-4fa7-abcf-60c4d03a6370 | None | None          | power off   | available          | False       |
+--------------------------------------+------+---------------+-------------+--------------------+-------------+
```

While understanding the reason why a server is not as listed above, you may use ironic APIs to change the state if they are not in the desired state:

After sourcing stackrc file;

```
ironic node-set-power-state <uuid> off
ironic node-set-provision-state <uuid> provide
ironic node-set-maintenance <uuid> false
In case of larger deployments, the default values of max resource per stack may
not be sufficient.
```

Reboot the Undercloud node.

## Deploying Overcloud

With the templates in place, Overcloud deploy can run the command mentioned in the Appendix. OpenStack help Overcloud deploy will show all the arguments that can be passed to the deployment command.

A snippet is provided below:

```
[stack@osp8-director ~]$ cat run.sh
#!/bin/bash
openstack overcloud deploy --templates \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-
isolation.yaml \
-e /home/stack/templates/network-environment.yaml \
-e /home/stack/templates/network-management.yaml \
-e /home/stack/templates/storage-environment.yaml \
-e /home/stack/templates/timezone.yaml \
-e /home/stack/templates/cisco-plugins.yaml \
--control-flavor control --compute-flavor compute --ceph-storage-flavor ceph-
storage \
--compute-scale 4 --control-scale 3  --ceph-storage-scale 3 \
--libvirt-type kvm \
--ntp-server 171.68.38.66 \
--neutron-network-type vlan \
--neutron-bridge-mappings datacentre:br-ex, \
physnet-tenant:br-tenant,floating:br-floating \
--neutron-network-vlan-ranges physnet-tenant:250:700,floating:160:160 \
--neutron-disable-tunneling --timeout 90 \
--verbose --debug --log-file overcloud_new.log
```

The following are a few parameters that need to be noted:

```
--control-flavor control --compute-flavor compute --ceph-storage-flavor ceph-
storage
```

ntp server is the server name to be used in the overcloud /etc/ntp.conf file
neutron-network-type is vlan.
neutron-network-vlan-ranges is physnet-tenant:250:700,floating:160:160. Here VLAN ranges from 250 to 700 were reserved for tenants, while VLAN 160 is for floating ip network.
Verbose, debug and log files are self-explanatory.

After successful deployment, the deploy command should show you the following:

```
DEBUG: os_cloud_config.utils.clients Creating nova client.
overcloud Endpoint: http://172.22.215.91:5000/v2.0/
overcloud Deployed
DEBUG: openstackclient.shell clean_up DeployOvercloud
```

Write down the endpoint URL to launch the dashboard later. This completes Overcloud deployment.

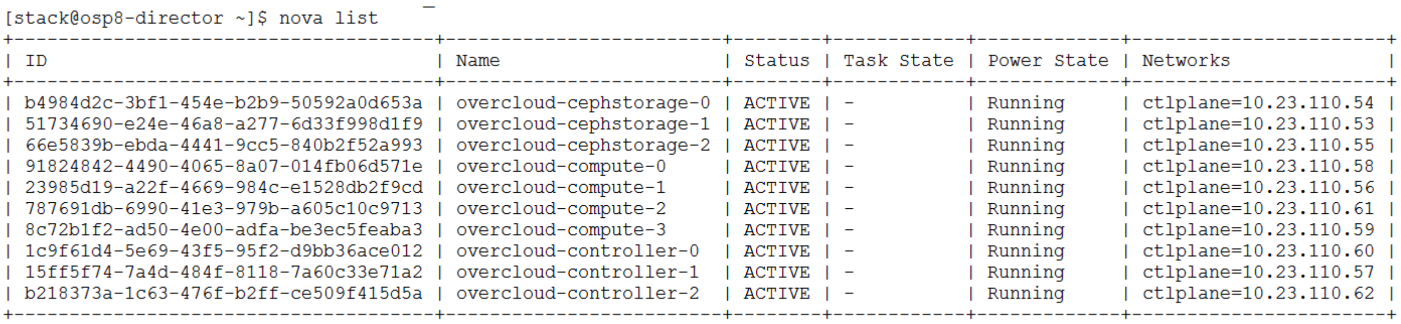
## Debugging Overcloud Failures

Overcloud deployment may fail for several reasons. Either because of a human error, for example, passing incorrect parameters or erroneous yaml configuration files or timeouts or bug. It is beyond the scope of this document to cover all of the possible failures. However, a few scenarios that were encountered on the configuration with explanations are provided in the Troubleshooting section of this document.

## Overcloud Post Deployment Steps

To perform the post deployment process, complete the following steps:

Run nova list and login as heat-admin to each host:



```
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks              |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| b4984d2c-3bf1-454e-b2b9-50592a0d653a | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.54 |
| 51734690-e24e-46a8-a277-6d33f998d1f9 | overcloud-cephstorage-1 | ACTIVE | -          | Running     | ctlplane=10.23.110.53 |
| 66e5839b-ebda-4441-9cc5-840b2f52a993 | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.55 |
| 91824842-4490-4065-8a07-014fb06d571e | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.58 |
| 23985d19-a22f-4669-984c-e1528db2f9cd | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.56 |
| 787691db-6990-41e3-979b-a605c10c9713 | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.61 |
| 8c72b1f2-ad50-4e00-adfa-be3ec5feaba3 | overcloud-compute-3     | ACTIVE | -          | Running     | ctlplane=10.23.110.59 |
| 1c9f61d4-5e69-43f5-95f2-d9bb36ace012 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.60 |
| 15ff5f74-7a4d-484f-8118-7a60c33e71a2 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.57 |
| b218373a-1c63-476f-b2ff-ce509f415d5a | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.62 |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
```

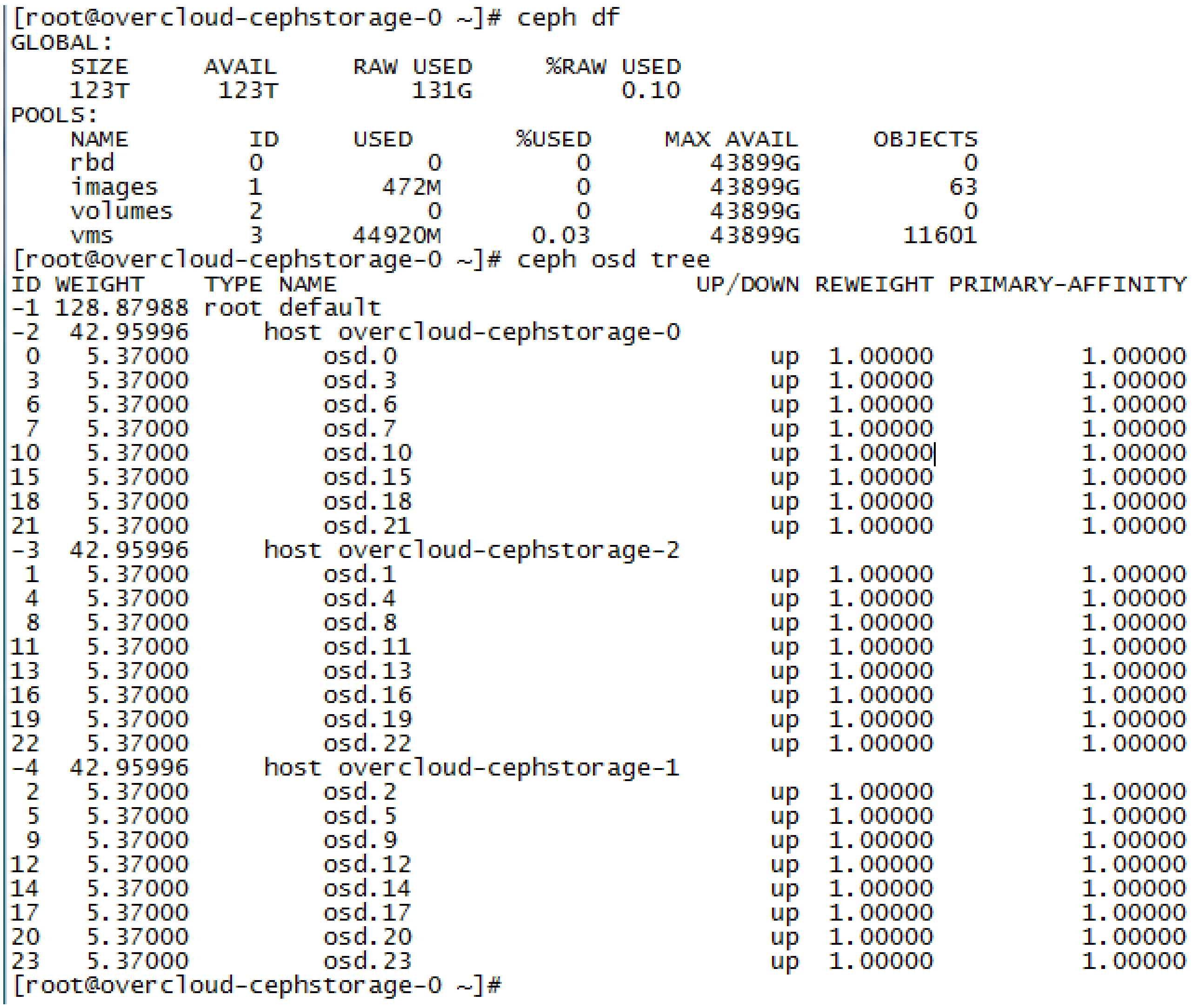
```
[stack@osp8-director ~]$ for i in $(nova list | awk '/ACTIVE/ {print $12}'|cut -d
"=" -f2 );
> do
> ssh -l heat-admin -o StrictHostKeyChecking=no $i "touch /tmp/abc; ls -l /tmp/abc"
> done
```

A command like the one listed above will validate that all the servers are up and running.

Check that the servers are registered with Red Hat Network.

subscription-manager status should reveal the status of this registration.

Query the Ceph pools and tree.



```
[root@overcloud-cephstorage-0 ~]# ceph df
GLOBAL:
    SIZE      AVAIL      RAW USED      %RAW USED
    123T      123T           131G           0.10
POOLS:
    NAME         ID      USED        %USED      MAX AVAIL      OBJECTS
    rbd          0          0            0         43899G            0
    images       1       472M            0         43899G           63
    volumes      2          0            0         43899G            0
    vms          3     44920M         0.03         43899G        11601
[root@overcloud-cephstorage-0 ~]# ceph osd tree
ID WEIGHT    TYPE NAME                           UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 128.87988 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                            up  1.00000          1.00000
 3   5.37000         osd.3                            up  1.00000          1.00000
 6   5.37000         osd.6                            up  1.00000          1.00000
 7   5.37000         osd.7                            up  1.00000          1.00000
10   5.37000         osd.10                           up  1.00000          1.00000
15   5.37000         osd.15                           up  1.00000          1.00000
18   5.37000         osd.18                           up  1.00000          1.00000
21   5.37000         osd.21                           up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                            up  1.00000          1.00000
 4   5.37000         osd.4                            up  1.00000          1.00000
 8   5.37000         osd.8                            up  1.00000          1.00000
11   5.37000         osd.11                           up  1.00000          1.00000
13   5.37000         osd.13                           up  1.00000          1.00000
16   5.37000         osd.16                           up  1.00000          1.00000
19   5.37000         osd.19                           up  1.00000          1.00000
22   5.37000         osd.22                           up  1.00000          1.00000
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2                            up  1.00000          1.00000
 5   5.37000         osd.5                            up  1.00000          1.00000
 9   5.37000         osd.9                            up  1.00000          1.00000
12   5.37000         osd.12                           up  1.00000          1.00000
14   5.37000         osd.14                           up  1.00000          1.00000
17   5.37000         osd.17                           up  1.00000          1.00000
20   5.37000         osd.20                           up  1.00000          1.00000
23   5.37000         osd.23                           up  1.00000          1.00000
[root@overcloud-cephstorage-0 ~]#
```

Also check the status of pcs resources as:

```
pcs resource cleanup
sleep 15
pcs status
pcs status | egrep -i “error|stop”
```

## Overcloud Post-Deployment Configuration

To perform the post-deployment configuration, complete the following steps:

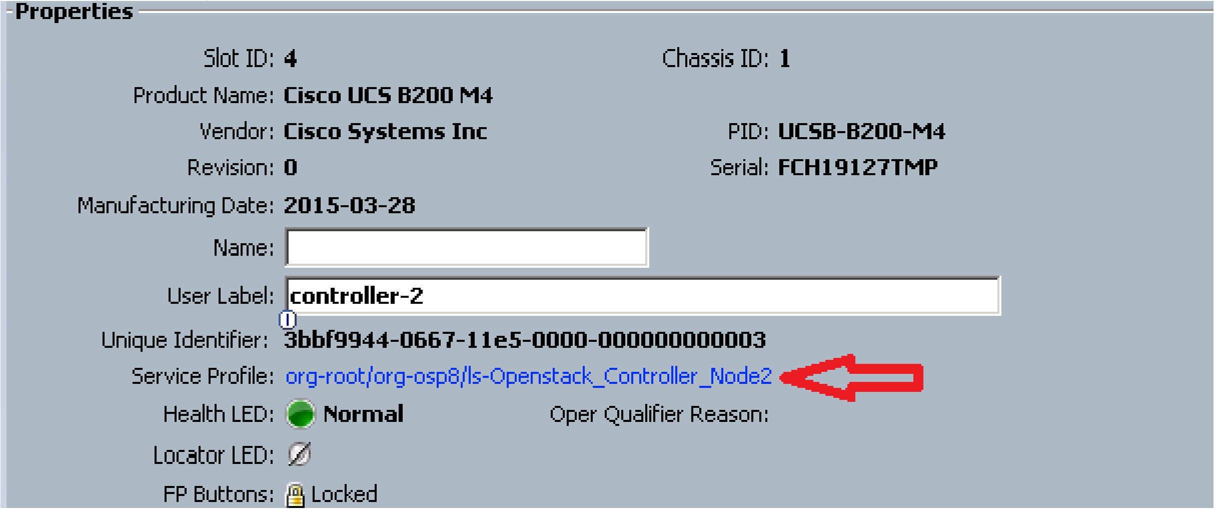
Start fence_cisco_ucs.

Run fence_cisco_ucs and pass the UCSM IP and passwords to it. Openstack_Controller_Node[1,2,3] are the service profile names for the controllers. Replace the string accordingly.

```
for i in 1 2 3
do
fence_cisco_ucs --ip=<UCSM IP> --username=admin --password=<password> \
--plug="Openstack_Controller_Node${i}" --suborg="/org-osp8/" --missing-as-off
--action=on --ssl-insecure -z;
done
```

```
Success: Powered ON
Success: Powered ON
Success: Powered ON
```

Replace the name of the controller service profile and your org name accordingly.
You can pick this up from General tab of UCS too.



Properties

Slot ID: 4
Chassis ID: 1
Product Name: Cisco UCS B200 M4
Vendor: Cisco Systems Inc
PID: UCSB-B200-M4
Revision: 0
Serial: FCH19127TMP
Manufacturing Date: 2015-03-28
Name:
User Label: controller-2
Unique Identifier: 3bbf9944-0667-11e5-0000-000000000003
Service Profile: org-root/org-osp8/ls-Openstack_Controller_Node2
Health LED: Normal
Oper Qualifier Reason:
Locator LED:
FP Buttons: Locked

Here the Service Profile name is Openstack_Controller_Node2 and the Sub-Organization is org-osp8

Configuring PaceMaker.

Before proceeding with pacemaker configuration, it is necessary to understand the relationship between the service profile names in UCS with the node names dynamically created by OpenStack as part of Overcloud deployment.

Either login through the Console or extract from /etc/neutron/plugin.ini from any of the controller nodes.

Plugin.ini will be updated by Cisco Plugins that have this information. Open /etc/neutron/plugin.in file and go to the end of the file. Extract the controller syntax.

The following is an example of extraction from plugin.ini file:

ucsm_host_list, could be populated as below in plugin.ini. In case it has appended localdomain, it needs to be removed with the current set of patches, details provided later.

```
ucsm_host_list=overcloud-compute-2:org-root/org-osp8/ls-
Openstack_Compute_Node1, overcloud-compute-3:org-root/org-osp8/ls-
Openstack_Compute_Node2, overcloud-compute-0:org-root/org-osp8/ls-
Openstack_Compute_Node3, overcloud-compute-1:org-root/org-osp8/ls-
Openstack_Compute_Node4, overcloud-controller-2:org-root/org-osp8/ls-
Openstack_Controller_Node1, overcloud-controller-1:org-root/org-osp8/ls-
Openstack_Controller_Node2, overcloud-controller-0:org-root/org-osp8/ls-
Openstack_Controller_Node3
```

Leave the org-root/<organization-name>. Instead extract just name of the host and the service profile name. There is no need to add organization here because fencing packages take the org-name as input during startup.

overcloud-controller-2:Openstack_Controller_Node1,
overcloud-controller-1:Openstack_Controller_Node2,

overcloud-controller-0:Openstack_Controller_Node3

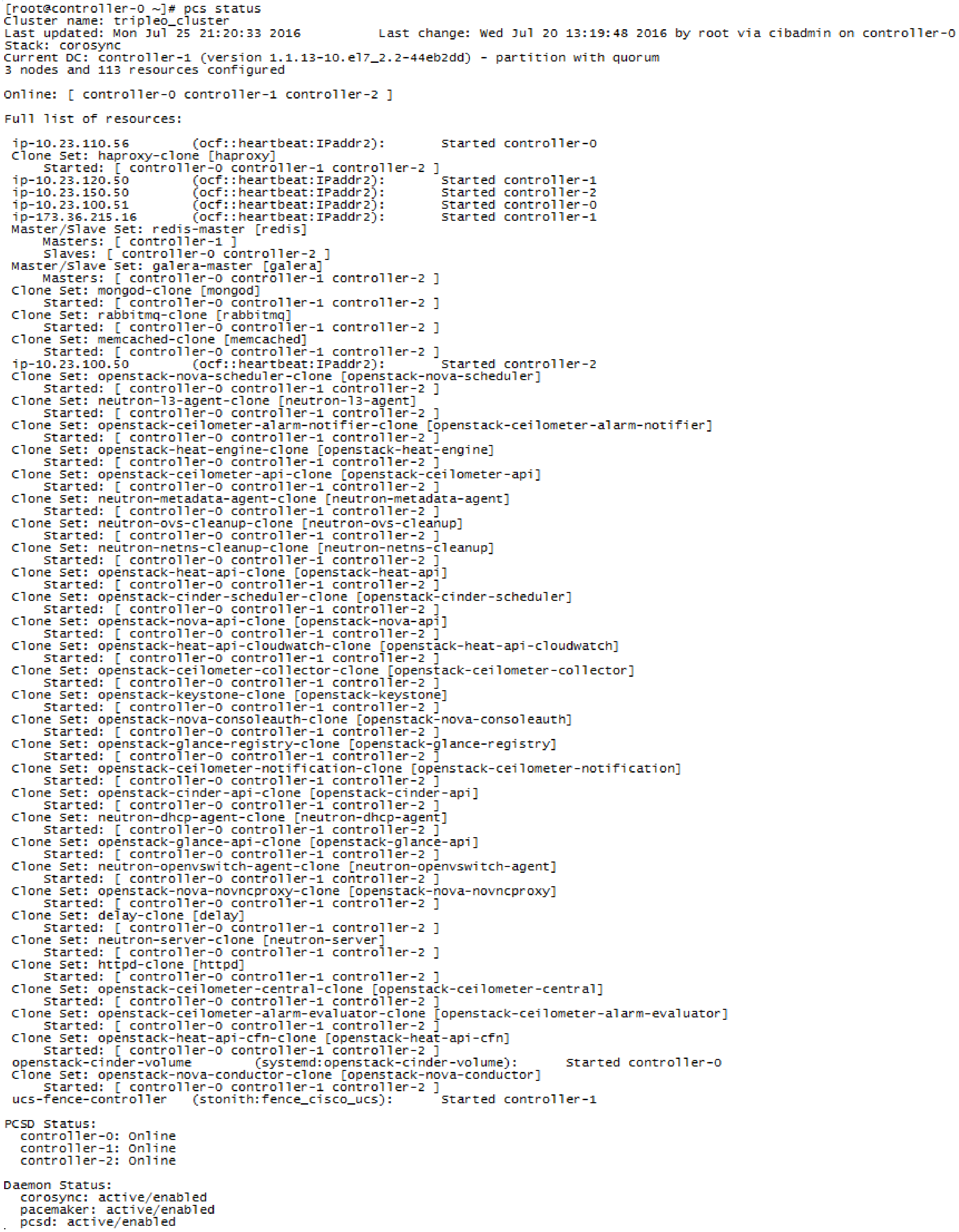
The mapping is controller-0 is mapped to Service Profile Controller_Node2 and so on. No need to extract the compute hosts as fencing packages run only on controller nodes to form the quorum.

Create a shell script as below with the following information and execute it

```
#/bin/bash
# Note that ';' as a separator instead of ',' from plugin.ini
sudo pcs stonith create ucs-fence-controller fence_cisco_ucs \
pcmk_host_map="overcloud-controller-1:Openstack_Controller_Node2;overcloud-
controller-0:Openstack_Controller_Node3;overcloud-controller-
2:Openstack_Controller_Node1" suborg="/org-osp8/" \
ipaddr=<UCSM IP> login=admin passwd=<password> ssl=1 ssl_insecure=1 op monitor
interval=60s
sleep 5;
pcs stonith update ucs-fence-controller power_timeout=60
pcs stonith update ucs-fence-controller meta failure-timeout=300s
pcs property set cluster-recheck-interval=300s
sleep 5;
pcs property set cluster-recheck-interval=300s
sudo pcs property set stonith-enabled=true
pcs property set stonith-timeout=300s
pcs resource cleanup
sleep 10;
sudo pcs stonith show ucs-fence-controller
sudo pcs property show
```

Querying ucs-fence-controller will reveal the mappings created.

```
[root@overcloud-controller-0 ~]# sudo pcs stonith show ucs-fence-controller
 Resource: ucs-fence-controller (class=stonith type=fence_cisco_ucs)
  Attributes: pcmk_host_map=overcloud-controller-
2:Openstack_Controller_Node1;overcloud-controller-
1:Openstack_Controller_Node2;overcloud-controller-
0:Openstack_Controller_Node3 suborg=/org-osp8/ ipaddr=10.23.10.5 log-
in=admin passwd=whatever password ssl=1 ssl_insecure=1 power_timeout=60
  Meta Attrs: failure-timeout=300s
  Operations: monitor interval=60s (ucs-fence-controller-monitor-interval-
60s)
```



```
[root@controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Mon Jul 25 21:20:33 2016          Last change: Wed Jul 20 13:19:48 2016 by root via cibadmin on controller-0
Stack: corosync
Current DC: controller-1 (version 1.1.13-10.el7_2.2-44eb2dd) - partition with quorum
3 nodes and 113 resources configured

Online: [ controller-0 controller-1 controller-2 ]

Full list of resources:

 ip-10.23.110.56        (ocf::heartbeat:IPaddr2):       Started controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ controller-0 controller-1 controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started controller-0
 ip-173.36.215.16       (ocf::heartbeat:IPaddr2):       Started controller-1
 Master/Slave Set: redis-master [redis]
     Masters: [ controller-1 ]
     Slaves: [ controller-0 controller-2 ]
 Master/Slave Set: galera-master [galera]
     Masters: [ controller-0 controller-1 controller-2 ]
 Clone Set: mongod-clone [mongod]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: rabbitmq-clone [rabbitmq]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: memcached-clone [memcached]
     Started: [ controller-0 controller-1 controller-2 ]
 ip-10.23.100.50        (ocf::heartbeat:IPaddr2):       Started controller-2
 Clone Set: openstack-nova-scheduler-clone [openstack-nova-scheduler]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-l3-agent-clone [neutron-l3-agent]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-alarm-notifier-clone [openstack-ceilometer-alarm-notifier]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-heat-engine-clone [openstack-heat-engine]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-api-clone [openstack-ceilometer-api]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-metadata-agent-clone [neutron-metadata-agent]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-ovs-cleanup-clone [neutron-ovs-cleanup]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-netns-cleanup-clone [neutron-netns-cleanup]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-heat-api-clone [openstack-heat-api]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-cinder-scheduler-clone [openstack-cinder-scheduler]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-nova-api-clone [openstack-nova-api]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-heat-api-cloudwatch-clone [openstack-heat-api-cloudwatch]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-collector-clone [openstack-ceilometer-collector]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-keystone-clone [openstack-keystone]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-nova-consoleauth-clone [openstack-nova-consoleauth]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-glance-registry-clone [openstack-glance-registry]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-notification-clone [openstack-ceilometer-notification]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-cinder-api-clone [openstack-cinder-api]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-dhcp-agent-clone [neutron-dhcp-agent]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-glance-api-clone [openstack-glance-api]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-openvswitch-agent-clone [neutron-openvswitch-agent]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-nova-novncproxy-clone [openstack-nova-novncproxy]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: delay-clone [delay]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: neutron-server-clone [neutron-server]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: httpd-clone [httpd]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-central-clone [openstack-ceilometer-central]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-ceilometer-alarm-evaluator-clone [openstack-ceilometer-alarm-evaluator]
     Started: [ controller-0 controller-1 controller-2 ]
 Clone Set: openstack-heat-api-cfn-clone [openstack-heat-api-cfn]
     Started: [ controller-0 controller-1 controller-2 ]
 openstack-cinder-volume        (systemd:openstack-cinder-volume):      Started controller-0
 Clone Set: openstack-nova-conductor-clone [openstack-nova-conductor]
     Started: [ controller-0 controller-1 controller-2 ]
 ucs-fence-controller   (stonith:fence_cisco_ucs):      Started controller-1

PCSD Status:
  controller-0: Online
  controller-1: Online
  controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

**Overcloud post deployment fixes for UCSM and Nexus Plugins.**

The following two patches are needed for UCSM and Nexus plugins. Download the patches from zipfile packaged in https://communities.cisco.com/docs/DOC-70256

Extract the zip file and copy config.py, mech_cisco_ucsm.py and nexus_network_driver.py from cisco-osp8-cvd/plugin_patches/ into a temporary directory on directory node.

Copy config.py to all the 3 controllers to /usr/lib/python2.7/site-packages/networking_cisco/plugins/ml2/drivers/cisco/ucsm/

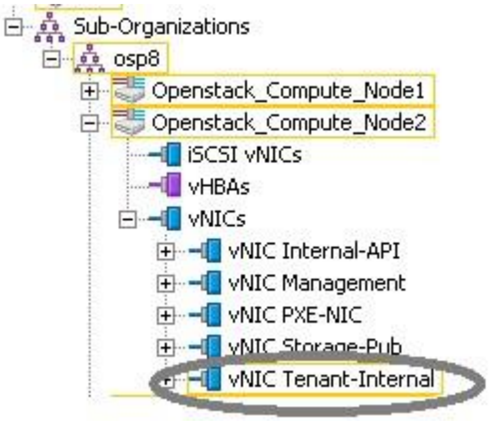
Copy mech_cisco_ucsm.py to all the 3 controllers to /usr/lib/python2.7/site-packages/networking_cisco/plugins/ml2/drivers/cisco/ucsm/

Copy nexus_network_driver.py to all the 3 controllers to /usr/lib/python2.7/site-packages/networking_cisco/plugins/ml2/drivers/cisco/nexus/

Check /etc/neutron/plugin.ini file and remove '.localdomain' entries from ucsm_host_list something like below:

ucsm_host_list=overcloud-compute-1:org-root/org-osp8/ls-Openstack_Compute_Node1, overcloud-compute-2:org-root/org-osp8/ls-Openstack_Compute_Node2, overcloud-compute-0:org-root/org-osp8/ls-Openstack_Compute_Node3, overcloud-compute-3:org-root/org-osp8/ls-Openstack_Compute_Node4, overcloud-controller-2:org-root/org-osp8/ls-Openstack_Controller_Node1, overcloud-controller-1:org-root/org-osp8/ls-Openstack_Controller_Node2, overcloud-controller-0:org-root/org-osp8/ls-Openstack_Controller_Node3

Add ucsm_virtio_eth_ports='Tenant-Internal' in /etc/neutron/plugin.ini to the UCSM section at the end of the file. Tenant-Internal is the UCS Configured interface for Tenant Traffic.
ucsm_virtio_eth_ports='Tenant-Internal'



Restart neutron

```
pcs resource restart neutron-server
```

Please check the readme file in plugin-patches directory of the zip file.

## Health Checks

To launch the dashboard URL created after successful installation of Overcloud, complete the following steps:

Go to http://172.22.215.91 (URL provided after Overcloud deployment) and login as admin and use the password created in the overcloudrc file (under $HOME of stack user).

Log into the system and navigate the tabs for any errors.

Update the system defaults.

# Functional Validation

Functional Validation includes the following:

- Navigating the dashboard across the admin, project, users tab to spot any issues
- Creating Tenants, Networks, Routers and Instances.
- Create Multiple Tenants, multiple networks and instances within different networks for the same tenant and with additional volumes with the following criteria:
  - Successful creation of Instances through CLI and validated through dashboard
  - Login to VM from the console.
  - Login to VMs through Floating IP's.
  - Reboot VMs
  - Check for the VLANs created both in UCSM and also on the Nexus switches. The VLANs should be available globally and also on the both port-channels created on each switch:

    ```
    Login to Nexus switch
    conf term
    show vlan | grep q-
    show running-config interfrace port-channel 17-18
    ```

The basic flow of creating and deleting instances through command line and horizon dashboard were tested. Creating multiple tenants and VLAN provisioning across Nexus switches and Cisco UCS Manager were verified while adding and deleting the instances.

For detailed information about validating Overcloud, refer to the Red Hat OpenStack Platform guide.

# Upscaling the POD

Scaling up the POD with growing business needs is a must. As business grows we need to add both compute and storage as needed by adding more hosts.

An attempt is made to scale up compute and storage. You may have to follow the steps below with the documented workarounds to add compute and storage nodes to the cluster.

## Scale Up Storage Nodes

### Provision the New Server in Cisco UCS

To provision the new server in UCS, complete the following steps:

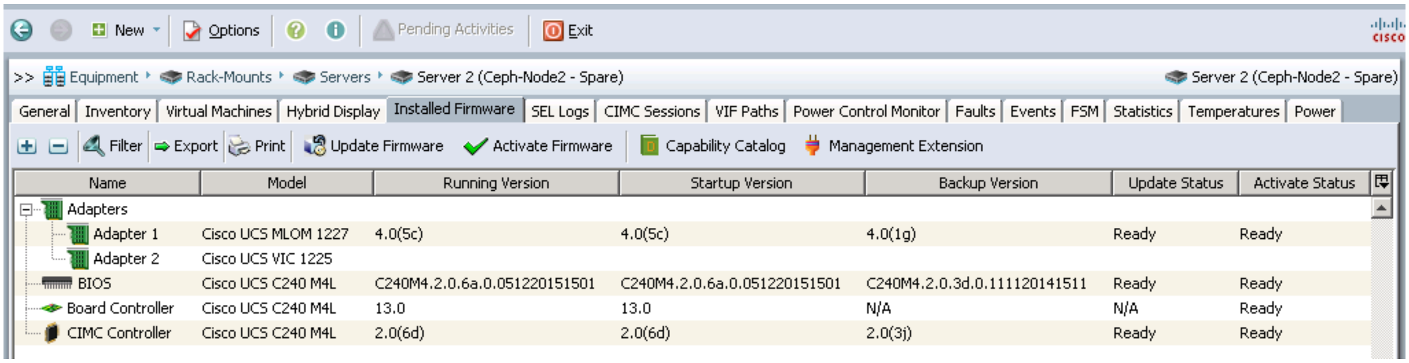
Rack the new C240M4 server(s). There is a single ceph.yaml in the current OpenStack version. Populate the hard disks in these storage servers in the same order as they exist in other servers.

Attach Console and discover the storage server(s) in UCS. Factory reset to defaults if needed and make them UCS managed.

Refer to this section for creating service profiles from Storage template. Create a new service profile from the template. Unbind the template and remove the storage policy that was attached to it earlier and associate the service profile to the server.

Upgrade firmware if needed.

Check the installed firmware on the new node and make sure that it is upgraded to the same version as other storage servers.

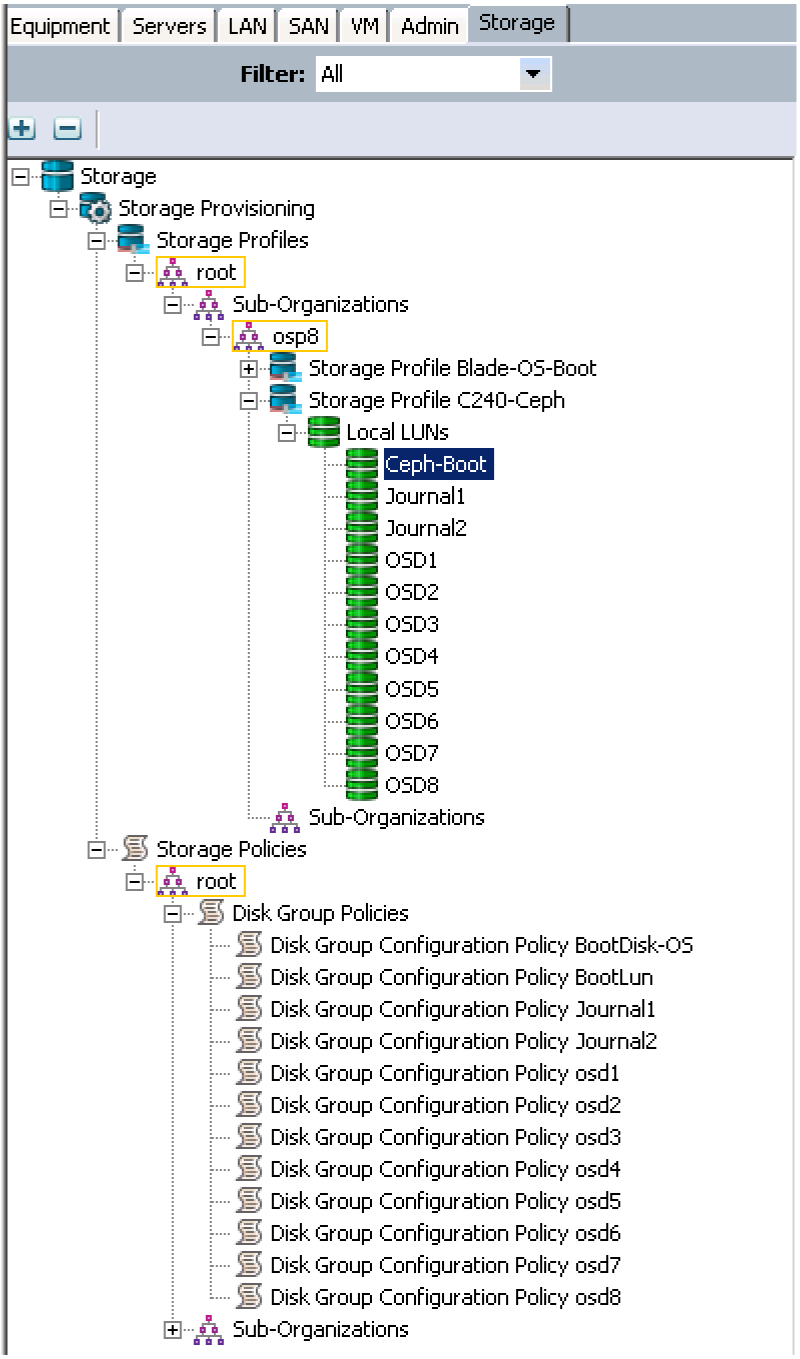


| Name | Model | Running Version | Startup Version | Backup Version | Update Status | Activate Status |
|---|---|---|---|---|---|---|
| Adapters | | | | | | |
| Adapter 1 | Cisco UCS MLOM 1227 | 4.0(5c) | 4.0(5c) | 4.0(1g) | Ready | Ready |
| Adapter 2 | Cisco UCS VIC 1225 | | | | | |
| BIOS | Cisco UCS C240 M4L | C240M4.2.0.6a.0.051220151501 | C240M4.2.0.6a.0.051220151501 | C240M4.2.0.3d.0.111120141511 | Ready | Ready |
| Board Controller | Cisco UCS C240 M4L | 13.0 | 13.0 | N/A | N/A | Ready |
| CIMC Controller | Cisco UCS C240 M4L | 2.0(6d) | 2.0(6d) | 2.0(3j) | Ready | Ready |

Create a new Storage Profile for Disks.

Before creating the storage profile, login to the equipment tab and make sure that all the new storage servers have the disks in place and they are physically on the same slots at par with other storage servers.

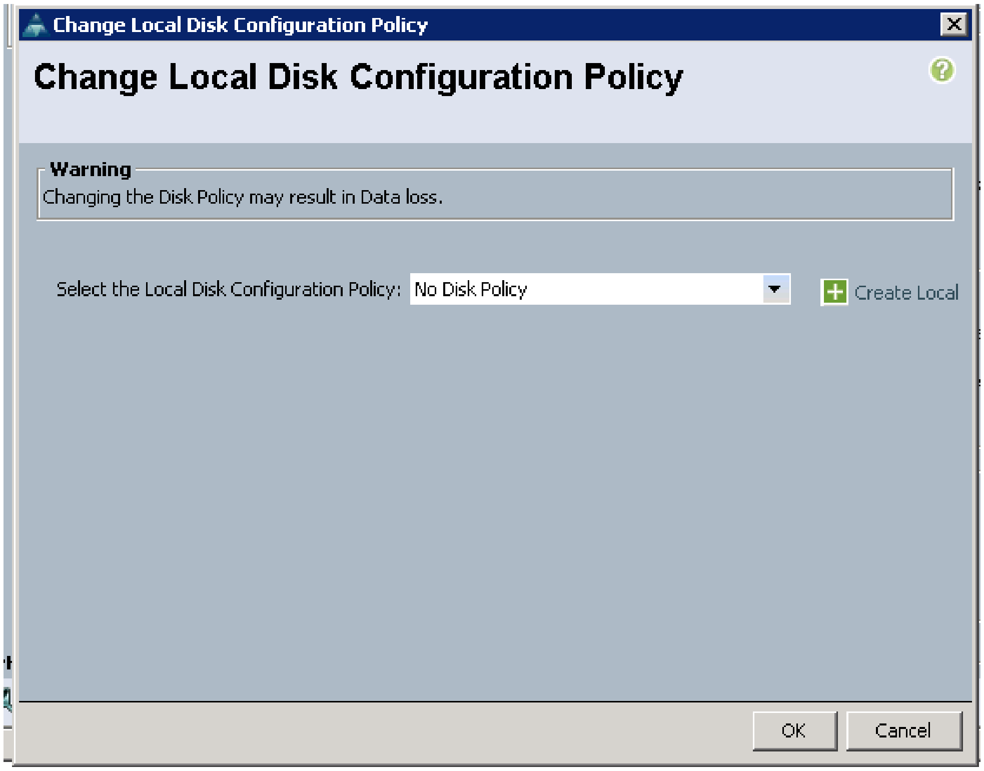
Since we used the storage profile earlier with other servers we cannot use them right away. The reason being the luns have to be added to the server in the same way as was done earlier. In case you are discovering more than one storage server at this stage, a single new profile created as below will serve the purpose. While creating this new storage profile, you can reuse the existing disk group configuration policies created earlier.



Filter: All

- Storage
  - Storage Provisioning
    - Storage Profiles
      - root
        - Sub-Organizations
          - osp8
            - Storage Profile Blade-OS-Boot
            - Storage Profile C240-Ceph
              - Local LUNs
                - Ceph-Boot
                - Journal1
                - Journal2
                - OSD1
                - OSD2
                - OSD3
                - OSD4
                - OSD5
                - OSD6
                - OSD7
                - OSD8
            - Sub-Organizations
    - Storage Policies
      - root
        - Disk Group Policies
          - Disk Group Configuration Policy BootDisk-OS
          - Disk Group Configuration Policy BootLun
          - Disk Group Configuration Policy Journal1
          - Disk Group Configuration Policy Journal2
          - Disk Group Configuration Policy osd1
          - Disk Group Configuration Policy osd2
          - Disk Group Configuration Policy osd3
          - Disk Group Configuration Policy osd4
          - Disk Group Configuration Policy osd5
          - Disk Group Configuration Policy osd6
          - Disk Group Configuration Policy osd7
          - Disk Group Configuration Policy osd8
        - Sub-Organizations

Go to the service profile of the new server and to the storage tab to create a new storage profile as shown below. Make sure that the local disk config policy is set to No Disk Policy.

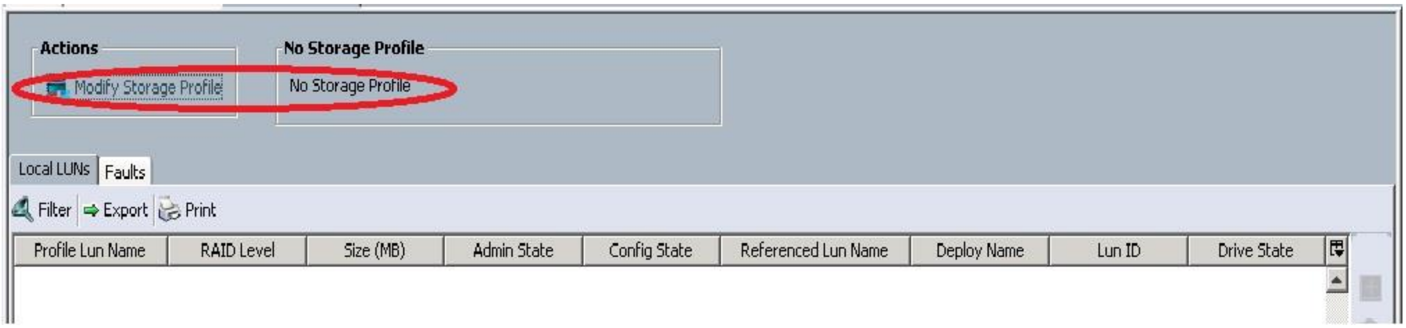


Change Local Disk Configuration Policy

## Change Local Disk Configuration Policy

**Warning**
Changing the Disk Policy may result in Data loss.

Select the Local Disk Configuration Policy: No Disk Policy

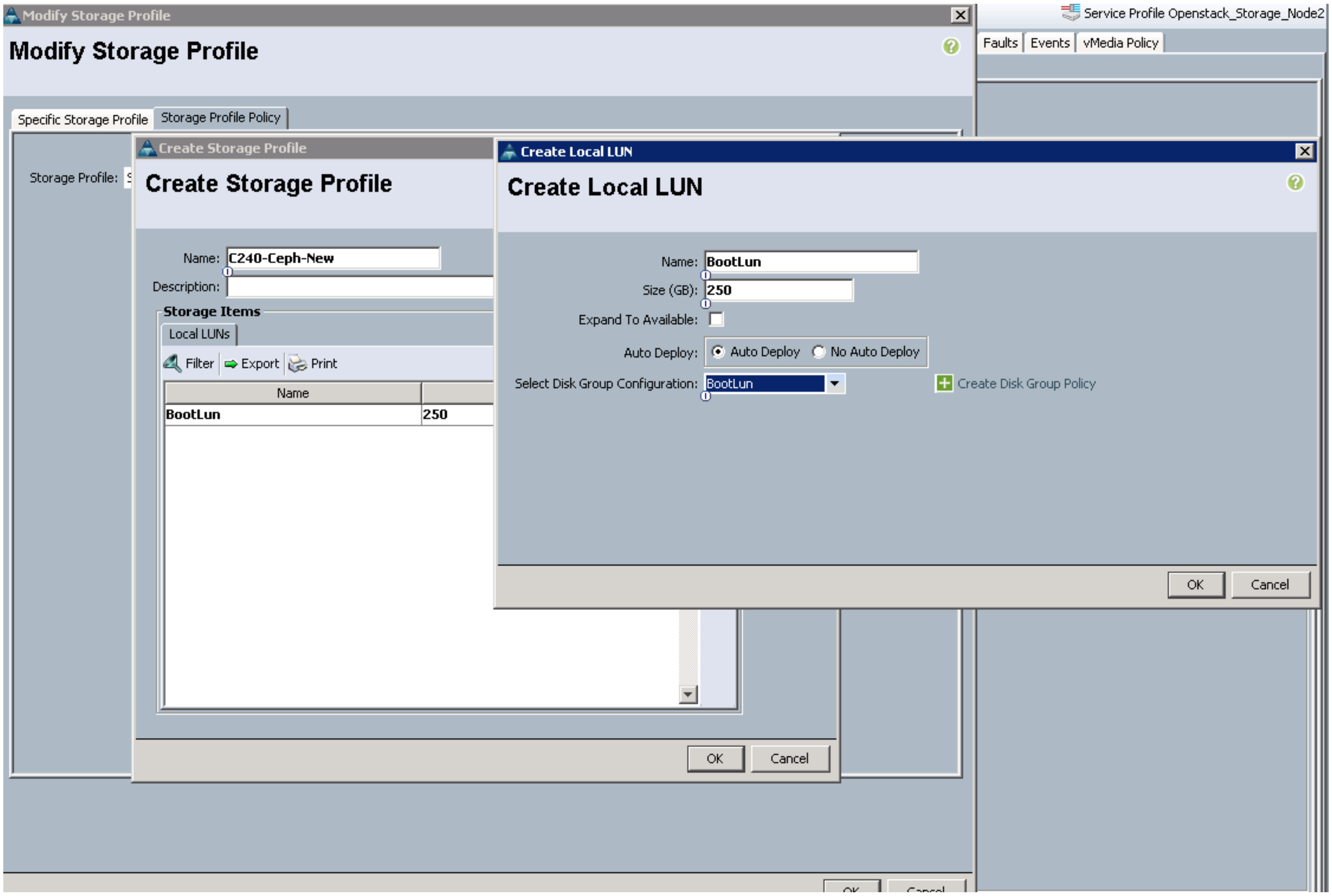


**Actions**

**No Storage Profile**

No Storage Profile

| Profile Lun Name | RAID Level | Size (MB) | Admin State | Config State | Referenced Lun Name | Deploy Name | Lun ID | Drive State |
|---|---|---|---|---|---|---|---|---|



Create Storage Profile

Name: C240-Ceph-New

Description:

| Name | |
|---|---|
| BootLun | 250 |

Create Local LUN

Name: BootLun

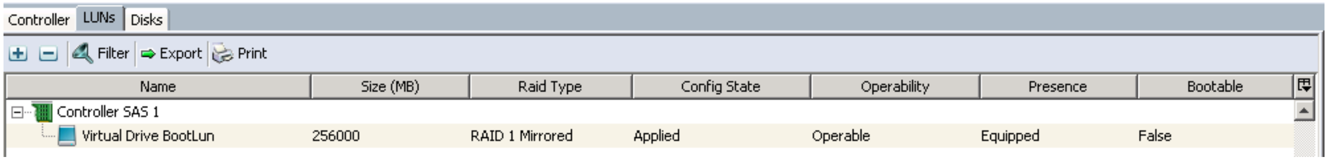
Size (GB): 250

Expand To Available: □

Auto Deploy: ● Auto Deploy ○ No Auto Deploy

Select Disk Group Configuration: BootLun

Attach this storage profile to the service profile. This will create the first boot lun LUN-0 on the server. Go back to the equipment tab and inventory/storage to check that this is the first Lun is added. This will be the boot lun LUN-0 that will be visible to the server bios. In case of multiple servers being added in this step, attach the new storage profile created above to all these service profiles. This in turn will create LUN-0 in all the nodes.



| Name | Size (MB) | Raid Type | Config State | Operability | Presence | Bootable |
|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | |
| Virtual Drive BootLun | 256000 | RAID 1 Mirrored | Applied | Operable | Equipped | False |

A subsequent update to this storage profile will be propagated across all these new service profiles.

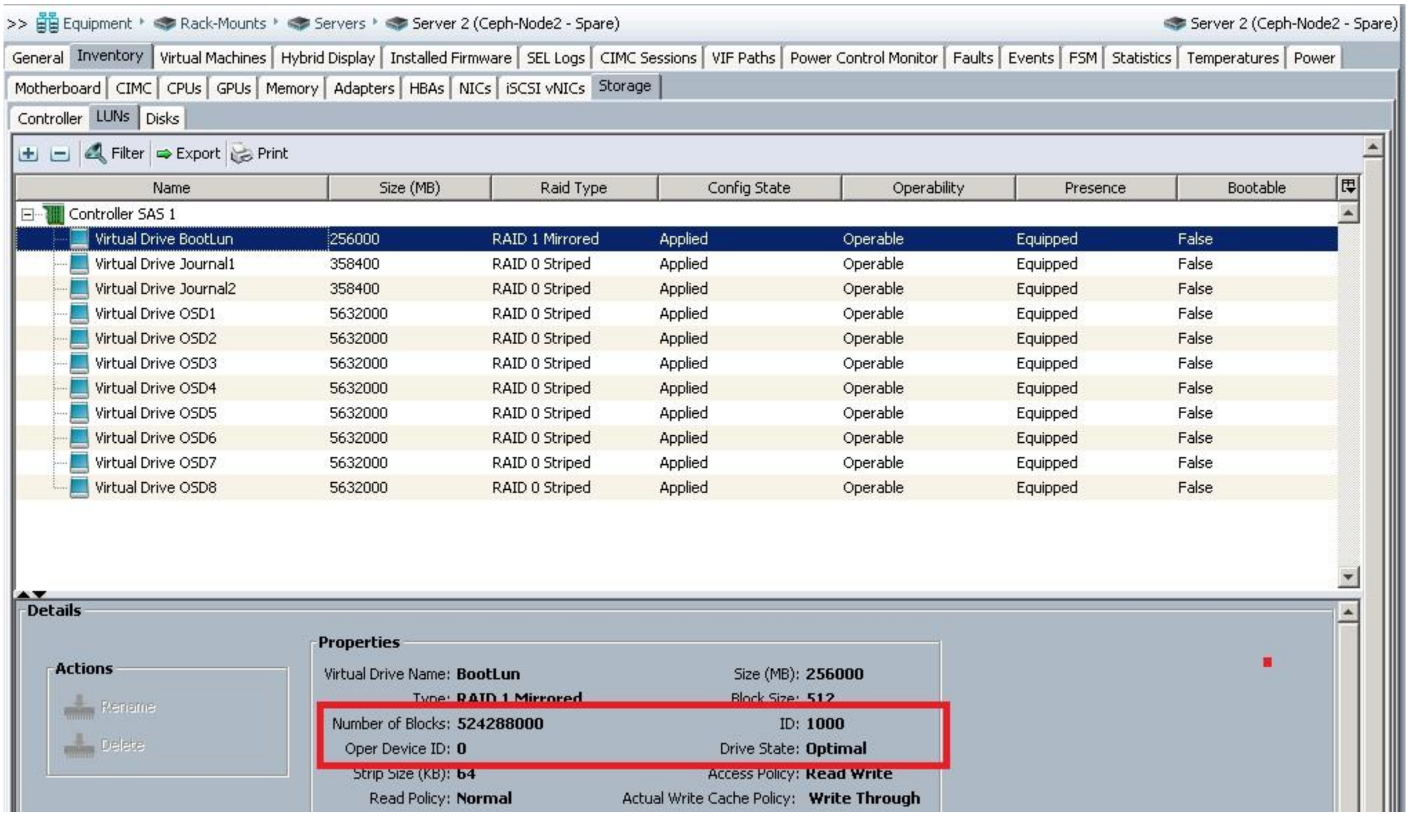
Go to Storage tab in UCSM and update the storage profile.

Create and attach SSD luns, which will be LUN-1 and LUN-2. Wait few minutes to make sure that all the new servers get these luns in the same order, boot as LUN-0, Journal1 as LUN-1 and Journal2 as LUN-2.

Verify from the equipment tab.

This will be consistent with other servers and we can expect sda for boot lun and sdb and sdc for SSD LUNs being used with the journals.

Add all the HDD LUNs later.



| Name | Size (MB) | Raid Type | Config State | Operability | Presence | Bootable |
|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | |
| Virtual Drive BootLun | 256000 | RAID 1 Mirrored | Applied | Operable | Equipped | False |
| Virtual Drive Journal1 | 358400 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive Journal2 | 358400 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD1 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD2 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD3 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD4 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD5 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD6 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD7 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |
| Virtual Drive OSD8 | 5632000 | RAID 0 Striped | Applied | Operable | Equipped | False |

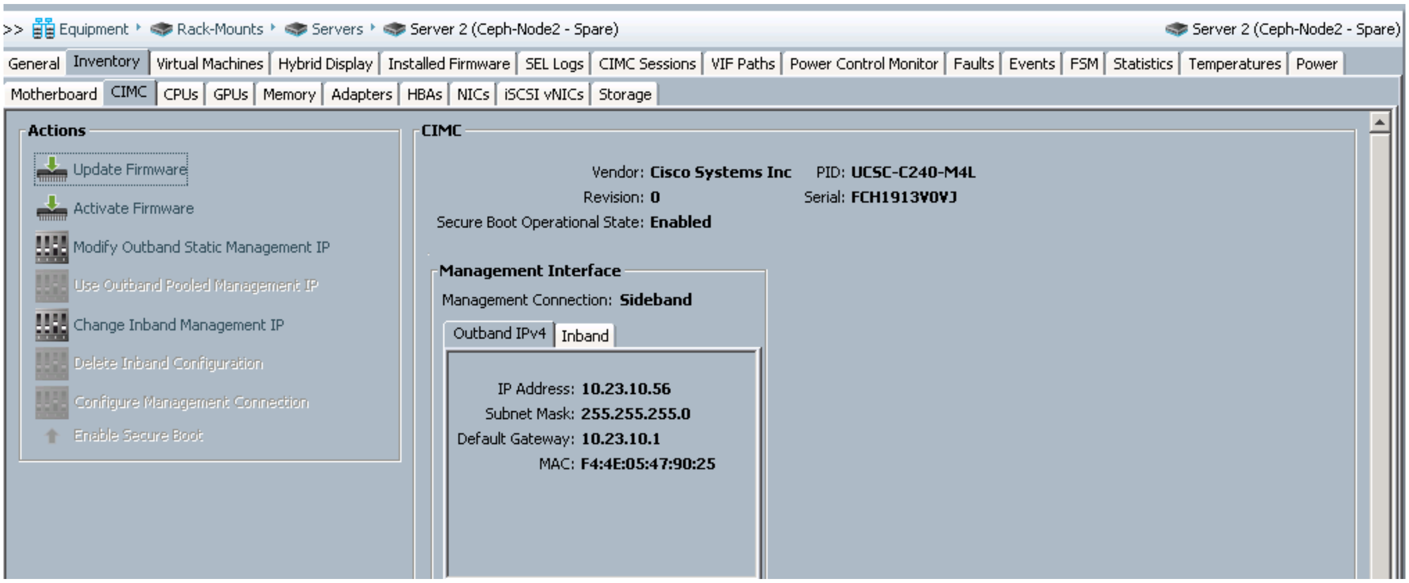
Virtual Drive Name: **BootLun** Size (MB): **256000**
Type: **RAID 1 Mirrored** Block Size: **512**
Number of Blocks: **524288000** ID: **1000**
Oper Device ID: **0** Drive State: **Optimal**
Strip Size (KB): **64** Access Policy: **Read Write**
Read Policy: **Normal** Actual Write Cache Policy: **Write Through**

The steps above do not represent the actual boot order. You may have to observe the actual boot order from KVM console to verify.

If the boot disks are being repurposed and are not new, go ahead and re-initialize the boot lun through bios. Boot server, CTR-R, F2 and reinitialize the VD for the boot LUNs.

Get the hardware inventory needed introspection.

Go to the Equipment tab > Inventory > CIMC and get the IPMI address.



Vendor: **Cisco Systems Inc** PID: **UCSC-C240-M4L**
Revision: **0** Serial: **FCH1913V0VJ**
Secure Boot Operational State: **Enabled**

Management Connection: **Sideband**

IP Address: **10.23.10.56**
Subnet Mask: **255.255.255.0**
Default Gateway: **10.23.10.1**
MAC: **F4:4E:05:47:90:25**

Under the same Inventory tab go to NIC subtab and get the pxe mac address of the server. The same inventory should have the CPU and memory details.

Specify the NIC order in the service profile. This should be the same as the other storage servers with provisioning interface as the first one.

Check the boot policy of the server. Validate that this is same as other storage servers. It should be LAN PXE first followed by local LUN.

## Run Introspection

To run Introspection, complete the following steps:

Prepare json file for introspection:

```
[stack@osp8-director ~]$ cat storage-new.json
{
  "nodes": [
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.56",
      "mac": [
        "00:25:b5:00:00:33"
      ],
      "memory": "131072",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "24"
    }
  ]
}
```

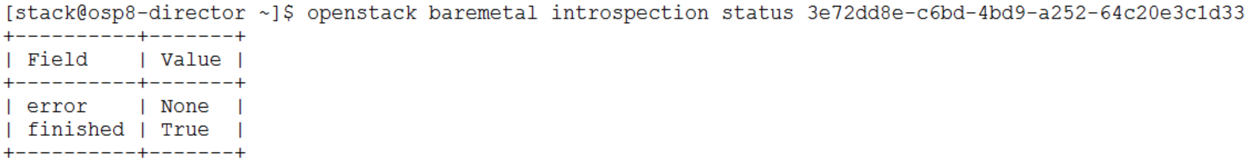
Check IPMI Connectivity:

```
[stack@osp8-director ~]$ ipmitool -I lanplus -H 10.23.10.56 -U admin -P <passwd>
chassis power off
Chassis Power Control: Down/Off
```

Initialize Boot Luns; in case you are reusing old disks it is recommended to initialize the boot luns.

Run discovery and introspection.

```
[stack@osp8-director ~]$ openstack baremetal import --json ~/storage-node.json
openstack baremetal configure boot
ironic node-list
[stack@osp8-director ~]$ ironic node-set-maintenance 948d704b-c82b-4b9a-8d01-
ad4899ce725f true
[stack@osp8-director ~]$ openstack baremetal introspection start 948d704b-c82b-
4b9a-8d01-ad4899ce725f
[stack@osp8-director ~]$ openstack baremetal introspection status 948d704b-c82b-
4b9a-8d01-ad4899ce725f
```



```
[stack@osp8-director ~]$ openstack baremetal introspection status 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33
+----------+-------+
| Field    | Value |
+----------+-------+
| error    | None  |
| finished | True  |
+----------+-------+
```

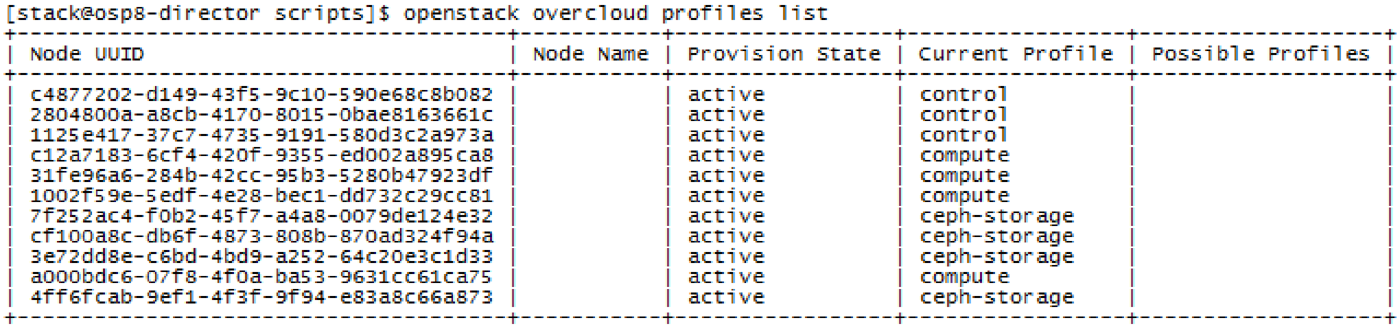
Repeat the steps above if you want to add multiple nodes.

Wait till the introspection is complete. The status command should yield finished as True and Error as none. Set the maintenance flag as false.

```
ironic node-set-maintenance 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 False
```

Update node properties:

```
[stack@osp8-director ~]$ ironic node-update 948d704b-c82b-4b9a-8d01-ad4899ce725f
\
> add properties/capabilities='profile:ceph-storage,boot_option:local'
```



```
[stack@osp8-director scripts]$ openstack overcloud profiles list
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| Node UUID                            | Node Name | Provision State | Current Profile | Possible Profiles |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| c4877202-d149-43f5-9c10-590e68c8b082 |           | active          | control         |                   |
| 2804800a-a8cb-4170-8015-0bae8163661c |           | active          | control         |                   |
| 1125e417-37c7-4735-9191-580d3c2a973a |           | active          | control         |                   |
| c12a7183-6cf4-420f-9355-ed002a895ca8 |           | active          | compute         |                   |
| 31fe96a6-284b-42cc-95b3-5280b47923df |           | active          | compute         |                   |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 |           | active          | compute         |                   |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 |           | active          | ceph-storage    |                   |
| cf100a8c-db6f-4873-808b-870ad324f94a |           | active          | ceph-storage    |                   |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 |           | active          | ceph-storage    |                   |
| a000bdc6-07f8-4f0a-ba53-9631cc61ca75 |           | active          | compute         |                   |
| 4ff6fcab-9ef1-4f3f-9f94-e83a8c66a873 |           | active          | ceph-storage    |                   |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
```

```
There are 4 ceph-storage nodes now.
```

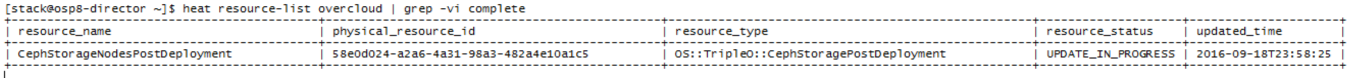
## Run Overcloud Deployment

The number of storage nodes has been incremented to 4 from 3. Here the number '4' indicates the total number of storage nodes in Overcloud.

```
#!/bin/bash
openstack overcloud deploy --templates --ceph-storage-scale 4 \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-
isolation.yaml \
-e /home/stack/templates/network-environment.yaml \
-e /home/stack/templates/network-management.yaml \
-e /home/stack/templates/storage-environment.yaml \
-e /home/stack/templates/timezone.yaml \
-e /home/stack/templates/cisco-plugins.yaml \
--log-file overcloud_storage-add.log
```

During the addition of nodes, Ceph health observed to be fine.

Heat Resource shows that the node is being added.



```
[stack@osp8-director ~]$ heat resource-list overcloud | grep -vi complete
+--------------------------------+--------------------------------------+------------------------------------+--------------------+---------------------+
| resource_name                  | physical_resource_id                 | resource_type                      | resource_status    | updated_time        |
+--------------------------------+--------------------------------------+------------------------------------+--------------------+---------------------+
| CephStorageNodesPostDeployment | 58e0d024-a2a6-4a31-98a3-482a4e10a1c5 | OS::TripleO::CephStoragePostDeployment | UPDATE_IN_PROGRESS | 2016-09-18T23:58:25 |
+--------------------------------+--------------------------------------+------------------------------------+--------------------+---------------------+
```

Addition of nodes complete with the following message:

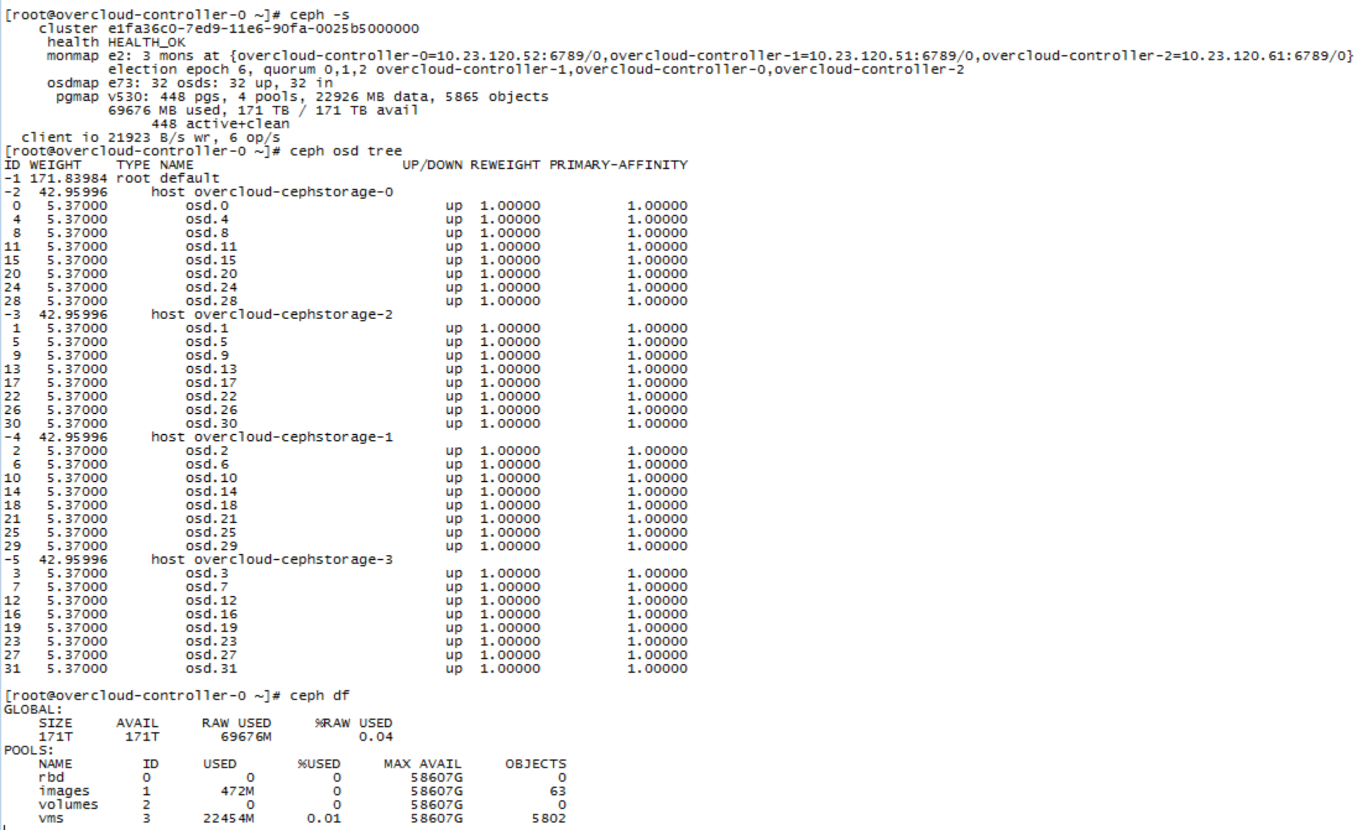
```
2016-09-19 00:11:09 [overcloud-CephStorageNodesPostDeployment-khpaqsy6woeh]:
UPDATE_COMPLETE  Stack UPDATE completed successfully
2016-09-19 00:11:09 [CephStorageNodesPostDeployment]: UPDATE_COMPLETE  state
changed
Stack overcloud UPDATE_COMPLETE
Overcloud Endpoint: http://172.22.215.16:5000/v2.0
Overcloud Deployed
```

## Post Deployment Health Checks

To perform the post-deployment health checks, complete the following steps:

Check with ironic and nova commands, the existence of the new node.

Check status of Ceph cluster.



```
[root@overcloud-controller-0 ~]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-controller-0,overcloud-controller-2
     osdmap e73: 32 osds: 32 up, 32 in
      pgmap v530: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69676 MB used, 171 TB / 171 TB avail
                 448 active+clean
  client io 21923 B/s wr, 6 op/s
[root@overcloud-controller-0 ~]# ceph osd tree
ID WEIGHT    TYPE NAME                            UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 171.83984 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                             up  1.00000          1.00000
 4   5.37000         osd.4                             up  1.00000          1.00000
 8   5.37000         osd.8                             up  1.00000          1.00000
11   5.37000         osd.11                            up  1.00000          1.00000
15   5.37000         osd.15                            up  1.00000          1.00000
20   5.37000         osd.20                            up  1.00000          1.00000
24   5.37000         osd.24                            up  1.00000          1.00000
28   5.37000         osd.28                            up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                             up  1.00000          1.00000
 5   5.37000         osd.5                             up  1.00000          1.00000
 9   5.37000         osd.9                             up  1.00000          1.00000
13   5.37000         osd.13                            up  1.00000          1.00000
17   5.37000         osd.17                            up  1.00000          1.00000
22   5.37000         osd.22                            up  1.00000          1.00000
26   5.37000         osd.26                            up  1.00000          1.00000
30   5.37000         osd.30                            up  1.00000          1.00000
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2                             up  1.00000          1.00000
 6   5.37000         osd.6                             up  1.00000          1.00000
10   5.37000         osd.10                            up  1.00000          1.00000
14   5.37000         osd.14                            up  1.00000          1.00000
18   5.37000         osd.18                            up  1.00000          1.00000
21   5.37000         osd.21                            up  1.00000          1.00000
25   5.37000         osd.25                            up  1.00000          1.00000
29   5.37000         osd.29                            up  1.00000          1.00000
-5  42.95996     host overcloud-cephstorage-3
 3   5.37000         osd.3                             up  1.00000          1.00000
 7   5.37000         osd.7                             up  1.00000          1.00000
12   5.37000         osd.12                            up  1.00000          1.00000
16   5.37000         osd.16                            up  1.00000          1.00000
19   5.37000         osd.19                            up  1.00000          1.00000
23   5.37000         osd.23                            up  1.00000          1.00000
27   5.37000         osd.27                            up  1.00000          1.00000
31   5.37000         osd.31                            up  1.00000          1.00000

[root@overcloud-controller-0 ~]# ceph df
GLOBAL:
    SIZE     AVAIL     RAW USED     %RAW USED
    171T      171T       69676M          0.04
POOLS:
    NAME        ID     USED       %USED     MAX AVAIL     OBJECTS
    rbd         0          0          0        58607G           0
    images      1       472M          0        58607G          63
    volumes     2          0          0        58607G           0
    vms         3     22454M       0.01        58607G        5802
```

This completes the addition of storage node in the cluster.

# Scale Up Compute Nodes

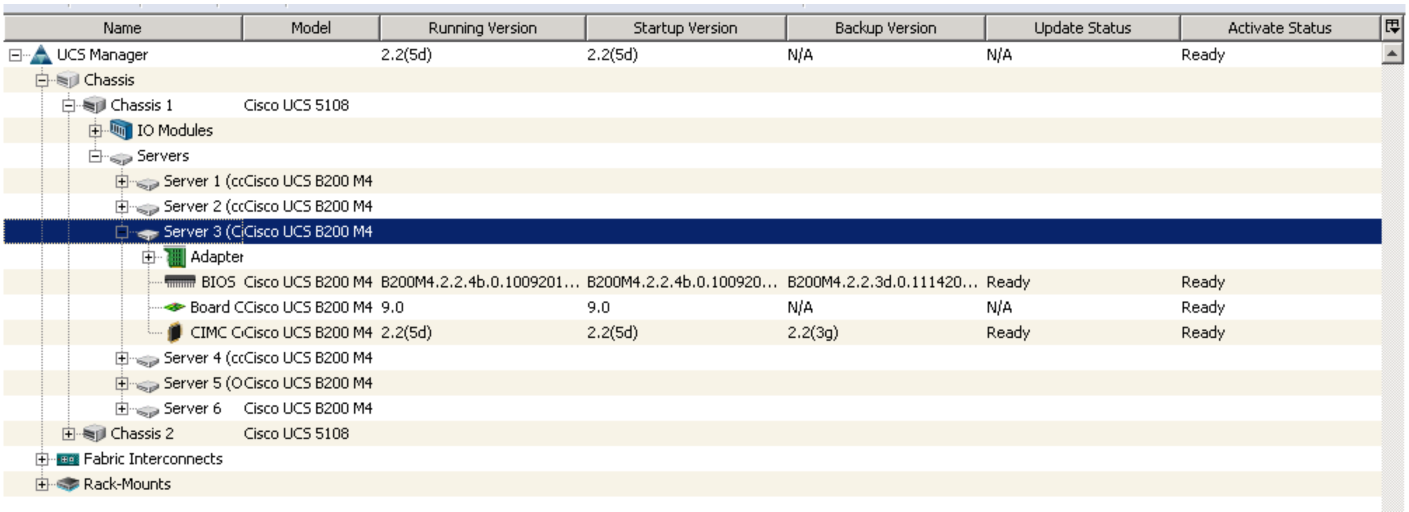
## Provision the New Blade Server in Cisco UCS

Insert the new Cisco UCS B200 M4 blade server into an empty slot in the chassis with similar configuration of local disks.

Refer to this section above for creating service profiles from Compute template. Create a new service profile from the template. Unbind the template and remove the storage policy that was attached to it earlier and associate the service profile to the server.

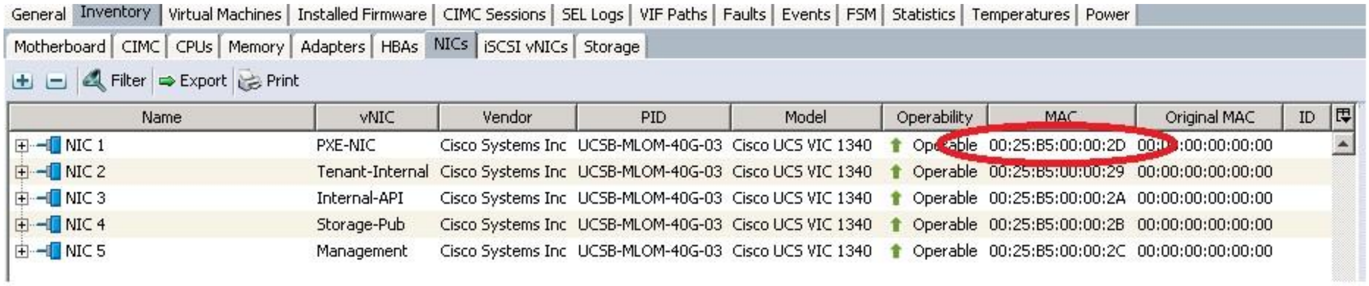
Upgrade firmware if needed.

Check the installed firmware on the new node and make sure that it is upgraded to the same version as other compute nodes.

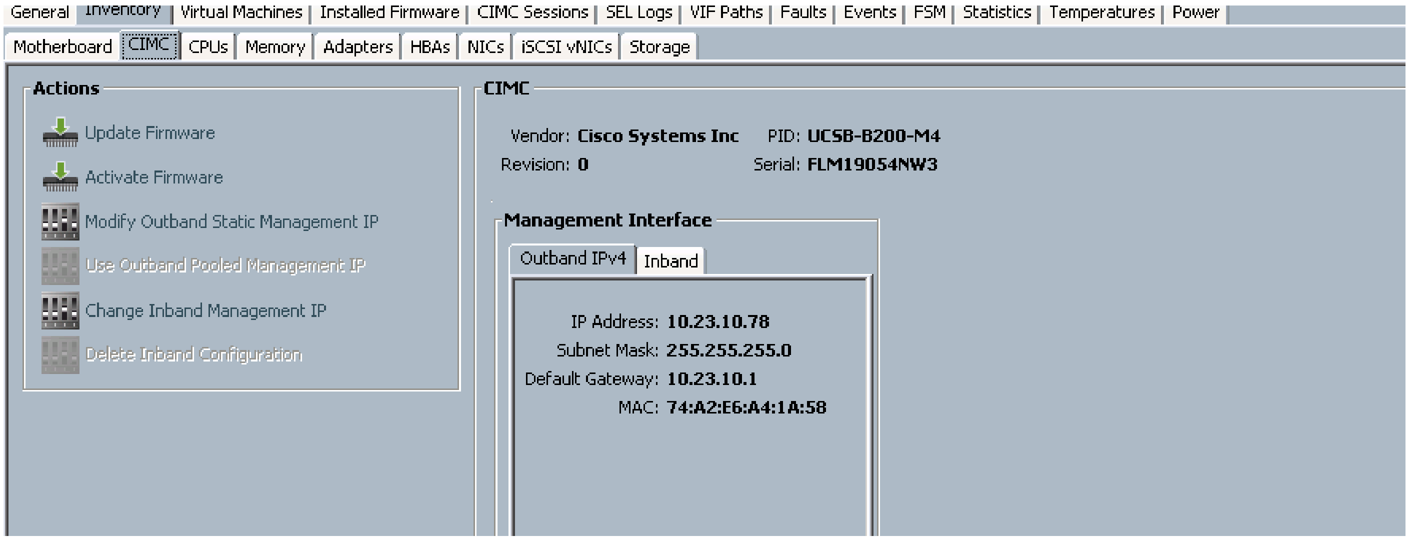


| Name | Model | Running Version | Startup Version | Backup Version | Update Status | Activate Status |
|---|---|---|---|---|---|---|
| UCS Manager | | 2.2(5d) | 2.2(5d) | N/A | N/A | Ready |
| Chassis | | | | | | |
| Chassis 1 | Cisco UCS 5108 | | | | | |
| IO Modules | | | | | | |
| Servers | | | | | | |
| Server 1 (cc | Cisco UCS B200 M4 | | | | | |
| Server 2 (cc | Cisco UCS B200 M4 | | | | | |
| Server 3 (C | Cisco UCS B200 M4 | | | | | |
| Adapter | | | | | | |
| BIOS | Cisco UCS B200 M4 | B200M4.2.2.4b.0.1009201… | B200M4.2.2.4b.0.100920… | B200M4.2.2.3d.0.111420… | Ready | Ready |
| Board C | Cisco UCS B200 M4 | 9.0 | 9.0 | N/A | N/A | Ready |
| CIMC C | Cisco UCS B200 M4 | 2.2(5d) | 2.2(5d) | 2.2(3g) | Ready | Ready |
| Server 4 (cc | Cisco UCS B200 M4 | | | | | |
| Server 5 (O | Cisco UCS B200 M4 | | | | | |
| Server 6 | Cisco UCS B200 M4 | | | | | |
| Chassis 2 | Cisco UCS 5108 | | | | | |
| Fabric Interconnects | | | | | | |
| Rack-Mounts | | | | | | |

Get the hardware inventory details needed for introspection. This include IPMI address, Provisioning MAC address, Boot Lun size, CPU and memory.


General | Inventory | Virtual Machines | Installed Firmware | CIMC Sessions | SEL Logs | VIF Paths | Faults | Events | FSM | Statistics | Temperatures | Power
Motherboard | CIMC | CPUs | Memory | Adapters | HBAs | NICs | iSCSI vNICs | Storage
Filter
Export
Print

| Name | vNIC | Vendor | PID | Model | Operability | MAC | Original MAC | ID |
|---|---|---|---|---|---|---|---|---|
| NIC 1 | PXE-NIC | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:2D | 00:00:00:00:00:00 | |
| NIC 2 | Tenant-Internal | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:29 | 00:00:00:00:00:00 | |
| NIC 3 | Internal-API | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:2A | 00:00:00:00:00:00 | |
| NIC 4 | Storage-Pub | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:2B | 00:00:00:00:00:00 | |
| NIC 5 | Management | Cisco Systems Inc | UCSB-MLOM-40G-03 | Cisco UCS VIC 1340 | Operable | 00:25:B5:00:00:2C | 00:00:00:00:00:00 | |



CIMC

Vendor: **Cisco Systems Inc** PID: **UCSB-B200-M4**
Revision: **0** Serial: **FLM19054NW3**

Management Interface

IP Address: **10.23.10.78**
Subnet Mask: **255.255.255.0**
Default Gateway: **10.23.10.1**
MAC: **74:A2:E6:A4:1A:58**

Collect the hardware inventory to create the json file for introspection.

Check the boot policy of the server. Validate that this is same as other compute nodes too. Should be LAN PXE first followed by local lun.

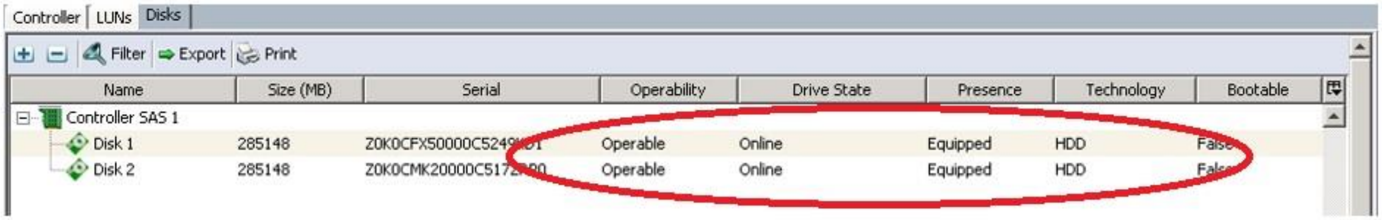
Update cisco-plugins.yaml file with details about this new server. Update the Tenant NIC address as below.

Append an entry in UCSM Host list and Nexus Switches entries in Cisco Plugins as below

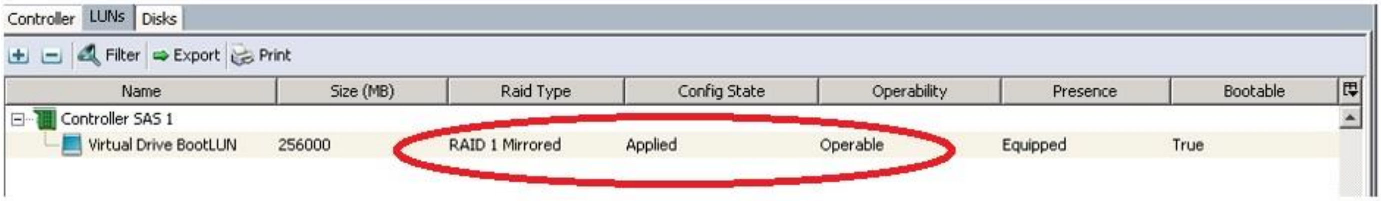
```
00:25:b5:00:00:29:org-root/org-osp8/ls-Openstack_Compute_Node4,
            },
            "00:25:b5:00:00:29": {
                "ports": "port-channel:17,port-channel:18"
            }
```

Make sure that this Service Profile is not bound to its template and check the order of NICs as below.

Check the status of the boot lun and make sure that the local disk config policy is no-disk policy.



| Name | Size (MB) | Serial | Operability | Drive State | Presence | Technology | Bootable |
|---|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | | |
| Disk 1 | 285148 | Z0K0CFX50000C5249[illegible] | Operable | Online | Equipped | HDD | False |
| Disk 2 | 285148 | Z0K0CMK20000C5172[illegible] | Operable | Online | Equipped | HDD | False |



| Name | Size (MB) | Raid Type | Config State | Operability | Presence | Bootable |
|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | |
| Virtual Drive BootLUN | 256000 | RAID 1 Mirrored | Applied | Operable | Equipped | True |

With the above the server is ready for introspection and Overcloud deploy.

## Run Introspection

To run Introspection, complete the following steps:

Prepare json file for introspection:

```
[stack@osp8-director ~]$ cat compute-new.json
{
  "nodes": [
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.78",
      "mac": [
        "00:25:b5:00:00:2d"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "40"
    }
  ]
}
```

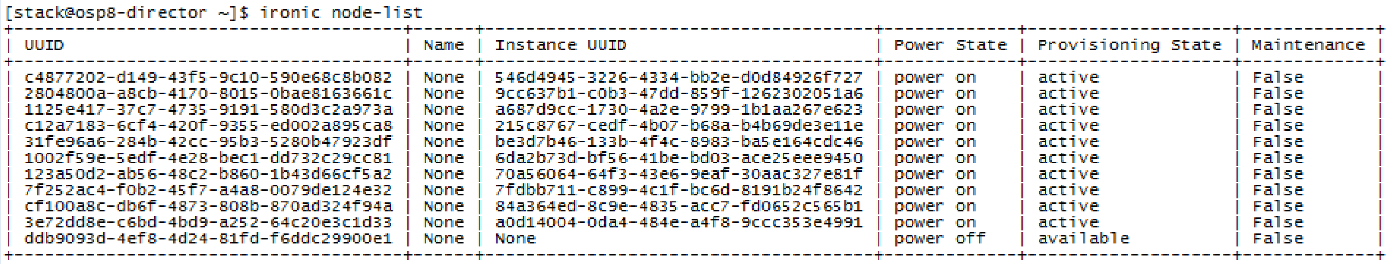
Check IPMI Connectivity:

[stack@osp8-director ~]$ ipmitool -I lanplus -H 10.23.10.78 -U admin -P <passwd> chassis power off

Chassis Power Control: Down/Off

`[stack@osp8-director ~]$` Run discovery and introspection:

```
[stack@osp8-director ~]$ openstack baremetal import --json ~/compute-new.json
[stack@osp8-director ~]$ openstack baremetal configure boot
[stack@osp8-director ~]$ ironic node-list
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-1262302051a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
| ddb9093d-4ef8-4d24-81fd-f6ddc29900e1 | None | None                                 | power off   | available          | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
```

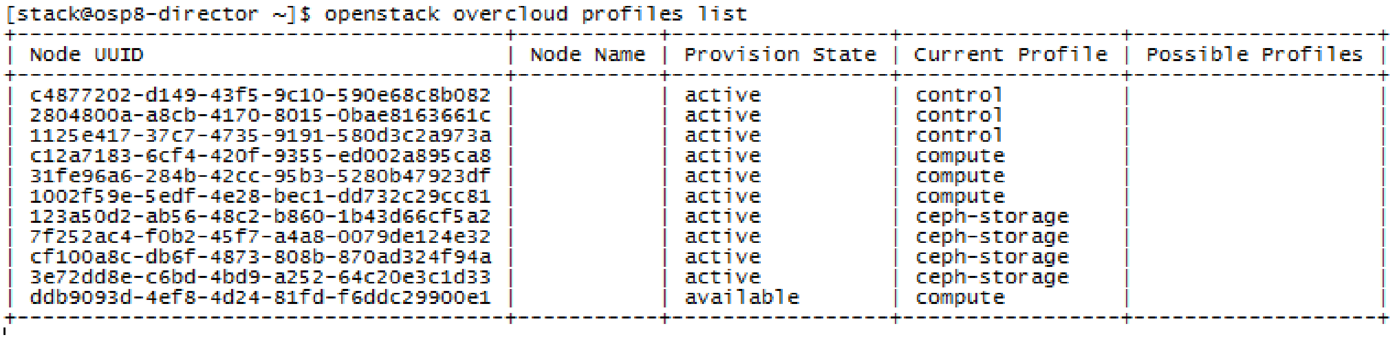
```
[stack@osp8-director ~]$ ironic node-set-maintenance ddb9093d-4ef8-4d24-81fd-
f6ddc29900e1 true
[stack@osp8-director ~]$ openstack baremetal introspection start ddb9093d-4ef8-
4d24-81fd-f6ddc29900e1
 [stack@osp8-director ~]$ openstack baremetal introspection status ddb9093d-4ef8-
4d24-81fd-f6ddc29900e1
+----------+-------+
| Field    | Value |
+----------+-------+
| error    | None  |
| finished | True  |
+----------+-------+
```

Wait till the introspection is complete. The status command should yield finished as True and Error as none. Alternatively open a KVM console to observe the status of introspection.

```
[stack@osp8-director ~]$ ironic node-set-maintenance ddb9093d-4ef8-4d24-81fd-
f6ddc29900e1 false
```

Update node properties

```
[stack@osp8-director ~]$ ironic node-update ddb9093d-4ef8-4d24-81fd-f6ddc29900e1
\
> add properties/capabilities='profile:compute,boot_option:local'
```



```
[stack@osp8-director ~]$ openstack overcloud profiles list
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| Node UUID                            | Node Name | Provision State | Current Profile | Possible Profiles |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
| c4877202-d149-43f5-9c10-590e68c8b082 |           | active          | control         |                   |
| 2804800a-a8cb-4170-8015-0bae8163661c |           | active          | control         |                   |
| 1125e417-37c7-4735-9191-580d3c2a973a |           | active          | control         |                   |
| c12a7183-6cf4-420f-9355-ed002a895ca8 |           | active          | compute         |                   |
| 31fe96a6-284b-42cc-95b3-5280b47923df |           | active          | compute         |                   |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 |           | active          | compute         |                   |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 |           | active          | ceph-storage    |                   |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 |           | active          | ceph-storage    |                   |
| cf100a8c-db6f-4873-808b-870ad324f94a |           | active          | ceph-storage    |                   |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 |           | active          | ceph-storage    |                   |
| ddb9093d-4ef8-4d24-81fd-f6ddc29900e1 |           | available       | compute         |                   |
+--------------------------------------+-----------+-----------------+-----------------+-------------------+
'
```

Check the status of added entries with ironic node-show. Repeat the above to all nodes that you would like to add as overcloud deploy can add all of these in a single go.

## Run Overcloud Deploy

Run the Overcloud deployment command. The number of compute nodes has been incremented to 4 from 3 earlier. Here the number '4' indicates the total number of storage nodes in Overcloud.

```
[stack@osp8-director ~]$ cat run_compute.sh
#!/bin/bash
openstack overcloud deploy --templates --compute-scale 4 \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-
isolation.yaml \
-e /home/stack/templates/network-environment.yaml \
-e /home/stack/templates/network-management.yaml \
-e /home/stack/templates/storage-environment.yaml \
-e /home/stack/templates/timezone.yaml \
-e /home/stack/templates/cisco-plugins.yaml \
--log-file overcloud_compute.log

2016-09-19 18:40:56 [overcloud]: UPDATE_COMPLETE  Stack UPDATE completed
successfully
Stack overcloud UPDATE_COMPLETE
Overcloud Endpoint: http://172.22.215.16:5000/v2.0
Overcloud Deployed
```

## Post Deployment and Health Checks

To perform the deployment and health checks, complete the following steps:

Login to each controller node and check for the existence of the new compute node in /etc/neutron/plugin.ini. If not please add in each Nexus Switch section and also in UCSM host list in plugin.ini file. Make sure to make the changes across all the controller nodes.

```
[root@overcloud-controller-0 ~]# grep compute-3 /etc/neutron/plugin.ini
overcloud-compute-3.localdomain=port-channel:17,port-channel:18
overcloud-compute-3.localdomain=port-channel:17,port-channel:18
```

```
ucsm_host_list=overcloud-compute-3.localdomain:org-root/org-osp8/ls-
Openstack_Compute_Node4,…
```

Restart Neutron

```
pcs resource restart neutron-server
```

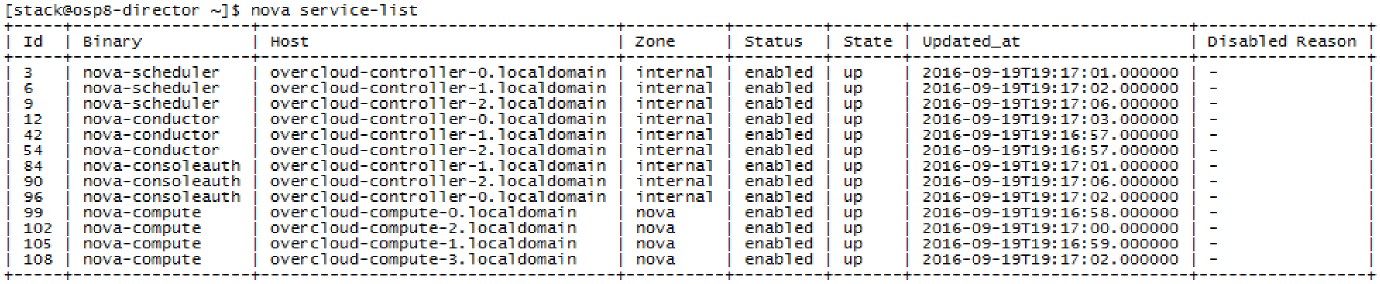
Restart nova-services as a post deployment.

```
pcs resource restart openstack-nova-scheduler
pcs resource restart openstack-nova-consoleauth
pcs resource restart openstack-nova-api
```

Check the status of PCS cluster and restart if needed with pcs resource cleanup.

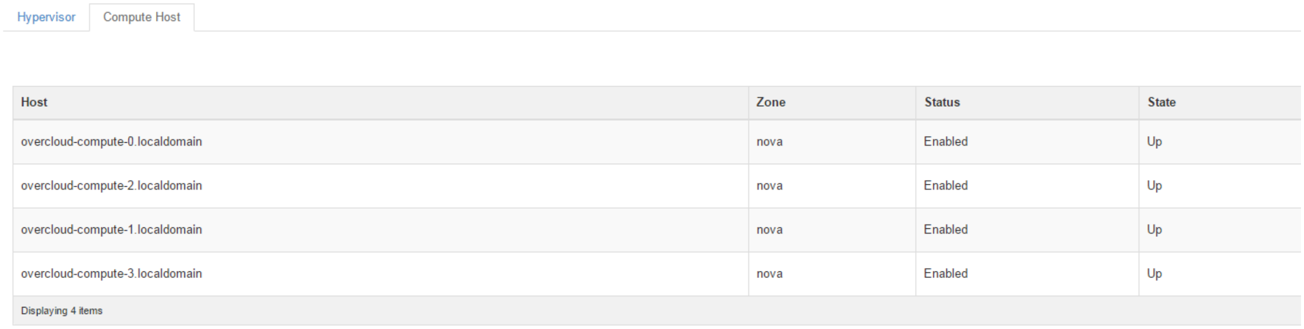
Check the status through ironic node-list and nova list.

Check with nova service-list after sourcing overcloudrc.



```
[stack@osp8-director ~]$ nova service-list
+-----+------------------+-----------------------------------+----------+---------+-------+----------------------------+-----------------+
| Id  | Binary           | Host                              | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+-----+------------------+-----------------------------------+----------+---------+-------+----------------------------+-----------------+
| 3   | nova-scheduler   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T19:17:01.000000 | -               |
| 6   | nova-scheduler   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T19:17:02.000000 | -               |
| 9   | nova-scheduler   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T19:17:06.000000 | -               |
| 12  | nova-conductor   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T19:17:03.000000 | -               |
| 42  | nova-conductor   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T19:16:57.000000 | -               |
| 54  | nova-conductor   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T19:16:57.000000 | -               |
| 84  | nova-consoleauth | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T19:17:01.000000 | -               |
| 90  | nova-consoleauth | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T19:17:06.000000 | -               |
| 96  | nova-consoleauth | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T19:17:02.000000 | -               |
| 99  | nova-compute     | overcloud-compute-0.localdomain    | nova     | enabled | up    | 2016-09-19T19:16:58.000000 | -               |
| 102 | nova-compute     | overcloud-compute-2.localdomain    | nova     | enabled | up    | 2016-09-19T19:17:00.000000 | -               |
| 105 | nova-compute     | overcloud-compute-1.localdomain    | nova     | enabled | up    | 2016-09-19T19:16:59.000000 | -               |
| 108 | nova-compute     | overcloud-compute-3.localdomain    | nova     | enabled | up    | 2016-09-19T19:17:02.000000 | -               |
+-----+------------------+-----------------------------------+----------+---------+-------+----------------------------+-----------------+
```

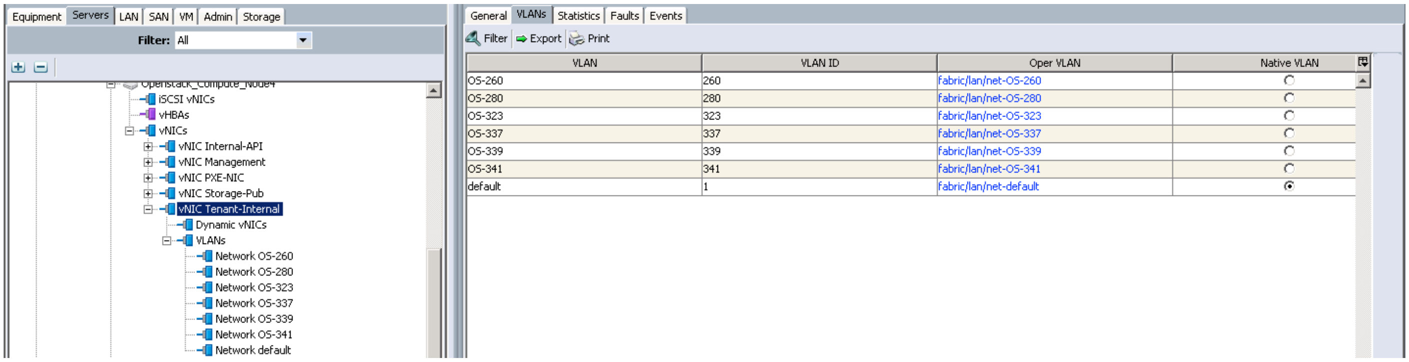
Log into dashboard to check the status of the new node added.



| Host | Zone | Status | State |
|---|---|---|---|
| overcloud-compute-0.localdomain | nova | Enabled | Up |
| overcloud-compute-2.localdomain | nova | Enabled | Up |
| overcloud-compute-1.localdomain | nova | Enabled | Up |
| overcloud-compute-3.localdomain | nova | Enabled | Up |

Displaying 4 items

The System should be up and running and will deploy VMs on the newly added node.
Create few VMs to make sure that the newly added compute host gets few VMs and the plugins are working fine.

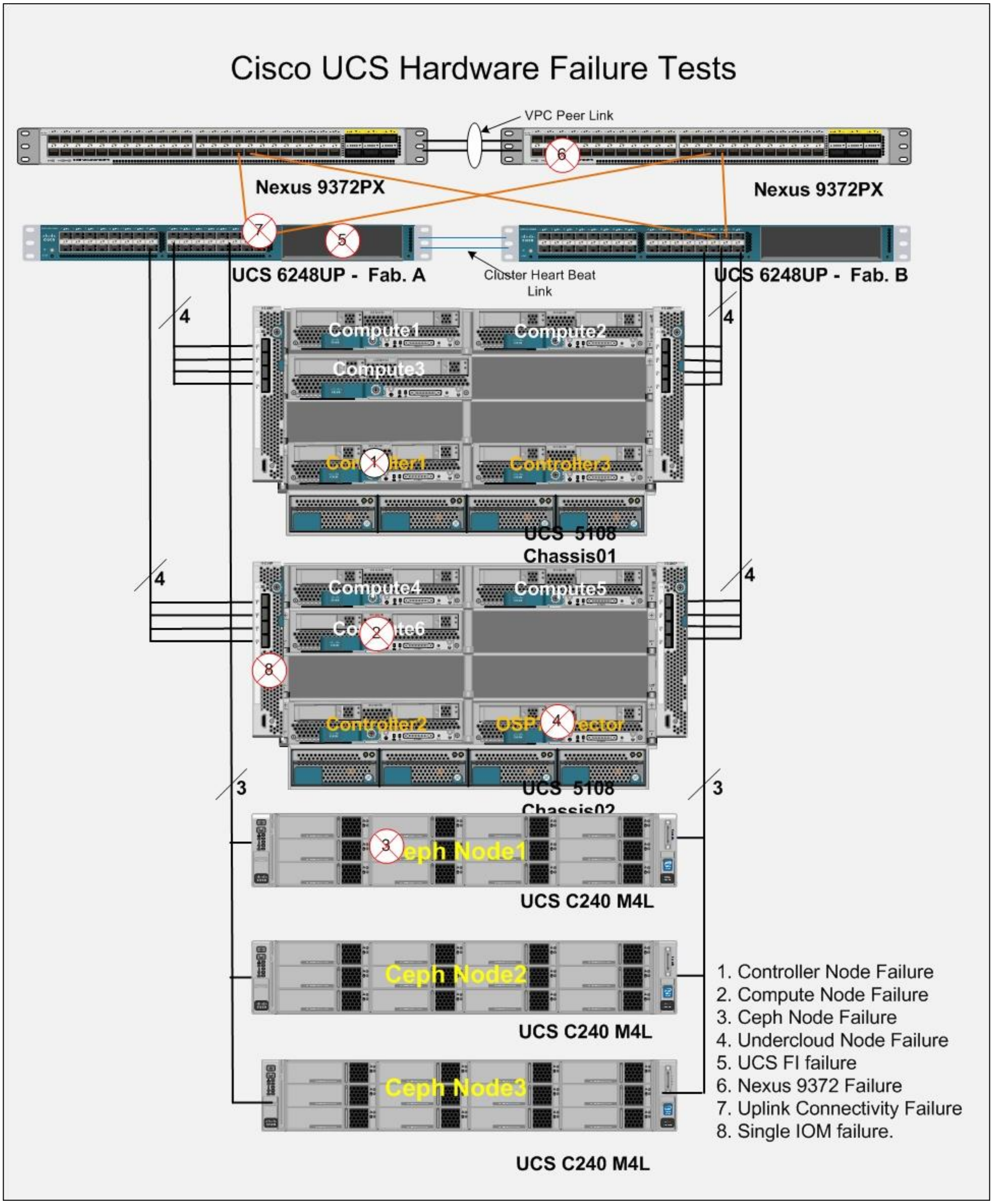


| VLAN | VLAN ID | Oper VLAN | Native VLAN |
|---|---|---|---|
| OS-260 | 260 | fabric/lan/net-OS-260 | ○ |
| OS-280 | 280 | fabric/lan/net-OS-280 | ○ |
| OS-323 | 323 | fabric/lan/net-OS-323 | ○ |
| OS-337 | 337 | fabric/lan/net-OS-337 | ○ |
| OS-339 | 339 | fabric/lan/net-OS-339 | ○ |
| OS-341 | 341 | fabric/lan/net-OS-341 | ○ |
| default | 1 | fabric/lan/net-default | ● |

```
[root@overcloud-controller-0 tmp]# openstack server list --all-projects
+--------------------------------------+--------------------+--------+----------------------------------------+
| ID                                   | Name               | Status | Networks                               |
+--------------------------------------+--------------------+--------+----------------------------------------+
| 2d178908-a72a-41ea-985e-782a924145de | tenant305_155_inst4 | ACTIVE | tenant305-155=10.20.155.4, 10.23.160.54 |
| bb14db94-1fe7-4c0e-a65a-88afbe36a6a0 | tenant305_155_inst3 | ACTIVE | tenant305-155=10.20.155.3, 10.23.160.53 |
| 9cb357c6-6f99-476f-a19b-0ebc128a189a | tenant305_105_inst2 | ACTIVE | tenant305-105=10.20.105.4, 10.23.160.52 |
| 327ab308-be2a-4012-8105-55af58887add | tenant305_105_inst1 | ACTIVE | tenant305-105=10.20.105.3, 10.23.160.51 |
| b3c59d3c-3ab5-4394-a32c-2e274f63c067 | tenant304_154_inst4 | ACTIVE | tenant304-154=10.20.154.4, 10.23.160.49 |
| 943f75fd-efc3-45f8-88b0-0370dda683ae | tenant304_154_inst3 | ACTIVE | tenant304-154=10.20.154.5, 10.23.160.48 |
| 4639ed99-7c84-46b5-a720-d35839f4028f | tenant304_104_inst2 | ACTIVE | tenant304-104=10.20.104.4, 10.23.160.47 |
| 5b674eeb-d7a6-4f42-bbca-95287eaedefd | tenant304_104_inst1 | ACTIVE | tenant304-104=10.20.104.3, 10.23.160.46 |
| 7d65c810-4316-411e-a19c-26a15d1c7536 | tenant303_153_inst4 | ACTIVE | tenant303-153=10.20.153.4, 10.23.160.44 |
| 9633037c-48b2-4eda-9a90-97a34684f9c6 | tenant303_153_inst3 | ACTIVE | tenant303-153=10.20.153.3, 10.23.160.43 |
| 535842d7-ede2-4cb2-a016-3eeb6ee16513 | tenant303_103_inst2 | ACTIVE | tenant303-103=10.20.103.4, 10.23.160.42 |
| 7ec6687b-e05c-451f-8e9c-c83a16e5bcfd | tenant303_103_inst1 | ACTIVE | tenant303-103=10.20.103.3, 10.23.160.41 |
| 412b6207-c8d8-4fc6-84d0-a70e647aafaf | tenant302_152_inst4 | ACTIVE | tenant302-152=10.20.152.4, 10.23.160.39 |
| f7f894e8-8298-4790-b3b3-f05bc8504967 | tenant302_152_inst3 | ACTIVE | tenant302-152=10.20.152.5, 10.23.160.38 |
| 1f283b9a-c4da-4e45-a359-0f15d1564777 | tenant302_102_inst2 | ACTIVE | tenant302-102=10.20.102.4, 10.23.160.37 |
| 0c65fe52-61e8-4cdf-8874-48bf00bf677c | tenant302_102_inst1 | ACTIVE | tenant302-102=10.20.102.3, 10.23.160.36 |
| d6fc1555-7a9d-4729-a6da-1b2a6413a87b | tenant301_151_inst4 | ACTIVE | tenant301-151=10.20.151.4, 10.23.160.34 |
| e0466971-3b9e-4693-9cac-4845e54b446c | tenant301_151_inst3 | ACTIVE | tenant301-151=10.20.151.3, 10.23.160.33 |
| 0b056f6b-d88f-42de-a2af-9fe70055559e | tenant301_101_inst2 | ACTIVE | tenant301-101=10.20.101.4, 10.23.160.32 |
| 3e236c78-4f51-4bae-92ae-cfdfbf964317 | tenant301_101_inst1 | ACTIVE | tenant301-101=10.20.101.3, 10.23.160.31 |
+--------------------------------------+--------------------+--------+----------------------------------------+

[root@overcloud-controller-0 tmp]# nova-manage vm list | grep overcloud-compute-3 | grep active | awk '{print $1" "$2" "$3" "$4}'
No handlers could be found for logger "oslo_config.cfg"
tenant305_155_inst4 overcloud-compute-3.localdomain m1.small active
tenant304_154_inst4 overcloud-compute-3.localdomain m1.small active
tenant303_153_inst4 overcloud-compute-3.localdomain m1.small active
tenant302_152_inst3 overcloud-compute-3.localdomain m1.small active
tenant302_102_inst1 overcloud-compute-3.localdomain m1.small active
tenant301_101_inst1 overcloud-compute-3.localdomain m1.small active
```

# High Availability

Both the hardware and software stack are injected with faults to trigger a failure of a running process on a node or an unavailability of hardware for a short or extended period of time. With the fault in place the functional validations are done as mentioned above. The purpose is to achieve business continuity without interruption to the clients. However, performance degradation is inevitable and has been documented wherever it was captured as part of the tests.



# High Availability of Software Stack

## OpenStack Services

The status of OpenStack services were checked with pcs status as below on Controller Node:

```
[root@overcloud-controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Mon Sep 19 06:37:03 2016          Last change: Sun Sep 18 16:56:59
2016 by root via cibadmin on overcloud-controller-0
Stack: corosync
Current DC: overcloud-controller-2 (version 1.1.13-10.el7_2.2-44eb2dd) - partition
with quorum
3 nodes and 113 resources configured

Online: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]

Full list of resources:

 ip-10.23.110.75        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
 Master/Slave Set: redis-master [redis]
     Masters: [ overcloud-controller-2 ]
     Slaves: [ overcloud-controller-0 overcloud-controller-1 ]
 Master/Slave Set: galera-master [galera]
     Masters: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: mongod-clone [mongod]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: rabbitmq-clone [rabbitmq]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: memcached-clone [memcached]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 ip-10.23.100.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-2
 Clone Set: openstack-nova-scheduler-clone [openstack-nova-scheduler]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-l3-agent-clone [neutron-l3-agent]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-alarm-notifier-clone [openstack-ceilometer-alarm-
notifier]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
```

```
 Clone Set: openstack-heat-engine-clone [openstack-heat-engine]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-api-clone [openstack-ceilometer-api]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-metadata-agent-clone [neutron-metadata-agent]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-ovs-cleanup-clone [neutron-ovs-cleanup]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-netns-cleanup-clone [neutron-netns-cleanup]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-heat-api-clone [openstack-heat-api]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-cinder-scheduler-clone [openstack-cinder-scheduler]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-nova-api-clone [openstack-nova-api]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-heat-api-cloudwatch-clone [openstack-heat-api-cloudwatch]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-collector-clone [openstack-ceilometer-collector]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-notification-clone [openstack-ceilometer-
notification]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-dhcp-agent-clone [neutron-dhcp-agent]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-glance-api-clone [openstack-glance-api]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-openvswitch-agent-clone [neutron-openvswitch-agent]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-nova-novncproxy-clone [openstack-nova-novncproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: delay-clone [delay]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: httpd-clone [httpd]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-keystone-clone [openstack-keystone]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-nova-consoleauth-clone [openstack-nova-consoleauth]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-glance-registry-clone [openstack-glance-registry]
```

```
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-cinder-api-clone [openstack-cinder-api]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-central-clone [openstack-ceilometer-central]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: neutron-server-clone [neutron-server]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-ceilometer-alarm-evaluator-clone [openstack-ceilometer-alarm-
evaluator]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 Clone Set: openstack-heat-api-cfn-clone [openstack-heat-api-cfn]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 openstack-cinder-volume        (systemd:openstack-cinder-volume):      Started
overcloud-controller-0
 Clone Set: openstack-nova-conductor-clone [openstack-nova-conductor]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
 ucs-fence-controller   (stonith:fence_cisco_ucs):      Started overcloud-
controller-1

PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Online
  overcloud-controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

Few identified services running on these nodes were either restarted or killed and/or rebooted the nodes.
For eg.
Master/Slave Set: redis-master [redis]
  Masters: [ overcloud-controller-2 ]
  Slaves: [ overcloud-controller-0 overcloud-controller-1 ]

Per above redis master is overcloud-controller-2. This node was rebooted and observed the behavior while the node getting rebooted and any impact on VMs.

The Ceph node monitors and services were also restarted to test any interruption of volume creation and booting of the VMs, but no issues were observed.

# High Availability of Hardware Stack

## HA of Fabric Interconnects

### FI Reboot Tests

Cisco UCS Fabric Interconnects work in pair with inbuilt HA. While both of them serve traffic during a normal operation, a surviving member can still keep the system up and running. Depending on the overprovisioning used in the deployment a degradation in performance may be expected.

An effort is made to reboot the Fabric one after the other and do functional tests as mentioned earlier.

- Check the status of FI's

```
Check the status of the UCS Fabric Cluster before reboot
UCS-OSP8-FAB-B# show cluster extended-state
Cluster Id: 0x3bbf9944066711e5-0xa8888c604f640804

Start time: Wed May 18 08:38:27 2016
Last election time: Wed May 18 09:13:24 2016

B: UP, PRIMARY
A: UP, SUBORDINATE

B: memb state UP, lead state PRIMARY, mgmt services state: UP
A: memb state UP, lead state SUBORDINATE, mgmt services state: UP
   heartbeat state PRIMARY_OK

INTERNAL NETWORK INTERFACES:
eth1, UP
eth2, UP

HA READY ←--System should be in HA ready before invoking any of the HA
tests on Fabrics.
Detailed state of the device selected for HA storage:
Chassis 1, serial: FOX1832G67B, state: active
Chassis 2, serial: FOX1831G2L5, state: active
Server 2, serial: FCH1913V0VJ, state: active
```

- Check the status of OpenStack PCS Cluster before reboot

```
Current DC: overcloud-controller-2 (version 1.1.13-10.el7_2.2-44eb2dd) -
partition with quorum
3 nodes and 113 resources configured
Online: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]
Full list of resources:
 ip-10.23.110.75        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-
controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-2
```

```
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
```

- Check the status of VMs.
- Run script to login to all the VMs.

```
[stack@osp8-director scripts]$ date;./tenantips.sh;date
Tue Aug 23 19:38:42 PDT 2016
        inet 10.20.191.5  netmask 255.255.255.0  broadcast 10.20.191.255
        inet 10.20.191.4  netmask 255.255.255.0  broadcast 10.20.191.255
```

- Reboot Fabric B ( primary )

  Log into UCS Fabric Command Line Interface and reboot the Fabric

```
Connect Local Management
UCS-OSP8-FAB-B(local-mgmt)# reboot
Before rebooting, please take a configuration backup.
Do you still want to reboot? (yes/no):yes
nohup: ignoring input and appending output to `nohup.out'

Broadcast message from root (Tue Aug 23 19:13:21 2016):

All shells being terminated due to system /sbin/reboot
Connection to 10.23.10.7 closed
```

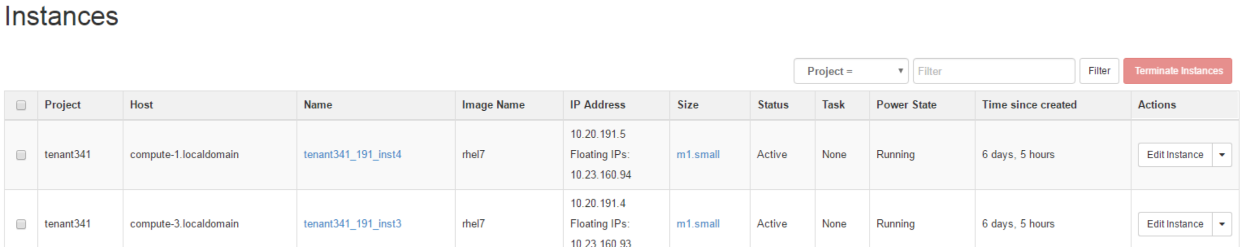
### Health Checks and Observations

The following is a list of health checks and observations:

- Check for VIP and Fabric A pings. Both should be down immediately. VIP recovers after a couple of minutes

```
show UCS-OSP8-FAB-A# show cluster extended-state
Cluster Id: 0x3bbf9944066711e5-0xa8888c604f640804

Start time: Wed May 18 09:12:15 2016
Last election time: Tue Aug 23 19:15:51 2016

A: UP, PRIMARY, (Management services: INIT IN PROGRESS)
B: DOWN, INAPPLICABLE

A: memb state UP, lead state PRIMARY, mgmt services state: INVALID
B: memb state DOWN, lead state INAPPLICABLE, mgmt services state: DOWN
   heartbeat state SECONDARY_FAILED

INTERNAL NETWORK INTERFACES:
eth1, UP
eth2, UP

HA NOT READY
Management services: initialization in progress on local Fabric Intercon-
nect
Detailed state of the device selected for HA storage:
Chassis 1, serial: FOX1832G67B, state: active
Chassis 2, serial: FOX1831G2L5, state: active
Server 2, serial: FCH1913V0VJ, state: active
```

- Login to dashboard. System could be slower, but works fine.



Instances

| □ | Project | Host | Name | Image Name | IP Address | Size | Status | Task | Power State | Time since created | Actions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| □ | tenant341 | compute-1.localdomain | tenant341_191_inst4 | rhel7 | 10.20.191.5<br>Floating IPs:<br>10.23.160.94 | m1.small | Active | None | Running | 6 days, 5 hours | Edit Instance ▾ |
| □ | tenant341 | compute-3.localdomain | tenant341_191_inst3 | rhel7 | 10.20.191.4<br>Floating IPs:<br>10.23.160.93 | m1.small | Active | None | Running | 6 days, 5 hours | Edit Instance ▾ |

- Check for PCS Cluster status on one of the controller nodes. System could be slow in the beginning but should respond as follows:

```
PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Online
  overcloud-controller-2: Online
Perform a quick health check on creating VMs, Check the sanity checks on Nexus
switches too for any impact on port-channels because of Fab B is down.
```

- Create Virtual Machines

Perform a quick health check on creating VMs and checking the status of existing instances.

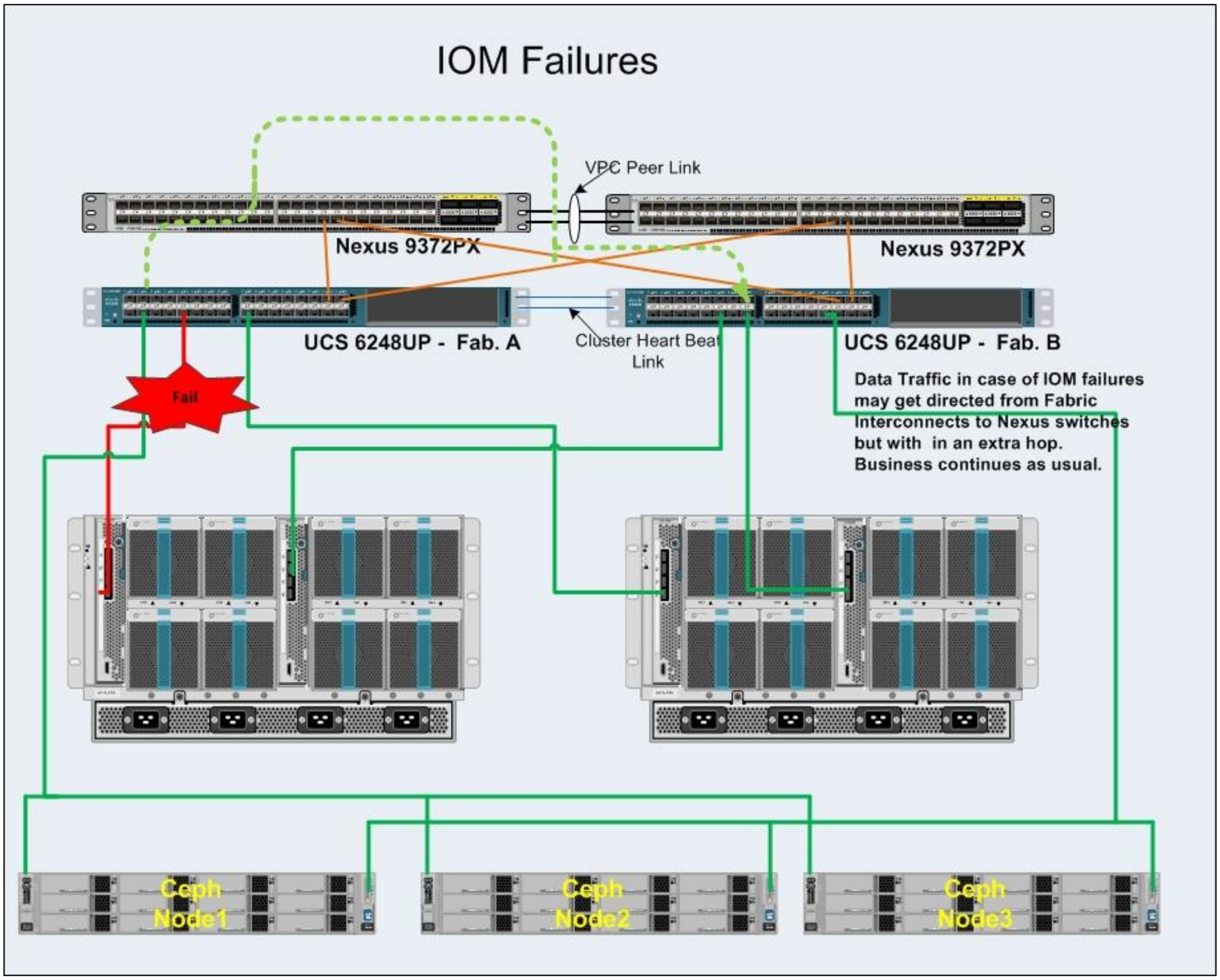
Fabric B might take around 15 minutes to come back online.

Reboot Fabric A

- Connect to the Fabric A now and check the cluster status. System should show HA READY before rebooting Fab A.
- Reboot Fab A by connecting to the local-mgmt similar to Fab B above.
- Perform the health check similar to the ones done on Fab B earlier.
- The test went fine with fence_cisco_ucs package ( patched ) in place.

## Hardware Failures of IO Modules

IO Module Failures seldom happen in UCS infrastructure and in most of the cases these are human mistakes. The failure tests were included just to validate the business continuity.


IOM Failures
VPC Peer Link
Nexus 9372PX
Nexus 9372PX
UCS 6248UP - Fab. A
Cluster Heart Beat Link
UCS 6248UP - Fab. B
Data Traffic in case of IOM failures may get directed from Fabric Interconnects to Nexus switches but with in an extra hop. Business continues as usual.
Fail
Ceph Node1
Ceph Node2
Ceph Node3

Multiple Tenants with multiple networks and Virtual machines were created. Identified the VMs belonging to the same tenant but with different networks and also going to different chassis.  One of the IO Modules was pulled out from the chassis and the L3 traffic validated.

## HA on Cisco Nexus Switches

Cisco Nexus switches are deployed in pairs and allow the upstream connectivity of the virtual machines to outside of the fabric. Cisco Nexus plugin creates VLANs on these switches both globally and on the port channel. The Nexus plugin replays these vlans or rebuilds the vlan information on the rebooted switch once it comes back up again. In order to test the HA of these switches multiple networks and instances were created and one of the switches were rebooted. The connectivity of the VMs through floating network checked and also the time it took for the plugins to replay was noted as below.

### Test Bed Setup before Injecting Fault

Nexus Switches

```
Software
  BIOS: version 07.17
  NXOS: version 7.0(3)I1(3)
  BIOS compile time:  09/10/2014
  NXOS image file is: bootflash:///n9000-dk9.7.0.3.I1.3.bin
  NXOS compile time:  8/21/2015 3:00:00 [08/21/2015 10:27:18]


Hardware
  cisco Nexus9000 C9372PX chassis
  Intel(R) Core(TM) i3-3227U C with 16402540 kB of memory.
  Processor Board ID SAL1913CBFP

  Device name: OSP8-N9K-FAB-B
  bootflash:   51496280 kB
Kernel uptime is 43 day(s), 19 hour(s), 18 minute(s), 29 second(s)

Last reset at 921588 usecs after  Tue Jun 21 23:31:14 2016

  Reason: Reset due to upgrade
  System version: 6.1(2)I3(2)
  Service:
```

N9K-A

```
interface port-channel17
  description OSP8-FAB-A
  switchport mode trunk
  switchport trunk allowed vlan 1,10,100,110,120,130,150,160,215
  switchport trunk allowed vlan add 255,257-258,262,264,266,268,271
…………………………………….
interface port-channel18
  description OSP8-FAB-A
  switchport mode trunk
  switchport trunk allowed vlan 1,10,100,110,120,130,150,160,215
  switchport trunk allowed vlan add 255,257-258,262,264,266,268,271
……………………………………..
```

N9K-B

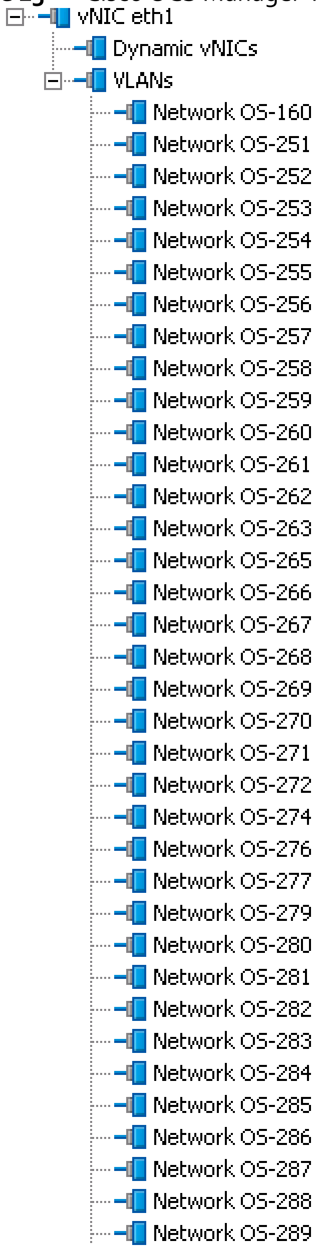
```
interface port-channel17
  description OSP8-FAB-B
  switchport mode trunk
  switchport trunk allowed vlan 1,10,100,110,120,130,150,160,215
  switchport trunk allowed vlan add 254-255,257-258,262,264,266,268
…………………….
interface port-channel18
  description OSP8-FAB-B
  switchport mode trunk
  switchport trunk allowed vlan 1,10,100,110,120,130,150,160,215
  switchport trunk allowed vlan add 254-255,257-258,262,264,266,268
  ………………………
```

Rebooted switch. The switch came up fine and the port-channel entries remained intact. The connectivity of the VMs went fine too.

```
Switch came up by Thu Aug  4 12:11:14 PDT 2016
Kernel uptime is 0 day(s), 0 hour(s), 2 minute(s), 1 second(s)

Last reset at 692395 usecs after  Thu Aug  4 18:53:23 2016
```

Repeated again on the other switch.

**Figure 13** Cisco UCS Manager VLANs



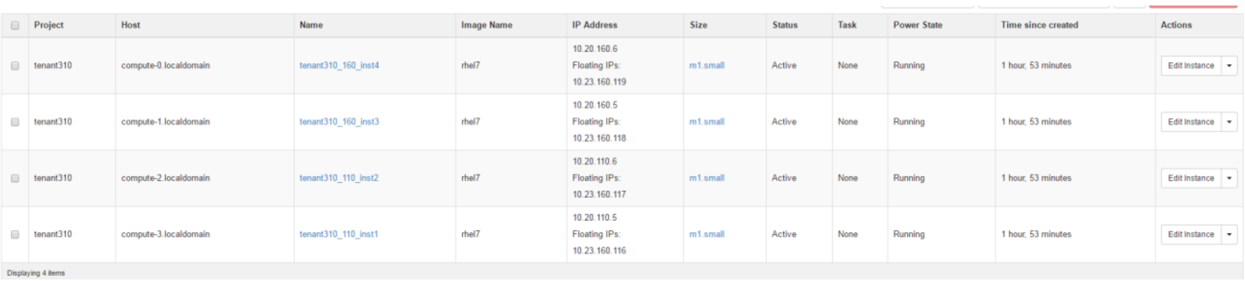
- vNIC eth1
  - Dynamic vNICs
  - VLANs
    - Network OS-160
    - Network OS-251
    - Network OS-252
    - Network OS-253
    - Network OS-254
    - Network OS-255
    - Network OS-256
    - Network OS-257
    - Network OS-258
    - Network OS-259
    - Network OS-260
    - Network OS-261
    - Network OS-262
    - Network OS-263
    - Network OS-265
    - Network OS-266
    - Network OS-267
    - Network OS-268
    - Network OS-269
    - Network OS-270
    - Network OS-271
    - Network OS-272
    - Network OS-274
    - Network OS-276
    - Network OS-277
    - Network OS-279
    - Network OS-280
    - Network OS-281
    - Network OS-282
    - Network OS-283
    - Network OS-284
    - Network OS-285
    - Network OS-286
    - Network OS-287
    - Network OS-288
    - Network OS-289

## Creating Virtual Machines

Creating tenant creates VLANs in compute nodes. However if a VM from one tenant is deleted, the VLAN on the computes will remain until the last vm of that tenant is deleted.

### Tenants

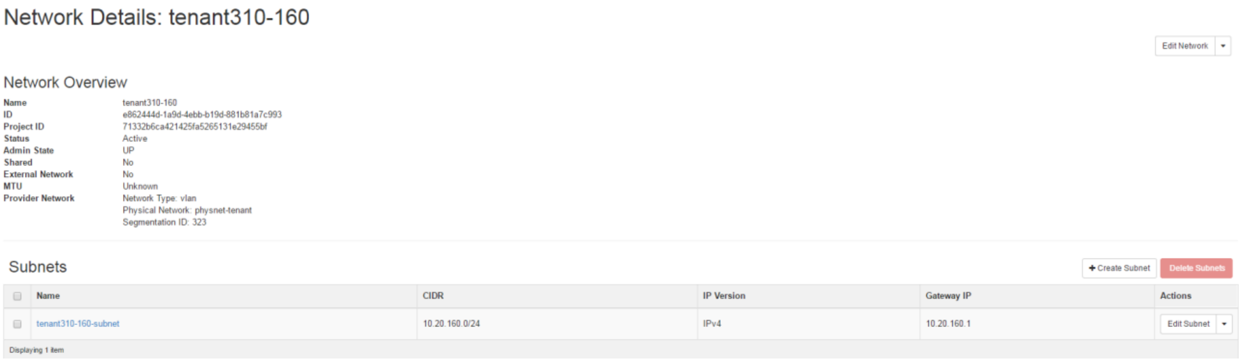
Delete one network from one tenant.



| ☐ | Project | Host | Name | Image Name | IP Address | Size | Status | Task | Power State | Time since created | Actions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ☐ | tenant310 | compute-0.localdomain | tenant310_160_inst4 | rhel7 | 10.20.160.6<br>Floating IPs:<br>10.23.160.119 | m1.small | Active | None | Running | 1 hour, 53 minutes | Edit Instance ▾ |
| ☐ | tenant310 | compute-1.localdomain | tenant310_160_inst3 | rhel7 | 10.20.160.5<br>Floating IPs:<br>10.23.160.118 | m1.small | Active | None | Running | 1 hour, 53 minutes | Edit Instance ▾ |
| ☐ | tenant310 | compute-2.localdomain | tenant310_110_inst2 | rhel7 | 10.20.110.6<br>Floating IPs:<br>10.23.160.117 | m1.small | Active | None | Running | 1 hour, 53 minutes | Edit Instance ▾ |
| ☐ | tenant310 | compute-3.localdomain | tenant310_110_inst1 | rhel7 | 10.20.110.5<br>Floating IPs:<br>10.23.160.116 | m1.small | Active | None | Running | 1 hour, 53 minutes | Edit Instance ▾ |

Displaying 4 items

Delete tenant310 with network 160.



Network Details: tenant310-160

Network Overview

| | |
|---|---|
| Name | tenant310-160 |
| ID | e862444d-1a9d-4ebb-b19d-881b81a7c993 |
| Project ID | 71332b6ca421425fa5265131e29455bf |
| Status | Active |
| Admin State | UP |
| Shared | No |
| External Network | No |
| MTU | Unknown |
| Provider Network | Network Type: vlan<br>Physical Network: physnet-tenant<br>Segmentation ID: 323 |

Subnets

| ☐ | Name | CIDR | IP Version | Gateway IP | Actions |
|---|---|---|---|---|---|
| ☐ | tenant310-160-subnet | 10.20.160.0/24 | IPv4 | 10.20.160.1 | Edit Subnet ▾ |

Displaying 1 item

```
The segmentation id is 323.

Delete instances tenant310_160_inst3 and tenant310_160_inst4

The router, network entries remain.
The global vlan remains
The nexus vlan remains
OSP8-N9K-FAB-A(config)# show vlan id 323

VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
323  VLAN0323                         active    Po17, Po18, Eth1/17, Eth1/18
```

## Connectivity Tests

Connectivity from the external client machine on floating IP to VMs.

Command used:

```
ssh -i tenant349kp.pem  -o StrictHostKeyChecking=no  cloud-user@10.23.160.77
/tmp/run.sh – for each VM created.

Host is tenant208-108-inst1 and Network is         inet 10.1.108.5  netmask
255.255.255.0  broadcast 10.1.108.255
Host is tenant208-108-inst2 and Network is         inet 10.1.108.6  netmask
255.255.255.0  broadcast 10.1.108.255
…….
…….
…….
Host is tenant348-148-inst1 and Network is         inet 10.2.148.5  netmask
255.255.255.0  broadcast 10.2.148.255
Host is tenant348-148-inst2 and Network is         inet 10.2.148.6  netmask
255.255.255.0  broadcast 10.2.148.255
```

```
Host is tenant348-198-inst3 and Network is         inet 10.2.198.5  netmask
255.255.255.0  broadcast 10.2.198.255
```

A script was created and pushed with 'scp' that in turn runs ifconfig on each VM and gets the details. This was validated for the VMs created above.

## HA on Controller Blades

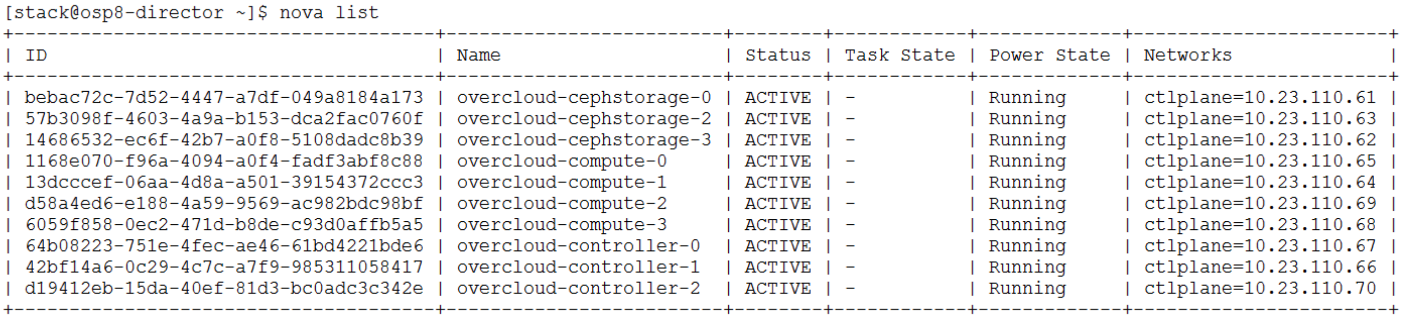
Controllers are key for the health of the cloud which hosts most of the OpenStack services. There are three types of controller failures that could happen.

Server reboot, pulling the blade out of the chassis while system is up and running and putting it back, pulling the blade from the chassis and replacing it simulating a total failure of the controller node.

### Server Reboot Tests

Run Health check before to make sure that system is healthy.

- Run nova list after sourcing stackrc as stack user on Undercloud node to verify that all the controllers are in healthy state as below



```
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks              |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| bebac72c-7d52-4447-a7df-049a8184a173 | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.61 |
| 57b3098f-4603-4a9a-b153-dca2fac0760f | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.63 |
| 14686532-ec6f-42b7-a0f8-5108dadc8b39 | overcloud-cephstorage-3 | ACTIVE | -          | Running     | ctlplane=10.23.110.62 |
| 1168e070-f96a-4094-a0f4-fadf3abf8c88 | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.65 |
| 13dcccef-06aa-4d8a-a501-39154372ccc3 | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.64 |
| d58a4ed6-e188-4a59-9569-ac982bdc98bf | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.69 |
| 6059f858-0ec2-471d-b8de-c93d0affb5a5 | overcloud-compute-3     | ACTIVE | -          | Running     | ctlplane=10.23.110.68 |
| 64b08223-751e-4fec-ae46-61bd4221bde6 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.67 |
| 42bf14a6-0c29-4c7c-a7f9-985311058417 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.66 |
| d19412eb-15da-40ef-81d3-bc0adc3c342e | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.70 |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
```

- Run pcs status on controller nodes and grep for error or stopped.

```
[root@overcloud-controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Tue Sep 20 11:19:36 2016          Last change: Mon Sep 19
20:47:46 2016 by root via cibadmin on overcloud-controller-0
Stack: corosync
Current DC: overcloud-controller-1 (version 1.1.13-10.el7_2.2-44eb2dd) -
partition with quorum
3 nodes and 113 resources configured

Online: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2
]

Full list of resources:

 ip-10.23.110.59        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-
controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-0
```

```
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started overcloud-
controller-1
 Master/Slave Set: redis-master [redis]
     Masters: [ overcloud-controller-1 ]
……
……
PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Online
  overcloud-controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

- Reboot the first controller node and check for pcs status and connectivity of the VMs.
- When the controller comes up, wait till all the services are through PCS are up and running.
- Repeat reboot of the second node and then the third node after the second comes up fully.

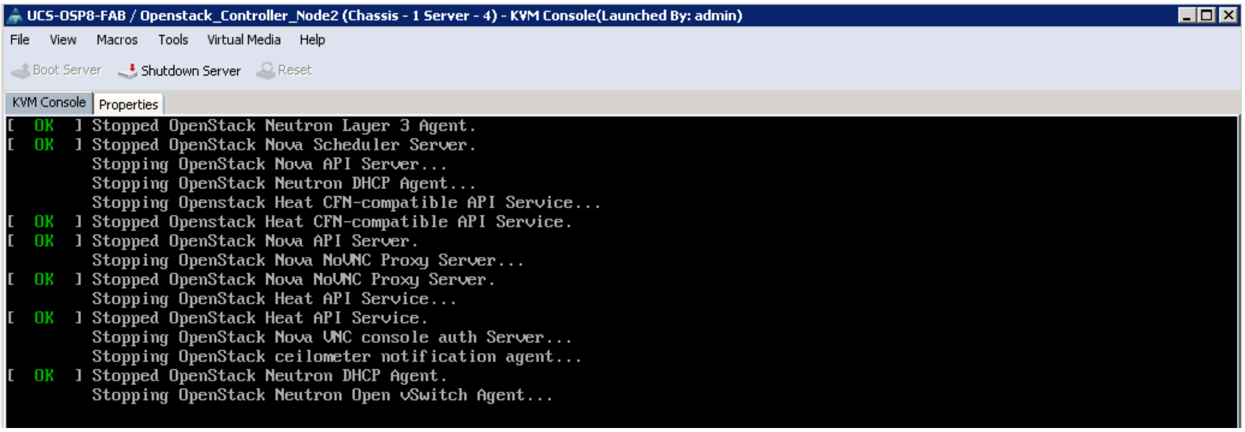
### Health Checks and Observations

The following is a list of health checks and observations:

- Do not reboot the second controller unless the prior one comes up first. Check pacemaker status, health of quorum ( corosync ), health of dashboard.
- Prior to reboot, check the connectivity of VMs

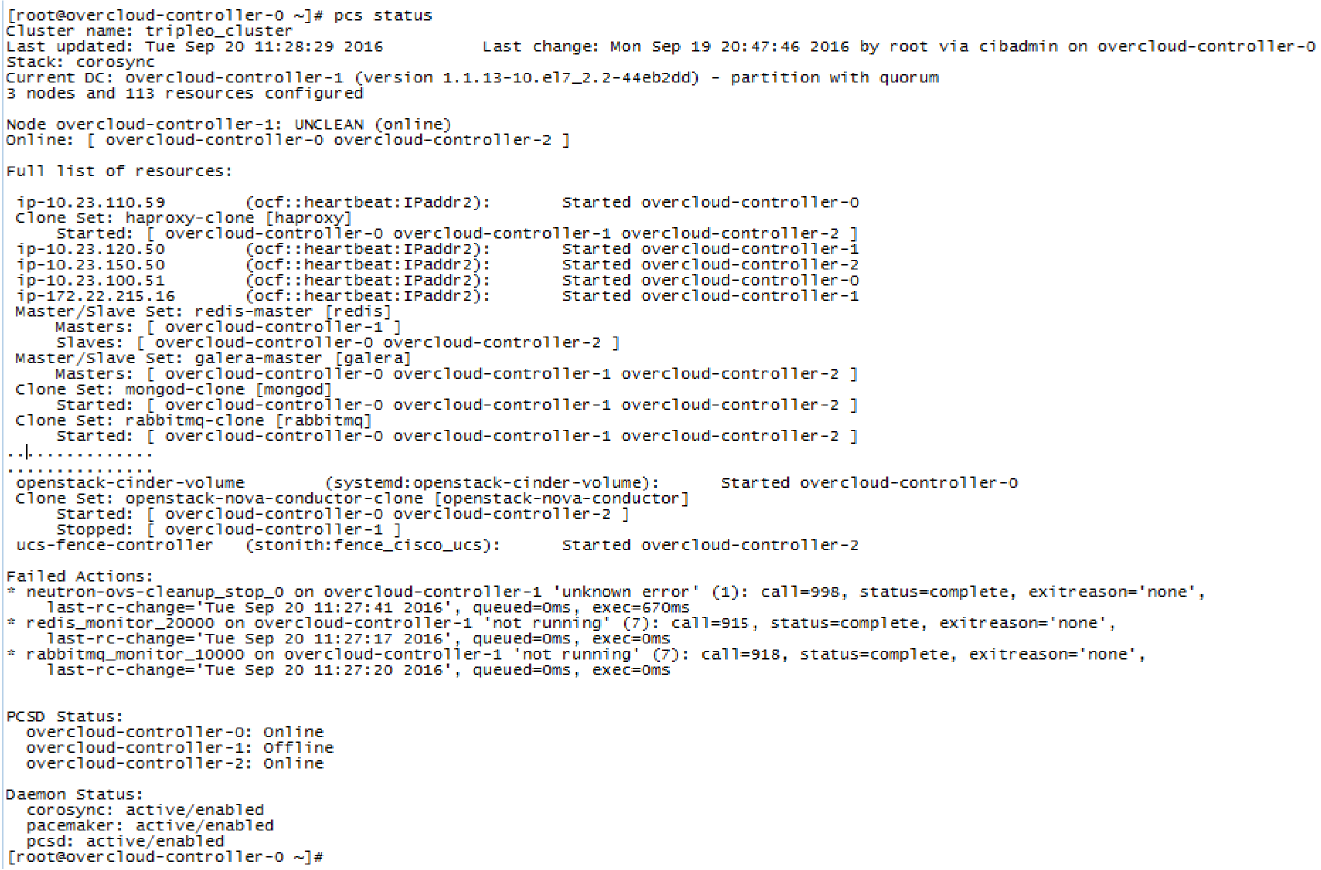
```
[stack@osp8-director scripts]$ date; ./tenantips.sh >> /dev/null; date;
Tue Sep 20 11:26:18 PDT 2016
Tue Sep 20 11:26:22 PDT 2016
[stack@osp8-director scripts]$

Took around 4 seconds to login, execute ifconfig and logout from all the
20VMs.
```

- Two controllers are minimum needed for healthy operation.
- While the first node is booting up, it takes time for pcs status command to complete.


UCS-OSP8-FAB / Openstack_Controller_Node2 (Chassis - 1 Server - 4) - KVM Console(Launched By: admin)
File
View
Macros
Tools
Virtual Media
Help
Boot Server
Shutdown Server
Reset
KVM Console
Properties

```
[  OK  ] Stopped OpenStack Neutron Layer 3 Agent.
[  OK  ] Stopped OpenStack Nova Scheduler Server.
         Stopping OpenStack Nova API Server...
         Stopping OpenStack Neutron DHCP Agent...
         Stopping Openstack Heat CFN-compatible API Service...
[  OK  ] Stopped Openstack Heat CFN-compatible API Service.
[  OK  ] Stopped OpenStack Nova API Server.
         Stopping OpenStack Nova NoVNC Proxy Server...
[  OK  ] Stopped OpenStack Nova NoVNC Proxy Server.
         Stopping OpenStack Heat API Service...
[  OK  ] Stopped OpenStack Heat API Service.
         Stopping OpenStack Nova VNC console auth Server...
         Stopping OpenStack ceilometer notification agent...
[  OK  ] Stopped OpenStack Neutron DHCP Agent.
         Stopping OpenStack Neutron Open vSwitch Agent...
```

```
PCS will report one server is offline.
```



```
[root@overcloud-controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Tue Sep 20 11:28:29 2016          Last change: Mon Sep 19 20:47:46 2016 by root via cibadmin on overcloud-controller-0
Stack: corosync
Current DC: overcloud-controller-1 (version 1.1.13-10.el7_2.2-44eb2dd) - partition with quorum
3 nodes and 113 resources configured

Node overcloud-controller-1: UNCLEAN (online)
Online: [ overcloud-controller-0 overcloud-controller-2 ]

Full list of resources:

 ip-10.23.110.59        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-0
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started overcloud-controller-1
 Master/Slave Set: redis-master [redis]
     Masters: [ overcloud-controller-1 ]
     Slaves: [ overcloud-controller-0 overcloud-controller-2 ]
 Master/Slave Set: galera-master [galera]
     Masters: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
 Clone Set: mongod-clone [mongod]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
 Clone Set: rabbitmq-clone [rabbitmq]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
...............
...............
 openstack-cinder-volume        (systemd:openstack-cinder-volume):      Started overcloud-controller-0
 Clone Set: openstack-nova-conductor-clone [openstack-nova-conductor]
     Started: [ overcloud-controller-0 overcloud-controller-2 ]
     Stopped: [ overcloud-controller-1 ]
 ucs-fence-controller   (stonith:fence_cisco_ucs):      Started overcloud-controller-2

Failed Actions:
* neutron-ovs-cleanup_stop_0 on overcloud-controller-1 'unknown error' (1): call=998, status=complete, exitreason='none',
    last-rc-change='Tue Sep 20 11:27:41 2016', queued=0ms, exec=670ms
* redis_monitor_20000 on overcloud-controller-1 'not running' (7): call=915, status=complete, exitreason='none',
    last-rc-change='Tue Sep 20 11:27:17 2016', queued=0ms, exec=0ms
* rabbitmq_monitor_10000 on overcloud-controller-1 'not running' (7): call=918, status=complete, exitreason='none',
    last-rc-change='Tue Sep 20 11:27:20 2016', queued=0ms, exec=0ms


PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Offline
  overcloud-controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
[root@overcloud-controller-0 ~]#
```

```
PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Offline
  overcloud-controller-2: Online
Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

Corosync will return that it gets only 2 votes out of 3 as below while the server is getting rebooted. This is normal.

- Ceph reports that one monitor is down

```
[root@overcloud-controller-0 cluster]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_WARN
            1 mons down, quorum 1,2 overcloud-controller-0,overcloud-
controller-2
     monmap e2: 3 mons at {overcloud-controller-
0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,overcloud-
controller-2=10.23.120.61:6789/0}
            election epoch 8, quorum 1,2 overcloud-controller-0,overcloud-
controller-2
     osdmap e105: 24 osds: 24 up, 24 in
      pgmap v15522: 448 pgs, 4 pools, 22936 MB data, 5865 objects
            69496 MB used, 128 TB / 128 TB avail
                 448 active+clean
```

- When the node comes up, the routers remain on the other 2 controllers and do not fall back. Can be queried with ip netns too.
- The login to VMs is observed to be slow.
- If controller node does not come up, check through KVM console to spot out any issues and hold off rebooting the second node before a healthy reboot operation of the first.

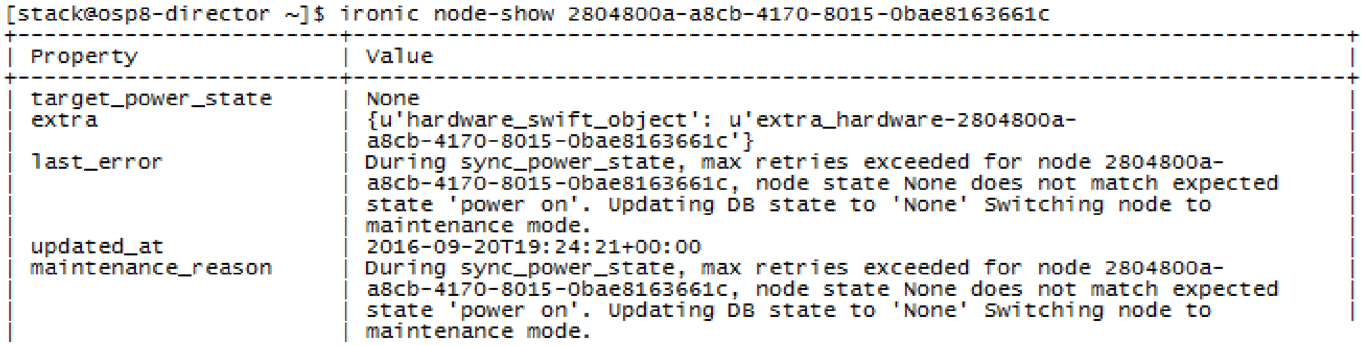
### Blade Pull Tests

One of the controller nodes blade was pulled out while the system is up and running. The validation tests like VM creation etc were done prior to the tests and to check the status when the blade is pulled from the chassis. This is like simulating a complete blade failure. After around 60 minutes the blade was re-inserted back in the chassis.

### Health Checks and Observations

The same behavior as observed during reboot were noticed during the blade pull tests. However unlike a reboot which completes in 5-10 minutes, this was for an extended period of time of 60 minutes to check the status of the cluster.

- Cisco UCS marks the blade as 'removed' and prompts to resolve the slot issue.
- Ironic gives up as it cannot bring the server back online and enables Maintenance mode to True for this node.



```
[stack@osp8-director ~]$ ironic node-show 2804800a-a8cb-4170-8015-0bae8163661c
+------------------------+----------------------------------------------------------------------+
| Property               | Value                                                                |
+------------------------+----------------------------------------------------------------------+
| target_power_state     | None                                                                 |
| extra                  | {u'hardware_swift_object': u'extra_hardware-2804800a-                |
|                        | a8cb-4170-8015-0bae8163661c'}                                        |
| last_error             | During sync_power_state, max retries exceeded for node 2804800a-     |
|                        | a8cb-4170-8015-0bae8163661c, node state None does not match expected |
|                        | state 'power on'. Updating DB state to 'None' Switching node to      |
|                        | maintenance mode.                                                    |
| updated_at             | 2016-09-20T19:24:21+00:00                                            |
| maintenance_reason     | During sync_power_state, max retries exceeded for node 2804800a-     |
|                        | a8cb-4170-8015-0bae8163661c, node state None does not match expected |
|                        | state 'power on'. Updating DB state to 'None' Switching node to      |
|                        | maintenance mode.
```

- Ceph storage will report that 1 out of 3 monitors are down similar to above. All the 3 controllers will be running one monitor each. However all the OSD's are up and running.
- After inserting the blade back into the same slot of the chassis, it needed a manual intervention to correct the above.

— Insert the blade back into the slot and resolve the slot issue in UCS.

— ironic node-set-power-state 2804800a-a8cb-4170-8015-0bae8163661c on

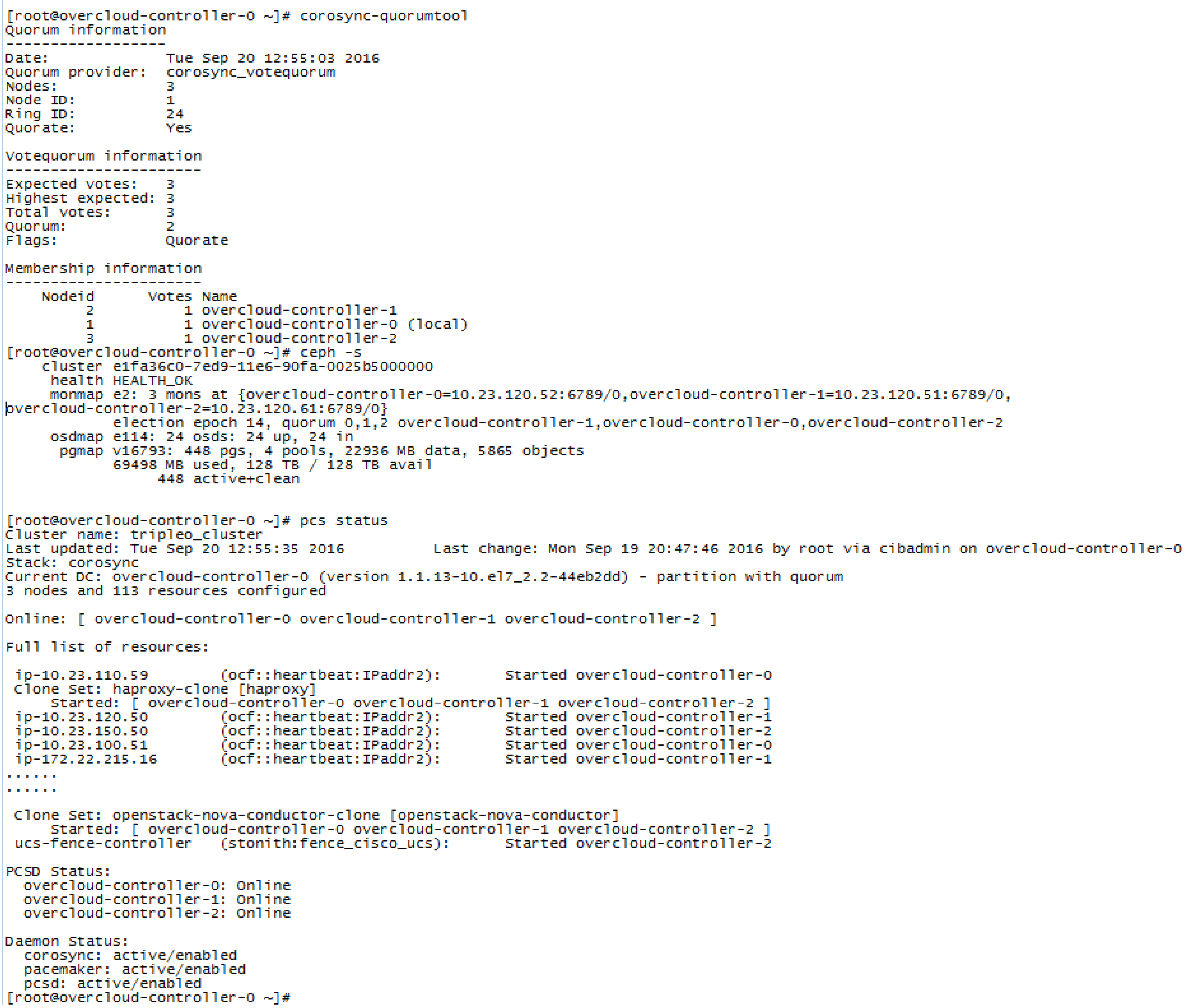
— ironic node-set-maintenance 2804800a-a8cb-4170-8015-0bae8163661c false

— Wait for a minute and check back for these columns with ironic node-list.

— You may observe that nova service-list is down on the new node. You may have to wait few minutes before it comes up.

```
[root@overcloud-controller-0 ~]# nova service-list | grep overcloud-controller-1
| 8  | nova-scheduler   | overcloud-controller-1.localdomain | internal | enabled | down  | 2016-09-20T19:20:16.000000 | -           |
| 11 | nova-conductor   | overcloud-controller-1.localdomain | internal | enabled | down  | 2016-09-20T19:20:18.000000 | -           |
| 71 | nova-consoleauth | overcloud-controller-1.localdomain | internal | enabled | down  | 2016-09-20T19:20:17.000000 | -           |
```

- Login to the controller node, check for pcs status and resolve any processes that were not brought up running 'pcs resource cleanup'
- Ocassional issue as reported in bug 1368594 is observed.



```
[root@overcloud-controller-0 ~]# corosync-quorumtool
Quorum information
------------------
Date:             Tue Sep 20 12:55:03 2016
Quorum provider:  corosync_votequorum
Nodes:            3
Node ID:          1
Ring ID:          24
Quorate:          Yes

Votequorum information
----------------------
Expected votes:   3
Highest expected: 3
Total votes:      3
Quorum:           2
Flags:            Quorate

Membership information
----------------------
    Nodeid      Votes Name
         2          1 overcloud-controller-1
         1          1 overcloud-controller-0 (local)
         3          1 overcloud-controller-2
[root@overcloud-controller-0 ~]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,
overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 14, quorum 0,1,2 overcloud-controller-1,overcloud-controller-0,overcloud-controller-2
     osdmap e114: 24 osds: 24 up, 24 in
      pgmap v16793: 448 pgs, 4 pools, 22936 MB data, 5865 objects
            69498 MB used, 128 TB / 128 TB avail
                 448 active+clean


[root@overcloud-controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Tue Sep 20 12:55:35 2016          Last change: Mon Sep 19 20:47:46 2016 by root via cibadmin on overcloud-controller-0
Stack: corosync
Current DC: overcloud-controller-0 (version 1.1.13-10.el7_2.2-44eb2dd) - partition with quorum
3 nodes and 113 resources configured

Online: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]

Full list of resources:

 ip-10.23.110.59        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started overcloud-controller-0
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started overcloud-controller-1
......
......

 Clone Set: openstack-nova-conductor-clone [openstack-nova-conductor]
     Started: [ overcloud-controller-0 overcloud-controller-1 overcloud-controller-2 ]
 ucs-fence-controller   (stonith:fence_cisco_ucs):      Started overcloud-controller-2

PCSD Status:
  overcloud-controller-0: Online
  overcloud-controller-1: Online
  overcloud-controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
[root@overcloud-controller-0 ~]#
```

### Blade Replacement

Unlike the above two types of failures, in this test the blade is completely removed and new one is added. There were few issues encountered while rebuilding the failed controller blade and adding it as a replacement. The fix for bug 1298430 will give business continuity but there is a need to fix the failed blade. While this issue is being investigated, an interim solution was developed to circumvent the above limitation. This is included in the Hardware failures section. Different types of hardware failures that can happen on a controller blade and how to mitigate these issues considering the dependency of Controller blade on IPMI and MAC addresses is addressed there.

## HA on Compute Blades

### Reboot Test

Tests and Observations are as follows:

- Many Instances were provisioned across the pod and reboot of the Compute Node was attempted.

```
[root@overcloud-compute-3 ~]# virsh list
 Id    Name                           State
----------------------------------------------------
 2     instance-0000000f              running
 3     instance-0000001b              running
 4     instance-00000021              running
 5     instance-00000030              running
 6     instance-0000003c              running
 7     instance-00000048              running
```

- Identified a compute host to be rebooted and the VMs that could be impacted

  About 6 VMs were up and running

- Identify the floating IP’s for these VMs from nova list —-all-tenants and capture data to login without password, run ifconfig script. The script sshs to all the VMs run’s ifconfig and returns serially.

  Running a script N-S  for all the VMs

```
[stack@osp8-director scripts]$ date;./tenantips.sh ; date
Mon Sep 19 14:26:55 PDT 2016
        inet 10.20.155.4  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.155.3  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.4  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.3  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.154.4  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.104.4  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.104.3  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.153.4  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.153.3  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.103.4  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.103.3  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.152.4  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.152.5  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.102.4  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.102.3  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.151.4  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.151.3  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.101.4  netmask 255.255.255.0  broadcast 10.20.101.255
        inet 10.20.101.3  netmask 255.255.255.0  broadcast 10.20.101.255
Mon Sep 19 14:26:59 PDT 2016

Took around 3 seconds to login to all the VMs.
The tenantips.sh does something like below:
[stack@osp8-director scripts]$ tail -3 tenantips.sh
ssh -i tenant301kp.pem -o StrictHostKeyChecking=no cloud-user@10.23.160.33
/usr/sbin/ifconfig | grep "10.20"
ssh -i tenant301kp.pem -o StrictHostKeyChecking=no cloud-user@10.23.160.32
/usr/sbin/ifconfig | grep "10.20"
```

```
ssh -i tenant301kp.pem -o StrictHostKeyChecking=no cloud-user@10.23.160.31
/usr/sbin/ifconfig | grep "10.20"
```

set resume_guests_state_on_host_boot=true in nova.conf to get the instances back online after reboot.

- Rebooted the Compute Node overcloud-compute-3
- Instances came up fine and the same script to validate the login with floating ips worked fine.
- By default guests will not come up unless resume_guests_state_on_host_boot is set to true. If this parameter isn't set before reboot:
- Now Reboot the node and check connectivity of VMs during and after reboot.

```
[root@overcloud-compute-3 ~]# reboot
PolicyKit daemon disconnected from the bus.
We are no longer a registered authentication agent.
Connection to 10.23.110.57 closed by remote host.
Connection to 10.23.110.57 closed.
[stack@osp8-director ~]$

Using the same script while server is rebooting yields connectivity fail-
ures to the 6 VMs on this host.
[stack@osp8-director scripts]$ date;./tenantips.sh ; date
Mon Sep 19 14:32:37 PDT 2016
ssh: connect to host 10.23.160.54 port 22: No route to host
        inet 10.20.155.3  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.4  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.3  netmask 255.255.255.0  broadcast 10.20.105.255
ssh: connect to host 10.23.160.49 port 22: No route to host
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.104.4  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.104.3  netmask 255.255.255.0  broadcast 10.20.104.255
ssh: connect to host 10.23.160.44 port 22: No route to host
        inet 10.20.153.3  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.103.4  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.103.3  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.152.4  netmask 255.255.255.0  broadcast 10.20.152.255
ssh: connect to host 10.23.160.38 port 22: No route to host
        inet 10.20.102.4  netmask 255.255.255.0  broadcast 10.20.102.255
ssh: connect to host 10.23.160.36 port 22: No route to host
        inet 10.20.151.4  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.151.3  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.101.4  netmask 255.255.255.0  broadcast 10.20.101.255
ssh: connect to host 10.23.160.31 port 22: No route to host
Mon Sep 19 14:34:27 PDT 2016

Once the node comes up fine, the VMs are accessible.
[stack@osp8-director ~]$ ssh -l heat-admin 10.23.110.57
Last login: Mon Sep 19 14:36:08 2016 from 10.23.110.26
[heat-admin@overcloud-compute-3 ~]$ sudo -i
[root@overcloud-compute-3 ~]# date
Mon Sep 19 14:36:49 PDT 2016
[root@overcloud-compute-3 ~]# virsh list
 Id    Name                           State
----------------------------------------------------
 2     instance-0000000f              running
 3     instance-0000001b              running
 4     instance-00000021              running
 5     instance-00000030              running
 6     instance-0000003c              running
 7     instance-00000048              running

All the VMs are accessible from floating IP’s too now.
[stack@osp8-director scripts]$ date;./tenantips.sh ; date
Mon Sep 19 14:37:05 PDT 2016
        inet 10.20.155.4  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.155.3  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.4  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.3  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.154.4  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.104.4  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.104.3  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.153.4  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.153.3  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.103.4  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.103.3  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.152.4  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.152.5  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.102.4  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.102.3  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.151.4  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.151.3  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.101.4  netmask 255.255.255.0  broadcast 10.20.101.255
        inet 10.20.101.3  netmask 255.255.255.0  broadcast 10.20.101.255
Mon Sep 19 14:37:09 PDT 2016
```

### Blade Pull Tests

One of the Compute blades was pulled out while the system is up and running. This was also an extended test for about 60 minutes and then the blade was re-inserted back in the chassis.

Observations:

- Results were similar to reboot tests above.
- UCS Manager complained to resolve the host as it was pulled out from the chassis. This was acknowledged and the blade was re-inserted.
- The guest VMs came up when resume_guests property was set to true at host level.
- Similar to Controller blade pull tests, nova made the state as ‘NOSTATE’ and ironic put the blade into Maintenance mode.
- Similar steps like setting up the maintenance mode to off through ironic and nova reset state were issued to the blade after getting ‘ok’ status in UCS manager.

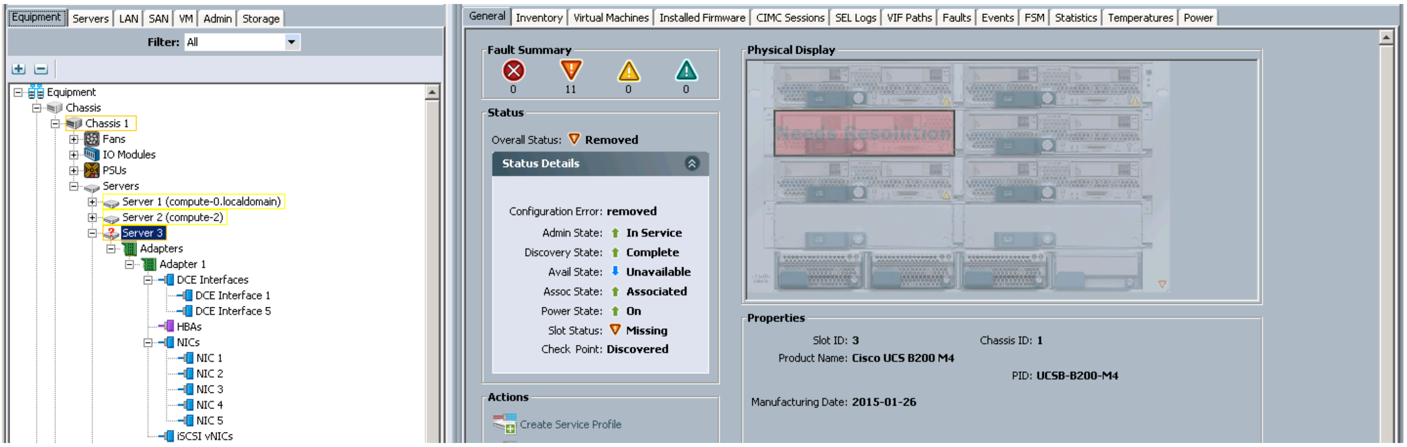
### Blade Replacement

Compute blade was pulled from the chassis completely and the server was decommissioned in UCS. This is to simulate a complete failure of a Compute blade. Then an attempt was made to remove this from OpenStack and add a new blade to the cloud. The service profile was reused in this method. The following were the tasks list and observations made during a Compute blade replacement test.

Blade replacement is actually a two phase process. First remove the faulty blade from the system, restore VMs and then add a new one.

To delete a Node, complete the following steps:

Pull blade from chassis.



Status

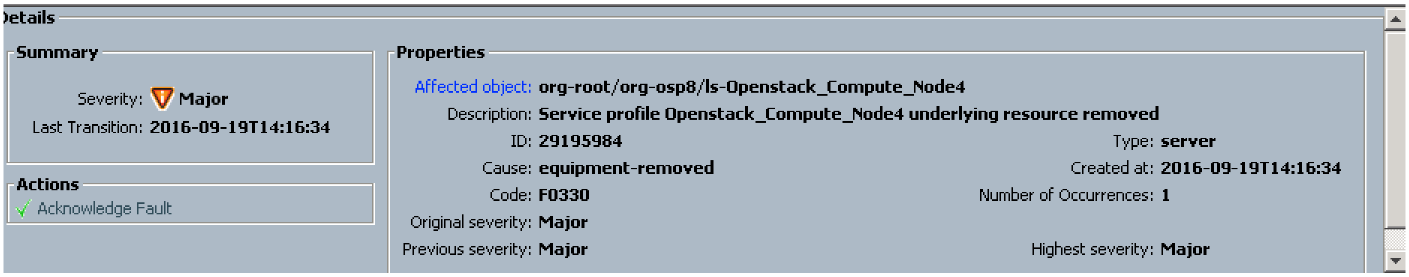
Overall Status: Removed

Status Details

Configuration Error: removed
Admin State: In Service
Discovery State: Complete
Avail State: Unavailable
Assoc State: Associated
Power State: On
Slot Status: Missing
Check Point: Discovered

Properties

Slot ID: 3 Chassis ID: 1
Product Name: Cisco UCS B200 M4
PID: UCSB-B200-M4
Manufacturing Date: 2015-01-26



etails

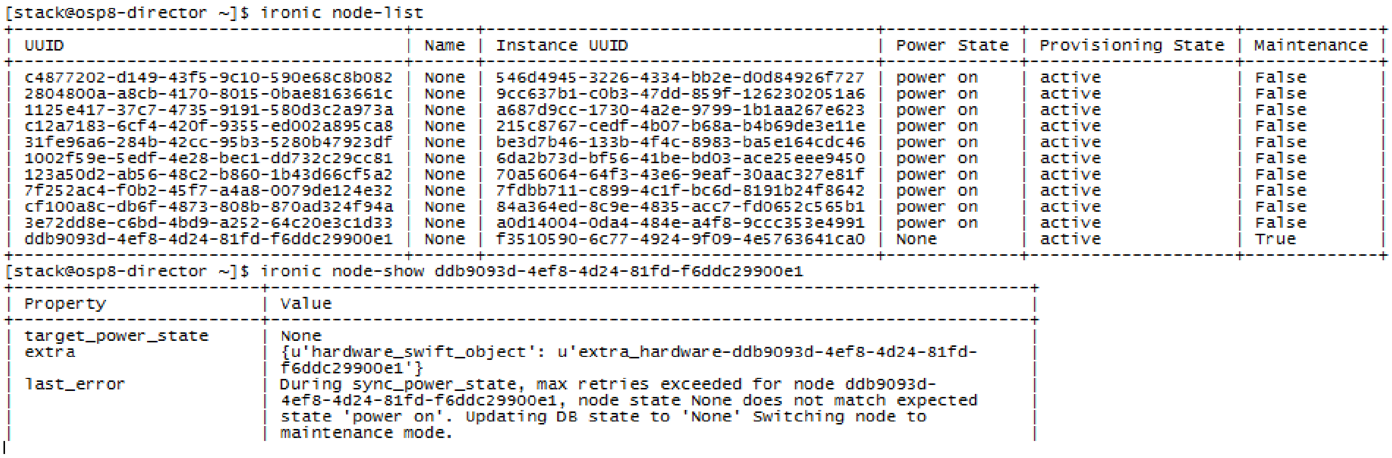
Summary

Severity: Major
Last Transition: 2016-09-19T14:16:34

Actions

Properties

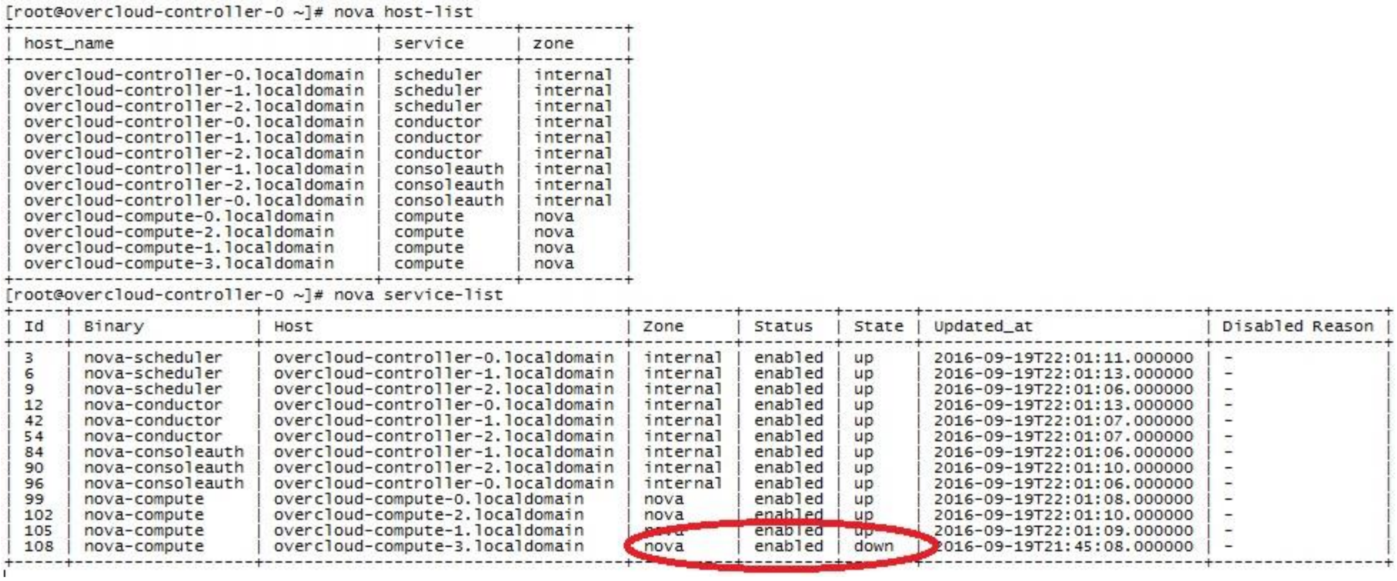
Affected object: org-root/org-osp8/ls-Openstack_Compute_Node4
Description: Service profile Openstack_Compute_Node4 underlying resource removed
ID: 29195984 Type: server
Cause: equipment-removed Created at: 2016-09-19T14:16:34
Code: F0330 Number of Occurrences: 1
Original severity: Major
Previous severity: Major Highest severity: Major



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-126230205 1a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
| ddb9093d-4ef8-4d24-81fd-f6ddc29900e1 | None | f3510590-6c77-4924-9f09-4e5763641ca0 | None        | active             | True        |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
[stack@osp8-director ~]$ ironic node-show ddb9093d-4ef8-4d24-81fd-f6ddc29900e1
+------------------------+------------------------------------------------------------------+
| Property               | Value                                                            |
+------------------------+------------------------------------------------------------------+
| target_power_state     | None                                                             |
| extra                  | {u'hardware_swift_object': u'extra_hardware-ddb9093d-4ef8-4d24-81fd- |
|                        | f6ddc29900e1'}                                                   |
| last_error             | During sync_power_state, max retries exceeded for node ddb9093d- |
|                        | 4ef8-4d24-81fd-f6ddc29900e1, node state None does not match expected |
|                        | state 'power on'. Updating DB state to 'None' Switching node to  |
|                        | maintenance mode.                                                |
```

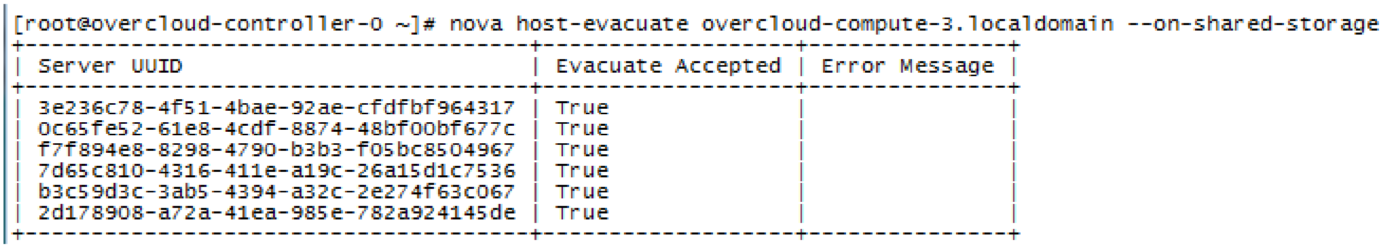
Evacuate the VMs from the failed node.

Issue nova host-list and service-list to find the status of this compute node and then evacuate the VMs from this node.



```
[root@overcloud-controller-0 ~]# nova host-list
+-----------------------------------+-------------+----------+
| host_name                         | service     | zone     |
+-----------------------------------+-------------+----------+
| overcloud-controller-0.localdomain | scheduler   | internal |
| overcloud-controller-1.localdomain | scheduler   | internal |
| overcloud-controller-2.localdomain | scheduler   | internal |
| overcloud-controller-0.localdomain | conductor   | internal |
| overcloud-controller-1.localdomain | conductor   | internal |
| overcloud-controller-2.localdomain | conductor   | internal |
| overcloud-controller-1.localdomain | consoleauth | internal |
| overcloud-controller-2.localdomain | consoleauth | internal |
| overcloud-controller-0.localdomain | consoleauth | internal |
| overcloud-compute-0.localdomain   | compute     | nova     |
| overcloud-compute-2.localdomain   | compute     | nova     |
| overcloud-compute-1.localdomain   | compute     | nova     |
| overcloud-compute-3.localdomain   | compute     | nova     |
+-----------------------------------+-------------+----------+
[root@overcloud-controller-0 ~]# nova service-list
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| Id  | Binary           | Host                               | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| 3   | nova-scheduler   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:01:11.000000 | -               |
| 6   | nova-scheduler   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:01:13.000000 | -               |
| 9   | nova-scheduler   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:01:06.000000 | -               |
| 12  | nova-conductor   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:01:13.000000 | -               |
| 42  | nova-conductor   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:01:07.000000 | -               |
| 54  | nova-conductor   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:01:07.000000 | -               |
| 84  | nova-consoleauth | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:01:06.000000 | -               |
| 90  | nova-consoleauth | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:01:10.000000 | -               |
| 96  | nova-consoleauth | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:01:06.000000 | -               |
| 99  | nova-compute     | overcloud-compute-0.localdomain    | nova     | enabled | up    | 2016-09-19T22:01:08.000000 | -               |
| 102 | nova-compute     | overcloud-compute-2.localdomain    | nova     | enabled | up    | 2016-09-19T22:01:10.000000 | -               |
| 105 | nova-compute     | overcloud-compute-1.localdomain    | nova     | enabled | up    | 2016-09-19T22:01:09.000000 | -               |
| 108 | nova-compute     | overcloud-compute-3.localdomain    | nova     | enabled | down  | 2016-09-19T21:45:08.000000 | -               |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
```

[root@overcloud-controller-0 ~]# nova host-evacuate overcloud-compute-3.localdomain --on-shared-storage



```
[root@overcloud-controller-0 ~]# nova host-evacuate overcloud-compute-3.localdomain --on-shared-storage
+--------------------------------------+-------------------+---------------+
| Server UUID                          | Evacuate Accepted | Error Message |
+--------------------------------------+-------------------+---------------+
| 3e236c78-4f51-4bae-92ae-cfdfbf964317 | True              |               |
| 0c65fe52-61e8-4cdf-8874-48bf00bf677c | True              |               |
| f7f894e8-8298-4790-b3b3-f05bc8504967 | True              |               |
| 7d65c810-4316-411e-a19c-26a15d1c7536 | True              |               |
| b3c59d3c-3ab5-4394-a32c-2e274f63c067 | True              |               |
| 2d178908-a72a-41ea-985e-782a924145de | True              |               |
+--------------------------------------+-------------------+---------------+
```

VMs are accessible through floating IPs.

```
[stack@osp8-director scripts]$ date; ./tenantips.sh ; date
Mon Sep 19 15:08:09 PDT 2016
        inet 10.20.155.4  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.155.3  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.4  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.3  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.154.4  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.104.4  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.104.3  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.153.4  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.153.3  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.103.4  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.103.3  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.152.4  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.152.5  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.102.4  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.102.3  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.151.4  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.151.3  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.101.4  netmask 255.255.255.0  broadcast 10.20.101.255
        inet 10.20.101.3  netmask 255.255.255.0  broadcast 10.20.101.255
      Mon Sep 19 15:08:12 PDT 2016

      No VMs on failed host now.
```

```
[root@overcloud-controller-0 ~]# nova-manage vm list | egrep -v "overcloud-
compute-3" | grep active | wc -l
No handlers could be found for logger "oslo_config.cfg"
20

[root@overcloud-controller-0 ~]# nova-manage vm list | egrep -i "overcloud-
compute-3"
No handlers could be found for logger "oslo_config.cfg"
[root@overcloud-controller-0 ~]#
```

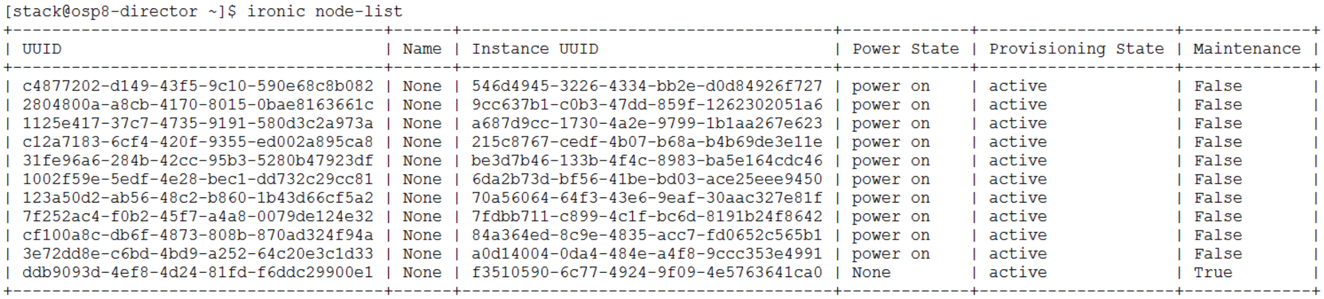
As the blade was completely pulled from the chassis ironic was unable to get to the power management and the above node-delete failed.

```
ipmitool -I lanplus -H <ipmi address> -U admin -P <password> chassis power
status
Error: Unable to establish IPMI v2 / RMCP+ session
Error: Unable to establish IPMI v2 / RMCP+ session
Error: Unable to establish IPMI v2 / RMCP+ session
Unable to get Chassis Power Status
```

Workarounds to delete the blade in the current status are as follows:

Update the error status to available status in ironic node-list

```
edit  /etc/ironic/ironic.conf
Update the enabled drivers temporarily as below
#enabled_drivers=pxe_ipmitool,pxe_ssh,pxe_drac
enabled_drivers=fake
Restart openstack-ironic-conductor
sudo service openstack-ironic-conductor restart

ironic node-update NODE_UUID replace driver=fake
[stack@osp8-director ~]$ ironic node-show ddb9093d-4ef8-4d24-81fd-
f6ddc29900e1| grep fake
| driver                 | fake
|
[stack@osp8-director ~]$
```

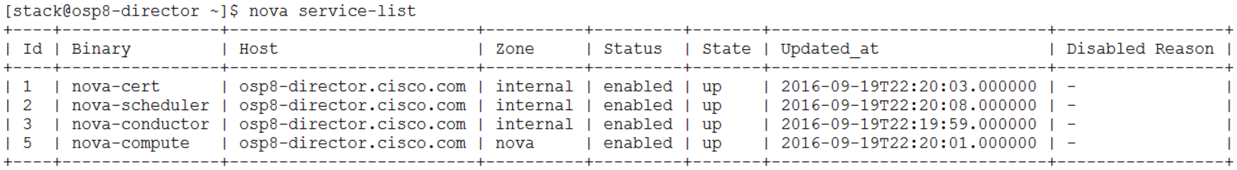


```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-1262302051a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
| ddb9093d-4ef8-4d24-81fd-f6ddc29900e1 | None | f3510590-6c77-4924-9f09-4e5763641ca0 | None        | active             | True        |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
```

Run nova service-list and identify the service ids

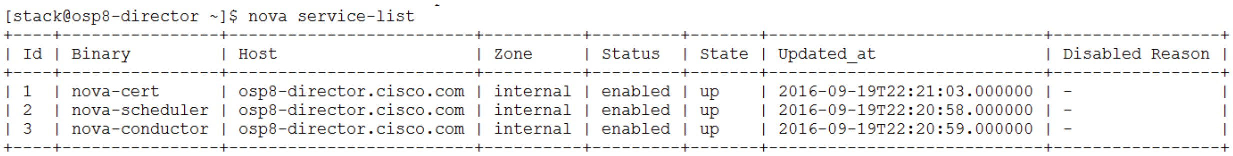
Delete the service ids associated with this node as:

```
nova service-list
```



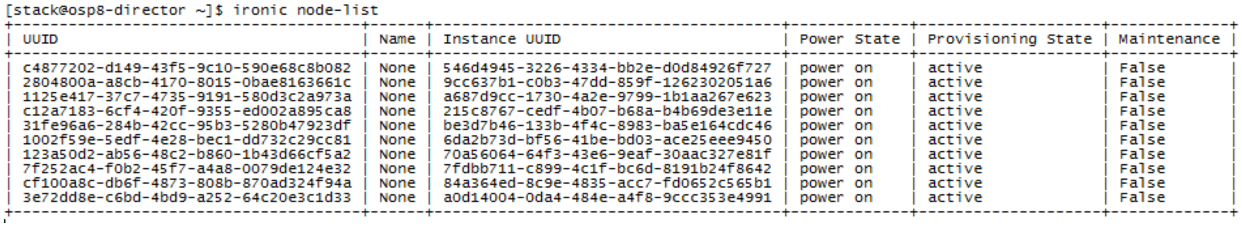
```
[stack@osp8-director ~]$ nova service-list
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
| Id | Binary           | Host                      | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
| 1  | nova-cert        | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:20:03.000000 | -               |
| 2  | nova-scheduler   | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:20:08.000000 | -               |
| 3  | nova-conductor   | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:19:59.000000 | -               |
| 5  | nova-compute     | osp8-director.cisco.com | nova     | enabled | up    | 2016-09-19T22:20:01.000000 | -               |
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
```

```
nova service-delete $id
```



```
[stack@osp8-director ~]$ nova service-list
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
| Id | Binary           | Host                      | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
| 1  | nova-cert        | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:21:03.000000 | -               |
| 2  | nova-scheduler   | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:20:58.000000 | -               |
| 3  | nova-conductor   | osp8-director.cisco.com | internal | enabled | up    | 2016-09-19T22:20:59.000000 | -               |
+----+------------------+---------------------------+----------+---------+-------+----------------------------+-----------------+
```

```
[stack@osp8-director ~]$ ironic node-set-provision-state ddb9093d-4ef8-
4d24-81fd-f6ddc29900e1 deleted

Delete the node from ironic

[stack@osp8-director ~]$ ironic node-delete ddb9093d-4ef8-4d24-81fd-
f6ddc29900e1
Deleted node ddb9093d-4ef8-4d24-81fd-f6ddc29900e1

[stack@osp8-director ~]$ nova delete f3510590-6c77-4924-9f09-4e5763641ca0
Request to delete server f3510590-6c77-4924-9f09-4e5763641ca0 has been ac-
cepted.
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-126230205 1a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
```

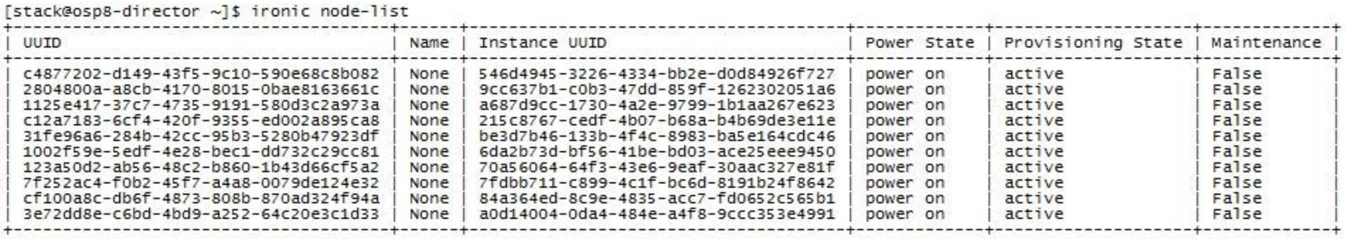
Revert back the "fake" driver from ironic.conf

```
edit vi /etc/ironic/ironic.conf.
enabled_drivers=pxe_ipmitool,pxe_ssh,pxe_drac
#enabled_drivers=fake
```

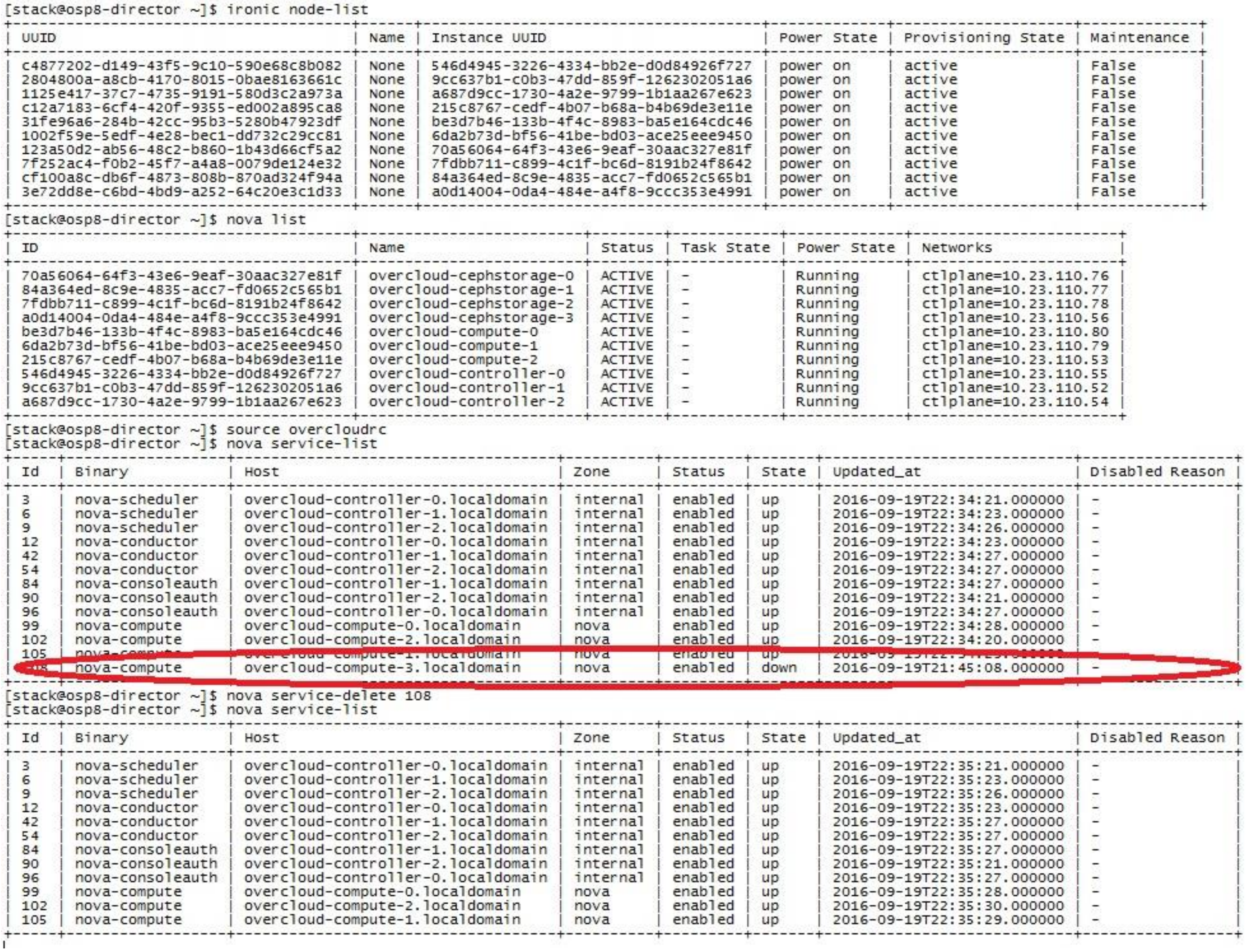
Restart ironic-conductor to pick up the drivers again.

```
service openstack-ironic-conductor restart
```

The deleted node should not exist anymore in ironic node-list or nova-list now.

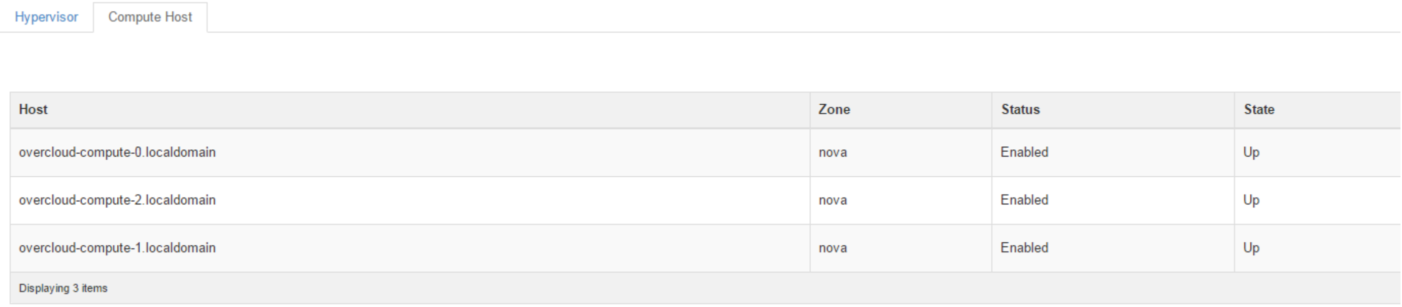


```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-1262302051a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | 546d4945-3226-4334-bb2e-d0d84926f727 | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 9cc637b1-c0b3-47dd-859f-1262302051a6 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | a687d9cc-1730-4a2e-9799-1b1aa267e623 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | 215c8767-cedf-4b07-b68a-b4b69de3e11e | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | be3d7b46-133b-4f4c-8983-ba5e164cdc46 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 6da2b73d-bf56-41be-bd03-ace25eee9450 | power on    | active             | False       |
| 123a50d2-ab56-48c2-b860-1b43d66cf5a2 | None | 70a56064-64f3-43e6-9eaf-30aac327e81f | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 84a364ed-8c9e-4835-acc7-fd0652c565b1 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | a0d14004-0da4-484e-a4f8-9ccc353e4991 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks             |
+--------------------------------------+-------------------------+--------+------------+-------------+----------------------+
| 70a56064-64f3-43e6-9eaf-30aac327e81f | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.76 |
| 84a364ed-8c9e-4835-acc7-fd0652c565b1 | overcloud-cephstorage-1 | ACTIVE | -          | Running     | ctlplane=10.23.110.77 |
| 7fdbb711-c899-4c1f-bc6d-8191b24f8642 | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.78 |
| a0d14004-0da4-484e-a4f8-9ccc353e4991 | overcloud-cephstorage-3 | ACTIVE | -          | Running     | ctlplane=10.23.110.56 |
| be3d7b46-133b-4f4c-8983-ba5e164cdc46 | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.80 |
| 6da2b73d-bf56-41be-bd03-ace25eee9450 | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.79 |
| 215c8767-cedf-4b07-b68a-b4b69de3e11e | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.53 |
| 546d4945-3226-4334-bb2e-d0d84926f727 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.55 |
| 9cc637b1-c0b3-47dd-859f-1262302051a6 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.52 |
| a687d9cc-1730-4a2e-9799-1b1aa267e623 | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.54 |
+--------------------------------------+-------------------------+--------+------------+-------------+----------------------+
[stack@osp8-director ~]$ source overcloudrc
[stack@osp8-director ~]$ nova service-list
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| Id  | Binary           | Host                               | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| 3   | nova-scheduler   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:34:21.000000 | -               |
| 6   | nova-scheduler   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:34:23.000000 | -               |
| 9   | nova-scheduler   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:34:26.000000 | -               |
| 12  | nova-conductor   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:34:23.000000 | -               |
| 42  | nova-conductor   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:34:27.000000 | -               |
| 54  | nova-conductor   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:34:27.000000 | -               |
| 84  | nova-consoleauth | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:34:27.000000 | -               |
| 90  | nova-consoleauth | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:34:21.000000 | -               |
| 96  | nova-consoleauth | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:34:27.000000 | -               |
| 99  | nova-compute     | overcloud-compute-0.localdomain    | nova     | enabled | up    | 2016-09-19T22:34:28.000000 | -               |
| 102 | nova-compute     | overcloud-compute-2.localdomain    | nova     | enabled | up    | 2016-09-19T22:34:20.000000 | -               |
| 105 | nova-compute     | overcloud-compute-1.localdomain    | nova     | enabled | up    | 2016-09-19T22:34:29.000000 | -               |
| 108 | nova-compute     | overcloud-compute-3.localdomain    | nova     | enabled | down  | 2016-09-19T21:45:08.000000 | -               |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
[stack@osp8-director ~]$ nova service-delete 108
[stack@osp8-director ~]$ nova service-list
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| Id  | Binary           | Host                               | Zone     | Status  | State | Updated_at                 | Disabled Reason |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
| 3   | nova-scheduler   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:35:21.000000 | -               |
| 6   | nova-scheduler   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:35:23.000000 | -               |
| 9   | nova-scheduler   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:35:26.000000 | -               |
| 12  | nova-conductor   | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:35:23.000000 | -               |
| 42  | nova-conductor   | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:35:27.000000 | -               |
| 54  | nova-conductor   | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:35:27.000000 | -               |
| 84  | nova-consoleauth | overcloud-controller-1.localdomain | internal | enabled | up    | 2016-09-19T22:35:27.000000 | -               |
| 90  | nova-consoleauth | overcloud-controller-2.localdomain | internal | enabled | up    | 2016-09-19T22:35:21.000000 | -               |
| 96  | nova-consoleauth | overcloud-controller-0.localdomain | internal | enabled | up    | 2016-09-19T22:35:27.000000 | -               |
| 99  | nova-compute     | overcloud-compute-0.localdomain    | nova     | enabled | up    | 2016-09-19T22:35:28.000000 | -               |
| 102 | nova-compute     | overcloud-compute-2.localdomain    | nova     | enabled | up    | 2016-09-19T22:35:30.000000 | -               |
| 105 | nova-compute     | overcloud-compute-1.localdomain    | nova     | enabled | up    | 2016-09-19T22:35:29.000000 | -               |
+-----+------------------+------------------------------------+----------+---------+-------+----------------------------+-----------------+
```

Login to the dashboard to check that overcloud-compute-3 does not exist anymore.



| Host | Zone | Status | State |
|---|---|---|---|
| overcloud-compute-0.localdomain | nova | Enabled | Up |
| overcloud-compute-2.localdomain | nova | Enabled | Up |
| overcloud-compute-1.localdomain | nova | Enabled | Up |

Displaying 3 items

### Node Addition

When the compute blade has been completely removed from OpenStack, a new blade can be added. The procedure for adding a new compute blade is same as how it was addressed earlier in upscaling the compute pod.

## HA on Storage Nodes

Ceph, the software stack deployed by Red Hat OpenStack Platform director, has its high availability built in. By default, the system will be replicating the placement groups and has 3 copies distributed across the hosts.

The parameter osd_pool_default_size = 3 in ceph.conf brings this feature by default when installed.

If we create a crushmap from the existing cluster as below it reveals what type of buckets are in and what mode of replication is being done by default in the cluster.

ceph osd getcrushmap -o /tmp/crushmap.bin

crushtool -d crushmap.bin -o /tmp/crushmap.txt

```
rule replicated_ruleset {
        ruleset 0
        type replicated ← Default to Replication mode
        min_size 1
        max_size 10
        step take default
        step chooseleaf firstn 0 type host ←Default distribution of PG copies
        step emit
}
```

Whenever a Ceph node goes down, the system will start rebuilding from the copies of replicas. While this is an expected behavior of Ceph, it causes some CPU and memory overhead too. This is one of the reasons to have a minimum of 3 nodes for ceph and leave some good amount of free space with in the storage cluster. This will help Ceph to move the blocks around in case of failures like this. More the nodes better it is, as this rebuild activity is distributed across the cluster. Though there are other parameters like osd_max_backfills to control this activity and its impact on CPU, it may not be feasible to cover all of these recovery parameters in this document.

What needs to be noted is that the recovery kicks in as part of the tests below. The ceph cluster status may show warnings while the tests are being conducted as it is moving the placement groups and may cause performance issues on the storage cluster. Hence checking the health of nodes while adding/rebuilding a new node is important.

### Reboot Test

Check the status of the cluster:

```
[root@overcloud-controller-0 ~]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-
controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-
controller-0,overcloud-controller-2
     osdmap e73: 32 osds: 32 up, 32 in
      pgmap v589: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69543 MB used, 171 TB / 171 TB avail
                 448 active+clean
  client io 5864 B/s wr, 1 op/s
```

Reboot one of the Ceph storage node:

```
[stack@osp8-director scripts]$ ping 10.23.110.62
```

```
PING 10.23.110.62 (10.23.110.62) 56(84) bytes of data.
From 10.23.110.26 icmp_seq=1 Destination Host Unreachable
From 10.23.110.26 icmp_seq=2 Destination Host Unreachable
From 10.23.110.26 icmp_seq=3 Destination Host Unreachable
From 10.23.110.26 icmp_seq=4 Destination Host Unreachable
```

ceph –w reports as OSD’s recovering

```
[root@overcloud-controller-0 ~]# ceph -w
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_WARN
            134 pgs stale
            8/32 in osds are down
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-
controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-
controller-0,overcloud-controller-2
     osdmap e75: 32 osds: 24 up, 32 in
      pgmap v633: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69516 MB used, 171 TB / 171 TB avail
                 314 active+clean
                 134 stale+active+clean


2016-09-19 21:18:41.668127 mon.0 [INF] pgmap v633: 448 pgs: 134
stale+active+clean, 314 active+clean; 22926 MB data, 69516 MB used, 171 TB / 171
TB avail

2016-09-19 21:18:45.783782 mon.0 [INF] pgmap v634: 448 pgs: 125 ac-
tive+undersized+degraded, 88 stale+active+clean, 235 active+clean; 22926 MB data,
69512 MB used, 171 TB / 171 TB avail; 25071 kB/s rd, 3 op/s; 1668/17595 objects
degraded (9.480%)

2016-09-19 21:18:46.973329 mon.0 [INF] pgmap v635: 448 pgs: 350 ac-
tive+undersized+degraded, 98 active+clean; 22926 MB data, 69515 MB used, 171 TB /
171 TB avail; 117 MB/s rd, 7 op/s; 4997/17595 objects degraded (28.400%)

2016-09-19 21:18:50.736079 mon.0 [INF] pgmap v636: 448 pgs: 350 ac-
tive+undersized+degraded, 98 active+clean; 22926 MB data, 69515 MB used, 171 TB /
171 TB avail; 99151 kB/s rd, 4 op/s; 4997/17595 objects degraded (28.400%)

2016-09-19 21:18:51.816130 mon.0 [INF] pgmap v637: 448 pgs: 350 ac-
tive+undersized+degraded, 98 active+clean; 22926 MB data, 69515 MB used, 171 TB /
171 TB avail; 414 B/s wr, 0 op/s; 4997/17595 objects degraded (28.400%)

2016-09-19 21:18:55.730472 mon.0 [INF] pgmap v638: 448 pgs: 350 ac-
tive+undersized+degraded, 98 active+clean; 22926 MB data, 69515 MB used, 171 TB /
171 TB avail; 409 B/s wr, 0 op/s; 4997/17595 objects degraded (28.400%)
```

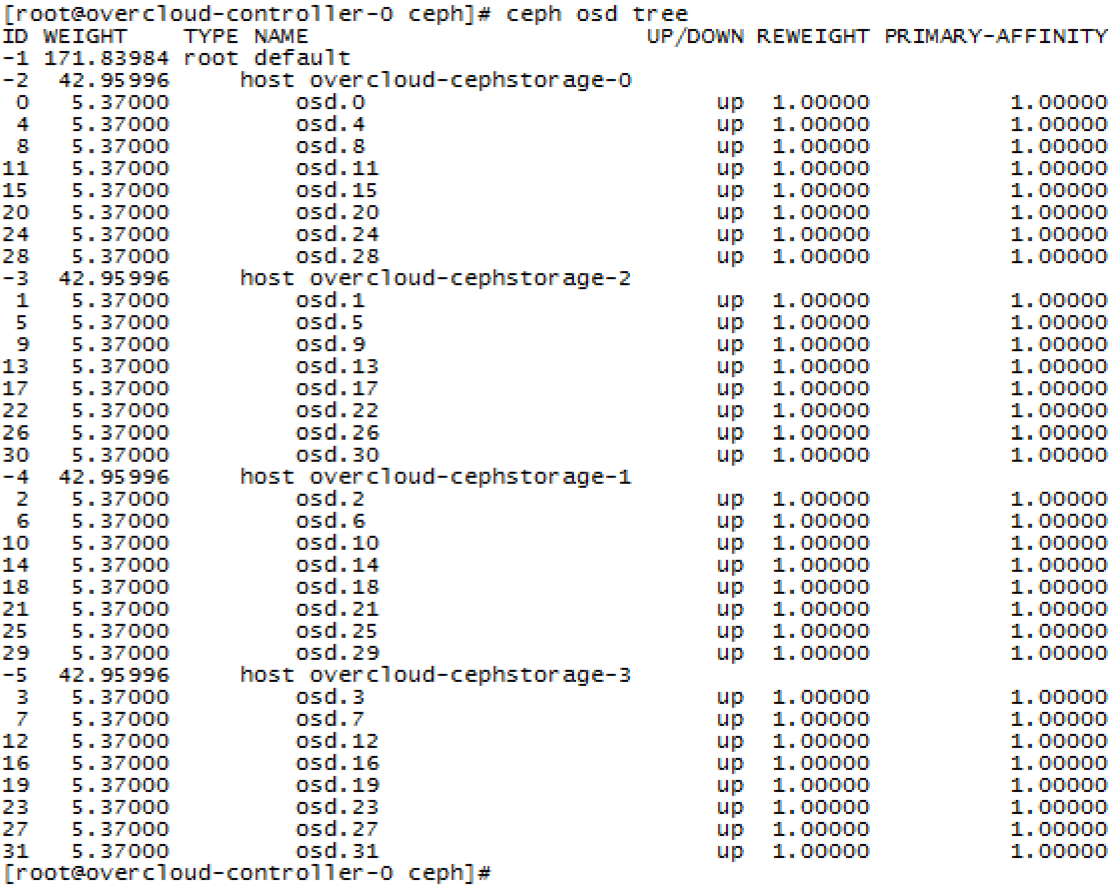
ceph –s reports Warning as below during reboot exercise

```
[root@overcloud-controller-0 ~]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_WARN
            350 pgs degraded
            323 pgs stuck unclean
            350 pgs undersized
            recovery 4997/17595 objects degraded (28.400%)
            8/32 in osds are down
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-
controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-
controller-0,overcloud-controller-2
     osdmap e75: 32 osds: 24 up, 32 in
      pgmap v672: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69515 MB used, 171 TB / 171 TB avail
            4997/17595 objects degraded (28.400%)
                 350 active+undersized+degraded
                  98 active+clean
  client io 16914 B/s wr, 5 op/s
```

After few minutes the cluster recovers as shown below:

```
[root@overcloud-cephstorage-0 ceph]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-
controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-
controller-0,overcloud-controller-2
     osdmap e81: 32 osds: 32 up, 32 in
      pgmap v765: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69499 MB used, 171 TB / 171 TB avail
                 448 active+clean
  client io 15035 B/s wr, 3 op/s
```

ceph osd tree reports after recovery



```
[root@overcloud-controller-0 ceph]# ceph osd tree
ID WEIGHT    TYPE NAME                         UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 171.83984 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                          up  1.00000          1.00000
 4   5.37000         osd.4                          up  1.00000          1.00000
 8   5.37000         osd.8                          up  1.00000          1.00000
11   5.37000         osd.11                         up  1.00000          1.00000
15   5.37000         osd.15                         up  1.00000          1.00000
20   5.37000         osd.20                         up  1.00000          1.00000
24   5.37000         osd.24                         up  1.00000          1.00000
28   5.37000         osd.28                         up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                          up  1.00000          1.00000
 5   5.37000         osd.5                          up  1.00000          1.00000
 9   5.37000         osd.9                          up  1.00000          1.00000
13   5.37000         osd.13                         up  1.00000          1.00000
17   5.37000         osd.17                         up  1.00000          1.00000
22   5.37000         osd.22                         up  1.00000          1.00000
26   5.37000         osd.26                         up  1.00000          1.00000
30   5.37000         osd.30                         up  1.00000          1.00000
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2                          up  1.00000          1.00000
 6   5.37000         osd.6                          up  1.00000          1.00000
10   5.37000         osd.10                         up  1.00000          1.00000
14   5.37000         osd.14                         up  1.00000          1.00000
18   5.37000         osd.18                         up  1.00000          1.00000
21   5.37000         osd.21                         up  1.00000          1.00000
25   5.37000         osd.25                         up  1.00000          1.00000
29   5.37000         osd.29                         up  1.00000          1.00000
-5  42.95996     host overcloud-cephstorage-3
 3   5.37000         osd.3                          up  1.00000          1.00000
 7   5.37000         osd.7                          up  1.00000          1.00000
12   5.37000         osd.12                         up  1.00000          1.00000
16   5.37000         osd.16                         up  1.00000          1.00000
19   5.37000         osd.19                         up  1.00000          1.00000
23   5.37000         osd.23                         up  1.00000          1.00000
27   5.37000         osd.27                         up  1.00000          1.00000
31   5.37000         osd.31                         up  1.00000          1.00000
[root@overcloud-controller-0 ceph]#
```

VMs connectivity through floating IP continues without interruption.

```
[stack@osp8-director scripts]$ ./tenantips.sh
        inet 10.20.155.6  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.155.5  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.6  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.5  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.154.6  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.104.6  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.104.5  netmask 255.255.255.0  broadcast 10.20.104.255
        inet 10.20.153.6  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.153.5  netmask 255.255.255.0  broadcast 10.20.153.255
        inet 10.20.103.6  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.103.5  netmask 255.255.255.0  broadcast 10.20.103.255
        inet 10.20.152.6  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.152.5  netmask 255.255.255.0  broadcast 10.20.152.255
        inet 10.20.102.6  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.102.5  netmask 255.255.255.0  broadcast 10.20.102.255
        inet 10.20.151.6  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.151.5  netmask 255.255.255.0  broadcast 10.20.151.255
        inet 10.20.101.6  netmask 255.255.255.0  broadcast 10.20.101.255
        inet 10.20.101.5  netmask 255.255.255.0  broadcast 10.20.101.255
[stack@osp8-director scripts]$
```

The node comes after few minutes, while the cluster shows warning issues during the reboot period.

The status of the cluster observed fine after few minutes of reboot. The warning message continues until the recovery activity is complete.

### System Power Off

The behavior in system power off is very similar to what observed on Controller and Compute blade pull tests.

System took around 6 minutes to come back to OK status. The time system takes to recover depends on the active number of placement group and copies the system was attempting to move around.

There is a more detailed description and symptoms observed during power off that are listed in Node Replacement section below.

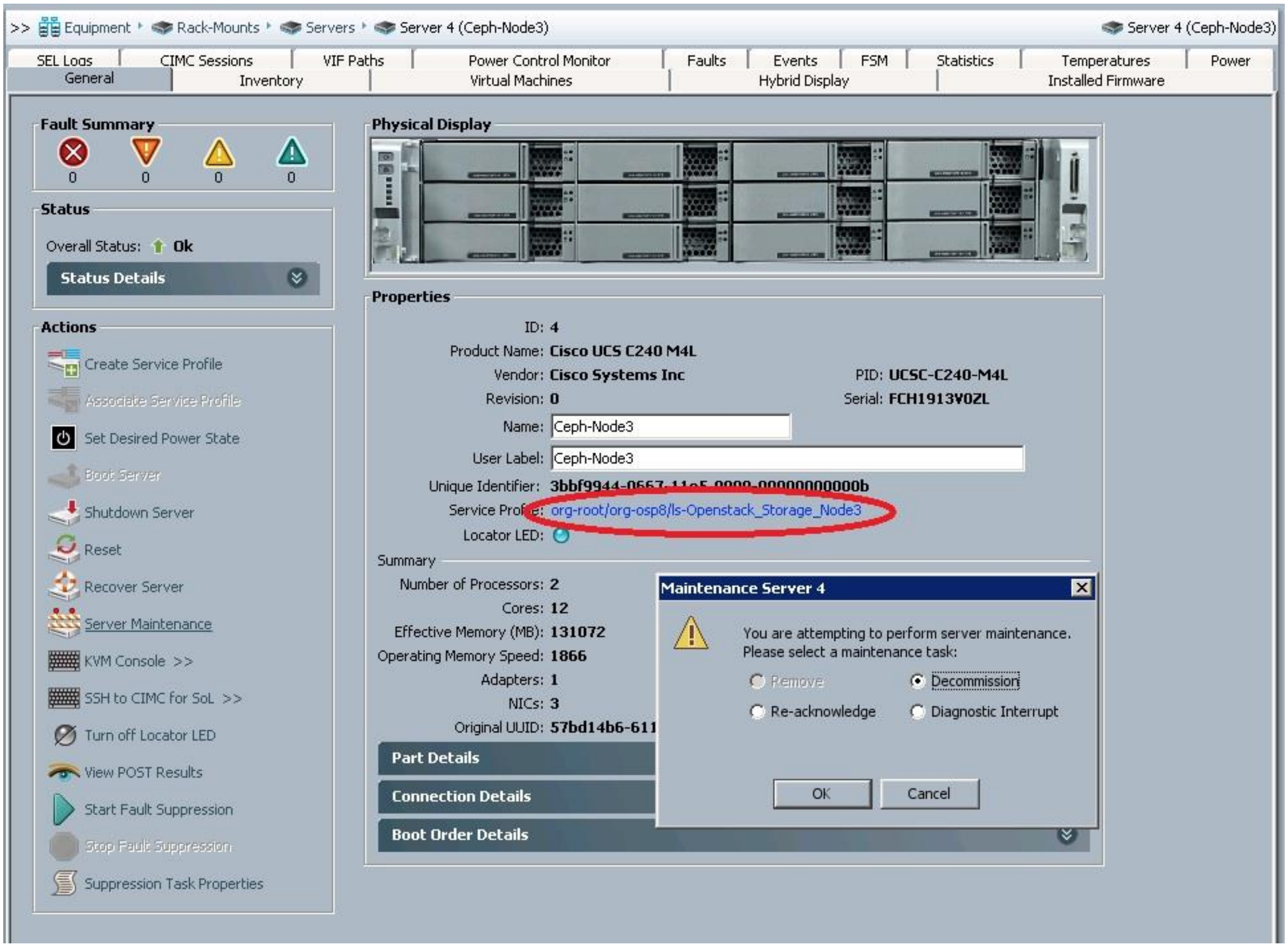
### Node Replacement

One of the storage servers was powered off ( pull the power cord ) completely and the server was decommissioned in UCS. This is to simulate a complete failure of the storage server. Then an attempt was made to remove this node from OpenStack and add a new one to the cloud. The following were the tasks list and observations made during a Storage node replacement test.

Node replacement is actually a two phase process. First remove the server from the system and then add a new one.

To delete a node, complete the following steps:

Decommission one of the storage nodes as below:



Overall Status: Ok

**Properties**

ID: 4
Product Name: Cisco UCS C240 M4L
Vendor: Cisco Systems Inc
PID: UCSC-C240-M4L
Revision: 0
Serial: FCH1913V0ZL
Name: Ceph-Node3
User Label: Ceph-Node3
Unique Identifier: 3bbf9944-0667-11e5-0000-00000000000b
Service Profile: org-root/org-osp8/ls-Openstack_Storage_Node3
Locator LED:

Summary

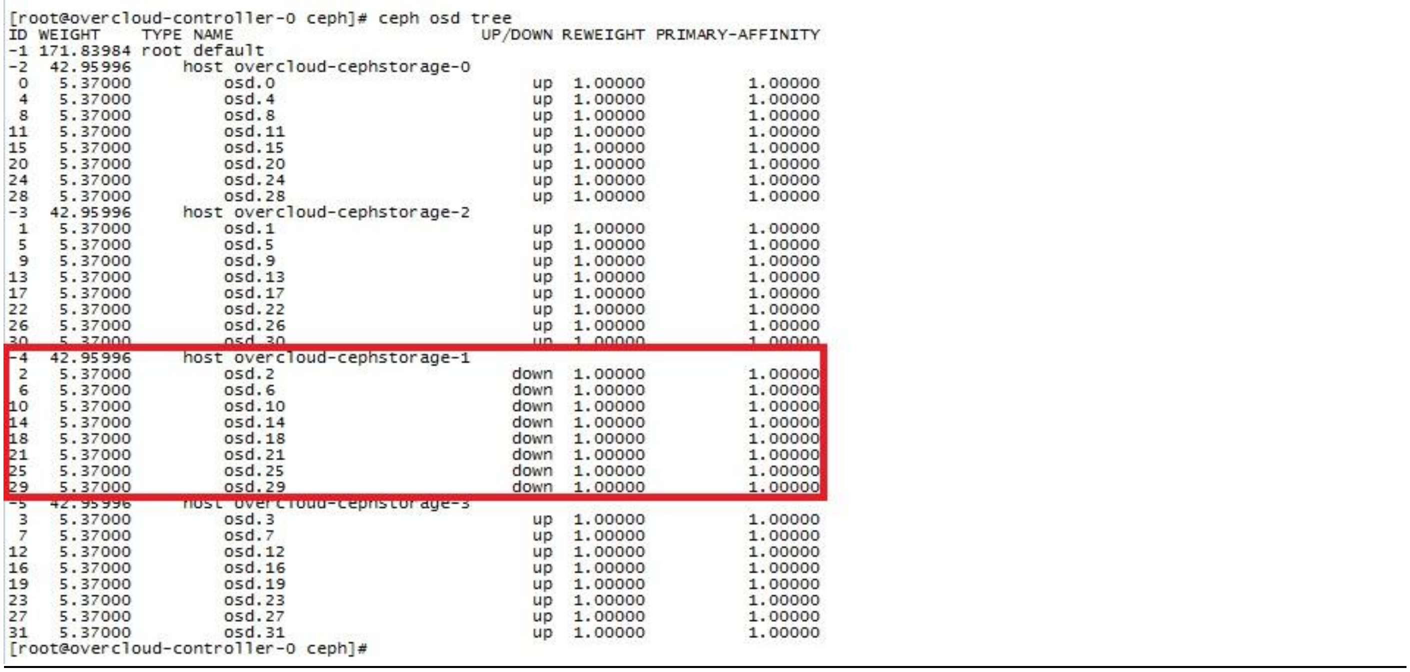
Number of Processors: 2
Cores: 12
Effective Memory (MB): 131072
Operating Memory Speed: 1866
Adapters: 1
NICs: 3
Original UUID: 57bd14b6-611

**Maintenance Server 4**

You are attempting to perform server maintenance.
Please select a maintenance task:

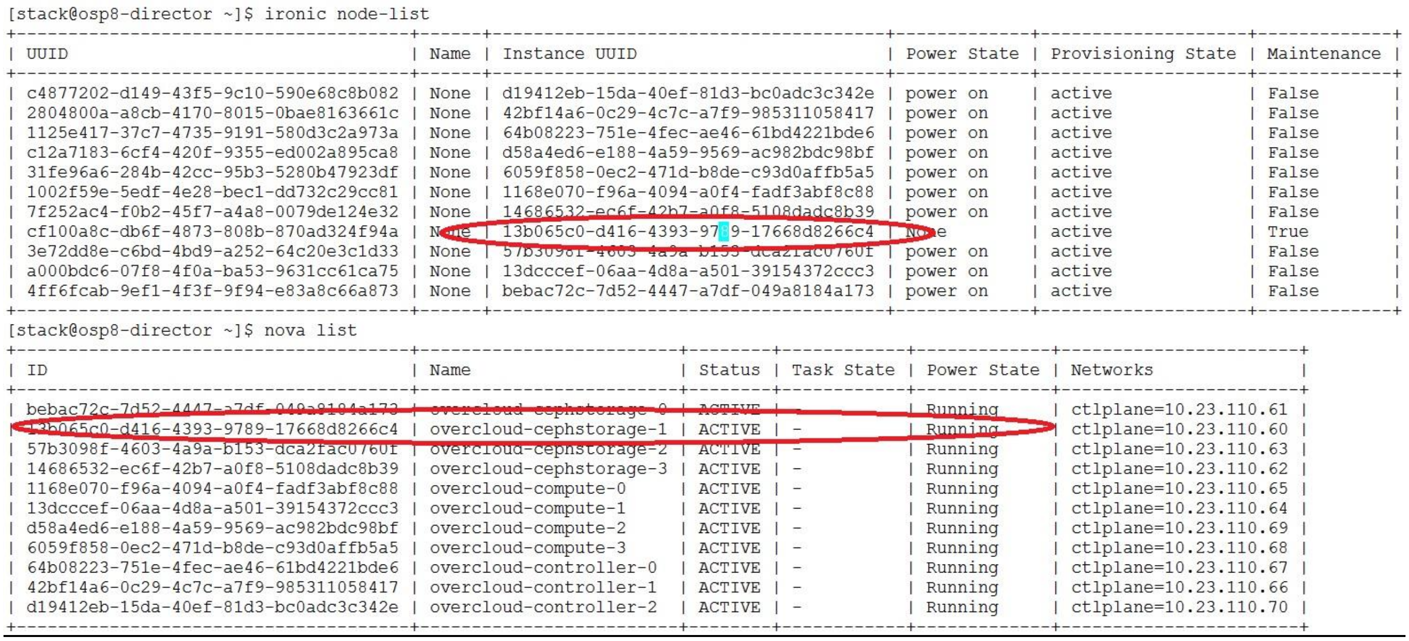
○ Remove ● Decommission
○ Re-acknowledge ○ Diagnostic Interrupt

OK Cancel

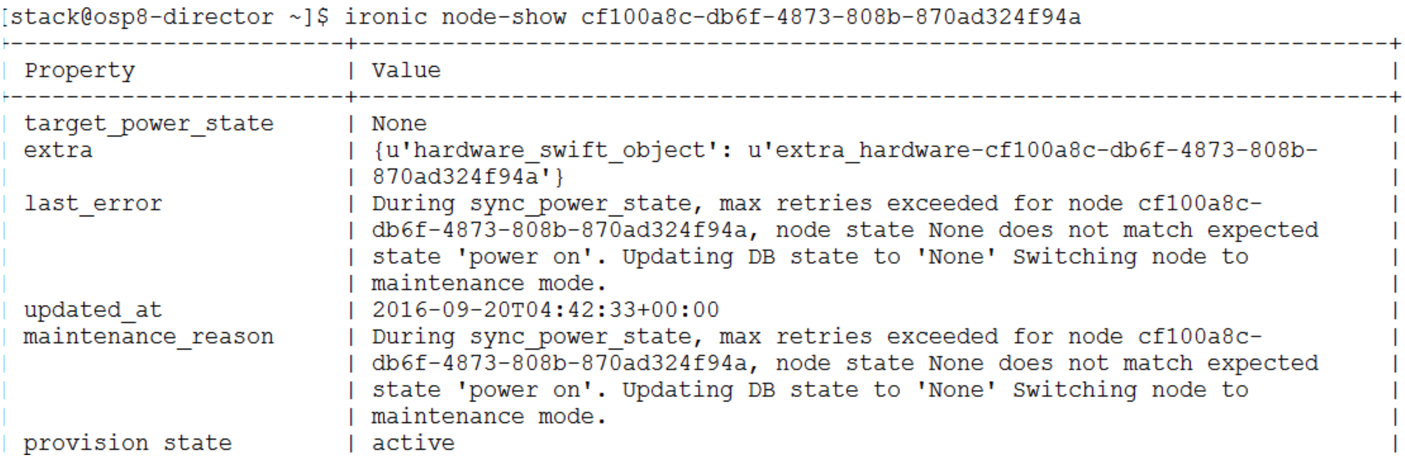


```
[root@overcloud-controller-0 ceph]# ceph osd tree
ID WEIGHT    TYPE NAME                          UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 171.83984 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                           up  1.00000          1.00000
 4   5.37000         osd.4                           up  1.00000          1.00000
 8   5.37000         osd.8                           up  1.00000          1.00000
11   5.37000         osd.11                          up  1.00000          1.00000
15   5.37000         osd.15                          up  1.00000          1.00000
20   5.37000         osd.20                          up  1.00000          1.00000
24   5.37000         osd.24                          up  1.00000          1.00000
28   5.37000         osd.28                          up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                           up  1.00000          1.00000
 5   5.37000         osd.5                           up  1.00000          1.00000
 9   5.37000         osd.9                           up  1.00000          1.00000
13   5.37000         osd.13                          up  1.00000          1.00000
17   5.37000         osd.17                          up  1.00000          1.00000
22   5.37000         osd.22                          up  1.00000          1.00000
26   5.37000         osd.26                          up  1.00000          1.00000
30   5.37000         osd.30                          up  1.00000          1.00000
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2                         down  1.00000          1.00000
 6   5.37000         osd.6                         down  1.00000          1.00000
10   5.37000         osd.10                        down  1.00000          1.00000
14   5.37000         osd.14                        down  1.00000          1.00000
18   5.37000         osd.18                        down  1.00000          1.00000
21   5.37000         osd.21                        down  1.00000          1.00000
25   5.37000         osd.25                        down  1.00000          1.00000
29   5.37000         osd.29                        down  1.00000          1.00000
-5  42.95996     host overcloud-cephstorage-3
 3   5.37000         osd.3                           up  1.00000          1.00000
 7   5.37000         osd.7                           up  1.00000          1.00000
12   5.37000         osd.12                          up  1.00000          1.00000
16   5.37000         osd.16                          up  1.00000          1.00000
19   5.37000         osd.19                          up  1.00000          1.00000
23   5.37000         osd.23                          up  1.00000          1.00000
27   5.37000         osd.27                          up  1.00000          1.00000
31   5.37000         osd.31                          up  1.00000          1.00000
[root@overcloud-controller-0 ceph]#
```

```
[root@overcloud-controller-0 ceph]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_WARN
            338 pgs degraded
            313 pgs stuck unclean
            338 pgs undersized
            recovery 3741/17595 objects degraded (21.262%)
            8/32 in osds are down
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-controller-0,overcloud-controller-2
     osdmap e83: 32 osds: 24 up, 32 in
      pgmap v1119: 448 pgs, 4 pools, 22926 MB data, 5865 objects
            69505 MB used, 171 TB / 171 TB avail
            3741/17595 objects degraded (21.262%)
                 338 active+undersized+degraded
                 110 active+clean
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | d19412eb-15da-40ef-81d3-bc0adc3c342e | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 42bf14a6-0c29-4c7c-a7f9-985311058417 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | 64b08223-751e-4fec-ae46-61bd4221bde6 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | d58a4ed6-e188-4a59-9569-ac982bdc98bf | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | 6059f858-0ec2-471d-b8de-c93d0affb5a5 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 1168e070-f96a-4094-a0f4-fadf3abf8c88 | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 14686532-ec6f-42b7-a0f8-5108dadc8b39 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | 13b065c0-d416-4393-9789-17668d8266c4 | None        | active             | True        |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | 57b3098f-4603-4a9a-b153-dca2fac0760f | power on    | active             | False       |
| a000bdc6-07f8-4f0a-ba53-9631cc61ca75 | None | 13dcccef-06aa-4d8a-a501-39154372ccc3 | power on    | active             | False       |
| 4ff6fcab-9ef1-4f3f-9f94-e83a8c66a873 | None | bebac72c-7d52-4447-a7df-049a8184a173 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks              |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| bebac72c-7d52-4447-a7df-049a8184a173 | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.61 |
| 13b065c0-d416-4393-9789-17668d8266c4 | overcloud-cephstorage-1 | ACTIVE | -          | Running     | ctlplane=10.23.110.60 |
| 57b3098f-4603-4a9a-b153-dca2fac0760f | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.63 |
| 14686532-ec6f-42b7-a0f8-5108dadc8b39 | overcloud-cephstorage-3 | ACTIVE | -          | Running     | ctlplane=10.23.110.62 |
| 1168e070-f96a-4094-a0f4-fadf3abf8c88 | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.65 |
| 13dcccef-06aa-4d8a-a501-39154372ccc3 | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.64 |
| d58a4ed6-e188-4a59-9569-ac982bdc98bf | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.69 |
| 6059f858-0ec2-471d-b8de-c93d0affb5a5 | overcloud-compute-3     | ACTIVE | -          | Running     | ctlplane=10.23.110.68 |
| 64b08223-751e-4fec-ae46-61bd4221bde6 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.67 |
| 42bf14a6-0c29-4c7c-a7f9-985311058417 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.66 |
| d19412eb-15da-40ef-81d3-bc0adc3c342e | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.70 |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
```



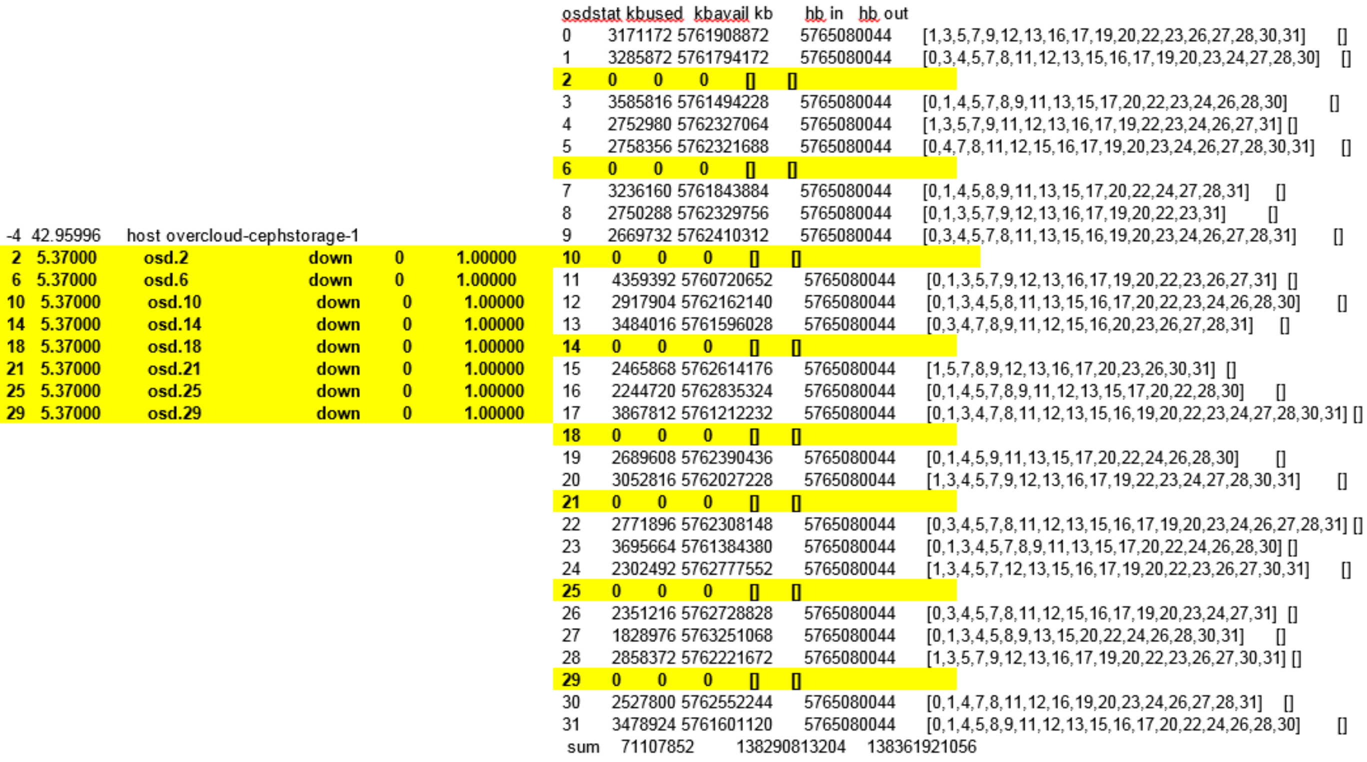
```
[stack@osp8-director ~]$ ironic node-show cf100a8c-db6f-4873-808b-870ad324f94a
+------------------------+-----------------------------------------------------------------+
| Property               | Value                                                           |
+------------------------+-----------------------------------------------------------------+
| target_power_state     | None                                                            |
| extra                  | {u'hardware_swift_object': u'extra_hardware-cf100a8c-db6f-4873-808b- |
|                        | 870ad324f94a'}                                                  |
| last_error             | During sync_power_state, max retries exceeded for node cf100a8c- |
|                        | db6f-4873-808b-870ad324f94a, node state None does not match expected |
|                        | state 'power on'. Updating DB state to 'None' Switching node to |
|                        | maintenance mode.                                               |
| updated_at             | 2016-09-20T04:42:33+00:00                                       |
| maintenance_reason     | During sync_power_state, max retries exceeded for node cf100a8c- |
|                        | db6f-4873-808b-870ad324f94a, node state None does not match expected |
|                        | state 'power on'. Updating DB state to 'None' Switching node to |
|                        | maintenance mode.                                               |
| provision_state        | active                                                          |
```

Check the health of placement groups before removing the server completely from the cluster.

```
[root@overcloud-controller-0 ceph]# ceph pg dump_stuck stale
ok
[root@overcloud-controller-0 ceph]# ceph pg dump_stuck inactive
ok
[root@overcloud-controller-0 ceph]# ceph pg dump_stuck unclean
ok
[root@overcloud-controller-0 ceph]#

[root@overcloud-controller-0 ceph]#  ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-controller-0,overcloud-controller-2
     osdmap e85: 32 osds: 24 up, 24 in
      pgmap v1426: 448 pgs, 4 pools, 22927 MB data, 5865 objects
            69434 MB used, 128 TB / 128 TB avail
                 448 active+clean
[root@overcloud-controller-0 ceph]# ceph health detail
HEALTH_OK
```

Run Ceph PG dump to validate that the OSD's do not have any copies.

From ceph osd tree, the OSD's in overcloud-cephstorage-1 are down. Hence make sure that there are no PG's in these OSD's.



```
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2              down    0      1.00000
 6   5.37000         osd.6              down    0      1.00000
10   5.37000         osd.10             down    0      1.00000
14   5.37000         osd.14             down    0      1.00000
18   5.37000         osd.18             down    0      1.00000
21   5.37000         osd.21             down    0      1.00000
25   5.37000         osd.25             down    0      1.00000
29   5.37000         osd.29             down    0      1.00000
```

```
osdstat kbused  kbavail kb       hb in   hb out
0       3171172 5761908872      5765080044      [1,3,5,7,9,12,13,16,17,19,20,22,23,26,27,28,30,31]      []
1       3285872 5761794172      5765080044      [0,3,4,5,7,8,11,12,13,15,16,17,19,20,23,24,27,28,30]    []
2       0       0       0       []      []
3       3585816 5761494228      5765080044      [0,1,4,5,7,8,9,11,13,15,17,20,22,23,24,26,28,30]        []
4       2752980 5762327064      5765080044      [1,3,5,7,9,11,12,13,16,17,19,22,23,24,26,27,31] []
5       2758356 5762321688      5765080044      [0,4,7,8,11,12,15,16,17,19,20,23,24,26,27,28,30,31]     []
6       0       0       0       []      []
7       3236160 5761843884      5765080044      [0,1,4,5,8,9,11,13,15,17,20,22,24,27,28,31]     []
8       2750288 5762329756      5765080044      [0,1,3,5,7,9,12,13,16,17,19,20,22,23,31]        []
9       2669732 5762410312      5765080044      [0,3,4,5,7,8,11,13,15,16,19,20,23,24,26,27,28,31]       []
10      0       0       0       []      []
11      4359392 5760720652      5765080044      [0,1,3,5,7,9,12,13,16,17,19,20,22,23,26,27,31] []
12      2917904 5762162140      5765080044      [0,1,3,4,5,8,11,13,15,16,17,20,22,23,24,26,28,30]       []
13      3484016 5761596028      5765080044      [0,3,4,7,8,9,11,12,15,16,20,23,26,27,28,31]     []
14      0       0       0       []      []
15      2465868 5762614176      5765080044      [1,5,7,8,9,12,13,16,17,20,23,26,30,31]  []
16      2244720 5762835324      5765080044      [0,1,4,5,7,8,9,11,12,13,15,17,20,22,28,30]      []
17      3867812 5761212232      5765080044      [0,1,3,4,7,8,11,12,13,15,16,19,20,22,23,24,27,28,30,31] []
18      0       0       0       []      []
19      2689608 5762390436      5765080044      [0,1,4,5,9,11,13,15,17,20,22,24,26,28,30]       []
20      3052816 5762027228      5765080044      [1,3,4,5,7,9,12,13,16,17,19,22,23,24,27,28,30,31]       []
21      0       0       0       []      []
22      2771896 5762308148      5765080044      [0,3,4,5,7,8,11,12,13,15,16,17,19,20,23,24,26,27,28,31] []
23      3695664 5761384380      5765080044      [0,1,3,4,5,7,8,9,11,13,15,17,20,22,24,26,28,30] []
24      2302492 5762777552      5765080044      [1,3,4,5,7,12,13,15,16,17,19,20,22,23,26,27,30,31]      []
25      0       0       0       []      []
26      2351216 5762728828      5765080044      [0,3,4,5,7,8,11,12,15,16,17,19,20,23,24,27,31] []
27      1828976 5763251068      5765080044      [0,1,3,4,5,8,9,13,15,20,22,24,26,28,30,31]      []
28      2858372 5762221672      5765080044      [1,3,5,7,9,12,13,16,17,19,20,22,23,26,27,30,31] []
29      0       0       0       []      []
30      2527800 5762552244      5765080044      [0,1,4,7,8,11,12,16,19,20,23,24,26,27,28,31]    []
31      3478924 5761601120      5765080044      [0,1,4,5,8,9,11,12,13,15,16,17,20,22,24,26,28,30]       []
 sum    71107852        138290813204    138361921056
```

This makes sure that there is nothing in osds 2,6,10,14,18,21,25 and 29. These are the OSD's that are part of the node that was deleted. Ceph moved all the copies from this node to other node.

Making sure that no placement groups are attached to the OSD's using ceph pg dump or ceph osd stat makes sure of data integrity. The above command confirms that all the data has been moved out of the OSD's. It is not recommended to delete a node with any placement groups residing in these OSD's. Please wait till the recovery activity is complete. Do not let the Ceph cluster reach its full ratio when removing nodes or OSD's. Removing OSD's could cause the cluster to reach full ratio and could cause data integrity issues.

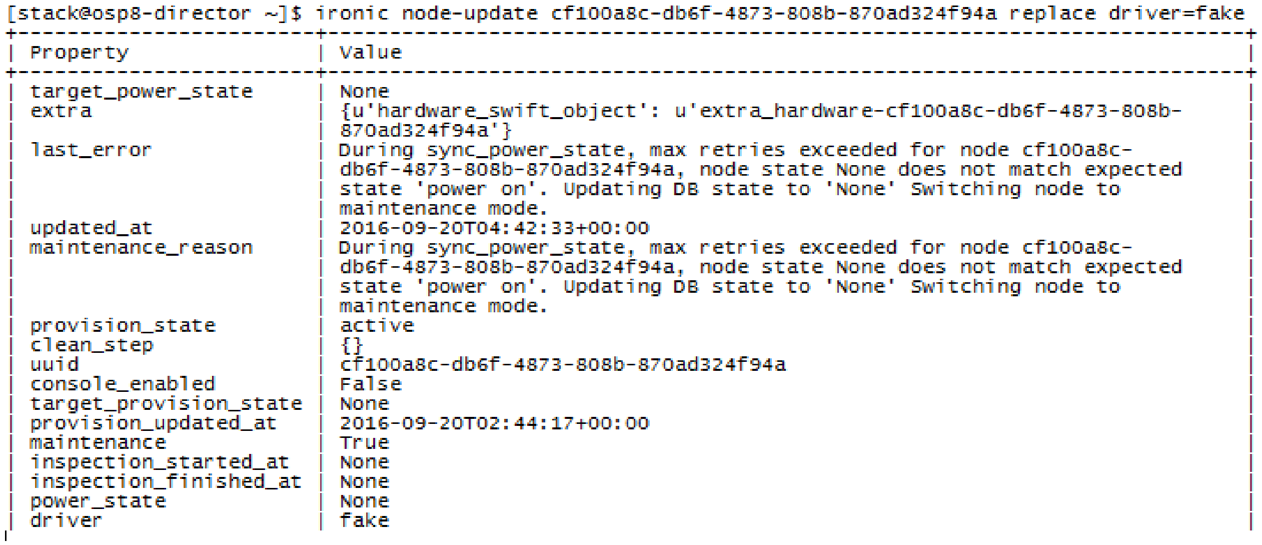
IPMI status.

As the node is switched off it is not reachable through IPMI.

```
[stack@osp8-director ~]$ ipmitool -I lanplus -H 10.23.10.75 -U admin -P <passwd>
chassis power status
Error: Unable to establish IPMI v2 / RMCP+ session
[stack@osp8-director ~]$
```

Update the driver entries to work around the issue.

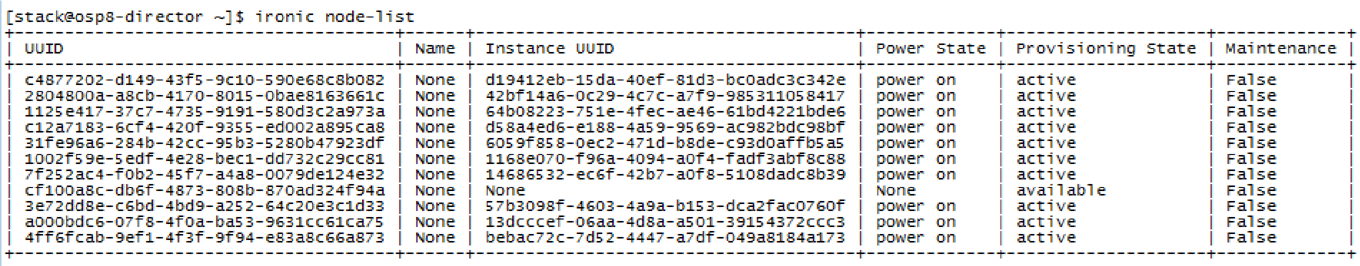
```
edit  vi /etc/ironic/ironic.conf
Update the enabled drivers temporarily as below
#enabled_drivers=pxe_ipmitool,pxe_ssh,pxe_drac
enabled_drivers=fake
Restart openstack-ironic-conductor
sudo service openstack-ironic-conductor restart

ironic node-update NODE_UUID replace driver=fake
```



```
[stack@osp8-director ~]$ ironic node-update cf100a8c-db6f-4873-808b-870ad324f94a replace driver=fake
+------------------------+-----------------------------------------------------------------------+
| Property               | Value                                                                 |
+------------------------+-----------------------------------------------------------------------+
| target_power_state     | None                                                                  |
| extra                  | {u'hardware_swift_object': u'extra_hardware-cf100a8c-db6f-4873-808b-  |
|                        | 870ad324f94a'}                                                        |
| last_error             | During sync_power_state, max retries exceeded for node cf100a8c-      |
|                        | db6f-4873-808b-870ad324f94a, node state None does not match expected  |
|                        | state 'power on'. Updating DB state to 'None' Switching node to       |
|                        | maintenance mode.                                                     |
| updated_at             | 2016-09-20T04:42:33+00:00                                             |
| maintenance_reason     | During sync_power_state, max retries exceeded for node cf100a8c-      |
|                        | db6f-4873-808b-870ad324f94a, node state None does not match expected  |
|                        | state 'power on'. Updating DB state to 'None' Switching node to       |
|                        | maintenance mode.                                                     |
| provision_state        | active                                                                |
| clean_step             | {}                                                                    |
| uuid                   | cf100a8c-db6f-4873-808b-870ad324f94a                                  |
| console_enabled        | False                                                                 |
| target_provision_state | None                                                                  |
| provision_updated_at   | 2016-09-20T02:44:17+00:00                                             |
| maintenance            | True                                                                  |
| inspection_started_at  | None                                                                  |
| inspection_finished_at | None                                                                  |
| power_state            | None                                                                  |
| driver                 | fake                                                                  |
```

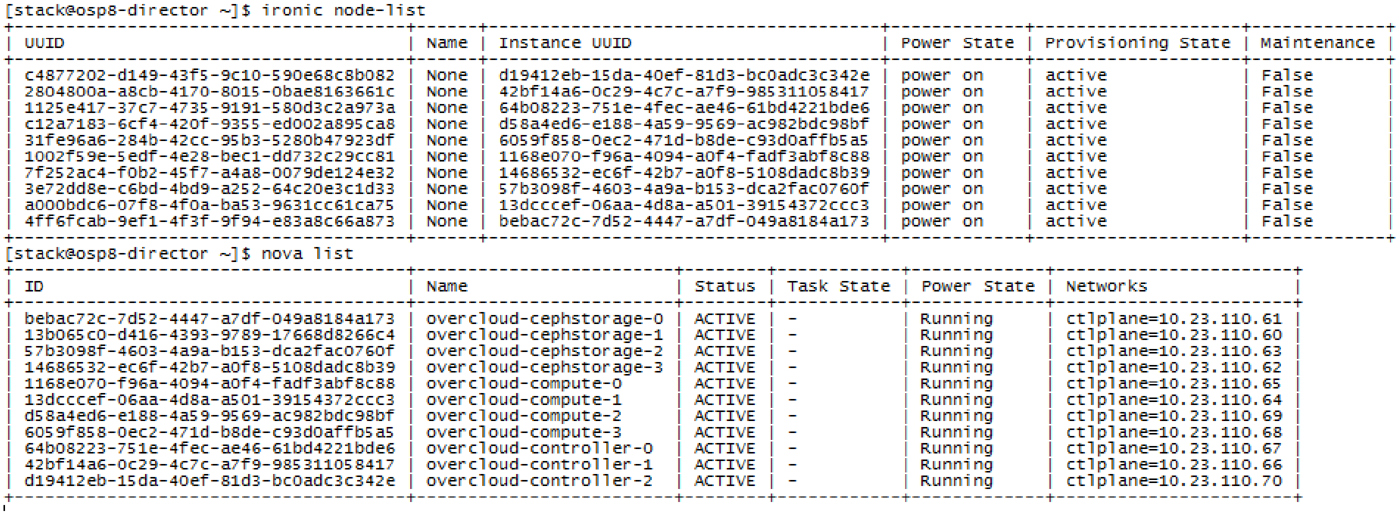
The node in ironic node-list should be with provision-state=active and maintenance=false
If not

```
[stack@osp8-director ~]$ ironic node-set-maintenance cf100a8c-db6f-4873-808b-
870ad324f94a false
[stack@osp8-director ~]$ ironic node-set-provision-state cf100a8c-db6f-4873-808b-
870ad324f94a deleted
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | d19412eb-15da-40ef-81d3-bc0adc3c342e | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 42bf14a6-0c29-4c7c-a7f9-985311058417 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | 64b08223-751e-4fec-ae46-61bd4221bde6 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | d58a4ed6-e188-4a59-9569-ac982bdc98bf | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | 6059f858-0ec2-471d-b8de-c93d0affb5a5 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 1168e070-f96a-4094-a0f4-fadf3abf8c88 | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 14686532-ec6f-42b7-a0f8-5108dadc8b39 | power on    | active             | False       |
| cf100a8c-db6f-4873-808b-870ad324f94a | None | None                                 | None        | available          | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | 57b3098f-4603-4a9a-b153-dca2fac0760f | power on    | active             | False       |
| a000bdc6-07f8-4f0a-ba53-9631cc61ca75 | None | 13dcccef-06aa-4d8a-a501-39154372ccc3 | power on    | active             | False       |
| 4ff6fcab-9ef1-4f3f-9f94-e83a8c66a873 | None | bebac72c-7d52-4447-a7df-049a8184a173 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
```

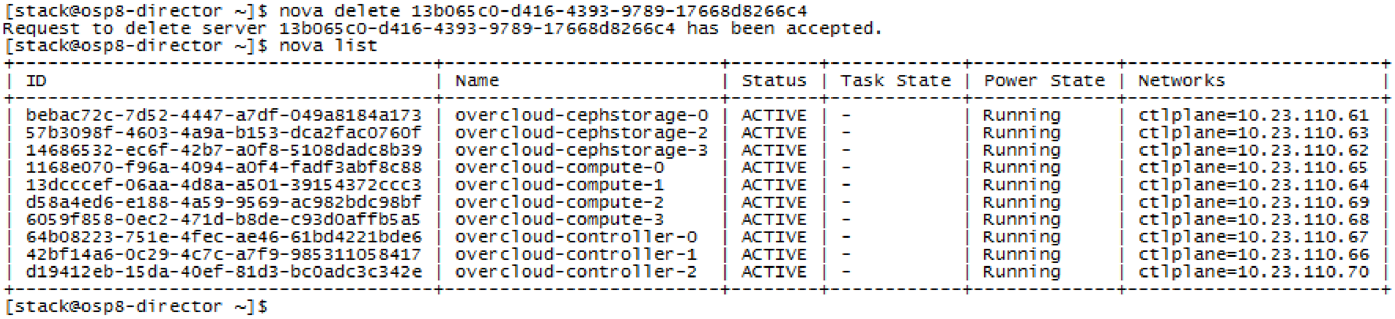
```
[stack@osp8-director ~]$ ironic node-delete cf100a8c-db6f-4873-808b-870ad324f94a
Deleted node cf100a8c-db6f-4873-808b-870ad324f94a
[stack@osp8-director ~]$
```



```
[stack@osp8-director ~]$ ironic node-list
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| UUID                                 | Name | Instance UUID                        | Power State | Provisioning State | Maintenance |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
| c4877202-d149-43f5-9c10-590e68c8b082 | None | d19412eb-15da-40ef-81d3-bc0adc3c342e | power on    | active             | False       |
| 2804800a-a8cb-4170-8015-0bae8163661c | None | 42bf14a6-0c29-4c7c-a7f9-985311058417 | power on    | active             | False       |
| 1125e417-37c7-4735-9191-580d3c2a973a | None | 64b08223-751e-4fec-ae46-61bd4221bde6 | power on    | active             | False       |
| c12a7183-6cf4-420f-9355-ed002a895ca8 | None | d58a4ed6-e188-4a59-9569-ac982bdc98bf | power on    | active             | False       |
| 31fe96a6-284b-42cc-95b3-5280b47923df | None | 6059f858-0ec2-471d-b8de-c93d0affb5a5 | power on    | active             | False       |
| 1002f59e-5edf-4e28-bec1-dd732c29cc81 | None | 1168e070-f96a-4094-a0f4-fadf3abf8c88 | power on    | active             | False       |
| 7f252ac4-f0b2-45f7-a4a8-0079de124e32 | None | 14686532-ec6f-42b7-a0f8-5108dadc8b39 | power on    | active             | False       |
| 3e72dd8e-c6bd-4bd9-a252-64c20e3c1d33 | None | 57b3098f-4603-4a9a-b153-dca2fac0760f | power on    | active             | False       |
| a000bdc6-07f8-4f0a-ba53-9631cc61ca75 | None | 13dcccef-06aa-4d8a-a501-39154372ccc3 | power on    | active             | False       |
| 4ff6fcab-9ef1-4f3f-9f94-e83a8c66a873 | None | bebac72c-7d52-4447-a7df-049a8184a173 | power on    | active             | False       |
+--------------------------------------+------+--------------------------------------+-------------+--------------------+-------------+
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks              |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| bebac72c-7d52-4447-a7df-049a8184a173 | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.61 |
| 13b065c0-d416-4393-9789-17668d8266c4 | overcloud-cephstorage-1 | ACTIVE | -          | Running     | ctlplane=10.23.110.60 |
| 57b3098f-4603-4a9a-b153-dca2fac0760f | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.63 |
| 14686532-ec6f-42b7-a0f8-5108dadc8b39 | overcloud-cephstorage-3 | ACTIVE | -          | Running     | ctlplane=10.23.110.62 |
| 1168e070-f96a-4094-a0f4-fadf3abf8c88 | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.65 |
| 13dcccef-06aa-4d8a-a501-39154372ccc3 | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.64 |
| d58a4ed6-e188-4a59-9569-ac982bdc98bf | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.69 |
| 6059f858-0ec2-471d-b8de-c93d0affb5a5 | overcloud-compute-3     | ACTIVE | -          | Running     | ctlplane=10.23.110.68 |
| 64b08223-751e-4fec-ae46-61bd4221bde6 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.67 |
| 42bf14a6-0c29-4c7c-a7f9-985311058417 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.66 |
| d19412eb-15da-40ef-81d3-bc0adc3c342e | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.70 |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
```

From the above it appears that nova list still has an entry for the deleted node.

```
[stack@osp8-director ~]$ nova delete 13b065c0-d416-4393-9789-17668d8266c4
Request to delete server 13b065c0-d416-4393-9789-17668d8266c4 has been accepted.
```



```
[stack@osp8-director ~]$ nova delete 13b065c0-d416-4393-9789-17668d8266c4
Request to delete server 13b065c0-d416-4393-9789-17668d8266c4 has been accepted.
[stack@osp8-director ~]$ nova list
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| ID                                   | Name                    | Status | Task State | Power State | Networks              |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
| bebac72c-7d52-4447-a7df-049a8184a173 | overcloud-cephstorage-0 | ACTIVE | -          | Running     | ctlplane=10.23.110.61 |
| 57b3098f-4603-4a9a-b153-dca2fac0760f | overcloud-cephstorage-2 | ACTIVE | -          | Running     | ctlplane=10.23.110.63 |
| 14686532-ec6f-42b7-a0f8-5108dadc8b39 | overcloud-cephstorage-3 | ACTIVE | -          | Running     | ctlplane=10.23.110.62 |
| 1168e070-f96a-4094-a0f4-fadf3abf8c88 | overcloud-compute-0     | ACTIVE | -          | Running     | ctlplane=10.23.110.65 |
| 13dcccef-06aa-4d8a-a501-39154372ccc3 | overcloud-compute-1     | ACTIVE | -          | Running     | ctlplane=10.23.110.64 |
| d58a4ed6-e188-4a59-9569-ac982bdc98bf | overcloud-compute-2     | ACTIVE | -          | Running     | ctlplane=10.23.110.69 |
| 6059f858-0ec2-471d-b8de-c93d0affb5a5 | overcloud-compute-3     | ACTIVE | -          | Running     | ctlplane=10.23.110.68 |
| 64b08223-751e-4fec-ae46-61bd4221bde6 | overcloud-controller-0  | ACTIVE | -          | Running     | ctlplane=10.23.110.67 |
| 42bf14a6-0c29-4c7c-a7f9-985311058417 | overcloud-controller-1  | ACTIVE | -          | Running     | ctlplane=10.23.110.66 |
| d19412eb-15da-40ef-81d3-bc0adc3c342e | overcloud-controller-2  | ACTIVE | -          | Running     | ctlplane=10.23.110.70 |
+--------------------------------------+-------------------------+--------+------------+-------------+-----------------------+
[stack@osp8-director ~]$
```

```
[stack@osp8-director ~]$ nova list | grep ACTIVE | wc -l
10
[stack@osp8-director ~]$ ironic node-list | grep None | wc -l
10
```

Revert back the “fake” driver from ironic.conf.

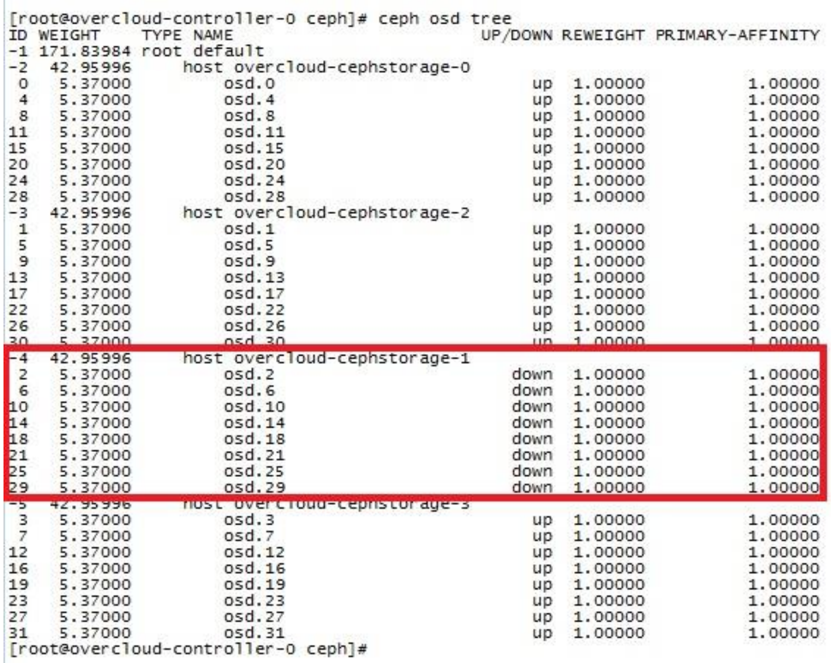
```
edit vi /etc/ironic/ironic.conf.
enabled_drivers=pxe_ipmitool,pxe_ssh,pxe_drac
#enabled_drivers=fake
Restart ironic-conductor to pick up the drivers again.
Sudo service openstack-ironic-conductor restart
```

Storage node deletion differs from compute node deletion here. In both the cases we have deleted the nodes from UCS and OpenStack so far. However ceph entries still remain and these have to be cleaned up.

### Clean Up Ceph after Node Deletion

To clean up Ceph after a node deletion, complete the following steps:

Check the details from ceph health and osd tree:



```
[root@overcloud-controller-0 ceph]# ceph osd tree
ID WEIGHT    TYPE NAME                          UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 171.83984 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                           up  1.00000          1.00000
 4   5.37000         osd.4                           up  1.00000          1.00000
 8   5.37000         osd.8                           up  1.00000          1.00000
11   5.37000         osd.11                          up  1.00000          1.00000
15   5.37000         osd.15                          up  1.00000          1.00000
20   5.37000         osd.20                          up  1.00000          1.00000
24   5.37000         osd.24                          up  1.00000          1.00000
28   5.37000         osd.28                          up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                           up  1.00000          1.00000
 5   5.37000         osd.5                           up  1.00000          1.00000
 9   5.37000         osd.9                           up  1.00000          1.00000
13   5.37000         osd.13                          up  1.00000          1.00000
17   5.37000         osd.17                          up  1.00000          1.00000
22   5.37000         osd.22                          up  1.00000          1.00000
26   5.37000         osd.26                          up  1.00000          1.00000
30   5.37000         osd.30                          up  1.00000          1.00000
-4  42.95996     host overcloud-cephstorage-1
 2   5.37000         osd.2                         down  1.00000          1.00000
 6   5.37000         osd.6                         down  1.00000          1.00000
10   5.37000         osd.10                        down  1.00000          1.00000
14   5.37000         osd.14                        down  1.00000          1.00000
18   5.37000         osd.18                        down  1.00000          1.00000
21   5.37000         osd.21                        down  1.00000          1.00000
25   5.37000         osd.25                        down  1.00000          1.00000
29   5.37000         osd.29                        down  1.00000          1.00000
-5  42.95996     host overcloud-cephstorage-3
 3   5.37000         osd.3                           up  1.00000          1.00000
 7   5.37000         osd.7                           up  1.00000          1.00000
12   5.37000         osd.12                          up  1.00000          1.00000
16   5.37000         osd.16                          up  1.00000          1.00000
19   5.37000         osd.19                          up  1.00000          1.00000
23   5.37000         osd.23                          up  1.00000          1.00000
27   5.37000         osd.27                          up  1.00000          1.00000
31   5.37000         osd.31                          up  1.00000          1.00000
[root@overcloud-controller-0 ceph]#
```

```
[root@overcloud-controller-0 ceph]# ceph osd stat
     osdmap e85: 32 osds: 24 up, 24 in
```

Remove OSD's from Ceph. Change the OSD ID's to your setup and from the output of osd tree above.

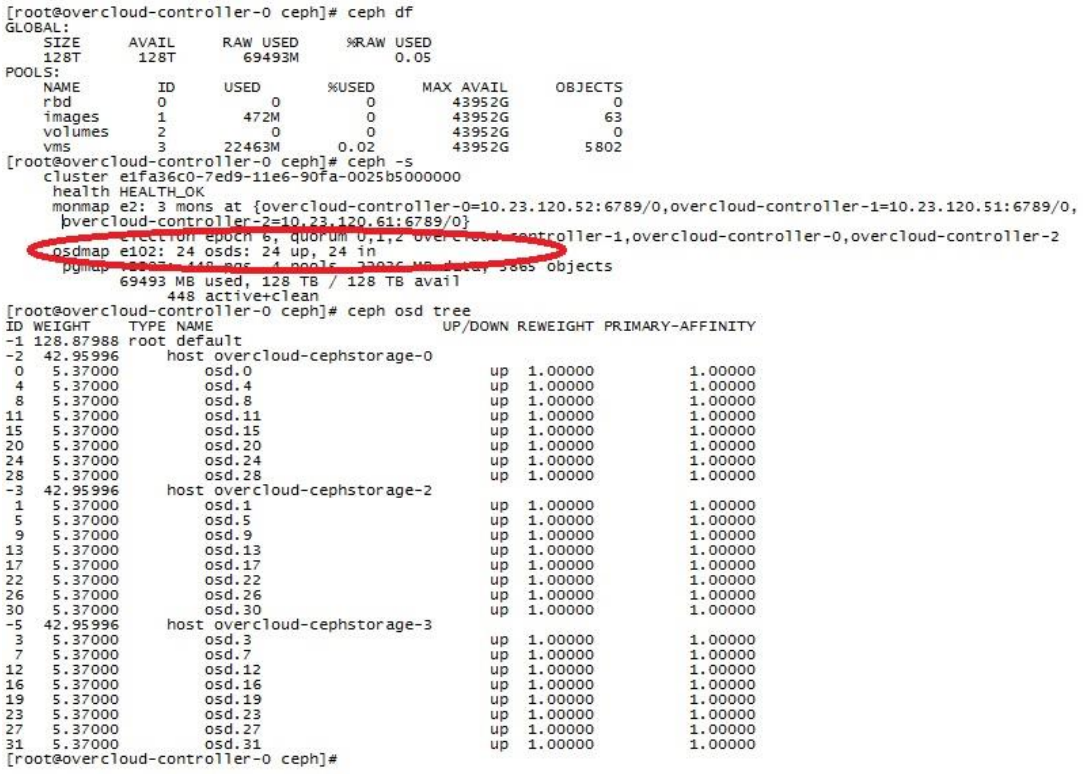
```
for i in 2 6 10 14 18 21 25 29
do
ceph osd out $i
ceph osd crush remove osd.$i
ceph auth del osd.$i
ceph osd rm $i
done
      osd.2 is already out.
      removed item id 2 name 'osd.2' from crush map
      updated
      removed osd.2
      osd.6 is already out.
      removed item id 6 name 'osd.6' from crush map
      updated
      removed osd.6
      osd.10 is already out.
      removed item id 10 name 'osd.10' from crush map
      updated
      removed osd.10
      osd.14 is already out.
      removed item id 14 name 'osd.14' from crush map
      updated
```

```
removed osd.14
osd.18 is already out.
removed item id 18 name 'osd.18' from crush map
updated
removed osd.18
osd.21 is already out.
removed item id 21 name 'osd.21' from crush map
updated
removed osd.21
osd.25 is already out.
removed item id 25 name 'osd.25' from crush map
updated
removed osd.25
osd.29 is already out.
removed item id 29 name 'osd.29' from crush map
updated
removed osd.29
[root@overcloud-controller-0 ceph]#
```

Clean up ceph crush host entries:

```
[root@overcloud-controller-0 ceph]# ceph osd crush remove overcloud-
cephstorage-1
removed item id -4 name 'overcloud-cephstorage-1' from crush map
```

Health checks after deletion:



```
[root@overcloud-controller-0 ceph]# ceph df
GLOBAL:
    SIZE     AVAIL     RAW USED     %RAW USED
    128T      128T       69493M          0.05
POOLS:
    NAME         ID     USED       %USED     MAX AVAIL     OBJECTS
    rbd          0           0         0        43952G           0
    images       1        472M         0        43952G          63
    volumes      2           0         0        43952G           0
    vms          3      22463M      0.02        43952G        5802
[root@overcloud-controller-0 ceph]# ceph -s
    cluster e1fa36c0-7ed9-11e6-90fa-0025b5000000
     health HEALTH_OK
     monmap e2: 3 mons at {overcloud-controller-0=10.23.120.52:6789/0,overcloud-controller-1=10.23.120.51:6789/0,
      overcloud-controller-2=10.23.120.61:6789/0}
            election epoch 6, quorum 0,1,2 overcloud-controller-1,overcloud-controller-0,overcloud-controller-2
     osdmap e102: 24 osds: 24 up, 24 in
      pgmap v[illegible]: 448 pgs, 4 pools, 22936 MB data, 5865 objects
            69493 MB used, 128 TB / 128 TB avail
                 448 active+clean
[root@overcloud-controller-0 ceph]# ceph osd tree
ID WEIGHT    TYPE NAME                        UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 128.87988 root default
-2  42.95996     host overcloud-cephstorage-0
 0   5.37000         osd.0                         up  1.00000          1.00000
 4   5.37000         osd.4                         up  1.00000          1.00000
 8   5.37000         osd.8                         up  1.00000          1.00000
11   5.37000         osd.11                        up  1.00000          1.00000
15   5.37000         osd.15                        up  1.00000          1.00000
20   5.37000         osd.20                        up  1.00000          1.00000
24   5.37000         osd.24                        up  1.00000          1.00000
28   5.37000         osd.28                        up  1.00000          1.00000
-3  42.95996     host overcloud-cephstorage-2
 1   5.37000         osd.1                         up  1.00000          1.00000
 5   5.37000         osd.5                         up  1.00000          1.00000
 9   5.37000         osd.9                         up  1.00000          1.00000
13   5.37000         osd.13                        up  1.00000          1.00000
17   5.37000         osd.17                        up  1.00000          1.00000
22   5.37000         osd.22                        up  1.00000          1.00000
26   5.37000         osd.26                        up  1.00000          1.00000
30   5.37000         osd.30                        up  1.00000          1.00000
-5  42.95996     host overcloud-cephstorage-3
 3   5.37000         osd.3                         up  1.00000          1.00000
 7   5.37000         osd.7                         up  1.00000          1.00000
12   5.37000         osd.12                        up  1.00000          1.00000
16   5.37000         osd.16                        up  1.00000          1.00000
19   5.37000         osd.19                        up  1.00000          1.00000
23   5.37000         osd.23                        up  1.00000          1.00000
27   5.37000         osd.27                        up  1.00000          1.00000
31   5.37000         osd.31                        up  1.00000          1.00000
[root@overcloud-controller-0 ceph]#
```

## Node Addition

When the storage node has been completely removed from OpenStack and the ceph entries cleaned, a new server can be added. The procedure for adding a new storage node is same as how it was addressed earlier in upscaling the storage pod.

## HA on Undercloud Node

RHOSP 8 supports one Undercloud Node as of the date this document was first published. Also in the test bed, the compute and storage nodes are NATed through Undercloud node. Though this does not pose any challenges during Overcloud operation, any future heat stack or overcloud deploys could be impacted.

The following backup and recovery method has been documented on Red Hat web site for reference. This procedure has not been validated in this CVD. It is strongly recommended to test the below procedure in a test environment and document the process to restore the Undercloud node from backup. Subsequently take a backup of the Undercloud node and store the back up for an easy retrieval later in case of failures.

https://access.redhat.com/webassets/avalon/d/Red_Hat_OpenStack_Platform-8-Back_Up_and_Restore_Red_Hat_OpenStack_Platform-en-US/Red_Hat_OpenStack_Platform-8-Back_Up_and_Restore_Red_Hat_OpenStack_Platform-en-US.pdf

# Hardware Failures of Blade Servers

The hardware failures of blade servers are infrequent and happen very rarely. Cisco stands behind the customers to support in such conditions. There is also a Return Material Authorization (RMA) process in place. Depending on the types of failure, either the parts or the entire blade may be replaced. This section at a high level covers the types of failures that could happen on Cisco UCS blades running OpenStack and how to get the system up and running with little or no business interruption.

This section was validated specifically for Controller blades. The replacement of compute and storage blades are covered earlier in the High Availability section.

## Types of Failures

- CPU Failures
- Memory or DIMM Failures
- Virtual Interface Card Failures
- Motherboard Failures
- Hard Disk Failures
- Chassis Slot Issues

Any such failures happening on a blade either leads to degraded performance while the system continues to operate (like DIMM or disk Failures) or it could fail completely. In case of complete failures, OpenStack Nova and Ironic may also take them offline and there is a need to fix the errors.

A compute node failure will impact only the VMs running on the compute node and these can be evacuated to another node.

Ceph storage nodes are configured with replication factor of 3, and the system continues to operate though the recovery operation may cause slight degraded performance of the storage cluster.

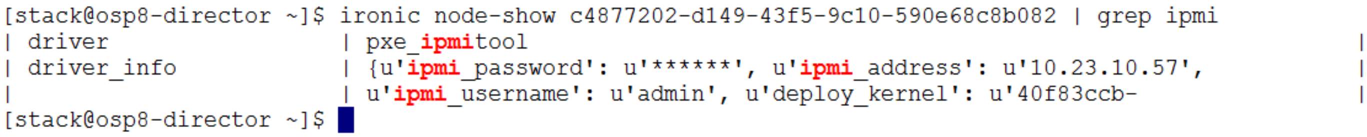
In case of total failure of controller blades, the fencing packages will fence the failed node. You may need a fix for this bug 1298430. Instructions provided earlier in the document on how to get the fix included in the overcloud image.

## OpenStack Dependency on Hardware

From OpenStack point of view, the following hardware variables are seeded into the system and these may have to be addressed in case of failures:

### IPMI Address

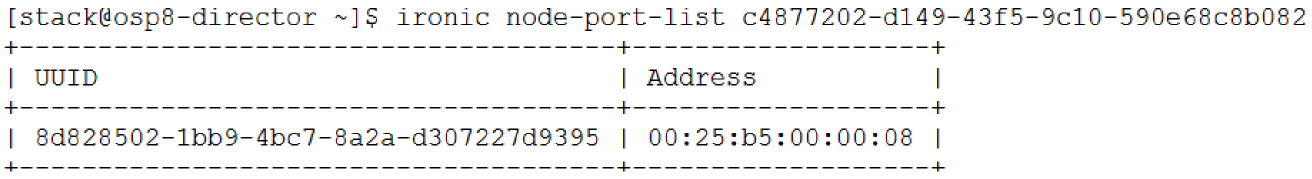
OpenStack uses IPMI address and it powers on/off the blades with this address. These can be queried through ironic API's as below.



```
[stack@osp8-director ~]$ ironic node-show c4877202-d149-43f5-9c10-590e68c8b082 | grep ipmi
| driver                     | pxe_ipmitool                                                           |
| driver_info                | {u'ipmi_password': u'******', u'ipmi_address': u'10.23.10.57',         |
|                            | u'ipmi_username': u'admin', u'deploy_kernel': u'40f83ccb-              |
[stack@osp8-director ~]$
```

## NICs and MAC addresses

The controller Ethernet interfaces and MAC addresses are available in the Local Disk of the failed blade. Hence failure cases of hard disks is also included above. Apart from this, the provisioning Interface MAC address is also stored in the Undercloud node.



```
[stack@osp8-director ~]$ ironic node-port-list c4877202-d149-43f5-9c10-590e68c8b082
+--------------------------------------+-------------------+
| UUID                                 | Address           |
+--------------------------------------+-------------------+
| 8d828502-1bb9-4bc7-8a2a-d307227d9395 | 00:25:b5:00:00:08 |
+--------------------------------------+-------------------+
```

Retain these addresses in case of failures.

## Local Disk

The local hard disk has all the configuration information and should be available. It is strongly recommended to have a pair of Local Disks in RAID-10 configuration to overcome against disk failures.

Post hardware failures, if all of the above are brought back, the system can be made operational and this is what is addressed in this section.

# Cisco UCS Failure Scenarios

As mentioned earlier there can be several types of failures including CPU or Memory and system may perform in a degraded fashion. Not all of these are covered in this document, but the ones which have hooks to OpenStack are covered here.

## Hard Disk Failure

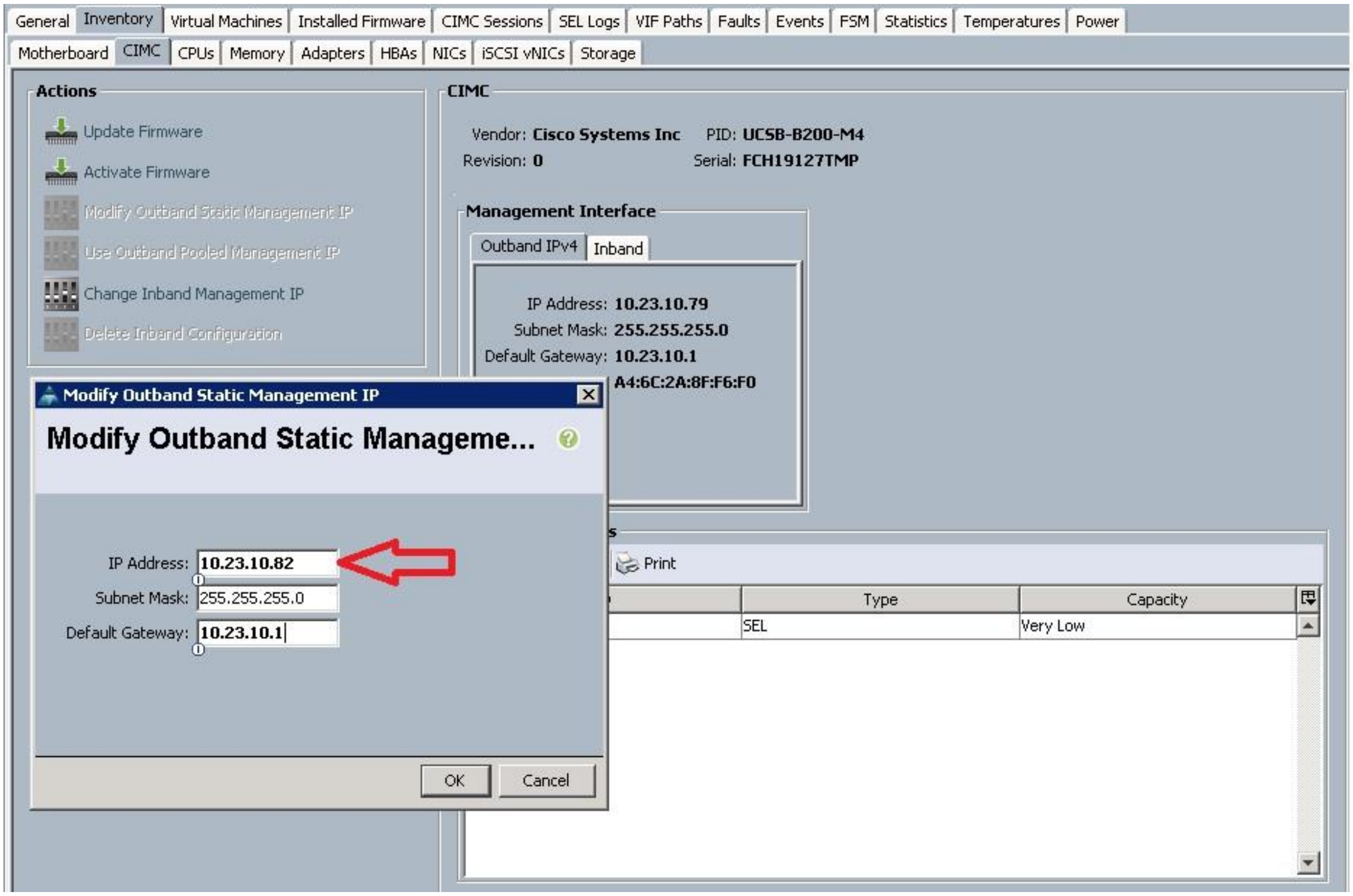
Assuming you have followed the recommendations to have RAID-10 configuration for the local disks, failure of one disk will be taken care by the RAID controller.

## Blade Server Replacement

In case there is a need to replace the blade, the ipmi address, MAC addresses and Local disks have to be restored. It is assumed that there is no double failure here.

### IPMI Address

The IPMI addresses are allocated from the KVM pool. When a blade fails system will hold the address until it has been decommissioned. If system is decommissioned, it will release the free IP to KVM Pool. We can allocate this free old IP to the new blade. The below figure shows how to change the IPMI address in UCS as an example.



Vendor: **Cisco Systems Inc** PID: **UCSB-B200-M4**
Revision: **0** Serial: **FCH19127TMP**

IP Address: **10.23.10.79**
Subnet Mask: **255.255.255.0**
Default Gateway: **10.23.10.1**
**A4:6C:2A:8F:F6:F0**

Modify Outbound Static Management IP

**Modify Outbound Static Manageme...**

IP Address: **10.23.10.82**
Subnet Mask: 255.255.255.0
Default Gateway: **10.23.10.1**

| Type | Capacity |
|---|---|
| SEL | Very Low |

### NICs and MAC Addresses

Service profiles are like SIM card of a phone that store all the hardware identity. Once the Service Profile is disassociated from the failed node and attached to the new node, all of the policies like Boot Policy and Network interfaces along with MAC addresses are available to the new blade.

### Local Disks

The two hard disks can be taken out from the failed blade and inserted into the new blade. You have to make sure that the new blade is identical and upgraded to the same firmware version as of the failed blade. The local disks have the controller binaries and the cluster configuration information. Associating the service profile will bring up all the hardware profiles on the new blade. Hence now system will be in sync from both hardware and software side and should be up and running.

## Case Study

The following case study describes the step by step process on how to replace the controller blade.

**PCS Status**

```
[root@controller-0 ~]# pcs status
Cluster name: tripleo_cluster
Last updated: Wed Aug 17 10:37:28 2016          Last change: Wed Aug 17 08:01:46
2016 by root via cibadmin on controller-0
Stack: corosync
Current DC: controller-0 (version 1.1.13-10.el7_2.2-44eb2dd) - partition with
quorum
3 nodes and 113 resources configured

Online: [ controller-0 controller-1 controller-2 ]
```

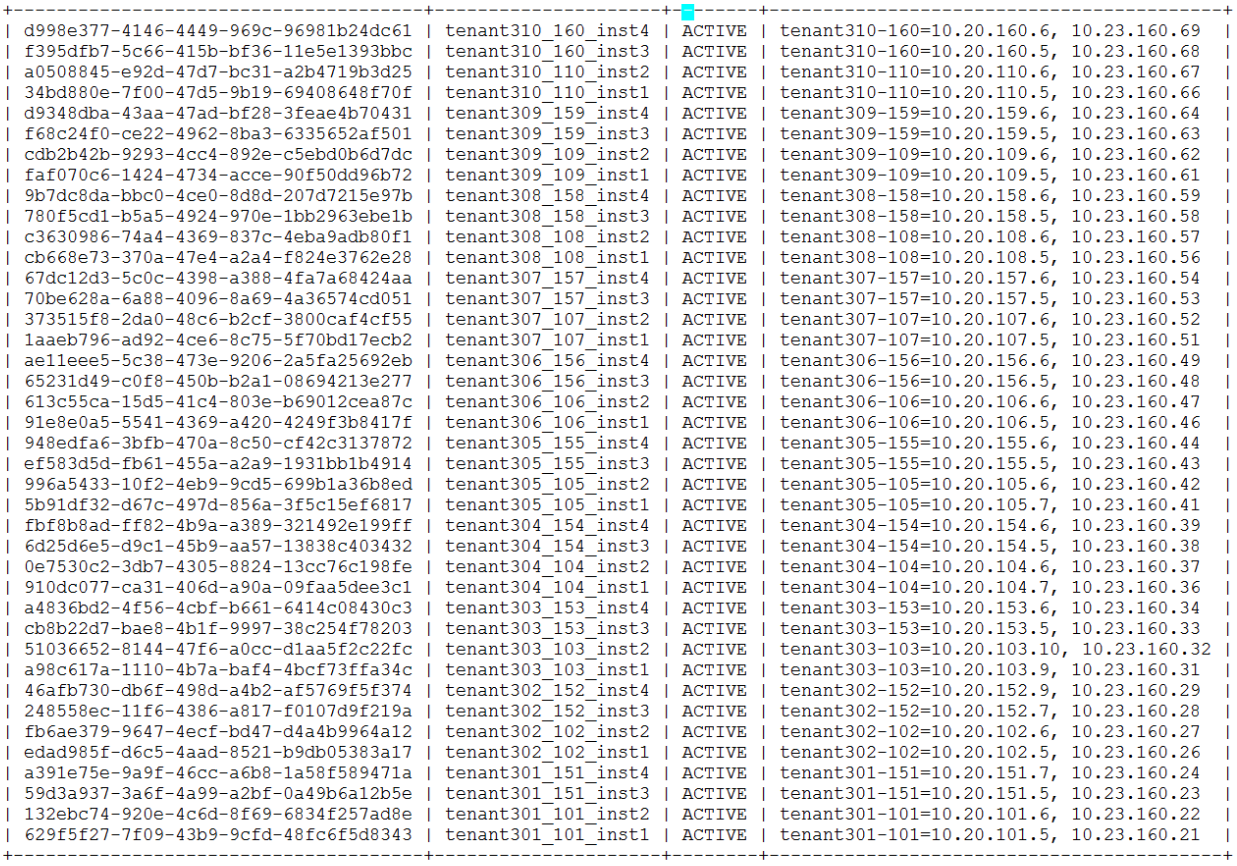
```
Full list of resources:

 ip-10.23.110.56        (ocf::heartbeat:IPaddr2):       Started controller-0
 Clone Set: haproxy-clone [haproxy]
     Started: [ controller-0 controller-1 controller-2 ]
 ip-10.23.120.50        (ocf::heartbeat:IPaddr2):       Started controller-1
 ip-10.23.150.50        (ocf::heartbeat:IPaddr2):       Started controller-2
 ip-10.23.100.51        (ocf::heartbeat:IPaddr2):       Started controller-0
 ip-172.22.215.16       (ocf::heartbeat:IPaddr2):       Started controller-1
```

### Quorum and CRM Node Information

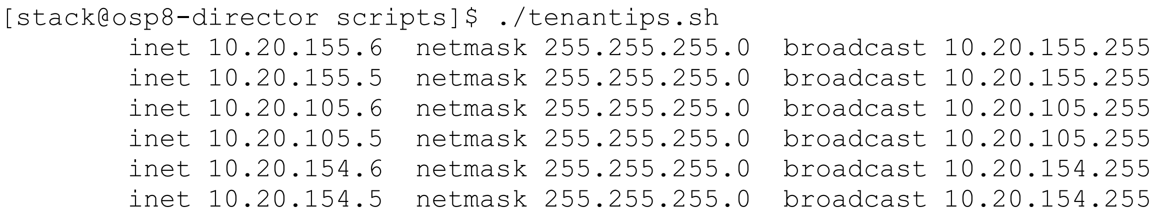
```
[root@overcloud-controller-0 ~]# corosync-quorumtool
Quorum information
------------------
Nodes:            3
Node ID:          1
Ring ID:          24
Quorate:          Yes

Votequorum information
----------------------
Expected votes:   3
Highest expected: 3
Total votes:      3
Quorum:           2
Flags:            Quorate

Membership information
----------------------
    Nodeid      Votes Name
         2          1 overcloud-controller-1
         1          1 overcloud-controller-0 (local)
         3          1 overcloud-controller-2
```

### Tenants Availability and Checks

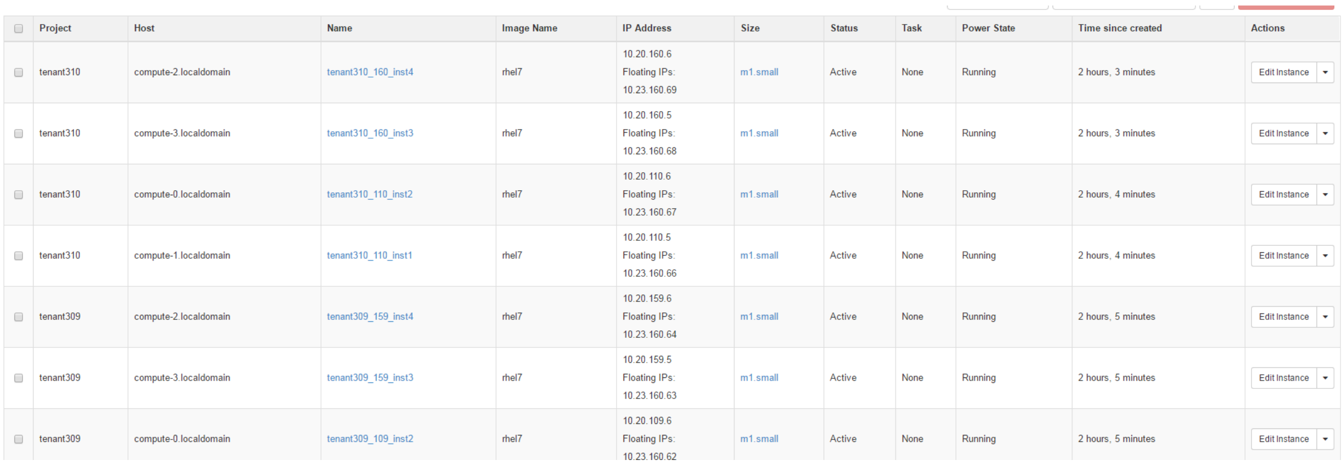
Make sure that the connectivity to the tenants work.



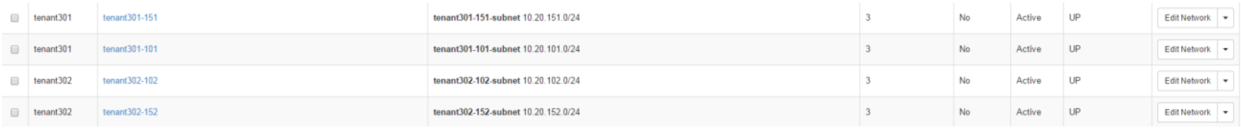
```
+--------------------------------------+---------------------+--------+-------------------------------------------+
| d998e377-4146-4449-969c-96981b24dc61 | tenant310_160_inst4 | ACTIVE | tenant310-160=10.20.160.6, 10.23.160.69  |
| f395dfb7-5c66-415b-bf36-11e5e1393bbc | tenant310_160_inst3 | ACTIVE | tenant310-160=10.20.160.5, 10.23.160.68  |
| a0508845-e92d-47d7-bc31-a2b4719b3d25 | tenant310_110_inst2 | ACTIVE | tenant310-110=10.20.110.6, 10.23.160.67  |
| 34bd880e-7f00-47d5-9b19-69408648f70f | tenant310_110_inst1 | ACTIVE | tenant310-110=10.20.110.5, 10.23.160.66  |
| d9348dba-43aa-47ad-bf28-3feae4b70431 | tenant309_159_inst4 | ACTIVE | tenant309-159=10.20.159.6, 10.23.160.64  |
| f68c24f0-ce22-4962-8ba3-6335652af501 | tenant309_159_inst3 | ACTIVE | tenant309-159=10.20.159.5, 10.23.160.63  |
| cdb2b42b-9293-4cc4-892e-c5ebd0b6d7dc | tenant309_109_inst2 | ACTIVE | tenant309-109=10.20.109.6, 10.23.160.62  |
| faf070c6-1424-4734-acce-90f50dd96b72 | tenant309_109_inst1 | ACTIVE | tenant309-109=10.20.109.5, 10.23.160.61  |
| 9b7dc8da-bbc0-4ce0-8d8d-207d7215e97b | tenant308_158_inst4 | ACTIVE | tenant308-158=10.20.158.6, 10.23.160.59  |
| 780f5cd1-b5a5-4924-970e-1bb2963ebe1b | tenant308_158_inst3 | ACTIVE | tenant308-158=10.20.158.5, 10.23.160.58  |
| c3630986-74a4-4369-837c-4eba9adb80f1 | tenant308_108_inst2 | ACTIVE | tenant308-108=10.20.108.6, 10.23.160.57  |
| cb668e73-370a-47e4-a2a4-f824e3762e28 | tenant308_108_inst1 | ACTIVE | tenant308-108=10.20.108.5, 10.23.160.56  |
| 67dc12d3-5c0c-4398-a388-4fa7a68424aa | tenant307_157_inst4 | ACTIVE | tenant307-157=10.20.157.6, 10.23.160.54  |
| 70be628a-6a88-4096-8a69-4a36574cd051 | tenant307_157_inst3 | ACTIVE | tenant307-157=10.20.157.5, 10.23.160.53  |
| 373515f8-2da0-48c6-b2cf-3800caf4cf55 | tenant307_107_inst2 | ACTIVE | tenant307-107=10.20.107.6, 10.23.160.52  |
| 1aaeb796-ad92-4ce6-8c75-5f70bd17ecb2 | tenant307_107_inst1 | ACTIVE | tenant307-107=10.20.107.5, 10.23.160.51  |
| ae11eee5-5c38-473e-9206-2a5fa25692eb | tenant306_156_inst4 | ACTIVE | tenant306-156=10.20.156.6, 10.23.160.49  |
| 65231d49-c0f8-450b-b2a1-08694213e277 | tenant306_156_inst3 | ACTIVE | tenant306-156=10.20.156.5, 10.23.160.48  |
| 613c55ca-15d5-41c4-803e-b69012cea87c | tenant306_106_inst2 | ACTIVE | tenant306-106=10.20.106.6, 10.23.160.47  |
| 91e8e0a5-5541-4369-a420-4249f3b8417f | tenant306_106_inst1 | ACTIVE | tenant306-106=10.20.106.5, 10.23.160.46  |
| 948edfa6-3bfb-470a-8c50-cf42c3137872 | tenant305_155_inst4 | ACTIVE | tenant305-155=10.20.155.6, 10.23.160.44  |
| ef583d5d-fb61-455a-a2a9-1931bb1b4914 | tenant305_155_inst3 | ACTIVE | tenant305-155=10.20.155.5, 10.23.160.43  |
| 996a5433-10f2-4eb9-9cd5-699b1a36b8ed | tenant305_105_inst2 | ACTIVE | tenant305-105=10.20.105.6, 10.23.160.42  |
| 5b91df32-d67c-497d-856a-3f5c15ef6817 | tenant305_105_inst1 | ACTIVE | tenant305-105=10.20.105.7, 10.23.160.41  |
| fbf8b8ad-ff82-4b9a-a389-321492e199ff | tenant304_154_inst4 | ACTIVE | tenant304-154=10.20.154.6, 10.23.160.39  |
| 6d25d6e5-d9c1-45b9-aa57-13838c403432 | tenant304_154_inst3 | ACTIVE | tenant304-154=10.20.154.5, 10.23.160.38  |
| 0e7530c2-3db7-4305-8824-13cc76c198fe | tenant304_104_inst2 | ACTIVE | tenant304-104=10.20.104.6, 10.23.160.37  |
| 910dc077-ca31-406d-a90a-09faa5dee3c1 | tenant304_104_inst1 | ACTIVE | tenant304-104=10.20.104.7, 10.23.160.36  |
| a4836bd2-4f56-4cbf-b661-6414c08430c3 | tenant303_153_inst4 | ACTIVE | tenant303-153=10.20.153.6, 10.23.160.34  |
| cb8b22d7-bae8-4b1f-9997-38c254f78203 | tenant303_153_inst3 | ACTIVE | tenant303-153=10.20.153.5, 10.23.160.33  |
| 51036652-8144-47f6-a0cc-d1aa5f2c22fc | tenant303_103_inst2 | ACTIVE | tenant303-103=10.20.103.10, 10.23.160.32 |
| a98c617a-1110-4b7a-baf4-4bcf73ffa34c | tenant303_103_inst1 | ACTIVE | tenant303-103=10.20.103.9, 10.23.160.31  |
| 46afb730-db6f-498d-a4b2-af5769f5f374 | tenant302_152_inst4 | ACTIVE | tenant302-152=10.20.152.9, 10.23.160.29  |
| 248558ec-11f6-4386-a817-f0107d9f219a | tenant302_152_inst3 | ACTIVE | tenant302-152=10.20.152.7, 10.23.160.28  |
| fb6ae379-9647-4ecf-bd47-d4a4b9964a12 | tenant302_102_inst2 | ACTIVE | tenant302-102=10.20.102.6, 10.23.160.27  |
| edad985f-d6c5-4aad-8521-b9db05383a17 | tenant302_102_inst1 | ACTIVE | tenant302-102=10.20.102.5, 10.23.160.26  |
| a391e75e-9a9f-46cc-a6b8-1a58f589471a | tenant301_151_inst4 | ACTIVE | tenant301-151=10.20.151.7, 10.23.160.24  |
| 59d3a937-3a6f-4a99-a2bf-0a49b6a12b5e | tenant301_151_inst3 | ACTIVE | tenant301-151=10.20.151.5, 10.23.160.23  |
| 132ebc74-920e-4c6d-8f69-6834f257ad8e | tenant301_101_inst2 | ACTIVE | tenant301-101=10.20.101.6, 10.23.160.22  |
| 629f5f27-7f09-43b9-9cfd-48fc6f5d8343 | tenant301_101_inst1 | ACTIVE | tenant301-101=10.20.101.5, 10.23.160.21  |
+--------------------------------------+---------------------+--------+-------------------------------------------+
```



```
[stack@osp8-director scripts]$ ./tenantips.sh
        inet 10.20.155.6  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.155.5  netmask 255.255.255.0  broadcast 10.20.155.255
        inet 10.20.105.6  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.105.5  netmask 255.255.255.0  broadcast 10.20.105.255
        inet 10.20.154.6  netmask 255.255.255.0  broadcast 10.20.154.255
        inet 10.20.154.5  netmask 255.255.255.0  broadcast 10.20.154.255
```



| □ | Project | Host | Name | Image Name | IP Address | Size | Status | Task | Power State | Time since created | Actions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| □ | tenant310 | compute-2.localdomain | tenant310_160_inst4 | rhel7 | 10.20.160.6<br>Floating IPs:<br>10.23.160.69 | m1.small | Active | None | Running | 2 hours, 3 minutes | Edit Instance ▾ |
| □ | tenant310 | compute-3.localdomain | tenant310_160_inst3 | rhel7 | 10.20.160.5<br>Floating IPs:<br>10.23.160.68 | m1.small | Active | None | Running | 2 hours, 3 minutes | Edit Instance ▾ |
| □ | tenant310 | compute-0.localdomain | tenant310_110_inst2 | rhel7 | 10.20.110.6<br>Floating IPs:<br>10.23.160.67 | m1.small | Active | None | Running | 2 hours, 4 minutes | Edit Instance ▾ |
| □ | tenant310 | compute-1.localdomain | tenant310_110_inst1 | rhel7 | 10.20.110.5<br>Floating IPs:<br>10.23.160.66 | m1.small | Active | None | Running | 2 hours, 4 minutes | Edit Instance ▾ |
| □ | tenant309 | compute-2.localdomain | tenant309_159_inst4 | rhel7 | 10.20.159.6<br>Floating IPs:<br>10.23.160.64 | m1.small | Active | None | Running | 2 hours, 5 minutes | Edit Instance ▾ |
| □ | tenant309 | compute-3.localdomain | tenant309_159_inst3 | rhel7 | 10.20.159.5<br>Floating IPs:<br>10.23.160.63 | m1.small | Active | None | Running | 2 hours, 5 minutes | Edit Instance ▾ |
| □ | tenant309 | compute-0.localdomain | tenant309_109_inst2 | rhel7 | 10.20.109.6<br>Floating IPs:<br>10.23.160.62 | m1.small | Active | None | Running | 2 hours, 5 minutes | Edit Instance ▾ |



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| □ | tenant301 | tenant301-151 | **tenant301-151-subnet** 10.20.151.0/24 | 3 | No | Active | UP | Edit Network ▾ |
| □ | tenant301 | tenant301-101 | **tenant301-101-subnet** 10.20.101.0/24 | 3 | No | Active | UP | Edit Network ▾ |
| □ | tenant302 | tenant302-102 | **tenant302-102-subnet** 10.20.102.0/24 | 3 | No | Active | UP | Edit Network ▾ |
| □ | tenant302 | tenant302-152 | **tenant302-152-subnet** 10.20.152.0/24 | 3 | No | Active | UP | Edit Network ▾ |

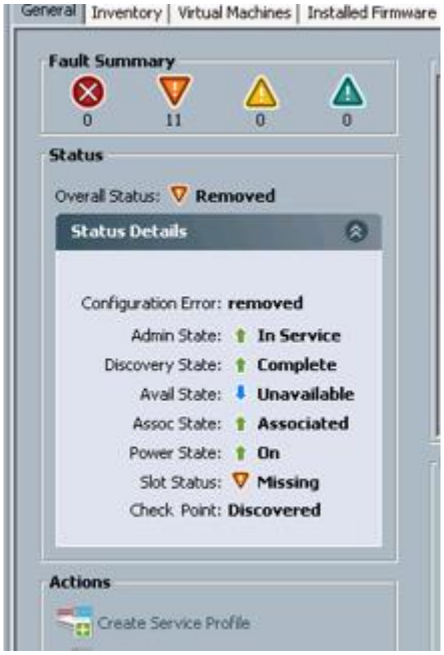
## Insert the New Blade into the Chassis

Insert the new spare blade into the chassis.

## Fault Injection

Identify the overcloud-controller and UCS Service Profile mapping from /etc/neutron/plugin.ini on any other controller node.

### Remove the Blade from the Chassis



Fault Summary: 0 11 0 0

Status

Overall Status: Removed

Status Details

Configuration Error: removed
Admin State: In Service
Discovery State: Complete
Avail State: Unavailable
Assoc State: Associated
Power State: On
Slot Status: Missing
Check Point: Discovered

```
PCSD Status:
  controller-0: Online

  controller-1: Offline
  controller-2: Online

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

## Health Checks

**Nova and Ironic Status**

After few minutes ironic makes the node as not available.

**Tenant and VMs Status**

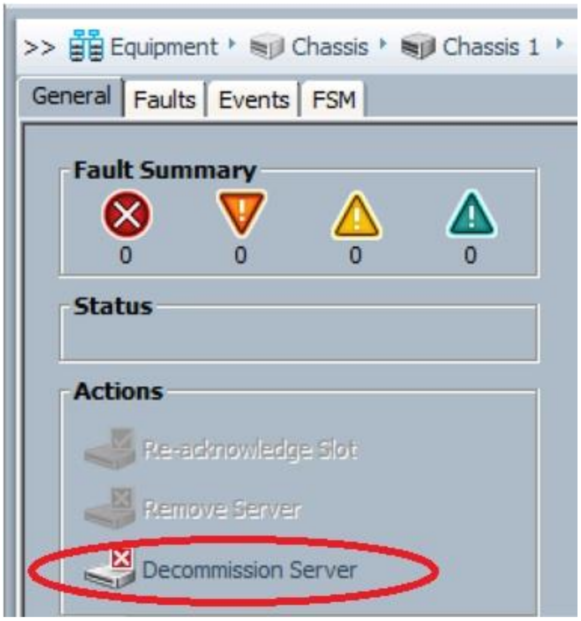
Make sure that you can login to dashboard, create new VMs and North-South and East-West traffic between VMs is uninterrupted. You may observe slowness in creating the VMs.

**Dashboard**

Dashboard starts working fine after few minutes. VNC handshake observed to be slow, but lets you in.

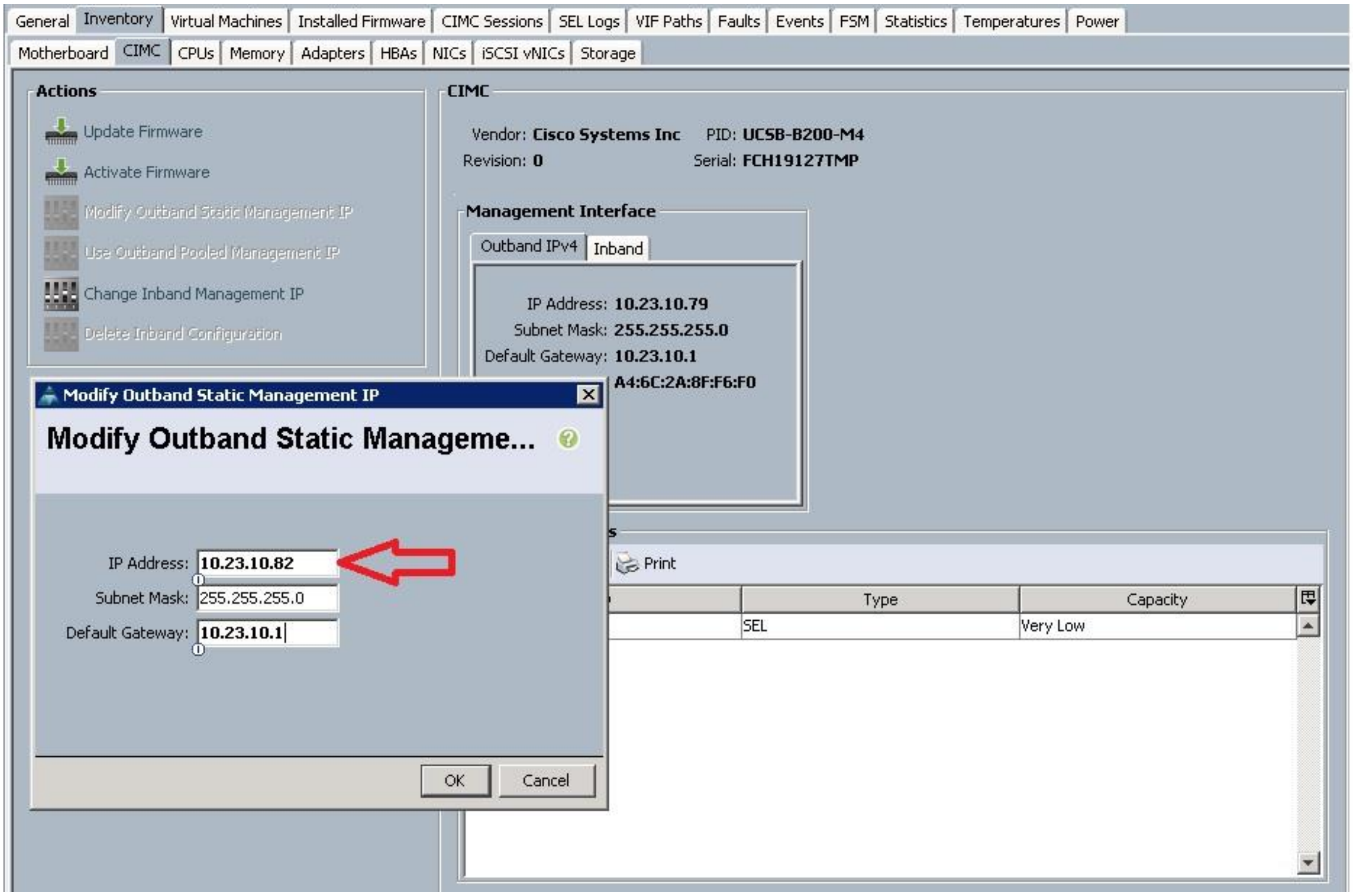
## Remove Failed Blade from Inventory

Go to the Equipment tab and decommission the blade from the chassis.



## Change IPMI Address

Change the IPMI address of the newly inserted blade to that of the old blade. As the failed blade has been decommissioned from Cisco UCS, the IPMI address is released into the free pool. Use this address back into the new blade. We are assuring that the IPMI address stored in OpenStack for the controller is back in sync with Cisco UCS.



CIMC

Vendor: Cisco Systems Inc PID: UCSB-B200-M4
Revision: 0 Serial: FCH19127TMP

Management Interface

IP Address: 10.23.10.79
Subnet Mask: 255.255.255.0
Default Gateway: 10.23.10.1
A4:6C:2A:8F:F6:F0

Modify Outband Static Management IP

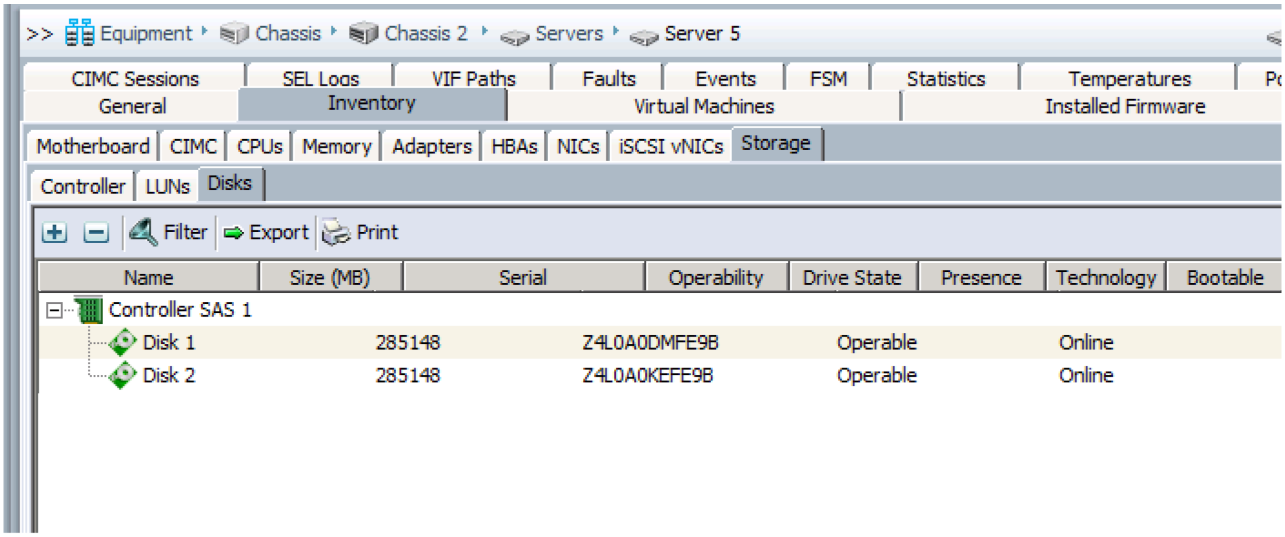
Modify Outband Static Manageme...

IP Address: 10.23.10.82
Subnet Mask: 255.255.255.0
Default Gateway: 10.23.10.1

| Type | Capacity |
|---|---|
| SEL | Very Low |

## Insert Old Disks

Remove the boot disks from the failed blade and insert them into the new blade. Make sure that the old disks appear in the server inventory.



| Name | Size (MB) | Serial | Operability | Drive State | Presence | Technology | Bootable |
|---|---|---|---|---|---|---|---|
| Controller SAS 1 | | | | | | | |
| Disk 1 | 285148 | Z4L0A0DMFE9B | Operable | | Online | | |
| Disk 2 | 285148 | Z4L0A0KEFE9B | Operable | | Online | | |

## Associate Service Profile

Associate the existing service profile that was disassociated earlier from the failed blade to the new or replacement blade.

Make sure that Config is in progress and monitor the status in FSM tab of the server.

## Reboot the Server

The association should boot up the server based on the desired power state, otherwise boot it up. It should show you the login prompt as below.

## Post Replacement Steps

Even though the server is up and running, you may need the following steps to let the server join back the cluster again.

The server is in maintenance mode

Use ironic commands to bring them to normal state.

```
ironic node-set-power-state <UUID> on
ironic node-set-maintenance <UUID> false
```

## Health Checks Post Replacement

Log into any one of the controllers and check the status from pacemaker. If any services observed to be down, you may restart them with pcs resource cleanup.

# Frequently Asked Questions

## Cisco Unified Computing System

- Can we use IO Modules 2208 instead of IOM 2204 as shown in the topology diagram?

  Yes, both IOM 2204 and IOM 2208 are supported.

- When should we use C240M4S and when C240M4L for storage servers?

  This boils down as a design question and depends on the requirement. The C240M4 SFF, small form factor offers more spindles and hence higher IOPS with reasonable bandwidth capacity. The C240M4 LFF, the large form factor has a higher storage capacity but may not be as good as SFF on total IOPS per node. Validation has been done and the performance metrics provided that should help you choose the right hardware.

- Can I use different hardware like Cisco M3 blades and different VIC adapters in the solution?

  Cisco hardware higher than the version in the BOM are supported. While lower versions may still work, they have not been validated.

- How many chassis or blades and servers can I scale horizontally?

  Usually the limits are imposed by the FI ports and scalability of the controllers from OpenStack. Earlier validations did not reveal any issues with 3 fully loaded blade chassis and 12 Storage nodes

- How can I connect my OpenStack to an existing Ceph installation?

  Please refer Red Hat documentation for external ceph here.

## OpenStack

- My network topology differs from what mentioned in this document. What changes I need to do to the configuration?

  The network topology verified in the configuration is included in the Appendix. There were limited IP's and the floating network was used. It is not necessary to have the same settings. However you may have to change yaml files accordingly and tweaks may be necessary. Please refer Red Hat documentation on how to accommodate these changes in the template files.

- Why version lock directives have been delegated in this document?

  OpenStack is continuously updated and changes in binaries and configurations go neck and neck. The purpose of providing lock file is to lock and provide binaries as close as possible to the validated design. This ensures consistency with minimal deviations from the validate design and adoption of configuration files like the yaml files. You can always install a higher version than mentioned but the specifics needed on configuration files may vary and/or some of the validations that were done in this document may have to be redone to avoid any regressions.

# Troubleshooting

This section details some troubleshooting tips for Red Hat OpenStack Platform 8 on Cisco UCS servers. Troubleshooting in OpenStack is exhaustive and this section is limited.

## Cisco Unified Computing System

- The provisioning interface should be enabled as native across all the blades and rack servers for successful introspection and Overcloud deploy.
- The native flag for external network shouldn't be enabled on Overcloud nodes as observed on the test bed.
- Specify the PCI order for network interfaces. This ensures that they are enumerated in the same way as specified in the templates.
- In case of using updating templates make sure that the service profiles are unbound from the service profile template for successful operation of UCS Manager Plugin.
- Before applying service profiles, you should make sure that all the disks are in 'Unconfigured Good' status. The storage profile, that is attached to these service profiles will then successfully get applied and then will make the boot lun in operable mode.

## Undercloud Install

Undercloud install observed to be straight forward and very few issues observed. Mostly these were human mistakes like typos in the configuration file.

- Make sure that the server is registered with Red Hat Content Delivery Network for downloading the packages. In case the server is behind proxy, update /etc/rhsm/rhsm.conf file with appropriate proxy server values.
- Double check the entries in Undercloud configuration file. Provide enough room for discovery_iprange and dhcp start/end, also considering the future expansion or upscaling of the servers later. Most of these parameters are explained in the sample file provided in /usr/share.
- Leave the value of undercloud_debug=true as default to check for failures. The log file install-undercloud.log is created as part of Undercloud install in /home/stack/.instack. This will be handy to browse through on issues encountered during the install.
- A repeat of Undercloud install preferably has to be done in a cleaner environment after reinstalling the base operating system.

## Introspection

Failure of introspecting the nodes can by many. Make sure that you have verified all the post undercloud and pre-introspection steps mentioned earlier in this document.

- A correct value of ipmi and mac addresses and powering on/off with ipmitool as mentioned earlier in this document should isolate the issues. Check with ironic node-list and ironic node-show to ensure that the registered values are correct.
- The boot luns configured in UCS through storage profile should be in available state before starting introspection. The size of the lun specified in the instack.json file should be equal or less that the size of the lun seen in UCS.
- The best way to debug introspection failures is to open KVM console on the server and check for issues.
- In case system takes you to the shell prompt and dumps /run/initramfs/sosreport.txt provides some insight as well.

- dnsmasq is the dhcp process that pxe uses to discover. Within the provisioning subnet configured you should have only one dhcp process or this dnsmasq process running on the Undercloud node. Any overlap will cause discovery failures.

- Running ‘sudo –u journalctl -u openstack-ironic-inspector -u openstack-ironic-inspector-dnsmasq will show issues encountered by discovered and dnsmasq.

- Monitoring introspection with ‘openstack bare metal introspection bulk status’ will show if any few servers have failed.

- At times if the status of the node(s) becomes available, you may have to update the status to manageable with ironic API before running introspection.

- The default value of introspection is 60 minutes.  This may have to be changed as mentioned earlier in case introspection is taking longer time.

## Running Introspection on Failed Nodes

At times it may not be feasible to do bulk introspection of all the nodes because of say lun issue on one single node, in particular if you have large number of nodes in the cloud.

```
ironic node-delete <uuid>
Create a json file for the failed node.
openstack baremetal import --json ~/add-node.json
openstack baremetal configure boot
ironic node-list
ironic node-set-maintenance <uuid> true
openstack baremetal introspection start <uuid>
openstack baremetal introspection status <uuid>
ironic node-set-maintenance <uuid> false
```

# Overcloud Install

Debugging Overcloud failures sometimes is a daunting task. The issues could be as simple as passing incorrect parameters to Overcloud deploy while some could be bugs as well. Here is an attempt to narrow down the problems. It is difficult to cover all the failure scenarios here. Few of them found out on the configuration are mentioned here. The best place is to debug from here and then move forward with Red Hat and OpenStack documentation.

- Check for the flavors pre-defined and verify that they match correctly. Incorrect flavors and/or the number of nodes passed may error with insufficient number of nodes while running Overcloud deploy command. Run instack-ironic-deployment --show-profile to confirm.

- Before running Overcloud deployment, run OpenStack Overcloud profiles list to confirm the nodes available and attached to the respective profiles.

- Make sure that you have ntp server configured and check with ntpdate -d -y <ntp server> to check the drift. Preferably should be less than 20ms for ceph monitors.

- Run in debug mode to capture the errors while running Overcloud deploy.

## Debug Network Issues

- Overcloud image has been customized with root passwords. This will allow us to login to the node directly through KVM console in case of failures even if heat-admin user is still not setup.

- journalctl -u os-collect-config | egrep -i "error|trace|fail" should shed some light around any errors or failures happened during Overcloud deploy.

- Incorrect configuration of yaml files may result in network configuration issues. Run journalctl as above to start with. Validate the yaml files with online yaml parsers.
- Run ifconfig and ovs-vsctl show the mappings.
- cd /etc/os-net-config. jq . config.json will spill out the actual parameters that went on to that node.
- Login from director node to the other nodes and check for the routes. There should be one static route either externally or to the Undercloud node, depending on the way masquadering was configured

### Debug Ceph Storage Issues

- Run journalctl as above and check for dmesg and /var/log/messages to reveal any failures related with partitioning and/or network.
- Ceph partitions are pre-created with wipe-disk.yaml file. Validate this with /root/wipe-disk.txt file and running cat /proc/partitions. Only the journal partitions are pre-created. The OSD partitions are created by Red Hat OpenStack Platform director.
- Checking the partitions in /proc/partitions and the existence of /var/log/ceph/*, /var/lib/ceph* and /etc/ceph/keyring and other files reveal at what stage it failed.
- The monitors should be setup before creating Ceph OSD's. Existence of /etc/ceph/* on controller nodes, followed by that in storage nodes will reveal whether monitor setup was successful or not.
- Run ceph -s to check the health and observe for how many total OSD's, how many are up etc.
- Run ceph osd tree to reveal issues with any individual OSD's.
- If you detect clock skew issues on monitors, check for ntp daemon, sync up the time on monitors running on controller nodes and restart the monitors /etc/init.d/ceph mon restart

### Debug Heat Stack Issues

The following sequence may be followed to debug heat stack create/update issues.

```
heat stack-list
heat resource-list overcloud | grep -vi complete
heat resource-list -n5 overcloud | grep -vi complete
heat resource-show overcloud Controller
heat deployment-show <deployment id obtained above>
```

Sometimes logging onto the KVM console directly will reveal the issues. As Overcloud image has been customized with root password, you should be able to login once the OS is deployed, but there are failures thereafter.

## Overcloud Post-Deployment Issues

Check for errors with pcs status on controller nodes. If some resources are not up or running, then this needs to be addressed first.

```
pcs resource cleanup will restart all the services.
pcs resource restart <resource name obtained from pcs status>
nova list, nova service-list and keystone endpoint-list could be handy to debug.
nova service-list and hypervisor-list or show will reveal details of the
hypervisors configured on the system. If any nodes are missing than expected,
that may have to be addressed too.
```

## Nexus Plugin Checks

Validate entries in /etc/neutron/plugin.ini on all the controllers.

Any VMs created should have VLAN entries globally in the switch and also in both the port-channels and both the switches. Any missing entry will raise an alarm here.

**Nexus Global VLANs**

```
301  VLAN0301                         active    Po17, Po18, Eth1/17, Eth1/18
307  VLAN0307                         active    Po17, Po18, Eth1/17, Eth1/18
312  VLAN0312                         active    Po17, Po18, Eth1/17, Eth1/18
318  VLAN0318                         active    Po17, Po18, Eth1/17, Eth1/18
321  VLAN0321                         active    Po17, Po18, Eth1/17, Eth1/18
```

**Nexus Port Channel VLANs**

```
show running-config interface port-channel 17-18
interface port-channel18
  description OSP8-FAB-B
  switchport mode trunk
  switchport trunk allowed vlan 1,10,100,110,120,150,160,215,252-253
  switchport trunk allowed vlan add 260,265,274,284,293,301,307,312
  switchport trunk allowed vlan add 318,321-322,347,362
  spanning-tree port type edge trunk
  mtu 9216
  vpc 18
....................................
....................................
```

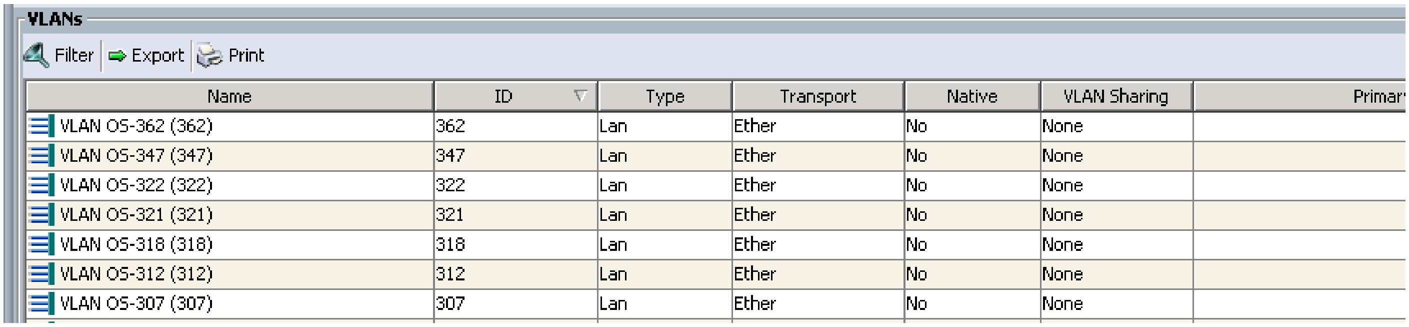
The above output is truncated for readability purposes.

## Cisco UCS Manager Plugin Checks

Validate entries in /etc/neutron/plugin.ini on all the controllers.

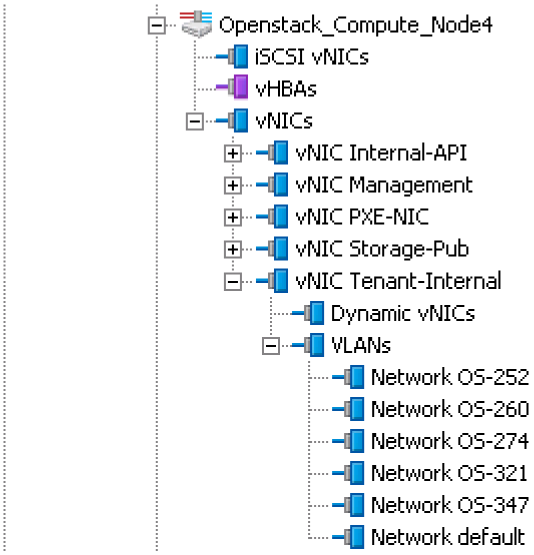
**Cisco UCS Manager Global VLANs**

Log into UCS Manager and check for the VLANs both globally and on the hypervisor where the VM(s) is provisioned. Check the host names from CLI or horizon.



| Name | ID | Type | Transport | Native | VLAN Sharing | Primar |
|---|---|---|---|---|---|---|
| VLAN OS-362 (362) | 362 | Lan | Ether | No | None | |
| VLAN OS-347 (347) | 347 | Lan | Ether | No | None | |
| VLAN OS-322 (322) | 322 | Lan | Ether | No | None | |
| VLAN OS-321 (321) | 321 | Lan | Ether | No | None | |
| VLAN OS-318 (318) | 318 | Lan | Ether | No | None | |
| VLAN OS-312 (312) | 312 | Lan | Ether | No | None | |
| VLAN OS-307 (307) | 307 | Lan | Ether | No | None | |

Hypervisor VLANs



- Openstack_Compute_Node4
  - iSCSI vNICs
  - vHBAs
  - vNICs
    - vNIC Internal-API
    - vNIC Management
    - vNIC PXE-NIC
    - vNIC Storage-Pub
    - vNIC Tenant-Internal
      - Dynamic vNICs
      - VLANs
        - Network OS-252
        - Network OS-260
        - Network OS-274
        - Network OS-321
        - Network OS-347
        - Network default

## Run Time Issues

Operational issues can be many but a brief overview of where to check in case of failures around VM Creation is provided below:

Nova commands like nova list --all-tenants, nova-manage vm list and virsh list on compute nodes could be a starting point.

Check /var/log/neutron and grep -i "error\|trace" server.log. Few may be informational and probably ignorable.

Check the following files to spot any errors

/var/log/neutron/server.log

/etc/neutron/plugin.ini

/etc/neutron/neutron.conf

/var/log/nova/*

Execute the following on controller nodes.

ip netns

ip netns exec <ns> <arguments>

# Best Practices

- While you will have business continuity you may have degraded performance during the period. Hence, it is strongly recommended to have one or two spare servers.
- Plan your networks beforehand, prepare check list of items and make they are in place before working on the actual installation. It is suggested to proof read the complete document once before attempting the installation.
- Capacity planning is another important factor to be considered for the organic growth. This not only includes the physical resources like data center space and servers but also the network subnet sizing etc.
- Follow the operational best practices like housekeeping, purging the log and archives, etc. In larger installation environments  you may have to size the /var/log separately.

# Reference Documents

- http://www.cisco.com/c/en/us/products/servers-unified-computing/index.html
- http://www.cisco.com/c/en/us/support/servers-unified-computing/unified-computing-system/products-technical-reference-list.html
- https://access.redhat.com/webassets/avalon/d/Red_Hat_OpenStack_Platform-8-Director_Installation_and_Usage-en-US/Red_Hat_OpenStack_Platform-8-Director_Installation_and_Usage-en-US.pdf
- http://docs.openstack.org/developer/tripleo-docs/troubleshooting/troubleshooting.html
- http://docs.openstack.org/developer/tripleo-docs/advanced_deployment/network_isolation.html
- http://docs.ceph.com/docs/master/rbd/rbd-openstack/

# Conclusion

This Cisco Validated Design is a joint contribution from Cisco Systems, Inc., Red Hat, Inc. and Intel Corporation. The solution combines the technologies, expertise, and contributions to the OpenStack community and experience from the field and will provide a rich experience to the end users both on installation and day to day operational aspects of OpenStack.

# Appendix

Sample Network and other yaml files can also be downloaded from https://communities.cisco.com/docs/DOC-70256

## Undercloud instackenv.json

```
[stack@osp8-director ~]$ cat instackenv.json
{
  "nodes": [
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.57",
      "mac": [
        "00:25:b5:00:00:08"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "32"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.79",
      "mac": [
        "00:25:b5:00:00:10"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "32"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.76",
      "mac": [
        "00:25:b5:00:00:18"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "32"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.69",
      "mac": [
        "00:25:b5:00:00:1e"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "40"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.67",
      "mac": [
        "00:25:b5:00:00:23"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "40"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.59",
      "mac": [
        "00:25:b5:00:00:28"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "40"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.78",
      "mac": [
        "00:25:b5:00:00:2d"
      ],
      "memory": "262144",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "40"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.66",
      "mac": [
        "00:25:b5:00:00:30"
      ],
      "memory": "131072",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "24"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.74",
      "mac": [
        "00:25:b5:00:00:36"
      ],
      "memory": "131072",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "24"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.75",
      "mac": [
        "00:25:b5:00:00:3a"
      ],
      "memory": "131072",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "24"
    },
    {
      "pm_user": "admin",
      "pm_password": "<passwd>",
      "pm_type": "pxe_ipmitool",
      "pm_addr": "10.23.10.56",
      "mac": [
        "00:25:b5:00:00:33"
      ],
      "memory": "131072",
      "disk": "250",
      "arch": "x86_64",
      "cpu": "24"
    }
  ]
}

[stack@osp8-director ~]$
```

# Overcloud Templates

## network-environment.yaml

```
resource_registry:
  OS::TripleO::NodeUserData:
    /home/stack/templates/wipe-disk.yaml
  OS::TripleO::Compute::Net::SoftwareConfig:
    /home/stack/templates/nic-configs/compute.yaml
  OS::TripleO::Controller::Net::SoftwareConfig:
    /home/stack/templates/nic-configs/controller.yaml
  OS::TripleO::CephStorage::Net::SoftwareConfig:
    /home/stack/templates/nic-configs/ceph-storage.yaml

parameter_defaults:
  # This section is where deployment-specific configuration is done
  # Customize the IP subnets to match the local environment
  InternalApiNetCidr: 10.23.100.0/24
  StorageNetCidr: 10.23.120.0/24
  StorageMgmtNetCidr: 10.23.150.0/24
  TenantNetCidr: 10.20.20.0/24
  ExternalNetCidr: 172.22.215.0/24
  # CIDR subnet mask length for provisioning network
  ControlPlaneSubnetCidr: '24'
  # Customize the IP ranges on each network to use for static IPs and VIPs
  InternalApiAllocationPools: [{'start': '10.23.100.50', 'end': '10.23.100.250'}]
  StorageAllocationPools: [{'start': '10.23.120.50', 'end': '10.23.120.250'}]
  StorageMgmtAllocationPools: [{'start': '10.23.150.50', 'end': '10.23.150.250'}]
  TenantAllocationPools: [{'start': '10.20.20.10', 'end': '10.20.20.250'}]
  ExternalAllocationPools: [{'start': '172.22.215.16', 'end': '172.22.215.20'}]
  ExternalInterfaceDefaultRoute: "172.22.215.1"
  ControlPlaneDefaultRoute: 10.23.110.26
  EC2MetadataIp: 10.23.110.26
  DnsServers: ["8.8.8.8","8.8.4.4"]
  StorageNetworkVlanID: 120
  StorageMgmtNetworkVlanID: 150
  InternalApiNetworkVlanID: 100
  ExternalNetworkVlanID: 215
  NeutronExternalNetworkBridge: "''"
```

## storage-environment.yaml

```
parameters:
  CinderEnableIscsiBackend: false
  CinderEnableRbdBackend: true
  NovaEnableRbdBackend: true
  GlanceBackend: rbd
```

## timezone.yaml

```
parameter_defaults:
  TimeZone: 'US/Pacific'
```

## management.yaml

```
heat_template_version: 2015-04-30
```

```
description: >
  Management network. System administration, SSH, DNS, NTP, etc. This network
  would usually be the default gateway for the non-controller nodes.

parameters:
  # the defaults here work for static IP assignment (IPAM) only
  ManagementNetCidr:
    default: '10.23.10.0/24'
    description: Cidr for the management network.
    type: string
  ManagementNetValueSpecs:
    default: {'provider:physical_network': 'management', 'provider:network_type':
'flat'}
    description: Value specs for the management network.
    type: string
  ManagementNetAdminStateUp:
    default: true
    description: This admin state of of the network.
    type: boolean
  ManagementNetEnableDHCP:
    default: false
    description: Whether to enable DHCP on the associated subnet.
    type: boolean
  ManagementNetShared:
    default: false
    description: Whether this network is shared across all tenants.
    type: boolean
  ManagementNetName:
    default: management
    description: The name of the management network.
    type: string
  ManagementSubnetName:
    default: management_subnet
    description: The name of the management subnet in Neutron.
    type: string
  ManagementAllocationPools:
    default: [{'start': '10.23.10.101', 'end': '10.23.10.150'}]
    description: Ip allocation pool range for the management network.
    type: json

resources:
  ManagementNetwork:
    type: OS::Neutron::Net
    properties:
      admin_state_up: {get_param: ManagementNetAdminStateUp}
      name: {get_param: ManagementNetName}
      shared: {get_param: ManagementNetShared}
      value_specs: {get_param: ManagementNetValueSpecs}

  ManagementSubnet:
    type: OS::Neutron::Subnet
    properties:
      cidr: {get_param: ManagementNetCidr}
      enable_dhcp: {get_param: ManagementNetEnableDHCP}
      name: {get_param: ManagementSubnetName}
      network: {get_resource: ManagementNetwork}
      allocation_pools: {get_param: ManagementAllocationPools}

outputs:
  OS::stack_id:
    description: Neutron management network
    value: {get_resource: ManagementNetwork}
```

## controller.yaml

```
heat_template_version: 2015-04-30

description: >
  Software Config to drive os-net-config to configure VLANs for the
  controller role.

parameters:
  ControlPlaneIp:
    default: ''
    description: IP address/subnet on the ctlplane network
    type: string
  ExternalIpSubnet:
    default: ''
    description: IP address/subnet on the external network
    type: string
  InternalApiIpSubnet:
    default: ''
    description: IP address/subnet on the internal API network
    type: string
  TenantIpSubnet:
    default: ''
    description: IP address/subnet on the tenant network
    type: string
  StorageIpSubnet:
    default: ''
    description: IP address/subnet on the storage network
    type: string
  StorageMgmtIpSubnet:
    default: ''
    description: IP address/subnet on the storage mgmt network
    type: string
  ManagementIpSubnet: # Only populated when including environments/network-
management.yaml
    default: ''
    description: IP address/subnet on the management network
    type: string
  ExternalNetworkVlanID:
    default: 215
    description: Vlan ID for the external network traffic.
    type: number
  InternalApiNetworkVlanID:
    default: 100
    description: Vlan ID for the internal_api network traffic.
    type: number
  StorageNetworkVlanID:
    default: 120
    description: Vlan ID for the storage network traffic.
    type: number
  StorageMgmtNetworkVlanID:
    default: 150
    description: Vlan ID for the storage mgmt network traffic.
    type: number
  ManagementNetworkVlanID:
    default: 10
    description: Vlan ID for the management network traffic.
    type: number
  ExternalInterfaceDefaultRoute:
    default: '173.26.215.1'
    description: default route for the external network
    type: string
  ControlPlaneSubnetCidr: # Override this via parameter_defaults
    default: '24'
    description: The subnet CIDR of the control plane network.
    type: string
  DnsServers: # Override this via parameter_defaults
    default: ['8.8.8.8','8.8.4.4']
    description: A list of DNS servers (2 max for some implementations) that will be
added to resolv.conf.
    type: comma_delimited_list
  EC2MetadataIp: # Override this via parameter_defaults
    description: The IP address of the EC2 metadata server.
    type: string


resources:
  OsNetConfigImpl:
    type: OS::Heat::StructuredConfig
    properties:
      group: os-apply-config
      config:
        os_net_config:
          network_config:
             -
              type: interface
              name: nic1
              use_dhcp: false
              dns_servers: {get_param: DnsServers}
              addresses:
                -
                  ip_netmask:
                    list_join:
                      - '/'
                      - - {get_param: ControlPlaneIp}
                        - {get_param: ControlPlaneSubnetCidr}
              routes:
                 -
                  ip_netmask: 169.254.169.254/32
                  next_hop: {get_param: EC2MetadataIp}
             -
              type: ovs_bridge
              name: {get_input: bridge_name}
              use_dhcp: false
              members:
                -
                 type: interface
                 name: nic4
                 use_dhcp: false
                 primary: true
                -
                 type: vlan
                 vlan_id: {get_param: ExternalNetworkVlanID}
                 addresses:
```

```
       -
        ip_netmask: {get_param: ExternalIpSubnet}
     routes:
       -
        ip_netmask: 0.0.0.0/0
        next_hop: {get_param: ExternalInterfaceDefaultRoute}
-
  type: ovs_bridge
  name: br-intapi
  use_dhcp: false
  members:
    -
      type: interface
      name: nic3
      use_dhcp: false
      use_dhcp: false
      primary: true
    -
      type: vlan
      vlan_id: {get_param: InternalApiNetworkVlanID}
      addresses:
      -
        ip_netmask: {get_param: InternalApiIpSubnet}
-
  type: ovs_bridge
  name: br-storage-pub
  use_dhcp: false
  mtu: 9000
  members:
    -
      type: interface
      name: nic5
      use_dhcp: false
      mtu: 9000
      primary: true
    -
      type: vlan
      mtu: 9000
      vlan_id: {get_param: StorageNetworkVlanID}
      addresses:
      -
        ip_netmask: {get_param: StorageIpSubnet}
-
  type: ovs_bridge
  name: br-storage-clus
  use_dhcp: false
  mtu: 9000
  members:
    -
      type: interface
      name: nic6
      use_dhcp: false
      mtu: 9000
      primary: true
    -
      type: vlan
      mtu: 9000
      vlan_id: {get_param: StorageMgmtNetworkVlanID}
      addresses:
      -
```

```
                    ip_netmask: {get_param: StorageMgmtIpSubnet}
            -
              type: ovs_bridge
              name: br-tenant
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic2
                  use_dhcp: false
                  # force the MAC address of the bridge to this interface
                  primary: true
            -
              type: ovs_bridge
              name: br-floating
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic7
                  use_dhcp: false
                  primary: true
            -
              type: ovs_bridge
              name: br-mgmt
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic8
                  use_dhcp: false
                  primary: true
                -
                  type: vlan
                  vlan_id: {get_param: ManagementNetworkVlanID}
                  addresses:
                  -
                    ip_netmask: {get_param: ManagementIpSubnet}

outputs:
  OS::stack_id:
    description: The OsNetConfigImpl resource.
    value: {get_resource: OsNetConfigImpl}
```

## compute.yaml

```
heat_template_version: 2015-04-30

description: >
  Software Config to drive os-net-config with 2 bonded nics on a bridge
  with a VLANs attached for the compute role.

parameters:
  ControlPlaneIp:
    default: ''
    description: IP address/subnet on the ctlplane network
    type: string
  ExternalIpSubnet:
    default: ''
    description: IP address/subnet on the external network
    type: string
  InternalApiIpSubnet:
    default: ''
    description: IP address/subnet on the internal API network
    type: string
  TenantIpSubnet:
    default: ''
    description: IP address/subnet on the tenant network
    type: string
  StorageIpSubnet:
    default: ''
    description: IP address/subnet on the storage network
    type: string
  StorageMgmtIpSubnet:
    default: ''
    description: IP address/subnet on the storage mgmt network
    type: string
  ManagementIpSubnet: # Only populated when including environments/network-
management.yaml
    default: ''
    description: IP address/subnet on the management network
    type: string
  ExternalNetworkVlanID:
    default: 215
    description: Vlan ID for the external network traffic.
    type: number
  InternalApiNetworkVlanID:
    default: 100
    description: Vlan ID for the internal_api network traffic.
    type: number
  StorageNetworkVlanID:
    default: 120
    description: Vlan ID for the storage network traffic.
    type: number
  StorageMgmtNetworkVlanID:
    default: 150
    description: Vlan ID for the storage mgmt network traffic.
    type: number
  ManagementNetworkVlanID:
    default: 10
    description: Vlan ID for the management network traffic.
    type: number
  ExternalInterfaceDefaultRoute:
    default: '173.26.215.1'
    description: default route for the external network
    type: string
  ControlPlaneSubnetCidr: # Override this via parameter_defaults
    default: '24'
    description: The subnet CIDR of the control plane network.
    type: string
  ControlPlaneDefaultRoute:
    default: '10.23.100.26'
    description: default route for the external network
    type: string
  DnsServers: # Override this via parameter_defaults
    default: ['8.8.8.8','8.8.4.4']
    description: A list of DNS servers (2 max for some implementations) that
will be added to resolv.conf.
    type: comma_delimited_list
  EC2MetadataIp: # Override this via parameter_defaults
    description: The IP address of the EC2 metadata server.
    type: string

resources:
  OsNetConfigImpl:
    type: OS::Heat::StructuredConfig
    properties:
      group: os-apply-config
      config:
        os_net_config:
          network_config:
             -
              type: interface
              name: nic1
              use_dhcp: false
              dns_servers: {get_param: DnsServers}
              addresses:
                -
                  ip_netmask:
                    list_join:
                      - '/'
                      - - {get_param: ControlPlaneIp}
                        - {get_param: ControlPlaneSubnetCidr}
              routes:
                -
                  ip_netmask: 169.254.169.254/32
                  next_hop: {get_param: EC2MetadataIp}
                -
                  default: true
                  next_hop: {get_param: ControlPlaneDefaultRoute}
             -
              type: ovs_bridge
              name: br-storage-pub
              use_dhcp: false
              mtu: 9000
              members:
                -
                  type: interface
                  name: nic4
                  use_dhcp: false
                  mtu: 9000
                  primary: true
                 -
                  type: vlan
                  mtu: 9000
                  vlan_id: {get_param: StorageNetworkVlanID}
                  addresses:
                 -
                    ip_netmask: {get_param: StorageIpSubnet}
              -
              type: ovs_bridge
              name: br-mgmt
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic5
                  use_dhcp: false
                  primary: true
                -
                  type: vlan
                  vlan_id: {get_param: ManagementNetworkVlanID}
                  addresses:
                   -
                    ip_netmask: {get_param: ManagementIpSubnet}
             -
              type: ovs_bridge
              name: br-intapi
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic3
                  use_dhcp: false
                  primary: true
                -
                  type: vlan
                  vlan_id: {get_param: InternalApiNetworkVlanID}
                  addresses:
                  -
                    ip_netmask: {get_param: InternalApiIpSubnet}
             -
              type: ovs_bridge
              name: br-tenant
              use_dhcp: false
              members:
                -
                  type: interface
                  name: nic2
                  use_dhcp: false
                  primary: true

outputs:
  OS::stack_id:

outputs:
  OS::stack_id:
    description: The OsNetConfigImpl resource.
    value: {get_resource: OsNetConfigImpl}
```

## ceph-storage.yaml

```
heat_template_version: 2015-04-30

description: >
  Software Config to drive os-net-config with 2 bonded nics on a bridge
  with a VLANs attached for the ceph storage role.

parameters:
  ControlPlaneIp:
    default: ''
    description: IP address/subnet on the ctlplane network
    type: string
  ExternalIpSubnet:
    default: ''
    description: IP address/subnet on the external network
    type: string
  InternalApiIpSubnet:
    default: ''
    description: IP address/subnet on the internal API network
    type: string
  TenantIpSubnet:
    default: ''
    description: IP address/subnet on the tenant network
    type: string
  StorageIpSubnet:
    default: ''
    description: IP address/subnet on the storage network
    type: string
  StorageMgmtIpSubnet:
    default: ''
    description: IP address/subnet on the storage mgmt network
    type: string
  ManagementIpSubnet: # Only populated when including environments/network-
management.yaml
    default: ''
    description: IP address/subnet on the management network
    type: string
  ExternalNetworkVlanID:
    default: 215
    description: Vlan ID for the external network traffic.
    type: number
  InternalApiNetworkVlanID:
    default: 100
    description: Vlan ID for the internal_api network traffic.
    type: number
  StorageNetworkVlanID:
    default: 120
    description: Vlan ID for the storage network traffic.
    type: number
  StorageMgmtNetworkVlanID:
    default: 150
    description: Vlan ID for the storage mgmt network traffic.
    type: number
  ManagementNetworkVlanID:
    default: 10
    description: Vlan ID for the management network traffic.
    type: number
  ExternalInterfaceDefaultRoute:
    default: '173.26.215.1'
    description: default route for the external network
    type: string
  ControlPlaneSubnetCidr: # Override this via parameter_defaults
    default: '24'
    description: The subnet CIDR of the control plane network.
    type: string
  ControlPlaneDefaultRoute:
    default: '10.23.100.26'
    description: default route for the external network
    type: string
  DnsServers: # Override this via parameter_defaults
    default: ['8.8.8.8','8.8.4.4']
    description: A list of DNS servers (2 max for some implementations) that will be
added to resolv.conf.
    type: comma_delimited_list
  EC2MetadataIp: # Override this via parameter_defaults
    description: The IP address of the EC2 metadata server.
    type: string

resources:
  OsNetConfigImpl:
    type: OS::Heat::StructuredConfig
    properties:
      group: os-apply-config
      config:
        os_net_config:
          network_config:
            -
              type: interface
              name: nic1
              use_dhcp: false
              dns_servers: {get_param: DnsServers}
              addresses:
                -
                  ip_netmask:
                    list_join:
                      - '/'
                      - - {get_param: ControlPlaneIp}
                        - {get_param: ControlPlaneSubnetCidr}
              routes:
                -
                  ip_netmask: 169.254.169.254/32
                  next_hop: {get_param: EC2MetadataIp}
                -
                  default: true
                  next_hop: {get_param: ControlPlaneDefaultRoute}
              -
              type: ovs_bridge
              name: br-storage-pub
              use_dhcp: false
              mtu: 9000
              members:
                 -
                  type: interface
                  name: nic2
                  use_dhcp: false
                  mtu: 9000
                  primary: true
                 -
                  type: vlan
                  mtu: 9000
                  vlan_id: {get_param: StorageNetworkVlanID}
                  addresses:
                 -
                    ip_netmask: {get_param: StorageIpSubnet}

             -
              type: ovs_bridge
              name: br-storage-clus
              use_dhcp: false
              mtu: 9000
              members:
                 -
                  type: interface
                  name: nic3
                  use_dhcp: false
                  mtu: 9000
                  primary: true
                 -
                  type: vlan
                  mtu: 9000
                  vlan_id: {get_param: StorageMgmtNetworkVlanID}
                  addresses:
                 -
                    ip_netmask: {get_param: StorageMgmtIpSubnet}
outputs:
  OS::stack_id:
    description: The OsNetConfigImpl resource.
    value: {get_resource: OsNetConfigImpl}
```

## ceph.yaml (C240M4L)

```
ceph::profile::params::osd_journal_size: 20000
ceph::profile::params::osd_pool_default_pg_num: 128
ceph::profile::params::osd_pool_default_pgp_num: 128
ceph::profile::params::osd_pool_default_size: 3
ceph::profile::params::osd_pool_default_min_size: 1
ceph::profile::params::manage_repo: false
ceph::profile::params::authentication_type: cephx
ceph::profile::params::osds:
    '/dev/sdd':
        journal: '/dev/sdb1'
    '/dev/sde':
        journal: '/dev/sdb2'
    '/dev/sdf':
        journal: '/dev/sdb3'
    '/dev/sdg':
        journal: '/dev/sdb4'
    '/dev/sdh':
        journal: '/dev/sdc1'
    '/dev/sdi':
        journal: '/dev/sdc2'
    '/dev/sdj':
        journal: '/dev/sdc3'
    '/dev/sdk':
        journal: '/dev/sdc4'

ceph_classes: []

ceph_osd_selinux_permissive: true
```

## ceph.yaml (C240M4S)

```
ceph::profile::params::osd_journal_size: 20000
ceph::profile::params::osd_pool_default_pg_num: 128
ceph::profile::params::osd_pool_default_pgp_num: 128
ceph::profile::params::osd_pool_default_size: 3
ceph::profile::params::osd_pool_default_min_size: 1
ceph::profile::params::manage_repo: false
ceph::profile::params::authentication_type: cephx
ceph::profile::params::osds:
    '/dev/sdf':
        journal: '/dev/sdb1'
    '/dev/sdg':
        journal: '/dev/sdb2'
    '/dev/sdh':
        journal: '/dev/sdb3'
    '/dev/sdi':
        journal: '/dev/sdb4'
    '/dev/sdj':
        journal: '/dev/sdb5'
    '/dev/sdk':
        journal: '/dev/sdc1'
    '/dev/sdl':
        journal: '/dev/sdc2'
    '/dev/sdm':
        journal: '/dev/sdc3'
    '/dev/sdn':
        journal: '/dev/sdc4'
    '/dev/sdo':
        journal: '/dev/sdc5'
    '/dev/sdp':
        journal: '/dev/sdd1'
    '/dev/sdq':
        journal: '/dev/sdd2'
    '/dev/sdr':
        journal: '/dev/sdd3'
    '/dev/sds':
        journal: '/dev/sdd4'
    '/dev/sdt':
        journal: '/dev/sde1'
    '/dev/sdu':
        journal: '/dev/sde2'
    '/dev/sdv':
        journal: '/dev/sde3'
    '/dev/sdw':
        journal: '/dev/sde4'

ceph_classes: []

ceph_osd_selinux_permissive: true
```

## cisco-plugins.yaml

```
resource_registry:
  OS::TripleO::AllNodesExtraConfig: /usr/share/openstack-tripleo-heat-
templates/puppet/extraconfig/all_nodes/neutron-ml2-cisco-nexus-ucsm.yaml

parameter_defaults:
  NetworkUCSMIp: '10.23.10.5'
  NetworkUCSMUsername: 'admin'
  NetworkUCSMPassword: <passwd>
  NetworkUCSMHostList: 00:25:b5:00:00:1a:org-root/org-osp8/ls-
Openstack_Compute_Node1,00:25:b5:00:00:1f:org-root/org-osp8/ls-
Openstack_Compute_Node2,00:25:b5:00:00:24:org-root/org-osp8/ls-
Openstack_Compute_Node3,00:25:b5:00:00:29:org-root/org-osp8/ls-
Openstack_Compute_Node4,00:25:b5:00:00:02:org-root/org-osp8/ls-
Openstack_Controller_Node1,00:25:b5:00:00:0a:org-root/org-osp8/ls-
Openstack_Controller_Node2,00:25:b5:00:00:12:org-root/org-osp8/ls-
Openstack_Controller_Node3
  NetworkNexusConfig: {
    "UCSO-N9K-FAB-A": {
        "ip_address": "10.23.10.3",
        "nve_src_intf": 0,
        "password": "<passwd>",
        "physnet": "",
        "servers": {
            "00:25:b5:00:00:1a": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:1f": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:24": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:29": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:02": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:0a": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:12": {
                "ports": "port-channel:17,port-channel:18"
            },
        },
        "ssh_port": 22,
        "username": "admin"
     },
    "UCSO-N9K-FAB-B": {
        "ip_address": "10.23.10.4",
        "nve_src_intf": 0,
        "password": "<passwd>",
        "physnet": "",
        "servers": {
            "00:25:b5:00:00:1a": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:1f": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:24": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:29": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:02": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:0a": {
                "ports": "port-channel:17,port-channel:18"
            },
            "00:25:b5:00:00:12": {
                "ports": "port-channel:17,port-channel:18"
            },
        },
        "ssh_port": 22,
        "username": "admin"
     }
}
  NetworkNexusManagedPhysicalNetwork: physnet-tenant
  NetworkNexusVlanNamePrefix: 'q-'
  NetworkNexusSviRoundRobin: 'false'
  NetworkNexusProviderVlanNamePrefix: 'p-'
  NetworkNexusPersistentSwitchConfig: 'false'
  NetworkNexusSwitchHeartbeatTime: 30
  NetworkNexusSwitchReplayCount: 1000
  NetworkNexusProviderVlanAutoCreate: 'true'
  NetworkNexusProviderVlanAutoTrunk: 'true'
  NetworkNexusVxlanGlobalConfig: 'false'
  NetworkNexusHostKeyChecks: 'false'
  EnablePackageInstall: false
  NeutronMechanismDrivers: 'openvswitch,cisco_nexus,cisco_ucsm'
  NeutronServicePlugins: 'router'
  NeutronTypeDrivers: 'vlan'
  NeutronCorePlugin: 'ml2'
  NeutronNetworkVLANRanges: 'physnet-tenant:250:700,floating:160:160'
  NetworkNexusVxlanVniRanges: '0:0'
  NetworkNexusVxlanMcastRanges: '0.0.0.0:0.0.0.0'
parameters:
  controllerExtraConfig:
    neutron::server::api_workers: 1
    neutron::agents::metadata::metadata_workers: 1
    neutron::server::rpc_workers: 1
```

## wipe-disk.yaml (C240M4L)

```
heat_template_version: 2015-04-30
#This configuraion is only for C240M4L server. Change for C240M4S
#The first for loop zapdisks all the storage disks
#The second for loop creates aligned partitions, only for the journals. Change the val-
ues of sdb and sdc accordingly
#In case of more journal disks for C240M4S add the device names for the journals in the
second for loop.
#Do not create partitions for data disks
resources:
  userdata:
    type: OS::Heat::MultipartMime
    properties:
      parts:
      - config: {get_resource: clean_disk}

  clean_disk:
    type: OS::Heat::SoftwareConfig
    properties:
      config: |
        #!/bin/bash
        DATA_DISKS="sdd sde sdf sdg sdh sdi sdj sdk"
        JOURNAL_DISKS="sdb sdc"
        JOURNAL_SIZE=20G
        { for disk in $DATA_DISKS $JOURNAL_DISKS
        do
           sgdisk -Z /dev/$disk
           sgdisk -g /dev/$disk
        done } > /root/wipe-disk.txt
        { for disk in $JOURNAL_DISKS
        do
          export ptype1=45b0969e-9b03-4f30-b4c6-b4b80ceff106
          for i in $(seq 1 $(( ($(echo $DATA_DISKS|wc -w)+$(echo $JOURNAL_DISKS|wc -w)-
1) / $(echo $JOURNAL_DISKS|wc -w) )) )
          do
          sgdisk --new=$i::+$JOURNAL_SIZE --change-name="$i:ceph journal" --
typecode="$i:$ptype1" /dev/$disk
          done
        done } >> /root/wipe-disk.txt

outputs:
  OS::stack_id:
    value: {get_resource: userdata}
```

### wipe-disk.yaml (C240M4S)

```
heat_template_version: 2015-04-30
#This configuraion is only for C240M4L server. Change for C240M4S
#The first for loop zapdisks all the storage disks
#The second for loop creates aligned partitions, only for the journals. Change the val-
ues of sdb and sdc accordingly
#In case of more journal disks for C240M4S add the device names for the journals in the
second for loop.
#Do not create partitions for data disks
resources:
  userdata:
    type: OS::Heat::MultipartMime
    properties:
      parts:
      - config: {get_resource: clean_disk}

  clean_disk:
    type: OS::Heat::SoftwareConfig
    properties:
      config: |
        #!/bin/bash
        DATA_DISKS="sdf sdg sdh sdi sdj sdk sdl sdm sdn sdo sdp sdq sdr sds sdt sdu sdv
sdw"
        JOURNAL_DISKS="sdb sdc sdd sde"
        JOURNAL_SIZE=20G
        { for disk in $DATA_DISKS $JOURNAL_DISKS
        do
           sgdisk -Z /dev/$disk
           sgdisk -g /dev/$disk
        done } > /root/wipe-disk.txt
        { for disk in $JOURNAL_DISKS
        do
          export ptype1=45b0969e-9b03-4f30-b4c6-b4b80ceff106
          for i in $(seq 1 $(( ($(echo $DATA_DISKS|wc -w)+$(echo $JOURNAL_DISKS|wc -w)-
1) / $(echo $JOURNAL_DISKS|wc -w) )) )
          do
          sgdisk --new=$i::+$JOURNAL_SIZE --change-name="$i:ceph journal" --
typecode="$i:$ptype1" /dev/$disk
          done
        done } >> /root/wipe-disk.txt

outputs:
  OS::stack_id:
    value: {get_resource: userdata}
```

## run.sh

```
#!/bin/bash
openstack overcloud deploy --templates \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-isolation.yaml \
-e /home/stack/templates/network-environment.yaml \
-e /home/stack/templates/network-management.yaml \
-e /home/stack/templates/storage-environment.yaml \
-e /home/stack/templates/timezone.yaml \
-e /home/stack/templates/cisco-plugins.yaml \
--control-flavor control --compute-flavor compute --ceph-storage-flavor ceph-storage \
--compute-scale 4 --control-scale 3  --ceph-storage-scale 4 \
--libvirt-type kvm \
--ntp-server 171.68.38.66 \
--neutron-network-type vlan \
--neutron-bridge-mappings datacentre:br-ex,physnet-tenant:br-tenant,floating:br-
floating \
--neutron-network-vlan-ranges physnet-tenant:250:700,floating:160:160 \
--neutron-disable-tunneling --timeout 90 \
--verbose --debug --log-file overcloud_new.log
```

## create_network_router.sh

```
source /home/stack/overcloudrc

neutron net-create ext-net-160 --router:external true --provider:physical_network
floating --provider:network_type vlan --shared --provider:segmentation_id 160

neutron subnet-create --name ext-subnet-160 --enable_dhcp=False \
--allocation-pool start=10.23.160.20,end=10.23.160.249 --gateway 10.23.160.1 ext-net-
160 10.23.160.0/24

#Create public image
export IMAGE=$(ls /home/stack/images/rhel-guest-image-7.2-20160302.0.x86_64.qcow2)
echo $IMAGE
openstack image create --disk-format qcow2 --public --container-format bare --file
${IMAGE} rhel7
```

## create_vm.sh

The following script creates 1 Tenant, with 2 Networks and 4VMs for the tenant. This can be looped to created multiple tenants. Please proofread before running the script. There are few hardcodings done.

```
#!/bin/bash
export NW1=$1
export NW2=$(($1+50))
export id=$2
inst1=tenant${id}_${NW1}_inst1
inst2=tenant${id}_${NW1}_inst2
inst3=tenant${id}_${NW2}_inst3
inst4=tenant${id}_${NW2}_inst4
export TENANT_CIDR1=10.20.${NW1}.0/24
export TENANT_CIDR2=10.20.${NW2}.0/24
source /home/stack/overcloudrc
KEYSTONE_URL=$OS_AUTH_URL
if [[ ! -f keystonerc_tenant${id} ]]
then
# create tenantdemo environment
openstack user create --password tenant${id} tenant${id}
openstack project create tenant${id}
openstack role add --user tenant${id} --project tenant${id} _member_
cat > keystonerc_tenant${id} << EOF
export OS_USERNAME=tenant${id}
export OS_TENANT_NAME=tenant${id}
export OS_PASSWORD=tenant${id}
export OS_CLOUDNAME=overcloud
export OS_AUTH_URL=${KEYSTONE_URL}
EOF
fi
source keystonerc_tenant${id}
env | grep OS_
# create network
neutron net-list
neutron net-create tenant${id}-${NW1}
neutron net-create tenant${id}-${NW2}
neutron subnet-create --name tenant${id}-${NW1}-subnet tenant${id}-${NW1}
${TENANT_CIDR1}
neutron subnet-create --name tenant${id}-${NW2}-subnet tenant${id}-${NW2}
${TENANT_CIDR2}
neutron router-create tenant${id}
subID1=$(neutron subnet-list | awk "/tenant${id}-${NW1}-subnet/ {print \$2}")
neutron router-interface-add tenant${id} $subID1
subID2=$(neutron subnet-list | awk "/tenant${id}-${NW2}-subnet/ {print \$2}")
neutron router-interface-add tenant${id} $subID2

for i in $(neutron security-group-list | awk ' /default/ { print $2 } ')
do
  # add ssh and icmp to default security groups
  neutron security-group-rule-create --direction ingress --protocol icmp  $i
  #openstack security group rule create --proto icmp $i
  neutron security-group-rule-create --direction ingress --protocol tcp --
port_range_min 22 --port_range_max 22 $i
  #neutron security group rule create --proto tcp --src-port 22 --dst-ip 22 $i
  neutron security-group-show $i
  openstack security group show $i
done

openstack keypair create tenant${id}kp > tenant${id}kp.pem
chmod 600 tenant${id}kp.pem

netname1=`neutron net-list | grep tenant${id}-${NW1} | awk '{print $2}'`
openstack server create --flavor m1.small --image rhel7 \
--key-name tenant${id}kp --nic net-id=${netname1} ${inst1}
openstack server create --flavor m1.small --image rhel7 \
--key-name tenant${id}kp --nic net-id=${netname1} ${inst2}

netname2=`neutron net-list | grep tenant${id}-${NW2} | awk '{print $2}'`
openstack server create --flavor m1.small --image rhel7 \
--key-name tenant${id}kp --nic net-id=${netname2} ${inst3}
openstack server create --flavor m1.small --image rhel7 \
--key-name tenant${id}kp --nic net-id=${netname2} ${inst4}

while [[ $(openstack server  list | grep BUILD) ]]
do
  sleep 3
done
openstack server  list
source /home/stack/overcloudrc
netid=$(neutron net-list | awk "/ext-net-160/ { print \$2 }")
neutron router-gateway-set tenant${id} ${netid}

source keystonerc_tenant${id}

openstack ip floating create ext-net-160
sleep 10
float_ip=$(openstack ip floating list | grep ext-net-160 | grep None | awk
'{print $6}' | sort -u | head -1)
openstack ip floating add ${float_ip} ${inst1}
openstack ip floating create ext-net-160
sleep 10
float_ip=$(openstack ip floating list | grep ext-net-160 | grep None | awk
'{print $6}' | sort -u | head -1)
openstack ip floating add ${float_ip} ${inst2}
openstack ip floating create ext-net-160
sleep 10
float_ip=$(openstack ip floating list | grep ext-net-160 | grep None | awk
'{print $6}' | sort -u | head -1)
openstack ip floating add ${float_ip} ${inst3}
openstack ip floating create ext-net-160
sleep 10
float_ip=$(openstack ip floating list | grep ext-net-160 | grep None | awk
'{print $6}' | sort -u | head -1)
openstack ip floating add ${float_ip} ${inst4}
```

# About the Authors

**Ramakrishna Nishtala, Cisco Systems, Inc.**

Ramakrishna Nishtala is a Technical Leader in Cisco UCS and Data Center solutions group and has over 20 years of experience in IT infrastructure, Automation, Virtualization and Cloud computing. In his current role at Cisco Systems, he works on best practices, optimization and performance tuning on OpenStack and other Open Source Storage solutions on Cisco UCS platforms. Prior to this he was involved in data center migration strategies, Compute and Storage Consolidation, end to end performance optimization on databases, Application and Web Servers and solutions engineering.

**Steven Reichard, Red Hat**

Steven Reichard is a consulting engineer and manager in Red Hat's System's Design and Engineering group. This team's mission is to eliminate roadblocks to the wider adoption & ease-of-use of our product portfolio to solve ever more demanding customer/partner solutions. Most recently Steve has focused on Red Hat OpenStack Platform including enabling partner solutions based on Red Hat OpenStack Platform and Red Hat Ceph Storage.  Steve is a Red Hat Certified Engineer (RHCE) who has more than 20 years of computer industry experience.

**Dariusz Komla, Intel**

Dariusz Komła is a Data Center Engineer at Intel. He works in Reference Implementation team, with focus on Software Defined Infrastructure and cloud solutions design. Dariusz also works with Intel partners on building cloud architectures based on OpenStack. In his previous role as IaaS Architect he was responsible for designing and implementing complex platforms based on OpenStack and Ceph storage in distributed data centers. Before his role as architect he spent more than 10 years as a Data Center Administrator gaining broad experience in different Linux/Windows systems, cloud solutions and system designs.

## Acknowledgements

- Cisco Systems: Brian Bowen, Carol Bouchard, Muhammad Afzal, Sandhya Dasu, Shanthi Kumar Adloori, Timothy Swanson, Vijay Durairaj, Vishwanath Jakka
- Intel: Kamil Rogon, Patryk Wolsza
- Red Hat: Andrew Beekhof, Dan Sneddon, David Critch, Guil Barros, Marek Grac