Cisco UCS Common Platform Architecture Version 2 (CPAv2) for Big Data with Hortonworks

Building a 64 Node Hadoop Cluster

Last Updated: April 9, 2014

Cisco Validated Design

Building Architectures to Solve Business Problems

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.

About the Authors

Raghunath Nambiar

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

Manankumar Trivedi

Manankumar Trivedi is a Performance Engineer in the Data Center Solutions Group at Cisco Systems. He is part of solution engineering team focusing on big data infrastructure and performance. He holds masters of Science degree from Stratford University.

Karthik Kulkarni

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.

Ajay Singh

Ajay Singh, Director, Technology Alliances at Hortonworks. Ajay is responsible for design and validation of ecosystem solutions to optimally integrate, deploy and operate Hortonworks Data Platform.

Acknowledgments

The authors acknowledge contributions of Ashwin Manjunatha and Lisa DeRuyter for their contributions in developing this document.

Cisco UCS Common Platform Architecture Version 2 (CPAv2) for Big Data with Hortonworks Data Platform 2.0 (HDP 2.0)

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. The Hortonworks Data Platform (HDP) is a 100% open source distribution of Apache Hadoop that is truly enterprise grade having been built, tested and hardened with enterprise rigor. The combination of HDP and Cisco UCS provides industry-leading platform for Hadoop based applications.

Audience

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. The Hortonworks Data Platform (HDP) is a 100% open source distribution of Apache Hadoop that is truly enterprise grade having been built, tested and hardened with enterprise rigor. The combination of HDP and Cisco UCS provides industry-leading platform for Hadoop based applications.

Cisco UCS Common Platform Architecture Version 2 (CPAv2) for Big Data

The Cisco UCS solution for Hortonworks Data Platform 2.0 is based on Cisco UCS Common Platform Architecture Version 2 (CPAv2) 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:

- **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.
- **Cisco UCS 2200 Series Fabric Extenders** extend the network into each rack, acting as remote line cards for fabric interconnects and providing highly scalable and extremely cost-effective connectivity for a large number of nodes.

- **Cisco UCS C-Series Rack Mount Servers** are 2-socket servers based on Intel Xeon E-2600 v2 series processors and supporting up to 768 GB of main memory. 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 option, along with 4 Gigabit Ethernet LAN-on-motherboard (LOM) ports.
- **Cisco UCS Virtual Interface Cards (VICs)** are 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.
- **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 Unified Computing System, 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.

Hortonworks Data Platform

The Hortonworks Data Platform 2.0 (HDP 2.0) is an enterprise-grade, hardened Apache Hadoop distribution that enables you to store, process, and manage large data sets.

Apache Hadoop is an open-source software framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed for high-availability and fault-tolerance, and can scale from a single server up to thousands of machines.

The Hortonworks Data Platform combines the most useful and stable versions of Apache Hadoop and its related projects into a single tested and certified package. Hortonworks offers the latest innovations from the open source community, along with the testing and quality you expect from enterprise-quality software.

The Hortonworks Data Platform is designed to integrate with and extend the capabilities of your existing investments in data applications, tools, and processes. With Hortonworks, you can refine, analyze, and gain business insights from both structured and unstructured data - quickly, easily, and economically.

Hortonworks Key Features and Benefits

With the Hortonworks Data Platform, enterprises can retain and process more data, join new and existing data sets, and lower the cost of data analysis. Hortonworks enables enterprises to implement the following data management principles:

- **Retain as much data as possible.** Traditional data warehouses age, and over time will eventually store only summary data. Analyzing detailed records is often critical to uncovering useful business insights.
- **Join new and existing data sets.** Enterprises can build large-scale environments for transactional data with analytic databases, but these solutions are not always well suited to processing nontraditional data sets such as text, images, machine data, and online data. Hortonworks enables enterprises to incorporate both structured and unstructured data in one comprehensive data management system.

- **Archive data at low cost.** It is not always clear what portion of stored data will be of value for future analysis. Therefore, it can be difficult to justify expensive processes to capture, cleanse, and store that data. Hadoop scales easily, so you can store years of data without much incremental cost, and find deeper patterns that your competitors may miss.
- **Access all data efficiently.** Data needs to be readily accessible. Apache Hadoop clusters can provide a low-cost solution for storing massive data sets while still making the information readily available. Hadoop is designed to efficiently scan all of the data, which is complimentary to databases that are efficient at finding subsets of data.
- **Apply data cleansing and data cataloging.** Categorize and label all data in Hadoop with enough descriptive information (metadata) to make sense of it later, and to enable integration with transactional databases and analytic tools. This greatly reduces the time and effort of integrating with other data sets, and avoids a scenario in which valuable data is eventually rendered useless.
- **Integrate with existing platforms and applications.** There are many business intelligence (BI) and analytic tools available, but they may not be compatible with your particular data warehouse or DBMS. Hortonworks connects seamlessly with many leading analytic, data integration, and database management tools.

The Hortonworks Data Platform is the foundation for the next-generation enterprise data architecture—one that addresses both the volume and complexity of today's data.

Solution Overview

The current version of the Cisco UCS Common Platform Architecture (CPAv2) Version 2 for Big Data offers the following configuration depending on the compute and storage requirements:

Table 1 Cisco UCS CPAv2 Configuration Details

Performance and Capacity Balanced	Capacity Optimized	Capacity Optimized with Flash Memory
16 Cisco UCS C240 M3 Rack Servers, each with: · 2 Intel Xeon processors E5-2660 v2 · 256 GB of memory · LSI MegaRaid 9271CV 8i card · 24 1-TB 7.2K SFF SAS drives (384 TB total)	16 Cisco UCS C240 M3 Rack Servers, each with: · 2 Intel Xeon processors E5-2640 v2 · 128 GB of memory · LSI MegaRaid 9271CV 8i card · 12 4-TB 7.2 LFF SAS drives (768 TB total)	16 Cisco UCS C240 M3 Rack Servers, each with: · 2 Intel Xeon processors E5-2660 v2 · 128 GB of memory · Cisco UCS Nytro MegaRAID 200-GB Controller · 12 4-TB 7.2K LFF SAS drives (768 TB total)

Note This CVD describes the installation process for a 64 node Performance and Capacity Balanced Cluster configuration.

The Performance and capacity balanced cluster configuration consists of the following:

- Two Cisco UCS 6296UP Fabric Interconnects
- Eight Cisco Nexus 2232PP Fabric Extenders (two per rack)
- 64 Cisco UCS C240 M3 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, two Cisco Nexus 2232PP Fabric Extenders and sixteen Cisco UCS C240M3 Servers, connected to each of the vertical PDUs for redundancy; thereby, ensuring availability during power source failure. The expansion racks consists of two Cisco Nexus 2232PP Fabric Extenders and sixteen Cisco UCS C240M3 Servers are 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	Cisco Nexus FEX 2232PP
37	Cisco Nexus FEX 2232PP
36	Unused
35	Unused
34	Unused
33	Unused
32	Cisco UCS C240M3
31	
30	Cisco UCS C240M3
29	
28	Cisco UCS C240M3
27	
26	Cisco UCS C240M3
25	
24	Cisco UCS C240M3
23	
22	Cisco UCS C240M3
21	

20	Cisco UCS C240M3
19	
18	Cisco UCS C240M3
17	
16	Cisco UCS C240M3
15	
14	Cisco UCS C240M3
13	
12	Cisco UCS C240M3
11	
10	Cisco UCS C240M3
9	
8	Cisco UCS C240M3
7	
6	Cisco UCS C240M3
5	
4	Cisco UCS C240M3
3	
2	Cisco UCS C240M3
1	

Table 3 Rack 2-4 (Expansion Racks)

Cisco 42URack	Expansion Rack
42	Unused
41	Unused
40	Unused
39	Unused
38	Cisco Nexus FEX 2232PP
37	Cisco Nexus FEX 2232PP
36	Unused
35	Unused
34	Unused
33	Unused
32	Cisco UCS C240M3
31	
30	Cisco UCS C240M3
29	
28	Cisco UCS C240M3
27	
26	Cisco UCS C240M3
25	
24	Cisco UCS C240M3
23	
22	Cisco UCS C240M3
21	
20	Cisco UCS C240M3
19	
18	Cisco UCS C240M3
17	
16	Cisco UCS C240M3
15	

14	Cisco UCS C240M3
13	
12	Cisco UCS C240M3
11	
10	Cisco UCS C240M3
9	
8	Cisco UCS C240M3
7	
6	Cisco UCS C240M3
5	
4	Cisco UCS C240M3
3	
2	Cisco UCS C240M3
1	

Server Configuration and Cabling

The Cisco UCS C240 M3 rack server is equipped with Intel Xeon E5-2660 v2 processors, 256 GB of memory, Cisco UCS Virtual Interface Card 1225 Cisco, Cisco LSI MegaRAID SAS 9271 CV-8i storage controller and 24 x 1TB 7.2K SATA disk drives.

Figure 1 illustrates the ports on the Cisco Nexus 2232PP fabric extender connecting to the Cisco UCS C240M3 Servers. Sixteen Cisco UCS C240M3 servers are used in Master rack configurations.

Figure 1 Fabric Topology

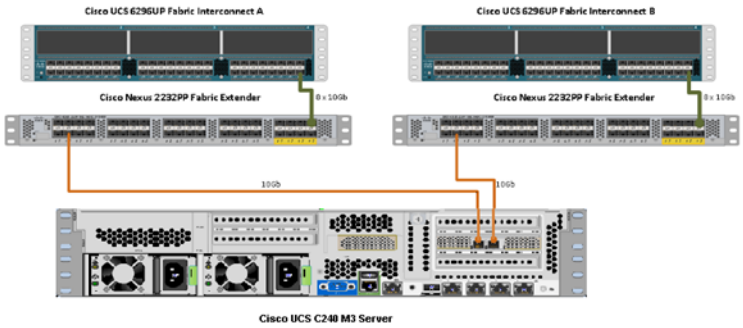
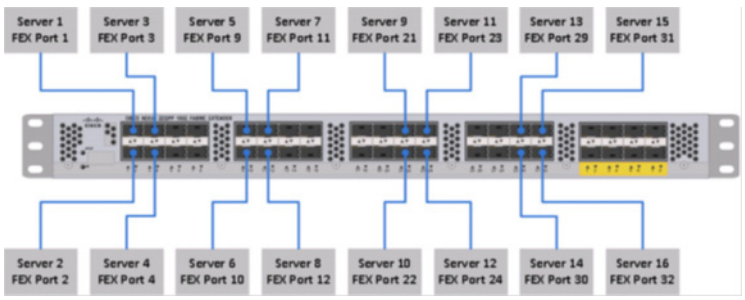


Figure 2 illustrates the port connectivity between the Cisco Nexus 2232PP fabric extender and Cisco UCS C240M3 server.

Figure 2 Connectivity Diagram of Cisco Nexus 2232PP FEX and Cisco UCS C240M3 Servers



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

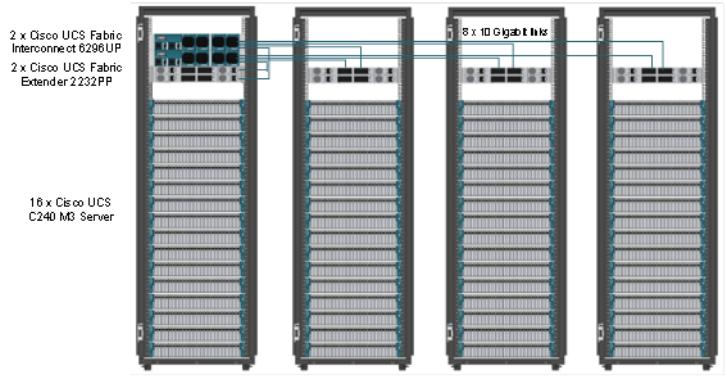
http://www.cisco.com/en/US/docs/unified_computing/ucs/c-series_integration/ucsm2.1/b_UCSM2-1_C-Integration_chapter_010.html

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 3 depicts a 64 node cluster. Each link in the figure represents 8 x 10 Gigabit links.

Figure 3 64 Nodes Cluster Configuration



Software Distributions and Versions

The software distributions required versions are listed below.

Hortonworks Data Platform (HDP 2.0)

The Hortonworks Data Platform supported is HDP 2.0. For more information visit http://www.hortonworks.com.

Red Hat Enterprise Linux (RHEL)

The operating system supported is Red Hat Enterprise Linux 6.4. For more information visit http://www.redhat.com.

Software Versions

The software versions tested and validated in this document are shown in Table 4.

Table 4 Software Version

Layer	Component	Version or Release
Compute	Cisco UCS C240-M3	1.5.4f
Network	Cisco UCS 6296UP	UCS 2.2(1b)A
	Cisco UCS VIC1225 Firmware	2.2(1b)
	Cisco UCS VIC1225 Driver	2.1.1.41
	Cisco Nexus 2232PP	5.2(3)N2(2.21b)
Storage	LSI 9271-8i Firmware	23.12.0-0021
	LSI 9271-8i Driver	06.602.03.00
Software	Red Hat Enterprise Linux Server	6.4 (x86_64)
	Cisco UCS Manager	2.2(1b)
	HDP	2.0

The latest drivers can be downloaded from the link below:

http://software.cisco.com/download/release.html?mdfid=284296254&flowid=31743&softwareid=283853158&release=1.5.1&relind=AVAILABLE&rellifecycle=&reltype=latest

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 IP address of Fabric Interconnect A using web browser.
3. Launch UCS Manager.
4. Edit the chassis discovery policy.
5. Enable server and uplink ports.
6. Create pools and polices for service profile template.
7. Create Service Profile template and 64 Service profiles.
8. Start discover process.

9. Associate to server.

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 username 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(1b)

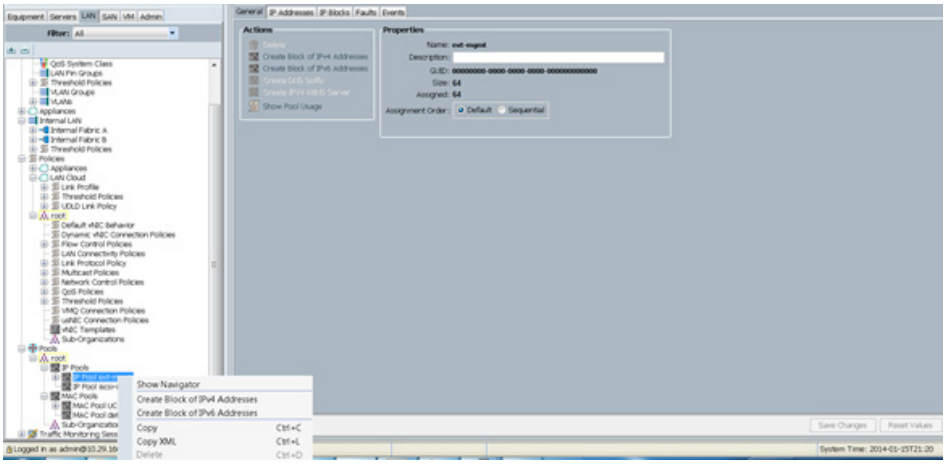
This document assumes the use of UCS 2.2(1b). Refer to Upgrading between Cisco UCS 2.0 Releases to upgrade the Cisco UCS Manager software and Cisco UCS 6296 Fabric Interconnect software to version 2.2(1b). Also, make sure the Cisco UCS C-Series version 2.2(1b) software bundles is installed on the Fabric 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.

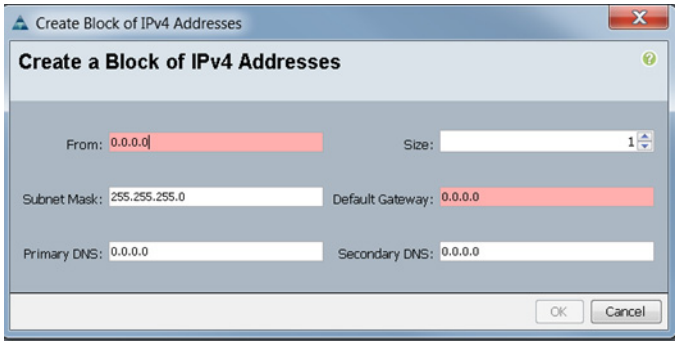
1. Select the LAN tab at the top of the left window.
2. Select `Pools > IpPools >` Ip Pool ext-mgmt.
3. Right-click `IP Pool ext-mgmt.`
4. Select `Create Block of IPv4 Addresses.`

Figure 4 Adding Block of IPv4 Addresses for KVM Access Part 1



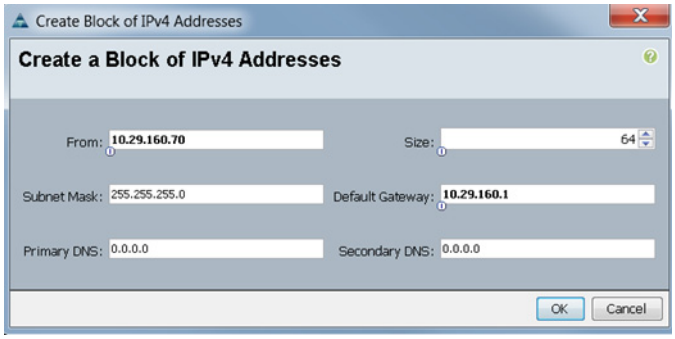
5. Enter the starting IP address of the block and number of IPs needed, as well as the subnet and gateway information.

Figure 5 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 6 Adding Block of IPv4 Addresses for KVM Access Part 3

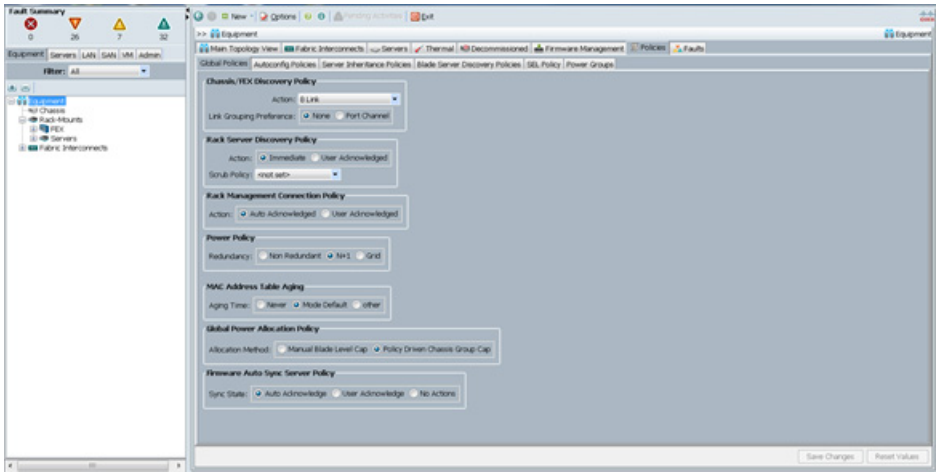


Editing Chassis and FEX Discovery Policy

These steps provide details for modifying the chassis discovery policy. Setting the discovery policy now will simplify the addition of future Cisco UCS B-Series Chassis and additional Fabric Extenders for further Cisco UCS C-Series connectivity.

1. Navigate to the `Equipment` tab in the left pane.
2. In the right pane, click the `Policies` tab.
3. Under `Global Policies`, change the `Chassis/FEX Discovery Policy` to `8-link`.
4. Click `Save Changes` in the bottom right hand corner.
5. Click `OK`.

Figure 7 Chasses and FEX Discovery Policy



Enabling Server Ports and Uplink Ports

Follow these steps to enable server and 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 all the ports that are connected to the Cisco 2232 FEX (8 per FEX), right-click them, and select `Reconfigure > Configure as a Server Port`.
5. Select port 1 that is connected to the uplink switch, right-click, then select `Reconfigure > Configure as Uplink Port`.
6. Select Show Interface and select 10GB for Uplink Connection.
7. A pop-up window appears to confirm your selection. Click `Yes` then `OK` to continue.
8. Select `Equipment > Fabric Interconnects > Fabric Interconnect B (subordinate) > Fixed Module`.
9. Expand the `UnConfigured Ethernet Ports` section.
10. Select all the ports that are connected to the Cisco 2232 Fabric Extenders (8 per Fex), right-click them, and select `Reconfigure > Configure as Server Port`.
11. A prompt displays asking if this is what you want to do. Click `Yes` then `OK` to continue.
12. Select port number 1, which is connected to the uplink switch, right-click, then select `Reconfigure > Configure as Uplink Port`.
13. Select Show Interface and select 10GB for Uplink Connection.
14. A pop-up window appears to confirm your selection. Click `Yes` then OK to continue.

Figure 8 Enabling Server Ports

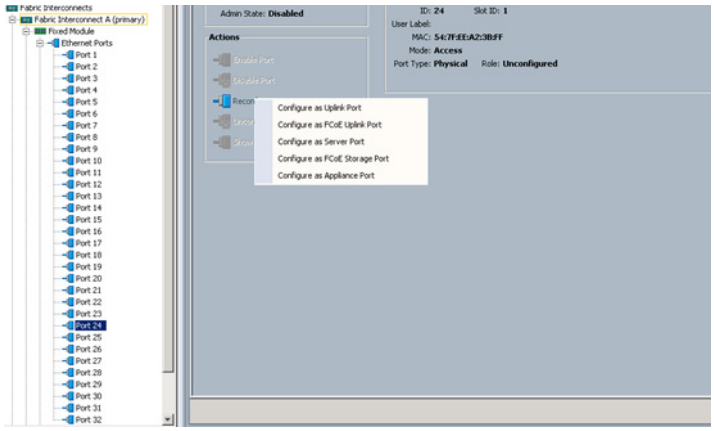
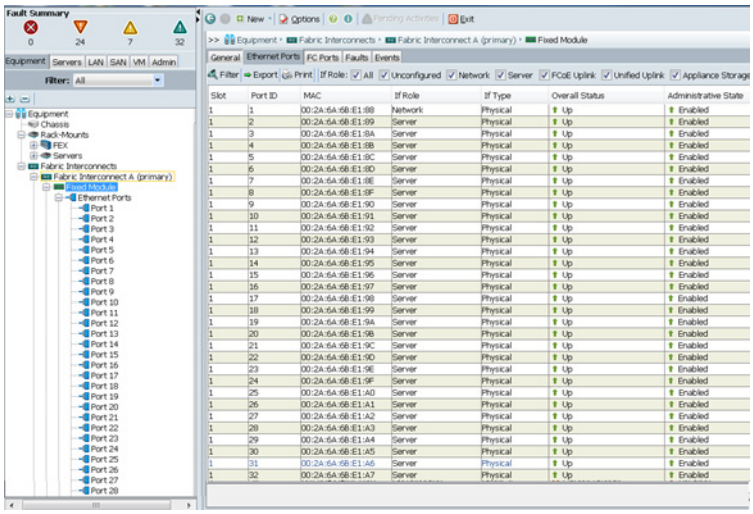


Figure 9 Showing Servers and Uplink Ports



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 Cisco UCS Manager GUI.
2. Select Create Organization from the options
3. Enter a name for the organization.

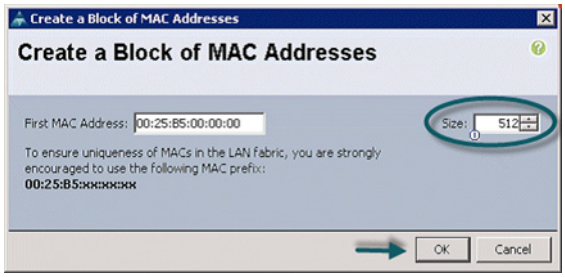
4. (Optional) Enter a description for the organization.
5. Click OK.
6. Click OK in the success message box.

Creating MAC Address Pools

Follow these steps to create MAC address pools:

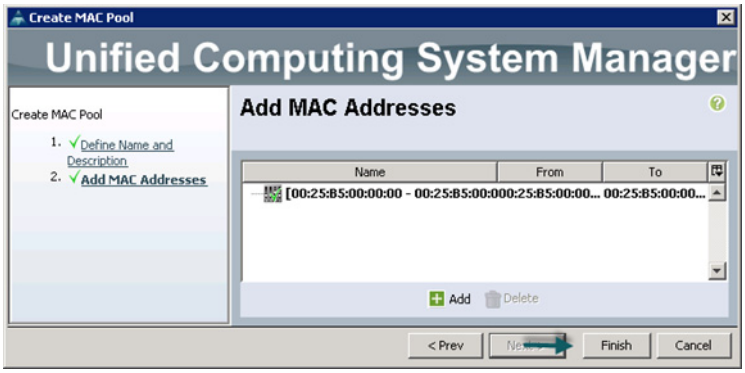
1. Select th `LAN` tab on the left of the window.
2. Select `Pools > root`.
3. Right-click `MAC Pools` under the root organization.
4. Select `Create MAC Pool` to create the MAC address pool. Enter `ucs` for the name of the MAC pool.
5. (Optional) Enter a description of the MAC pool.
6. Click `Next`.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

Figure 10 Specifying First MAC Address and Size



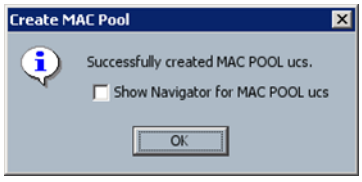
11. Click `Finish`.

Figure 11 Adding MAC Addresses



12. When the message box displays, click OK.

Figure 12 Confirming Newly Added MAC Pool



Configuring VLANs

VLANs are configured as in shown in Table 5.

Table 5 VLAN Configurations

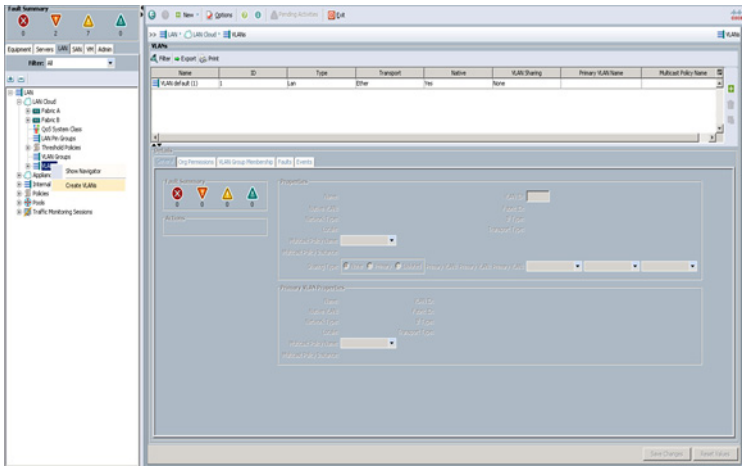
VLAN	Fabric	NIC Port	Function	Failover
vlan160_mgmt	A	eth0	Management, User connectivity	Fabric Failover to B
vlan12_HDFS	B	eth1	Hadoop	Fabric Failover to A
vlan11_DATA	A	eth2	Hadoop and/or SAN/NAS access, ETL	Fabric Failover to B

All of the VLANs created need to be trunked to the upstream distribution switch connecting the fabric interconnects. For this deployment vlan160_mgmt is configured for management access and user connectivity, vlan12_HDFS is configured for Hadoop interconnect traffic and vlan11_DATA is configured for optional secondary interconnect and/or SAN/NAS access, heavy ETL, etc.

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

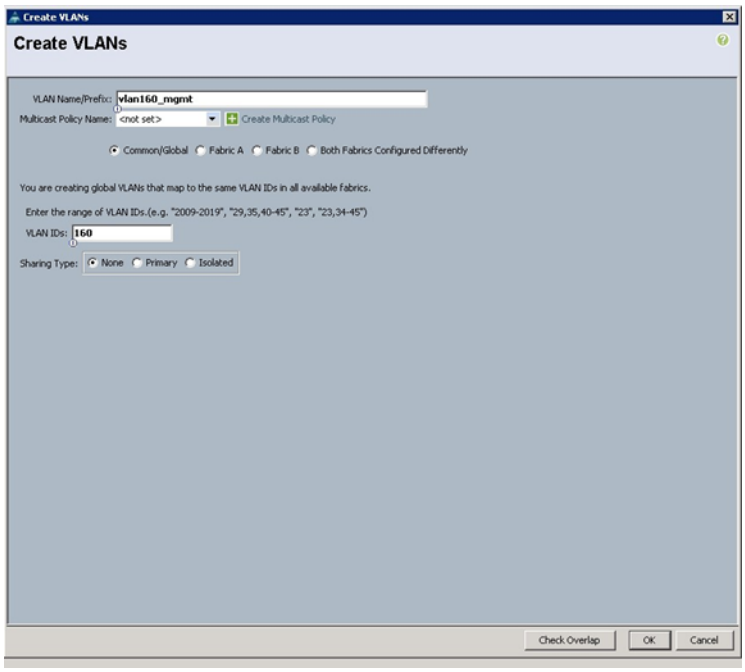
1. Select the LAN tab in the left pane in the Cisco UCS Manager GUI.
2. Select LAN > VLANs.
3. Right-click the VLANs under the root organization.
4. Select Create VLANs to create the VLAN.

Figure 13 Creating VLAN



5. Enter vlan160_mgmt for the VLAN Name.
6. Select Common/Global for vlan160_mgmt.
7. Enter 160 on VLAN IDs of the Create VLAN IDs.
8. Click OK and then, click Finish.
9. Click OK in the success message box.

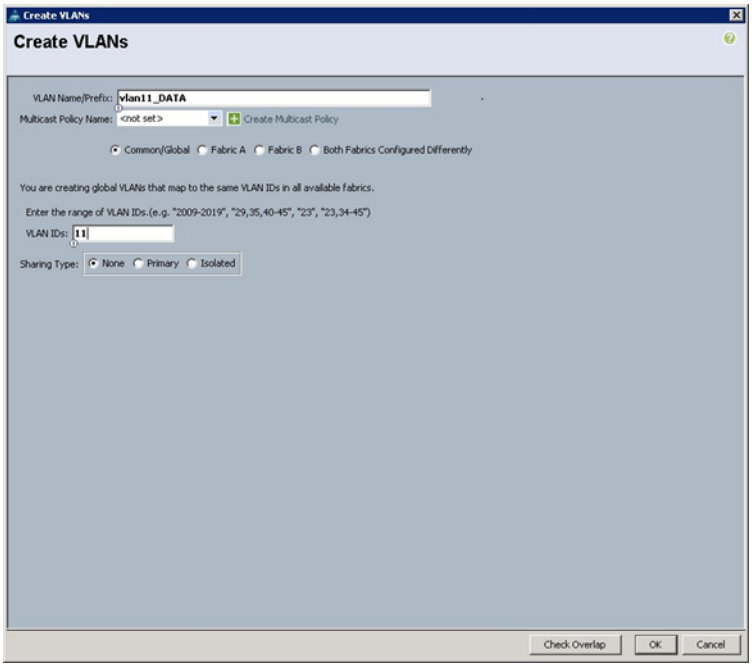
Figure 14 Creating Management VLAN



10. Select the LAN tab in the left pane again.
11. Select LAN > VLANs.

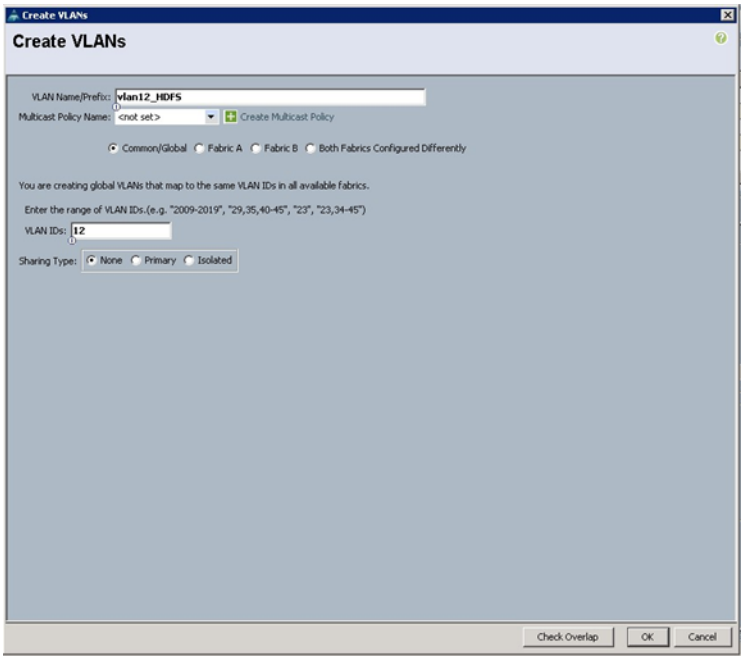
12. Right-click the VLANs under the root organization.
13. Select Create VLANs to create the VLAN.
14. Enter vlan11_DATA for the VLAN Name.
15. Select Common/Global for the vlan11_DATA.
16. Enter 11 on VLAN IDs of the Create VLAN IDs.
17. Click OK and then, click Finish.
18. Click OK in the success message box.

Figure 15 Creating VLAN for Data



19. Select the LAN tab in the left pane again.
20. Select LAN > VLANs.
21. Right-click the VLANs under the root organization.
22. Select Create VLANs to create the VLAN.
23. Enter vlan12_HDFS for the VLAN Name.
24. Select Common/Global for the vlan12_HDFS.
25. Enter 12 on VLAN IDs of the Create VLAN IDs.
26. Click Ok and then, click Finish.

Figure 16 Creating VLAN for Hadoop Data



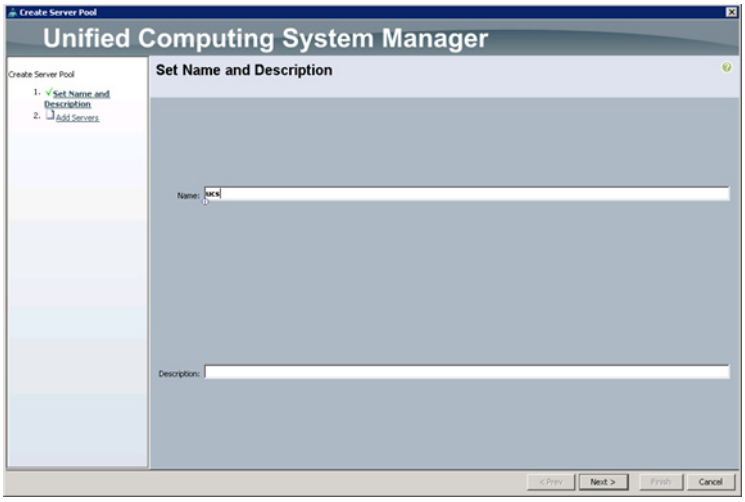
Creating a Server Pool

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:

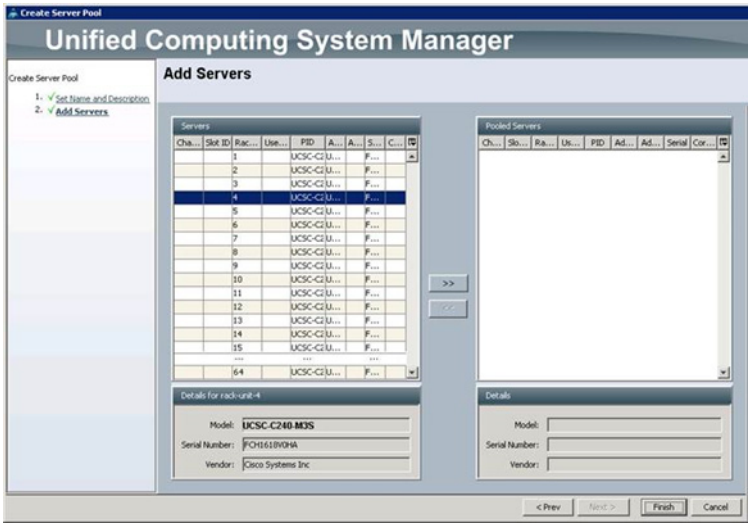
1. Select the `Servers` tab in the left pane in the UCS Manager GUI.
2. Select `Pools > root`.
3. Right-click the `Server Pools`.
4. Select `Create Server Pool`.
5. Enter your required name (ucs) for the Server Pool in the name text box.
6. (Optional) enter a description for the organization.
7. Click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

Figure 17 Setting Name and Description of the Server Pool



8. Select all the Cisco UCS C240M3S 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 18 Adding Servers to the Server Pool



Creating Policies for Service Profile Templates

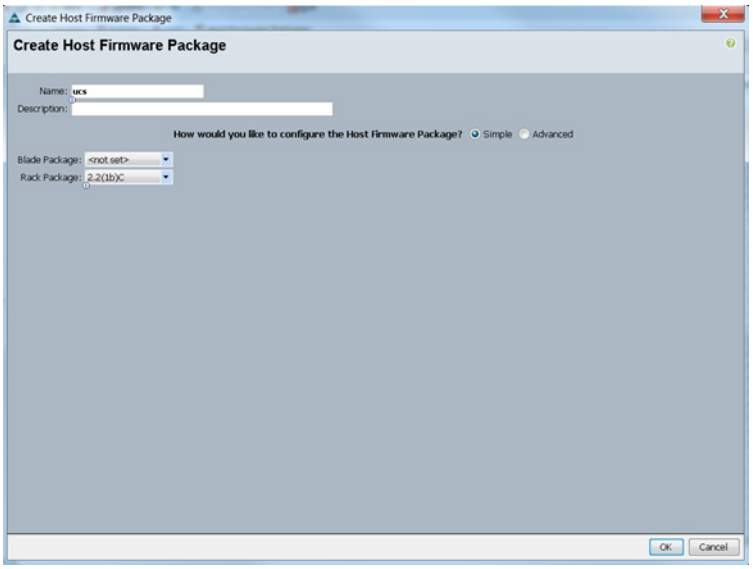
Creating Host Firmware Package Policy

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 Cisco 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. Select 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 19 Creating Host Firmware Package



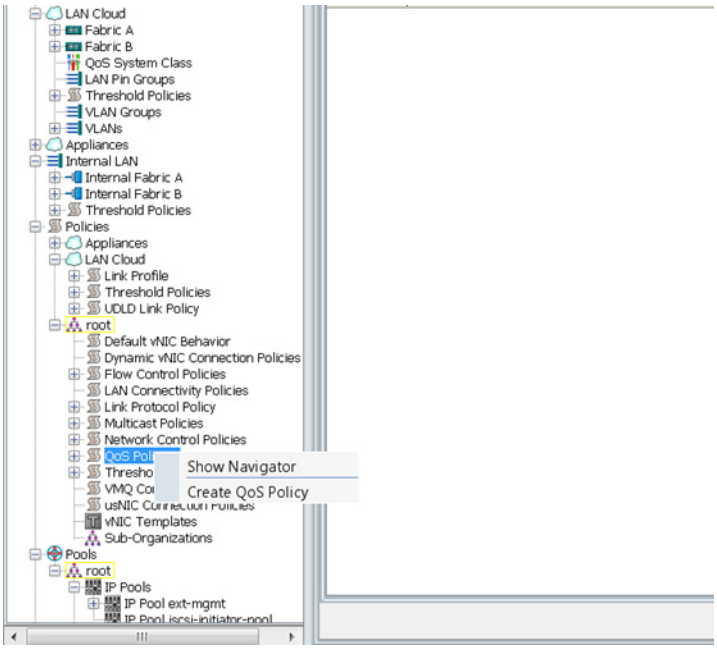
Creating QoS Policies

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

Best Effort Policy

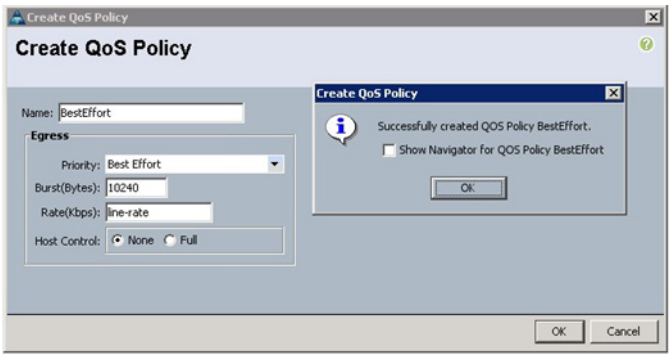
1. Select the `LAN` tab in the left pane in the Cisco UCS Manager GUI.
2. Select `Policies > root`.
3. Right-click QoS Policies.
4. Select `Create QoS Policy`.

Figure 20 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. When the pop-up window appears, click OK to complete the creation of the Policy.

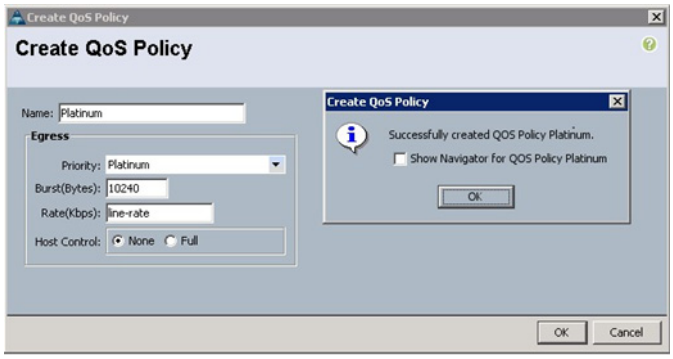
Figure 21 Creating Best Effort Qos Policy



Platinum Policy

1. Select the `LAN` tab in the left pane in the UCSM 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. When the pop-up window appears, click OK to complete the creation of the Policy.

Figure 22 Creating Platinum QoS Policy



Setting Jumbo Frames

Follow these steps for setting Jumbo frames and enabling QoS:

1. Select the `LAN` tab in the left pane in the UCSM 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 23 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. Uncheck` 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 24 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 Cisco 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 25 Creating Server BIOS Policy

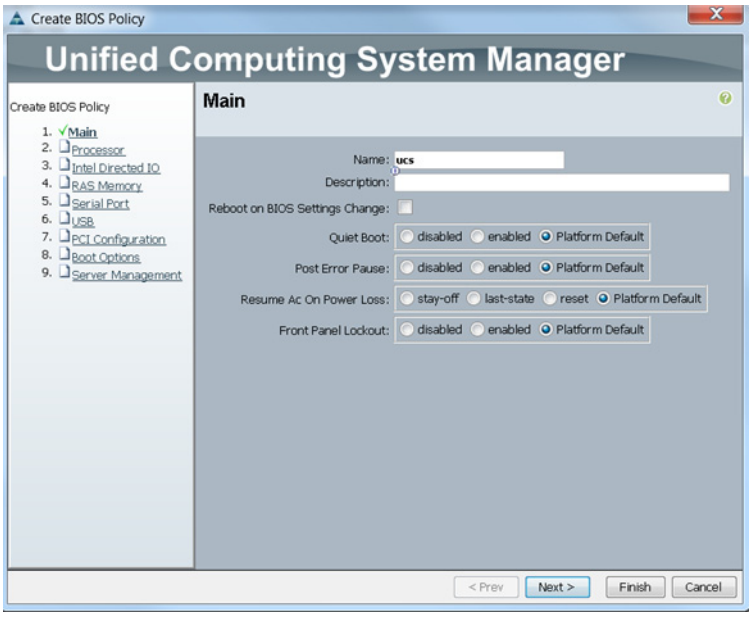


Figure 26 Creating Server BIOS Policy for Processor

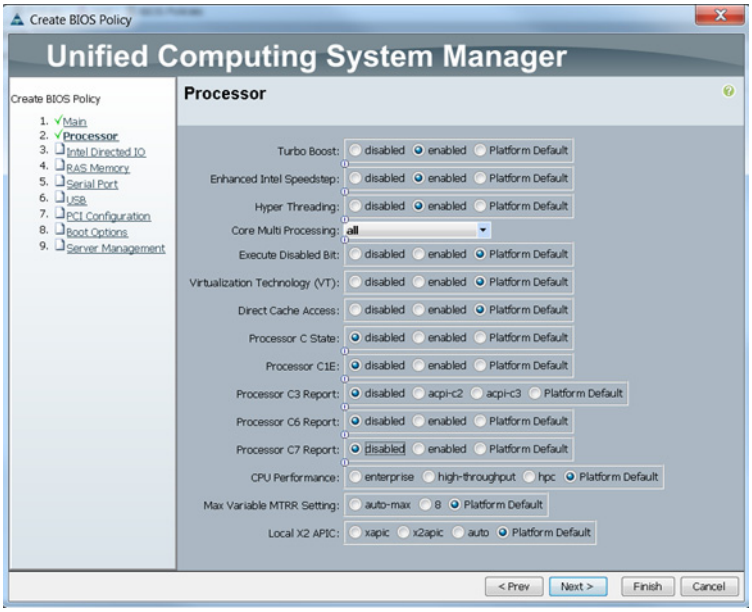
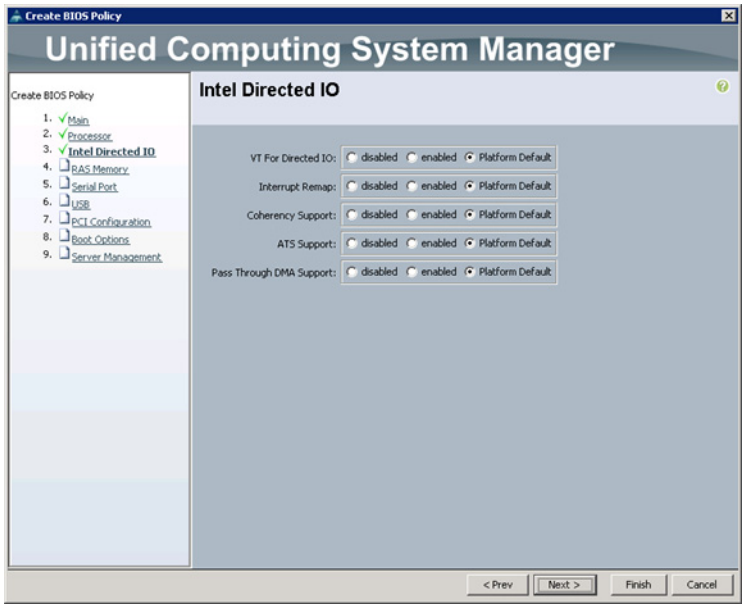
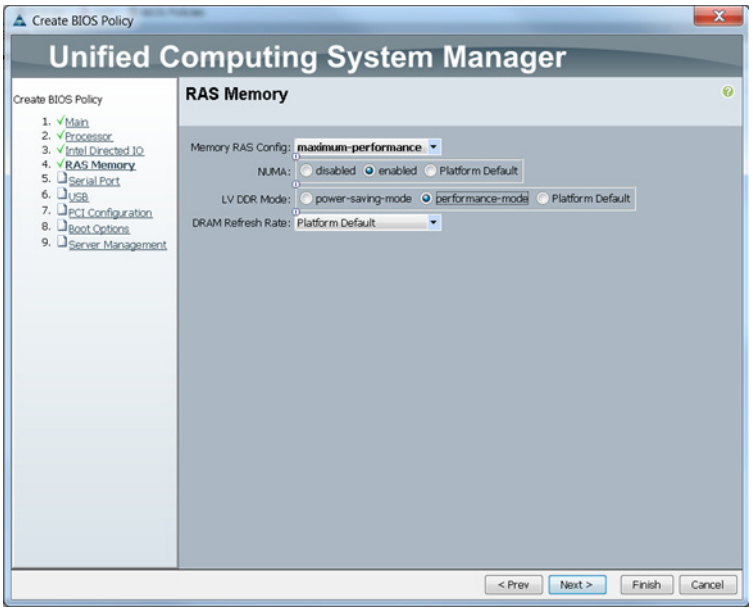


Figure 27 Creating Server BIOS Policy for Intel Directed IO



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

Figure 28 Creating Server BIOS Policy for Memory

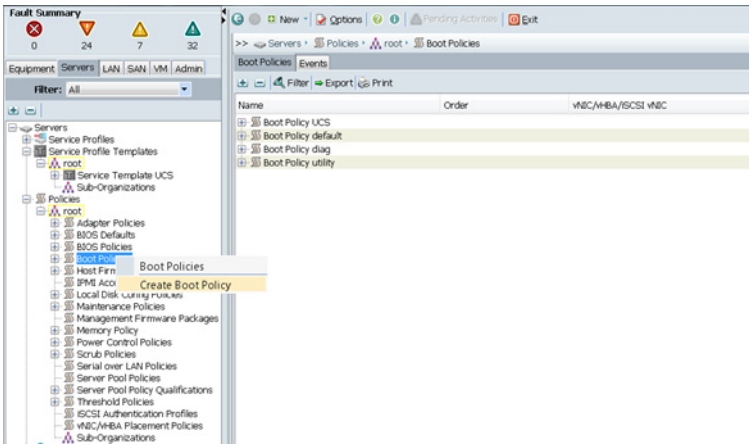


Creating a Boot Policy

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

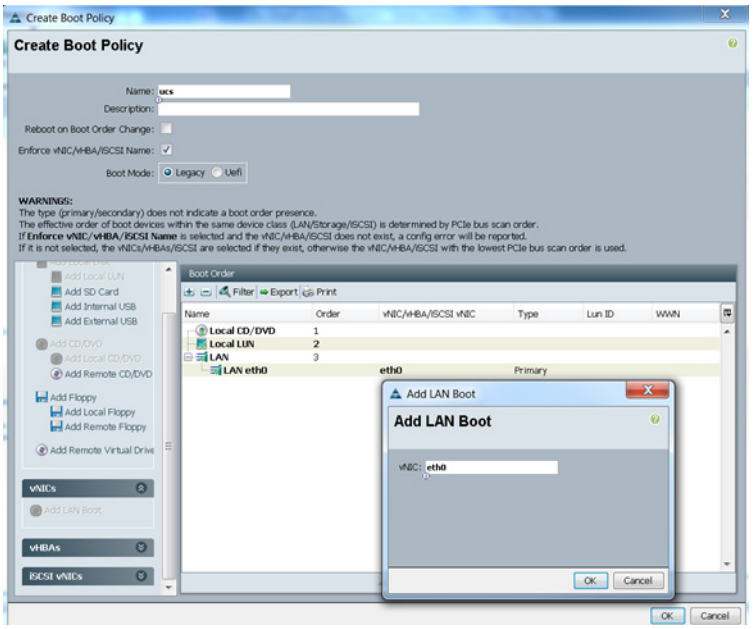
1. Select the `Servers` tab in the left pane in the Cisco UCS Manager GUI.
2. Select `Policies > root`.
3. Right-click the `Boot Policies`.
4. Select `Create Boot Policy`.

Figure 29 Creating Bot 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 > Add Local Disk and select Add Local LUN.
12. Expand vNICs and select Add LAN Boot and enter eth0.
13. Click OK to add the Boot Policy.
14. Click OK.

Figure 30 Creating Boot Policy Part 2

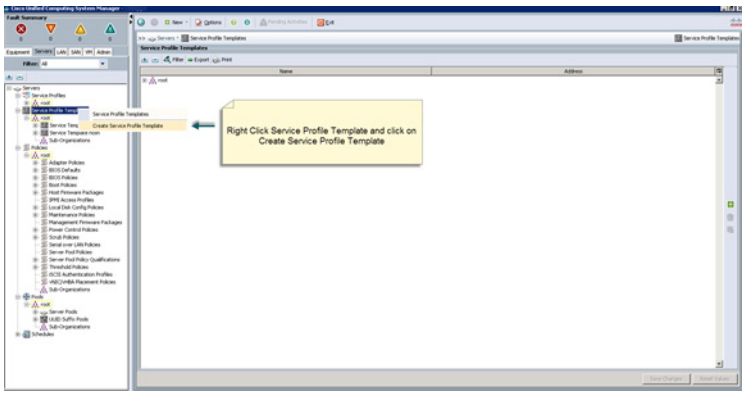


Creating a Service Profile Template

Follow these steps to create a service profile template:

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

Figure 31 Creating a Service Profile Template



The Create Service Profile Template window appears.

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

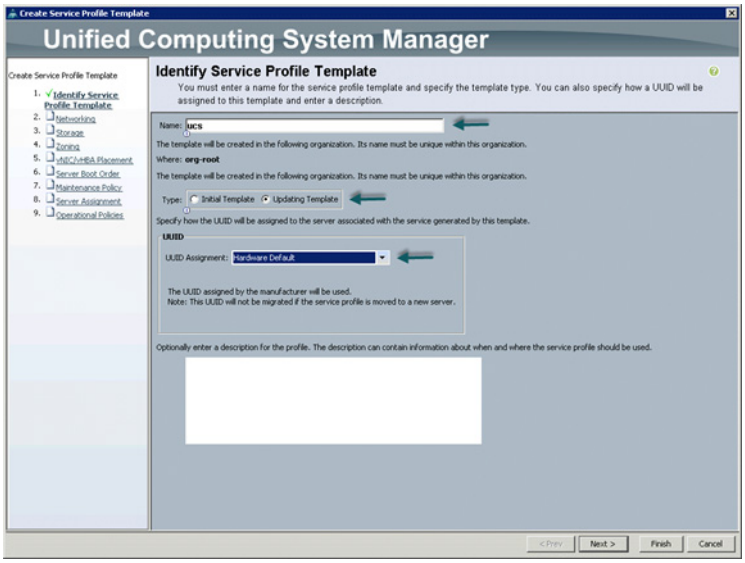
a. Name the service profile template as `ucs`.

b. Select the `Updating Template` radio button.

c. In the UUID section, select `Hardware Default` as the UUID pool.

d. Click `Next` to continue to the next section.

Figure 32 Identify Service Profile Template



Configuring Network Settings for a Template

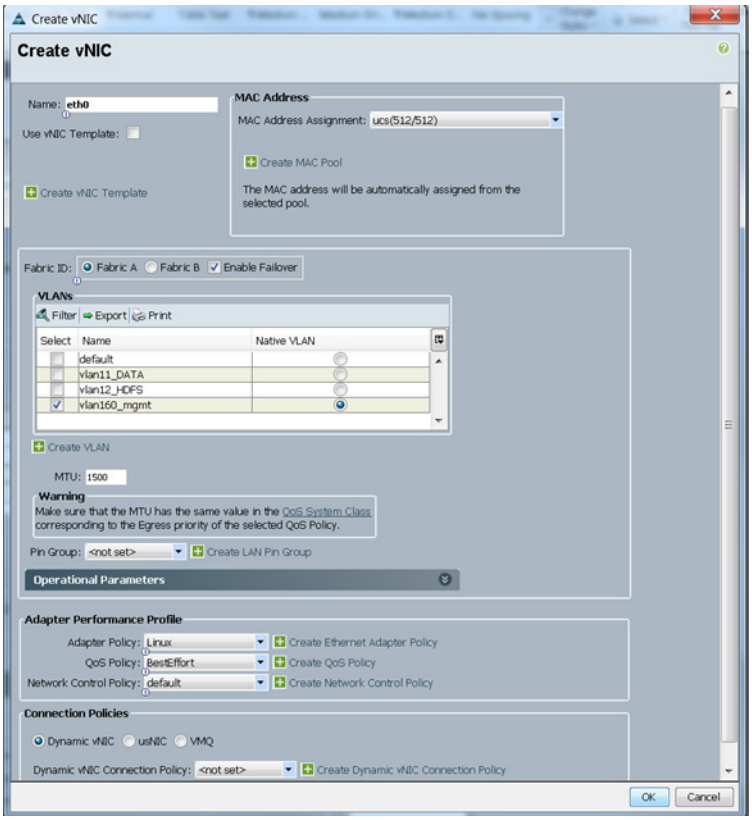
Follow these steps to configure the network settings for a template:

1. Keep the Dynamic vNIC Connection Policy field at the default.
2. Select 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 33 Configuring Network Settings for a Template

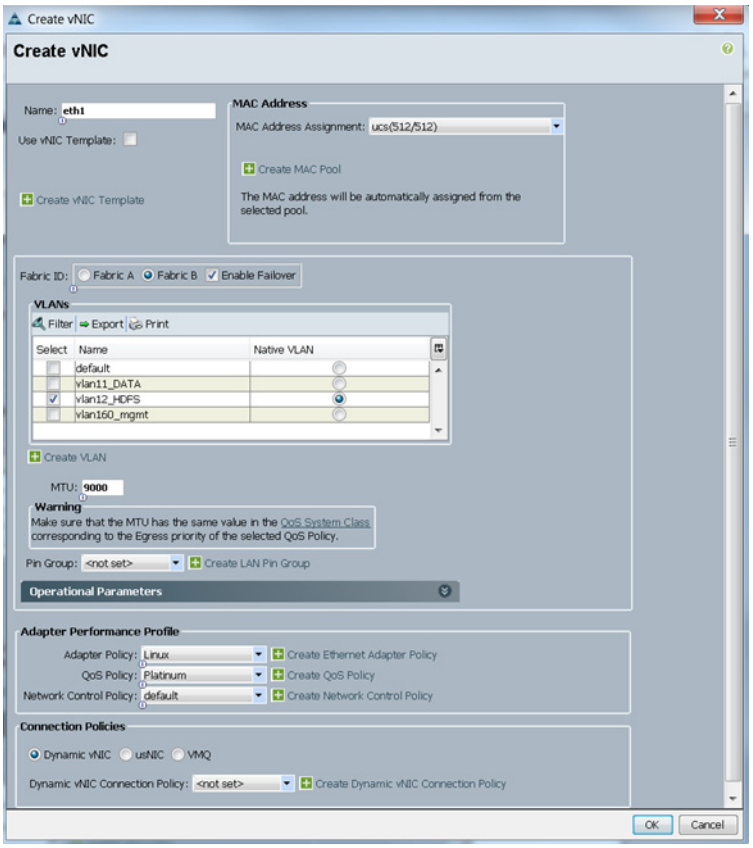
4. The Create vNIC window displays. Name the vNIC as eth0.
5. Select `ucs` in the Mac Address Assignment pool.
6. Select the Fabric A radio button and check the Enable failover check box for the Fabric ID.
7. Check the vlan160_mgmt check box for VLANs and select the Native VLAN radio button.
8. Select 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 34 Configuring vNIC eth0



15. The Create vNIC window appears. Name the vNIC eth1.
16. Select ucs in the Mac Address Assignment pool.
17. Select Fabric B radio button and check the Enable failover check box for the Fabric ID.
18. Check the vlan12_HDFS check box for VLANs and select the Native VLAN radio button
19. Select MTU size as 9000
20. Select adapter policy as Linux
21. Select QoS Policy as Platinum.
22. Keep the Network Control Policy as Default.
23. Keep the Connection Policies as Dynamic vNIC.
24. Keep the Dynamic vNIC Connection Policy as <not set>.
25. Click OK.

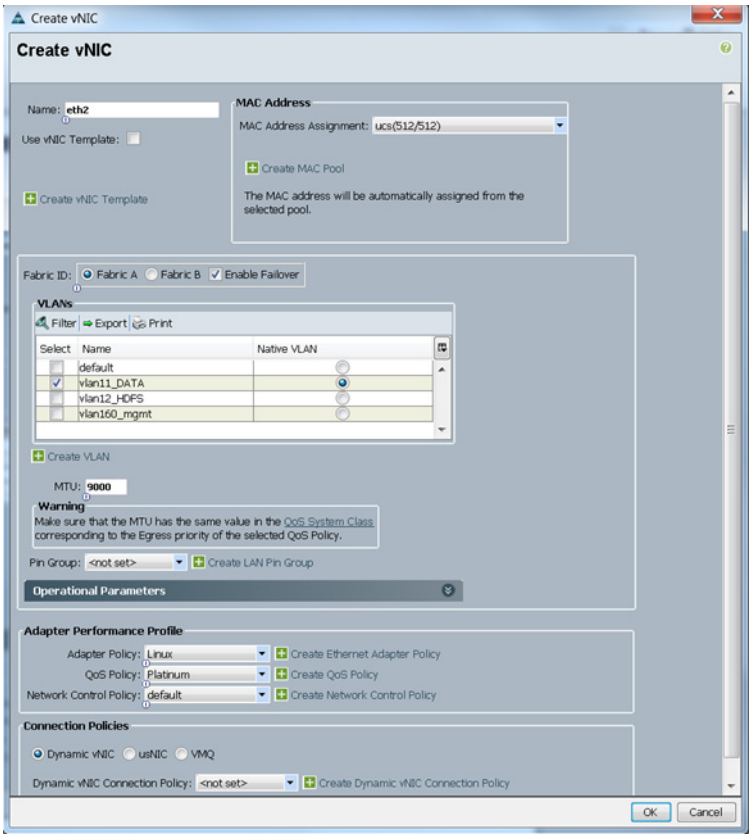
Figure 35 Configuring vNIC eth1



The Create vNIC window appears.

26. Name the vNIC eth2.
27. Select `ucs` in the Mac Address Assignment pool.
28. Select Fabric A radio button and check the Enable failover check box for the Fabric ID.
29. Check the vlan11_DATA check box for VLANs and select the Native VLAN radio button
30. Select MTU size as 9000
31. Select adapter policy as Linux
32. Select QoS Policy as Platinum.
33. Keep the Network Control Policy as Default.
34. Keep the Connection Policies as Dynamic vNIC.
35. Keep the Dynamic vNIC Connection Policy as <not set>.
36. Click OK.

Figure 36 Configuring vNIC eth2



Configuring a Storage Policy for a Template

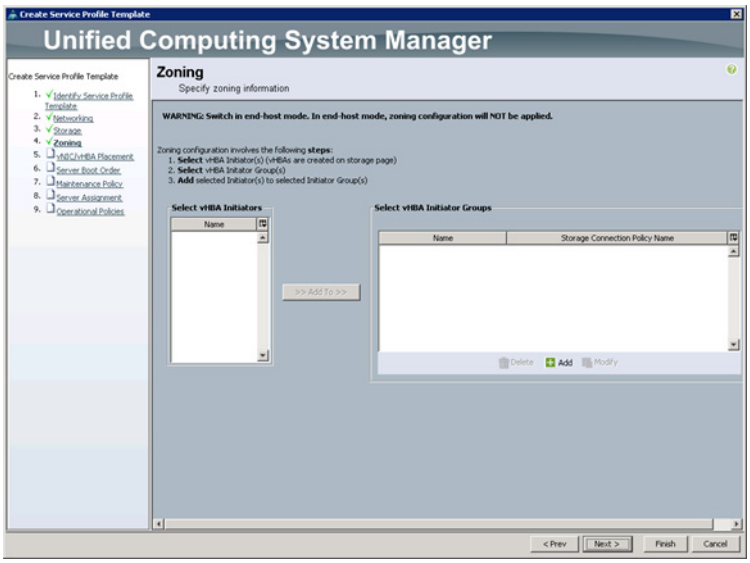
Follow these steps to configure storage policies:

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

Figure 37 Configuring Storage Settings

4. Click `Next` when the zoning window appears to go to the next section.

Figure 38 Configure Zoning



Configuring vNIC/vHBA Placement for a Template

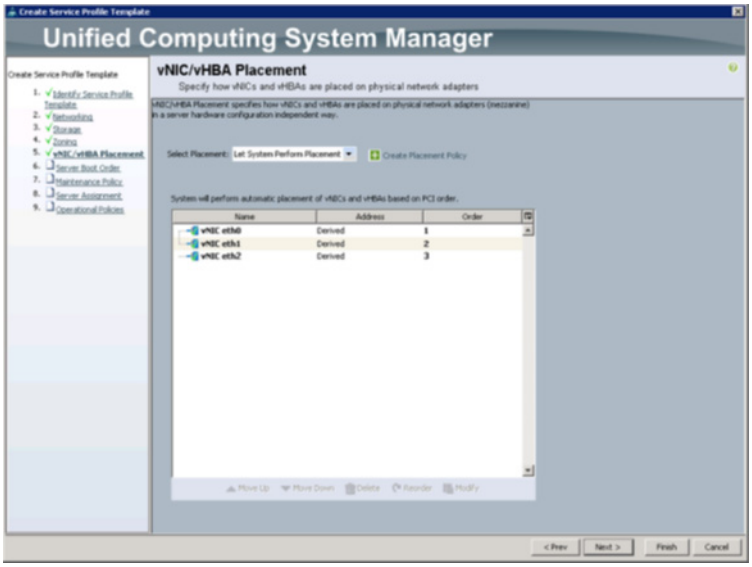
Follow these steps to configure vNIC/vHBA placement policy:

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 39 vNIC/vHBA Placement

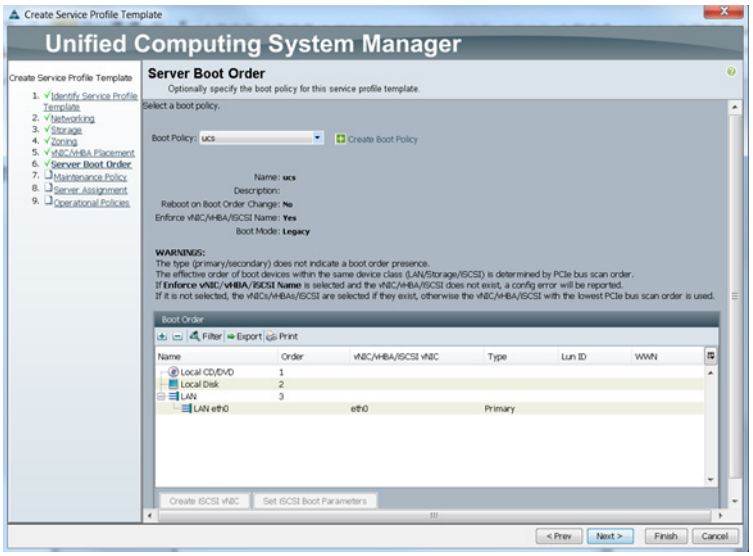


Configuring the Server Boot Order for a Template

Follow these steps to set the boot order for servers:

1. Select `ucs` in the Boot Policy name field.
2. Check the `Enforce vNIC/vHBA/iSCSI Name` check box.
3. Review to make sure that all of the boot devices were created and identified.
4. Verify that the boot devices are in the correct boot sequence.
5. Click `OK`.
6. Click `Next` to continue to the next section.

Figure 40 Creating Boot Policy



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

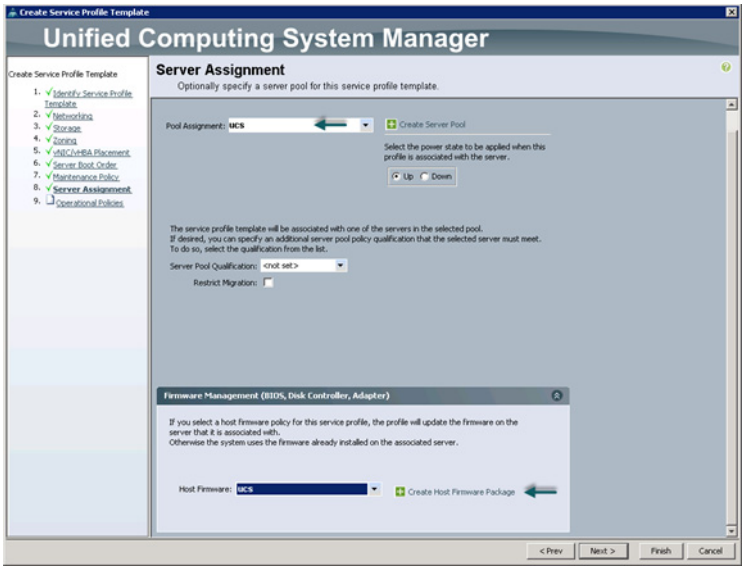
1. Keep the Maintenance policy at `no policy used` by default.
2. Click `Next` to continue to the next section.

Configuring the Server Assignment for a Template

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

1. Select `ucs` for the Pool Assignment field.
2. Keep the Server Pool Qualification field at `default`.
3. Select `ucs` in Host Firmware Package.

Figure 41 Server Assignment

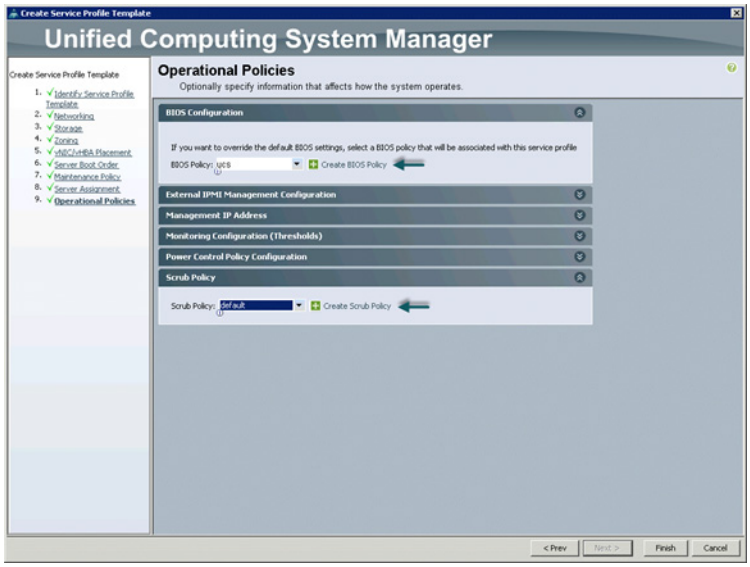


Configuring Operational Policies for a Template

In the Operational Policies Window, follow these steps:

1. Select `ucs` in the `BIOS Policy` field.
2. Click `Finish` to create the Service Profile template.
3. Click `OK` in the pop-up window to proceed.

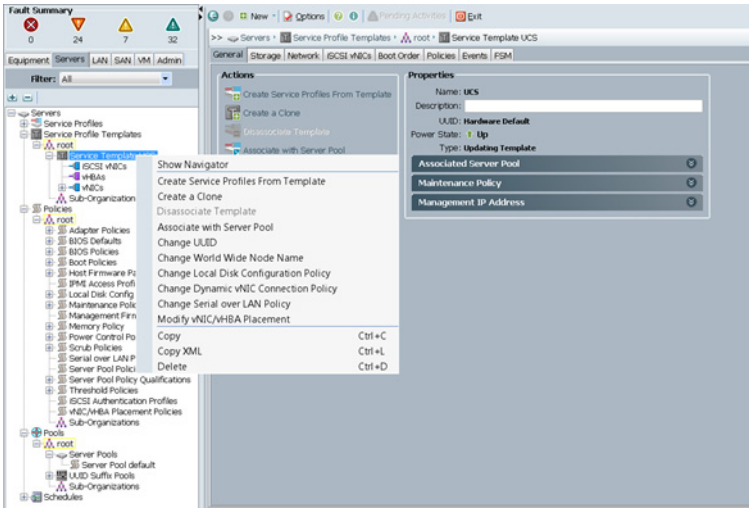
Figure 42 Selecting BIOS Policy



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

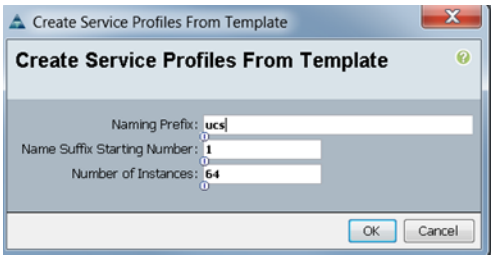
1. Go to `Service Profile Templates > root.`
2. Right-click `Service Profile Templates ucs.`
3. Select `Create Service Profiles From Template.`

Figure 43 Creating Service Profiles from Template



The Create Service Profile from Template window appears.

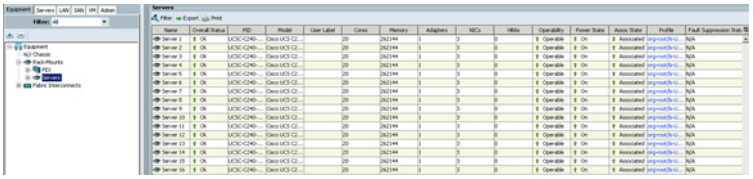
Figure 44 Selecting Name and Total Number of Service Profiles



Cisco UCS Manager will discover the servers. The association of the Service Profiles will take place automatically.

The Final Cisco UCS Manager window is shown in Figure 45.

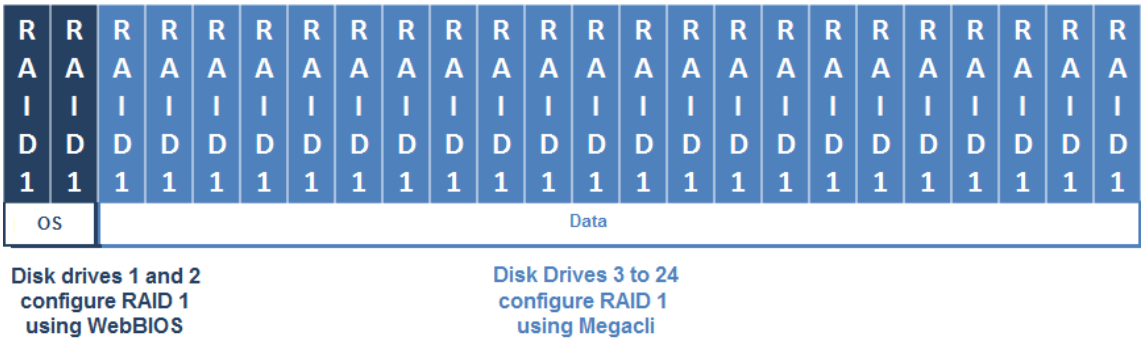
Figure 45 Cisco UCS Manager Displaying all Notes



Configuring Disk Drives for OS on Name Nodes

Namenode and Secondary Namenode have a different RAID configuration compared to Data nodes. This section details the configuration of disk drives for OS on these nodes (rhel1 and rhel2). The disk drives are configured as RAID1, read ahead cache is enabled and write cache is enabled while battery is present. The first two disk drives are used for operating system and remaining 22 disk drives are using for HDFS as described in the following sections.

There are several ways to configure RAID: using LSI WebBIOS Configuration Utility embedded in the MegaRAID BIOS, booting DOS and running MegaCLI commands, using Linux based MegaCLI commands, or using third party tools that have MegaCLI integrated. For this deployment, the first two disk drives are configured using LSI WebBIOS Configuration Utility and rests are configured using Linux based MegaCLI commands after the OS is installed.



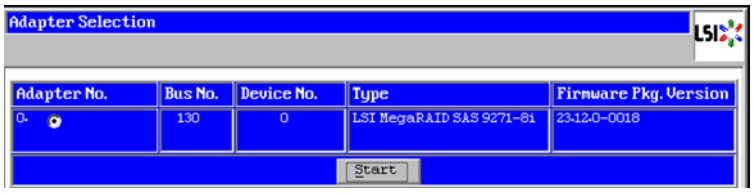
Follow these steps to create RAID1 on the first two disk drives to install the operating system:

1. When the server is booting, the following text appears on the screen:
2. Press `<Ctrl><H>` to launch the WebBIOS.
3. Press `Ctrl+H` immediately.

The Adapter Selection window appears.

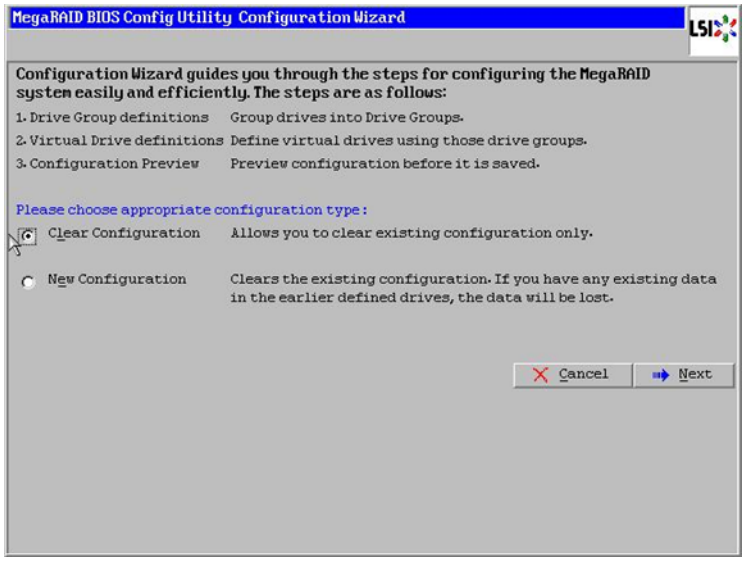
1. Click `Start` to continue.
2. Click `Configuration Wizard`.

Figure 46 Adapter Selection for RAID Configuration



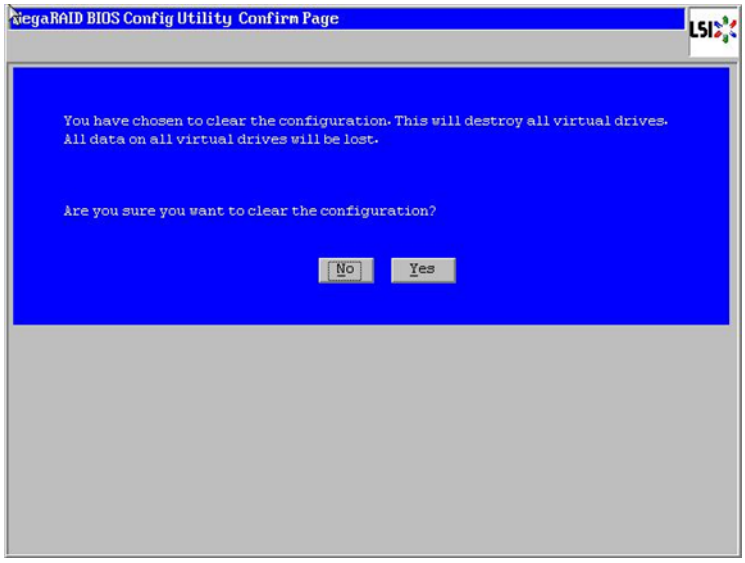
3. In the configuration wizard window, choose `Clear Configuration` and click `Next` to clear the existing configuration.

Figure 47 Clearing Current Configuration on the Controller



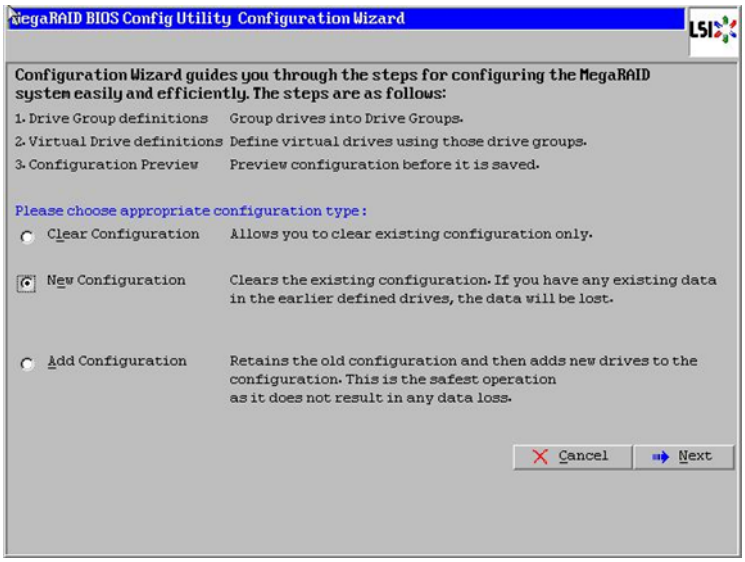
4. Choose `Yes` when asked to confirm the wiping of the current configuration.
5. In the Physical View, make sure that all the drives are `Unconfigured Good`.
6. Click `Configuration Wizard`.

Figure 48 Confirming Clearance of the Previous Configuration on the Controller



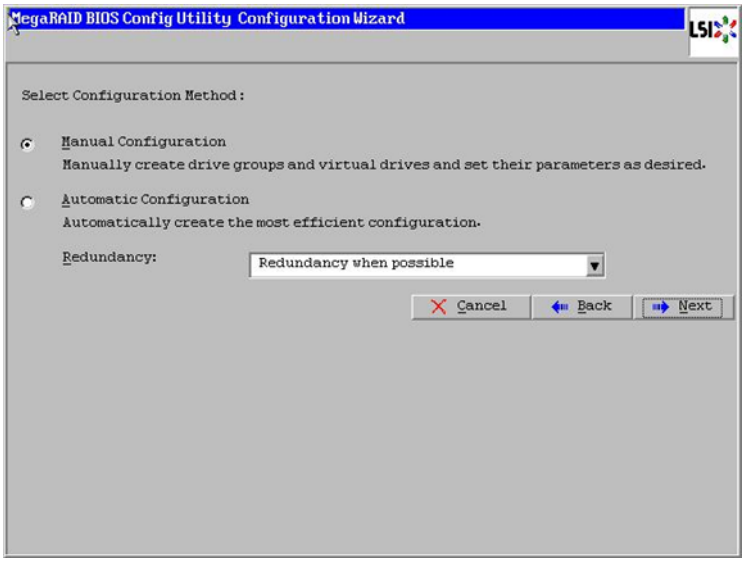
7. In the Configuration Wizard window choose the configuration type to be `New Configuration` and click `Next`.

Figure 49 Choosing to Create a New Configuration



8. Select the configuration method to be `Manual Configuration;` this enables you to have complete control over all attributes of the new storage configuration, such as, the drive groups, virtual drives and the ability to set their parameters.
9. Click `Next`.

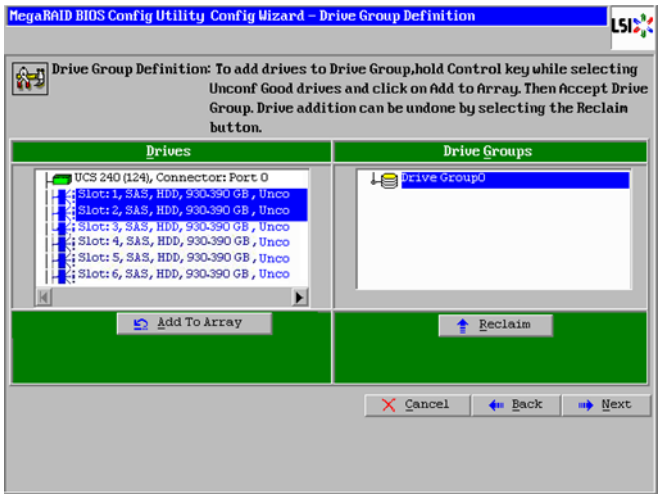
Figure 50 Choosing the Manual Configuration Method



The Drive Group Definition window appears. Use this window to choose the first two drives to create drive group.

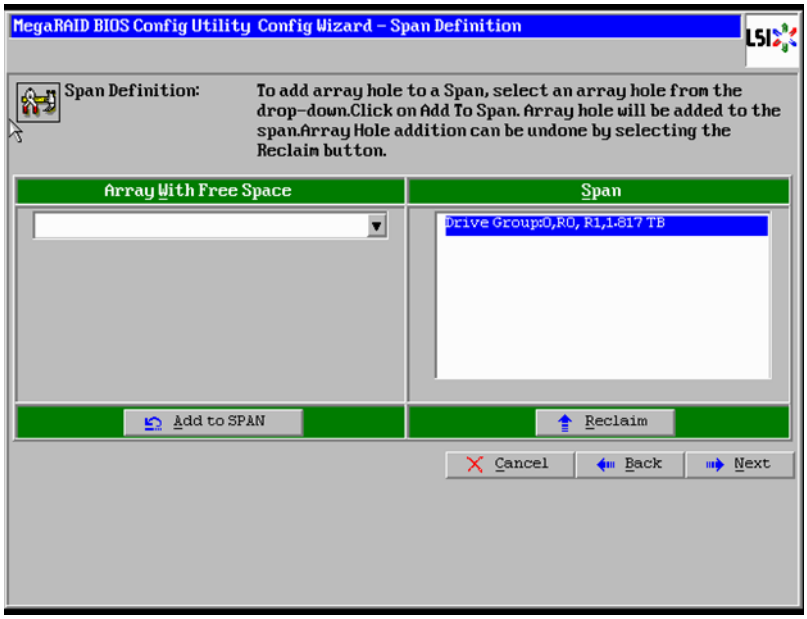
10. Click `Add to Array` to move the drives to a proposed drive group configuration in the Drive Groups pane.
11. Click `Accept DG` and then, click `Next`.

Figure 51 Selecting First Drive and Adding to the Drive Group



12. In the Span definitions Window, click `Add to SPAN` and click `Next`.

Figure 52 Span Definition Window

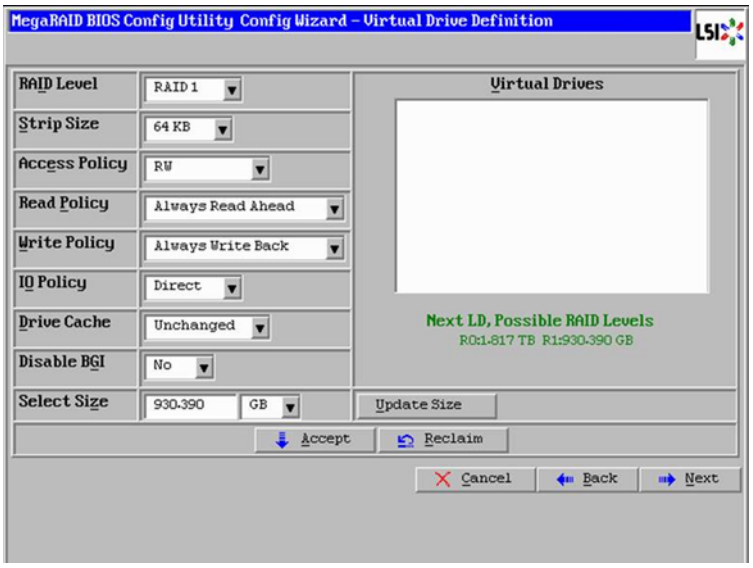


13. In the Virtual Drive definitions window;
    a. Click `Update Size`.
    b. Change `Strip Size` to 64 KB. A larger strip size produces higher read performance
    c. From the Read Policy drop-down list, choose `Always Read Ahead`.
    d. From the Write Policy drop-down list, choose `Write Back with BBU`.
    e. Make sure RAID Level is set to `RAID1`.
    f. Click `Accept` to accept the changes to the virtual drive definitions.

g. Click Next.

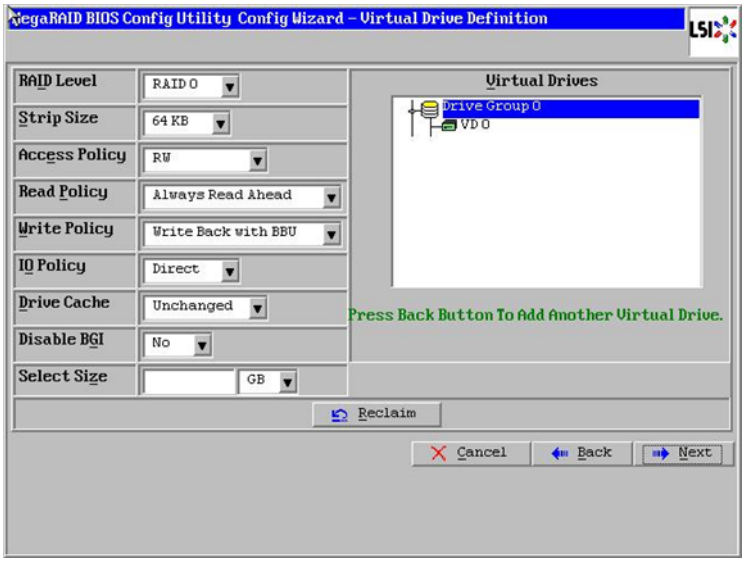
Note Clicking Update Size might change some of the settings in the window. Make sure all settings are correct before accepting.

Figure 53 Virtual Drive Definition Window



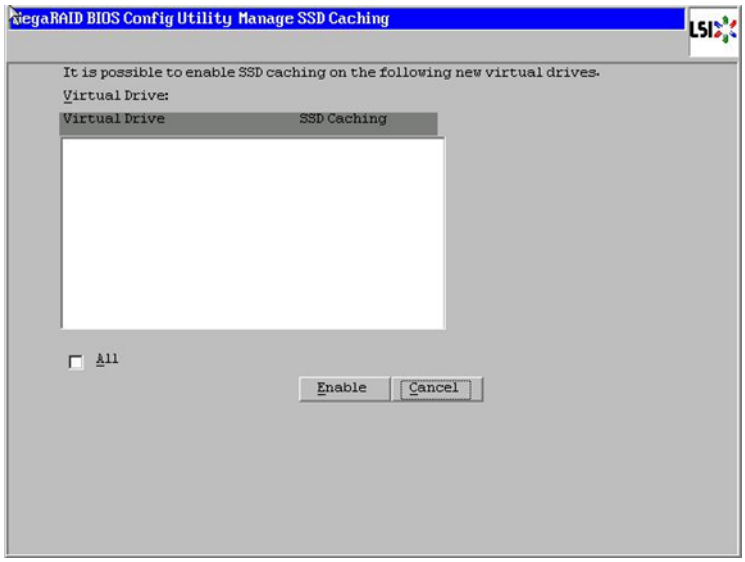
14. After you finish the virtual drive definitions, click `Next`. The Configuration Preview window appears showing VD0.
15. Check the virtual drive configuration in the Configuration Preview window and click `Accept` to save the configuration.

Figure 54 Completed Virtual Drive Definition



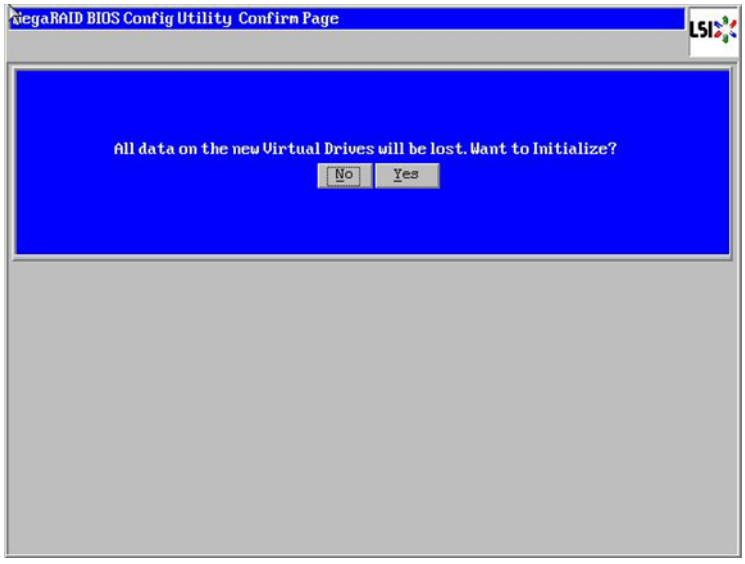
16. Click `Yes` to save the configuration.
17. In the managing SSD Caching Window, click `Cancel`.

Figure 55 SSD Caching Window



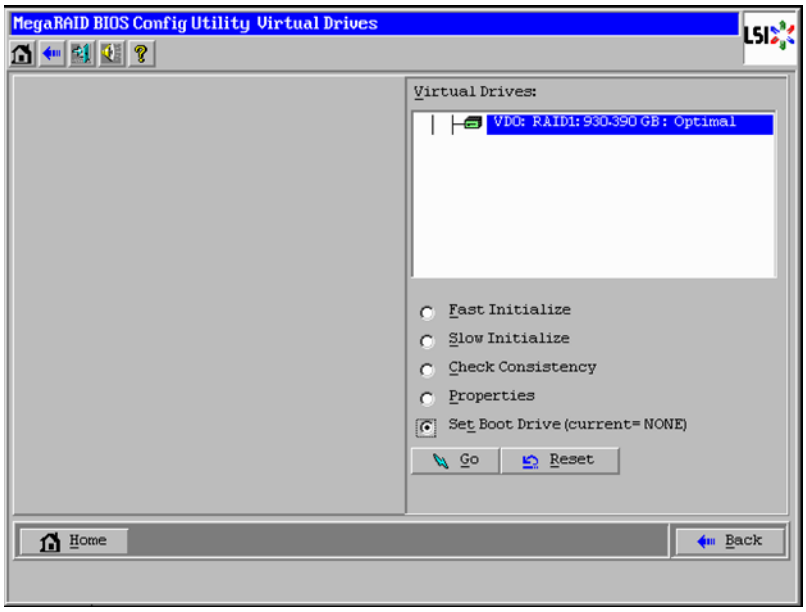
18. Click `Yes` when prompted to confirm the initialization.

Figure 56 Initializing Virtual Drive Window



19. Set VD0 as the Boot Drive and click `Go`.
20. Click `Home`.
21. Review the Configuration and click `Exit`.

Figure 57 Setting Virtual Drive as Boot Drive



Configuring disks 3-24 are done using Linux based MegaCLI command as described in the section about Configuring Data Drives for Namenode later in this document.

Configuring Disk Drives for OS on Data Nodes

Nodes 3 through 64 are configured as data nodes. This section details the configuration of disk drives for OS on the data nodes. As stated above, the focus of this CVD is the High Performance Configuration featuring 24 1TB SFF disk drives. The disk drives are configured as individual RAID0 volumes with 1MB stripe size. Read ahead cache is enabled and write cache is enabled while battery is present. The first disk drive is used for operating system and remaining 23 disk drives are using for HDFS as described in the following sections.

Note In the case of High Capacity Configuration featuring 12 4TB LFF disk drives, the disk drives are configured as individual RAID0 volumes with 1MB stripe size. Read ahead cached is enable and write cache is enabled while battery is present. Two partitions of 1TB and 3TB are created on the first disk drive, the 1TB partition is used for operating system and the 3TB partition is used for HDFS along with disk drives 2 through 12.

There are several ways to configure RAID: using LSI WebBIOS Configuration Utility embedded in the MegaRAID BIOS, booting DOS and running MegaCLI commands, using Linux based MegaCLI commands, or using third party tools that have MegaCLI integrated. For this deployment, the first disk drive is configured using LSI WebBIOS Configuration Utility and rest is configured using Linux based MegaCLI commands after the OS is installed.

Follow these steps to create RAID0 on the first disk drive to install the operating system:

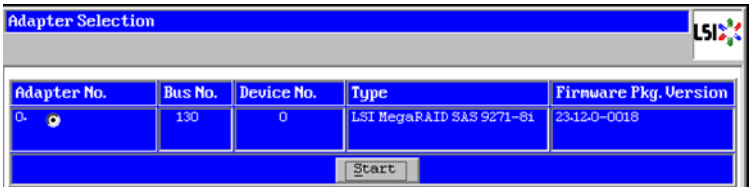
1. When the server is booting, the following text appears on the screen:
   a. Press <Ctrl><H> to launch the WebBIOS.

b. Press Ctrl+H immediately.

The Adapter Selection window appears.

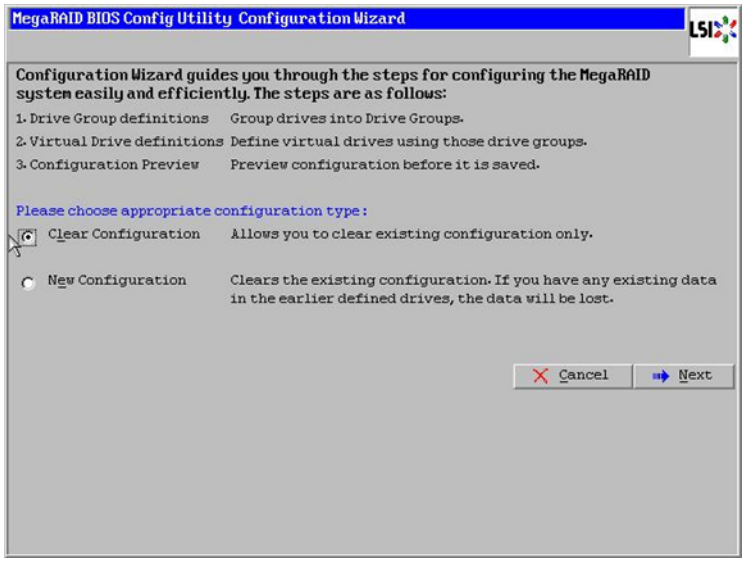
2. Click Start to continue.
3. Click Configuration Wizard.

Figure 58 Adapter Selection for RAID Configuration



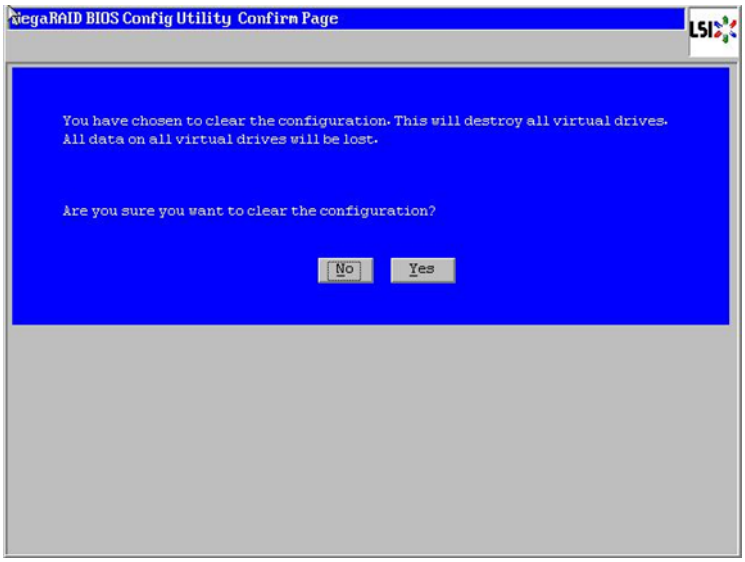
4. In the configuration wizard window, choose Clear Configuration and click Next to clear the existing configuration.

Figure 59 Clearing Current Configuration on the Controller



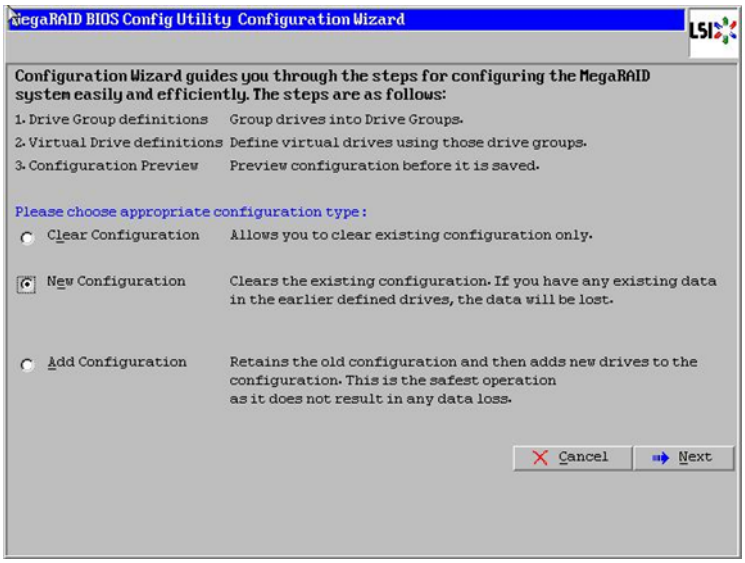
5. Choose Yes when asked to confirm the wiping of the current configuration.
6. In the Physical View, make sure that all the drives are Unconfigured Good.
7. Click Configuration Wizard.

Figure 60 Confirming Clearance of the Previous Configuration on the Controller



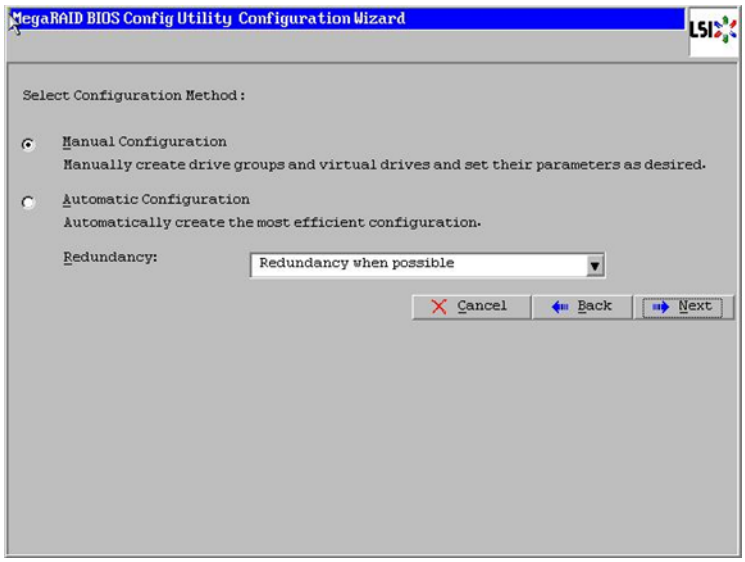
8. In the Configuration Wizard window choose the configuration type to be New Configuration and click Next.

Figure 61 Choosing to Create a New Configuration



9. Select the configuration method to be Manual Configuration; this enables you to have complete control over all attributes of the new storage configuration, such as, the drive groups, virtual drives and the ability to set their parameters.
10. Click Next.

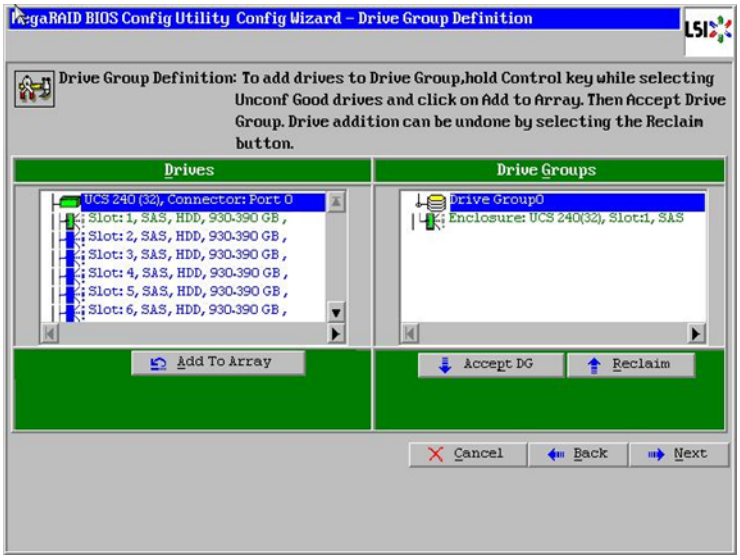
Figure 62 Choosing Manual Configuration Method



The Drive Group Definition window appears. Use this window to choose the first drive to create drive groups.

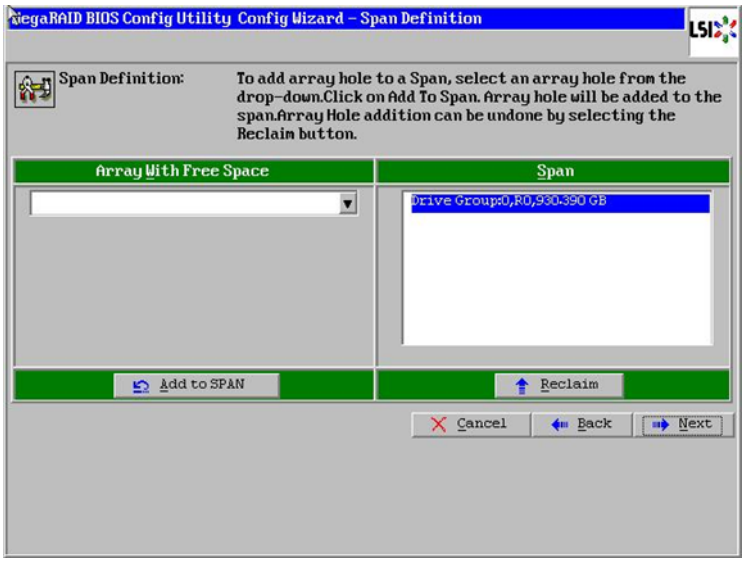
11. Click Add to Array to move the drives to a proposed drive group configuration in the Drive Groups pane.
12. Click Accept DG.
13. Click Next.

Figure 63 Selecting the First Drive and Adding to Drive Group



14. In the Span definitions Window, click Add to SPAN and click Next.

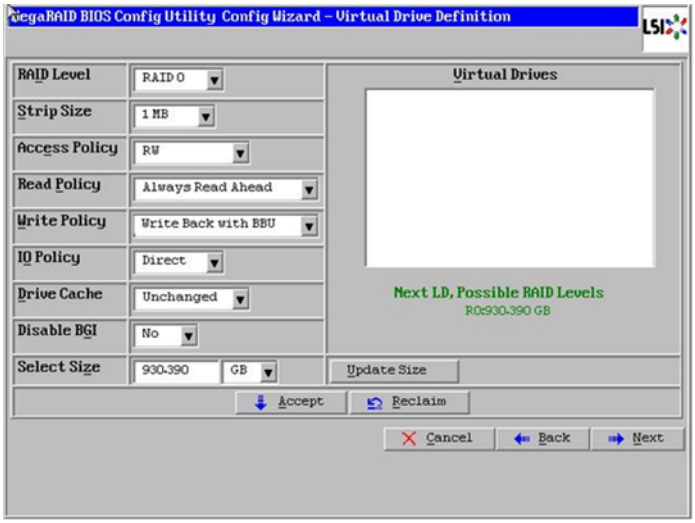
Figure 64 Span Definition Window



15. In the Virtual Drive definitions window;
    a. Click on Update Size.
    b. Change Strip Size to 1MB. A larger strip size produces higher read performance
    c. From the Read Policy drop-down list, choose Always Read Ahead.
    d. From the Write Policy drop-down list, choose Write Back with BBU.
    e. Make sure RAID Level is set to RAID0.
    f. Click Accept to accept the changes to the virtual drive definitions.
    g. Click Next.

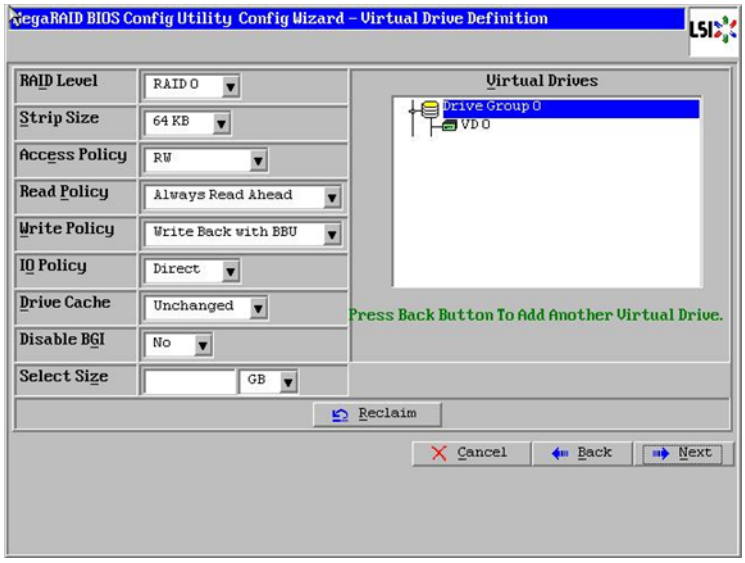
Note Clicking Update Size might change some of the settings in the window. Make sure all settings are correct before accepting.

Figure 65 Virtual Drive Definition Window



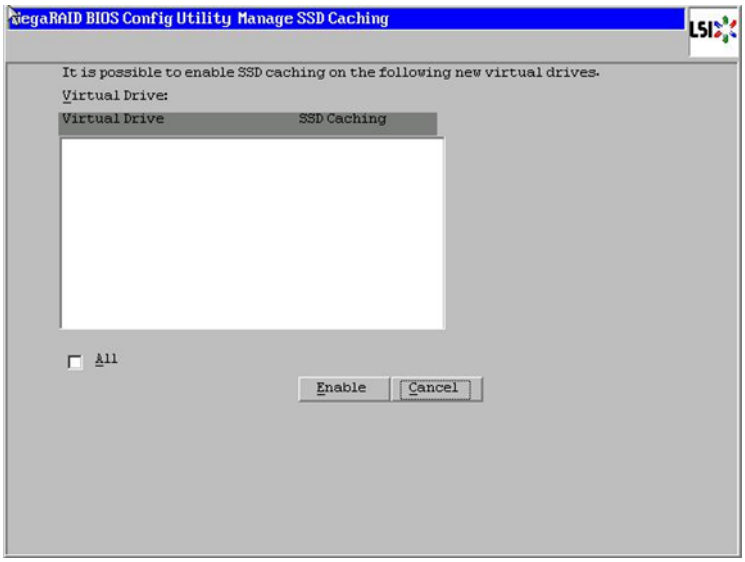
16. After you finish the virtual drive definitions, click Next. The Configuration Preview window appears showing VD0.
17. Check the virtual drive configuration in the Configuration Preview window and click Accept to save the configuration.

Figure 66 Completed Virtual Drive Definition



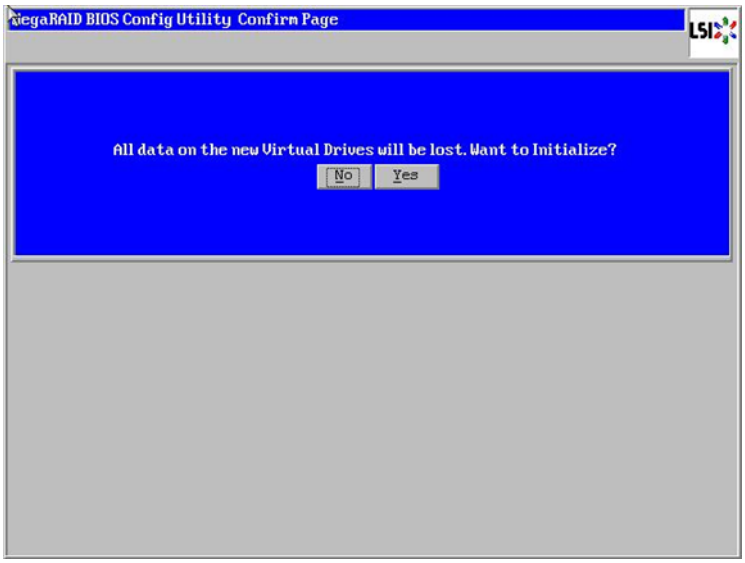
18. Click Yes to save the configuration.
19. In the managing SSD Caching Window, click Cancel.

Figure 67 SSD Caching Window



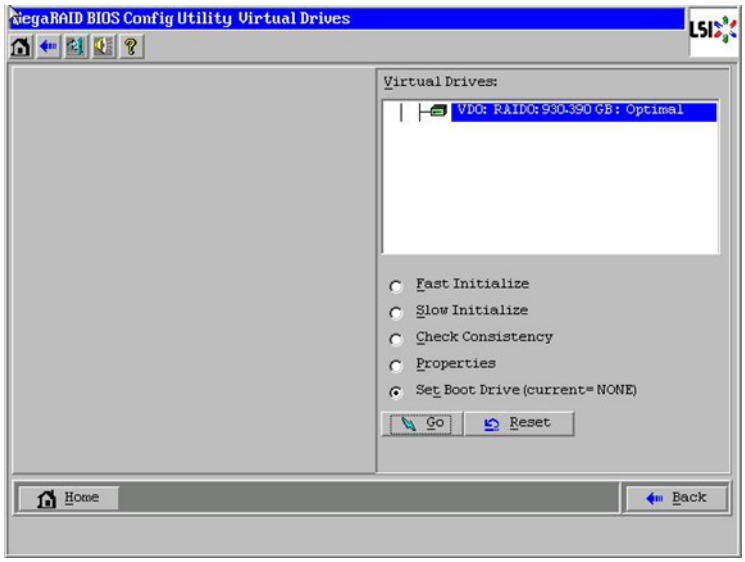
20. Click Yes.when prompted to confirm the initialization.

Figure 68 Initializing Virtual Drive Window



21. Set VD0 as the Boot Drive and click Go.

Figure 69 Setting Virtual Drive as Boot Drive



22. Click Home.
23. Review the Configuration and Click Exit.

The steps above can be repeated to configure disks 2-24 or using Linux based MegaCLI commands as described in the section about Configuring Data Drives later in this document.

Installing Red Hat Linux 6.4 with KVM

The following section provides detailed procedures for installing Red Hat Linux 6.4.

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.

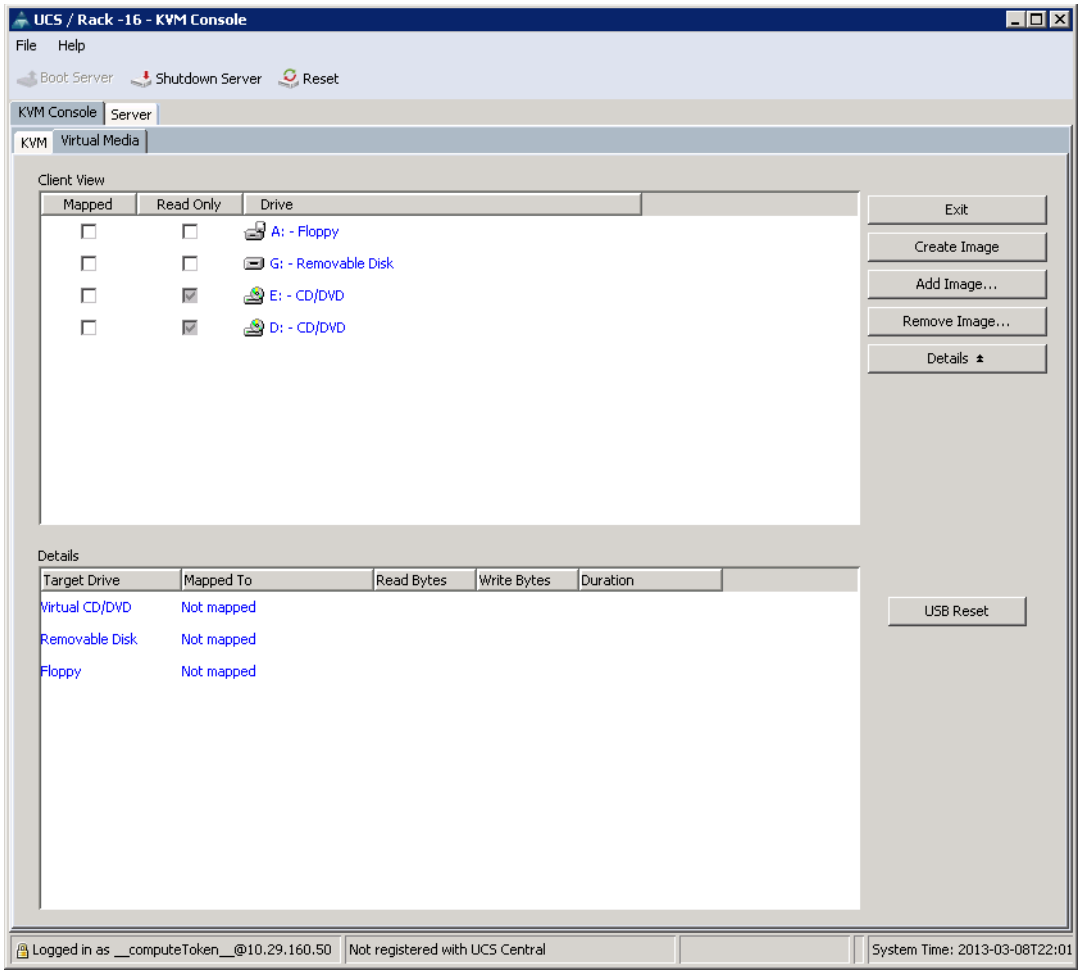
1. Log in to the Cisco UCS 6296 Fabric Interconnect and launch the Cisco UCS Manager application.
2. Select the Equipment tab.
3. In the navigation pane expand Rack-mounts and Servers.
4. Right click the server and select KVM Console.

Figure 70 Selecting KVM Console Option

5. In the KVM window, select the Virtual Media tab.
6. Click the Add Image button found in the right hand corner of the Virtual Media selection window.
7. Browse to the Red Hat Enterprise Linux Server 6.4 installer ISO image file.

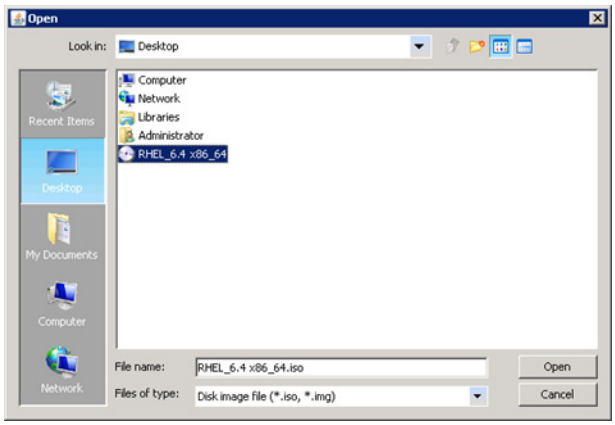
Note The Red Hat Enterprise Linux 6.4 DVD is assumed to be on the client machine.

Figure 71 Adding an ISO Image



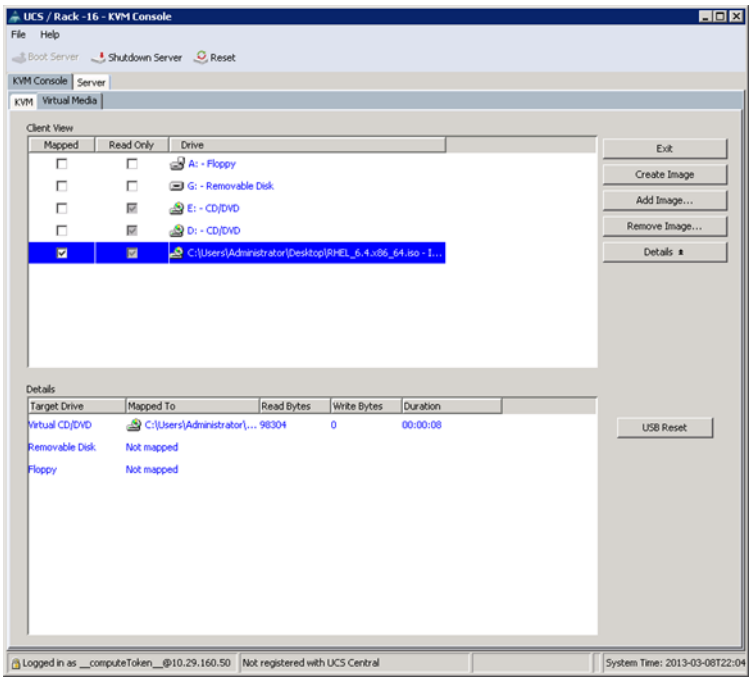
8. Click Open to add the image to the list of virtual media.

Figure 72 Browse to Red Hat Enterprise Linux ISO Image



9. Check the check box for Mapped, next to the entry corresponding to the image you just added.In the KVM window, select the KVM tab to monitor during boot.
10. In the KVM window, select the Boot Server button in the upper left corner.
11. Click OK.
12. Click OK to reboot the system.

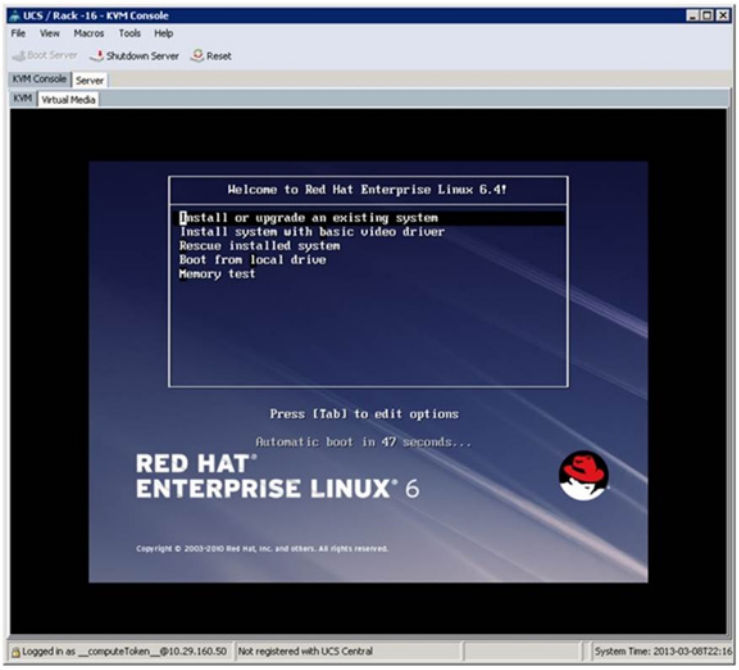
Figure 73 Mapping ISO Image



On reboot, the machine detects the presence of the Red Hat Enterprise Linux Server 6.4 install media.

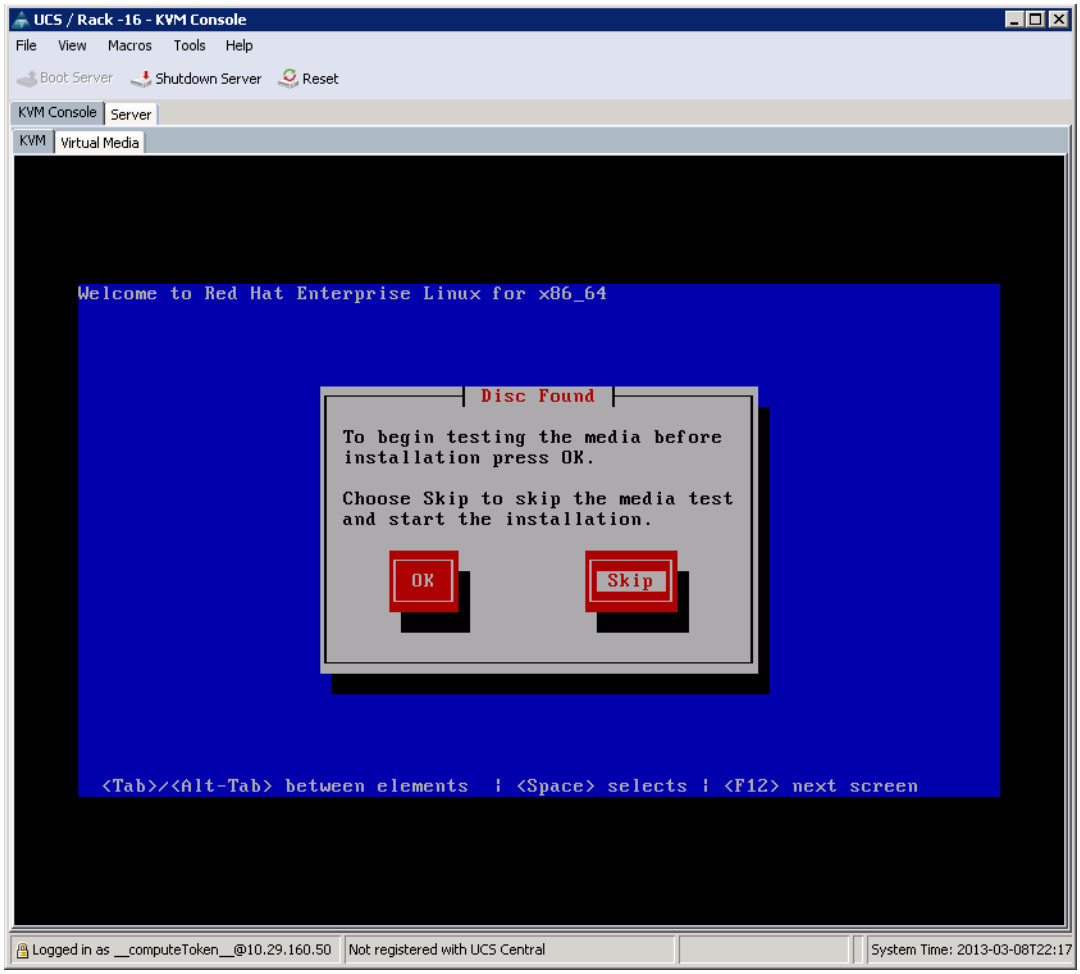
13. Select the Install or Upgrade an Existing System.

Figure 74 Select Install Option



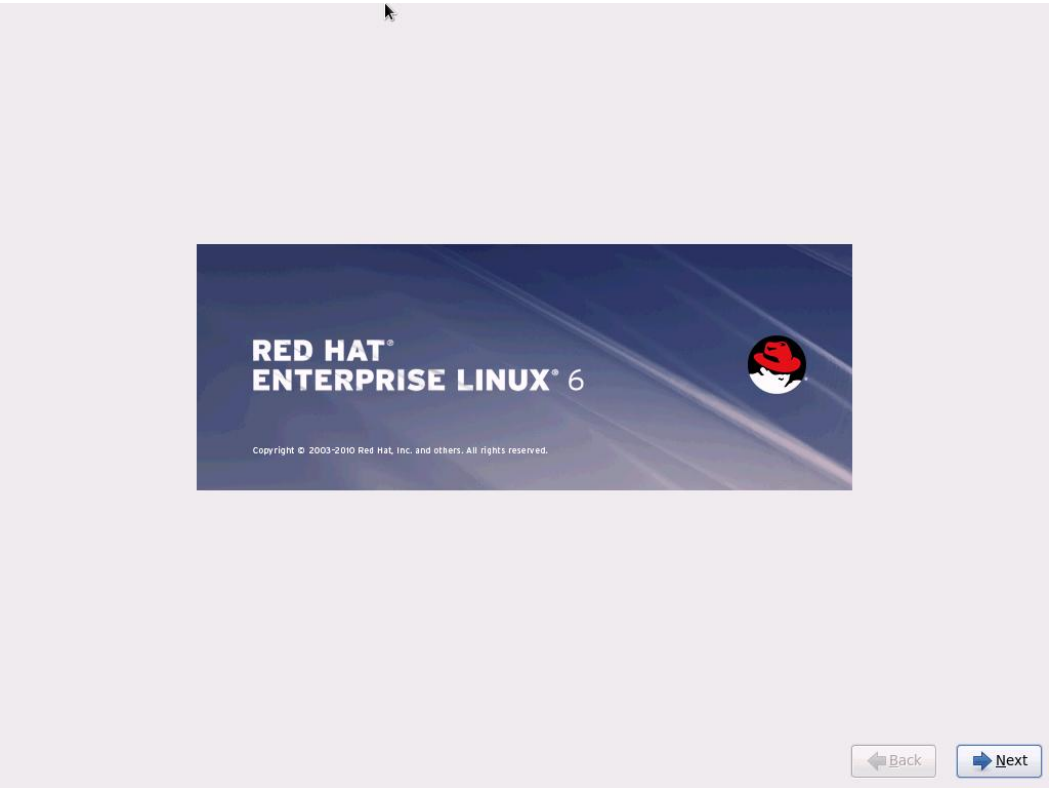
14. Skip the Media test since you are installing from an ISO Image.

Figure 75 Skipping Media Test



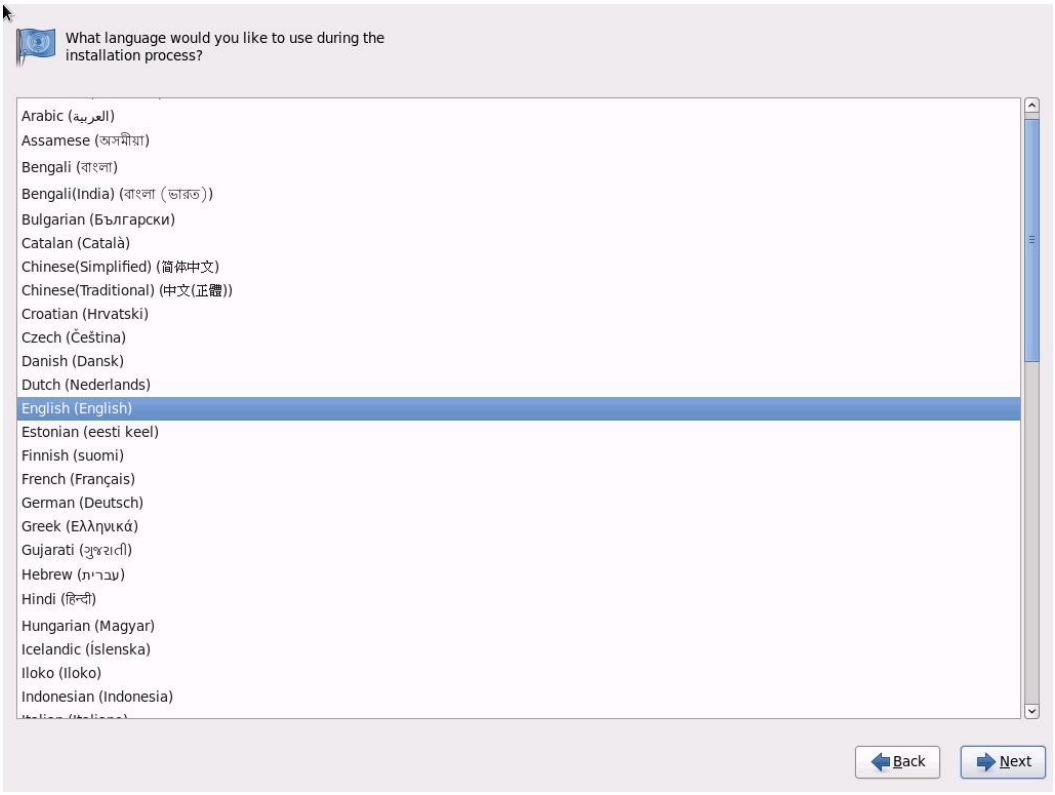
15. Click Next.

Figure 76 Red Hat Linux Welcome Screen



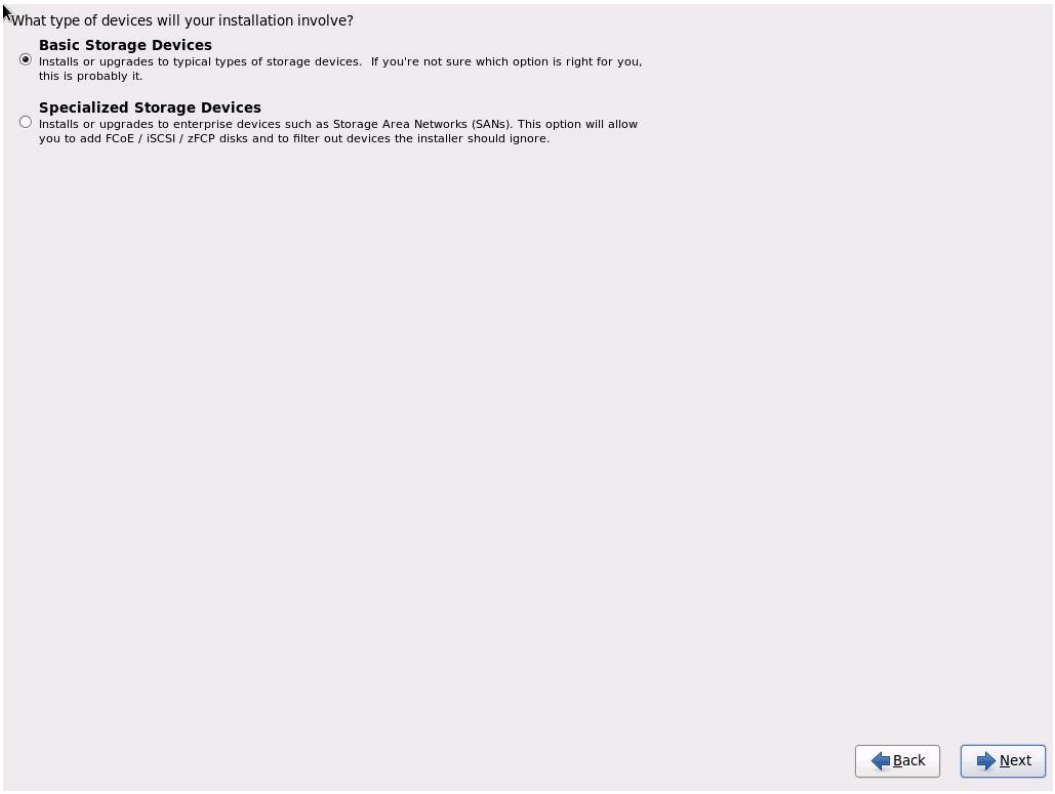
16. Select the Language for the Installation

Figure 77 Selecting Language for the Installation



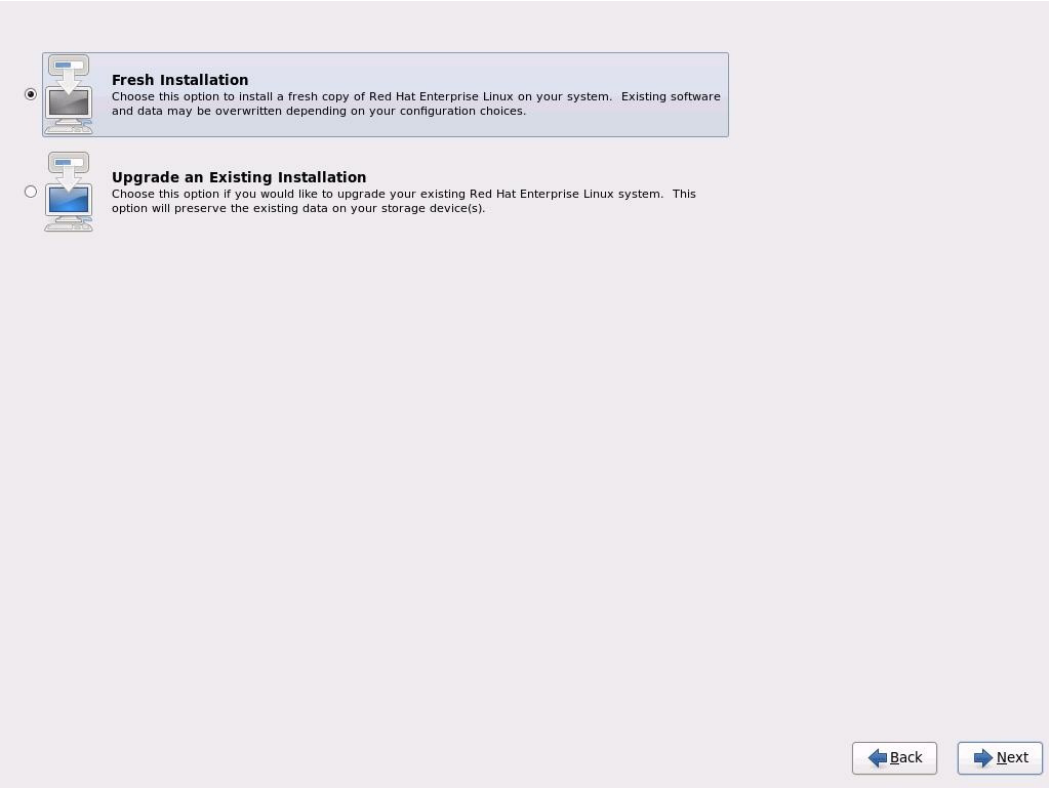
17. Select Basic Storage Devices.

Figure 78 Selecting Basic Storage Devices



18. Select Fresh Installation.
19. Enter the Host name of the server and click Next.

Figure 79 Selecting Fresh Installation

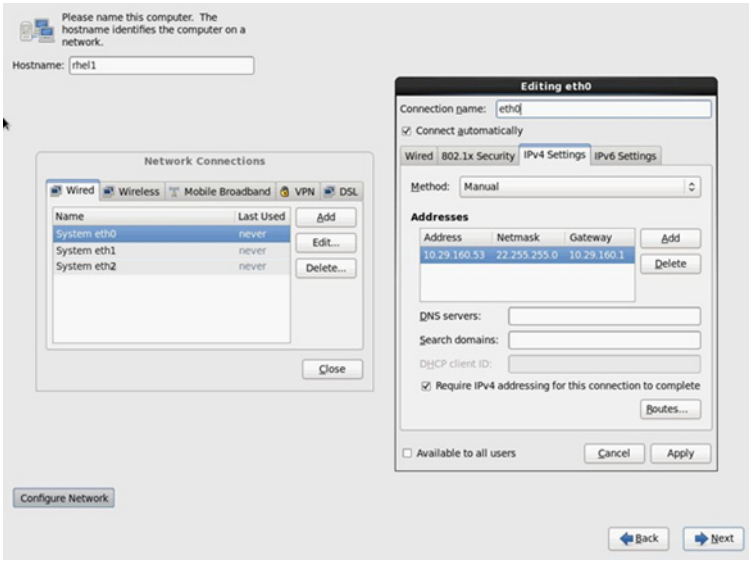


20. Click Configure Network. The Network Connections window should appear.
21. In the Network Connections window, select the Wired tab.
22. Select the interface System eth0 and click Edit.

Editing System eth0 appears.

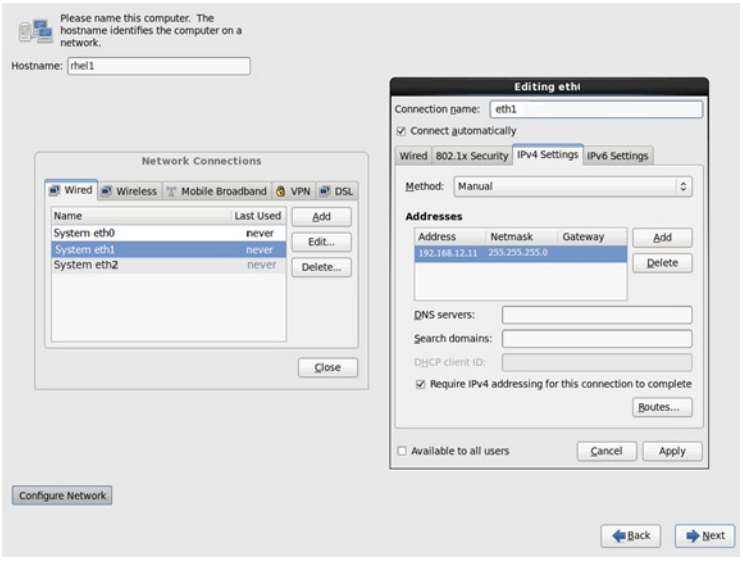
23. Check the check box Connect automatically.
24. In the drop down menu select Manual Method.
25. Click Add and enter IP Address, Netmask and the Gateway.
26. For this demonstration use the following:
    a. IP Address: 10.29.160.53
    b. Netmask: 255.255.255.0
    c. Gateway: 10.29.160.1
27. Add DNS servers (optional).
28. Click Apply.

Figure 80 Configuring Network for eth0



29. Repeat the steps 26 to steps 28 for system eth1 with the following:
    - IP Address: 192.168.12.11
    - Netmask: 255.255.255.0

Figure 81 Configuring Network for eth1

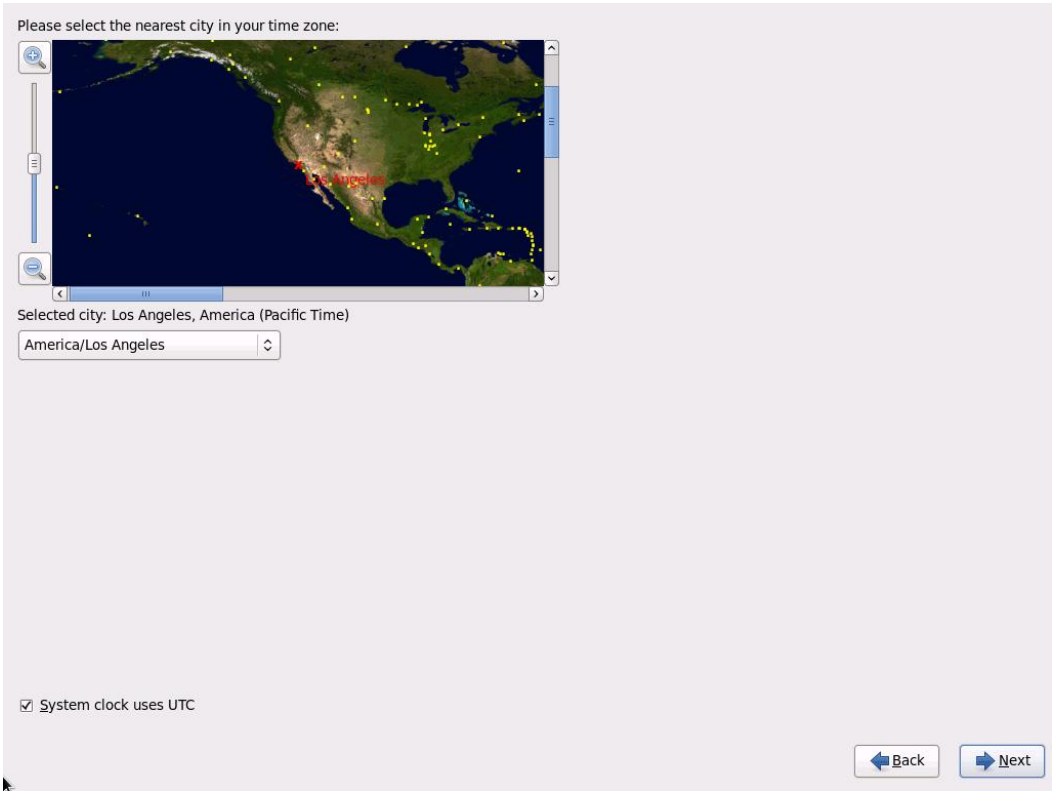


30. Repeat the steps 26 to steps 28 for system eth2 with the following:
    - IP Address: 192.168.11.11
    - Netmask: 255.255.255.0

Note Table 6 lists the IP address of the nodes in the cluster.

31. Select the appropriate Time Zone.

Figure 82 Select Time Zone

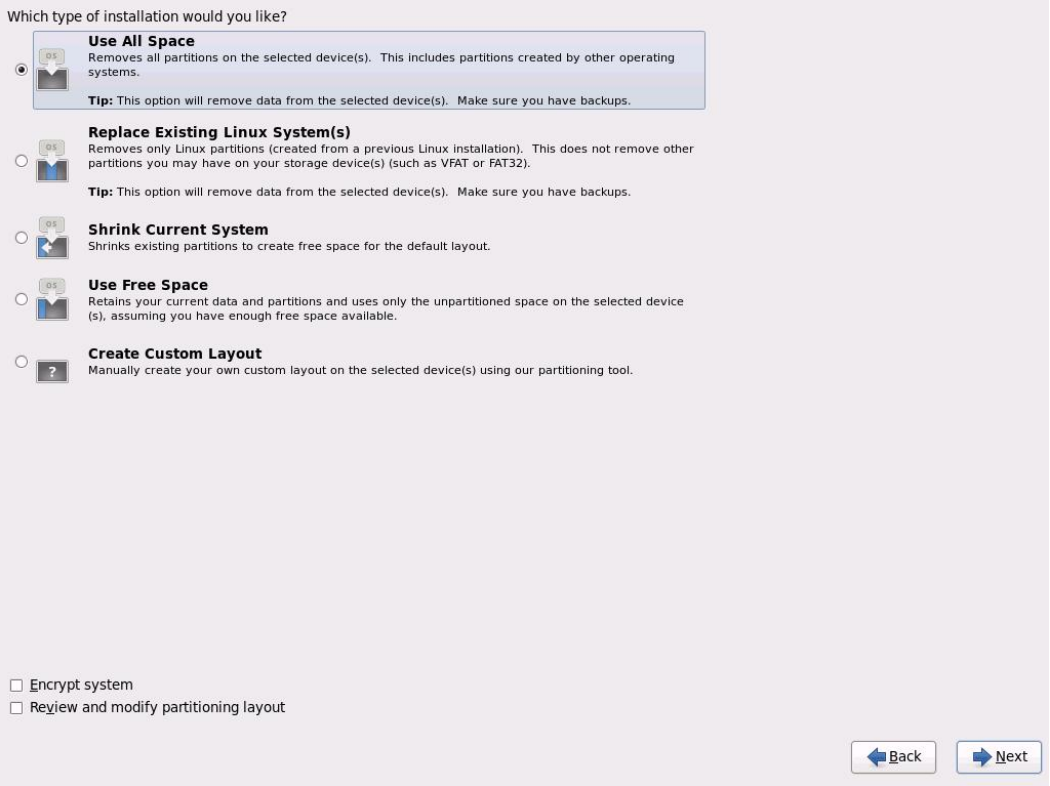


32. Enter the root Password and click Next.

Figure 83 Entering the Root Password

33. Select Use All Space and click Next.
34. Choose an appropriate boot drive.

Figure 84 Selecting Install Options

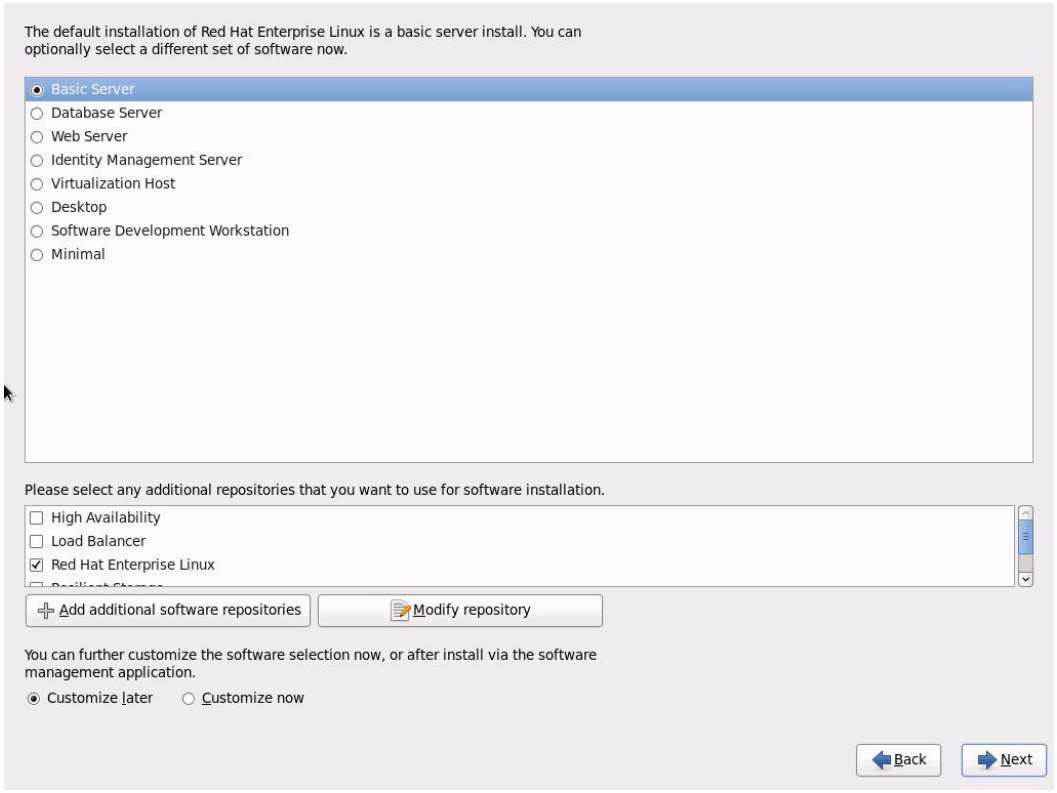


35. Click Write changes to disk and click Next.

Figure 85 Confirming Formatting of Disk

36. Select Basic Server Installation and click Next.

Figure 86 Selecting Type of Installation



After the installer is finished loading, it will continue with the installation as shown in the figures below:

Figure 87 Starting the Installation

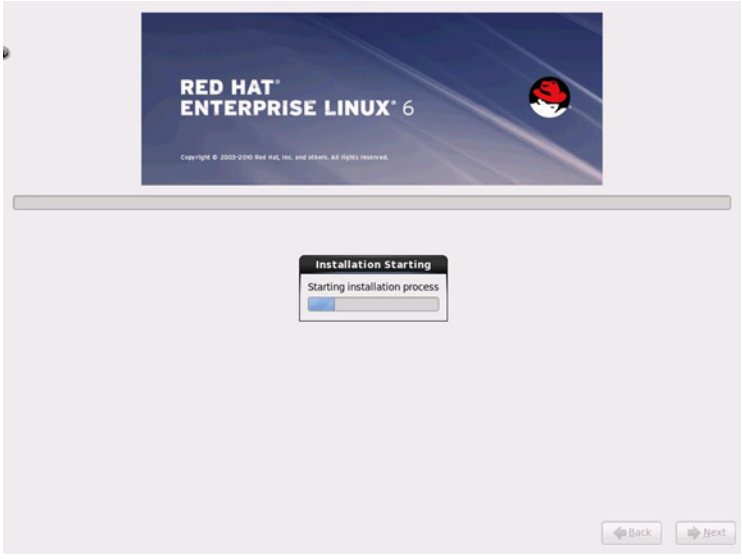
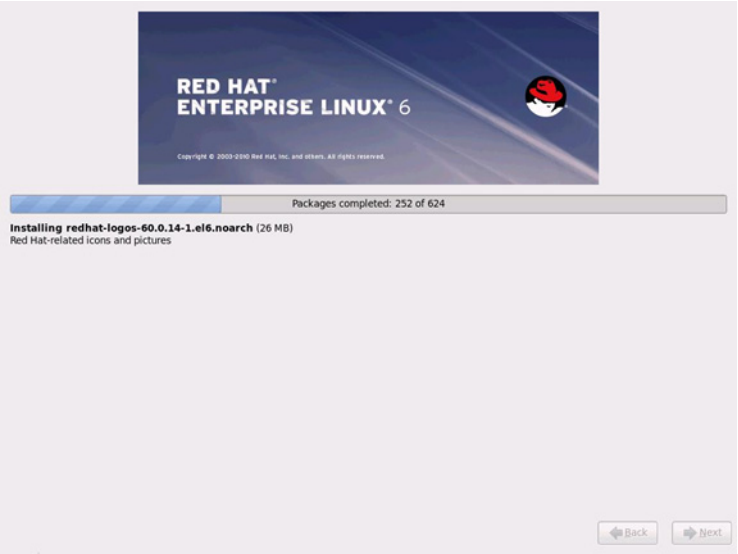


Figure 88 Installation in Progress



37. When the installation is complete, reboot the system.
38. Repeat steps 1 through 37 to install Red Hat Linux 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 hostnames and their corresponding IP addresses are shown in Table 6.

Table 6 Hostnames and IP Addresses

Hostname	eth0	eth1	eth2
rhel1	10.29.160.53	192.168.12.11	192.168.11.11
rhel2	10.29.160.54	192.168.12.12	192.168.11.12
rhel3	10.29.160.55	192.168.12.13	192.168.11.13
rhel4	10.29.160.56	192.168.12.14	192.168.11.14
rhel5	10.29.160.57	192.168.12.15	192.168.11.15
rhel6	10.29.160.58	192.168.12.16	192.168.11.16
rhel7	10.29.160.59	192.168.12.17	192.168.11.17
rhel8	10.29.160.60	192.168.12.18	192.168.11.18
rhel9	10.29.160.61	192.168.12.19	192.168.11.19
rhel10	10.29.160.62	192.168.12.20	192.168.11.20
rhel11	10.29.160.63	192.168.12.21	192.168.11.21
rhel12	10.29.160.64	192.168.12.22	192.168.11.22
rhel13	10.29.160.65	192.168.12.23	192.168.11.23
rhel14	10.29.160.66	192.168.12.24	192.168.11.24

rhel15	10.29.160.67	192.168.12.25	192.168.11.25
rhel16	10.29.160.68	192.168.12.26	192.168.11.26
…	…	…	…
rhel64	10.29.160.116	192.168.12.74	192.168.11.74

Post OS Install Configuration

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

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 Parallel-SSH (pssh) and shell-scripts without having to use passwords.

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

1. Login to the Admin Node (rhel1)

   ssh 10.29.160.53

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. 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 {53..116}; do echo -n "$IP -> "; ssh-copy-id -i ~/.ssh/id_rsa.pub 10.29.160.$IP; done

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

5. Enter the password of the remote host.

Installing and Configuring Parallel Shell

PARALLEL-SSH

Parallel SSH is used to run commands on several hosts at the same time. It takes a file of hostnames and a bunch of common ssh parameters as parameters, executes the given command in parallel on the nodes specified.

1. From the system that is connected to the Internet, download pssh.

```
wget https://parallel-ssh.googlecode.com/files/pssh-2.3.1.tar.gz
```

```
[root@redhat ~]# wget  https://parallel-ssh.googlecode.com/files/pssh-2.3.1.tar.gz
--2013-04-24 05:39:42--  https://parallel-ssh.googlecode.com/files/pssh-2.3.1.tar.gz
Resolving parallel-ssh.googlecode.com... 74.125.129.82, 2607:f8b0:400e:c02::52
Connecting to parallel-ssh.googlecode.com|74.125.129.82|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 23427 (23K) [application/x-gzip]
Saving to: âpssh-2.3.1.tar.gz.1â

100%[======================================================================================

2013-04-24 05:39:43 (240 KB/s) - âpssh-2.3.1.tar.gz.1â
```

```
scp pssh-2.3.1.tar.gz rhel1:/root
```

2. Copy pssh-2.3.1.tar.gz to the Admin Node

```
ssh rhel1
tar xzf pssh-2.3.1.tar.gz
cd pssh-2.3.1
python setup.py install
```

3. Extract and Install pssh on the Admin node.

```
[root@redhat ~]# scp pssh-2.3.1.tar.gz rhel1:/root
The authenticity of host 'rhel1 (10.29.160.53)' can't be established.
RSA key fingerprint is 25:15:c9:7d:e0:db:78:2c:0d:ce:e5:2d:e3:e2:5e:44.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'rhel1' (RSA) to the list of known hosts.
root@rhel1's password:
pssh-2.3.1.tar.gz
[root@redhat ~]# ssh rhel1
root@rhel1's password:
Last login: Wed Apr 24 09:06:38 2013 from 10.29.160.90
[root@rhel1 ~]# tar xzf pssh-2.3.1.tar.gz
[root@rhel1 ~]# cd pssh-2.3.1
[root@rhel1 pssh-2.3.1]# python setup.py install
running install
running build
running build_py
running build_scripts
running install_lib
running install_scripts
changing mode of /usr/bin/pslurp to 755
changing mode of /usr/bin/pnuke to 755
changing mode of /usr/bin/prsync to 755
changing mode of /usr/bin/pscp to 755
changing mode of /usr/bin/pssh-askpass to 755
changing mode of /usr/bin/pssh to 755
running install_data
running install_egg_info
Removing /usr/lib/python2.6/site-packages/pssh-2.3.1-py2.6.egg-info
Writing /usr/lib/python2.6/site-packages/pssh-2.3.1-py2.6.egg-info
```

4. Create a host file containing the IP addresses of all the nodes and all the Datanodes in the cluster. This file is passed as a parameter to pssh to identify the nodes to run the commands on.

```
vi /root/allnodes
# This file contains ip address of all nodes of the cluster
#used by parallel-shell (pssh). For Details man pssh
10.29.160.53
10.29.160.54
10.29.160.55
10.29.160.56
10.29.160.57
10.29.160.58
10.29.160.59
10.29.160.60
10.29.160.61
10.29.160.62
10.29.160.63
10.29.160.64
10.29.160.65
10.29.160.66
10.29.160.67
10.29.160.68
...
10.29.160.116
```

```
vi /root/datanodes
10.29.160.55
10.29.160.56
10.29.160.57
10.29.160.58
10.29.160.59
10.29.160.60
10.29.160.61
10.29.160.62
10.29.160.63
10.29.160.64
10.29.160.65
10.29.160.66
10.29.160.67
10.29.160.68
...
10.29.160.116
```

CLUSTER SHELL

1. From the system connected to the Internet, download the 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/
   ```

2. Login to rhel1 and install cluster shell

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

3. Edit /etc/clustershell/groups file to include hostnames for all the nodes of the cluster.
4. For 64 node cluster all: rhel[1-64]

Configuring /etc/hosts

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

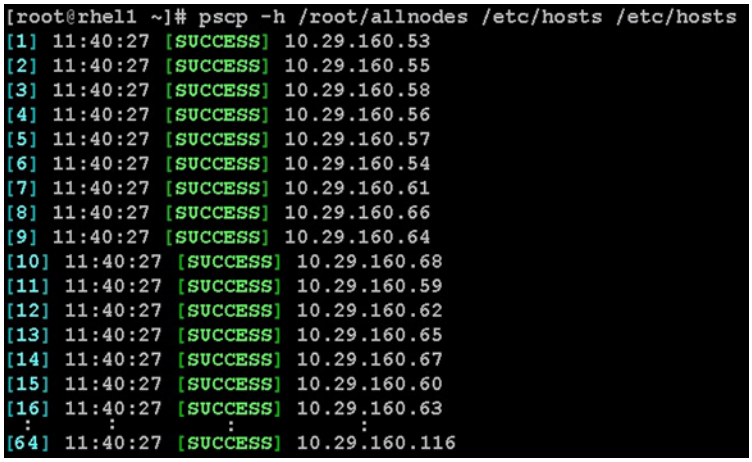
1. Populate the host file with IP addresses and corresponding hostnames on the Admin node (rhel1).

   ```
   vi /etc/hosts
   127.0.0.1 localhost localhost.localdomain localhost4
   localhost4.localdomain4
   ::1 localhost localhost.localdomain localhost6
   localhost6.localdomain6
   192.168.12.11 rhel1
   192.168.12.12 rhel2
   192.168.12.13 rhel3
   ```

```
192.168.12.14 rhel4
192.168.12.15 rhel5
192.168.12.16 rhel6
192.168.12.17 rhel7
192.168.12.18 rhel8
192.168.12.19 rhel9
192.168.12.20 rhel10
192.168.12.21 rhel11
192.168.12.22 rhel12
192.168.12.23 rhel13
192.168.12.24 rhel14
192.168.12.25 rhel15
192.168.12.26 rhel16
...
192.168.12.74 rhel64
```

1. Deploy /etc/hosts from the admin node (rhel1) to all the nodes via the following pscp command:

   ```
   pscp -h /root/allnodes /etc/hosts /etc/hosts
   ```



Creating Red Hat Local Repository

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
3. Alternatively, if you have access to a Red Hat ISO Image, Copy the ISO file to rhel1.

   ```
   scp rhel-server-6.4-x86_64-dvd.iso rhel1:/root
   ```

Note 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-6.4-x86_64-dvd.iso
/mnt/rheliso/
```

4. Copy the contents of the ISO to the /var/www/html/rhelrepo directory

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

5. On rhel1, create a repo file to enable the use of the yum command.

```
vi /var/www/html/rhelrepo/rheliso.repo
[rhel6.4]
name=Red Hat Enterprise Linux 6.4
baseurl=http://10.29.160.53/rhelrepo
gpgcheck=0
enabled=1
```

Note Based on this repository file, yum requires httpd to be running on rhel1 for other nodes to access the repository. Steps to install and configure httpd are in the following section.

6. Copy the rheliso.repo to all the nodes of the cluster.

```
pscp -h /root/allnodes /var/www/html/rhelrepo/rheliso.repo
/etc/yum.repos.d/
```

```
[root@rhel1 ~]# pscp -h /root/allnodes /var/www/html/rhelrepo/rheliso.rep
o /etc/yum.repos.d/
[1] 15:00:09 [SUCCESS] 10.29.160.57
[2] 15:00:09 [SUCCESS] 10.29.160.54
[3] 15:00:09 [SUCCESS] 10.29.160.53
[4] 15:00:09 [SUCCESS] 10.29.160.56
[5] 15:00:09 [SUCCESS] 10.29.160.58
[6] 15:00:09 [SUCCESS] 10.29.160.55
[7] 15:00:09 [SUCCESS] 10.29.160.60
[8] 15:00:09 [SUCCESS] 10.29.160.59
[9] 15:00:09 [SUCCESS] 10.29.160.65
[10] 15:00:09 [SUCCESS] 10.29.160.64
[11] 15:00:09 [SUCCESS] 10.29.160.61
[12] 15:00:09 [SUCCESS] 10.29.160.67
[13] 15:00:09 [SUCCESS] 10.29.160.62
[14] 15:00:09 [SUCCESS] 10.29.160.63
[15] 15:00:09 [SUCCESS] 10.29.160.66
[16] 15:00:09 [SUCCESS] 10.29.160.68
 :      :        :         :
[64] 15:00:09 [SUCCESS] 10.29.160.116
```

7. 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.

```
vi /etc/yum.repos.d/rheliso.repo
[rhel6.4]
name=Red Hat Enterprise Linux 6.4
baseurl=file:///var/www/html/rhelrepo
gpgcheck=0
enabled=1
```

8. pssh -h /root/allnodes "yum clean all"

```
[root@rhel1 ~]# pssh -h /root/allnodes "yum clean all"
[1] 12:14:09 [SUCCESS] 10.29.160.55
[2] 12:14:09 [SUCCESS] 10.29.160.53
[3] 12:14:09 [SUCCESS] 10.29.160.57
[4] 12:14:09 [SUCCESS] 10.29.160.54
[5] 12:14:09 [SUCCESS] 10.29.160.62
[6] 12:14:09 [SUCCESS] 10.29.160.59
[7] 12:14:09 [SUCCESS] 10.29.160.56
[8] 12:14:09 [SUCCESS] 10.29.160.58
[9] 12:14:09 [SUCCESS] 10.29.160.61
[10] 12:14:09 [SUCCESS] 10.29.160.65
[11] 12:14:09 [SUCCESS] 10.29.160.60
[12] 12:14:09 [SUCCESS] 10.29.160.68
[13] 12:14:09 [SUCCESS] 10.29.160.63
[14] 12:14:09 [SUCCESS] 10.29.160.64
[15] 12:14:10 [SUCCESS] 10.29.160.66
[16] 12:14:10 [SUCCESS] 10.29.160.67
 :       :        :          :
[64] 12:14:10 [SUCCESS] 10.29.160.116
```

Creating the Red Hat Repository Database

Install the createrepo package. Use it to regenerate the repository database(s) for the local copy of the RHEL DVD contents. Then purge the yum caches.

```
yum -y install createrepo
cd /var/www/html/rhelrepo
createrepo .
yum clean all
```

```
[root@rhel1 rhelrepo]# createrepo .
 368/3596 - Packages/pygobject2-doc-2.20.0-5.el6.x86_64.rpm
iso-8859-1 encoding on Ville Skyttä <ville.skytta@iki.fi> - 2.8.2-2

3596/3596 - Packages/lohit-bengali-fonts-2.4.3-6.el6.noarch.rpm
Saving Primary metadata
Saving file lists metadata
Saving other metadata
```

Upgrading the LSI driver

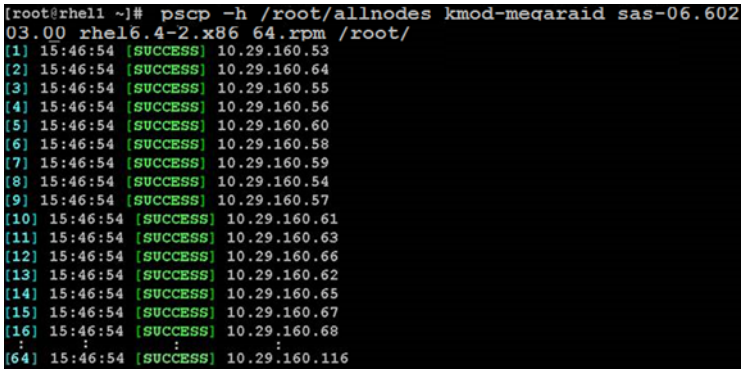
The latest LSI driver is required for performance and bug fixes. The latest drivers can be downloaded from the link below:

http://software.cisco.com/download/release.html?mdfid=284296254&flowid=31743&softwareid=283853158&release=1.5.1&relind=AVAILABLE&rellifecycle=&reltype=latest

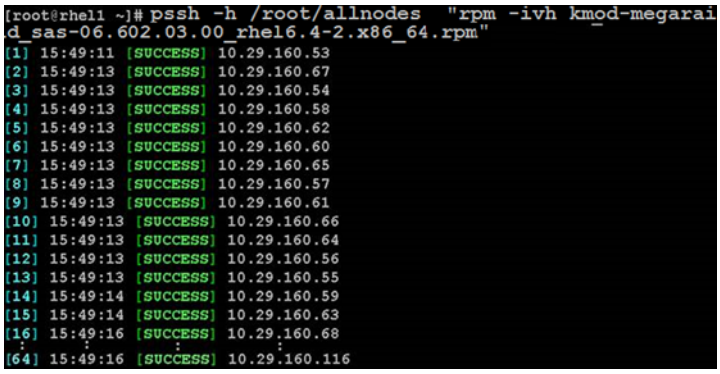
In the ISO image, the required driver kmod-megaraid_sas-06.602.03.00_rhel6.4-2.x86_64.rpm can be located at ucs-cxxx-drivers.1.5.1\Linux\Storage\LSI\92xx\RHEL\RHEL6.4

From a node connected to the Internet, download and transfer kmod-megaraid_sas-06.602.03.00_rhel6.4-2.x86_64.rpm to rhel1 (admin node). Install the rpm on all nodes of the cluster using the following pssh commands. For this example, the rpm is assumed to be in present working directory of rhel1.

```
pscp -h /root/allnodes
kmod-megaraid_sas-06.602.03.00_rhel6.4-2.x86_64.rpm /root/
```

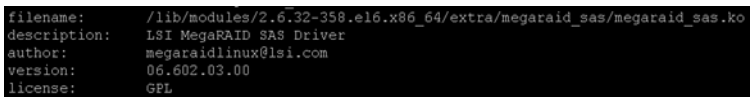



```
pssh -h /root/allnodes  "rpm -ivh
kmod-megaraid_sas-06.602.03.00_rhel6.4-2.x86_64.rpm "
```



Make sure that the above installed version of kmod-megaraid_sas driver is being used on all nodes by running the command "modinfo megaraid_sas" on all nodes.

```
pssh -h /root/allnodes  "modinfo megaraid_sas | head -5"
```



Installing httpd

Follow these steps to install httpd:

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.

```
/etc/httpd/conf/httpd.conf
   ServerName 10.29.160.53:80
```

```
#ServerName www.example.com:80
ServerName 10.29.160.53:80

#
# UseCanonicalName: Determines how Apache constructs self-referencing
# URLs and the SERVER_NAME and SERVER_PORT variables.
# When set "Off", Apache will use the Hostname and Port supplied
# by the client.  When set "On", Apache will use the value of the
# ServerName directive.
#
UseCanonicalName Off
```

3. Ensure httpd is able to read the repofiles

```
chcon -R -t httpd_sys_content_t /var/www/html/rhelrepo
```

4. Start httpd

```
service httpd start
chkconfig httpd on
```

Installing xfsprogs

Follow these steps to install xfsprogs;

1. Install xfsprogs on all the nodes for xfs filesystem.

```
pssh -h /root/allnodes "yum -y install xfsprogs"
```

```
[root@rhel1 ~]# pssh -h /root/allnodes "yum -y install xfsprogs"
[1] 12:26:34 [SUCCESS] 10.29.160.57
[2] 12:26:35 [SUCCESS] 10.29.160.56
[3] 12:26:35 [SUCCESS] 10.29.160.53
[4] 12:26:35 [SUCCESS] 10.29.160.59
[5] 12:26:35 [SUCCESS] 10.29.160.61
[6] 12:26:35 [SUCCESS] 10.29.160.63
[7] 12:26:35 [SUCCESS] 10.29.160.54
[8] 12:26:35 [SUCCESS] 10.29.160.62
[9] 12:26:35 [SUCCESS] 10.29.160.66
[10] 12:26:35 [SUCCESS] 10.29.160.60
[11] 12:26:35 [SUCCESS] 10.29.160.68
[12] 12:26:35 [SUCCESS] 10.29.160.58
[13] 12:26:35 [SUCCESS] 10.29.160.64
[14] 12:26:35 [SUCCESS] 10.29.160.55
[15] 12:26:35 [SUCCESS] 10.29.160.65
[16] 12:26:35 [SUCCESS] 10.29.160.67
```

Setting JAVA_HOME

Execute the following command on admin node (rhel1):

```
echo "export JAVA_HOME=/usr/lib/jvm/jre-1.7.0-openjdk.x86_64" >>
/etc/profile
```

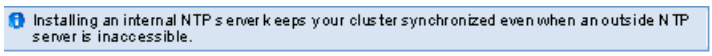
```
[root@rhel1 ~]# echo export JAVA_HOME=/usr/lib/jvm/jre-1.7.0-openjdk.x86_64 >> /etc/profile
```

Copy the profile file from admin node (rhel1) to all the nodes using the following command:

```
pscp -h /root/allnodes /etc/profile /etc/
```

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.



1. 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
```

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

```
vi /root/ntp.conf
server 10.29.160.53
driftfile /var/lib/ntp/drift
restrict 127.0.0.1
restrict -6 ::1
includefile /etc/ntp/crypto/pw
keys /etc/ntp/keys
```

```
[root@rhel1 ~]# for SERVER in {54..116}; do scp /root/ntp.conf 10.29.160.$SERVER:/etc/ntp.conf; done
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
ntp.conf                                      100%  142     0.1KB/s   00:00
```

3. 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 {54..116}; do scp /root/ntp.conf 10.29.160.$SERVER:/etc/ntp.conf; done

Note Do not use pssh /root/allnodes command without editing the host file allnodes as it overwrites /etc/ntp.conf from the admin node.

4. Syncronize the time and restart NTP daemon on all nodes:

```
pssh -h /root/allnodes "yum install -y ntpdate"
pssh -h /root/allnodes "service ntpd stop"
pssh -h /root/allnodes "ntpdate rhel1"
pssh -h /root/allnodes "service ntpd start"
```

```
[root@rhel1 ~]# pssh -h /root/allnodes "service ntpd restart"
[1] 13:38:55 [SUCCESS] 10.29.160.54
[2] 13:38:55 [SUCCESS] 10.29.160.53
[3] 13:38:55 [SUCCESS] 10.29.160.56
[4] 13:38:55 [SUCCESS] 10.29.160.57
[5] 13:38:55 [SUCCESS] 10.29.160.55
[6] 13:38:55 [SUCCESS] 10.29.160.58
[7] 13:38:55 [SUCCESS] 10.29.160.60
[8] 13:38:55 [SUCCESS] 10.29.160.59
[9] 13:38:55 [SUCCESS] 10.29.160.64
[10] 13:38:55 [SUCCESS] 10.29.160.62
[11] 13:38:55 [SUCCESS] 10.29.160.61
[12] 13:38:55 [SUCCESS] 10.29.160.66
[13] 13:38:55 [SUCCESS] 10.29.160.63
[14] 13:38:55 [SUCCESS] 10.29.160.65
[15] 13:38:55 [SUCCESS] 10.29.160.67
[16] 13:38:55 [SUCCESS] 10.29.160.68
 :     :       :          :
[64] 13:38:55 [SUCCESS] 10.29.160.116
```

Make sure restart of NTP daemon across reboots:

```
pssh -h /root/allnodes "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
```

```
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.

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

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

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

Verify the ulimit setting with the following step:

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

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

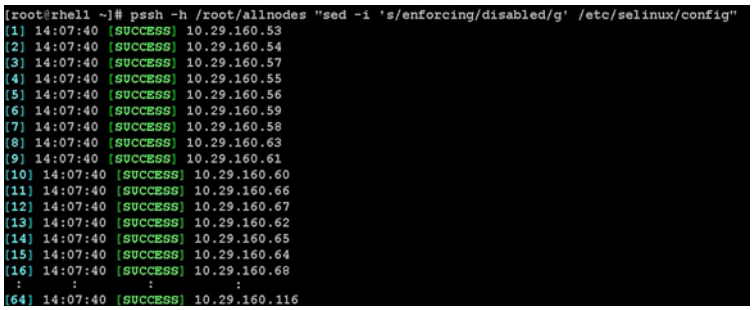
```
clush -B -a ulimit -n
```

Disabling SELinux

SELinux must be disabled during the HDP 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.

```
pssh -h /root/allnodes "sed -i
's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config "
```




```
pssh -h /root/allnodes "setenforce 0"
```

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 add the following line:

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

2. Save the file and run:

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

Disabling the Linux Firewall

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

```
pssh -h /root/allnodes  "service iptables stop"
```

```
[root@rhel1 ~]# pssh -h /root/allnodes "service iptables stop"
[1] 14:13:25 [SUCCESS] 10.29.160.54
[2] 14:13:25 [SUCCESS] 10.29.160.55
[3] 14:13:25 [SUCCESS] 10.29.160.57
[4] 14:13:25 [SUCCESS] 10.29.160.59
[5] 14:13:25 [SUCCESS] 10.29.160.56
[6] 14:13:25 [SUCCESS] 10.29.160.62
[7] 14:13:25 [SUCCESS] 10.29.160.60
[8] 14:13:25 [SUCCESS] 10.29.160.66
[9] 14:13:25 [SUCCESS] 10.29.160.61
[10] 14:13:25 [SUCCESS] 10.29.160.63
[11] 14:13:25 [SUCCESS] 10.29.160.67
[12] 14:13:25 [SUCCESS] 10.29.160.58
[13] 14:13:25 [SUCCESS] 10.29.160.53
[14] 14:13:25 [SUCCESS] 10.29.160.68
[15] 14:13:25 [SUCCESS] 10.29.160.65
[16] 14:13:25 [SUCCESS] 10.29.160.64
[64] 14:13:25 [SUCCESS] 10.29.160.116
```

```
pssh -h /root/allnodes "chkconfig iptables off"
```

```
[root@rhel1 ~]# pssh -h /root/allnodes "chkconfig iptables off"
[1] 14:13:25 [SUCCESS] 10.29.160.54
[2] 14:13:25 [SUCCESS] 10.29.160.55
[3] 14:13:25 [SUCCESS] 10.29.160.57
[4] 14:13:25 [SUCCESS] 10.29.160.59
[5] 14:13:25 [SUCCESS] 10.29.160.56
[6] 14:13:25 [SUCCESS] 10.29.160.62
[7] 14:13:25 [SUCCESS] 10.29.160.60
[8] 14:13:25 [SUCCESS] 10.29.160.66
[9] 14:13:25 [SUCCESS] 10.29.160.61
[10] 14:13:25 [SUCCESS] 10.29.160.63
[11] 14:13:25 [SUCCESS] 10.29.160.67
[12] 14:13:25 [SUCCESS] 10.29.160.58
[13] 14:13:25 [SUCCESS] 10.29.160.53
[14] 14:13:25 [SUCCESS] 10.29.160.68
[15] 14:13:25 [SUCCESS] 10.29.160.65
[16] 14:13:25 [SUCCESS] 10.29.160.64
 :      :         :         :
[64] 14:13:25 [SUCCESS] 10.29.160.116
```

Configuring Data Drives on Name Node

The section Configuring Disk Drives for OS on Name Nodes describes the steps to configure the first two disk drives for the operating system for nodes rhel1 and rhel2. The remaining disk drives can also be configured Raid 1 similarly or by using MegaCli as described below.

Follow these steps to configure data drive on Name Node:

1. From the LSI website
   http://www.lsi.com/support/Pages/Download-Results.aspx?keyword=9271-8i
2. Download MegaCli and its dependencies and transfer to Admin node.

   ```
   scp  /root/MegaCli64  rhel1:/root/
   scp  /root/Lib_Utils-1.00-08.noarch.rpm rhel1:/root/
   scp  /root/Lib_Utils2-1.00-01.noarch.rpm rhel1:/root/
   ```

3. Copy all three files to all the nodes using the following commands:

   ```
   pscp -h  /root/allnodes  /root/MegaCli64  /root/
   ```

```
[root@rhel1 ~]# pscp -h /root/allnodes /root/MegaCli64 /root/
[1] 13:00:40 [SUCCESS] 10.29.160.53
[2] 13:00:40 [SUCCESS] 10.29.160.61
[3] 13:00:40 [SUCCESS] 10.29.160.58
[4] 13:00:40 [SUCCESS] 10.29.160.62
[5] 13:00:40 [SUCCESS] 10.29.160.56
[6] 13:00:40 [SUCCESS] 10.29.160.57
[7] 13:00:40 [SUCCESS] 10.29.160.66
[8] 13:00:40 [SUCCESS] 10.29.160.59
[9] 13:00:40 [SUCCESS] 10.29.160.60
[10] 13:00:40 [SUCCESS] 10.29.160.55
[11] 13:00:40 [SUCCESS] 10.29.160.68
[12] 13:00:40 [SUCCESS] 10.29.160.54
[13] 13:00:40 [SUCCESS] 10.29.160.63
[14] 13:00:40 [SUCCESS] 10.29.160.64
[15] 13:00:40 [SUCCESS] 10.29.160.65
[16] 13:00:40 [SUCCESS] 10.29.160.67
 :      :         :         :
[64] 13:00:40 [SUCCESS] 10.29.160.116
```

```
pscp -h  /root/allnodes  /root/Lib_Utils*  /root/
```

```
[root@rhel1 ~]# pscp -h /root/allnodes /root/Lib_Utils* /root/
[1] 13:01:26 [SUCCESS] 10.29.160.53
[2] 13:01:26 [SUCCESS] 10.29.160.58
[3] 13:01:26 [SUCCESS] 10.29.160.59
[4] 13:01:26 [SUCCESS] 10.29.160.60
[5] 13:01:26 [SUCCESS] 10.29.160.67
[6] 13:01:26 [SUCCESS] 10.29.160.63
[7] 13:01:26 [SUCCESS] 10.29.160.61
[8] 13:01:26 [SUCCESS] 10.29.160.57
[9] 13:01:26 [SUCCESS] 10.29.160.54
[10] 13:01:26 [SUCCESS] 10.29.160.56
[11] 13:01:26 [SUCCESS] 10.29.160.62
[12] 13:01:26 [SUCCESS] 10.29.160.55
[13] 13:01:26 [SUCCESS] 10.29.160.64
[14] 13:01:26 [SUCCESS] 10.29.160.66
[15] 13:01:26 [SUCCESS] 10.29.160.65
[16] 13:01:26 [SUCCESS] 10.29.160.68
 :     :        :        :
[64] 13:01:26 [SUCCESS] 10.29.160.116
```

4. Run the following command to install the rpms on all the nodes:

   ```
   pssh -h /root/allnodes "rpm -ivh Lib_Utils*"
   ```

```
[root@rhel1 ~]# pssh -h /root/allnodes "rpm -ivh Lib_Utils*"
[1] 13:02:05 [SUCCESS] 10.29.160.64
[2] 13:02:05 [SUCCESS] 10.29.160.62
[3] 13:02:05 [SUCCESS] 10.29.160.57
[4] 13:02:05 [SUCCESS] 10.29.160.66
[5] 13:02:05 [SUCCESS] 10.29.160.58
[6] 13:02:05 [SUCCESS] 10.29.160.59
[7] 13:02:05 [SUCCESS] 10.29.160.54
[8] 13:02:05 [SUCCESS] 10.29.160.67
[9] 13:02:05 [SUCCESS] 10.29.160.60
[10] 13:02:05 [SUCCESS] 10.29.160.65
[11] 13:02:05 [SUCCESS] 10.29.160.56
[12] 13:02:05 [SUCCESS] 10.29.160.55
[13] 13:02:05 [SUCCESS] 10.29.160.63
[14] 13:02:05 [SUCCESS] 10.29.160.61
[15] 13:02:05 [SUCCESS] 10.29.160.68
[16] 13:02:05 [SUCCESS] 10.29.160.53
 :     :        :        :
[64] 13:02:05 [SUCCESS] 10.29.160.116
```

5. Run the following script as root user on NameNode and Secondary NameNode to create the virtual drives.

   ```
   vi /root/raid1.sh

   ./MegaCli64 -cfgldadd
   r1[$1:3,$1:4,$1:5,$1:6,$1:7,$1:8,$1:9,$1:10,$1:11,$1:12,$1:13,$1:14,$
   1:15,$1:16,$1:17,$1:18,$1:19,$1:20,$1:21,$1:22,$1:23,$1:24] wb ra
   nocachedbadbbu strpsz1024 -a0
   ```

The above script requires enclosure ID as a parameter.

6. Run the following command to get enclousure id:

```
./MegaCli64 pdlist -a0 | grep Enc | grep -v 252 | awk '{print $4}' |
sort | uniq -c | awk '{print $2}'

chmod 755 raid1.sh
```

7. Run MegaCli script as follows:

```
./raid1.sh <EnclosureID> obtained by running the command above

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

Configuring the Filesystem for NameNodes

```
vi /root/driveconf.sh
#!/bin/bash
#disks_count=`lsblk -id | grep sd  | wc -l`
#if [ $disks_count -eq 2 ];  then
#   echo "Found 2 disks"
#else
#    echo "Found $disks_count disks. Expecting 2. Exiting.."
#    exit 1
fi
[[ "-x" == "${1}" ]] && set -x && set -v && shift 1
for X in /sys/class/scsi_host/host?/scan
do
echo '- - -' > ${X}
done
for X in /dev/sd?
do
echo $X
if [[ -b ${X} && `/sbin/parted -s ${X} print quit|/bin/grep -c boot`
-ne 0 ]]
then
echo "$X bootable - skipping."
continue
else
Y=${X##*/}1
/sbin/parted  -s  ${X} mklabel gpt quit
/sbin/parted  -s  ${X} mkpart 1 6144s 100% quit
/sbin/mkfs.xfs -f -q -l size=65536b,lazy-count=1,su=256k -d
sunit=1024,swidth=6144 -r extsize=256k -L ${Y} ${X}1
(( $? )) && continue
```

```
/bin/mkdir -p /DATA/${Y}
(( $? )) && continue
/bin/mount  -t xfs  -o allocsize=128m,noatime,nobarrier,nodiratime
${X}1  /DATA/${Y}
(( $? )) && continue
echo "LABEL=${Y} /DATA/${Y} xfs
allocsize=128m,noatime,nobarrier,nodiratime 0 0" >> /etc/fstab
fi
done
```

Configuring Data Drives on Data Nodes

The section Configuring Disk Drives for OS on Data Nodesdescribes the setps to configure the first disk drive for the operating system for nodes rhel3 to rhel64. Remaining disk drives can also be configured similarly or using MegaCli as described below.

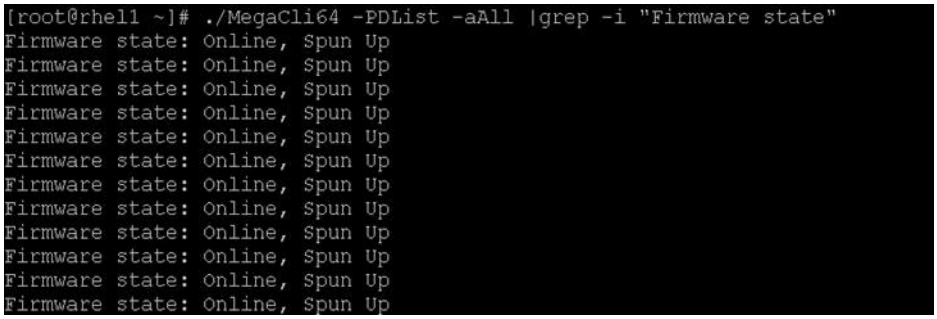
Issue the following command from the admin node to create the virtual drives with RAID 0 configurations on all the datanodes:

```
pssh -h /root/datanodes "./MegaCli64 -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 existing configurations. To clear and reconfigure existing configurations refer to Embedded MegaRAID Software Users Guide available at www.lsi.com.

Note Make sure all the drives are up by running the following command: `./MegaCli64 -PDList -aAll |grep -i "Firmware state"`



Configuring the Filesystem for Data Nodes

On the Admin node, create a file containing the following script.

1. To create partition tables and file systems on the local disks supplied to each of the nodes, run the following script as the root user on each node:

```
vi /root/driveconf.sh
#!/bin/bash
#disks_count=`lsblk -id | grep sd  | wc -l`
#if [ $disks_count -eq 24 ];  then
#    echo "Found 24 disks"
#else
#    echo "Found $disks_count disks. Expecting 24. Exiting.."
#    exit 1
fi
[[ "-x" == "${1}" ]] && set -x && set -v && shift 1
for X in /sys/class/scsi_host/host?/scan
do
echo '- - -' > ${X}
done
for X in /dev/sd?
do
echo $X
if [[ -b ${X} && `/sbin/parted -s ${X} print quit|/bin/grep -c boot`
-ne 0 ]]
then
echo "$X bootable - skipping."
continue
else
Y=${X##*/}1
/sbin/parted  -s  ${X} mklabel gpt quit
/sbin/parted  -s  ${X} mkpart 1 6144s 100% quit
/sbin/mkfs.xfs -f -q -l size=65536b,lazy-count=1,su=256k -d
sunit=1024,swidth=6144 -r extsize=256k -L ${Y} ${X}1
(( $? )) && continue
/bin/mkdir  -p  /DATA/${Y}
(( $? )) && continue
/bin/mount  -t xfs  -o allocsize=128m,noatime,nobarrier,nodiratime
${X}1  /DATA/${Y}
(( $? )) && continue
echo "LABEL=${Y} /DATA/${Y} xfs
allocsize=128m,noatime,nobarrier,nodiratime 0 0" >> /etc/fstab
fi
done
```

2. Run the following command to copy driveconf.sh to all the data nodes:

```
pscp -h /root/datanodes /root/driveconf.sh /root/
```

3. Run the following command from the admin node to run the script across all data nodes:

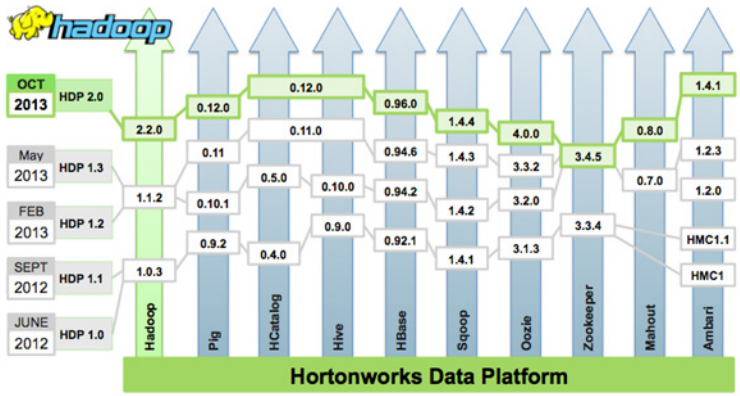
```
pssh -h /root/datanodes "./driveconf.sh"
```

```
[root@rhel1 ~]# pssh -h /root/allnodes  "./driveconf.sh"
[1] 16:15:24 [SUCCESS] 10.29.160.67
[2] 16:15:24 [SUCCESS] 10.29.160.54
[3] 16:15:24 [SUCCESS] 10.29.160.63
[4] 16:15:24 [SUCCESS] 10.29.160.66
[5] 16:15:24 [SUCCESS] 10.29.160.65
[6] 16:15:24 [SUCCESS] 10.29.160.62
[7] 16:15:24 [SUCCESS] 10.29.160.61
[8] 16:15:24 [SUCCESS] 10.29.160.60
[9] 16:15:24 [SUCCESS] 10.29.160.59
[10] 16:15:24 [SUCCESS] 10.29.160.58
[11] 16:15:24 [SUCCESS] 10.29.160.57
[12] 16:15:24 [SUCCESS] 10.29.160.64
[13] 16:15:25 [SUCCESS] 10.29.160.56
[14] 16:15:25 [SUCCESS] 10.29.160.55
[15] 16:15:25 [SUCCESS] 10.29.160.53
[16] 16:15:35 [SUCCESS] 10.29.160.68
 :       :       :
[64] 16:15:35 [SUCCESS] 10.29.160.116
```

Installing Hortonworks Data Platform 2.0

HDP is an enterprise grade, hardened Hadoop distribution. HDP combines Apache Hadoop and its related projects into a single tested and certified package. It offers the latest innovations from the open source community with the testing and quality you expect from enterprise quality software. HPD components are depicted in figure 10.

Figure 89 HDP Components



Pre-Requisites for HDP Installation

This section details the pre-requisites for HDP Installation, such as setting up the HDP Repository.

Hortonworks Repository

Follow these steps to install the Hortonworks Repository:

1. From a host connected to the Internet, download the Hortonworks repositories as shown below and transfer it to the admin node:

```
mkdir -p /tmp/Hortonworks
```

2. Download Hortonworks HDP Repo:

```
wget
http://public-repo-1.hortonworks.com/HDP/centos6/HDP-2.0.6.0-centos6-
rpm.tar.gz
```

3. Download Hortonworks HDP-Utils Repo:

```
wget
http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.16/repos/centos6
/HDP-UTILS-1.1.0.16-centos6.tar.gz
```

4. Download Ambari Repo:

```
wget
http://public-repo-1.hortonworks.com/ambari/centos6/ambari-1.4.3.38-c
entos6.tar.gz
```

5. Copy the repository directory to the admin node:

```
scp -r /tmp/Hortonworks/ rhel1:/var/www/html
```

6. Extract the files:

```
login to rhel1
cd /var/www/html/Hortonworks
tar -zxvf HDP-2.0.6.0-centos6-rpm.tar.gz
tar -zxvf HDP-UTILS-1.1.0.16-centos6.tar.gz
tar -zxvf ambari-1.4.3.38-centos6.tar.gz
```

7. Create the hdp.repo file with following contents:

```
vi /etc/yum.repos.d/hdp.repo

[HDP-UTILS-1.1.0.16]
name=Hortonworks Data Platform Utils Version - HDP-UTILS-1.1.0.16
baseurl= http://rhel1/Hortonworks/HDP-UTILS-1.1.0.16/repos/centos6
gpgcheck=0
enabled=1
priority=1
```

```
[HDP]
name=Hortonworks Data Platform HDP
baseurl= http://rhel1/Hortonworks/HDP/centos6/2.x/updates/2.0.6.0
gpgcheck=0
enabled=1
priority=1
```

8. Create the Ambari repo file with following contents:

```
vi /etc/yum.repos.d/ambari.repo

   [Updates-ambari-1.4.3.38]
name=ambari-1.4.3.38 - Updates
baseurl= http://rhel1/Hortonworks/ambari/centos6/1.x/updates/1.4.3.38
gpgcheck=0
enabled=1
priority=1
```

9. From the admin node copy the repo files to /etc/yum.repos.d/ of all the nodes of the cluster:

```
pscp -h /root/allnodes /etc/yum.repos.d/hdp.repo /etc/yum.repos.d/

pscp -h /root/allnodes /etc/yum.repos.d/ambari.repo /etc/yum.repos.d/
```

Hortonworks Data Platform Installation

Follow these steps to install HDP:

1. Install and Setup Ambari Server on rhel1

```
yum install ambari-server
```

2. Setup Ambari Server

```
ambari-server setup -j $JAVA_HOME -s
```

Start Ambari Server

```
ambari-server start
```

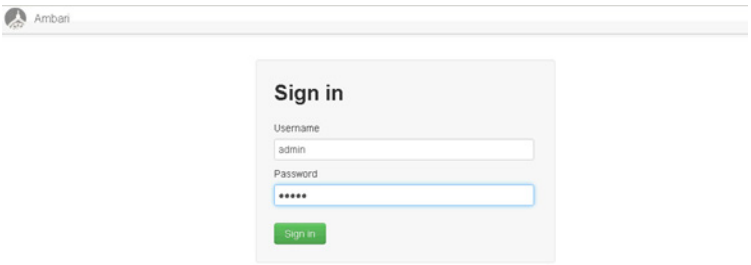
Confirm Ambari Server Startup

```
ps -ef | grep ambari-server
```

Log into Ambari Server

When the Ambari service has been started, access the Ambari Install Wizard through the browser.

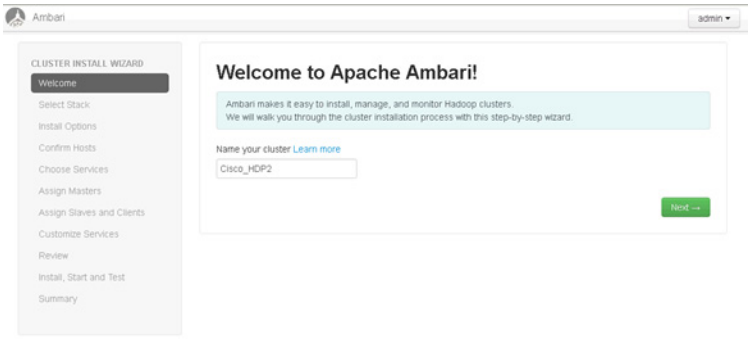
1. Point the browser to http://<ip address for rhel1>:8080
2. Log in to the Ambari Server using the default username/password: admin/admin. This can be changed at a later period of time.



Create Cluster Name

Following these steps to create the cluster name:

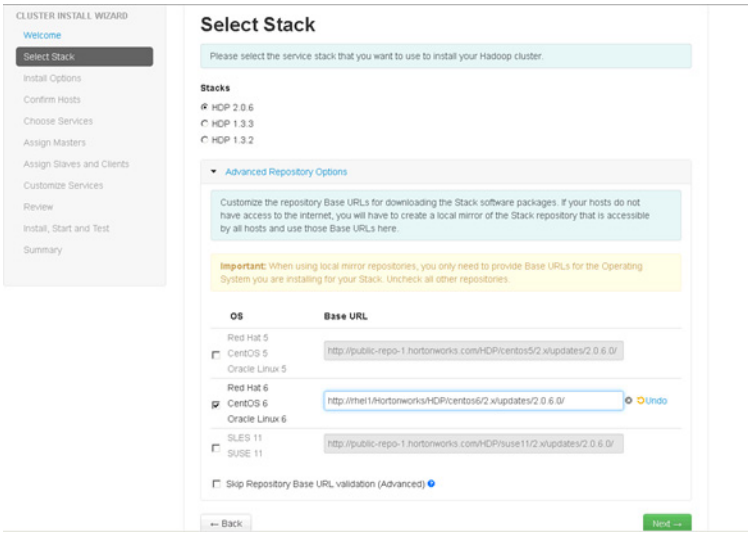
1. At the Welcome page, type "Cisco_HDP2" as the name for the cluster in the text box.
2. Click Next.



Select Stack

Follow these steps to select stack:

1. Select HDP 2.0.6 stack.
2. Expand "Advanced Repository Options".
3. Under the advanced repository option:
   - Select RedHat 6 checkbox
   - Update the RedHat 6 Base URL to http://rhel1/Hortonworks/rhel1/Hortonworks/HDP/centos6/2.x/updates/2.0.6.0/



Hortonworks Data Platform Installation

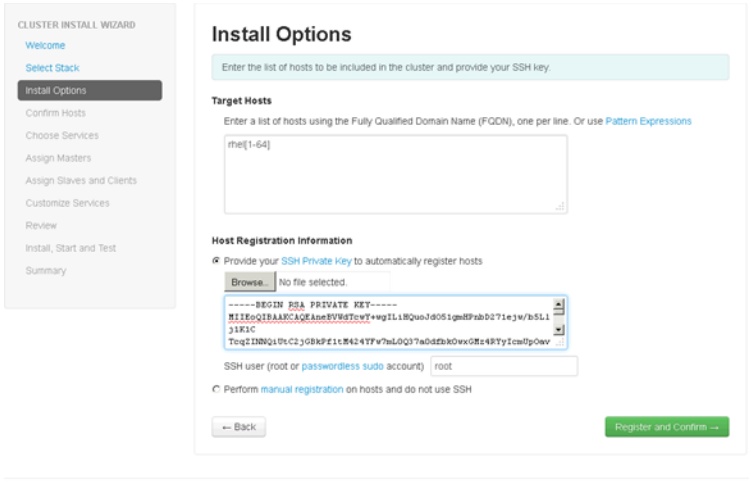
In order to build up the cluster, the install wizard needs to know general information about how the cluster has to be set up. This requires providing the Fully Qualified Domain Name (FQDN) of each of the host. The wizard also needs to access the private key file that was created in Setting Up Password-less Login. It uses these to locate all the hosts in the system and to access and interact with them securely.

1. Use the Target Hosts text box to enter the list of host names, one per line. One can also use ranges inside brackets to indicate larger sets of hosts.
2. Select the option Provide your SSH Private Key in the Ambari cluster install wizard.
3. Copy the contents of the file /root/.ssh/id_rsa on rhel1 and paste it in the text area provided by the Ambari cluster install wizard.

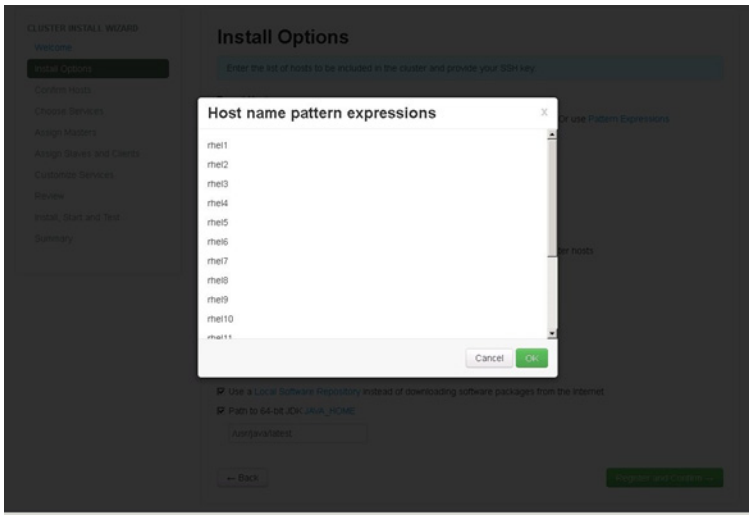
Note Make sure that no extra white space after the text-----END RSA PRIVATE KEY-----

4. Click the Register and Confirm button to continue.

```
[root@rhel1 ~]# cat /root/.ssh/id_rsa
-----BEGIN RSA PRIVATE KEY-----
MIIEoQIBAAKCAQEAyDOIRbk4mBZrizc0/gOM2iYT2h4vxkIxA/uvQVPthFreUdgT
Zehw/Qtdk7meeqhgqsHmb1CriF0m6SxvPEXW2cGoAx75hZwTuDIR3Qlvk6oYUmDW
BKq5TMfUMKfD7tknkGkg5N+YHsPCoNILlz/Wqc01hZZ0tiCmrxeRnPGSlJY74/Db
A0BewMuNajAoVppPD6cLGF6/NKORpEDUnCuwe5pCRV5tko+gzBeBF5oeCS6Ya6I7
nS0HplJXV0Mv23SNUwl3cswbqLdrr3atG6YRieVrmmr/PlrKMp192tzQ1mHZMBqG
w1RJTILjygW0gp5g7NQBGeM7sX4V6Omzv4vmzwIBIwKCAQEAg4+UEI+o2PjKVCuX
2h+XEwMUXCJ3KoNEyBpr2nj7KxckYas/8oLN6B1pYR0UB3X2YZVc6hBwuLI+JDMk
hrGNMALqwDjtHUlOyX/9HDlmlDyTo9k8LvPY2q8zqvHnJ+3Jisi92Dspc01xRRxQ
wnpofjAmlCDx5WXp4MZYX9HynCcKmheFefobLys6gloxd84eHW1y6b0xU1dh7hsQ
pcK+xpdFWlsHYFbvckTuCHUAezF4+uBT5F0PMiD7PwzrvbXKA65ABuezv9gg2/I1
PekIkRvbosniFbBUi2ZOS1uN/gsaZgmSQ9gTarJlV8zMy6K31LETcOckl2LZHRX2
5sEx6wKBgQD9CiKc0HFiulrQWW5cLTDJU8wzTiNK4M9lQb2LOhfFuZfluiAl3Ref
yiL9MjE3A5Mnn9pcRxMmXXPF4t9iuLh3+3tCsr1TzPml4WT+Fipa9sh+3JZ2HKgm
pCquAEdoFRK4oP3/yYQg95gie2SC9sBOz6zVohdyNUvnkiMb9vwi3wKBgQDKiyTi
Yu421OwsYKfZ7YjomjRKUFaH4CKtnyJy1SM3wFPRnZJd4BUaMq0DaTxr2tW4si+4
t88M8XS6FHGHymSqRtLOtYzMlmmwUtjCLNZQfqSeg1NovekXxXL0iUzel8PL3ZOH
AeBj0/GLQ3SF/PGWMokCwNtaJoV/xldBdIsqEQKBgEERPBmx8UVF3N29ZYVqtMYO
O9KtsU3Ex52x0ad1VpHt5TsSmo1kvO6TEE+8cw4lfZx5j+vXwxh+bjozBj30/Dwc
GGGbrQbrkKscs5HLL3Z5+QqtwEpB4hiQnUKvnVVHP1QMJA6S53YxCdz7KHlypnqq
bkWQFKhW2QEIUivDKuRlAoGASzr/EkIAtUfFb5Gdbj0n4V3Y6Gb7kY3DvNS1BhSm
rk7ADAdTnzX5NZ3LO8gAf9TwS+ppfx+zTfNIn0MFmNYlY9EpyJs0S/1adLEOroWu
sC8J8bu/5RNWk8z+z9s5zwUrd5txT2cYlJ8t1KQGtWyUPxoVoe/ccfENA5LP872S
xnsCgYAFRE4SbB416p9miiR1+gNCiihM9N+FmHMmcP/y80QL/MoAYoHB1Tn8cwVu
l+sju4bWGUZvnGMWXwpEU5zVBra+yShh309IwjP/1kpCNWz7CX+/uI6FY+slZxTr
t5P/AvhOvUKMhRFjXFQoY5yqNUkasvIu6S8Q1unl8N2IhEgw1g==
-----END RSA PRIVATE KEY-----
```



Hostname Pattern Expressions

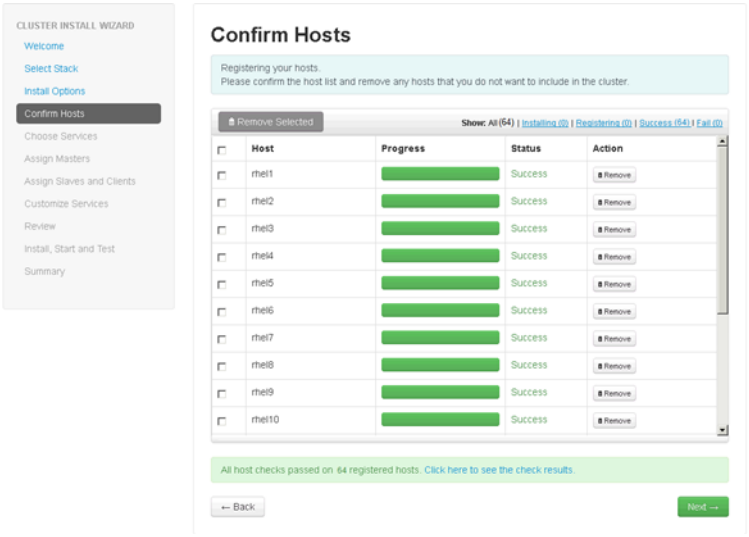


Confirm Hosts

The following screen allows you to make sure that Ambari has located the correct hosts for the cluster and to check those hosts to make sure they have the correct directories, packages, and processes to continue the install.

If any hosts were selected in error, it can be removed by selecting the appropriate checkboxes and clicking the grey Remove Selected button. To remove a single host, click the small white Remove button in the Action column.

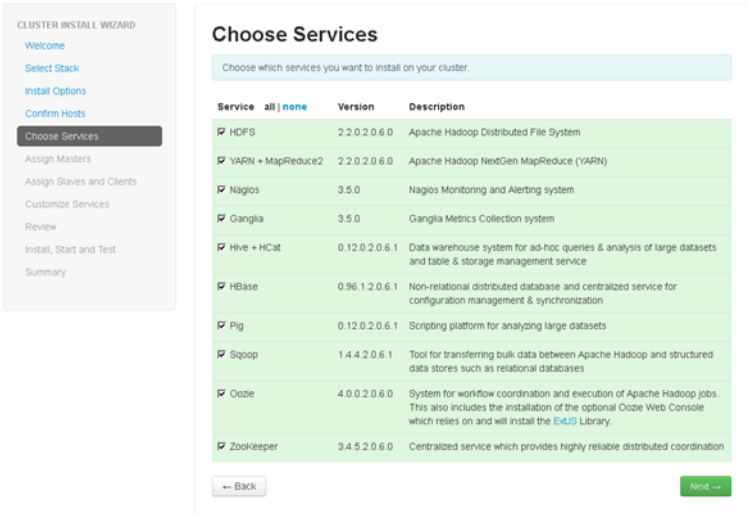
When the lists of hosts are confirmed, click Next.



Choose Services

HDP is made up of a number of components. See Understand the Basics for more information.

1. Select all to preselect all items
2. When you have made your selections, click Next.

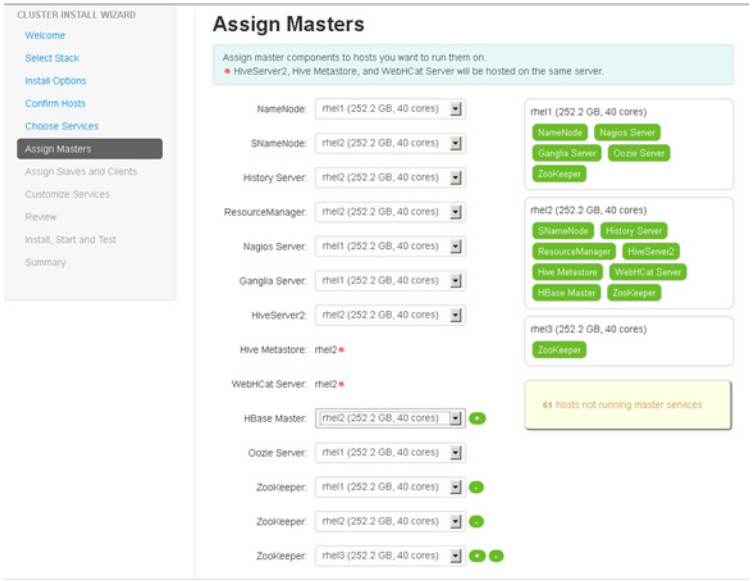


Assign Masters

The Ambari install wizard attempts to assign the master nodes for various services that have been selected to appropriate hosts in the cluster. The right column shows the current service assignments by host, with the hostname and its number of CPU cores and amount of RAM indicated.

1. Reconfigure the service assignment to match the table shown below.

Service Name	Host
NameNode	rhel1
SNameNode	rhel2
HistoryServer	rhel2
ResouceManager	rhel2
Nagios Server	rhel1
Ganglia Collector	rhel1
HiveServer2	rhel2
HBase Master	rhel2
Oozie Server	rhel1
Zookeeper	rhel1, rhel2, rhel3



Note On a small cluster (<16 nodes), consolidate all master services to run on a single node. For large clusters (> 64 nodes), deploy master services across 4 nodes.

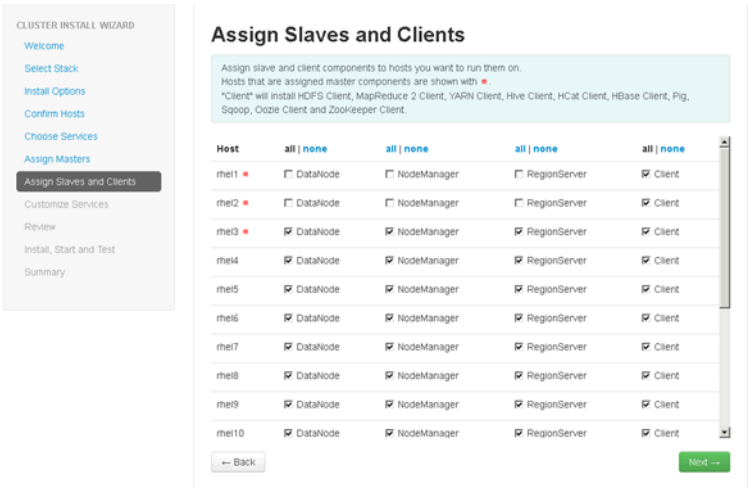
2. Click Next.

Assign Slaves and Clients

The Ambari install wizard attempts to assign the slave components (DataNodes, NodeManagers , and RegionServers) to appropriate hosts in the cluster. Reconfigure the service assignment to match the table below:

Client Service Name	Host
DataNode	rhel3 to rhel64
NodeManager	rhel3 to rhel64
RegionServer	rhel3 to rhel64
Client	All nodes, rhel1-rhel64

1. Assign DataNode, NodeManager, RegionServer nodes to rhel3- rhel64.
2. Assign Client to all nodes.
3. Click Next.



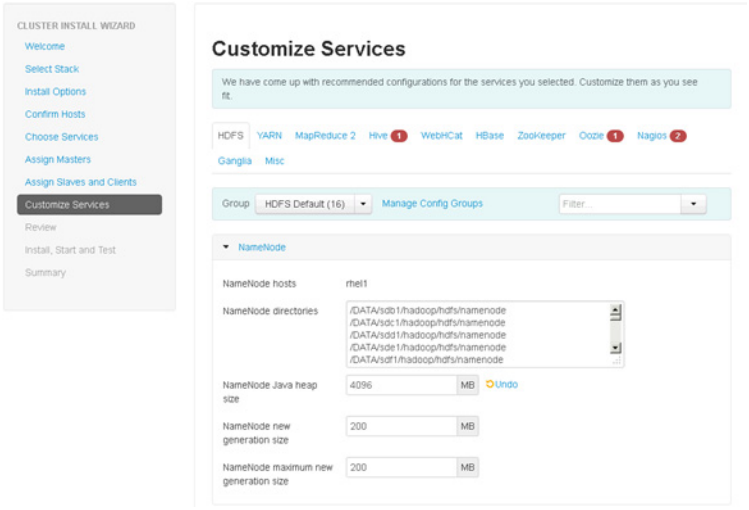
Customize Services

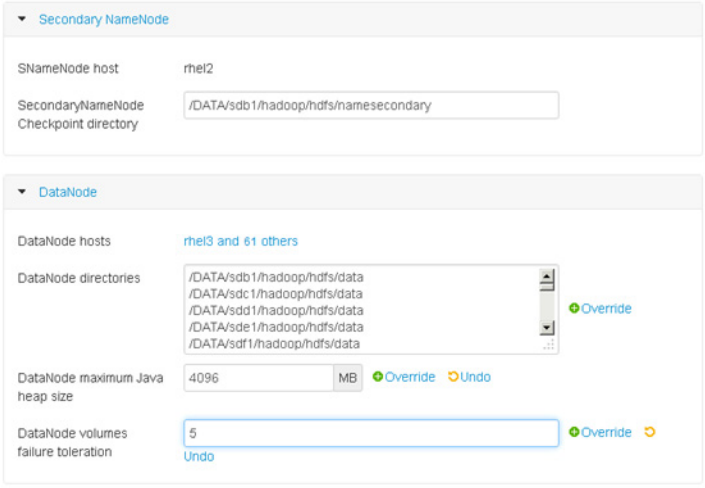
This section details the configuration settings for Hadoop components. The wizard attempts to set reasonable defaults for each of the options here, but this can be modified to meet specific requirements. The following sections provide configuration guidance that should be refined to meet specific use case requirements.

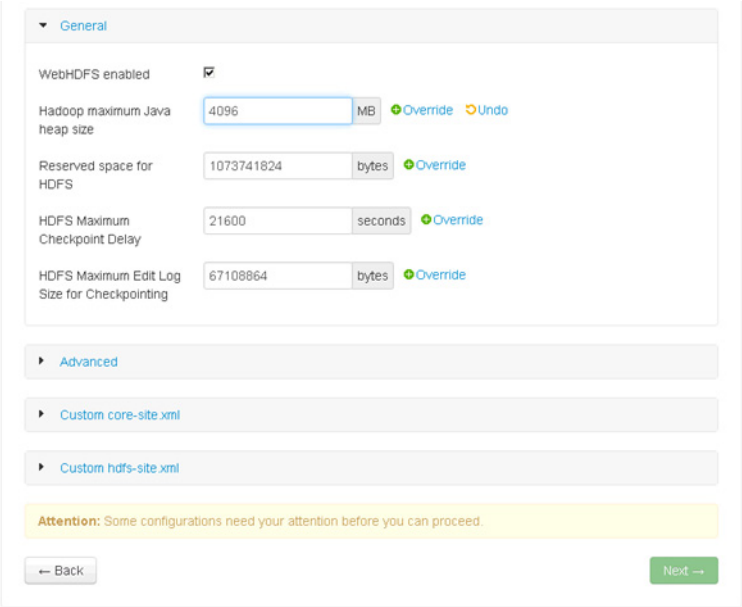
HDFS

Update the following HDFS configurations:

Property Name	Value
NameNode Java Heap Size	4096
Hadoop maximum Java heap size	4096
DataNode maximum Java heap size	4096
Datanode Volumes Failure Toleration	5


CLUSTER INSTALL WIZARD
Welcome
Select Stack
Install Options
Confirm Hosts
Choose Services
Assign Masters
Assign Slaves and Clients
Customize Services
Review
Install, Start and Test
Summary
Customize Services
We have come up with recommended configurations for the services you selected. Customize them as you see fit.
HDFS
YARN
MapReduce 2
Hive 1
WebHCat
HBase
ZooKeeper
Oozie 1
Nagios 2
Ganglia
Misc
Group
HDFS Default (16)
Manage Config Groups
Filter...
NameNode
NameNode hosts
rhel1
NameNode directories
/DATA/sdb1/hadoop/hdfs/namenode
/DATA/sdc1/hadoop/hdfs/namenode
/DATA/sdd1/hadoop/hdfs/namenode
/DATA/sde1/hadoop/hdfs/namenode
/DATA/sdf1/hadoop/hdfs/namenode
NameNode Java heap size
4096
MB
Undo
NameNode new generation size
200
MB
NameNode maximum new generation size
200
MB

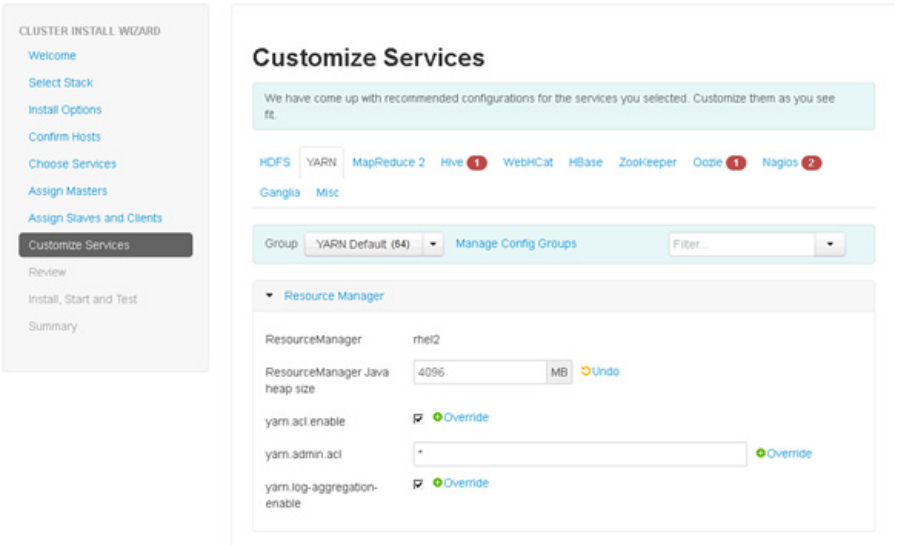

Secondary NameNode
SNameNode host
rhel2
SecondaryNameNode Checkpoint directory
/DATA/sdb1/hadoop/hdfs/namesecondary
DataNode
DataNode hosts
rhel3 and 61 others
DataNode directories
/DATA/sdb1/hadoop/hdfs/data
/DATA/sdc1/hadoop/hdfs/data
/DATA/sdd1/hadoop/hdfs/data
/DATA/sde1/hadoop/hdfs/data
/DATA/sdf1/hadoop/hdfs/data
Override
DataNode maximum Java heap size
4096
MB
Override
Undo
DataNode volumes failure toleration
5
Override
Undo

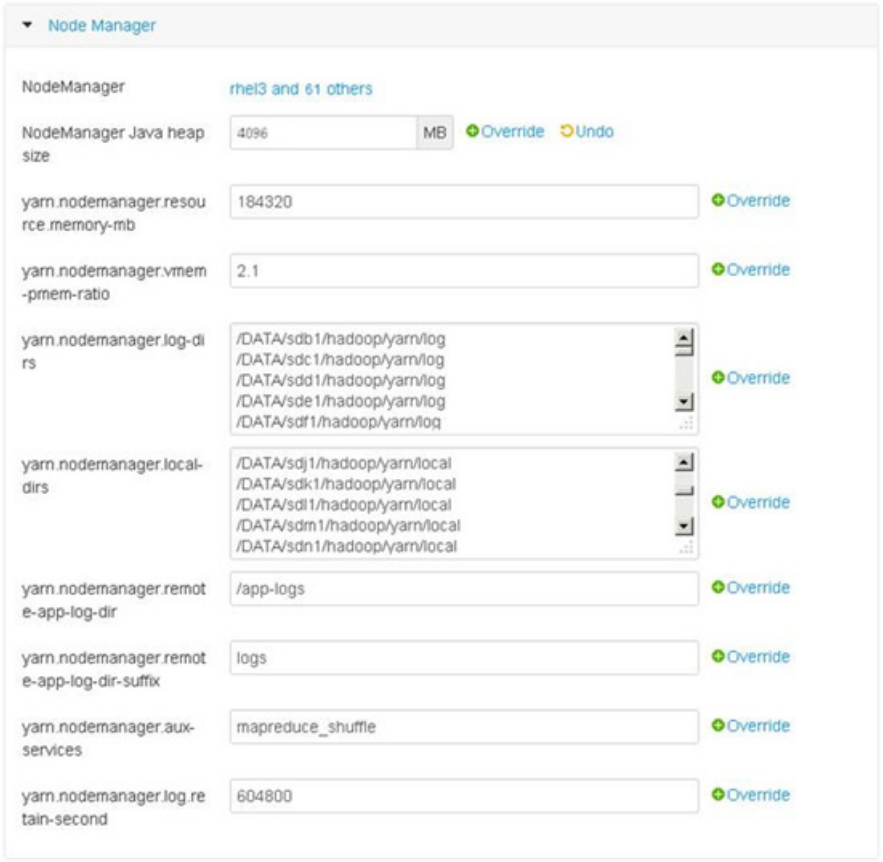


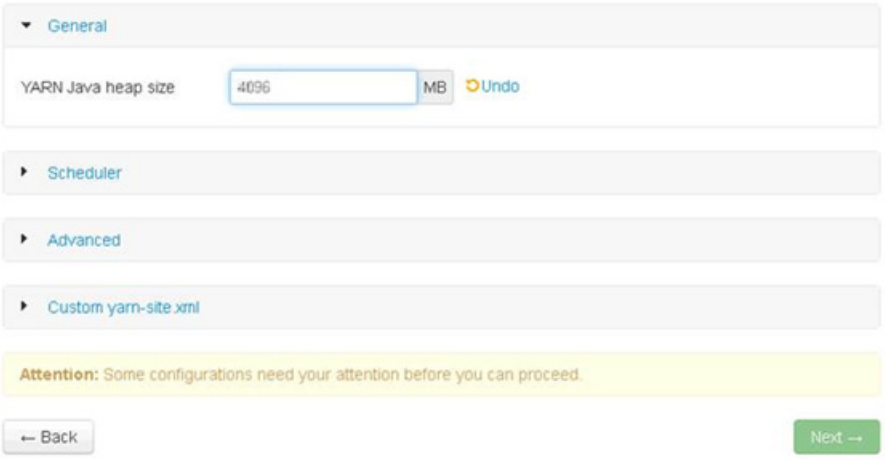
YARN

Update the following YARN configurations:

Property Name	Value
ResourceManager Java heap size	4096
NodeManager Java heap size	2048
yarn.nodemanager.resource.memory-mb	184320
YARN Java heap size	4096
yarn.scheduler.minimum-allocation-mb	4096
yarn.scheduler.maximum-allocation-mb	184320



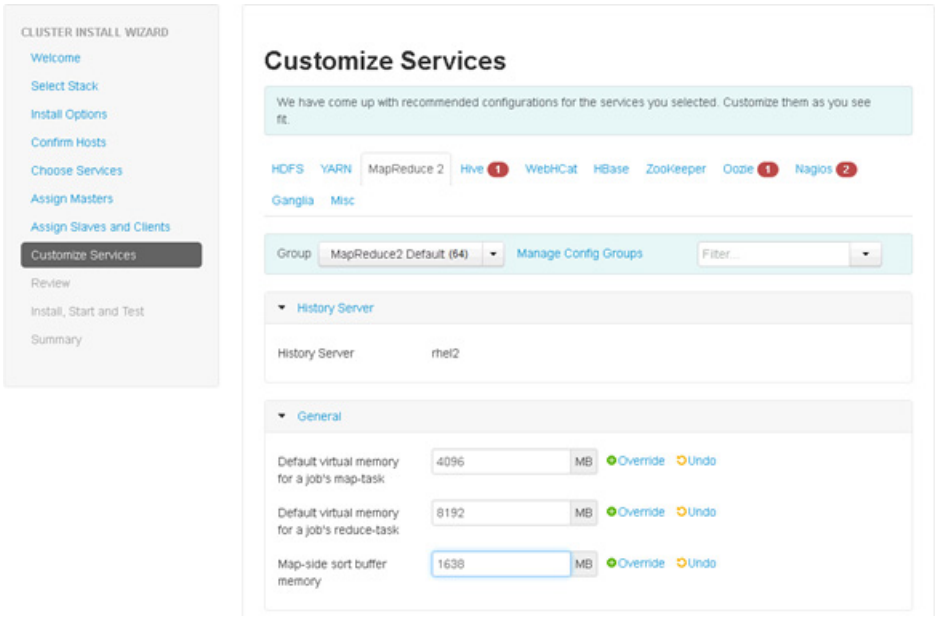

Node Manager
NodeManager
rhel3 and 61 others
NodeManager Java heap size
4096
MB
Override
Undo
yarn.nodemanager.resource.memory-mb
184320
Override
yarn.nodemanager.vmem-pmem-ratio
2.1
Override
yarn.nodemanager.log-dirs
/DATA/sdb1/hadoop/yarn/log
/DATA/sdc1/hadoop/yarn/log
/DATA/sdd1/hadoop/yarn/log
/DATA/sde1/hadoop/yarn/log
/DATA/sdf1/hadoop/yarn/log
Override
yarn.nodemanager.local-dirs
/DATA/sdj1/hadoop/yarn/local
/DATA/sdk1/hadoop/yarn/local
/DATA/sdl1/hadoop/yarn/local
/DATA/sdm1/hadoop/yarn/local
/DATA/sdn1/hadoop/yarn/local
Override
yarn.nodemanager.remote-app-log-dir
/app-logs
Override
yarn.nodemanager.remote-app-log-dir-suffix
logs
Override
yarn.nodemanager.aux-services
mapreduce_shuffle
Override
yarn.nodemanager.log.retain-second
604800
Override

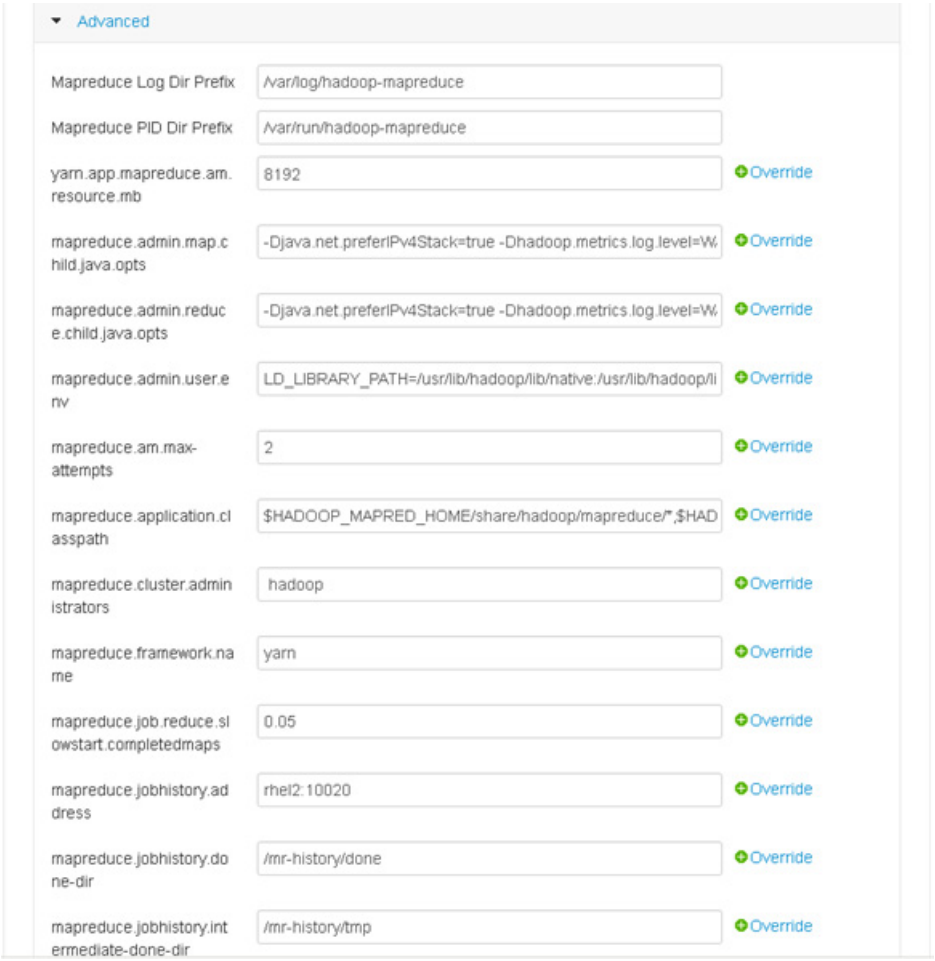

General
YARN Java heap size
4096
MB
Undo
Scheduler
Advanced
Custom yarn-site.xml
Attention: Some configurations need your attention before you can proceed.
← Back
Next →

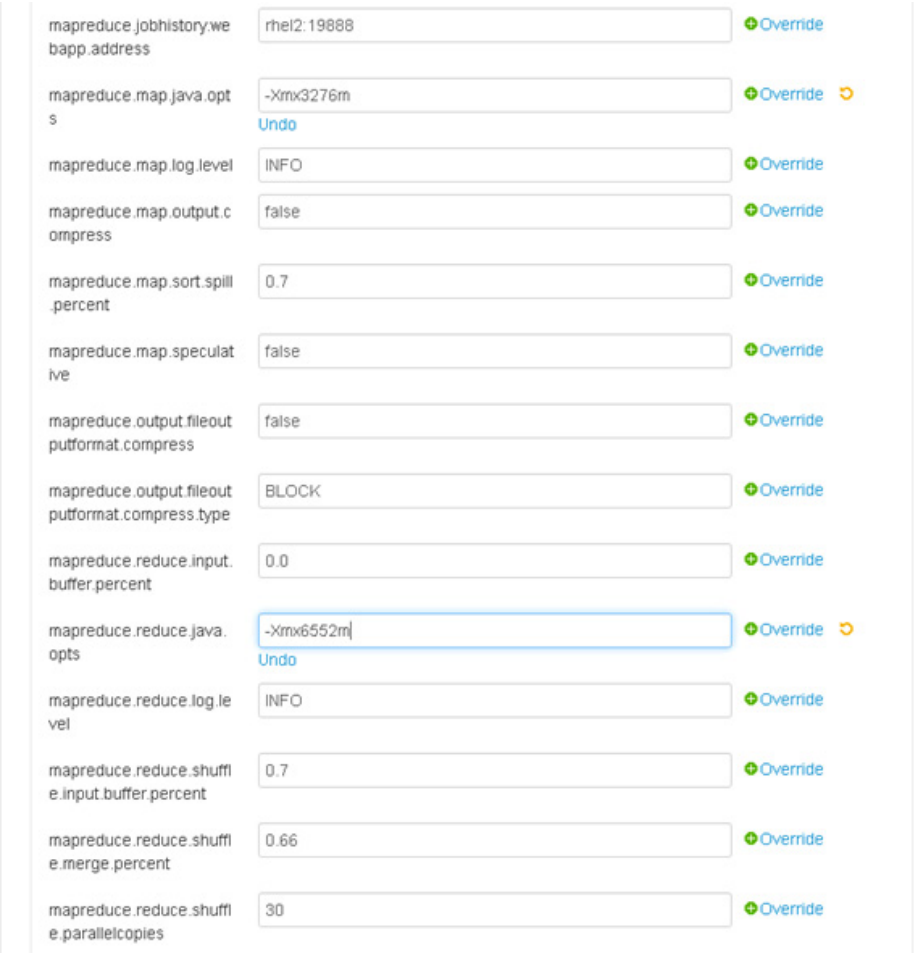
MapReduce

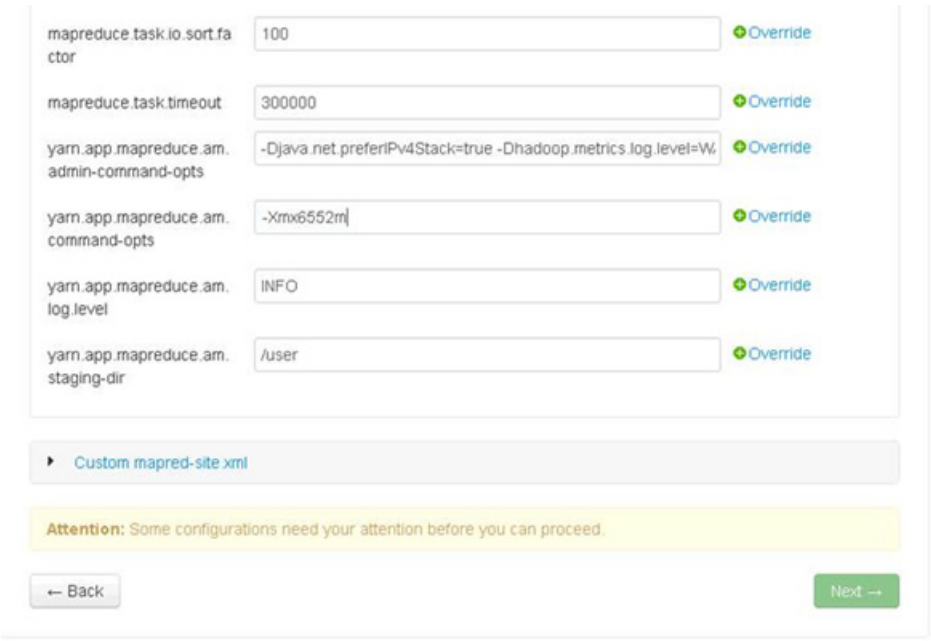
Update the following MapReduce configurations:

Property Name	Value
Default virtual memory for a job's map-task	4096
Default virtual memory for a job's reduce-task	8192
Map-side sort buffer memory	1638
yarn.app.mapreduce.am.resource.mb	8192
mapreduce.map.java.opts	-Xmx3276m
mapreduce.reduce.java.opts	-Xmx6552m
yarn.app.mapreduce.am.command-opts	-Xmx6552m



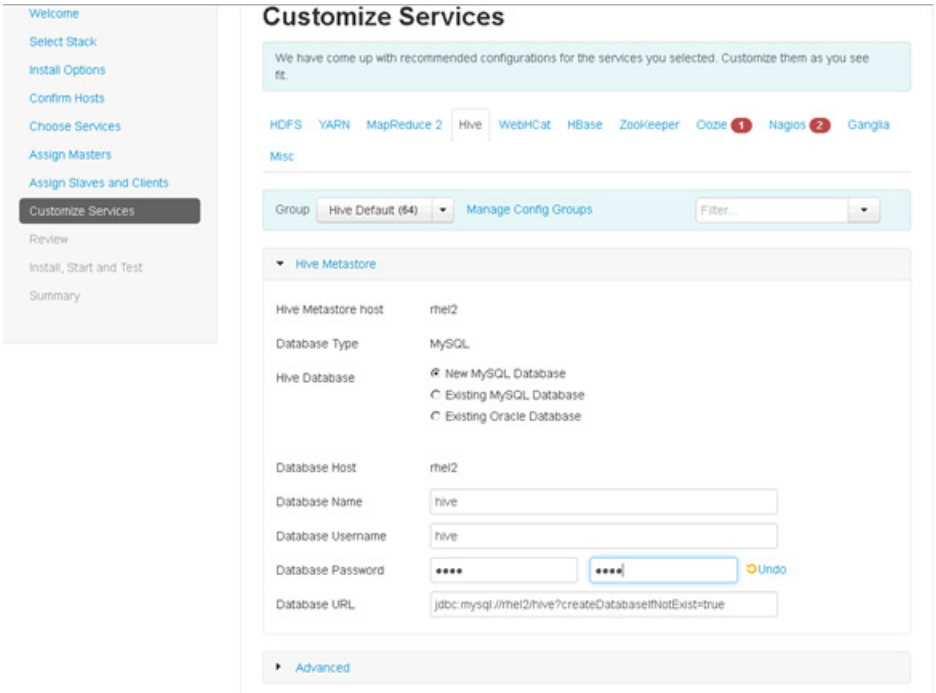

Advanced
Mapreduce Log Dir Prefix
/var/log/hadoop-mapreduce
Mapreduce PID Dir Prefix
/var/run/hadoop-mapreduce
yarn.app.mapreduce.am.resource.mb
8192
Override
mapreduce.admin.map.child.java.opts
-Djava.net.preferIPv4Stack=true -Dhadoop.metrics.log.level=W
Override
mapreduce.admin.reduce.child.java.opts
-Djava.net.preferIPv4Stack=true -Dhadoop.metrics.log.level=W
Override
mapreduce.admin.user.env
LD_LIBRARY_PATH=/usr/lib/hadoop/lib/native:/usr/lib/hadoop/li
Override
mapreduce.am.max-attempts
2
Override
mapreduce.application.classpath
$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*,$HAD
Override
mapreduce.cluster.administrators
hadoop
Override
mapreduce.framework.name
yarn
Override
mapreduce.job.reduce.slowstart.completedmaps
0.05
Override
mapreduce.jobhistory.address
rhel2:10020
Override
mapreduce.jobhistory.done-dir
/mr-history/done
Override
mapreduce.jobhistory.intermediate-done-dir
/mr-history/tmp
Override


mapreduce.jobhistory.webapp.address
rhel2:19888
Override
mapreduce.map.java.opts
-Xmx3276m
Undo
Override
mapreduce.map.log.level
INFO
Override
mapreduce.map.output.compress
false
Override
mapreduce.map.sort.spill.percent
0.7
Override
mapreduce.map.speculative
false
Override
mapreduce.output.fileoutputformat.compress
false
Override
mapreduce.output.fileoutputformat.compress.type
BLOCK
Override
mapreduce.reduce.input.buffer.percent
0.0
Override
mapreduce.reduce.java.opts
-Xmx6552m
Undo
Override
mapreduce.reduce.log.level
INFO
Override
mapreduce.reduce.shuffle.input.buffer.percent
0.7
Override
mapreduce.reduce.shuffle.merge.percent
0.66
Override
mapreduce.reduce.shuffle.parallelcopies
30
Override



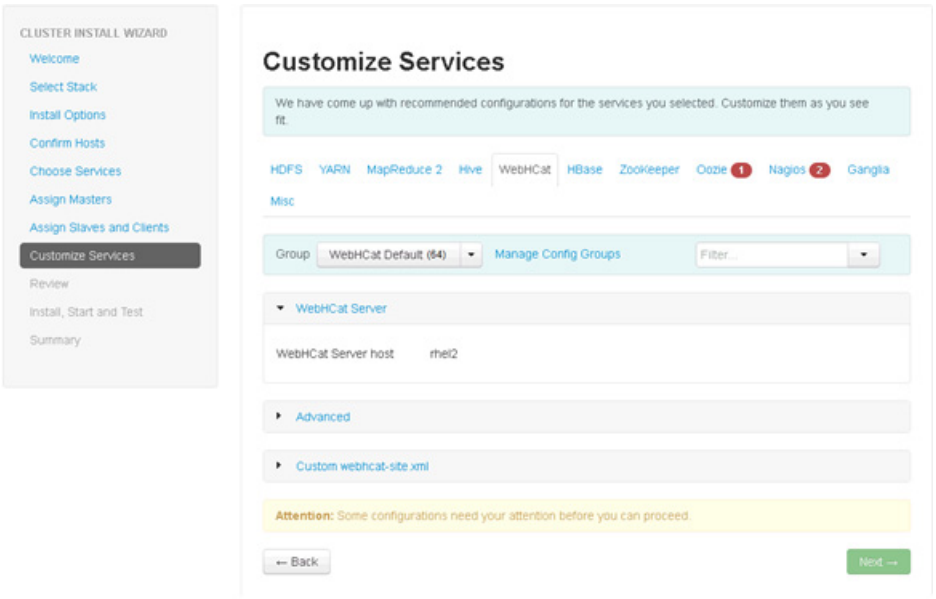
Hive/HCat

Enter the hive database password as per the organizational policy.



WebHCat

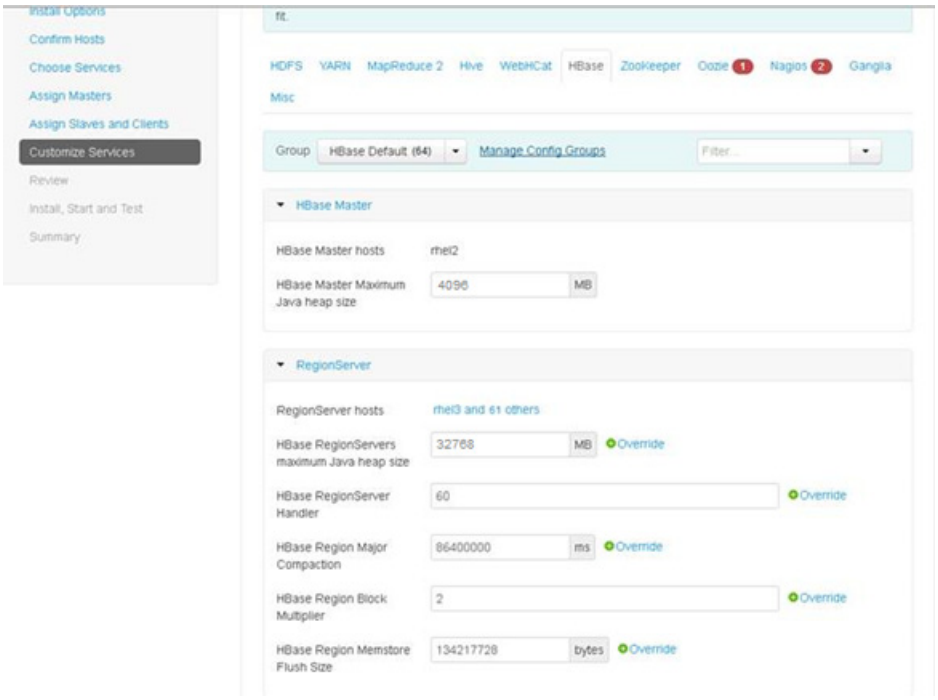
No changes are required.



HBase

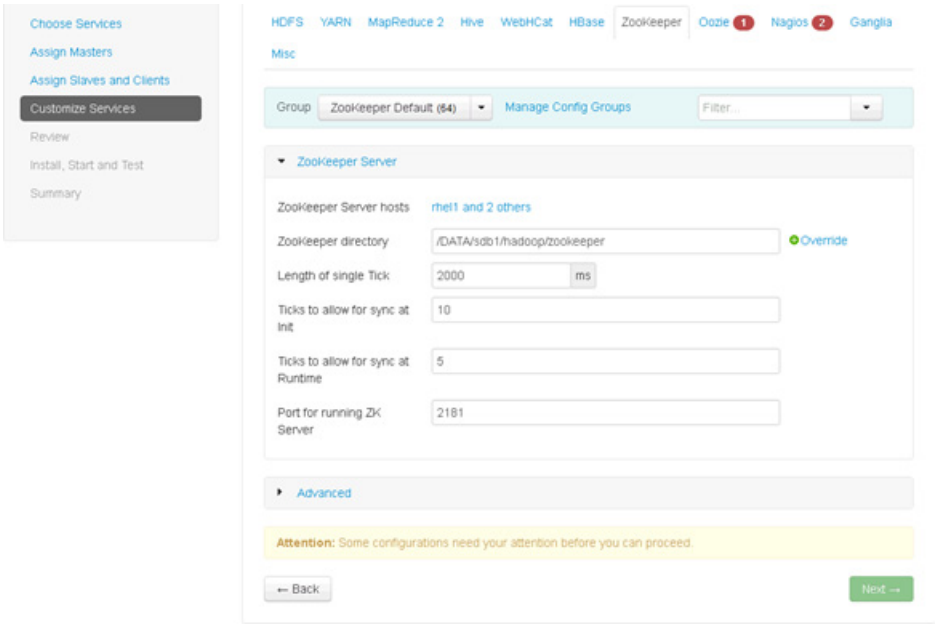
Update the following HBASE configurations:

Property Name	Value
HBase Master Maximum Java Heap Size	4096
HBase RegionServers Maximum Java Heap Size	32768



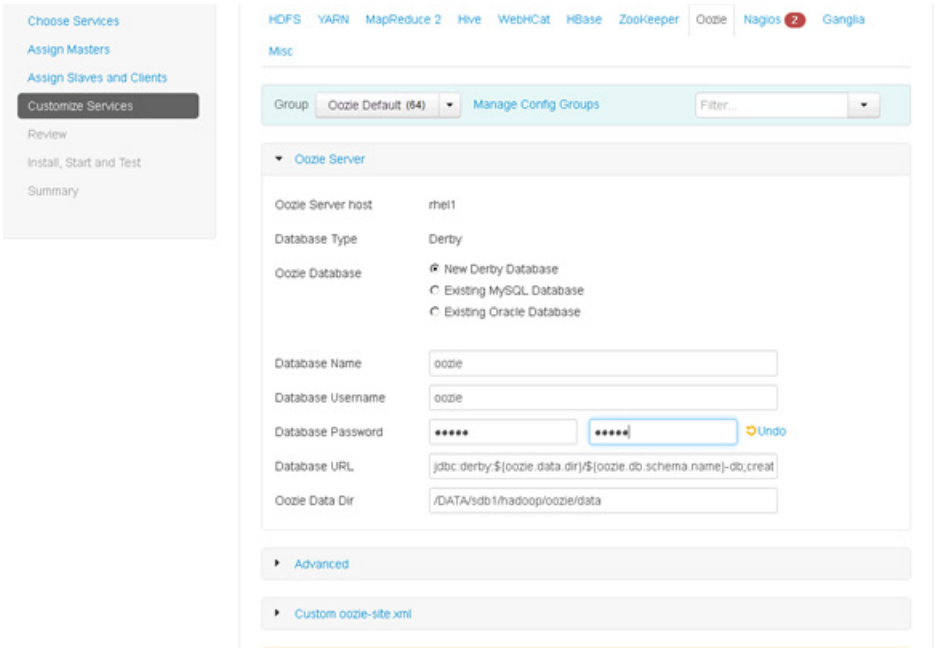
Zookeeper

No changes are required.



Oozie

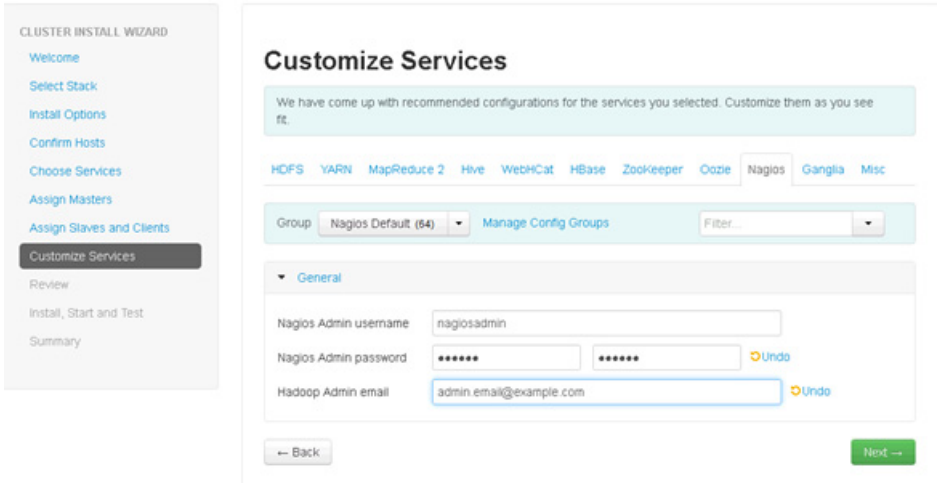
Enter the oozie database password as per organizational policy.



Nagio

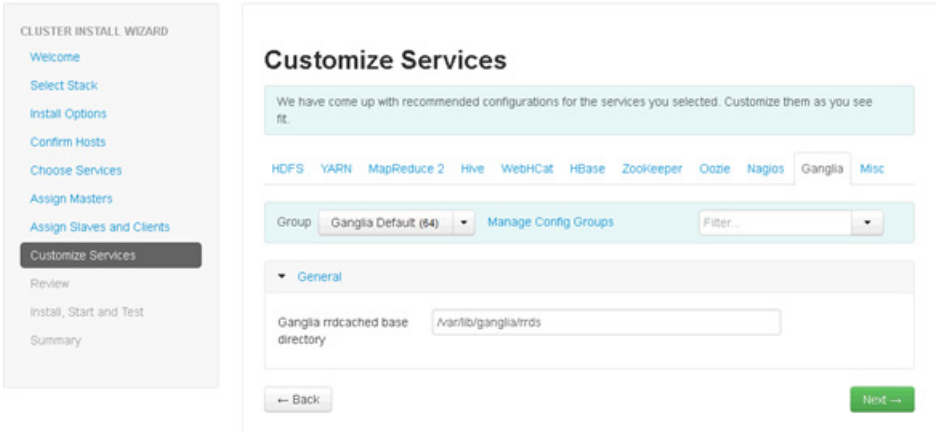
Provide the following:

- Nagios admin password as per organizational policy
- Hadoop admin email



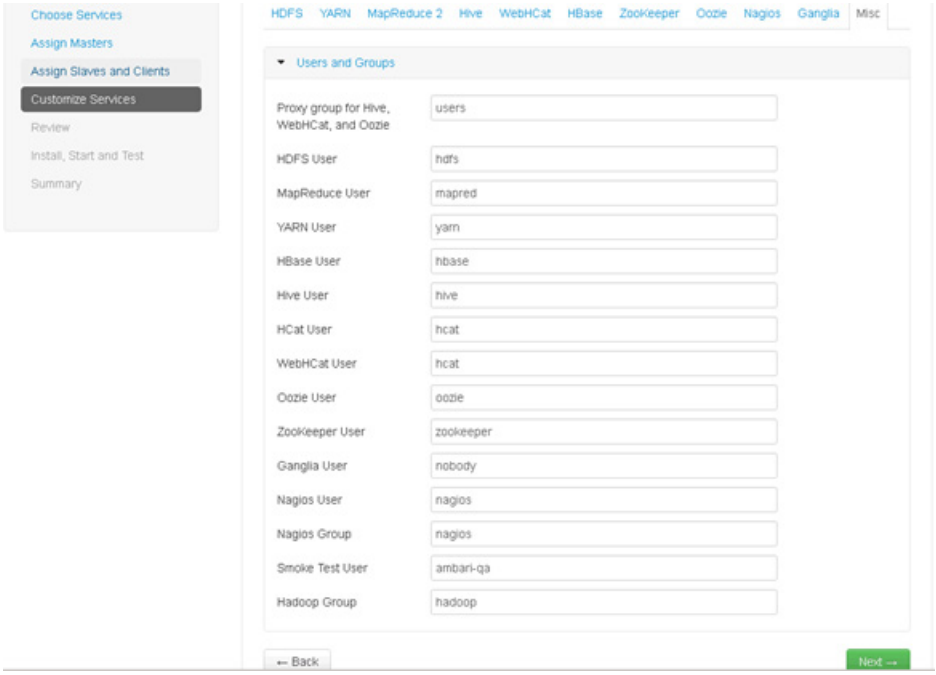
Ganglia

No changes are required.



Misc

No changes are required.

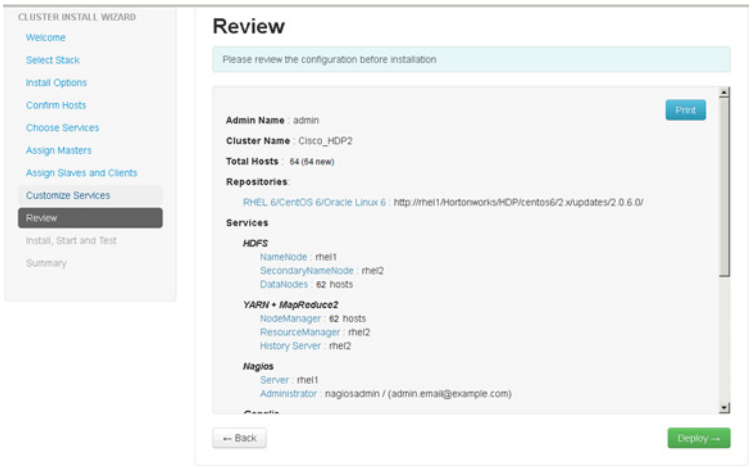


Review

The assignments that have been made are displayed. Check to make sure everything is correct before clicking the Deploy button. If any changes are to be made, use the left navigation bar to return to the appropriate screen.

Deploy

When the review is complete, click the Deploy button.

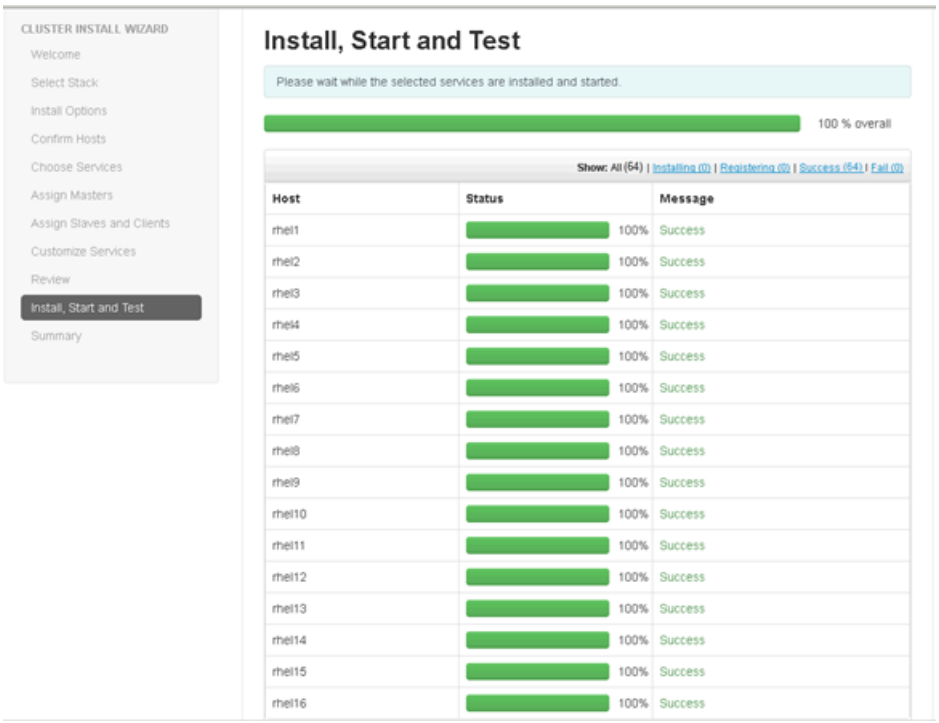


The progress of the install is shown on the screen. Each component is installed, started and validated. The next screen displays the overall status of the install in the progress bar at the top of the screen and a host-by-host status in the main section.

To see specific information on what tasks have been completed per host, click the link in the Message column for the appropriate host. In the Tasks pop-up, click the individual task to see the related log files. Select filter conditions by using the Show dropdown list. To see a larger version of the log contents, click the Open icon or to copy the contents to the clipboard, use the Copy icon.

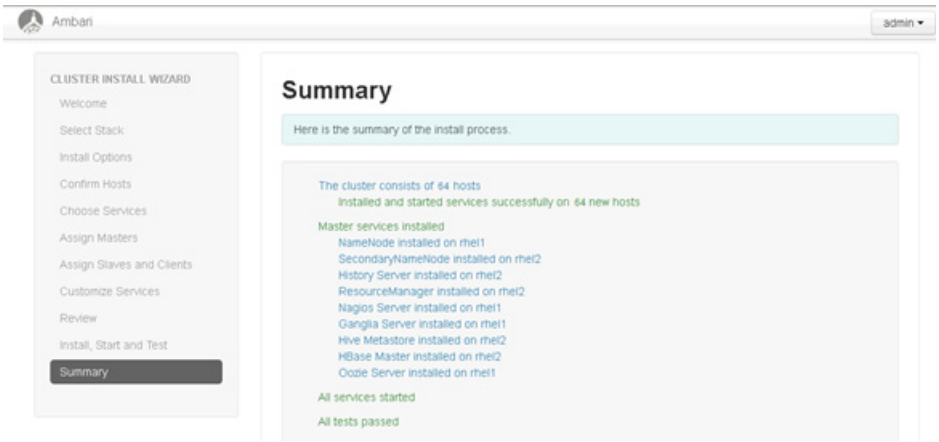
Note Depending on which components are installing, the entire process may take 10 or more minutes.

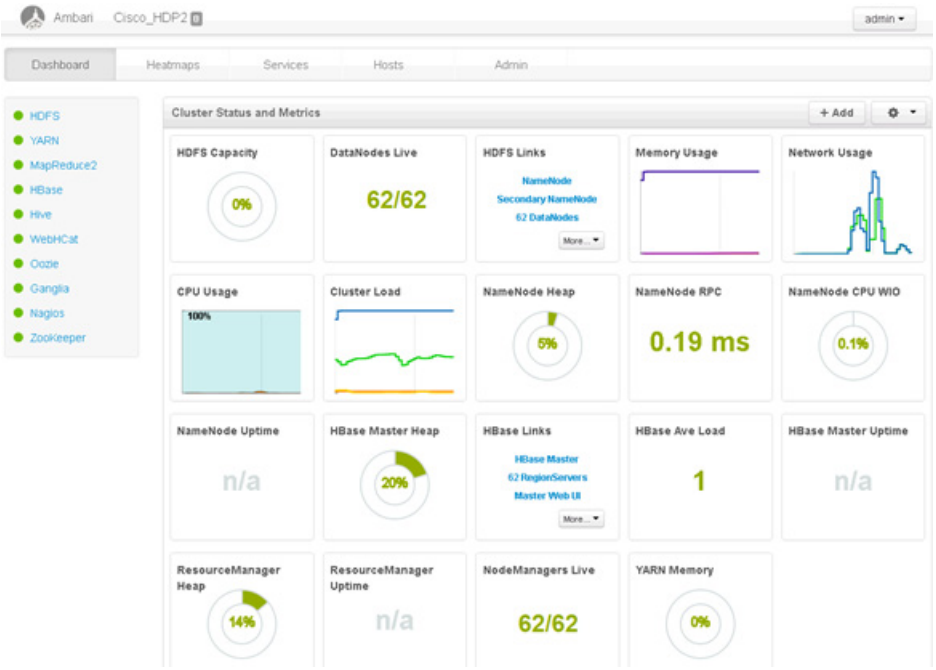
When successfully installed and started the services appears, click Next.



Summary of Installation Process

The Summary page gives a summary of the accomplished tasks. Click Complete.





Conclusion

Hadoop has evolved into a leading data management platform across all verticals. The Cisco CPAv2 for Big Data for HDP 2.0 offers a dependable deployment model for enterprise Hadoop that offer 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 160 servers (10 racks) can be supported with no additional switching in a single UCS domain. Each additional rack requires two Cisco Nexus 2232PP 10GigE Fabric Extenders and 16 Cisco UCS C240 M3 Rack-Mount Servers. Scaling beyond 10 racks (160 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 the 64 node Performance and Capacity Balanced Cluster.

Table 7 Bill of Materials for Base Rack

Part Number	Description	Quantity
UCS-SL-CPA2-PC	Performance and Capacity Balanced Cluster	1
UCSC-C240-M3S	UCS C240 M3 SFF w/o CPU mem HD PCIe w/ rail kit expdr	16
UCS-RAID9271CV-8I	MegaRAID 9271CV with 8 internal SAS/SATA ports with Supercap	16

UCSC-PCIE-CSC-02	Cisco VIC 1225 Dual Port 10Gb SFP+ CNA	16
CAB-9K12A-NA	Power Cord 125VAC 13A NEMA 5-15 Plug North America	32
UCSC-PSU2-1200	1200W 2u Power Supply For UCS	32
UCSC-RAIL-2U	2U Rail Kit for UCS C-Series servers	16
UCSC-HS-C240M3	Heat Sink for UCS C240 M3 Rack Server	32
UCSC-PCIF-01F	Full height PCIe filler for C-Series	48
UCS-CPU-E52660B	2.20 GHz E5-2660 v2/95W 10C/25MB Cache/DDR3 1866MHz	128
UCS-MR-1X162RZ-A	16GB DDR3-1866-MHz RDIMM/PC3-14900/dual rank/x4/1.5v	256
UCS-HD1T7KS2-E	1TB SAS 7.2K RPM 2.5 inch HDD/hot plug/drive sled mounted	384
UCS-SL-BD-FI96	Cisco UCS 6296 FI w/ 18p LIC, Cables Bundle	2
N2K-UCS2232PF	Cisco Nexus 2232PP with 16 FET (2 AC PS, 1 FAN (Std Airflow)	2
SFP-H10GB-CU3M=	10GBASE-CU SFP+ Cable 3 Meter	28
RACK-UCS2	Cisco R42610 standard rack w/side panels	1
RP208-30-1P-U-2=	Cisco RP208-30-U-2 Single Phase PDU 20x C13 4x C19 (Country Specific)	2
CON-UCW3-RPDUX	UC PLUS 24X7X4 Cisco RP208-30-U-X Single Phase PDU 2x (Country Specific)	6

Table 8 Bill of Materials for Expansion Racks

Part Number	Description	Quantity
UCSC-C240-M3S	UCS C240 M3 SFF w/o CPU mem HD PCIe w/ rail kit expdr	48
UCS-RAID9271CV-8I	MegaRAID 9271CV with 8 internal SAS/SATA ports with Supercap	48
UCSC-PCIE-CSC-02	Cisco VIC 1225 Dual Port 10Gb SFP+ CNA	48
CAB-9K12A-NA	Power Cord 125VAC 13A NEMA 5-15 Plug North America	96
UCSC-PSU2-1200	1200W 2u Power Supply For UCS	96
UCSC-RAIL-2U	2U Rail Kit for UCS C-Series servers	48
UCSC-HS-C240M3	Heat Sink for UCS C240 M3 Rack Server	96
UCSC-PCIF-01F	Full height PCIe filler for C-Series	144
UCS-CPU-E52660B	2.20 GHz E5-2660 v2/95W 10C/25MB Cache/DDR3 1866MHz	96
UCS-MR-1X162RZ-A	16GB DDR3-1866-MHz RDIMM/PC3-14900/dual rank/x4/1.5v	768
UCS-HD1T7KS2-E	1TB SAS 7.2K RPM 2.5 inch HDD/hot plug/drive sled mounted	1152
N2K-UCS2232PF	Cisco Nexus 2232PP with 16 FET (2 AC PS, 1 FAN (Std Airflow)	6
CON-SNTP-UCS2232	SMARTNET 24X7X4 Cisco Nexus 2232PP	6
SFP-H10GB-CU3M=	10GBASE-CU SFP+ Cable 3 Meter	84
RACK-UCS2	Cisco R42610 standard rack w/side panels	3

RP208-30-1P-U-2=	Cisco RP208-30-U-2 Single Phase PDU 20x C13 4x C19 (Country Specific)	6
CON-UCW3-RPDUX	UC PLUS 24X7X4 Cisco RP208-30-U-X Single Phase PDU 2x (Country Specific)	18

Table 9 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 10 Hortonworks Software

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