Cisco UCS Integrated Infrastructure for Big Data with MapR

With Optional Multi-Tenancy Extension

Last Updated: May 27, 2015

Cisco Validated Design

Building Architectures to Solve Business Problems

About the Authors

Raghunath Nambiar, Distinguished Engineer, Data Center Business Group (Cisco Systems)

Raghunath Nambiar is a Distinguished Engineer at Cisco's Data Center Business Group. His current responsibilities include emerging technologies and big data strategy.

Karthik Kulkarni

Karthik Kulkarni, Technical Marketing Engineer, Data Center Solutions Group (Cisco Systems)

Karthik Kulkarni is a Technical Marketing Engineer in the Data Center Solutions Group at Cisco Systems. He is part of solution engineering team focusing on big data infrastructure and performance.

Ashwin Manjunatha

Ashwin Manjunatha, Technical Marketing Engineer, Data Center Solutions Group (Cisco Systems)

Ashwin Manjunatha works in the Data Center Solutions Group at Cisco Systems. Ashwin holds Master of Science Degree in Computer Engineering from Wright State University. He specializes in Hadoop and Distributed Computing.

James Sun

James Sun, Senior Solutions Architect (MapR Technologies)

James Sun manages the technological relationship with worldwide alliances at MapR Technologies. James has over 15 years of experience in information technology. Prior to MapR, he held several senior technical positions at technological companies such as NetApp, Yahoo and EMC. He holds a PhD. from Stanford University.

About Cisco Validated Design (CVD) Program

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.

Acknowledgment

The authors acknowledge contributions of Manankumar Trivedi, Karthik Karupasamy and Sindhu Sudhir in developing this document.

Cisco UCS Integrated Infrastructure for Big Data with MapR

Audience

This document describes the architecture and deployment procedures for MapR on a 64 Cisco UCS C240 M4 node cluster based on Cisco UCS Integrated Infrastructure for Big Data. The intended audience of this document includes, but is not limited to, sales engineers, field consultants, professional services, IT managers, partner engineering and customers who want to deploy MapR on Cisco UCS Integrated Infrastructure for Big Data.

Introduction

Big data technology and Apache Hadoop in particular, are finding use in an enormous number of applications and are being evaluated and adopted by enterprises of all kinds. As this important technology helps transform large volumes of data into actionable information, many organizations are struggling to deploy effective and reliable Hadoop infrastructure that performs and scales and is appropriate for mission–critical applications in the enterprise. Deployed as part of comprehensive data center architecture, the Cisco UCS with MapR delivers a powerful and flexible infrastructure that increases business and IT agility, reduces total cost of ownership (TCO), and delivers exceptional return on investment (ROI) at scale, while fundamentally transforming the way that organizations do business with Hadoop technology.

Cisco UCS Integrated Infrastructure for Big Data with MapR and Multi–Tenancy

The Cisco UCS solution for MapR is based on Cisco UCS Integrated Infrastructure for Big Data, a highly scalable architecture designed to meet a variety of scale-out application demands with seamless data integration and management integration capabilities built using the following components:

Corporate Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

Cisco UCS 6200 Series Fabric Interconnects

Cisco UCS 6200 Series Fabric Interconnects provide high-bandwidth, low-latency connectivity for servers, with integrated, unified management provided for all connected devices by Cisco UCS Manager. Deployed in redundant pairs, Cisco fabric interconnects offer the full active-active redundancy, performance, and exceptional scalability needed to support the large number of nodes that are typical in clusters serving big data applications. Cisco UCS Manager enables rapid and consistent server configuration using service profiles, automating ongoing system maintenance activities such as firmware updates across the entire cluster as a single operation. Cisco UCS Manager also offers advanced monitoring with options to raise alarms and send notifications about the health of the entire cluster.

Figure 1 ***Cisco UCS 6296UP 96-Port Fabric Interconnect***

Cisco UCS C-Series Rack Mount Servers

Cisco UCS C240 M4 High-Density Rack servers (Small Form Factor Disk Drive Model) are enterprise-class systems that support a wide range of computing, I/O, and storage-capacity demands in compact designs. Cisco UCS C-Series Rack-Mount Servers are based on Intel Xeon E5-2600 v3 product family and 12-Gbps SAS throughput, delivering significant performance and efficiency gains over the previous generation of servers. The servers use dual Intel Xeon processor E5-2600 v3 series CPUs and support up to 768 GB of main memory (128 or 256 GB is typical for big data applications) and a range of disk drive and SSD options. 24 Small Form Factor (SFF) disk drives are supported in performance-optimized option and 12 Large Form Factor (LFF) disk drives are supported in capacity-optimized option, along with 4 Gigabit Ethernet LAN-on-motherboard (LOM) ports. Cisco UCS virtual interface cards 1227 (VICs) designed for the M4 generation of Cisco UCS C-Series Rack Servers are optimized for high-bandwidth and low-latency cluster connectivity, with support for up to 256 virtual devices that are configured on demand through Cisco UCS Manager.

Figure 2 Cisco UCS C240 M4 Rack Server

Cisco UCS Virtual Interface Cards (VICs)

Cisco UCS Virtual Interface Cards (VICs), unique to Cisco, Cisco UCS Virtual Interface Cards incorporate next-generation converged network adapter (CNA) technology from Cisco, and offer dual 10-Gbps ports designed for use with Cisco UCS C-Series Rack-Mount Servers. Optimized for virtualized networking, these cards deliver high performance and bandwidth utilization and support up to 256 virtual devices. The Cisco UCS Virtual Interface Card (VIC) 1227 is a dual-port, Enhanced Small Form-Factor Pluggable (SFP+), 10 GigabitEthernet Ethernet and Fiber Channel over Ethernet (FCoE)-capable, PCI Express (PCIe) modular LAN on motherboard (mLOM) adapter. It is designed exclusively for the M4 generation of Cisco UCS C-Series Rack Servers and the C3160 dense storage servers.

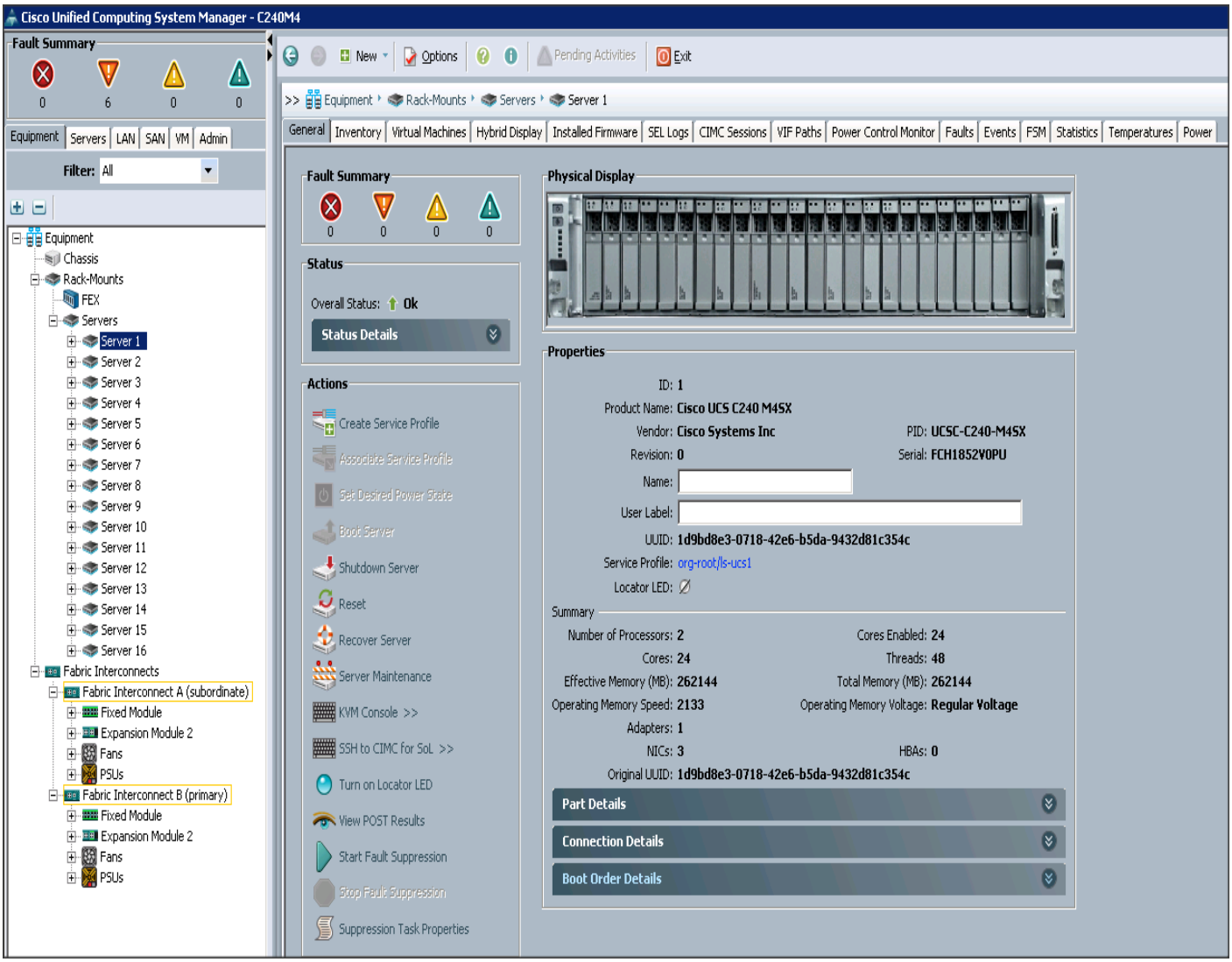
Figure 3 Cisco UCS VIC 1227

Cisco UCS Manager

Cisco UCS Manager resides within the Cisco UCS 6200 Series Fabric Interconnects. It makes the system self-aware and self-integrating, managing all of the system components as a single logical entity. Cisco UCS Manager can be accessed through an intuitive graphical user interface (GUI), a command-line interface (CLI), or an XML application-programming interface (API). Cisco UCS Manager uses service profiles to define the personality, configuration, and connectivity of all resources within Cisco UCS,

radically simplifying provisioning of resources so that the process takes minutes instead of days. This simplification allows IT departments to shift their focus from constant maintenance to strategic business initiatives.

Figure 4 Cisco UCS Manager



Cisco UCS Director Express for Big Data

Cisco UCS Director Express for Big Data provides a single-touch solution that automates deployment of Hadoop Distributions on leading Cisco UCS Integrated Infrastructure for Big Data.

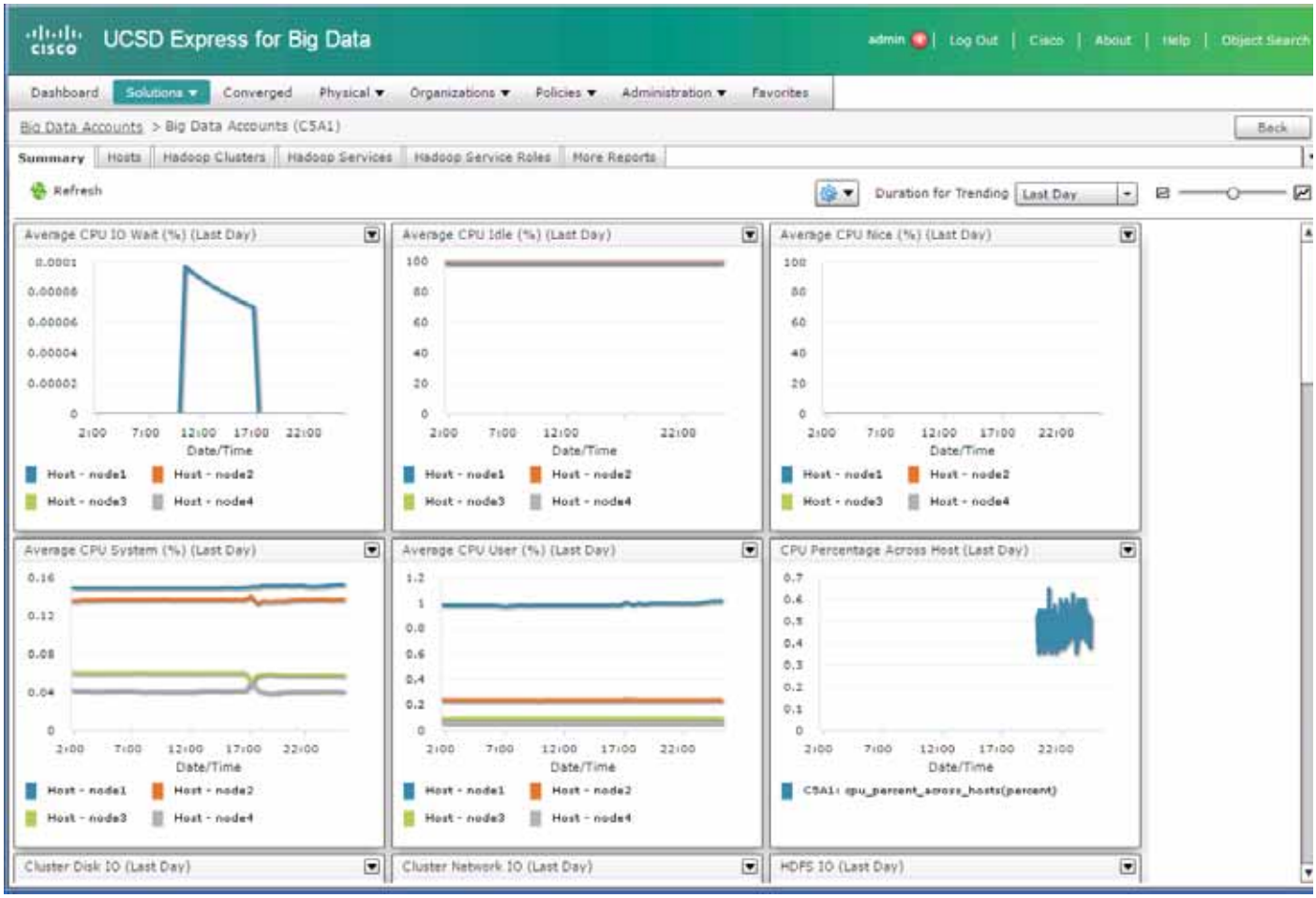
It also provides a single management pane across both physical infrastructure and Hadoop software. All elements of the infrastructure are handled automatically with little need for user input. Through this approach, configuration of physical computing, internal storage, and networking infrastructure is integrated with the deployment of operating systems, Java packages, and Hadoop along with the

provisioning of Hadoop services. Cisco UCS Director Express for Big Data is integrated with major Hadoop distributions from Hortonworks, Cloudera, and MapR, providing single-pane management across the entire infrastructure.

It complements and communicates with Hadoop managers, providing a system wide perspective and enabling administrators to correlate Hadoop activity with network and computing activity on individual Hadoop nodes.

The appendix section describes on how to go about configuring Cisco UCS Director Express for Big Data and deploying popular Hadoop distributions such as Cloudera, MapR and Hortonworks on the Cisco UCS Integrated Infrastructure for Big Data cluster.

Figure 5 Cisco USCD Express for Big Data



MapR Distribution Including Apache Hadoop: A Complete Hadoop Platform

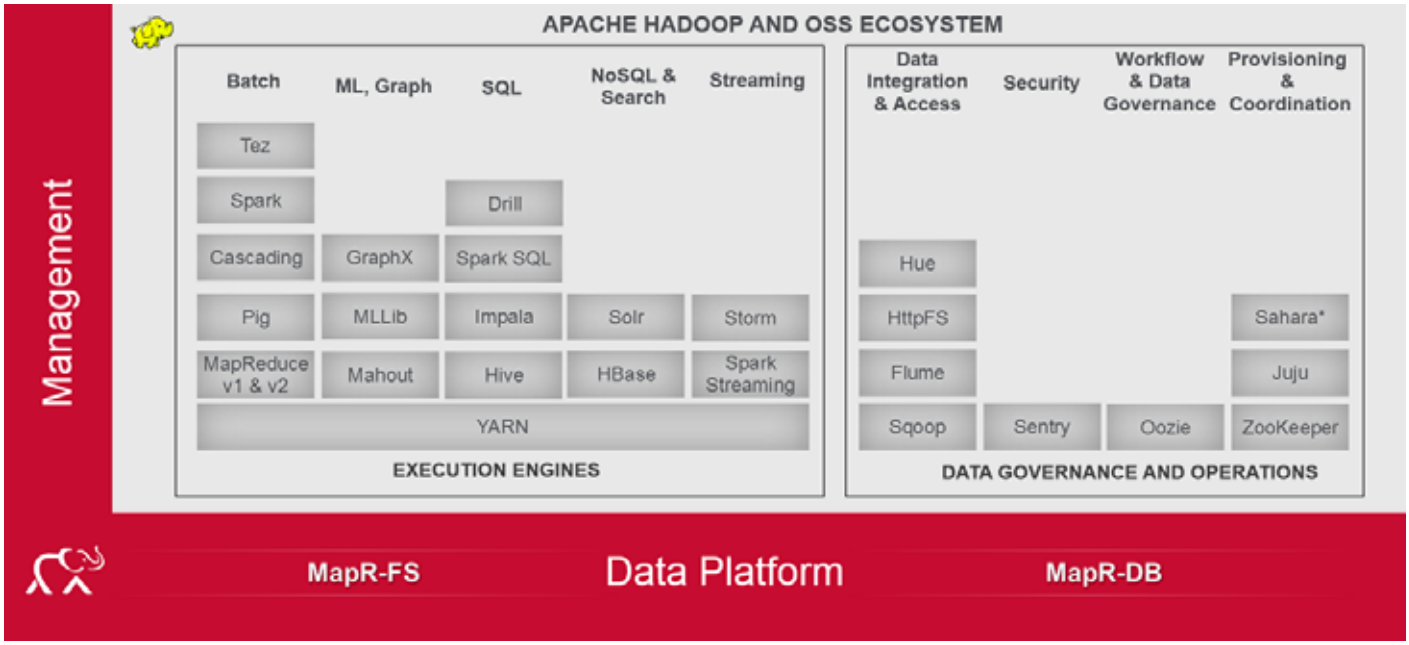
As one of the technology leaders in Hadoop, the MapR Distribution Including Apache Hadoop provides enterprise–class Hadoop solutions that are fast to develop and easy to administer. With significant investment in critical technologies, MapR offers industry's one of the most comprehensive Hadoop

platforms, fully optimized for performance scalability. MapR's distribution delivers more than a dozen tested and validated Hadoop software modules over a fortified data platform, offering exceptional ease of use, reliability, and performance for Hadoop solutions.

Key highlights of MapR Distribution Including Apache Hadoop are:

- **Performance** – Ultra–fast performance and throughput
- **Scalability** – Up to a trillion files, with no restrictions on the number of nodes in a cluster
- **Standards–based API's and tools** – Standard Hadoop API's, ODBC, JDBC, LDAP, Linux PAM, and more
- **MapR Direct Access NFS** – Random read/write, real–time data flows, existing non–Java applications work seamlessly
- **Manageability** – Advanced management console, rolling upgrades, REST API support
- **Integrated security** – Kerberos and non–Kerberos options with wire–level encryption
- **Advanced multi–tenancy** – Volumes, data placement control, job placement control, queues, and more
- **Consistent snapshots** – Full data protection with point–in–time recovery
- **High availability** – Ubiquitous HA with no NameNode architecture, YARN HA, NFS HA
- **Disaster recovery** – Cross–site replication with mirroring
- **MapR–DB** – Integrated Enterprise–Grade NoSQL database

Figure 6 MapR Data Platform



Benefits of Multi-Tenancy in MapR

The MapR Distribution Including Apache Hadoop offers multi–tenancy out–of–the box. It provides powerful features to logically partition a physical cluster to provide separate administrative control, data placement, job execution, and user quotas. Volumes, which is a unique feature in MapR is one of the features that contribute to multi–tenancy. Volumes provide a way to organize data and apply different policies to different data sets, applications and users/groups. A single cluster can have more than thousands of volumes.

In a typical deployment, the data for each user, group, application or business unit is grouped into a single volume so that it can be managed separately from the data of other users, groups, applications, and business units.

Other Hadoop distributions do not support volumes, so policies can only be defined at the file or directory level (too granular) or at the cluster level (too coarse). As a workaround, organizations using other Hadoop distributions create separate physical clusters for each tenant, which results in increased architectural complexities, and higher error/ failure rate.

Multi–tenancy in MapR also has significant total cost of ownership (TCO) advantages, allowing organizations to leverage a single cluster for multiple use cases rather than maintaining a large number of isolated clusters. This reduces overall administrative overhead, and also enables the higher utilization efficiency of a common resource pool.

In a typical deployment, the data for each user, group, application, or business unit is grouped into a single volume so that it can be managed separately from the data of the other users, groups, applications, and business units.

Data Placement Control

With MapR, a volume can be restricted to a subset of a cluster's nodes. This provides the ability to isolate sensitive data/applications, as well as the ability to leverage heterogeneous hardware.

For example, data placement control can be used to keep data such as the personally identifiable information (PII) on separate nodes, or Apache Spark data on nodes that have SSDs. Further, it can be used for more advanced storage tiering policies, such as keeping old data on nodes having a higher storage capacity (such as Cisco UCS C3160) and less compute power, thus providing lower cost per TB storage. In combination with the MapR warden pluggable services, data placement control also enables administrators to designate specific nodes for an application or service, such as for Spark, effectively creating a "mini–cluster" within the larger cluster to guarantee SLAs, and resource availability.

Job Placement Control

MapR provides the ability to restrict a specific job, or jobs from a specific user or group, to a subset of the nodes in the cluster. This enables administrators to guarantee SLAs for specific applications, and to create separation between different applications or business units. This also allows administrators to designate a small subset of the nodes for low–priority jobs or jobs that require access to external systems through the corporate firewall.

Access Control and Security

MapR provides fine–grained access control based on Access Control Expressions (ACEs) for tables, column families, and columns; POSIX access control lists for files; and strong role–based access control (RBAC) for tables, column families, and columns.

MapR provides cryptographically secure wire–level authentication and encryption. Organizations that have a Kerberos infrastructure can leverage it for authentication, while organizations that do not have a Kerberos infrastructure can leverage an integrated, key–based scheme that provides the same security without the complexity associated with deploying and managing Kerberos.

Administration and Reporting

From an administrative perspective, MapR allows organizations to define and enforce storage, CPU, and memory quotas at the volume, user, and group levels. For service providers to provide accurate usage and billing information, MapR offers reporting on resource usage for over 60 different metrics. These metrics are available via the MapR Control System (MCS) browser–based user interface, and—for up–stream integration—via the command–line interface and the REST API.

Solution Overview

This CVD describes architecture and deployment procedures for MapR Distribution on a 64 Cisco UCS C240 M4SX node cluster based on Cisco UCS Integrated Infrastructure for Big Data. The solution describes in detail configuring MapR on the infrastructure. Further, this CVD describes steps to setup Multi–Tenancy on the same MapR installation.

The current version of the Cisco UCS Integrated Infrastructure for Big Data offers the following configuration depending on the compute and storage requirements:

Table 1 Cisco UCS Integrated Infrastructure for Big Data Configuration Details

Performance Optimized	Capacity Optimized
16 Cisco UCS C240 M4 Rack Servers (SFF), each with: • 2 Intel Xeon processors E5-2680 v3 CPUs • 256 GB of memory • Cisco 12-Gbps SAS Modular Raid Controller with 2-GB flash-based write cache (FBWC) • 24 1.2-TB 10K SFF SAS drives (460 TB total) • 2 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for Boot • Cisco UCS VIC 1227 (with 2 10 GE SFP+ ports)	16 Cisco UCS C240 M4 Rack Servers (LFF), each with: • 2 Intel Xeon processors E5-2620 v3 CPU • 128 GB of memory • Cisco 12-Gbps SAS Modular Raid Controller with 2-GB FBWC • 12 4-TB 7.2K LFF SAS drives (768 TB total) • 2 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for Boot • Cisco UCS VIC 1227 (with 2 10 GE SFP+ ports)

Note This CVD describes the install process of MapR for a 64 node of Performance Optimized cluster configuration.

The Performance cluster configuration consists of the following:

- Two Cisco UCS 6296UP Fabric Interconnects
- 64 UCS C240 M4 Rack-Mount servers (16 per rack)
- Four Cisco R42610 standard racks
- Eight vertical power distribution units (PDUs) (Country Specific)

Rack and PDU Configuration

Each rack consists of two vertical PDUs. The master rack consists of two Cisco UCS 6296UP Fabric Interconnects, sixteen Cisco UCS C240 M4 Servers connected to each of the vertical PDUs for redundancy; thereby, ensuring availability during power source failure. The expansion racks consists of sixteen Cisco UCS C240 M4 Servers connected to each of the vertical PDUs for redundancy; thereby, ensuring availability during power source failure, similar to the master rack.

Note Please contact your Cisco representative for country specific information.

Table 2 and Table 3 describe the rack configurations of rack 1 (master rack) and racks 2-4 (expansion racks).

Table 2 Rack 1 (Master Rack)

Cisco 42URack	Master Rack
42	Cisco UCS FI 6296UP
41	
40	Cisco UCS FI 6296UP
39	
38	Unused
37	Unused
36	Unused
35	
34	Unused
33	
32	Cisco UCS C240 M4
31	
30	Cisco UCS C240 M4
29	
28	Cisco UCS C240 M4
27	
26	Cisco UCS C240 M4
25	

Table 2 Rack 1 (Master Rack)

Cisco 42URack	Master Rack
24	Cisco UCS C240 M4
23	
22	Cisco UCS C240 M4
21	
20	Cisco UCS C240 M4
19	
18	Cisco UCS C240 M4
17	
16	Cisco UCS C240 M4
15	
14	Cisco UCS C240 M4
13	
12	Cisco UCS C240 M4
11	
10	Cisco UCS C240 M4
9	
8	Cisco UCS C240 M4
7	
6	Cisco UCS C240 M4
5	
4	Cisco UCS C240 M4
3	
2	Cisco UCS C240 M4
1	

Table 3 Rack 2-4 (Expansion Racks)

Cisco 42URack	Expansion Rack
42	Unused
41	Unused
40	Unused
39	Unused
38	Unused
37	Unused
36	Unused
35	Unused

Table 3 Rack 2-4 (Expansion Racks)

Cisco 42URack	Expansion Rack
34	Unused
33	Unused
32	Cisco UCS C240 M4
31	
30	Cisco UCS C240 M4
29	
28	Cisco UCS C240 M4
27	
26	Cisco UCS C240 M4
25	
24	Cisco UCS C240 M4
23	
22	Cisco UCS C240 M4
21	
20	Cisco UCS C240 M4
19	
18	Cisco UCS C240 M4
17	
16	Cisco UCS C240 M4
15	
14	Cisco UCS C240 M4
13	
12	Cisco UCS C240 M4
11	
10	Cisco UCS C240 M4
9	
8	Cisco UCS C240 M4
7	
6	Cisco UCS C240 M4
5	
4	Cisco UCS C240 M4
3	
2	Cisco UCS C240 M4
1	

Port Configuration on Fabric Interconnects

Table 4 Port Types and Port Numbers

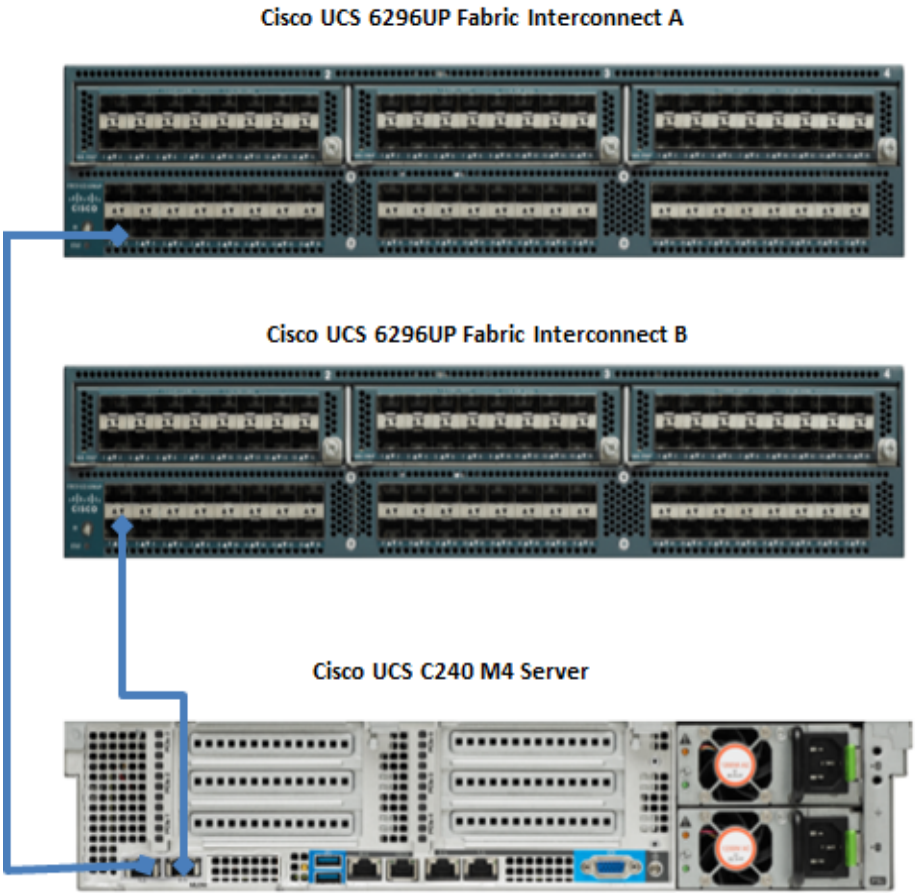
Port Type	Port Number
Network	1
Server	2 to 65

Server Configuration and Cabling for C240M4

The C240 M4 rack server is equipped with Intel Xeon E5–2680 v3 processors, 256 GB of memory, Cisco UCS Virtual Interface Card 1227, Cisco 12–Gbps SAS Modular Raid Controller with 2–GB FBWC,24 1.2–TB 10K SFF SAS drives, 2 120–GB SATA SSD for Boot.

Figure 8, illustrates the port connectivity between the Fabric Interconnect and Cisco UCS C240 M4 server. Sixteen Cisco UCS C240 M4 servers are used in Master rack configurations.

Figure 7 Fabric Topology for C240 M4



For more information on physical connectivity and single–wire management, see:

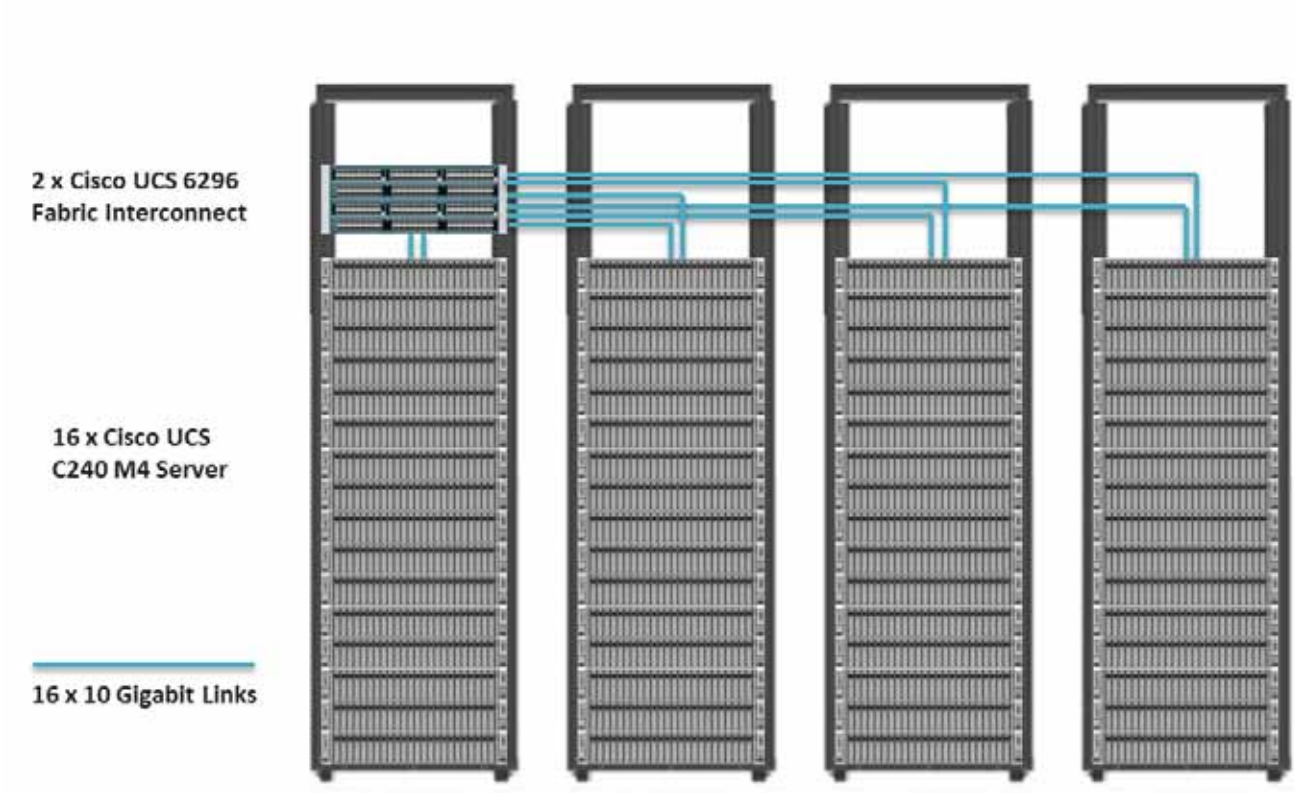
http://www.cisco.com/en/US/docs/unified_computing/ucs/c–series_integration/ucsm2.1/b_UCSM2–1_C–Integration_chapter_010.html

For more information on physical connectivity illustrations and cluster setup, see:

http://www.cisco.com/en/US/docs/unified_computing/ucs/c–series_integration/ucsm2.1/b_UCSM2–1_C–Integration_chapter_010.html#reference_FE5B914256CB4C47B30287D2F9CE3597

Figure 8 depicts a 64–node cluster. Every rack has 16 Cisco UCS C240 M4 servers. Each link in the figure represents 16 x 10 Gigabit Ethernet link from each of the 16 servers connecting to a Fabric Interconnect as a Direct Connect. Every server is connected to both Fabric Interconnect represented with dual link.

Figure 8 64 Nodes Cluster Configuration



Software Distributions and Versions

The software distributions required versions are listed below.

MapR

MapR Hadoop is API–compatible and includes or works with the family of Hadoop ecosystem components such as Spark, Hive, Pig, Flume, and others. For more information, see: https://www.mapr.com/

Red Hat Enterprise Linux (RHEL)

The operating system supported is Red Hat Enterprise Linux 6.5. For more information visit http://www.redhat.com

Software Versions

The software versions tested and validated in this document are shown in table 5.

Table 5 Software Versions

Layer	Component	Version or Release
Compute	Cisco UCS C240-M4	C240M4.2.0.3d
Network	Cisco UCS 6296UP	UCS 2.2(3d)A
	Cisco UCS VIC1227 Firmware	4.0(1d)
	Cisco UCS VIC1227 Driver	2.1.1.66
Storage	LSI SAS 3108	24.5.0-0020
Software	Red Hat Enterprise Linux Server	6.5 (x86_64)
	Cisco UCS Manager	2.2(3d)
	MapR	Enterprise Edition (M5) and Enterprise Database Edition (M7)

Note

- The latest drivers can be downloaded from the link below: https://software.cisco.com/download/release.html?mdfid=283862063&flowid=25886&softwareid=283853158&release=1.5.7d&relind=AVAILABLE&rellifecycle=&reltype=latest
- The latest supported RAID controller driver is already included with the RHEL 6.5 operating system.
- Cisco UCS C240 M4 Rack Servers are supported from UCS firmware 2.2(3d) onwards.

Fabric Configuration

This section provides details for configuring a fully redundant, highly available Cisco UCS 6296 fabric configuration.

1. Initial setup of the Fabric Interconnect A and B.
2. Connect to UCS Manager using virtual IP address of using the web browser.
3. Launch UCS Manager.
4. Enable server, uplink and appliance ports.
5. Start discovery process.

6. Create pools and polices for Service profile template.
7. Create Service Profile template and 64 Service profiles.
8. Associate Service Profiles to servers.

Performing Initial Setup of Cisco UCS 6296 Fabric Interconnects

This section describes the steps to perform initial setup of the Cisco UCS 6296 Fabric Interconnects A and B.

Configure Fabric Interconnect A

1. Connect to the console port on the first Cisco UCS 6296 Fabric Interconnect.
2. At the prompt to enter the configuration method, enter console to continue.
3. If asked to either perform a new setup or restore from backup, enter setup to continue.
4. Enter **y** to continue to set up a new Fabric Interconnect.
5. Enter **y** to enforce strong passwords.
6. Enter the password for the admin user.
7. Enter the same password again to confirm the password for the admin user.
8. When asked if this fabric interconnect is part of a cluster, answer y to continue.
9. Enter A for the switch fabric.
10. Enter the cluster name for the system name.
11. Enter the Mgmt0 IPv4 address.
12. Enter the Mgmt0 IPv4 netmask.
13. Enter the IPv4 address of the default gateway.
14. Enter the cluster IPv4 address.
15. To configure DNS, answer y.
16. Enter the DNS IPv4 address.
17. Answer **y** to set up the default domain name.
18. Enter the default domain name.
19. Review the settings that were printed to the console, and if they are correct, answer **yes** to save the configuration.
20. Wait for the login prompt to make sure the configuration has been saved.

Configure Fabric Interconnect B

1. Connect to the console port on the second Cisco UCS 6296 Fabric Interconnect.
2. When prompted to enter the configuration method, enter console to continue.
3. The installer detects the presence of the partner Fabric Interconnect and adds this fabric interconnect to the cluster. Enter y to continue the installation.
4. Enter the admin password that was configured for the first Fabric Interconnect.
5. Enter the Mgmt0 IPv4 address.
6. Answer **yes** to save the configuration.
7. Wait for the login prompt to confirm that the configuration has been saved.

For more information on configuring Cisco UCS 6200 Series Fabric Interconnect, see:

http://www.cisco.com/en/US/docs/unified_computing/ucs/sw/gui/config/guide/2.0/b_UCSM_GUI_Configuration_Guide_2_0_chapter_0100.html

Logging Into Cisco UCS Manager

Follow these steps to login to Cisco UCS Manager.

1. Open a web browser and navigate to the Cisco UCS 6296 Fabric Interconnect cluster address.
2. Click the **Launch** link to download the Cisco UCS Manager software.
3. If prompted to accept security certificates, accept as necessary.
4. When prompted, enter **admin** for the user-name and enter the administrative password.
5. Click **Login** to log in to the Cisco UCS Manager.

Upgrading Cisco UCS Manager Software to Version 2.2(3d)

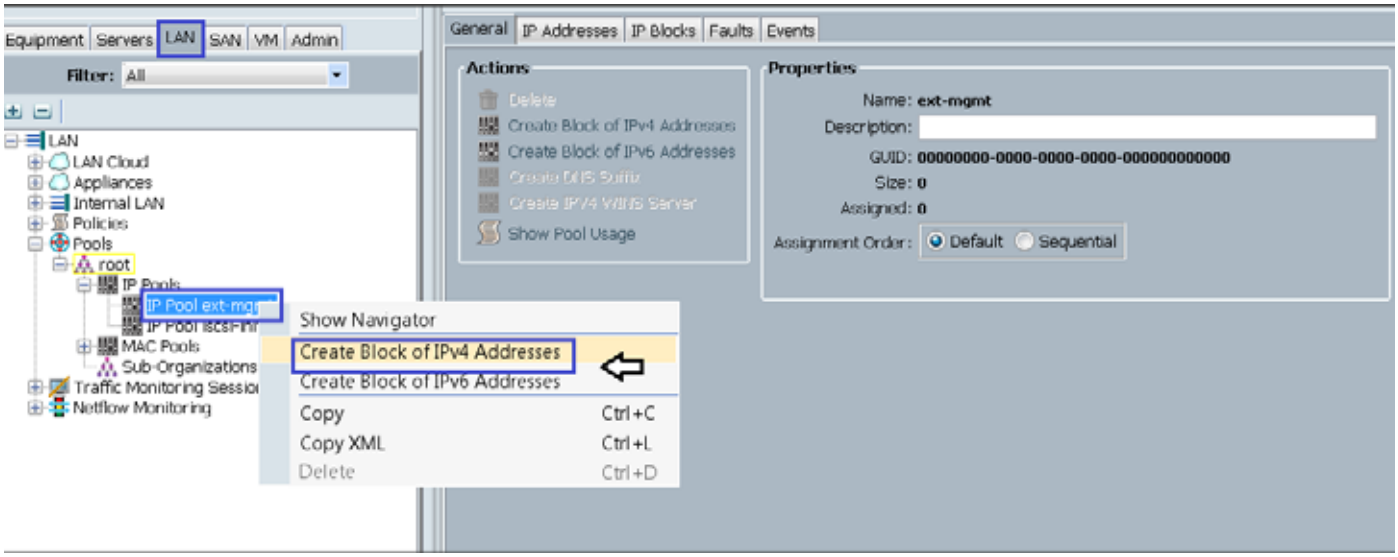
This document assumes the use of UCS 2.2(3d). Refer to Upgrading between Cisco UCS 2.0 Releases to upgrade the Cisco UCS Manager software and UCS 6296 Fabric Interconnect software to version 2.2(3d). Also, make sure the UCS C-Series version 2.2(3d) software bundles is installed on the Fabric Interconnects.

Adding Block of IP Addresses for KVM Access

These steps provide details for creating a block of KVM IP addresses for server access in the Cisco UCS environment.

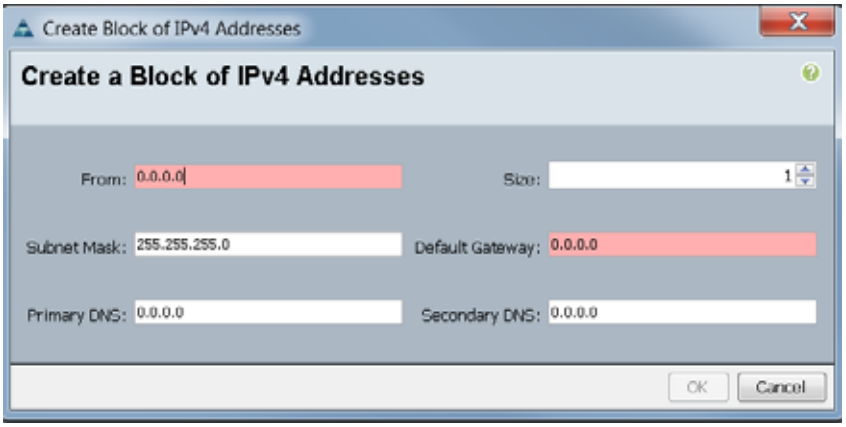
1. Select the **LAN** tab at the top of the left window.
2. Select **Pools > IP Pools > IP Pool ext-mgmt**.
3. Right-click **IP Pool ext-mgmt**
4. Select **Create Block of IPv4 Addresses**.

Figure 9 Adding Block of IPv4 Addresses for KVM Access Part 1



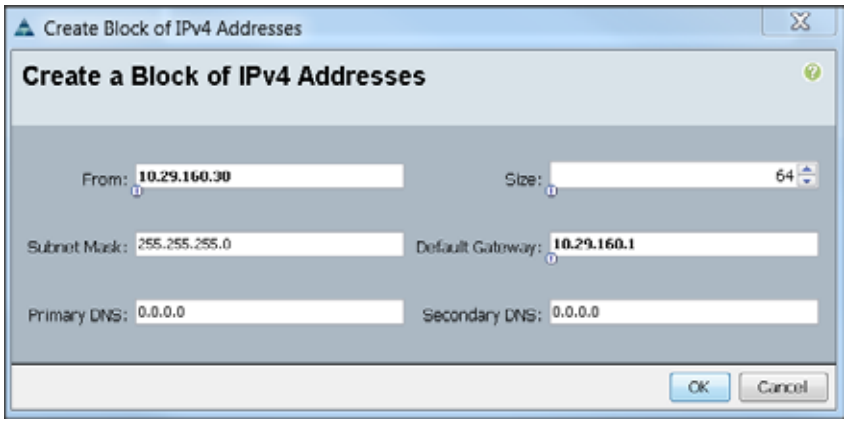
5. Enter the starting IP address of the block and number of IPs needed, as well as the subnet and gateway information.

Figure 10 Adding Block of IPv4 Addresses for KVM Access Part 2



6. Click **OK** to create the IP block.
7. Click **OK** in the message box.

Figure 11 Adding Block of IPv4 Addresses for KVM Access Part 3



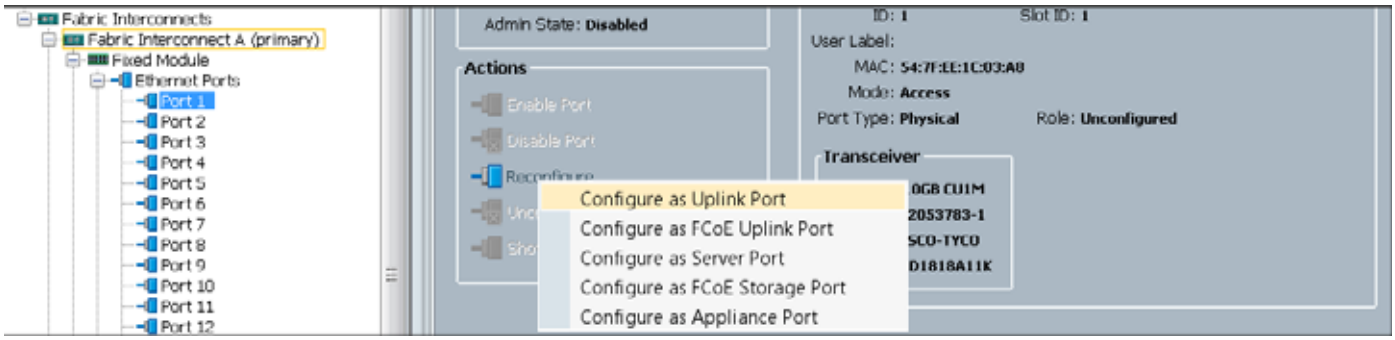
Enabling Uplink Port

These steps provide details for enabling uplinks ports.

1. Select the Equipment tab on the top left of the window.
2. Select **Equipment > Fabric Interconnects > Fabric Interconnect A (primary) > Fixed Module.**
3. Expand the Unconfigured Ethernet Ports section.
4. Select **port 1**, that is connected to the uplink switch, right-click, then select **Reconfigure > Configure as Uplink Port**.
5. Select **Show Interface** and select 10GB for Uplink Connection.
6. A pop-up window appears to confirm your selection. Click **Yes**, then click **OK** to continue.

7. Select **Equipment > Fabric Interconnects > Fabric Interconnect B (subordinate) > Fixed Module**.
8. Expand the Unconfigured Ethernet Ports section.
9. Select **port 1**, that is connected to the uplink switch, right-click, then select **Reconfigure > Configure as Uplink Port**.
10. Select **Show Interface** and select 10GB for Uplink Connection.
11. A pop-up window appears to confirm your selection. Click **Yes**, then click **OK** to continue.

Figure 12 Enabling Uplink Ports



Configuring VLANs

VLANs are configured as in shown in table 6.

Table 6 VLAN Configurations

VLAN	Fabric	NIC Port	Function	Failover
default(VLAN1)	A	eth0	Management, User connectivity	Fabric Failover to B
vlan11_DATA1	B	eth1	MapR	Fabric Failover to A
vlan12_DATA2	A	eth2	MapR	Fabric Failover to B

All of the VLANs created need to be trunked to the upstream distribution switch connecting the fabric interconnects. For this deployment default VLAN1 is configured for management access (Installing and configuring OS, clustershell commands, setup NTP, user connectivity, etc) and vlan11_DATA1 and vlan12_DATA2 is configured for Hadoop Data traffic.

With MapR supporting multiple NICs, where Hadoop uses multiple IP subnets for its data traffic, vlan12_DATA2 can be configured to carry Hadoop Data traffic allowing use of both the Fabrics (10 GigE on each Fabric allowing 20Gbps active–active connectivity).

Further, if there are other distributed applications co–existing in the same Hadoop cluster, then these applications could use vlan12_DATA2 providing full 10GigE connectivity to this application on a different fabric without affecting Hadoop Data traffic on vlan11_DATA1

Note

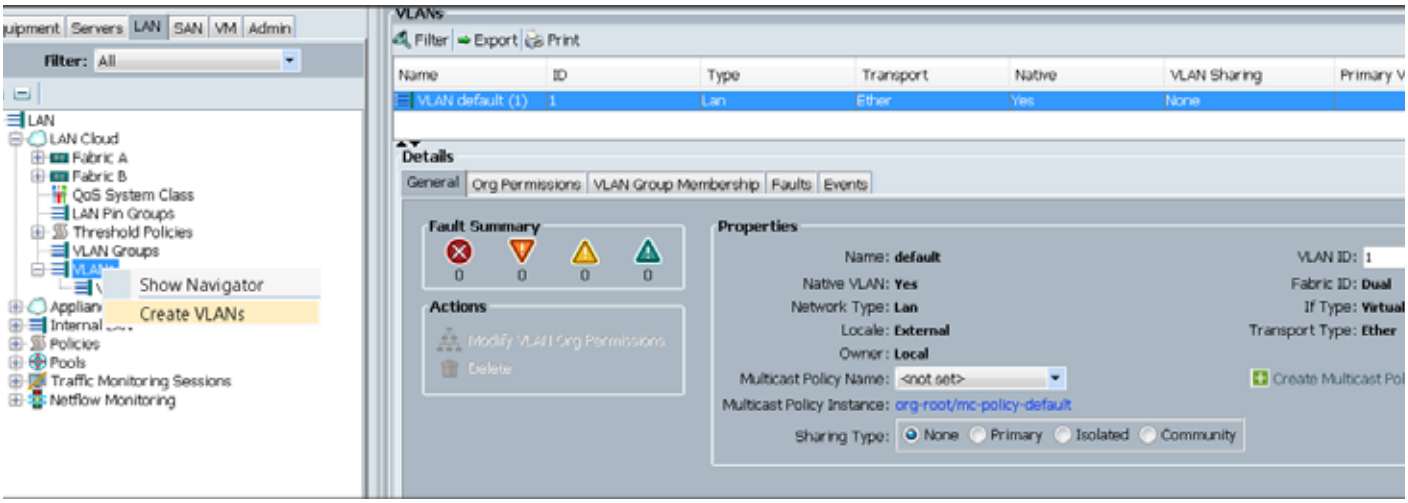
- All applications talking to Hadoop should be able to reach Hadoop VLAN. That is, all applications should be able to access all the Hadoop nodes.

- We are using default VLAN1 for management traffic.

Follow these steps to configure the VLANs in the Cisco UCS Manager GUI:

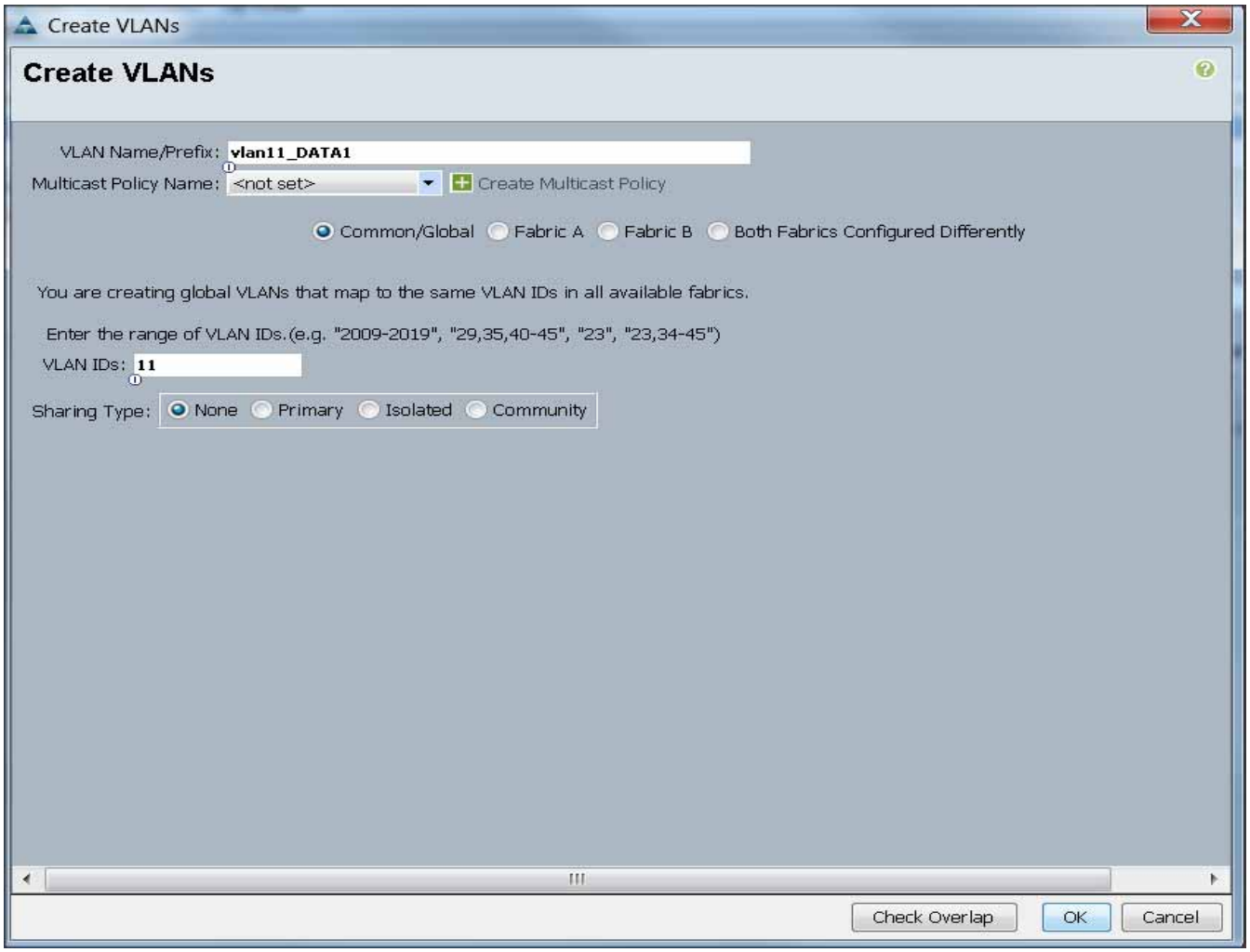
1. Select the **LAN** tab in the left pane in the UCS Manager GUI.
2. Select **LAN** > **VLANs**.
3. Right-click the **VLANs** under the root organization.
4. Select **Create VLANs** to create the VLAN.

Figure 13 Creating VLAN



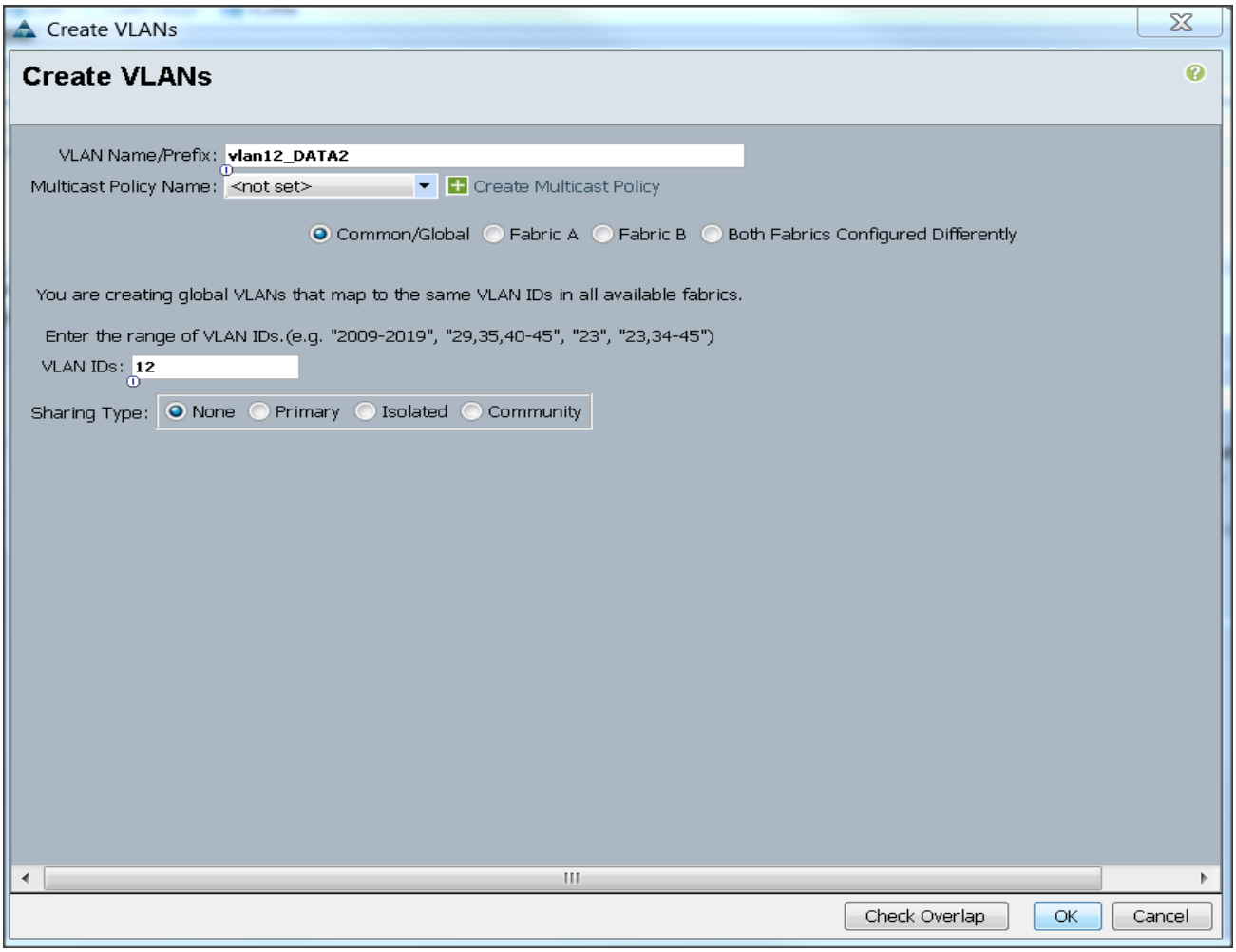
5. Enter vlan11_DATA1 for the VLAN Name.
6. Click the **Common/Global** radio button for the vlan11_DATA1.
7. Enter 11 on VLAN IDs of the Create VLAN IDs.
8. Click **OK** and then, click **Finish**.
9. Click **OK** in the success message box.

Figure 14 Creating VLAN for Data



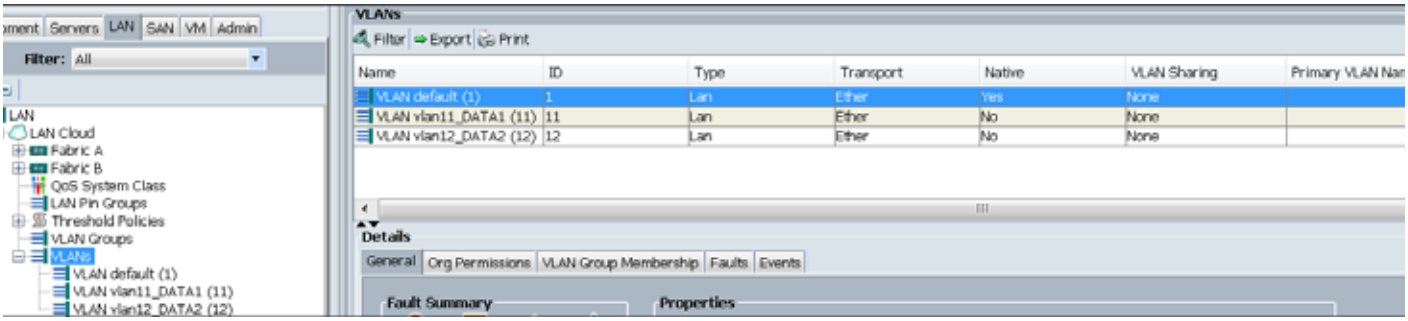
10. Select the **LAN** tab in the left pane again
11. Select **LAN > VLANs**.
12. Right-click the **VLANs** under the root organization.
13. Select Create **VLANs** to create the VLAN.
14. Enter vlan12_DATA2 for the VLAN Name.
15. Click the **Common/Global** radio button for the vlan12_DATA2.
16. Enter 12 on VLAN IDs of the Create VLAN IDs.
17. Click **OK** and then, click **Finish**.

Figure 15 Creating VLAN for Hadoop Data



18. The below screenshot shows the created VLANs.

Figure 16 List of VLANs created for Hadoop Data

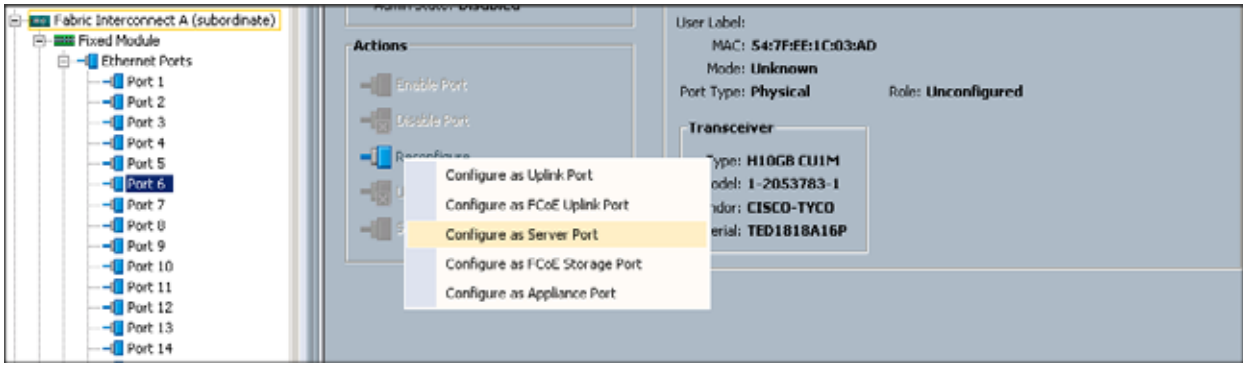


Enabling Server Ports

These steps provide details for enabling server ports.

19. Select the **Equipment** tab on the top left of the window.
20. Select **Equipment > Fabric Interconnects > Fabric Interconnect A (primary) > Fixed Module**.
21. Expand the Unconfigured Ethernet Ports section.
22. Select all the ports that are connected to the Servers right-click them, and select **Reconfigure > Configure as a Server Port**.
23. A pop-up window appears to confirm your selection. Click **Yes** then **OK** to continue.
24. Select **Equipment > Fabric Interconnects > Fabric Interconnect B (subordinate) > Fixed Module**.
25. Expand the Unconfigured Ethernet Ports section.
26. Select all the ports that are connected to the Servers right-click them, and select **Reconfigure > Configure as a Server Port**.
27. A pop-up window appears to confirm your selection. Click **Yes** then **OK** to continue.

Figure 17 Enabling Server Ports



Creating Pools for Service Profile Templates

Creating an Organization

Organizations are used as a means to arrange and restrict access to various groups within the IT organization, thereby enabling multi-tenancy of the compute resources. This document does not assume the use of Organizations; however the necessary steps are provided for future reference.

Follow these steps to configure an organization within the Cisco UCS Manager GUI:

1. Click **New** on the top left corner in the right pane in the UCS Manager GUI.
2. Select Create Organization from the options
3. Enter a name for the organization.
4. (Optional) Enter a description for the organization.

5. Click **OK**.
6. Click **OK** in the success message box.

Creating MAC Address Pools

Follow these steps to create MAC address pools:

1. Select the **LAN** tab on the left of the window.
2. Select **Pools** > **root**.
3. Right-click **MAC** Pools under the root organization.
4. Select **Create MAC Pool** to create the MAC address pool. Enter ucs for the name of the MAC pool.
5. (Optional) Enter a description of the MAC pool.
6. Select Assignment Order Sequential.
7. Click **Next**.
8. Click **Add**.
9. Specify a starting MAC address.
10. Specify a size of the MAC address pool, which is sufficient to support the available server resources.
11. Click **OK**.

Figure 18 Creating MAC Pool Window

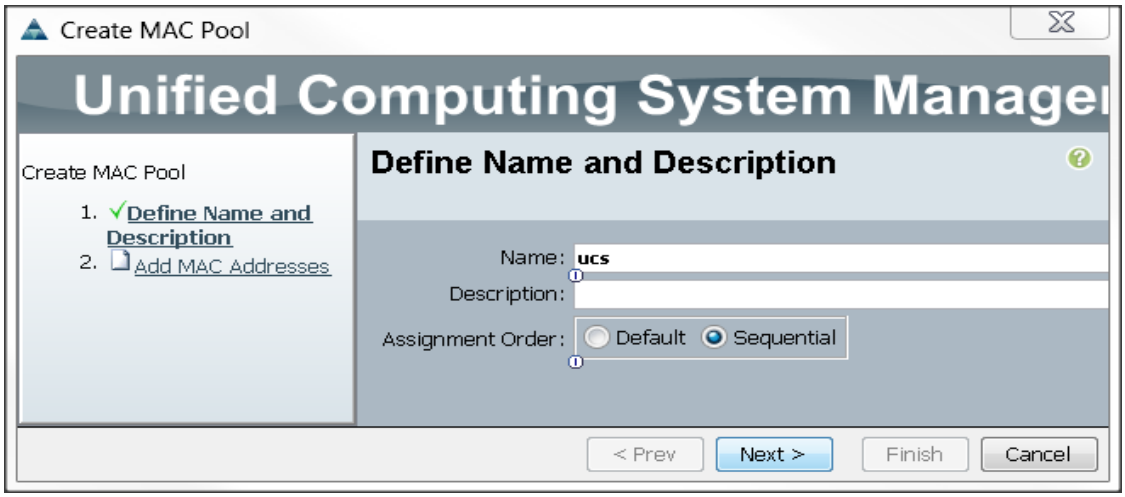
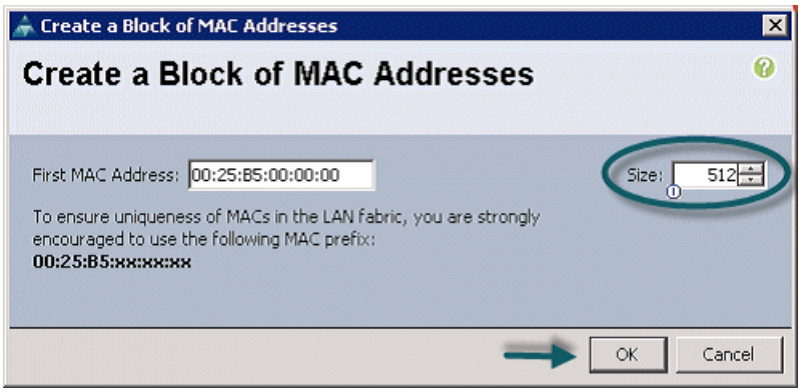
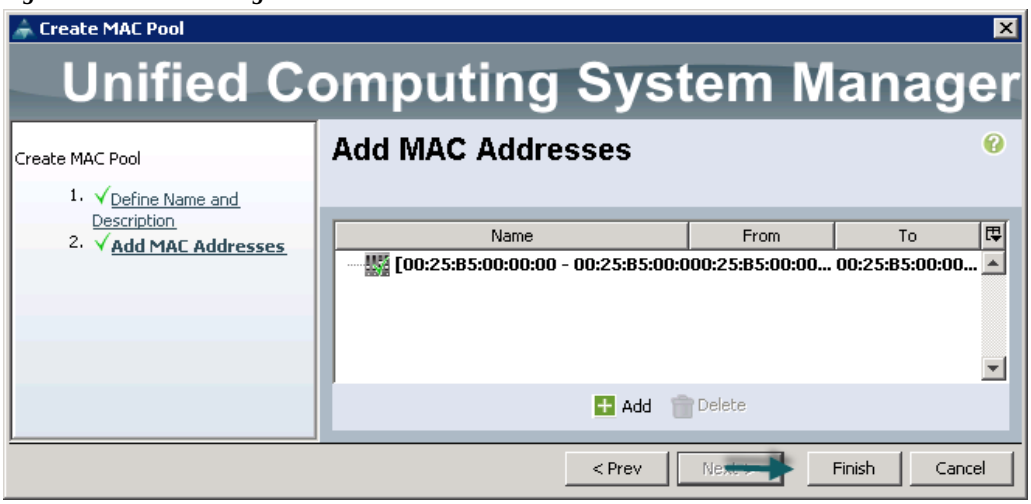


Figure 19 Specifying First MAC Address and Size



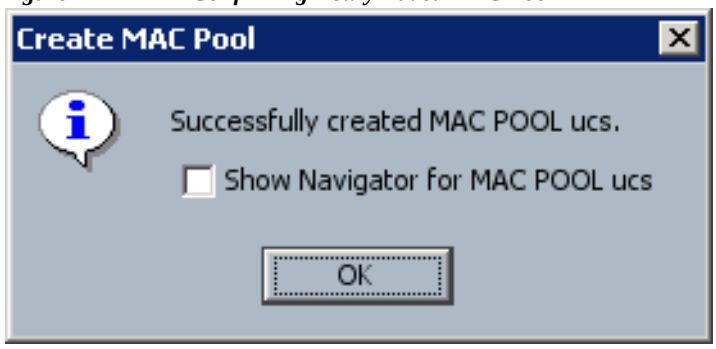
12. Click **Finish**.

Figure 20 Adding MAC Addresses



13. When the message box displays, click **OK**.

Figure 21 Confirming Newly Added MAC Pool



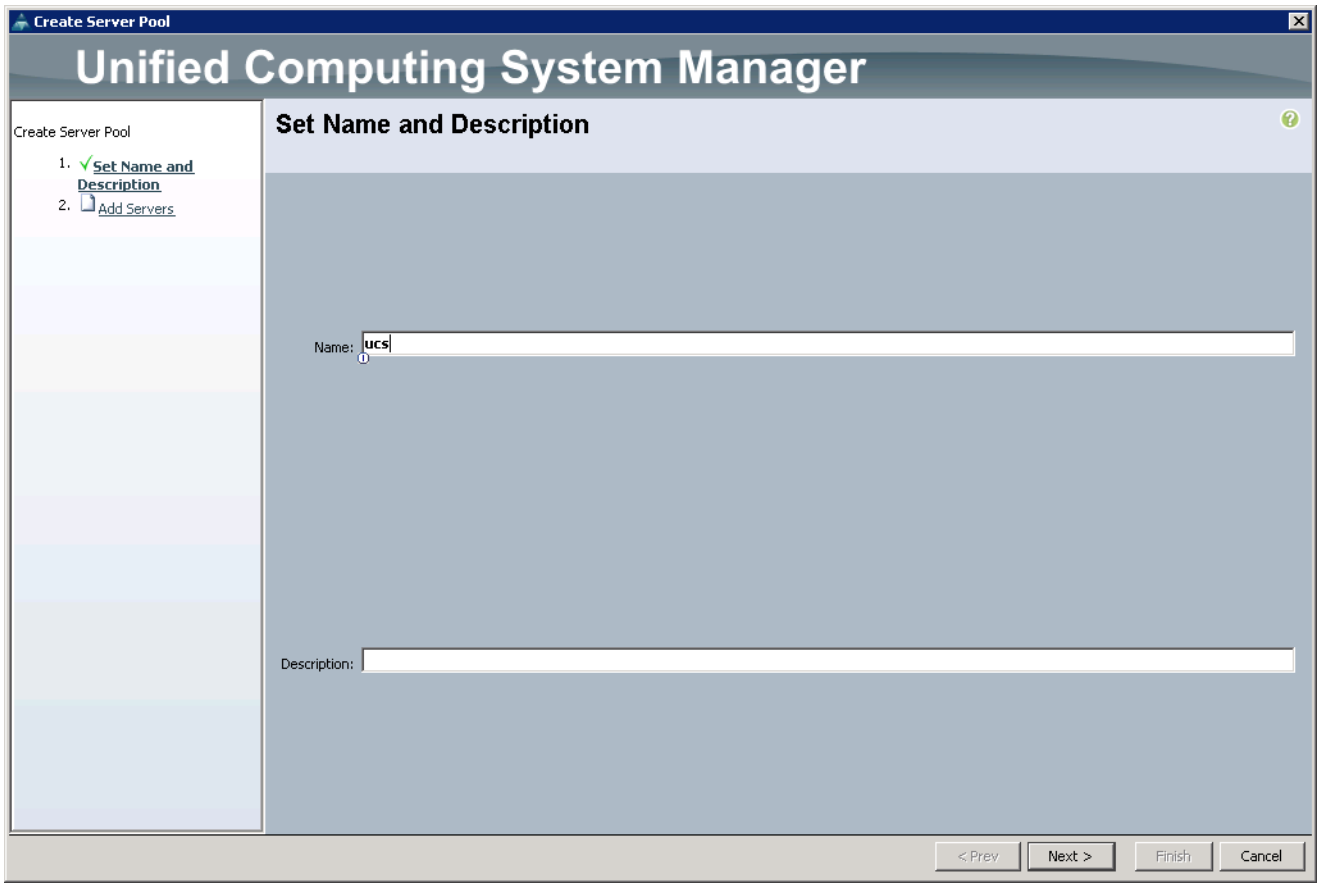
Creating Server Pools

A server pool contains a set of servers. These servers typically share the same characteristics. Those characteristics can be their location in the chassis, or an attribute such as server type, amount of memory, local storage, type of CPU, or local drive configuration. You can manually assign a server to a server pool, or use server pool policies and server pool policy qualifications to automate the assignment

Follow these steps to configure the server pool within the Cisco UCS Manager GUI:

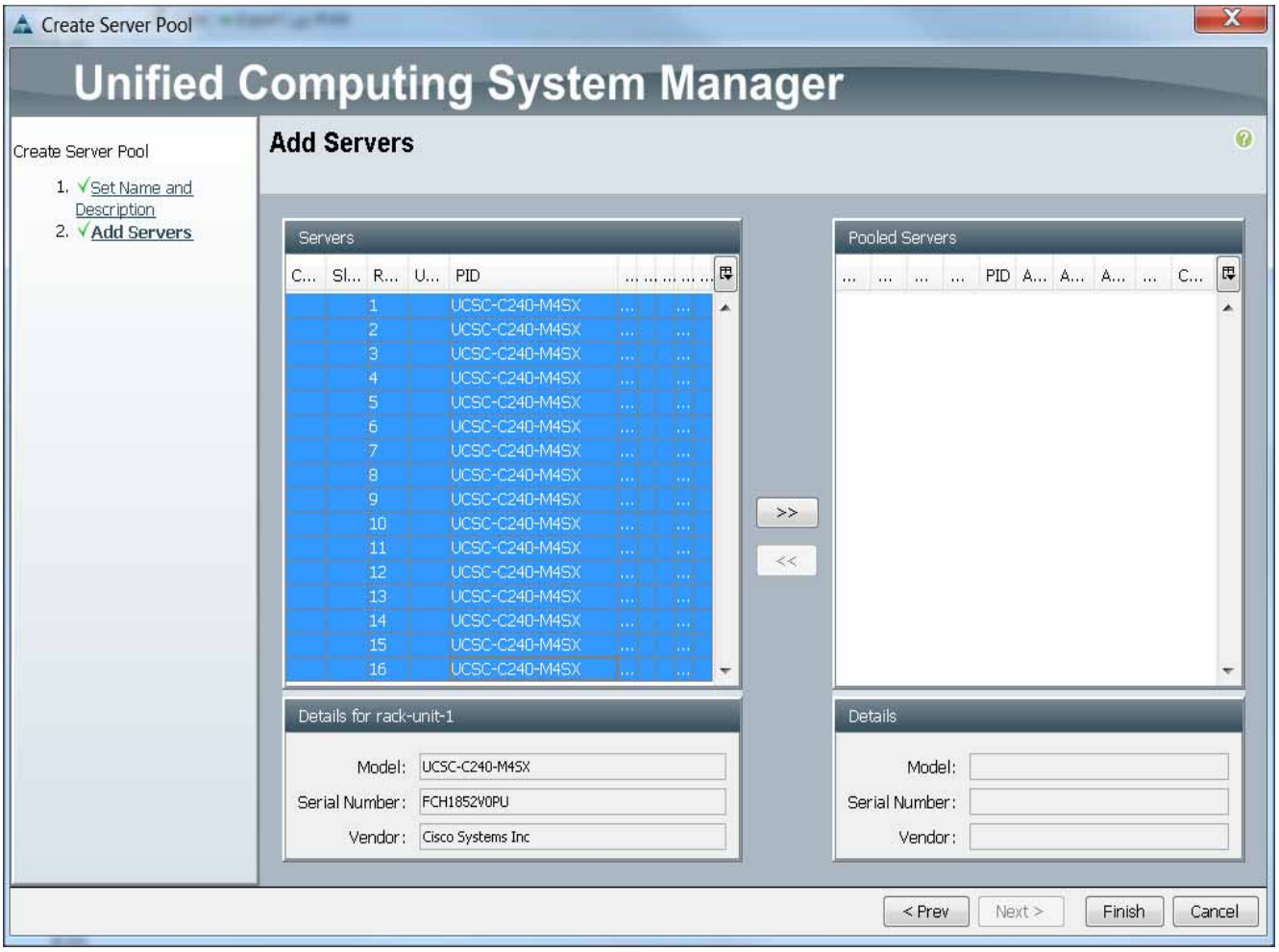
1. Select the **Servers** tab in the left pane in the UCS Manager GUI.
2. Select **Pools** > **root**.
3. Right-click the **Server Pools**.
4. Select **Create Server Pool**.
5. Enter your required name (ucs) for the Server Pool in the name text box.
6. (Optional) enter a description for the organization
7. Click **Next** to add the servers.

Figure 22 Setting Name and Description of Server Pool



8. Select all the Cisco UCS C240M4SX servers to be added to the server pool you previously created (ucs), then Click >> to add them to the pool.
9. Click **Finish**.
10. Click **OK**, and then click **Finish**.

Figure 23 Adding Servers to the Server Pool



Creating Policies for Service Profile Templates

Creating Host Firmware Package Policy

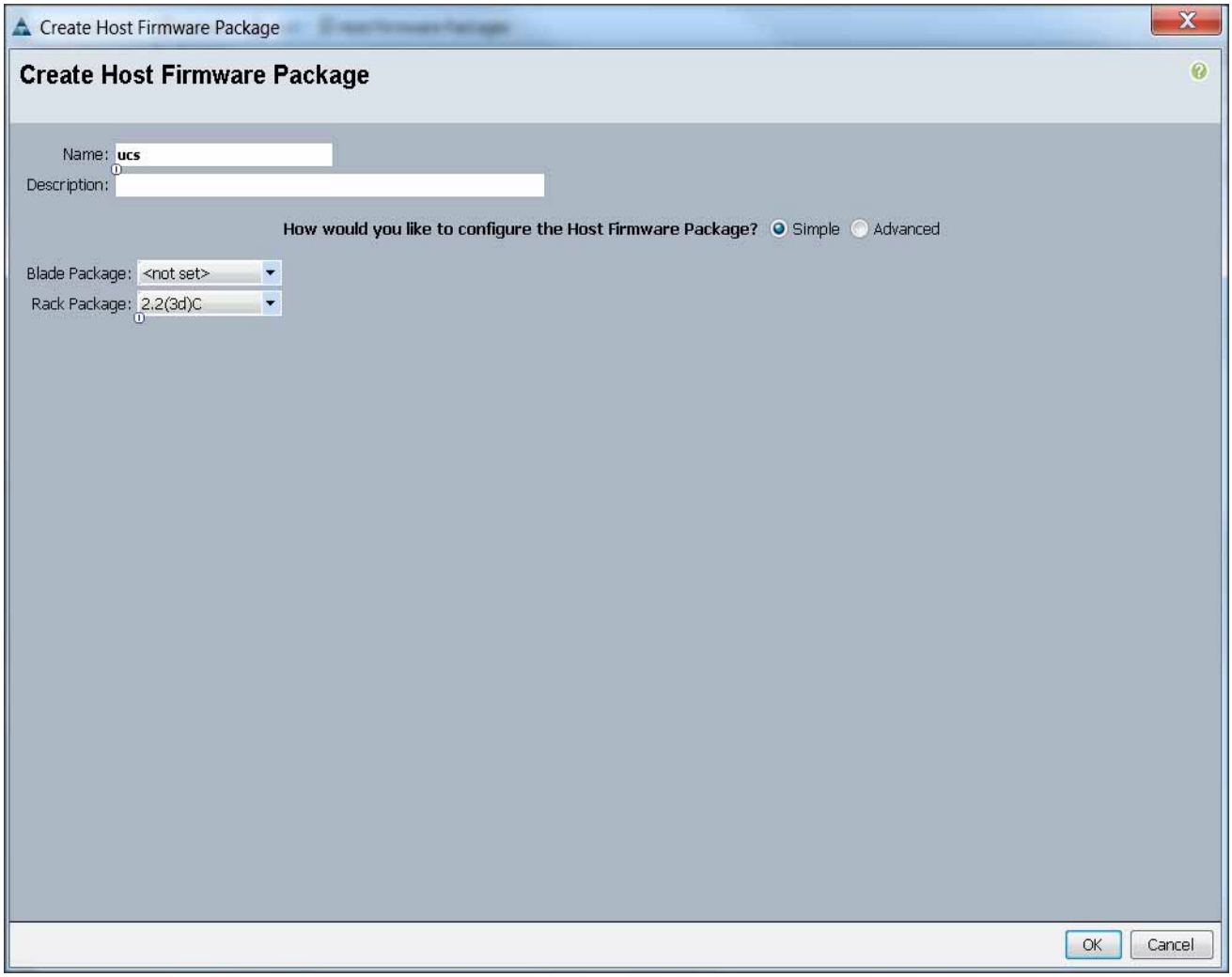
Firmware management policies allow the administrator to select the corresponding packages for a given server configuration. These include adapters, BIOS, board controllers, FC adapters, HBA options, ROM and storage controller properties as applicable.

Follow these steps to create a firmware management policy for a given server configuration using the Cisco UCS Manager GUI:

1. Select the **Servers** tab in the left pane in the UCS Manager GUI.
2. Select **Policies > root**.
3. Right-click **Host Firmware Packages**.
4. Select **Create Host Firmware Package**.
5. Enter your required Host Firmware package name (ucs).

6. Click the **Simple** radio button to configure the Host Firmware package.
7. Select the appropriate Rack package that you have.
8. Click **OK** to complete creating the management firmware package.
9. Click **OK**.

Figure 24 Creating Host Firmware Package



Creating QoS Policies

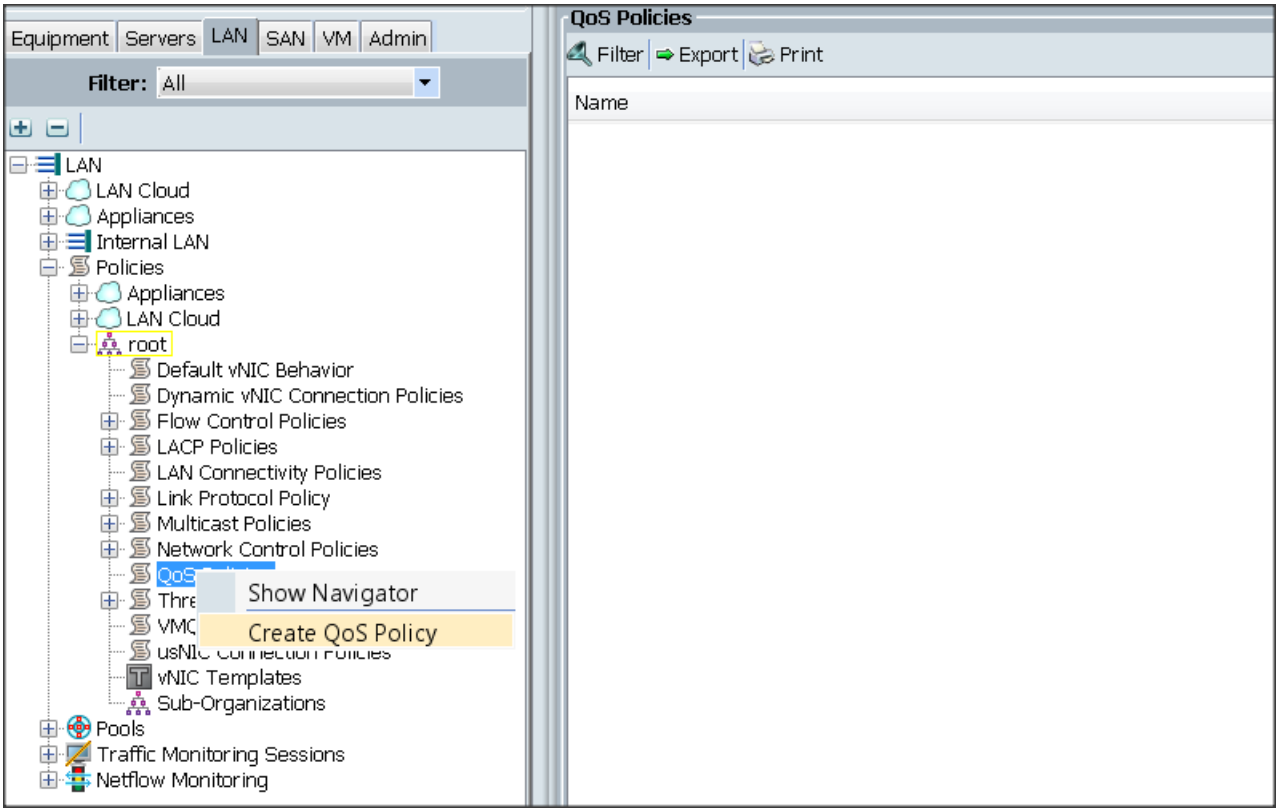
Follow these steps to create the QoS policy for a given server configuration using the Cisco UCS Manager GUI:

Best Effort Policy

1. Select the **LAN** tab in the left pane in the UCS Manager GUI.
2. Select **Policies > root**.

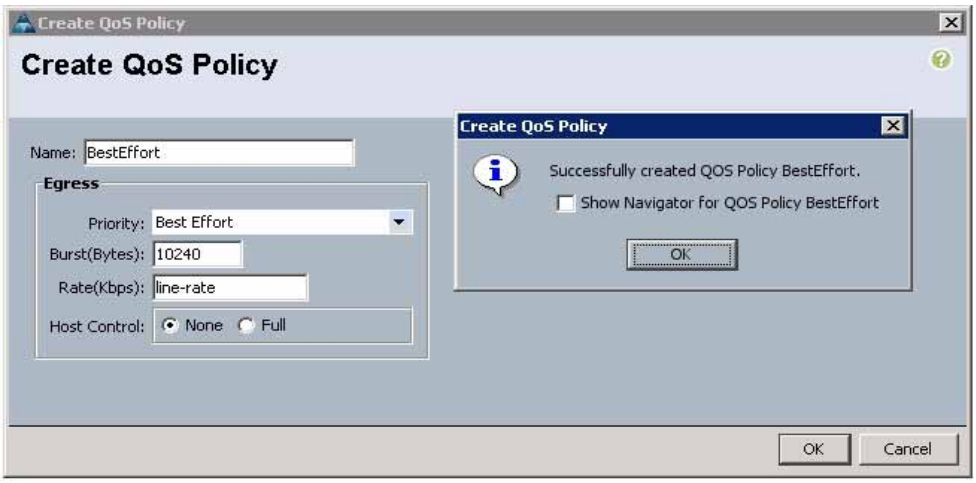
3. Right-click **QoS** Policies.
4. Select **Create QoS Policy**.

Figure 25 Creating QoS Policy



5. Enter BestEffort as the name of the policy.
6. Select BestEffort from the drop down menu.
7. Keep the Burst (Bytes) field as default (10240).
8. Keep the Rate (Kbps) field as default (line-rate).
9. Keep Host Control radio button as default (none).
10. Once the pop-up window appears, click **OK** to complete the creation of the Policy.

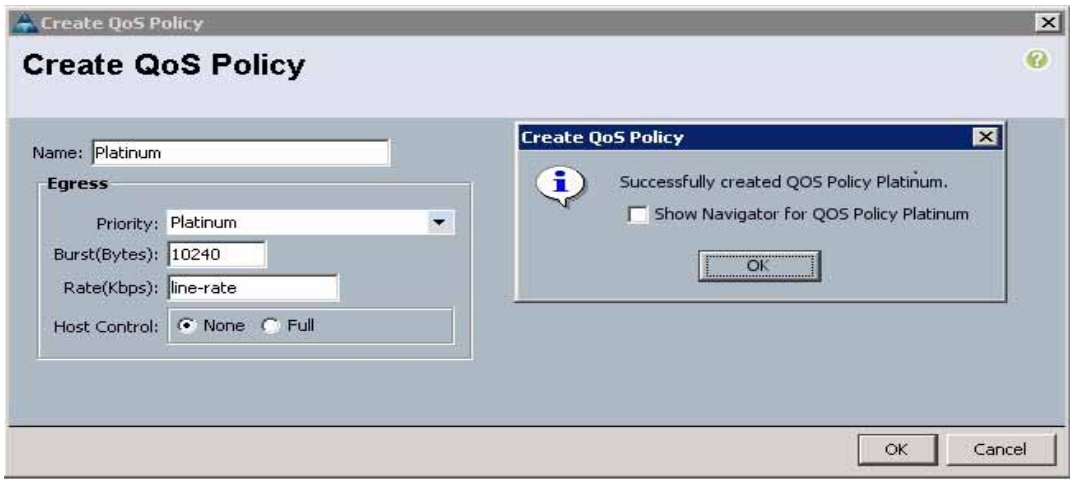
Figure 26 Creating BestEffort QoS Policy



Platinum Policy

1. Select the **LAN** tab in the left pane in the UCS Manager GUI.
2. Select **Policies > root**.
3. Right-click **QoS Policies**.
4. Select **Create QoS Policy**.
5. Enter Platinum as the name of the policy.
6. Select Platinum from the drop down menu.
7. Keep the Burst (Bytes) field as default (10240).
8. Keep the Rate (Kbps) field as default (line-rate).
9. Keep Host Control radio button as default (none).
10. Once the pop-up window appears, click **OK** to complete the creation of the Policy.

Figure 27 Creating Platinum QoS Policy



Setting Jumbo Frames

Follow these steps for setting up the Jumbo frames and enabling QoS:

1. Select the **LAN** tab in the left pane in the UCS Manager GUI.
2. Select **LAN Cloud** > **QoS System Class**.
3. In the right pane, select the **General** tab
4. In the Platinum row, enter 9000 for MTU.
5. Check the **Enabled** Check box next to Platinum.
6. In the Best Effort row, select best-effort for weight.
7. In the Fiber Channel row, select none for weight.
8. Click **Save Changes**.
9. Click **OK**.

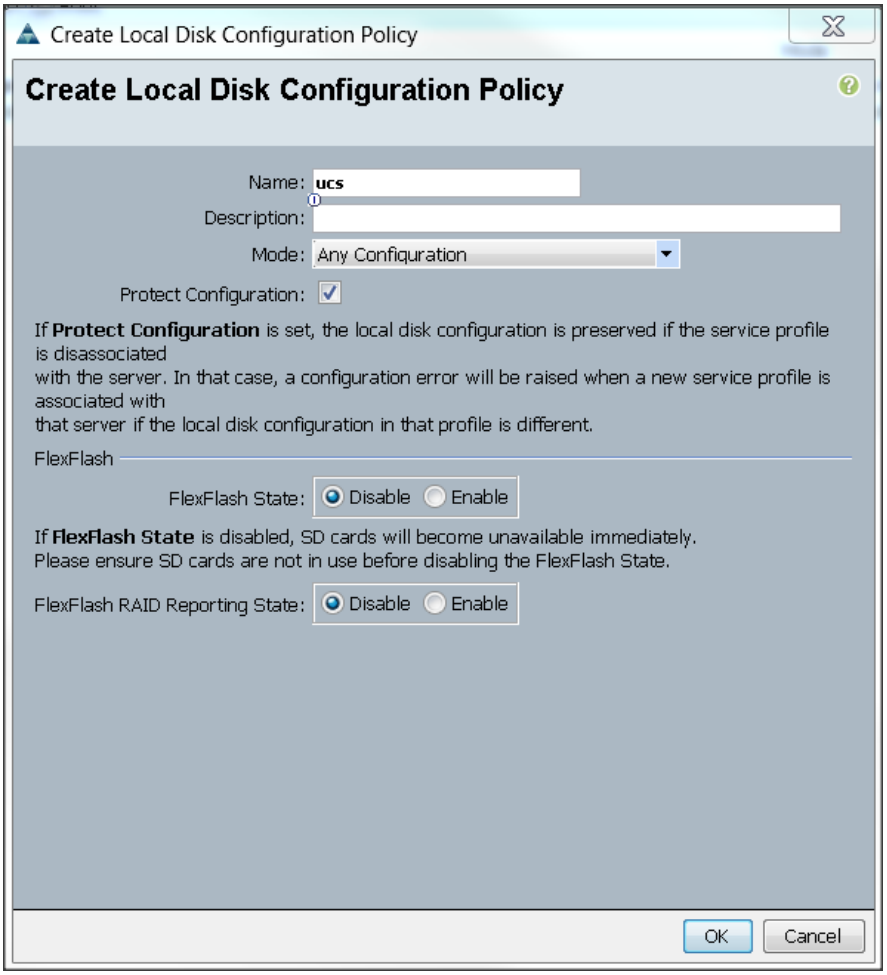
Figure 28 Setting Jumbo Frames

Creating Local Disk Configuration Policy

Follow these steps to create local disk configuration in the Cisco UCS Manager GUI:

1. Select the **Servers** tab on the left pane in the UCS Manager GUI.
2. Go to **Policies** > **root**.
3. Right-click Local Disk Config Policies.
4. Select Create Local Disk Configuration Policy.
5. Enter ucs as the local disk configuration policy name.
6. Change the Mode to Any Configuration. Check the **Protect Configuration** box.
7. Keep the FlexFlash State field as default (Disable).
8. Keep the FlexFlash RAID Reporting State field as default (Disable).
9. Click **OK** to complete the creation of the Local Disk Configuration Policy.
10. Click **OK**.

Figure 29 Configuring Local Disk Policy



Creating Server BIOS Policy

The BIOS policy feature in Cisco UCS automates the BIOS configuration process. The traditional method of setting the BIOS is done manually and is often error-prone. By creating a BIOS policy and assigning the policy to a server or group of servers, you can enable transparency within the BIOS settings configuration.

Note BIOS settings can have a significant performance impact, depending on the workload and the applications. The BIOS settings listed in this section is for configurations optimized for best performance which can be adjusted based on the application, performance and energy efficiency requirements.

Follow these steps to create a server BIOS policy using the Cisco UCS Manager GUI:

1. Select the **Servers** tab in the left pane in the UCS Manager GUI.
2. Select **Policies** > **root**.
3. Right-click **BIOS Policies**.

4. Select Create BIOS Policy.
5. Enter your preferred BIOS policy name (ucs).
6. Change the BIOS settings as per the following figures:

Figure 30 Creating Server BIOS Policy

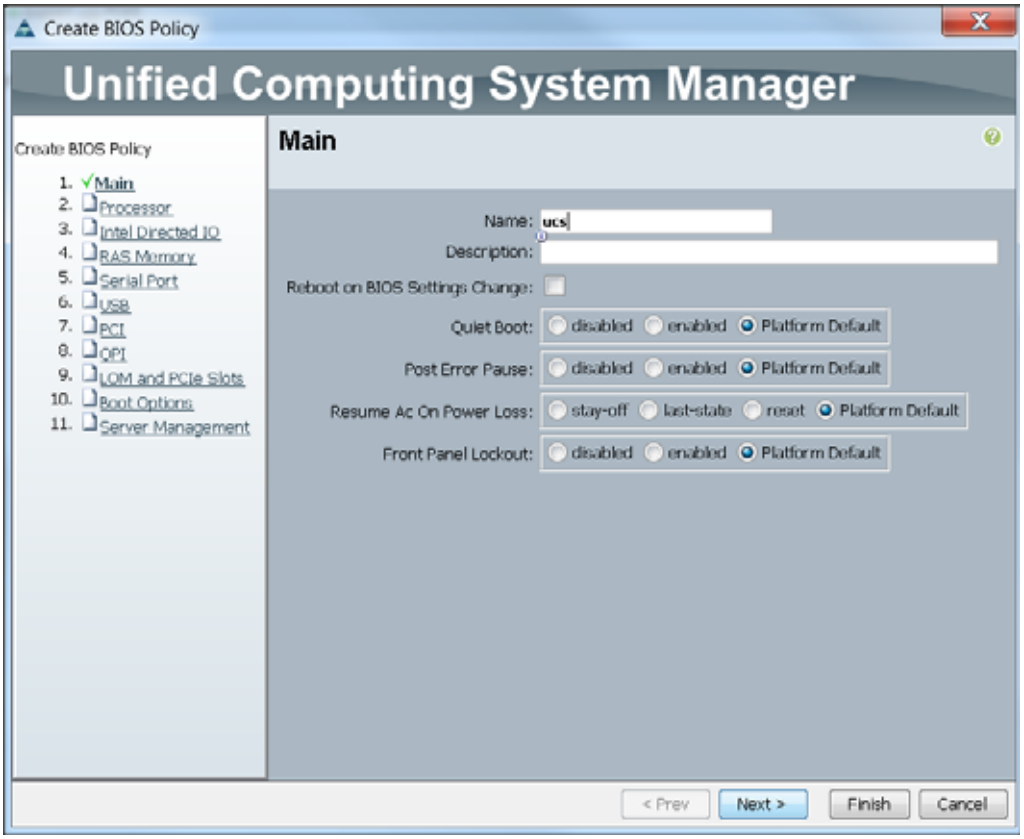


Figure 31 Creating Server BIOS Policy for Processor

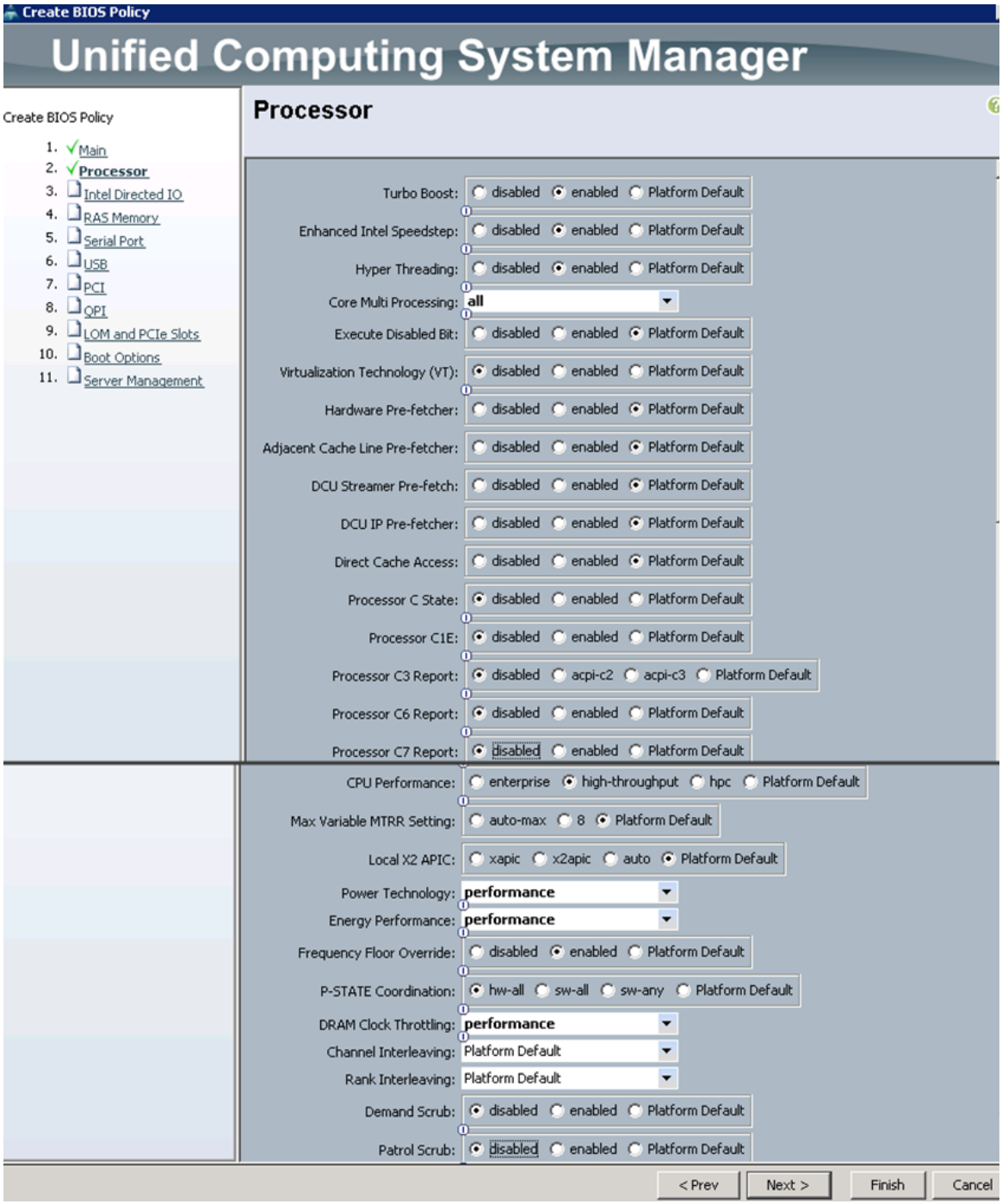


Figure 32 Creating Server BIOS Policy for Intel Directed IO

7. Click **Finish** to complete creating the BIOS policy.
8. Click **OK**.

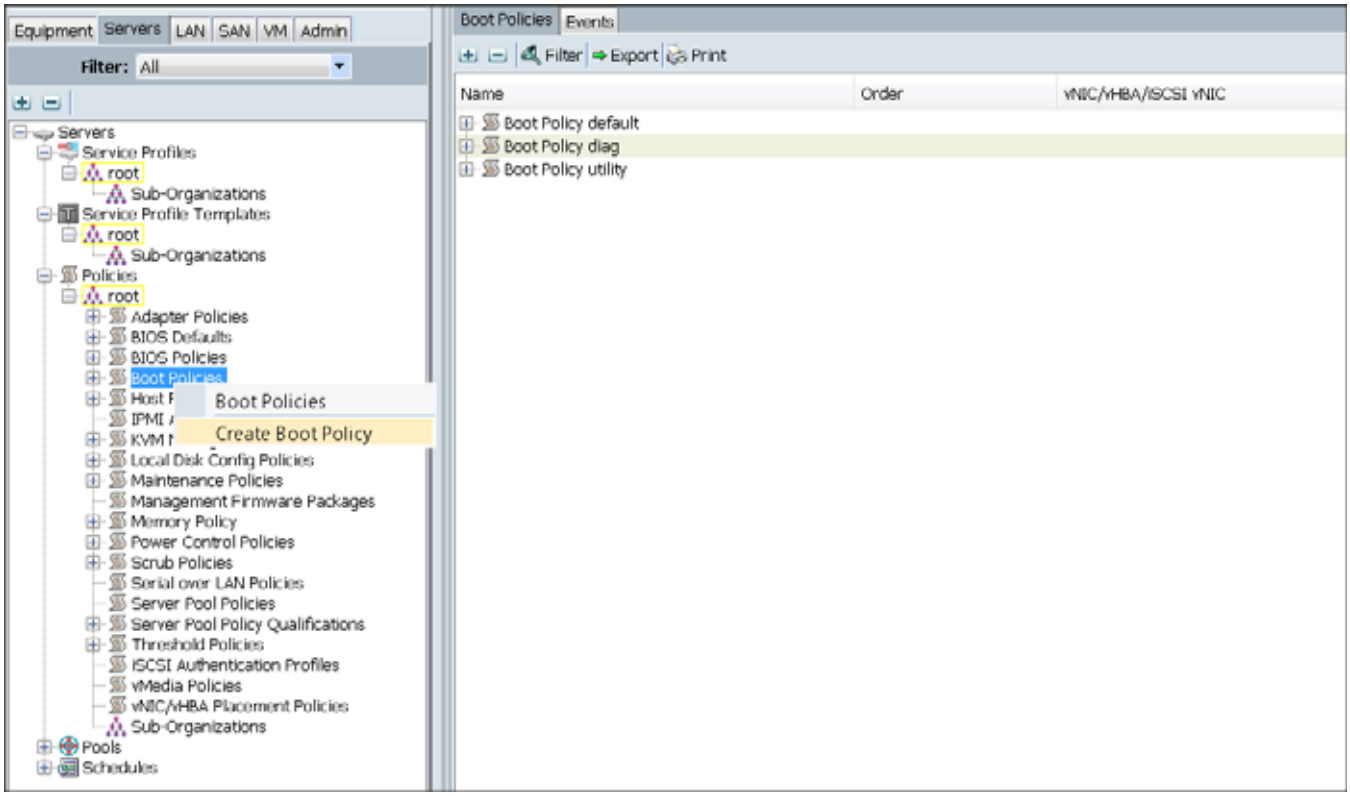
Figure 33 Creating Server BIOS Policy for Memory

Creating Boot Policy

Follow these steps to create boot policies within the Cisco UCS Manager GUI:

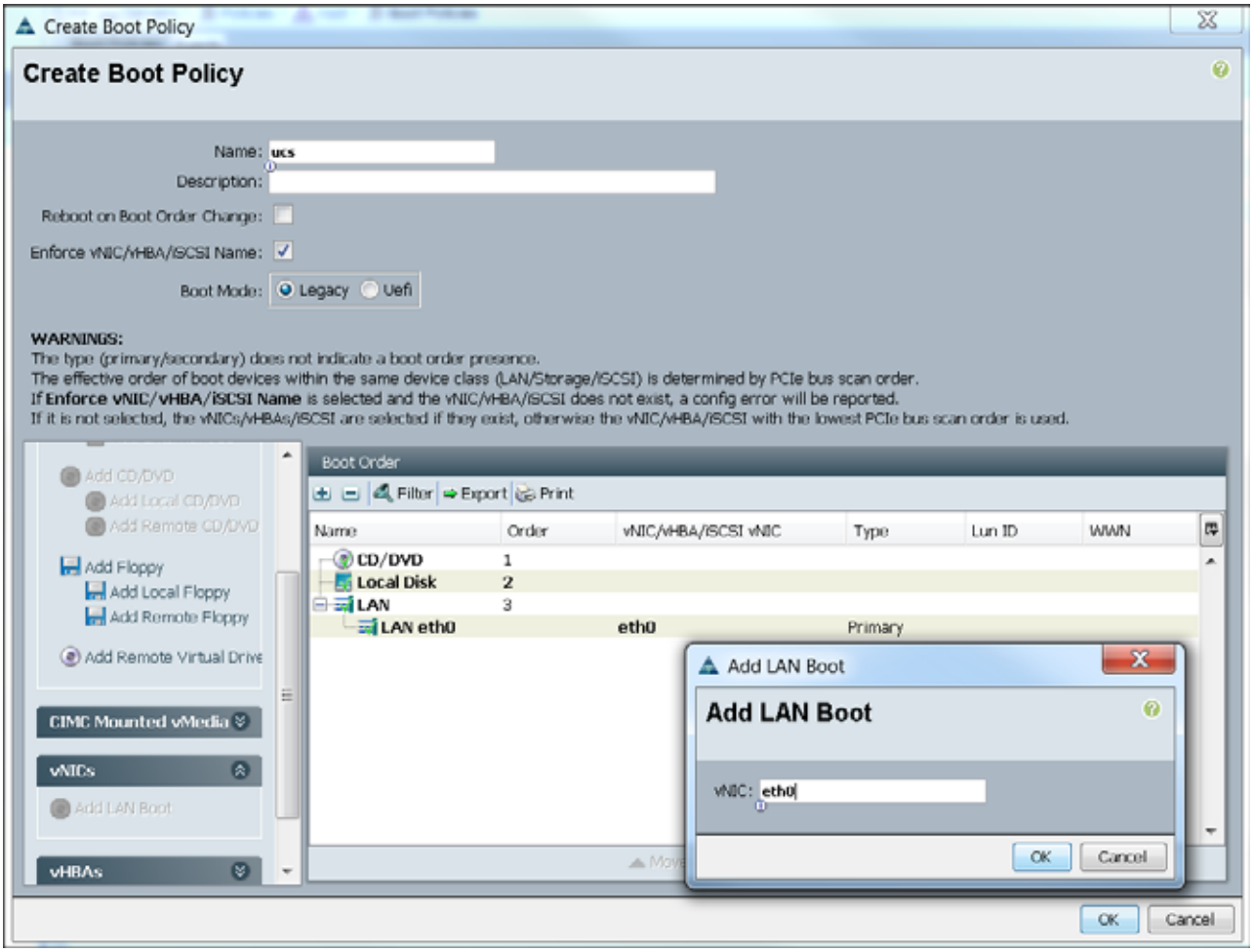
1. Select the **Servers** tab in the left pane in the UCS Manager GUI.
2. Select **Policies > root**.
3. Right-click the **Boot Policies**.
4. Select **Create Boot Policy**.

Figure 34 Creating Boot Policy Part 1



5. Enter ucs as the boot policy name.
6. (Optional) enter a description for the boot policy.
7. Keep the Reboot on Boot Order Change check box unchecked.
8. Keep Enforce vNIC/vHBA/iSCSI Name check box checked.
9. Keep Boot Mode Default (Legacy).
10. Expand **Local Devices > Add CD/DVD** and select **Add Local CD/DVD**.
11. Expand Local Devices and select **Add Local Disk**.
12. Expand vNICs and select Add LAN Boot and enter eth0.
13. Click **OK** to add the Boot Policy.
14. Click **OK**.

Figure 35 Creating Boot Policy Part 2

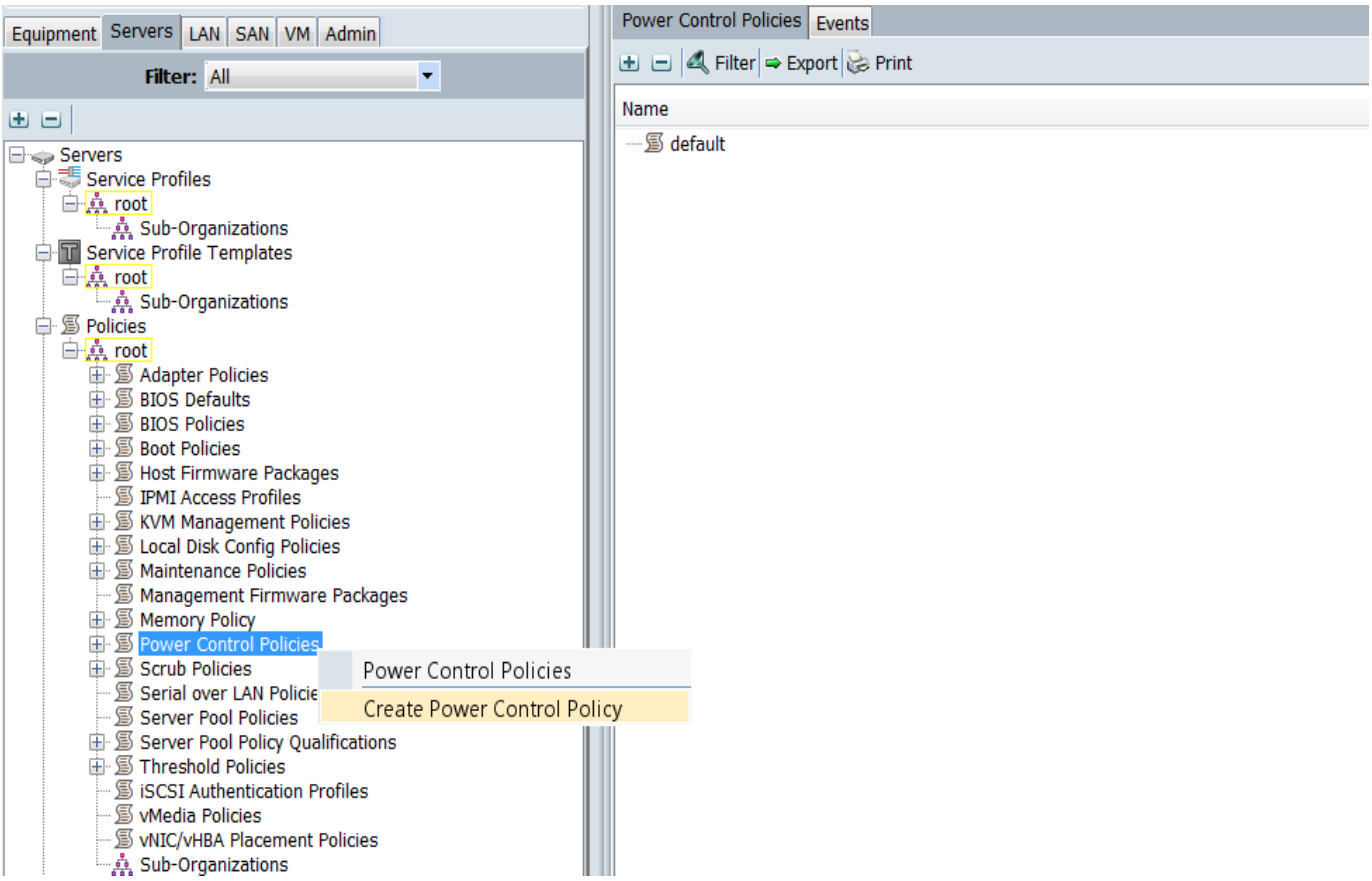


Creating Power Control Policy

Follow these steps to create the Power Control policies within the Cisco UCS Manager GUI:

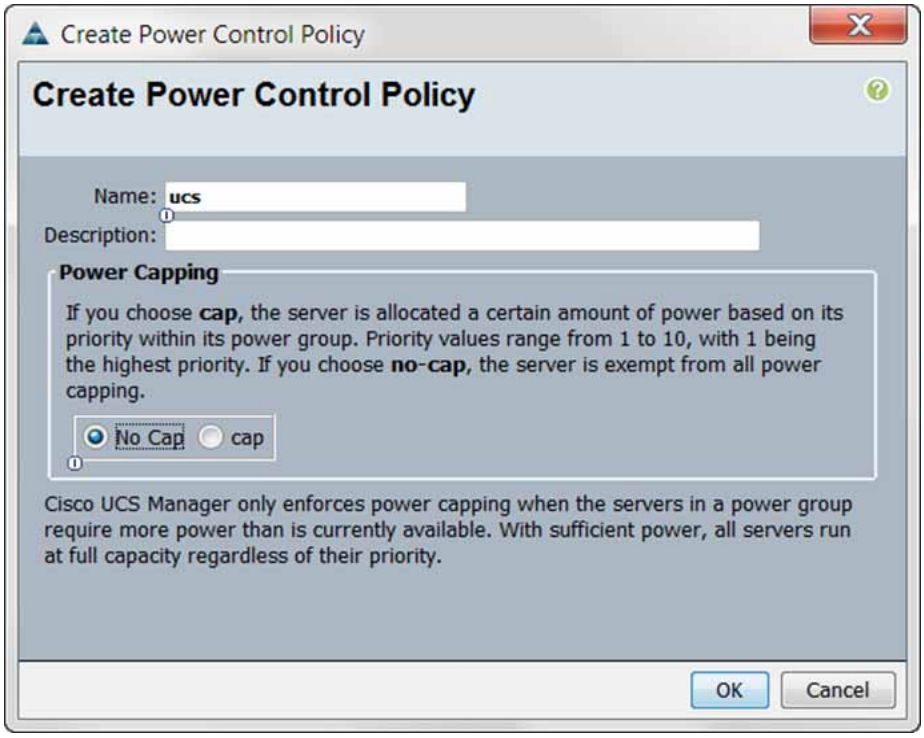
15. Select the **Servers** tab in the left pane in the UCS Manager GUI.
16. Select **Policies > root**.
17. Right-click the **Power Control Policies**.
18. Select **Create Power Control Policy.**

Figure 36 Creating Power Control Policy Part 1



19. Enter ucs as the Power Control policy name.
20. (Optional) enter a description for the boot policy.
21. Select **No cap** for Power Capping selection.
22. Click **OK** to the Power Control Policy.
23. Click **OK**.

Figure 37 Creating Power Control Policy Part 2

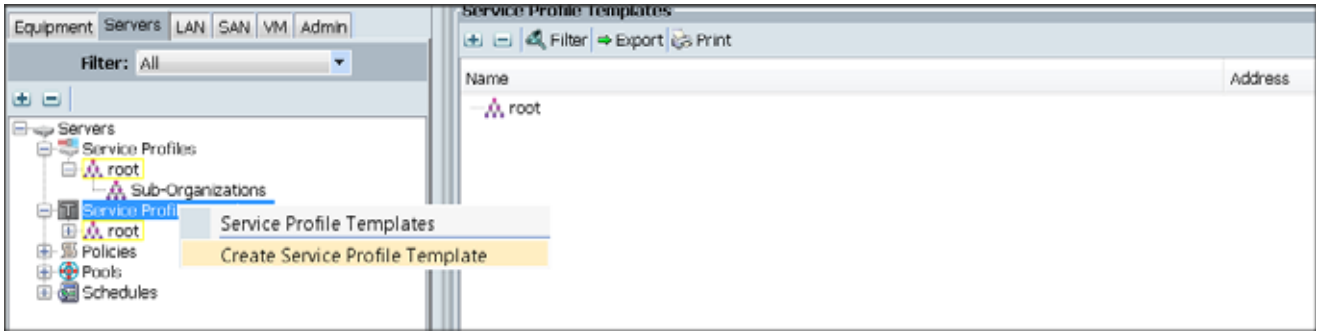


Creating Service Profile Template

To create a service profile template, follow these steps:

1. Select the **Servers** tab in the left pane in the UCS Manager GUI.
2. Right-click **Service Profile Templates**.
3. Select **Create Service Profile Template**.

Figure 38 Creating Service Profile Template



4. The Create Service Profile Template window appears.

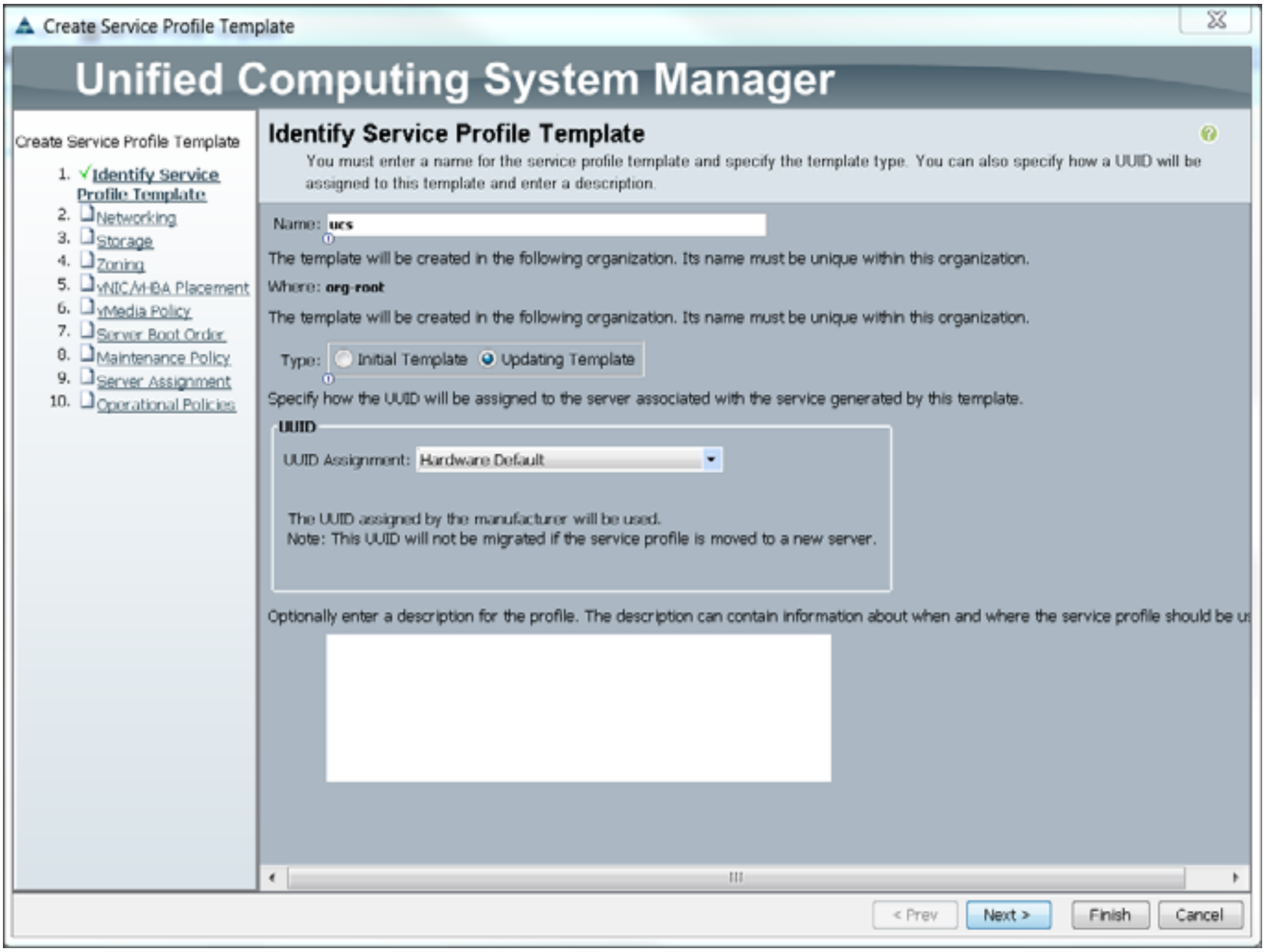
   These steps below provide a detailed configuration procedure to identify the service profile template:

   a. Name the service profile template as ucs. Click the **Updating Template** radio button.

b. In the UUID section, select **Hardware Default** as the UUID pool.

c. Click **Next** to continue to the next section.

Figure 39 Identify Service Profile Template



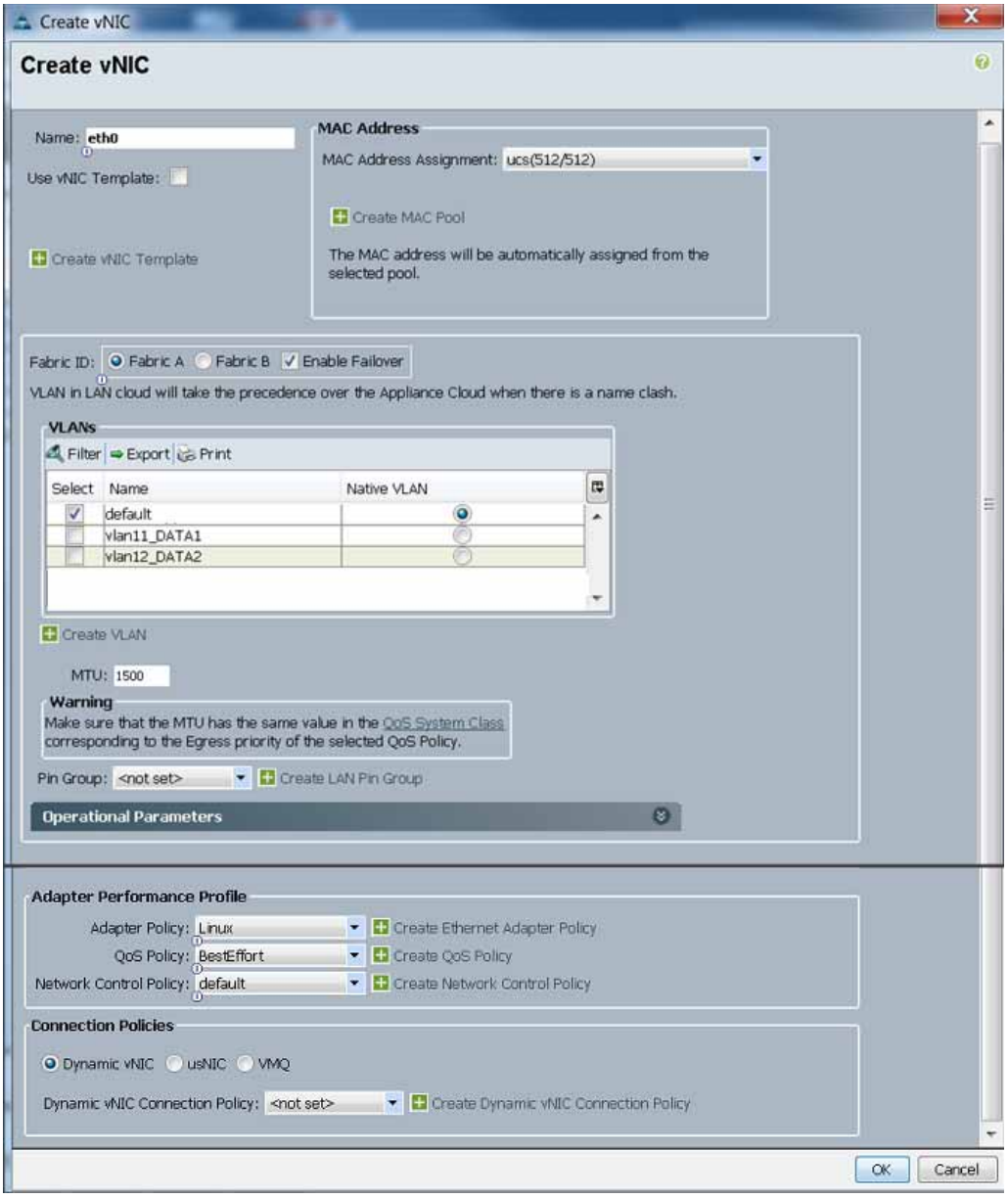
Configuring Network Settings for the Template

1. Keep the Dynamic vNIC Connection Policy field at the default.
2. Click the **Expert** radio button for the option, **how would you like to configure LAN connectivity?**
3. Click **Add** to add a vNIC to the template.

Figure 40 Configuring Network Settings for the Template

4. The Create vNIC window displays. Name the vNIC as eth0.
5. Select UCS in the Mac Address Assignment pool.
6. Click the **Fabric A** radio button and Check the **Enable failover** check box for the Fabric ID.
7. Check the **default** check box for VLANs and click the **Native VLAN** radio button.
8. Select MTU size as 1500.
9. Select adapter policy as Linux.
10. Select QoS Policy as BestEffort.
11. Keep the Network Control Policy as Default.
12. Keep the Connection Policies as Dynamic vNIC.
13. Keep the Dynamic vNIC Connection Policy as <not set>.
14. Click **OK**.

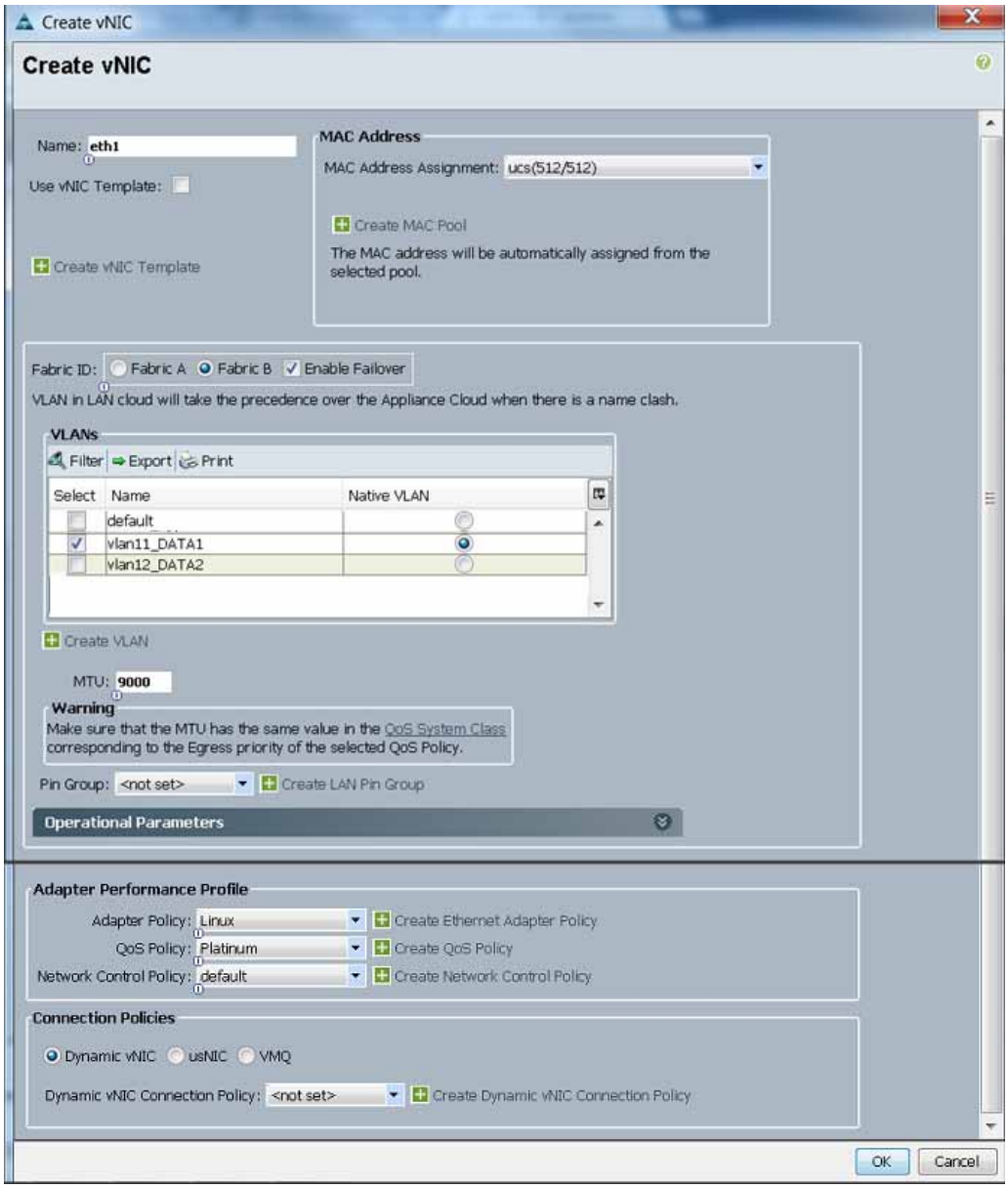
Figure 41 Configuring vNIC eth0



15. Click **Add** to add a vNIC to the template.
16. The Create vNIC window appears. Name the vNIC eth1.
17. Select ucs in the Mac Address Assignment pool.
18. Click the **Fabric B** radio button and Check the **Enable failover** check box for the Fabric ID.
19. Check the **vlan11_DATA1** check box for VLANs, and click the **Native VLAN** radio button.
20. Select MTU size as 9000.
21. Select adapter policy as Linux.
22. Select QoS Policy as Platinum.
23. Keep the Network Control Policy as Default.

24. Keep the Connection Policies as Dynamic vNIC.
25. Keep the Dynamic vNIC Connection Policy as <not set>.
26. Click **OK**.

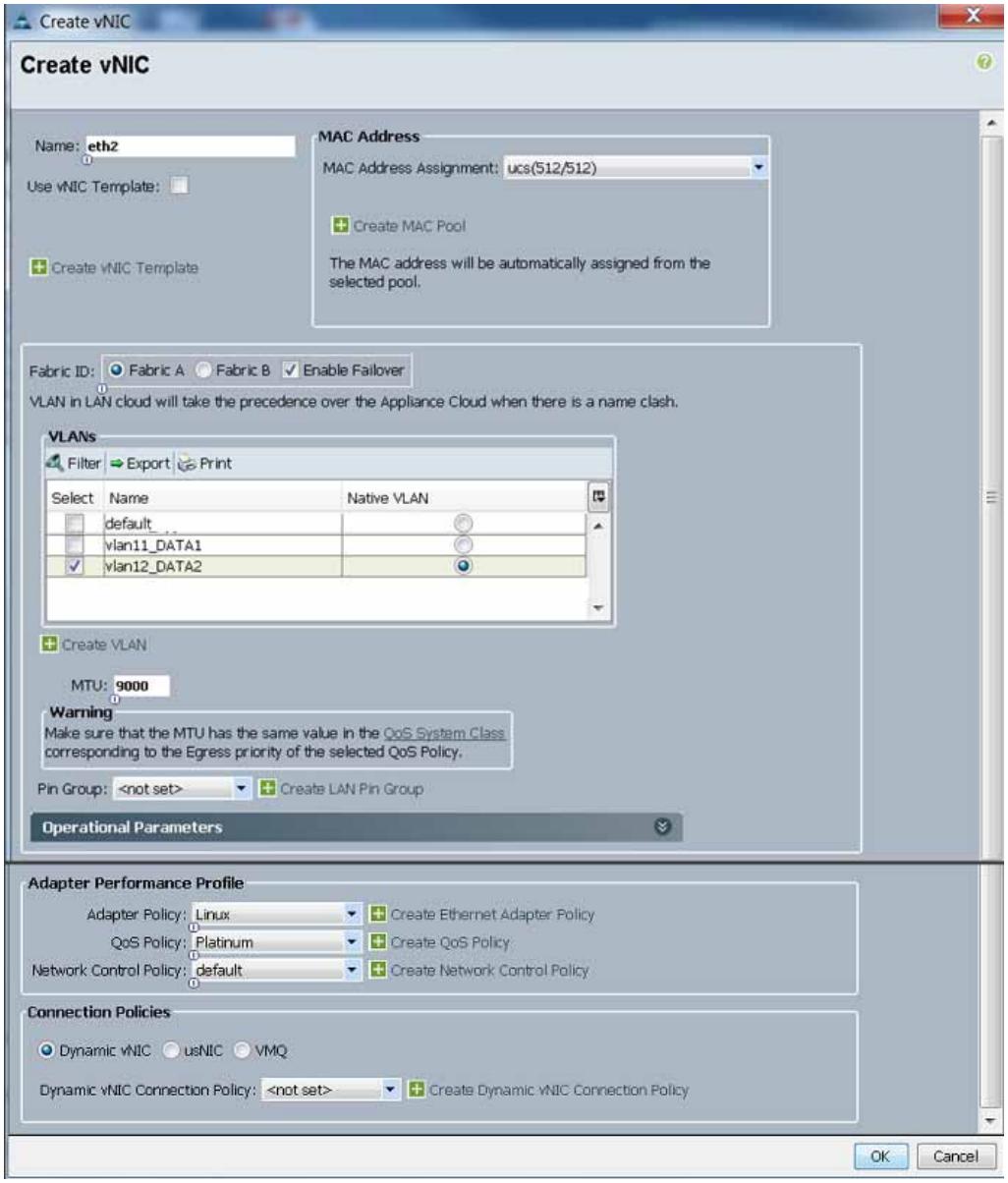
Figure 42 Configuring vNIC eth1



27. Click **Add** to add a vNIC to the template.
28. The Create vNIC window appears. Name the vNIC eth2.
29. Select ucs in the Mac Address Assignment pool.
30. Click the **Fabric A** radio button, and then Check the **Enable failover** check box for the Fabric ID.
31. Check the **vlan12_DATA2** check box for VLANs, and then click the **Native VLAN** radio button.
32. Select MTU size as 9000.

33. Select adapter policy as Linux.
34. Select QoS Policy as Platinum.
35. Keep the Network Control Policy as Default.
36. Keep the Connection Policies as Dynamic vNIC.
37. Keep the Dynamic vNIC Connection Policy as <not set>.
38. Click **OK**.

Figure 43 Configuring vNIC eth2

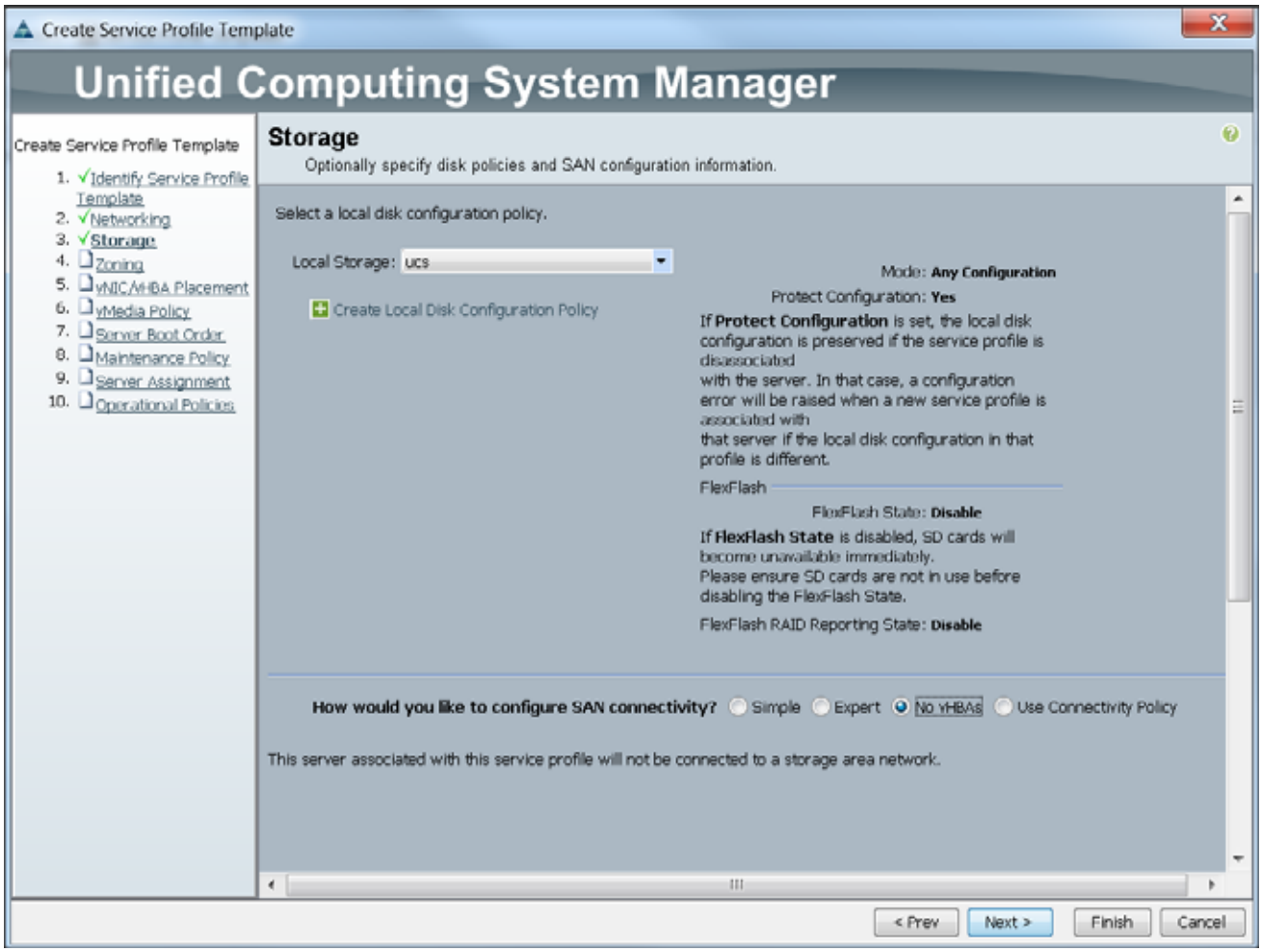


Configuring Storage Policy for the Template

Follow these steps to configure storage policies:

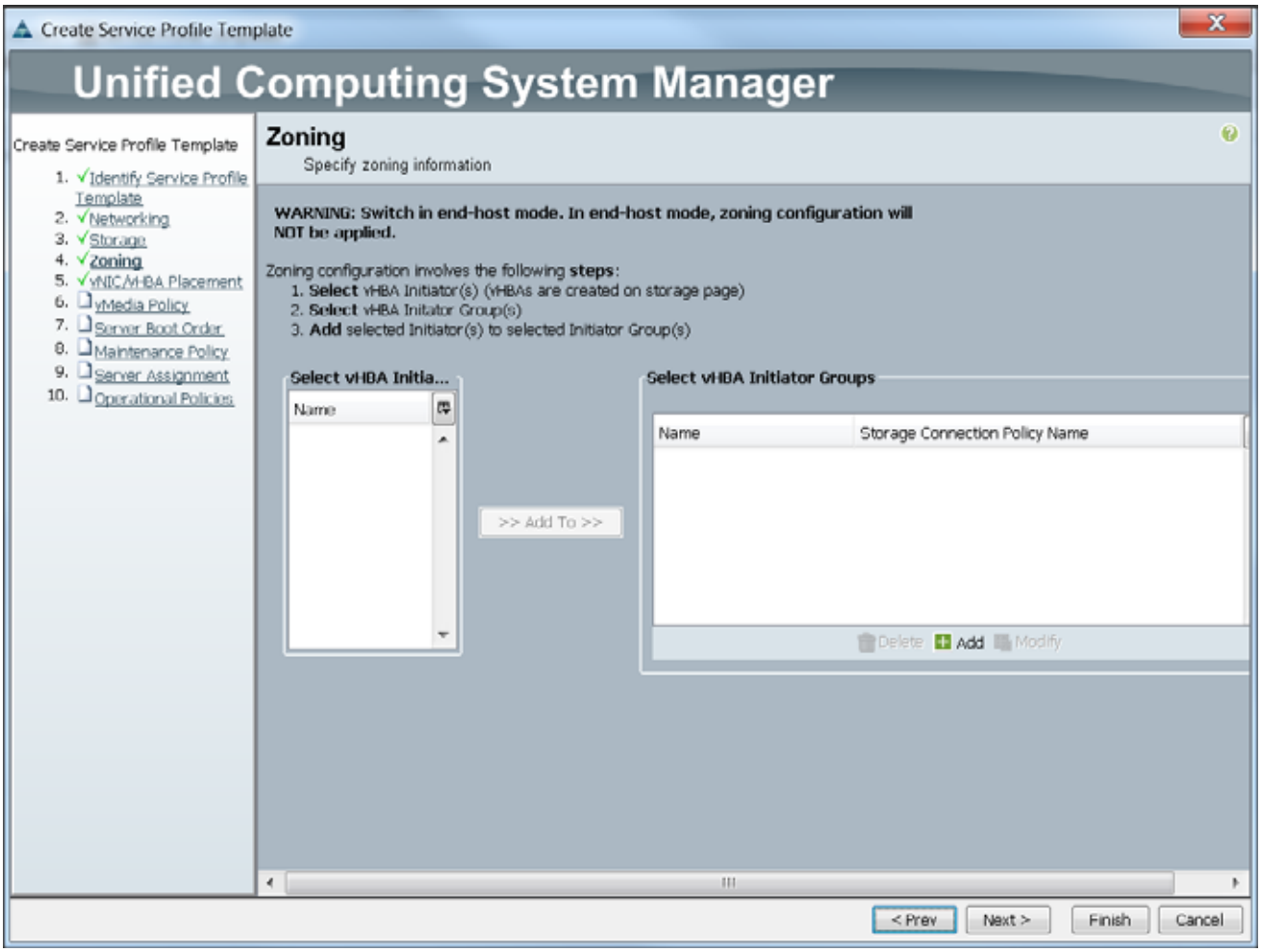
1. Select ucs for the local disk configuration policy.
2. Click the **No vHBAs** radio button for the option, **How would you like to configure SAN connectivity?**
3. Click **Next** to continue to the next section.

Figure 44 Configuring Storage Settings



4. Click **Next** once the zoning window appears to go to the next section.

Figure 45 Configure Zoning

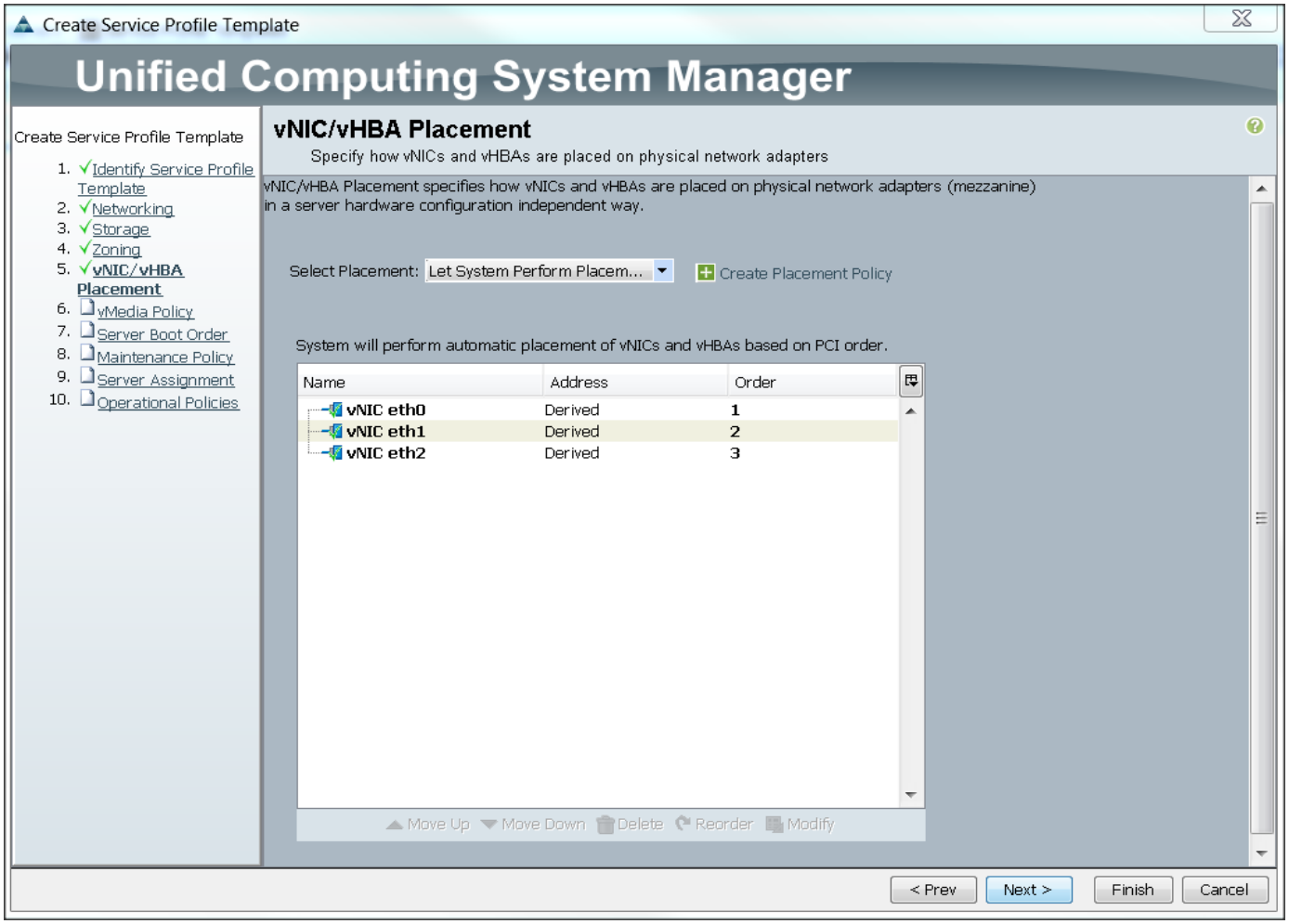


Configuring vNIC/vHBA Placement for the Template

Follow these steps to configure vNIC/vHBA placement policy:

1. Select the Default Placement Policy option for the Select Placement field.
2. Select eth0, eth1 and eth2 assign the vNICs in the following order:
   a. eth0
   b. eth1
   c. eth2
3. Review to make sure that all of the vNICs were assigned in the appropriate order.
4. Click **Next** to continue to the next section.

Figure 46 vNIC/vHBA Placement



Configuring vMedia Policy for the Template

1. Click **Next** once the vMedia Policy window appears to go to the next section.

Figure 47 UCSM vMedia Policy Window

Configuring Server Boot Order for the Template

Follow these steps to set the boot order for servers:

1. Select ucs in the Boot Policy name field.
2. Review to make sure that all of the boot devices were created and identified.
3. Verify that the boot devices are in the correct boot sequence.
4. Click **OK**.
5. Click **Next** to continue to the next section.

Figure 48 Creating Boot Policy

In the Maintenance Policy window, follow these steps to apply the maintenance policy:

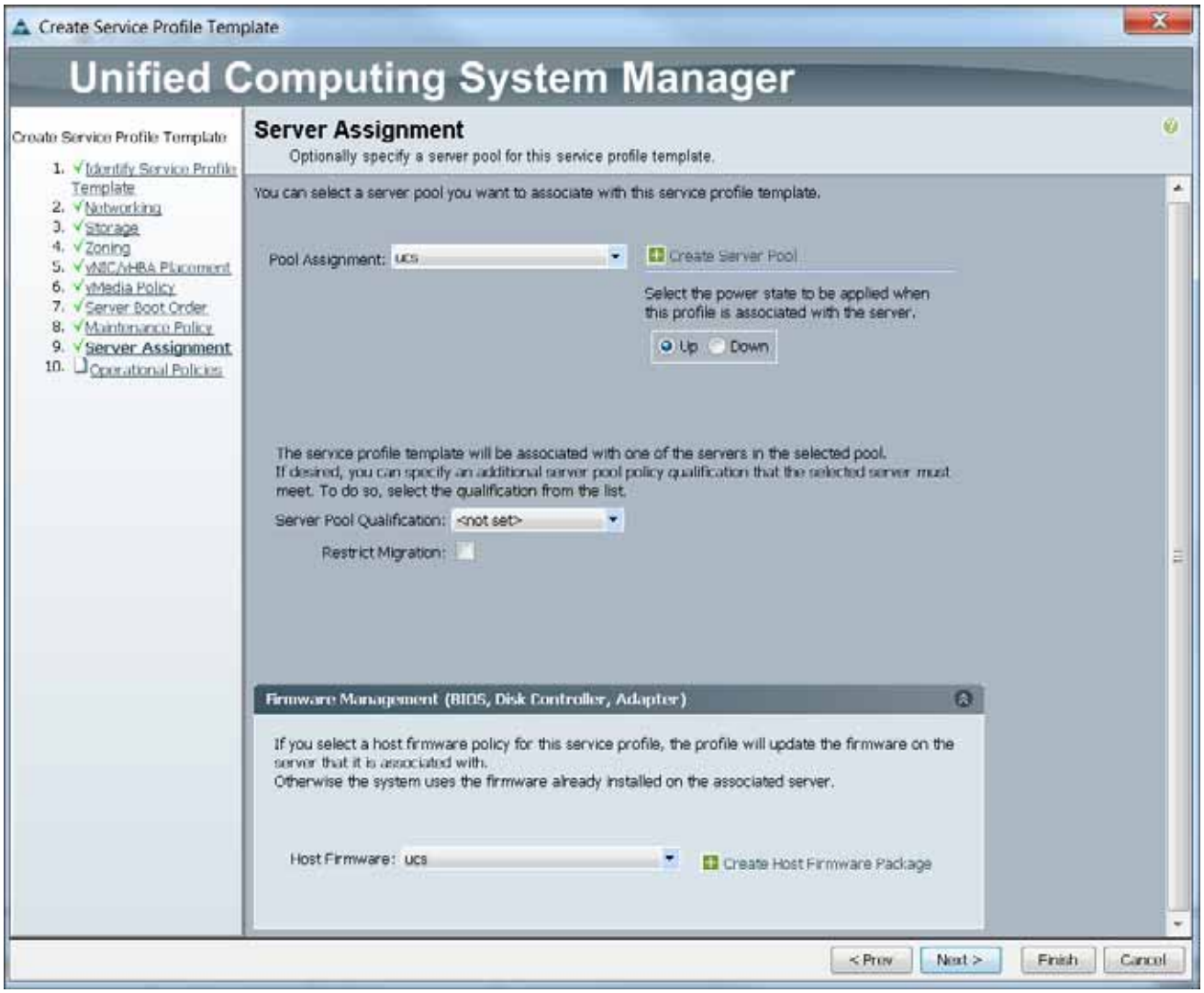
1. Keep the Maintenance policy at no policy used by default.
2. Click **Next** to continue to the next section.

Configuring Server Assignment for the Template

In the Server Assignment window, follow these steps to assign the servers to the pool:

3. Select ucs for the Pool Assignment field.
4. Keep the Server Pool Qualification field at default.
5. Select ucs in Host Firmware Package.

Figure 49 Server Assignment

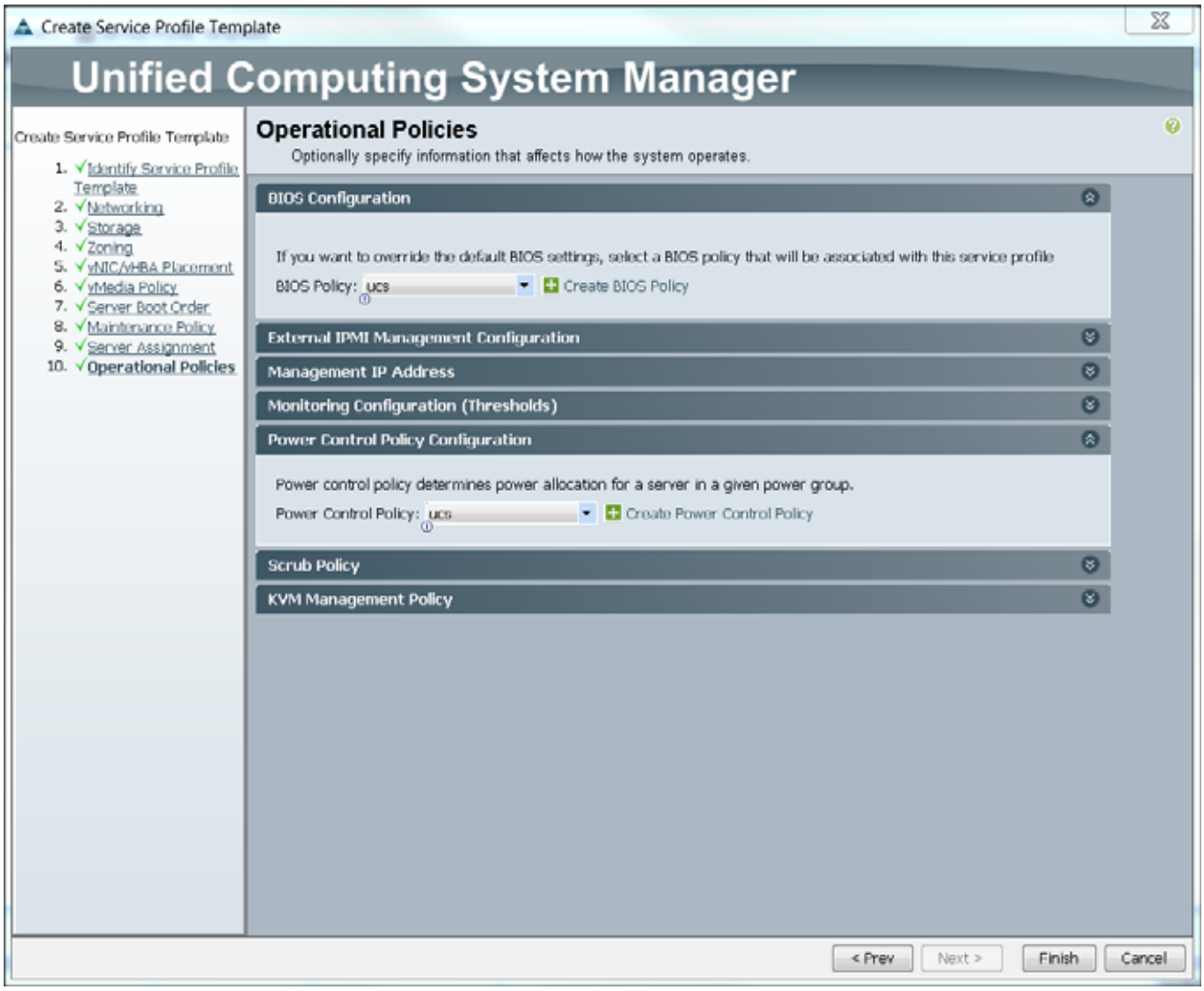


Configuring Operational Policies for the Template

In the Operational Policies Window, follow these steps:

6. Select ucs in the BIOS Policy field.
7. Select ucs in the Power Control Policy field.
8. Click **Finish** to create the Service Profile template.
9. Click **OK** in the pop-up window to proceed.

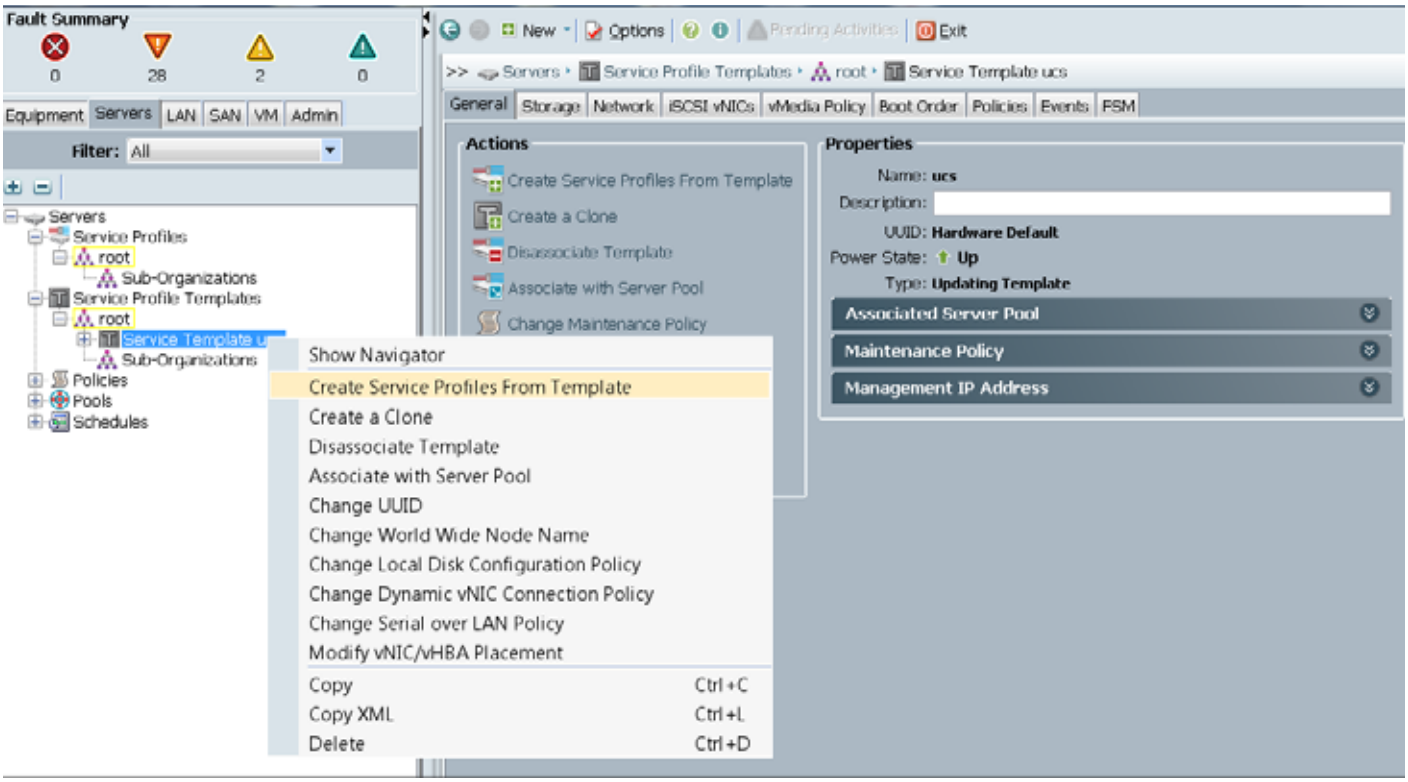
Figure 50 Selecting BIOS and Power Control Policy



Select the **Servers** tab in the left pane of the UCS Manager GUI.

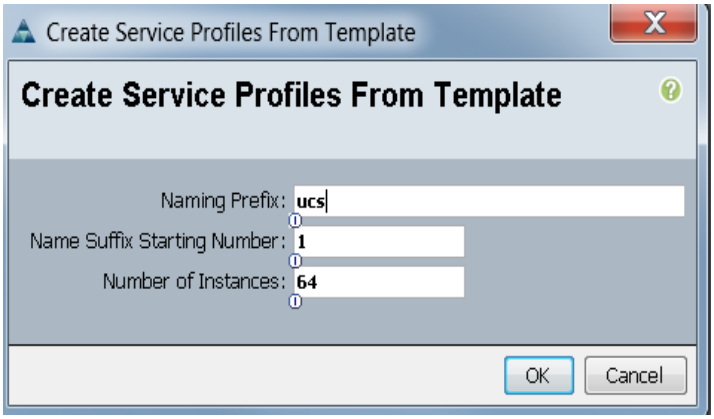
1. Go to Service Profile **Templates > root**.
2. Right-click Service Profile Templates ucs.
3. Select Create Service Profiles From Template.

Figure 51 Creating Service Profiles from Template



4. The Create Service Profile from Template window appears.

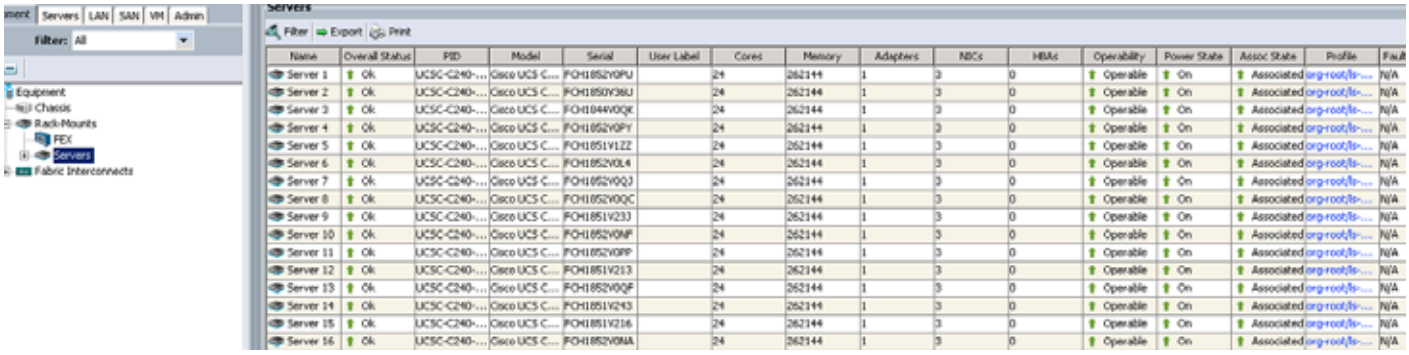
Figure 52 Selecting Name and Total number of Service Profiles



Association of the Service Profiles will take place automatically.

The Final Cisco UCS Manager window is shown in Figure 46.

Figure 53 UCS Manager showing all Nodes



Installing Red Hat Enterprise Linux 6.5 using software RAID on C240 M4 Systems

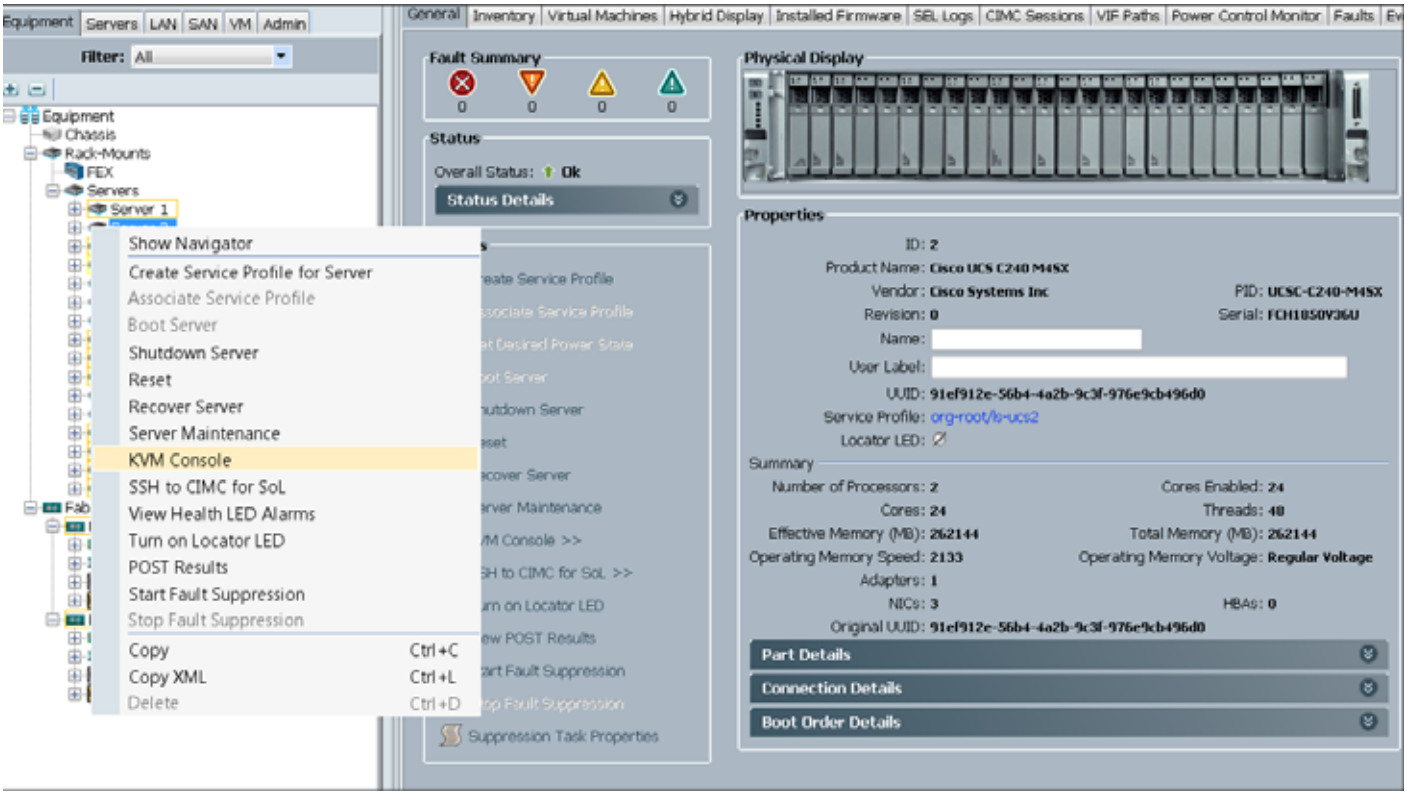
The following section provides detailed procedures for installing Red Hat Enterprise Linux 6.5 using Software RAID (OS based Mirroring) on Cisco UCS C240 M4 servers.

There are multiple methods to install Red Hat Linux operating system. The installation procedure described in this deployment guide uses KVM console and virtual media from Cisco UCS Manager.

Note This requires RHEL 6.5 DVD/ISO for the installation.

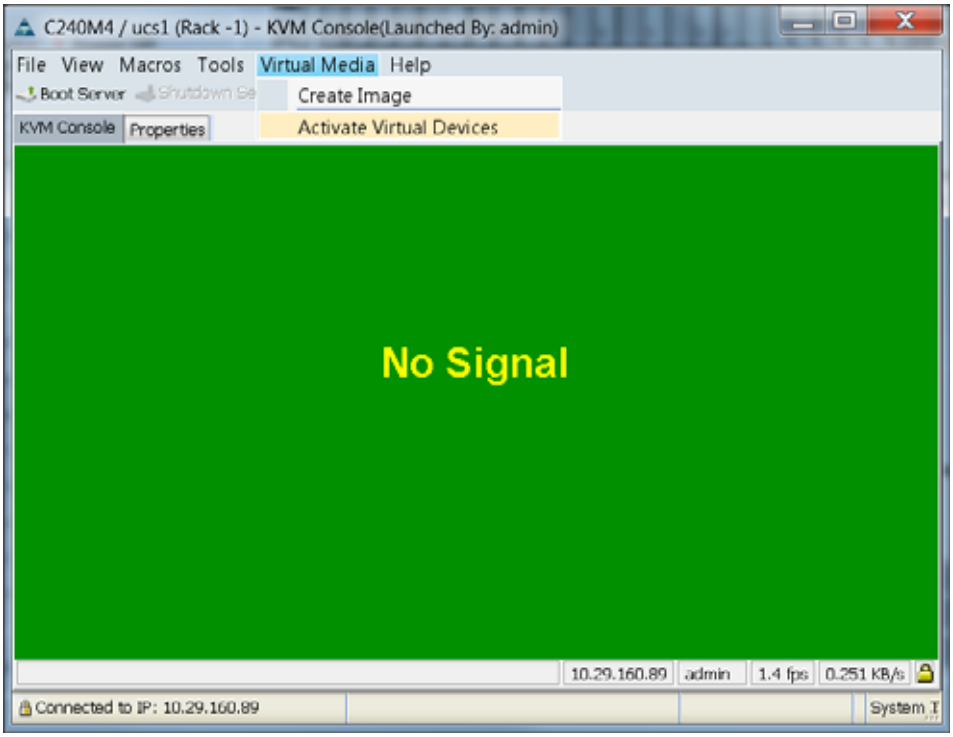
1. Log in to the Cisco UCS 6296 Fabric Interconnect and launch the Cisco UCS Manager application.
2. Select the **Equipment tab**.
3. In the navigation pane expand Rack-Mounts and then Servers.
4. Right click on the server and select KVM Console.

Figure 54 Selecting KVM Console Option



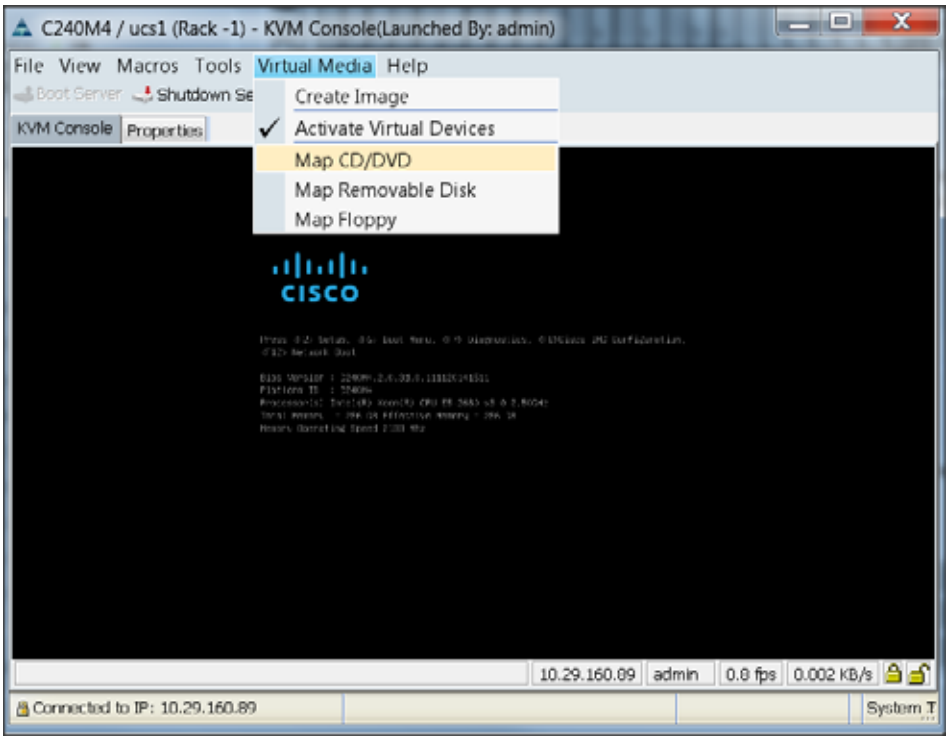
5. In the KVM window, select the **Virtual Media** tab.
6. Click the **Activate Virtual Devices** from the **Virtual Media** tab.

Figure 55 KVM Console



7. In the **KVM** window, select the **Virtual Media** tab and Click the **Map CD/DVD**.

Figure 56 KVM Console

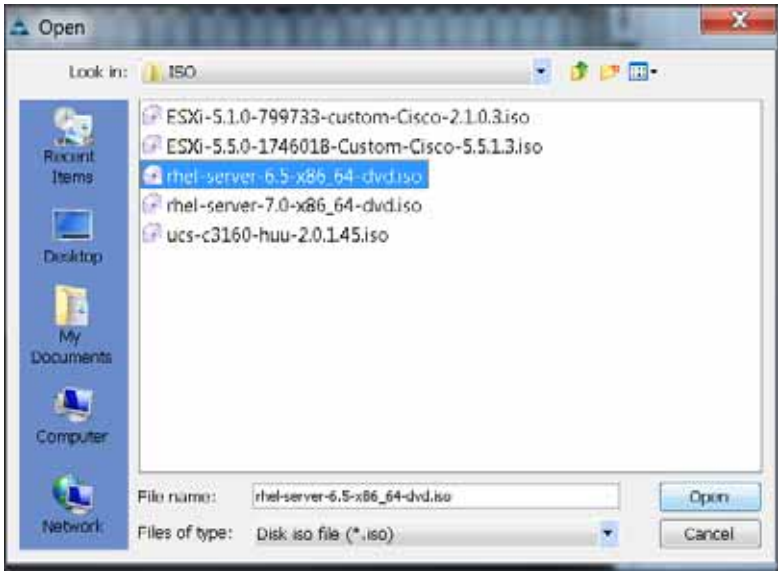


8. Browse to the Red Hat Enterprise Linux Server 6.5 installer ISO image file.

Note The Red Hat Enterprise Linux 6.5 DVD is assumed to be on the client machine.

9. Click **Open** to add the image to the list of virtual media.

Figure 57 Browse to Red Hat Enterprise Linux ISO Image

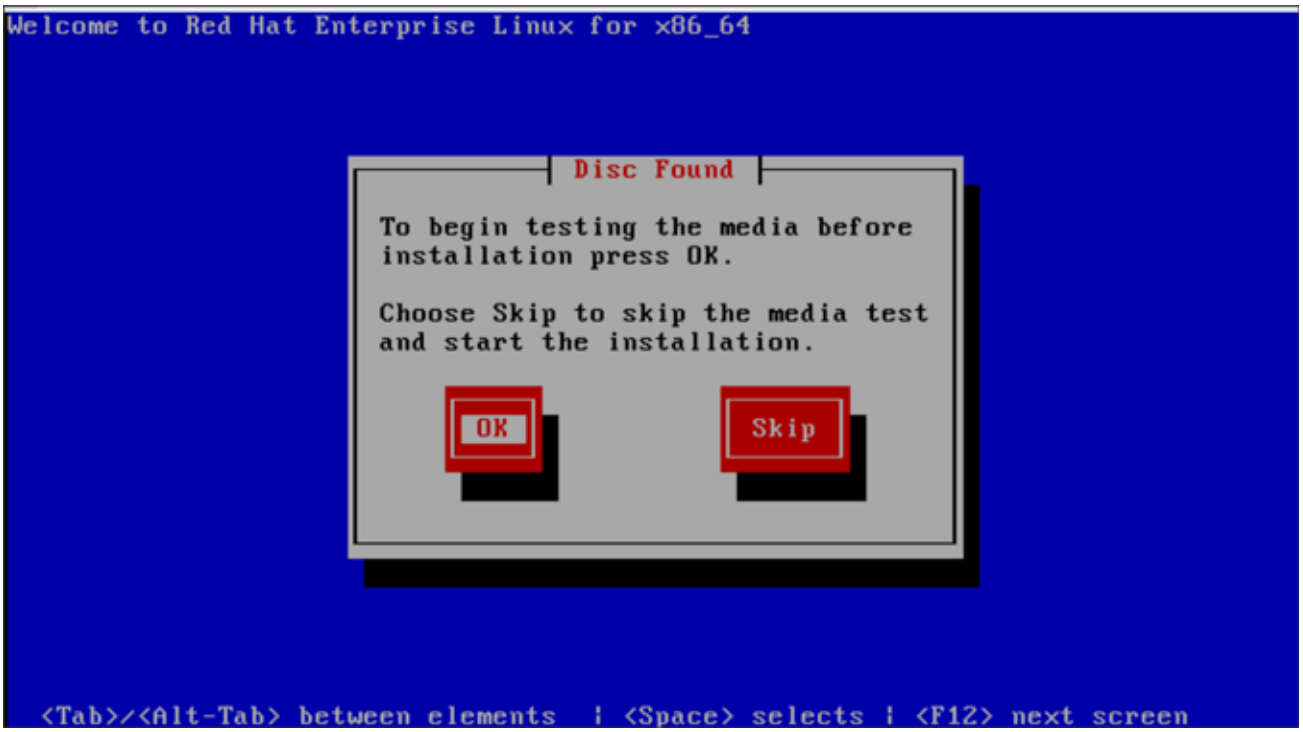


10. In the KVM window, select the **KVM** tab to monitor during boot.
11. In the KVM window, select the **Macros > Static Macros > Ctrl-Alt-Del** button in the upper left corner.
12. Click **OK**.
13. Click **OK** to reboot the system.
14. On reboot, the machine detects the presence of the Red Hat Enterprise Linux Server 6.5 install media.
15. Select the **Install or Upgrade an Existing System**

Figure 58 RHEL Installation

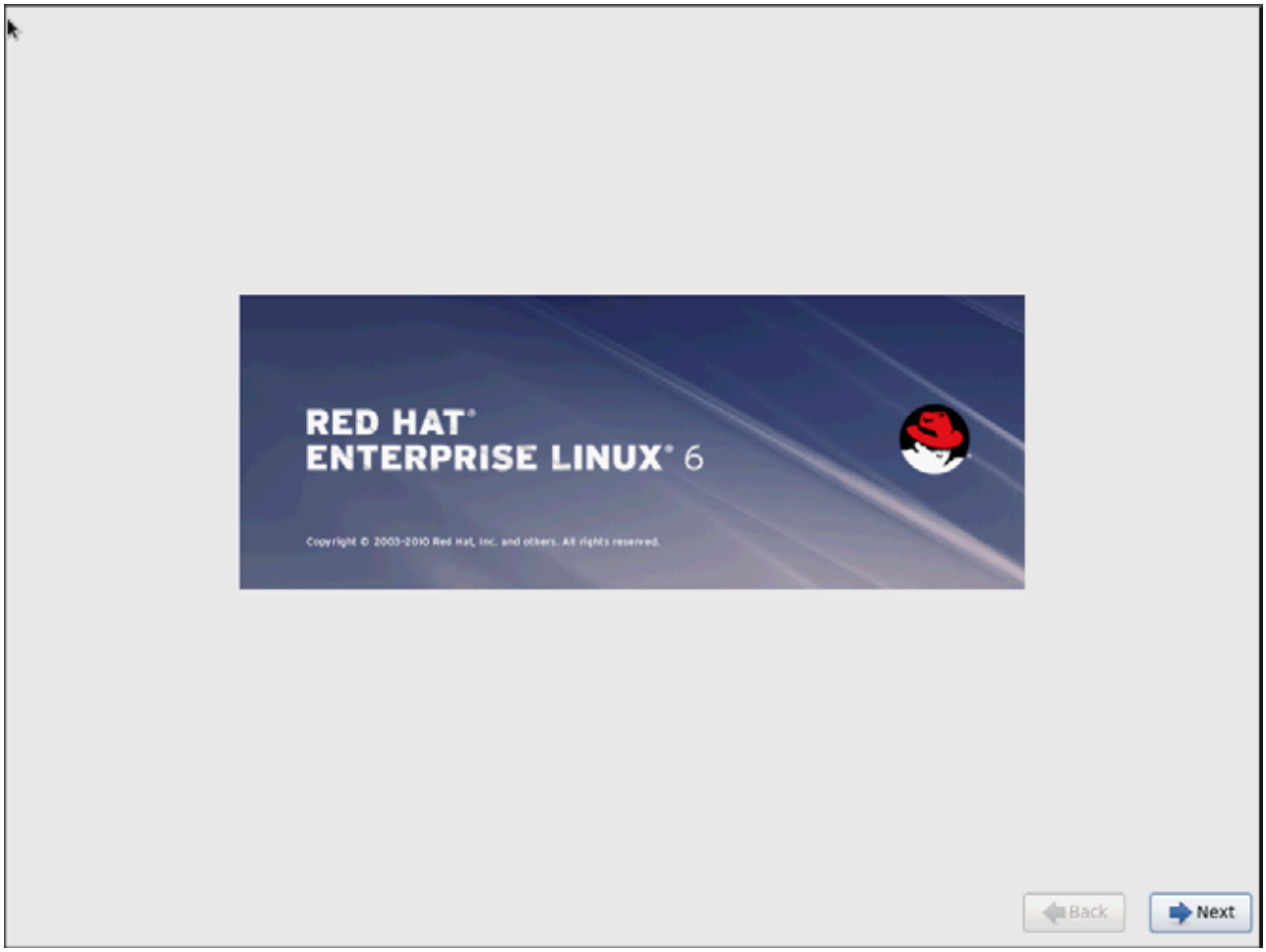
16. Skip the Media test and start the installation.

Figure 59 RHEL Installation: Media Test



17. Click **Next**

Figure 60 RHEL Installation: Installation Wizard



18. Select language of installation, and then Click **Next**

Figure 61 RHEL Installation: Language Selection

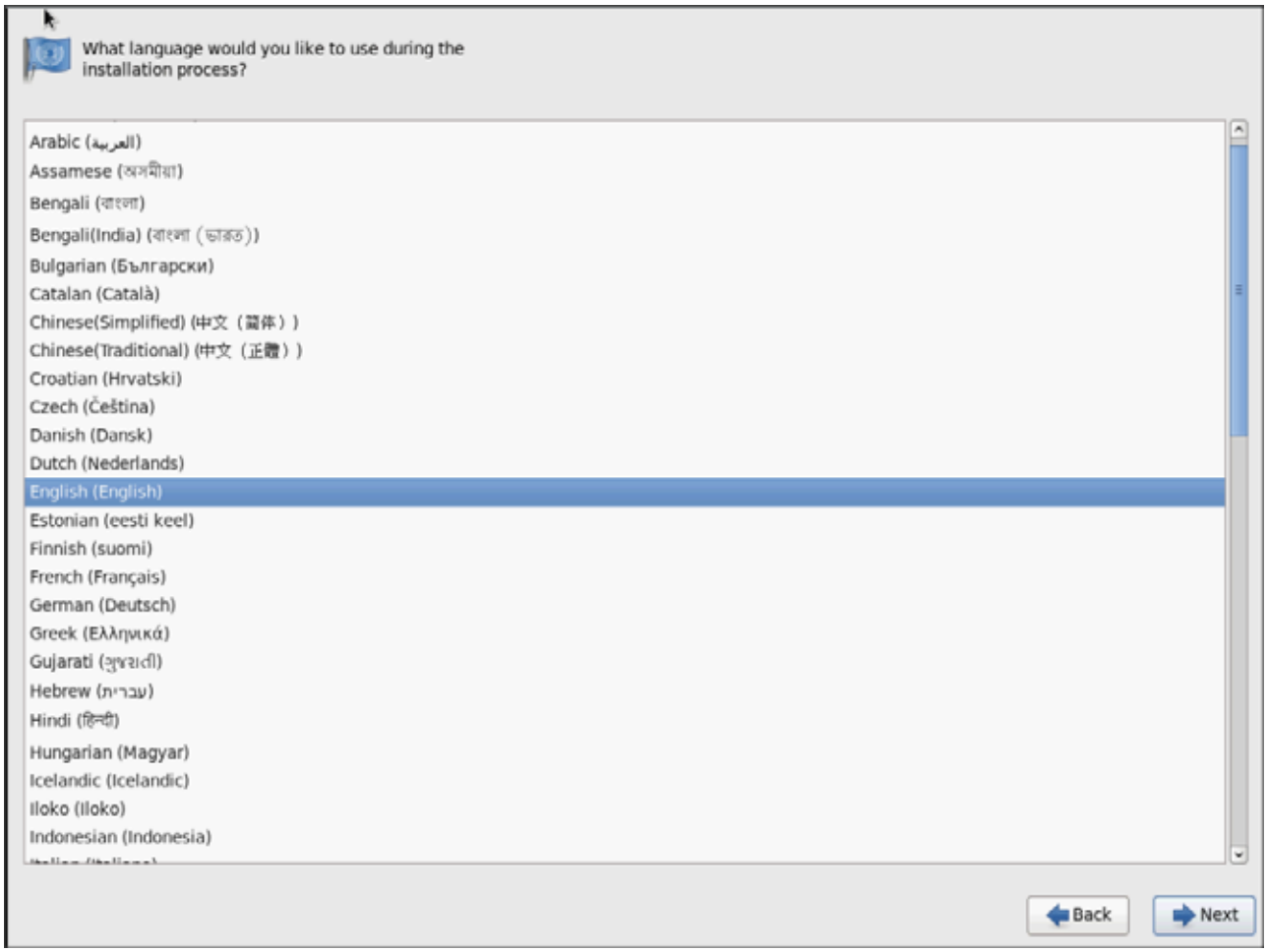
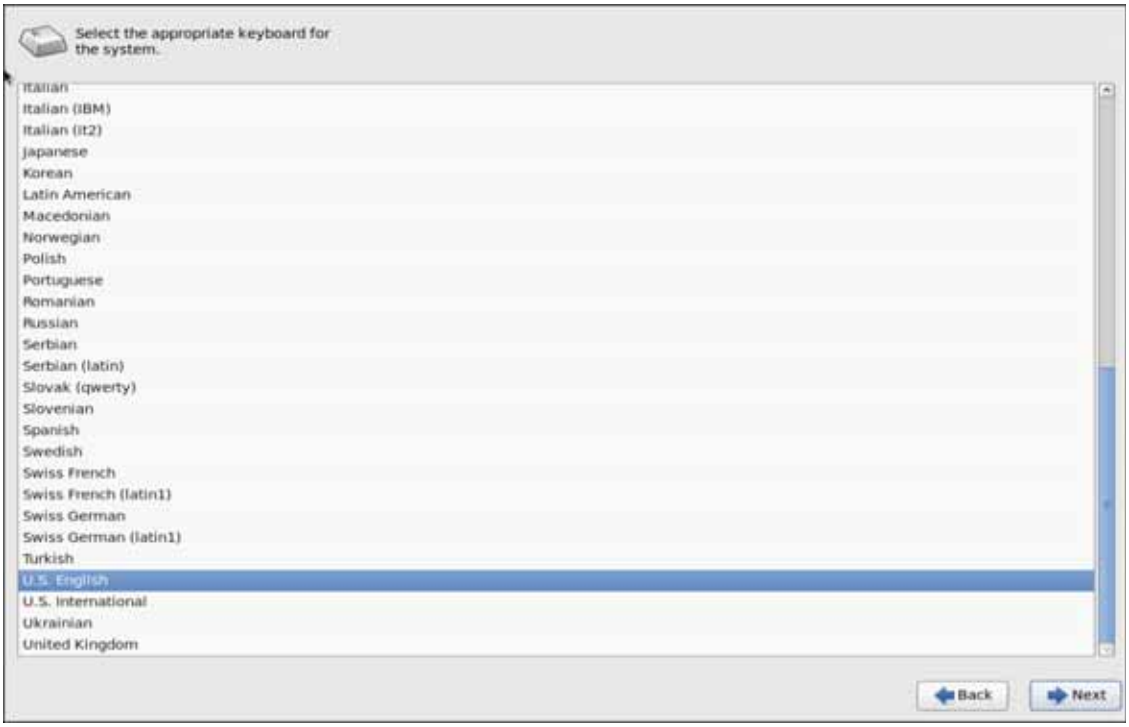
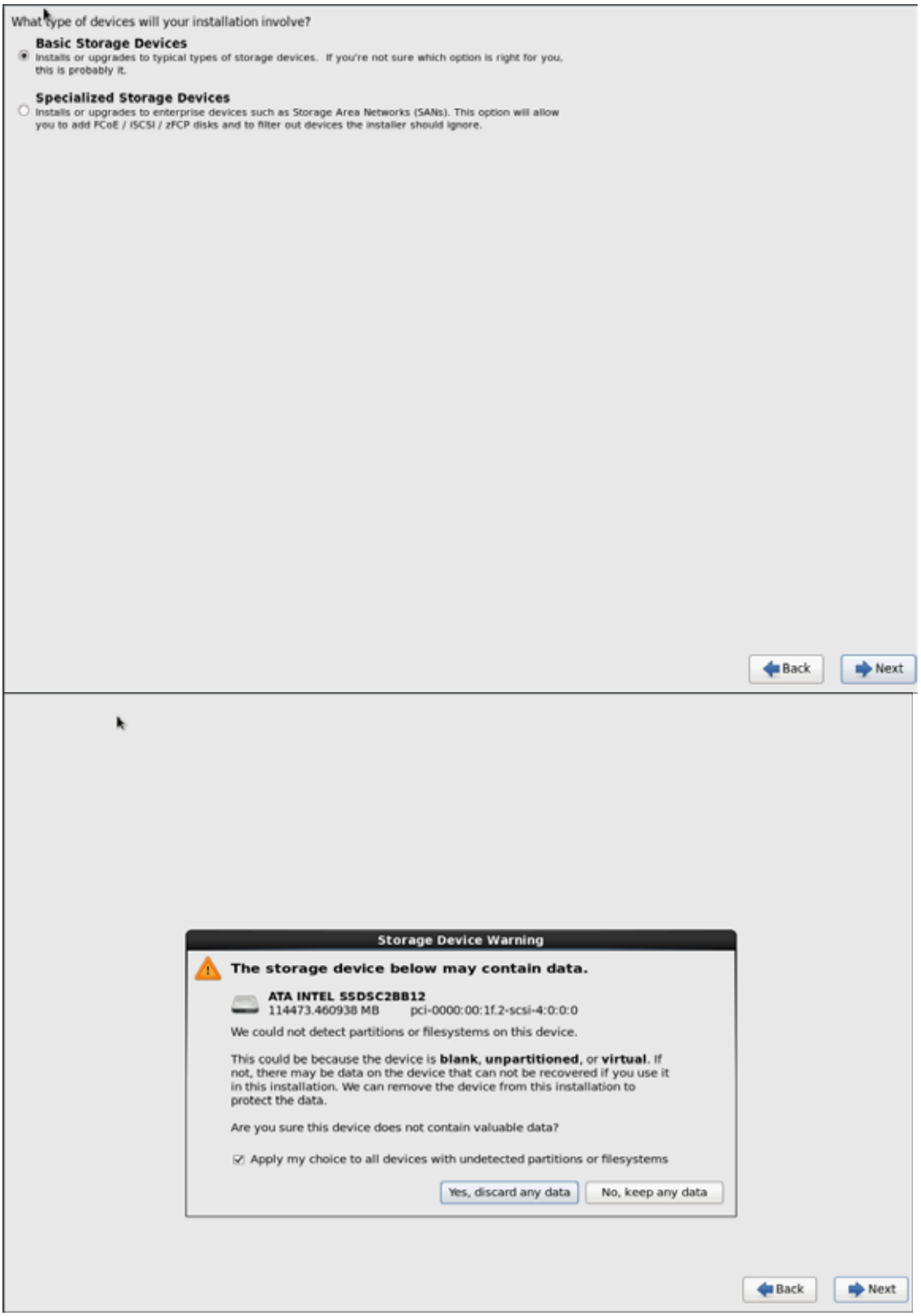


Figure 62 RHEL Installation: Language Selection



19. Select Basic Storage Devices and Click **Next**

Figure 63 RHEL Installation: Installation Type



20. Provide hostname and configure Network for the host.

Figure 64 RHEL Installation: Provide Host Name

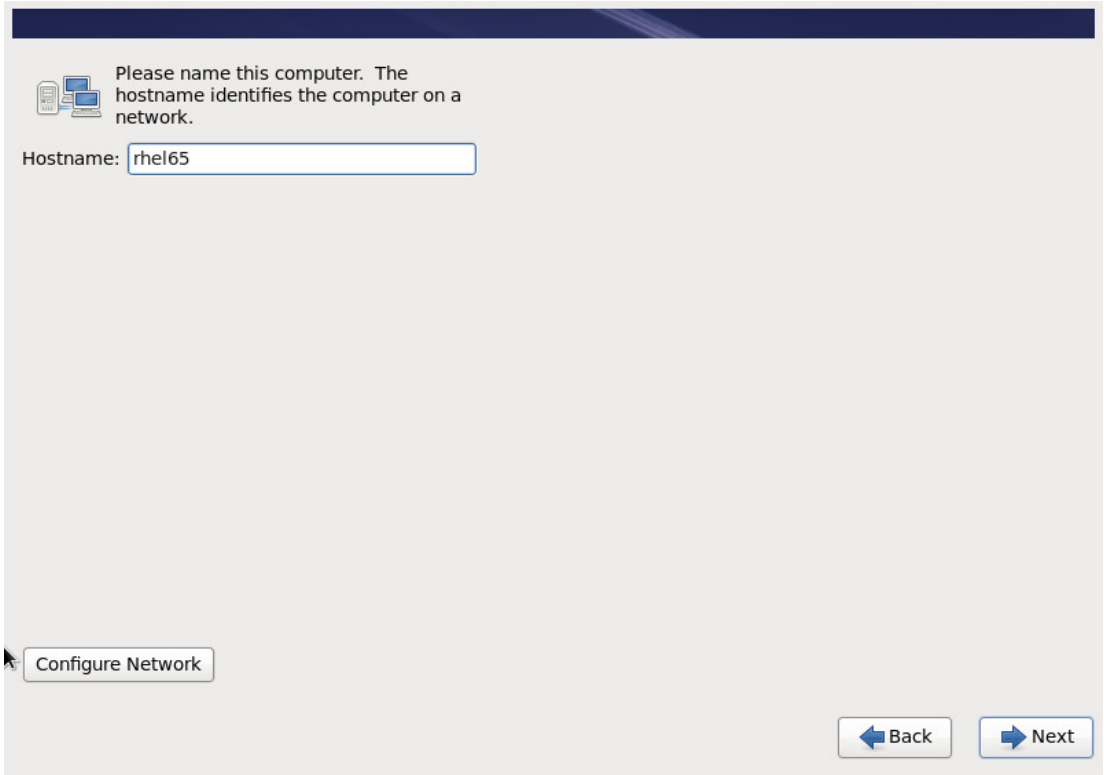


Figure 65 RHEL Installation: IPV4 Setting for eth0

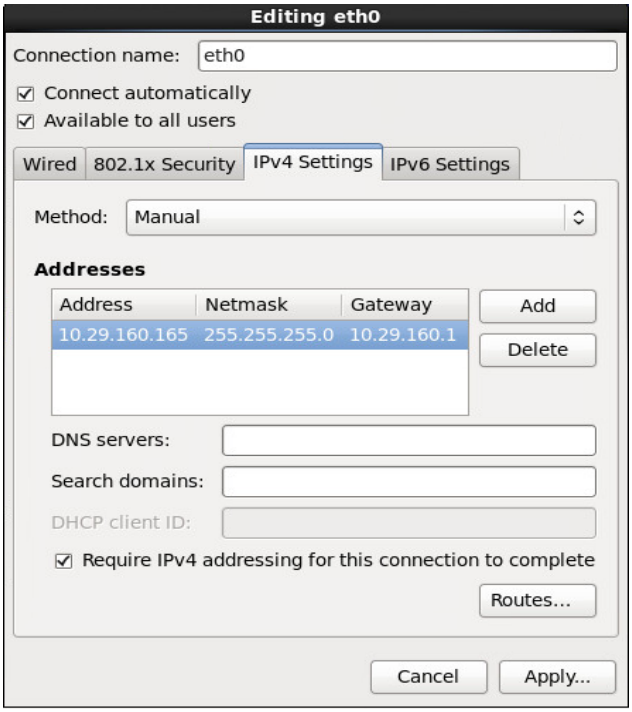


Figure 66 RHEL Installation: IPV4 Setting for eth1

Figure 67 RHEL Installation: Selecting Location

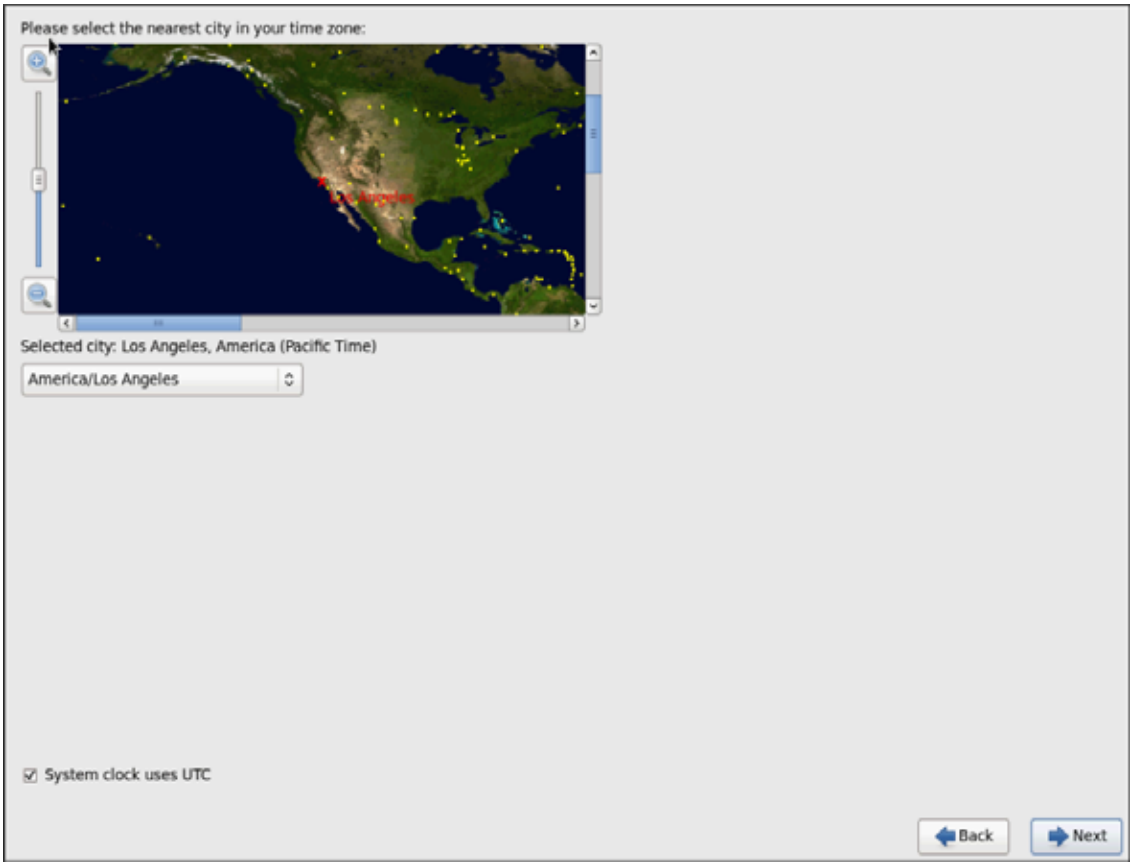
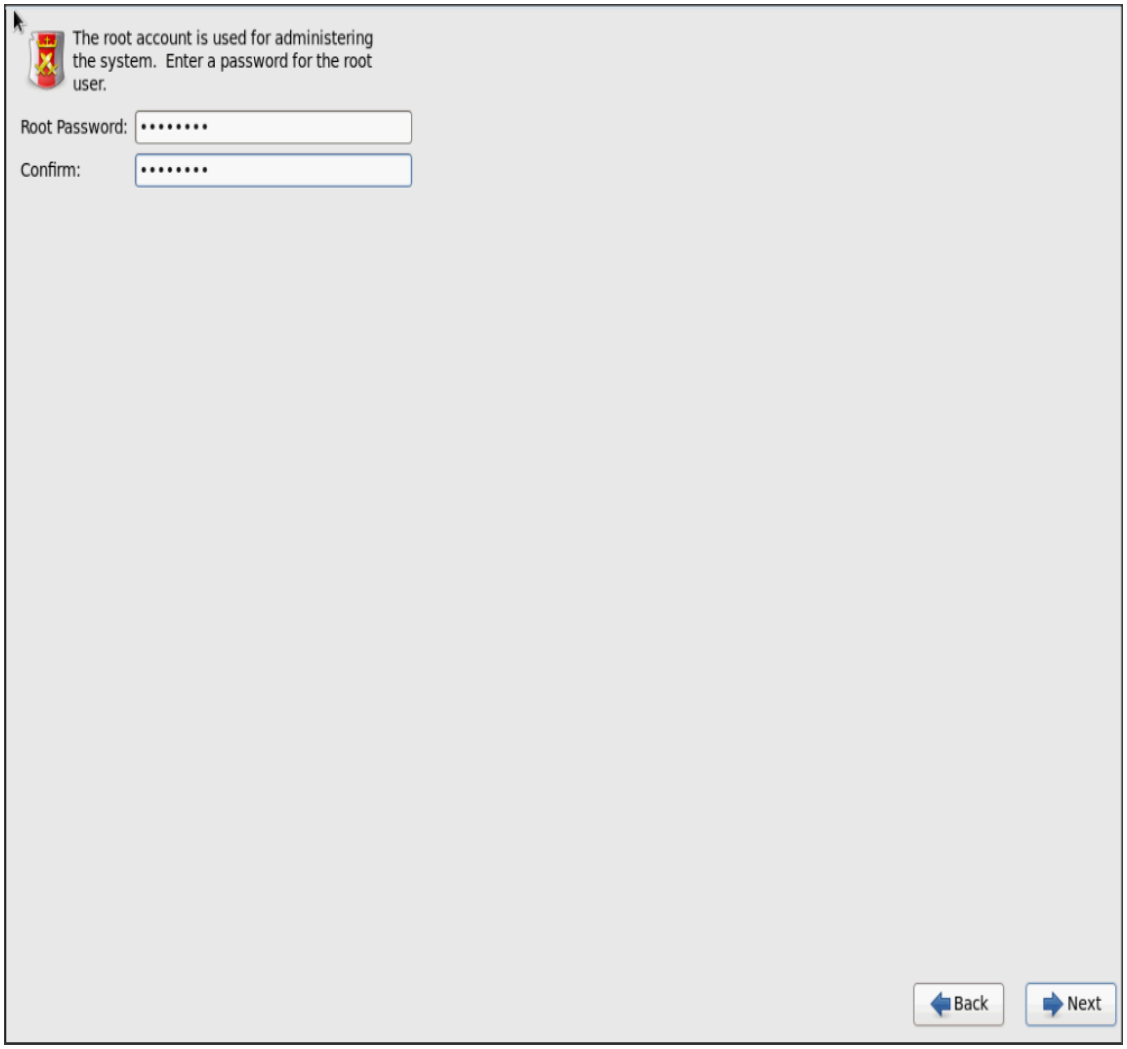
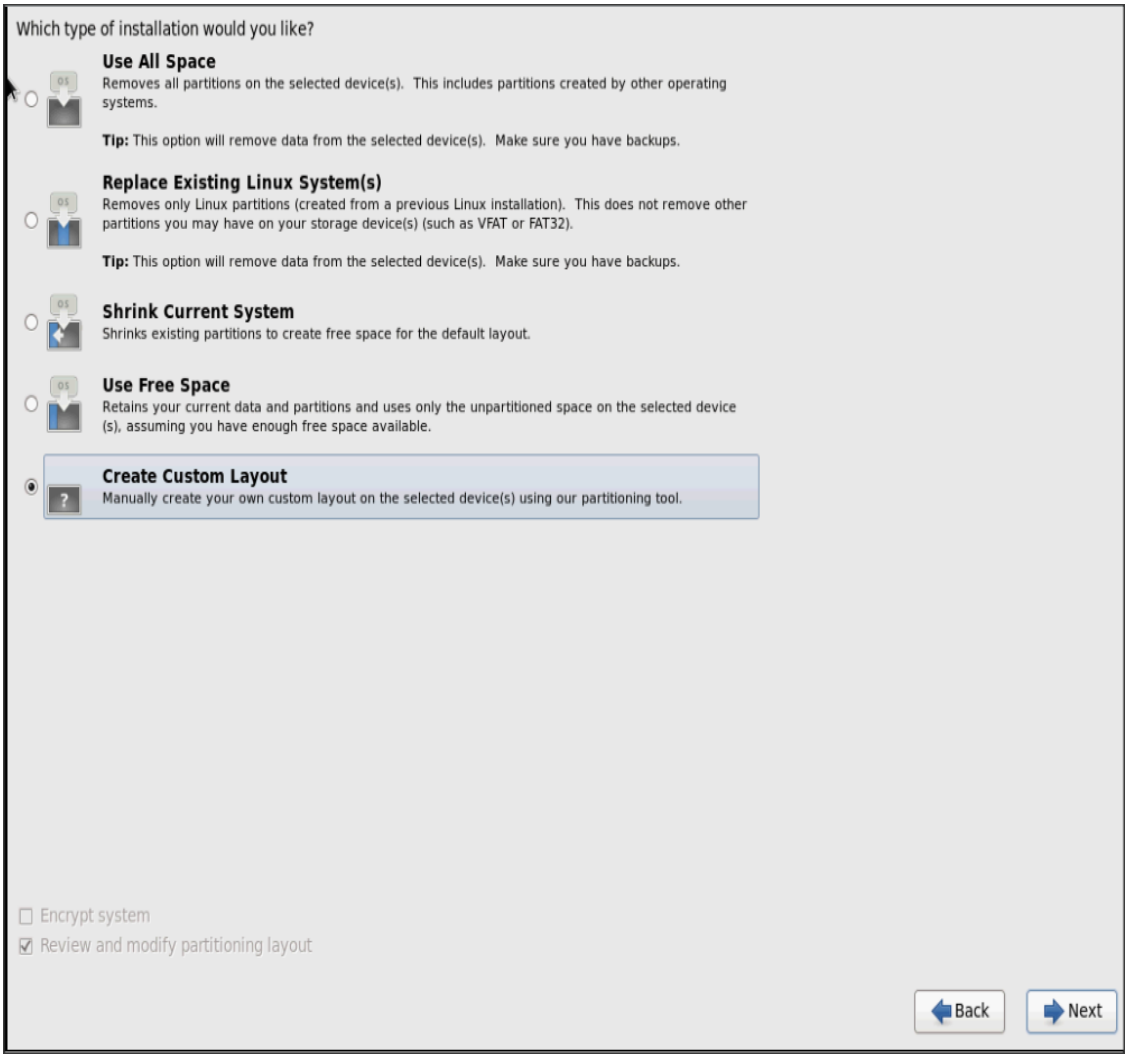


Figure 68 RHEL Installation: Enter Root Credentials



21. Choose **Create custom layout** for Installation type.

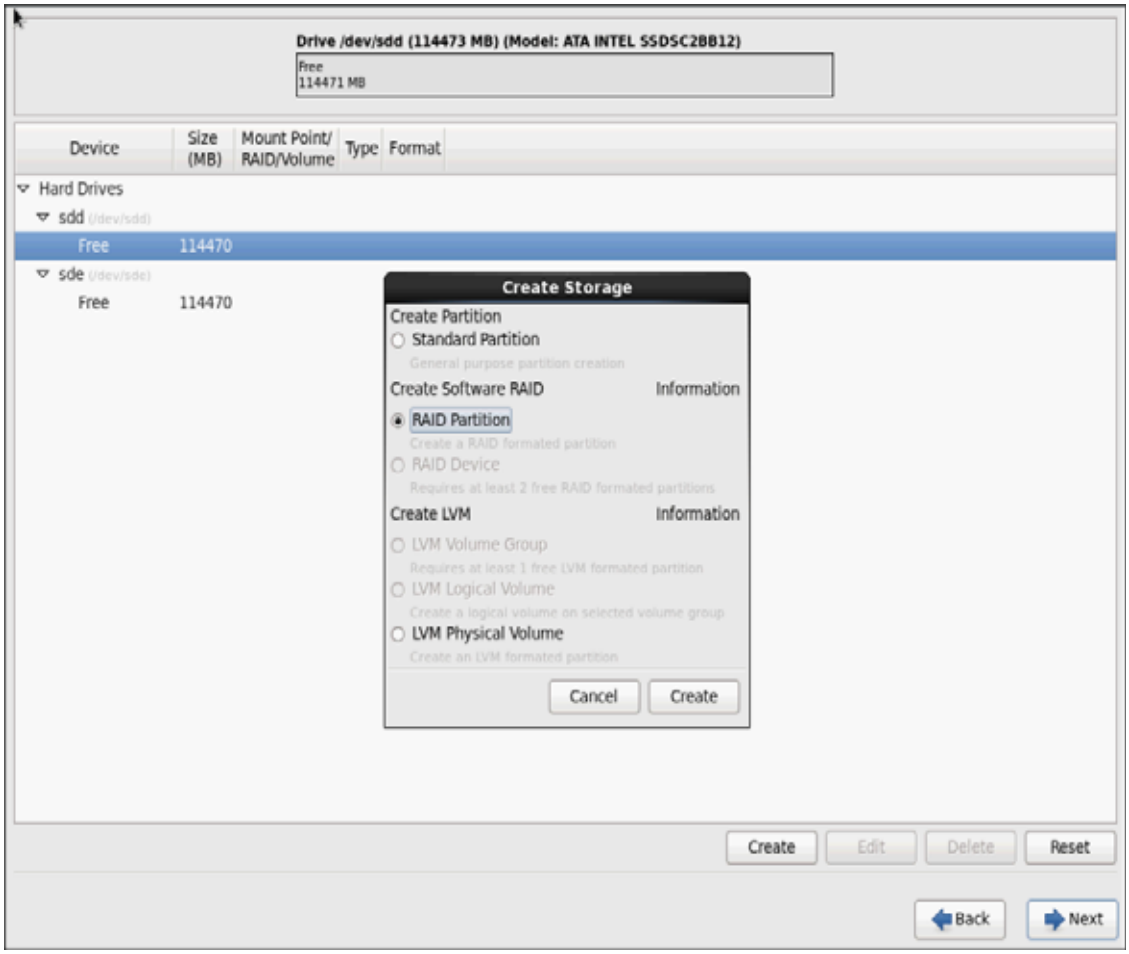
Figure 69 RHEL Installation: Create Custom Layout



Following steps can be used to create two software RAID 1 partitions for boot and / (root) partitions.

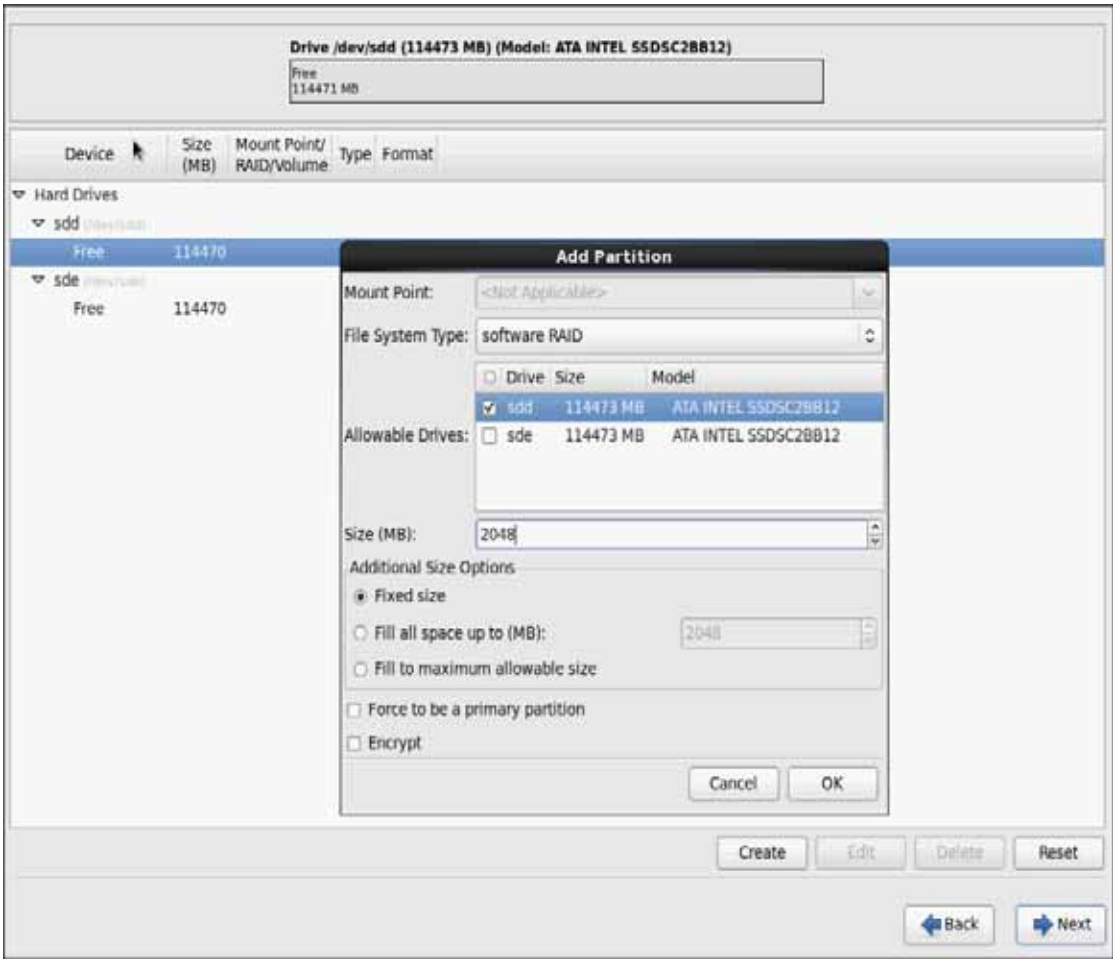
22. Choose free volume and click on **Create** and choose **RAID Partition.**

Figure 70 RHEL Installation: Create RAID Partition



23. Choose “Software RAID” for File system Type and set size for Boot volume.

Figure 71 RHEL Installation: Add Partition



24. Similarly, do the RAID partitioning for the other free volume.

Figure 72 RHEL Installation: Create RAID Partition

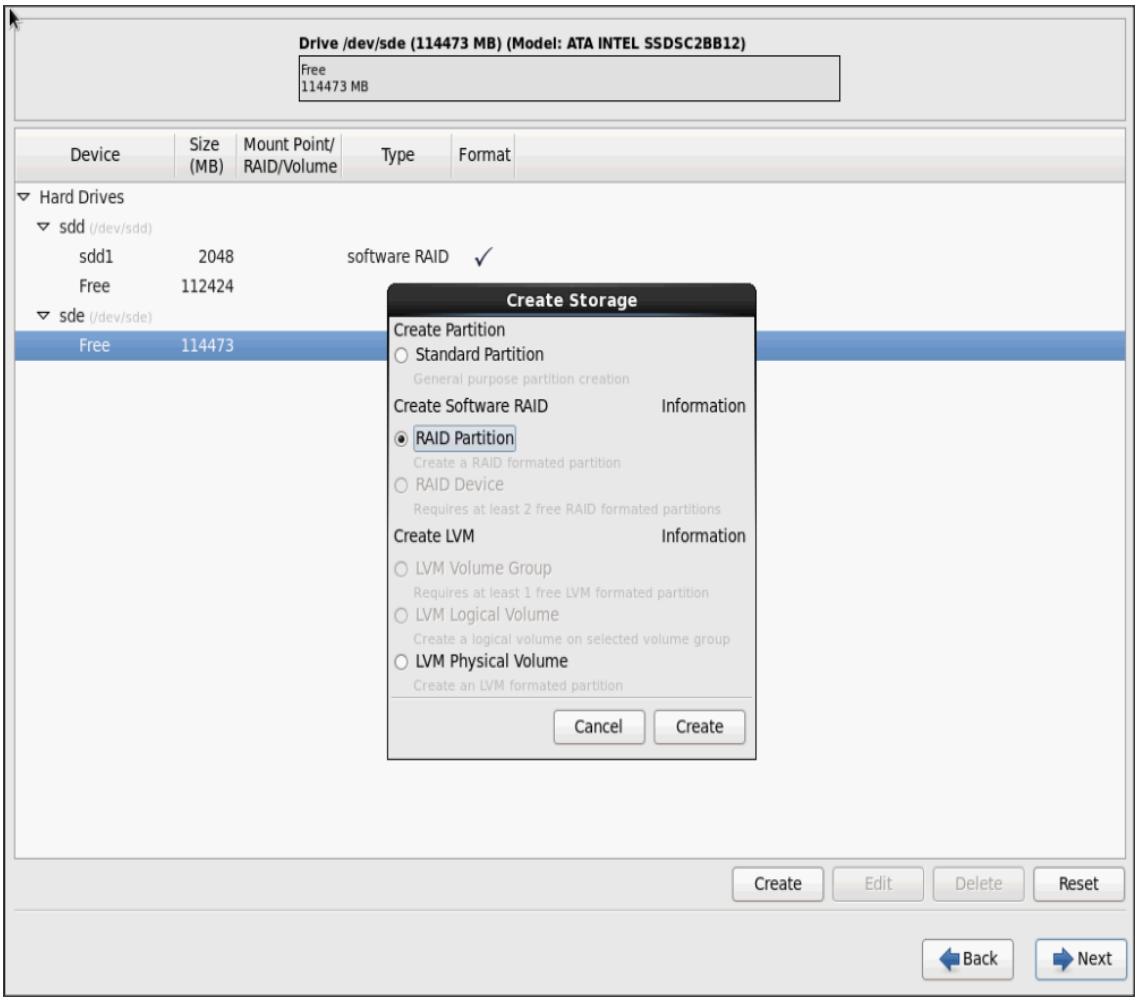
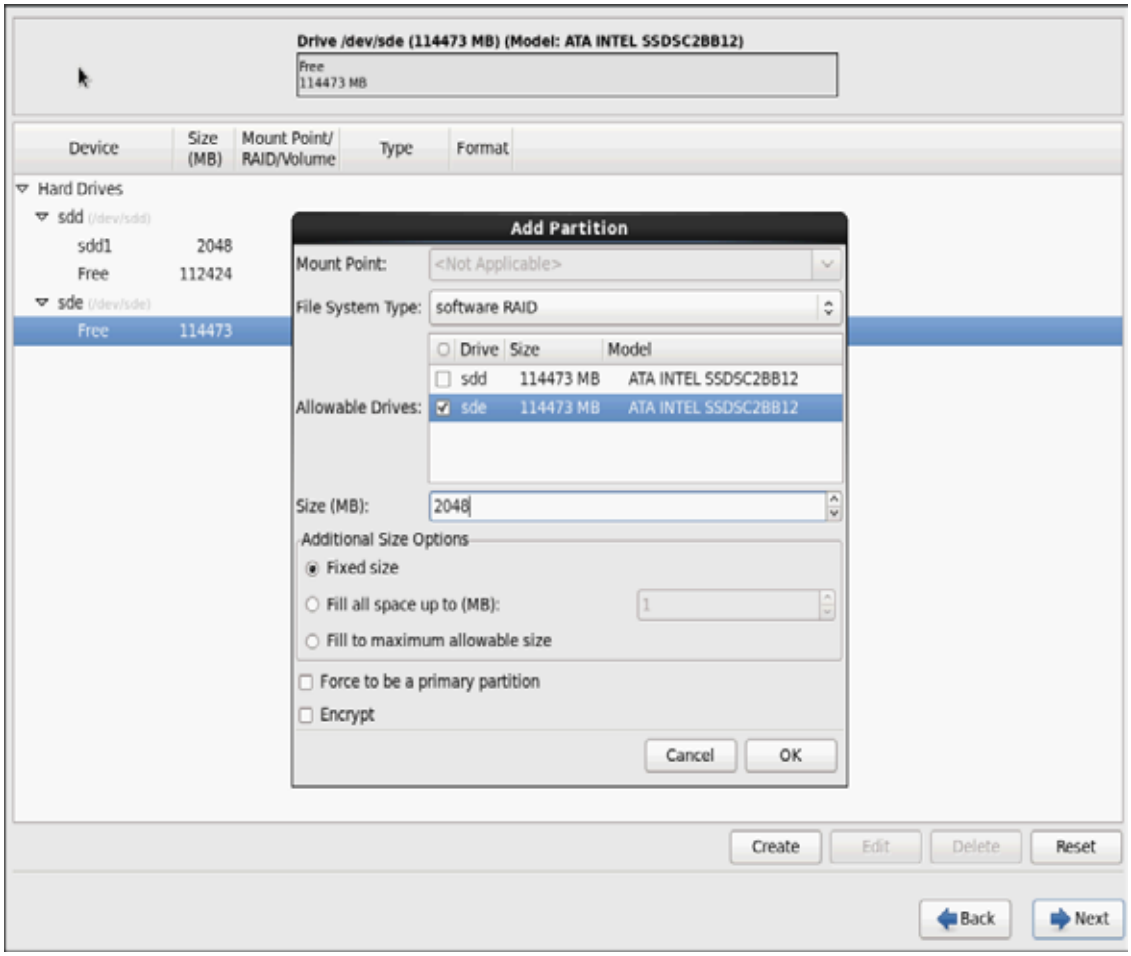


Figure 73 RHEL Installation: Add Partition



25. Now similarly create RAID partitions for root (/) partition on both the devices and use rest of the available space.

Figure 74 RHEL Installation: Create RAID Partition

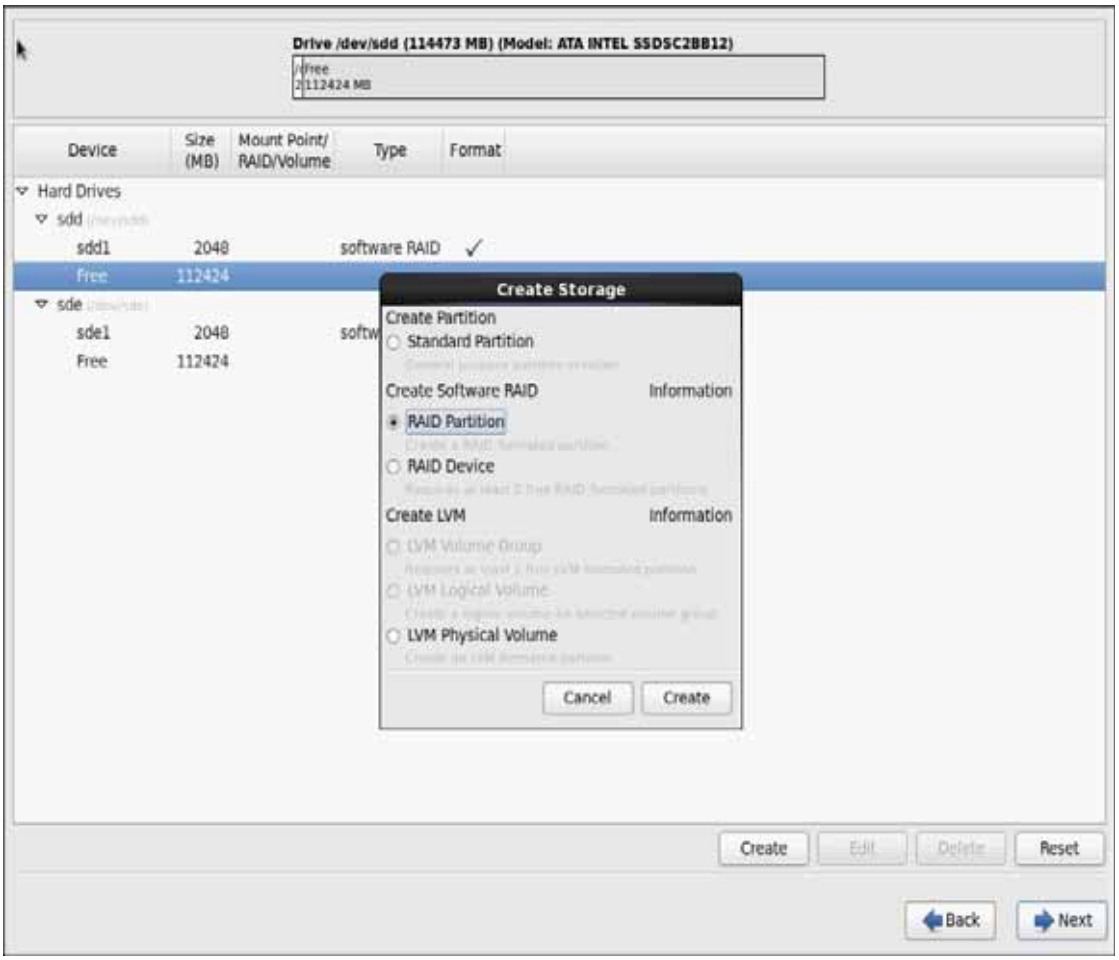


Figure 75 RHEL Installation: Add Partition

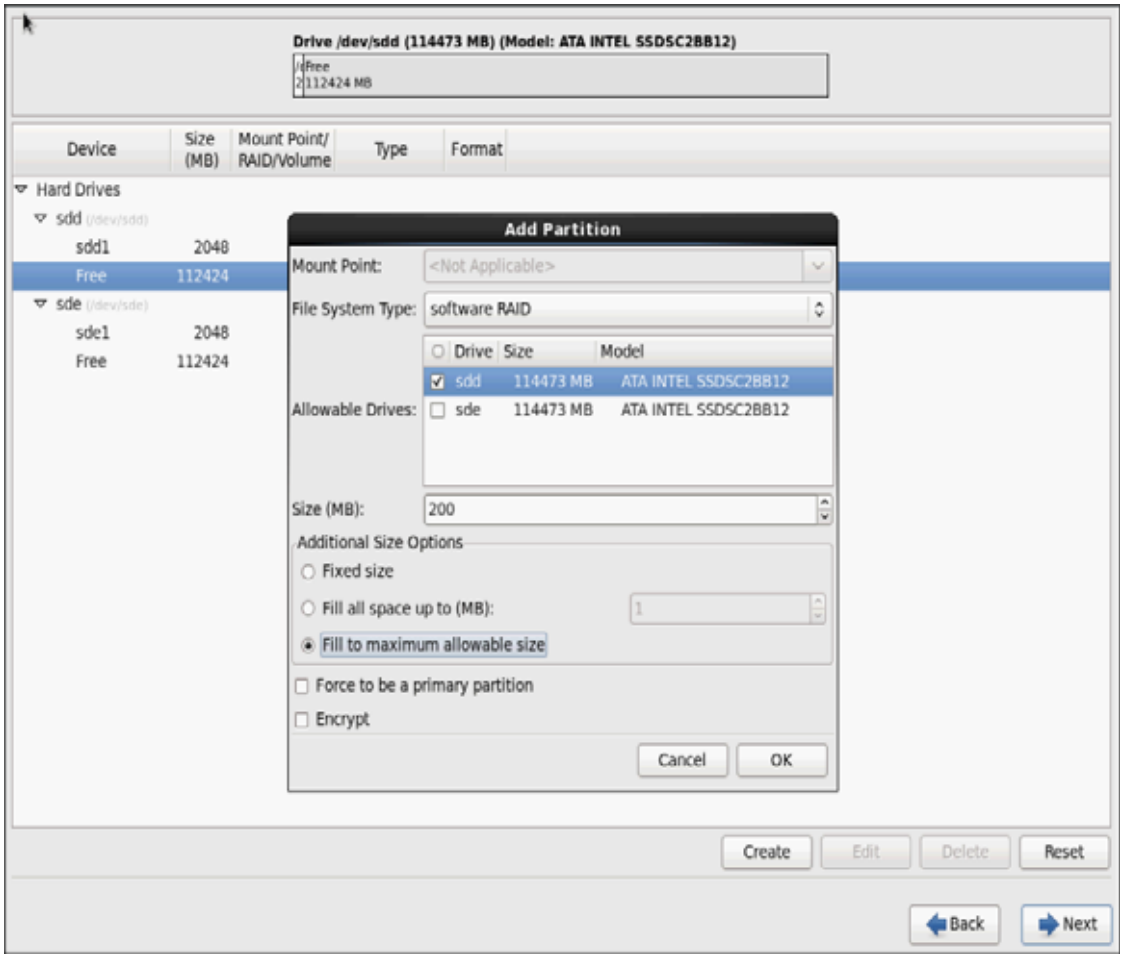


Figure 76 RHEL Installation: Create RAID Partition

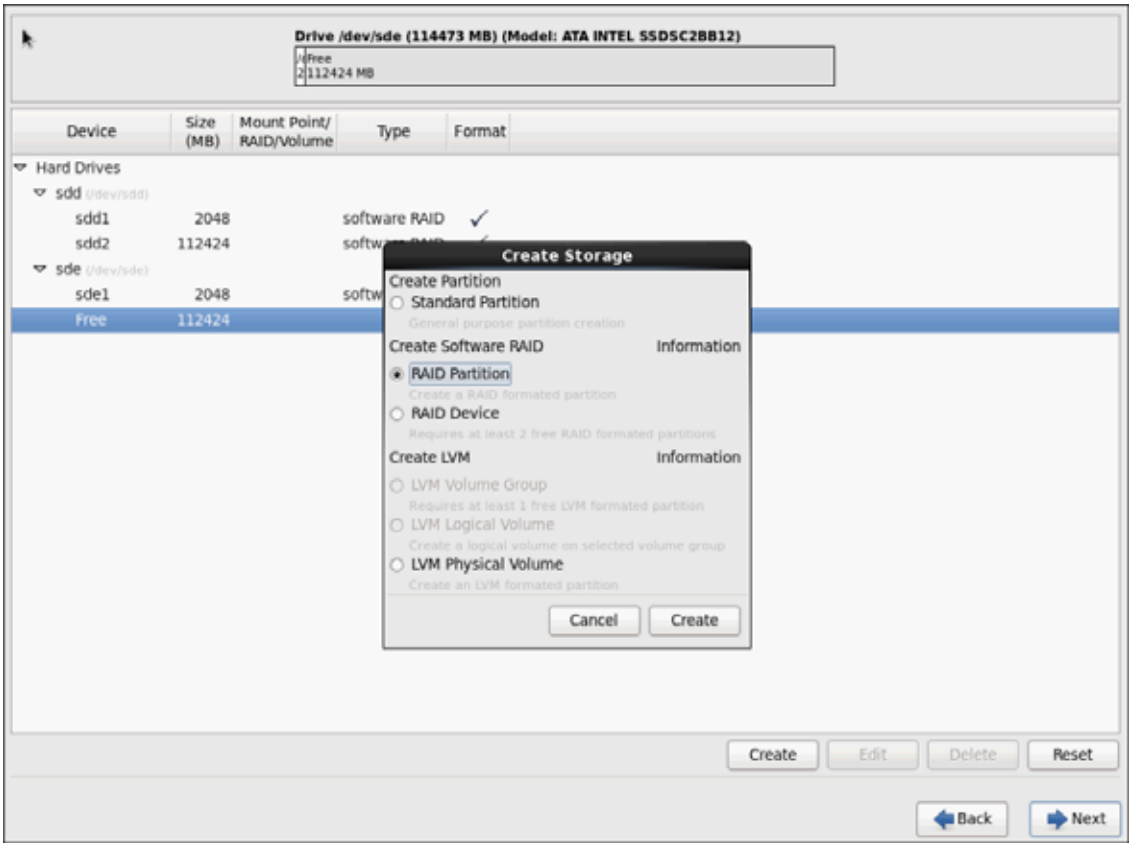
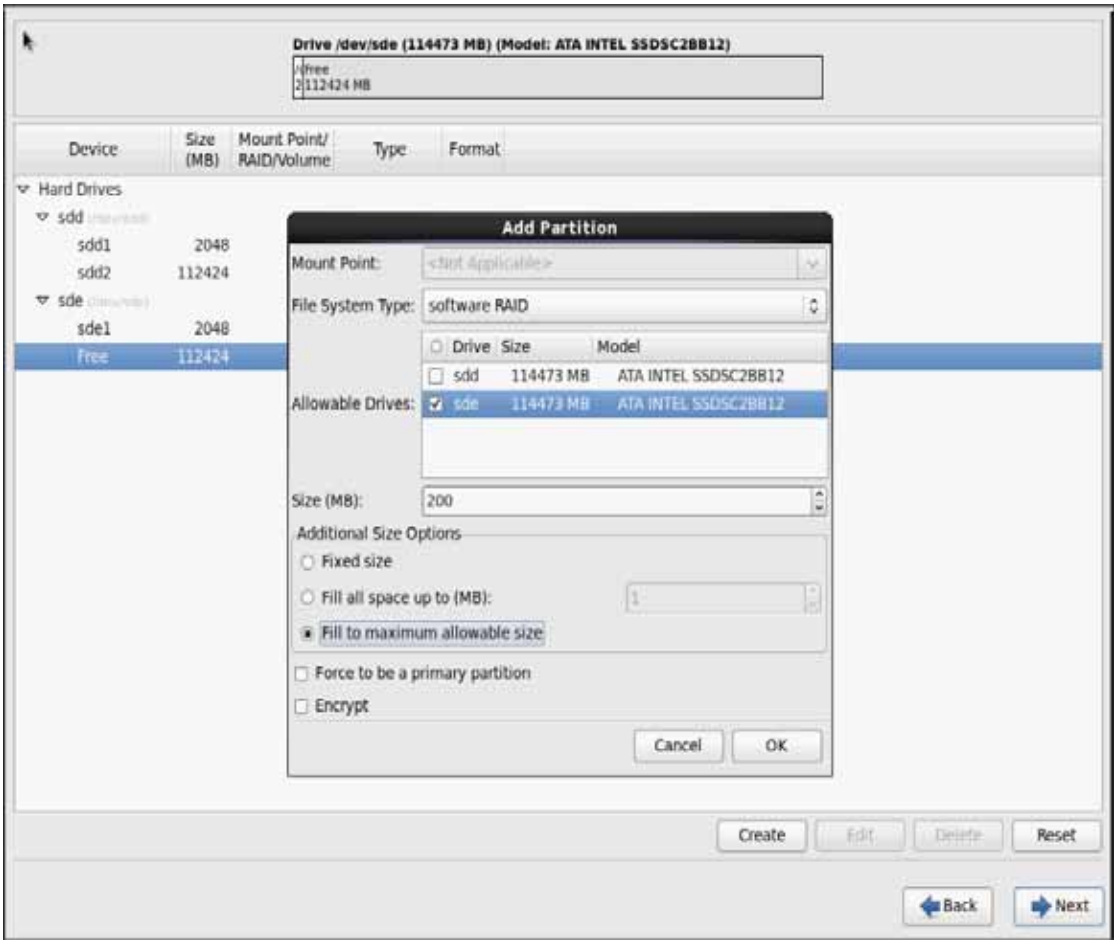


Figure 77 RHEL Installation: Add Partition

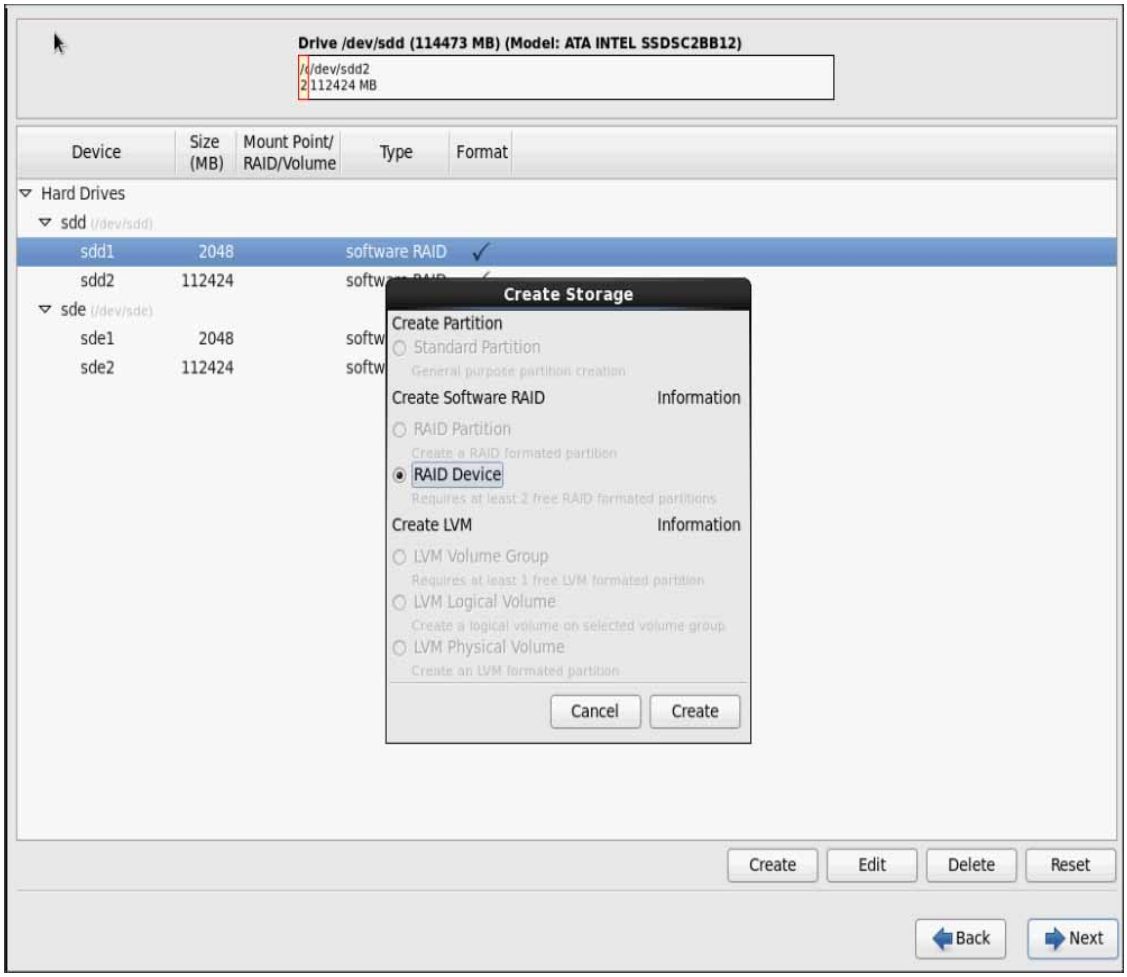


26. The above steps created 2 boot and 2 root (/) partitions. Following steps will RAID1 devices.

Figure 78 RHEL Installation: Selected RAID Devices

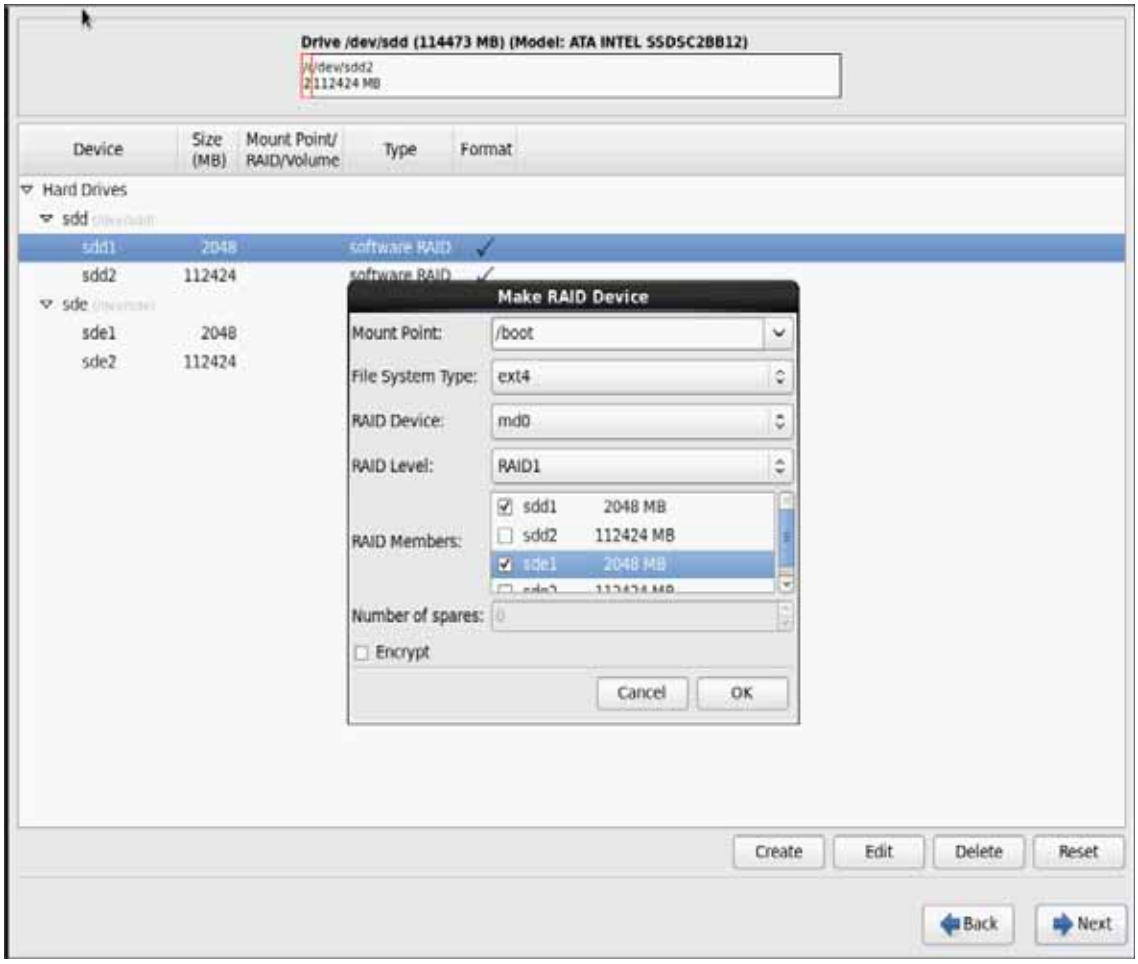
27. Choose one of the boot partitions and click on **Create > RAID Device**

Figure 79 RHEL Installation: Create RAID Device



28. Choose this as /boot (boot device) and in RAID members, choose all the boot partitions created above in order to create a software RAID 1 for boot

Figure 80 RHEL Installation: Make RAID Device



29. Similarly repeat for / partitions created above choosing both members with mount point as “/”.

Figure 81 RHEL Installation: Create RAID Device

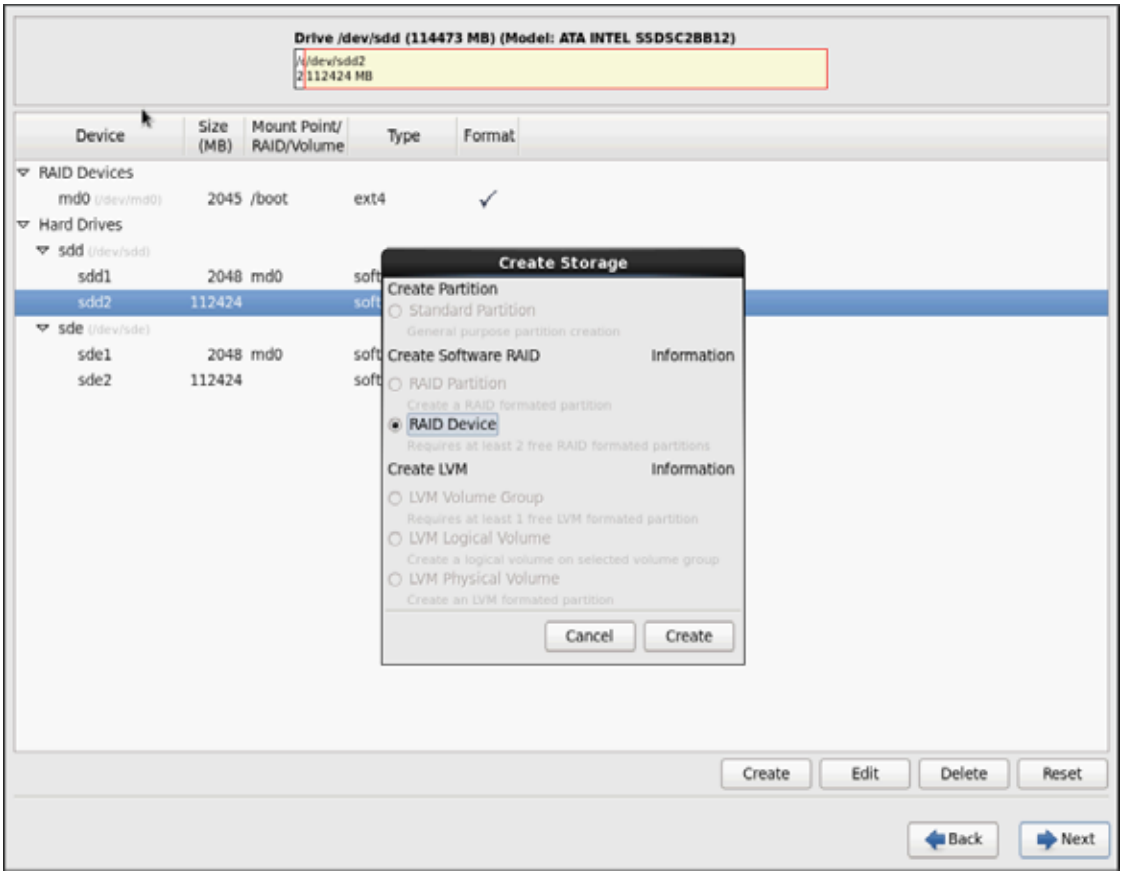


Figure 82 RHEL Installation: Make RAID Device

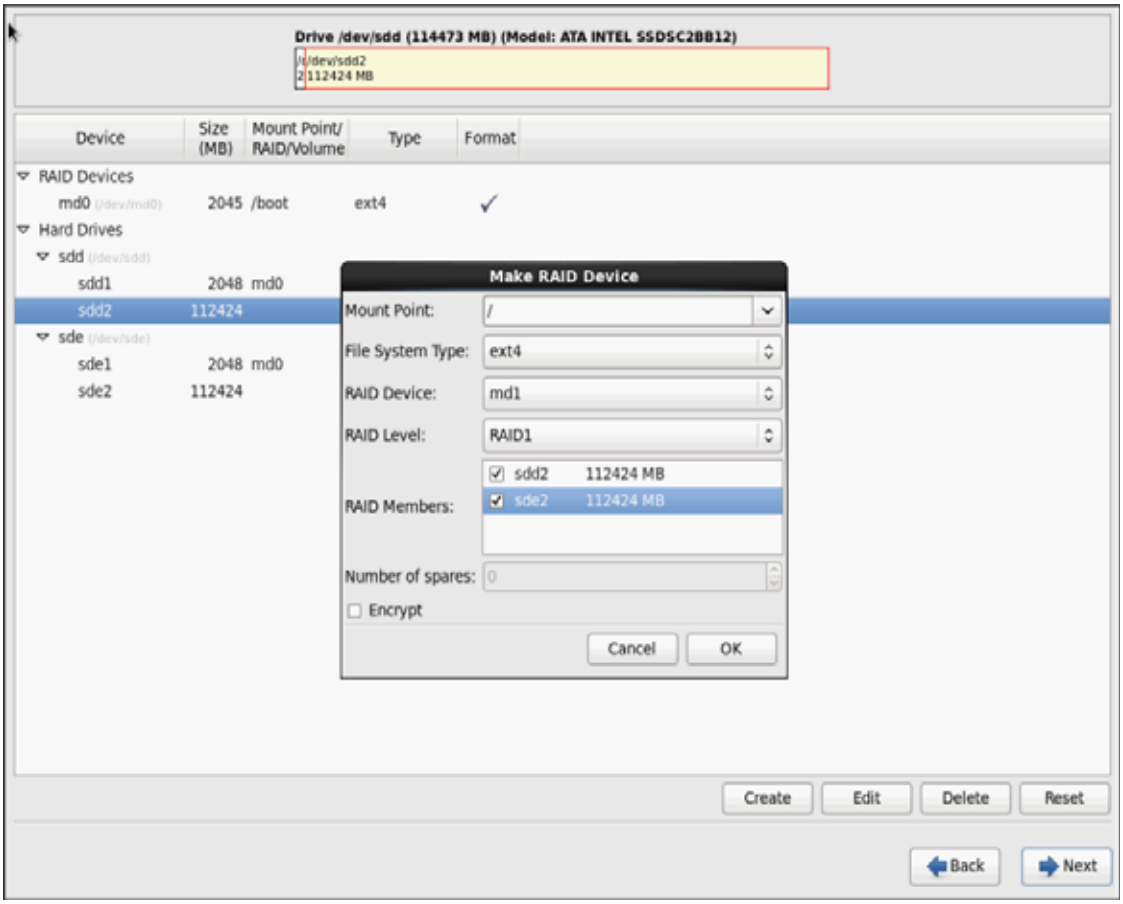


Figure 83 RHEL Installation: Selected RAID Devices

30. Click on **Next.**

Figure 84 RHEL Installation: Partitioning Warning

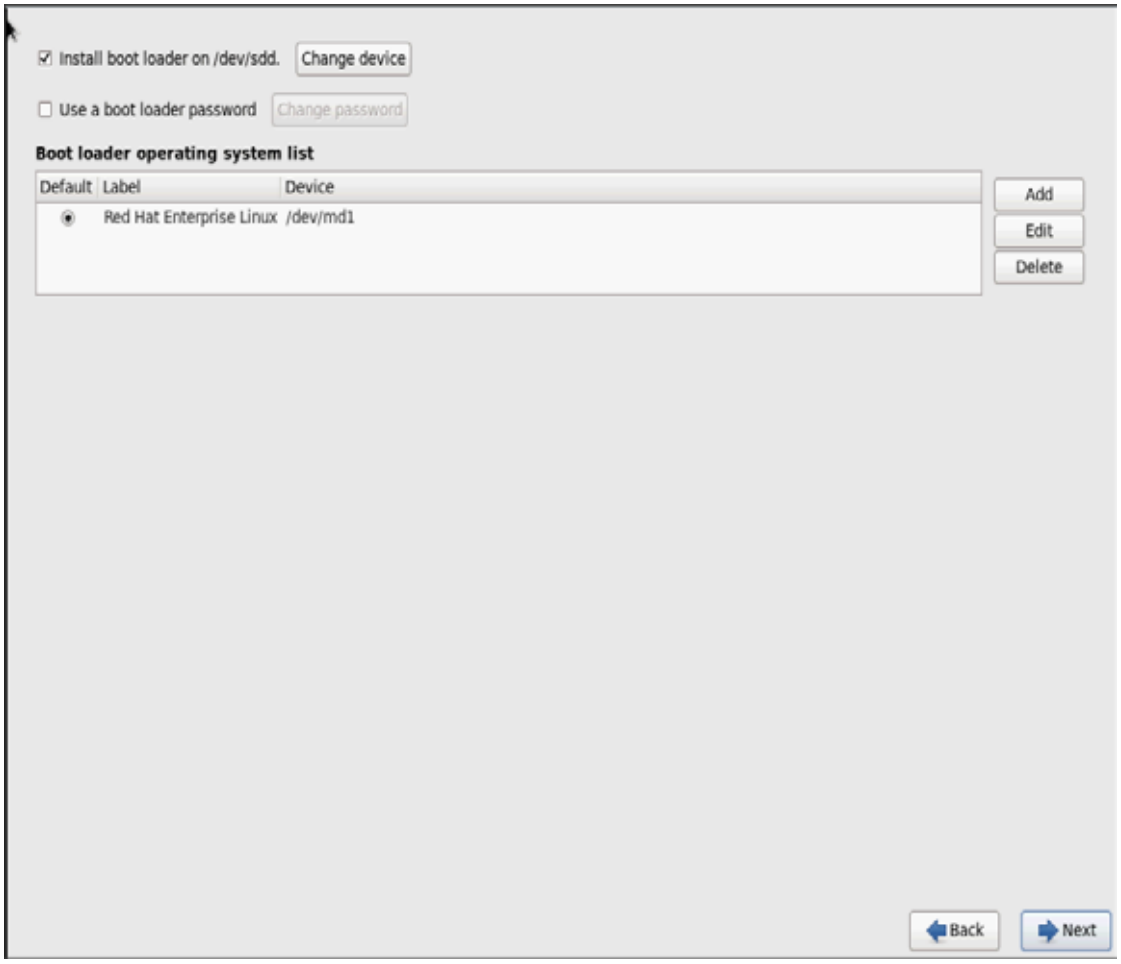
Note Swap partition can be created using the similar steps, however, since these systems are high in memory, this step is skipped (click **Yes**)

31. Click **Next**, and **Format.**

Figure 85 RHEL Installation: Format Warning

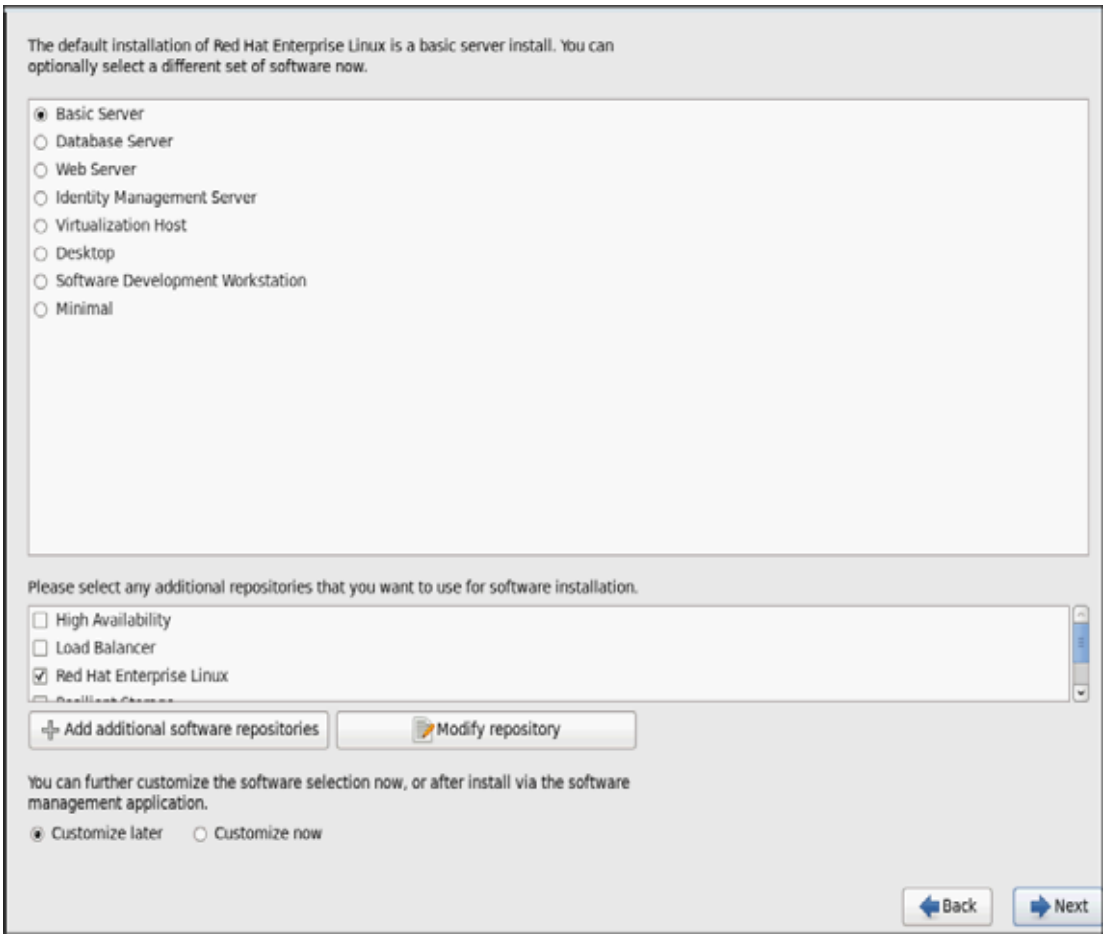
32. Select default settings and click **Next.**

Figure 86 RHEL Installation: Install Boot Loader



33. Continue with RHEL Installation as shown below.

Figure 87 RHEL Installation: Keep the Default Installation



34. Once the installation is complete reboot the system.

Repeat the steps 1 through 34, to install Red Hat Enterprise Linux 6.5 on Servers 2 through 64.

Note The OS installation and configuration of the nodes that is mentioned above can be automated through PXE boot or third party tools.

The host-names and their corresponding IP addresses are shown in Table 7.

Table 7 Host-names and IP Addresses

Hostname	eth0	eth1	eth2
rhel1	10.29.160.101	192.168.11.101	192.168.12.101
rhel2	10.29.160.102	192.168.11.102	192.168.12.102
rhel3	10.29.160.103	192.168.11.103	192.168.12.103
rhel4	10.29.160.104	192.168.11.104	192.168.12.104
rhel5	10.29.160.105	192.168.11.105	192.168.12.105
rhel6	10.29.160.106	192.168.11.106	192.168.12.106

Table 7 Host-names and IP Addresses

Hostname	eth0	eth1	eth2
rhel7	10.29.160.107	192.168.11.107	192.168.12.107
rhel8	10.29.160.108	192.168.11.108	192.168.12.108
rhel9	10.29.160.109	192.168.11.109	192.168.12.109
rhel10	10.29.160.110	192.168.11.110	192.168.12.110
rhel11	10.29.160.111	192.168.11.111	192.168.12.111
rhel12	10.29.160.112	192.168.11.112	192.168.12.112
rhel13	10.29.160.113	192.168.11.113	192.168.12.113
rhel14	10.29.160.114	192.168.11.114	192.168.12.114
rhel15	10.29.160.115	192.168.11.115	192.168.12.115
rhel16	10.29.160.116	192.168.11.116	192.168.12.116
rhel64	10.29.160.164	192.168.11.164	192.168.12.164

Post OS Install Configuration

Choose one of the nodes of the cluster or a separate node as Admin Node for management such as MapR installation, cluster parallel shell, creating a local Red Hat repo and others. In this document, we use rhel1 for this purpose.

Setting Up Password-less Login

To manage all of the clusters nodes from the admin node we need to setup password-less login. It assists in automating common tasks with cluster-shell (clush, a cluster wide parallel shell), and shell-scripts without having to use passwords.

Once Red Hat Linux is installed across all the nodes in the cluster, follow these steps in order to enable password-less login across all the nodes.

1. Login to the Admin Node (rhel1)

```
ssh 10.29.160.101
```

2. Run the ssh-keygen command to create both public and private keys on the admin node.

```
[root@rhel1 ~]# ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Created directory '/root/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
ab:4e:78:10:54:81:4e:04:8d:af:4f:a4:b2:c4:bb:88 root@rhel1
The key's randomart image is:
+--[ RSA 2048]----+
|  .=ooo.         |
|  ..+            |
|   +.            |
|    +.           |
|.  +.   S        |
|.oo .o   .       |
|.o.o. o .        |
|+.  .o .         |
|E..  .o          |
+-----------------+
```

3. Then run the following command from the admin node to copy the public key id_rsa.pub to all the nodes of the cluster. ssh-copy-id appends the keys to the remote-host's .ssh/authorized_key.

```
for IP in {101..168}; do echo -n "$IP -> "; ssh-copy-id -i ~/.ssh/id_rsa.pub
10.29.160.$IP; done
```

Enter **yes** for **Are you sure you want to continue connecting (yes/no)?**

Enter the password of the remote host.

Configuring /etc/hosts

Setup /etc/hosts on the Admin node and other nodes as follows; this is a pre-configuration to setup DNS as shown in the further section.

Follow these steps to create the host file across all the nodes in the cluster:

1. Populate the host file with IP addresses and corresponding hostnames on the Admin node (rhel1) and other nodes as follows

On Admin Node (rhel1)

```
vi /etc/hosts
127.0.0.1 local host localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6

10.29.160.101 rhel1.mgmt
10.29.160.102 rhel2.mgmt
10.29.160.103 rhel3.mgmt
10.29.160.104 rhel4.mgmt
10.29.160.105 rhel5.mgmt
10.29.160.106 rhel6.mgmt
10.29.160.107 rhel7.mgmt
10.29.160.108 rhel8.mgmt
10.29.160.109 rhel9.mgmt
10.29.160.110 rhel10.mgmt
10.29.160.111 rhel11.mgmt
```

```
10.29.160.112 rhel12.mgmt
10.29.160.113 rhel13.mgmt
10.29.160.114 rhel14.mgmt
10.29.160.115 rhel15.mgmt
10.29.160.116 rhel16.mgmt
...
10.29.160.164 rhel64.mgmt


192.168.11.101 rhel1
192.168.11.102 rhel2
192.168.11.103 rhel3
192.168.11.104 rhel4
192.168.11.105 rhel5
192.168.11.106 rhel6
192.168.11.107 rhel7
192.168.11.108 rhel8
192.168.11.109 rhel9
192.168.11.110 rhel10
192.168.11.111 rhel11
192.168.11.112 rhel12
192.168.11.113 rhel13
192.168.11.114 rhel14
192.168.11.115 rhel15
192.168.11.116 rhel16
...
192.168.11.164 rhel64

192.168.12.101 rhel1-2
192.168.12.102 rhel2-2
192.168.12.103 rhel3-2
192.168.12.104 rhel4-2
192.168.12.105 rhel5-2
192.168.12.106 rhel6-2
192.168.12.107 rhel7-2
192.168.12.108 rhel8-2
192.168.12.109 rhel9-2
192.168.12.120 rhel10-2
192.168.12.121 rhel12-2
192.168.12.122 rhel12-2
192.168.12.123 rhel13-2
192.168.12.124 rhel14-2
192.168.12.125 rhel15-2
192.168.12.126 rhel16-2
...
192.168.12.164 rhel64-2
```

Setup ClusterShell

ClusterShell (or clush) is cluster wide shell to run commands on several hosts in parallel.

From the system connected to the Internet download Cluster shell (clush) and install it on rhel1. Cluster shell is available from EPEL (Extra Packages for Enterprise Linux) repository.

```
wget http://dl.fedoraproject.org/pub/epel//6/x86_64/clustershell-1.6-1.el6.noarch.rpm

scp clustershell-1.6-1.el6.noarch.rpm rhel1:/root/
```

Login to rhel1 and install cluster shell

```
yum –y install clustershell-1.6-1.el6.noarch.rpm
```

```
Edit /etc/clustershell/groups file to include host-names for all the nodes of the cluster.
These set of hosts are taken when running clush with '-a' option
```

For 68 node cluster as in our CVD, set groups file as follows,

```
vi /etc/clustershell/groups
all: rhel[1-64].mgmt
```

Note For more information and documentation on ClusterShell, visit https://github.com/cea-hpc/clustershell/wiki/UserAndProgrammingGuide

Note clustershell will not work if not ssh to the machine earlier (as it requires to be in known_hosts file), for instance, as in the case below for rhel<host> and rhel<host>.mgmt.

```
[root@rhel1 ~]# ssh rhel2
The authenticity of host 'rhel2 (192.168.11.102)' can't be established.
RSA key fingerprint is 9e:4d:91:3d:b9:ef:eb:97:b4:80:dc:3b:85:f5:ad:20.
Are you sure you want to continue connecting (yes/no)?
```

```
[root@rhel1 ~]# ssh rhel5.mgmt
The authenticity of host 'rhel5.mgmt (10.29.160.105)' can't be established.
RSA key fingerprint is 7a:98:75:9a:6a:1a:80:a4:97:43:6c:8a:12:57:db:74.
Are you sure you want to continue connecting (yes/no)?
```

Creating Red Hat Enterprise Linux (RHEL) 6.5 Local Repo

To create a repository using RHEL DVD or ISO on the admin node (in this deployment rhel1 is used for this purpose), create a directory with all the required RPMs, run the createrepo command and then publish the resulting repository.

1. Log on to rhel1. Create a directory that would contain the repository.

   ```
   mkdir -p /var/www/html/rhelrepo
   ```

2. Copy the contents of the Red Hat DVD to **/var/www/html/rhelrepo** directory.
3. Alternatively, if you have access to a Red Hat ISO Image, Copy the ISO file to rhel1.

   ```
   scp rhel-server-6.5-x86_64-dvd.iso rhel1:/root/
   ```

Here we assume you have the Red Hat ISO file located in your present working directory.

```
mkdir -p /mnt/rheliso
mount -t iso9660 -o loop /root/rhel-server-6.5-x86_64-dvd.iso /mnt/rheliso/
```

4. Next, copy the contents of the ISO to the **/var/www/html/rhelrepo** directory

   ```
   cp -r /mnt/rheliso/* /var/www/html/rhelrepo
   ```

```
[root@rhel1 ~]# mkdir -p /var/www/html/rhelrepo
[root@rhel1 ~]# mkdir -p /mnt/rheliso
[root@rhel1 ~]#
[root@rhel1 ~]# mount -t iso9660 -o loop /root/rhel-server-6.5-x86_64-dvd.iso /mnt/rheliso/
[root@rhel1 ~]# cp -r /mnt/rheliso/* /var/www/html/rhelrepo/
```

5. Now on rhel1 create a.repo file to enable the use of the yum command.

```
vi /var/www/html/rhelrepo/rheliso.repo
[rhel6.5]
name=Red Hat Enterprise Linux 6.5
baseurl=http://10.29.160.101/rhelrepo
gpgcheck=0
enabled=1
```

6. Now copy rheliso.repo file from **/var/www/html/rhelrepo to /etc/yum.repos.d** on rhel1

```
cp /var/www/html/rhelrepo/rheliso.repo /etc/yum.repos.d/
```

Note Based on this repo file yum requires httpd to be running on rhel1 for other nodes to access the repository.

7. Copy the **rheliso.repo** to all the nodes of the cluster.

```
clush -a -b -c /etc/yum.repos.d/rheliso.repo --dest=/etc/yum.repos.d/
```

```
[root@rhel1 ~]# clush -a -b -c /etc/yum.repos.d/rheliso.repo --dest=/etc/yum.repos.d/
```

8. To make use of repository files on rhel1 without httpd, edit the baseurl of repo file **/etc/yum.repos.d/rheliso.repo** to point repository location in the file system.

Note This step is needed to install software on Admin Node (rhel1) using the repo (such as httpd, createrepo, etc).

```
vi /etc/yum.repos.d/rheliso.repo
[rhel6.5]
name=Red Hat Enterprise Linux 6.5
baseurl=file:///var/www/html/rhelrepo
gpgcheck=0
enabled=1
```

9. Creating the Red Hat Repository Database.

Install the createrepo package on admin node (rhel1). Use it to regenerate the repository database(s) for the local copy of the RHEL DVD contents.

```
yum -y install createrepo
```

```
[root@rhel1 ~]# yum -y install createrepo
Loaded plugins: product-id, refresh-packagekit, security, subscription-manager
This system is not registered to Red Hat Subscription Management. You can use subscription-manager to register.
rhel6.5                                                                                  | 3.9 kB     00:00
rhel6.5/primary_db                                                                       | 3.1 MB     00:00
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package createrepo.noarch 0:0.9.9-18.el6 will be installed
--> Processing Dependency: python-deltarpm for package: createrepo-0.9.9-18.el6.noarch
--> Running transaction check
---> Package python-deltarpm.x86_64 0:3.5-0.5.20090913git.el6 will be installed
--> Processing Dependency: deltarpm = 3.5-0.5.20090913git.el6 for package: python-deltarpm-3.5-0.5.20090913git.el6.x86_64
--> Running transaction check
```

10. Run createrepo on the RHEL repository to create the repo database on admin node

```
cd /var/www/html/rhelrepo
createrepo .
```

```
[root@rhel1 rhelrepo]# createrepo .
Spawning worker 0 with 3763 pkgs
Workers Finished
Gathering worker results

Saving Primary metadata
Saving file lists metadata
Saving other metadata
Generating sqlite DBs
Sqlite DBs complete
```

11. Finally, purge the yum caches after httpd is installed (steps in section “Install Httpd”).

Configuring DNS

This section details setting up DNS using dnsmasq as an example based on the /etc/hosts configuration setup in the earlier section.

Follow these steps to create the host file across all the nodes in the cluster:

1. Disable Network manager on all nodes

```
clush -a -b service NetworkManager stop
clush -a -b chkconfig NetworkManager off
```

2. Update /etc/resolv.conf file to point to Admin Node

```
vi /etc/resolv.conf
nameserver 192.168.11.101
```

Note This step is needed if setting up dnsmasq on Admin node. Else this file should be updated with the correct nameserver.

3. Install and Start dnsmasq on Admin node

```
yum -y install dnsmasq
service dnsmasq start
chkconfig dnsmasq on
```

4. Deploy /etc/resolv.conf from the admin node (rhel1) to all the nodes via the following clush command:

```
clush -a -B -c /etc/resolv.conf
```

Note A clush copy without - –dest copies to the same directory location as the source-file directory.

5. Ensure DNS is working fine by running the following command on Admin node and any datanode

```
[root@rhel2 ~]# nslookup rhel1
Server:192.168.11.101
Address:192.168.11.101#53

Name: rhel1
Address: 192.168.11.101 •

[root@rhel2 ~]# nslookup rhel1.mgmt
Server: 192.168.11.101
Address: 192.168.11.101#53

Name: rhel1.mgmt
Address: 10.29.160.101 •

[root@rhel2 ~]# nslookup 10.29.160.101
Server: 192.168.11.101
Address: 192.168.11.101#53

101.160.29.10.in-addr.arpa name = rhel1.mgmt. •
```

Installing httpd

Setting up RHEL repo on the admin node requires httpd. This section describes the process of setting up one

1. Install httpd on the admin node to host repositories.

   The Red Hat repository is hosted using HTTP on the admin node, this machine is accessible by all the hosts in the cluster.

```
yum –y install httpd
```

2. Add ServerName and make the necessary changes to the server configuration file.

```
vi /etc/httpd/conf/httpd.conf
ServerName 10.29.160.101:80
```

```
[root@rhel1 ~]# vi /etc/httpd/conf/httpd.conf
[root@rhel1 ~]# cat /etc/httpd/conf/httpd.conf | grep ServerName
# ServerName gives the name and port that the server uses to identify itself.
#ServerName www.example.com:80
ServerName 10.29.160.101:80
# ServerName directive.
#    ServerName dummy-host.example.com
```

3. Start httpd

```
service httpd start
```

```
chkconfig httpd on
```

4. Purge the yum caches after httpd is installed (step followed from section Setup Red Hat Repo)

```
clush -a -B yum clean all
clush –a –B yum repolist
```

```
[root@rhel1 ~]# clush -a -B yum clean all
---------------
rhel[1-17] (17)
---------------
Loaded plugins: product-id, refresh-packagekit, security, subscription-manager
This system is not registered to Red Hat Subscription Management. You can use subscription-manager to register.
Cleaning repos: rhel6.5
Cleaning up Everything
```

Note While suggested configuration is to disable SELinux as shown below, if for any reason SELinux needs to be enabled on the cluster, then ensure to run the following to make sure that the httpd is able to read the Yum repofiles `chcon -R -t httpd_sys_content_t /var/www/html/`

Upgrading Cisco Network driver for VIC1227

The latest Cisco Network driver is required for performance and updates. The latest drivers can be downloaded from the link below:

https://software.cisco.com/download/release.html?mdfid=283862063&flowid=25886&softwareid=283853158&release=1.5.7d&relind=AVAILABLE&rellifecycle=&reltype=latest

In the ISO image, the required driver `kmod-enic-2.1.1.66-rhel6u5.el6.x86_64.rpm` can be located at `\Linux\Network\Cisco\12x5x\RHEL\RHEL6.5`

From a node connected to the Internet, download, extract and transfer `kmod-enic-2.1.1.66-rhel6u5.el6.x86_64.rpm` to rhel1 (admin node).

Install the rpm on all nodes of the cluster using the following clush commands. For this example the rpm is assumed to be in present working directory of rhel1.

```
[root@rhel1 ~]# clush -a -b -c kmod-enic-2.1.1.66-rhel6u5.el6.x86_64.rpm
[root@rhel1 ~]# clush -a -b "rpm -ivh kmod-enic-2.1.1.66-rhel6u5.el6.x86_64.rpm "
```

Ensure that the above installed version of kmod-enic driver is being used on all nodes by running the command “modinfo enic” on all nodes

```
[root@rhel1 ~]# clush -a -B "modinfo enic | head -5"
```

```
filename:       /lib/modules/2.6.32-431.el6.x86_64/extra/enic/enic.ko
version:        2.1.1.66
license:        GPL v2
author:         Scott Feldman <scofeldm@cisco.com>
description:    Cisco VIC Ethernet NIC Driver
```

Setting up JAVA

MapR requires JAVA 7, download jdk-7u75–linux-x64.rpm from oracle.com (http://www.oracle.com/technetwork/java/javase/downloads/jdk7–downloads-1880260.html) to admin node (rhel1).

Create the following files java-set-alternatives.sh and java-home.sh on admin node (rhel1)

vi java-set-alternatives.sh

```
#!/bin/bash
for item in java javac javaws jar jps javah javap jcontrol jconsole jdb; do
 rm -f /var/lib/alternatives/$item
 alternatives --install /usr/bin/$item $item /usr/java/jdk1.7.0_75/bin/$item 9
 alternatives --set $item /usr/java/jdk1.7.0_75/bin/$item
done
```

vi java-home.sh

```
export JAVA_HOME=/usr/java/jdk1.7.0_75
```

Run the following commands on admin node (rhel1) to install and setup java on all nodes

1. Copying JDK rpm to all nodes

   ```
   clush -b -a -c /root/jdk-7u75-linux-x64.rpm --dest=/root/
   ```

2. Make the two java scripts created above executable

   ```
   chmod 755 ./java-set-alternatives.sh ./java-home.sh
   ```

3. Copying java-set-alternatives.sh to all nodes

   ```
   clush -b -a -c ./java-set-alternatives.sh --dest=/root/
   ```

4. Extract and Install JDK on all nodes

   ```
   clush -a -b rpm -ivh /root/jdk-7u75-linux-x64.rpm
   ```

5. Setup Java Alternatives

   ```
   clush -b -a ./java-set-alternatives.sh
   ```

6. Ensure correct java is setup on all nodes (should point to newly installed java path)

   ```
   clush -b -a "alternatives --display java | head -2"
   ```

7. Setup JAVA_HOME on all nodes

   ```
   clush -b -a -c ./java-home.sh --dest=/etc/profile.d
   ```

8. Display JAVA_HOME on all nodes

   ```
   clush -a -b "echo \$JAVA_HOME"
   ```

9. Display current java -version

   ```
   clush -B -a java -version
   ```

NTP Configuration

The Network Time Protocol (NTP) is used to synchronize the time of all the nodes within the cluster. The Network Time Protocol daemon (ntpd) sets and maintains the system time of day in synchronism with the timeserver located in the admin node (rhel1). Configuring NTP is critical for any Hadoop Cluster. If server clocks in the cluster drift out of sync, serious problems will occur with HBase and other services.

> Installing an internal NTP server keeps your cluster synchronized even when an outside NTP server is inaccessible.

Configure `/etc/ntp.conf` on the admin node with the following contents:

```
vi /etc/ntp.conf
driftfile /var/lib/ntp/drift
restrict 127.0.0.1
restrict -6 ::1
server 127.127.1.0
fudge 127.127.1.0 stratum 10
includefile /etc/ntp/crypto/pw
keys /etc/ntp/keys
```

Create `/root/ntp.conf` on the admin node and copy it to all nodes

```
vi /root/ntp.conf
server 10.29.160.101
driftfile /var/lib/ntp/drift
restrict 127.0.0.1
restrict -6 ::1
includefile /etc/ntp/crypto/pw
keys /etc/ntp/keys
```

Copy `ntp.conf` file from the admin node to `/etc` of all the nodes by executing the following command in the admin node (rhel1)

```
for SERVER in {102..168}; do scp /root/ntp.conf
10.29.160.$SERVER:/etc/ntp.conf; done
```

```
[root@rhel1 ~]# for SERVER in {102..168}; do scp /root/ntp.conf 10.29.160.$SERVER:/etc/ntp.conf; done
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
ntp.conf                                                    100%  136     0.1KB/s   00:00
```

Note Instead of the above for loop, this could be run as a clush command with "–w" option.

```
clush -w rhel[2-68]  -b -c /root/ntp.conf --dest=/etc
```

Do not use `clush –a –b –c /root/ntp.conf --dest=/etc` command as it overwrites **/etc/ntp.conf** on the admin node.

Run thef following to syncronize the time and restart NTP daemon on all nodes

```
clush -a -B "yum install -y ntpdate"
clush -a -b "service ntpd stop"
clush -a -b "ntpdate rhel1"
clush -a -b "service ntpd start"
```

Ensure restart of NTP daemon across reboots

```
clush -a -b "chkconfig ntpd on"
```

Enabling Syslog

Syslog must be enabled on each node to preserve logs regarding killed processes or failed jobs. Modern versions such as syslog-ng and rsyslog are possible, making it more difficult to be sure that a syslog daemon is present. One of the following commands should suffice to confirm that the service is properly configured:

```
clush -B -a rsyslogd -v
```

```
[root@rhel1 ~]# clush -B -a rsyslogd -v
---------------
rhel[1-17] (17)
---------------
rsyslogd 5.8.10, compiled with:
        FEATURE_REGEXP:                         Yes
        FEATURE_LARGEFILE:                      No
        GSSAPI Kerberos 5 support:              Yes
        FEATURE_DEBUG (debug build, slow code): No
        32bit Atomic operations supported:      Yes
        64bit Atomic operations supported:      Yes
        Runtime Instrumentation (slow code):    No

See http://www.rsyslog.com for more information.
```

```
clush -B -a service rsyslog status
```

Setting ulimit

On each node, **ulimit -n** specifies the number of inodes that can be opened simultaneously. With the default value of 1024, the system appears to be out of disk space and shows no inodes available. This value should be set to 64000 on every node.

Higher values are unlikely to result in an appreciable performance gain.

For setting ulimit on Redhat, edit `/etc/security/limits.conf` on admin node rhel1 and add the following lines:

```
root soft nofile 64000
root hard nofile 64000
```

```
[root@rhel1 ~]# cat /etc/security/limits.conf | grep 64000
root soft nofile 64000
root hard nofile 64000
```

Copy the `/etc/security/limits.conf` file from admin node (rhel1) to all the nodes using the following command.

```
clush -a -b -c /etc/security/limits.conf --dest=/etc/security/
```

```
[root@rhel1 ~]# clush -a -b -c /etc/security/limits.conf --dest=/etc/security/
```

Verify the **ulimit** setting with the following steps:

Note Ulimit values are applied on a new shell, running the command on a node on an earlier instance of a shell will show old values

Run the following command at a command line. The command should report 64000.

```
clush -B -a ulimit -n
```

Disabling SELinux

SELinux must be disabled during the install procedure and cluster setup. SELinux can be enabled after installation and while the cluster is running.

SELinux can be disabled by editing `/etc/selinux/config` and changing the SELINUX line to SELINUX=disabled. The following command will disable SELINUX on all nodes.

```
clush -a -b "sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config "
clush –a –b "setenforce 0"
```

```
[root@rhel1 ~]# clush -a -b "sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config "
```

Note The above command may fail if SELinux is already disabled.

Set TCP Retries

Adjusting the tcp_retries parameter for the system network enables faster detection of failed nodes. Given the advanced networking features of UCS, this is a safe and recommended change (failures observed at the operating system layer are most likely serious rather than transitory). On each node, set the number of TCP retries to 5 can help detect unreachable nodes with less latency.

1. Edit the file **/etc/sysctl.conf** and on admin node rhel1 and add the following lines:

```
net.ipv4.tcp_retries2=5
```

Copy the **/etc/sysctl.conf** file from admin node (rhel1) to all the nodes using the following command.

```
clush -a -b -c /etc/sysctl.conf --dest=/etc/
```

2. Load the settings from default sysctl file **/etc/sysctl.conf** by running the below command.

```
clush -B -a sysctl -p
```

Disabling the Linux Firewall

The default Linux firewall settings are far too restrictive for any Hadoop deployment. Since the UCS Big Data deployment will be in its own isolated network, there's no need to leave the IP tables service running.

```
clush -a -b "service iptables stop"
clush -a -b "chkconfig iptables off"
```

```
[root@rhel1 ~]# clush -a -b "service iptables stop"
[root@rhel1 ~]# clush -a -b "chkconfig iptables off"
```

Swapping

Lowering vm.swappiness reduces anonymous paging and minimizes OOM killer invocations. With vm.swappiness set to 1, the kernel will try to reclaim from the page cache instead of application (anonymous) pages.

In order to reduce Swapping, run the following on all nodes. Variable vm.swappiness defines how often swap should be used. 0 is No Swapping, 60 is the default value.

```
clush -a -b " echo \'vm.swappiness=0\' >> /etc/sysctl.conf"
```

Load the settings from default sysctl file /etc/sysctl.conf

```
clush –a –b “sysctl –p”
```

Disable Transparent Huge Pages

Disabling Transparent Huge Pages (THP) reduces elevated CPU usage caused by THP. From the admin node, run the following commands

```
clush -a -b "echo never >
/sys/kernel/mm/redhat_transparent_hugepage/enabled"
clush -a -b "echo never >
/sys/kernel/mm/redhat_transparent_hugepage/defrag"
```

The above command needs to be run for every reboot, hence, copy this command to /etc/rc.local so they are executed automatically for every reboot.

On Admin node, run the following commands

```
rm –f /root/thp_disable
echo "echo never > /sys/kernel/mm/redhat_transparent_hugepage/enabled" >>
/root/thp_disable
echo "echo never > /sys/kernel/mm/redhat_transparent_hugepage/defrag " >>
/root/thp_disable
```

Copy file to each node

```
clush –a –b –c /root/thp_disable
```

Append the content of file thp_disable to /etc/rc.local

```
clush -a -b “cat /root/thp_disable >> /etc/rc.local”
```

Install Openssl

Install Openssl and Openssl-devel version 1.0.1e-30 and above for RHEL6.5. If openssl is already installed (generally the case), use the following command to upgrade openssl.

```
clush –a –b  -c /root/openssl-*
clush –a –b  rpm –Uvh openssl-1.0.1e-*.rpm openssl-devel-1.0.1e-*.rpm
```

```
[root@rhel1 ~]# rpm -Uvh openssl-1.0.1e-30.el6_6.5.x86_64.rpm openssl-devel-1.0.1e-30.el6_6.5.x86_64.rpm
warning: openssl-1.0.1e-30.el6_6.5.x86_64.rpm: Header V3 RSA/SHA1 Signature, key ID c105b9de: NOKEY
Preparing...                ########################################### [100%]
   1:openssl                ########################################### [ 50%]
   2:openssl-devel          ########################################### [100%]
```

(RPMs are available at:
http://mirror.centos.org/centos/6/updates/x86_64/Packages/openssl-1.0.1e-30.el6_6.5.x86_64.rpm and

http://mirror.centos.org/centos/6/updates/x86_64/Packages/openssl-devel-1.0.1e-30.el6_6.5.x86_64.rpm)

Note This requires krb5-devel and zlib-devel as dependencies. If not installed, install it as follows on the nodes throwing error "yum –y install krb5-devel zlib-devel"

Disable IPv6 Defaults

Disable IPv6 as the addresses used are IPv4.

```
clush -a -b "echo \'net.ipv6.conf.all.disable_ipv6 = 1\' >> /etc/sysctl.conf"
clush -a -b "echo \'net.ipv6.conf.default.disable_ipv6 = 1\' >> /etc/sysctl.conf"
clush -a -b "echo \'net.ipv6.conf.lo.disable_ipv6 = 1\' >> /etc/sysctl.conf"
```

Load the settings from default sysctl file /etc/sysctl.conf

```
clush –a –b "sysctl –p"
```

Configuring Data Drives

This section describes steps to configure non–OS disk drives as individual RAID0 volumes using StorCli command as described below. These volumes are going to be used for MapRFS (HDFS supported) Data.

Issue the following command from the admin node to create the virtual drives with individual RAID 0 configurations on all the nodes.

From the website download storcli:
http://www.lsi.com/downloads/Public/RAID%20Controllers/RAID%20Controllers%20Common%20Files/1.14.12_StorCLI.zip

Extract the zip file and copy storcli-1.14.12-1.noarch.rpm from the linux directory.

1. Download storcli and its dependencies and transfer to Admin node.

   ```
   scp storcli-1.14.12-1.noarch.rpm rhel1:/root/
   ```

2. Copy storcli rpm to all the nodes using the following commands:

   ```
   clush -a -b -c /root/storcli-1.14.12-1.noarch.rpm --dest=/root/
   ```

3. Run the below command to install storcli on all the nodes

   ```
   clush -a -b rpm -ivh storcli-1.14.12-1.noarch.rpm
   ```

4. Run the below command to copy storcli64 to root directory.

   ```
   cd /opt/MegaRAID/storcli/
   cp storcli64 /root/
   ```

```
[root@rhel1 ~]# cd /opt/MegaRAID/storcli/
[root@rhel1 storcli]# ls
install.log  libstorelibir-2.so  libstorelibir-2.so.14.07-0  storcli64
[root@rhel1 storcli]# cp storcli64 /root/
```

5. Copy storcli64 to all the nodes using the following commands:

   ```
   clush -a -b -c /root/storcli64 --dest=/root/
   clush –a –B ./storcli64 –cfgeachdskraid0 WB RA direct NoCachedBadBBU strpsz1024 –a0
   ```

WB: Write back

RA: Read Ahead

NoCachedBadBBU: Do not write cache when the BBU is bad.

Strpsz1024: Strip Size of 1024K

Note The command above will not override any existing configuration. To clear and reconfigure existing configurations refer to Embedded MegaRAID Software Users Guide available at www.lsi.com

Cluster Verification and Micro-Benchmark

This section provides a set of micro–benchmarks and prerequisites scripts to verify that all the systems are configured correctly:

- Prerequisite script to verify configuration across the cluster
- STREAM benchmark to test memory bandwidth
- RPCtest to test network bandwidth
- IOzone to test I/O

Running these tests is optional. Test results can vary based on topology and configuration.

Running Cluster Verification Script

The section describes the steps to create the script cluster_verification.sh that helps to verify CPU, memory, NIC, storage adapter settings across the cluster on all nodes. This script also checks additional prerequisites such as NTP status, SELinux status, ulimit settings, JAVA_HOME settings and JDK version, IP address and hostname resolution, Linux version and firewall settings.

Create script cluster_verification.sh as follows on the Admin node (rhel1)

```
vi cluster_verification.sh

#!/bin/bash

shopt -s expand_aliases

# Setting Color codes
green='\e[0;32m'
red='\e[0;31m'
NC='\e[0m' # No Color

echo -e "${green} === Cisco UCS Integrated Infrastructure for Big Data \ Cluster
Verification === ${NC}"
echo ""
echo ""
echo -e "${green} ==== System Information  ==== ${NC}"
echo ""
echo ""
echo -e "${green}System ${NC}"
clush -a -B " `which dmidecode` |grep -A2 '^System Information'"
echo ""
echo ""
echo -e "${green}BIOS ${NC}"
clush -a -B " `which dmidecode` | grep -A3 '^BIOS I'"
echo ""
```

```
echo “”
echo -e “${green}Memory ${NC}”
clush -a -B “cat /proc/meminfo | grep -i ^memt | uniq”
echo “”
echo “”
echo -e “${green}Number of Dimms ${NC}”
clush -a -B “echo -n 'DIMM slots: ';  `which dmidecode` |grep -c \
'^[[:space:]]*Locator:'”
clush -a -B “echo -n 'DIMM count is: ';  `which dmidecode` | grep \ “Size“| grep -c
“MB””
clush -a -B “ `which dmidecode` | awk '/Memory Device$/,/^$/ {print}' |\ grep -e
'^Mem' -e Size: -e Speed: -e Part | sort -u | grep -v -e 'NO \ DIMM' -e 'No Module
Installed' -e Unknown”
echo “”
echo “”
# probe for cpu info #
echo -e “${green}CPU ${NC}”
clush -a -B “grep '^model name' /proc/cpuinfo | sort -u”
echo “”
clush -a -B “`which lscpu` | grep -v -e op-mode -e ^Vendor -e family -e\ Model: -e
Stepping: -e BogoMIPS -e Virtual -e ^Byte -e '^NUMA node(s)'“
echo “”
echo “”
# probe for nic info #
echo -e “${green}NIC ${NC}”
clush -a -B “`which ifconfig` | egrep '(^e|^p)' | awk '{print \$1}' | \ xargs -l
`which ethtool` | grep -e ^Settings -e Speed”
echo “”
clush -a -B “`which lspci` | grep -i ether”
echo “”
echo “”
# probe for disk info #
echo -e “${green}Storage ${NC}”
clush -a -B “echo 'Storage Controller: '; `which lspci` | grep -i -e \ raid -e storage
-e lsi”
echo “”
clush -a -B “dmesg | grep -i raid | grep -i scsi”
echo “”
clush -a -B “lsblk -id | awk '{print \$1,\$4}'|sort | nl”
echo “”
echo “”

echo -e “${green} ================ Software  ======================== ${NC}“
echo “”
echo “”
echo -e “${green}Linux Release ${NC}”
clush -a -B “cat /etc/*release | uniq”
echo “”
echo “”
echo -e “${green}Linux Version ${NC}”
clush -a -B “uname -srvm | fmt”
echo “”
echo “”
echo -e “${green}Date ${NC}”
clush -a -B date
echo “”
echo “”
echo -e “${green}NTP Status ${NC}”
clush -a -B “ntpstat 2>&1 | head -1”
echo “”
echo “”
echo -e “${green}SELINUX ${NC}”
clush -a -B “echo -n 'SElinux status: '; grep ^SELINUX= \ /etc/selinux/config 2>&1”
echo “”
```

```
echo “”
echo -e “${green}IPTables ${NC}”
clush -a -B “`which chkconfig` --list iptables 2>&1”
echo “”
clush -a -B “ `which service` iptables status 2>&1 | head -10”
echo “”
echo “”
echo -e “${green}Transparent Huge Pages ${NC}”
clush -a -B “ cat /sys/kernel/mm/*transparent_hugepage/enabled”
echo “”
echo “”
echo -e “${green}CPU Speed${NC}”
clush -a -B “echo -n 'CPUspeed Service: ';  `which service` cpuspeed \ status 2>&1”
clush -a -B “echo -n 'CPUspeed Service: '; `which chkconfig` --list \ cpuspeed 2>&1”
echo “”
echo “”
echo -e “${green}Java Version${NC}”
clush -a -B 'java -version 2>&1; echo JAVA_HOME is ${JAVA_HOME:-Not \ Defined!}'
echo “”
echo “”
echo -e “${green}Hostname Lookup${NC}”
clush -a -B “ ip addr show“
echo “”
echo “”
echo -e “${green}Open File Limit${NC}”
clush -a -B 'echo -n “Open file limit(should be >32K): “; ulimit -n'
# MapR related RPMs
clush -a -B 'rpm -qa | grep -i nfs |sort'

clush -a -B 'rpm -qa | grep -i nfs |sort'
clush -a -B 'echo Missing RPMs: ; for each in make patch redhat-lsb irqbalance
syslinux hdparm sdparm dmidecode nc; do rpm -q $each | grep “is not installed“;  done'
clush -a -B “ls -d /opt/mapr/* | head”
# mapr login for hadoop
clush -a -B 'echo “mapr login for Hadoop”; getent passwd mapr'
clush -a -B 'echo “Root login”; getent passwd root'
exit
```

Change permissions to executable

chmod 755 cluster_verification.sh

Run the Cluster Verification tool from the admin node. This can be run before starting Hadoop to identify any discrepancies in Post OS Configuration between the servers or during troubleshooting of any cluster / Hadoop issues.

./cluster_verification.sh

Running STREAM Benchmark

The STREAM benchmark measures sustainable memory bandwidth (in MB/s) and the corresponding computation rate for simple vector kernels. To download the STREAM benchmark, see:

http://www.cs.virginia.edu/stream/

Follow these steps to run the STREAM benchmark:

1. Log on to the admin node. Copy and extract STREAM file to each node (/root/).

```
clush -B -a “tar -xvf stream.tgz”
```

2. Run the following command to run the STREAM benchmark on all nodes:

```
clush –B –a "/root/stream/runme.sh > /root/stream.log"
```

3. Run the following command to verify the results:

```
Extract the five lines of the result as shown and verify it on all the nodes.
$clush –B –a "grep –A5 \"Function     \" stream.log"
---------------
rhel1
---------------
Function      Rate (MB/s)   Avg time     Min time     Max time
Copy:        53289.0222       0.0241       0.0240       0.0243
Scale:        73664.0430       0.0175       0.0174       0.0177
Add:          75339.0246       0.0257       0.0255       0.0259
Triad:        76845.8770       0.0252       0.0250       0.0254
```

Note Results can vary based on the configuration.

Running RPCtest

MapR RPCtest is network bandwidth measurement test. In this solution the methodology adopted to verify the network bandwidth across the cluster requires configuring half the nodes as senders and remaining half as receivers. This test is included in MapR software available at /opt/mapr/servers/tools/rpctest as part of the installation.

Follow the steps below to run RPCtest:

1. Log on to the admin node and run the following commands to create the script:

```
#!/bin/bash
# Define sender nodes
# 8 servers in each rack act as servers and the other half as clients
senders=( 192.168.12.11 192.168.12.12 192.168.12.13 192.168.12.14
192.168.12.19 192.168.12.20 192.168.12.21 192.168.12.22
192.168.12.27 192.168.12.28 192.168.12.29 192.168.12.30
192.168.12.35 192.168.12.36 192.168.12.37 192.168.12.38
192.168.12.43 192.168.12.44 192.168.12.45 192.168.12.46
192.168.12.51 192.168.12.52 192.168.12.53 192.168.12.54
192.168.12.59 192.168.12.60 192.168.12.61 192.168.12.62
192.168.12.67 192.168.12.68 192.168.12.69 192.168.12.70 )
for node in "${half1[@]}"; do
  ssh –n $node /opt/mapr/servers/tools/rpctest –server &
done
sleep 9 # let the servers set up
# Define receiver nodes
receivers=( 192.168.12.15 192.168.12.16 192.168.12.17 192.168.12.18
192.168.12.23 192.168.12.24 192.168.12.25 192.168.12.26
192.168.12.31 192.168.12.32 192.168.12.33 192.168.12.34
192.168.12.39 192.168.12.40 192.168.12.41 192.168.12.42
192.168.12.47 192.168.12.48 192.168.12.49 192.168.12.50
192.168.12.55 192.168.12.56 192.168.12.57 192.168.12.58
192.168.12.63 192.168.12.64 192.168.12.65 192.168.12.66
192.168.12.71 192.168.12.72 192.168.12.73 192.168.12.74 )
i=0
for node in "${receivers[@]}"; do
  ssh –n $node "/opt/mapr/servers/tools/rpctest –client 5000 \ ${senders[$i]} >
rpctest.log" &
  ((i++))
```

```
done
#wait $!   # Comment/uncomment this to make it sequential/concurrent
sleep 5
tmp=${half1[@]}
clush -w ${tmp// /,} pkill rpctest
```

2. Run runRPCtest.sh command from the admin node.
3. Results are generated on receiver nodes. Verify results for all the nodes.

```
$clush -B -w 192.168.12. [19-26, 35-42, 51-58,67-74] cat rpctest.log
---------------
Rhel19
---------------
23:49:42  rpcs 17620,  mb 1150.6
23:49:43  rpcs 17772,  mb 1164.7
23:49:44  rpcs 17771,  mb 1164.6
23:49:45  rpcs 17024,  mb 1115.7
Rate: 1108.93 mb/s, time: 4.73158 sec, #rpcs 80063, rpcs/sec 16921
---------------
```

Note Results can vary based on the topology and configuration.

Running IOzone Benchmark

IOzone is a filesystem benchmark that measures the performance of various I/O operations, such as read, write, re–read, re–write, fread, fwrite, random read and random write.

Warning **IOzone is data destructive. Do not run the test on disks with data.**

Follow these steps to run IOzone benchmark test:

1. Download IOzone from http://www.iozone.org/ and copy to all nodes at /root/.
2. Create the following script, run IOzone.sh on the admin node.

```
#!/bin/bash
# Parallel IOzone tests to stress/measure disk controller
# These tests are destructive therefore
# Test must be run BEFORE MapR filesystem is formatted with disksetup
# Run iozone command once on a single device to verify iozone command
D=$(dirname "$0")
abspath=$(cd "$D" 2>/dev/null && pwd || echo "$D")
# run iozone with -h option for usage, adjust path below for iozone location
# Set list of device names for the 'for' loop
lsblk -id | grep -o ^sd. | sort > /tmp/iozone.disks
for i in `lsblk -i | grep -B2 md[0-1] | grep -v '-' | awk '{print $1}'`;  do sed
-i "/$i/d" /tmp/iozone.disks; done
disks=`cat /tmp/iozone.disks | xargs`
echo $disks
set -x
for disk in $disks; do
echo   $abspath/iozone -I -r 1M -s 80G -i 0 -i 1 -i 2 -f /dev/$disk >
$disk-iozone.log&
   sleep 3 #Some controllers seem to lockup without a sleep
done
```

3. Copy runIOzone.sh to all the nodes at location /root/.
4. Run the following command to start the test:

```
clush -B -a  runIOzone.sh
```

5. Verify that the tests are running and wait for its completion.

```
clush -B -a "ps -aef | grep iozone | wc -l"
---------------
rhel[1-64] (64)
---------------
```

6. Run the following command to verify the test results.

```
Result is generated for each disk as sd<x>-iozone.log, where <x> is the device id.
These logs have sequential and random write and read latencies from each disks.
$ grep " 83886080    " sd*.log
sdb-iozone.log: 83886080    1024 97978 97951 100673 99254 49002 66552
sdc-iozone.log: 83886080   1024 101290 100745 97803 97006 48863 66671
sdd-iozone.log:  83886080    1024 94286 94937 96752 95872 48871 65605
```

Note Results can vary based on configuration.

Installing MapR

Installing MapR software across the cluster involves performing several steps on each node. To make the installation process simpler, start with the installation of core MapR components such as CLDB, MapR–FS, NFS gateway and Yarn. Any additional Hadoop ecosystem components can be easily installed by following instructions on http://doc.mapr.com/display/MapR/Ecosystem+Guide. Follow Figure 91 for role assignments for installation of services on the 64–node cluster.

The following sections describe the steps and options for installing MapR software:

- Preparing Packages and Repositories
- MapR Installation
  - Installing MapR packages
  - Verify successful installation
- Configure the Node with the `configure.sh` Script
- Formatting Disks with the `disksetup` Script

Planning the Cluster

The first step towards deploying the MapR is planning which nodes contribute to the cluster, and selecting the services that will run on each node.

MapR Services

In a typical cluster, most nodes are dedicated to data processing and storage, and a smaller number of nodes run services that provide cluster coordination and management. Some applications run on cluster nodes and others run on client nodes that can communicate with the cluster.

The following table shows some of the services that can be run on a node.

Figure 88 shows the list of MapR services and the corresponding descriptions.

Figure 88 MapR Services

MapReduce | Storage | Management | Application

Service	Description
Warden	Warden runs on every node, coordinating the node's contribution to the cluster.
TaskTracker (optional)	Hadoop TaskTracker starts and tracks MapReduce tasks on a node. The TaskTracker service receives task assignments from the JobTracker service and manages task execution.
NodeManager	Hadoop YARN NodeManager service. The NodeManager manages node resources and monitors the health of the node. It works with the ResourceManager to manage YARN containers that run on the node.
FileServer	FileServer is the MapR service that manages disk storage for MapR–FS and MapR–DB on each node.
CLDB	Maintains the container location database (CLDB) service. The CLDB service coordinates data storage services among MapR–FS FileServer nodes, MapR NFS gateways, and MapR clients.
NFS	Provides read–write MapR Direct Access NFS access to the cluster, with full support for concurrent read and write access.
MapR HBase Client (optional)	Provides access to MapR–DB tables via HBase APIs. Required on all client nodes that will access table data in MapR–FS
JobTracker (optional)	Hadoop JobTracker service. The JobTracker service coordinates the execution of MapReduce jobs by assigning tasks to TaskTracker nodes and monitoring task execution.
ResourceManager	Hadoop YARN ResourceManager service. The ResourceManager manages cluster resources, and tracks resource usage and node health.
ZooKeeper	Enables high availability (HA) and fault tolerance for MapR clusters by providing coordination.
HistoryServer	Archives MapReduce job metrics and metadata.
Web Server	Runs the MapR Control System.
Hue	Hue is Hadoop user interface that interacts with Apache Hadoop and its ecosystem components, such as Hive, Pig, and Oozie.
Pig	Pig is a high–level data–flow language and execution framework.
Hive	Hive is a data warehouse that supports SQL–like ad hoc querying and data summarization.
Flume	Flume is a service for aggregating large amounts of log data
Oozie	Oozie is a workflow scheduler system for managing Hadoop jobs.
Mahout	Mahout is a set of scalable machine–learning libraries that analyze user behavior.
Spark	Spark is a processing engine for large datasets.
Sqoop	Sqoop is a tool for transferring bulk data between Hadoop and relational databases.

Node Types

The MapR installer categorizes nodes as control nodes (which runs only cluster management services to manage the cluster), data nodes, control–as–data nodes (which combine the functions of control and data nodes), or client nodes. For deployment of MapR on Cisco UCS Integrated Infrastructure for Big Data, control services co–exist on data nodes (control–as–data node) as control services have a small footprint. Client node could be any node accessing the MapR cluster (all nodes in the MapR cluster are also client nodes).

Figure 89 Node Types

Node Type	Description
Data node	Used for processing data, so they have the FileServer and TaskTracker services installed. If MapR–DB or HBase is run on a data node, the HBase Client service is also installed. Data nodes are used for running YARN applications and MapReduce jobs, and for storing file and table data. These nodes run the FileServer service along with NodeManager (for YARN nodes), TaskTracker (for MapReduce nodes), and HBase client (for MapR–DB and HBase nodes).
Control–as–data node	Acts as both control and data nodes. They perform both functions and have both sets of services installed.
Client node	Provides access to the cluster so the user can communicate via the command line or the MapR Control System. Client nodes provide access to each node on the cluster so the user can submit jobs and retrieve data. A client node can be an edge node of the cluster, laptop, or any Windows machine.

Node Types and Associated Services

The following table shows which services are assigned to each node type. When deploying MapR on Cisco UCS Integrated Infrastructure for Big Data, all Control Node services are deployed on Control–as–data node. There are no nodes running purely as control nodes as they also run data node services.

Figure 90 Services Assigned on Various Node Types

Node Type	YARN Main Services	Core MapR Services	Additional MRv1 Services	Additional HBase Service
Control–as–data node	ResourceManager (RM) HistoryServer (HS) NodeManager (NM)	CLDB ZooKeeper FileServer NFS Webserver FileServer	JobTracker (optional) TaskTracker (optional)	
Data node	NodeManager (NM)	FileServer	TaskTracker (optional)	
Client node			MapR Client	HBase Client (optional)

Hostnames and Roles

This section describes the cluster plan of a 64–node cluster with hostnames and roles assignments for the following services as shown in Figure 91 below.

- ResourceManager (RM)
- HistoryServer (HS)
- NodeManager (NM)
- TaskTracker (TT, optional)
- JobTracker (JT, optional), FileServer (FS)
- Container Location Database (CLDB)

- Zookeeper, and
- Webserver

Note Starting with MapR version 4.0, both Yarn and MapReduce V1 are supported not only in the same cluster but also on the same node.

Figure 91 Host names and Role Assignment

Rack-1 Hostnames	MapR Roles	Rack-2 Hostnames	MapR Roles	Rack-3 Hostnames	MapR Roles	Rack-4 Hostnames	MapR Roles
rhel1	CLDB,FS, NM, NFS,HS	rhel17	CLDB, FS, NM, NFS, HS	rhel33	CLDB, FS, NM, NFS,HS	rhel49	FS, NM, NFS
rhel2	ZooKeeper FS, NM, NFS	rhel18	ZooKeeper FS, NM, NFS	rhel34	ZooKeeper FS, NM, NFS	rhel50	FS, NM, NFS
rhel3	Webserver, FS, NM, NFS	rhel19	Webserver, FS, NM, NFS	rhel35	Webserver, FS, NM, NFS	rhel51	Webserver, FS, NM, NFS
rhel4	FS, NM, NFS	rhel20	FS, NM, NFS	rhel36	FS, NM, NFS	rhel52	FS, NM, NFS
rhel5	FS, NM, NFS, RM	rhel21	FS, NM, NFS, RM	rhel37	FS, NM, NFS, RM	rhel53	FS, NM, NFS
rhel6	FS, NM, NFS	rhel22	FS, NM, NFS	rhel38	FS, NM, NFS	rhel54	FS, NM, NFS
rhel7	FS, NM, NFS	rhel23	FS, NM, NFS	rhel39	FS, NM, NFS	rhel55	FS, NM, NFS
rhel8	FS, NM, NFS	rhel24	FS, NM, NFS	rhel40	FS, NM, NFS	rhel56	FS, NM, NFS
rhel9	FS, NM, NFS	rhel25	FS, NM, NFS	rhel41	FS, NM, NFS	rhel57	FS, NM, NFS
rhel10	FS, NM, NFS	rhel26	FS, NM, NFS	rhel42	FS, NM, NFS	rhel58	FS, NM, NFS
rhel11	FS, NM, NFS	rhel27	FS, NM, NFS	rhel43	FS, NM, NFS	rhel59	FS, NM, NFS
rhel12	FS, NM, NFS	rhel28	FS, NM, NFS	rhel44	FS, NM, NFS	rhel60	FS, NM, NFS
rhel13	FS, NM, NFS	rhel29	FS, NM, NFS	rhel45	FS, NM, NFS	rhel61	FS, NM, NFS
rhel14	FS, NM, NFS	rhel30	FS, NM, NFS	rhel46	FS, NM, NFS	rhel62	FS, NM, NFS
rhel15	FS, NM, NFS	rhel31	FS, NM, NFS	rhel47	FS, NM, NFS	rhel63	FS, NM, NFS
rhel16	FS, NM, NFS	rhel32	FS, NM, NFS	rhel48	FS, NM, NFS	rhel64	FS, NM, NFS

Note All Job management are performed by Resource Manager and Node Manager. In this CVD, TaskTracker and JobTracker are not installed.

Preparing Packages and Repositories

A local repository on the admin node is set up to provide access to installation packages. With this method, the package manager on each node retrieves the installations package from the admin node (rhel1 is used as admin node as already mentioned) and installs the packages. Nodes do not need to have an Internet access.

Below are instructions on setting up a local repository for Red Hat Linux distribution. These instructions create a single repository that includes both MapR components and the Hadoop ecosystem components.

RPM Repositories for MapR Core Software

MapR hosts `rpm` repositories for installing the MapR core software using Linux package management tools. For every release of the core MapR software, a repository is created for each supported platform.

These platform–specific repositories are hosted at:
http://package.mapr.com/releases/<version>/<platform>

http://package.mapr.com/releases/v4.1.0/redhat/mapr–v4.1.0GA.rpm.tgz

http://archive.mapr.com/releases/ecosystem–all/redhat/mapr–ecosystem–20150503.rpm.tgz

RPM Repositories for Hadoop Ecosystem Tools

MapR hosts `rpm` repositories for installing Hadoop ecosystem tools, such as Spark, Flume, Hive, Mahout, Oozie, Pig and Sqoop. At any given time, MapR's recommended versions of ecosystem tools that work with the latest version of MapR core software are available in the link below.

These platform–specific repositories are hosted at:
http://package.mapr.com/releases/ecosystem/<platform>

To create the local repositories follow the steps below:

1. Login as `root` on the admin node (rhel1).
2. Create the following directory on rhel1

```
mkdir -p /var/www/html/mapr.local
```

3. On a node that is connected to the Internet, download the following files, substituting the appropriate `<version>` and `<datestamp>`:

```
wget
http://package.mapr.com/releases/v<version>/redhat/mapr-v<versi
on>GA.rpm.tgz

wget
http://package.mapr.com/releases/ecosystem/redhat/mapr-ecosystem-<datestamp>.rpm.t
gz
```

Note For this document we use the version 4.1.0. See MapR Repositories and Package Archives for the correct paths for all past releases at: http://archive.mapr.com/releases/

```
[root@internet-host ~]# wget
http://package.mapr.com/releases/v4.1.0/redhat/mapr-v4.1.0GA.rp
m.tgz

[root@internet-host ~]# wget
http://package.mapr.com/releases/ecosystem/redhat/mapr-ecosystem-20150420.rpm.tgz
```

Note The server `internet-host` is an edge host that has access to the Internet and to the admin node (rhel1). It is not a part of the MapR cluster. It is only used to download and transfer files to the admin node from the Internet as the admin node is not directly connected to the Internet.

4. Copy the files to `/var/www/html/mapr.local` on the admin node, and extract them there.

```
[root@internet-host ~]# scp mapr-v4.1.0GA.rpm.tgz
rhel1:/var/www/html/mapr.local/
[root@internet-host ~]# scp mapr-ecosystem-20150420.rpm.tgz
rhel1:/var/www/html/mapr.local/
```

```
[root@rhel1 mapr.local]# tar -xvzf mapr-v4.1.0GA.rpm.tgz
[root@rhel1 mapr.local]# tar -xvzf
mapr-ecosystem-20150420.rpm.tgz
```

5. Create the base repository headers:

```
[root@rhel1 mapr.local]# createrepo /var/www/html/mapr.local
```

```
[root@rhel1 mapr.local]# createrepo /var/www/html/mapr.local
Spawning worker 0 with 128 pkgs
Workers Finished
Gathering worker results

Saving Primary metadata
Saving file lists metadata
Saving other metadata
Generating sqlite DBs
Sqlite DBs complete
```

To add the repository on each node follow the steps below:

6. Create repo file `/etc/yum.repos.d/maprtech.repo` on the admin node (rhel1):

```
vi /etc/yum.repos.d/maprtech.repo
[maprtech]
name=MapR Technologies, Inc.
baseurl=http://10.29.160.101/mapr.local
enabled=1
gpgcheck=0
```

```
[maprtech]
name=MapR Technologies, Inc.
baseurl=http://10.29.160.101/mapr.local
enabled=1
gpgcheck=0
```

7. Copy the `maprtech.repo` specification to all the nodes of the cluster. Then, update the yum metadata cache so that the repository files will be properly accessed.

```
clush -a -c /etc/yum.repos.d/maprtech.repo
clush -a yum makecache
```

```
clush a -c /etc/yum.repos.d/maprtech.repo
clush -a yum makecache
```

8. Create mapr user across all nodes

Users of the cluster must have the same credentials and uid on every node in the cluster. Each user (or department) that runs the MapR jobs needs an account and must belong to a common group (gid). If a directory service, such as LDAP, is not used, this user is created on each node. Every user must have the same uid and primary gid on every node.

In addition, a MapR user with full privileges to administer the cluster is created. If a user named 'mapr' does not exist. It is recommended that the user named 'mapr' is created in advance in order to test the connectivity issues prior to the installation step.

```
clush -a groupadd -g 5000 mapr
```

```
clush -a "useradd -g 5000 -u 5000 mapr"

clush -a  -B "echo maprpassw0rd | passwd mapr --stdin"
```

```
clush -a groupadd -g 5000 mapr
clush -a "useradd -g 5000 -u 5000 mapr"
clush -a  -B "echo maprpassw0rd | passwd mapr --stdin"
```

Note Password of mapr user is set to maprpassw0rd.

9. Verify mapr user on all nodes.

```
clush -a -B id mapr
```

```
[root@rhel1 ~]# clush -a -B id mapr
uid=5000(mapr) gid=5000(mapr) groups=5000(mapr)
```

MapR Software Installation

Perform the following steps on each node:

1. **Install** the planned MapR services as shown in Figure 92.
2. Run the `configure.sh` script to **configure** the node.
3. **Format** raw drives and partitions allocated to MapR using the `disksetup` script.

Figure 92 shows services and corresponding packages.

Figure 92 MapR Services and Packages

Service	Package
MapR core	mapr-core
Cluster location DB (CLDB)	mapr-cldb
History server	mapr-historyserver
ResourceManager and/or JobTracker	mapr-resourcemanager and/or mapr-jobtracker
MapR Control System	mapr-webserver
MapR File Server	mapr-fileserver
NFS	mapr-nfs
NodeManager and/or TaskTracker	mapr-nodemanager and/or mapr-tasktracker
ZooKeeper	mapr-zookeeper
Hadoop Ecosystem Components	**Package**
Drill	mapr-drill
Spark	mapr-spark
Hive	mapr-hive
Mahout	mapr-mahout
Oozie	mapr-oozie
Pig	mapr-pig
Sqoop	mapr-sqoop

Installing MapR packages

Based on the Cluster Plan for which services to run on which nodes, as shown in Table 8 above, use the commands in this section to install the appropriate packages for each node. Configuring the local yum repository ensures that the package dependencies will be managed correctly.

Installing CLDB

```
clush –B -w rhel[1,17,33] 'yum –y install mapr-cldb'
```

Installing ResourceManager

```
clush –B -w rhel[5,21,37] 'yum –y install mapr-resourcemanager'
```

Installing Mapr Webserver

```
clush –B -w rhel[3,19,35,51] 'yum –y install mapr-webserver'
```

Note Make sure httpd is not installed on these nodes.

Installing Mapr–Zookeeper

```
clush –B -w rhel[2,18,34] 'yum –y install mapr-zookeeper'
```

Installing NFS, Fileserver and Nodemanager on all cluster nodes

```
clush –B -a 'yum –y install mapr-fileserver mapr–nfs mapr-nodemanager'
```

```
clush -B -w rhel[1,17,33] 'yum -y install mapr-cldb'
clush -B -w rhel[5,21,37] 'yum -y install mapr-resourcemanager'
clush -B -w rhel[3,19,35,51] 'yum -y install mapr-webserver'
clush -B -w rhel[2,18,34] 'yum -y install mapr-zookeeper'
clush -B -a 'yum -y install mapr-fileserver mapr-nfs mapr-nodemanager'
```

Configure mapr–nfs

Run the following commands from the admin node (rhel1):

```
clush -a mkdir -p /mapr
echo “localhost:/mapr  /mapr  hard,nolock” > /opt/mapr/conf/mapr_fstab
clush -a -c /opt/mapr/conf/mapr_fstab --dest /opt/mapr/conf/mapr_fstab
```

```
clush -a mkdir -p /mapr
echo "localhost:/mapr  /mapr  hard,nolock" > /opt/mapr/conf/mapr_fstab
clush -a -c /opt/mapr/conf/mapr_fstab --dest /opt/mapr/conf/mapr_fstab
```

Verification

To verify that the software has been installed successfully, check the `/opt/mapr/roles` directory on each node. The software is installed in directory `/opt/mapr` and a file is created in `/opt/mapr/roles` for every service that installs successfully. Examine this directory to verify installation for the node. For example:

```
# clush -a -B “ls -l /opt/mapr/roles”
```

Configure the Node with the configure.sh Script

⚠ Configure the node first, then prepare raw disks and partitions with the disksetup command.

The script `configure.sh` configures a node to be part of a MapR cluster, or modifies services running on an existing node in the cluster. The script creates (or updates) configuration files related to the cluster and the services running on the node. Before performing this step, make sure to have a list of the hostnames of the CLDB and ZooKeeper nodes, Optionally specify the ports for the CLDB and ZooKeeper nodes as well. If not specified, the default ports are assigned as:

- CLDB – 7222
- ZooKeeper – 5181

The script `configure.sh` takes an optional cluster name and log file, and comma–separated lists of CLDB and ZooKeeper host names or IP addresses (and optionally ports), using the following syntax:

```
/opt/mapr/server/configure.sh -C <host>[:<port>][,<host>[:<port>]...] -Z
<host>[:<port>][,<host>[:<port>]...] [-L <logfile>][-N <cluster name>]
```

Configure All nodes with CLDB and Zookeeper locations

```
clush -B -a '/opt/mapr/server/configure.sh -C rhel1,rhel17,rhe33 -Z
rhel2,rhel18,rhel34 -HS rhel1 -RM rhel5,rhel21,rhel37 -N ciscomapr -no-autostart'
```

```
clush -B -a '/opt/mapr/server/configure.sh -C rhel1,rhel17,rhe33 -Z rhel2,rhel18
,rhel34 -HS rhel1 -RM rhel5,rhel21,rhel37 -N ciscomapr -no-autostart'
```

Formatting Disks with the disksetup Script

If `mapr-fileserver` is installed on this node, use the following procedure to format disks and partitions to be used by MapR.

⛔ Run the configure.sh script (described above) before running disksetup.

The `disksetup` script is used to format disks to be used by the MapR cluster. Create a text file `/tmp/MapR.disks` listing the disks and partitions to be used by MapR on the node. Each line lists either a single disk or all applicable partitions on a single disk. When listing multiple partitions on a line, separate them by spaces.

Identify and format the data disks for MapR

- Create a list of disks to be formatted:

Create the following script on rhel1 and copy it to all the nodes:

```
vim mapr_disks.sh

#!/bin/bash
#This script creates file (MapR.disks) containing list of non-os disk #drives used
during MapR Installation
[[ "-x" == "${1}" ]] && set -x && set -v && shift 1
count=1
for HD in /sys/class/scsi_host/host?/scan
do
echo '- - -' > ${HD}
done
for HD in /dev/sd?
do
if [[ -b ${HD} && `/sbin/parted -s ${HD} print quit|/bin/grep -c boot` -ne 0
]]
then
continue
else
echo $HD >> /tmp/MapR.disks
fi
done
```

Copy the scipt to nodes:

```
chmod +x mapr_disks.sh
clush -a -c  mapr_disks.sh
clush -a -B /root/mapr_disks.sh
```

Verify the file on all nodes does not contain os drives:

```
clush -aB cat /tmp/MapR.disks
```

- Confirm that the disks are not in use
  The cfdisk, mount, and pvdisplay utilities can be used to confirm that the system is not using the disks listed in /tmp/MapR.disks. This confirmation is not necessary during the initial setup, but may be relevant when nodes are removed or re–added to the cluster.
  Format the disks to MapR-FS

```
clush -B -a '/opt/mapr/server/disksetup -F  /tmp/MapR.disks'
```

The script disksetup removes all data from the specified disks. Make sure to specify the disks correctly, and that all data has been backed up elsewhere.

This procedure assumes free, un-mounted physical partitions or hard disks for use by MapR.

After successful installation of MapR software on each node according to the cluster plan, bring up the cluster.

Updating environment variables in /opt/mapr/conf/env.sh

There are a few key environment variables for the MapR software saved in /opt/mapr/conf/env.sh. These values must be properly configured BEFORE launching the cluster software. The default file is as shown below:

```
#!/bin/bash
# Copyright (c) 2009 & onwards. MapR Tech, Inc., All rights reserved
# Please set all environment variable you want to be used during MapR cluster
# runtime here.
# namely MAPR_HOME, JAVA_HOME, MAPR_SUBNETS

#export JAVA_HOME=
#export MAPR_SUBNETS=
#export MAPR_HOME=
```

```
#export MAPR_ULIMIT_U=
#export MAPR_ULIMIT_N=
#export MAPR_SYSCTL_SOMAXCONN=
```

For this deployment, we need to explicitly set values for JAVA_HOME and MAPR_SUBNETS as shown below;

```
export JAVA_HOME=/usr/java/jdk1.7.0_75
export MAPR_SUBNETS=192.168.11.0/24,192.168.12.0/24
```

Note By mentioning MAPR_SUBNETS and providing the two vlans, this enables MapR to use both VLANs (NICs) for traffic and thus using full 20 GiGE for Hadoop traffic.

Make those changes in `rhel1:/opt/mapr/conf/env.sh` and then distribute them to the entire cluster with the command

```
$ clush -B -a -c /opt/mapr/conf/env.sh
```

Bringing Up the Cluster

The installation of software across a cluster of nodes will go more smoothly if the services have been pre-planned and each node has been validated. Referring to the cluster design developed in section "Planning the Cluster", ensure that each node has been prepared and that the MapR packages have been installed on each node in accordance with the plan. The process for launching the cluster can be broken down into several steps:

- Initialization Sequence
- Troubleshooting
- Installing the Cluster License
- Verifying Cluster Status
- Setting up a multi–tenancy MapR cluster

The initialization sequence involves starting the ZooKeeper service, starting the CLDB service, setting up the administrative user, and installing a MapR license. Once these initial steps are done, the cluster is functional on a limited set of nodes. Not all services are started yet, but the MapR Control System Dashboard, or the MapR Command Line Interface are available, to examine nodes and activity on the cluster.

Initialization Sequence

First, start the ZooKeeper service. It is important that all ZooKeeper instances start up, because the rest of the system cannot start unless a majority of ZooKeeper instances are up and running. Next, start the **warden** service on each node, or at least on the nodes that host the CLDB and webserver services. The warden service manages all MapR services on the node (except ZooKeeper) and helps coordinate communications. Starting the warden automatically starts the CLDB.

To bring up the cluster, follow these steps:

1. Start **ZooKeeper** on all nodes where it is installed, by issuing one of the following commands:

```
clush -B -w rhel[2,18,34] service mapr-zookeeper start
```

```
JMX enabled by default
Using config: /opt/mapr/zookeeper/zookeeper-3.4.5/conf/zoo.cfg
Starting zookeeper ... STARTED
```

2. Verify that the ZooKeeper service is running properly:

```
clush –B -w rhel[2,18,34] service mapr-zookeeper status
```

```
JMX enabled by default
Using config: /opt/mapr/zookeeper/zookeeper-3.4.5/conf/zoo.cfg
zookeeper running as process 1287.
```

The servers should display the running pid for the zookeeper process

3. On the nodes running CLDB or webserver, start the **warden** by issuing one of the following commands

```
clush -a service mapr-warden start
```

```
[root@rhel1 ~]# clush -a service mapr-warden start
rhel3.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel3.mgmt:
rhel3.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel2.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel2.mgmt:
rhel2.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel1.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel1.mgmt:
rhel1.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel4.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel4.mgmt:
rhel4.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel8.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel8.mgmt:
rhel8.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel7.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel7.mgmt:
rhel7.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel5.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel5.mgmt:
rhel5.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
rhel6.mgmt: Starting WARDEN, logging to /opt/mapr/logs/warden.log.
rhel6.mgmt:
rhel6.mgmt: For diagnostics look at /opt/mapr/logs/ for createsystemvolumes.log, warden.log and configure
 services log files
```

Before continuing, wait 30 to 60 seconds for the warden to start the CLDB service. Calls to MapR (such as maprcli) may fail if executed before the CLDB has started successfully.

1. Log in to rhel1 and issue the following command to give full permission to the chosen administrative user mapr:

```
clush -B –w rhel1 /opt/mapr/bin/maprcli acl edit -type cluster –user mapr:fc
```

2. Confirm that the cluster is up before bringing up other nodes

```
clush -B –w rhel[17,33] 'hadoop fs -ls /'
```

```
[root@rhel1 ~]# hadoop fs -ls /
Found 8 items
drwx------   - devuser1 dev          0 2015-05-06 00:08 /DEV
drwx------   - qauser1  qa           0 2015-05-05 22:30 /QA
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /apps
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /hbase
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /opt
drwxrwxrwx   - mapr     mapr         3 2015-05-05 22:05 /tmp
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /user
drwxr-xr-x   - mapr     mapr         1 2015-04-28 22:55 /var
```

Troubleshooting

Difficulty bringing up the cluster seems daunting, but most cluster problems are easily resolved. For the latest support tips, visit http://answers.mapr.com.

- Can each node connect with the others? For a list of ports that must be open, see http://answers.mapr.com.
- Is the warden running on each node? On the node, run the following command as root:

  ```
  $ service mapr-warden status
  WARDEN running as process 18732
  ```

  If the warden service is not running, check the warden log file, `/opt/mapr/logs/warden.log`, for clues.

  To restart the warden service:

  ```
  $ service mapr-warden start
  ```

- The ZooKeeper service is not running on one or more nodes
  - Check the warden log file for errors related to resources, such as low memory
  - Check the warden log file for errors related to user permissions
  - Check for DNS and other connectivity issues between ZooKeeper nodes
- The MapR CLI program `/opt/mapr/bin/maprcli` won't run
  - Did you configure this node? See Installing MapR Software.
- Permission errors appear in the log
  - Check that MapR changes to the following files have not been overwritten by automated configuration management tools:

Table 8 MapR Dependant Files

`/etc/sudoers`	Allows the `mapr` user to invoke commands as root
`/etc/security/limits.conf`	Allows MapR services to increase limits on resources such as memory, file handles, threads and processes, and maximum priority level
`/etc/udev/rules.d/99-mapr-disk.rules`	Covers permissions and ownership of raw disk devices

Before contacting Support, collect cluster's logs using the `mapr-support-collect` script.

Installing Cluster License

Note Contact MapR sales representative to obtain a valid MapR license key. This is necessary to enable the enterprise–class features of the MapR packages (eg MapR-DB, NFS, ResourceManager HA, storage snapshots and mirrors, etc.).

Using web–based MCS to install the license:

1. On a machine that is connected to the cluster and to the Internet, perform the following steps to open the MapR Control System and install the license:

2. In a browser, view the MapR Control System by navigating to the node that is running the MapR Control System. For example, `rhel3`.

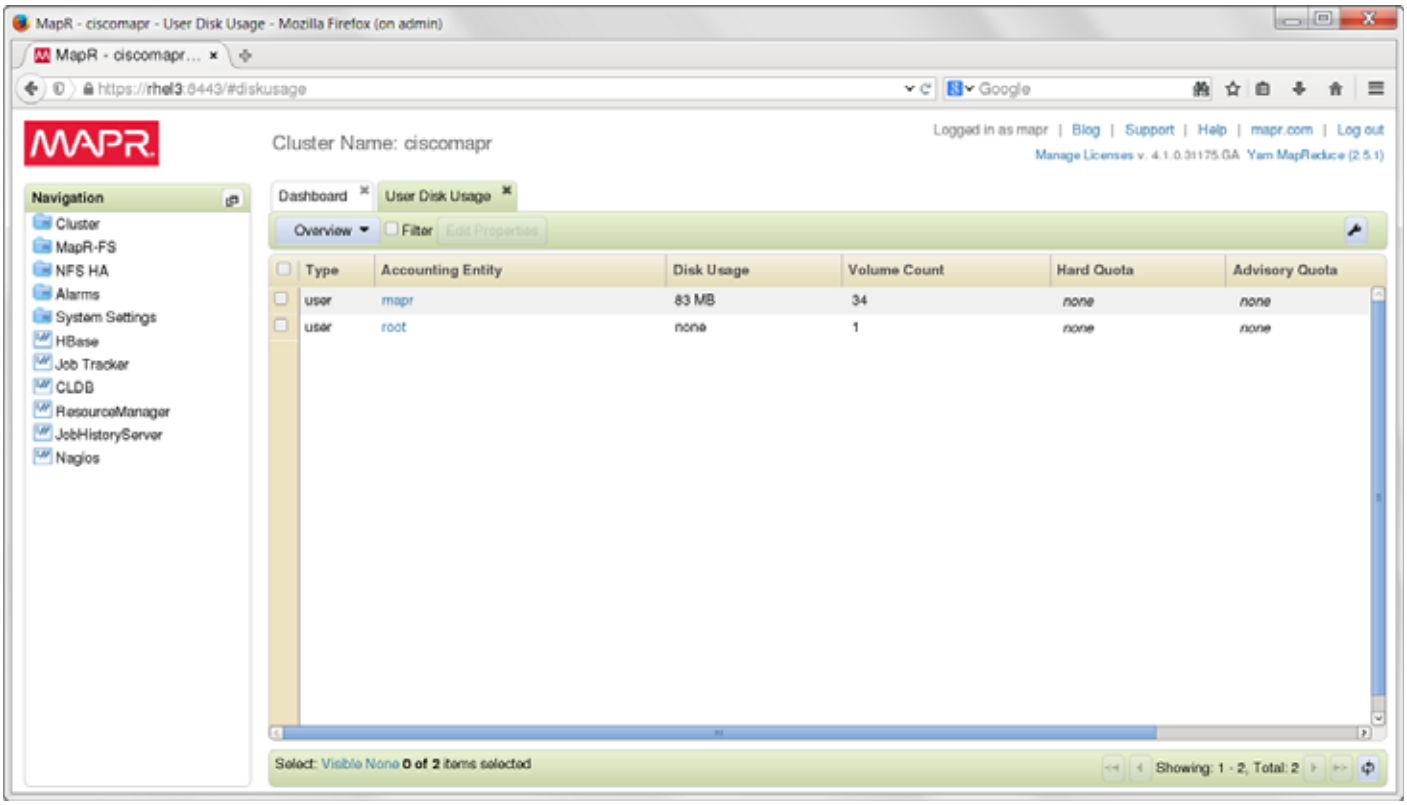
   ```
   https://<MCS node>:8443
   ```

   The node won't have an HTTPS certificate yet, so the browser will warn that the connection is not trustworthy. Ignore the warning this time. The first time MapR starts, accept the Terms of Use and choose whether to enable the MapR Dial Home service.

3. Log in to the MapR Control System as the administrative user. Until a license is applied, the MapR Control System dashboard might show some nodes in the amber "degraded" state.

Note The nodes health will be in amber until the license is applied. Once the license is applied, the node health should come up as green.

Figure 93 MapR Control System

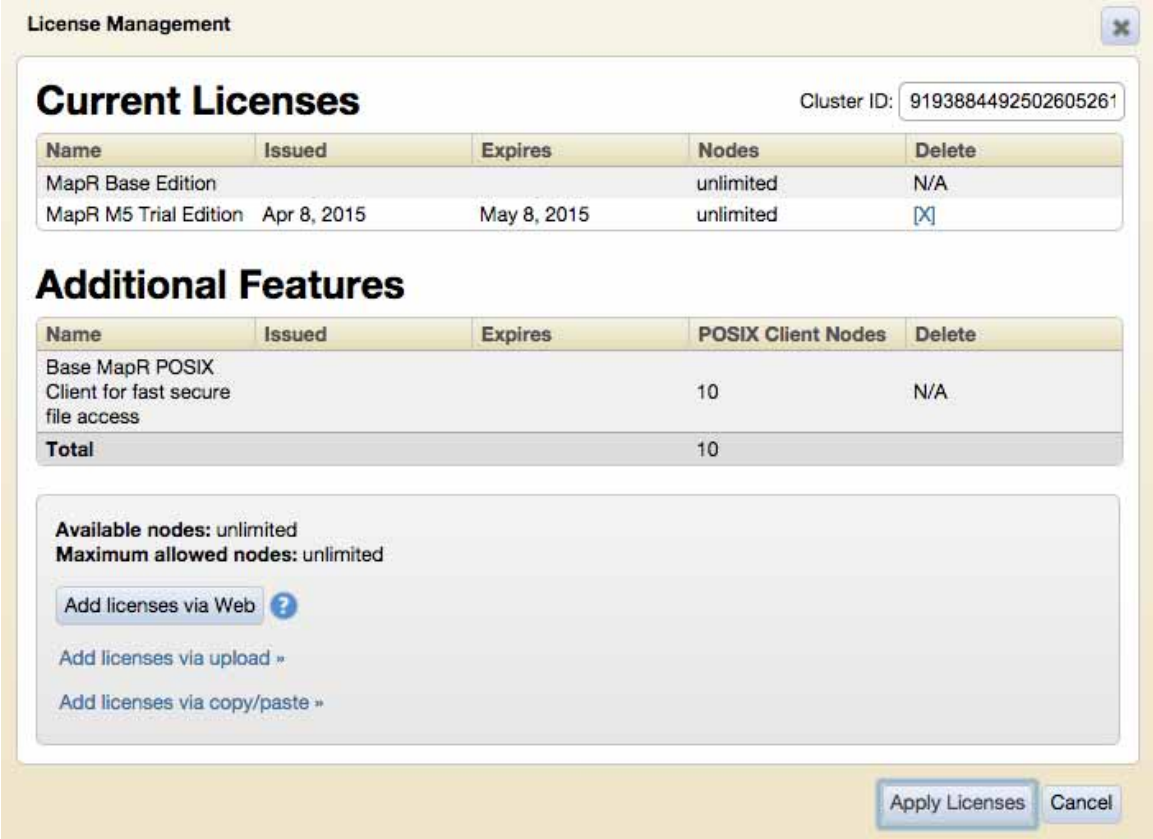


4. In the navigation pane of the MapR Control System, expand the System Settings Views group and click Manage Licenses to display the MapR License Management dialog.

   a. Click **Add Licenses via copy/paste and paste the license key.**

   b. If the cluster is already registered, the license is applied automatically. Otherwise, click **OK** to register the cluster on MapR.com and follow the instructions there.

Figure 94 MapR License Management Dialog



Installing a license from the command line (optional)

Use the following steps if the cluster and the Internet are not accessible at the same time.

1. Obtain a valid license file from MapR
2. Copy the license file to a cluster node
3. Run the following command to add the license:

```
maprcli license add [ -cluster <name> ] -license <filename> -is_file true
```

Restarting MapR services after license installation

Certain HA features of the MapR cluster will not start properly until a valid license is installed. Once successfully installed the trial license or a permanent one provided by mapr, restart the distributed CLDB services, as well as the ResourceManager service and the NFS service. This can be done from any node in the cluster with the following commands:

```
maprcli node services -name cldb -action start -filter "[csvc==cldb]"
maprcli node services -name resourcemanager -action start -filter
"[csvc==resourcemanager]"
maprcli node services -name nfs -action start -filter "[csvc==nfs]"
```

```
maprcli node services name cldb action start filter "[csvc==cldb]"
maprcli node services name resourcemanager action start filter "[csvc==resourcemanager]
maprcli node services name nfs action start filter "[csvc==nfs]"
```

The effect of those commands is to start the respective services on all nodes in the cluster configured with those services. Nodes on which the service is already running will not be affected.

Managing Log Files

Over time, the log directories increase in size. For regular maintenance/compliance, it is a good practice to move the files in the directories mentioned below to different location either in the cluster (on one of the MapR Volumes) or to a different location for any compliance/audit requirements. This action could be setup as a cron job.

On each node, the two directories that contain the log files are:

```
/opt/mapr/hadoop/hadoop-0.20.2/logs
/opt/mapr/hadoop/hadoop-2.5.1/logs
```

Verifying Cluster Status

Verify cluster status using the web interface:

1. Log in to the MapR Control System.
2. Under the **Cluster** group in the left pane, click **Dashboard**.
3. Check the **Services** pane and make sure each service is running the correct number of instances, according to the cluster plan.

Verify cluster status using the command line interface

1. Log in to a cluster node
2. Use the following command to list MapR services:

```
$maprcli service list
$maprcli license list
$maprcli disk list -host <name or IP address>
```

```
root@rhel1 ~]# maprcli disk list -host rhel3
liskname   powerstatus  status  vendor  hostname  modelnum       availablespace  storagepoolid  fstype   mount  firmwareversion
dev/sda1   running      0       Cisco   rhel3     UCSC-MRAID12G  3664                           ext4     1      4.25
dev/sda2   running      0       Cisco   rhel3     UCSC-MRAID12G  1049359                        ext4     1      4.25
dev/sda3   running      0       Cisco   rhel3     UCSC-MRAID12G                                 swap     0      4.25
dev/sdb    running      0       Cisco   rhel3     UCSC-MRAID12G  1143257         1              MapR-FS  0      4.25
dev/sdd    running      0       Cisco   rhel3     UCSC-MRAID12G  1143257         1              MapR-FS  0      4.25
dev/sde    running      0       Cisco   rhel3     UCSC-MRAID12G  1143249         2              MapR-FS  0      4.25
dev/sdc    running      0       Cisco   rhel3     UCSC-MRAID12G  1143257         1              MapR-FS  0      4.25
dev/sdf    running      0       Cisco   rhel3     UCSC-MRAID12G  1143249         2              MapR-FS  0      4.25
dev/sdg    running      0       Cisco   rhel3     UCSC-MRAID12G  1143249         2              MapR-FS  0      4.25
```

Installing Additional Hadoop Components

The final step in installing a MapR cluster is to install and bring up Hadoop ecosystem components such as the following and integrating them with a MapR cluster:

Please refer to MapR Install guide at http://doc.mapr.com/display/MapR/Ecosystem+Guide for detailed instructions on installation and configuration of Hadoop components.

- Apache Drill – Installing and using Drill on a MapR cluster
- Cascading - Installing and using Cascading on a MapR cluster
- Flume – Installing and using Flume on a MapR cluster
- HBase - Installing and using HBase on MapR
- Hive – Installing and using Hive on a MapR cluster, and setting up a MySQL metastore

- Hue - Installing and using Hue on MapR
- Impala – Installing and using Impala on a MapR cluster
- Mahout - Environment variable settings needed to run Mahout on MapR
- Oozie – Installing and using Oozie on a MapR cluster
- Pig - Installing and using Pig on a MapR cluster
- Spark and Shark – Installing and running Spark and Shark on MapR
- Sqoop - Installing and using Sqoop on a MapR cluster
- Storm (Version 0.9.3–1411) – Installing and using Storm on a MapR cluster

Setting up a Multi-Tenancy MapR cluster

After installing the MapR core and any desired Hadoop components, perform the optional steps to prepare the cluster for production with Multi–Tenancy. Review the topics below for next steps that might apply to the cluster.

- Setting up Topology
- Setting up Volumes (with ownership, quota, replication factor, QoS etc)
- Associate volumes with topologies
- Setting up ACL policies for volumes.
- Different Data placement policies for Different QoS with/without heterogeneous Servers based on SLA
- Different Replication Factor
- Different Latencies for Data replication
- Access control (two different user groups can't or can access data between the groups through Access control)
- Job Placement (current consideration) with/without heterogeneous Servers based on SLA
- Administration and Reporting

Note With Multi-Tenancy, the cluster could be a heterogeneous cluster with Cisco UCS C240 M4 with SFF (Small Form Factor) drives as in this CVD, Cisco UCS C240 M4 with LFF (Large Form Factor) drives and Cisco UCS C3160 (with 60 LFF drives). Based on the SLA requirements different servers can be part of the different tenants.

Creating user groups for Multi–Tenancy

Here we provide an example to setup tenents (QA) and (DEV) and assign it to a two user groups. Tenets are separated from each other through a concept of "Volumes" which will be described in detail in the following sections. Users of the cluster must have the same credentials and uid on every node in the cluster. Each user (or department) that runs the MapR jobs needs an account and must belong to a common group (gid). If a directory service, such as LDAP, is not used, this user is created on each node. Every user must have the same uid and primary gid on every node.

1. Create qauser1 user across all nodes.

```
clush -a groupadd -g 6001 qa

clush -a "useradd -g 6001 -u 6001 qauser1"
```

```
clush -a  –B “echo passw0rd | passwd qauser1  --stdin”
```

```
clush -a groupadd -g 6001 qa
clush -a "useradd -g 6001 -u 6001 qauser1"
clush -a  -B "echo passw0rd | passwd qauser1  --stdin"
```

Note Password of qauser1 user is set to passw0rd.

2. Create devuser1 user across all nodes

```
clush -a groupadd –g 6002 dev

clush -a “useradd –g 6002 -u 6002 devuser1”

clush –a  -B “echo passw0rd | passwd devuser1 --stdin”
```

```
clush -a groupadd -g 6002 dev
clush -a "useradd -g 6002 -u 6002 devuser1"
clush -a  -B "echo passw0rd | passwd devuser1 --stdin"
```

Note Password of devuser1 user is set to passw0rd.

3. Create devuser2 user across all nodes

```
clush –a “useradd -g 6002 –u 6004 devuser2”

clush -a  –B “echo passw0rd | passwd devuser2 --stdin”
```

```
clush -a "useradd -g 6002 -u 6004 devuser2"
clush -a  -B "echo passw0rd | passwd devuser2 --stdin"
```

Note Password of devuser2 user is set to passw0rd.

Setting Up Topologies

Setting the physical topology of a MapR cluster node is done via the maprcli command

```
maprcli node move -serverids <ids> –topology <topology>
```

For this UCS deployment, we’ll use the following simple script `/root/set_topology` on each node that allows the simple migration to a new topology

```
$cat /root/set_topology
#!/bin/bash

if [ –z “${1}” ] ; then
    echo “usage: $0 <topology>”
    exit 1
fi

hexid=`cat /opt/mapr/hostid`
myid=`printf “%d“ “0x$hexid“`
maprcli node move -serverids $myid –topology $1
```

Save the above commands into a script file (set_topology) on the administrative node. Distribute it to all the other nodes in the cluster

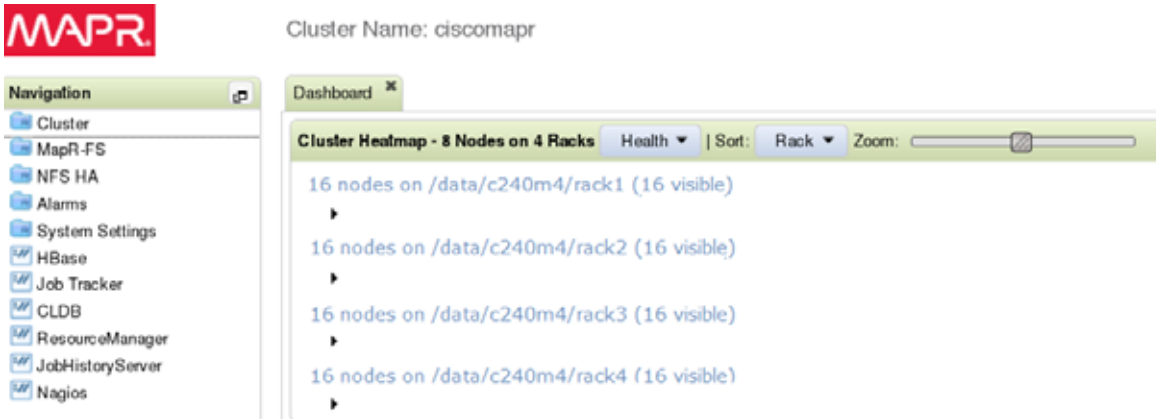
```
$ chmod a+x /root/set_topology
$ clush -B -a -c /root/set_topology
```

Now that it is distributed, we can set the topology appropriately for 4 racks of C240M4 nodes:

```
$ clush -B -w rhel[1-16] '/root/set_topology /data/c240m4/rack1'
$ clush -B -w rhel[17-32] '/root/set_topology /data/c240m4/rack2'
$ clush -B -w rhel[33-48] '/root/set_topology /data/c240m4/rack3'
$ clush -B -w rhel[49-64] '/root/set_topology /data/c240m4/rack4'
```

Verify the topologies by logging into MapR Control System (MCS) through the browser https://rhel3:8443

Figure 95 Verify the Topologies



Create and Associate Volumes to Topologies, setting QoS policies and Quota

Now we will create a volume named 'QA' associated with the `'/data/c240m4/rack1'` topology, have it mounted under path `/QA` in the root directory of MapR-FS, set replication factor as 2 (default 3) and minimum replication factor as 1 (default 2). Note that we can also assign quota to this volume, in our case, it is 1TB as hard limit and 800GB as soft limit. Also, the workload in this volume could be latency sensitive, we will use the low_latency QoS policy.

```
$ maprcli volume create -nodelay 1 -type rw -advisoryquota 800G -minreplication 1
-quota 1000G -replication 2 -name QA -path /QA -topology /data/c240m4/rack1
-rootdirperms 700 -replicationtype low_latency -readonly 0 -user mapr:fc
```

Now we will create a volume named 'DEV' associated with the `'/data/c240m4/rack2'` topology, have it mounted under path `/DEV` in the root directory of MapR-FS, set replication factor as 3 and minimum replication factor as 2. Note that we can also assign quota to this volume, in our case, it is 2TB as hard limit and 1200GB as soft limit. Also, the workload in this volume could be throughput sensitive, we will use the high_throughput QoS policy.

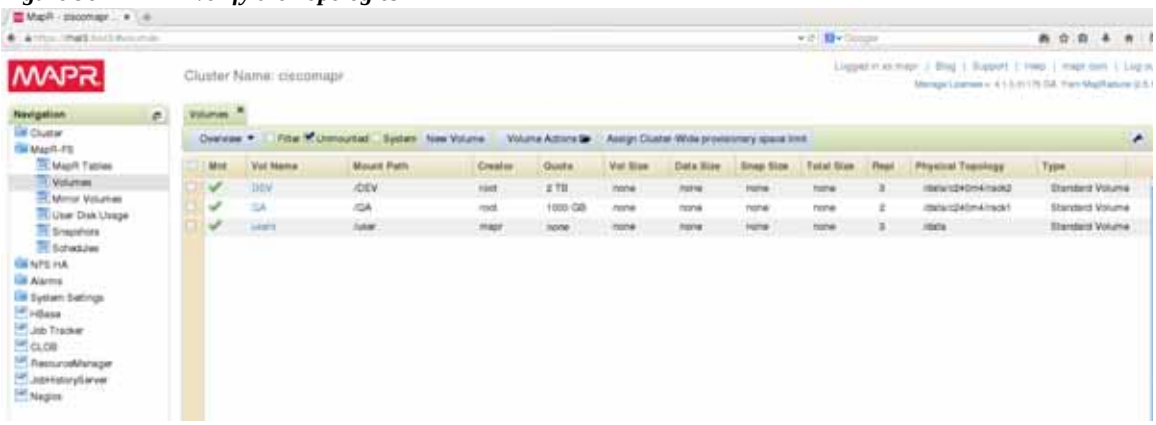
```
$ maprcli volume create -nodelay 1 -type rw -advisoryquota 1200G -minreplication 2
-quota 2000G -replication 3 -name DEV -path /DEV -topology /data/c240m4/rack2
-rootdirperms 700 -replicationtype high_throughput -readonly 0 -user mapr:fc
```

Table 9 Created and Associated Volume Details

Parameter	Description
Advisoryquota	The advisory quota for the volume as integer plus unit. Units: B, K, M, G, T, P
Minreplication	The minimum replication level. Default: 2 When the replication factor falls below this minimum, re-replication occurs as aggressively as possible to restore the replication level. If any containers in the CLDB volume fall below the minimum replication factor, writes are disabled until aggressive re–replication restores the minimum level of replication.
Name	The name of the volume to create.
Path	The path at which to mount the volume.
Quota	The quota for the volume as integer plus unit. Example: quota=500G; Units: B, K, M, G, T, P
Readonly	Specifies whether or not the volume is read-only: 0 = Volume is read/write. 1 = Volume is read–only.
Replication	The desired replication level. Default: 3 When the number of copies falls below the desired replication factor, but remains equal to or above the minimum replication factor, re-replication occurs after the timeout specified in the cldb.fs.mark.rereplicate.sec parameter.
Replicationtype	The desired replication type. You can specify low_latency (star replication) or high_throughput (chain replication). The default setting is high_throughput.
Rootdirperms	Permissions on the volume root directory.
Topology	The rack path to the volume.
Type	The type of volume to create. mirror – standard mirror (read-only) volume (promotable to standard read–write volume) rw - standard (read–write) volume (convertible to standard mirror volume) 0 - standard (read–write) volume (for backward compatibility) 1 - non–convertible mirror (read–only) volume (for backward compatibility)
User	Space-separated list of user:permission pairs.

Verify the topologies by logging into MapR Control System (MCS) through browser https://rhel3:8443.

Figure 96 Verify the Topologies



Setting up ACL policies for volumes

To Change the ownership and permission of the /QA and /DEV directories perform the following commands:

```
$ hadoop fs -chown -R qauser1 /QA
$ hadoop fs -chgrp -R qa /QA
$ hadoop fs -chmod -R 700 /QA
$ hadoop fs -chown -R devuser1 /DEV
$ hadoop fs -chgrp -R dev /DEV
$ hadoop fs -chmod -R 700 /DEV
```

```
[root@rhel1 ~]# hadoop fs -chown -R qauser1 /QA
[root@rhel1 ~]# hadoop fs -chgrp -R qa /QA
[root@rhel1 ~]# hadoop fs -chmod -R 700 /QA
```

```
[root@rhel1 ~]# hadoop fs -chown -R devuser1 /DEV
[root@rhel1 ~]# hadoop fs -chgrp -R dev /DEV
[root@rhel1 ~]# hadoop fs -chgrp -R dev /DEV
```

Now we can issue a hadoop fs -ls / to verify read–write permissions.

```
$ hadoop fs -ls /
```

```
[root@rhel1 ~]# hadoop fs -ls /
Found 8 items
drwx------   - devuser1 dev          0 2015-05-06 00:08 /DEV
drwx------   - qauser1  qa           0 2015-05-05 22:30 /QA
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /apps
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /hbase
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /opt
drwxrwxrwx   - mapr     mapr         3 2015-05-05 22:05 /tmp
drwxr-xr-x   - mapr     mapr         0 2015-04-28 22:55 /user
drwxr-xr-x   - mapr     mapr         1 2015-04-28 22:55 /var
```

Alternatively, the volume is mounted through NFS and is available under `/mapr/<cluster name>` as a directory.

```
$ ls -al /mapr/ciscomapr
```

```
[root@rhel1 ~]# ls -al /mapr/ciscomapr
total 2
drwxr-xr-x 10 mapr     mapr 9 May  6 00:24 .
dr-xr-xr-x  3 root     root 0 May  6 00:25 ..
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 apps
drwx------  2 devuser1 dev  0 May  6 00:08 DEV
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 hbase
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 opt
drwx------  2 qauser1  qa   0 May  5 22:30 QA
drwxrwxrwx  3 mapr     mapr 3 May  5 22:05 tmp
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 user
drwxr-xr-x  3 mapr     mapr 1 Apr 28 22:55 var
```

Now the volume looks like a Unix directory `chown` command can be used to change its ACLs.

```
$ chown -R devuser2:dev /mapr/ciscomapr/DEV
$ chmod -R 750 /mapr/ciscomapr/DEV
$ ls -al /mapr/ciscomapr/DEV
```

```
[root@rhel1 ~]# chown -R devuser2:dev /mapr/ciscomapr/DEV
[root@rhel1 ~]# chmod -R 750 /mapr/ciscomapr/DEV
[root@rhel1 ~]# ls -al /mapr/ciscomapr/
total 2
drwxr-xr-x 10 mapr     mapr 9 May  6 00:24 .
dr-xr-xr-x  3 root     root 0 May  6 00:39 ..
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 apps
drwxr-x---  2 devuser2 dev  0 May  6 00:08 DEV
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 hbase
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 opt
drwx------  2 qauser1  qa   0 May  5 22:30 QA
drwxrwxrwx  3 mapr     mapr 3 May  5 22:05 tmp
drwxr-xr-x  2 mapr     mapr 0 Apr 28 22:55 user
drwxr-xr-x  3 mapr     mapr 1 Apr 28 22:55 var
```

Job Placement

Assume a mixture of high and low processing power servers in the same cluster. Job placement allows users to submit important jobs that can only run on the high power servers to get fast turnaround time.

Follow these steps to set up Job Placement:

1. Edit `/opt/mapr/hadoop/hadoop-2.5.1/etc/hadoop/yarn-site.xml` on resource manager nodes. Add the following content between <configuration> and </configuration> tags.

```
<property>
 <name>node.labels.file</name>
 <value>/tmp/label.txt</value>
 <description>The path to the node labels file.</description>
</property>

<property>
 <name>node.labels.monitor.interval</name>
```

```
<value>120000</value>
```

```
property>
<name>node.labels.file</name>
<value>/tmp/label.txt</value>
<description>The path to the node labels file.</description>
/property>

property>
<name>node.labels.monitor.interval</name>
<value>120000</value>
<description>Interval for checking the labels file for updates (default is 120000 ms)</description>
/property>
```

2. Create label.txt with following content

```
vi  label.txt
rhel1 highCPU
rhel2 highCPU
rhel3 highCPU
rhel4 highCPU,highIO
rhel5 lowCPU,lowIO
rhel6 lowCPU
rhel7 lowCPU
rhel8 lowCPU
rhel9 highCPU
rhel10 highCPU
<upto>
rhel64 highCPU
```

```
rhel1 highCPU
rhel2 highCPU
rhel3 highCPU
rhel4 highCPU,highIO
rhel5 lowCPU,lowIO
rhel6 lowCPU
rhel7 lowCPU
rhel8 lowCPU
rhel9 highCPU
rhel10 highCPU
```

3. Change permission of that file and copy to MapR–FS.

```
chmod 755 label.txt
cp label.txt /mapr/ciscomapr/tmp
```

```
chmod 755 label.txt
cp label.txt /mapr/ciscomapr/tmp
```

4. Restart Resource manager

```
maprcli node services -name resourcemanager –action restart -filter
"[csvc==resourcemanager]"
```

```
maprcli node services name resourcemanager action restart filter
"[csvc==resourcemanager]"
```

5. Refresh Labels.

```
yarn rmadmin –refreshLabels
```

```
yarn rmadmin -refreshLabels
```

6. Verify the Labels.

```
yarn rmadmin -showLabels
```

```
[root@rhel5 ~]# vi /opt/mapr/hadoop/hadoop-2.5.1/etc/hadoop/yarn-site.xml
[root@rhel5 ~]# vi  label.txt
[root@rhel5 ~]# vi  label.txt
[root@rhel5 ~]# chmod 755 label.txt
[root@rhel5 ~]# cp label.txt /mapr/ciscomapr/tmp
[root@rhel5 ~]# maprcli node services -name resourcemanager -action restart -filter "[csvc==resourcemanager]"
[root@rhel5 ~]# yarn rmadmin -refreshLabels
15/05/05 22:08:50 INFO client.RMProxy: Connecting to ResourceManager at rhel5/192.168.11.105:8033
Refreshed labels for nodes in the cluster successfully.

[root@rhel5 ~]# yarn rmadmin -showLabels
15/05/05 22:09:56 INFO client.RMProxy: Connecting to ResourceManager at rhel5/192.168.11.105:8033
                  Nodes     Labels
                  rhel4  [highCPU]
                  rhel3  [highCPU]
                  rhel2  [highCPU]
                  rhel1  [highCPU]
                  rhel8   [lowCPU]
                  rhel7   [lowCPU]
                  rhel6   [lowCPU]
                  rhel5   [lowCPU]
```

7. Run a sample teragen jobs as two different users (qauser1 and devuser1):

   As qauser1, submit the job to nodes labled as “highCPU“ (i.e. rhel1 – rhel4):

```
 yarn jar
/opt/mapr/hadoop/hadoop-2.5.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.5
.1-mapr-1503.jar teragen -Dmapreduce.job.label=highCPU 10000000000 /QA/teragen
```

```
yarn jar /opt/mapr/hadoop/hadoop-2.5.1/share/hadoop/mapreduce/hadoop-mapreduce-
xamples-2.5.1-mapr-1503.jar teragen -Dmapreduce.job.label=highCPU 10000000000 /
A/teragen
```

   Also on another console, run this command to show CPU utilization of the nodes:

```
while :; do maprcli node list -column cpu; sleep 1; done
```

   The first horizontal screen displays CPU Utilization while the second displays the output of teragen. Observe the servers rhel1 through rhel4 have higher utilization.

```
utilization  hostname  ip
9            rhel1     192.168.11.101,192.168.12.101
4            rhel2     192.168.11.102,192.168.12.102
8            rhel3     192.168.11.103,192.168.12.103
8            rhel4     192.168.11.104,192.168.12.104
1            rhel5     192.168.11.105,192.168.12.105
1            rhel6     192.168.11.106,192.168.12.106
1            rhel7     192.168.11.107,192.168.12.107
1            rhel8     192.168.11.108,192.168.12.108

DEPRECATED: Use of this script to execute mapred command is deprecated.
Instead use the mapred command for it.

15/05/05 22:20:50 INFO client.RMProxy: Connecting to ResourceManager at rhel5/192.168.11.105:8032
Killed job job_1430878071477_0002
[qauser1@rhel1 ~]$ rm -rf /mapr/ciscomapr/QA/teragen/
[qauser1@rhel1 ~]$ yarn jar /opt/mapr/hadoop/hadoop-2.5.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.5.1-mapr-1503.jar teragen -Dmapreduce.job.l
bel=highCPU -Dmapred.mapr.tasks=192 10000000000 /QA/teragen
15/05/05 22:21:12 INFO client.RMProxy: Connecting to ResourceManager at rhel5/192.168.11.105:8032
15/05/05 22:21:13 INFO terasort.TeraSort: Generating 10000000000 using 2
15/05/05 22:21:13 INFO mapreduce.JobSubmitter: number of splits:2
15/05/05 22:21:13 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1430878071477_0003
15/05/05 22:21:13 INFO security.ExternalTokenManagerFactory: Initialized external token manager class - com.mapr.hadoop.yarn.security.MapRTicketManager
15/05/05 22:21:13 INFO impl.YarnClientImpl: Submitted application application_1430878071477_0003
15/05/05 22:21:13 INFO mapreduce.Job: The url to track the job: http://rhel5:8088/proxy/application_1430878071477_0003/
15/05/05 22:21:13 INFO mapreduce.Job: Running job: job_1430878071477_0003
15/05/05 22:21:18 INFO mapreduce.Job: Job job_1430878071477_0003 running in uber mode : false
15/05/05 22:21:18 INFO mapreduce.Job:  map 0% reduce 0%
15/05/05 22:21:50 INFO mapreduce.Job:  map 1% reduce 0%
15/05/05 22:22:39 INFO mapreduce.Job:  map 2% reduce 0%
```

The same command does not work if devuser1 runs it due to ACL restriction.

```
[devuser1@rhel1 ~]$ yarn jar /opt/mapr/hadoop/hadoop-2.5.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.5.1-mapr-1503.jar teragen -Dmapreduce.job.l
abel=highCPU 10000000000 /QA/teragen
org.apache.hadoop.security.AccessControlException: User devuser1(user id 6002)  does not have access to /QA/teragen
        at com.mapr.fs.MapRFileSystem.getMapRFileStatus(MapRFileSystem.java:1306)
        at com.mapr.fs.MapRFileSystem.getFileStatus(MapRFileSystem.java:928)
        at org.apache.hadoop.fs.FileSystem.exists(FileSystem.java:1434)
        at org.apache.hadoop.examples.terasort.TeraGen.run(TeraGen.java:292)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
        at org.apache.hadoop.examples.terasort.TeraGen.main(TeraGen.java:309)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:606)
        at org.apache.hadoop.util.ProgramDriver$ProgramDescription.invoke(ProgramDriver.java:72)
        at org.apache.hadoop.util.ProgramDriver.run(ProgramDriver.java:145)
        at org.apache.hadoop.examples.ExampleDriver.main(ExampleDriver.java:90)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:606)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:212)
```

Administration and Reporting

Administration and Reporting From an administrative perspective, MapR allows organizations to define and enforce storage, CPU, and memory quotas at the volume, user, and group levels. For service providers to provide accurate usage and billing information, MapR

offers reporting on resource usage on over 60 different metrics. These metrics are available via the MapR Control System (MCS) browser–based user interface, and for up-stream integration—via the command–line interface and the REST API.

The following command demonstrates how to save above metrics in a log file in json format for later analysis:

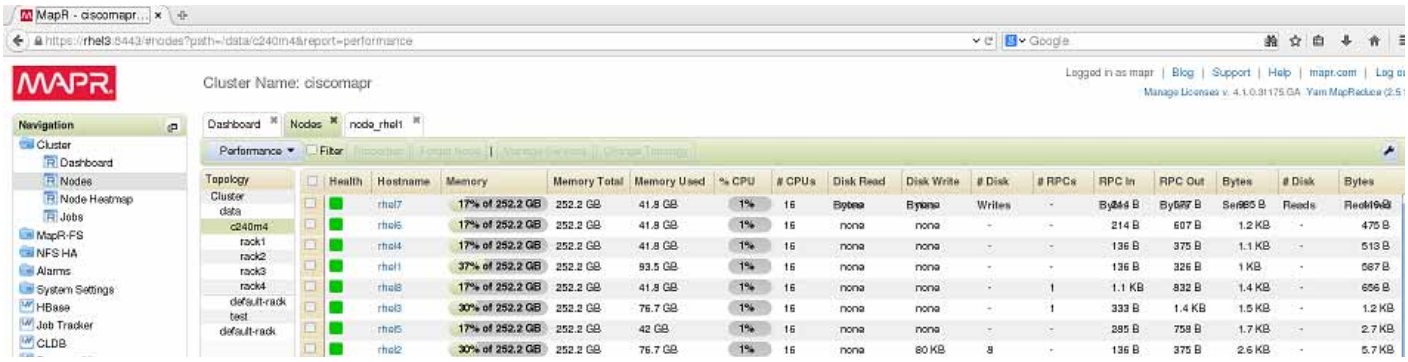
```
while :; do  maprcli node list -limit 50 -start 0  -json >>  maprlog.json; sleep
10;done
```

```
[root@rhel1 ~]# maprcli node list -limit 50 -start 0  -json
{
        "timestamp":1430971312393,
        "timeofday":"2015-05-07 12:01:52.393 GMT-0400",
        "status":"OK",
        "total":8,
        "data":[
                {
                        "id":"8865112475414864155",
                        "ip":[
                                "192.168.11.101",
                                "192.168.12.101"
                        ],
                        "hostname":"rhel1",
                        "racktopo":"/data/c240m4/rack1/rhel1",
                        "health":0,
                        "healthDesc":"Healthy",
                        "service":"historyserver,nodemanager,cldb,fileserver,nfs,hoststats",
                        "configuredservice":"historyserver,nodemanager,cldb,fileserver,nfs,hostst
ts",
                        "fs-heartbeat":0,
                        "jt-heartbeat":2,
                        "dtotal":3186,
                        "dused":0,
                        "davail":3185,
                        "rpcs":0,
                        "rpcin":285,
                        "rpcout":711,
                        "disks":7,
                        "MapRfs disks":3,
                        "faileddisks":0,
                        "dreads":0,
                        "dreadK":0,
                        "dwrites":0,
                        "dwriteK":0,
                        "cpus":16,
```

Note All the metrics are documented here: http://doc.mapr.com/display/MapR/node

The following figure shows all the performance metrics in the MapR Control System (MCS) at https://rhel3:8443/#nodes?path=/data/c240m4&report=performance

Figure 97 Performance Metrics in MapR Controller System



Conclusion

Hadoop has evolved into a leading data management platform across all verticals. The Cisco UCS Integrated Infrastructure for Big Data with MapR and Multi–Tenancy offers a dependable deployment model for enterprise Hadoop that offers a fast and predictable path for businesses to unlock value in Big Data.

The configuration detailed in the document can be extended to clusters of various sizes depending on what application demands. Up to 80 servers (5 racks) can be supported with no additional switching in a single UCS domain with no network over–subscription. Scaling beyond 5 racks (80 servers) can be implemented by interconnecting multiple UCS domains using Nexus 6000/7000 Series switches, scalable to thousands of servers and to hundreds of petabytes storage, and managed from a single pane using UCS Central.

Bill of Materials

This section provides the BOM for 64 nodes Performance Optimized Cluster.

Table 10 Bill of Materials for C240M4SX Base rack

Part Number	Description	Quantity
UCS-SL-CPA3-P	Performance Optimized Cluster	1
UCSC-C240-M4SX	UCS C240 M4 SFF 24 HD w/o CPU, mem, HD, PCIe, PS, railkt w/expndr	16
UCSC-MRAID12G	Cisco 12G SAS Modular Raid Controller	16
UCSC-MRAID12G-2GB	Cisco 12Gbps SAS 2GB FBWC Cache module (Raid 0/1/5/6)	16

Table 10 Bill of Materials for C240M4SX Base rack

Part Number	Description	Quantity
UCSC-MLOM-CSC-02	Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+	16
CAB-9K12A-NA	Power Cord 125VAC 13A NEMA 5-15 Plug North America	32
UCSC-PSU2V2-1200W	1200W V2 AC Power Supply for 2U C-Series Servers	32
UCSC-RAILB-M4	Ball Bearing Rail Kit for C220 M4 and C240 M4 rack servers	16
UCSC-HS-C240M4	Heat Sink for UCS C240 M4 Rack Server	32
UCSC-SCCBL240	Supercap cable 250mm	16
UCS-CPU-E52680D	2.50 GHz E5-2680 v3/120W 12C/30MB Cache/DDR4 2133MHz	32
UCS-MR-1X162RU-A	16GB DDR4-2133-MHz RDIMM/PC4-17000/dual rank/x4/1.2v	256
UCS-HD12T10KS2-E	1.2 TB 6G SAS 10K rpm SFF HDD	384
UCS-SD120G0KSB-EV	120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)	32
UCSC-PCI-1C-240M4	Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	16
UCS-FI-6296UP-UPG	UCS 6296UP 2RU Fabric Int/No PSU/48 UP/ 18p LIC	2
CON-SNTP-C240M4SX	SMARTNET 24X7X4 UCS C240 M4 SFF 24 HD w/o CPU, mem	16
CON-SNTP-FI6296UP	SMARTNET 24X7X4 UCS 6296UP 2RU Fabric Int/2 PSU/4 Fans	2
SFP-H10GB-CU3M	10GBASE-CU SFP+ Cable 3 Meter	32
UCS-ACC-6296UP	UCS 6296UP Chassis Accessory Kit	2
UCS-PSU-6296UP-AC	UCS 6296UP Power Supply/100-240VAC	4
N10-MGT012	UCS Manager v2.2	2
UCS-L-6200-10G-C	2rd Gen FI License to connect C-direct only	70
UCS-BLKE-6200	UCS 6200 Series Expansion Module Blank	6
UCS 6296UP Fan Module	UCS 6296UP Fan Module	8
CAB-9K12A-NA	Power Cord 125VAC 13A NEMA 5-15 Plug North America	4
UCS-FI-E16UP	UCS 6200 16-port Expansion module/16 UP/ 8p LIC	4
RACK-UCS2	Cisco R42610 standard rack w/side panels	1
CON-OS-R42610	ONSITE 8X5XNBD Cisco R42610 expansion rack no side panel	1

Table 10 Bill of Materials for C240M4SX Base rack

Part Number	Description	Quantity
RP208-30-1P-U-2=	Cisco RP208-30-U-2 Single Phase PDU 20x C13 4x C19 (Country Specific)	2
CON-OS-RPDUX	ONSITE 8X5XNBD Cisco RP208-30-U-X Single Phase PDU 2x	2

Table 11 Bill of Materials for Expansion Racks

Part Number	Description	Quantity
UCSC-C240-M4SX	UCS C240 M4 SFF 24 HD w/o CPU, mem, HD, PCIe, PS, railkt w/expndr	48
UCSC-MRAID12G	Cisco 12G SAS Modular Raid Controller	48
UCSC-MRAID12G-2GB	Cisco 12Gbps SAS 2GB FBWC Cache module (Raid 0/1/5/6)	48
UCSC-MLOM-CSC-02	Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+	48
CAB-9K12A-NA	Power Cord 125VAC 13A NEMA 5-15 Plug North America	96
UCSC-PSU2V2-1200W	1200W V2 AC Power Supply for 2U C-Series Servers	96
UCSC-RAILB-M4	Ball Bearing Rail Kit for C220 M4 and C240 M4 rack servers	48
UCSC-HS-C240M4	Heat Sink for UCS C240 M4 Rack Server	48
UCSC-SCCBL240	Supercap cable 250mm	48
UCS-CPU-E52680D	2.50 GHz E5-2680 v3/120W 12C/30MB Cache/DDR4 2133MHz	96
UCS-MR-1X162RU-A	16GB DDR4-2133-MHz RDIMM/PC4-17000/dual rank/x4/1.2v	768
UCS-HD12T10KS2-E	1.2 TB 6G SAS 10K rpm SFF HDD	1152
UCS-SD120G0KSB-EV	120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)	96
UCSC-PCI-1C-240M4	Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	48
SFP-H10GB-CU3M=	10GBASE-CU SFP+ Cable 3 Meter	96
CON-SNTP-C240M4SX	SMARTNET 24X7X4 UCS C240 M4 SFF 24 HD w/o CPU, mem	48
RACK-UCS2	Cisco R42610 standard rack w/side panels	3
CON-OS-R42610	ONSITE 8X5XNBD Cisco R42610 expansion rack no side panel	3

Part Number	Description	Quantity
RP208-30-1P-U-2=	Cisco RP208-30-U-2 Single Phase PDU 20x C13 4x C19 (Country Specific)	6
CON-OS-RPDUX	ONSITE 8X5XNBD Cisco RP208-30-U-X Single Phase PDU 2x	3

Table 12 Red Hat Enterprise Linux License

Red Hat Enterprise Linux		
RHEL-2S-1G-3A	Red Hat Enterprise Linux	64
CON-ISV1-RH2S1G3A	3 year Support for Red Hat Enterprise Linux	64

Table 13 Bill of Materials for MapR Software

Note Choose one of the part numbers.

Part Number	Description	Quantity
UCS–BD–M5–SL=	MapR M5 EDITION	64
UCS–BD–M7–SL=	MapR M7 EDITION	64

Appendix

Cisco UCS Director Express for Big Data

Introduction

Hadoop has become a strategic data platform embraced by mainstream enterprises as it offers the fastest path for businesses to unlock value in big data while maximizing existing investments.

As you consider Hadoop to meet your growing data and business needs, operational challenges often emerge. Despite its compelling advantages, Hadoop clusters can be difficult, complex, and time consuming to deploy. Moreover, with so much data increasing so quickly, there is a need to find ways to consistently deploy Hadoop clusters and manage them efficiently.

Note The UCSD Express appliances (UCSD Express VM and Baremetal Agent VM) can also be installed on an existing VMware ESXi server with proper network connectivity (See Figure 98) to the UCS domain that manages the Hadoop servers. In such a case, skip the sections until Downloading the UCS Director Express software components.

UCS Director Express for Big Data

Cisco UCS Director Express for Big Data provides a single-touch solution that automates deployment of Hadoop on Cisco UCS Common Platform Architecture (CPA) for Big Data infrastructure. It also provides a single management pane across both Cisco UCS integrated infrastructure and Hadoop software. All elements of the infrastructure are handled automatically with little need for user input. Through this approach, configuration of physical computing, internal storage, and networking infrastructure is integrated with the deployment of operating systems, Java packages, and Hadoop along with the provisioning of Hadoop services. Cisco UCS Director Express for Big Data is integrated with major Hadoop distributions from Cloudera, MapR, and Hortonworks, providing single-pane management across the entire infrastructure. It complements and communicates with Hadoop managers, providing a system wide perspective and enabling administrators to correlate Hadoop activity with network and computing activity on individual Hadoop nodes.

Key features of UCS Director (UCSD) Express for Big Data

- **Faster and Easier Big Data Infrastructure Deployment:** Cisco UCS Director Express for Big Data extends the Cisco UCS Integrated Infrastructure for Big Data with one-click provisioning, installation, and configuration, delivering a consistent, repeatable, flexible, and reliable end-to-end Hadoop deployment.
- **Massive Scalability and Performance:** Cisco’s unique approach provides appliance-like capabilities for Hadoop with flexibility that helps ensure that resources are deployed right the first time and can scale without arbitrary limitations.
- **Centralized Visibility:** Cisco UCS Director Express for Big Data provides centralized visibility into the complete infrastructure to identify potential failures and latent threats before they affect application and business performance.
- **Open and Powerful:** Provides open interfaces that allows further integration into third-party tools and services while allowing flexibility for your own add-on services.

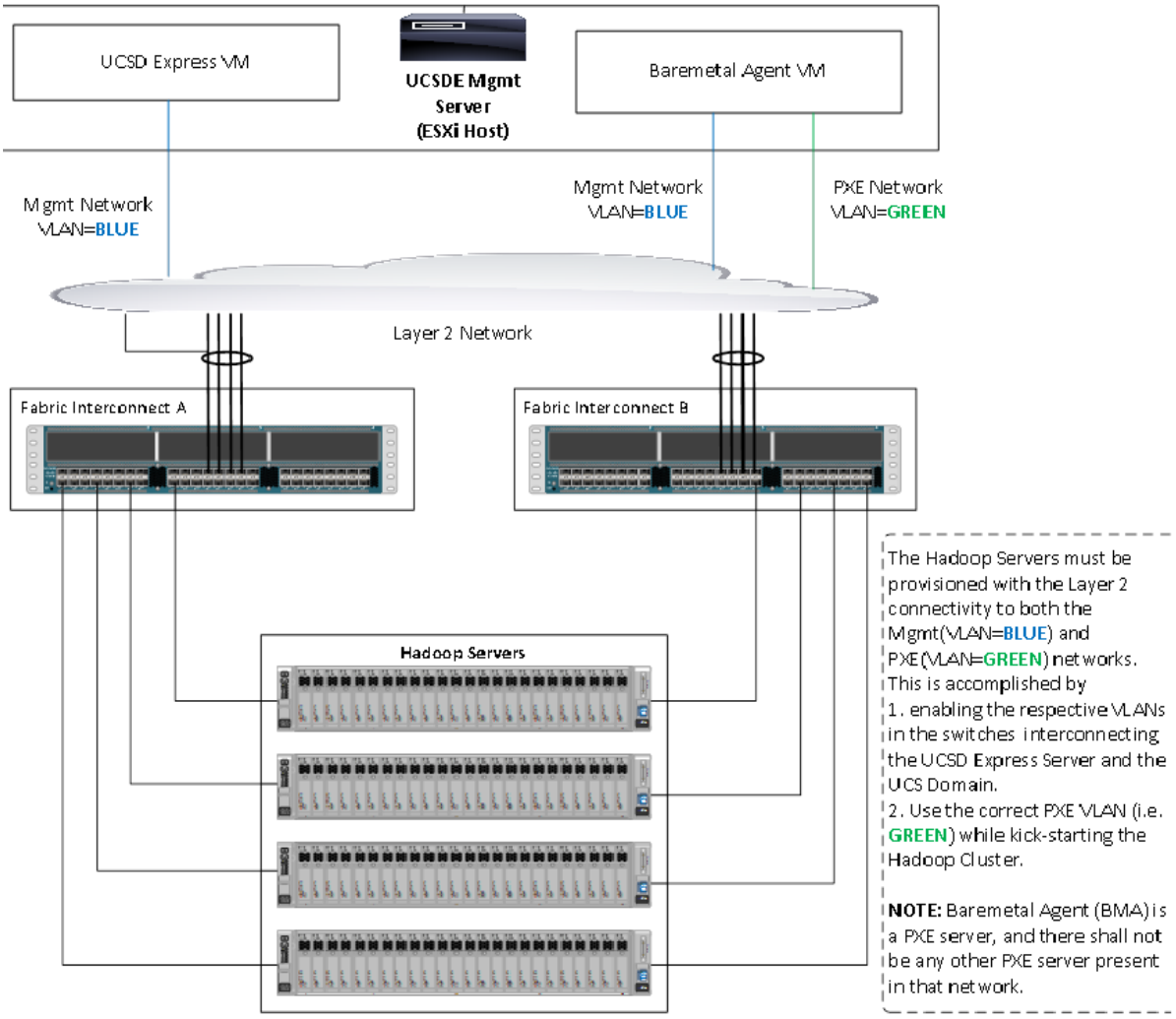
UCSD Express Management Server Configuration

The basic requirement for deploying and executing the UCSD Express software is a server with VMWare ESXi based virtualization environment. Such a physical server machine with ESXi must be connected to the target Hadoop servers in the UCS domain by means of the management network and a dedicated PXE network.

The following are the potential network topologies:

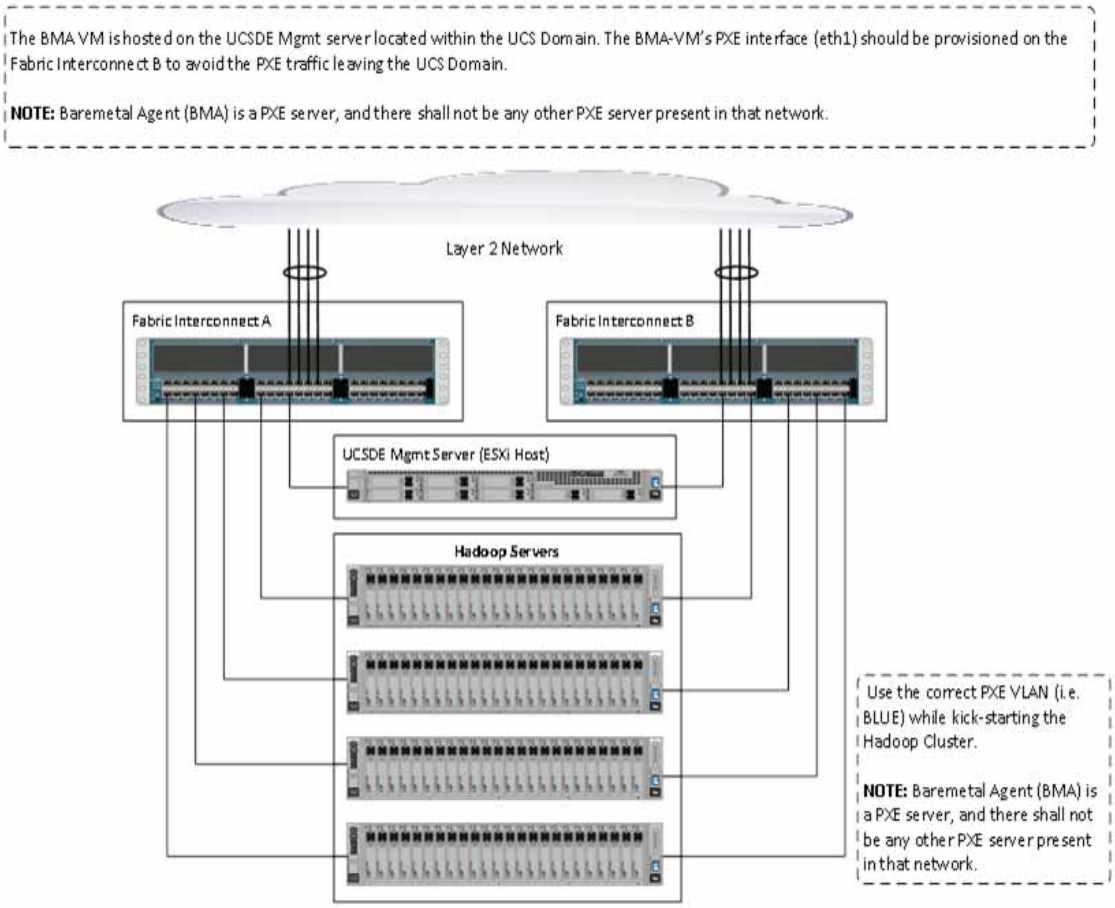
1. The UCSD Express Management server is outside of the UCS Domain containing the C-Series servers that would be used to form the Hadoop cluster. For example, a standalone (CIMC managed) C220 M4 rack server provisioned with UCSD Express VMs is connected to the UCS Domain

Figure 98 UCSD Express Management Server that lives outside the UCS Domain



2. The UCSD Express Management server is hosted on a C220 M4 rack server that is connected to and managed by the same UCS Domain. This is the method used in this document.

Figure 99 UCSD Express Management Server that is being managed as part of the same UCS Domain

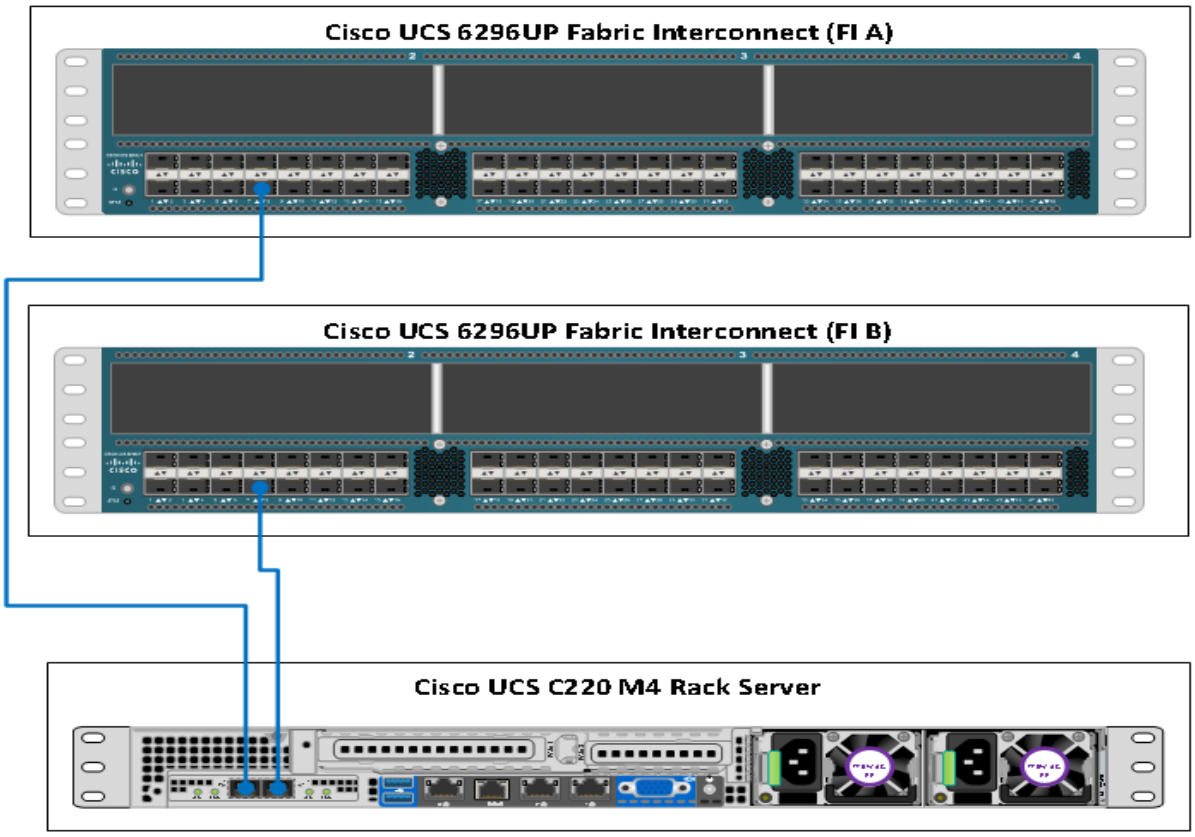


UCSD Management Server Cabling

For this deployment a C220 M4 rack server equipped with Intel Xeon E5-2620 v3 processors, 128 GB of memory, Cisco UCS Virtual Interface Card 1227, Cisco 12-Gbps SAS Modular Raid Controller with 512-MB FBWC, 4 X 600 GB 10K SFF SAS drives is used (any other Cisco UCS server can also be used for this purpose).

The C220 M4 server shall be connected to the UCS Fabric Interconnects as shown in Figure 100. The ports on the on the Fabric Interconnects must be configured as server ports.

Figure 100 Fabric Topology for C220 M4



Software Versions

The UCSD management server is a C220 M4 server that is managed by the UCS Manager. Refer to the software information section in the main part of this Cisco UCS Integrated Infrastructure for Big Data with . See Software Distributions and Versions. In addition, the following software distributions are necessary.

UCS Director Express for Big Data (1.1)

For more information visit
http://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-director-express-big-data-1-1/model.html

VMware vSphere 5.5

UCS Director express requires the VMware vSphere 5.5 hypervisor. For more information see
http://www.vmware.com

Fabric Configuration

The UCSD management server is a C220 M4 server that is managed by the UCS Manager. Refer to the Fabric Configuration section in the main part of this document for more details.

Configuring VLANs

UCSD Express management server requires two network interfaces. It's service profile need to be

- Management Network – default (VLAN 1)
- PXE Network

Table 14 UCSD Express Management Server vNIC configurations

VLAN	Fabric	NIC Port	Function	Failover
default(VLAN1)	A	eth0	Management, User connectivity	Fabric Failover to B
vlan85_PXE	B	eth1	PXE	Fabric Failover to A

PXE VLAN dedicated for PXE booting purpose. Follow these steps in Configuring VLANs to create a dedicated VLAN for PXE. The management network shall continue to be on the default VLAN.

Other UCS configurations

Perform all other UCS configurations such as QOS policy, necessary policies and service profile template by following the documentation above. See the section Creating Pools for Service Profile Templates onwards in this Cisco UCS Integrated Infrastructure for Big Data with cisco validated design.

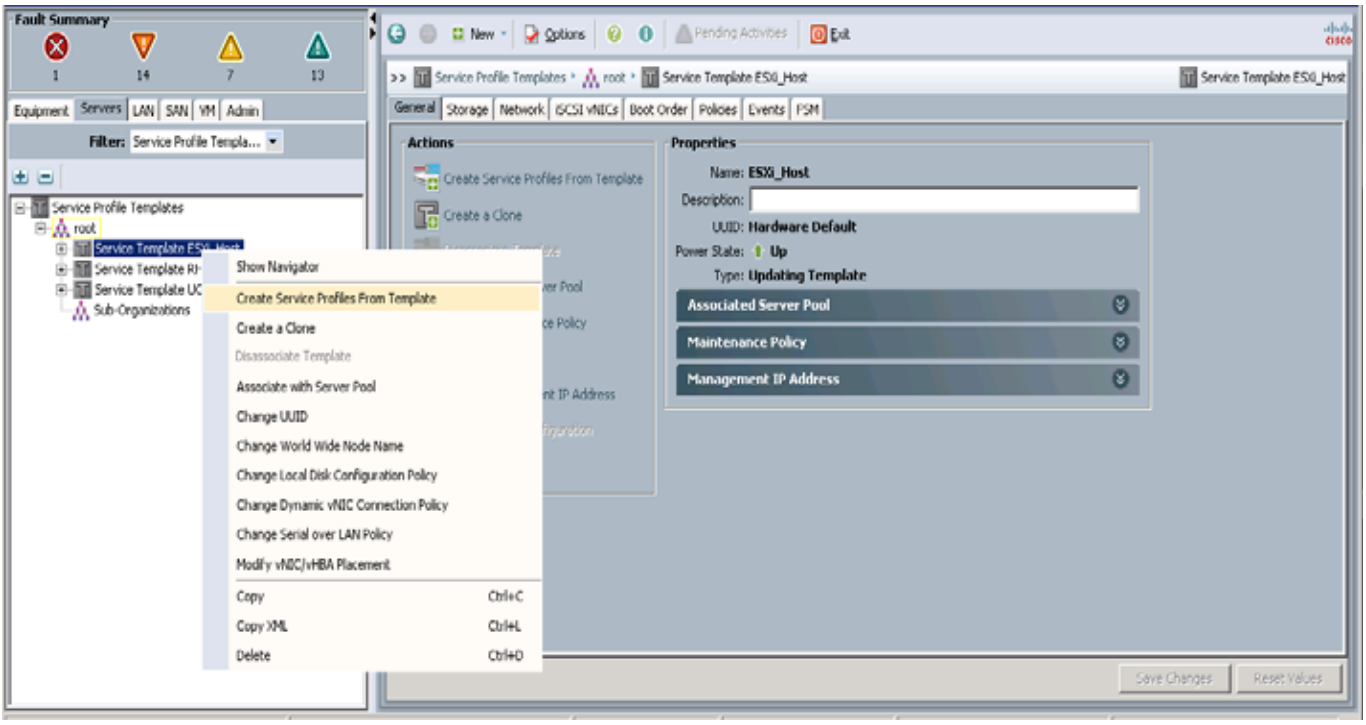
Note Create the service profile template named as ESXi_Host with two vNICs as shown in the above table. For vNIC eth0, select default VLAN as the native VLAN, and for vNIC eth1, select PXE VLAN (vlan85_PXE) as the native VLAN.

Creating Service Profile from the Template

Select the Servers tab in the left pane of the UCS Manager GUI.

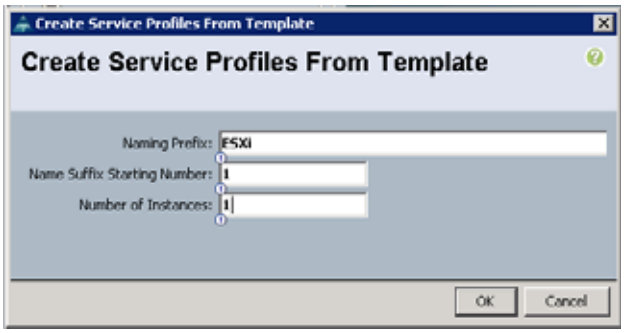
1. Go to Service Profile **Templates > root**.
2. Right-click **Service Profile Templates ESXi_Host**.
3. Select **Create Service Profiles From Template**.

Figure 101 Creating Service Profiles from Template



4. The Create Service Profile from Template window appears.

Figure 102 Selecting Name and Total number of Service Profiles



Association of the Service Profiles will take place automatically.

Installing VMware vSphere ESXi 5.5

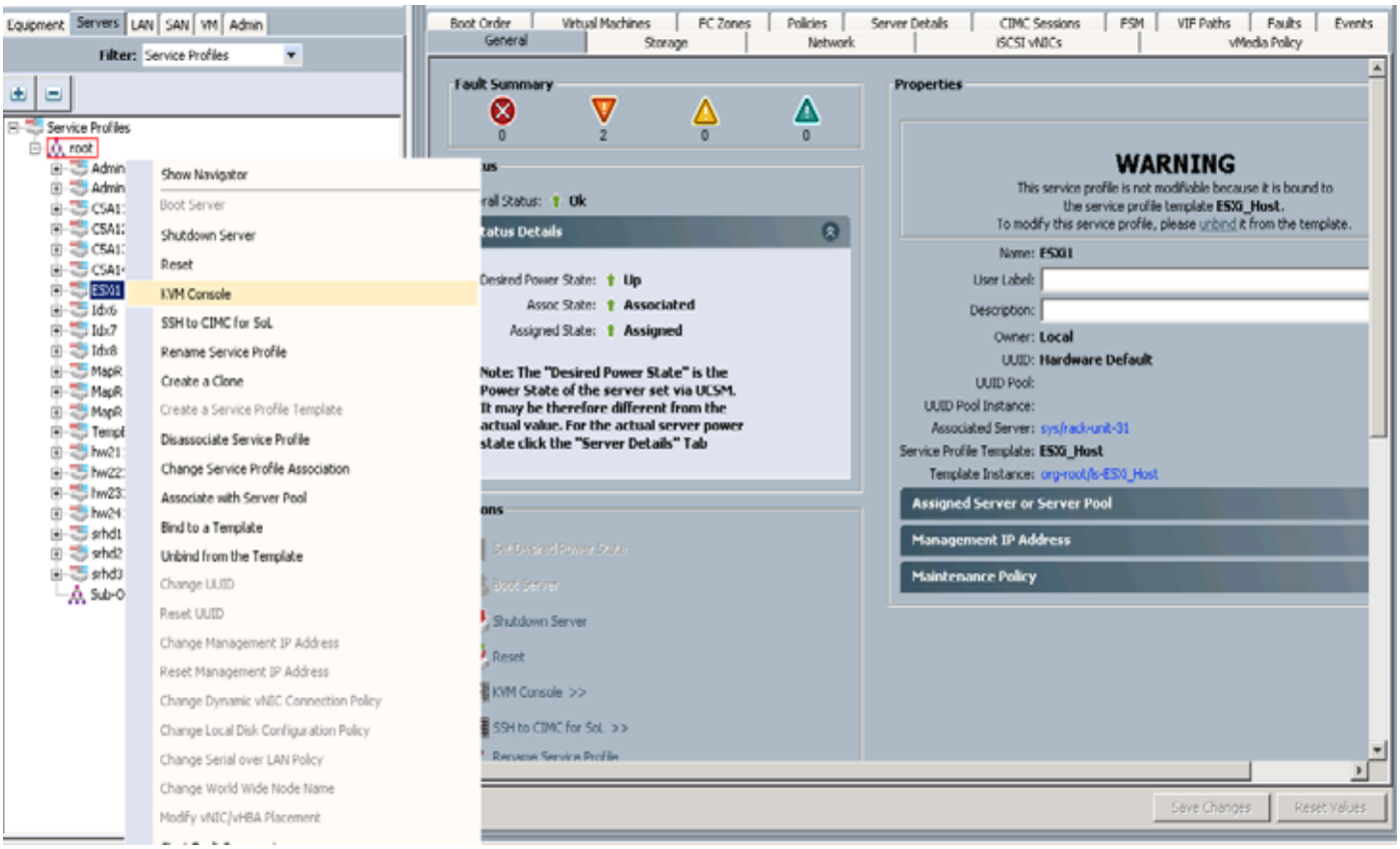
The following section provides detailed procedures for installing VMware vSphere ESXi 5.5.

There are multiple methods to install VMware vSphere ESXi 5.5. The installation procedure described in this deployment guide uses KVM console and virtual media from Cisco UCS Manager.

1. Log in to the Cisco UCS 6296 Fabric Interconnect and launch the Cisco UCS Manager application.

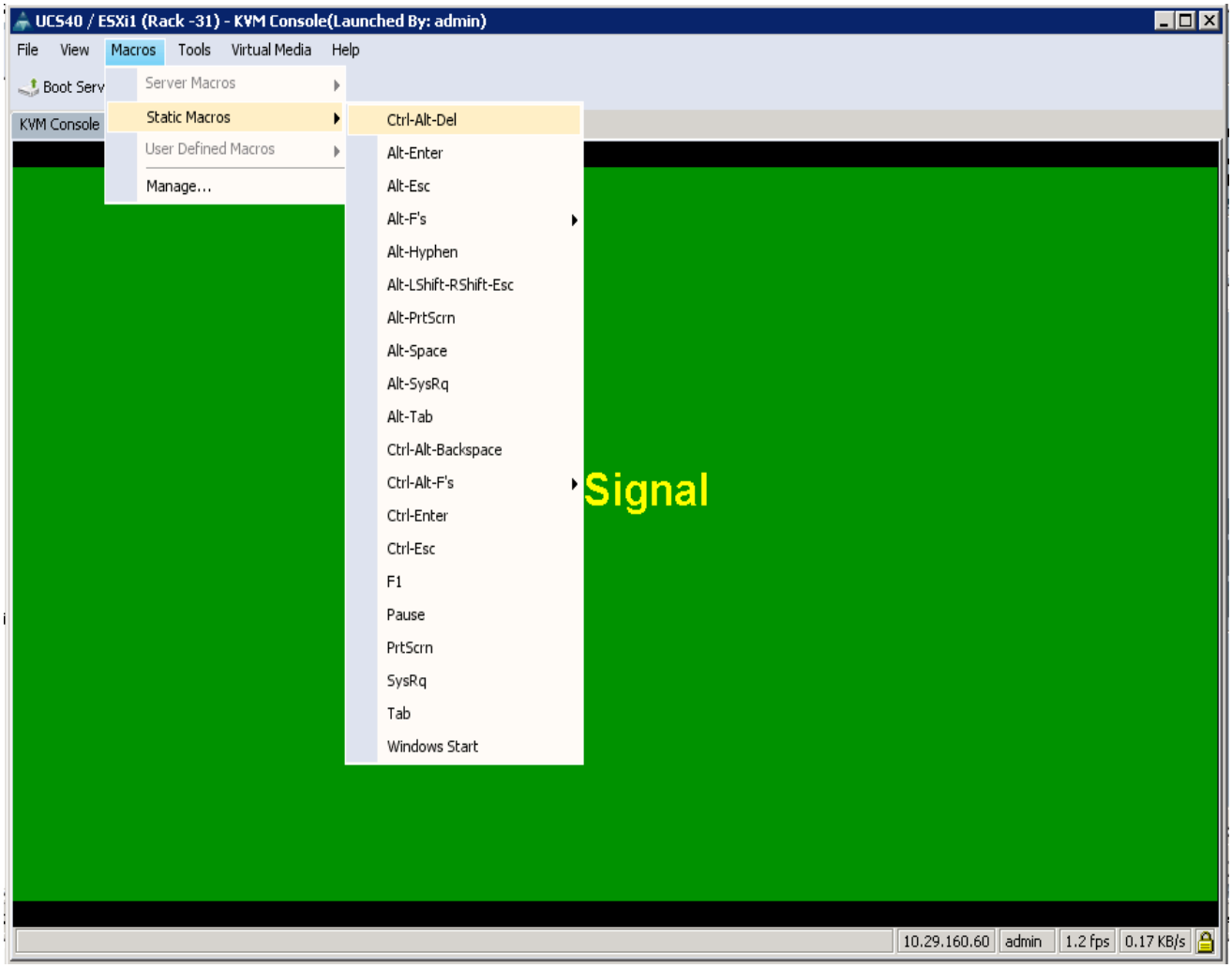
2. Select the **Servers** tab.
3. In the navigation pane expand Service Profiles.
4. Right click on the newly created service profile ESXi1 and select KVM Console.

Figure 103 Selecting KVM Console



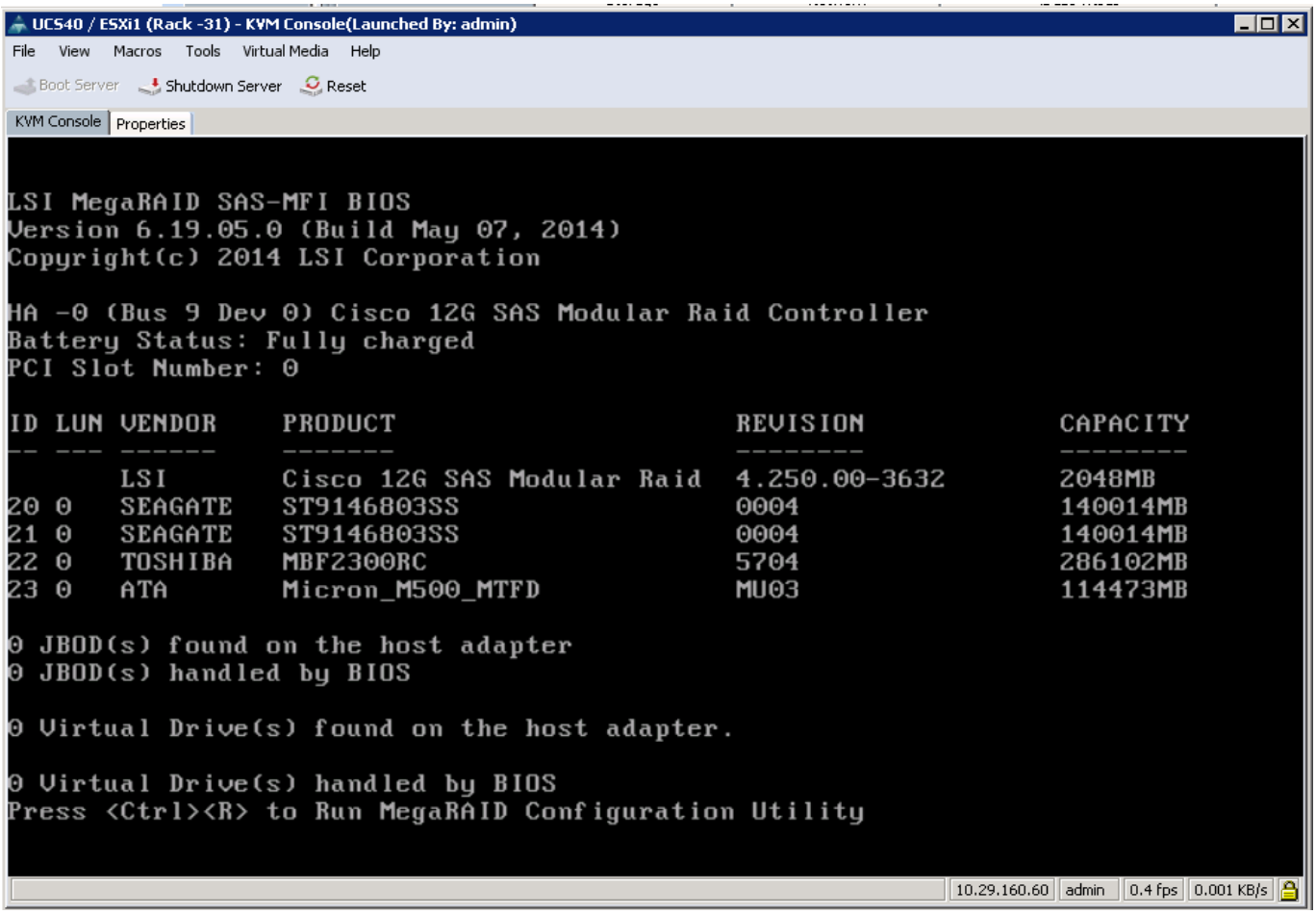
5. In the KVM window, force a reboot by executing the **Ctrl-Alt-Del** macro.

Figure 104 Sending Ctrl-Alt-Del to Reset the Server

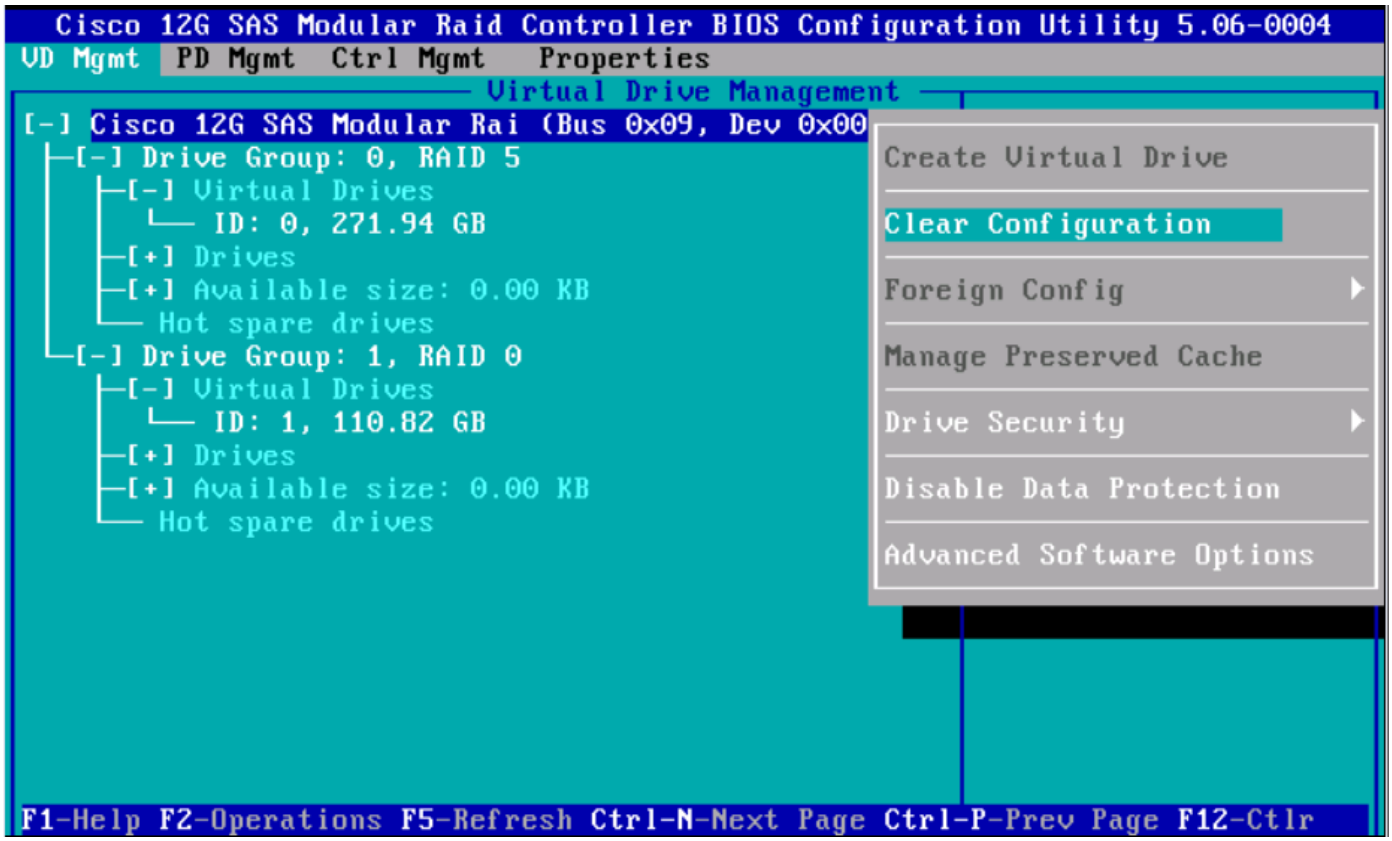


6. As the server goes through a reboot, monitor the progress via the KVM window. When the LSI MegaRAID SAS-MFI BIOS screen appears, press **Ctrl-R** to Enter the Cisco 12G SAS Modular Raid Controller BIOS Configuration Utility.

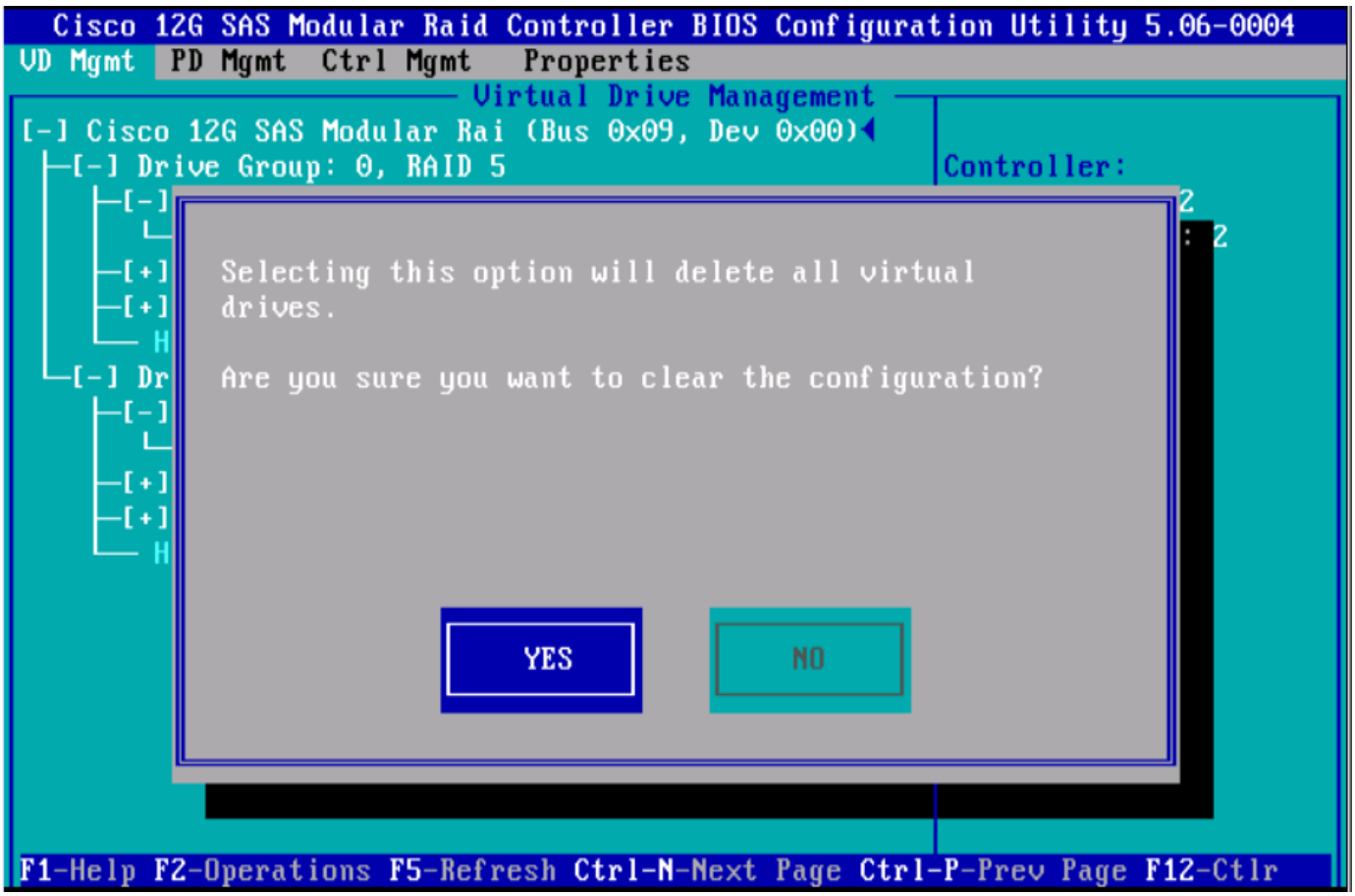
Figure 105 KVM Window displaying the LSI MegaRAID SAS-MFI BIOS screen



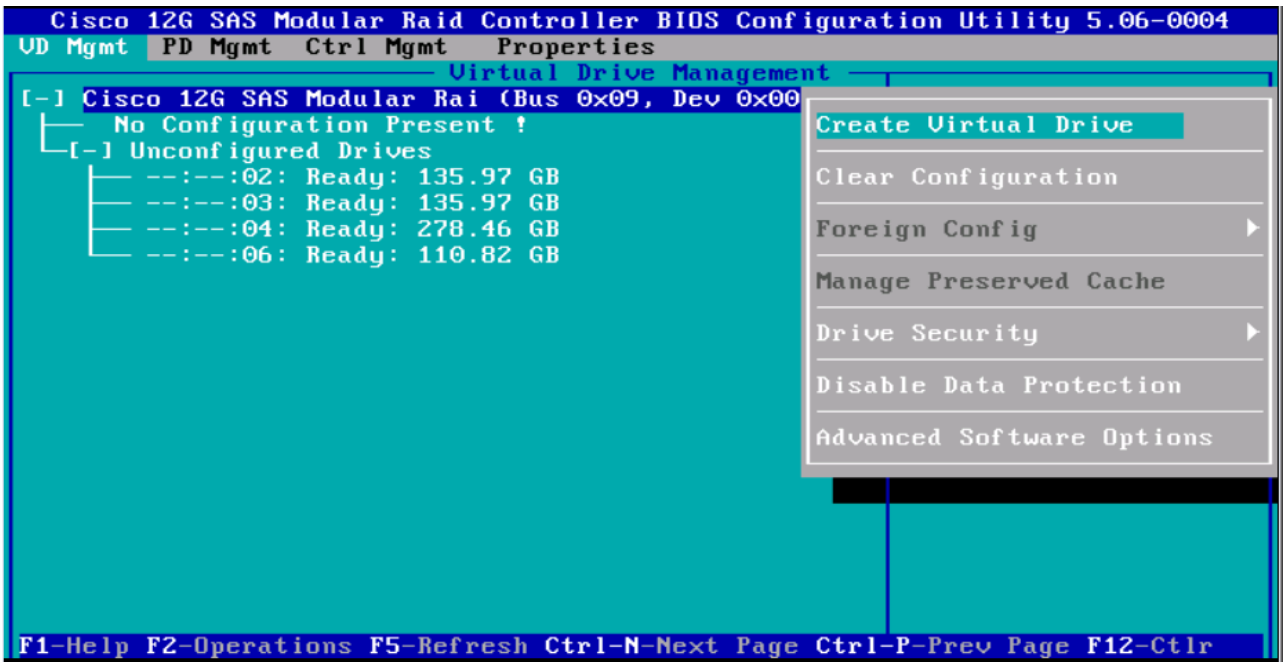
7. In the MegaRAID configuration utility, under VD Mgmt section, use the arrow keys to select the Cisco 12G SAS Modular RAID (Bus 0xNN, Dev 0xNN) line item.
8. Press the function key **F2**.
9. Select the option Clear Configuration, and press **ENTER.**



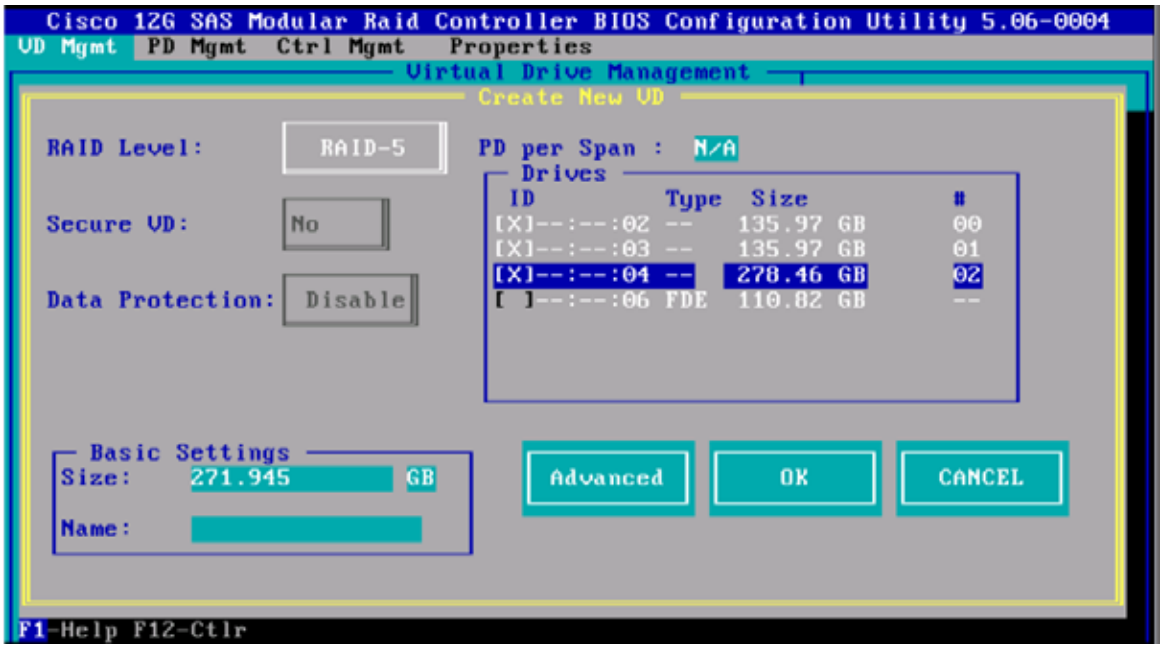
10. To the question Are you sure you want to clear the configuration? click **YES** and press **ENTER** key.



11. In the VD Mgmt section, use the arrow keys to select the Cisco 12G SAS Modular RAID (Bus 0xNN, Dev 0xNN) line item.
12. Press the function key **F2**, select Create Virtual Drive and press **ENTER**.

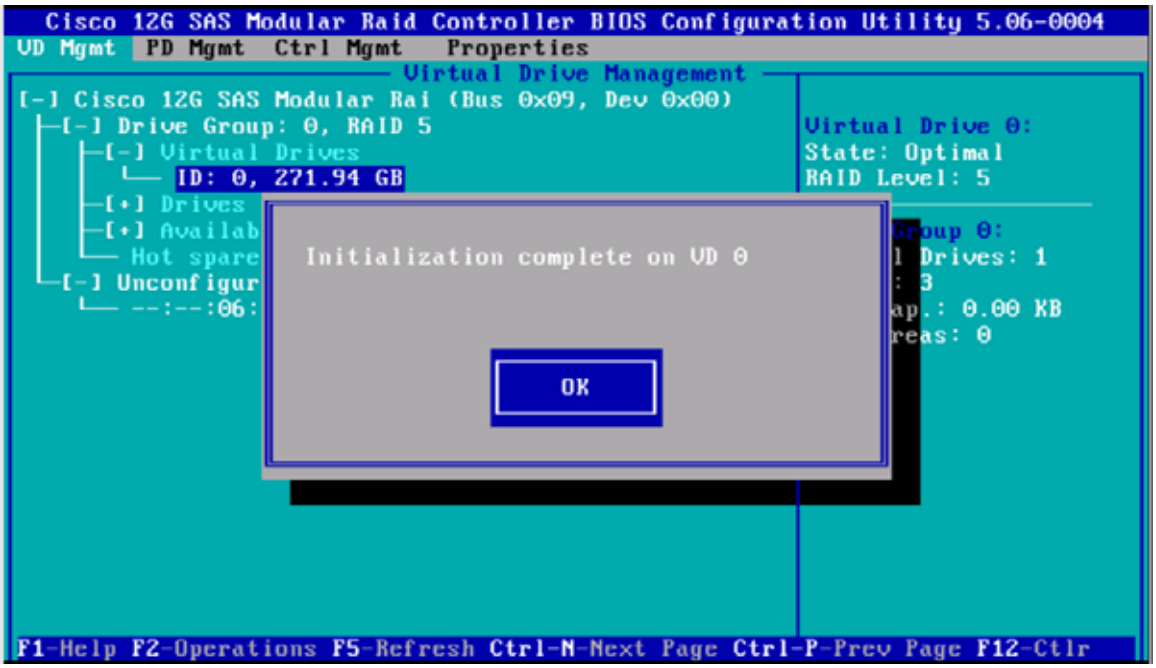


13. In the RAID Level: press **ENTER** and choose **RAID-5**.
14. In the Drives section, press **SPACE** on the desired number of drives to select them to be part of the RAID group. Use the Up and Down arrow keys to navigate.

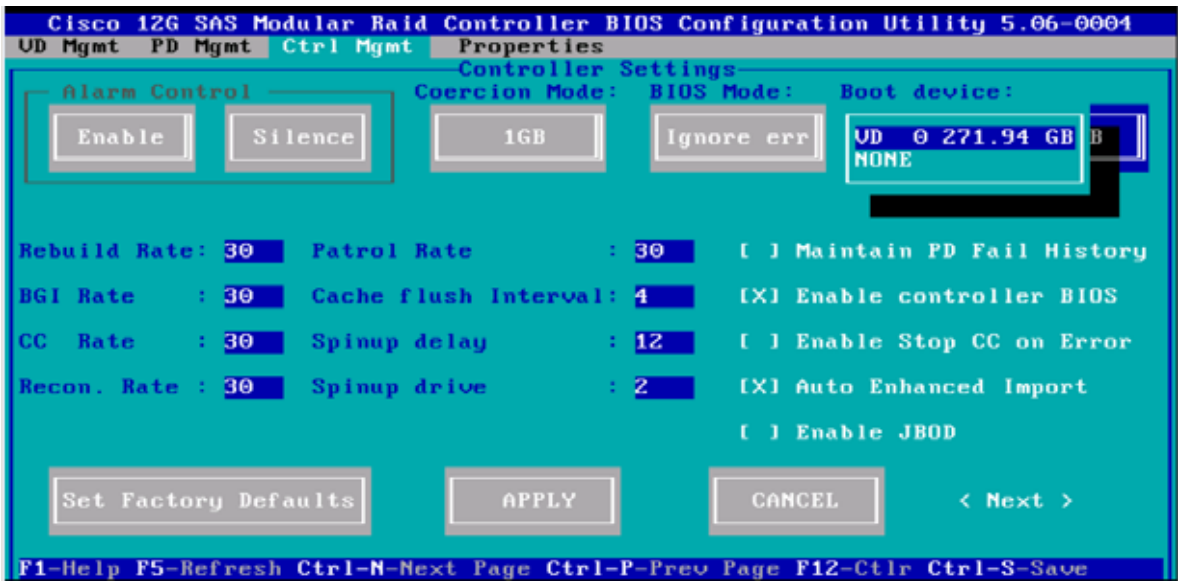


15. Select the **Advanced** button, and Check the Initialize checkbox.
16. Press **OK** to continue with initialization.

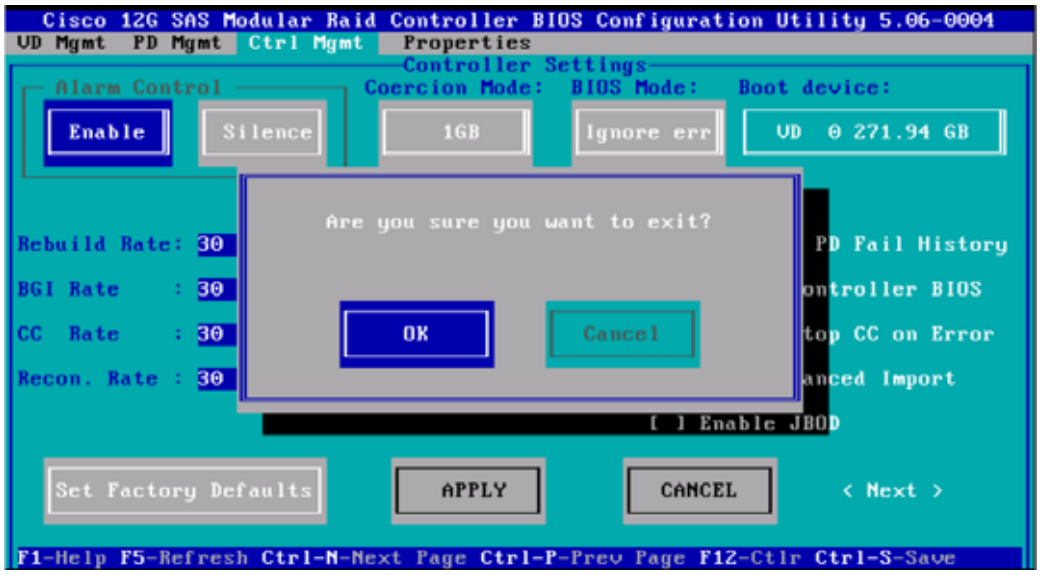
17. After the initialization is complete, the following message appears. Press **OK** to continue.



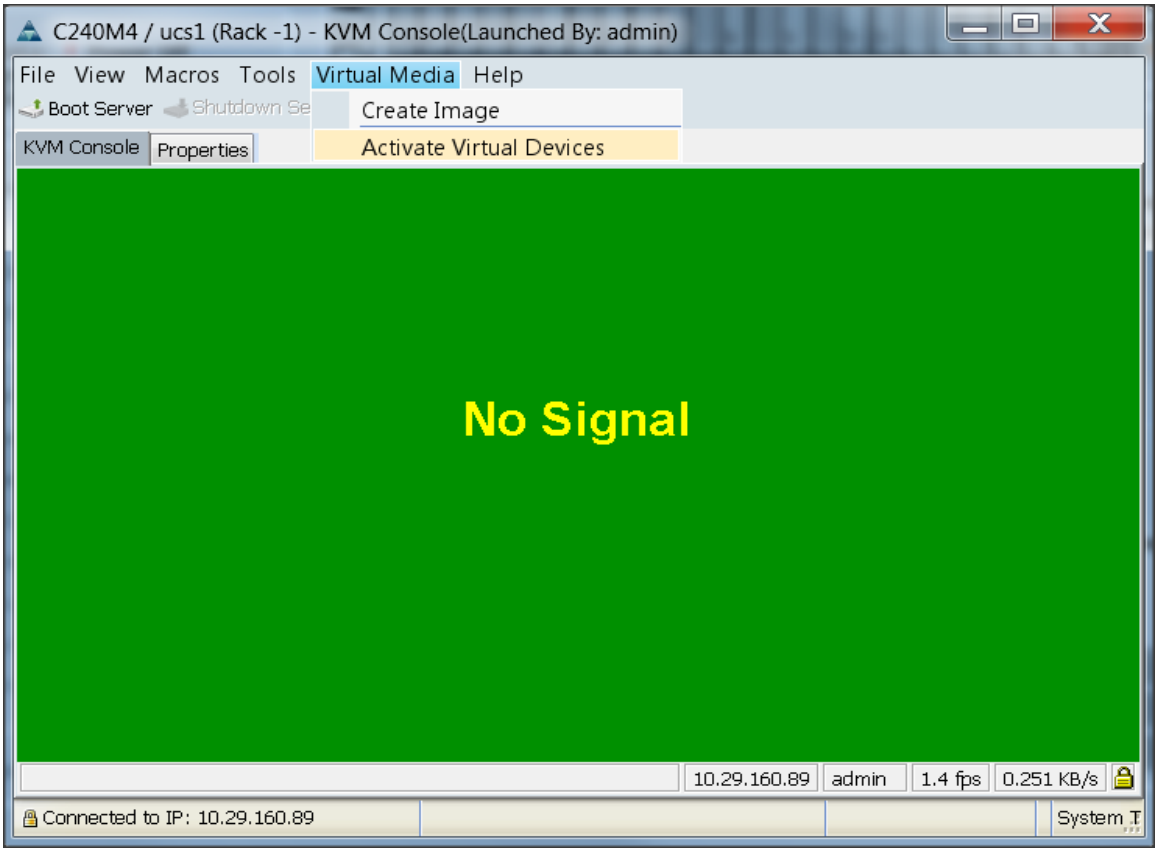
18. Press **Ctrl-N** twice to navigate to the Ctrl Mgmt screen.
19. Select Boot device field and press **ENTER**.



20. Select the **VD 0**, and press **ENTER** again.
21. Press **Ctrl+S** to save the configuration.
22. Press **ESC** to exit the MegaRAID configuration utility.

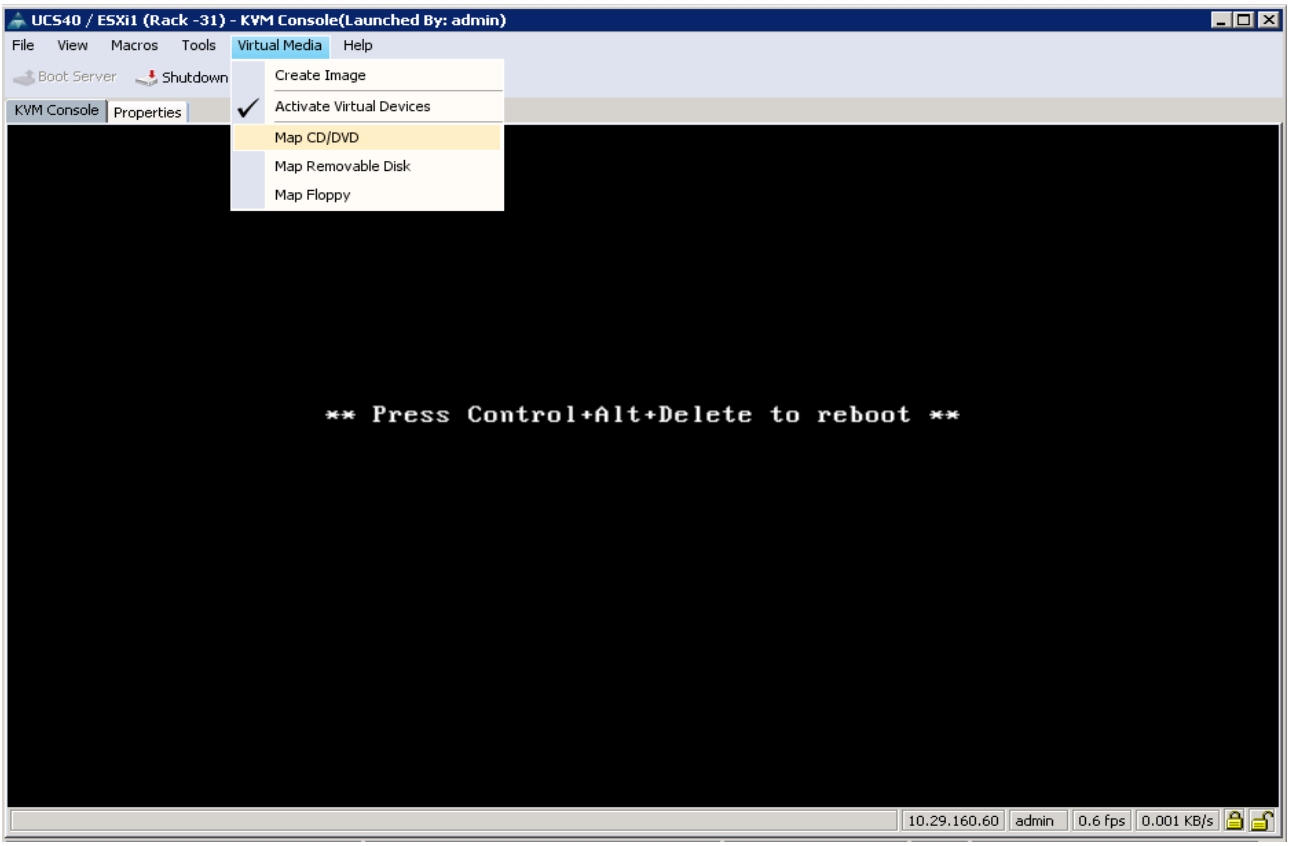


23. In the KVM window, select the Virtual Media menu.
24. Click the Activate Virtual Devices found in the right hand corner of the Virtual Media selection menu.



25. In the KVM window, select the Virtual Media menu and Select **Map CD/DVD**.

Figure 106 Mapping the CD/DVD Virtual Media

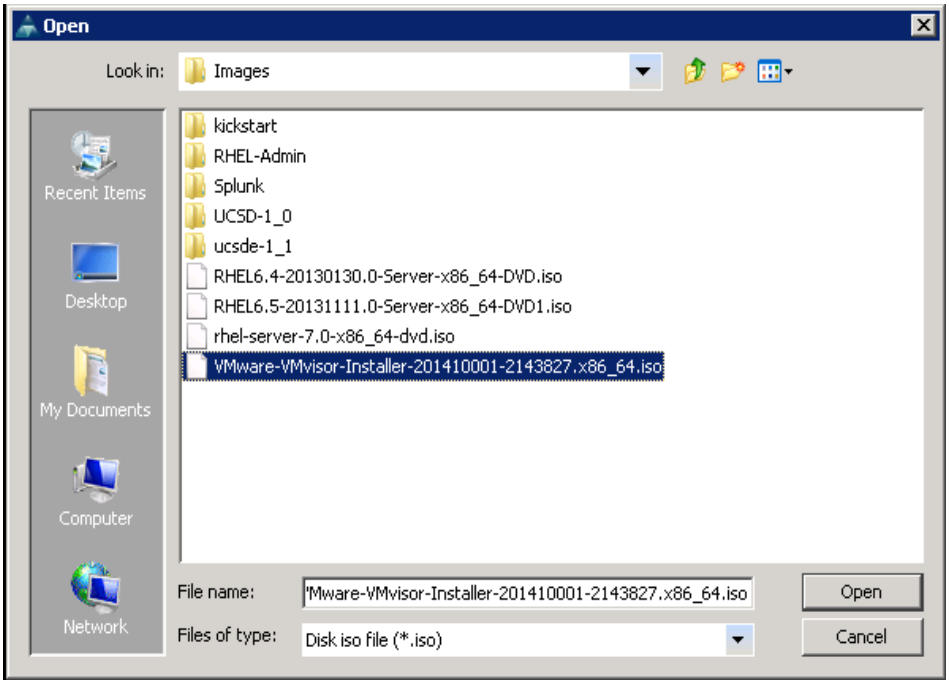


26. Browse to the VMware vSphere ESXi 5.5 installer ISO image file.

Note The VMware vSphere ESXi 5.5 installable ISO is assumed to be on the client machine.

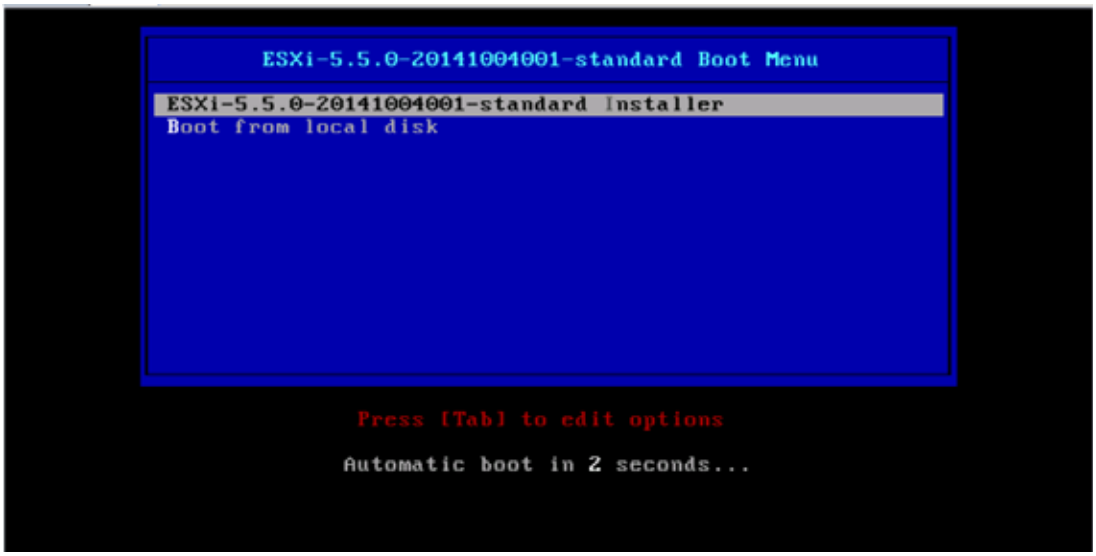
27. Click **Open** to add the image to the list of virtual media.

Figure 107 Browse to VMWare ESXi Hypervisor ISO Image



28. In the KVM window, select the **KVM** tab to monitor during boot.
29. In the KVM window, select the **Macros** > **Static Macros** > **Ctrl-Alt-Del** button in the upper left corner.
30. Click **OK** to reboot the system.
31. On reboot, the machine detects the presence of the VMWare ESXi install media.

Figure 108 ESXi Standard Boot Menu



32. Select the ESXi-5.5.0-yyyymmddnnnn-standard Installer. The installer begins automatically.

Figure 109 Loading the ESXi Installer

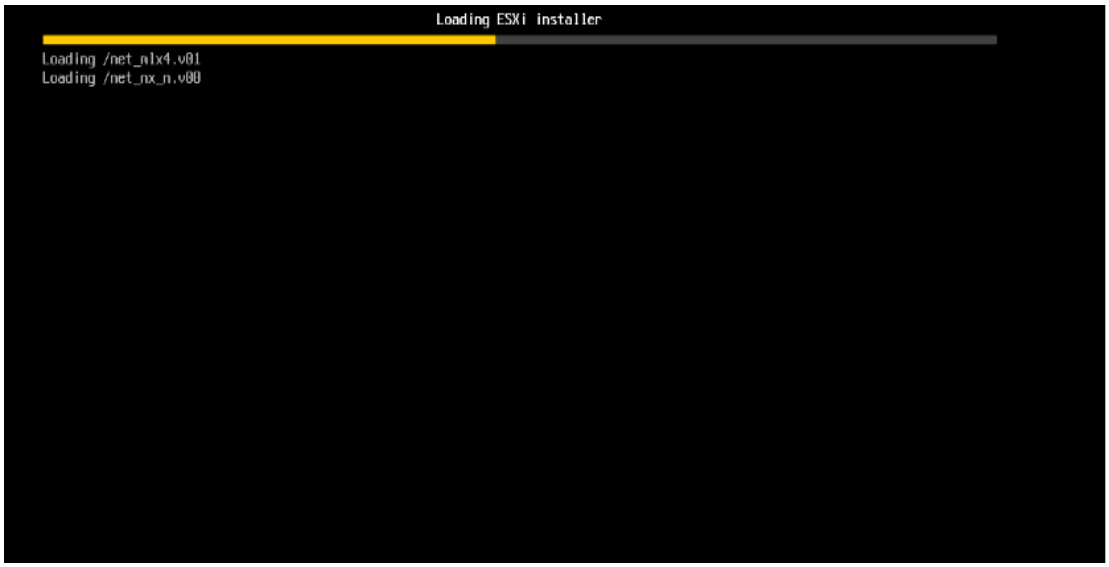
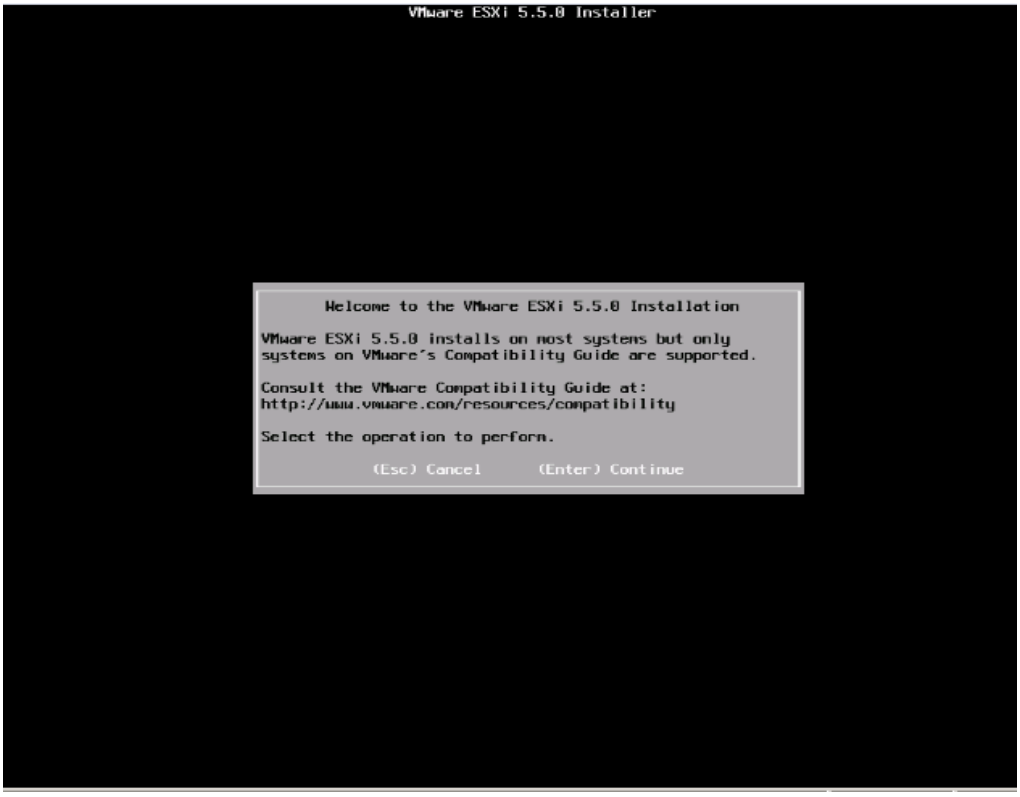


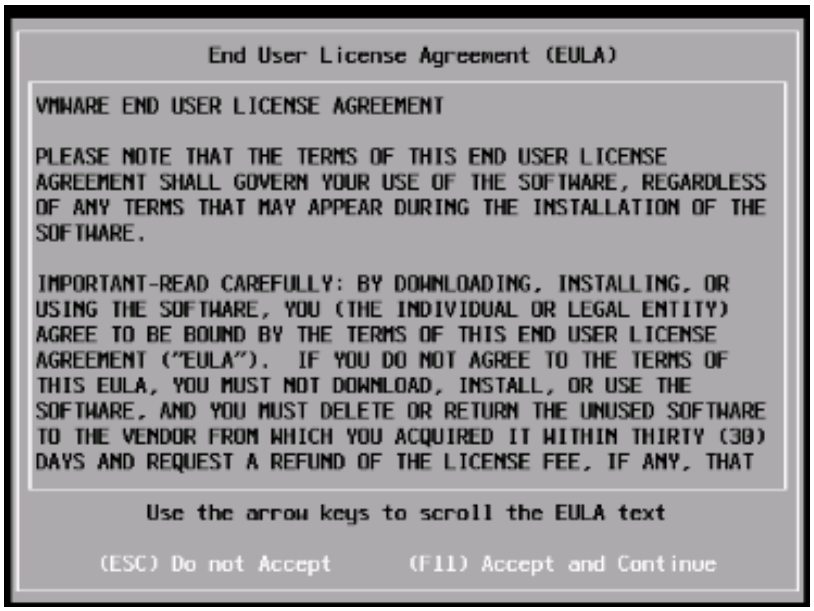
Figure 110 VMWare ESXi Installation screen



33. Press **ENTER** to continue.

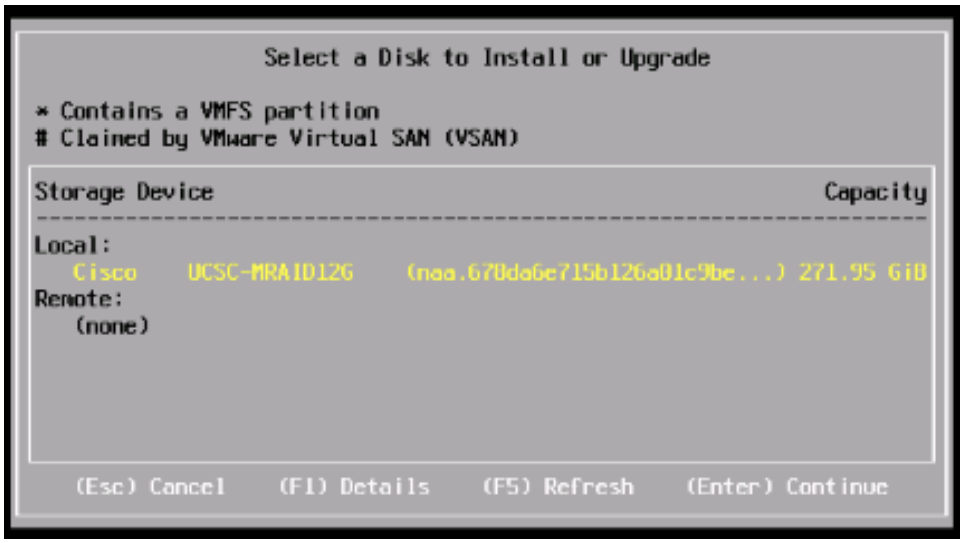
34. Press **F11** to accept End user License Agreement (EULA) and continue.

Figure 111 Accept End User License Agreement (EULA)



35. Select the storage device. Press **ENTER** to proceed with the installation.

Figure 112 Selecting the Storage Device for installing the ESXi operating system.



36. Select the Keyboard US Default. Press **ENTER** to continue.

Figure 113 Choose the Keyboard layout

37. Choose the root password and confirm it. Press **ENTER** to continue.

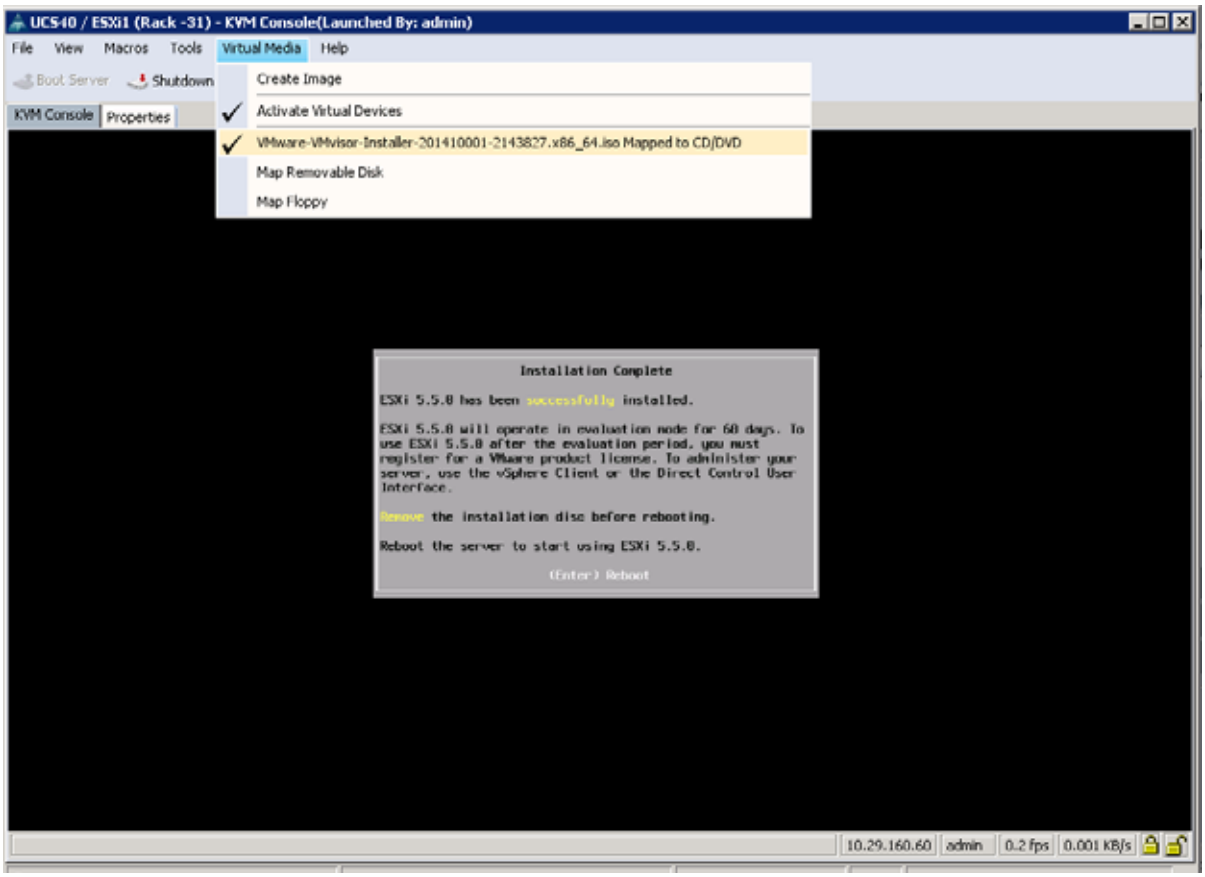
Figure 114 Choose the root password

38. Press **F11** to confirm and begin installation.

39. Once the installation completes, the following message is displayed in the KVM.

40. Remove the VMWare vSphere Hypervisor's ISO from the Virtual Media menu, by selecting it as shown.

Figure 115 ESXi installation complete – Unmount the Virtual Media



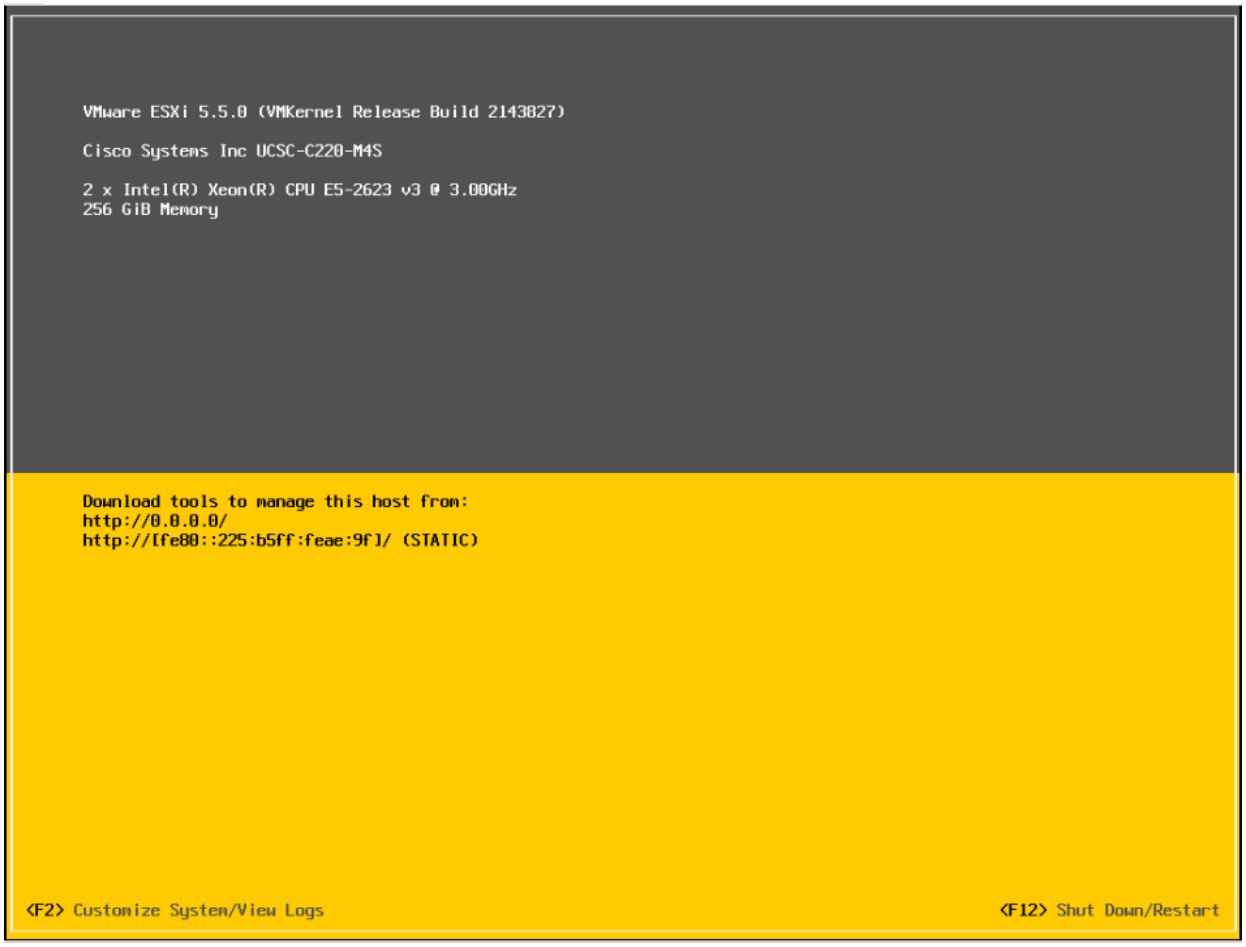
41. Click **Yes** to proceed with un-mapping of the ISO.
42. Press **ENTER** to reboot the server.

The VMWare vSphere ESXi installation is complete.

Configuring the Management Network

1. Once the server reboots, press **F2** to log on.
2. Enter username as root, and the password chosen above.

Figure 116 VMWare ESXi initial screen as seen via the KVM Console

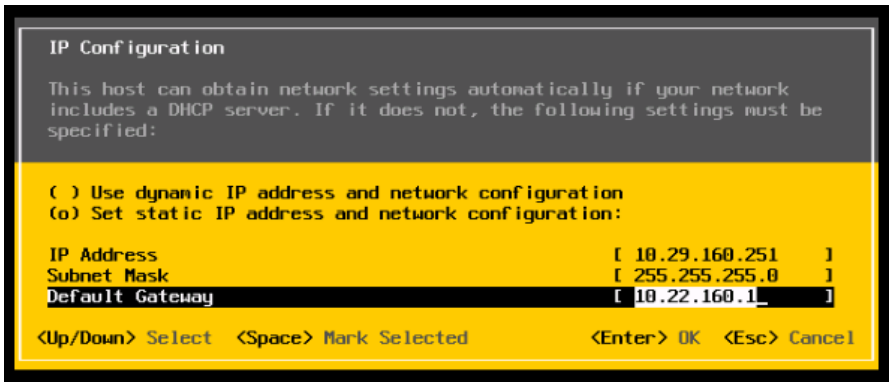


3. Press **F2** to continue
4. Select Configure Management Network, and press **ENTER.**
5. Select **IP Configuration** option.

Figure 117 Enter the IP configuration option of the Management Network

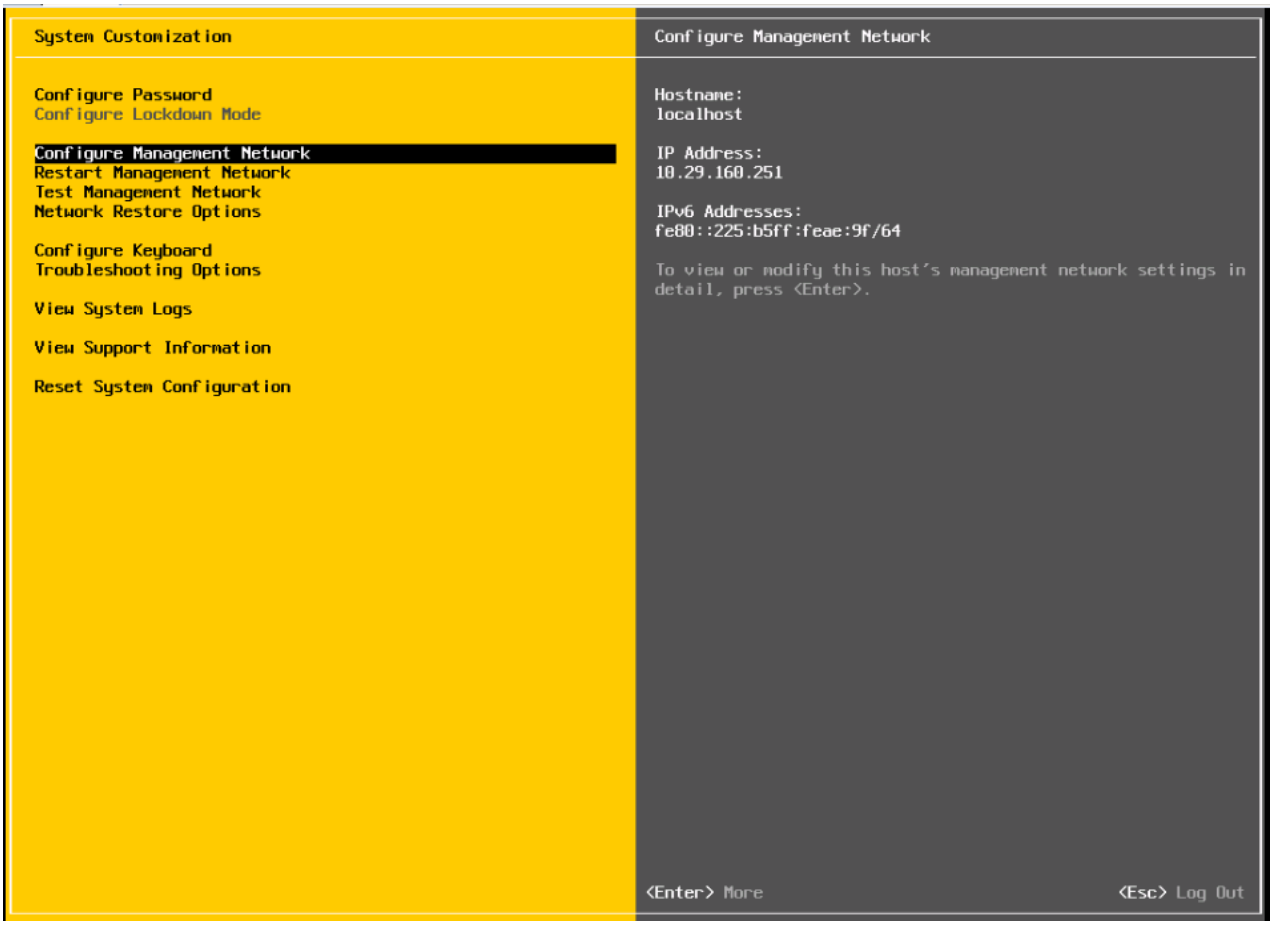
6. Press **ENTER** to continue.
7. Use the Up/Down arrow keys to highlight the Set Static IP address and network configuration option, and press **SPACE** key to select it.
8. Enter the static IP address, Subnet Mask and Default Gateway.

Figure 118 Enter the IP Address configuration details



9. Press **OK** to submit the changes.
10. Press **ESC** key exit the Management Network Screen.
11. In the Configure Management Network: Confirm dialog box, Press **Y** to restart the Management Network.
12. Verify the IP address settings in the System Customization screen.

Figure 119 Verify the IP address details in the System Customization screen



Installing the VMWare ESXi client software

1. Using a web browser, visit the url: https://10.29.160.251/
2. Click on Download vSphere Client.

Figure 120 Accessing the ESXi web interface

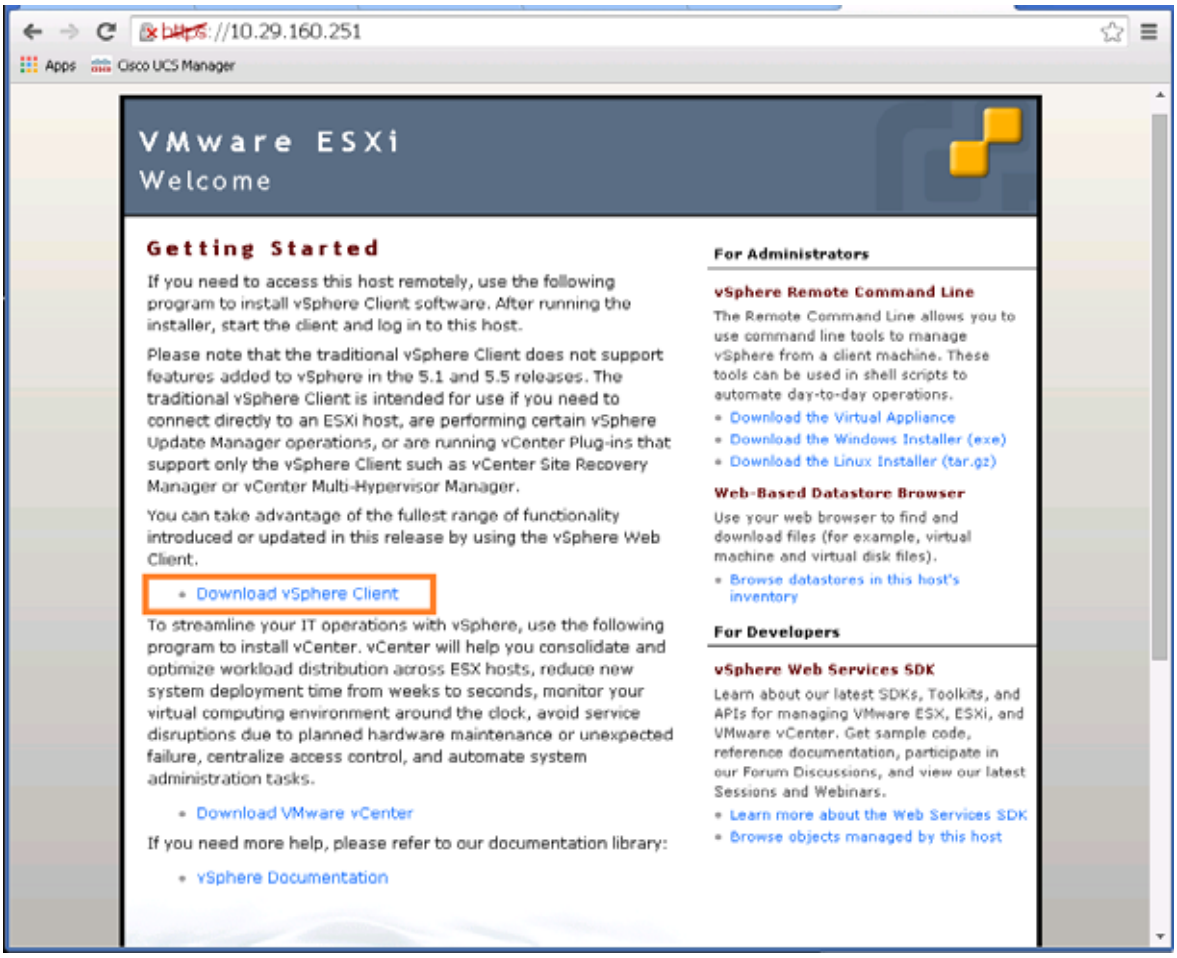
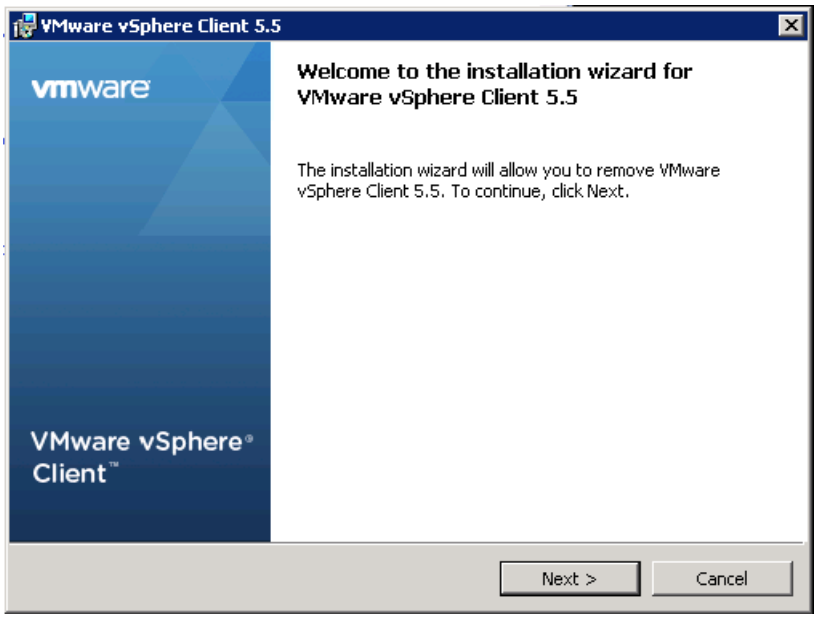


Figure 121 Download the VMWare vSphere ESXI client software

3. Proceed to install the downloaded VMWare client software.

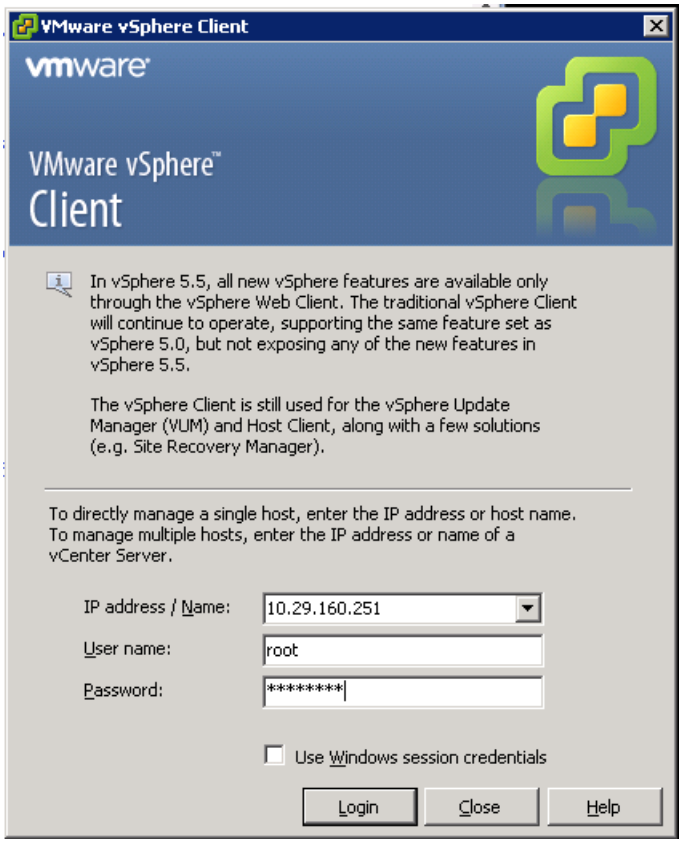
Figure 122 Installing the vSphere Client software



Configuring the vSphere ESXi hypervisor

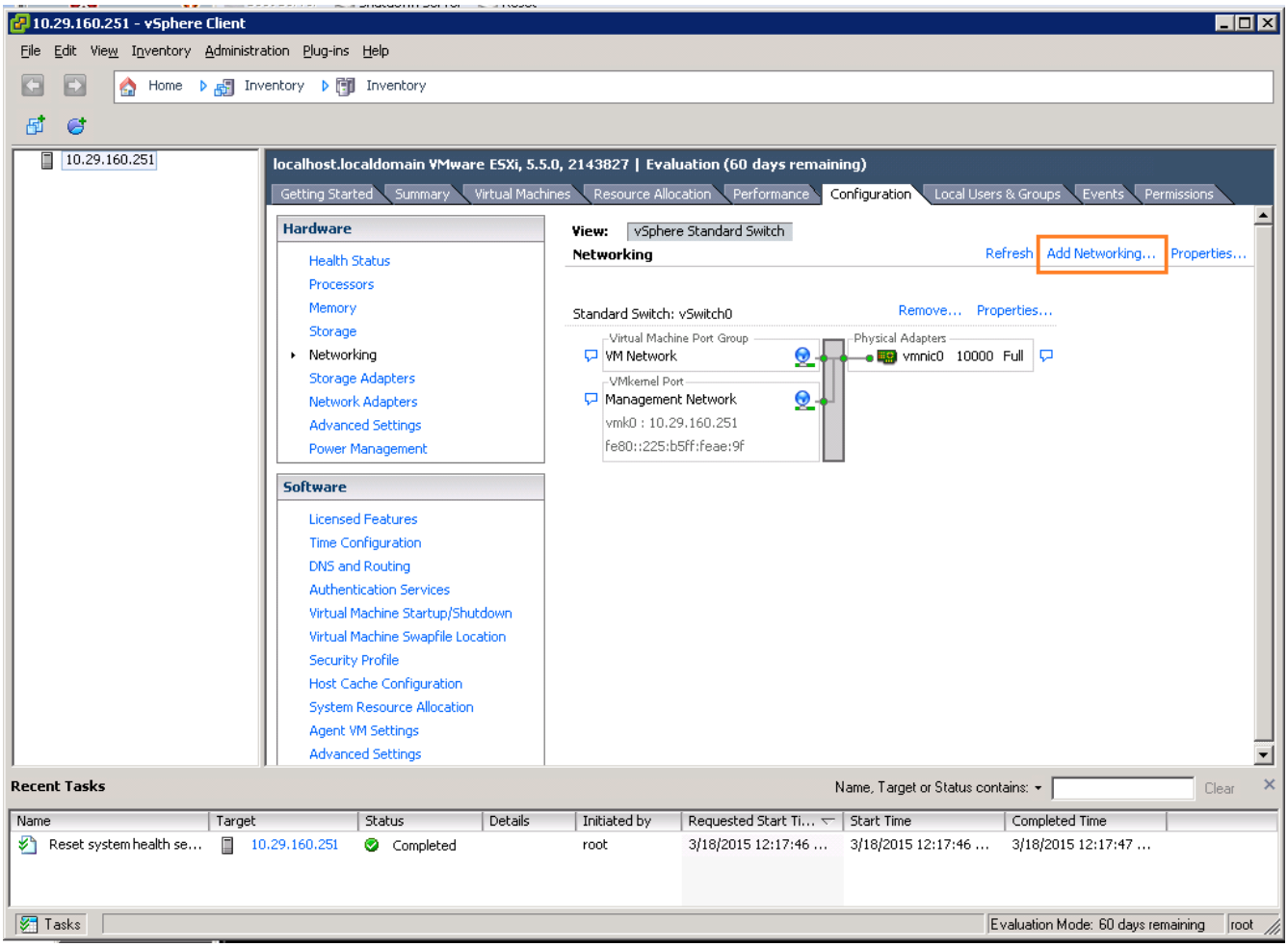
1. After the installation is complete, launch the VMWare vSphere client.
2. Enter the chosen IP address, the username as root, and the chosen password.
3. Click on **Login** to continue.

Figure 123 Logging into the ESXi using vSphere Client



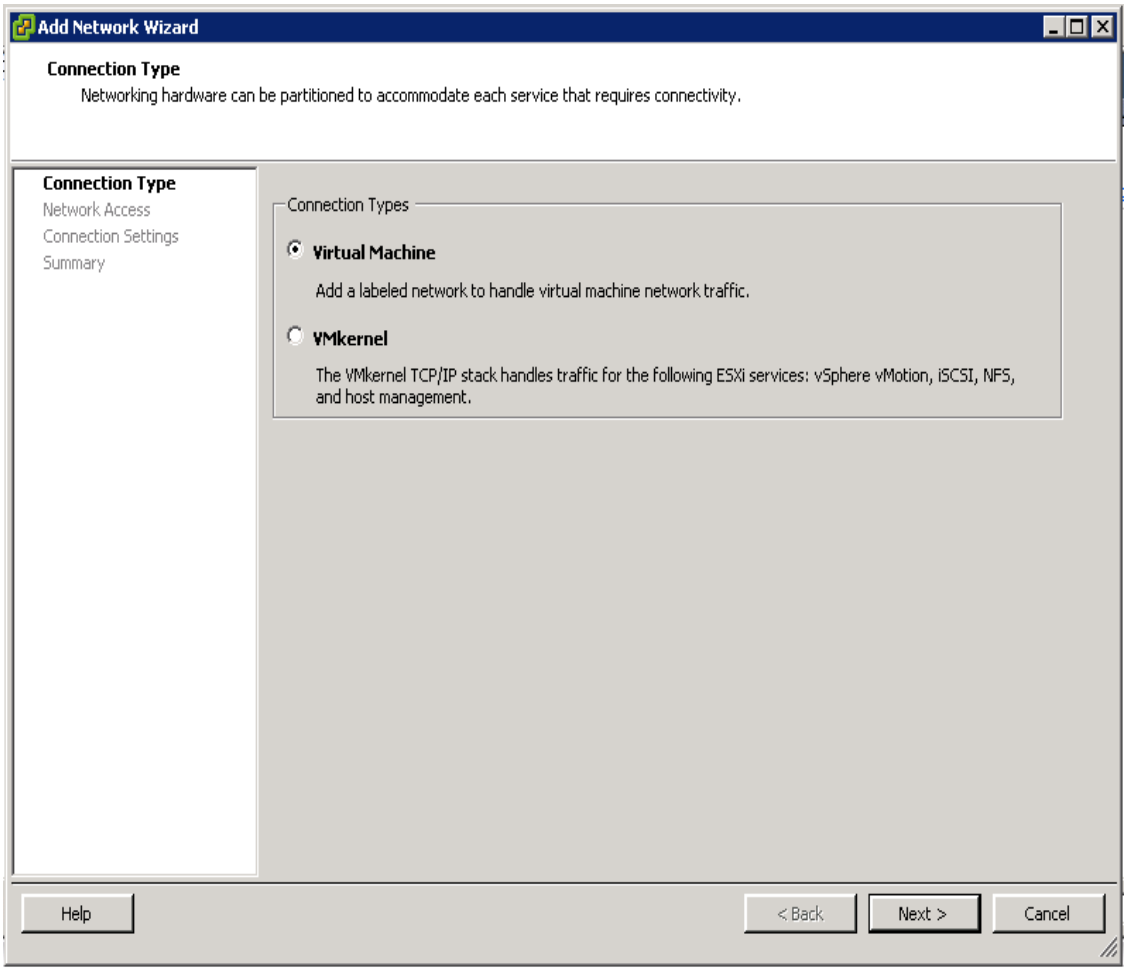
4. In the vSphere Client, click on the Configuration tab on the right, and within the Hardware section, click on Networking.
5. Click on Add Networking link on the upper right hand side.

Figure 124 vSphere Client Networking screen



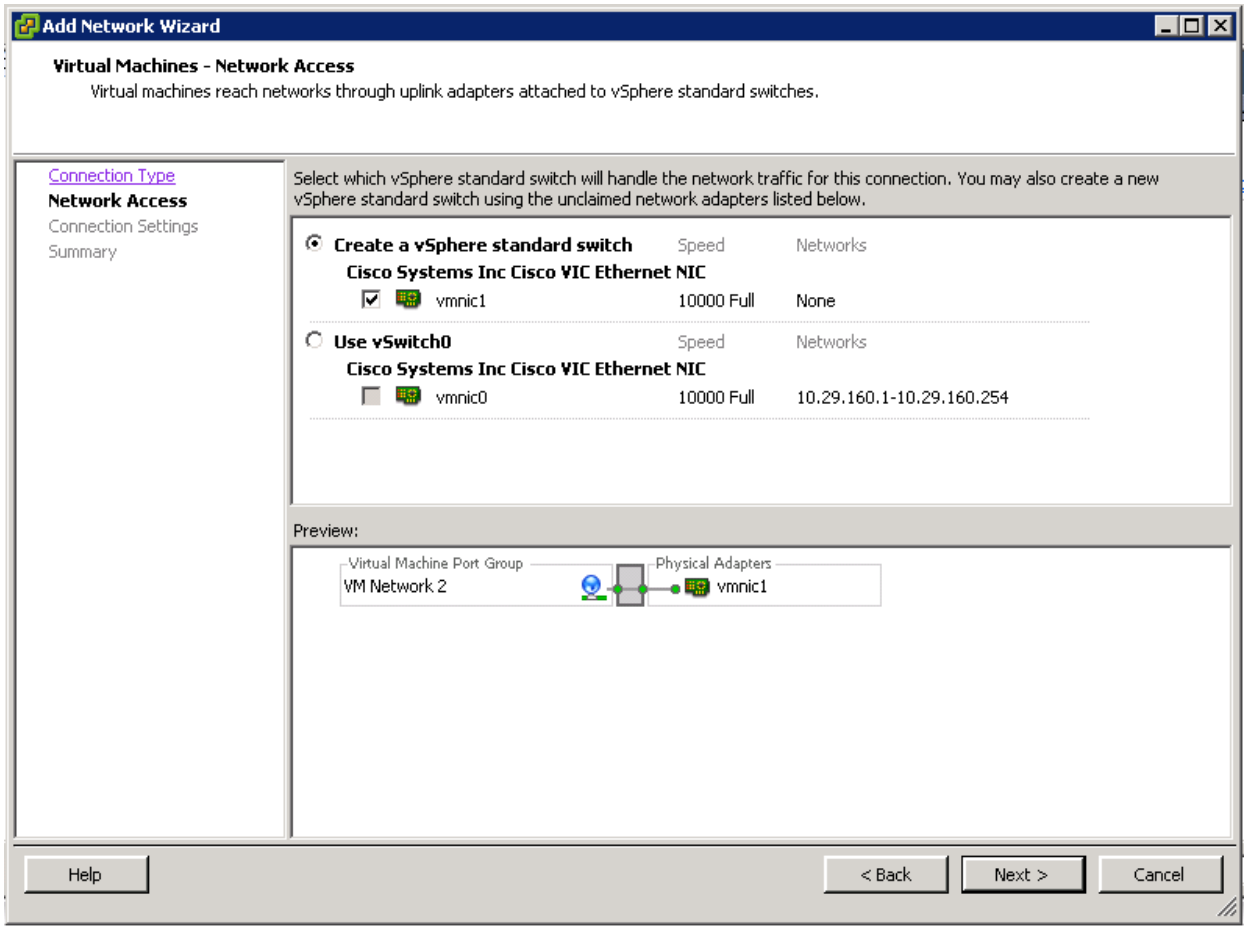
6. In the Add Networking dialog box, click the **Virtual Machine** radio button and click **Next**.

Figure 125 Adding a new Virtual Machine Network



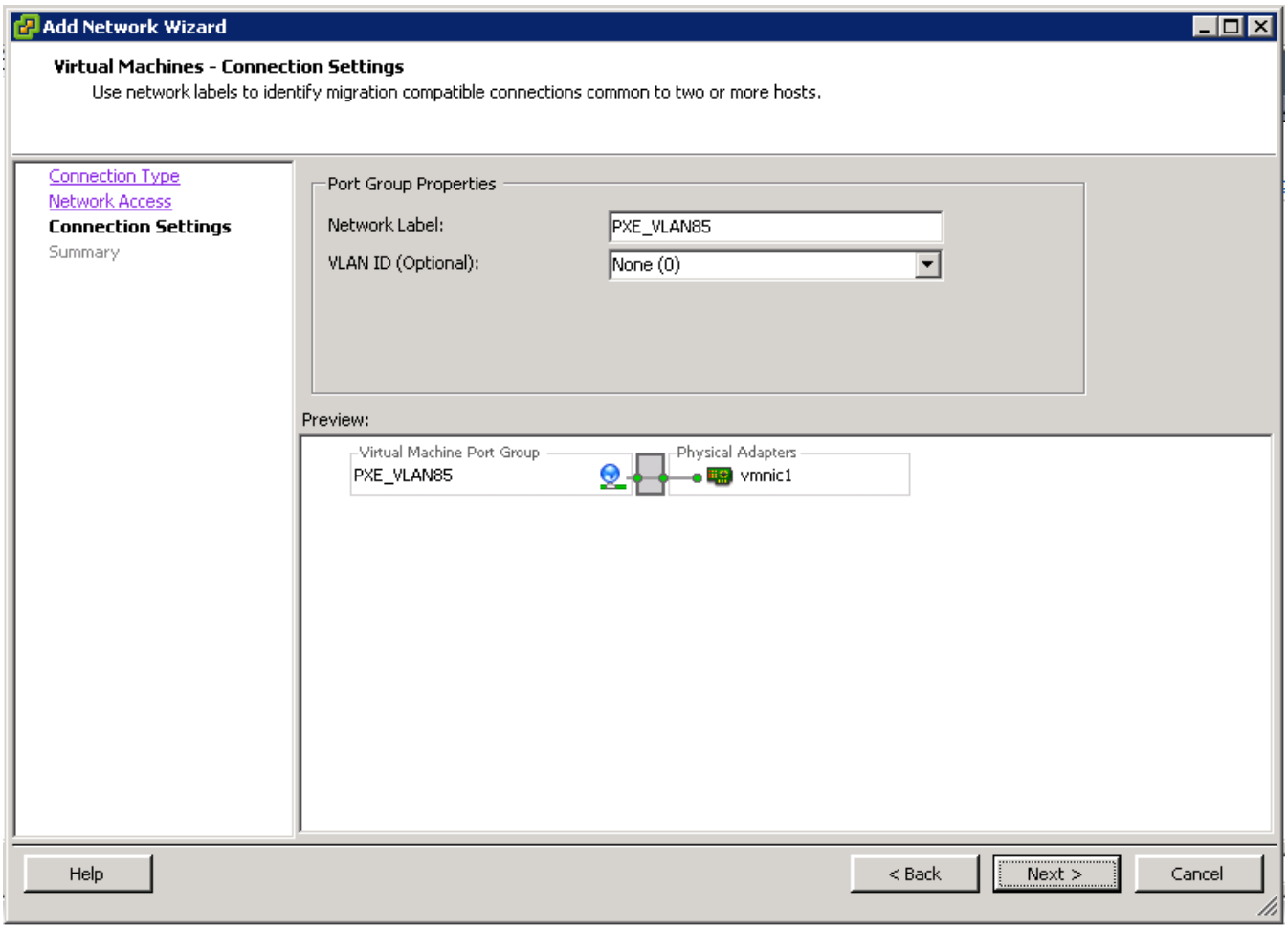
7. Click the **Create a vSphere standard switch** radio button and make sure that the checkbox next to vmnic1 is checked.
8. Click **Next**.

Figure 126 Creating a new vSphere Standard Switch



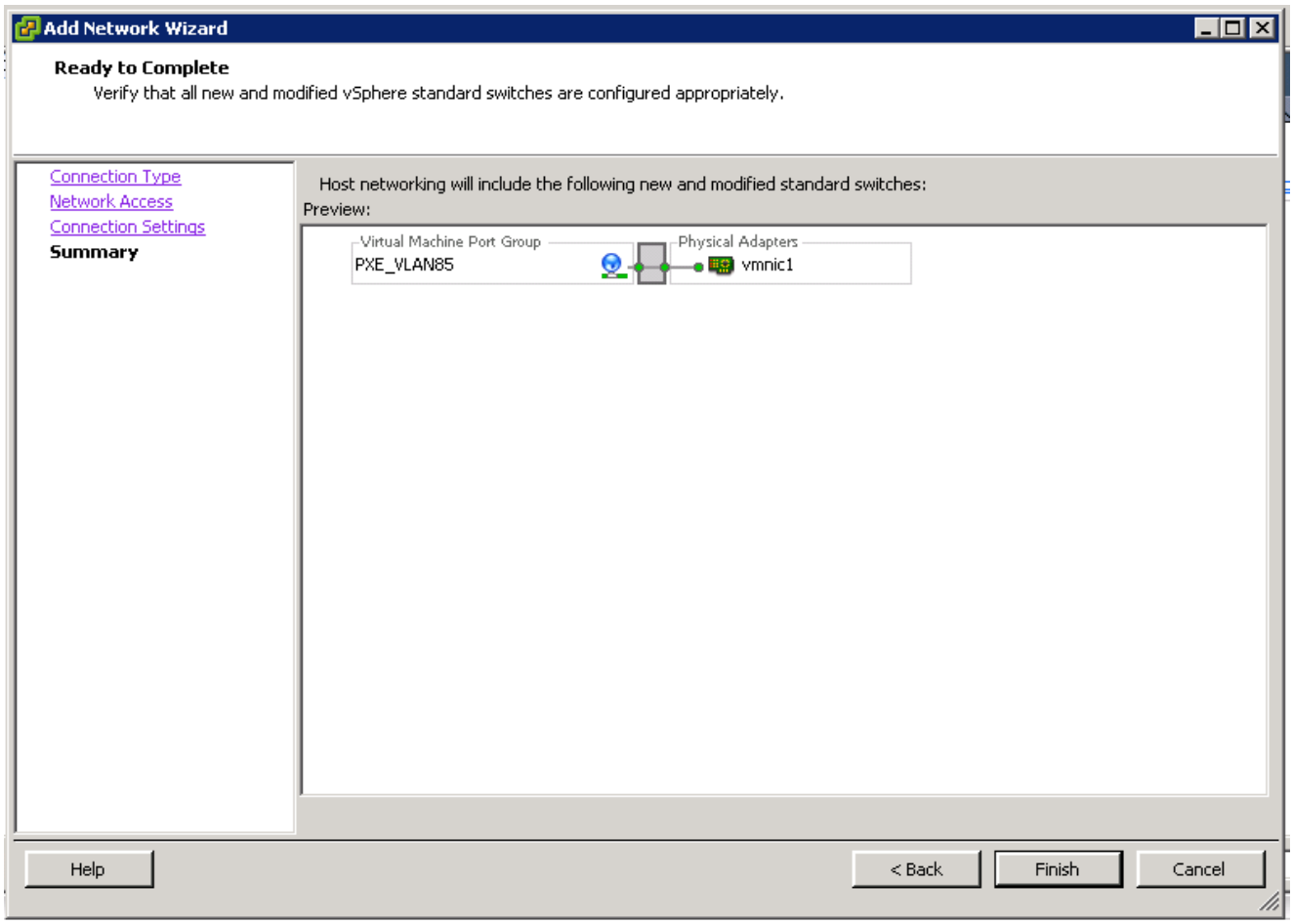
9. In the Port Group Properties, change the Network Label field to PXE_VLAN85.
10. Leave the VLAN ID(Optional) field as None(0).
11. Click **Next**.

Figure 127 Creating the Port Group for the PXE VLAN



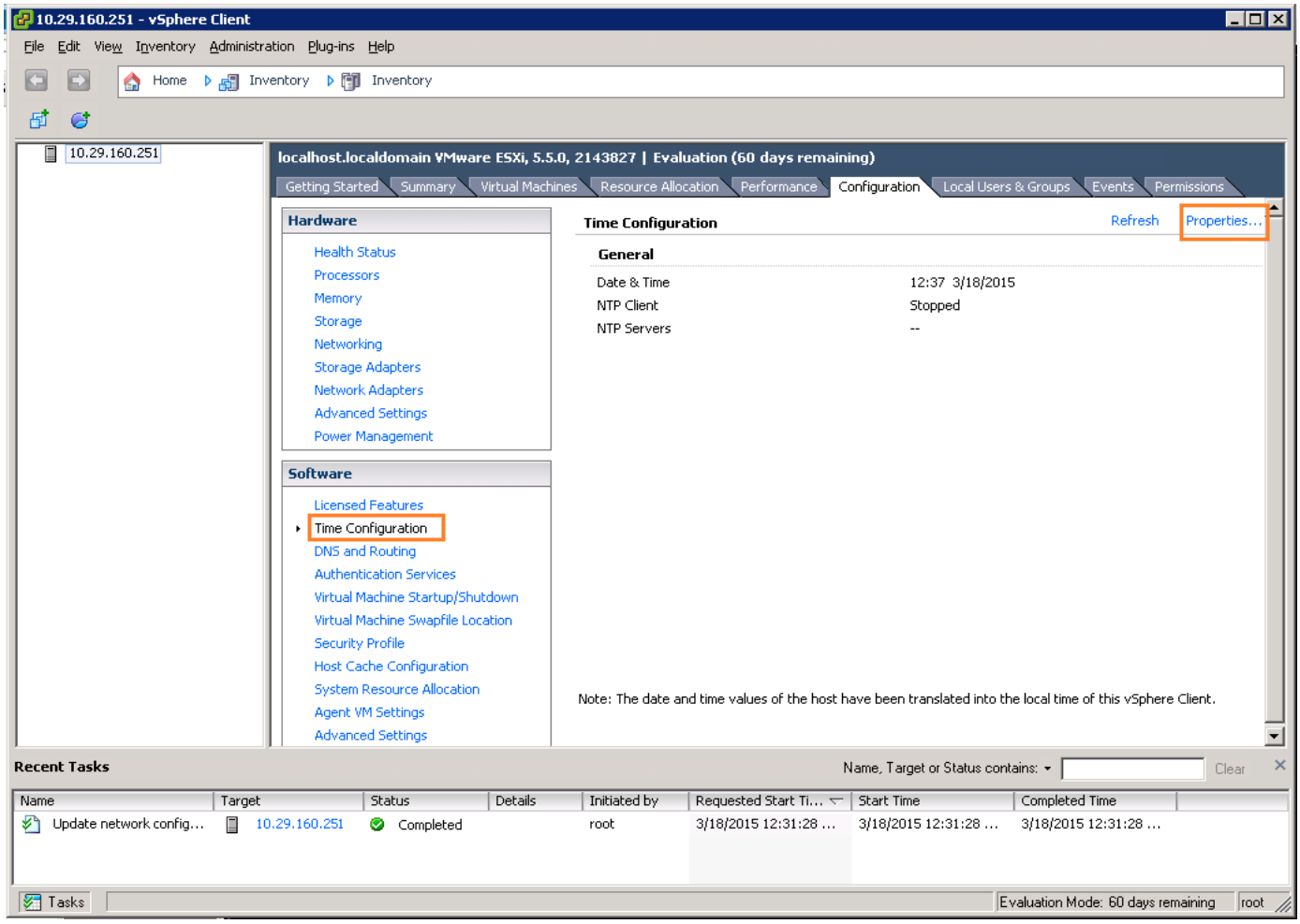
12. Click **Finish** to complete adding the Network.

Figure 128 Verify the Created vSphere Standard Switches



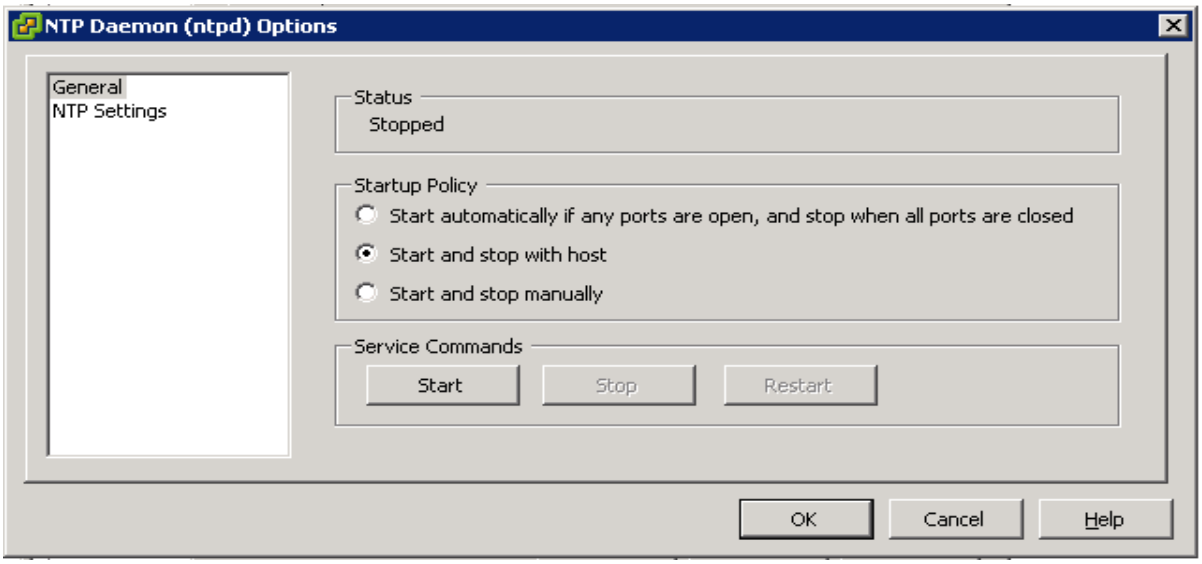
13. Click on the **Time Configuration** under the Software section.
14. Click on **Properties** at the upper right hand corner.

Figure 129 Enabling the NTP Client on the ESXi



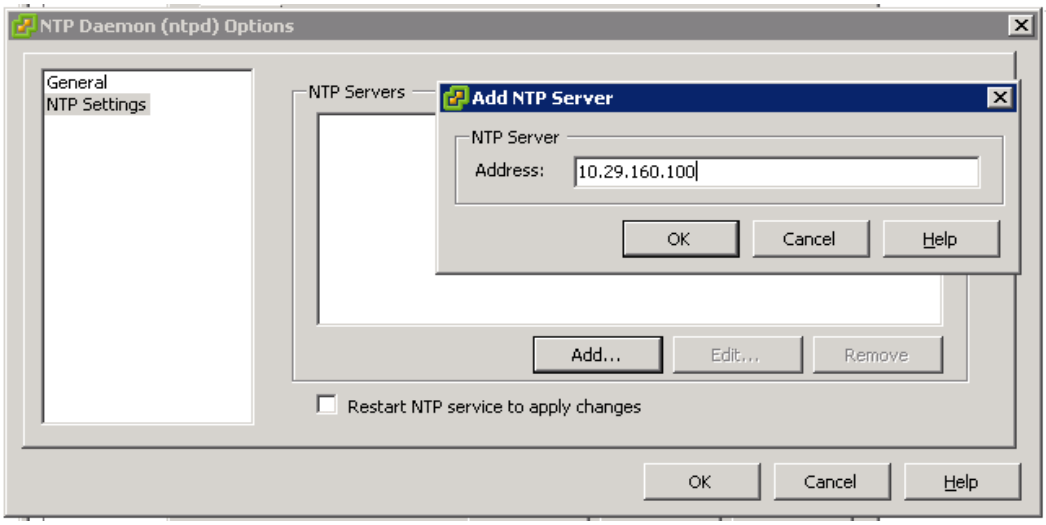
15. In the NTP Daemon (ntpd) Options dialog box, click **Options**.
16. Click on the **General** options.
17. Click to select the start and stop with **host** radio button.

Figure 130 NTP Daemon



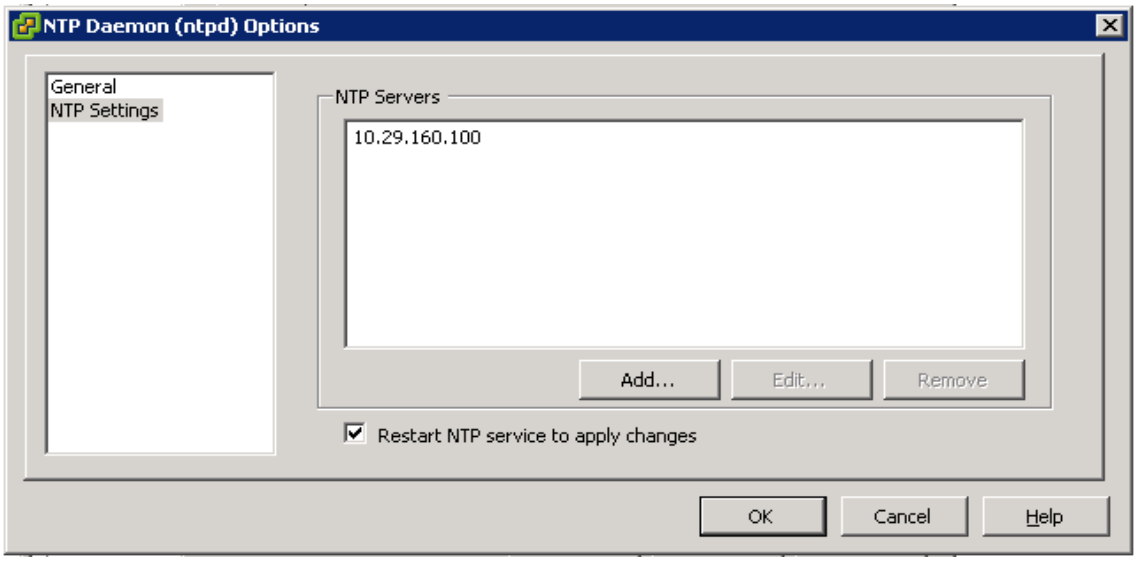
18. Click on **NTP** Settings option.
19. Click on **Add** button to add the NTP server's IP address.
20. Press **OK** to continue.

Figure 131 Adding a new NTP Server to the ESXi NTP Settings



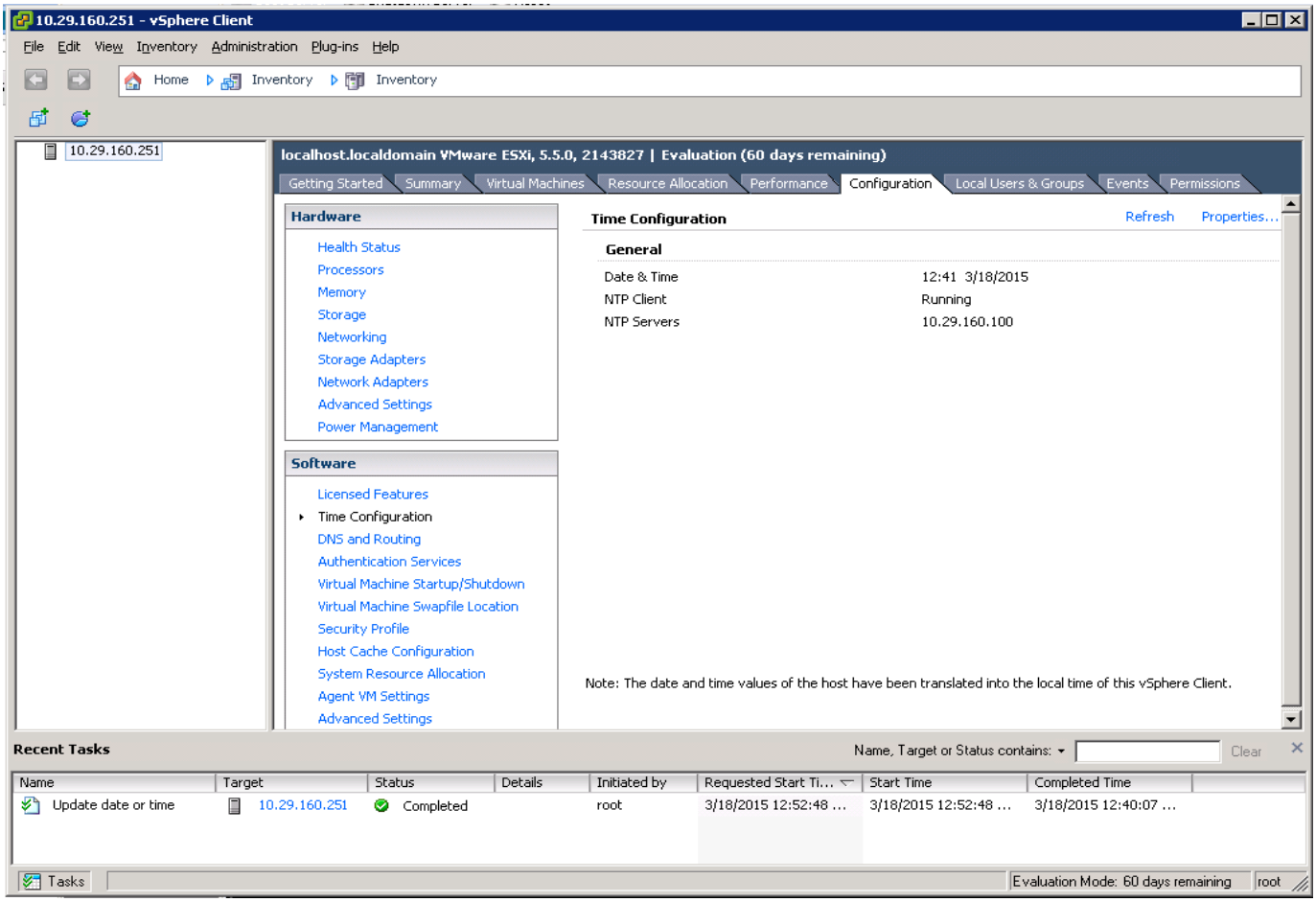
21. In the next screen, verify the IP-address in the NTP Servers list.
22. Click on the checkbox **Restart NTP service to apply changes**.
23. Press the button **OK** twice to complete the time configurations.

Figure 132 Restart NTP Service



24. Time configuration option would now show that the NTP client is running, along with the IP address of the NTP client.

Figure 133 Verifying the NTP Client



Downloading the UCS Director Express Software Components

The software components of UCS Director Express for Big Data need to be downloaded from three different locations.

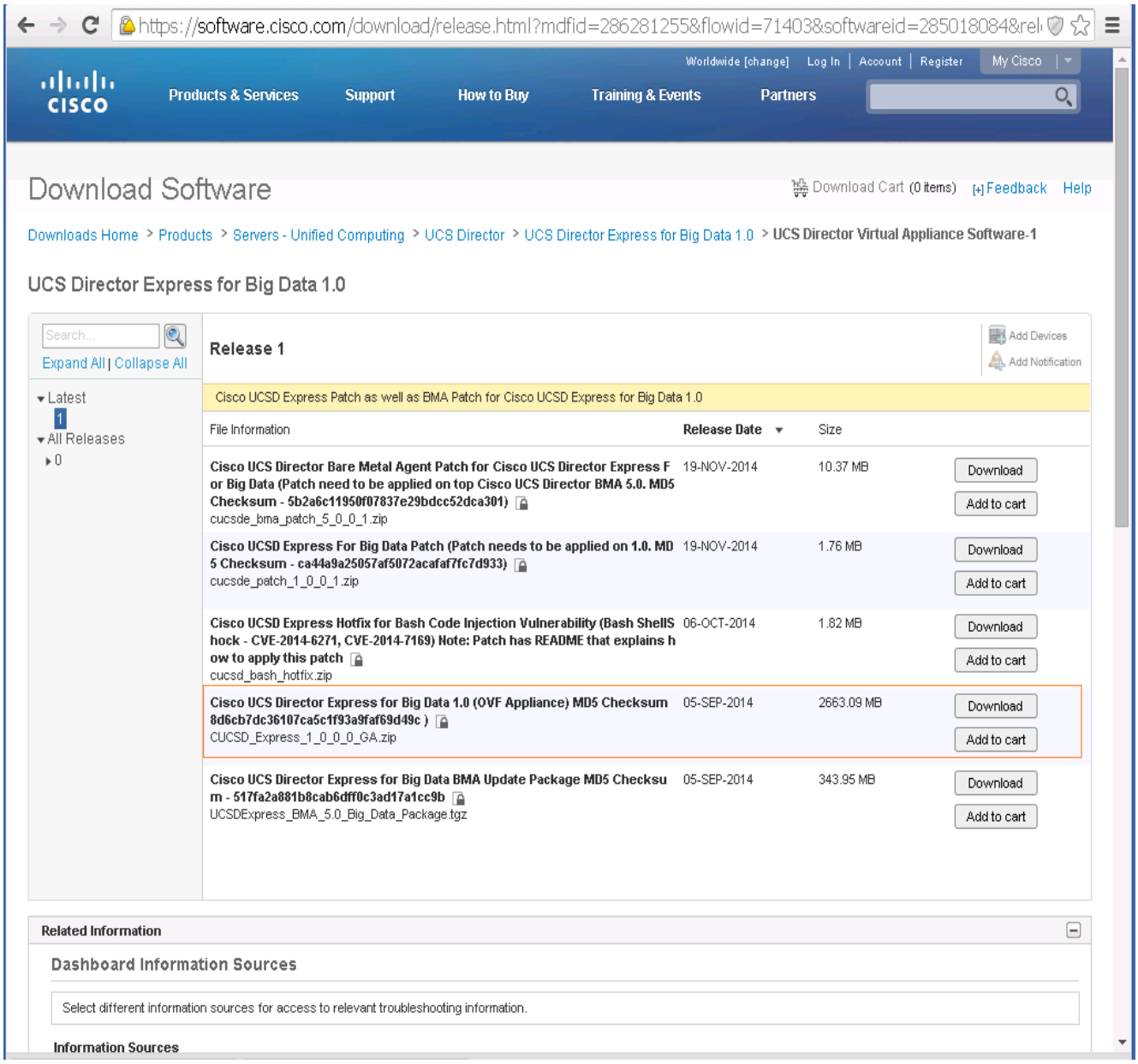
Table 15 Cisco UCS Director Express Big Data 1.1 Software Components

Software component	File Names	Link to Download
Cisco UCS Director Express 1.0 OVF	CUCSD_Express_1_0_0_0_GA.zip	https://software.cisco.com/download/release.html?mdfid=286281255&flowid=71403&softwareid=285018084&release=1&relind=AVAILABLE&rellifecycle=&reltype=latest
Cisco UCS Director 5.2.0.1 patch	cucsd_patch_5_2_0_1.zIP	https://software.cisco.com/download/release.html?mdfid=286283454&flowid=72903&softwareid=285018084&release=5&relind=AVAILABLE&rellifecycle=&reltype=latest
Cisco UCS Director Baremetal Agent 5.2 OVF	CUCSD_BMA_5_2_0_0_VMWARE_GA.zip	
Cisco UCS Director Express for Big Data 1.1 Upgrade Package	UCSDExpress_Big_Data_1.1_Upgrade_Package.zip	https://software.cisco.com/download/release.html?mdfid=286284995&flowid=73724&softwareid=285018084&release=1&relind=AVAILABLE&rellifecycle=&reltype=latest
25. Cisco UCS Director Express for Big Data BMA Update Package	UCSDExpress_BMA_Big_Data_1.1_Upgrade_Package.zip	

Download the Software Components

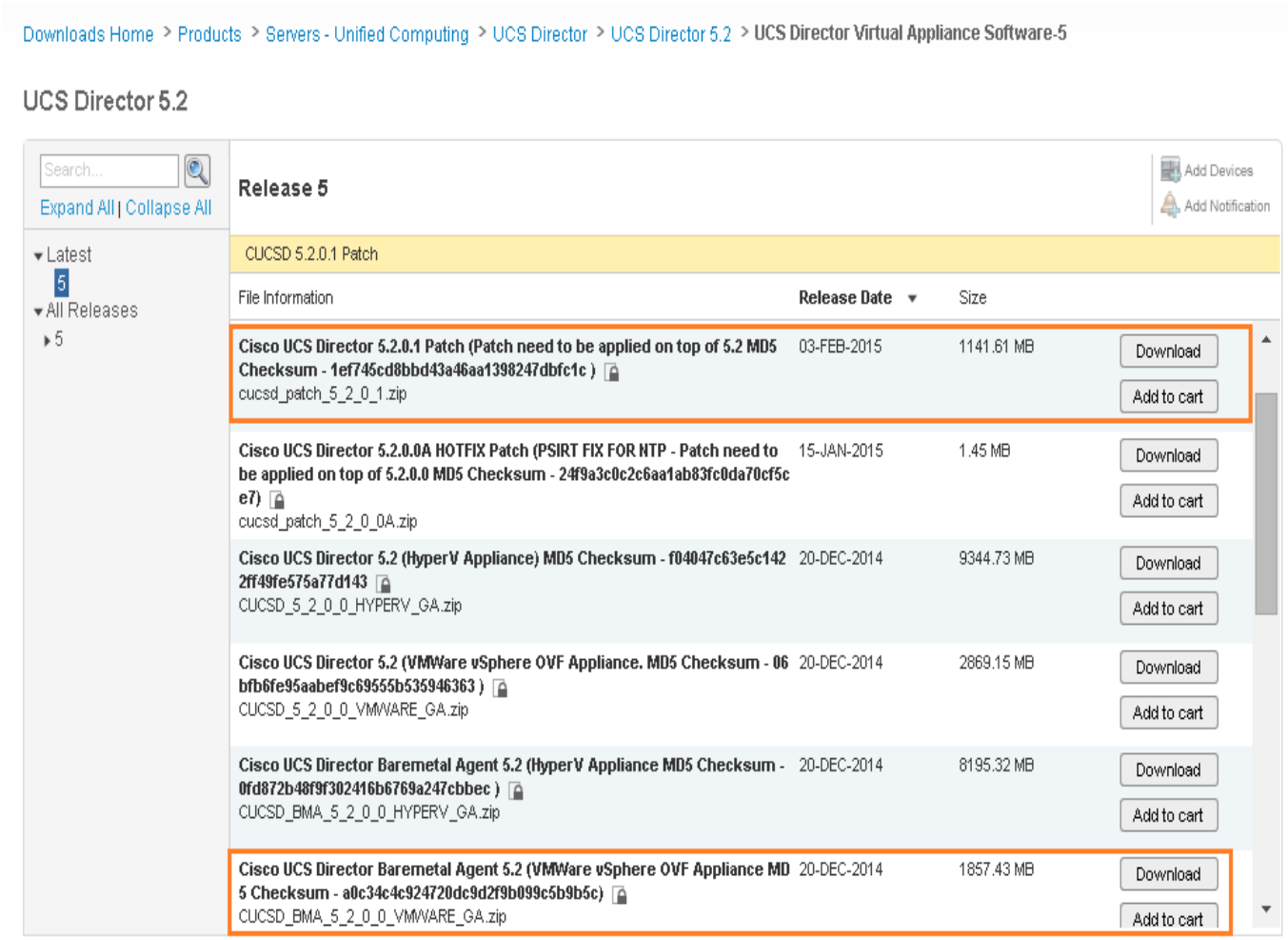
1. Using the links provided Table 15 above, download the Cisco UCS Director Express for Big Data 1.0 OVF Appliance zip file.

Figure 134 Cisco UCS Director Express for Big Data 1.0 Download Page



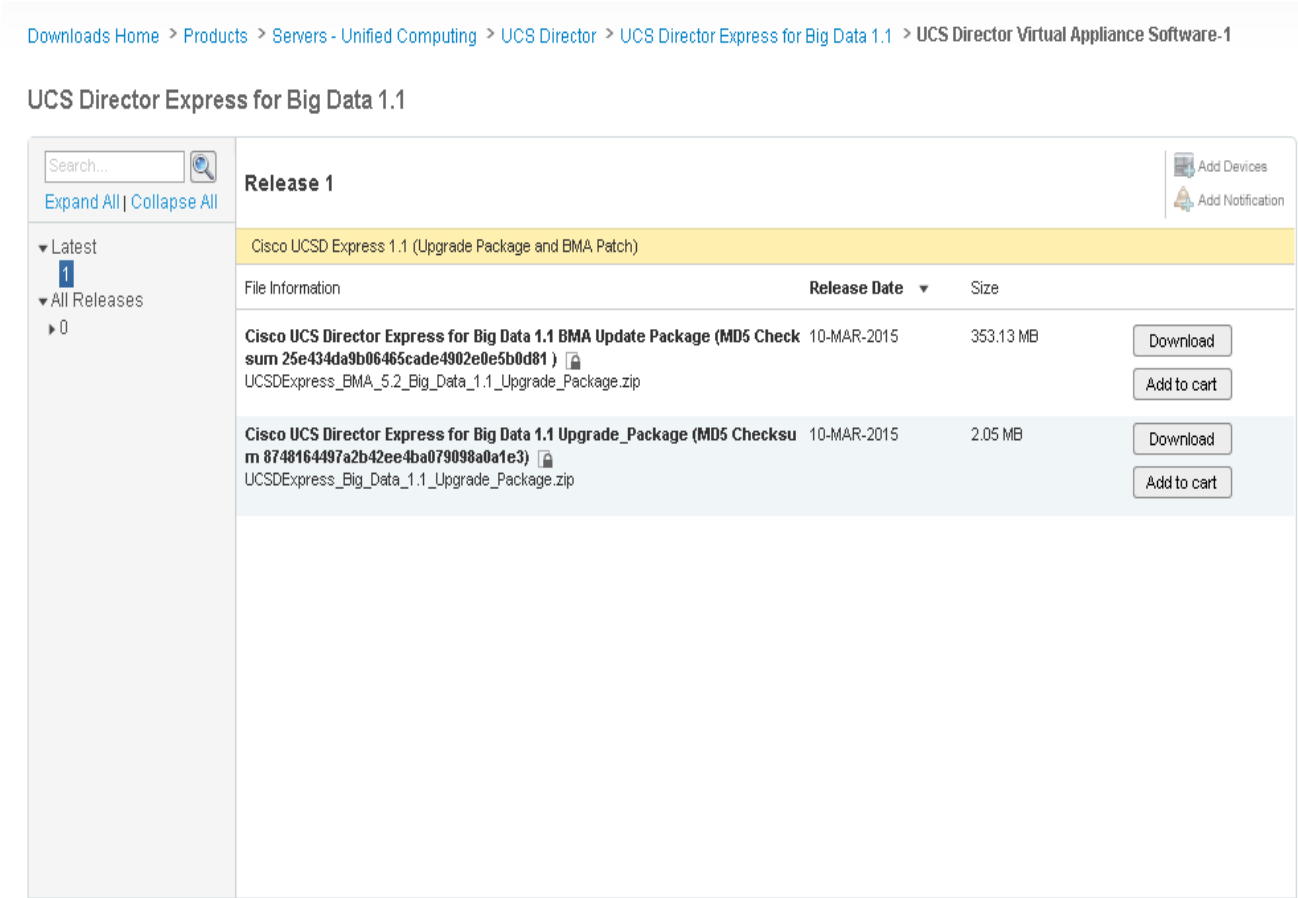
2. Using the links provided Table 15 above; download the Cisco UCS Director 5.2.0.1 Patch zip file, and Cisco UCS Director Baremetal Agent 5.2 VMware vSphere OVF Appliance zip file.

Figure 135 Cisco UCS Director 5.2 Download Page



3. Using the links provided Table 21 above; download the Cisco UCS Director 5.2.0.1 Patch zip file, and the Cisco UCS Director Baremetal Agent 5.2 VMWare vSphere OVF Appliance zip file.

Figure 136 Cisco UCS Director Express for Big Data 1.1 Download Page



4. Place all the files in a directory in the client windows workstation.
5. Unzip the contents of the CUCSD_Express_1_0_0_0_GA.zip and CUCSD_BMA_5_2_0_0_VMWARE_GA.zip.

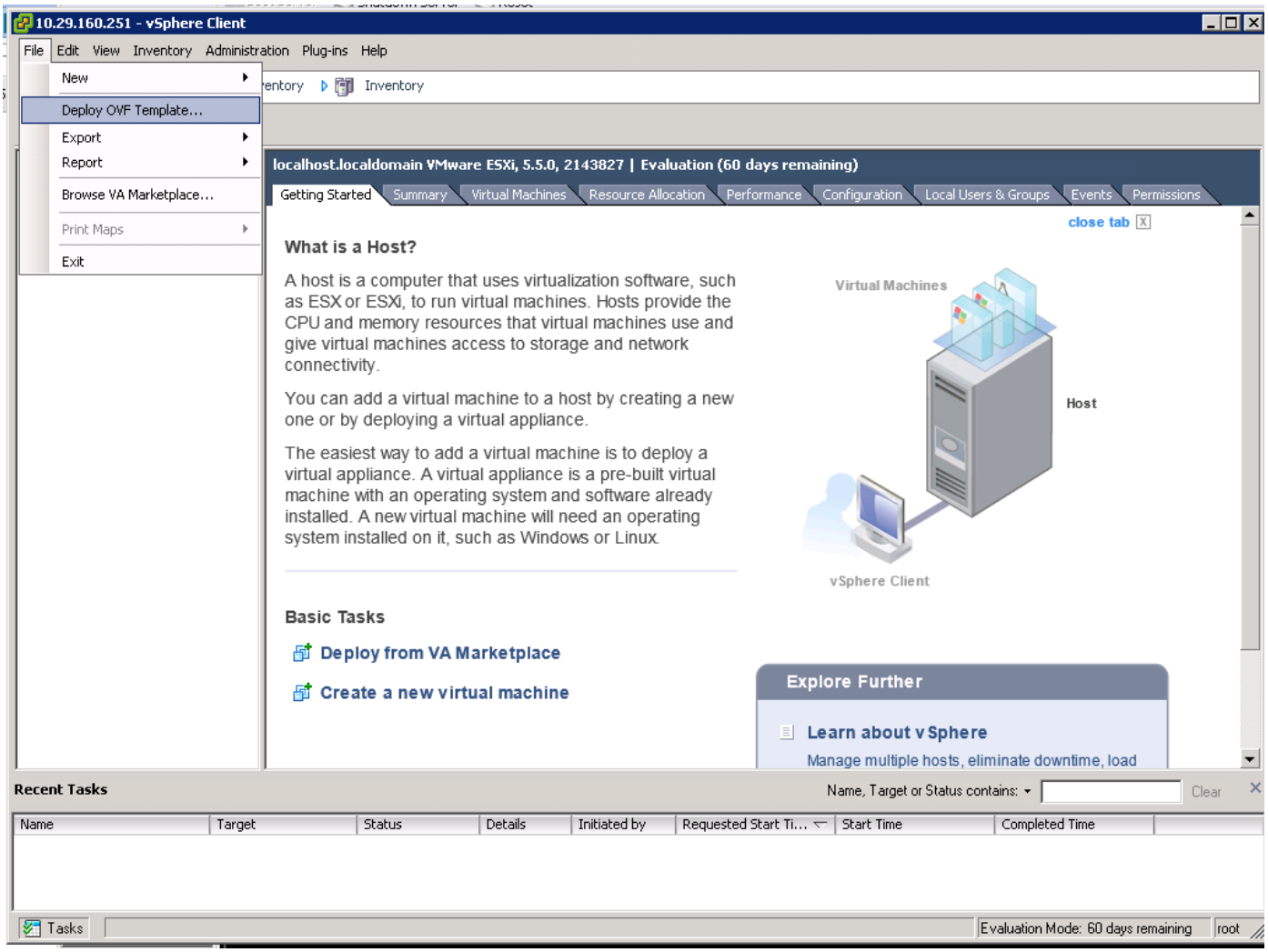
Installing Cisco UCS Director Express for Big Data

The Cisco UCS Director Express for Big Data shall be installed on the VMWare vSphere hypervisor using the vSphere Client software.

Deploying the Cisco UCS Director Baremetal Agent OVF

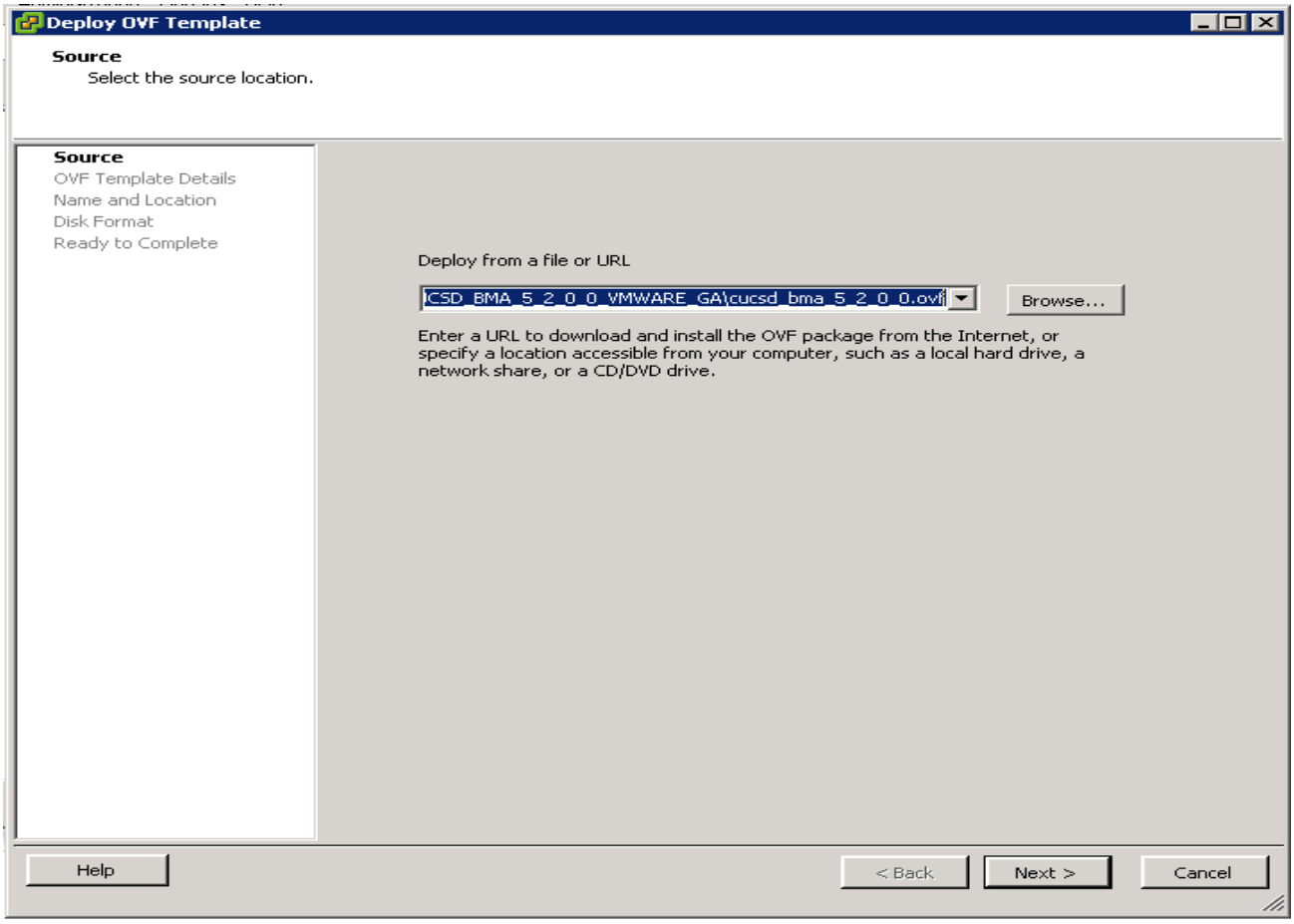
1. Launch the VMWare vSphere client software
2. Enter the chosen IP address, the username as root, and the chosen password.
3. Click on **Login** to continue.
4. From the **File** menu, Select **Deploy OVF Template**.

Figure 137 Deploy OVF in the vSphere Client



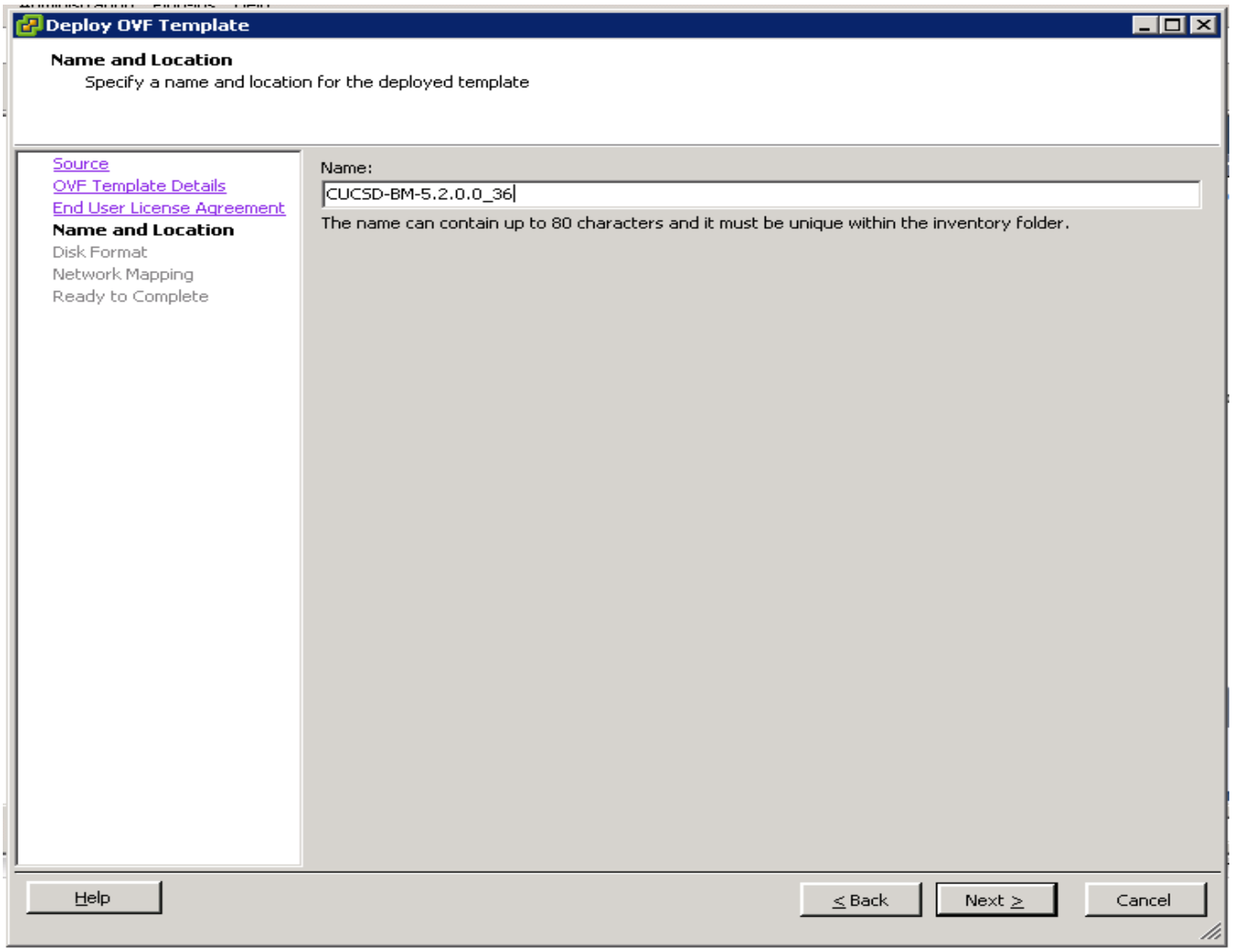
5. Choose the Cisco UCS Director Baremetal Agent 5.2.0.0 OVF template. Click **Open**.
6. Click **Next** to continue.

Figure 138 Select the Cisco UCS Director Baremetal Agent OVF file



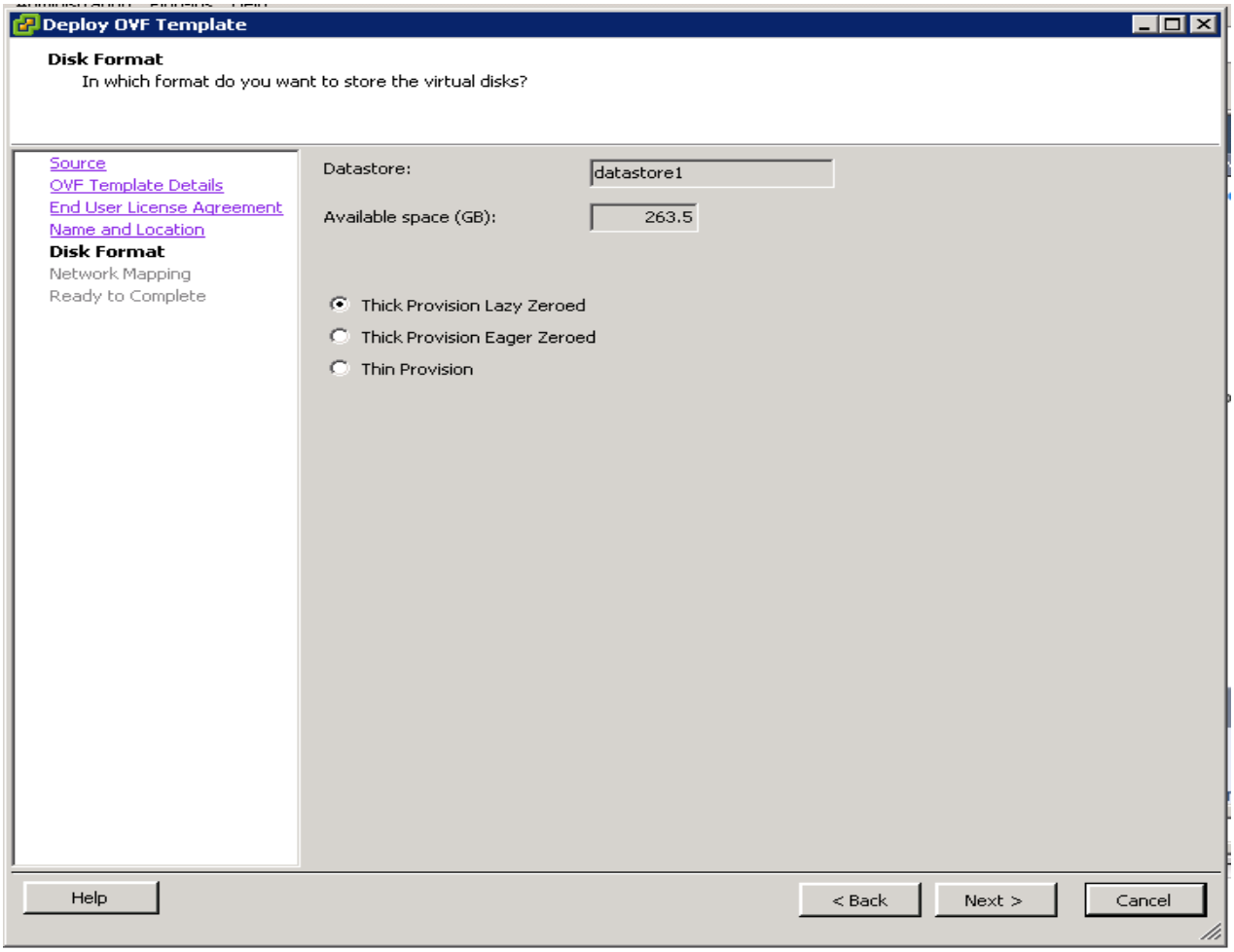
7. Review the details of the OVF template, Click **Next**.
8. Accept the End User License Agreement. Click **Next** to continue.
9. In the **Name and Location** option, Enter the name of the VM. Click **Next** to continue.

Figure 139 Enter Cisco UCS Director Baremetal Agent VM Name



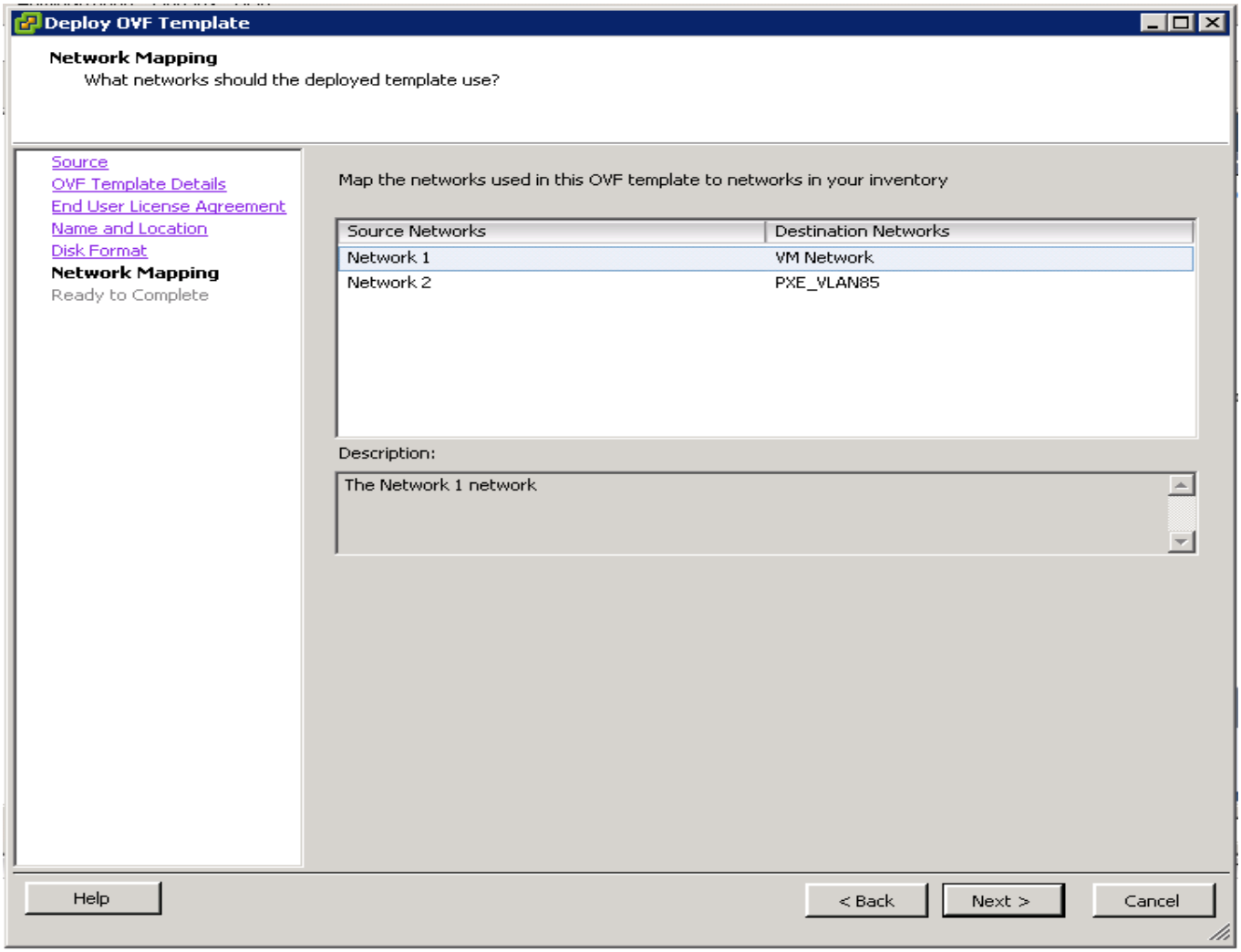
10. In the Disk Format option, click the **Thick Provision Lazy Zeroed** radio button. Click **Next** to continue.

Figure 140 Select the Disk Format for the VM



11. In the Network Mapping option,
   - Choose **VM Network** as the destination network for source Network 1.
   - Choose **PXE_VLAN85** as the destination network for source Network 2.
12. Click **Next** to continue.

Figure 141 Network Mapping for Deployed Template



13. Review the details of the VM, click the check box **Power on after deployment** and click **Finish** to proceed with the VM deployment.

Figure 142 Deploy the Cisco UCS Director Baremetal Agent VM

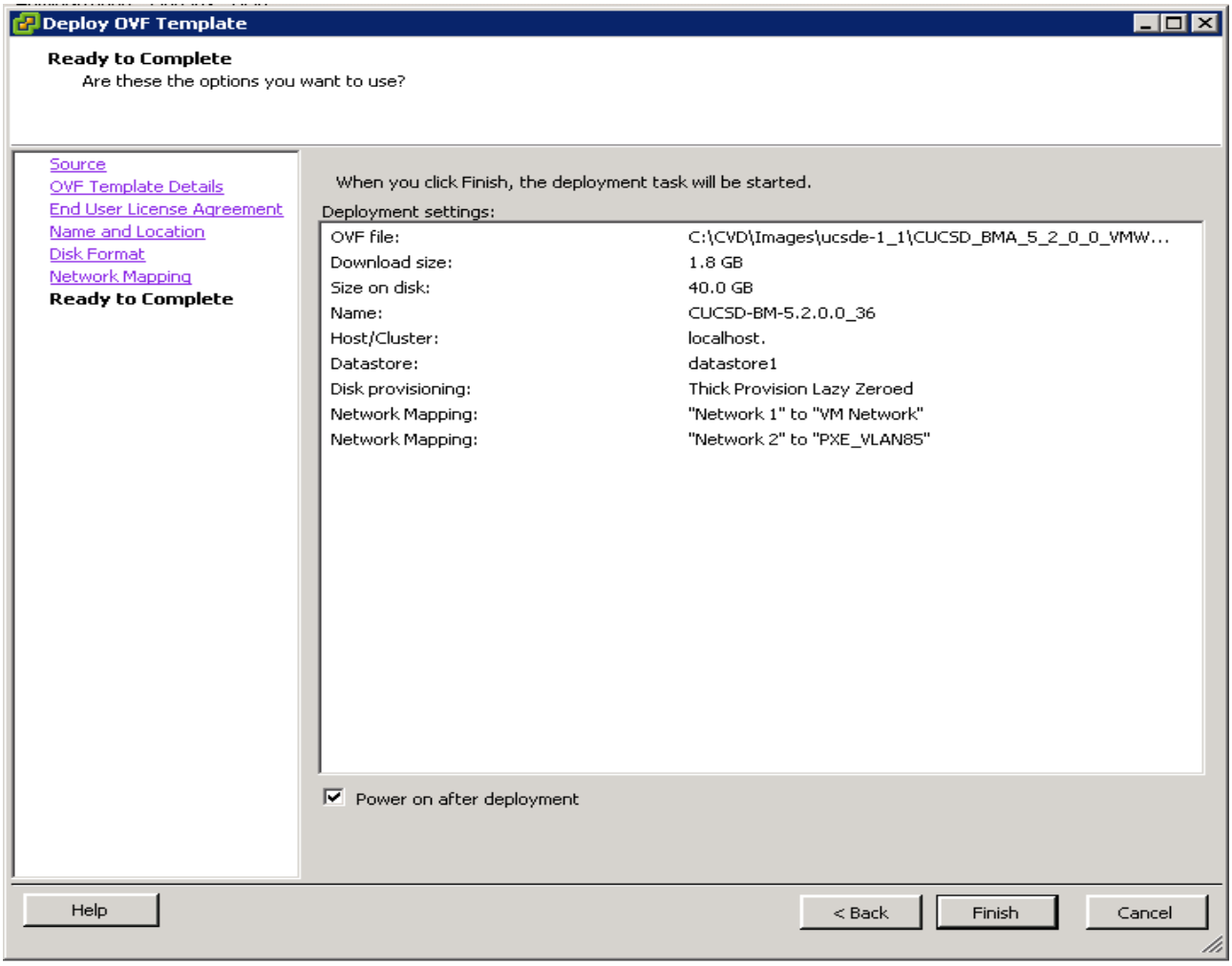
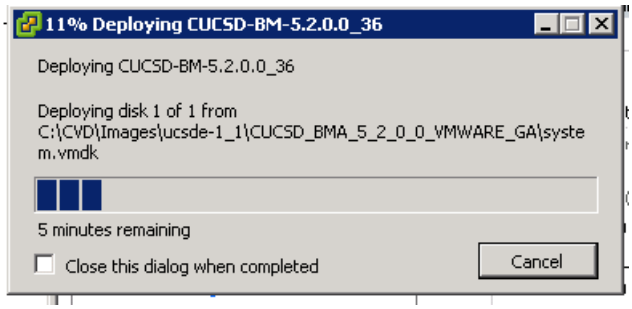


Figure 143 Cisco UCS Director Baremetal Agent VM Deployment in Progress

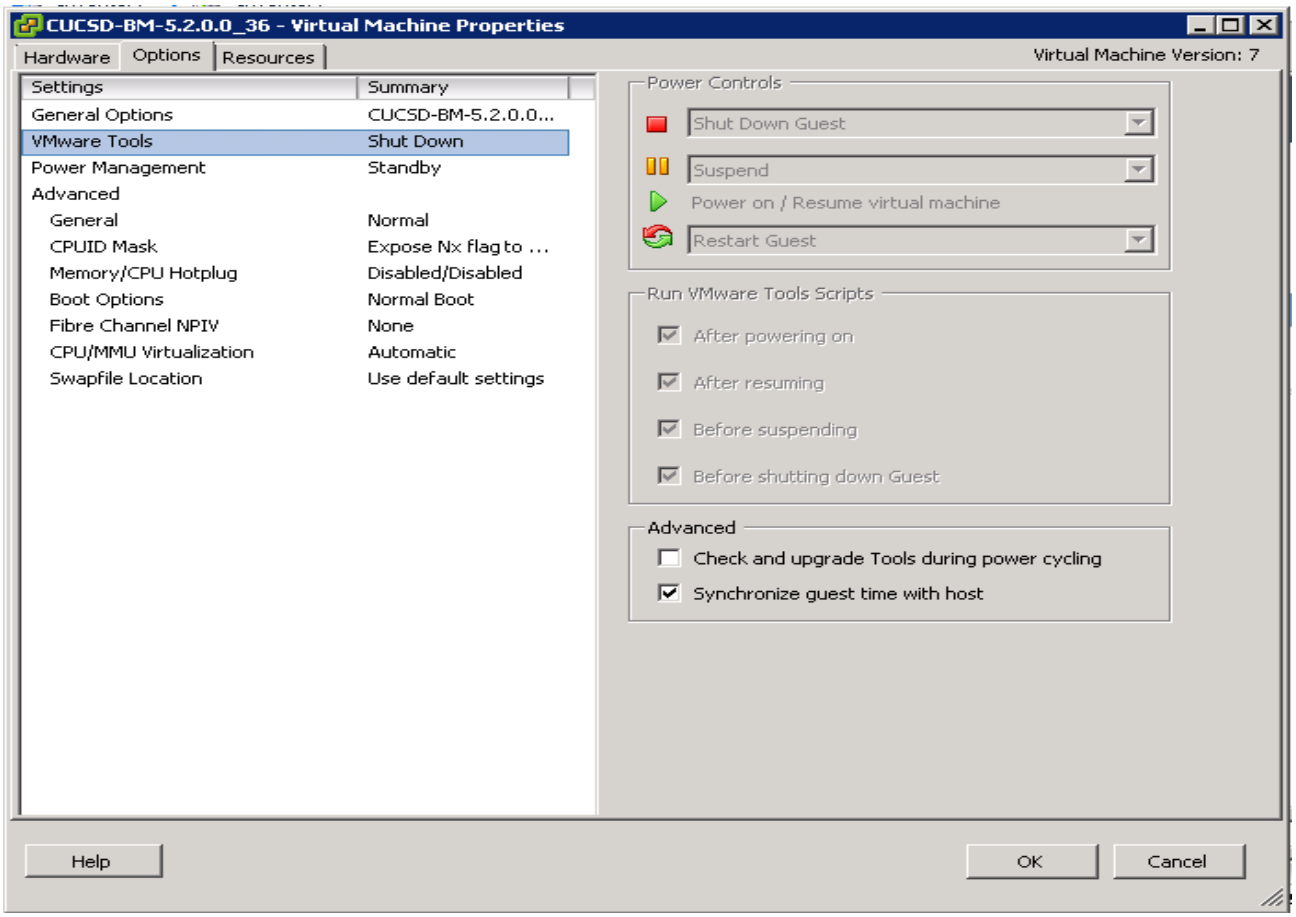


Configuring the Cisco UCS Director Baremetal Agent VM (BMA-VM)

The Cisco UCS Director Baremetal Agent VM named as CUCSD-BM-5.2.0.0_36 shall be known as BMA-VM here onwards.

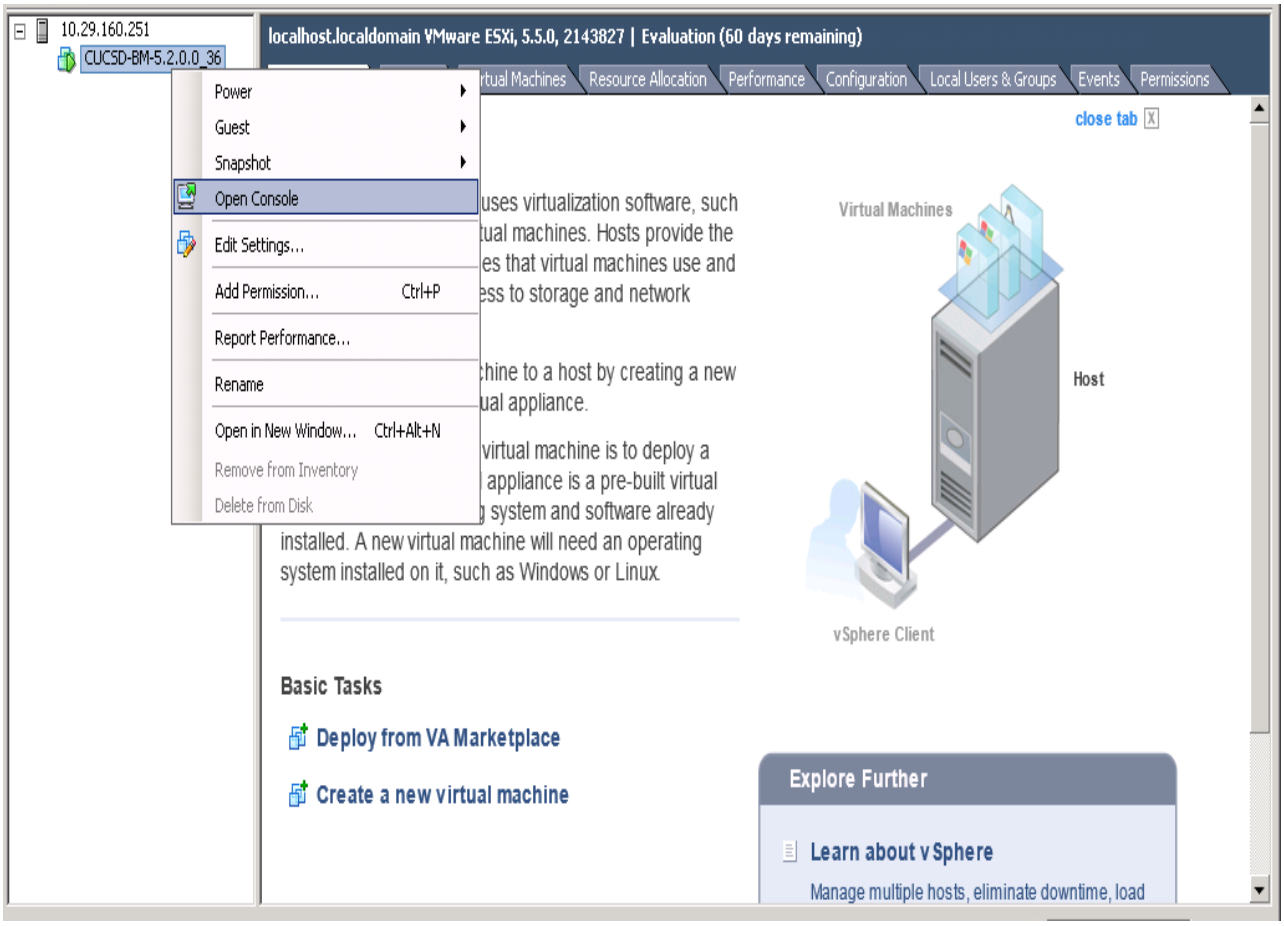
1. Right click on the BMA-VM, and select **Edit Settings**.
2. In the Virtual Machine Properties dialog box, click on the Options Tab.
3. Click on the VMWare **Tools**, Click on the **Synchronize guest time with host** option in the Advanced **section**.
4. Click on **OK** button to accept the changes.

Figure 144 Edit VM Settings to Synchronize the Guest Time with the ESXi Host



5. Right click on the BMA-VM, and select **Open Console**.

Figure 145 Access the VM Console of the BMA-VM



6. In the console accept the End User License Agreement by typing **yes** and press **ENTER**.

Figure 146 Accept the EULA

```
Do you agree with the terms of the End User License Agreement?
yes/no [no]: yes_
```

7. Login as **root** user using the default password **pxeboot**.
8. Configure the network interfaces by editing the ifcfg-eth0 and ifcfg-eth1 files located at **/etc/sysconfig/network-scripts/** directory, as follows:

Table 16 BMA-VM network configurations

Network Interface	Configuration
eth0	IP Address: 10.29.160.36, Subnet Mask: 255.255.255.0
eth1	IP Address: 192.168.85.36, Subnet Mask: 255.255.255.0

Figure 147 Editing the BMA-VM NIC eth0

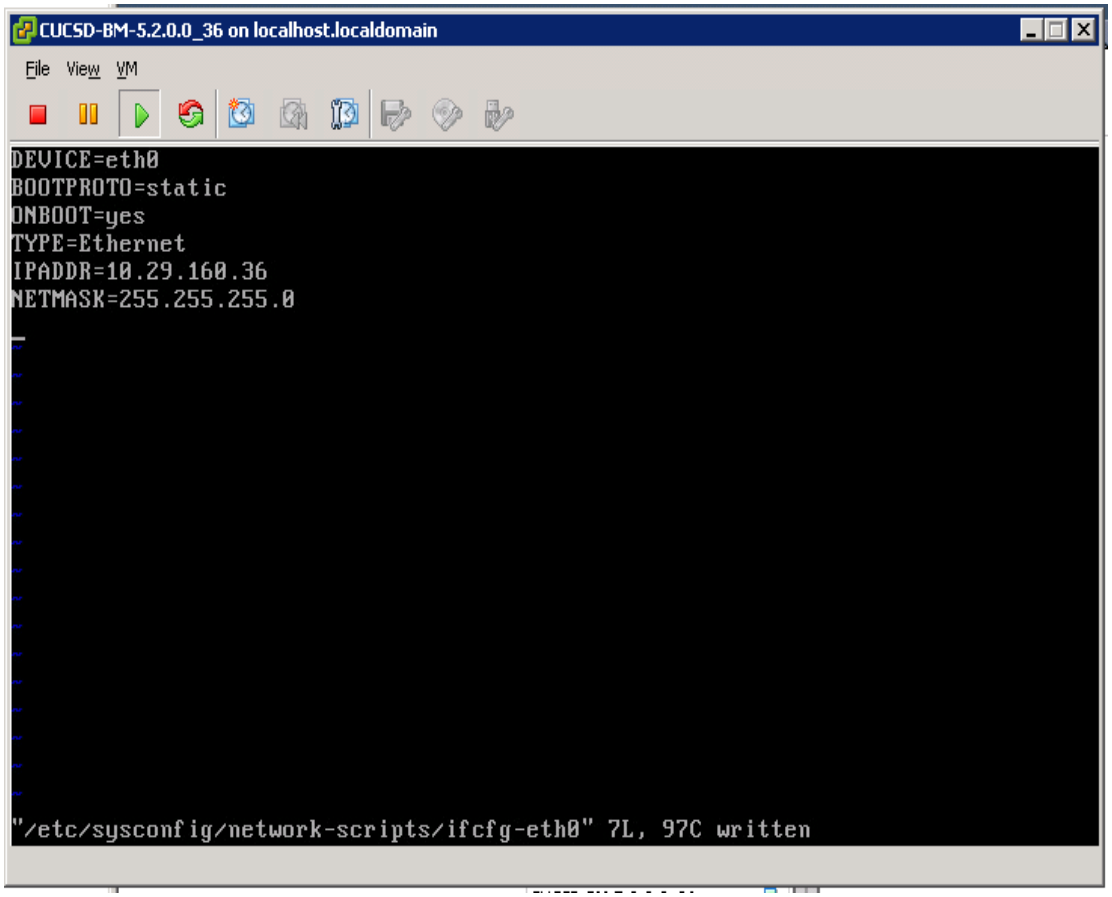
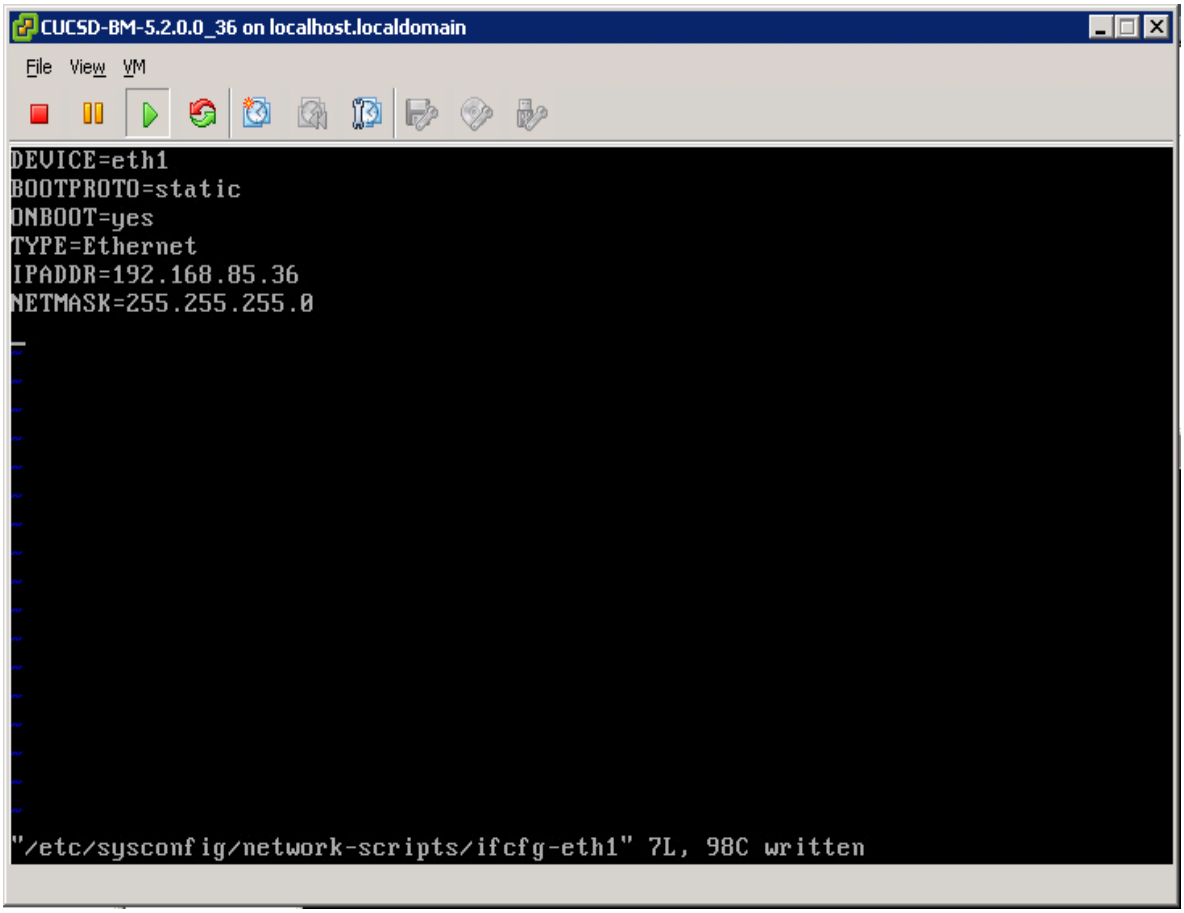


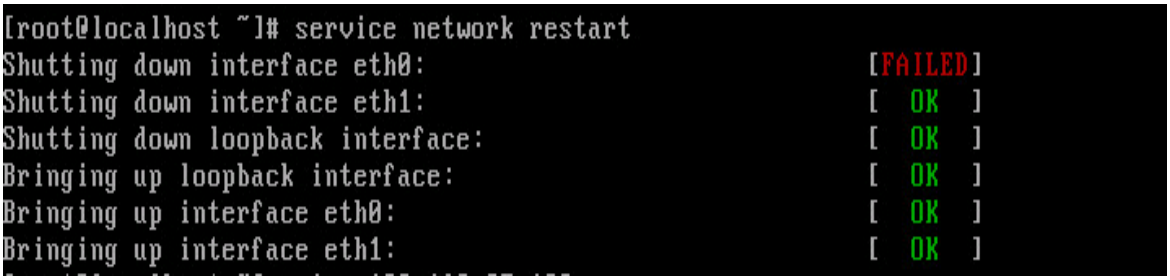
Figure 148 Editing the BMA-VM NIC eth1



9. Restart the network service by using the service command.

service network restart

Figure 149 Restart the network



Installing the Cisco UCS Director Express Big Data Upgrade Package

1. Copy over the **UCSDExpress_BMA_5.2_Big_Data_1.1_Upgrade_Package.zip** that was downloaded from cisco.com to this VM, by using a secure shell FTP session.
2. Unzip the contents in a temporary staging directory.
3. Change directory into the scripts/bin directory.
4. Change the permissions to add execute permissions to the copyfiles.sh script file and execute it.

chmod +x copyfiles.sh

Figure 150 Install the Cisco UCS Director Express Big Data Upgrade Package

```
[root@localhost stage]# ls
CentOSLive        bd_bma_version.info  feature-bigdata-intel.jar
Hortonworks-2.1   cloudera-5.0.1       mapr_common_templates
Hortonworks-2.2   cloudera-5.2.0       ntp_server_config.sh
MapR-3.1.1        cloudera-5.2.1       run.sh.template
MapR-4.0.1        cloudera-5.3.0       scripts
bd-sw-rep         common_templates     templates
[root@localhost stage]# cd scripts/bin
[root@localhost bin]# chmod +x ./copyfiles.sh
```

5. Execute the copyfiles.sh script.

./copyfiles.sh

This script copies the number of software modules such as CentOSLive image into the BMA-VM and creates a new repository directory by name **bd-sw-rep** under the **/opt/cnsaroot** directory. This new directory acts as the repository of all the Big Data specific 3rd party hadoop distribution directories.

Configuring the Big Data software repositories

Copy the Contents of RHEL6.5 ISO into the BMA-VM

1. Copy over the contents of the RHEL6.5 ISO into the directory **/opt/cnsaroot/images/RHEL6.5** on the BMA-VM.
2. Copy the contents of the directory **/opt/cnsaroot/images/RHEL6.5/isolinux** into the directory **/opt/cnsaroot/RHEL6.5**.

Figure 151 Copy the Contents of RHEL6.5 ISO into the BMA-VM

```
[root@localhost ~]# cd /opt/cnsaroot/RHEL6.5
[root@localhost RHEL6.5]# cp /opt/cnsaroot/images/RHEL6.5/isolinux/* .
[root@localhost RHEL6.5]# ls
TRANS.TBL  boot.msg   initrd.img    isolinux.cfg  splash.jpg    vmlinuz
boot.cat   grub.conf  isolinux.bin  memtest       vesamenu.c32
[root@localhost RHEL6.5]#
```

Download and Place the Common Utility files in BMA-VM

3. From a host connected to the Internet, download the Parallel-SSH and Cluster-Shell utility tools and copy them over to the **/opt/cnsaroot/bd-sw-rep** directory.

- Download Parallel SSH archive from https://pypi.python.org/packages/source/p/pssh/pssh-2.3.1.tar.gz
- Download Cluster-Shell RPM package from http://dl.fedoraproject.org/pub/epel/6/x86_64/clustershell-1.6-1.el6.noarch.rpm

Figure 152 Copy the Cluster-Shell and Passwordless-SSH Utilities

```
-rw-r--r-- 1 root root  250400 Feb 18 21:18 clustershell-1.6-1.el6.noarch.rpm
-rw-r--r-- 1 root root   23427 Feb 18 21:17 pssh-2.3.1.tar.gz
[root@localhost bd-sw-rep]# pwd
/opt/cnsaroot/bd-sw-rep
[root@localhost bd-sw-rep]#
```

4. By following the instructions on this page of the BMA-Install guide, download and copy over the Hadoop Distro RPMs into their respective directories under **/opt/cnsaroot/bd-sw-rep**. http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-director-express/bma-install-config/1-1/b_ucsd_express_bma_install_config_guide_1-1/b_ucsd_express_bma_install_config_guide_chapter_0101.html#reference_F3FE769E6A114DAD8CD5E3296556B70E
5. Upload the appropriate License files to the Hadoop distribution directories

- Place the Cloudera License in a file called ClouderaEnterpriseLicense.lic and place it under the **/opt/cnsaroot/bd-sw-rep/cloudera05.x.y**.
- Place the MapR license in a file called license.txt MapR License and place it under the directory **/opt/cnsaroot/bd-sw-rep/MapR-X.Y.Z**.

Note distribution does not require any license file.

Figure 153 Copy the RPM Packages for the Hadoop Distributions

```
[root@localhost ~]# cd /opt/cnsaroot/bd-sw-rep/
[root@localhost bd-sw-rep]# ls cloudera-5.3.0/
ClouderaEnterpriseLicense.lic  cm5.3.0-centos6.tar.gz
catalog.properties             mysql-connector-java-5.1.26.tar.gz
cdh5.3.0-centos6.tar.gz        userrpmlist.txt
[root@localhost bd-sw-rep]# ls Hortonworks-2.2/
HDP-2.2.0.0-centos6-rpm.tar.gz     catalog.properties
HDP-UTILS-1.1.0.20-centos6.tar.gz  openssl-1.0.1e-30.el6.x86_64.rpm
ambari-1.7.0-centos6.tar.gz        userrpmlist.txt
[root@localhost bd-sw-rep]# ls MapR-4.0.2
catalog.properties                          mapr-v4.0.2GA.rpm.tgz
catalog.properties.txt                      mapr-whirr-0.7.0.16780-1.noarch.rpm
ext-2.2.zip                                 pdsh-2.27-1.el6.rf.x86_64.rpm
libgenders-1.14-2.el6.rf.x86_64.rpm         soci-3.2.1-1.el6.x86_64.rpm
libgenders-devel-1.14-2.el6.rf.x86_64.rpm   soci-mysql-3.2.1-1.el6.x86_64.rpm
license.txt                                 sshpass-1.05-1.el6.x86_64.rpm
mapr-drill-0.7.0.29434-1.noarch.rpm         userrpmlist.txt
mapr-ecosystem-20150205.rpm.tgz
[root@localhost bd-sw-rep]#
```

Setup a UCSD Patch Directory in the BMA-VM

Cisco UCS Director Express for Big Data VM which will be installed in the next section, requires the patches to be kept in a web server. The BMA-VM comes pre-configured with a web-server used during PXE booting process. This section walks through the steps to create a directory to hold these patches in the BMA-VM.

1. In BMA-VM, create a directory by name patches under /var/www/html.

mkdir /var/www/html/patches

2. Copy over the Cisco UCS Director Express for Big Data 1.1 specific patch files (See Table 3) to this patch directory.

Figure 154 Setup a UCSD Patch Directory in the HTTP Root Path

```
[root@localhost ~]# ls -l /var/www/html/patches
total 1172256
-rw-r--r-- 1 root root    2139421 Feb 18 04:52 UCSDExpress_Big_Data_1.1_Upgrade_Package.zip
-rw-r--r-- 1 root root 1197064934 Feb  3 13:16 cucsd_patch_5_2_0_1.zip
```

3. Start the HTTPD server in the BMA-VM.

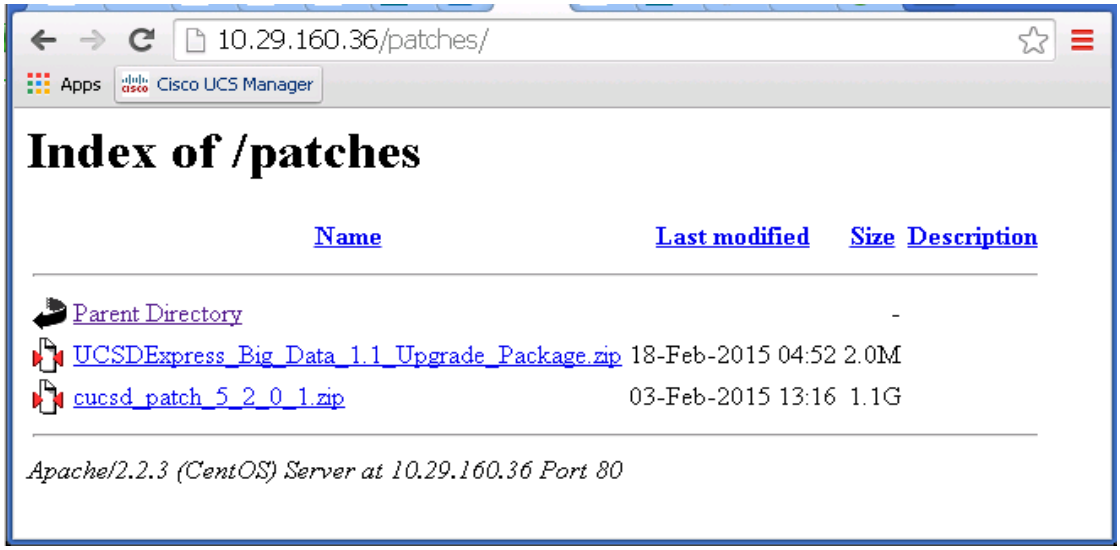
service httpd start

Figure 155 Start the HTTPD

```
[root@localhost bd-sw-rep]# service httpd start
Starting httpd:                                            [  OK  ]
```

4. Verify if these files are accessible by visiting the URL **http://<BMA-VM's >IP address/patches/**.

Figure 156 Verify the Accessibility of the Cisco UCS Director Express Patches

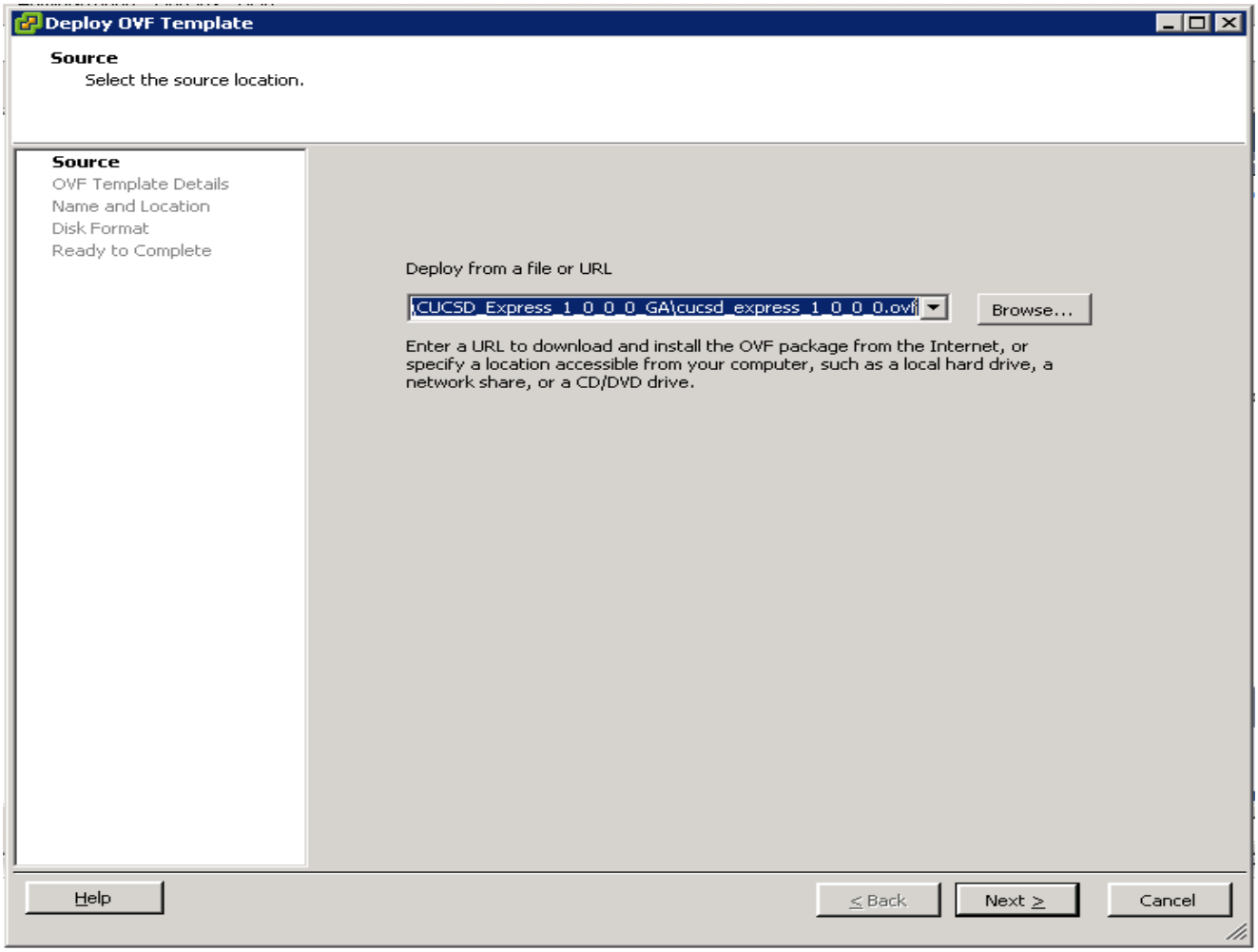


BMA-VM configurations are complete.

Deploying the Cisco UCS Director Express OVF

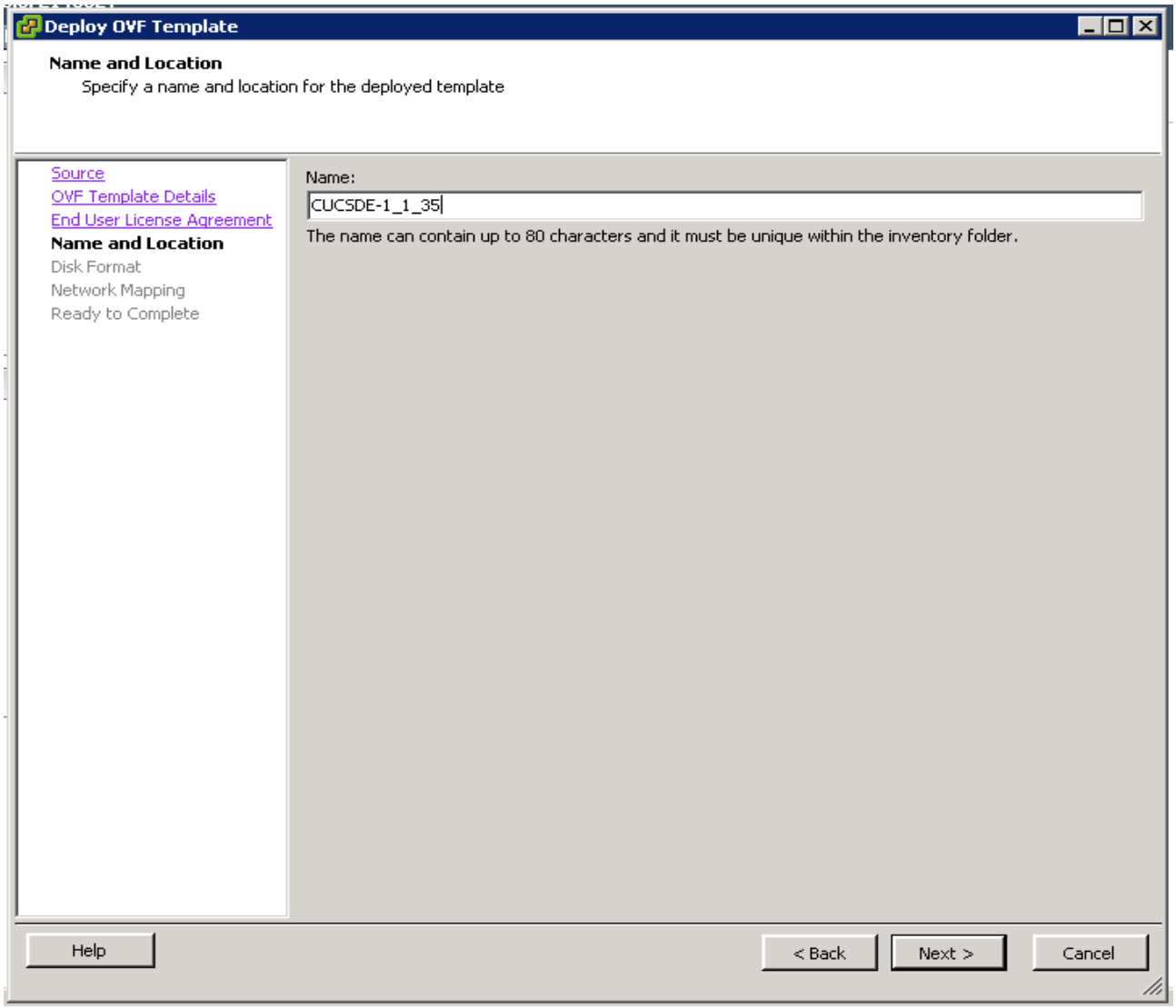
1. Launch the VMWare vSphere client software
2. Enter the chosen IP address, the username as root, and the chosen password.
3. Click **Login** to continue.
4. From the **File** menu, Select **Deploy OVF Template**.
5. Choose the Cisco UCS Director Express for Big Data 1.0 OVF template. Click **Open**.

Figure 157 Deploy the Cisco UCSD Express 1.0 OVF



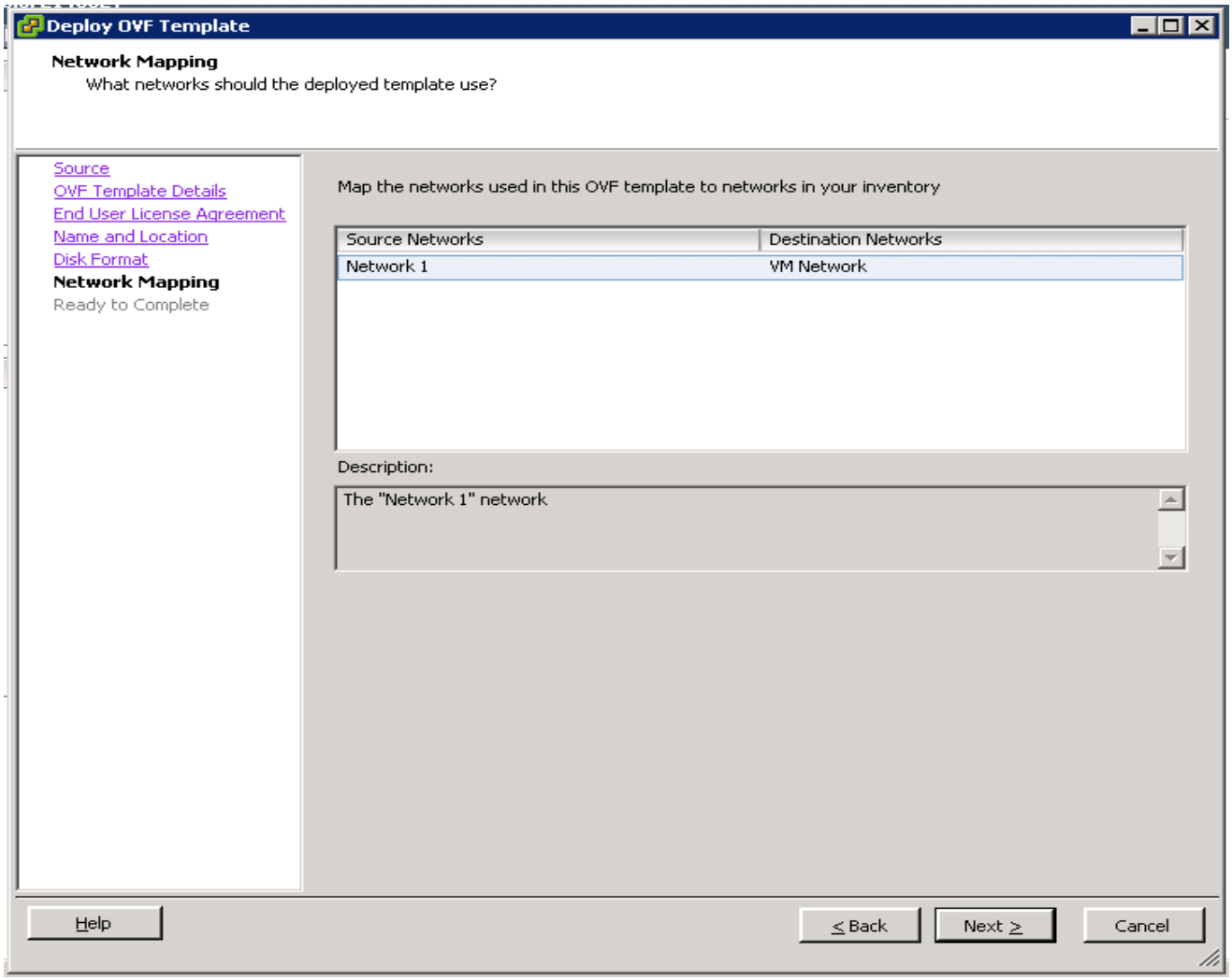
6. Review the details of the OVF, and Click **Next** to continue.
7. Accept the EULA, Click **Next** to continue.
8. Name the VM, Click **Next** to continue.

Figure 158 Name the Cisco UCS Director Express VM



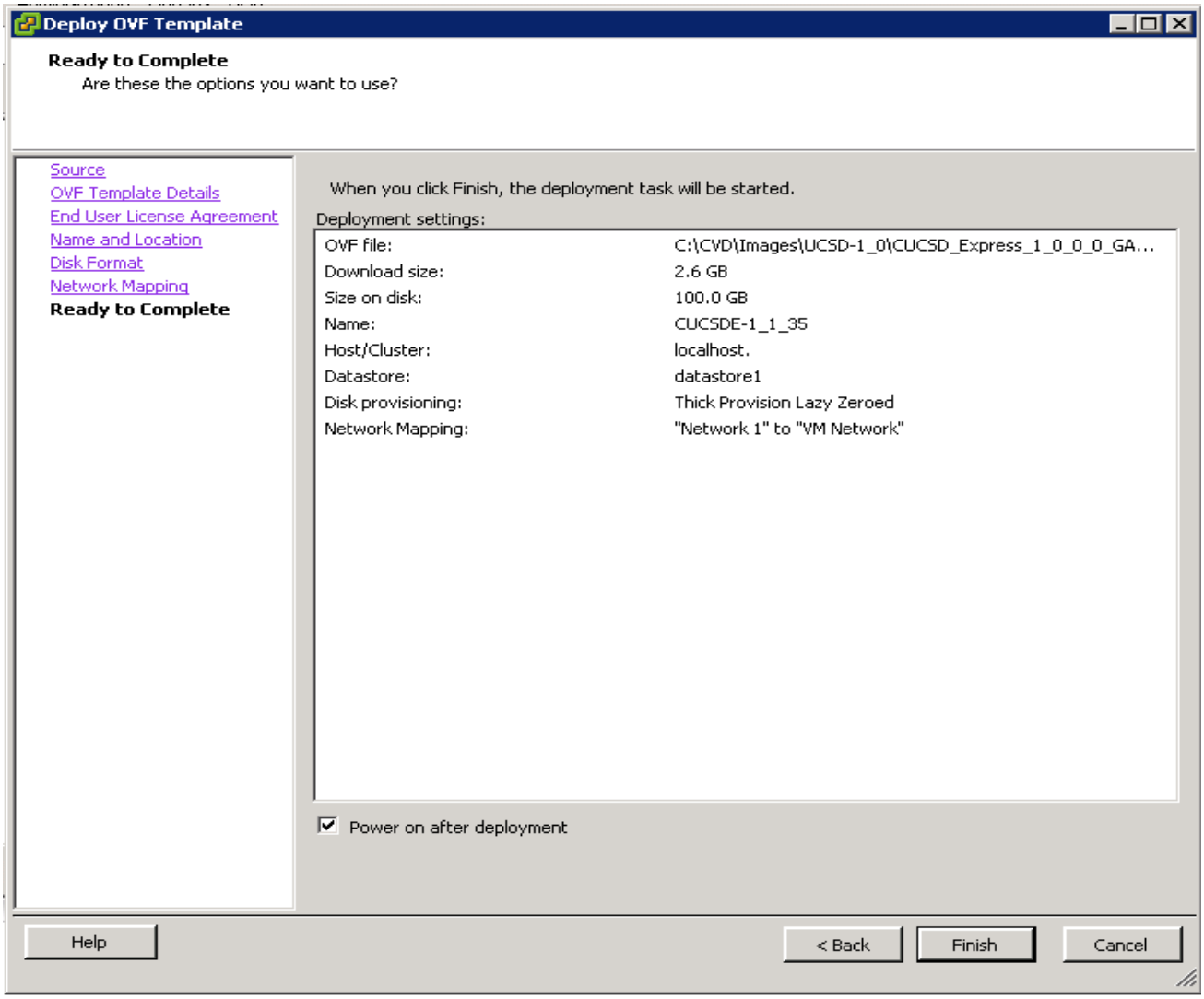
9. Choose the destination network **VM Network** for the source network **Network 1**. Click **Next** to continue.

Figure 159 Cisco UCS Director Express VM Network Configuration



10. In the Disk Format option, click the **Thick Provision Lazy Zeroed** radio button. Click **Next** to continue.
11. Review the details of the VM, Check the checkbox **Power On after deployment**.
12. Click **Finish** to proceed with deployment.

Figure 160 Deploy the Cisco UCS Director Express VM

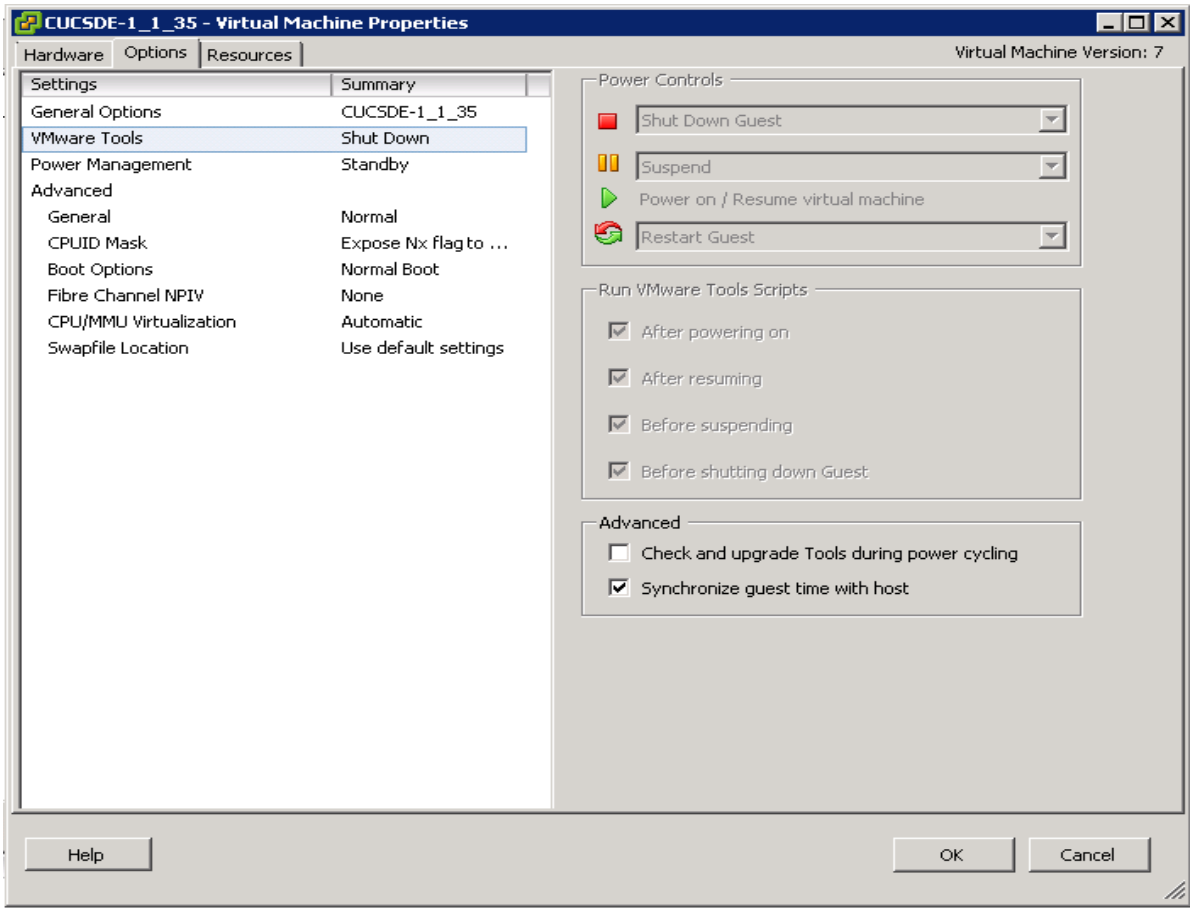


Configuring the Cisco UCS Director Express VM (UCSD-VM)

The Cisco UCS Director Express VM named as CUCSDE-1_1_35 shall be known as UCSD-VM here onwards.

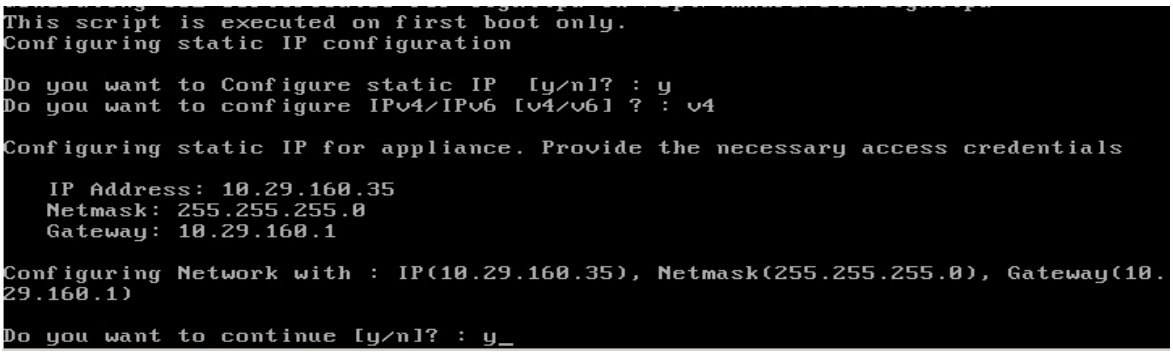
1. Right click on the UCSD-VM, and select **Edit** Settings.
2. In the Virtual Machine Properties dialog box, click on the **Options** tab.
3. Click on the **VMware** Tools, Click on the **Synchronize guest time with host** option in the **Advanced** section.
4. Click on **OK** button to accept the changes.

Figure 161 Edit VM Settings to Synchronize the Guest Time with the ESXi Host



5. Right-click on the UCSD-VM and select **Open Console**.
6. Accept the End User License Agreement by typing **yes** and press the **ENTER**.
7. In the prompt to configure the static IP for the network interface, enter the IP address, Netmask and Gateway information.
8. Enter **y** to continue with the network configuration.

Figure 162 Assigning the Static IP Address to the UCSD-VM eth0



9. Configure the UCSD Express as the personality by entering the number 2.
10. **At the prompt Switching personality to UCSD Express, Are you sure to continue [y/n]?** Type **y** and hit **ENTER**.

Figure 163 Choose the UCSD Express Personality

```
Configuring Personality
          Select the personality

              1)  Default - UCSD
              2)  UCSD Express
              3)  Cirrus

Personality : [1/2/3]? 2
Switching personality to UCSD Express. Are you sure to continue [y/n]? y_
```

11. The UCSD-VM goes through a personality change configuration as shown below.

Figure 164 UCSD-VM First-Boot Initializations

```
completed db privileges
copying my.cnf.template
Completed copying my.cnf.template
Forcing it to a login prompt
Completed forcing it to a login prompt
starting database
started database
sleep 1m
JRE Copy Start
JRE Copy End
Installing native files
Unzip of native files completed
Installing native (/usr/lib) files
Installed native (/usr/lib) files
Installing native (/usr/include) files
Installed native (/usr/include) files
Installing native (/usr/bin) files
Installed native (/usr/bin) files
Installing native (/etc) files
Installed native (/etc) files
Installing CUIC-vix files
Installed CUIC-vix files
JRE_HOME is
Wed Feb 18 09:31:47 UTC 2015 : Initializing CUIC Database schema
```

Note This step takes about 10-15 minutes to complete.

Applying the Upgrade Patches

1. Open a SSH/Putty session to the UCSD-VM.
2. Login as the user **shelladmin** with password **changeme**.

Figure 165 Logging onto the UCSD-VM Shell Administration Tool

```
login as: shelladmin
shelladmin@10.29.160.35's password:
```

3. In the Shell Admin Menu, enter 3 to stop the services.
4. At the prompt, **Do you want to stop services [y/n]?** Type **y** to confirm and hit **ENTER** to continue.

Figure 166 Issuing the Command to Stop all the Services Via Shell Administration Tool.

```
                    Standalone Node

        Select a number from the menu below

            1)  Change ShellAdmin Password
            2)  Display Services Status
            3)  Stop Services
            4)  Start Services
            5)  Stop Database
            6)  Start Database
            7)  Backup Database
            8)  Restore Database
            9)  Time Sync
            10) Ping Hostname/IP Address
            11) Show Version
            12) Import CA Cert (JKS) File
            13) Import CA Cert(PEM) File for VNC
            14) Configure Network Interface
            15) Display Network Details
            16) Enable Database for Cisco UCS Director Baremetal Agent
            17) Add Cisco UCS Director Baremetal Agent Hostname/IP
            18) Tail Inframgr Logs
            19) Apply Patch
            20) Shutdown Appliance
            21) Reboot Appliance
            22) Manage Root Access
            23) Login as Root
            24) Configure Multi Node Setup (Advanced Deployment)
            25) Clean-up Patch Files
            26) Collect logs from a Node
            27) Collect Diagnostics
            28) Change Personality
            29) Quit

            SELECT> 3

      Do you want to stop services [y/n]? : y
```

5. In the Shell Admin menu, type 2 to view the status of the services. They all should be **NOT-RUNNING** as shown below.

Figure 167 Verifying the Status of the UCSD-VM Services

```
              SELECT> 2
 Service                    Status        PID
 ----------             ----------      -----
 broker                 NOT-RUNNING         -
 controller             NOT-RUNNING         -
 eventmgr               NOT-RUNNING         -
 client                 NOT-RUNNING         -
 idaccessmgr            NOT-RUNNING         -
 inframgr               NOT-RUNNING         -
 TOMCAT                 NOT-RUNNING         -
 websock                NOT-RUNNING         -

 3467 ?        00:00:00 mysqld_safe
 3888 ?        00:03:05 mysqld
Press return to continue ...
```

6. In the Shell Admin menu, type **19** and **ENTER** to start the patching process.
7. Type **n** to the prompt **Do you want to take database backup before applying patch[y/n]?**.
8. At the prompt, Patch URL: **enter http://<BMA_IP>/patches/cucsd_patch_5_2_0_1.zip**
9. Hit **ENTER** to continue.

Figure 168 Cisco UCS Director 5.2.0.1 Patch Application Process

```
          Select a number from the menu below

               1)  Change ShellAdmin Password
               2)  Display Services Status
               3)  Stop Services
               4)  Start Services
               5)  Stop Database
               6)  Start Database
               7)  Backup Database
               8)  Restore Database
               9)  Time Sync
               10) Ping Hostname/IP Address
               11) Show Version
               12) Import CA Cert (JKS) File
               13) Import CA Cert(PEM) File for VNC
               14) Configure Network Interface
               15) Display Network Details
               16) Enable Database for Cisco UCS Director Baremetal Agent
               17) Add Cisco UCS Director Baremetal Agent Hostname/IP
               18) Tail Inframgr Logs
               19) Apply Patch
               20) Shutdown Appliance
               21) Reboot Appliance
               22) Manage Root Access
               23) Login as Root
               24) Configure Multi Node Setup (Advanced Deployment)
               25) Clean-up Patch Files
               26) Collect logs from a Node
               27) Collect Diagnostics
               28) Change Personality
               29) Quit

               SELECT> 19
Applying Patch...
Do you want to take database backup before applying patch[y/n]? n
User selected option not to take backup, proceeding with applying patch
   Applying Patch:
   Patch URL :http://10.29.160.36/patches/cucsd_patch_5_2_0_1.zip

Applying the Patch http://10.29.160.36/patches/cucsd_patch_5_2_0_1.zip [y/n]? y
```

This 5.2.0.1 patch that is being applied to the UCSD-VM's, upgrades all the core application software to the latest Cisco UCS Director's code base. After this step completes, the Big Data Upgrade package for release 1.1 needs to be applied.

10. In the Shell Admin menu, type **19** and **ENTER** to start the patching process.
11. Type n to the prompt **Do you want to take database backup before applying patch[y/n]?**.
12. At the prompt, Patch URL:, enter **http://<BMA_IP>/patches/ UCSDExpress_Big_Data_1.1_Upgrade_Package.zip**
13. Hit **ENTER** to continue.

Figure 169 Cisco UCS Director Express for Big Data 1.1 Upgrade Package Installation Process

```
                1)  Change ShellAdmin Password
                2)  Display Services Status
                3)  Stop Services
                4)  Start Services
                5)  Stop Database
                6)  Start Database
                7)  Backup Database
                8)  Restore Database
                9)  Time Sync
                10) Ping Hostname/IP Address
                11) Show Version
                12) Import CA Cert (JKS) File
                13) Import CA Cert(PEM) File for VNC
                14) Configure Network Interface
                15) Display Network Details
                16) Enable Database for Cisco UCS Director Baremetal Agent
                17) Add Cisco UCS Director Baremetal Agent Hostname/IP
                18) Tail Inframgr Logs
                19) Apply Patch
                20) Shutdown Appliance
                21) Reboot Appliance
                22) Manage Root Access
                23) Login as Root
                24) Configure Multi Node Setup (Advanced Deployment)
                25) Clean-up Patch Files
                26) Collect logs from a Node
                27) Collect Diagnostics
                28) Change Personality
                29) Quit

                SELECT> 19
Applying Patch...
Do you want to take database backup before applying patch[y/n]? n
User selected option not to take backup, proceeding with applying patch
   Applying Patch:
   Patch URL :http://10.29.160.36/patches/UCSDExpress_Big_Data_1.1_Upgrade_Package.z
ip

Applying the Patch http://10.29.160.36/patches/UCSDExpress_Big_Data_1.1_Upgrade_Pack
age.zip [y/n]? y
```

Figure 170 Cisco UCS Director Express for Big Data 1.1 Upgrade Package Application Complete

```
****************************
Wed Jan 21 22:10:45 UTC 2015 : Copying ui.properties file
****************************
Directory doesn't exit, continuing with installation process
*******************
Wed Jan 21 22:10:45 UTC 2015 : Copying SSL File
*******************
*****************************************
Wed Jan 21 22:10:45 UTC 2015 : Copying VMWare Files & scalability folder
*****************************************
Scalability folder exists, taking backup /opt/scalability-01-21-2015-22-10-45
Diagnostics folder exists, taking backup /opt/diagnostics-01-21-2015-22-10-45
*********************************************************
Wed Jan 21 22:10:45 UTC 2015 : Copying localization related files
*********************************************************
Japanese Directory exits.
TrueType folder is present
*********************************************************
Wed Jan 21 22:10:45 UTC 2015 : Copying sysmgr jar to T1 library locations if exist
*********************************************************
*********************************************************
Wed Jan 21 22:10:45 UTC 2015 : Personality specific changes for upgrade
*********************************************************
Personality details --> Product Name : UCSD Express for Big Data , Product Version :
.0.0.0
Restored account-type-exclusion-list.properties for UCSD Express for Big Data
Restored DefaultRoleMenuMappings.properties for UCSD Express for Big Data
Restored RegularSet_menu.xml for UCSD Express for Big Data
Restored AdminSet_menu.xml for UCSD Express for Big Data
Restored feature-exclusion-list.properties for UCSD Express for Big Data
Restored reports.xml for UCSD Express for Big Data
Restored about.json for UCSD Express for Big Data
Restored signed-sku-mapping.xml for UCSD Express for Big Data
***********************************************
Restart services and database for the changes to take effect
***********************************************

 INFO (FileUtil.java:958) *********
 INFO (FileUtil.java:963)
 INFO (FileUtil.java:967) 150121 22:10:45 [FileUtil] RunCommandThread: Completed thre
d:      Thread[Thread-1,5,main]

Completed installing package 0
*******************************
Press return to continue ...
```

14. After the successful application of the patch, type **4** and **ENTER** to start the services.

Note It takes about a few minutes for all the services to get started.

15. Type 2 to check on the services status. All the services should now be in **RUNNING** state.

Figure 171 Verify the Status of the Services in the UCSD-VM

```
SELECT> 2
Service                    Status        PID
----------              ----------      -----
broker                    RUNNING       7756
controller                RUNNING       7888
eventmgr                  RUNNING       7966
client                    RUNNING       8025
idaccessmgr               RUNNING       8113
inframgr                  RUNNING       8172
TOMCAT                    RUNNING       8240
websock                   RUNNING       8320

3467 ?        00:00:00 mysqld_safe
3888 ?        00:05:52 mysqld
Press return to continue ...
```

Note Even after all the services are in a RUNNING state, it would take an additional 3 to 5 minutes for the UCSD-VM client services to become available.

Configuring the Cisco UCS Director Express for Big Data (UCSD Express)

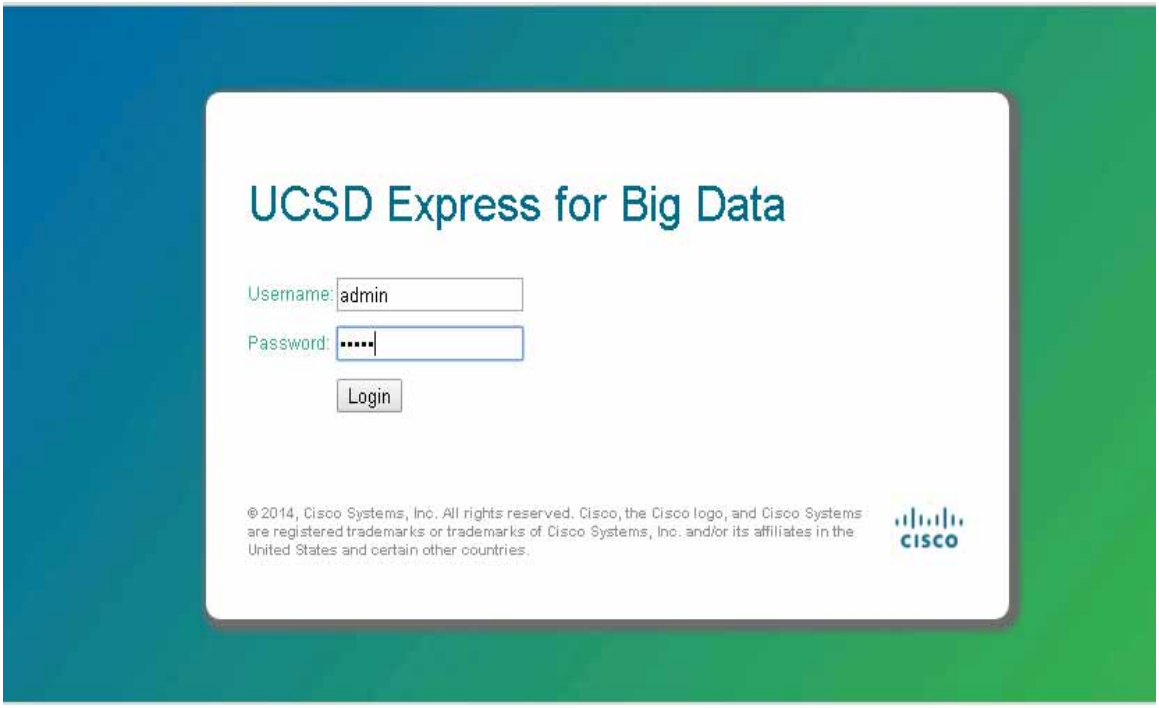
The Cisco UCS Director Express for Big Data, henceforth known as UCSD-Express, needs to be configured with the IP address to the UCS domain (i.e. UCS Manager's) physical account. This allows the UCSD-Express to query the UCS Manager and perform inventory collection.

The UCSD-Express will also need to be configured with the BMA's physical account and configure it's services such as DHCP.

Add the licenses to UCSD-Express

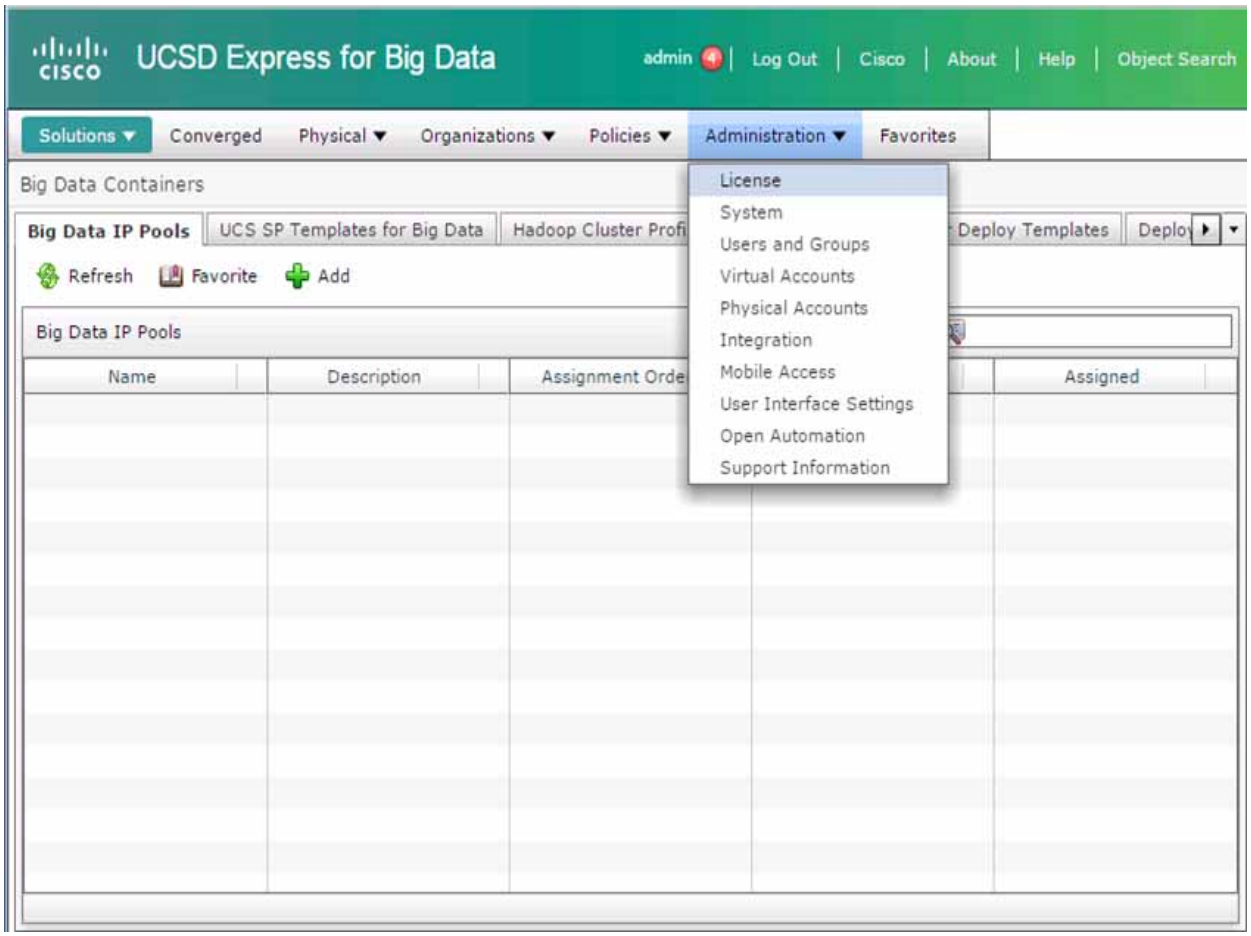
1. Using a web browser, visit the URL http://<UCSD-VM's IP>/.
2. Login as user **admin** with the default password **admin**.

Figure 172 Logging onto the Cisco UCS Director Express for Big Data



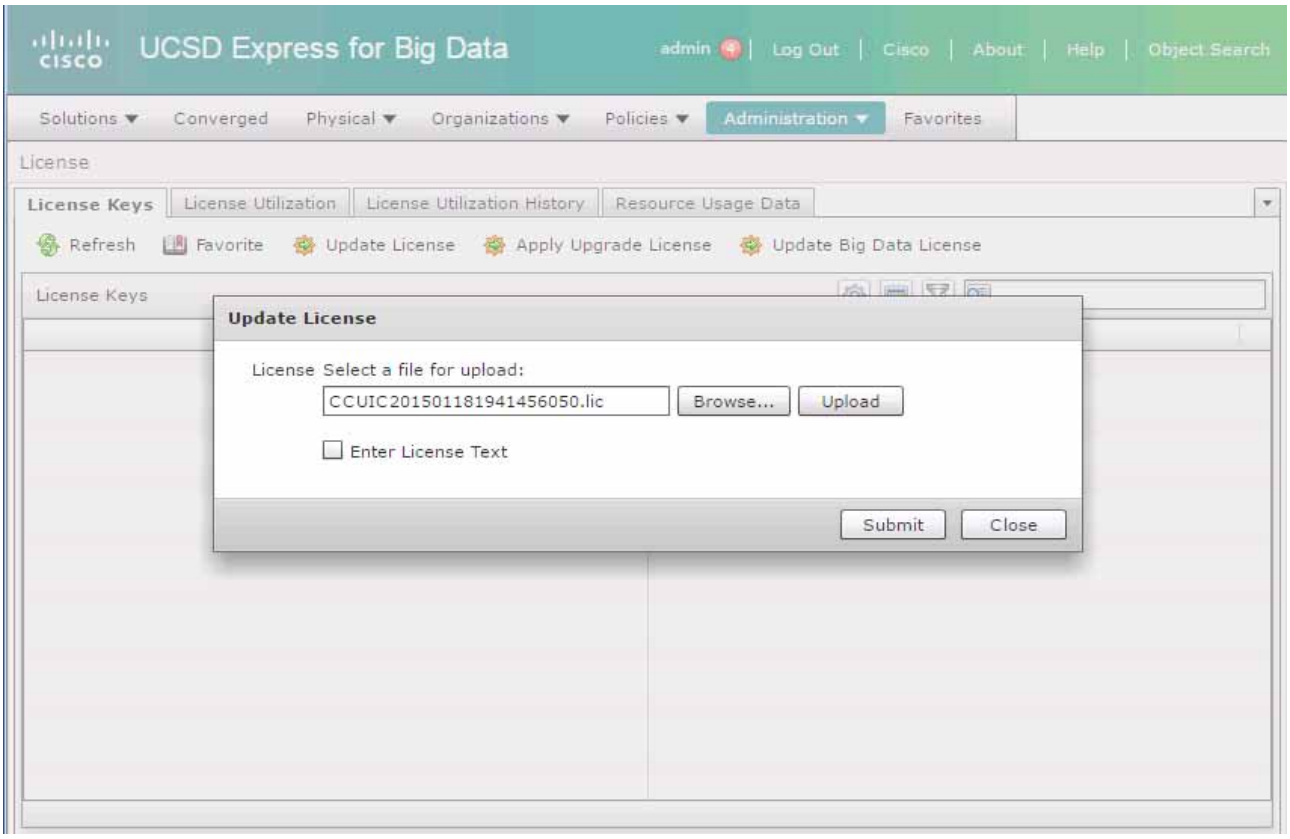
3. Navigate to **Administration > License screen.**

Figure 173 Accessing the License Administration Page



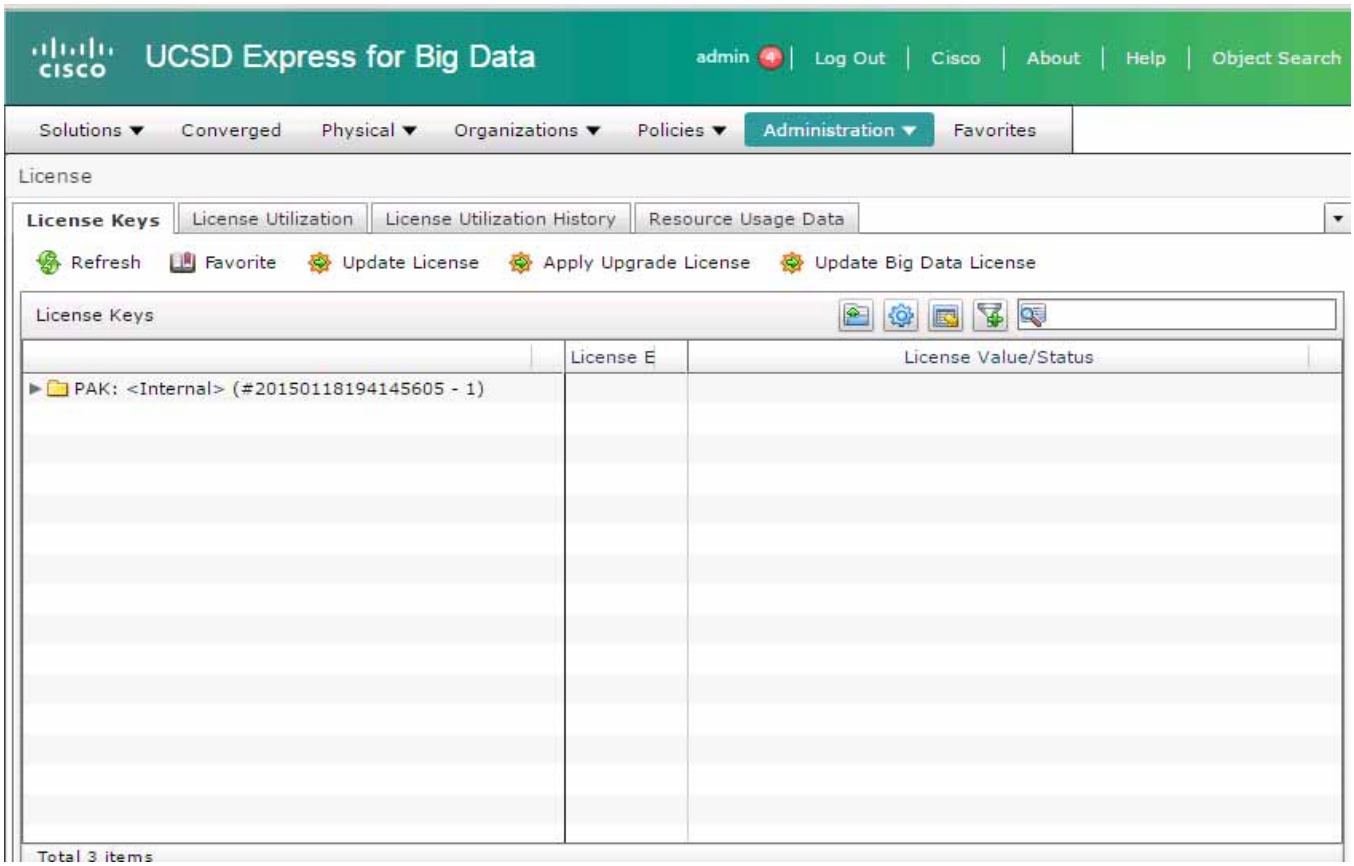
4. Click on **License Keys** tab.
5. Click on **Update License**.
6. In the **Update License** dialog box, click **Browse** to select the license file.
7. Click **Upload**.
8. After the license file gets uploaded, Click **Submit** to apply the license.

Figure 174 Applying the Base Cisco UCS Director License.



9. The license keys are displayed as shown below.

Figure 175 Cisco UCS Director Base Licenses got Applied Successfully



10. Click on **Update Big Data License**.
11. In the **Update Big Data Subscription** dialog box, click **Browse** to select the Big Data specific license file.
12. Click **Upload**.
13. After the license file gets uploaded, Click **Submit**.

Figure 176 Applying the Cisco UCS Director Express Big Data Subscription License

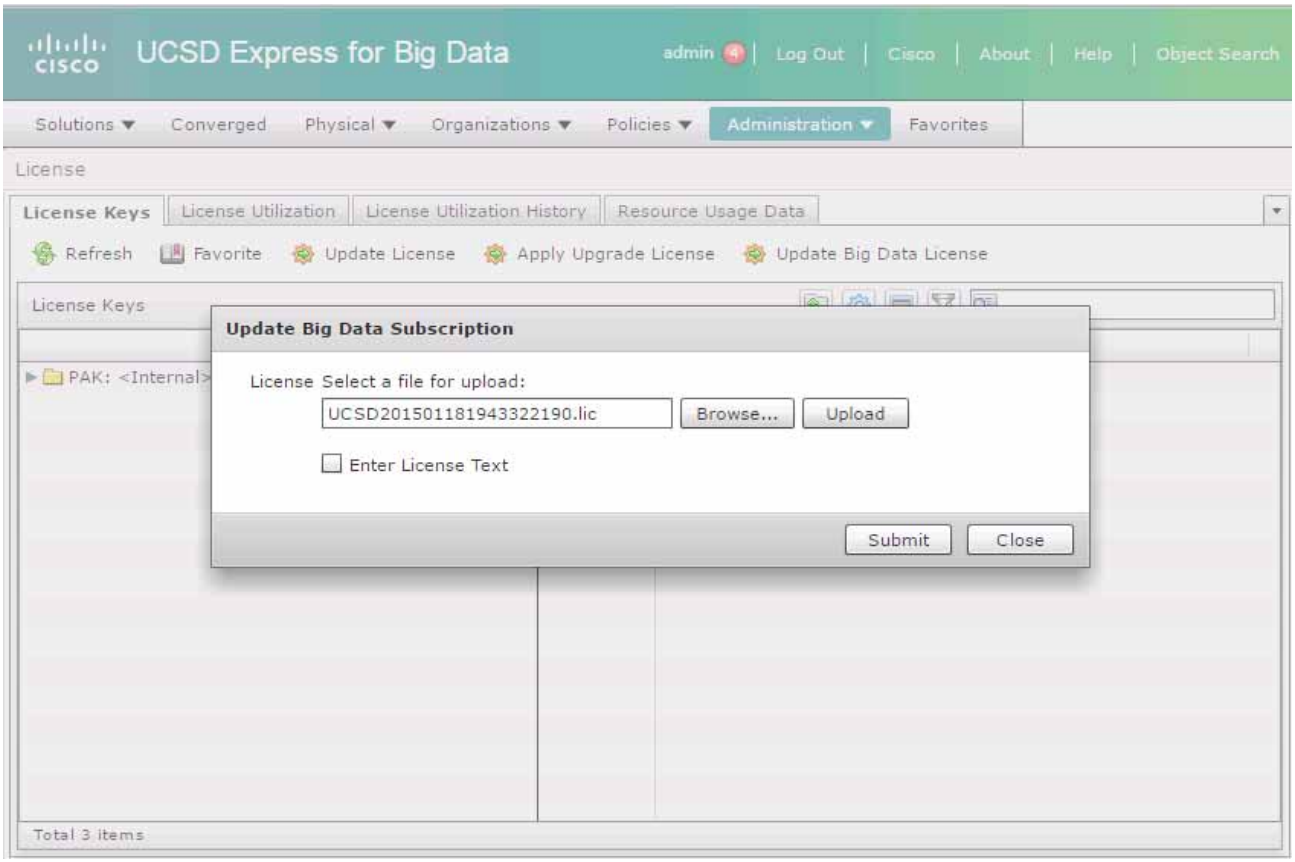
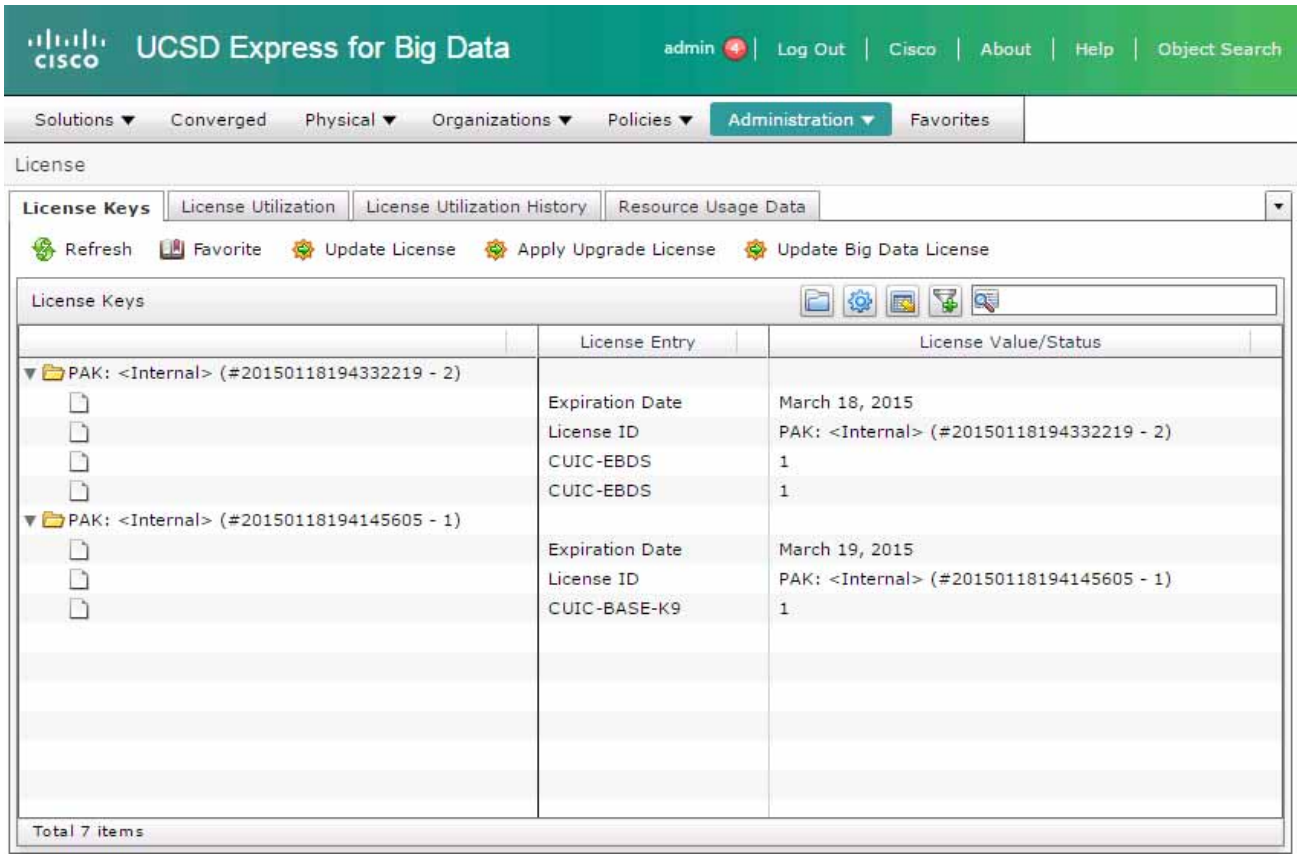


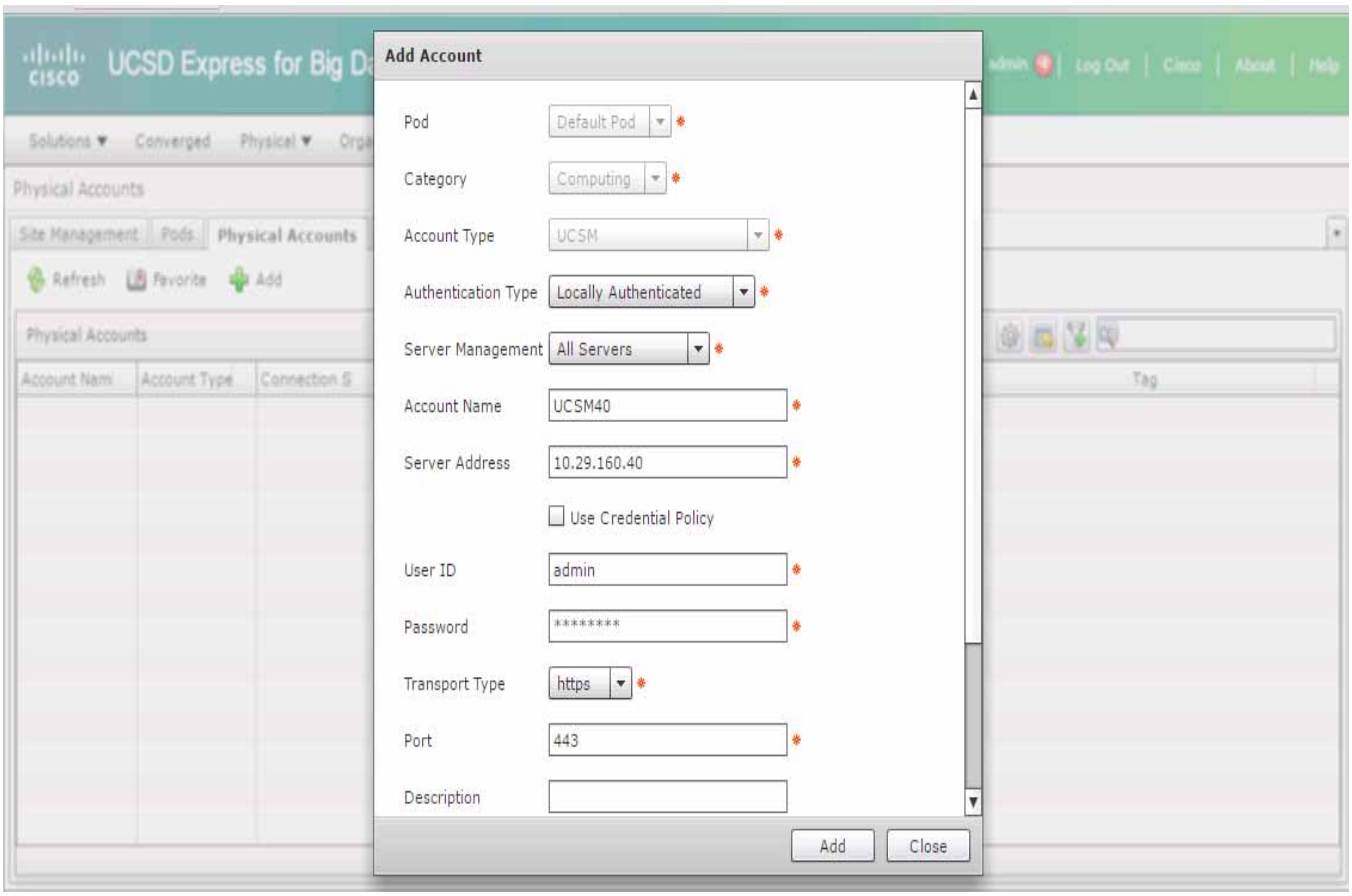
Figure 177 Completion of the License Application.



Add the UCS Manager physical account to the UCSD-Express

1. In the UCSD-Express web console, navigate to **Administration >Physical Accounts**.
2. Click **+ ADD** button
   a. Input the UCS Manager Account details as follows.
   b. In the Account Name field, enter a name to this UCS Manager account.
   c. In the Server Address field, enter the IP address of the UCS Manager.
   d. In the User ID field, enter admin.
   e. In the Password field, enter the password to the UCS Manager's admin user.
   f. In the Transport Type field, choose https.
3. Click **Add**.

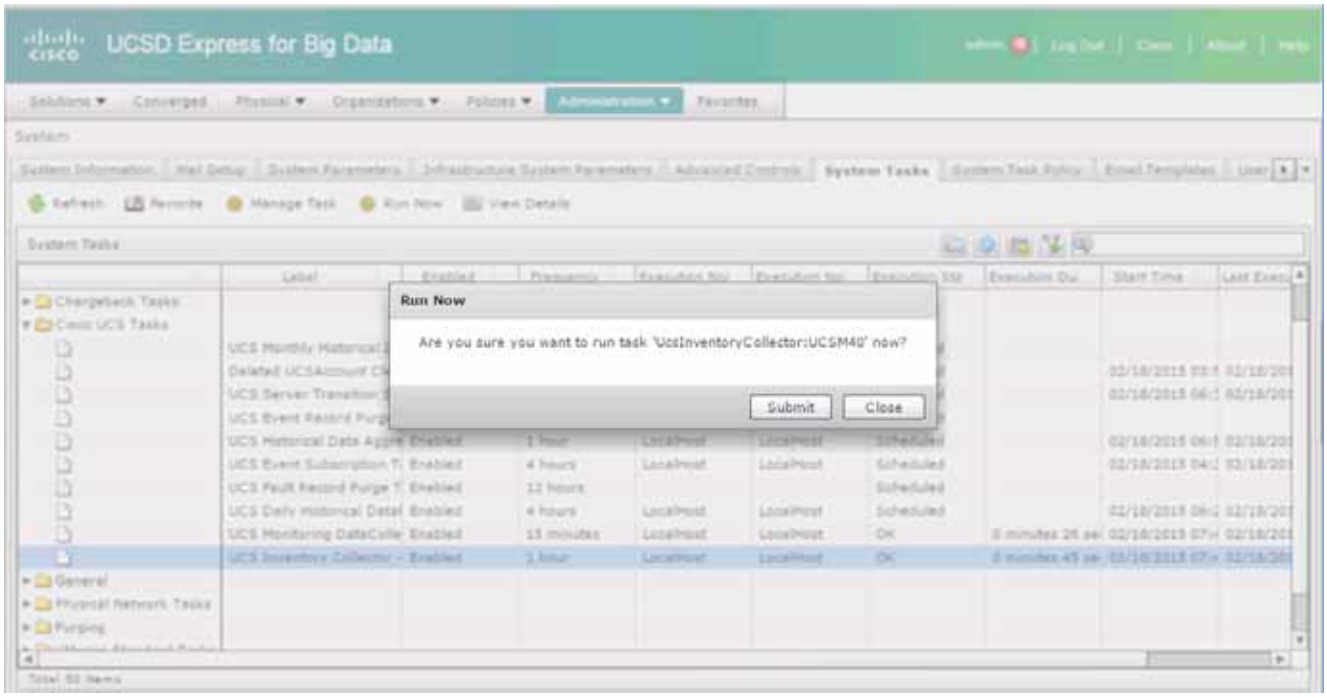
Figure 178 Adding the UCS Manager as a Physical Account in the UCSD-VM



Note After adding a physical account, the UCSD-Express will query the UCS Manager to perform the inventory collection. This process of inventory collection happens at scheduled intervals.Optionally, you may kick start the inventory collection process manually. These optional steps are described in the steps 4 to 8 below.

4. Goto **Administration > System**.
5. Click on **System Tasks** tab.
6. Open the folder Cisco UCS Tasks.
7. Click on UCS Inventory Collector Task.
8. Click **Run Now** button to execute the task.

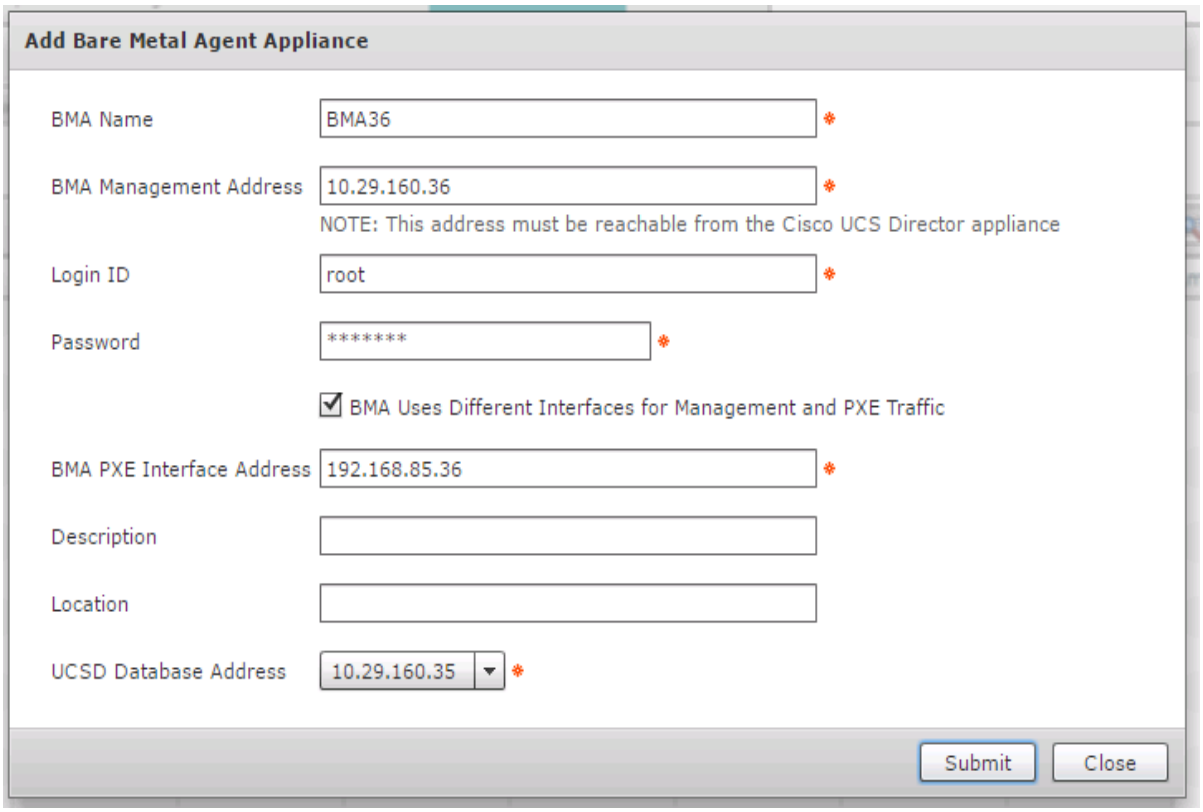
Figure 179 Start the UCS Inventory Collection System Task



Add the Bare Metal Agent physical account to the UCSD-Express

1. In the UCSD-Express web console, navigate to **Administration > Physical Accounts**.
2. Click on **Bare Metal Agents** tab; Click **+ Add**.
3. Enter the BMA physical account information details as follows:
4. In BMA Name field, enter a name to this BMA physical account.
5. In the **BMA Management Address** field, enter the BMA-VM's IP address assigned to **NIC eth0**.
6. In the **Login ID field**, enter **root**.
7. In the **Password** field, enter the password. Default password is **pxeboot**.
8. Check the checkbox **BMA Uses Different Interfaces for Management and PXE Traffic**.
9. In the BMA PXE Interface Address field, enter PXE IP address i.e. BMA-VM's IP address assigned to NIC **eth1**.
10. Click **Submit**.

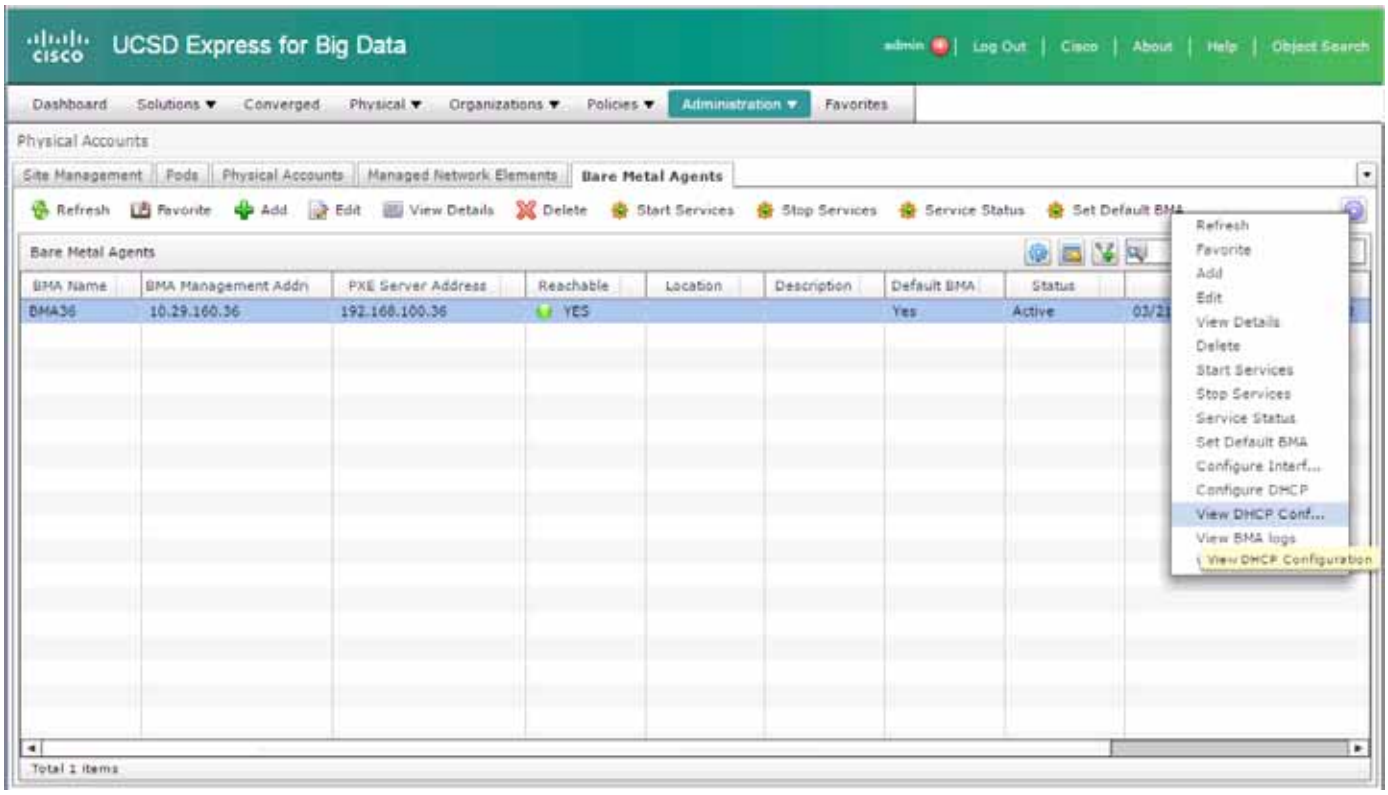
Figure 180 Adding the Bare Metal Agent Appliance Information



Configure the Bare Metal Agent's DHCP services

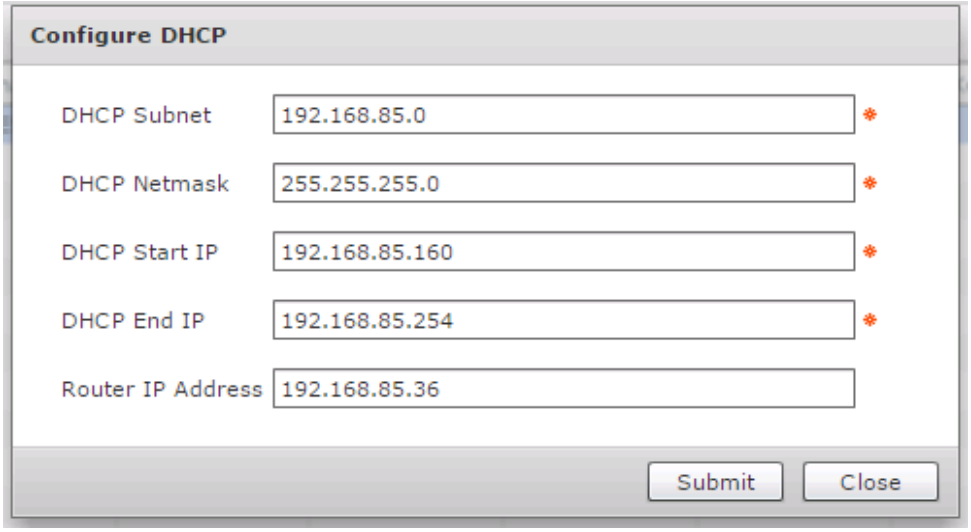
1. Navigate to **Administration > Physical Accounts >Bare Metal Agents**.
2. Select the **BMA** entry.
3. On the menu items row, click on the downward facing arrow located at the far right.
4. Select **Configure DHCP**.

Figure 181 Configuring the DHCP



5. In the **Configure DHCP** dialog box, enter the following
6. In the **DHCP Subnet** field, enter the subnet that's associated with the BMA-VM's **eth1** NIC.
7. In the **DHCP Netmask**, enter the appropriate subnet mask value for this network.
8. In the **DHCP Start IP**, enter a starting IP address in the same subnet.
9. In the **DHCP End IP**, enter a starting IP address in the same subnet.
10. In the **Router IP Address**, enter the IP address of the gateway router in the network if available, if not may be left as blank or input the IP address of the BMA-VM's **eth1** NIC.
11. Click **Submit**.

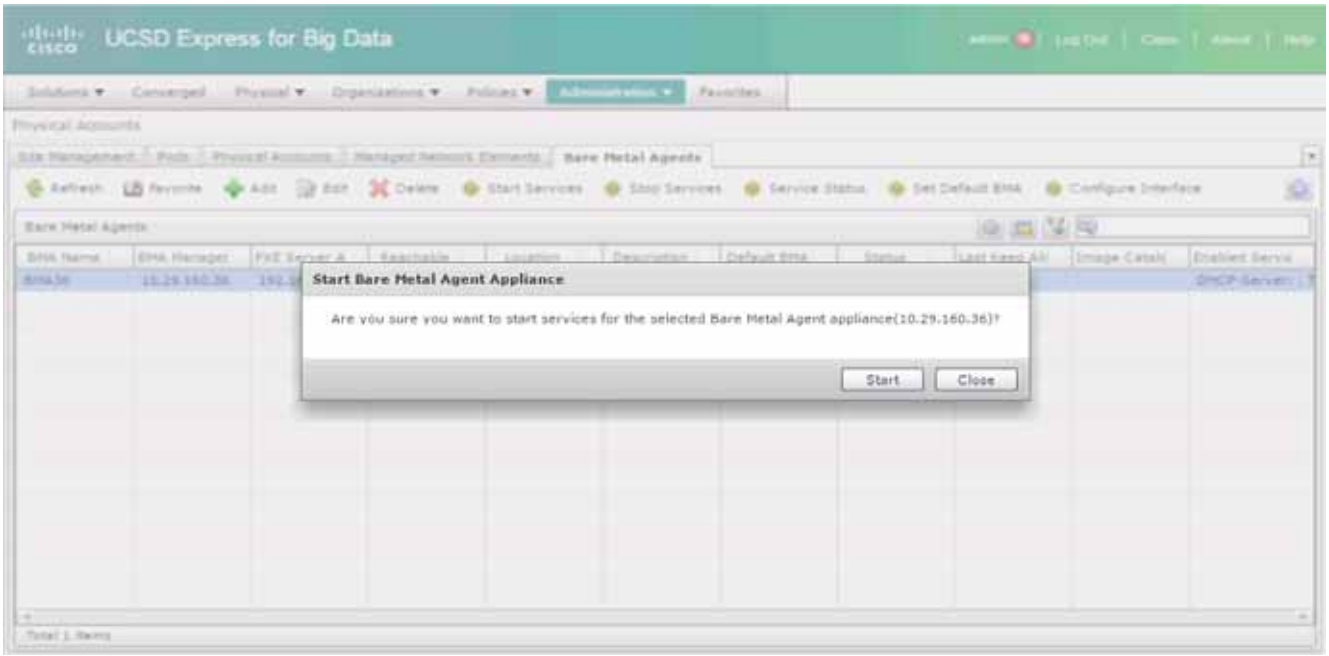
Figure 182 Configuring the DHCP services on the BMA



Start the BMA services

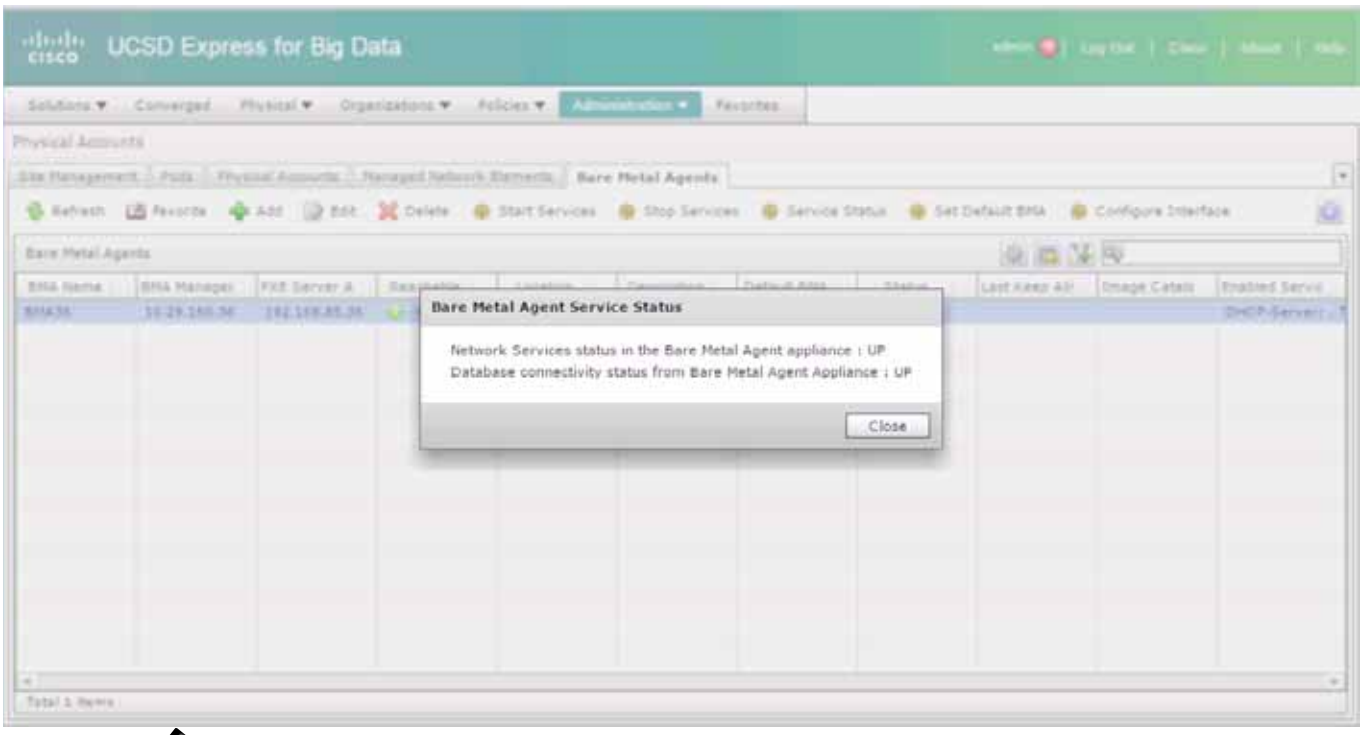
1. Navigate to **Administration >Physical Accounts >Bare Metal Agents**.
2. Select the BMA entry.
3. Click **Start Services**.
4. In the **Start Bare Metal Agent Appliance** dialog box, click **Start** to start the services.

Figure 183 Starting the BMA Services



5. Click on **Service Status**, to **check the status of the services**.
6. The Bare Metal Agent Service Status **message box should display both the** Network Services status and Database connectivity status as UP.

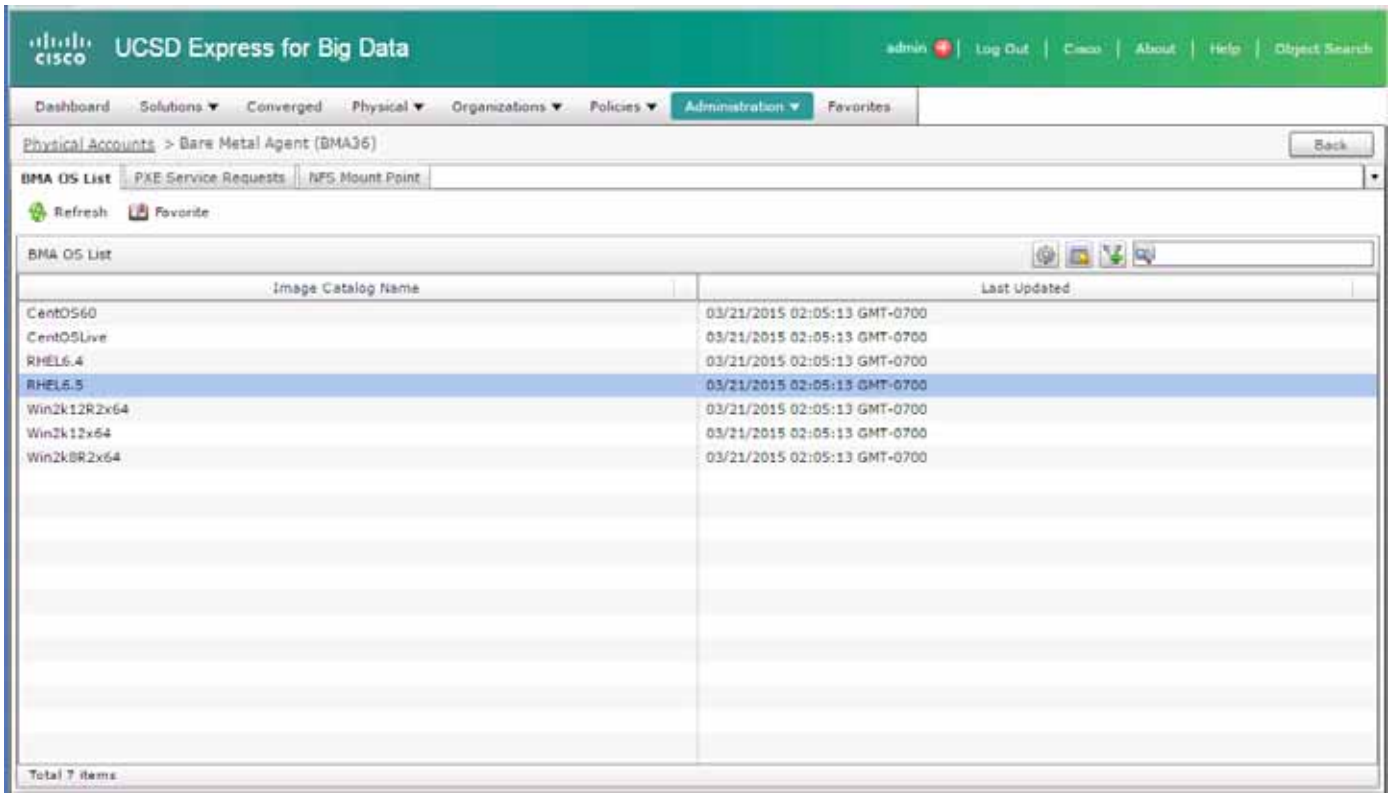
Figure 184 Verifying the Bare Metal Agent Services Status



Note It may take a little while for the service status and on the BMA entry to get updated. The UCSD-Express and the associated BMA parts are now ready.

7. Double click on the BMA entry to verify the RHEL operating system repository.

Figure 185 Verifying the RHEL Operating System Software



Note BMA-VM software periodically scan the /opt/cnsaroot directory to update the available list of operating system software repositories.

Creating the Hadoop Cluster using UCSD-Express

For creating a Hadoop cluster of a desired distribution, the UCS Manager that's managing the target servers must be pre-configured to meet the following requirements. For performing these configurations, refer to any Cisco UCS Integrated Infrastructure for Big Data Cisco Validated Designs found at http://www.cisco.com/go/bigdata_design

a. The uplink ports fabric Interconnects must be reachable to that the UCSD-Express appliances management network (i.e. eth0).

b. The UCS-Manager must be configured with a host firmware policy containing C-series rack mount server firmware packages.

c. UCS Manager must be configured to discover the Rack Servers in its domain, and the respective ports are configured as server ports.

d. The server pool must be configured with appropriate set of physical servers that are part of the UCS domain.

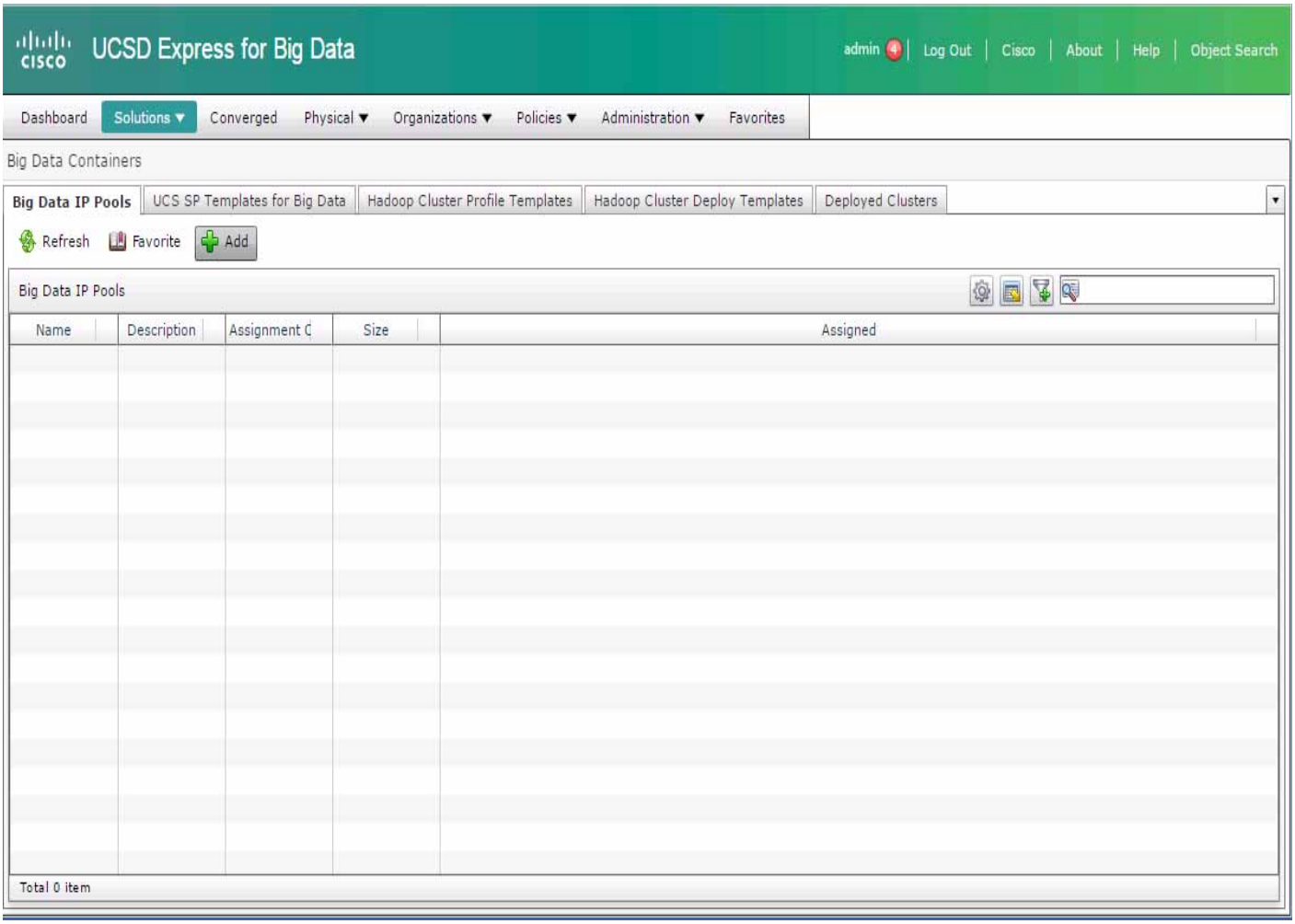
e. The QOS System Classes Platinum and Best Effort must be configured and enabled.

Note C240/C220 M4 Rack Servers are supported from UCS firmware 2.2(3d) onwards.

Create the IP Address pools

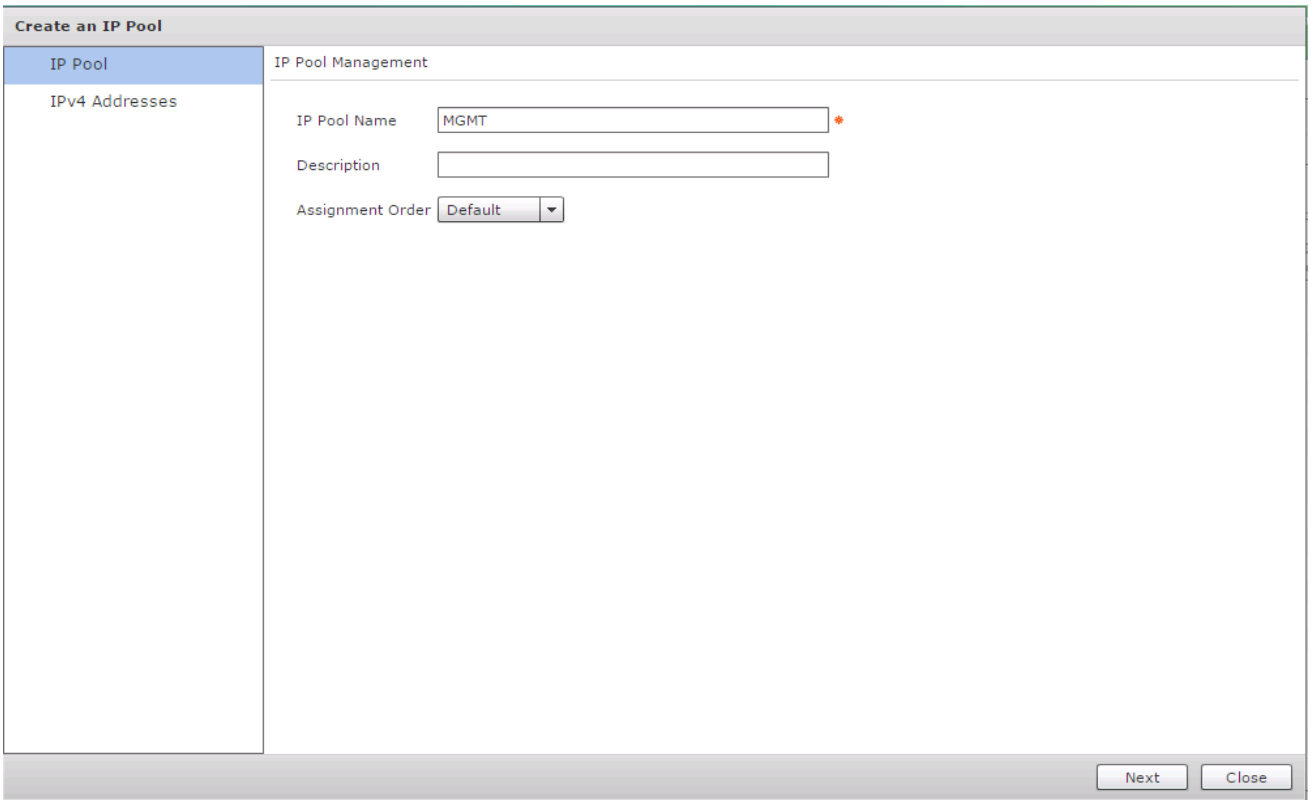
1. Using a web browser, visit the URL **http://<UCSD-VM's IP>/**.
2. Login as user admin with the default password **admin**.
3. Navigate to **Solutions** > **Big Data Containers**.
4. Click on the **Big Data IP Pools** Tab.
5. Click on **+ Add**.

Figure 186 Creating the IP Address Pools



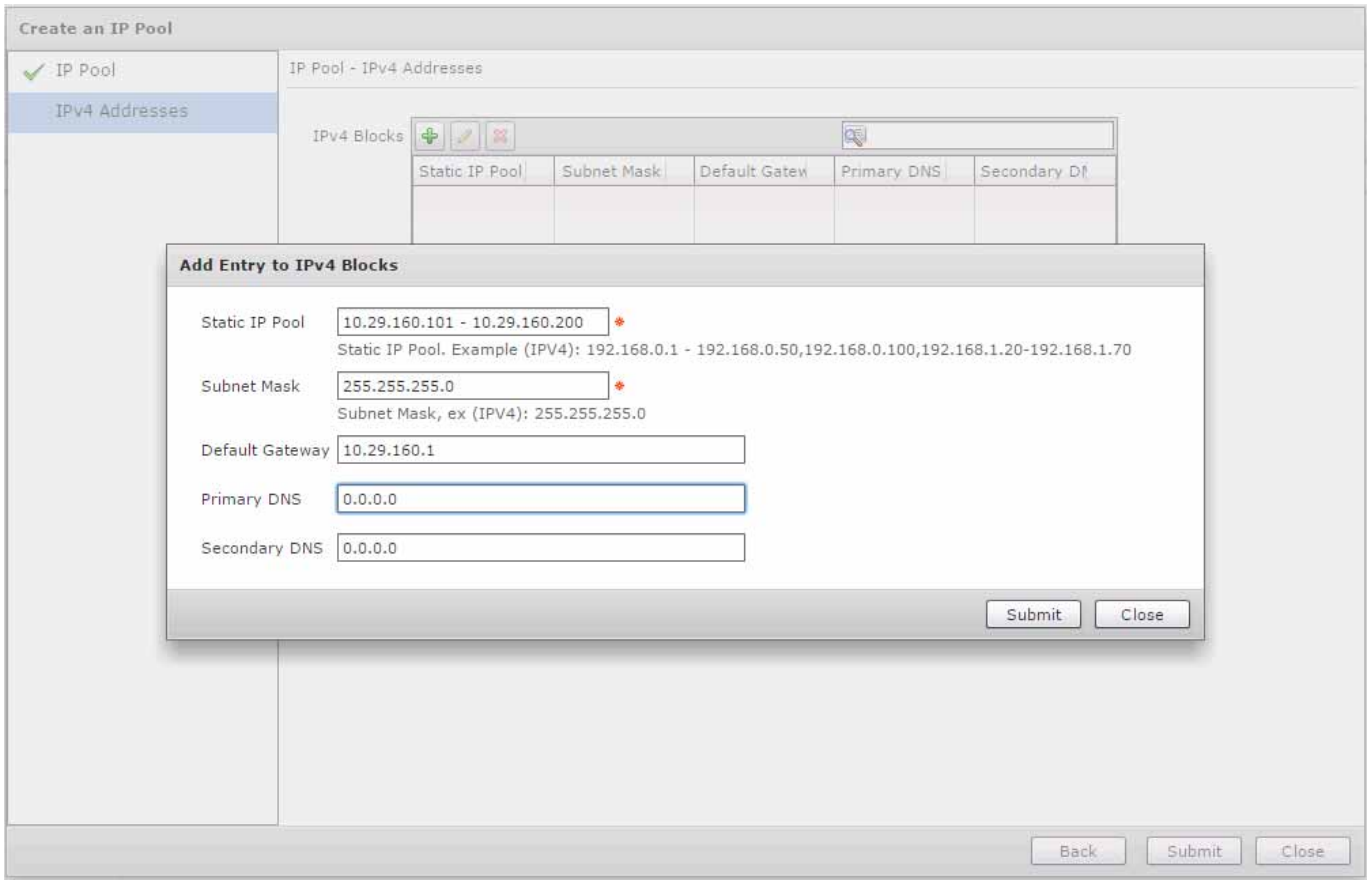
6. In the **Create an IP Pool** dialog box.
7. Enter the name **MGMT**. Click **Next** to continue.

Figure 187 Creating the IP Address pool for MGMT VLAN



8. In the IPv4 Blocks table, click on +.
9. In the Add Entry to IPv4 Blocks dialog box, enter the following.
   - In the Static IP Pool field, enter the Static IP Address pool range in the format A.B.C.X – A.B.C.Y.
   - In the Subnet Mask field, enter the appropriate subnet mask.
   - In the Default Gateway field, enter the IP address of the Gateway if present.
   - In the Primary DNS field, enter the IP address of the DNS server.
10. Click **Submit**.

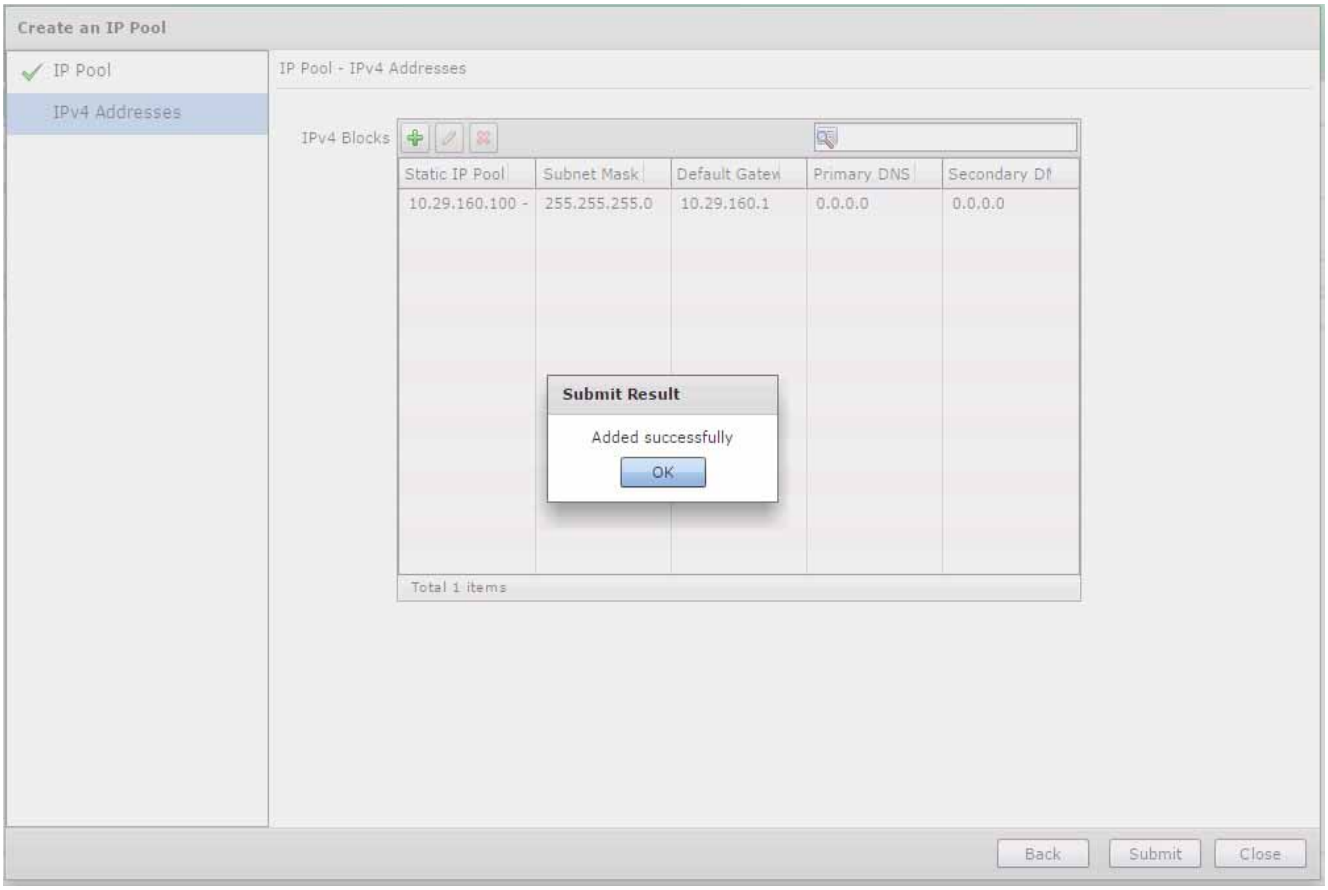
Figure 188 Adding a Block of IP Address to the MGMT IP Address Pool



Note The Default Gateway, Primary and Secondary DNS fields are optional.

11. Click **Submit** again to create the Big Data IP Pool.

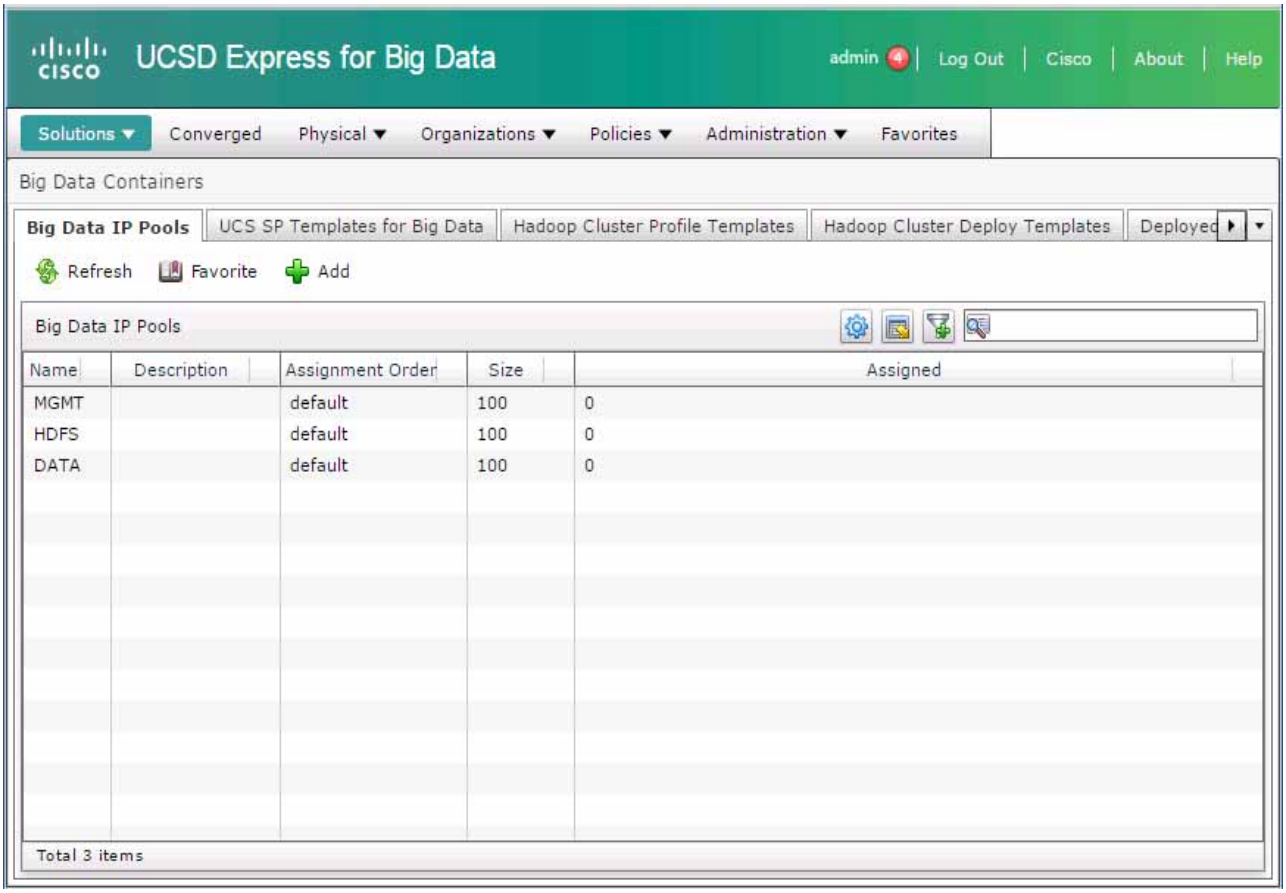
Figure 189 IP Address Pool Added Successfully



Repeat this process for two more interfaces, by creating an IP address pool by name HDFS for Hadoop configurations to be associated with vNIC eth1, and an IP address pool by name DATA to be associated with vNIC eth2 in the service profiles. Please refer to “Configuring VLAN Section” above in Cisco UCS Integrated Infrastructure for Big Data CVDs.

The following figure shows the UCSD-Express that is fully provisioned all the necessary Big Data IP address Pools.

Figure 190 All the IP Address Pools have been Configured Successfully



Creating a Hadoop Cluster

1. Using a web browser, visit the URL **http://<UCSD-VM's IP>/**.
2. Login as user **admin** with the default password **admin**.
3. Navigate to **Solutions >Big Data Containers**.
4. Click on the **Hadoop Cluster Deploy Templates** Tab.
5. Click on **Create Instant Hadoop Cluster**.
6. In the Instant Hadoop Cluster Creation dialog box, enter the following.
7. In Big Data Account Name field, enter a preferred name.
8. In the UCS Manager Policy Name Prefix field, enter a prefix that is less than equal to 5 letters long.
9. In the Hadoop Cluster Name field, enter a preferred name of the cluster – this will be the name assigned to the Hadoop cluster within the context of selected Hadoop Manager.
10. In the Hadoop Node Count filed, enter the desired number of nodes.

The minimum number of nodes allowed for Cloudera and Hortonworks Hadoop cluster is 4 and for MapR cluster it is 3.

Note There should be sufficient number of servers available in the server pool.

11. In the password fields, enter the preferred passwords and confirm them.
12. Choose the OS Version from the drop-down box. For C220 M4/C240 M4 rack servers, only OS supported is RHEL 6.5.

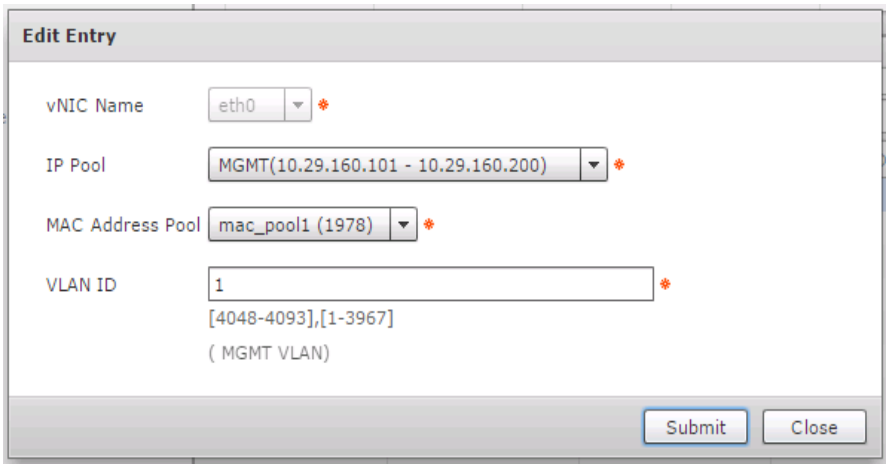
Note At the time of this writing, RHEL6.5 is the only OS that is supported on C220 M4/C240 M4 rack servers.

13. In the Hadoop Distribution field, select **MapR** from the drop-down list.
14. In the Hadoop Distribution Version field, select **MapR-4.0.2** from the drop down list.

Figure 191 Selecting the Hadoop Distribution Version

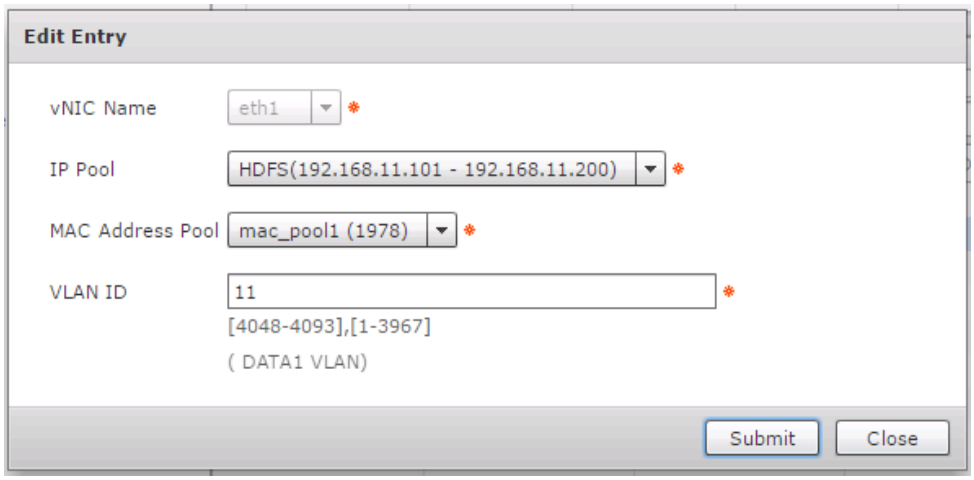
15. In the UCS Manager Account, select the appropriate UCS-Manager account.
16. Select the organization.
17. vNIC Template Entry
18. Double-click on row eth0 and select appropriate Mgmt IP-pool, MAC Address Pool and enter the MGMT VLAN id. Click Submit.

Figure 192 Editing the vNIC Template to Provide the MGMT Network Configurations



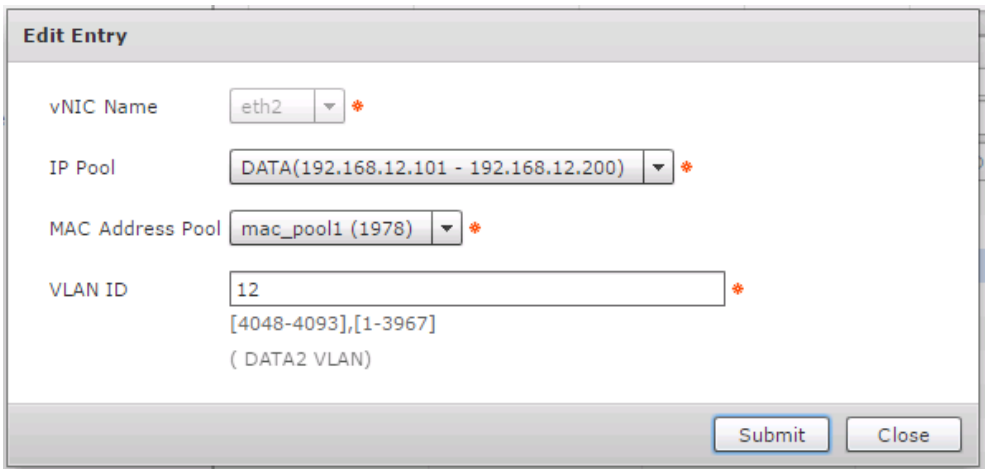
19. Double-click on **eth1** and select appropriate IP-pool, MAC Address Pool and enter the DATA1 VLAN ID. Click **Submit**.

Figure 193 Editing the vNIC Template to Provide the DATA1 Network Configurations



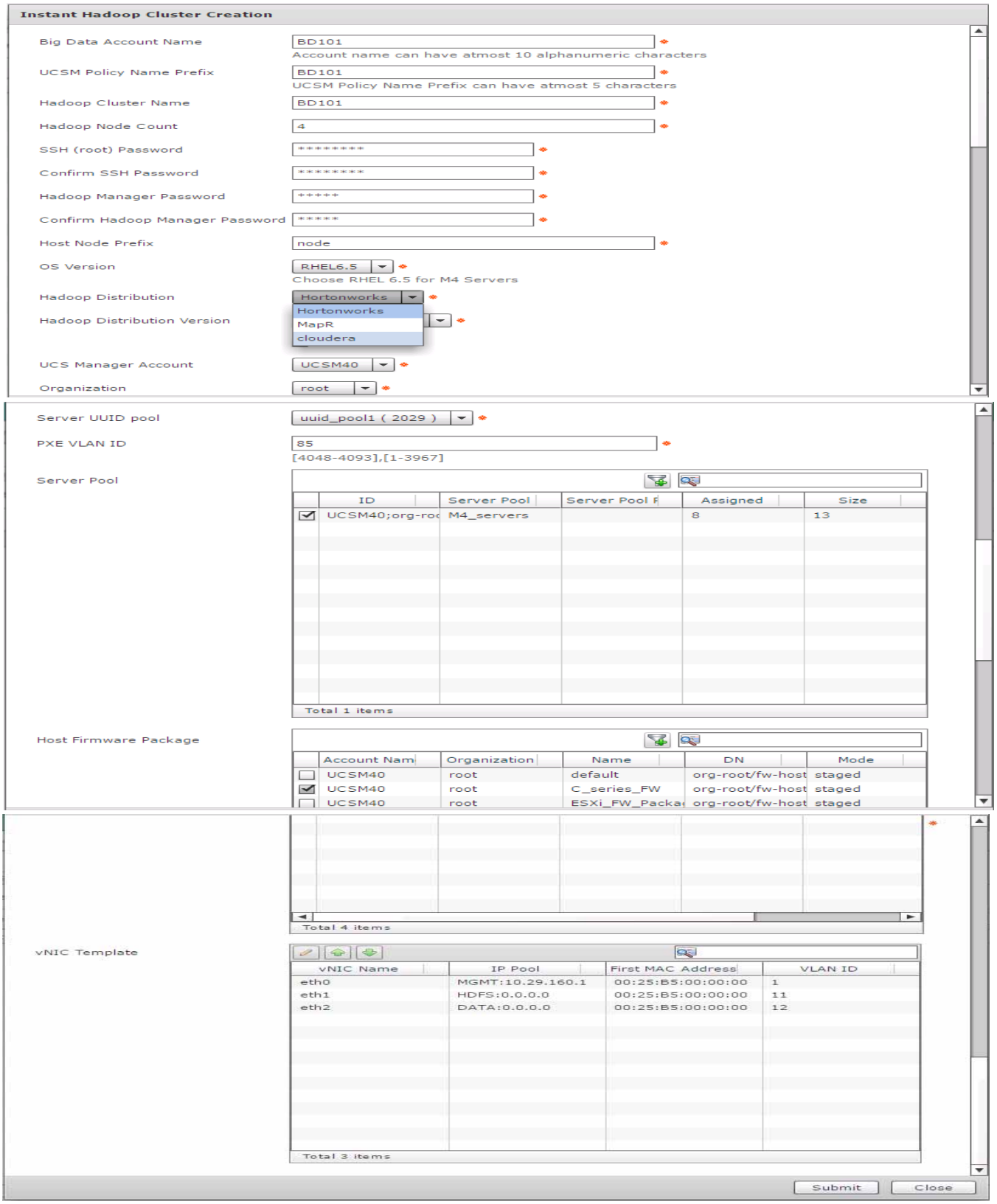
20. Double-click on **eth2** and select appropriate IP-pool, MAC Address Pool and enter the DATA VLAN ID. Click **Submit**.

Figure 194 Editing the vNIC Template to Provide the DATA2 Network Configurations



Note The following figure show the expanded version of the Instant Hadoop Cluster Creation dialog box with all the fields filed in.

Figure 195 Creating an Instant Hadoop Cluster



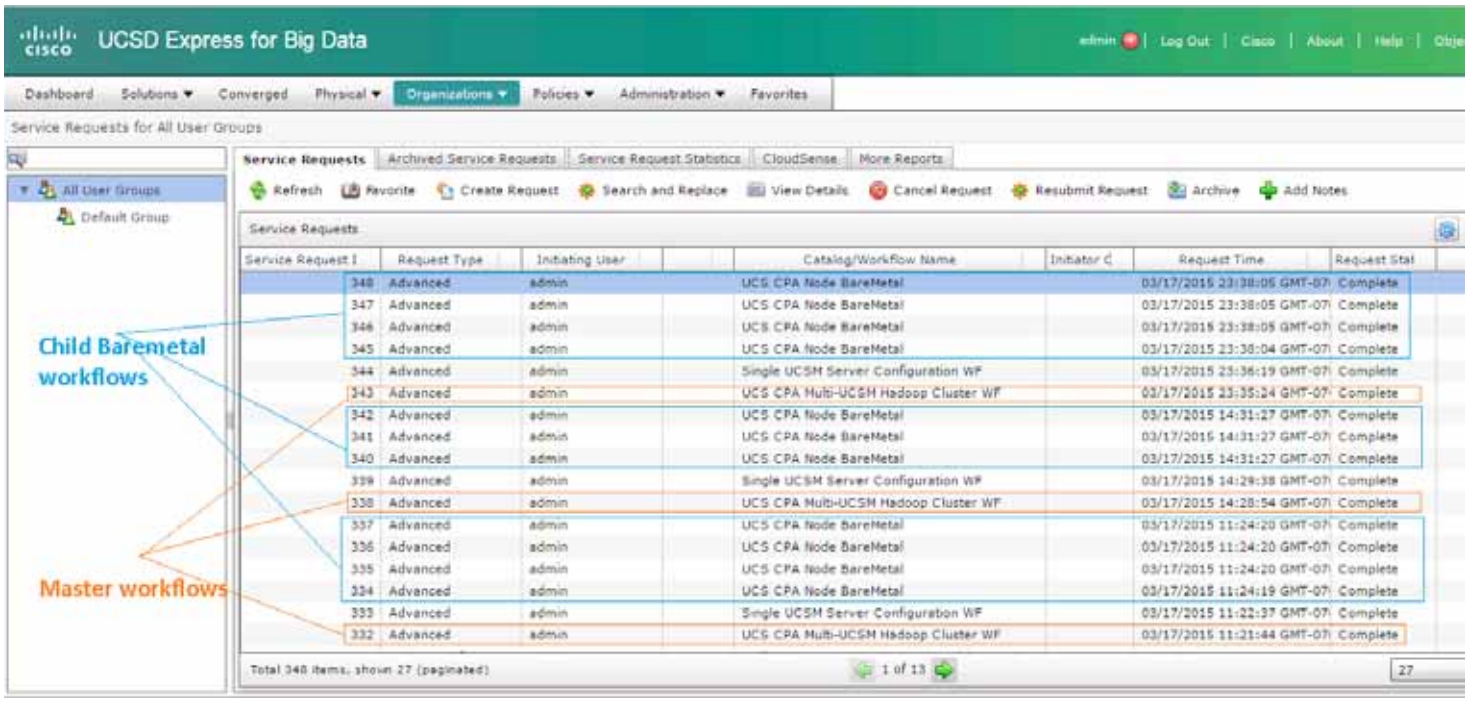
21. Click **Submit**.

Monitoring the Hadoop Cluster Creation

1. In the UCSD-Express web console, navigate to Organization ? Service Requests.
2. Browse through the workflows. There are 3 types of workflows executed.

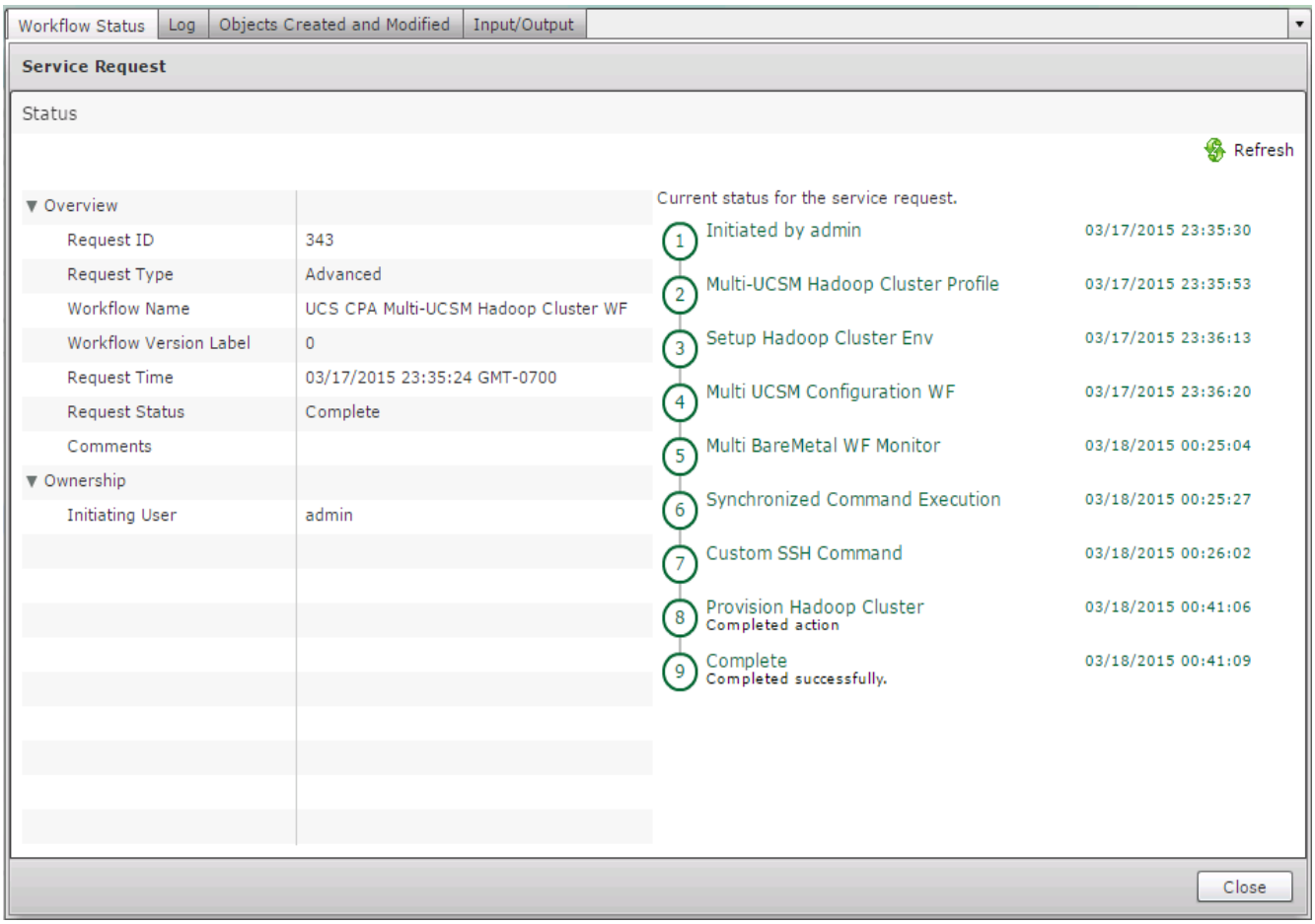
- There would be one Master Workflows i.e. UCS CPA Multi-UCS Manager Hadoop cluster WF, per the Hadoop cluster creation request. Master workflow kick starts one or more UCS Manager-specific workflows. Besides that, this master workflow is responsible for Hadoop cluster provisioning.
- UCS Manager specific workflows i.e. Single UCS Manager Server Configuration WF, would in turn kick start one or more UCS CPA Node Baremetal workflows.
- UCS CPA Baremetal workflows provision the UCS service profiles and perform OS installation and custom configuration per node.

Figure 196 List of Workflows Recently Complete



3. Double-click on one of the master workflows i.e. UCS CPA Multi-UCS Manager Hadoop Cluster to view the various steps undertaken to provision a Hadoop cluster.

Figure 197 Viewing a Completed Master Workflow



Note If necessary click on the Log tab to view the logs generated during the provisioning of the Hadoop Cluster.

4. Double-click on one of the child workflows: i.e. UCS CPA Node Baremetal.

Figure 198 A Completed UCS CPA Node Baremetal workflow.

Workflow Status | Log | Objects Created and Modified | Input/Output

Service Request

Status

Refresh

▼ Overview	
Request ID	345
Request Type	Advanced
Workflow Name	UCS CPA Node BareMetal
Workflow Version Label	0
Request Time	03/17/2015 23:38:04 GMT-0700
Request Status	Complete
Comments	
▼ Ownership	
Initiating User	admin

Current status for the service request.

1	Initiated by admin	03/17/2015 23:38:05
2	Modify Workflow Priority (High)	03/17/2015 23:38:08
3	Assign BareMetal SR ID	03/17/2015 23:38:11
4	Create UCS Service Profile from template	03/17/2015 23:38:17
5	Service Profile unbind/rebind Action	03/17/2015 23:39:21
6	Modify UCS Service Profile Boot Policy	03/17/2015 23:40:23
7	Associate UCS Service Profile	03/17/2015 23:45:59
8	Assign ServerIdentity	03/17/2015 23:46:00
9	Bind/Unbind vNIC Template	03/17/2015 23:46:09
10	Bind/Unbind vNIC Template	03/17/2015 23:46:13
11	Setup PXE Boot (OS Type: CentOSLive)	03/17/2015 23:46:38
12	Setup RAID Commands	03/17/2015 23:46:50

▼ Overview	
Request ID	345
Request Type	Advanced
Workflow Name	UCS CPA Node BareMetal
Workflow Version Label	0
Request Time	03/17/2015 23:38:04 GMT-0700
Request Status	Complete
Comments	
▼ Ownership	
Initiating User	admin

Current status for the service request.

13	UCS Blade Power ON Action	03/17/2015 23:47:34
14	Monitor PXE Boot	03/17/2015 23:53:16
15	Monitor RAID Configuration	03/17/2015 23:53:17
16	UCS Blade Power OFF Action	03/17/2015 23:53:31
17	Setup PXE Boot (OS Type: RHEL6.5)	03/17/2015 23:53:54
18	Setup RAID Commands	03/17/2015 23:53:57
19	UCS Blade Power ON Action	03/17/2015 23:57:17
20	Monitor PXE Boot	03/18/2015 00:04:19
21	Modify UCS Service Profile Boot Policy Server has Local Disks	03/18/2015 00:04:20
22	Service Profile unbind/rebind Action	03/18/2015 00:05:23
23	UCS Blade Power ON Action	03/18/2015 00:11:08
24	Assign IP Status	03/18/2015 00:11:08

Request Status	Complete
Comments	
▼ Ownership	
Initiating User	admin

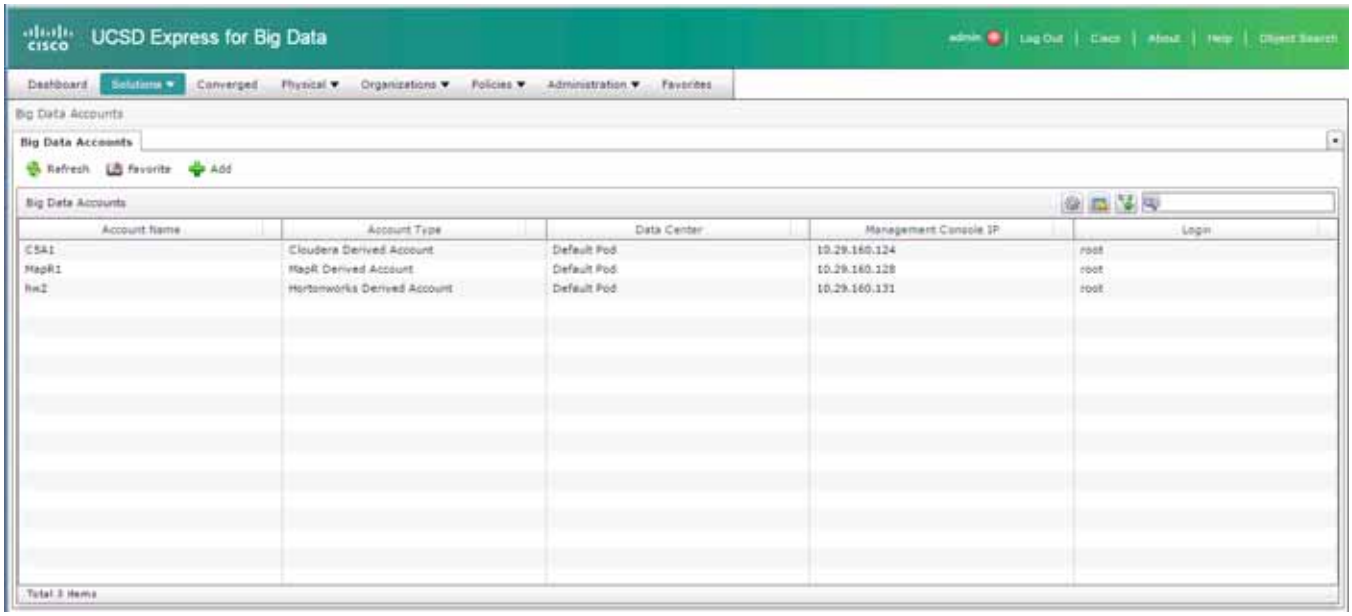
24	Assign IP Status	03/18/2015 00:11:08
25	Custom SSH Command	03/18/2015 00:16:37
26	Custom SSH Command	03/18/2015 00:17:10
27	Synchronized Command Execution	03/18/2015 00:18:14
28	UCS Blade Power OFF Action	03/18/2015 00:18:27
29	UCS Blade Power ON Action	03/18/2015 00:19:40
30	Synchronized Command Execution Completed action	03/18/2015 00:24:29
31	Complete Completed successfully.	03/18/2015 00:24:32

Close

Host and Cluster Performance Monitoring

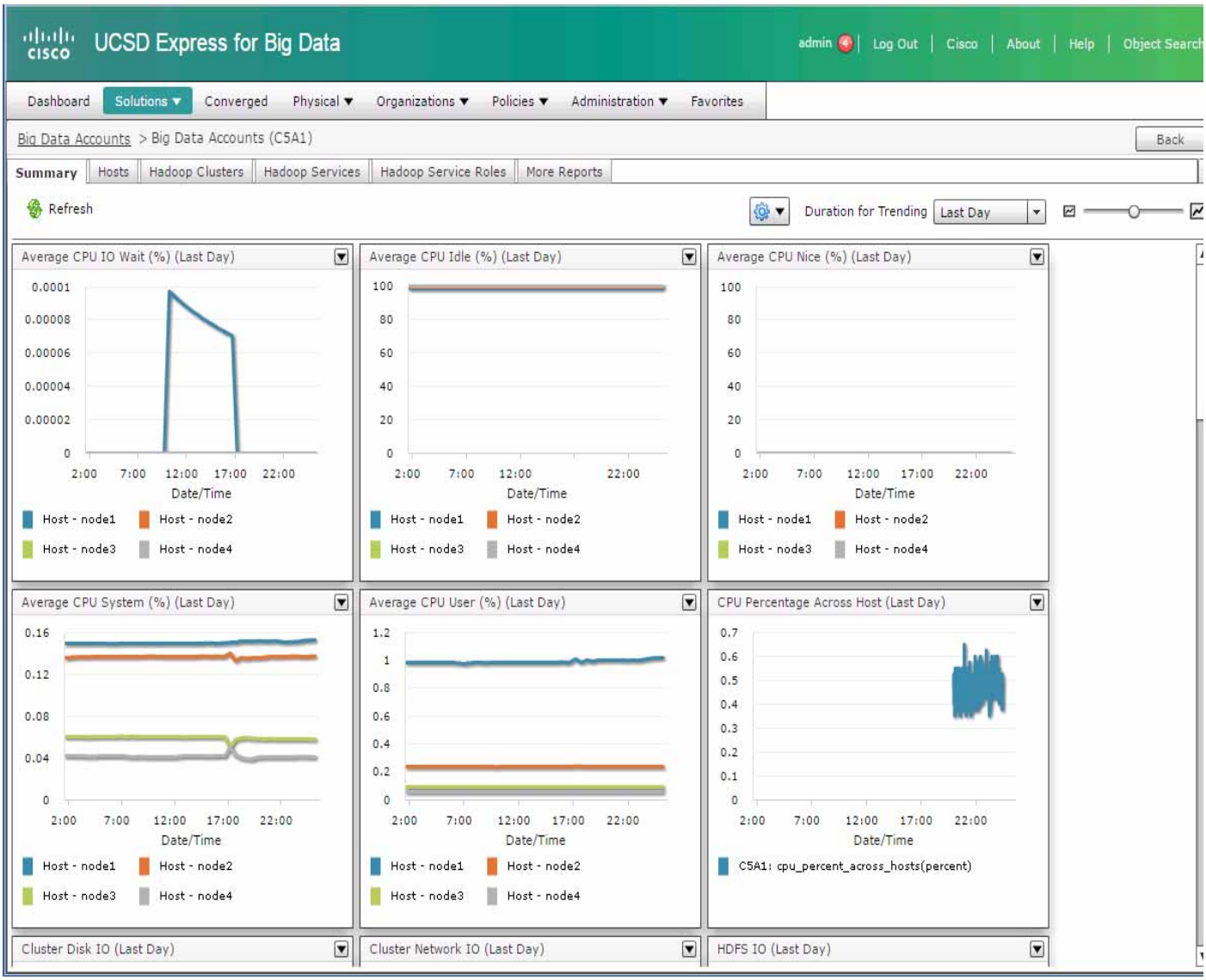
1. In the UCSD-Express web console, navigate to **Solutions > Big Data Accounts** for viewing the Hadoop cluster accounts.

Figure 199 Big Data Accounts Summary Screen



2. Double-click on one of the accounts to view the cluster-wide performance charts.

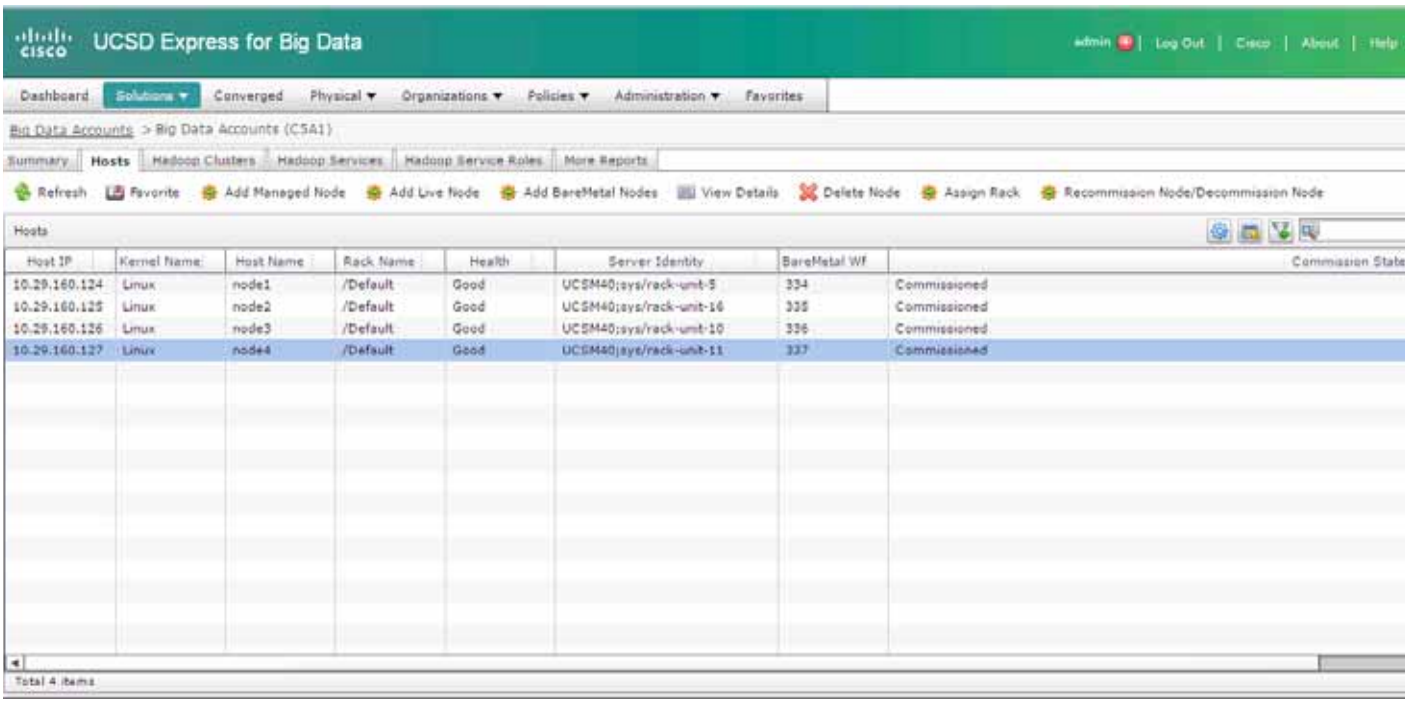
Figure 200 Hadoop Cluster Statistics



Cluster Management

1. In the UCSD-Express web console, navigate to **Solutions > Big Data Accounts** for viewing the Hadoop cluster accounts.
2. Double-Click on one of the accounts to drill into the cluster.
3. Click on the **Hosts** tab.

Figure 201 Big Data Accounts – Viewing the List of Hosts of a Particular Hadoop Cluster

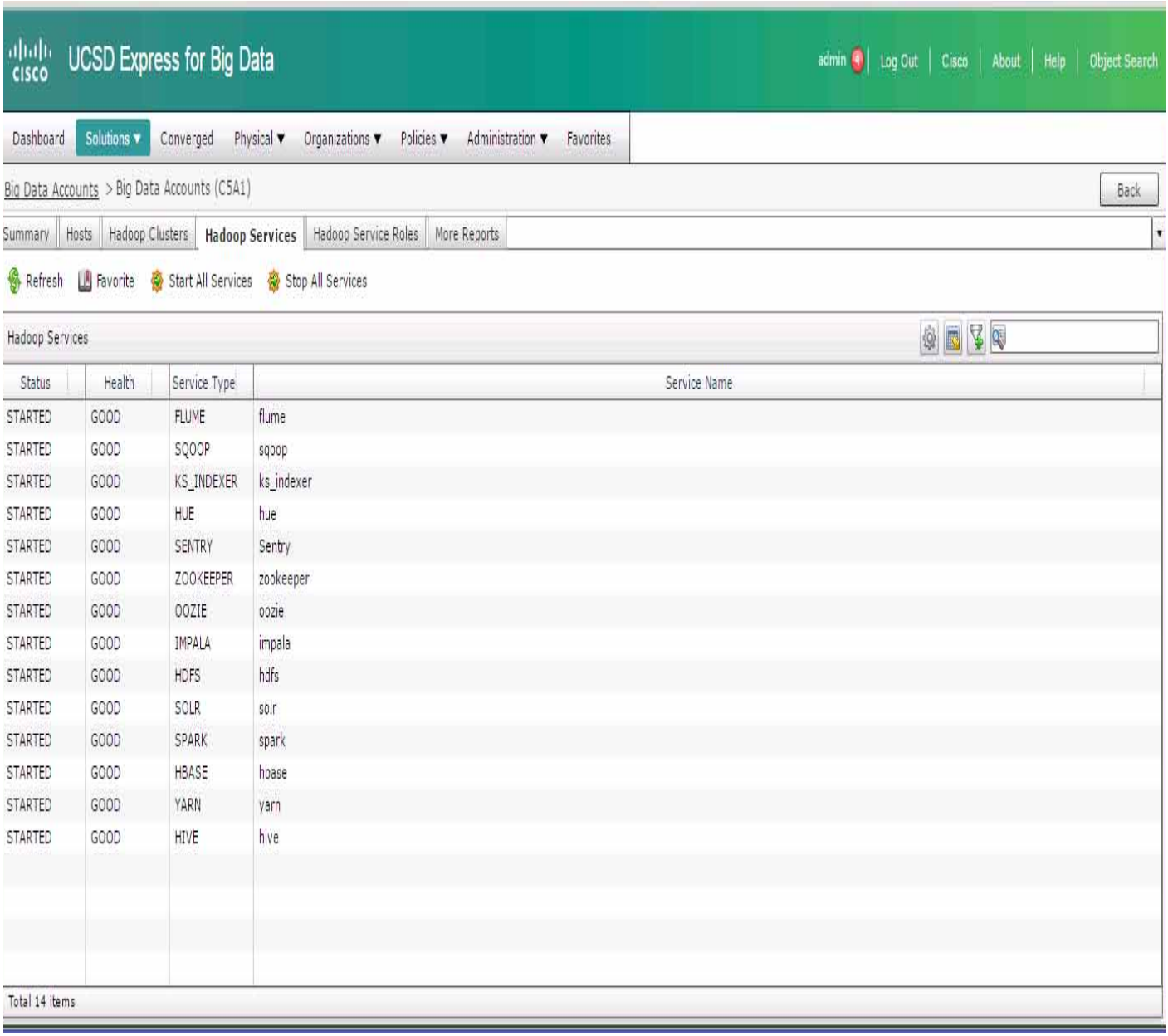


In this screen, the user can perform various management operations such as,

- Add one/more Baremetal nodes to the cluster.
- Delete a node back to Baremetal
- Decommision/Recommission

4. Click on the **Services** tab, where one could Start/Stop the Hadoop services.

Figure 202 Viewing the Services Provisioned in Specific Hadoop Cluster

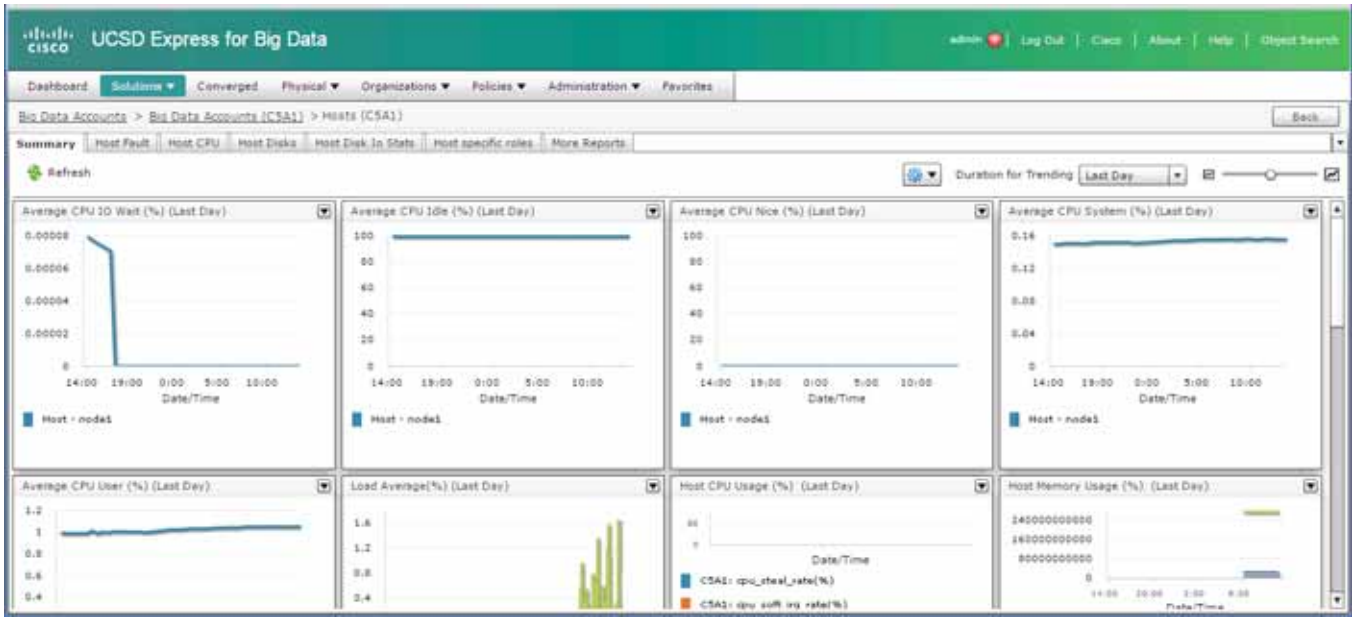


Status	Health	Service Type	Service Name
STARTED	GOOD	FLUME	flume
STARTED	GOOD	SQOOP	sqoop
STARTED	GOOD	KS_INDEXER	ks_indexer
STARTED	GOOD	HUE	hue
STARTED	GOOD	SENTRY	Sentry
STARTED	GOOD	ZOOKEEPER	zookeeper
STARTED	GOOD	OOZIE	oozie
STARTED	GOOD	IMPALA	impala
STARTED	GOOD	HDFS	hdfs
STARTED	GOOD	SOLR	solr
STARTED	GOOD	SPARK	spark
STARTED	GOOD	HBASE	hbase
STARTED	GOOD	YARN	yarn
STARTED	GOOD	HIVE	hive

Host level Monitoring

In the **Hosts** tab, double-click on one of the hosts to view the host's statistics.

Figure 203 Summary Statistics Screen of a Specific Host in a Hadoop Cluster



The user may monitor various resource utilization metrics of the particular host by clicking on the other tabs in this screen.

Reference

For details on managing the Hadoop clusters deployed on the Cisco UCS Integrated Infrastructure for Big Data, see the *Cisco UCS Director Express for Big Data Management Guide* at:

http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-director-express/management-guide/1-1/b_Management_Guide_for_Cisco_UCS_Director_Express_1_1.html

Bill of Materials

Table 23 provides the BOM for Cisco UCSD Big Data subscription licenses for up to 64 servers and Table 24 provides the BOM for the various Hadoop platforms.

Table 17 Bill of Material for UCSD for Big Data Subscription Licenses for up to 64 Servers

CUIC-SVR-OFFERS=	Cisco UCS Director Server Offerings	1
CON-SAU-SVROFFERS	Cisco UCS Director Server Offerings Software Application Sup	1
CUIC-BASE-K9	Cisco UCS Director Software License	1
CON-SAU-CUICBASE	SW APP SUPP + UPGR Cisco UCS Director Base Software	1
CUIC-TERM	Acceptance of Cisco UCS Director License Terms	1

Table 17 Bill of Material for UCSD for Big Data Subscription Licenses for up to 64 Servers

CUIC-EBDS-LIC=	UCSD Express for Big Data - Standard Edition (SE)	1
CUIC-EBDS-LIC	UCSD Express for Big Data - Standard Edition (SE)	64
CUIC-EBDS-S1-3YR	UCSD Express for Big Data - SE 3 year	64
CUIC-TERM	Acceptance of Cisco UCS Director License Terms	1

Table 18 Bill of Material for Various Hadoop Platforms

Part Number	Description
UCS-BD-CEBN=	CLOUDERA ENTERPRISE BASIC EDITION
UCS-BD-CEFN=	CLOUDERA ENTERPRISE FLEX EDITION
UCS-BD-CEDN=	CLOUDERA ENTERPRISE DATA HUB EDITION
UCS-BD-HDP-ENT=	HORTONWORKS ENTERPRISE EDITION
UCS-BD-HDP-EPL=	HORTONWORKS ENTERPRISE PLUS EDITION
UCS-BD-M5-SL=	MAPR M5 EDITION
UCS-BD-M7-SL=	MAPR M7 EDITION