FlexPod Select with Hortonworks Data Platform (HDP)

Last Updated: December 16, 2013

Cisco Validated Design

Building Architectures to Solve Business Problems

About the Authors

Raghunath Nambiar

Raghunath Nambiar, Strategist, Data Center Solutions, Cisco Systems

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

Ajay Singh

Ajay Singh, Director, Technology Alliances, Hortonworks

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

Prem Jain

Prem Jain, Senior Solutions Architect, Big Data team, NetApp Systems

Prem Jain is a Senior Solutions Architect with the NetApp big data team. Prem's 20+ year career in technology is comprised of solution development for data migration, virtualization, HPC and big data initiatives. He has architected innovative big data and data migration solutions and authored several reference architectures and technical white papers.

Karthik Kulkarni

Karthik Kulkarni, Technical Marketing Engineer, Data Center Solution, Cisco Systems

Karthik Kulkarni is a Technical Marketing Engineer at Cisco Data Center Business Group focusing on Big Data and Hadoop technologies.

Acknowledgments

The authors acknowledge the contributions of Manankumar Trivedi, Nicolas Panayiotou, Nathan Reynolds, and Ashwin Manjunatha in developing this document.

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

FlexPod Select with Hortonworks Data Platform (HDP)

Overview

Apache Hadoop, a software framework is gaining importance in IT portfolios. The FlexPod Select for Hadoop is an extension of FlexPod initiative built based on Cisco Common Platform Architecture (CPA) for Big Data for deployments that need enterprise class external storage array features. The solution offers a comprehensive analytic stack for big data that includes compute, storage, connectivity, enterprise Hadoop distribution with a full range of services to manage heavy workloads. The offer is a pre-validated solution for enterprise Hadoop deployments with breakthroughs around Hadoop stability, operations, and storage efficiency. By integrating all the hardware and software components and using highly reliable products, businesses can meet their tight SLAs around data performance while reducing the risk of deploying Hadoop.

Audience

The intended audience of this document includes, but is not limited to, sales engineers, field consultants, professional services, IT managers, partner engineering, and customers who want to deploy FlexPod Select for Hadoop with Hortonworks.

Big Data Challenges and Opportunities

Big data is defined as data that is so high in volume and high in speed that it cannot be affordably processed and analyzed using traditional relational database tools. Typically, machine generated data combined with other data sources creates challenges for both businesses and their IT organizations. With data in organizations growing explosively and most of that new data unstructured, companies and their IT groups are facing a number of extraordinary issues related to scalability and complexity.

Lines of business are motivated by top line business benefits to work on unsolvable or unaffordable problems involving machine generated data, often combined with other traditional data sources. They exploit big data to derive competitive advantage, provide better customer experiences and help make decisions faster. Big data can be used to prevent fraud, improve business logistics by correlating buyer

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

behavior with inventory, correlate patient treatments to their cures, improve homeland security and government intelligence, cross correlating very huge data sets from credit card transactions, RFID scans, video surveillance, and many other sources. More specifically to cater to the big data needs, an Apache Hadoop workload or cluster is required.

Big data is more about business opportunities than reducing costs. To address these challenges and risks of big data, companies need analytical solutions that meet the following criteria:

- Provide resilient and reliable storage for Hadoop.
- Implement high-performance Hadoop clusters.
- Build on an open partner-based, ecosystem.
- Allow efficient Hadoop clustering.
- Scale compute and storage independently and quickly as data grows in volume.
- Cost effectiveness.

The FlexPod Select for Hadoop is designed to address these challenges.

FlexPod Select for Hadoop Benefits

The FlexPod Select for Hadoop combines leading edge technologies from Cisco and NetApp to deliver a solution that exceeds the requirements of emerging big data analytics so that businesses can manage, process, and unlock the value of new and large volume data types that they generate. Designed for enterprises in data-intensive industries with business critical SLAs, the solution offers pre-sized storage, networking, and compute in a highly reliable, ready to deploy Apache Hadoop stack.

The key benefits of this solution are described in Table 1.

Table 1 Key benefits of FlexPod Select for Hadoop

Enterprise-Class Big Data Architecture	Accelerate Time to Value	Co-existence with Enterprise Applications
• Easy manageability, more reliability, scalability and high performance. • Fully redundant architecture. • Superior reliability and stability. • Lower cluster downtime. • Faster recovery from drive failure. • Fewer copies of Hadoop data means less storage to manage, higher storage efficiency. • Dynamically add capacity to as data grows, expand storage while cluster is running. • Protection of namenode Single Point of Failure.	• Reduced risk, better power and floor space foot print, pre-validated solution. • Validated, pre-tested, reference architecture (Cisco Verified Design). • Pre-sized, leading-edge storage, compute, networking with Hadoop, Hortonworks Data Platform (HDP). • Higher performance with faster interconnects, lower latency and less network congestion. • Well-established FlexPod channel. • Independent scaling of compute and storage or scale together. • Allocate more/less storage capacity to compute node.	• Seamless data and management integration with enterprise applications and traditional FlexPod deployments. • Global support and services. • Open analytical stack for higher interoperability within infrastructure. • Hadoop handles data well, in any kind of schema. • Open solution with best in class components. • Proven at PB scale. • Lower TCO, less rack-space needed, lower power. required (180TB in 4U).

Hortonworks Data Platform (HDP)

The Hortonworks Data Platform (HDP) 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.

FlexPod Select for Hadoop with Hortonworks Architecture

This section provides an architectural overview on the FlexPod Select for Hadoop with Hortonworks. In this section you will find information on solution components and their configuration brief:

- Solution Overview, page 9
- Configuration Overview, page 11

Solution Overview

Building upon the success of FlexPod, market leaders, Cisco and NetApp deliver the enterprise-class solution FlexPod Select for Hadoop with a pre-validated, faster Time to Value (*TtV) Hadoop solution for enterprises that provides control of and insights from big data. The FlexPod Select for Hadoop is based on a highly scalable architecture, that can scale from single rack to multiple racks, built using the following components:

*TtV is the time to realize a quantifiable business goal.

Connectivity and Management

- Cisco UCS 6200 Series Fabric Interconnects provide high speed, low latency connectivity for servers and centralized management for all connected devices with UCS Manager. Deployed in redundant pairs they offer full redundancy, performance (active-active), and exceptional scalability for large number of nodes typical in big data clusters. UCS Manger enables rapid and consistent server integration using service profile, ongoing system maintenance activities such as firmware update operations across the entire cluster as a single operation, advanced monitoring, and option to raise alarms and send notifications about the health of the entire cluster.
- Cisco Nexus 2200 Series Fabric Extenders, act as remote line cards for Fabric Interconnects providing a highly scalable and extremely cost-effective connectivity for large number of nodes.
- 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 GUI, a command-line interface (CLI), or an XML API. Cisco UCS Manager uses service profiles to define the personality, configuration, and connectivity of all resources within Cisco UCS, radically simplifying provisioning of resources so that the process takes minutes instead of days. This simplification allows IT departments to shift their focus from constant maintenance to strategic business initiatives. It also provides the most streamlined, simplified approach commercially available today for updating firmwares of all server components.

Compute

- Cisco UCS C220M3 Rack-Mount Servers, 2-socket server based on Intel® Xeon® E-2600 series processors optimized for performance and density. This server is expandable to 512 GB of main memory and has 8 small factor internal front-accessible, hot-swappable disk drives and two PCIe Gen 3.0 slots.
- Cisco UCS Virtual Interface Card 1225, unique to Cisco UCS is a dual-port PCIe Gen 2.0 x8 10-Gbps adapter designed for unified connectivity for Cisco UCS C-series Rack-Mount Servers.

Storage

- NetApp E5460 storage array provides increased performance and bandwidth for Hadoop clusters along with higher storage efficiency and scalability.
- NetApp FAS2220 and the Data ONTAP storage operating system provides high reliability for Hadoop with reduced single points of failure, faster recovery time and namenode metadata protection with hardware RAID.

Software

- Hortonworks Data Platform (HDP) is 100% open source distribution for Apache Hadoop. Tested at scale, leading enterprises rely on HDP for it’s modern data architecture.
- Apache Ambari is the fully open sourced Hadoop management platform included in HDP that allow you to provision, manage and monitor a Hadoop cluster of any size.
- Red Hat® Enterprise Linux® Server, the leading enterprise Linux distribution.

Configuration Overview

The solution is offered in a single rack and in multiple racks. The architecture consists of:

- Master Rack, page 11
- Expansion Rack, page 12

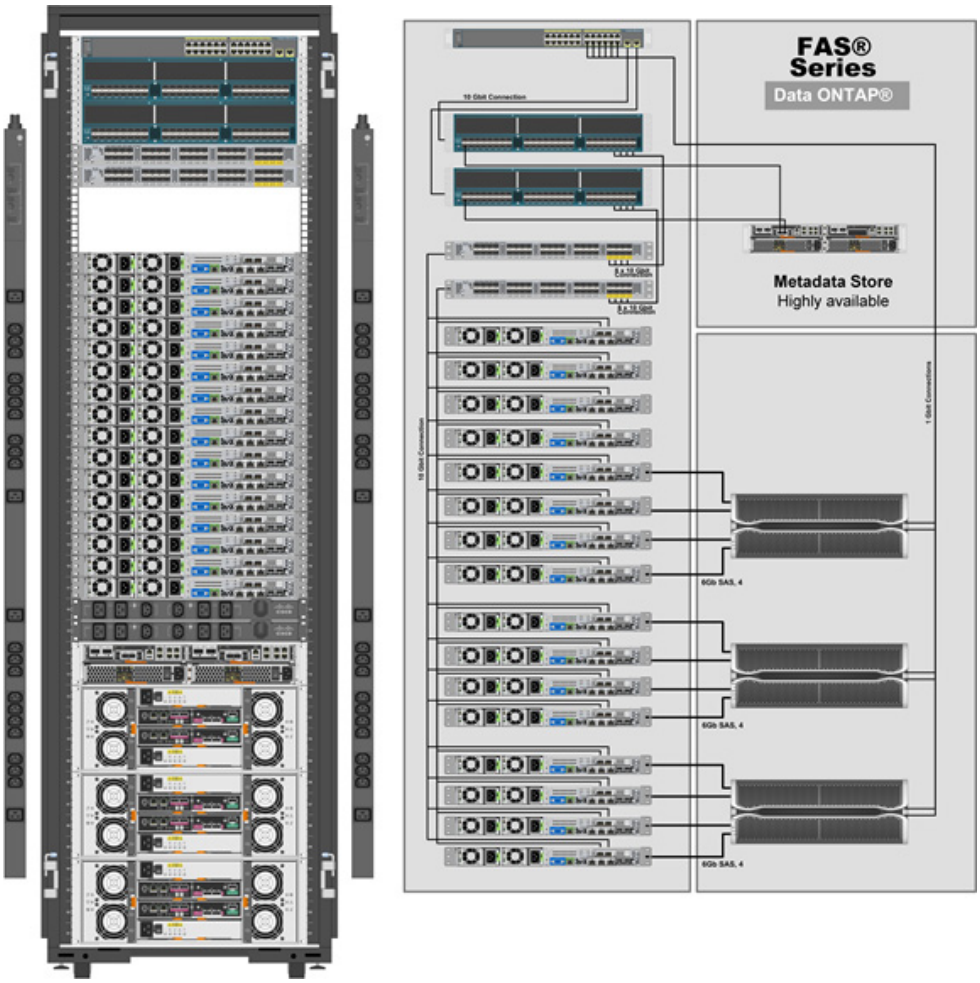
Master rack (single rack) is a standalone solution. The multiple rack solution consists of a master rack and one or more expansion racks. In a single UCS management domain, up to 9 expansion racks are supported. Deployments requiring more than 10 racks can be implemented by interconnecting multiple domains using Cisco Nexus 6000/7000 series switches and managed by UCS Central. Figure 1 shows the FlexPod Select for Hadoop master rack model.

Master Rack

The master rack consists of the following:

- Two Cisco UCS 6296UP Fabric Interconnects
- Two Cisco Nexus 2232PP Fabric Extenders
- Sixteen Cisco UCS C220M3 Rack-Mount Servers
- One Cisco Catalyst 2960S
- One NetApp FAS2220
- Three NetApp E5460
- Two vertical PDUs
- Two horizontal PDUs
- Cisco 42U Rack

Figure 1 Cisco Master Rack

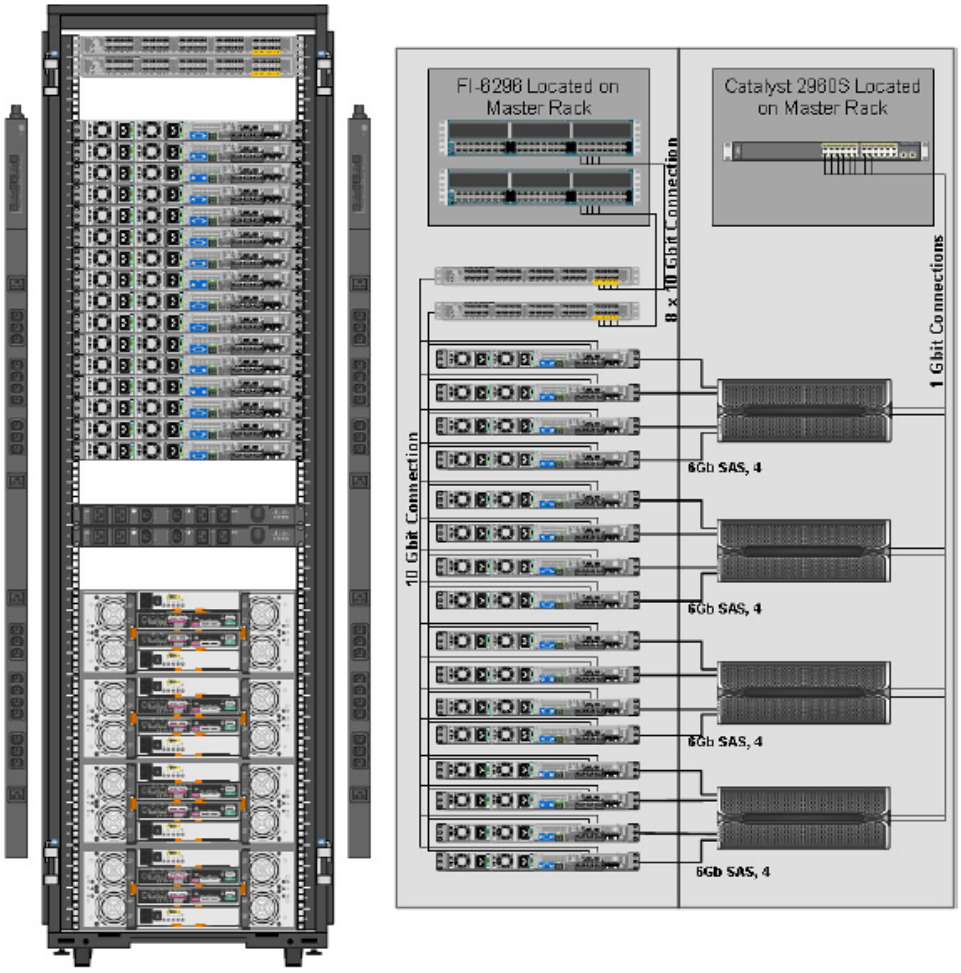


Expansion Rack

Figure 2 shows the FlexPod Select for Hadoop expansion rack model. The expansion rack consists of the following:

- Two Cisco Nexus 2232PP Fabric Extenders
- Sixteen UCS C220M3 Rack-Mount Servers
- Four NetApp E5460
- Two vertical PDUs
- Two horizontal PDUs
- Cisco 42U Rack

Figure 2 Cisco Expansion Rack



Rack and PDU Configuration

The rack configurations of the master rack and expansion rack are shown in Table 2 based on a Cisco 42U rack.

Table 2 Rack configuration details

Cisco 42U Racks	Master Rack	Expansion Rack
1	Cisco UCS FI 6296UP	
2		
3	Cisco UCS FI 6296UP	
4		
5	Cisco Nexus FEX 2232PP	Cisco Nexus FEX 2232PP
6	Cisco Nexus FEX 2232PP	Cisco Nexus FEX 2232PP

Table 2 Rack configuration details

Cisco 42U Racks	Master Rack	Expansion Rack
7	Cisco UCS C220M3	Cisco UCS C220M3
8	Cisco UCS C220M3	Cisco UCS C220M3
9	Cisco UCS C220M3	Cisco UCS C220M3
10	Cisco UCS C220M3	Cisco UCS C220M3
11	Cisco UCS C220M3	Cisco UCS C220M3
12	Cisco UCS C220M3	Cisco UCS C220M3
13	Cisco UCS C220M3	Cisco UCS C220M3
14	Cisco UCS C220M3	Cisco UCS C220M3
15	Cisco UCS C220M3	Cisco UCS C220M3
16	Cisco UCS C220M3	Cisco UCS C220M3
17	Cisco UCS C220M3	Cisco UCS C220M3
18	Cisco UCS C220M3	Cisco UCS C220M3
19	Cisco UCS C220M3	Cisco UCS C220M3
20	Cisco UCS C220M3	Cisco UCS C220M3
21	Cisco UCS C220M3	Cisco UCS C220M3
22	Cisco UCS C220M3	Cisco UCS C220M3
23		
24	PDU	PDU
25	PDU	PDU
26	NetApp FAS 2220	
27		NetApp E5460
28	Cisco Catalyst 2960S	
29		
30		
31	NetApp E5460	NetApp E5460
32		
33		
34		
35	NetApp E5460	NetApp E5460
36		
37		
38		
39	NetApp E5460	NetApp E5460
40		
41		
42		

The configuration consists of two vertical PDUs and two horizontal PDUs. The Cisco UCS 6296UP Fabric Interconnects, NetApp E5460s and NetApp FAS2220 are connected to each of the horizontal PDUs. The Cisco Nexus 2232PP Fabric Extenders and Cisco UCS C220M3 Servers are connected to each of the vertical PDUs for redundancy; thereby, ensuring availability during power source failure.

Note Contact your Cisco representative for country specific information.

Fabric Configuration

The master rack configuration consists of two Cisco UCS 6296UP Fabric Interconnects and two Cisco Nexus Fabric Extender 2232PP forming two fabrics, Fabric A and Fabric B topology. The Cisco UCS C220M3 Servers 1 to 16 are connected to Fabric A and Fabric B using 10Gb Ethernet connectivity through Cisco Nexus 2232PP Fabric Extenders, with eight uplinks.

The configuration details of the master rack and expansion racks are shown in Figure 1 and Figure 2 respectively.

Storage Configuration

NetApp E5460 belong to the NetApp E5400 modular data storage system family that support big-bandwidth datasets requiring high sequential throughput. The NetApp E5460s are configured with dual SAS controllers and 60 3TB 7.2K RPM SAS disk drives.

For more information, see:

http://www.netapp.com/us/products/storage-systems/e5400/index.aspx

NetApp FAS2200 offers powerful, affordable, flexible data storage for midsized businesses and distributed enterprises. The NetApp FAS2220 has 6 drives (600GB, 10K rpm, SAS) and 4 x 1GbE ports and 2 x 10GbE ports.

For more information, see:

http://www.netapp.com/us/products/storage-systems/fas2200/

Server Configuration and Cabling

Figure 3 illustrates the physical connectivity of Cisco UCS 6296UP Fabric Interconnects, Cisco Nexus 2232PP Fabric Extenders, and Cisco UCS C220M3 Servers.

Figure 3 Cisco Hardware Connectivity

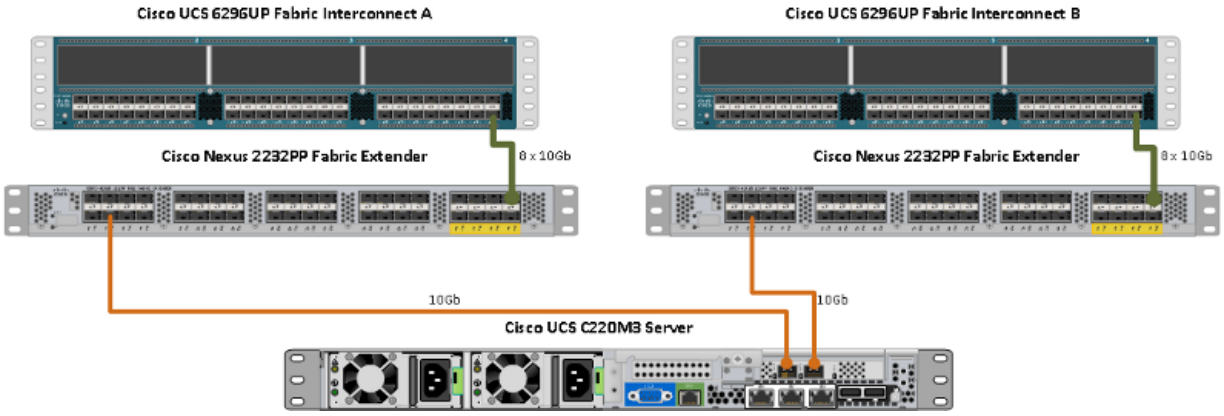
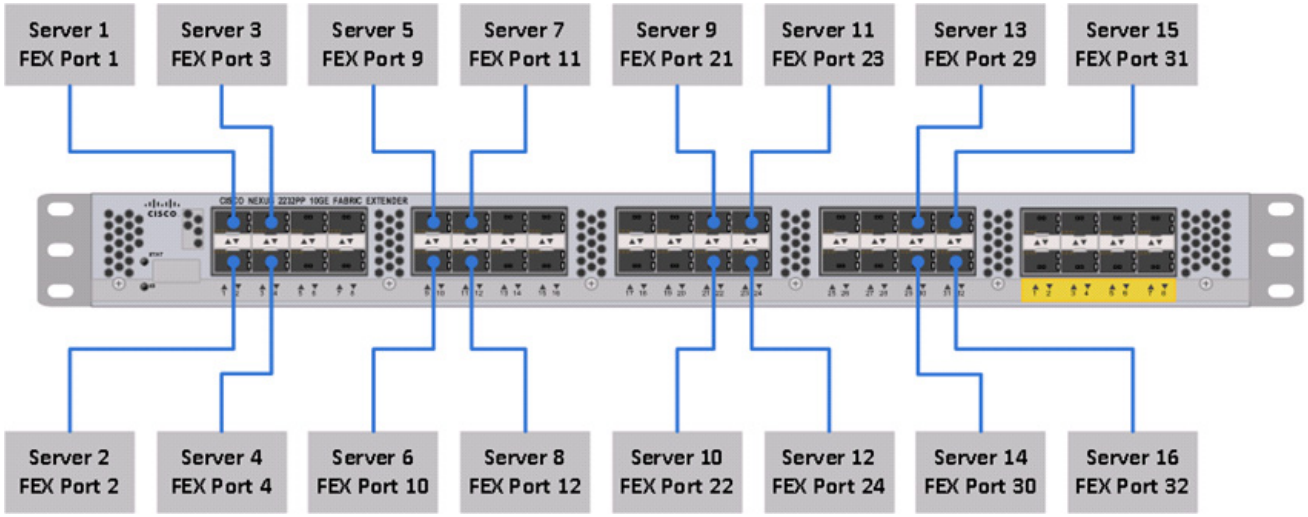


Figure 4 shows the ports of the Cisco Nexus 2232PP Fabric Extender connecting the Cisco UCS C220M3 Servers. Sixteen Cisco UCS C220M3 Servers are used in the master and expansion rack configurations offered by the FlexPod Select for Hadoop.

Figure 4 Connectivity Diagram of Cisco Nexus 2232PP FEX and Cisco UCS C220M3 Servers



Note Cisco UCS Manager version used for this deployment is UCS 2.1(1e).

For more information on configuring 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

Software Requirements

For this deployment we have used Hortonworks Data Platform and Red Hat Enterprise Linux Server.

Hortonworks Data Platform (HDP)

The Hortonworks Data Platform supported is HDP 1.3. For more information, see:
http://www.hortonworks.com

RHEL

The operating system supported is Red Hat Enterprise Linux Server 6.2. For more information on the Linux support, see:

www.redhat.com.

Software Versions

Table 3 provides the software version details of all the software requirements for this model.

Table 3 Software version details

Layer	Components	Version or Release	Details
Compute	Cisco UCS C220M3	C220M3.1.4.7b.0.1005 20120256	Hardware BIOS version
Network	Cisco UCS 6296UP	UCS 2.1(1e)	Fabric Interconnect
	Cisco Nexus 2232PP	5.1(3)N2(2.11a)	Fabric Extender
Storage	NetApp FAS 2220	Data ONTAP 8.1.2 7-mode	FAS Storage
	NetApp E35460	07.84	E-Series Storage
Software	Red Hat Enterprise Linux Server	6.2 (x86_64)	Linux Distribution
	Cisco UCSM	2.1(1e)	UCS Embedded Management Software
	NetApp OnCommand System Manager	2.1	FAS Management Software
	NetApp SANtricity	10.84	E-Series Management Software
	HDP	1.3	Hortonworks Data Platform

Fabric Configuration

This section provides details for configuring a fully redundant, highly available configuration for a FlexPod Select for Hadoop. Follow these steps to configure Cisco 6296UP Fabric Interconnect.

1. Configure FI A
2. Configure FI B
3. Connect to IP address of FI A using web browser. Launch Cisco UCS Manger
4. Edit the chassis discovery policy.
5. Enable server and Uplink Ports
6. Create pools and polices for service profile template.
7. Create SP template, 16 profiles
8. Start discover process
9. Associate to server
10. FI Configuration for NetApp FAS2220

Performing an Initial Setup of Cisco UCS 6296UP Fabric Interconnects

Follow these steps for initial setup of the Cisco UCS 6296 Fabric Interconnects:

Cisco UCS 6296 FI A

1. Connect to the console port on the first Cisco UCS 6296 Fabric Interconnect.
2. At the configuration method prompt, enter **console**.
3. If asked to either do 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, enter **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 for management port on the fabric interconnect.
12. Enter the Mgmt0 IPv4 subnet mask for the management port on the fabric interconnect.
13. Enter the IPv4 address of the default gateway.
14. Enter the cluster IPv4 address.
15. To configure DNS, enter **y**.
16. Enter the DNS IPv4 address.
17. Enter **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, enter **yes** to save the configuration.

20. Wait for the login prompt to make sure the configuration is saved successfully.

Cisco UCS 6296UP FI B

1. Connect to the console port on the second Cisco UCS 6296 Fabric Interconnect.
2. At the configuration method prompt, enter **console**.
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 for the first fabric interconnects.
5. Enter the Mgmt0 IPv4 address for the management port on the subordinate fabric interconnect.
6. Enter **y** to save the configuration.
7. Wait for the login prompt to make sure the configuration is saved successfully.

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 log into Cisco UCS Manager:

1. Open a Web browser and type the Cisco UCS 6296UP Fabric Interconnect cluster address.
2. If a Security Alert dialog box appears, click **Yes** to accept the security certificate and continue.
3. In the Cisco UCS Manager launch page, click **Launch UCS Manager**.
4. When prompted, enter admin for the user name and enter the administrative password and click **Login** to log in to the Cisco UCS Manager GUI.

Upgrade Cisco UCS Manager Software to Version 2.1(1e)

This document assumes the use of UCS 2.1(1e). For more information on upgrading the software version to Cisco UCS 2.0 release, see:

http://www.cisco.com/en/US/docs/unified_computing/ucs/sw/upgrading/from2.0/to2.1/b_UpgradingCiscoUCSFrom2.0To2.1.pdf

This link provides you information on upgrading Cisco UCS Manager software and Cisco UCS 6296 Fabric Interconnect software to version 2.1(1e).

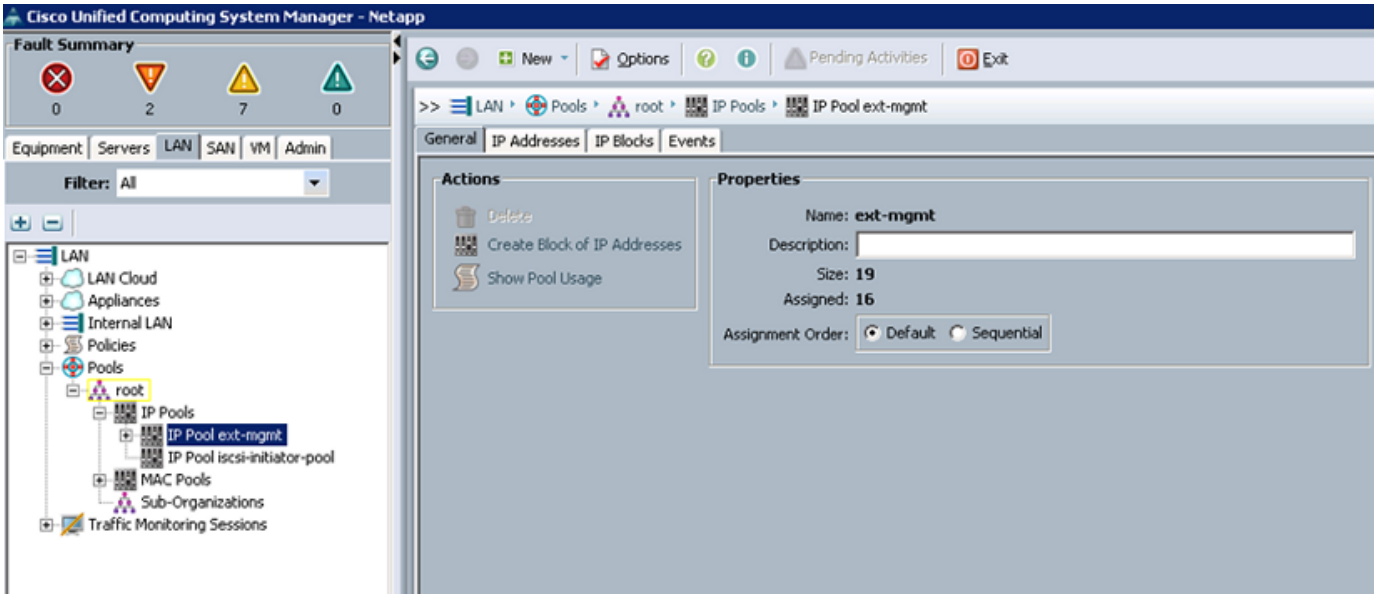
Note Make sure the Cisco UCS C-Series version 2.1(1e) software bundle is loaded on the Fabric Interconnects.

Adding a Block of IP Addresses for KVM Console

Follow these steps to create a block of KVM IP addresses for server access in the Cisco UCS Manager GUI:

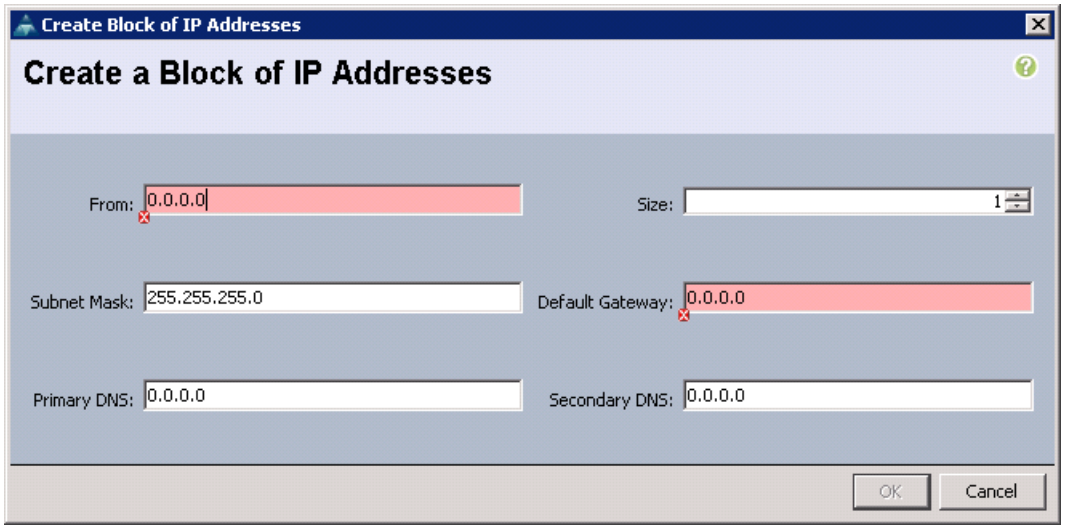
1. Select the LAN tab at the top in the left pane in the UCSM GUI.
2. Select **Pools > IP Pools > IP Pool ext-mgmt** as shown in Figure 5.

Figure 5 Management IP Pool in Cisco UCS Manager



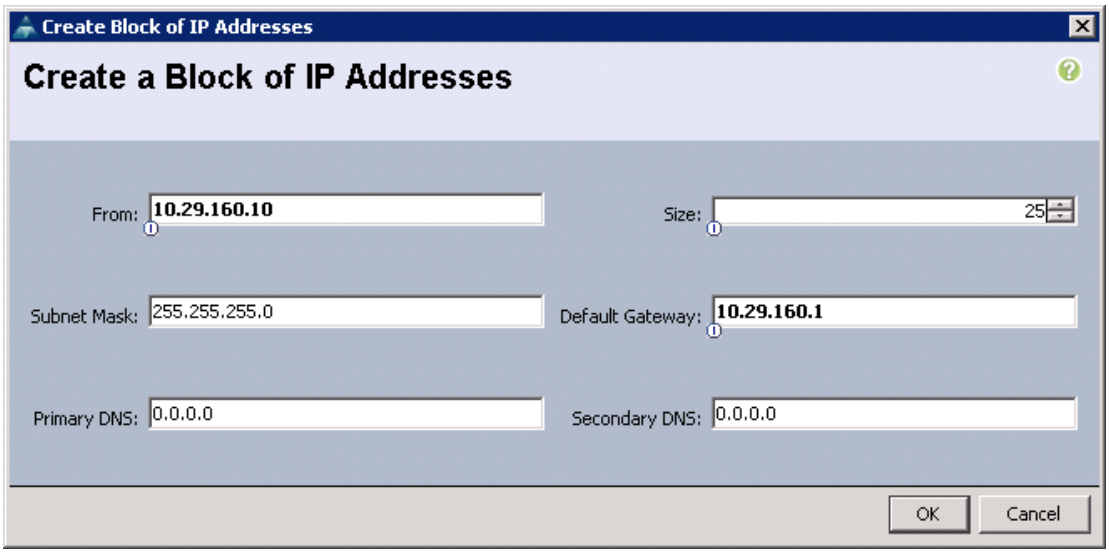
3. Right-click the IP Pool ext-mgmt.
4. Select Create Block of IP Addresses. Create Block of IP Address window appears as shown in Figure 6.

Figure 6 Creating a Block of IP Addresses



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

Figure 7 Entering the Block of IP Addresses



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

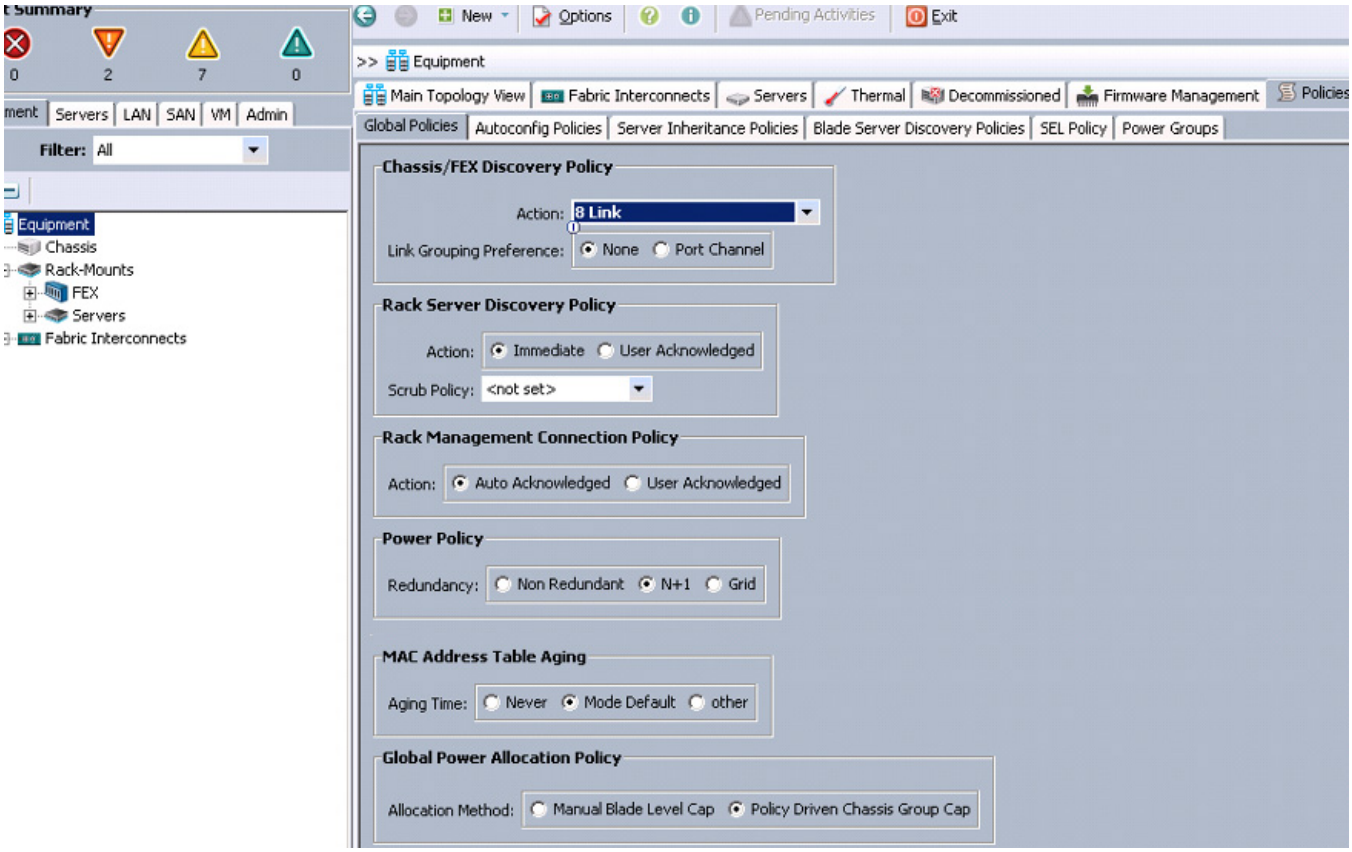
Editing the Chassis Discovery Policy

Setting the discovery policy now will simplify the addition of Cisco UCS B-Series Chassis in the future and additional fabric extenders for further C-Series connectivity.

To modify the chassis discovery policy, follow these steps:

1. Navigate to the Equipment tab in the left pane in the UCSM GUI.
2. In the right pane, select the Policies tab.
3. Under Global Policies, change the Chassis Discovery Policy to 8-link as shown in Figure 8.

Figure 8 Editing the Chassis Discovery Policy



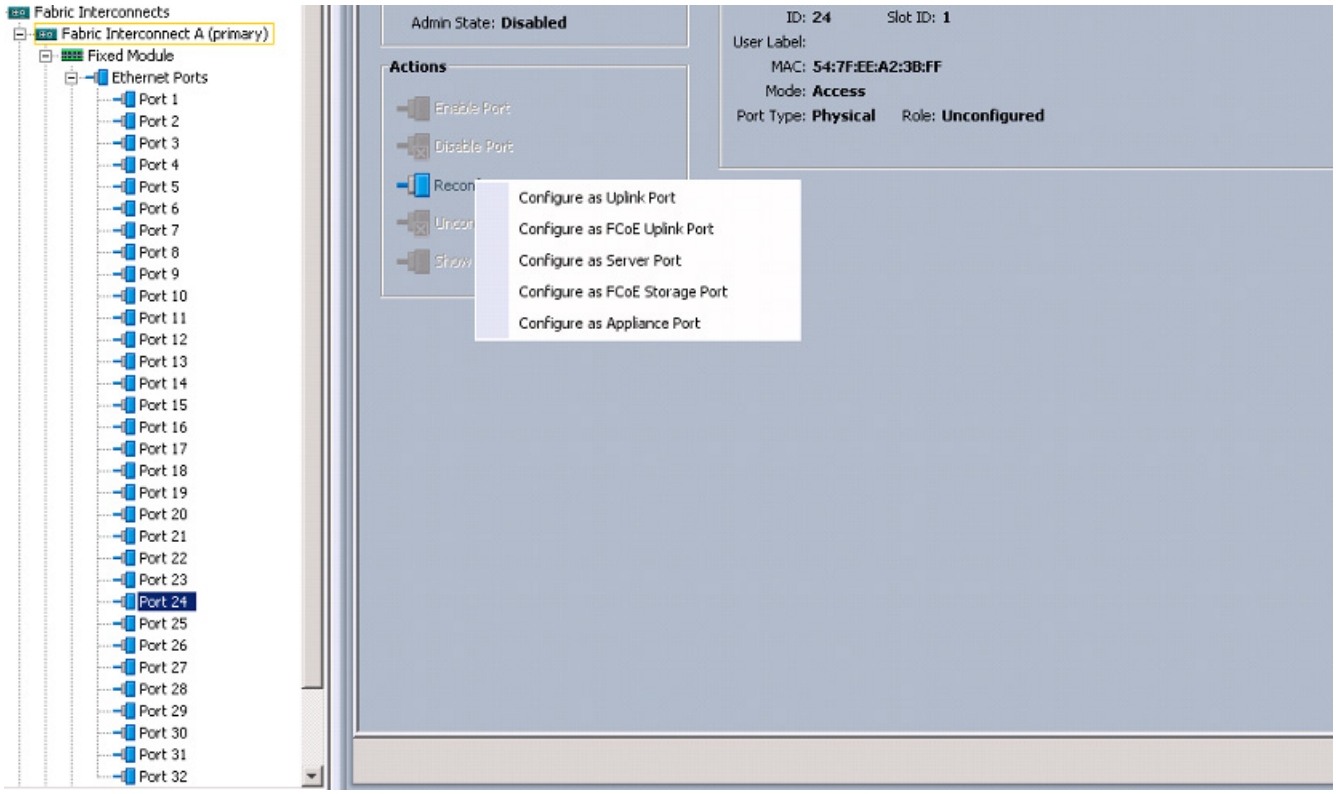
4. Click **Save Changes** in the bottom right corner in the Cisco UCSM GUI.
5. Click **OK**.

Enabling Server and Uplink Ports

To enable the server ports and uplink ports, follow these steps:

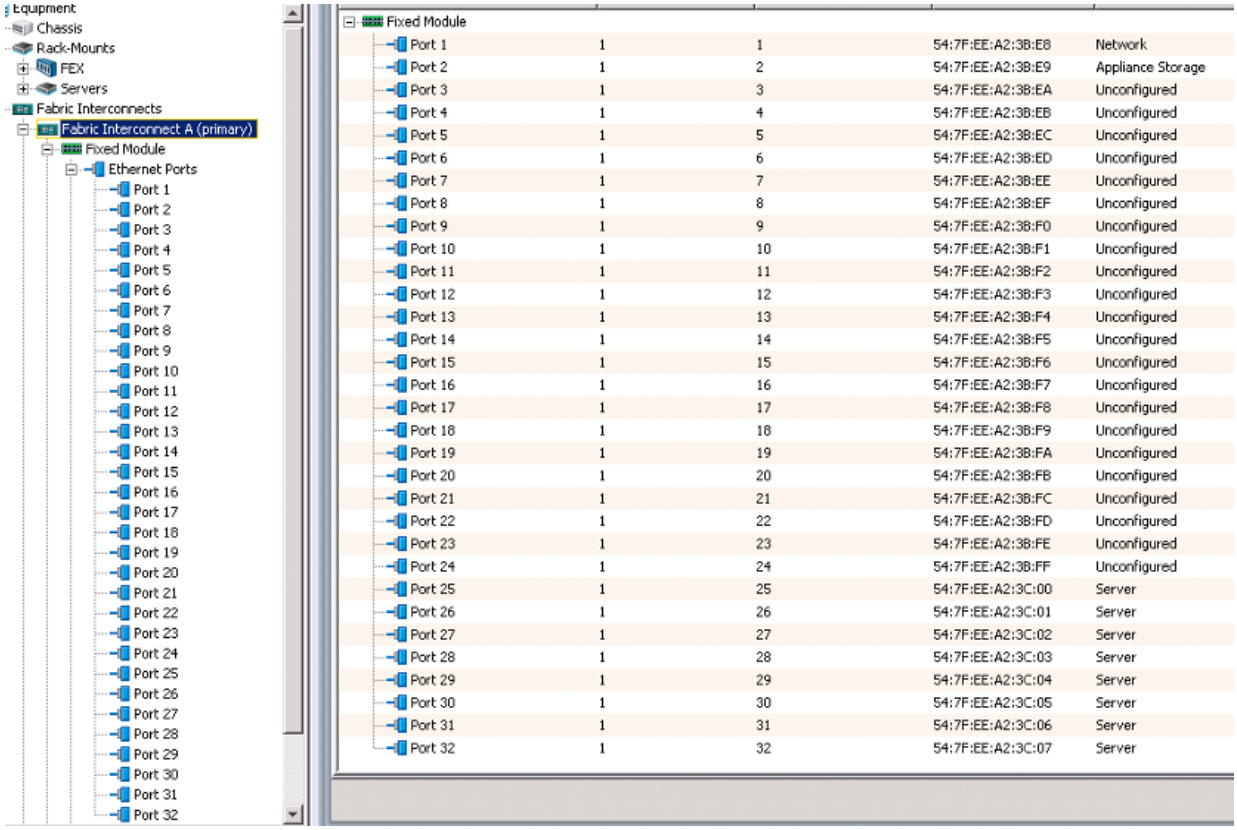
1. Select the Equipment tab on the top left corner in the left pane in the UCSM GUI.
2. Select **Equipment > Fabric Interconnects > Fabric Interconnect A (primary) > Fixed Module**.
3. Expand the Unconfigured Ethernet Ports.
4. Select the number of ports that are connected to the Cisco Nexus 2232PP FEXs (8 per FEX), right-click them, and select **Reconfigure > Configure as a Server Port** as shown in Figure 9.

Figure 9 Enabling Server Ports



5. Select port 1 that is connected to the Cisco Catalyst 2960-S switches, right-click them, and 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.
10. Select the number of ports that are connected to the Cisco Nexus 2232 FEXs (8 per FEX), right-click them, and select **Reconfigure > Configure as Server Port**.
11. A pop-up window appears to confirm your selection. Click **Yes**, then **OK** to continue.
12. Select port 1 that is connected to the Cisco Catalyst 2960-S switches, right-click and 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 10 Window Showing Server Ports and Uplink Ports



Creating Pools for Service Profile Template

Creating an Organization

Organizations are used as a means to organize 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 in the Cisco UCS Manager GUI:

1. Click **New** on the top left corner in the right pane in the UCSM 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 configure the necessary MAC address pools in the Cisco UCS Manager GUI:

1. Select the LAN tab in the left pane in the UCSM GUI.
2. Select **Pools > root**.
3. Right-click the MAC Pools under the root organization.
4. Select Create MAC Pool to create the MAC address pool.
5. Enter nosh for the name of the MAC pool.
6. (Optional) Enter a description of the MAC pool.
7. Click **Next**.
8. Click **Add**.
9. Specify a starting MAC address.
10. Specify a size of the MAC address pool sufficient to support the available server resources. See Figure 11, Figure 12, and Figure 13.

Figure 11 Specifying the First MAC Address and Size

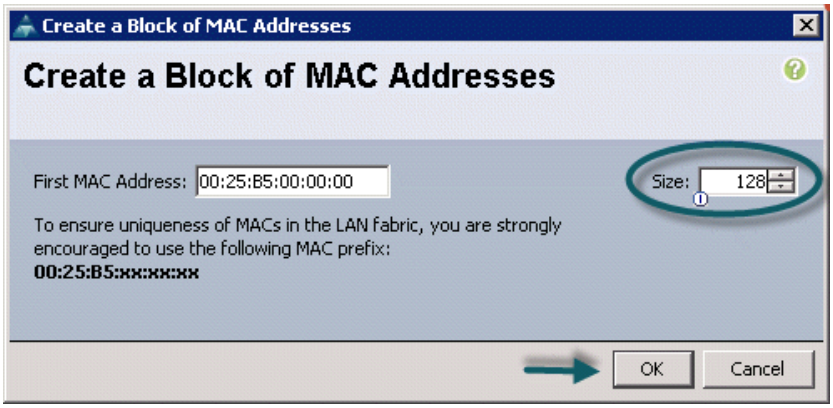


Figure 12 Range of MAC Addresses

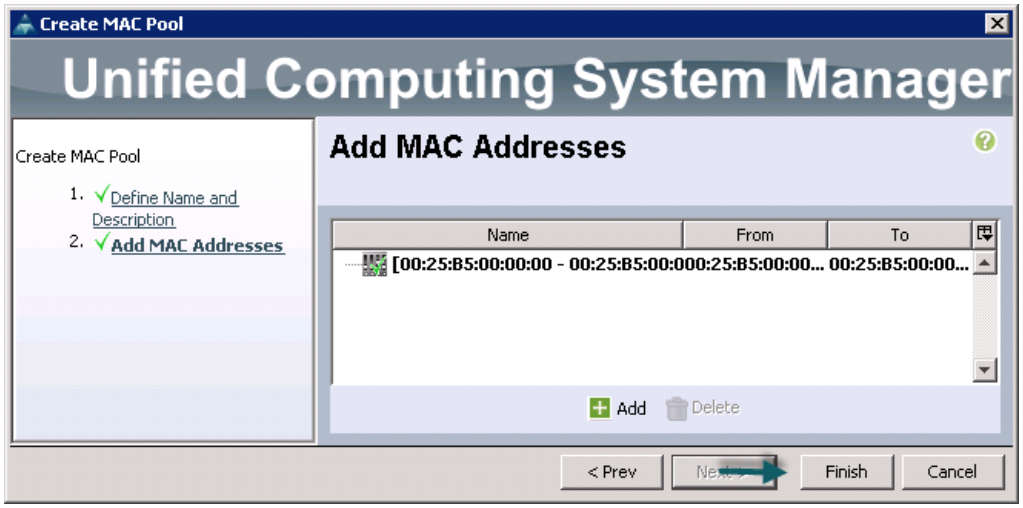
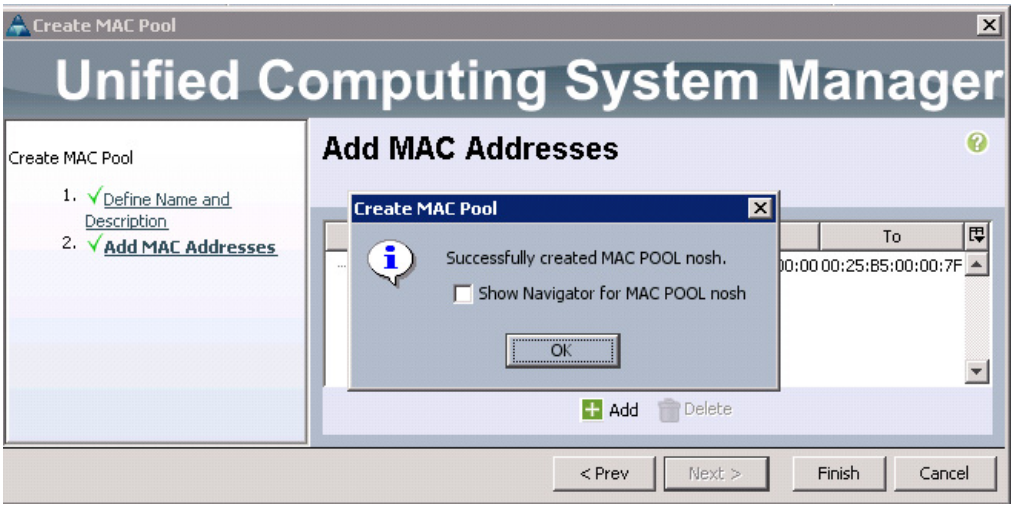


Figure 13 Created MAC Pool



11. Click **OK**.
12. Click **Finish**.
13. Click **OK** in the success message box.

Configuring VLANs

VLANs are configured as shown in Table 4.

Table 4 VLAN Configurations

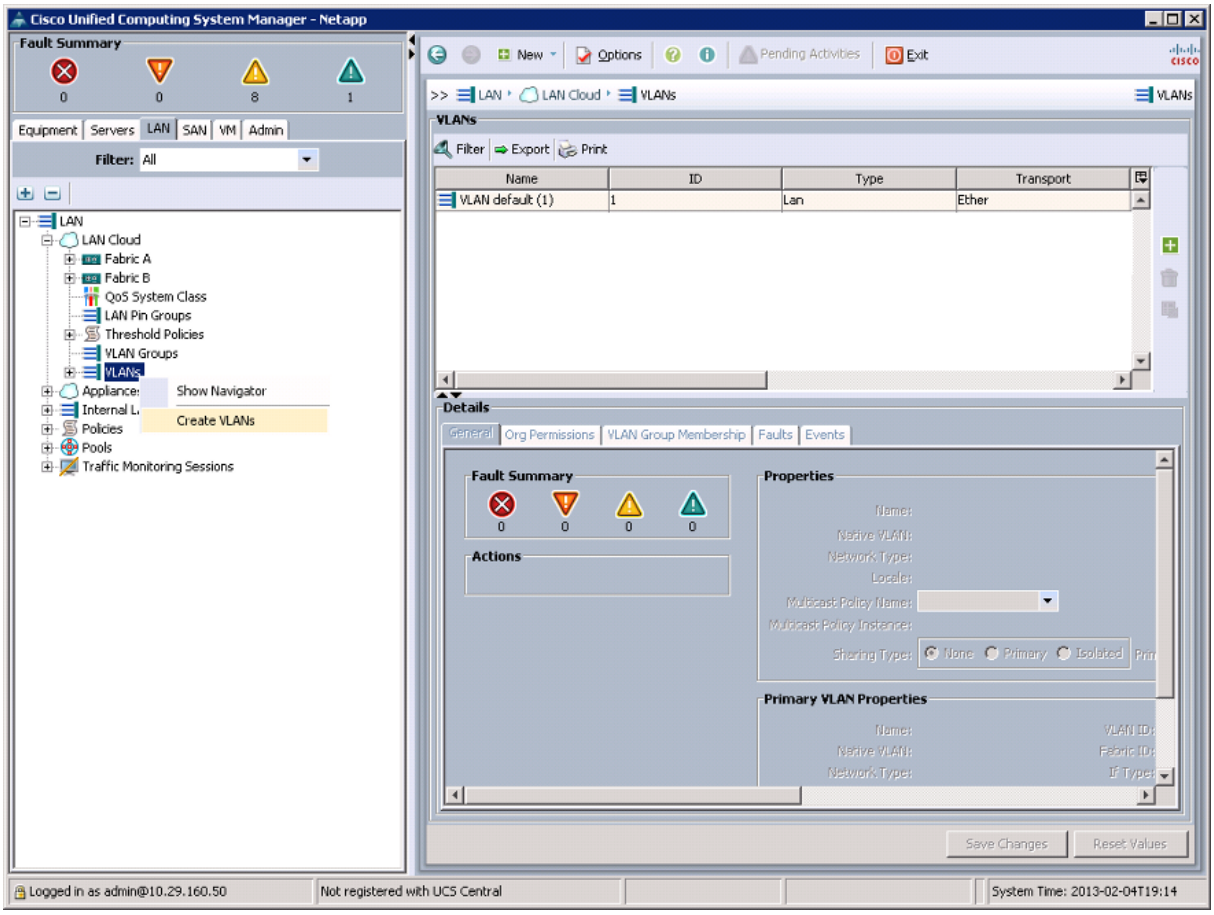
VLAN	Fabric	NIC Port	Function	Failover
vlan160_mgmt	A	eth0	Management, User connectivity	Fabric Failover to B
vlan11_NFS	A	eth1	NFS Traffic	Fabric Failover to B
vlan12_HDFS	B	eth2	HDFS Traffic	Fabric Failover to A

For this deployment we are using eth0 (vlan160_mgmt) for management packets, eth1 (vlan11_NFS) for NFS data traffic and eth2 (vlan12_HDFS) for HDFS data traffic.

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

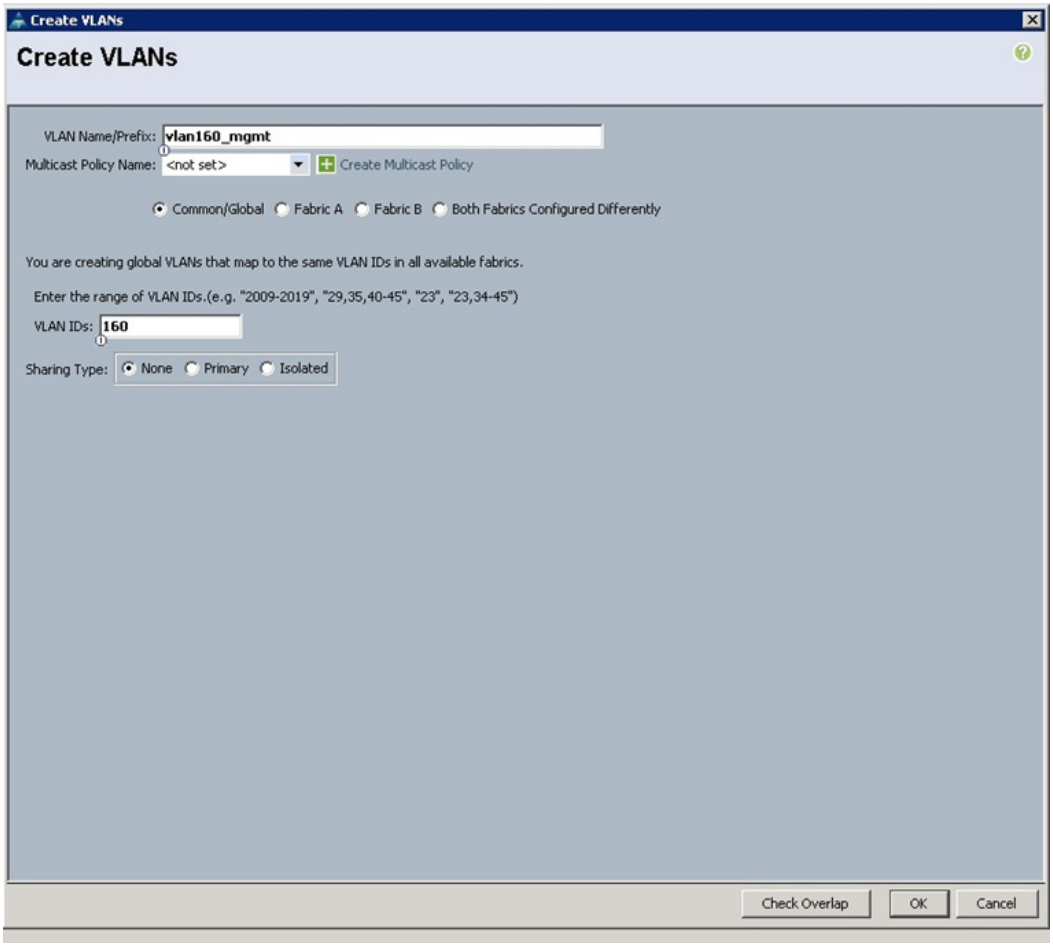
1. Select the LAN tab in the left pane in the UCSM GUI.
2. Select **LAN > VLANs**.
3. Right-click the VLANs under the root organization.
4. Select Create VLANs to create the VLAN.

Figure 14 Creating VLANs



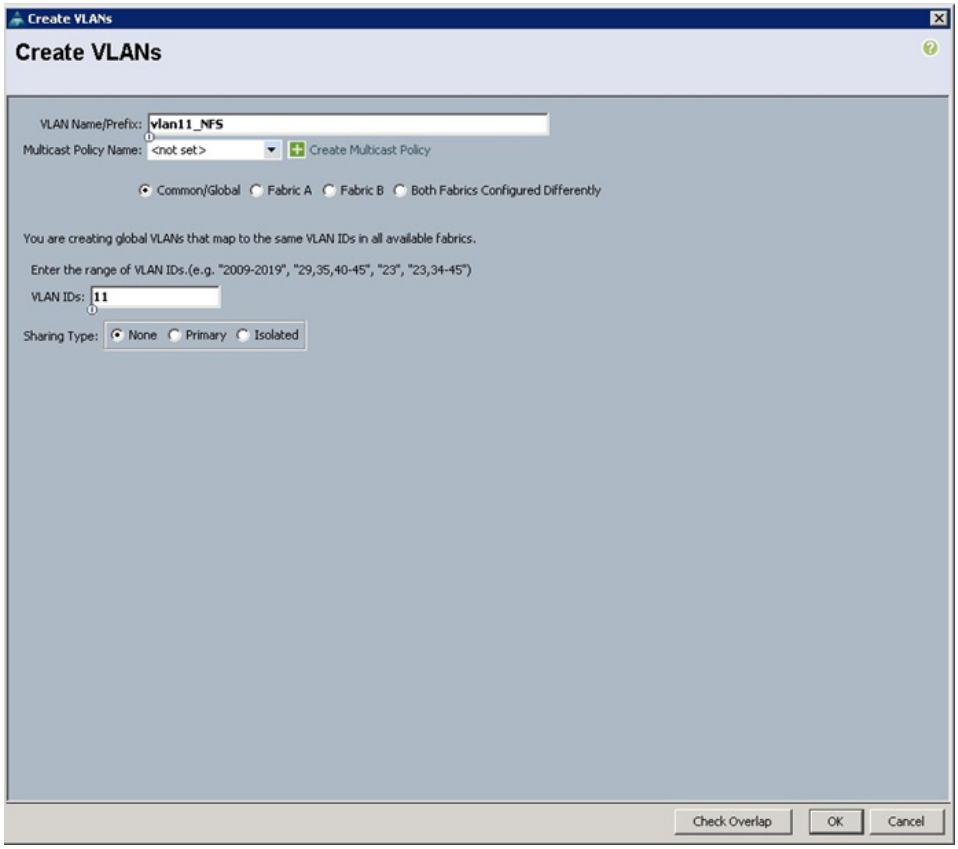
5. Enter vlan160_mgmt for the VLAN Name.
6. Select Common/Global for vlan160_mgmt.
7. Enter 160 on VLAN IDs of the Create VLAN IDs.

Figure 15 Creating VLAN for Fabric A



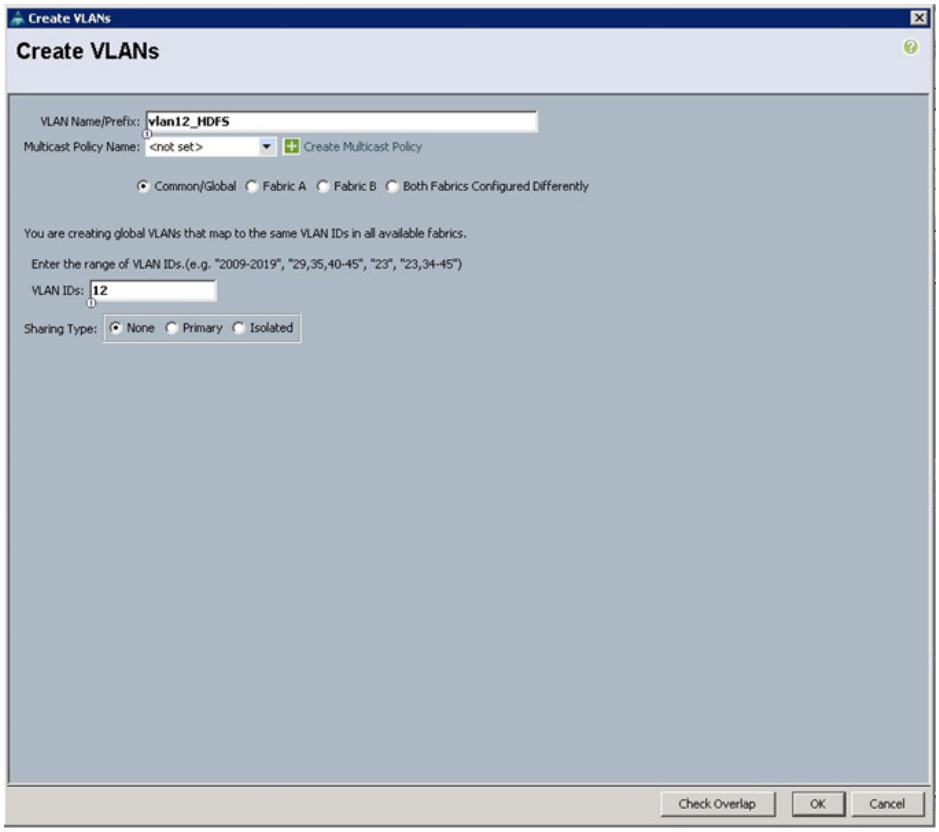
8. Click **OK** and then, click **Finish**.
9. Click **OK** in the success message box.
10. Select the LAN tab in the left pane again.
11. Select **LAN > VLANs**.
12. Right-click the VLANs under the root organization.
13. Select Create VLANs to create the VLAN.
14. Enter vlan11_NFS for the VLAN Name.
15. Select Common/Global for vlan11_NFS.
16. Enter 11 on VLAN IDs of the Create VLAN IDs.

Figure 16 Creating VLAN for Fabric B



17. Click **OK** and then, click **Finish**.
18. Click **OK** in the success message box.
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.

Figure 17 Creating Global HDFS VLAN



26. Click **OK** then click **Finish**.

Note All of the VLANs created need to be trunked to the upstream distribution switch connecting the fabric interconnects.

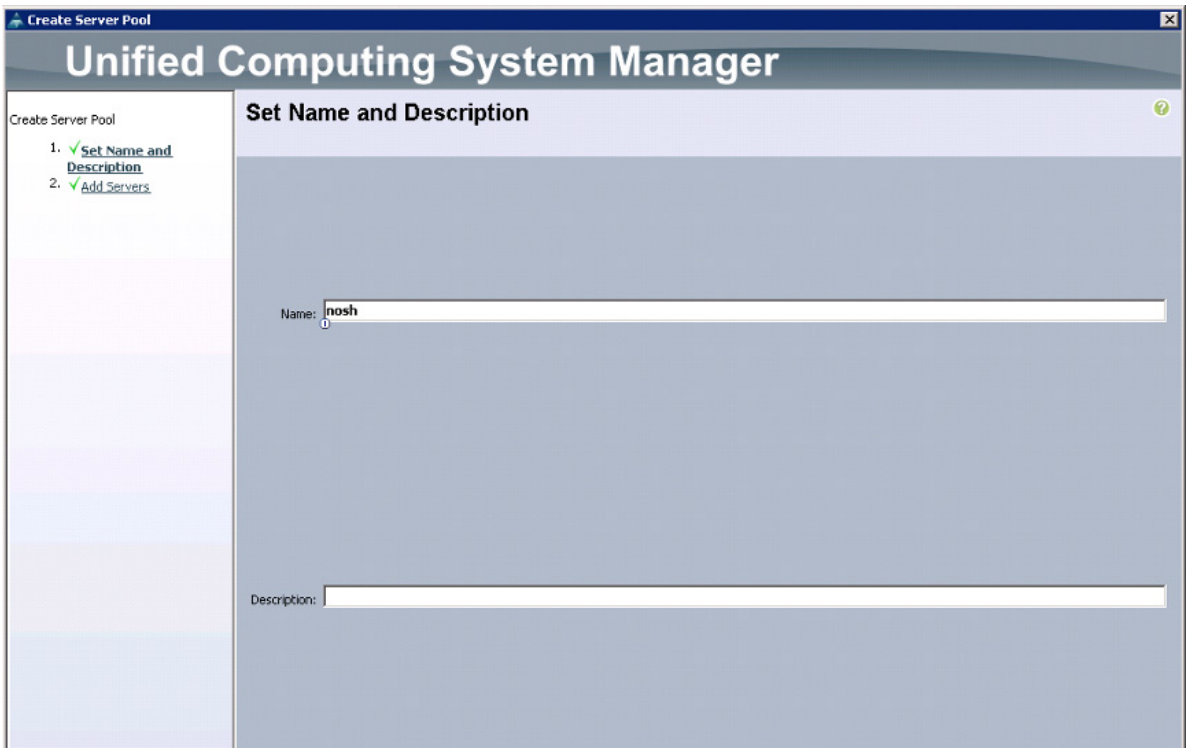
Creating Server Pool

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 in the Cisco UCS Manager GUI:

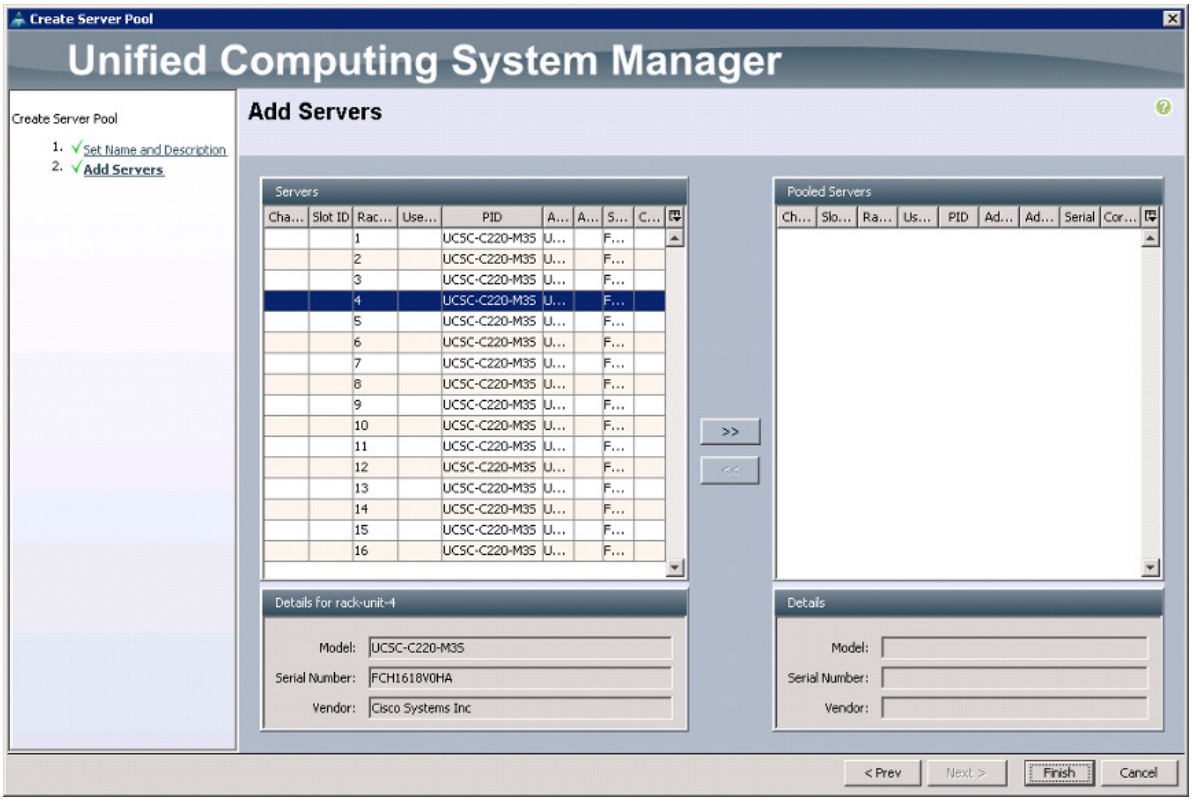
1. Select the Servers tab in the left pane in the Cisco UCS Manager GUI.
2. Select **Pools** > **root**.
3. Right-click the Server Pools.
4. Select Create Server Pool.
5. Enter nosh for the Server Pool Name.
6. (Optional) Enter a description for the organization.

Figure 18 Creating Server Pool



7. Click **Next** to add servers.
8. Select all the Cisco UCS C220M3 servers to be added to the nosh server pool. Click **>>** to add them to the pool.

Figure 19 Adding Server Pool



9. Click **Finish**.
10. Click **OK** and then click **Finish**.

Creating Policies for Service Profile Template

Creating Host Firmware Package Policy

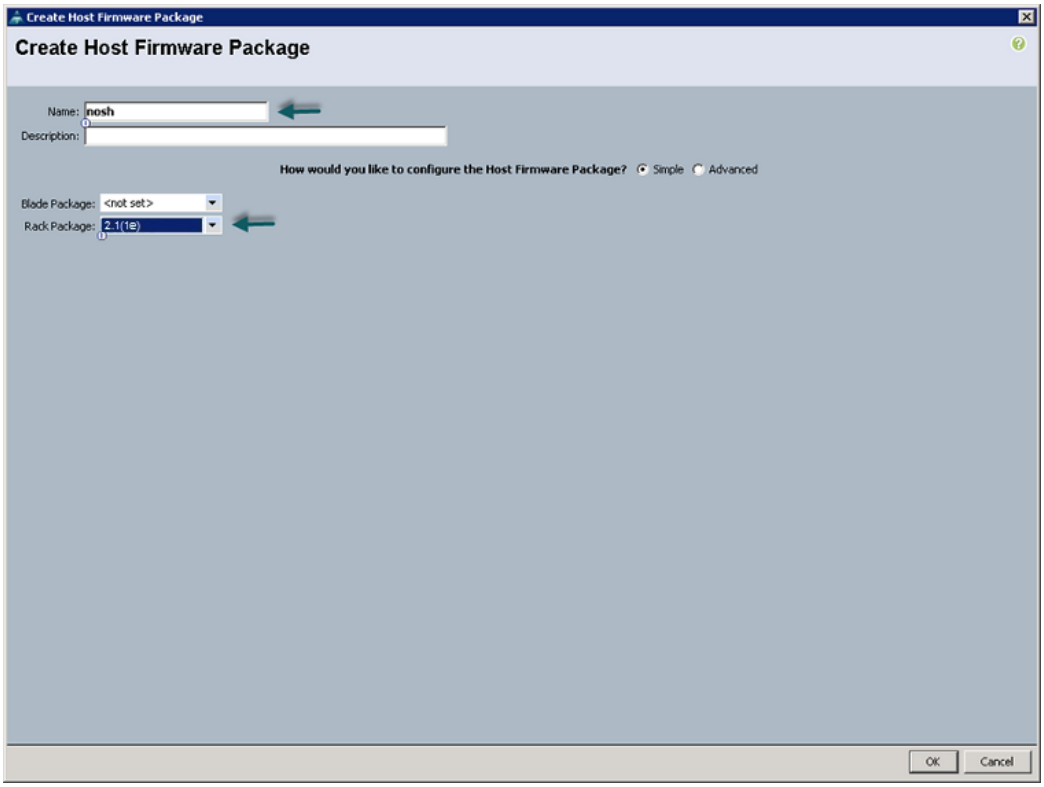
Firmware management policies allow the administrator to select the corresponding packages for a given server configuration. These often include adapter, BIOS, board controller, FC adapters, HBA option ROM, and storage controller properties.

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

1. Select the Servers tab in the left pane in the UCSM GUI.
2. Select **Policies > root**.
3. Right-click Host Firmware Packages.
4. Select Create Host Firmware Package.
5. Enter nosh as the Host firmware package name.
6. Select **Simple** radio button to configure the Host Firmware package.

7. Select the appropriate Rack package that you have.

Figure 20 Creating Host Firmware Package



8. Click **OK** to complete creating the management firmware package.
9. Click **OK**.

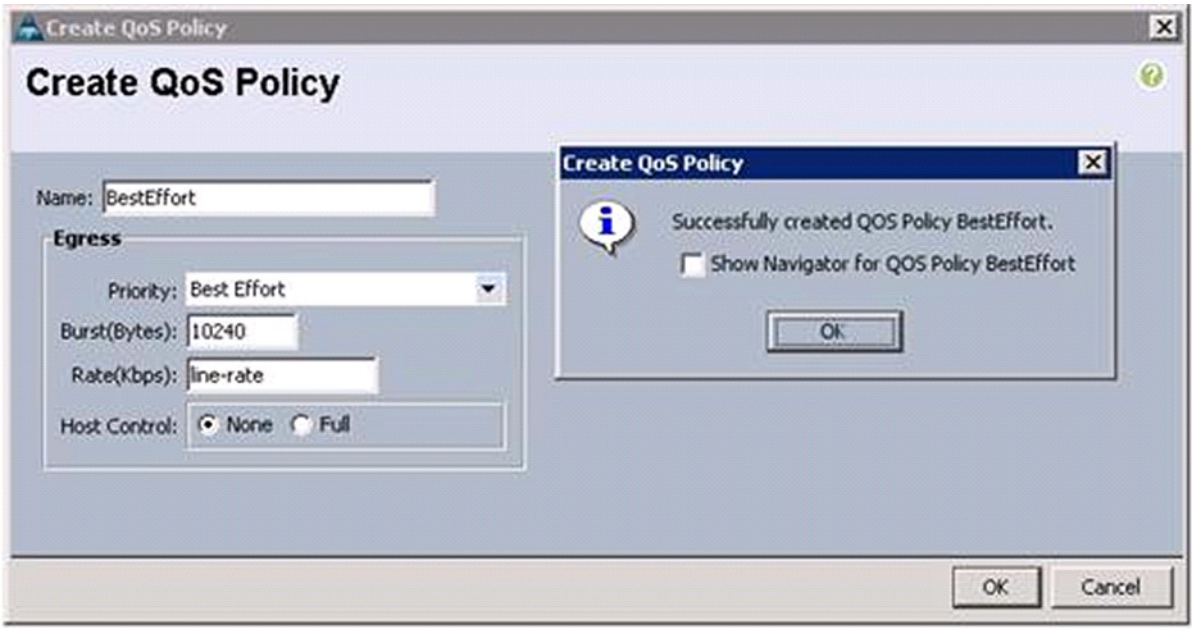
Creating QoS Policies

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

BestEffort Policy

1. Select the LAN tab in the left pane in the UCSM GUI.
2. Select **Policies** > **root**.
3. Right-click QoS Policies and select Create QoS Policy.
4. Enter BestEffort as the name of the policy.
5. Select Best Effort for Priority from the drop down menu.
6. Keep the Burst (Bytes) field as default, which is 10240.
7. Keep the Rate (Kbps) field as default, which is line-rate.
8. Make sure the Host Control radio button is **None**.
9. Click **OK**.

Figure 21 Creating QoS Policy - BestEffort

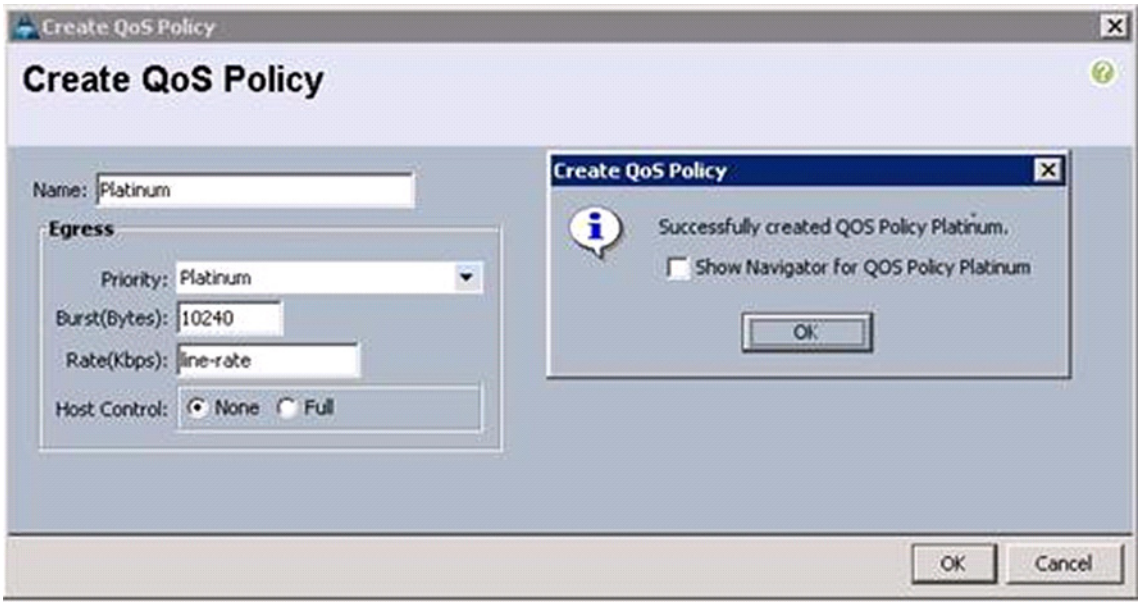


10. In the pop-up window, click **OK** to complete the QoS policy creation.

Platinum Policy

1. Select the LAN tab in the left pane in the UCSM GUI.
2. Select **Policies** > **root**.
3. Right-click QoS Policies and select Create QoS Policy.
4. Enter Platinum as the name of the policy.
5. Select Platinum for Priority from the drop down menu.
6. Keep the Burst (Bytes) field as default, which is 10240.
7. Keep the Rate (Kbps) field as default, which is line-rate.
8. Make sure the Host Control radio button is **None**.
9. Click **OK**.

Figure 22 Creating QoS Policy - Platinum



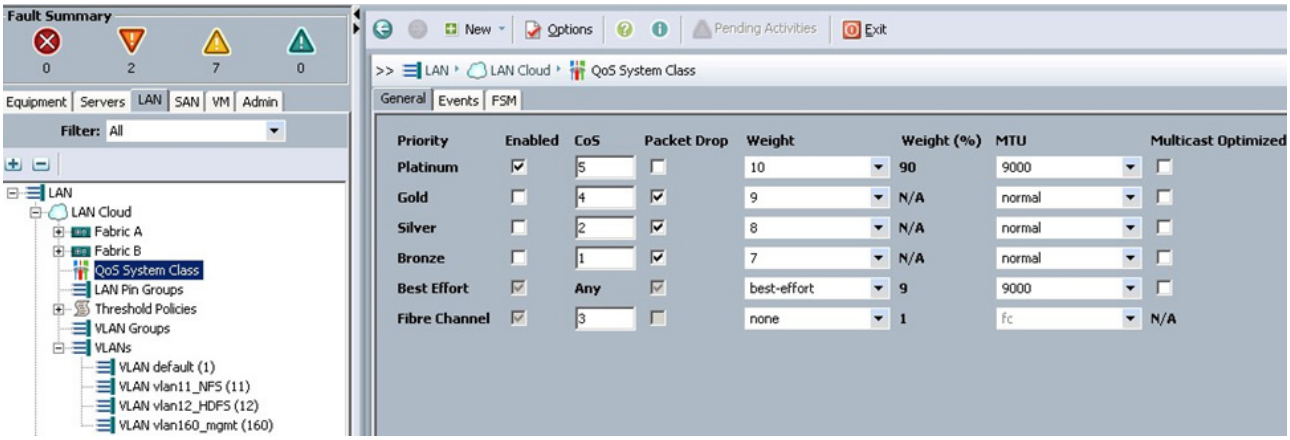
10. In the pop-up window, click **OK** to complete the QoS policy creation.

Setting Jumbo Frames

These steps provide details for setting Jumbo frames and enabling the quality of service in the Cisco UCS Fabric:

1. Select the Servers 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. In the Best Effort row, enter 9000 for MTU.
6. Check the Enabled check box next to Platinum.

Figure 23 Setting Jumbo Frame in Cisco UCS Fabric



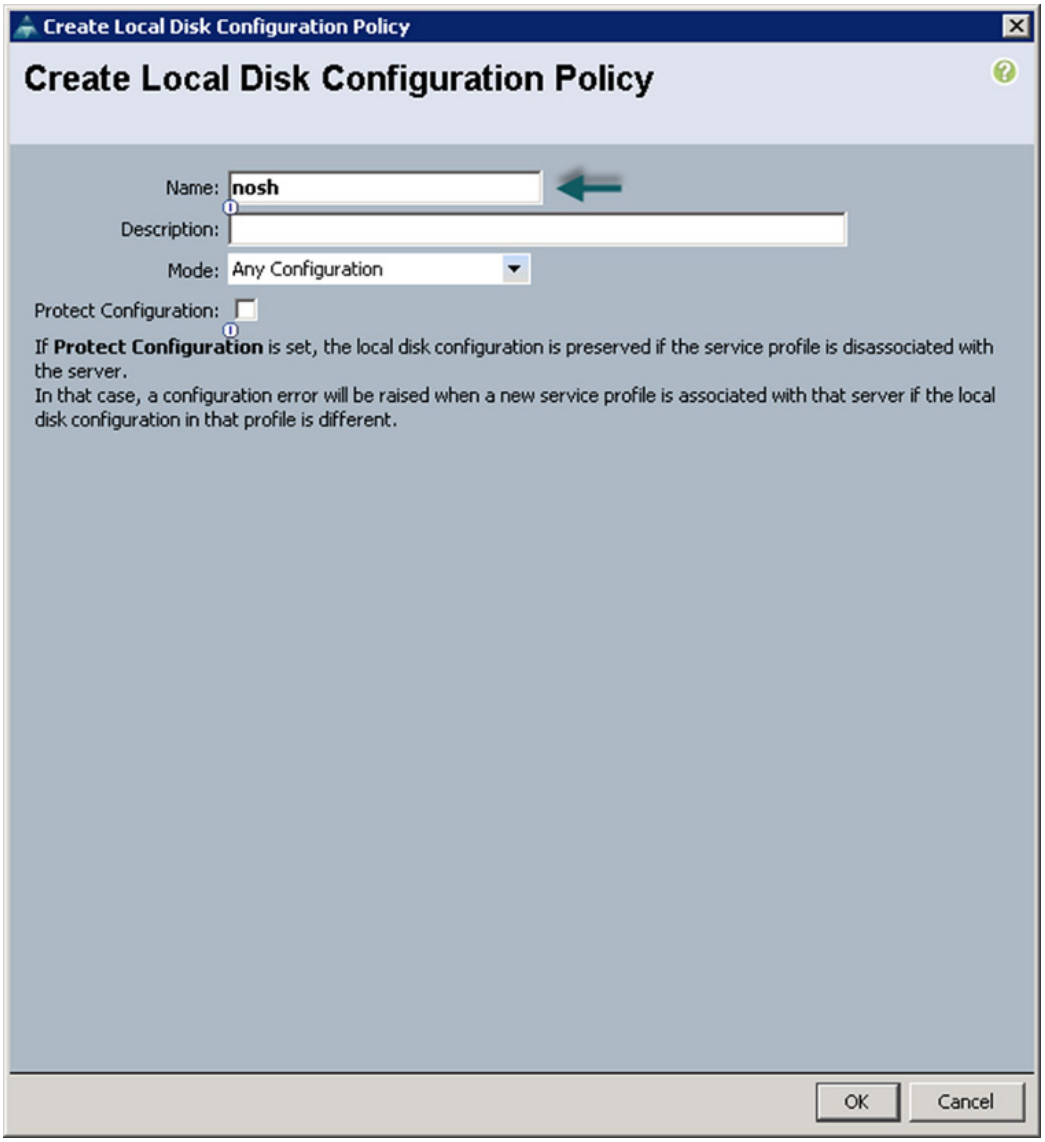
7. Click **Save Changes**.
8. Click **OK**.

Create a Local Disk Configuration Policy

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

1. Select the Servers tab in the left pane in the UCSM GUI.
2. Select **Policies > root**.
3. Right-click Local Disk Config Policies.
4. Select Create Local Disk Configuration Policy.
5. Enter nosh as the local disk configuration policy name.
6. Change the Mode to Any Configuration. Uncheck the Protect Configuration check box.

Figure 24 Configuring Local Disk Policy



7. Click **OK** to create the Local Disk Configuration Policy.
8. Click **OK**.

Create a Server BIOS Policy

The BIOS policy feature in Cisco UCS automates the BIOS configuration process.

The traditional method of setting the BIOS is manual and often error-prone. By creating a BIOS policy and assigning the policy to a server or group of servers, you can have the transparency in BIOS settings and configuration.

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

1. Select the Servers tab in the left pane in the UCSM GUI.

2. Select **Policies > root**.
3. Right-click BIOS Policies.
4. Select Create BIOS Policy.
5. Enter nosh as the BIOS policy name.
6. Change the BIOS settings as per Figure 25, Figure 26, Figure 27, and Figure 28.

Figure 25 Creating BIOS Policy

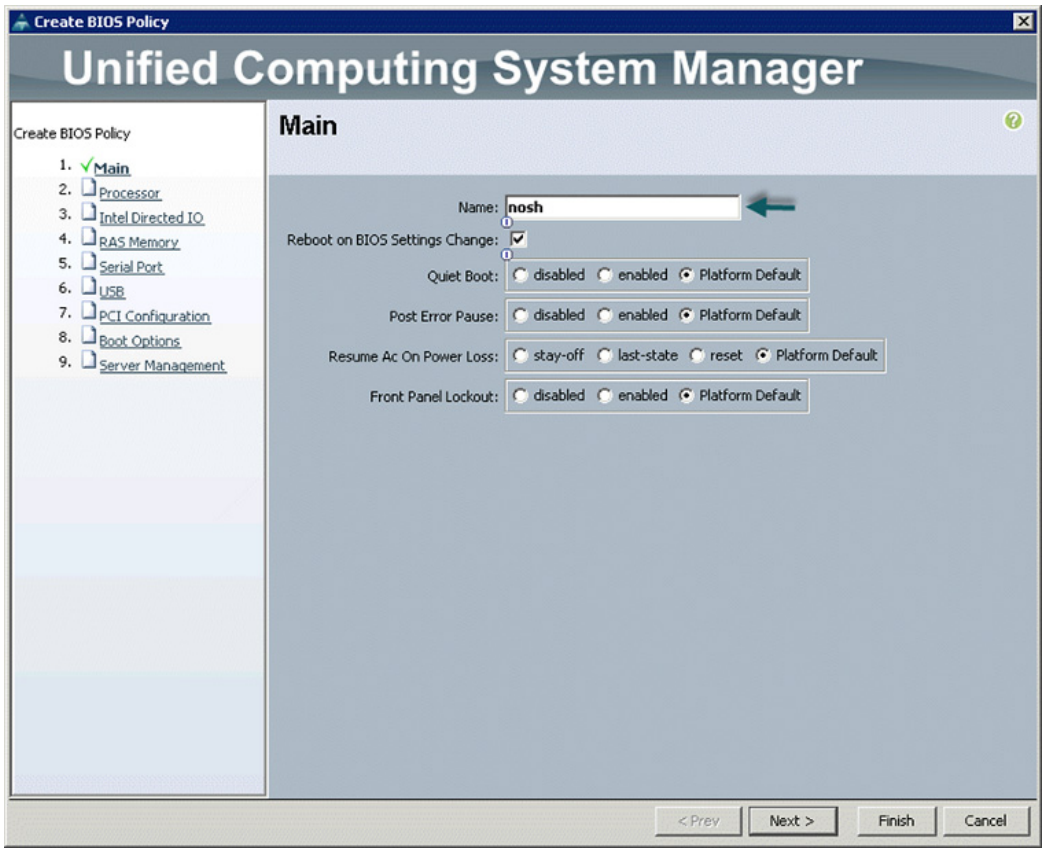


Figure 26 Processor Settings

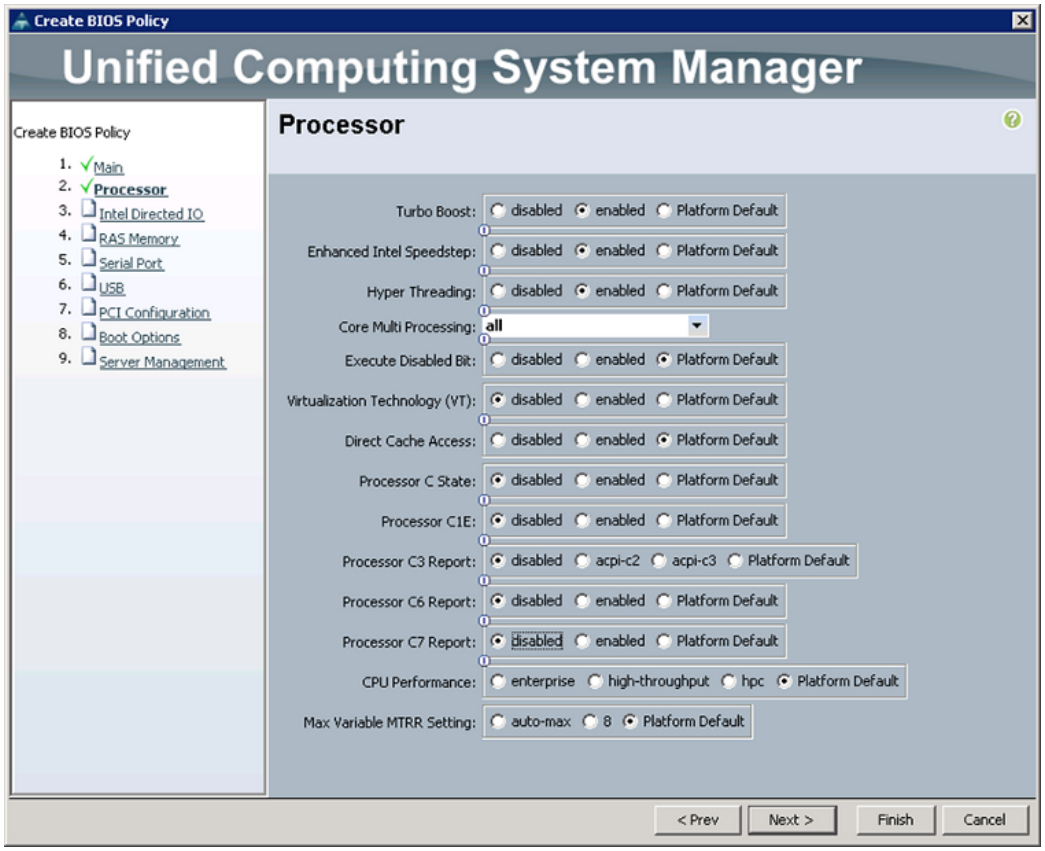


Figure 27 Intel Direct IO Settings

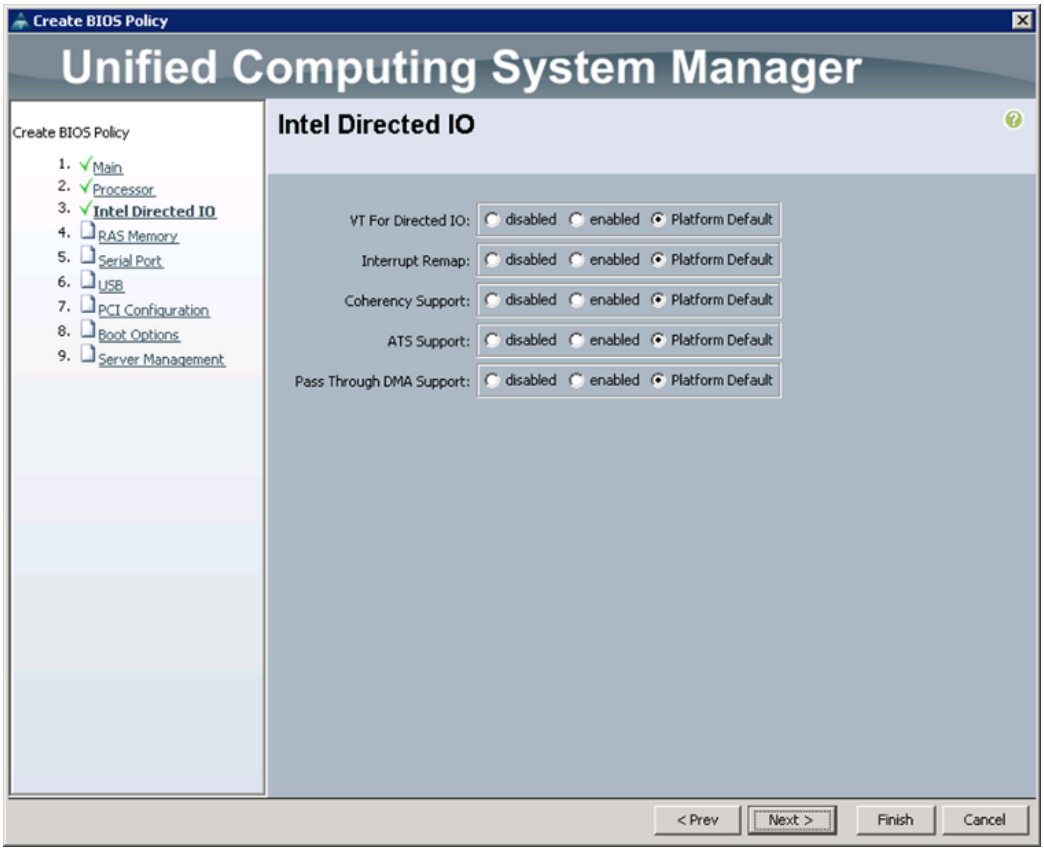
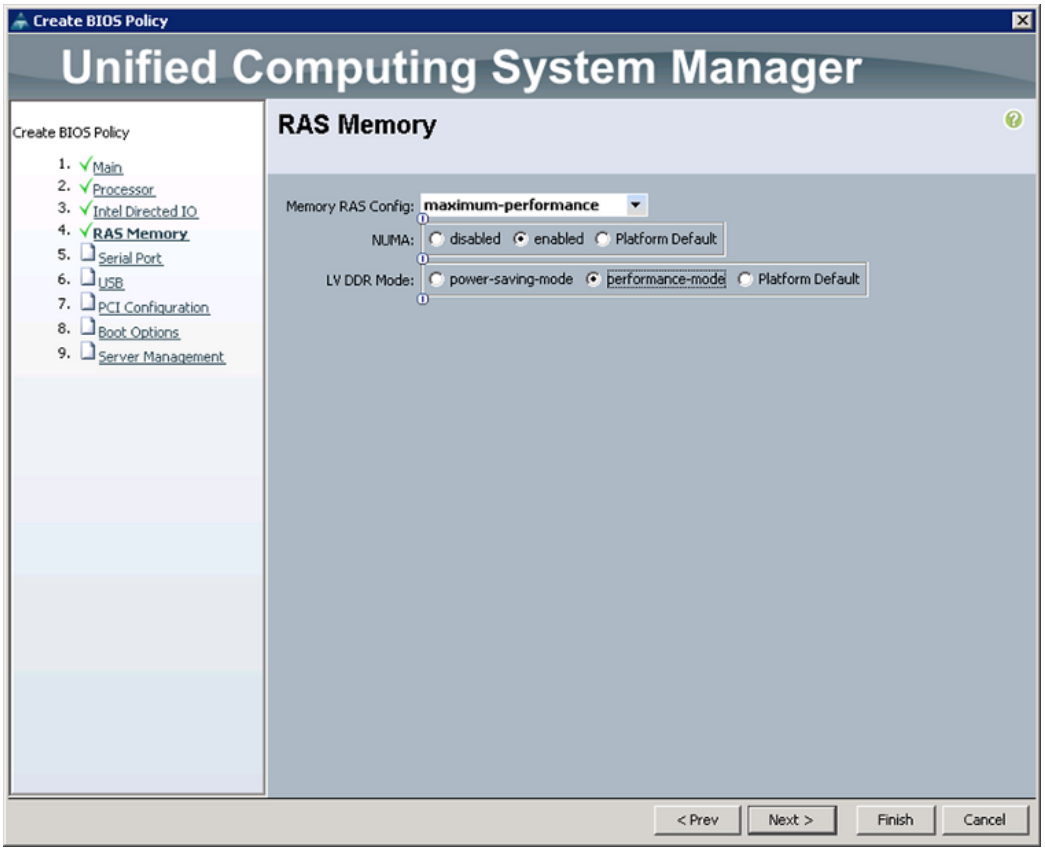


Figure 28 Memory Settings



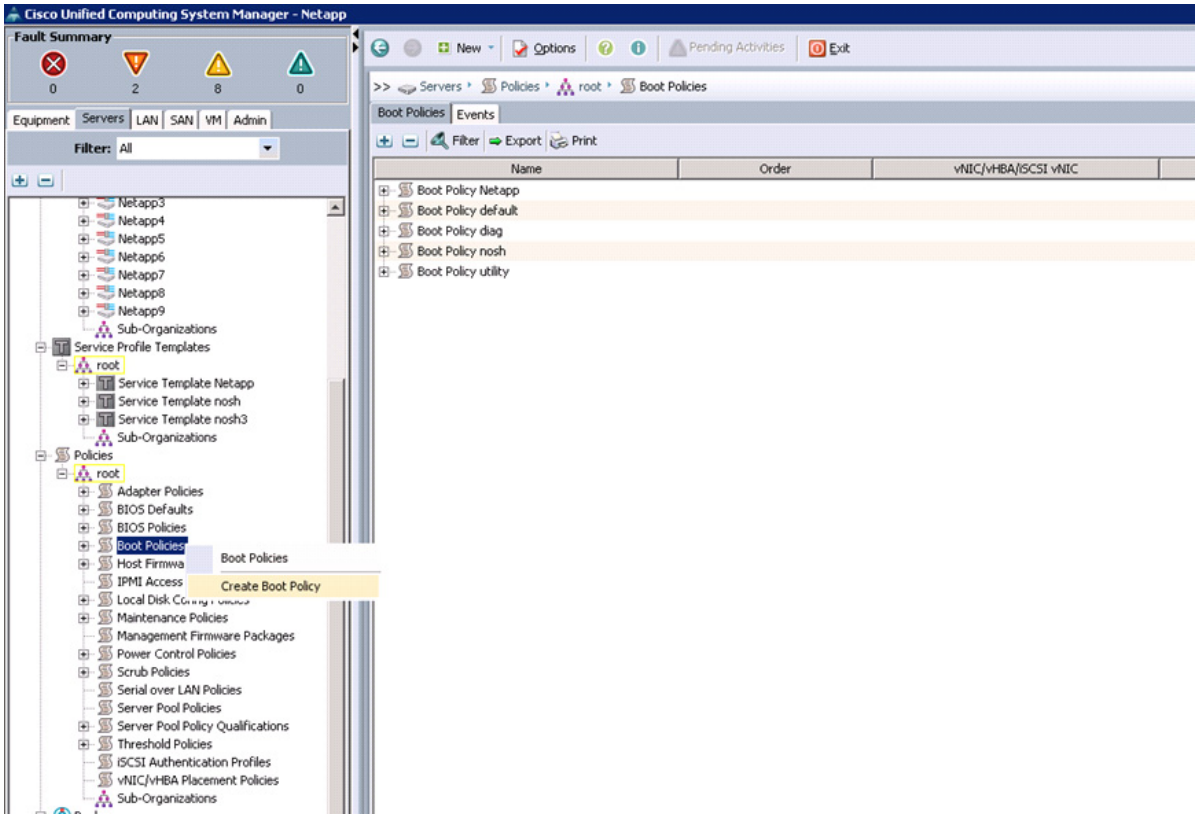
7. Click **Finish** to complete creating the BIOS policy.
8. Click **OK**.

Creating Boot Policies

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

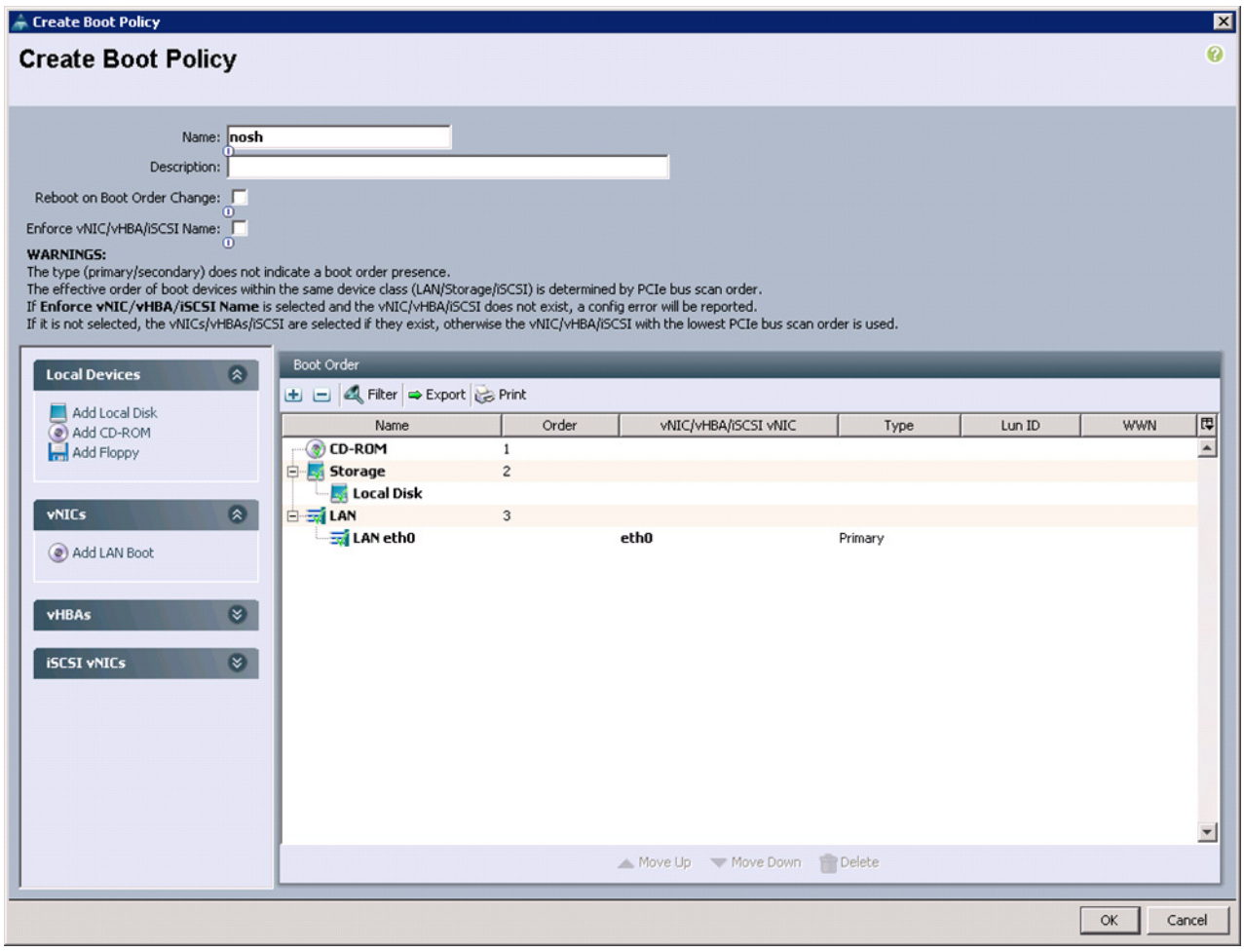
1. Select the Servers tab in the left pane in the UCSM GUI.
2. Select **Policies** > **root**.
3. Right-click the Boot Policies.
4. Select Create Boot Policy.

Figure 29 Creating Boot Policy



5. Enter nosh 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. Expand Local Devices and select Add CD-ROM.
9. Expand Local Devices and select Add Local Disk.
10. Expand vNICs and select Add LAN Boot and enter eth0.

Figure 30 Creating Boot Order



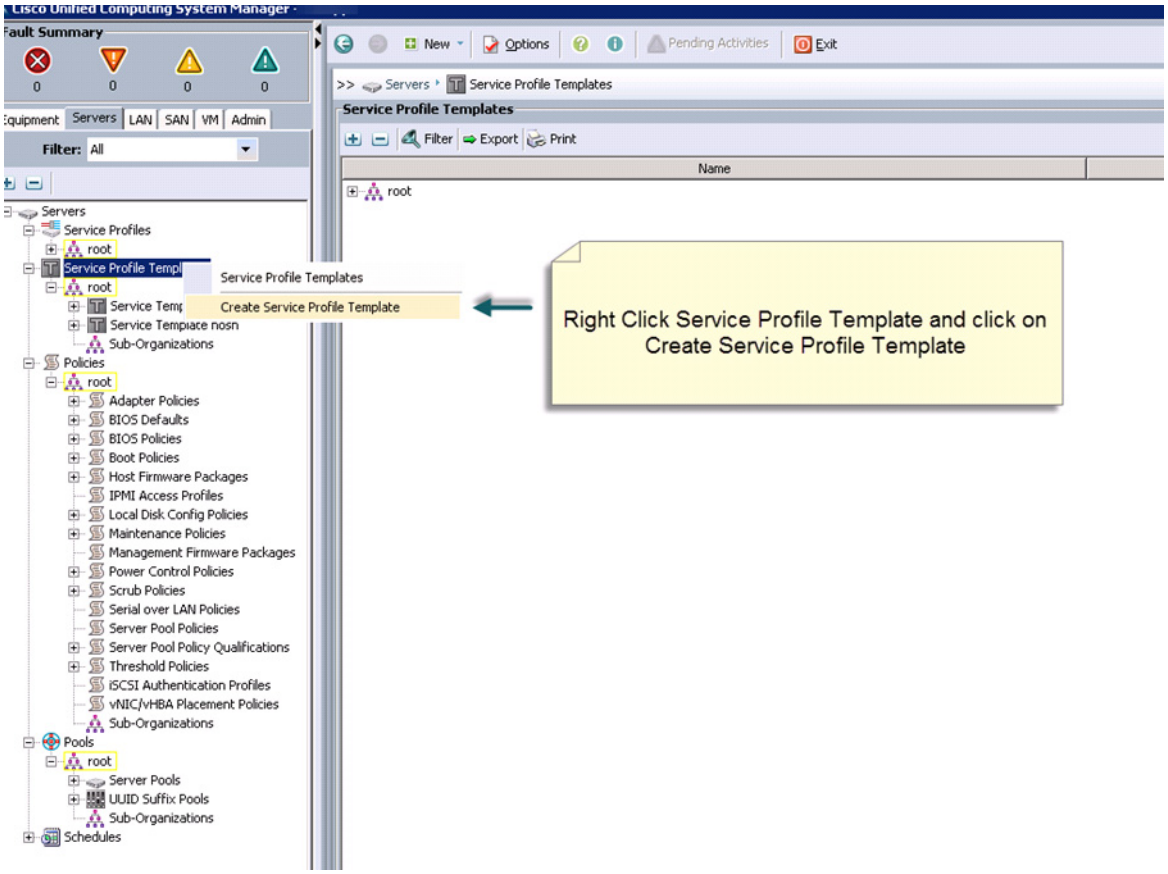
11. Click **OK** to add the Boot Policy.
12. Click **OK**.

Creating Service Profile Template

To create a service profile template, follow these steps:

1. Select the Servers tab in the left pane in the UCSM GUI.
2. Select **Policies** > **root**.
3. Right-click root.
4. Select Create Service Profile Template.

Figure 31 Creating Service Profile Template



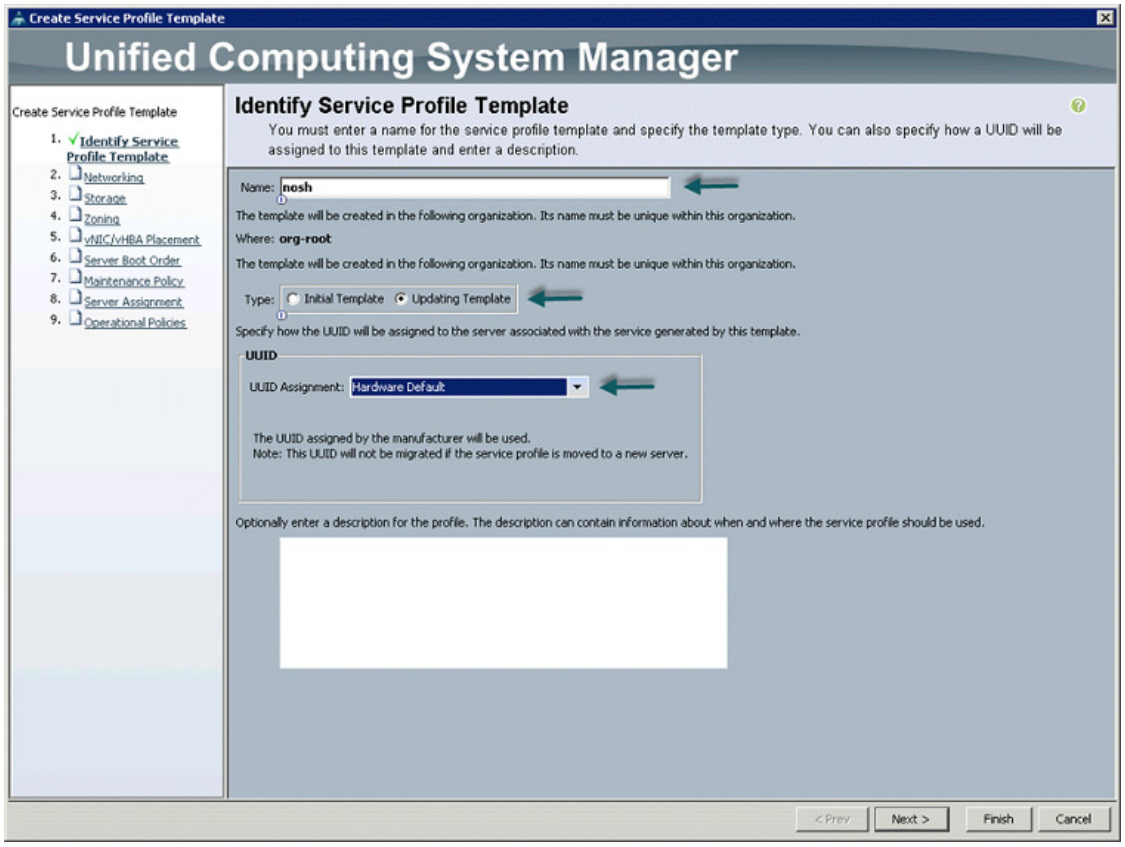
5. The Create Service Profile Template window appears.

   The following steps provide the detailed configuration procedure used to create a service profile template:

   a. Name the service profile template as nosh. Select the **Updating Template** radio button.

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

Figure 32 Identifying Service Profile Template



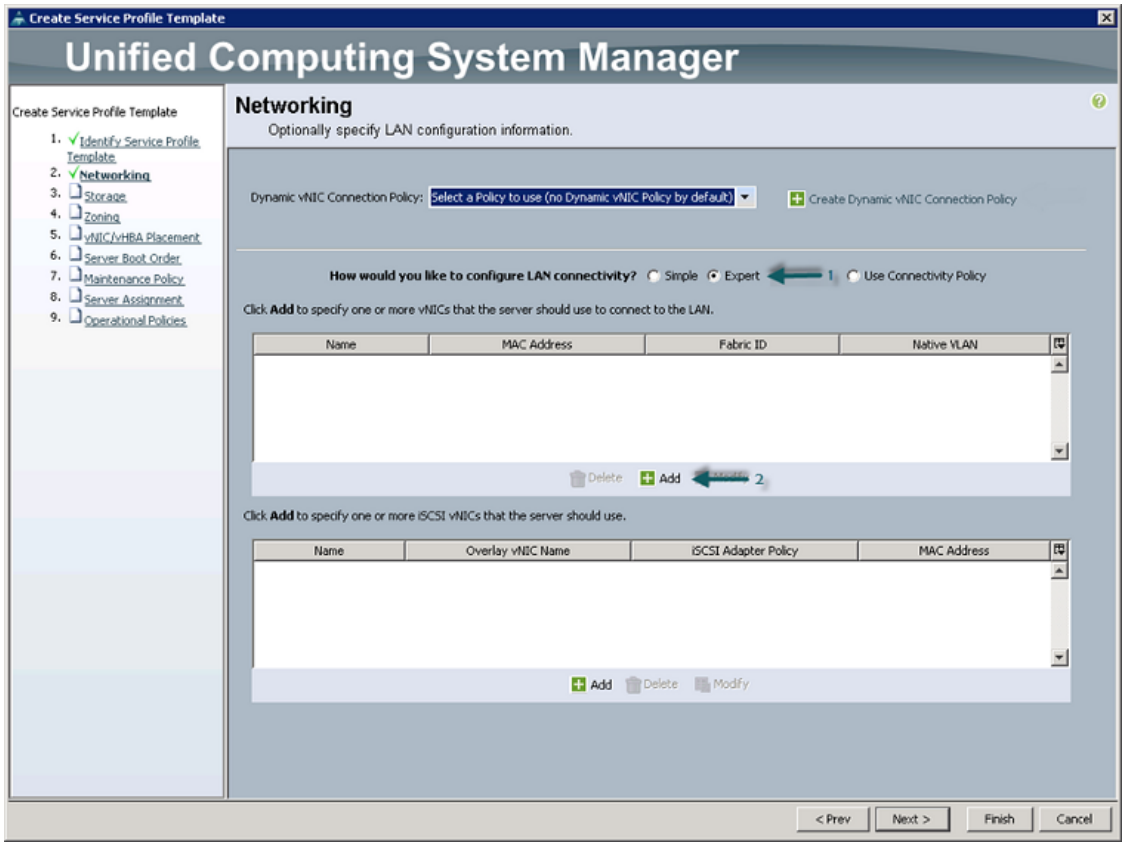
c. Click **Next** to continue to the next section.

Configuring Network Settings for the Template

In the Networking window, follow these steps to create vNICs:

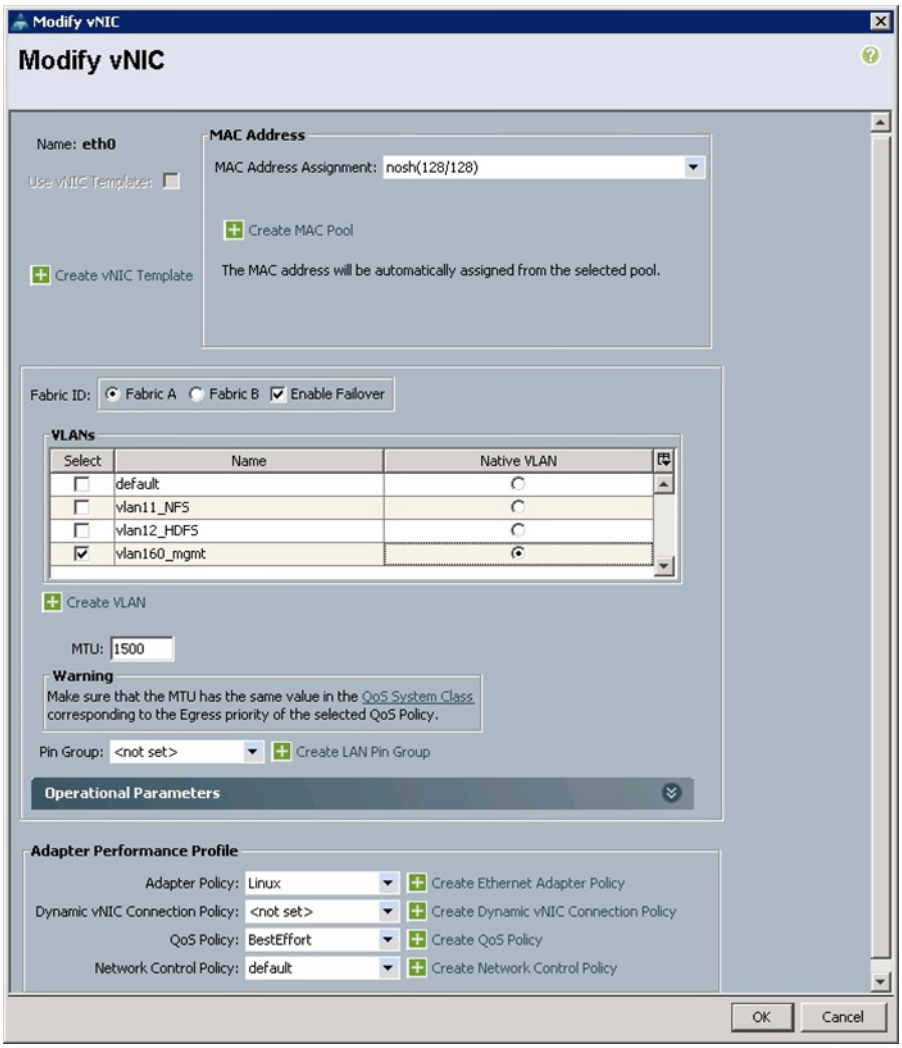
1. Keep the Dynamic vNIC Connection Policy field as default.
2. Select the **Expert** radio button for the option How would you like to configure LAN connectivity?
3. Click **Add** to add a vNIC to the template.

Figure 33 Adding vNICs



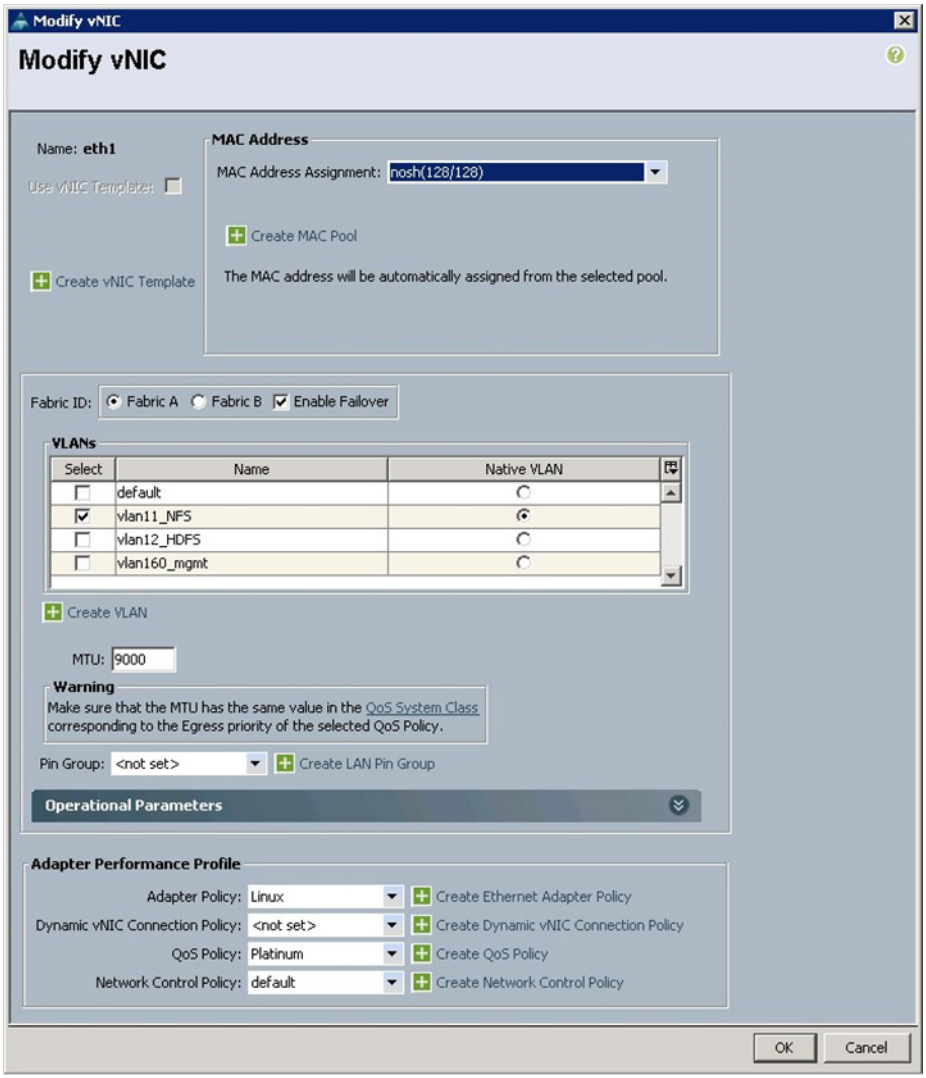
4. The Create vNIC window displays. Name the vNIC as eth0.
5. Select nosh in the Mac Address Assignment pool.
6. Select the **Fabric A** radio button and check the Enable failover check box for the Fabric ID.
7. Check the vlan160_mgmt check box for VLANs and select the **Native VLAN** radio button.
8. Select MTU size as 1500.
9. Select adapter policy as Linux.
10. Keep the Dynamic vNIC connection policy as <no set>.
11. Select QoS Policy as BestEffort.
12. Keep the Network Control Policy as Default.

Figure 34 Creating Management vNIC



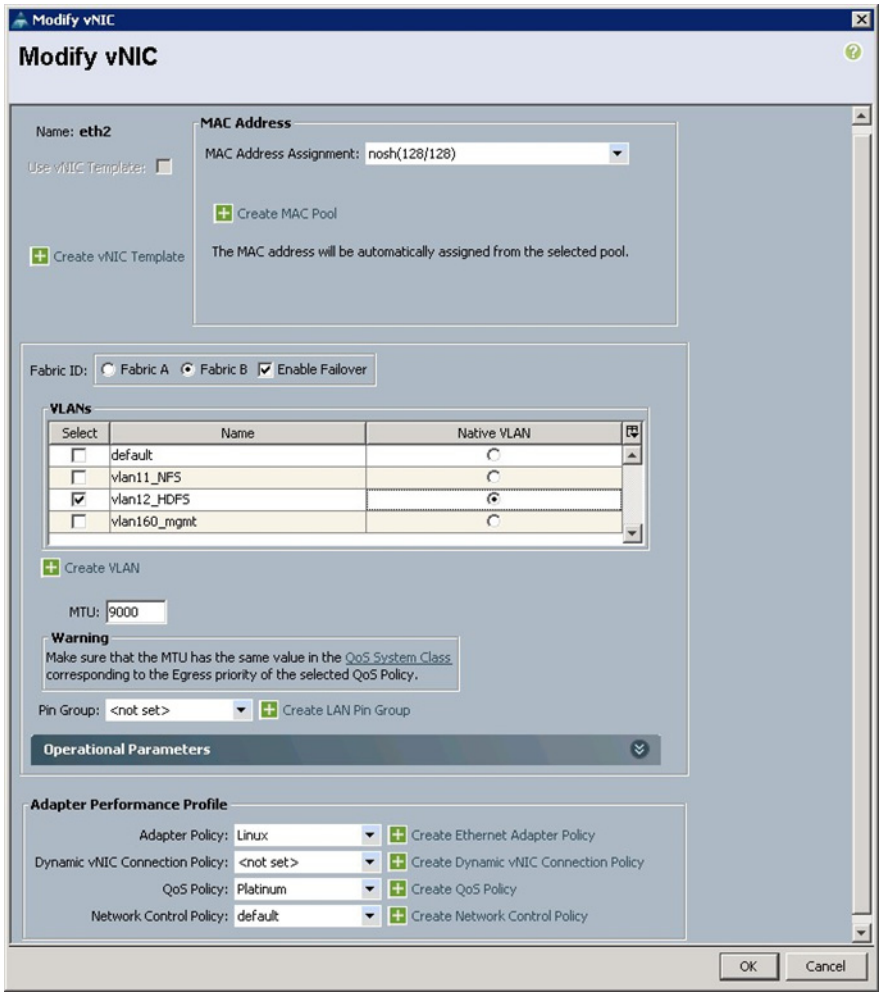
13. Click **OK**.
14. Click **Add** to add another vNIC to the template.
15. The Create vNIC window appears. Name the vNIC as eth1.
16. Select nosh in the Mac Address Assignment pool.
17. Select the **Fabric A** radio button and check the Enable failover check box for the Fabric ID.
18. Check the Default and vlan11 check boxes for VLANs and select the **vlan11** radio button for Native VLAN.
19. Select MTU size as 9000.
20. Select Adapter Policy as Linux.
21. Keep the Dynamic vNIC Connection Policy as <not set>.
22. Select QoS Policy to Platinum.
23. Keep the Network Control Policy as Default.

Figure 35 Creating NFS vNIC



24. Click **OK**.
25. Click **Add** to add another vNIC to the template.
26. The Create vNIC window appears. Name the vNIC as eth2.
27. Select nosh in the Mac Address Assignment pool.
28. Select the **Fabric B** radio button and check the Enable failover check box for the Fabric ID.
29. Check the vlan12_HDFS check box for VLANs and select the **Native VLAN** radio button.
30. Select MTU size as 9000.
31. Select adapter policy as Linux.
32. Keep the Dynamic vNIC Connection Policy as <no set>.
33. Select QoS Policy as Platinum.
34. Keep the Network Control Policy as Default.

Figure 36 Creating HDFS vNIC



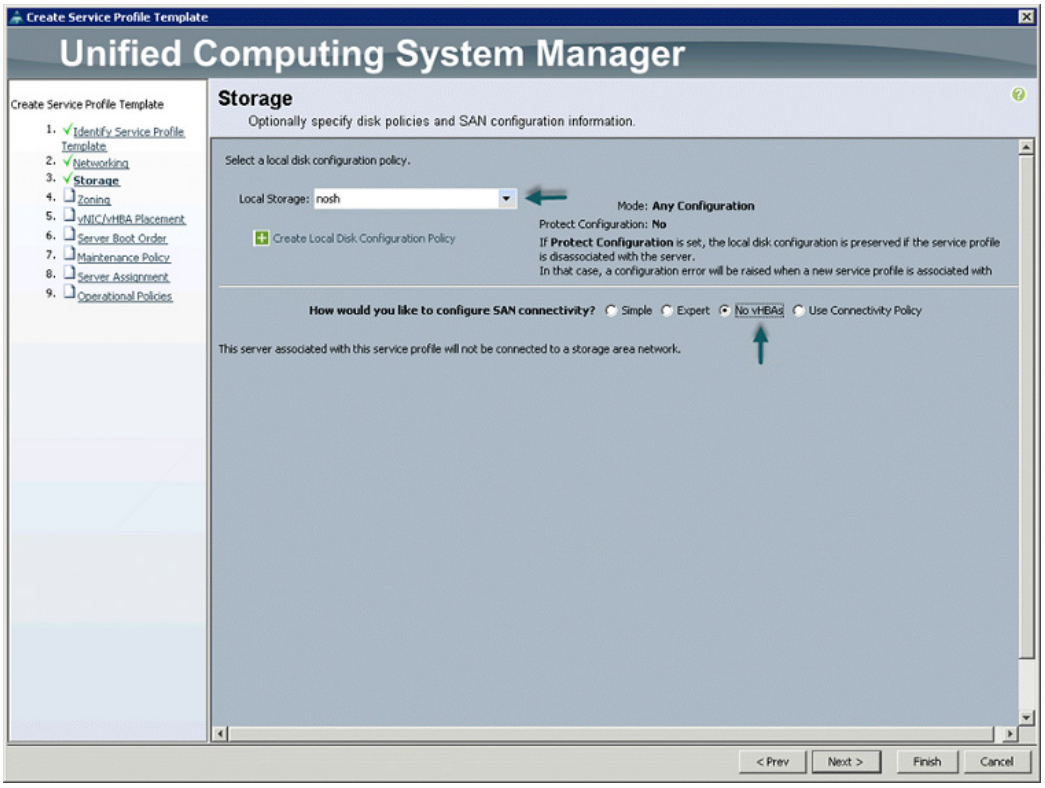
35. Click **OK**.
36. Click **Next** to continue to the next section.

Configuring Storage Policy for the Template

In the Storage window, follow these steps to configure storage:

1. Select nosh for the local disk configuration policy.
2. Select the **No vHBAs** radio button for the option How would you like to configure SAN connectivity?

Figure 37 Storage Settings



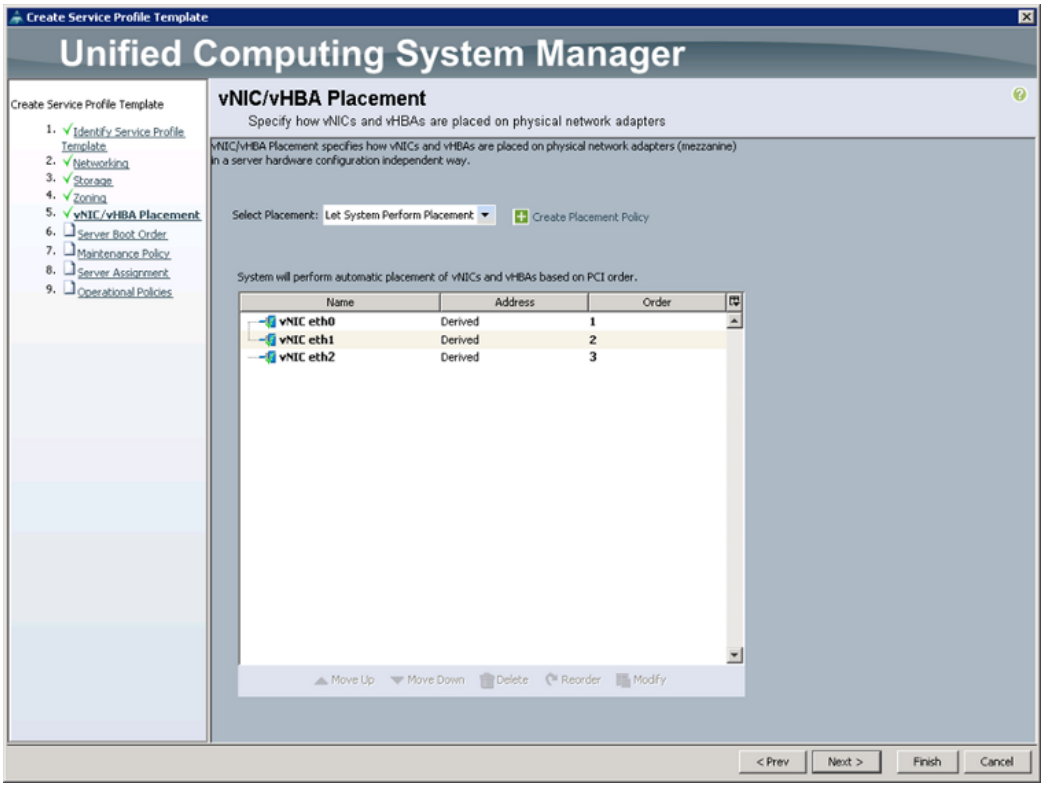
3. Click **Next** to continue to the next section.
4. Click **Next** in the Zoning Window to go to the next section.

Configuring vNIC/vHBA Placement for the Template

In the vNIC/vHBA Placement Section, follow these steps to configure placement policy:

1. Select the Default Placement Policy option for Select Placement field.
2. Select eth0, eth1, and eth2 assign the vNICs in the following order:
   - eth0
   - eth1
   - eth2
3. Review the table to make sure that all of the vNICs were assigned in the appropriate order.

Figure 38 Creating vNIC and vHBA Policy



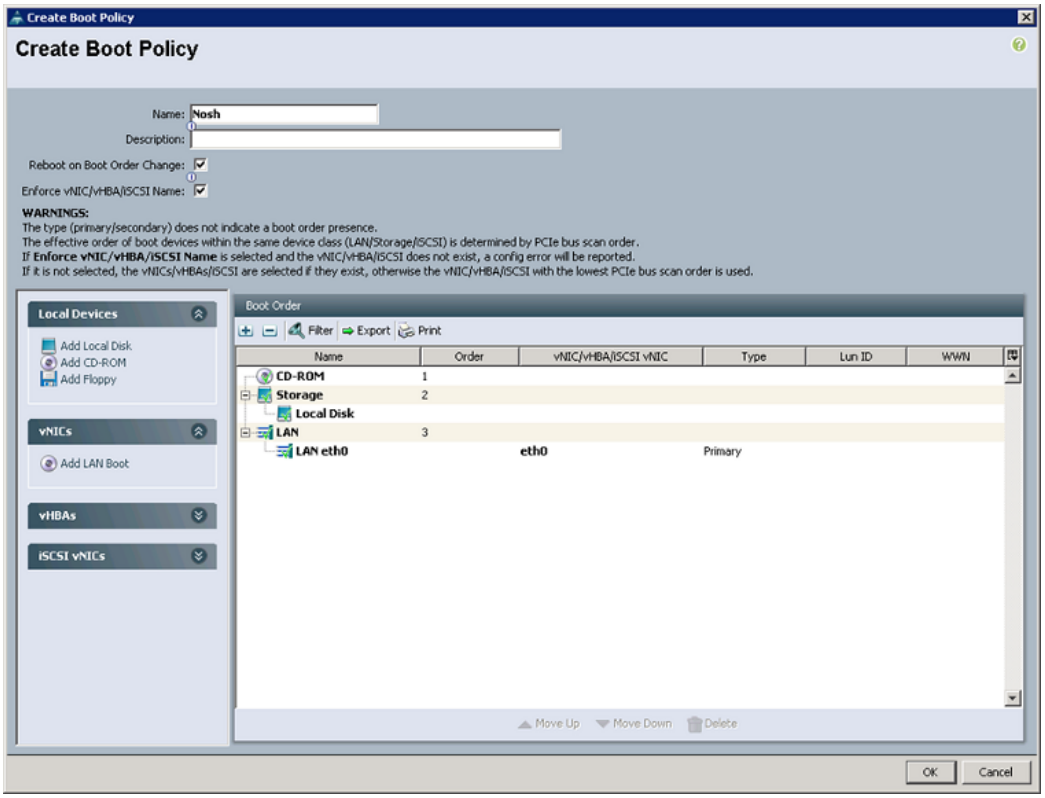
4. Click **Next** to continue to the next section.

Configuring Server Boot Order for the Template

In the Server Boot Order Section, follow these steps to set the boot order for servers:

1. Select nosh for the Boot Policy Name field.
2. Check the Reboot on Boot Order Change check box.
3. Check the Enforce vNIC/vHBA/iSCSI Name check box.
4. Review the table to make sure that all of the boot devices were created and identified. Verify that the boot devices are in the correct boot sequence.

Figure 39 Creating Boot Policy



5. Click **OK**.
6. Click **Next** to continue to the next section.

Configuring Maintenance Policy for the Template

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

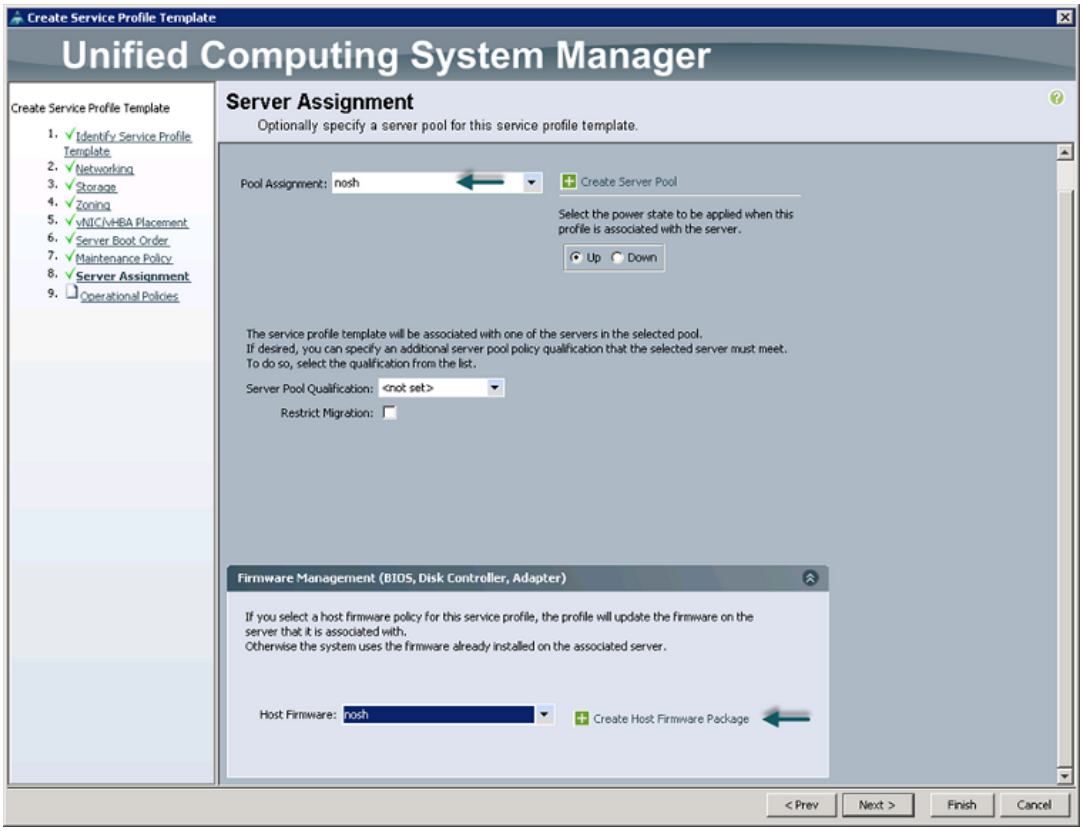
1. Keep the Maintenance Policy at no policy used by default.
2. Click **Next** to continue to the next section.

Configuring Server Assignment for the Template

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

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

Figure 40 Assigning Sever Pool for Service Profile Template



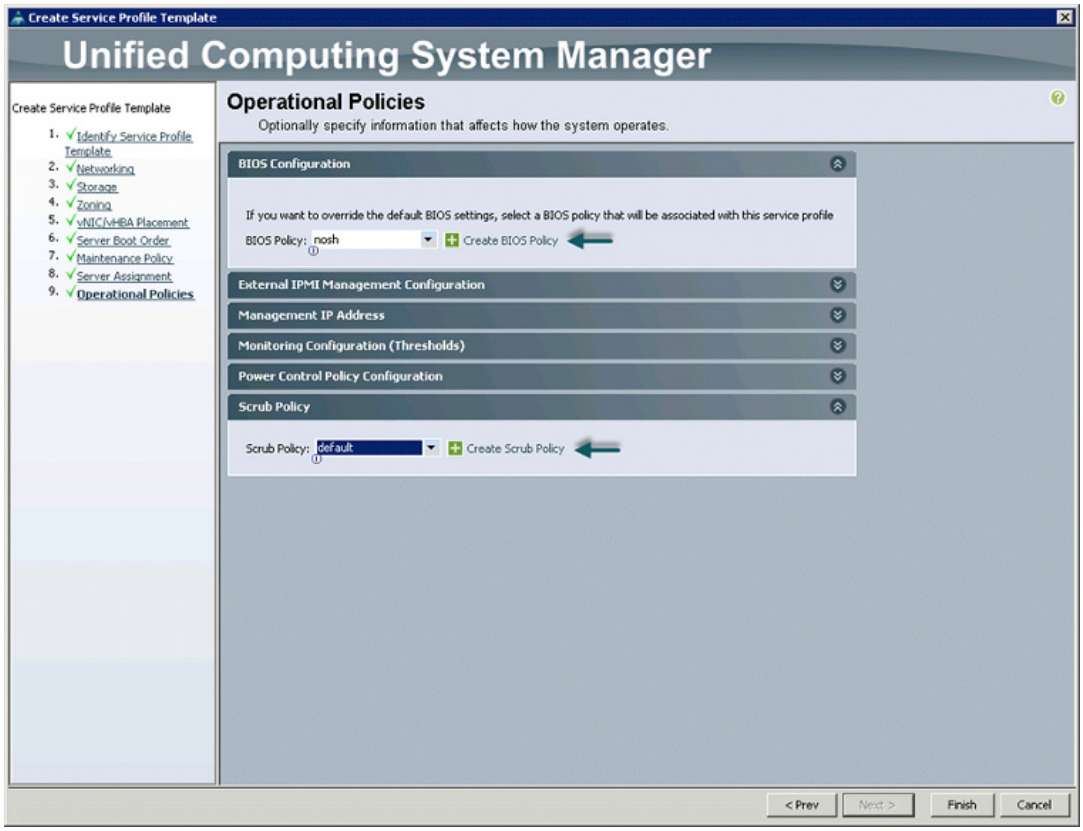
4. Click **Next** to continue to the next section.

Configuring Operational Policies for the Template

In the Operational Policies window, follow these steps:

1. Select nosh in the BIOS Policy field.

Figure 41 Creating Operational Policies

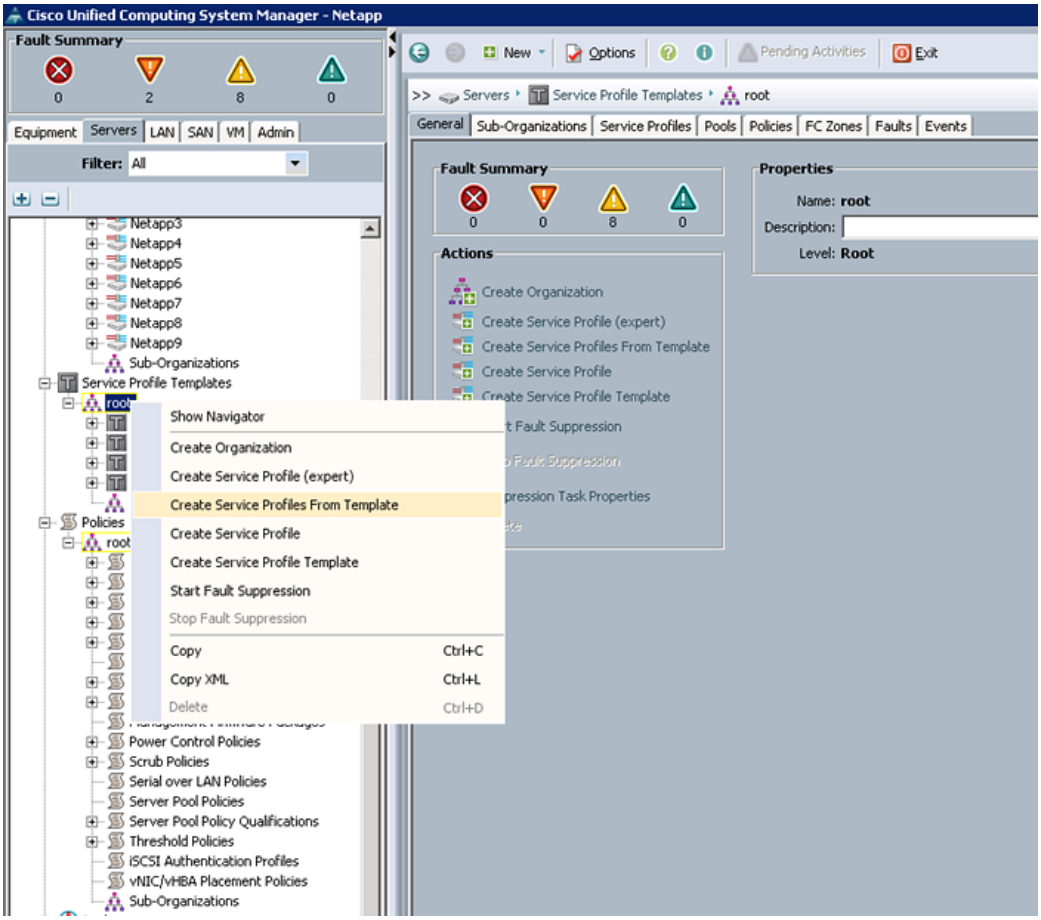


2. Click **Finish** to create the Service Profile template.
3. Click **OK** in the pop-up window to exit the wizard.

Select the Servers tab in the left pane in the UCSM GUI.

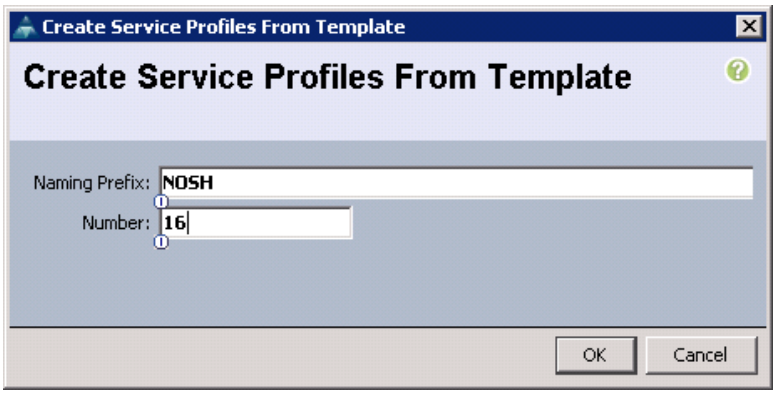
1. Select **Service Profile Templates > root.**
2. Right-click the root.
3. Select Create Service Profile Template.

Figure 42 Creating Service Profile



4. The Create Service Profile from Template window appears.

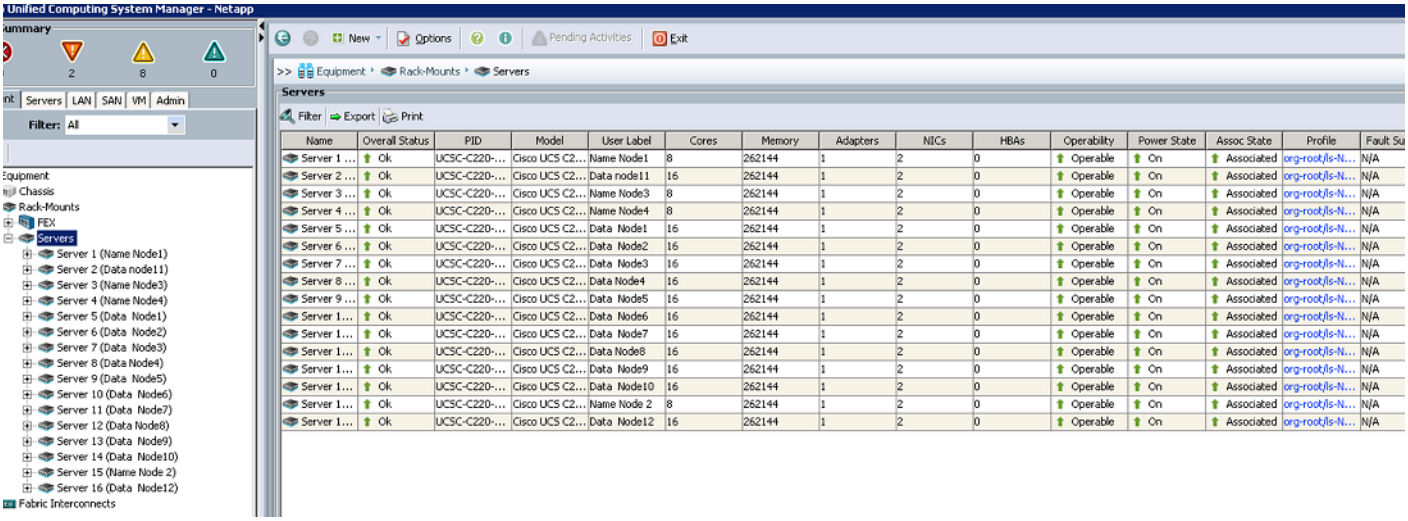
Figure 43 Creating Service Profile from Template



5. Now connect the power cable to the servers.
6. Servers will the be discovered by UCS Manager.
7. Association of Service Profile will take place automatically.

8. The final Cisco UCS Manager window is shown in Figure 44.

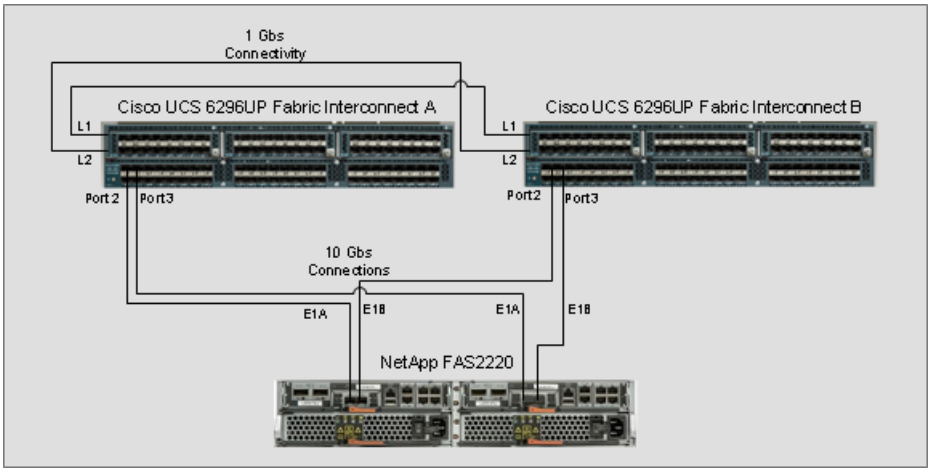
Figure 44 UCS Manager Showing Sixteen Nodes



Cisco UCS 6296UP FI Configuration for NetApp FAS 2220

The Cisco UCS 6296UP Fabric Interconnects are deployed in pairs with L1 to L1 and L2 to L2 connectivity for redundancy. NetApp FAS 2220 has one storage controllers. FAS controller port E1A is connected to FI A Port 2 as the Appliance port and E1B is connected to FI B Port 2 as the Appliance port with 10Gbs connectivity as shown in Figure 45.

Figure 45 Cisco UCS 6296UP FIs and NetApp FAS 2220 Connectivity



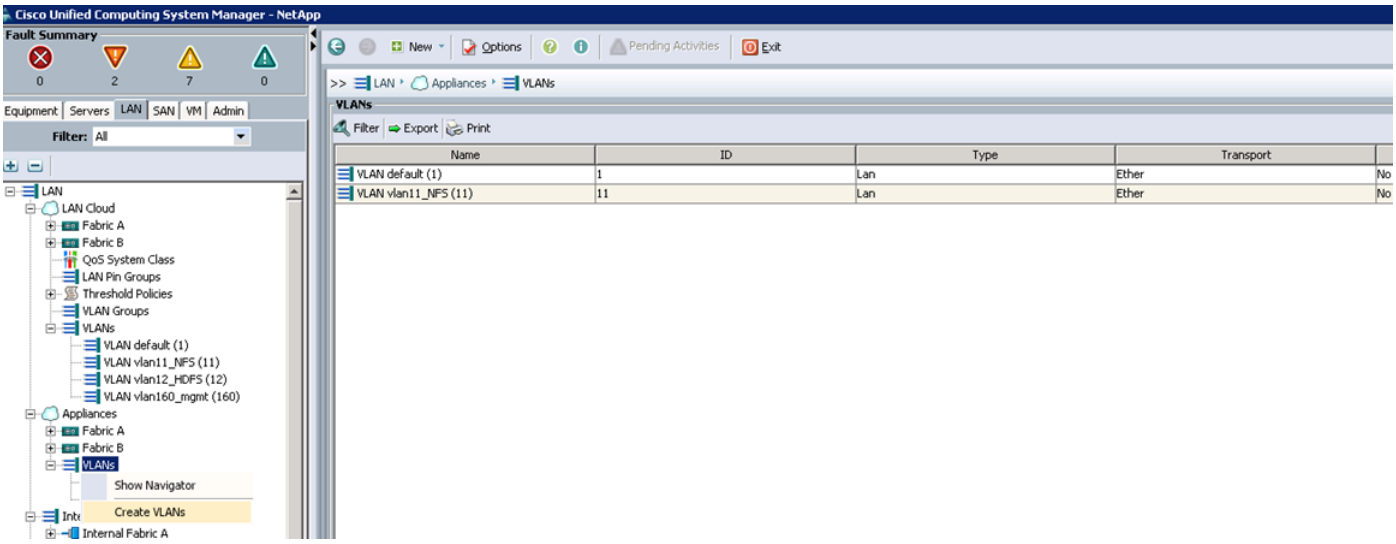
Configuring VLAN for Appliance Port

Follow these steps to configure VLAN appliance cloud:

1. Select the LAN tab in the left pane in the UCSM GUI.

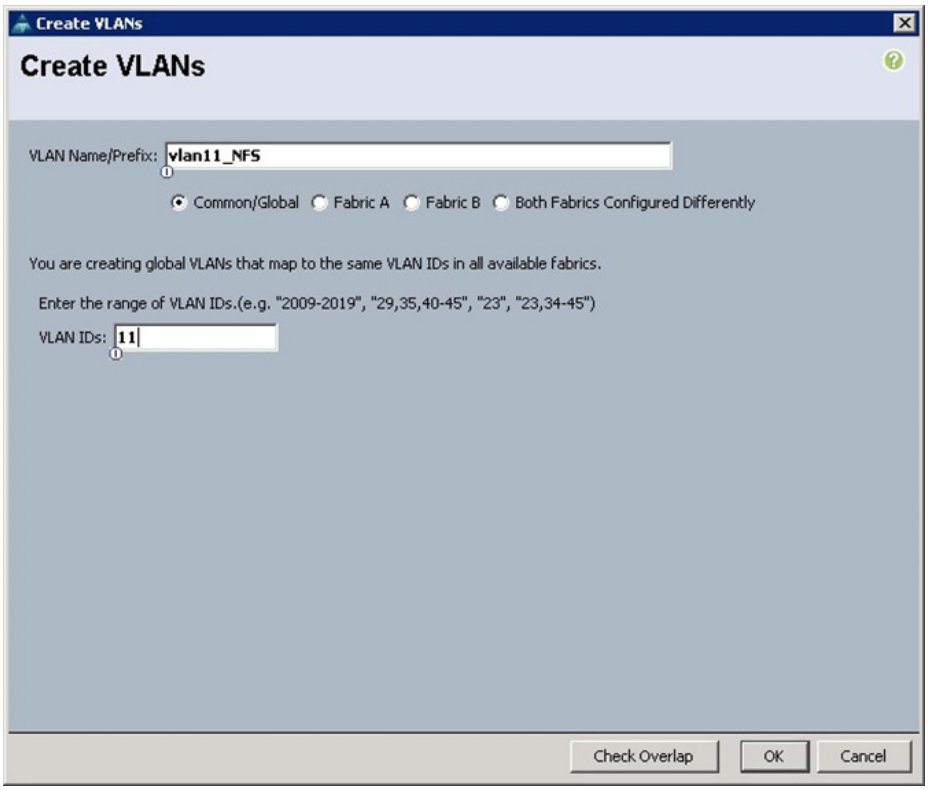
2. Select **LAN > Appliances > VLANs**.
3. Right-click VLANs under the root organization.
4. Select Create VLANs to create the VLAN.

Figure 46 Creating VLANs for Appliance Cloud



5. Enter vlan11_NFS for the VLAN Name.
6. Select the **Common/Global** radio button.
7. Enter 11 for VLAN ID.

Figure 47 Creating VLAN for Fabric A



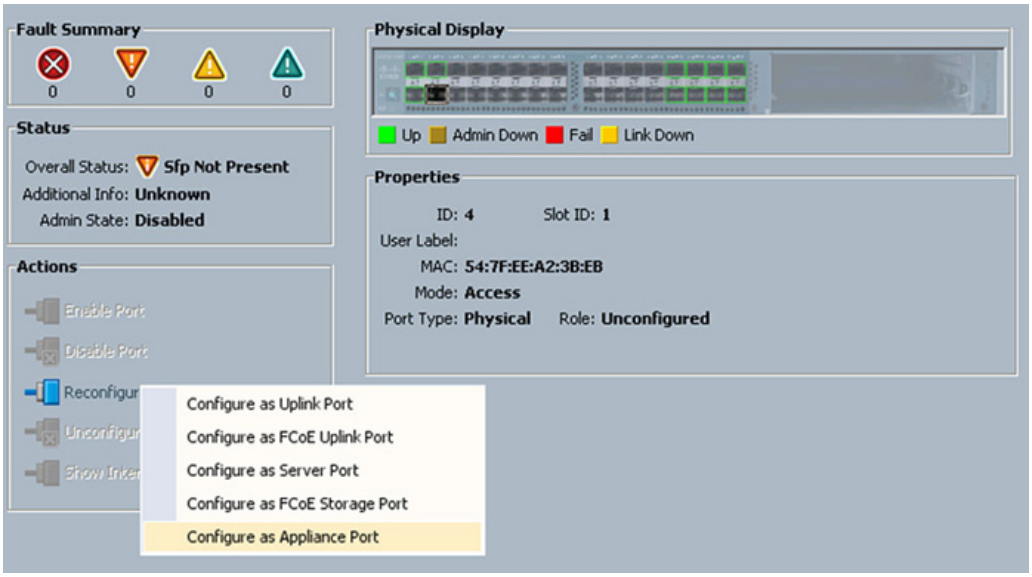
8. Click **OK** and then, click **Finish**.

Configure Appliance Port

Follow these steps to configure appliance ports:

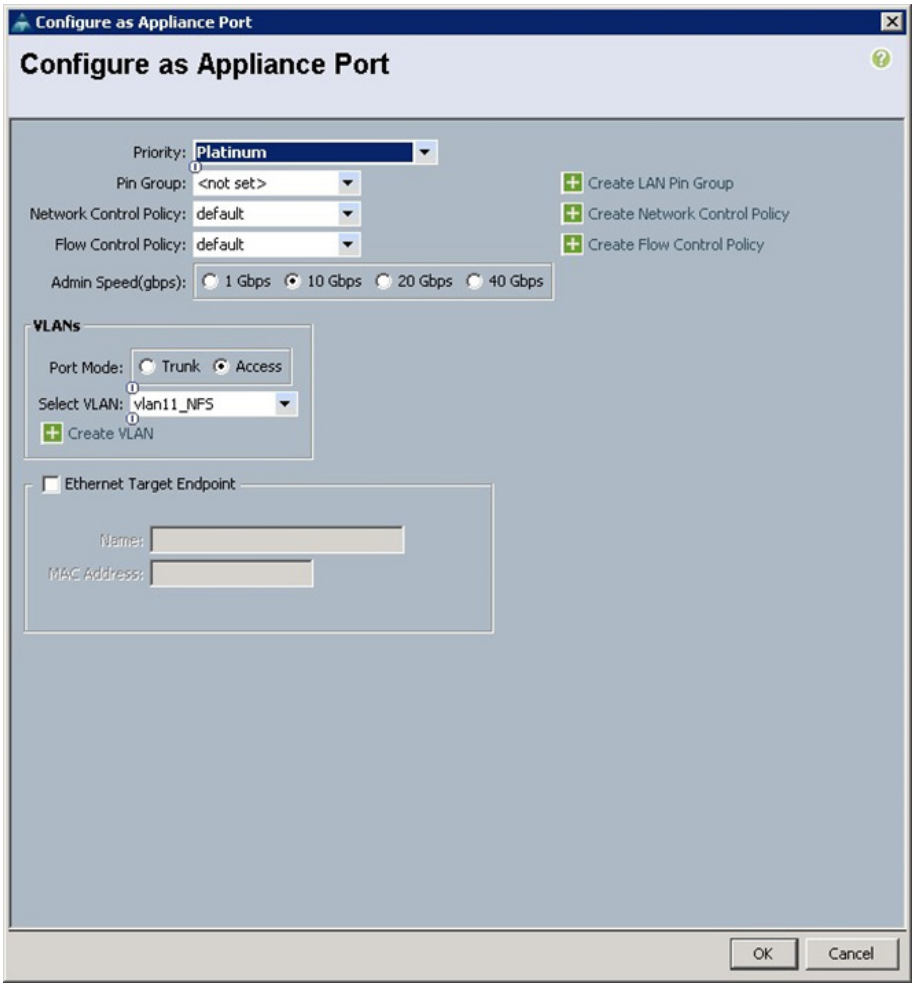
1. Select the Equipment tab in the left pane in the UCSM GUI.
2. Select **Equipment > Fabric Interconnects > Fabric Interconnect A (primary) > Fixed Module**.
3. Expand the Unconfigured Ethernet Ports.
4. Select the port number 2, and select **Reconfigure > Configure as an Appliance Port**.

Figure 48 Configuring Fabric A Appliance Port



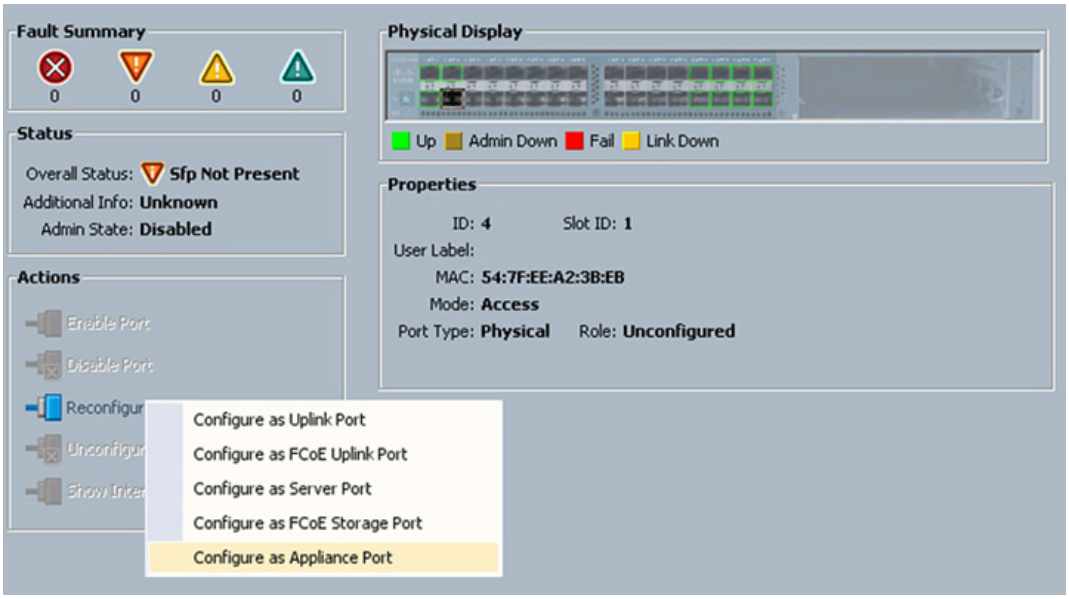
5. A confirmation message box appears. Click **Yes**, then **OK** to continue.
6. Select Platinum for the Priority.
7. Keep the Pin Group as <not set>.
8. Keep the Network Control Policy as Default.
9. Keep the Flow Control Policy as Default.
10. Select the **10Gbps** radio button for the Admin Speed.
11. Select the **Access** radio button for the Port Mode.
12. Select vlan11_NFS from the drop-down menu for the Select VLAN.

Figure 49 Configuring Appliance Port



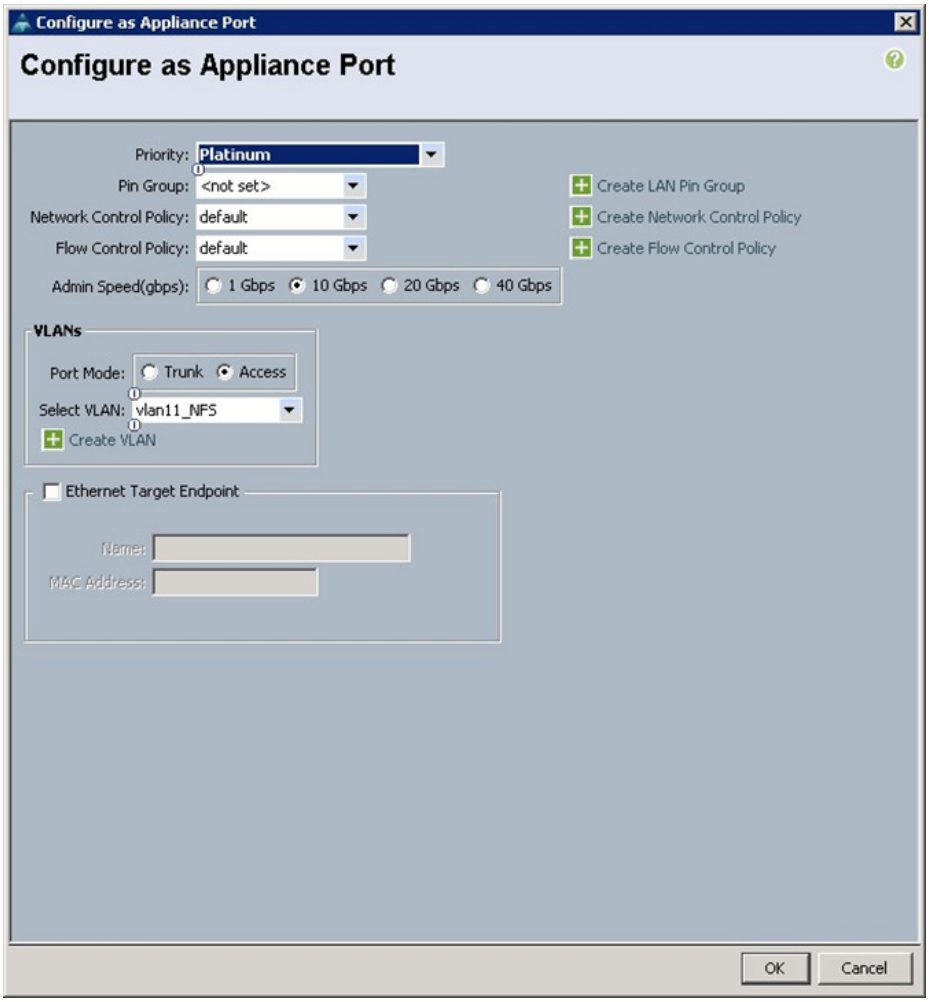
13. Click **OK**.
14. In the message box that appears, click **OK.**
15. Select **Equipment > Fabric Interconnects > Fabric Interconnect B (secondary) > Fixed Module**.
16. Expand the Unconfigured Ethernet Ports.
17. Select the port number 2, and select **Reconfigure > Configure as an Appliance Port**.

Figure 50 Configuring Fabric B Appliance Port



18. A confirmation message box appears. Click **Yes**, then **OK** to continue.
19. Select Platinum for the Priority.
20. Keep the Pin Group as <not set>.
21. Keep the Network Control Policy as Default.
22. Keep the Flow Control Policy as Default.
23. Select the **10Gbps** radio button for the Admin Speed.
24. Select the **Access** radio button for the Port Mode.
25. Select vlan11_NFS from the drop-down menu for the Select VLAN.

Figure 51 Configuring Appliance Port



26. Click **OK**.
27. In the message box that appears, click **OK**.

Server and Software Configuration

Service profile template creation is explained in "Fabric Configuration" section on page 18.

The following sections provide a detailed configuration procedure of the Cisco UCS C-Series Servers. These steps should be followed precisely because a failure to do so could result in an improper configuration.

Performing Initial Setup of C-Series Servers

These steps provide details for initial setup of the Cisco UCS C-Series Servers. It is important to get the systems to a known state with the appropriate firmware package.

Logging into the Cisco UCS 6200 Fabric Interconnects

To log into the Cisco UCS Manager application through Cisco UCS 6200 Series Fabric Interconnect, follow these steps:

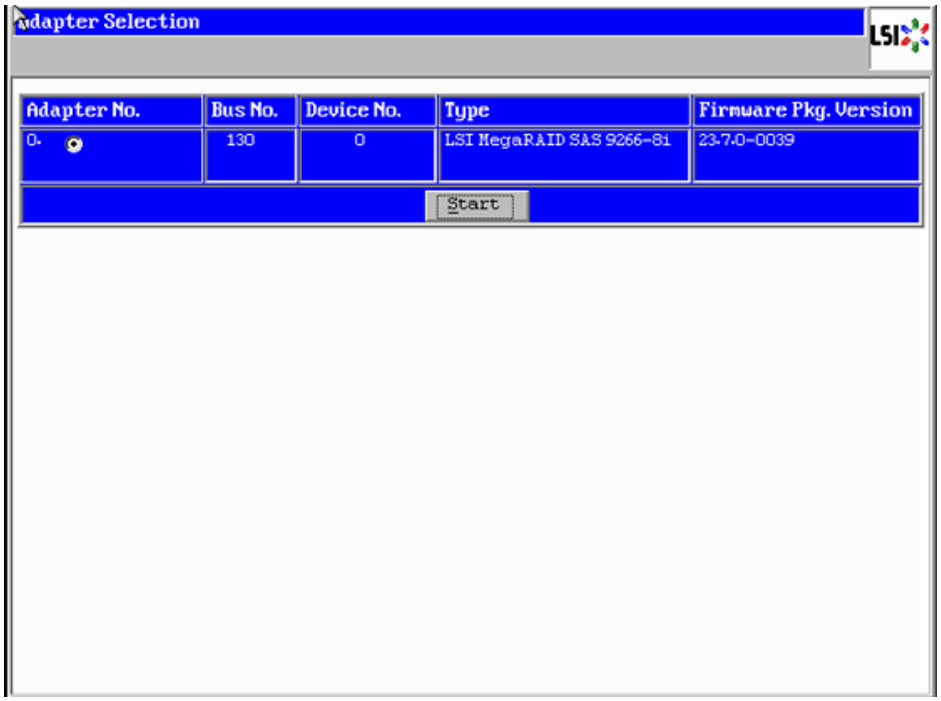
1. Log in to the Cisco UCS 6200 fabric interconnects and launch the Cisco UCS Manager application.
2. In the UCSM GUI, select the Servers tab.
3. Select Servers, right-click on Servers and Open KVM Console.
4. Navigate to the Actions section and click **KVM Console**.

Configuring Disk Drives for OS

There are several ways to configure RAID: using LSI WebBIOS Configuration Utility embedded in the MegaRAID BIOS, booting DOS and running 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. A RAID1 volume of two disk drives is configured for the operating system:

1. Once the server has booted and the MegaRAID Controller has been detected, the following will appear on the screen:
   - Press <Ctrl><H> for WebBIOS.
   - Press Ctrl+H immediately.
   - The Adapter Selection window appears.
2. Click **Start** to continue.

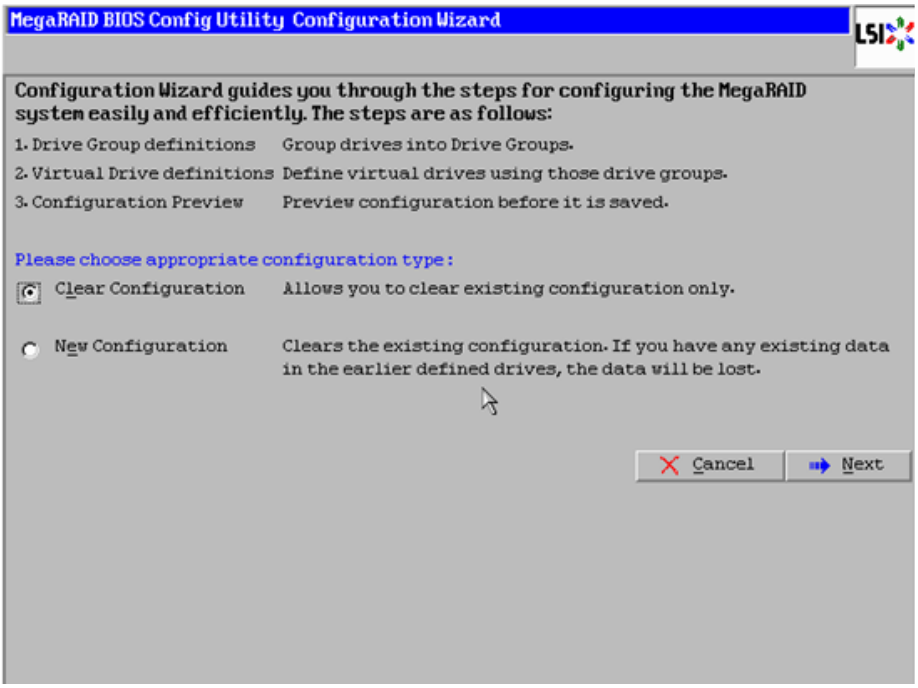
Figure 52 RAID Configuration for LSI MegaRAID SAS Controllers



3. Click **Configuration Wizard**.

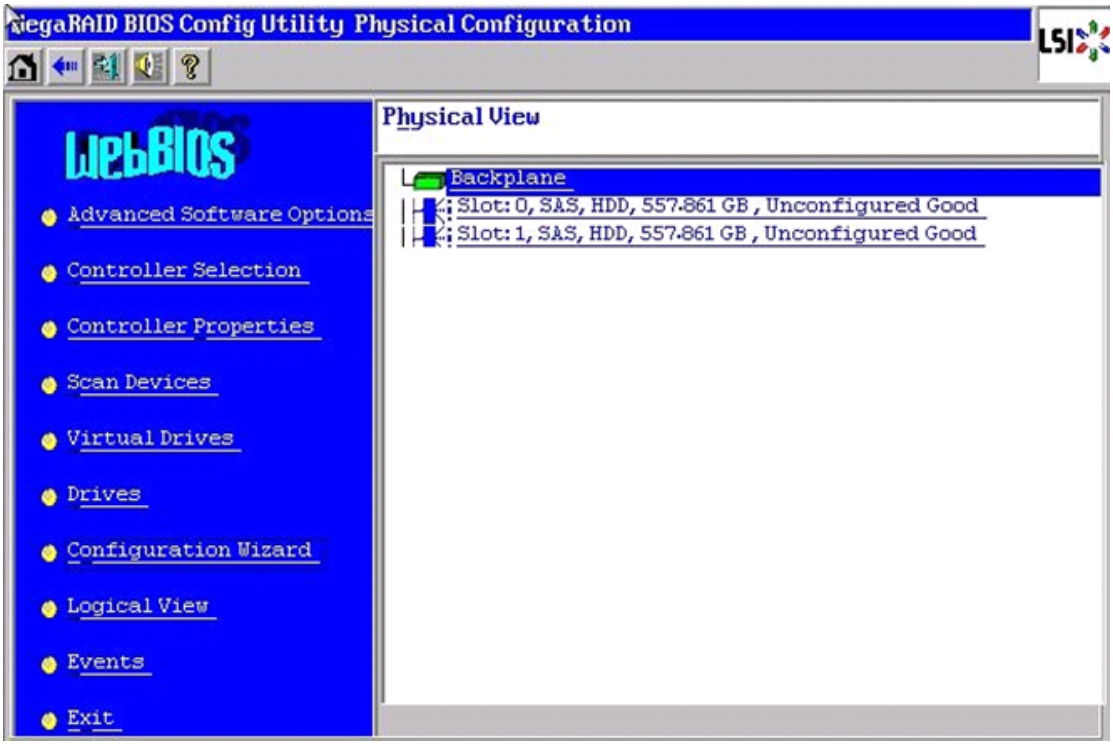
4. In the configure wizard window, select the configuration type as Clear Configuration and click **Next** to clear the existing configuration.

Figure 53 Clearing Existing Configuration



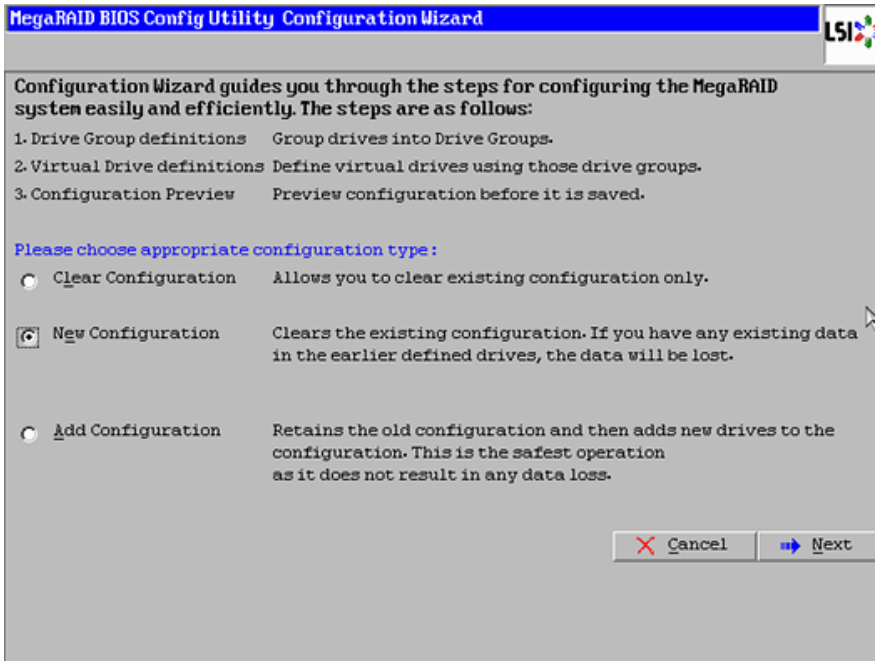
5. Click **Yes** when asked to confirm clear configuration.
6. In the Physical View, make sure all the drives are Unconfigured Good.

Figure 54 Physical View of Unconfigured Drives



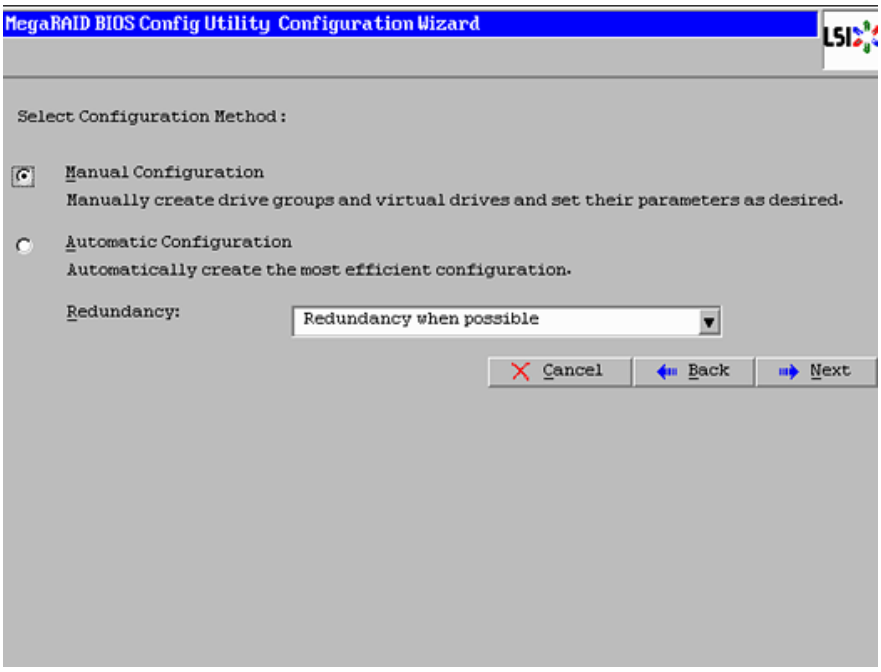
7. Click **Configuration Wizard**.
8. In the configure wizard window, select the configuration type as New Configuration and click **Next**.

Figure 55 Selecting New Configuration



9. Select the configuration method to be Manual Configuration to have control over all attributes of the new storage configuration such as drive groups, virtual drives, and to set their parameters.

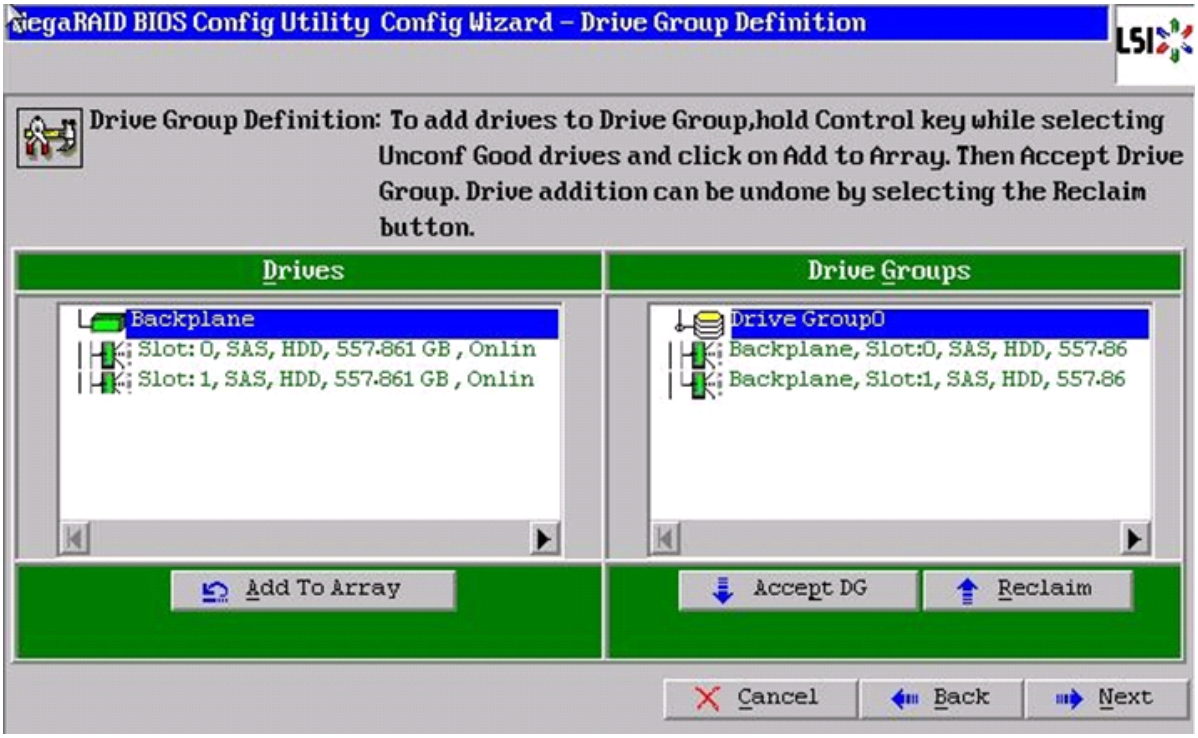
Figure 56 Selecting Manual Configuration



10. Click **Next**.

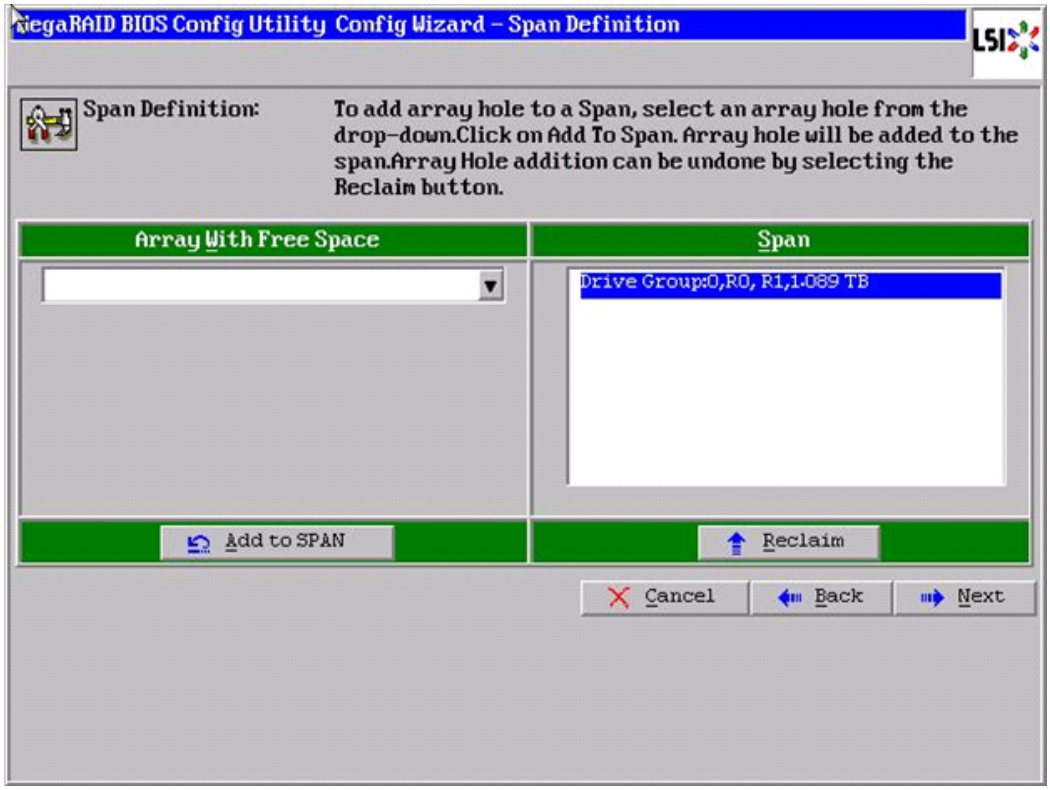
11. The Drive Group Definition window appears. In this window select the two drives to create drive groups.
12. Click **Add to Array** to move the drives to a proposed drive group configuration in the Drive Groups pane. Click **Accept DG** and then, click **Next**.

Figure 57 Moving Drives to Drive Groups



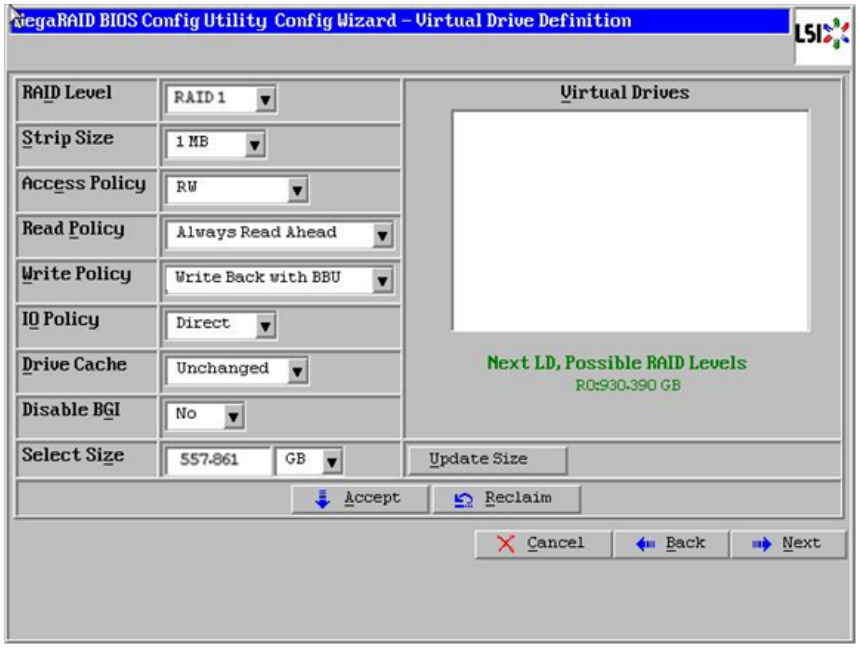
13. In the Span definitions Window, Click **Add to SPAN** and then, click **Next**.

Figure 58 Adding Arrayhole to Span



14. In Virtual Drive definitions window, follow these steps to configure read normal and write through modes:
    a. Change Strip Size to 1MB. A larger strip size produces higher read performance. If your server regularly performs random read requests, select a smaller strip size.
    b. From the read Policy drop down list, choose Read Normal.
    c. From the Write Policy drop down list, choose Write Back with BBU.
    d. Make Sure RAID Level is set to RAID1.

Figure 59 Defining Virtual Drive

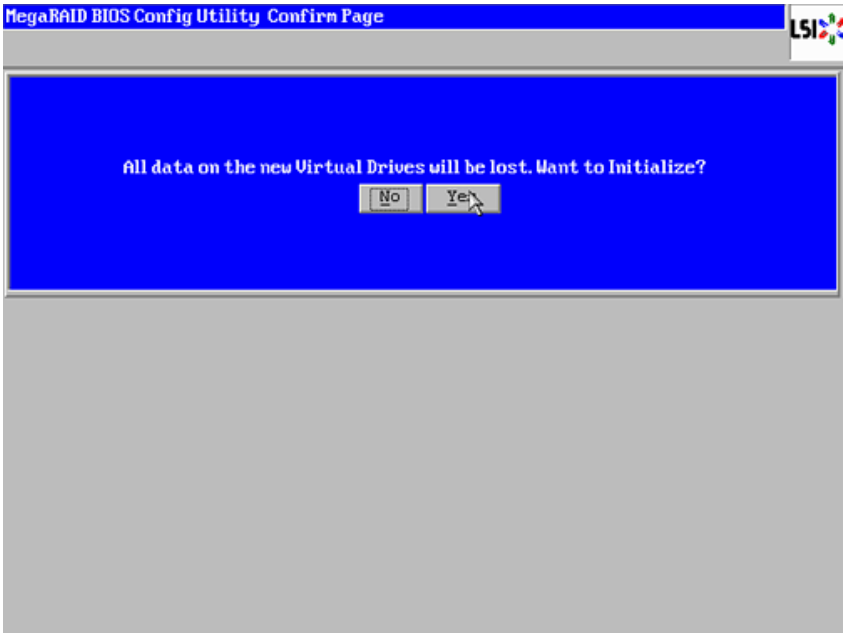


e. Click **Accept** to accept the changes to the virtual drive definitions.

f. Click **Next**.

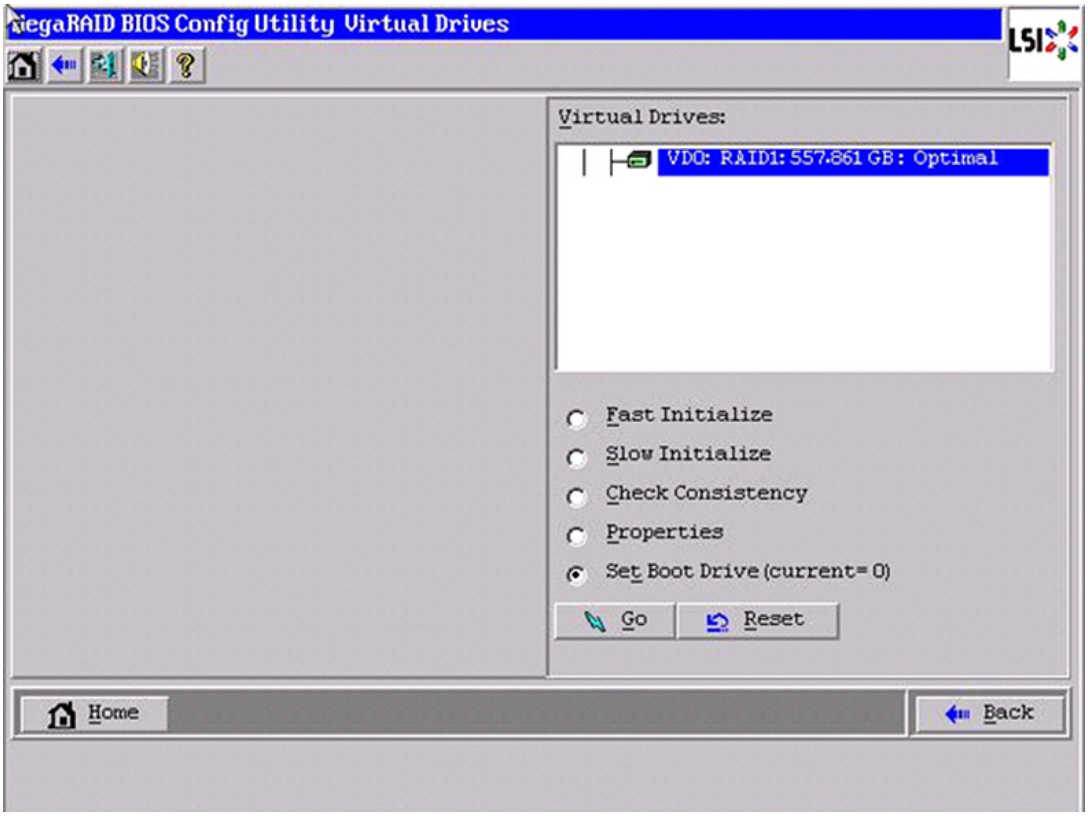
15. After you finish the virtual drive definitions, click **Next**. The Configuration Preview window appears.
16. Review the virtual drive configuration in the Configuration Preview window and click **Accept** to save the configuration.
17. Click **Yes** to save the configuration.
18. Click **Yes**. When asked to confirm to initialize.

Figure 60 Confirmation to Initialize



19. Set VD0 as the Boot Drive and click **Go**.

Figure 61 Setting Boot Drive



20. Click **Home**.
21. Review the Configuration and Click **Exit**.

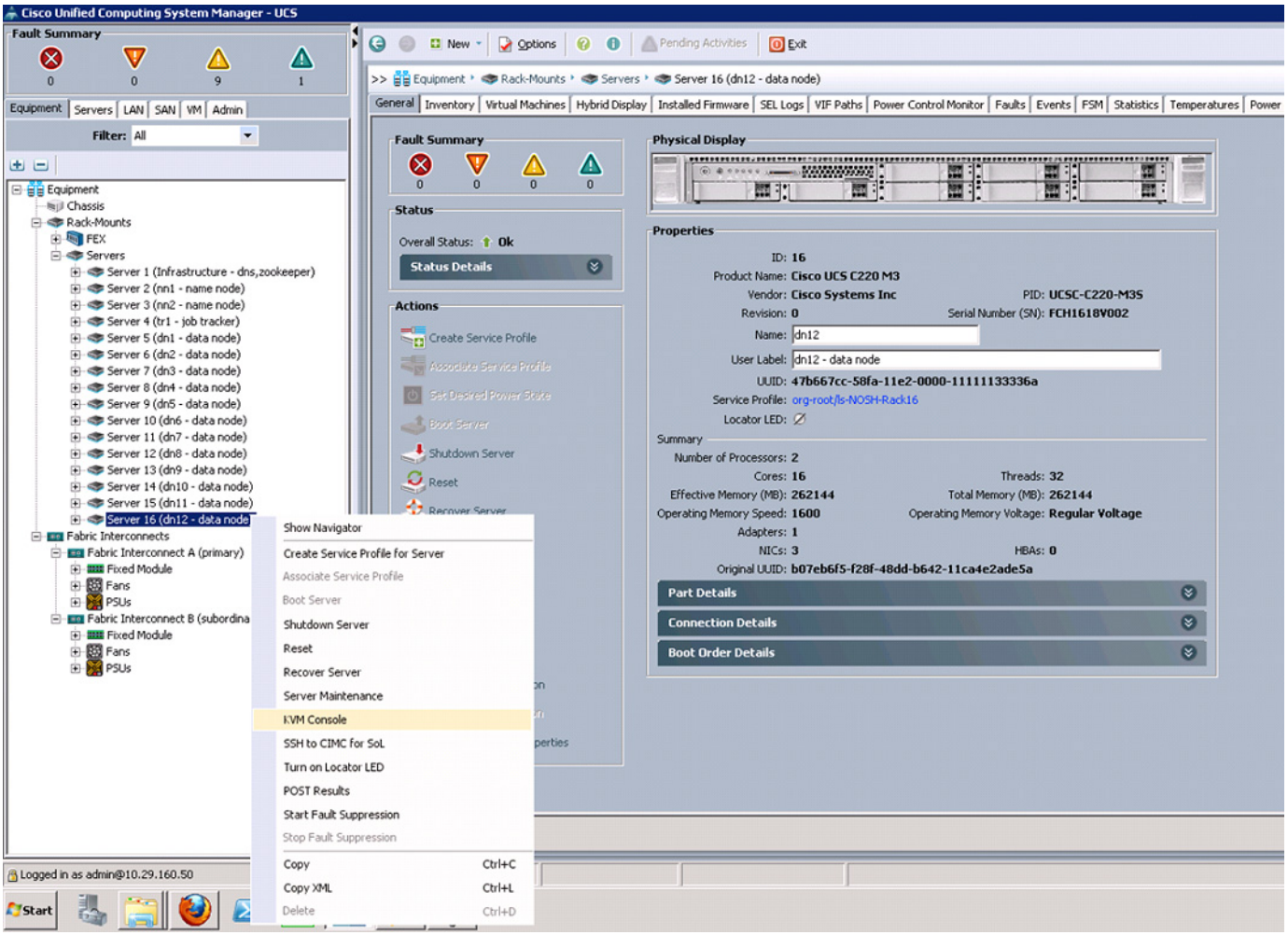
Installing Red Hat Enterprise Linux Server 6.2 using KVM

One of the options to install RHEL is explained in this section.

You can install Red Hat Enterprise Linux Server 6.2 using the KVM console of Cisco UCS Manager. To open the KVM console, follow these steps:

1. Login to Cisco UCS Manager.
2. Select Equipment tab and navigate to the servers.
3. Right-click on the server and select KVM Console.

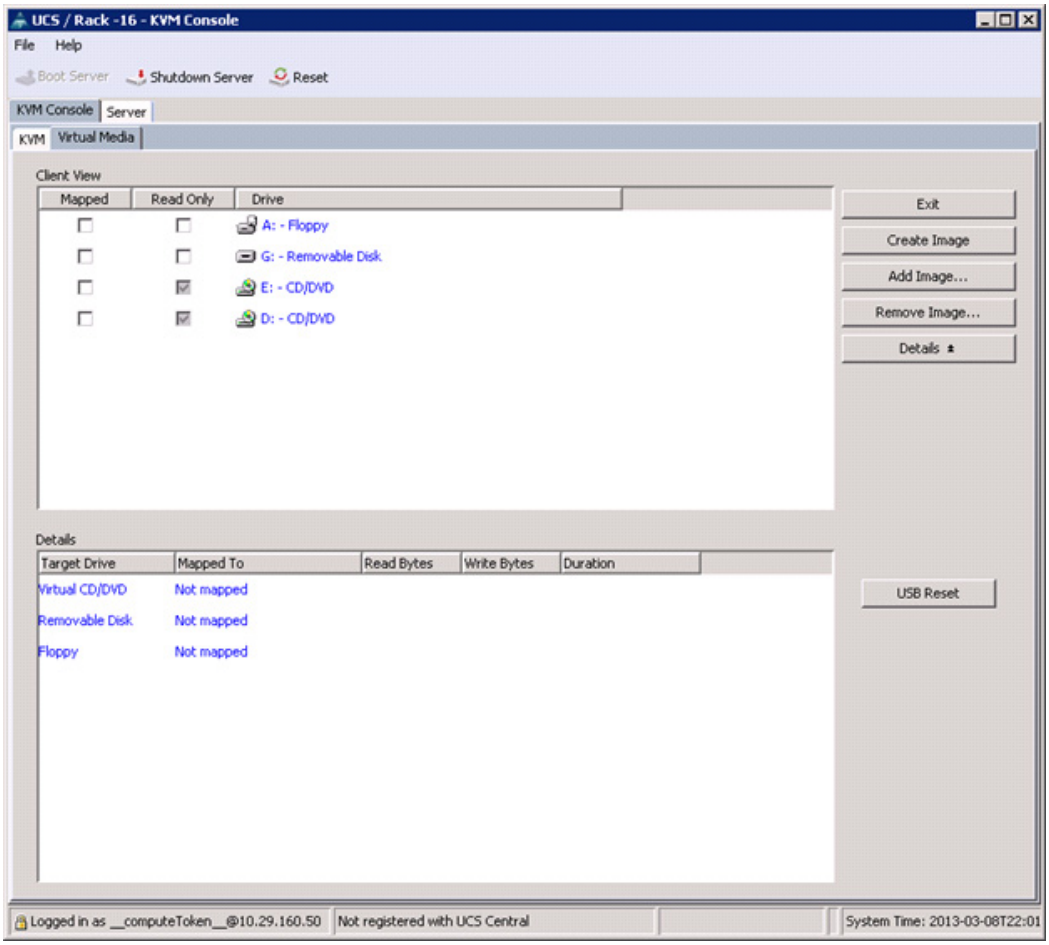
Figure 62 Launching KVM Console



To install Red Hat Linux Server 6.2, follow these steps:

1. In the KVM window, select the Virtual Media tab.

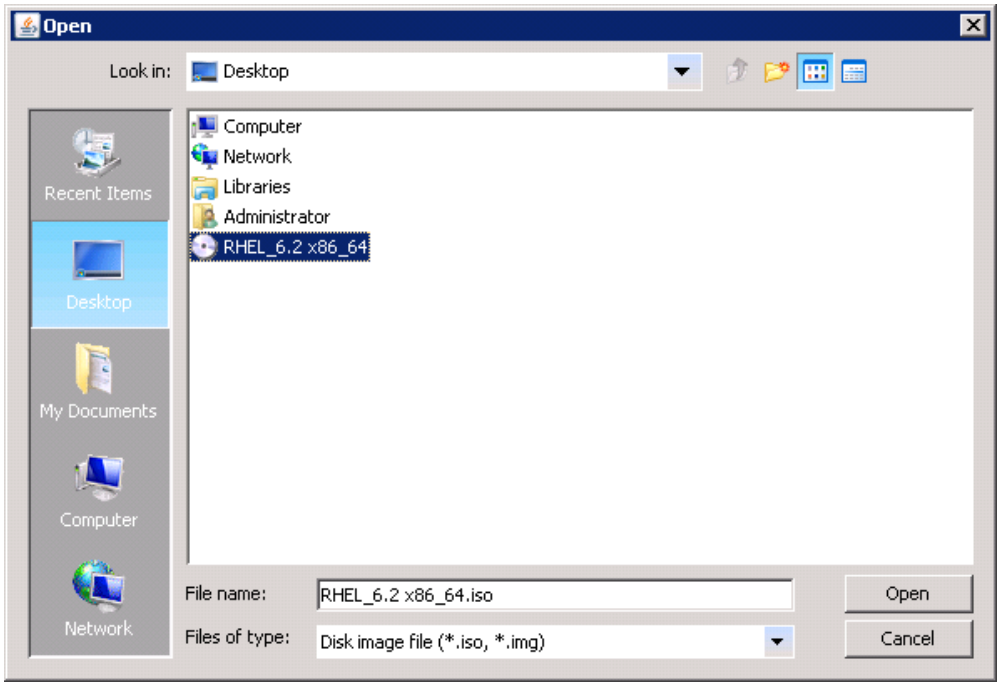
Figure 63 Adding ISO Image



2. Click **Add Image** in the window that appeared.
3. Browse to the Red Hat Enterprise Linux Server 6.2 installer ISO image file.

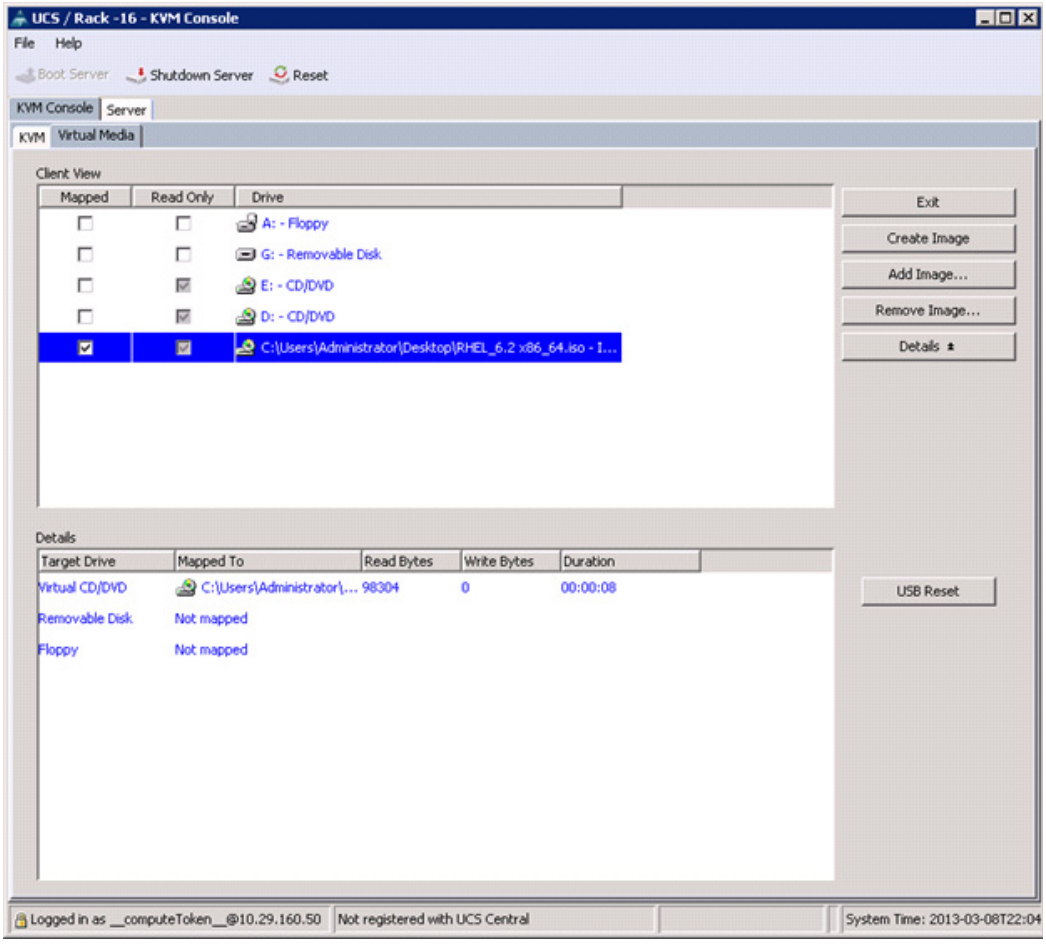
Note The Red Hat Enterprise Linux 6.2 DVD is assumed to be on the client machine. If not, create an ISO Image of Red Hat Enterprise Linux DVD using software such as ImgBurn or MagicISO.

Figure 64 Selecting the Red Hat Enterprise Linux ISO Image



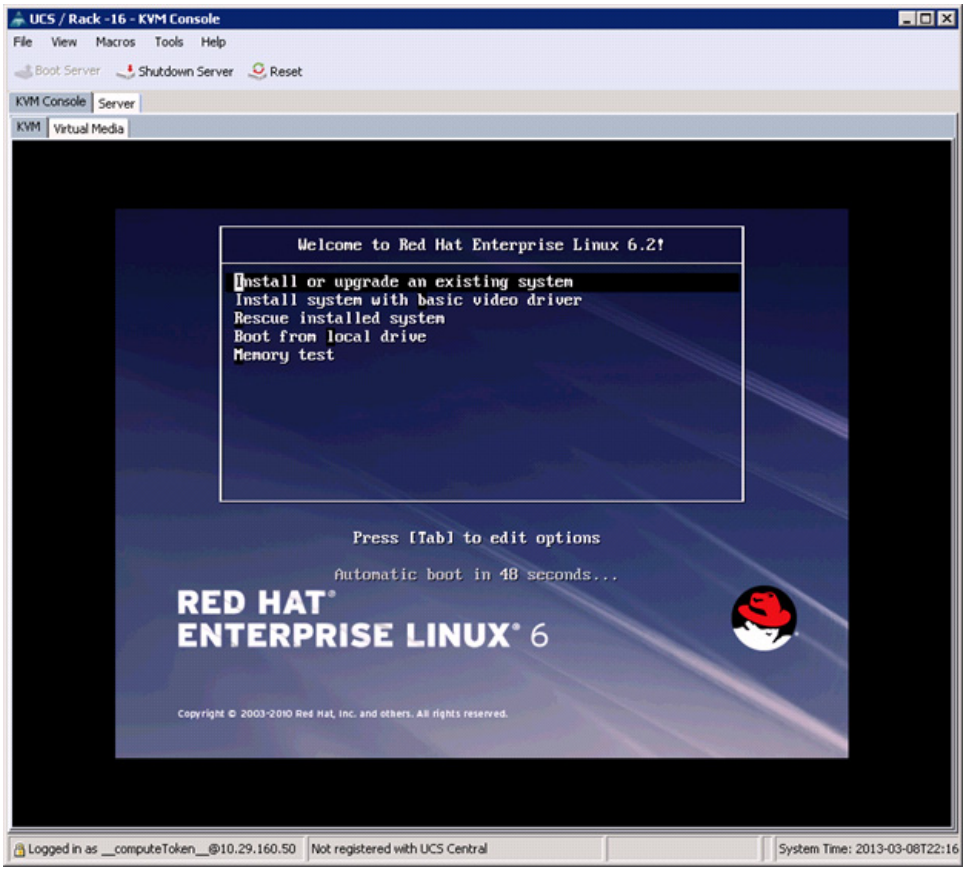
4. Click **Open** to add the image to the list of virtual media.
5. Check the check box for Mapped, next to the entry corresponding to the image you just added.

Figure 65 Mapping the ISO Image



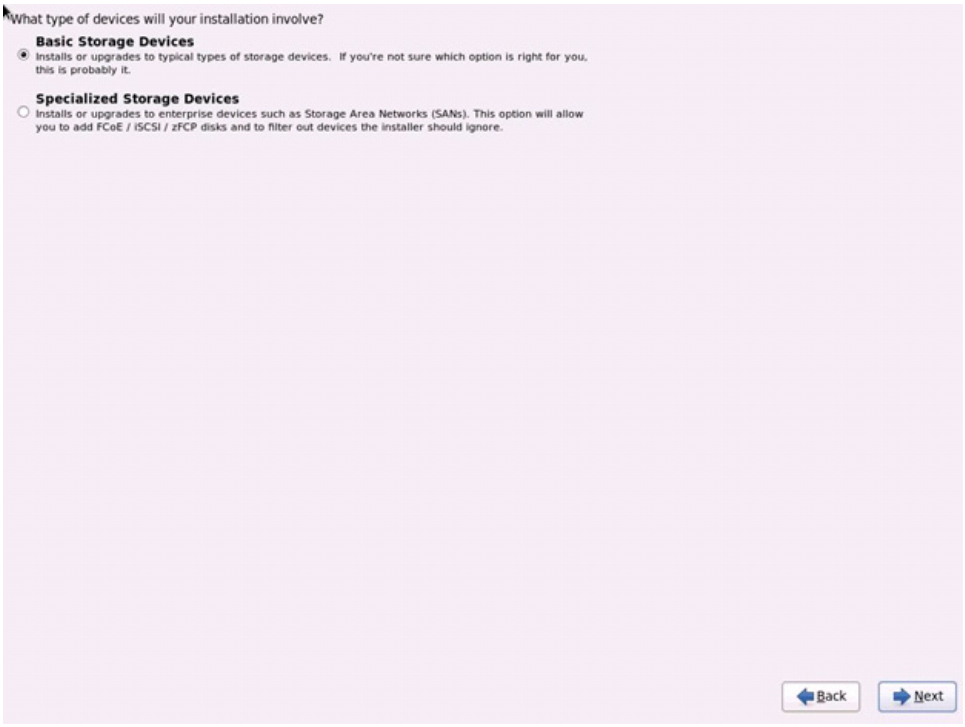
6. In the KVM window, select the KVM tab to monitor during boot.
7. In the KVM window, select the Boot Server button in the upper left corner.
8. Click **OK**.
9. Click **OK** to reboot the system.
10. On reboot, the machine detects the presence of the Red Hat Enterprise Linux Server 6.2 install media.
11. Select the Install or Upgrade an Existing System option.

Figure 66 Selecting the RHEL Installation Option



12. Skip the Media test as we are installing from ISO Image, click **Next** to continue.
13. Select Language for the Installation and click **Next**.
14. Select Basic Storage Devices and click **Next**.

Figure 67 Selecting Storage Device Type



15. Select Fresh Installation and click **Next**.

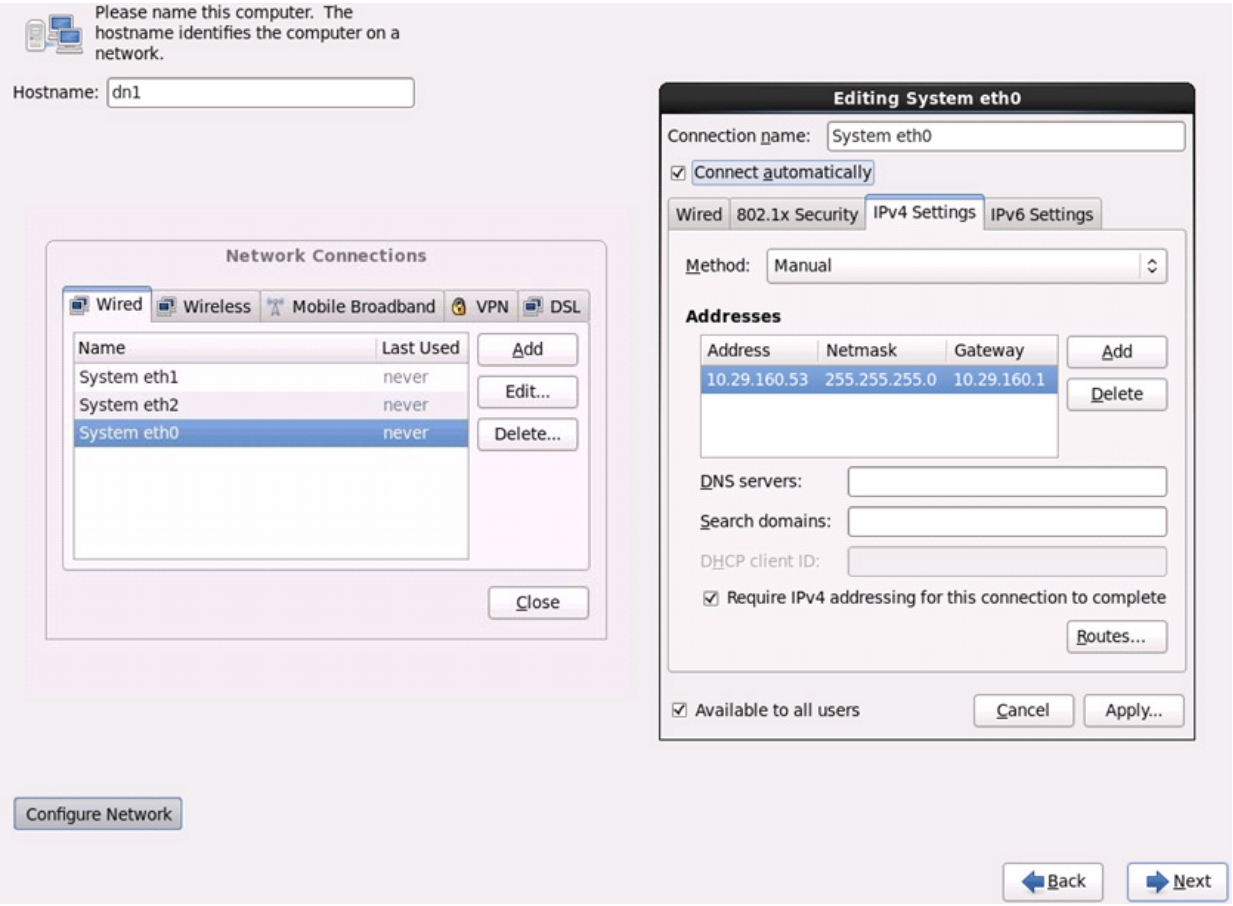
Figure 68 Selecting Installation Type

16. Enter the Host name of the server and click **Configure Network**.

Figure 69 Entering the Host Name

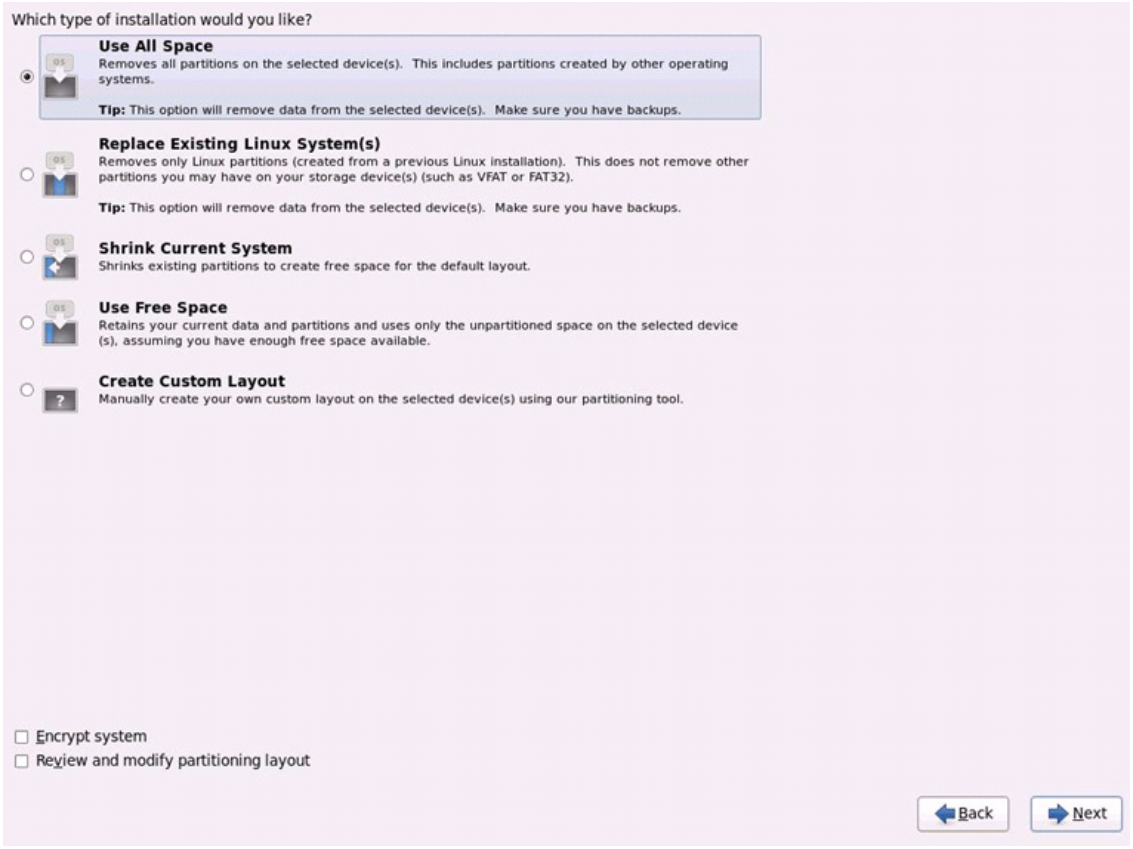
17. Network Connections window appears.
18. In the Network Connections window, Select the tab Wired.
19. Select the interface System eth0, and click **Edit**.
20. Editing System eth0 window appears.
21. Check the Connect automatically check box.
22. For the field Method, select Manual from the drop down list.
23. Click **Add** and enter IP Address, Netmask and Gateway.
24. Click **Apply**.

Figure 70 Configuring Network Connections



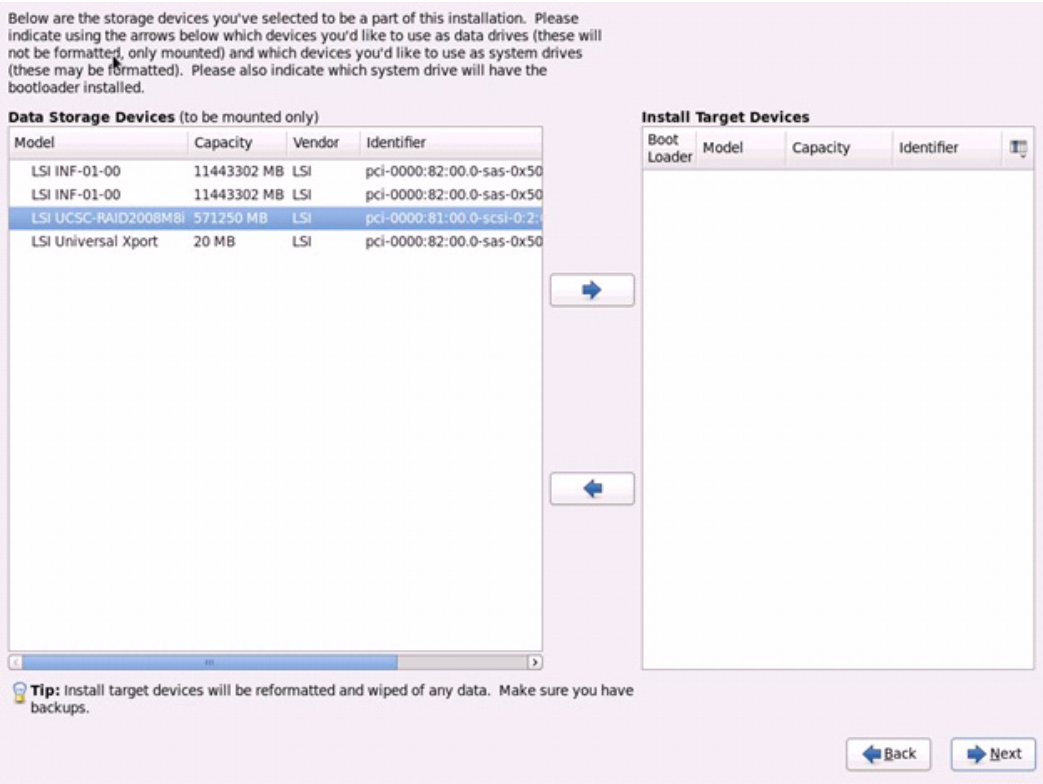
25. Select the Appropriate Time Zone and click **Next**.
26. Enter the root Password and click **Next**.
27. Select Use All Space and Click **Next**.

Figure 71 Selecting RHEL Install Type



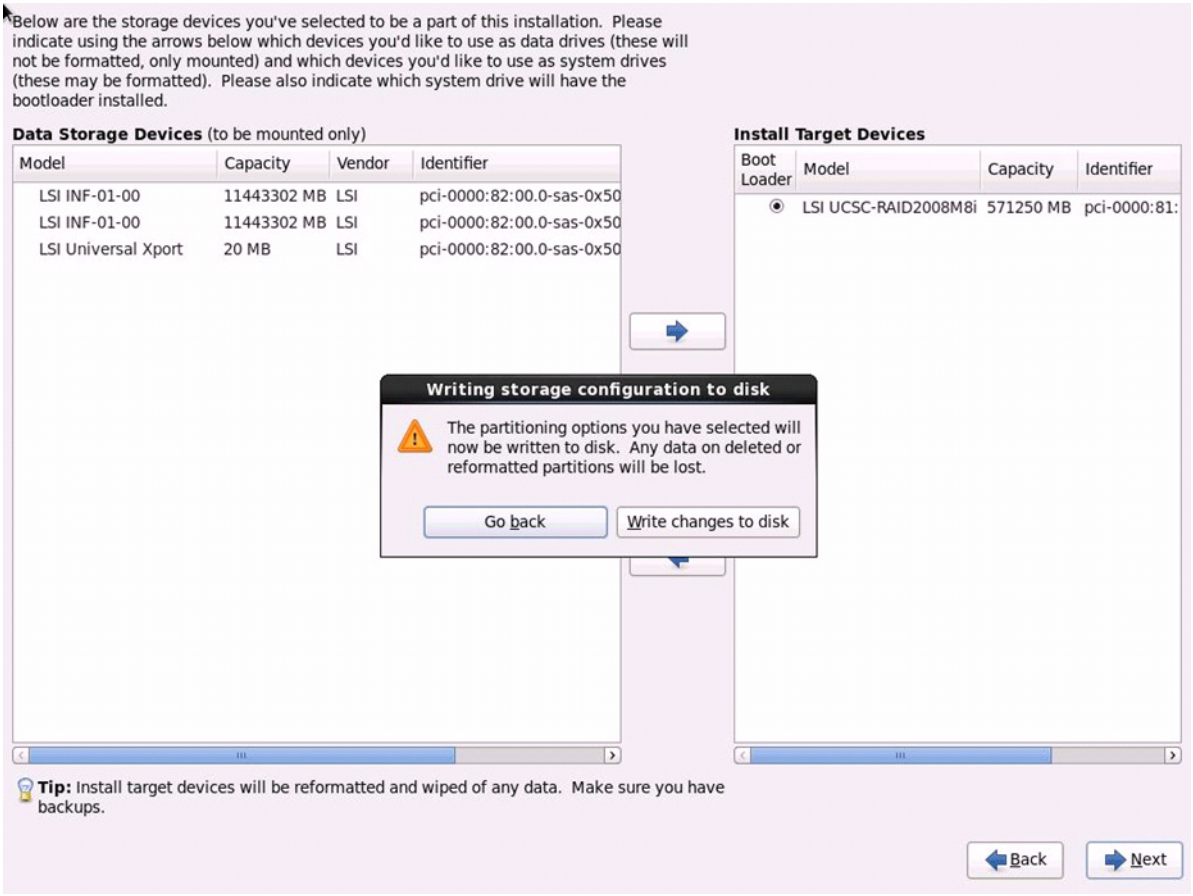
28. Select appropriate boot device. In this example, LSI UCSC-RAID 2008M-8i is selected. Click to add the selected boot device to appear in the right pane under Install Target Devices and click **Next**.

Figure 72 Selecting the Data Storage Device



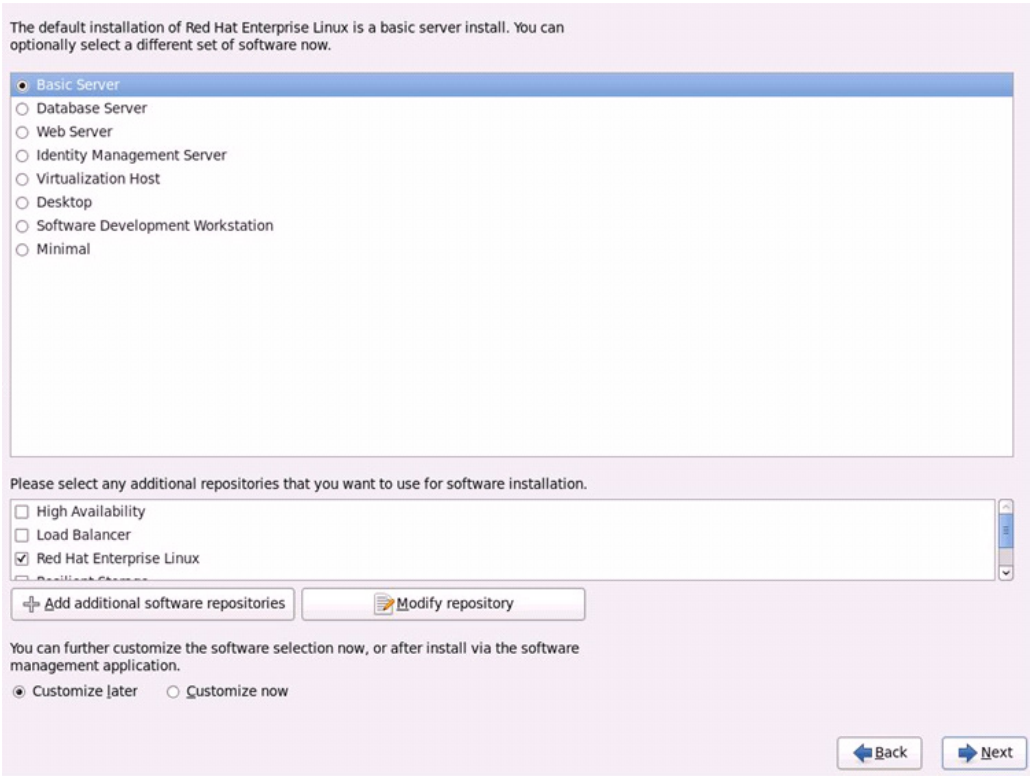
29. Click **write changes to the disks** and then, click **Next**.

Figure 73 Writing Partitioning Options into the Disk



30. Select Basic Server Installation and Click **Next.**

Figure 74 Selecting RHEL Installation Option



31. After the installer is finished loading, press **Enter** to continue with the install.

Figure 75 Installation Process in Progress

Post OS Install Configuration

Infrastructure Node

This section describes the steps needed to implement an infrastructure node for the cluster. The infrastructure node may provide some or all of the following services to the namenodes and datanodes in the cluster:

- NTP – time server
- DNS resolution for cluster-private hostnames
- DHCP IP address assignment for cluster-private NICs
- Local mirror of one or more software repository and/or distribution
- Management server

Note This section assumes that only the default basic server install has been done.

Setting Up Password-less Login

To manage all of the clusters nodes from the Infrastructure 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.

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

1. Login to the Infrastructure Node

```
ssh 10.29.160.53
```

2. Run the ssh-keygen command to create both public and private keys on the admin node.
3. Then run the following command from the Infrastructure 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..68}; do echo -n "$IP -> "; ssh-copy-id -i ~/.ssh/id_rsa.pub
10.29.160.$IP; done

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

Enter the password of the remote host.

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.

The tool can be downloaded from https://code.google.com/p/parallel-ssh/

Fetch and install this tool via the following commands:

```
cd /tmp/
curl https://parallel-ssh.googlecode.com/files/pssh-2.3.1.tar.gz -O -L
tar xzf pssh-2.3.1.tar.gz
cd pssh-2.3.1
python setup.py install
```

To make use of pssh, a file containing just only the IP addresses of the nodes in the cluster needs to be created. The following was used for the contents of the /root/pssh.hosts file on all of the nodes and will need to be customized to fit your implementation:

```
# /root/pssh.hosts - cluster node IPs or names
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
```

This file is used with pssh by specifying the -h option on the command line. For example, the following command will execute the hostname command on all of the nodes listed in the /root/pssh.hosts file:

```
pssh -h /root/pssh.hosts -A  hostname
```

For information on the -A option and other pssh options, use one or both of the following commands:

```
pssh -help
man pssh
```

Create Local Redhat Repo

If your infrastructure node and your cluster nodes have Internet access, you may be able to skip this section.

To create a repository using RHEL DVD or ISO on the infrastructure node (in this deployment 10.29.160.53 is as an infrastructure node), create a directory with all the required rpms, run the createrepo command and then publish the resulting repository.

1. Create the directories where the local copies of the software installation packages will reside. In this example, they are created under the /var/www/html/ directory.

   ```
   mkdir -p /var/www/html/JDK/
   mkdir -p /var/www/html/RHEL/6.2/
   ```

2. Then mount the RHEL DVD. This can be done by loading the DVD disc into a DVD drive attached to the server or by mounting the .iso image as in this example.

   ```
   mount /rhel-server-6.2-x86_64-dvd.iso/mnt -t iso9660 -o ro,loop=/dev/loop1
   ```

3. Next, copy the contents of the DVD to the /var/www/html/RHEL/6.2/ directory and then verify that the contents copied match their source.

   ```
   cd /mnt/;tar -c -p -b 128 -f - .
   cd /var/www/html/RHEL/6.2/;tar -x -p -b 128 -f - .
   diff -r /mnt/ /var/www/html/RHEL/6.2/
   ```

4. Now create a .repo file for the yum command.

   ```
   cat > /var/www/html/RHEL/6.2/rhel62copy.repo
   [rhel6.2]
   name=Red Hat Enterprise Linux 6.2
   baseurl=file:///var/www/html/RHEL/6.2/
   gpgcheck=0
   enabled=1
   ```

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

5. Copy the rhel62copy.repoto all the nodes of the cluster.

   ```
   pscp -h  /root/pssh.hosts \
   /var/www/html/RHEL/6.2/rhel62copy.repo /etc/yum.repos.d/
   ```

6. 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/RHEL/6.2/
   createrepo .
   yum clean all
   ```

7. Update Yum on all nodes.

```
pssh -h /root/allnodes "yum clean all"
```

Install Required Packages

This section assumes that only the default basic server install has been done.

Table 5 provides a list of packages that are required.

Table 5 Required list of packages

Package	Description
xfsprogs	Utilities for managing XFS filesystem
jdk	Java SE Development Kit6, Update 39(JDK 6u39) or more recent
Utilities	dnsmasq, httpd, lynx

Note Installation of Java and JDK is detailed in a separate section.

Create the following script install_packages.sh to install required packages:

Script install_packages.sh

```
yum –y install dnsmasq httpd lynx
# get and install xfsprogs from local repo on Infrastructure node
cd /tmp/
curl http://10.29.160.53/RHEL/6.2/Packages/xfsprogs-3.1.1-6.el6.x86_64.rpm  -O  -L
rpm -i /tmp/xfsprogs-3.1.1-6.el6.x86_64.rpm
```

Copy script disable_services.sh to all nodes and run the script on all nodes:

```
pscp –h /root/pssh.hosts /root/install_packages.sh /root/
pssh –h /root/pssh.hosts "/root/install_packages.sh"
```

Disable SELinux

Execute the following commands to disable SELinux on all the nodes:

```
pssh –h /root/pssh.hosts "setenforce 0"
pssh –h /root/pssh.hosts "sed -i -e 's/=enforcing/=disabled/g;'\ /etc/selinux/config"
Disable Unwanted Services
```

Execute the following commands as a script to disable and turn off unwanted services on all nodes:

Script disable_services.sh

```
$cat disable_services.sh

# disble/shutdown things we do not need
for X in bluetooth certmonger cgconfigd cgred cpuspeed cups dnsmasq \
ebtables fcoe fcoe-target ip6tables iptables iscsi iscsid ksm ksmtuned \
libvirtd-guests libvirtd postfix psacct qpidd rhnsd rhsmcertd \
sendmail smartd virt-who vsftpd winbind wpa_supplicant ypbind NetworkManager
do
        /sbin/service $X stop
```

```
        /sbin/chkconfig $X off
done
```

Copy script disable_services.sh to all nodes and run the script on all nodes:

```
pscp -h /root/pssh.hosts /root/disable_servicesh.h /root/
pssh -h /root/pssh.hosts "/root/disable_services.sh"
```

Enable and start the httpd service

Before starting the httpd service, you may need to edit the server configuration file (/etc/httpd/conf/httpd.conf) to change one or more of the following settings:

- Listen
- ServerName
- ExtendedStatus
- server-status

Ensure httpd is able to read the repofiles

```
chcon -R -t httpd_sys_content_t /var/www/html/RHEL/6.2/
```

Perform the following commands to enable and start the httpd service:

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

JDK Installation

Download Java SE 6 Development Kit (JDK)

Using a web browser, click on the following link:

http://www.oracle.com/technetwork/java/index.html

and download the latest Java™ SE 6 Development Kit (JDK™6).

Once the JDK6 package has been downloaded, place it in the /var/www/html/JDK/ directory.

Install JDK6 on All Node

Create the following script install_jdk.sh to install JDK:

Script install_jdk.sh

```
# Copy and install JDK
cd /tmp/
curl http://10.29.160.53/JDK/jdk-6u41-linux-x64.bin -O -L
sh ./jdk-6u41-linux-x64.bin -noregister
```

Copy script disable_services.sh to all nodes and run the script on all nodes:

```
pscp -h /root/pssh.hosts /root/install_jdk.sh /root/
pssh -h /root/pssh.hosts "/root/install_jdk.sh"
```

Ext js Installation

From infra-0, download extjs

```
wget
http://s3.amazonaws.com/public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.15/repos/centos6
/extjs/extjs-2.2-1.noarch.rpm
```

Copy the extjs rpm to all nodes from the infra-0 node

```
pscp -h /root/pssh.hosts /root/extjs-2.2-1.noarch.rpm /root/
```

Install extjs on all nodes

```
pssh -h /root/allnodes "yum  -y install /root/extjs-2.2-1.noarch.rpm"
```

Services to Configure On Infrastructure Node

These are some of the other services that you may want to configure on the infrastructure node. This is optional.

DHCP for Cluster Private Interfaces

If DHCP service is needed, it may be done via one of the following services:

- dnsmasq
- dhcp

DNS for Cluster Private Interfaces

Hostname resolution for cluster private interfaces may be done by one or two of the following services running on the infrastructure node:

- **/etc/hosts** file propagated to all nodes in the cluster
- **dnsmasq**
- **bind**

The configuration described in this document used both the /etc/hosts file and the dnsmasq service to provide DNS services. The FAS2220 is the main user of the DNS service in this configuration.

The following was used for the contents of the /etc/resolv.conf file on all of the nodes and will need to be customized to fit your implementation:

```
domain hadoop.local
search hadoop.local
nameserver 10.29.160.53
```

Once configured, the **/etc/resolv.conf** file can be pushed to all nodes via the following command:

```
pssh -h /root/pssh.hosts -A /etc/resolv.conf  /etc/resolv.conf
```

The following was used for the contents of the **/etc/nsswitch.conf** file on all of the nodes and may need to be customized to fit your implementation:

```
# /etc/nsswitch.conf - for all nodes
passwd:     files
```

```
shadow:     files
group:      files
#hosts:     db files nisplus nis dns
hosts:      files dns
ethers:     files
netmasks:   files
networks:   files
protocols:  files
rpc:        files
services:   files
automount:  files nisplus
aliases:    files nisplus
netgroup:   nisplus
publickey:  nisplus
bootparams: nisplus [NOTFOUND=return] files
```

Once configured, the **/etc/nsswitch.conf** file can be pushed to all nodes via the following command:

```
pssh –h /root/pssh.hosts –A /etc/nsswitch.conf  /etc/nsswitch.conf
```

The following was used for the contents of the **/etc/hosts file** on all of the nodes and will need to be customized to fit your implementation:

```
# /etc/hosts file for all nodes
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
localhost-stack
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
10.29.160.1     gateway
#
# NTAP FAS2220 unit
# 0.0.0.0       fas2220-e0P
10.29.160.43    fas2220-e0M.hadoop.local fas2220-e0M
#10.29.160.45    fas2220-e0a.hadoop.local fas2220-e0a
# 0.0.0.0       fas2220-e0b
# 0.0.0.0       fas2220-e0
# 0.0.0.0       fas2220-e0d
192.168.11.43   fas2220-e1a.hadoop.local fas2220-e1a
192.168.11.45   fas2220.hadoop.local fas2220-e1b fas2220
#192.168.11.45   vif-a
#
# NTAP E-Series E5460 units
10.29.160.33    e5460-2-A.hadoop.local e5460-2-A
10.29.160.34    e5460-2-B.hadoop.local e5460-2-B
10.29.160.37    e5460-1-A.hadoop.local e5460-1-A
10.29.160.38    e5460-1-B.hadoop.local e5460-1-B
10.29.160.35    e5460-3-A.hadoop.local e5460-3-A
10.29.160.36    e5460-3-B.hadoop.local e5460-3-B
#
# CISCO eth0 mappings  -VLAN160
10.29.160.53   infra.hadoop.local infra infra-0.hadoop.local infra-0 mailhost
infrastructure-0
10.29.160.54   nn1-0.hadoop.local nn1-0 namenode1-0 namenode-1-0 nn01-0
10.29.160.55   nn2-0.hadoop.local nn2-0 namenode2-0 namenode-2-0 nn02-0
10.29.160.56   tr1-0.hadoop.local tr1-0 tracker1-0 tracker-1-0 tr01-0
10.29.160.57   dn1-0.hadoop.local dn1-0 datanode1-0 datanode-1-0 dn01-0
10.29.160.58   dn2-0.hadoop.local dn2-0 datanode2-0 datanode-2-0 dn02-0
10.29.160.59   dn3-0.hadoop.local dn3-0 datanode3-0 datanode-3-0 dn03-0
10.29.160.60   dn4-0.hadoop.local dn4-0 datanode4-0 datanode-4-0 dn04-0
10.29.160.61   dn5-0.hadoop.local dn5-0 datanode5-0 datanode-5-0 dn05-0
10.29.160.62   dn6-0.hadoop.local dn6-0 datanode6-0 datanode-6-0 dn06-0
10.29.160.63   dn7-0.hadoop.local dn7-0 datanode7-0 datanode-7-0 dn07-0
10.29.160.64   dn8-0.hadoop.local dn8-0 datanode8-0 datanode-8-0 dn08-0
10.29.160.65   dn9-0.hadoop.local dn9-0 datanode9-0 datanode-9-0 dn09-0
10.29.160.66   dn10-0.hadoop.local dn10-0 datanode10-0 datanode-10-0
```

```
10.29.160.67   dn11-0.hadoop.local dn11-0 datanode11-0 datanode-11-0
10.29.160.68   dn12-0.hadoop.local dn12-0 datanode12-0 datanode-12-0
#
# CISCO eth1 mappings - VLAN11
192.168.11.11   infra-1 infra-1 infrastructure-1
192.168.11.12   nn1-1.hadoop.local nn1-1 namenode1-1 nn01-1
192.168.11.13   nn2-1.hadoop.local nn2-1 namenode2-1 nn02-1
192.168.11.14   tr1-1.hadoop.local tr1-1 tracker1-1 tracker-1-1 tr01-1
192.168.11.15   dn1-1.hadoop.local dn1-1 dn01-1
192.168.11.16   dn2-1.hadoop.local dn2-1 dn02-1
192.168.11.17   dn3-1.hadoop.local dn3-1 dn03-1
192.168.11.18   dn4-1.hadoop.local dn4-1 dn04-1
192.168.11.19   dn5-1.hadoop.local dn5-1 dn05-1
192.168.11.20   dn6-1.hadoop.local dn6-1 dn06-1
192.168.11.21   dn7-1.hadoop.local dn7-1 dn07-1
192.168.11.22   dn8-1.hadoop.local dn8-1 dn08-1
192.168.11.23   dn9-1.hadoop.local dn9-1 dn09-1
192.168.11.24   dn10-1.hadoop.local dn10-1
192.168.11.25   dn11-1.hadoop.local dn11-1
192.168.11.26   dn12-1.hadoop.local dn12-1
#
# eth2 mappings - VLAN12
192.168.12.11   infra-2.hadoop.local infra-2 infrastructure-2
192.168.12.12   nn1-2.hadoop.local nn1-2 namenode1-2 nn01-2
192.168.12.13   nn2-2.hadoop.local nn2-2 namenode2-2 nn02-2
192.168.12.14   tr1-2.hadoop.local tr1-2 tracker1-2 tracker-1-2 tr01-2
192.168.12.15   dn1-2.hadoop.local dn1-2 dn01-2
192.168.12.16   dn2-2.hadoop.local dn2-2 dn02-2
192.168.12.17   dn3-2.hadoop.local dn3-2 dn03-2
192.168.12.18   dn4-2.hadoop.local dn4-2 dn04-2
192.168.12.19   dn5-2.hadoop.local dn5-2 dn05-2
192.168.12.20   dn6-2.hadoop.local dn6-2 dn06-2
192.168.12.21   dn7-2.hadoop.local dn7-2 dn07-2
192.168.12.22   dn8-2.hadoop.local dn8-2 dn08-2
192.168.12.23   dn9-2.hadoop.local dn9-2 dn09-2
192.168.12.24   dn10-2.hadoop.local dn10-2
192.168.12.25   dn11-2.hadoop.local dn11-2
192.168.12.26   dn12-2.hadoop.local dn12-2
```

When configured, the **/etc/hosts** file can be pushed to all nodes through the following command:

```
pssh –h /root/pssh.hosts –A /etc/hosts  /etc/hosts
```

The following was used for the contents of the **/etc/dnsmasq.conf** file on the infrastructure node and will need to be customized to fit your implementation should you choose to use the **dnsmasq** service:

```
# Configuration file for dnsmasq.
#
# Format is one option per line, legal options are the same
# as the long options legal on the command line. See
# "/usr/sbin/dnsmasq --help" or "man 8 dnsmasq" for details.
domain-needed
bogus-priv
filterwin2k
no-resolv
local=/hadoop.local/
address=/doubleclick.net/127.0.0.1
address=/www.google-analytics.com/127.0.0.1
address=/google-analytics.com/127.0.0.1
interface=eth0
interface=eth1
interface=eth2

bind-interfaces
```

```
expand-hosts

domain=hadoop.local,10.29.160.0/24,local
domain=hadoop.local,192.168.11.0/24,local
domain=hadoop.local,192.168.12.0/24,local
#
dhcp-range=tag:mgmt,10.29.160.54,10.29.160.68,255.255.255.0,24h
dhcp-range=tag:csco_eth1,192.168.11.12,192.168.11.39,255.255.255.0,24h
dhcp-range=tag:csco_eth2,192.168.12.12,192.168.12.39,255.255.255.0,24h
#
dhcp-range=tag:data11,192.168.11.40,192.168.11.49,255.255.255.0,24h
dhcp-range=tag:data12,192.168.12.40,192.168.12.49,255.255.255.0,24h
#

# NTAP
# E-Series E5460 units
dhcp-host=net:mgmt,00:08:E5:1F:69:34,10.29.160.33,e5460-3-a
dhcp-host=net:mgmt,00:80:E5:1F:83:08,10.29.160.34,e5460-3-b
#
dhcp-host=net:mgmt,00:08:E5:1F:69:F4,10.29.160.35,e5460-2-a
dhcp-host=net:mgmt,00:08:E5:1F:9F:2C,10.29.160.36,e5460-2-b
#
dhcp-host=net:mgmt,00:08:E5:1F:6B:1C,10.29.160.37,e5460-1-a
dhcp-host=net:mgmt,00:08:E5:1F:67:A8,10.29.160.38,e5460-1-b
#
# NTAP
# FAS2220 unit
dhcp-host=net:mgmt,00:a0:98:30:58:1d,10.29.160.43,fas2220-e0M
dhcp-host=net:mgmt,00:a0:98:30:58:18,10.29.160.45,fas2220-e0a
dhcp-host=net:data11,00:a0:98:1a:19:6c,192.168.11.43,fas2220-e1a
dhcp-host=net:data11,00:a0:98:1a:19:6d,192.168.11.45,fas2220
#
# CISCO
# management (eth0)
# name nodes and tracker nodes
dhcp-host=net:mgmt,00:25:B5:02:20:6F,10.29.160.53,infra-0
dhcp-host=net:mgmt,00:25:B5:02:20:5F,10.29.160.54,nn1-0
dhcp-host=net:mgmt,00:25:B5:02:20:0F,10.29.160.55,nn2-0
dhcp-host=net:mgmt,00:25:B5:02:20:FF,10.29.160.56,tr1-0
dhcp-host=net:mgmt,00:25:B5:02:20:BF,10.29.160.57,dn1-0
dhcp-host=net:mgmt,00:25:B5:02:20:8E,10.29.160.58,dn2-0
dhcp-host=net:mgmt,00:25:B5:02:20:7E,10.29.160.59,dn3-0
dhcp-host=net:mgmt,00:25:B5:02:20:2E,10.29.160.60,dn4-0
dhcp-host=net:mgmt,00:25:B5:02:20:1E,10.29.160.61,dn5-0
dhcp-host=net:mgmt,00:25:B5:02:20:DE,10.29.160.62,dn6-0
dhcp-host=net:mgmt,00:25:B5:02:20:CE,10.29.160.63,dn7-0
dhcp-host=net:mgmt,00:25:B5:02:20:9D,10.29.160.64,dn8-0
dhcp-host=net:mgmt,00:25:B5:02:20:4D,10.29.160.65,dn9-0
dhcp-host=net:mgmt,00:25:B5:02:20:3D,10.29.160.66,dn10-0
dhcp-host=net:mgmt,00:25:B5:02:21:0D,10.29.160.67,dn11-0
#
# 10GbE cluster members (eth1)
# name nodes and tracker nodes
#
dhcp-host=net:data11,00:25:B5:02:20:9F,192.168.11.11,infra-1
dhcp-host=net:data11,00:25:B5:02:20:4F,192.168.11.12,nn1-1
dhcp-host=net:data11,00:25:B5:02:20:3F,192.168.11.13,nn2-1
dhcp-host=net:data11,00:25:B5:02:21:0F,192.168.11.14,tr1-1
dhcp-host=net:data11,00:25:B5:02:20:EF,192.168.11.15,dn1-1
dhcp-host=net:data11,00:25:B5:02:20:AF,192.168.11.16,dn2-1
dhcp-host=net:data11,00:25:B5:02:20:6E,192.168.11.17,dn3-1
dhcp-host=net:data11,00:25:B5:02:20:5E,192.168.11.18,dn4-1
dhcp-host=net:data11,00:25:B5:02:20:0E,192.168.11.19,dn5-1
dhcp-host=net:data11,00:25:B5:02:20:FE,192.168.11.20,dn6-1
```

```
dhcp-host=net:data11,00:25:B5:02:20:BE,192.168.11.21,dn7-1
dhcp-host=net:data11,00:25:B5:02:20:8D,192.168.11.22,dn8-1
dhcp-host=net:data11,00:25:B5:02:20:7D,192.168.11.23,dn9-1
dhcp-host=net:data11,00:25:B5:02:20:2D,192.168.11.24,dn10-1
dhcp-host=net:data11,00:25:B5:02:20:1D,192.168.11.25,dn11-1
dhcp-host=net:data11,00:25:B5:02:20:DD,192.168.11.26,dn12-1
#
# 10GbE cluster members (eth2)
# name nodes and tracker nodes
#
dhcp-host=net:data12,00:25:B5:02:20:8F,192.168.12.11,infra-2
dhcp-host=net:data12,00:25:B5:02:20:7F,192.168.12.12,nn1-2
dhcp-host=net:data12,00:25:B5:02:20:2F,192.168.12.13,nn2-2
dhcp-host=net:data12,00:25:B5:02:20:1F,192.168.12.14,tr1-2
dhcp-host=net:data12,00:25:B5:02:20:DF,192.168.12.15,dn1-2
dhcp-host=net:data12,00:25:B5:02:20:CF,192.168.12.16,dn2-2
dhcp-host=net:data12,00:25:B5:02:20:9E,192.168.12.17,dn3-2
dhcp-host=net:data12,00:25:B5:02:20:4E,192.168.12.18,dn4-2
dhcp-host=net:data12,00:25:B5:02:20:3E,192.168.12.19,dn5-2
dhcp-host=net:data12,00:25:B5:02:21:0E,192.168.12.20,dn6-2
dhcp-host=net:data12,00:25:B5:02:20:EE,192.168.12.21,dn7-2
dhcp-host=net:data12,00:25:B5:02:20:AE,192.168.12.22,dn8-2
dhcp-host=net:data12,00:25:B5:02:20:6D,192.168.12.23,dn9-2
dhcp-host=net:data12,00:25:B5:02:20:5D,192.168.12.24,dn10-2
dhcp-host=net:data12,00:25:B5:02:20:0D,192.168.12.25,dn11-2
dhcp-host=net:data12,00:25:B5:02:20:FD,192.168.12.26,dn12-2

dhcp-vendorclass=set:csco_eth1,Linux
dhcp-vendorclass=set:csco_eth2,Linux

dhcp-option=26,9000

# Set the NTP time server addresses to 192.168.0.4 and 10.10.0.5
dhcp-option=option:ntp-server,10.29.160.53

dhcp-lease-max=150
dhcp-leasefile=/var/lib/misc/dnsmasq.leases
dhcp-authoritative
local-ttl=5
```

Once the **/etc/dnsmasq.conf** file has been configured, the **dnsmasq** service must be started via the commands:

```
chkconfig dnsmasq on
service dnsmasq restart
```

NTP

If needed, the Infrastructure server can act as a time server for all nodes in the cluster via one of the following methods:

- **ntp** service
- **cron** or at job to push the time to the rest of the nodes in the cluster

The configuration described in this document used the **ntp** service running on the infrastructure node to provide time services for the other nodes in the cluster.

The following was used for the contents of the **/etc/ntp.conf** file on the infrastructure node and may need to be customized to fit your implementation should you choose to use the **ntp** service:

```
# /etc/ntp.conf - infrastructure node NTP config
```

```
# For more information about this file, see the man pages ntp.conf(5),
# ntp_acc(5), ntp_auth(5), ntp_clock(5), ntp_misc(5), ntp_mon(5).
driftfile /var/lib/ntp/drift
# Permit time synchronization with our time source, but do not
# permit the source to query or modify the service on this system.
restrict default kod nomodify notrap nopeer noquery
restrict -6 default kod nomodify notrap nopeer noquery
# Permit all access over the loopback interface.
restrict 127.0.0.1
restrict -6 ::1
# Hosts on local network are less restricted.
#restrict 192.168.1.0 mask 255.255.255.0 nomodify notrap
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
#server 0.rhel.pool.ntp.org
#server 1.rhel.pool.ntp.org
#server 2.rhel.pool.ntp.org
#broadcast 192.168.1.255 autokey        # broadcast server
#broadcastclient                        # broadcast client
#broadcast 224.0.1.1 autokey            # multicast server
#multicastclient 224.0.1.1              # multicast client
#manycastserver 239.255.254.254         # manycast server
#manycastclient 239.255.254.254 autokey # manycast client
# Undisciplined Local Clock. This is a fake driver intended for backup
# and when no outside source of synchronized time is available.
server  127.127.1.0     # local clock
fudge   127.127.1.0 stratum 10
includefile /etc/ntp/crypto/pw
keys /etc/ntp/keys
```

The following was used for the contents of the **/etc/ntp.conf** file on the other nodes and may need to be customized to fit your implementation should you choose to use the **ntp** service:

```
# /etc/ntp.conf – all other nodes
server 10.29.160.53
driftfile /var/lib/ntp/drift
restrict default kod nomodify notrap nopeer noquery
restrict -6 default kod nomodify notrap nopeer noquery
restrict 127.0.0.1
restrict -6 ::1
includefile /etc/ntp/crypto/pw
keys /etc/ntp/keys
```

Once all of the /etc/ntp.conf files have been configured, the ntpd service must be started by executing the following commands on then infrastructure node and then all of the other nodes:

```
chkconfig ntpd on
service ntpd restart
pssh –h /root/pssh.hosts –A chkconfig ntpd on
pssh –h /root/pssh.hosts –A service ntpd restart
```

System Tunings

/etc/sysctl.conf

The following should be appended to the /etc/sysctl.conf file on all of the nodes:

```
# ----------
# /etc/sysctl.conf -- append to the file on all nodes
# BEGIN: Hadoop tweaks
#
```

```
sunrpc.tcp_slot_table_entries=128
net.core.rmem_default=262144
net.core.rmem_max=16777216
net.core.wmem_default=262144
net.core.wmem_max=16777216
net.ipv4.tcp_window_scaling=1
fs.file-max=6815744
fs.xfs.rotorstep=254
vm.dirty_background_ratio=1
#
# END: Hadoop tweaks
# ----------
```

This can be accomplished via the following commands:

```
cat > /tmp/sysctl.cnf << _EOD
# ----------
# /etc/sysctl.conf -- append to the file on all nodes
# BEGIN: Hadoop tweaks
#
sunrpc.tcp_slot_table_entries=128
net.core.rmem_default=262144
net.core.rmem_max=16777216
net.core.wmem_default=262144
net.core.wmem_max=16777216
net.ipv4.tcp_window_scaling=1
fs.file-max=6815744
fs.xfs.rotorstep=254
vm.dirty_background_ratio=1
#
# END: Hadoop tweaks
# ----------
_EOD

cat /tmp/sysctl.cnf >> /etc/sysctl.conf
sysctl –p

pscp –h /root/pssh.hosts –A /tmp/sysctl.cnf /tmp/sysctl.cnf
pssh –h /root/pssh.hosts –A cat /tmp/sysctl.cnf >> /etc/sysctl.conf
pssh –h /root/pssh.hosts –A sysctl –p
```

/etc/rc.d/rc.local

```
The following should be appended to the /etc/rc.d/rc.local file on all of the nodes:
# ----------
# /etc/rc.d/rc.local – append to the file on all nodes
# BEGIN: Hadoop tweaks
#
svcpgm="/sbin/service"
svcact=" stop "
svctyp=""
queue_depth=128
nr_requests=128
read_ahead_kb=3072
max_sectors_kb=1024
scheduler="deadline"
dirty_background_ratio=1
dirty_ratio=20
dirty_expire_centisecs=3000
devsd="/dev/sd"

while (( ${#devsd} ))
do
  devsd="${devsd}[[:alpha:]]"
```

```
  for i in ${devsd}
  do
    [[ "${i}" != "${i##*]]}" ]] && devsd="" && break
    if [[ -b ${i} && `/sbin/parted -s ${i} print|/bin/grep -c boot` -eq 0 ]]
    then
      /sbin/parted -s ${i} print | /bin/grep xfs
      [[ 1 == $? ]] && continue
      /sbin/blockdev --setra 1024 ${i}
      dev=`echo  ${i} |/bin/cut -d/ -f 3`
      echo ${queue_depth}    > /sys/block/${dev}/device/queue_depth
      echo ${nr_requests}    > /sys/block/${dev}/queue/nr_requests
      echo ${read_ahead_kb}  > /sys/block/${dev}/queue/read_ahead_kb
      echo ${max_sectors_kb} > /sys/block/${dev}/queue/max_sectors_kb
      echo ${scheduler}      > /sys/block/${dev}/queue/scheduler
    fi
  done
done

echo $dirty_background_ratio > /proc/sys/vm/dirty_background_ratio
echo $dirty_ratio > /proc/sys/vm/dirty_ratio
echo ${dirty_expire_centisecs} > /proc/sys/vm/dirty_expire_centisecs
echo never > /sys/kernel/mm/redhat_transparent_hugepage/defrag
echo never > /sys/kernel/mm/redhat_transparent_hugepage/enabled
echo 0 > /proc/sys/vm/nr_hugepages

# Stop some services that may be problematic.
for i in cpuspeed irqbalance
do
  ${svcpgm} ${i}${svctyp} ${svcact}
done
#
# END: Hadoop tweaks
# ----------
```

This can be accomplished by copying the above to the file /tmp/rc.tmp and then executing the following commands:

```
cat /tmp/rc.tmp >> /etc/rc.d/rc.local
/etc/rc.d/rc.local

pscp –h /root/pssh.hosts –A /tmp/rc.tmp /tmp/rc.tmp
pssh –h /root/pssh.hosts –A cat /tmp/rc.tmp >> /etc/rc.d/rc.local
pssh –h /root/pssh.hosts –A m/etc/rc.d/rc.local
```

Storage Configuration

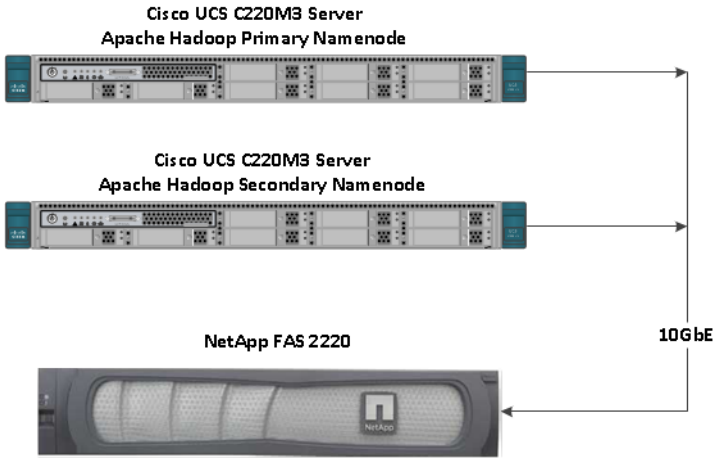
NetApp Storage Overview

The FlexPod Select for Hadoop leverages both the NetApp fabric-attached storage (FAS) and E-Series storage platforms to protect Hadoop Distributed File System (HDFS) metadata, and to provide HDFS data storage, respectively. The following subsections provide details of how both types of storage arrays are set up and provisioned, and how the provisioned storage is mapped to the servers in the Hadoop cluster.

FAS Overview

A NetApp FAS2220 storage system running Data ONTAP hosts a mirrored copy of the namenode metadata over a Network File System (NFS) mount, as shown in Figure 76. Notice that the secondary namenode is also connected to the FAS2220 to facilitate namenode recovery to the secondary namenode server, in the event of namenode failure.

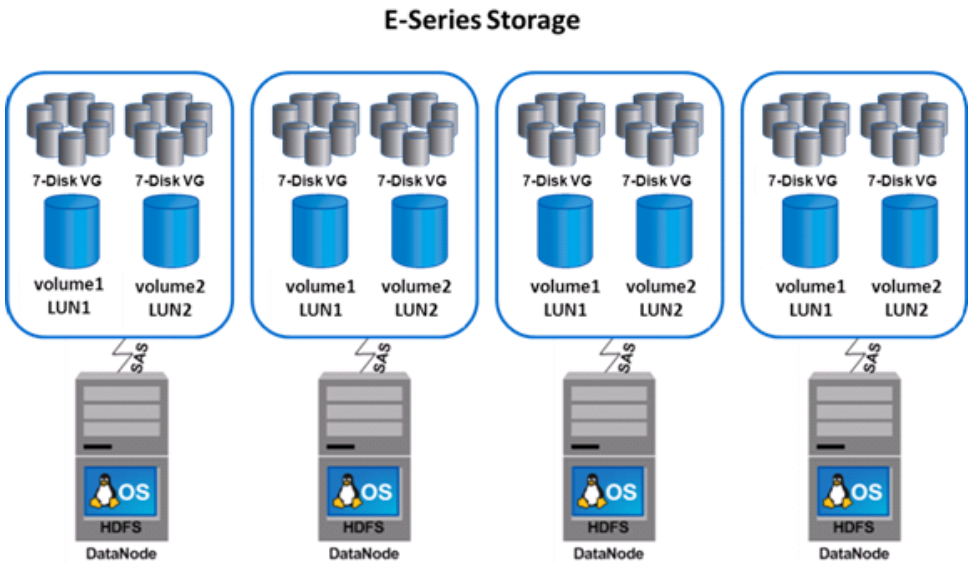
Figure 76 NFS Connectivity Between Namenodes and NetApp FAS 2220



E-Series Overview

NetApp recommends creating datanode servers on the E-Series array at a ratio of four nodes to one E5460 array, with each node having a single SAS path to one of the storage array controllers.

Figure 77 NetApp E5460 Array



FAS2220 and Data ONTAP 8

Initial setup of the FAS2220 is done with the Data ONTAP command line interface via the console serial port. Once the initial setup is done, further configuration and management of the FAS2220, the Data ONTAP 8 operating system and Data ONTAP features is done via the NetApp OnCommand System Manager management software.

FAS Initial Configuration

Table 6 lists the values for each parameter in the NetApp FAS2220 storage system configuration.

Table 6 NetApp FAS2220 storage system configuration template

Parameter Name	Sample Value
Hostname of the storage system	ntap-fas2220 (must resolve by DNS)
Hadoop Primary Namenode	nn1 (must resolve by DNS)
Hadoop secondary Namenode	nn2 (must resolve by DNS)
Aggregate size	Use existing aggr0 with the default RAID level of RAID-DP
Volumes	/vol/fsimage_bkp (100GB) Optional: /vol/common (100GB)
NFS share	/vol/fsimage_bkp (mounted on primary namenode and secondary namenode) Optional: /vol/common (mounted on all servers in the cluster)

Note The NetApp FAS2220 for Hadoop includes 6 disks in the 24-disk chassis. The factory default Raid DP aggr0 aggregate contains three of the six assigned disks. NetApp recommends adding the other three disks to the existing aggr0 to carve out data volumes instead of creating a new aggregate. The NetApp FAS2240-2 for Hadoop includes 12 disks in the 24-disk chassis. Out of these 12 disks, six are assigned to each controller. On each controller, the factory default Raid DP aggr0 contains three of the six assigned disks. Because of limited disk space in the FAS2220, NetApp recommends using the existing aggr0 to carve out data volumes instead of creating a new aggregate. This configuration is designed so that two additional disks will be added to aggr0, leaving one disk as a hot spare.

Data ONTAP 8.1.2 7-Mode

1. Complete Configuration Worksheet

   Before running the setup script, complete the configuration worksheet that is included in the Data ONTAP® 8.1 Software Setup Guide For 7-Mode, see:

   https://library.netapp.com/ecm/ecm_get_file/ECMP1119529

Note You must have access to the NetApp Support site to download the Software Setup Guide.

2. Run Setup Process

   Initial setup of the FAS2220 must be done via the serial console port using the Data ONTAP command line interface.

   When Data ONTAP is installed on a new storage system, the following files are not populated:

   - /etc/rc
   - /etc/exports
   - /etc/hosts
   - /etc/hosts.equiv

   To setup these files, follow these steps:

   a. Enter the configuration values the first time you power on the new system. The configuration values populate these files and configure the installed functionality of the system.

   b. Connect to the console port on the controller to access the CLI.

   c. Run the setup command at the storage system command prompt.

   d. Enter the following information:

```
Please enter the new hostname []: fas2220
Do you want to enable IPv6? [n]: RETURN
Do you want to configure interface groups? [n]: RETURN
Please enter the IP address for Network Interface e0a []:RETURN
```

   e. Press RETURN to accept the blank IP address

   f. Continue entering the following information:

```
Should interface e0a take over a partner IP address during failover? [n]: RETURN
Please enter the IP address for the Network Interface e0b []:Enter
Should interface e0b take over a partner IP address during failover? [n]: RETURN
Please enter the IP address for the Network Interface e0c []:Enter
Should interface e0c take over a partner IP address during failover? [n]: RETURN
Please enter the IP address for the Network Interface e0d []:Enter
Should interface e0d take over a partner IP address during failover? [n]: RETURN

Please enter the IP address for the Network Interface e1a []:Enter
Should interface e1a take over a partner IP address during failover? [n]: RETURN
Please enter the IP address for the Network Interface e1b []:Enter
Should interface e1b take over a partner IP address during failover? [n]: RETURN

Please enter the IP address for Network Interface e0M []: 10.29.160.43
Please enter the netmaskfor the Network Interface e0M [255.255.255.0]:
255.255.255.0

Should interface e0M take over a partner IP address during failover? [n]: y
Please enter the IPv4 address or interface name to be taken over by e0M []: e0M
Please enter flow control for e0M {none, receive, send, full} [full]: RETURN

Would you like to continue setup through the Web interface? [n]: RETURN

Please enter the name or IP address of the IPv4 default gateway: 10.29.160.1

The administration host is given root access to the storage system's / etc files
for system administration. To allow /etc root access to all NFS clients enter
RETURN below.
Please enter the name or IP address for administrative host: RETURN

Please enter timezone [GTM]: US/Pacific
```

Note Example time zone: America/New_York.

```
Where is the filer located? Hadoop Lab
Enter the root directory for HTTP files [home/http]: RETURN
Do you want to run DNS resolver? [n]: y
Please enter DNS domain name []: hadoop.local
Please enter the IP address for first nameserver []: 10.29.160.53
Do you want another nameserver? [n]: RETURN
```

Note Optionally, enter up to three name server IP addresses.

```
Do you want to run NIS client? [n]: RETURN
Press the Return key to continue through AutoSupport message
would you like to configure SP LAN interface [y]: RETURN
Would you like to enable DHCP on the SP LAN interface [y]: n
Please enter the IP address for the SP: RETURN
Please enter the netmask for the SP []:RETURN
Please enter the IP address for the SP gateway: RETURN
Please enter the name or IP address of the mail host [mailhost]: <<var_mailhost>>
Please enter the IP address for <<var_mailhost>> []:<<var_mailhost>>
New password: change_me
Retype new password: change_me
```

g. Enter the password for admin to login to the controller.

h. Enter reboot to activate the settings.

i. After the FAS unit reboots, you should be able to use OnCommand System Manager to connect to the FAS e0M interface that you configured above.

To configure the FAS2220 to serve NFS data by creating two NFS shares, follow these steps:

1. Log in to the FAS2220 as root, using putty or a similar SSH utility.
2. Add two additional disks to aggr0 to make it a total of five disks (three data disks and two parity disks) by running the aggr add command.

```
aggr add aggr0 2
```

3. Each controller contains a root volume named vol0. This root volume contains the file-system and files needed for the running Data ONTAP. From the factory, this volume is initially sized much larger than is required for array operation in the FlexPod Select for Hadoop environment. This volume should be resized, thereby releasing valuable storage space for use by other volumes being created on aggregate aggr0. By issuing the following command on each controller, the vol0 root volume will be down-sized to be smaller, but adequately sized for operations.

   ```
   vol size vol0 200g
   ```

4. Create two 100GB volumes.

```
vol create fsimage_bkp aggr0 100g
vol create common aggr0 100g
```

5. Share the directories through NFS.

```
exportfs -p
sec=sys,rw=192.168.11.0/24:192.168.12.0/24,root=192.168.11.0/24:192.168.12.0/24
/vol/fsimage_bkp

exportfs -p
sec=sys,rw=192.168.11.0/24:192.168.12.0/24,root=192.168.11.0/24:192.168.12.0/24
/vol/common
```

NetApp OnCommand System Manager 2.1

OnCommand® System Manager is the simple yet powerful browser-based management tool that enables administrators to easily configure and manage individual NetApp storage systems or clusters of systems.

System Manager is designed with wizards and workflows, simplifying common storage tasks such as creating volumes, LUNS, qtrees, shares, and exports, which saves time and prevents errors. System Manager works across all NetApp storage: FAS2000, 3000, and 6000 series as well as V-Series systems.

The following are NetApp OnCommand System Manager 2.1 prerequisites:

- If a proxy server is being used, it must be disabled.
- Sun JRE 6 must be installed.
- If running Windows or Linux guest OS on Mac OS using VMware Fusion:
  - The shared-folders feature must be disabled.
  - The desktop option of the mirrored-folders feature must be disabled.

Launch OnCommand System Manager

Double-click the System Manager icon on your desktop to launch System Manager. The NetApp OnCommand System Manager icon is shown in Figure 78.

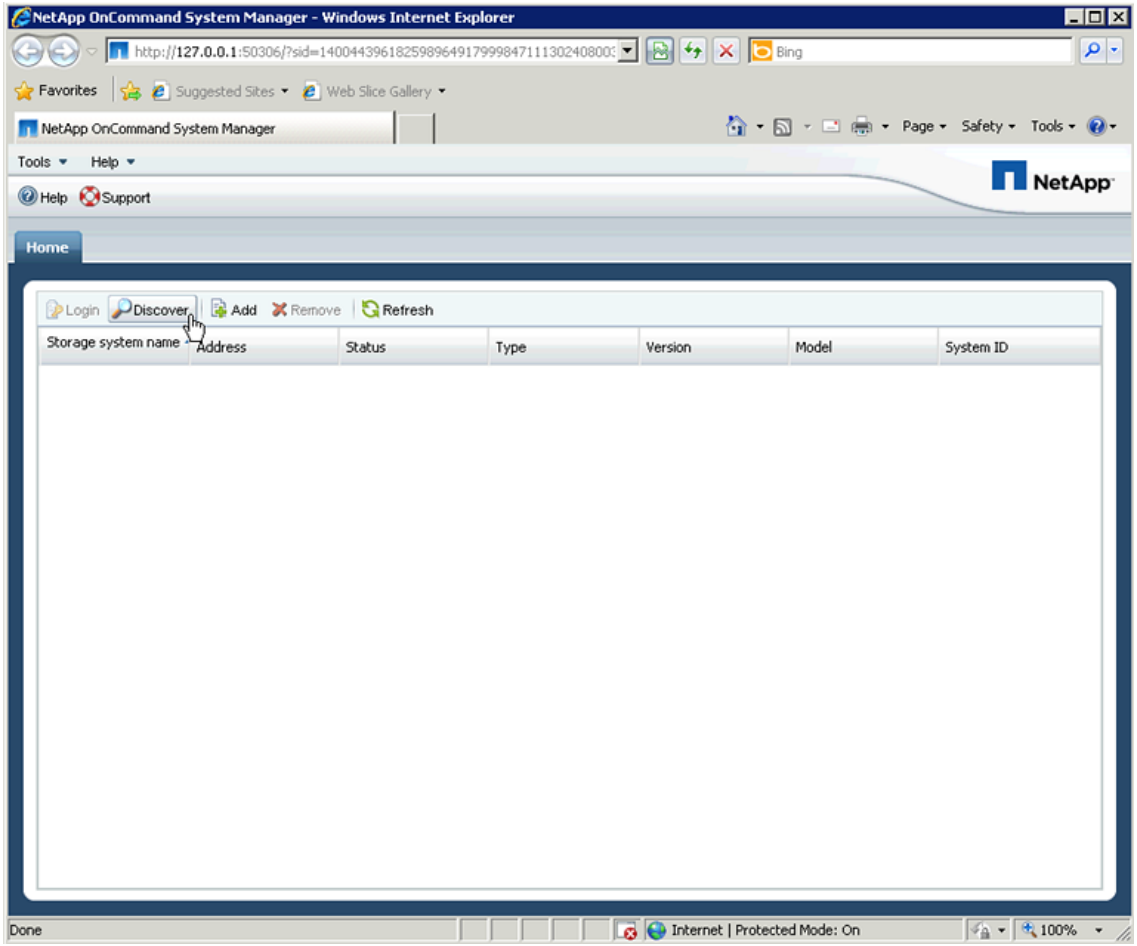
Figure 78 NetApp OnCommand System Manager Icon

Discover Storage Systems

To add storage system or HA pair, follow these steps:

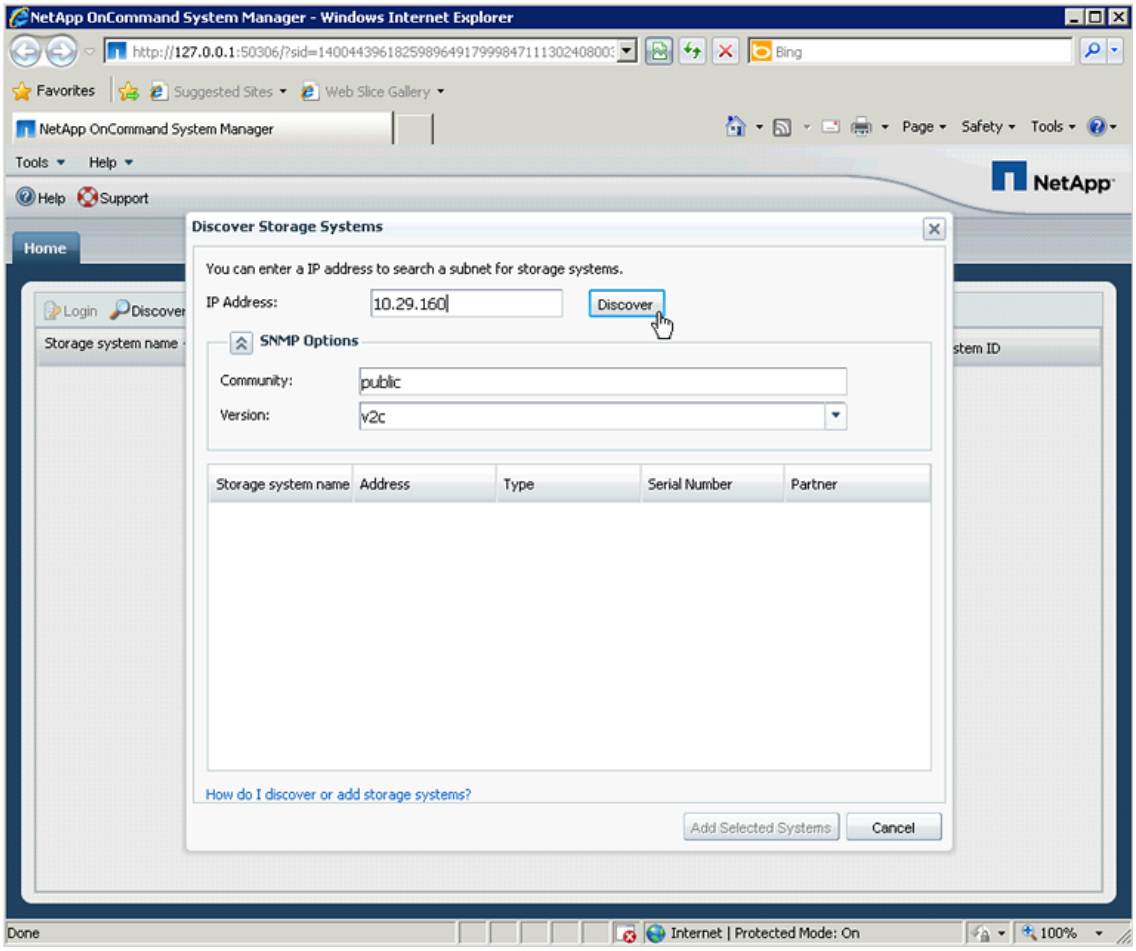
1. In the NetApp System Manager, select Home tab and click **Discover**.

Figure 79 Discovering Storage Systems



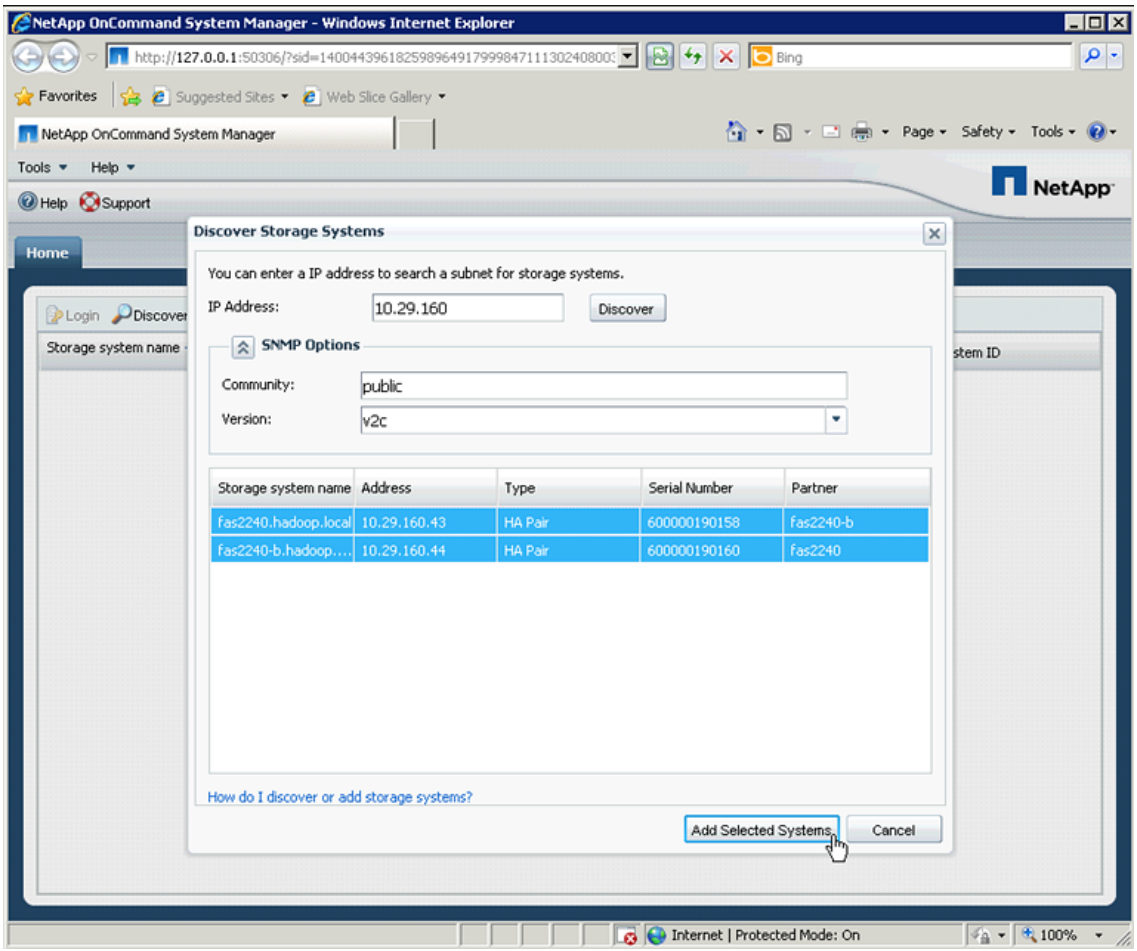
2. In the Discover Storage Systems dialog box, enter the subnet IP address and click **Discover**.

Figure 80 Entering the IP Address for Discovering Storage Systems



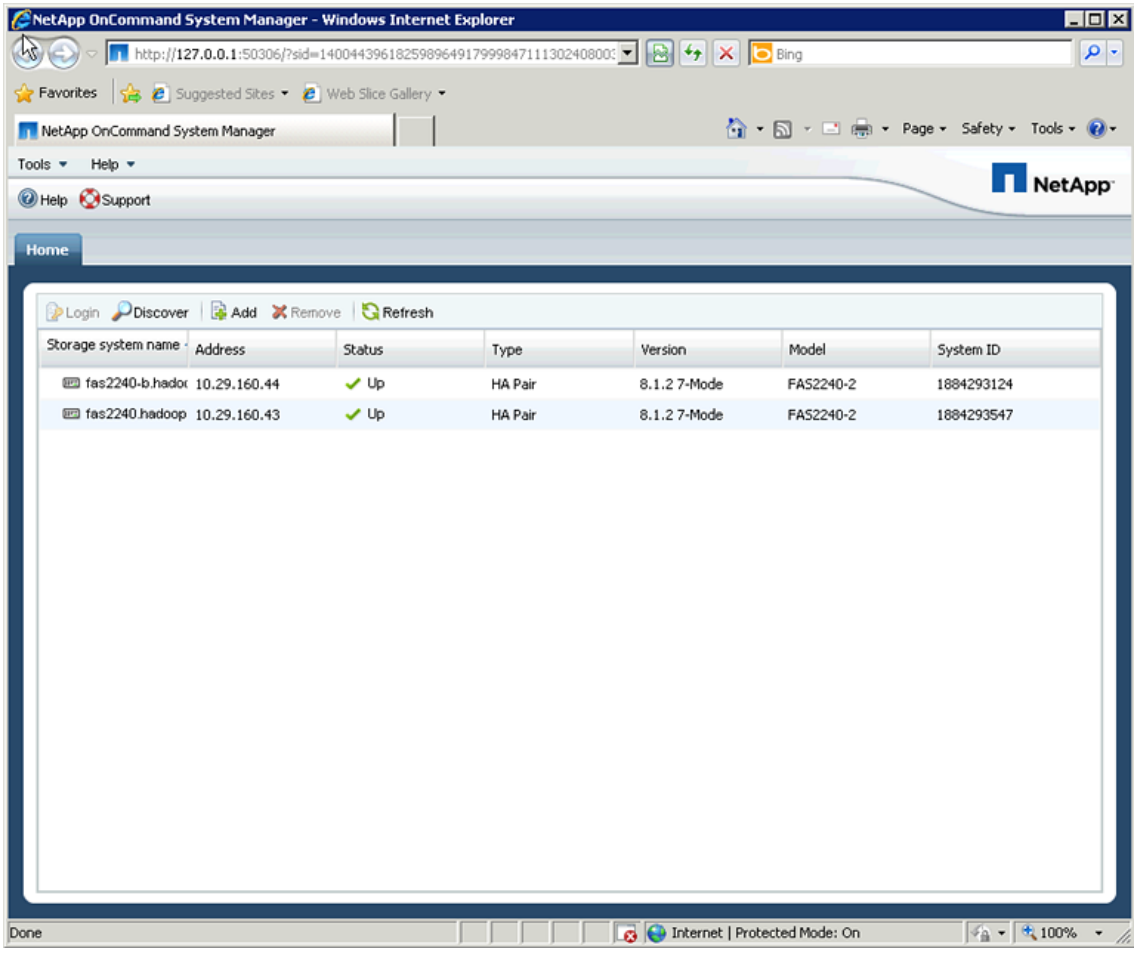
3. Select one or more storage systems from the list of discovered systems and click **Add Selected Systems**.

Figure 81 Adding Selected Storage Systems



4. Verify that the storage system or the HA pair that you added is included in the list of managed systems.

Figure 82 Verifying the Added Storage Systems

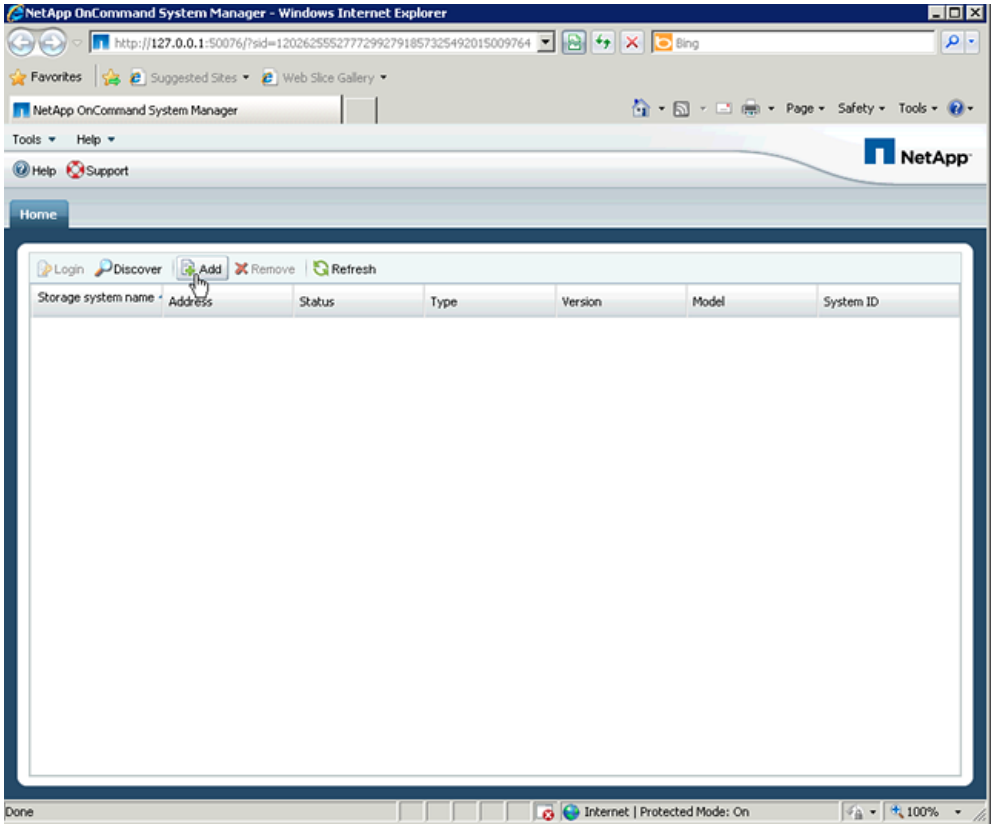


Adding Storage Systems

If you need to add a FAS unit to an existing System Manager setup, follow these steps within System Manager:

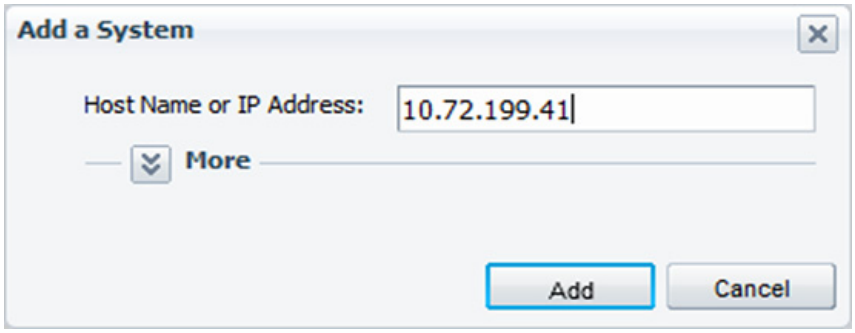
1. In the Home tab, click **Add**.

Figure 83 Adding FAS Storage Unit



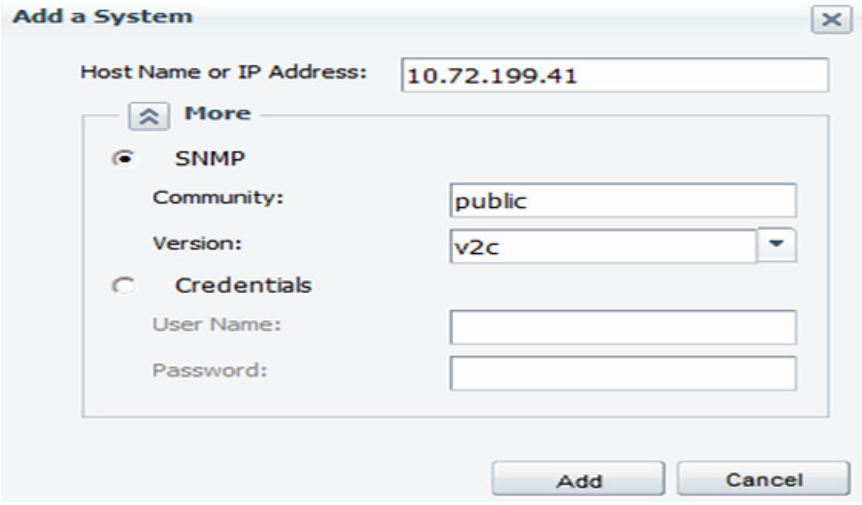
2. Type the fully qualified DNS host name or the IPv4 address of the storage system.

Figure 84 Entering Host IP Address



3. Click More.
4. Select the **SNMP** radio button method for discovering and adding the storage system.
5. You need to specify the SNMP community and the SNMP version.

Figure 85 Specifying SNMP Details



6. Enter user name and password.
7. Click **Add**.

E-Series Configuration & Management

The configuration and management of the E-Series E5460 storage array is done via the NetApp SANtricity management software.

Record Storage Configurations

Use the template in Table 7 to capture and keep a record of all volume groups, volumes, serving controllers, and Hadoop datanode hosts. The entries in the Table 7 are intended to serve as an example of a useful naming convention; however, individual customer requirements vary widely with respect to naming conventions. As a result, the specific names for individual projects should be substituted for those in the example template.

Note The recommended naming best practice is to associate the volume group name with the storage system and controller ID hosting the volume group. Similarly, the volume name should clearly reflect which datanode it serves.

Table 7 Storage provisioning details template

Volume	Volume Group	E-Series Storage System	Controller Slot	Datanode
vol1_datanode1	Vg1a_ntap01-A	Ntap01	A	Datanode1
vol2_datanode1	Vg1b_ntap01-A			
vol1_datanode2	Vg1a_ntap01-B	Ntap01	A	Datanode2
vol2_datanode2	Vg1b_ntap01-B			

Table 7 Storage provisioning details template

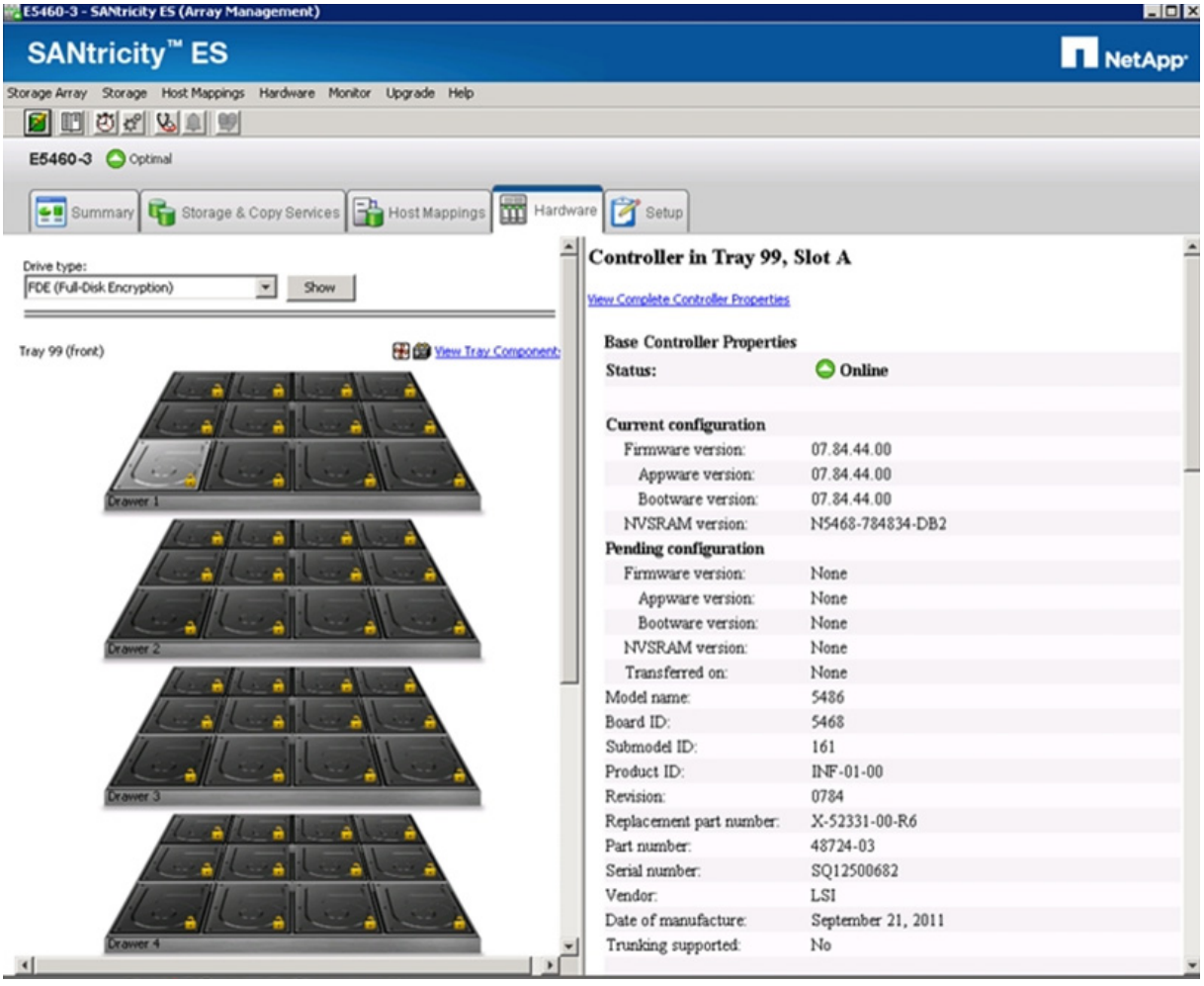
Volume	Volume Group	E-Series Storage System	Controller Slot	Datanode
vol1_datanode3	Vg2a_ntap01-A	Ntap01	B	Datanode3
vol2_datanode3	Vg2b_ntap01-A			
vol1_datanode4	Vg2a_ntap01-B	Ntap01	B	Datanode4
vol2_datanode4	Vg2b_ntap01-B			

Confirm That All Disks Are in Optimal Status

To confirm the health status of all disks, follow these steps:

1. Select the Hardware tab of SANtricity ES Storage Manager to view an inventory of all the hard disks in the array.

Figure 86 Hard Disk Details in SANtricity ES Manager



2. Check the Status Online, for the health of all drives and look for any errors. If there are no errors, proceed to the next step. If fault conditions are present, correct the faults before proceeding.

Selecting Drives for Volume Groups and Hot Spares

For highest performance, NetApp recommends balancing the use of even and odd disk slots for each controller, which provides balanced I/O across drive side channels. Table 8 defines the correct mapping of disk drives to volume groups.

Table 8 Mapping disk drives to volume groups

Host Name	Volume Group/Volume	Disk List Strings	Datanode
Host1	VG1/VOL1	[99,1,1],[99,2,1],[99,3,1],[99,4,1],[99,1,2], [99,2,2],[99,3,2]	A
Host1	VG2/VOL2	[99,5,1],[99,1,3],[99,2,3],[99,4,2],[99,5,2], [99,1,4],[99,2,4]	B
Host2	VG1/VOL1	[99,3,3],[99,4,3],[99,5,3],[99,1,5],[99,3,4], [99,4,4],[99,5,4]	A
Host2	VG2/VOL2	[99,2,5],[99,3,5],[99,4,5],[99,1,6],[99,2,6], [99,3,6],[99,4,6]	B
Host3	VG1/VOL1	[99,5,5],[99,1,7],[99,2,7],[99,3,7],[99,5,6], [99,1,8],[99,2,8]	A
Host3	VG2/VOL2	[99,4,7],[99,5,7],[99,1,9],[99,3,8],[99,4,8], [99,5,8],[99,1,10]	B
Host4	VG1/VOL1	[99,2,9],[99,3,9],[99,4,9],[99,5,9],[99,2,10], [99,3,10],[99,4,10]	A
Host4	VG2/VOL2	[99,1,11],[99,2,11],[99,4,11],[99,5,10],[99,1,12], [99,3,12], [99,5,12]	B

Note The remaining four drives ([99,2,12], [99,3,11], [99,4,12] and [99,5,11]) should be designated as hot spares.

Creating and Assigning Hot Spares

To allocate hot spares in the E-Series array, follow these steps:

1. Start the SANtricity ES Storage Manager client.
2. Select the Devices tab.
3. In the left pane (tree view), double-click the array where the volume group will be created. This step opens the Array Management Window (AMW) for that array.
4. In the AMW, select the Hardware tab and follow these steps:

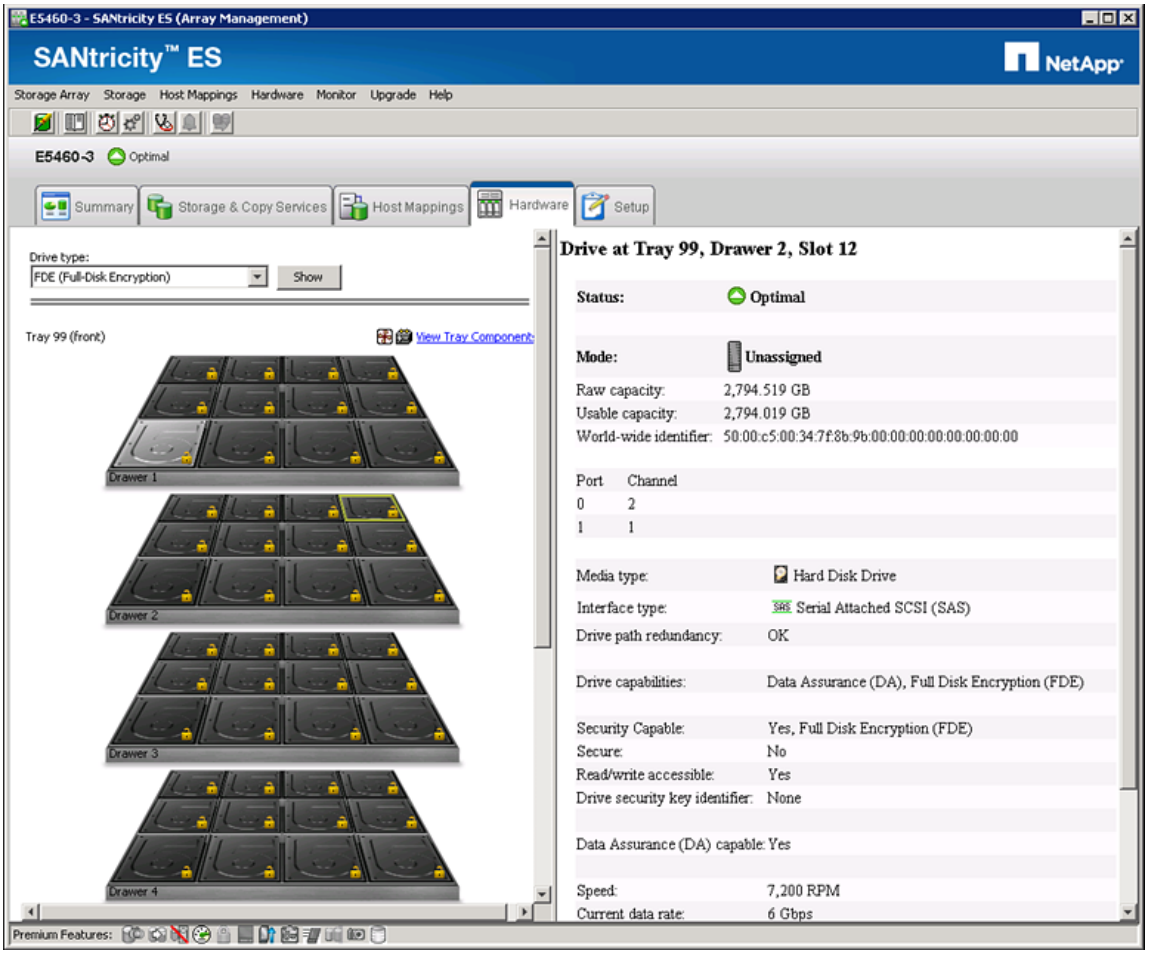
a. Select hot spares beginning with slot 12 in drawer 2.

b. Choose slot 11 in drawer 3 for the second spare disk.

c. Choose slot 12 in drawer 4 for the third spare disk.

d. Choose slot 11 in drawer 5 for the fourth spare disk.

Figure 87 Selecting a Drive as Spare



Note The drive selected in this step is allocated as a hot spare. The allocated hot spare drive is indicated by a red plus sign over the drive icon.

e. Right-click on the image of the chosen drive, and then select **Hot Spare Coverage**.

f. In the Hot Spare Drive Options dialog box, select the **Manually Assign Individual Drives** radio button and click **OK**.

Figure 88 Assigning Drives Manually

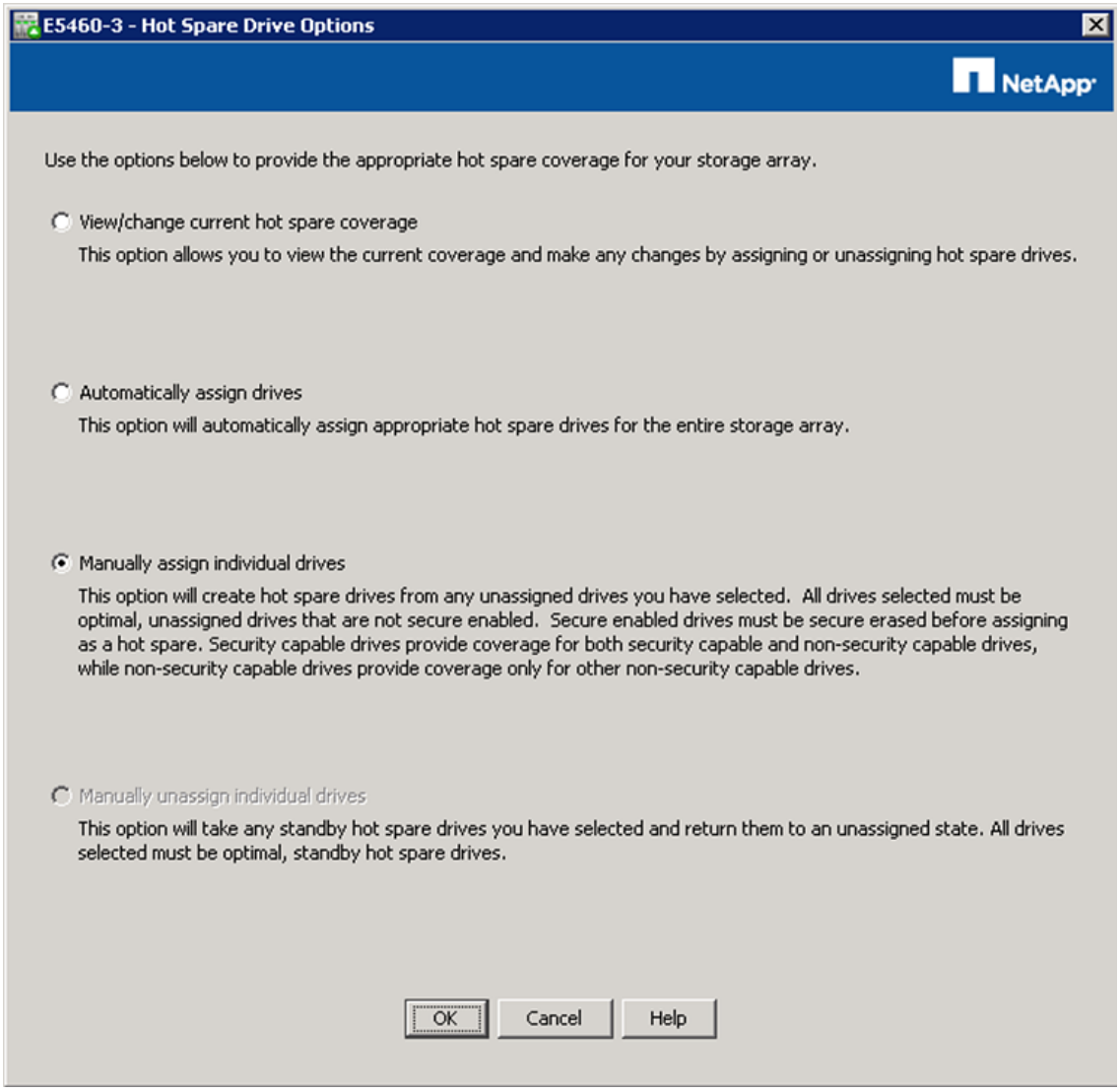
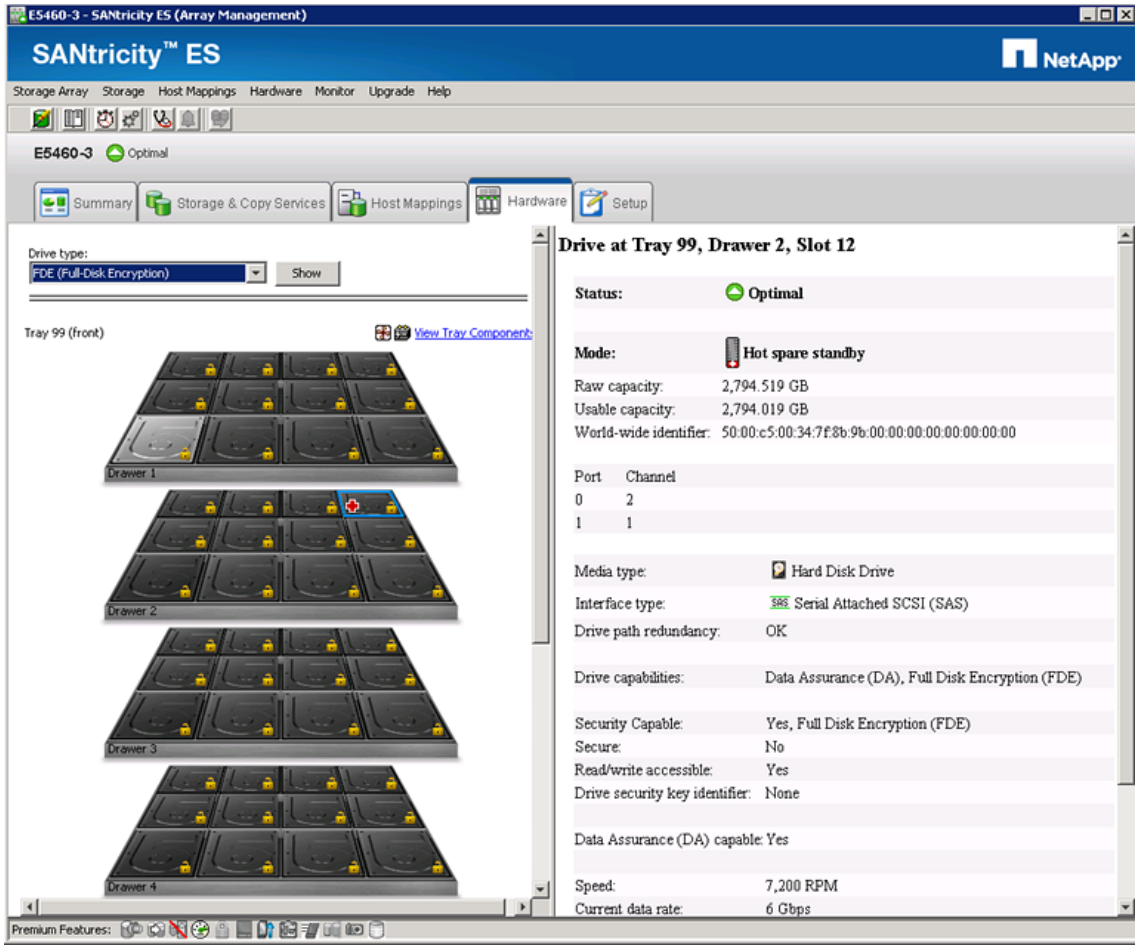


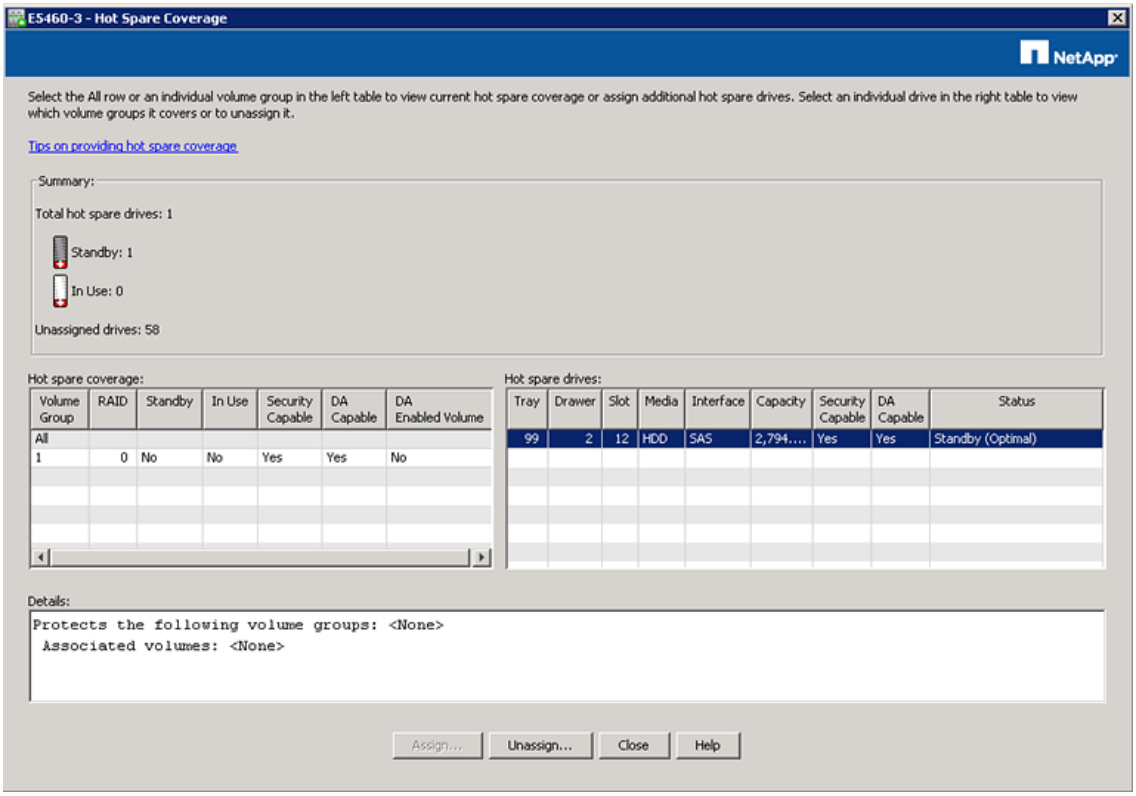
Figure 89 shows assigned Hot spare in standby mode in the SANtricity Manager window.

Figure 89 Hot Spare in Standby Mode



5. To change the drives that are allocated as hot spares, follow these steps:
   a. In the Hardware tab of the AMW, select the drive to be changed.
   b. From the main menu, select **Hardware > Hot Spare Coverage**.
   c. In the Hot Spare Drive Options dialog box, select View/Change Current Hot Spare Coverage. Click **OK**.

Figure 90 Summary Showing Total Hot Spare Drives



Creating Volume Groups

A volume group is a set of disk drives that are logically grouped together to provide storage with a single RAID level for all volumes in the group. Every E-Series storage array has eight RAID 5 volume groups of 6+1 disks (6 data disks and 1 parity disk). This leaves all 60 disks assigned in each shelf given the four hot spare drives previously prescribed.

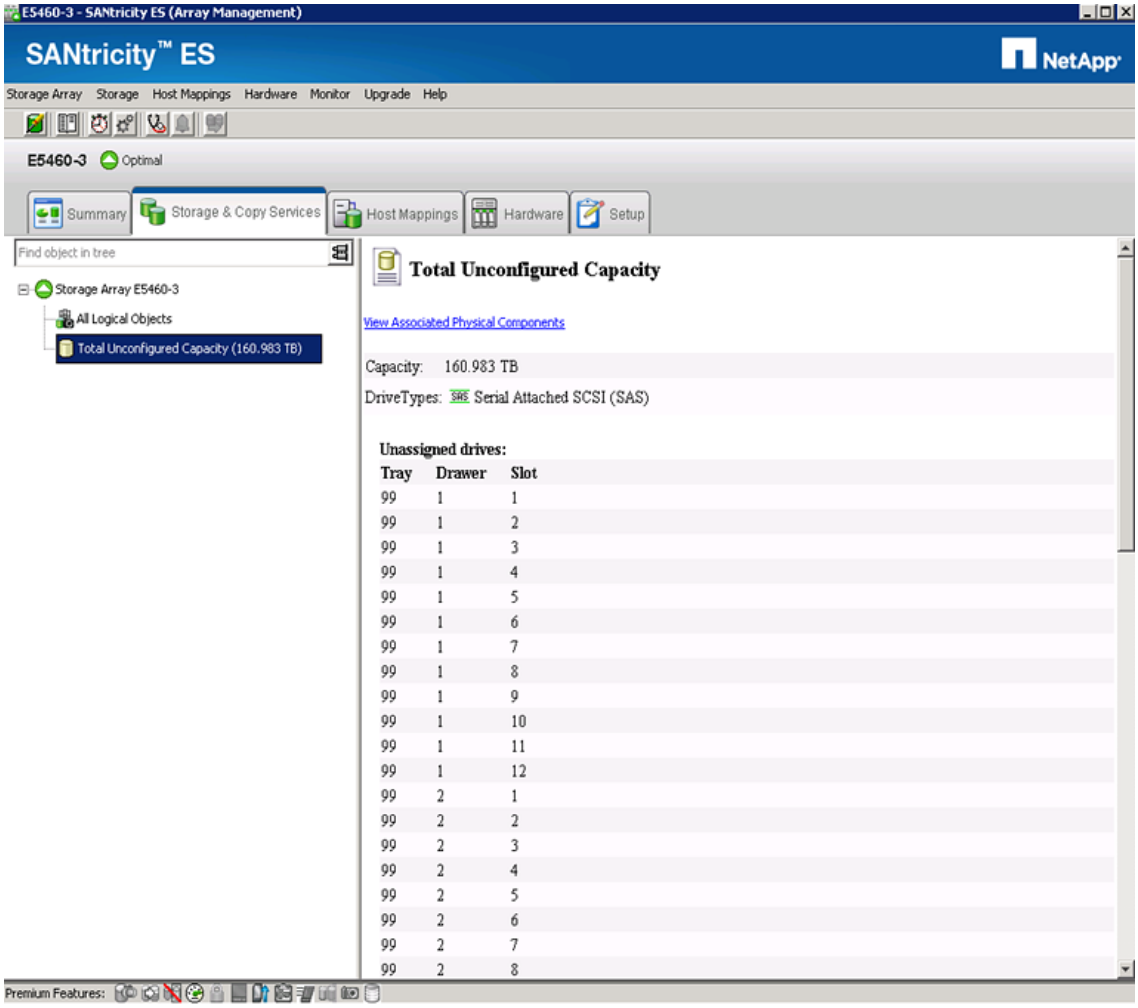
To create the volume groups, select the disks from across the five drawers starting with drawer one, slot one and alternating between odd and even slot numbers as the disks are selected in a round robin fashion. For more details on Hot spares, see "Selecting Drives for Volume Groups and Hot Spares" section on page 110.

Creating New Volume Groups

To create a volume group, follow these steps:

1. Log into SANtricity ES or Array Management window.
   a. In the Array Management window, select the Storage and Copy Services tab.
   b. Select Total Unconfigured Capacity in the volume group tree view.

Figure 91 Array Management Window Showing Unassigned Drive Details

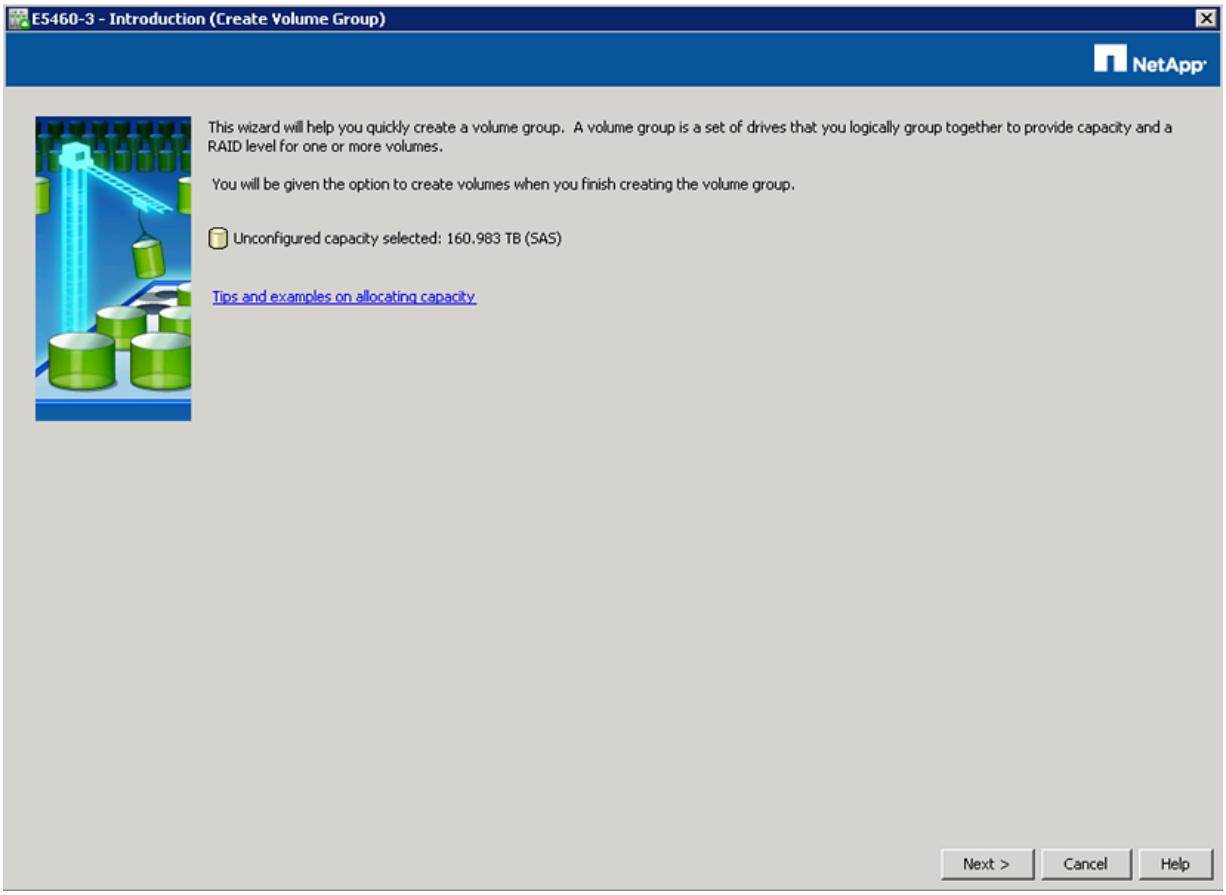


c. Right-click on Total Unconfigured Capacity, select Volume Group.

d. Click **Create,** to launch the Create Volume Group wizard.

2. In the Create Volume Group wizard, click **Next**.

Figure 92 Creating Volume Group Wizard

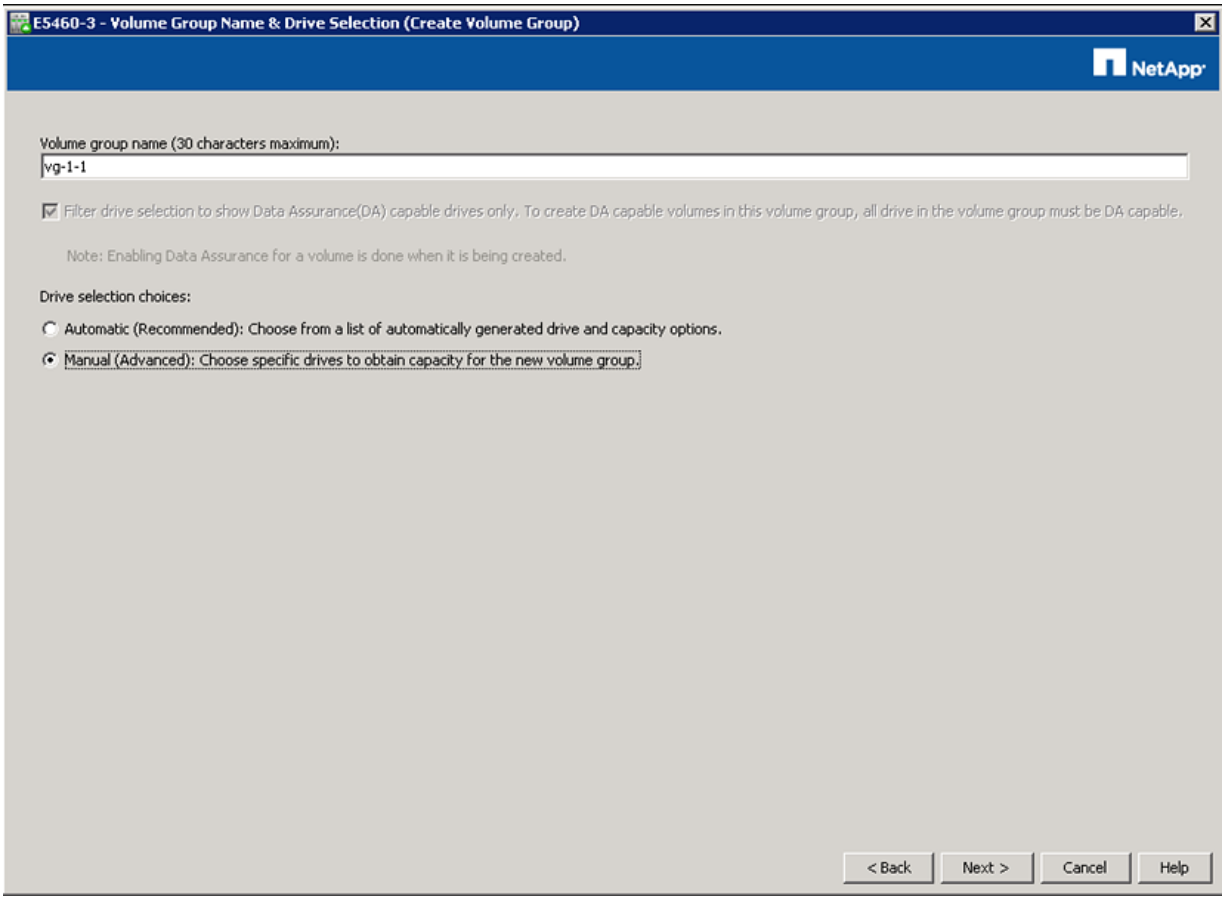


3. Specify a name and select drives for the volume group:
   a. Enter the volume group name.

Note The volume group name cannot contain spaces. Use underscores or dashes to separate the elements of the volume group name (for example, Test_Group_1 or Test-Group-1).

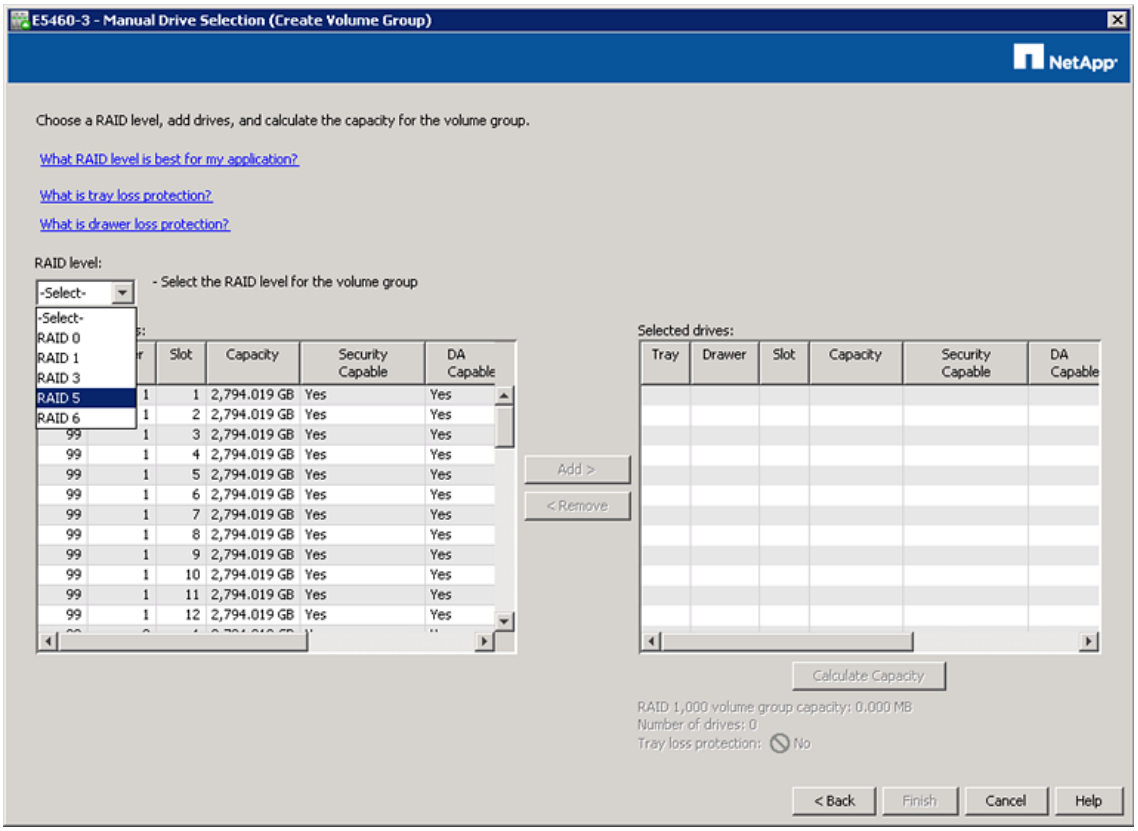
   b. Select the **Manual (Advanced)** radio button. Click **Next**.

Figure 93 Manual Drive Selection



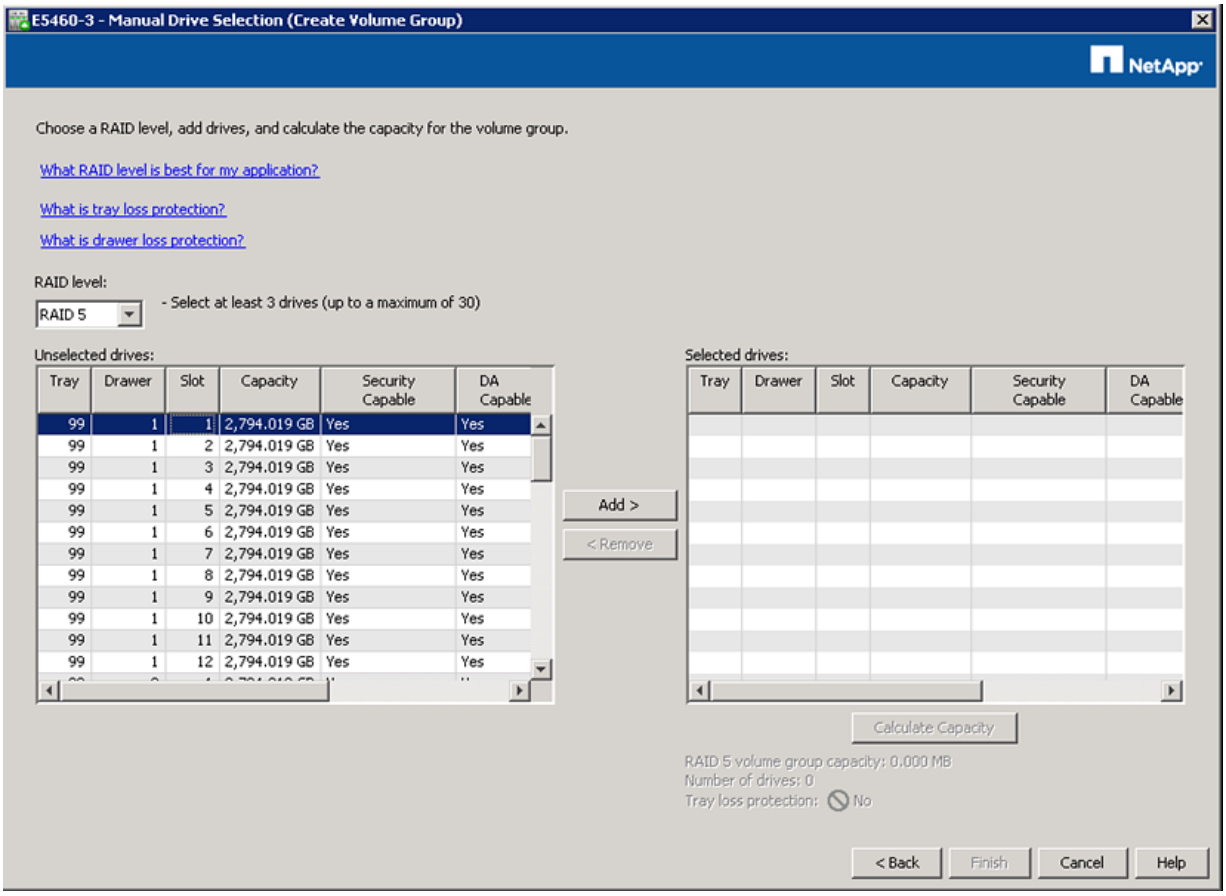
c. Select a RAID level for the volume group (in this example, RAID5 is selected).

Figure 94 Selecting RAID Levels for Volume Group



d. Using vertical striping, select the first available disk in the first drawer. Click Add > to add desired disks to the volume group.

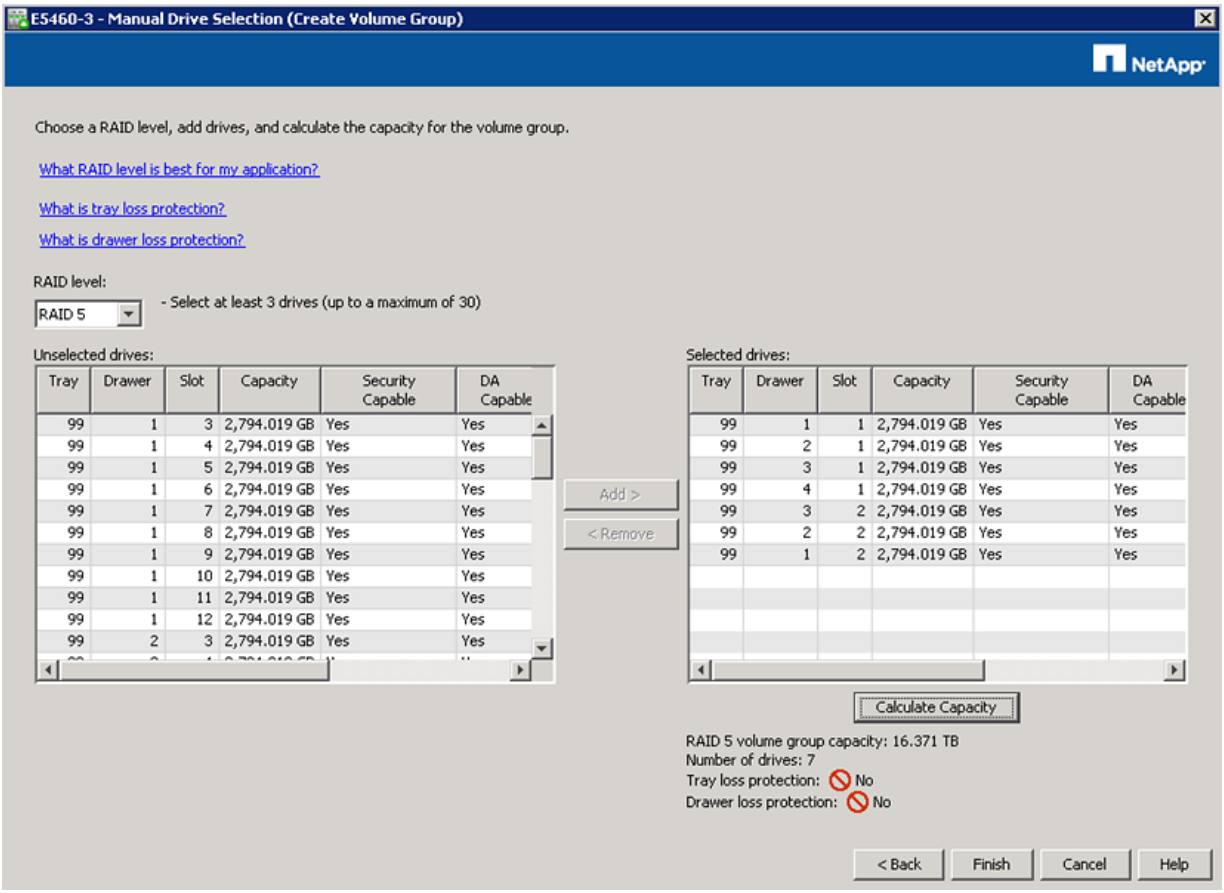
Figure 95 Adding Drives for Volume Group



e. Continue selecting disks until seven disks are selected. Disks should be selected across drawers in a vertical fashion. Avoid selecting more than two disks in the same drawer for a single volume group. Since seven disks are needed per volume group and there are only five drawers, three disks will be chosen from individual drawers, with four disks split evenly across the remaining two drawers (two per drawer).

f. After seven disks are selected, click **Calculate Capacity** to confirm that the capacity of the new volume group satisfies all the requirements for the group and that the desired RAID protection is achieved.

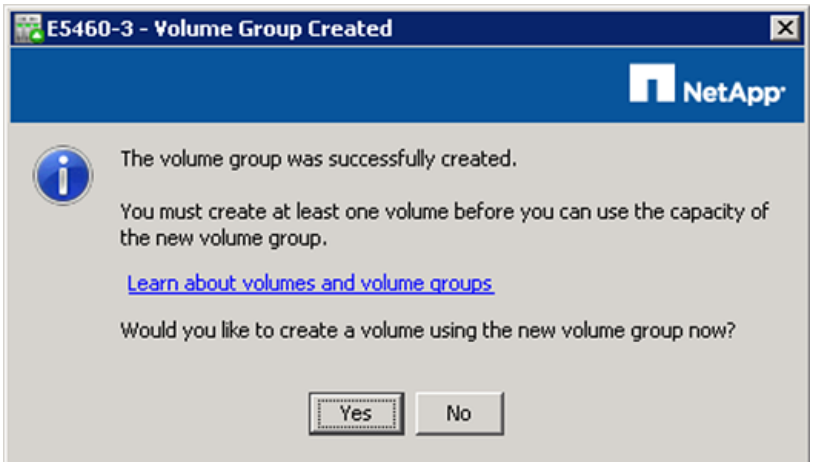
Figure 96 Calculating the Volume Group Capacity



g. Click **Finish** to create the new volume group. A message is displayed, confirming that the volume group was successfully created and providing the option to create a new volume.

h. Click **Yes** to create a volume in the new volume group. This will launch the Create Volume wizard.

Figure 97 Selected Drives Added Successfully to the Volume Group

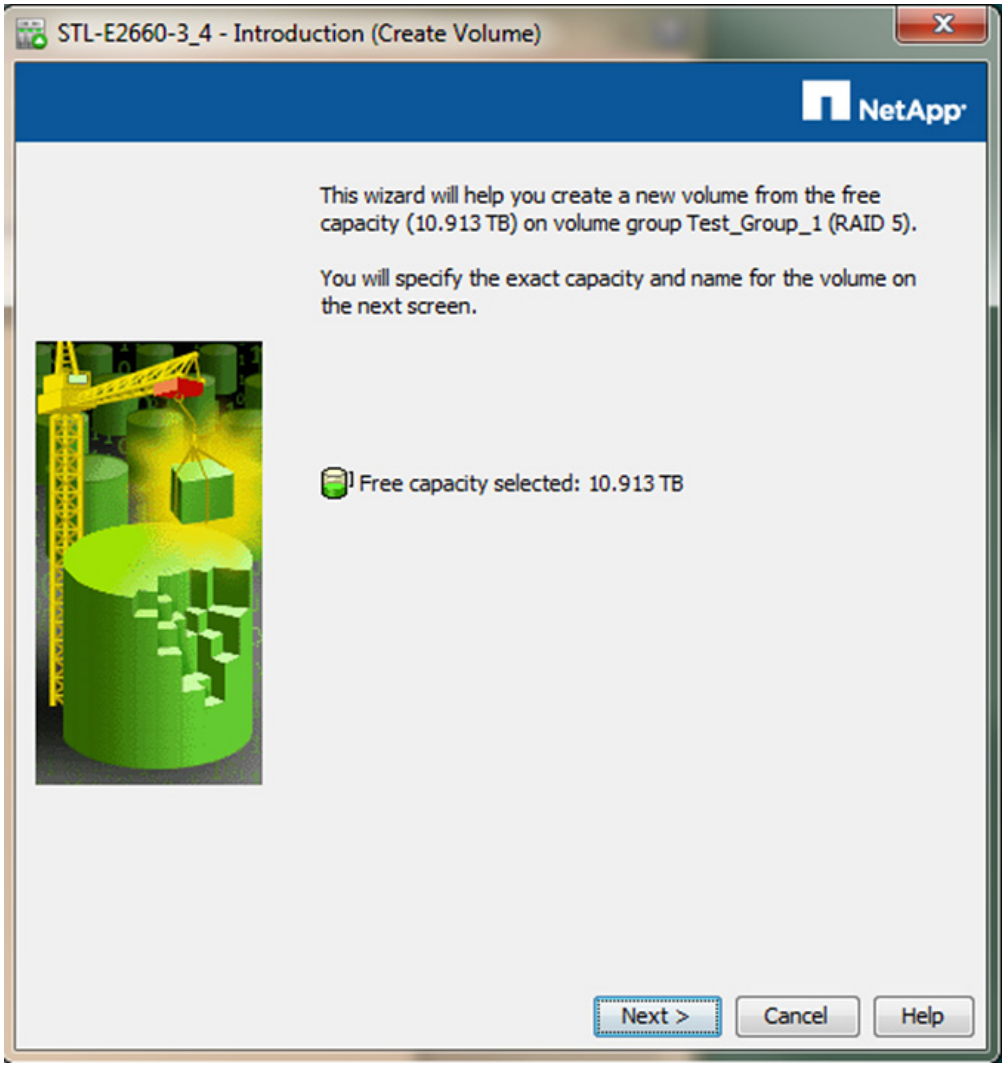


Create New Volume

To create new volume, follow these steps:

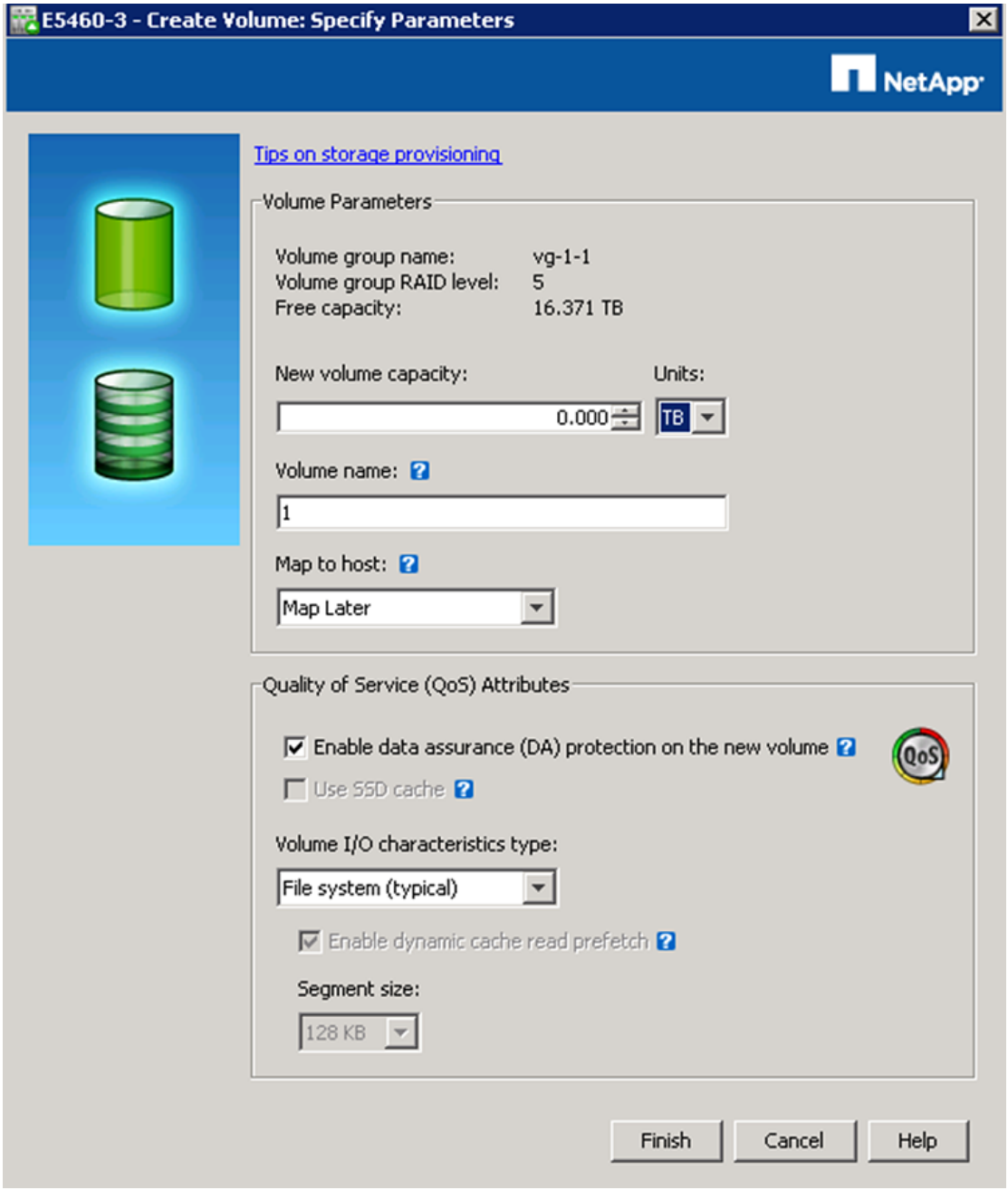
1. In the Create Volume wizard, click **Next** to continue.

Figure 98 Creating Volume



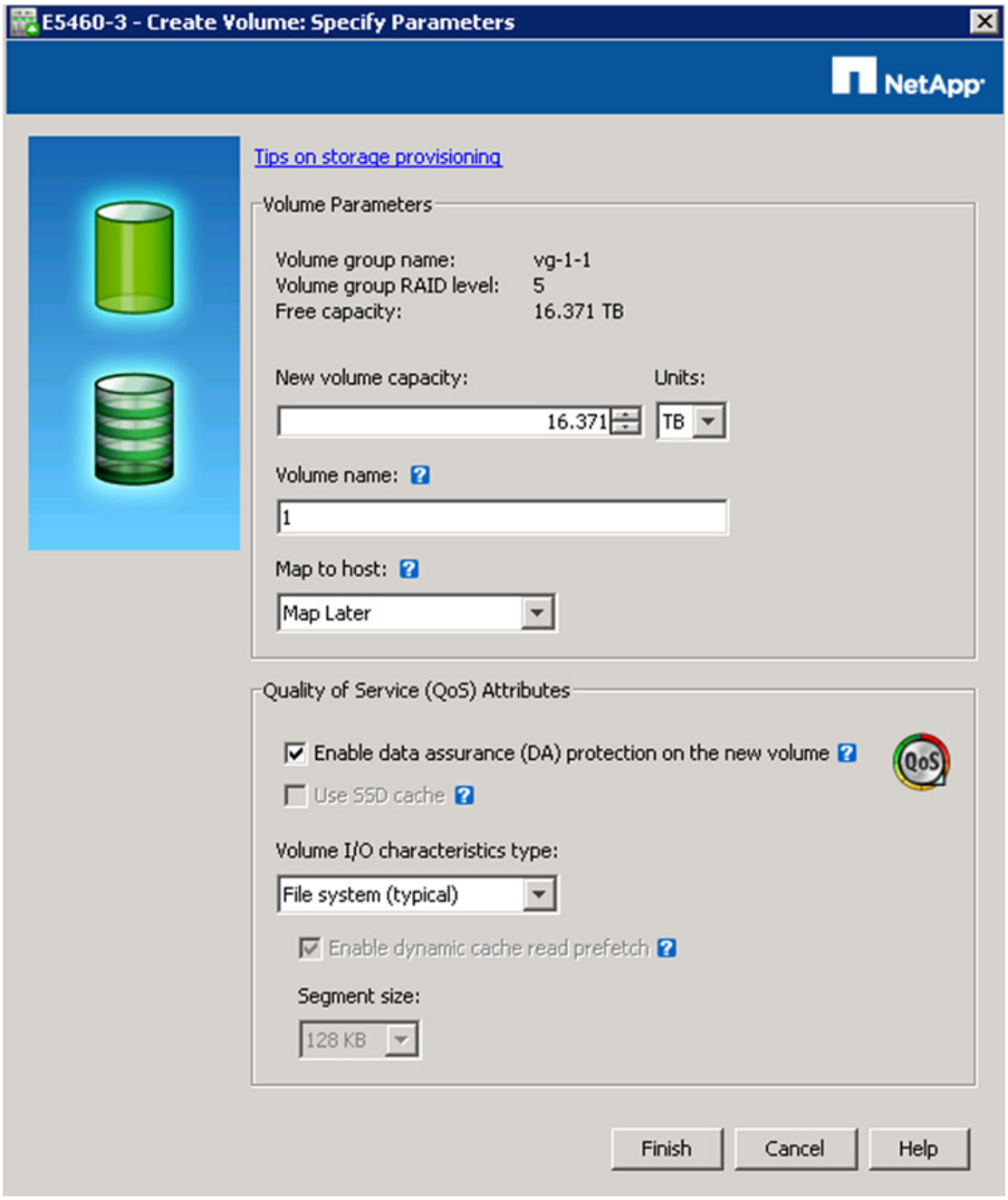
2. Specify the volume parameters.
   a. Select the appropriate unit for the new volume capacity. For the volume group free capacity shown in this example, the correct unit is TB.

Figure 99 Entering Volume Parameters



b. Enter the new volume capacity to be taken from the free capacity of the selected volume group. Because the entire volume group is used for this volume, enter the maximum size available for the available capacity (free capacity). In this example, the free capacity of 16.371 is entered for the new volume capacity.

Figure 100 Specifying New Volume Capacity



c. Enter the volume name.

d. Keep the Map to host field at default.

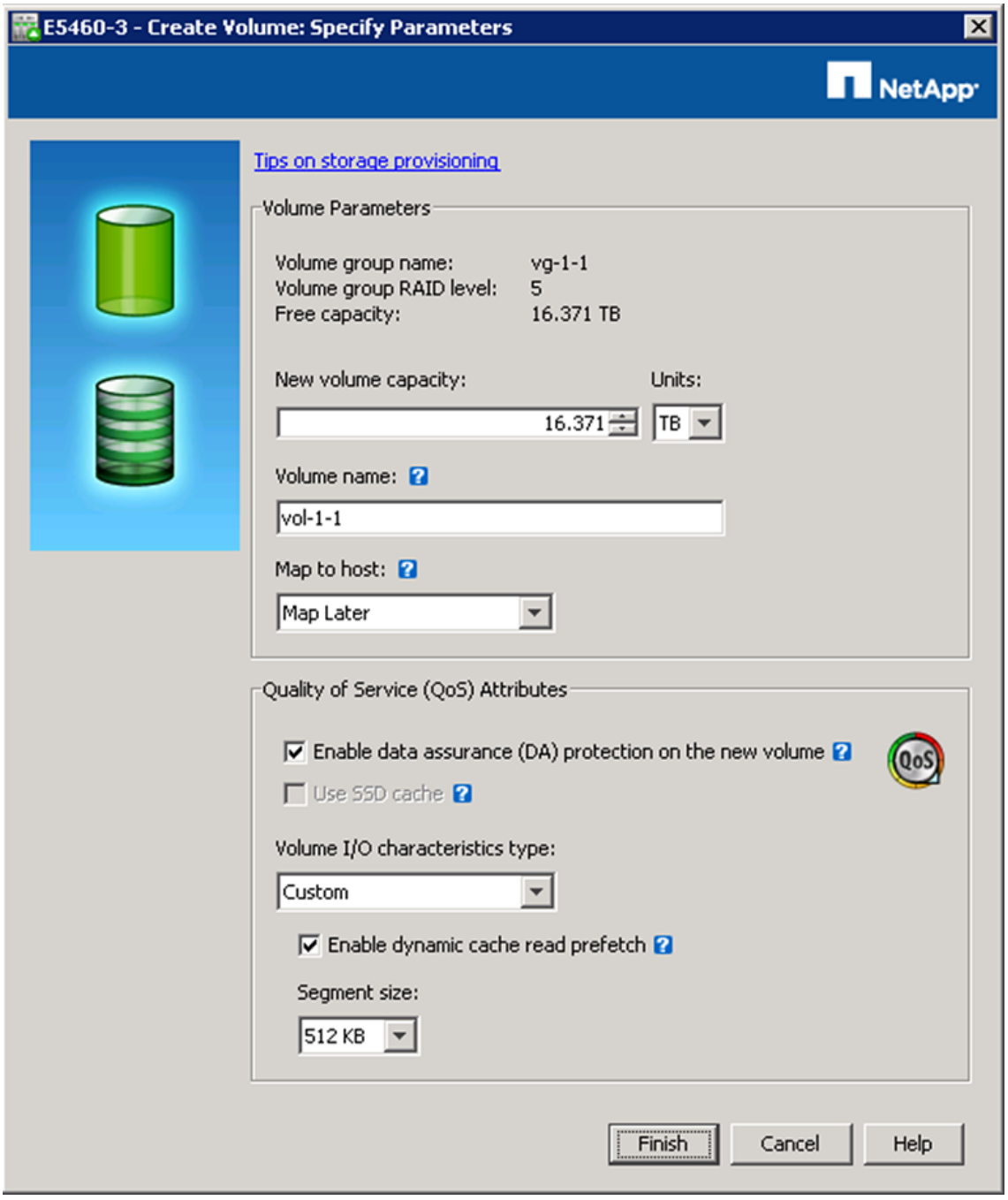
Note The volume name cannot contain spaces. Use underscores or dashes to separate elements within the volume name (for example, vol-1-1).

3. Configure the following quality of service attributes for the volume:

a. Select the volume I/O characteristics type from the drop down list. Select the type as Custom.

b. Check the Enable Dynamic Cache Read Prefetch check box and then select the 512KB segment size.

Figure 101 Setting QoS Attributes for Volume

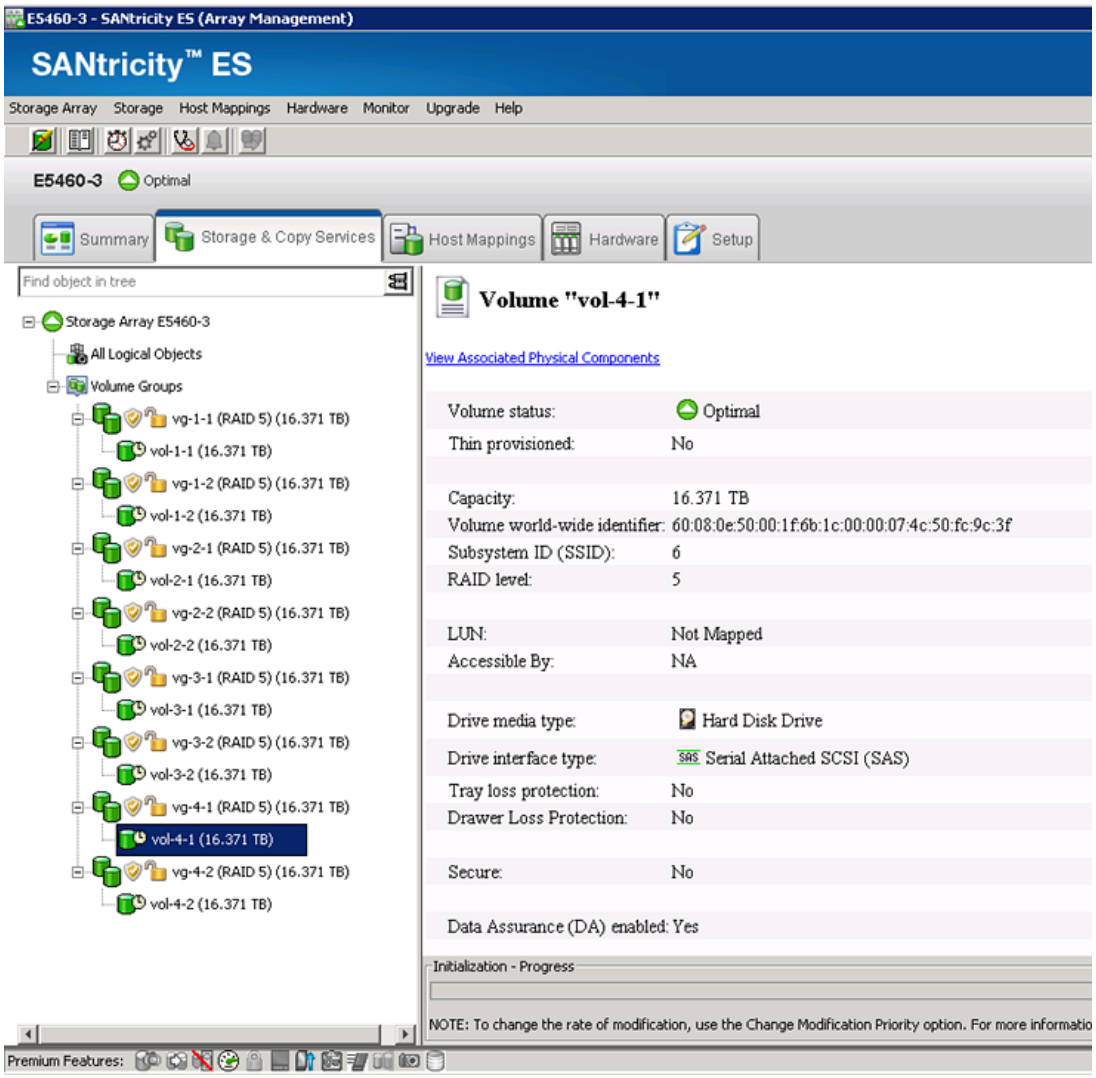


c. Click **Finish**.

d. On Create Volume Complete, a confirmation message box appears. Click **OK**.

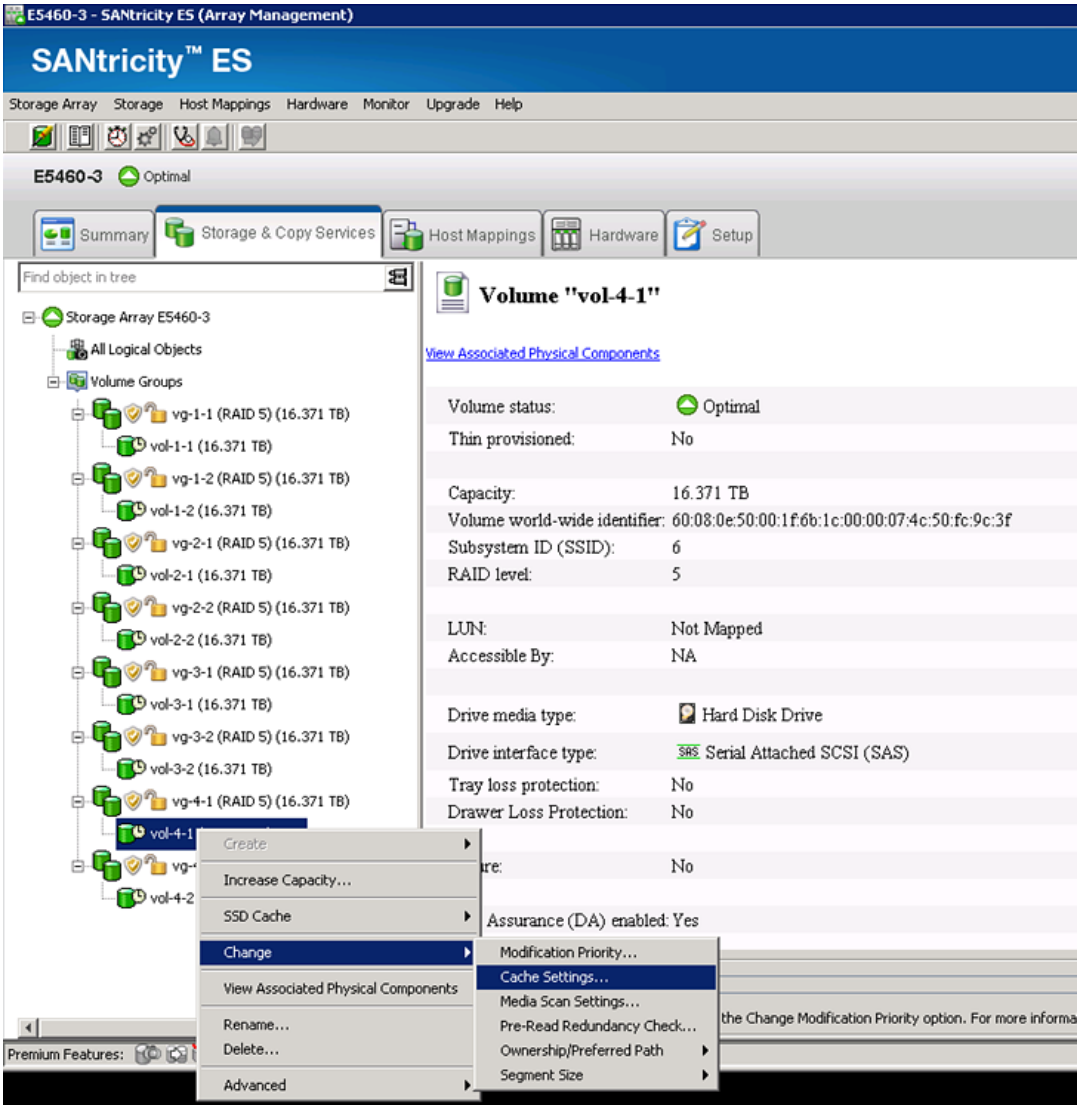
4. The volume group and the new volume are displayed in the Storage and Copy Services tab of the Array Management window.

Figure 102 Volume Group and Volume Capacity Details



5. Select the appropriate cache settings for the new volume:
   a. Right-click the volume name in the left pane in the Array Management window, and then select **Change > Cache Settings**.

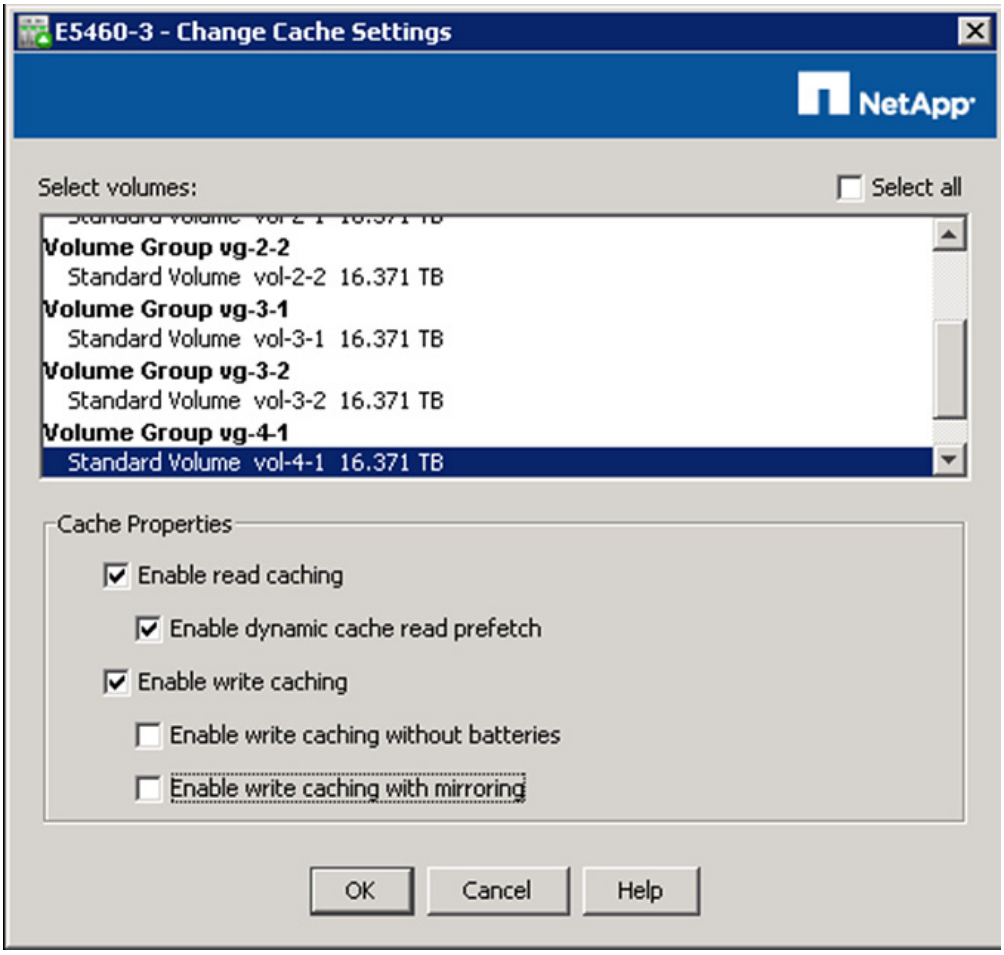
Figure 103 Cache Settings



6. Verify the following cache properties are enabled:
   a. Enable Read Caching
   b. Enable Write Caching
   c. Enable Dynamic Cache Read Prefetch

Note Verify that Enable Write Caching with Mirroring check box is deselected. This property is selected by default.

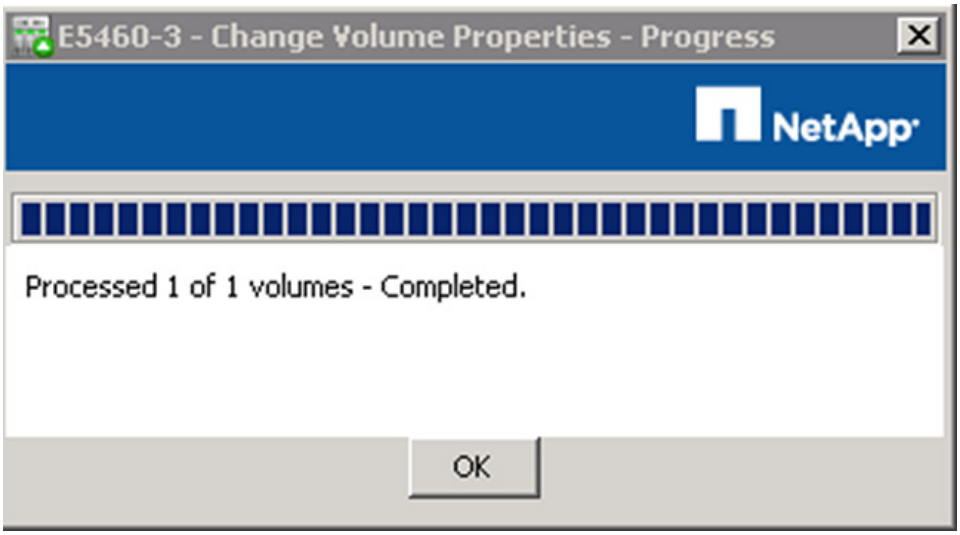
Figure 104 Verifying Cache Settings



d. Click **OK**.

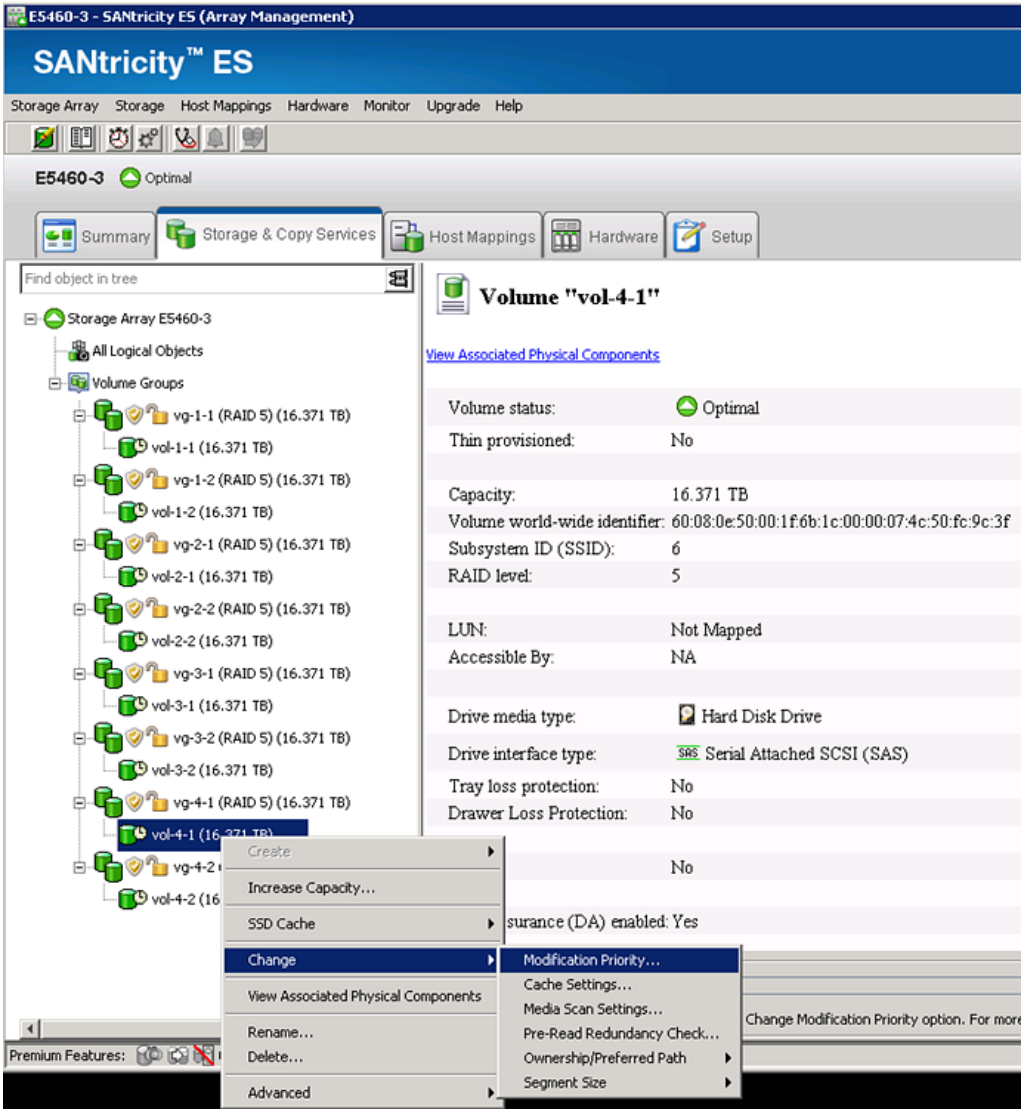
7. A pop-up message window appears displaying all the cache settings. Click **Yes** to confirm the settings.
8. Any Change in the Volume Properties on saving, will show a pop-up window showing progress indicator to indicate the change is completed successfully, click **OK**.

Figure 105 Changing Volume Properties



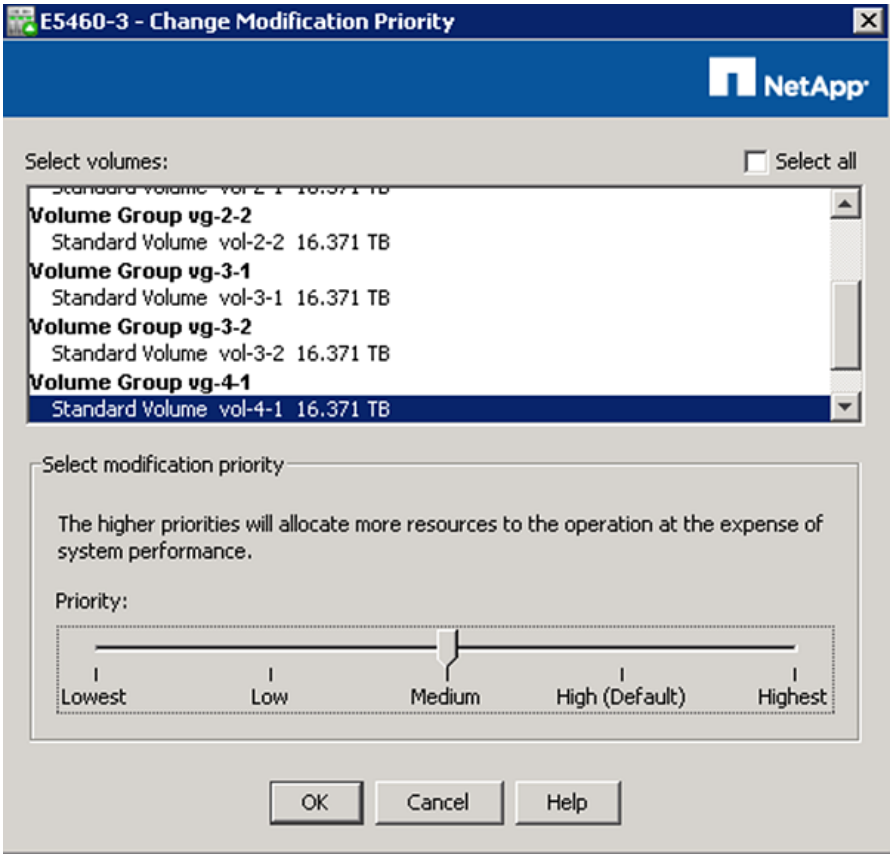
9. Right-click on the new volume. Select **Change > Modification Priority**.

Figure 106 Priority Settings



10. Use the priority slider to set the modification priority to the middle and click **OK**.

Figure 107 Allocating Resources Based on Priority Chosen



11. Click **Yes** in the confirmation message box, to confirm the change in the Priority.
12. Any Change in the Volume Properties on saving, will show a pop-up window showing progress indicator to indicate the change is completed successfully, click **OK**.

Table 9 provides the standard storage configuration:

Table 9 E-Series Hadoop standard storage configuration

Configuration	Value	Comments
Total number of disks	60	The DE6600 shelf provides 60 disk slots. When the E5400 controllers are installed in the shelf, the model is referred to as an E5460 shelf.
Number of volume groups	8	Stripe vertically across drawers in a shelf. There are two volume groups per host.
Volume group size	6+1 (RAID 5)	The volume group size is RAID 5.
Number of hot spares	4	One per drawer on drawers 2 through 5 for a total of 4 global hot spare drives per 60-disk shelf.
Number of volumes (LUNs)	8	Create two volumes of equal size (one volume per volume group) for each host.
Volumes per volume group	1	Create one volume per volume group.

Table 9 E-Series Hadoop standard storage configuration

Configuration	Value	Comments
Volume-to-host mapping	2:1	Each host maps exclusively to two volumes (LUNs).
Host groups	none	The Hadoop configuration does not share storage between hosts in the architecture.
Number of SAN share storage partitions	0	

After the volumes are created and the LUNs are available, the LUNs must be mapped to datanodes (hosts). It is critically important that each host has exclusive access to its LUNs through the assigned controller. NetApp strongly recommends using a naming convention that reflects the host-to-volume mappings.

Map Datanodes to E-Series SAS Ports

You need to determine the SAS port addresses of the datanodes and how they map to the SAS ports of the E-Series controllers. To map SAS ports of E-series controllers with the datanodes, follow these steps:

1. Identify all the hosts with their hostnames and SAS port addresses.
2. Identify hostnames and E-Series storage systems by name and label them accordingly.
3. Use the logic of physical proximity to assign hosts to E-Series controllers.
4. Use Table 10 to map datanodes to the E-Series SAS ports.

Table 10 SAS port-mapping template

Host Name	Storage Subsystem	Controller Port	Port Location	Datanode SAS ID
dn1	e5460-3	1	Controller A, Port 1	500605b002661880
dn2	e5460-3	4	Controller A, Port 4	
dn3	e5460-3	5	Controller B, port 5	
dn4	e5460-3	8	Controller B, Port 8	

Identify SAS Port ID Using LSI SAS2Flash Utility

The sas2flash utility is a product of LSI Corporation. It is designed to support the LSI SAS 9207-8e HBA installed in each host attached to the E5460 arrays, primarily to display the Port IDs assigned to the SAS host bus adapter (HBA) ports or to periodically update the BIOS and firmware on the HBA.

The sas2flash utility usually comes bundled with the following items:

- The sas2flash binary executable
- Version release notes in PDF format
- A reference guide in PDF format
- A BIOS binary file appropriate for the version (for example: mptsas2.rom)
- A firmware binary appropriate for the version (for example: 9207-8e.bin)

Issue the following command to list the port ID of the LSI SAS HBA:

```
./sas2flash -list
Output will look similar to this:
Version 11.00.00.00 (2011.08.22)
Copyright (c) 2008-2011 LSI Corporation. All rights reserved

        Adapter Selected is a LSI SAS: SAS2008(B2)

        Controller Number              : 0
        Controller                     : SAS2008(B2)
        PCI Address                    : 00:02:00:00
        SAS Address                    : 500605b-0-0266-1880
        NVDATA Version (Default)       : 0a.03.00.02
        NVDATA Version (Persistent)    : 0a.03.00.02
        Firmware Product ID            : 0x2213
        Firmware Version               : 11.00.00.00
        NVDATA Vendor                  : LSI
        NVDATA Product ID              : SAS9207-8e
        BIOS Version                   : 07.21.00.00
        UEFI BSD Version               : 04.30.03.00
        FCODE Version                  : N/A
        Board Name                     : 9200-3080
        Board Assembly                 : H3-25217-00C
        Board Tracer Number            : SP10414124

        Finished Processing Commands Successfully.
        Exiting SAS2Flash.
```

Note The line containing SAS Address provides the SAS port ID of the active port. Stripping out the hyphens allows the value to be used as the SAS port ID for that host when the host topology is created on the array being configured. In this case, the port ID is 500605b002661880.

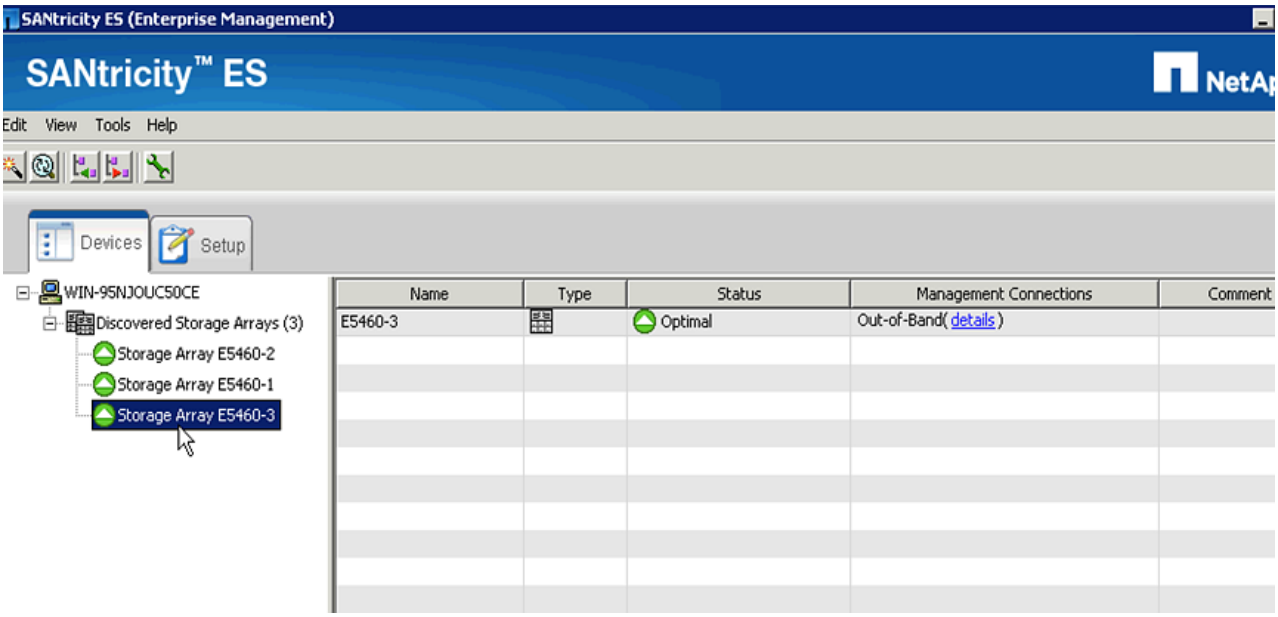
Map Datanodes to SAS Port IDs

After the SAS Port IDs for the datanode servers have been identified, map those IDs to the SAS ports on the E-Series storage array. Use the SANtricity ES Storage Manager to perform the mapping.

To map the datanodes, follow these steps:

1. Launch SANtricity ES Storage Manager on Windows or Linux as appropriate.
2. If the E-Series is already up and running, Storage Manager will list the array in the inventory. If not, add the appropriate storage controller IP addresses to SANtricity ES Storage Manager.

Figure 108 Discovered Storage Array



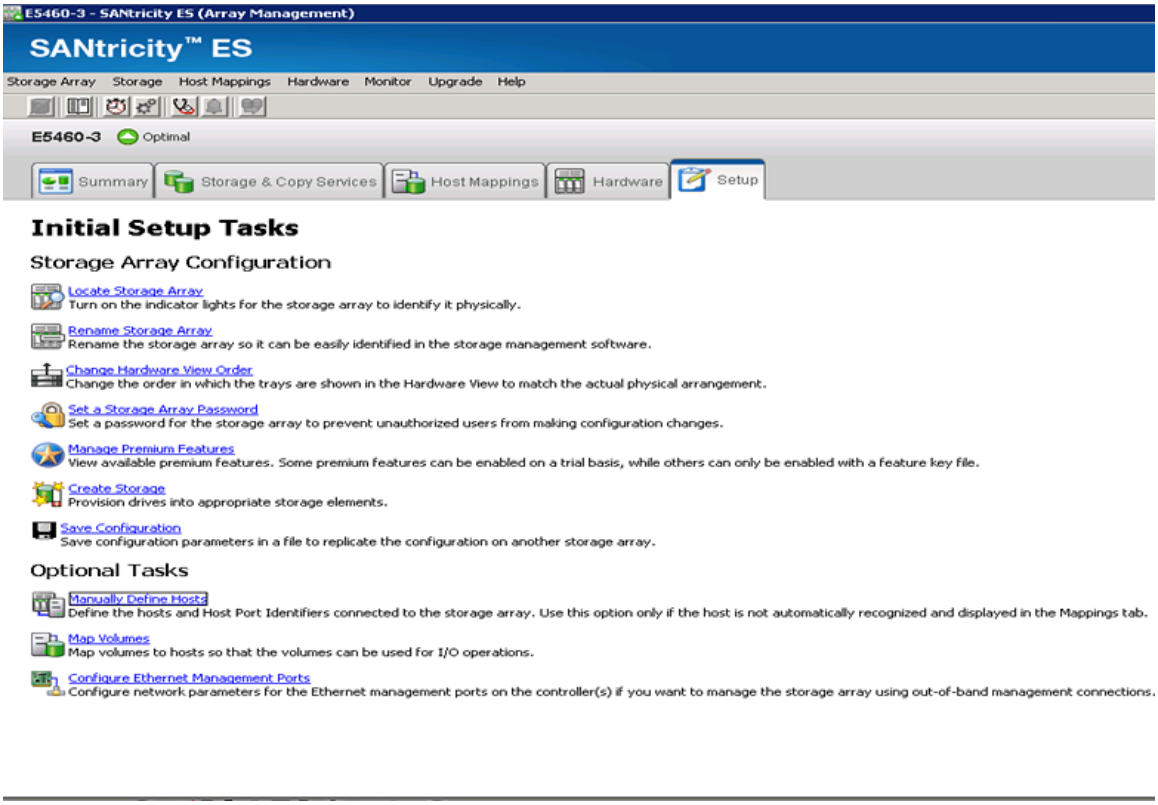
Note You must associate host port identifiers (World Wide Identifiers or WWIDs) with the applicable host before you can map storage to datanodes. Associating the wrong host ports can result in incorrect mapping of storage.

Hosts can be mapped using two different methods:

- The Define Host Wizard—In this example, we use the Define Host Wizard that is accessed via the Setup tab and then the Manually Define Hosts option on the Initial Setup Tasks page. With this method, hosts are matched to port identifiers from a list containing unassociated host port identifiers, which have been automatically detected by the storage controllers. Each host port identifier must be correctly associated with its host; otherwise, the host will access incorrect volumes and might fail to access any storage at all.
- The Manage Host Port Identifiers menu option—When adding host port identifiers, an alias or user label must be provided. The user label or the alias must not exceed 30 characters. Choose a meaningful user label or alias to easily identify the host port identifier. For example, include the host name and storage array name in the user label. As host port identifiers are associated with the applicable hosts, the identifiers are removed from the known unassociated host port identifiers list.

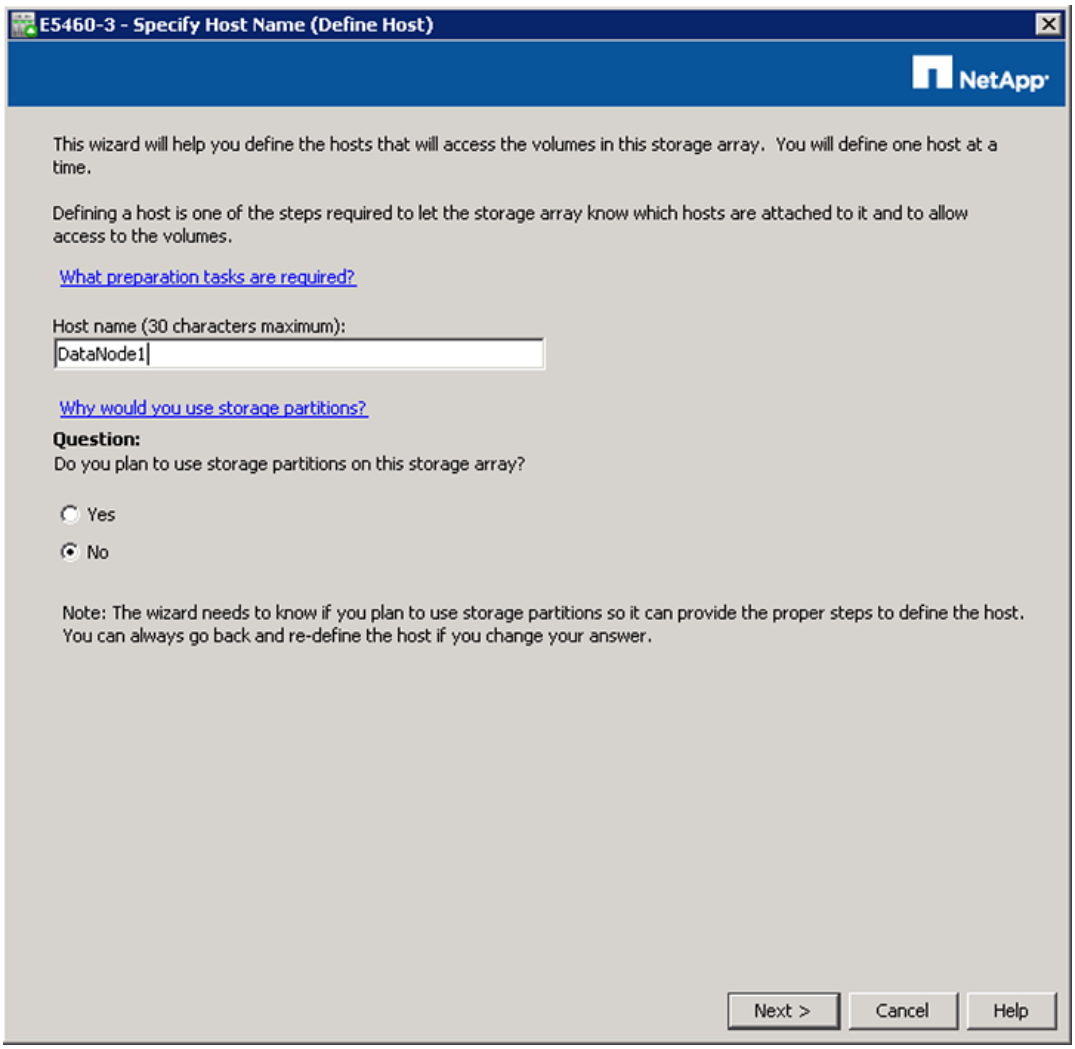
3. Select the Setup tab and then click **Manually Define Hosts**.

Figure 109 Defining Host Manually



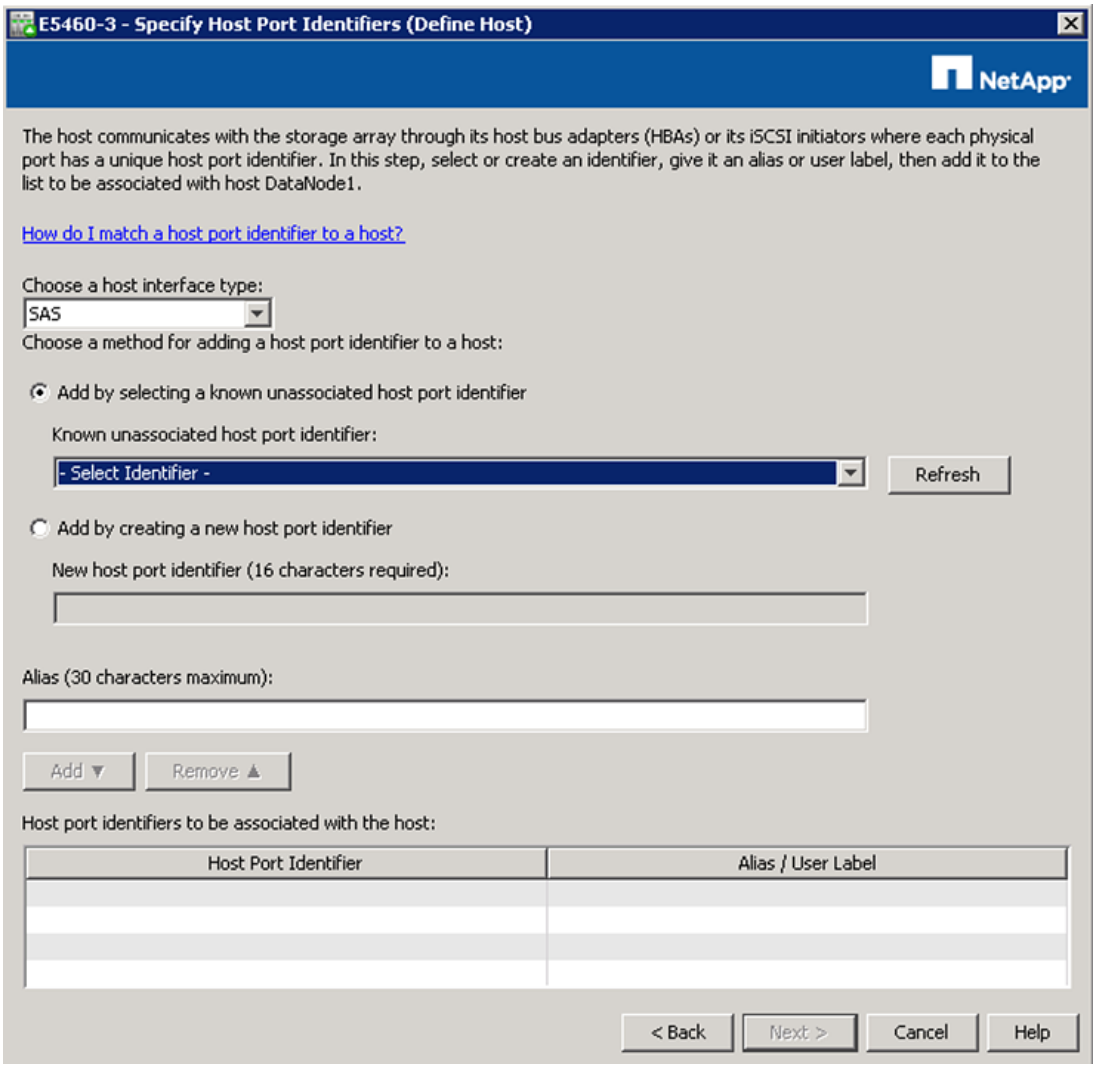
4. Enter the desired host name and click **Next**.

Figure 110 Entering Host Details



5. For Choose the host interface type option, select SAS from the drop down list.
6. Select the **Add by selecting a known unassociated host port identifier** radio button.
7. From the known unassociated host port identifier drop down list, select SAS ID of this host.

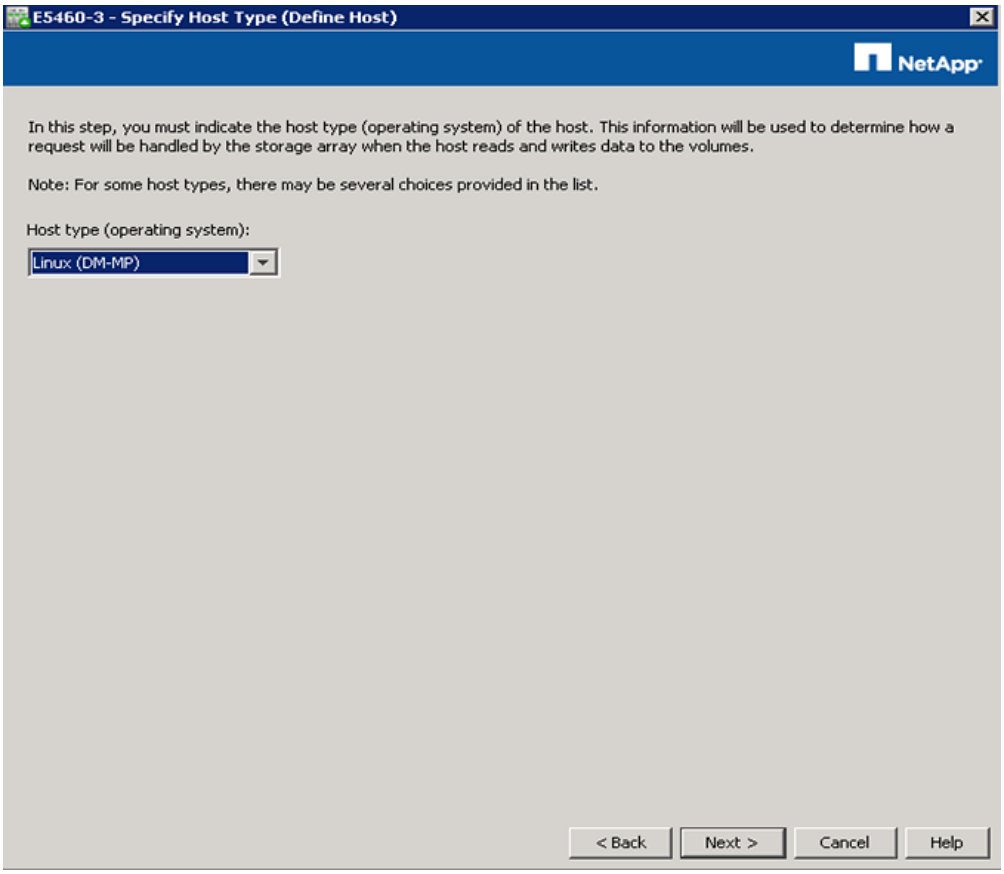
Figure 111 Entering Host Port Identifiers



Note Refer to the data gathered in, "Identify SAS Port ID Using LSI SAS2Flash Utility" section on page 132 that matches your host SAS connection.

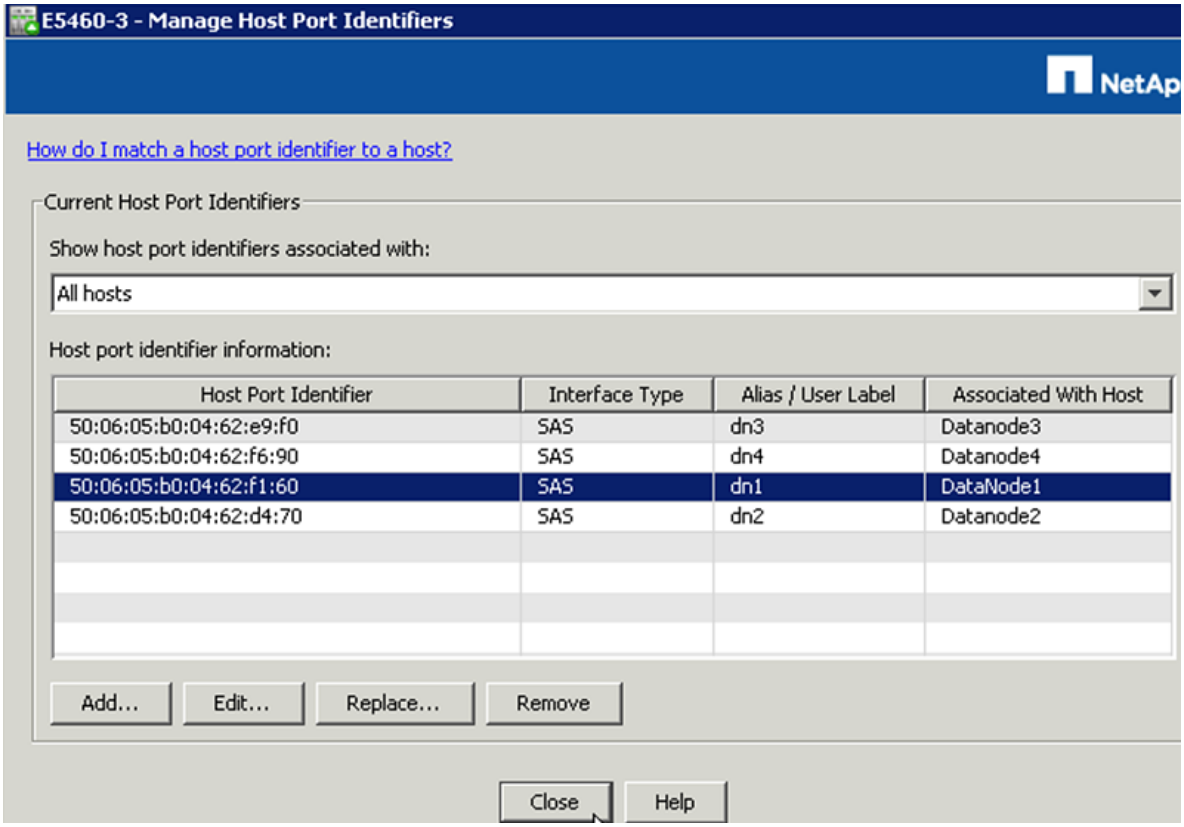
8. Select Add By Selecting a Known Unassociated Host Port Identifier and then from the drop down list select the appropriate host port identifier.
9. Enter the alias, click **Add**, and then click **Next**.
10. From the Host type (operating system) drop down list, select Linux (DM-MP) as the host type.

Figure 112 Specifying Host Type



11. The Preview (Define Host) pane appears. Click **Finish**.
12. Repeat steps 1 to11 for the remaining datanodes and host port identifiers.

Figure 113 Summary of Host Port Identifiers



Note To test, click **Host Mappings** at the top of the page and select Manage Host Port Identifiers from the list displayed. All of the hosts and port mappings are displayed. Verify that this list matches the information captured in Table 10.

E-Series Disk Initialization

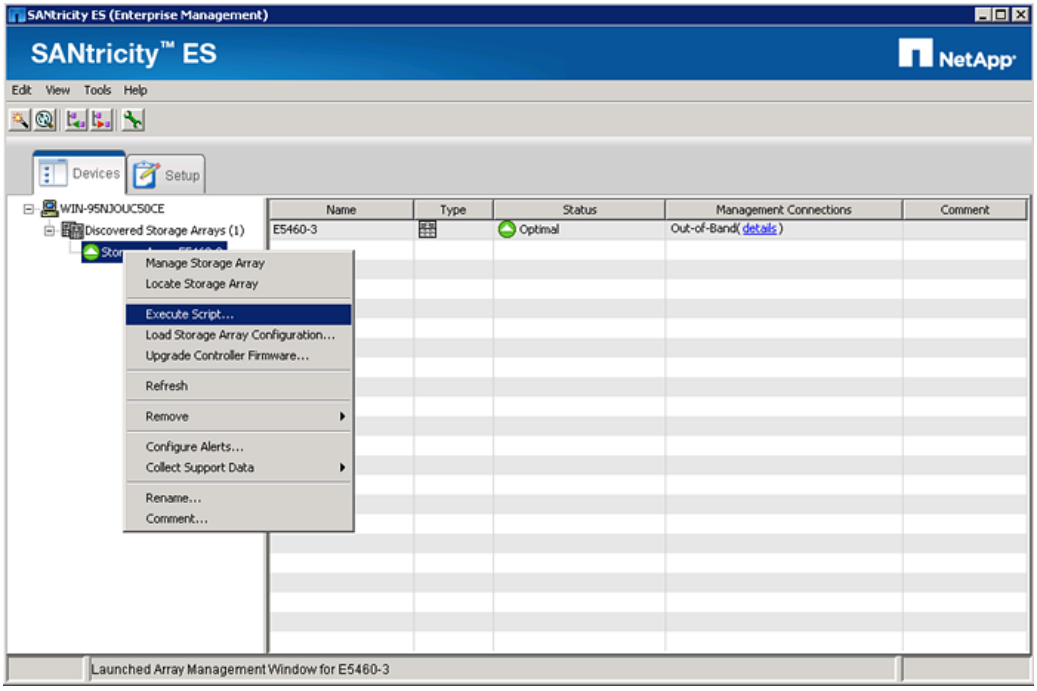
The disk initialization format is set to Immediate Availability Format (IAF) by default. When the disk initialization time, which can take more than 24 hours, can be a concern, NetApp recommends that IAF be disabled. This change blocks writes to the disks during the initialization process; however, the initialization process shortens dramatically. The decrease in initialization time ranges from minutes to several hours, depending on the size of the volume groups.

Note NetApp recommends that IAF be disabled when using 2TB and 3TB 7200-RPM disk drives; however, the setting should not be changed when Data Assurance (T10PI) is not enabled.

To disable the IAF disk initialization setting, follow these steps:

1. Open the SANtricity ES Management client, right-click the array where the setting need to be changed, and select Execute Script.

Figure 114 Selecting Execute Script to Verify Disk Initialization

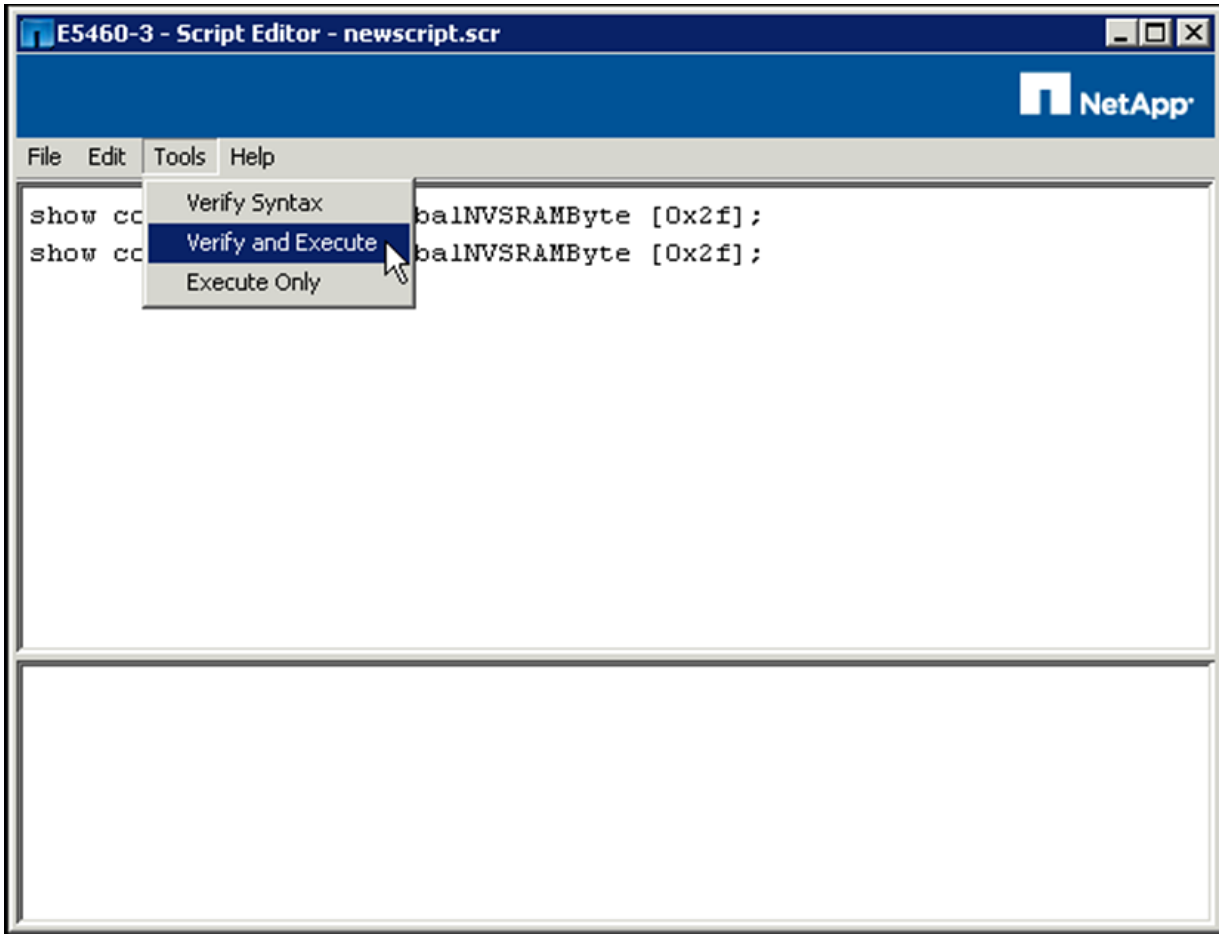


2. In the Script Editor – newscript.scr window, enter the show controller command to verify the current disk initialization method setting.

```
show controller [a] globalNVSRAMByte [0x2f];
show controller [b] globalNVSRAMByte [0x2f];
```

3. Select **Tools > Verify and Execute**. Confirm that the commands executed successfully.

Figure 115 Execute Commands and Verify

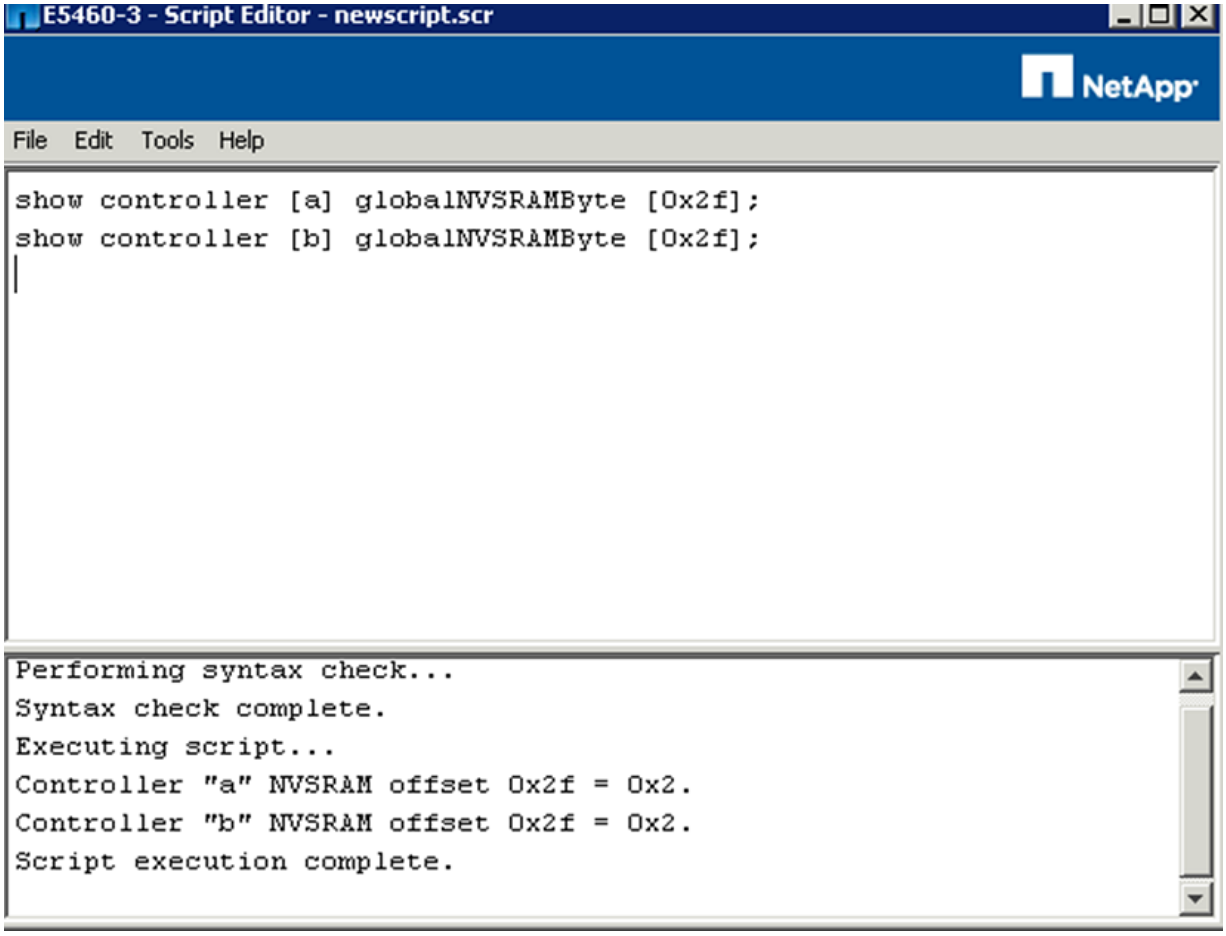


4. A setting of 0x0 indicates that IAF is enabled. To disable IAF, enter the set controller commands in the Script Editor window.

```
set controller [a] globalNVSRAMByte [0x2f] = 0x02, 0x02;
set controller [b] globalNVSRAMByte [0x2f] = 0x02, 0x02;
```

5. Select **Tools > Verify and Execute**. Confirm that the commands executed successfully.
6. Run the show controller commands again and confirm that the settings have been updated to 0x2 for both controllers.

Figure 116 Completion of Script Execution



Note With IAF disabled, disk initialization routines must complete before the array will allow data to be written to the newly created volume groups. If disk initialization cycles longer than 24 hours are acceptable, use the default IAF method. The disk initialization method must be selected before volume groups are created.

Disable E-Series Auto Volume Transfer

Auto volume transfer (AVT) is a feature of E-Series firmware that enables storage controller failover in the event that a controller becomes unavailable to datanodes. With the FlexPod Select for Hadoop, datanodes are directly connected to storage controllers using a single SAS cable. AVT is not applicable to this configuration, because connectivity to storage provides no redundancy. Failure of an E5460 controller is handled by the self-healing capability of Hadoop. Self-healing is supported by HDFS replication, which results in multiple copies of data being locally available to other nodes in the cluster. If a controller fails, MapReduce tasks using storage on that controller are reassigned to other healthy nodes in the Hadoop cluster that have access to another copy of the missing data. Since AVT adds no value to FlexPod Select for Hadoop, NetApp recommends that it be disabled.

To disable AVT, follow these steps:

1. Open the SANtricity ES Management client, right-click the array where the setting need to be changed, and select Execute Script as shown in Figure 114.
2. In the Script Editor – newscript.scr window, enter the show controller command to verify the current AVT setting.

```
show controller [a] hostNVSRAMByte [7,0x24];
show controller [b] hostNVSRAMByte [7,0x24];
```

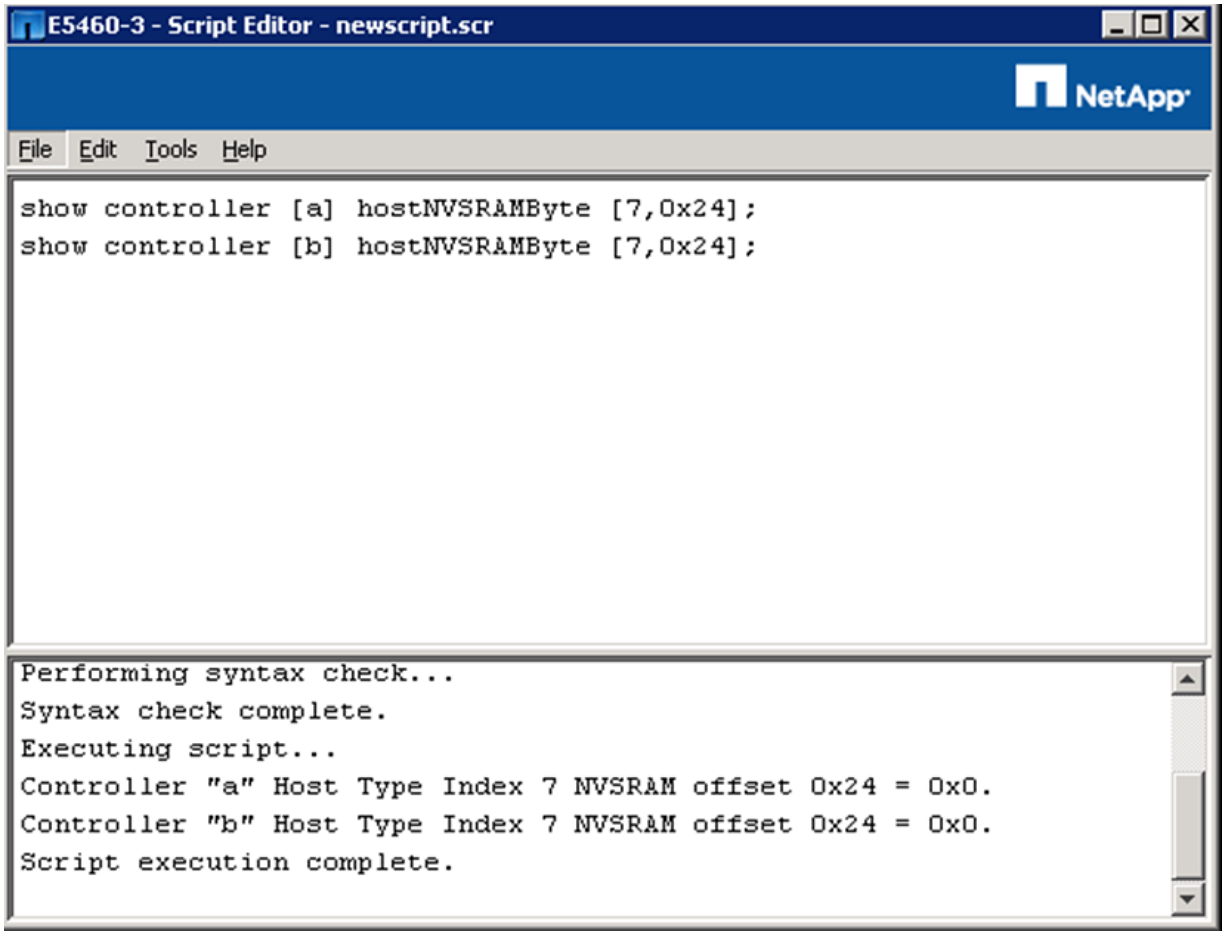
3. Select **Tools > Verify and Execute** as shown in Figure 115. Confirm that the commands executed successfully.
4. A setting of 0x1 indicates that AVT is enabled. To disable AVT, enter the set controller commands in the Script Editor window.

```
set controller [a] hostNVSRAMByte[7,0x24]=0x00;
set controller [b] hostNVSRAMByte[7,0x24]=0x00;
```

5. Select the **Tools > Verify and Execute**. Confirm that the commands executed successfully.
6. Rerun the show controller commands from step 2.

   Confirm the settings have been updated to 0x0 for both controllers.

Figure 117 Completion of Script Execution



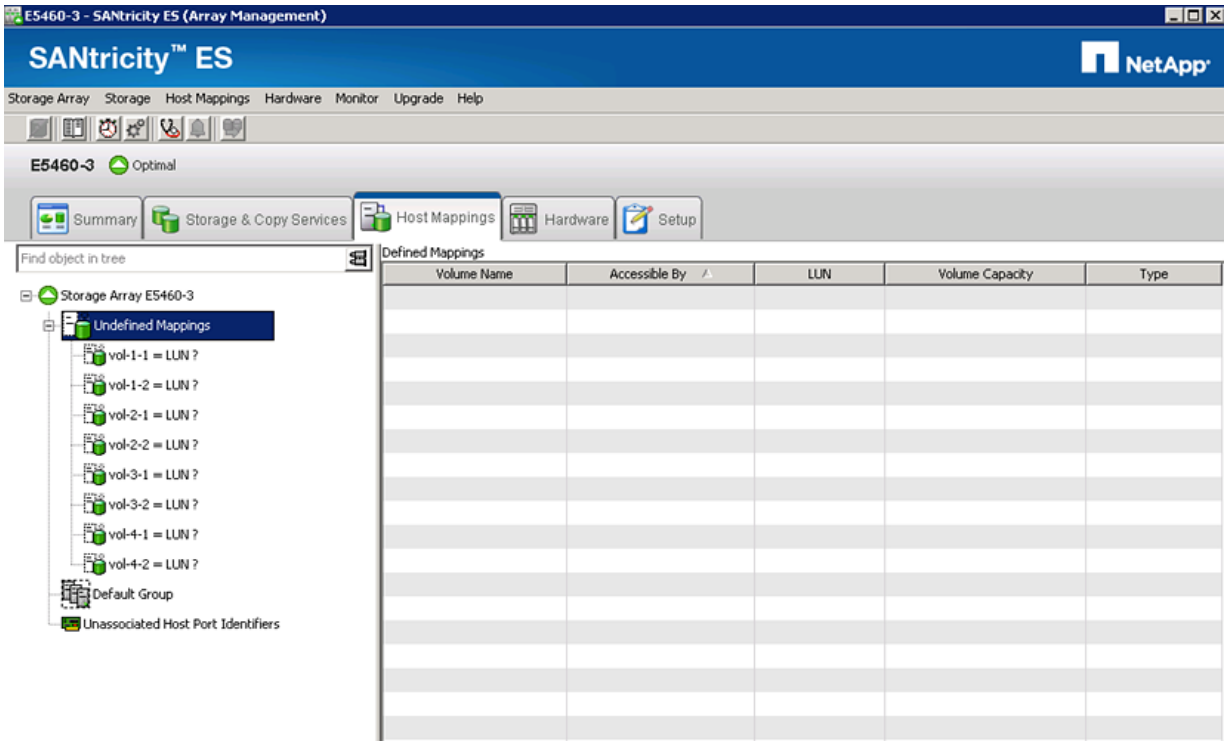
7. Close the Script Editor window and proceed with creating volume groups.

Map Volumes

To map the volumes to the assigned datanode, follow these steps:

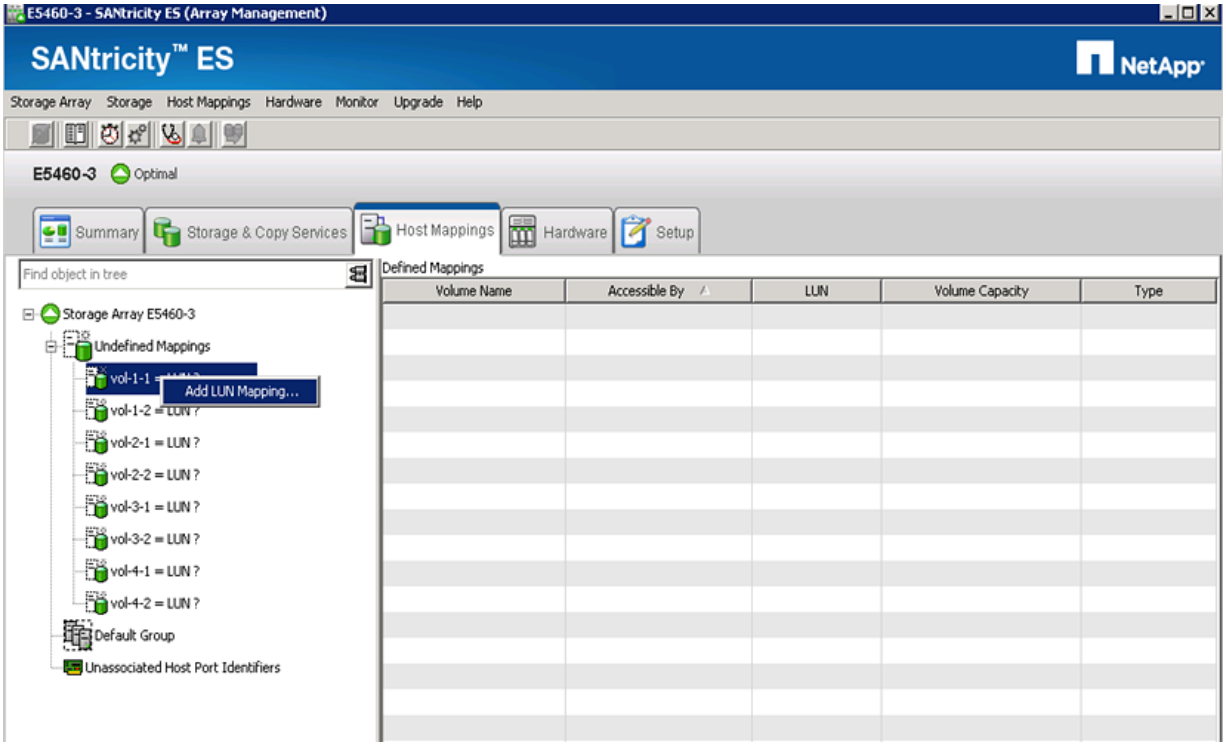
1. Select the Host Mappings tab at the top of the Array Management window. Expand the Undefined Mappings.

Figure 118 Unmapped LUNs



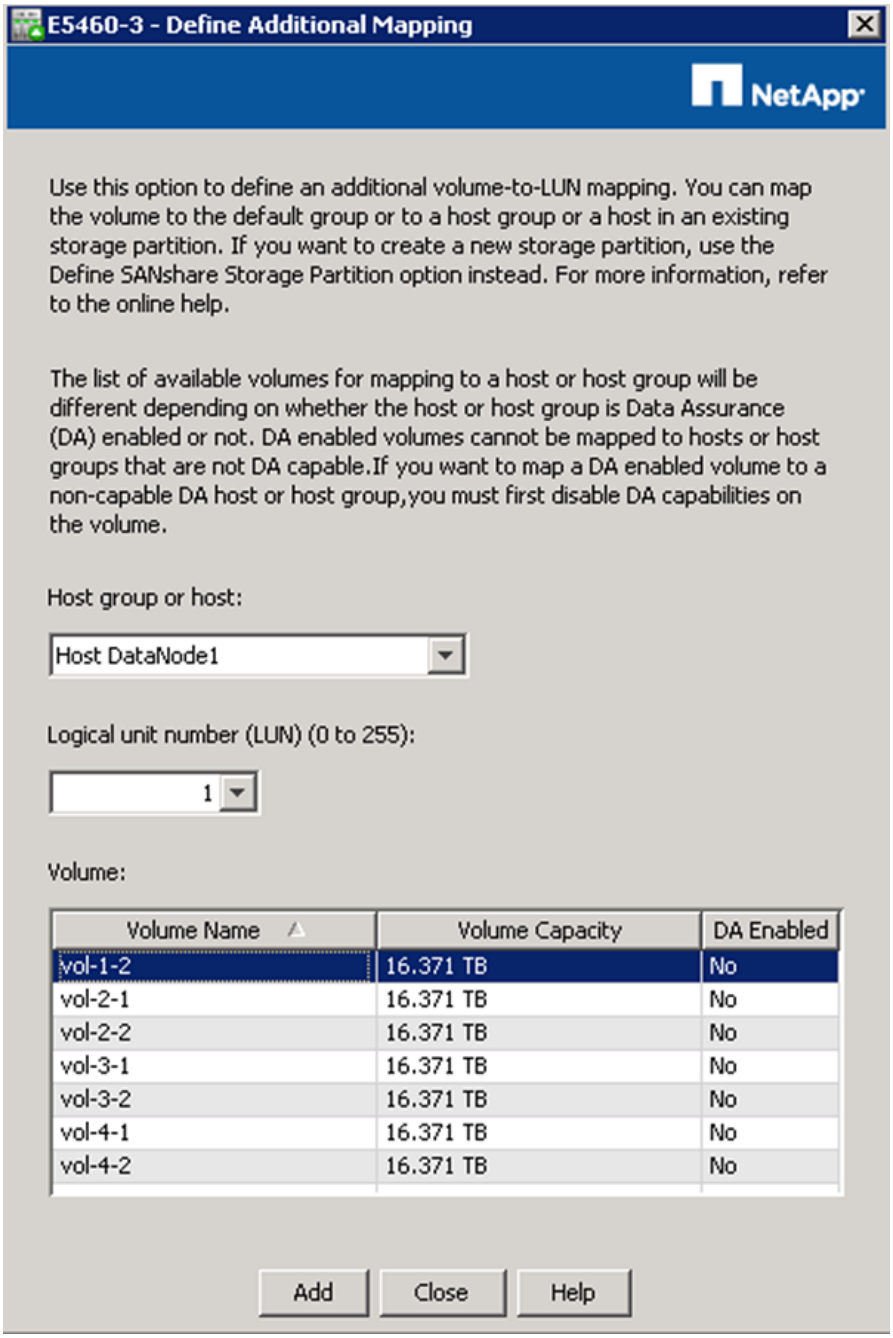
2. Right-click the new volume name under the Undefined Mappings. click **Add LUN Mapping**.

Figure 119 Adding LUN Mapping



3. In the Define Additional Mapping window, perform the following actions:
   a. Select the host to use for the mapping.
   b. Set the desired logical unit number (LUN) [in this example: 1].
   c. Select the desired volume.

Figure 120 Entering Additional Mapping Information



d. Click **Add**.

e. Repeat the above steps (a-d) for each volume to map to each host.

Note If a window opens giving the option to map more volumes, click **Close**. In order to maintain consistency with other applications that use E-Series storage, never use a LUN number of 0. NetApp recommends starting with a LUN number of 1.

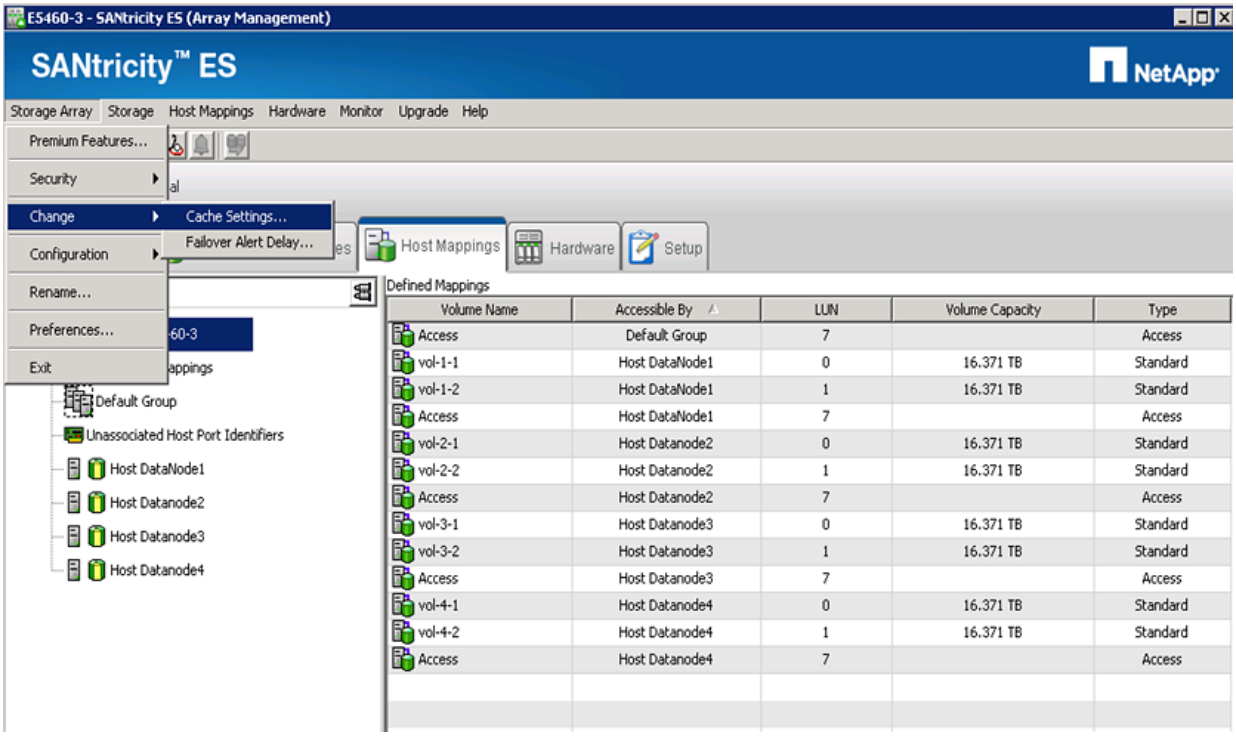
4. After all LUNs are mapped, reboot the hosts so that they recognize their LUNs.

Configure Cache Settings for Array

To configure caching for the entire storage array, follow these steps:

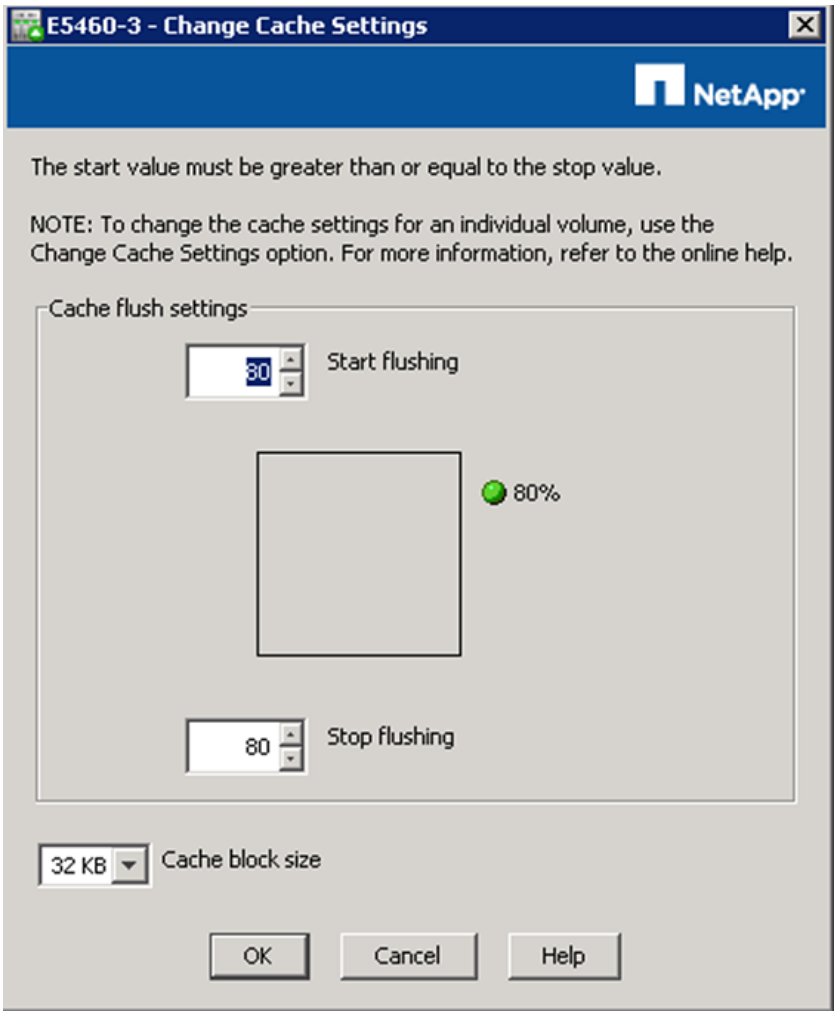
1. From the top menu of the SANtricity Array Management window, select Storage Array. Then select **Change > Cache Settings**.

Figure 121 Cache Settings



2. In the Change Cache Settings window, change the Cache Block Size setting to 32KB, verify that the Stop Flushing setting is 80, and click **OK**.

Figure 122 Changing the Cache Block Size



Note Do not make any other changes.

This completes the E5460 initialization and storage configuration.

DataNode File Systems on E5460 LUNs

Once the E5460 systems have finished their initialization, it is time to create partition tables and file systems on the LUNs supplied to each of the datanodes.

The following script should be run as root user on each of the datanodes:

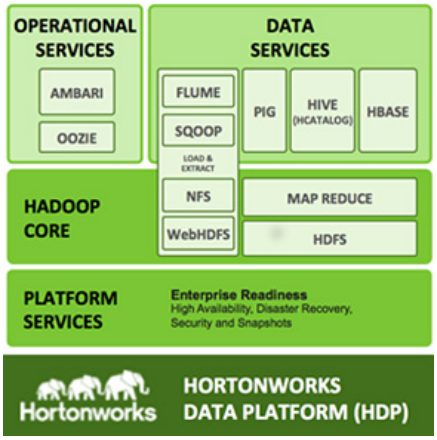
```
#!/bin/bash
[[ "-x" == "${1}" ]] && set -x && set -v && shift 1
# rescan scsi devs
for X in /sys/class/scsi_host/host?/scan
do
  echo '- - -' > ${X}
done
# find new LUNs
```

```
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
    (( $? )) && continue
    /bin/mkdir  -p  /HDP/${Y}
    (( $? )) && continue
    /bin/mount  -t xfs  -o allocsize=128m,noatime,nobarrier,nodiratime  ${X}1  /HDP/${Y}
    (( $? )) && continue
    echo "LABEL=${Y}  /HDP/${Y}   xfs   allocsize=128m,noatime,nobarrier,nodiratime  0  2"
>> /etc/fstab
  fi
done
```

Installing HDP

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 can expect from the enterprise level software. HDP components are shown in Figure 123.

Figure 123 HDP Components



Prerequisites for HDP Installation

This section details the prerequisites for HDP installation such as setting up of EPEL and HDP Repo.

Hortonworks and EPEL Repo

From a host connected to the Internet, download the EPEL and Hortonworks repositories as shown below and transfer it to infra-0.

1. Download EPEL Repository from the system connected to the Internet.

```
mkdir –p /tmp/Hortonworks
cd /tmp/Hortonworks
rpm -Uvh
http://download.fedoraproject.org/pub/epel/6/x86_64/epel-release-6-8.noarch.rpm
reposync  -r epel
```

2. Download Hortonworks HDP Repo

```
wget
http://public-repo-1.hortonworks.com/HDP/centos6/HDP-1.3.0.0-centos6-rpm.tar.gz
```

3. Download Hortonworks HDP-Utils Repo

```
wget
http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.15/repos/centos6/HDP-UTILS-1.
1.0.15-centos6.tar.gz
```

4. Download Ambari Repo

```
wget
http://public-repo-1.hortonworks.com/ambari/centos6/ambari-1.2.3.7-centos6.tar.gz
```

5. Copy the repository directory to infra-0

```
scp -r /tmp/Hortonworks/ infra-0:/var/www/html
```

6. Extract the files

```
login to infra-0
cd /var/www/html/Hortonworks
tar -zxvf HDP-1.3.0.0-centos6-rpm.tar.gz
tar -zxvf HDP-UTILS-1.1.0.15-centos6.tar.gz
tar -zxvf ambari-1.2.3.7-centos6.tar.gz
```

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

```
vi  /etc/yum.repos.d/hdp.repo
[HDP-1.3.0.0.]
name=Hortonworks Data Platform Version - HDP-1.3.0.0
baseurl=http://10.29.160.53/Hortonworks/HDP/centos6/1.x/GA/1.3.0.0/
gpgcheck=0
enabled=1
priority=1
```

8. Create the hdp-utils.repo file with following contents:

```
vi  /etc/yum.repos.d/hdp-utils.repo
[HDP-UTILS-1.1.0.15]
name=Hortonworks Data Platform Version -HDP-UTILS-1.1.0.15
baseurl=http://10.29.160.53/Hortonworks/HDP-UTILS-1.1.0.15/repos/centos6
gpgcheck=0
enabled=1
priority=1
```

9. Create the Ambari repo file with following contents:

```
vi  /etc/yum.repos.d/ambari.repo
[Updates-ambari-1.2.3.7]
name=ambari-1.2.3.7 - Updates
```

```
baseurl=http://10.29.160.53/Hortonworks/ambari/centos6/1.x/updates/1.2.3.7
gpgcheck=0
enabled=1
priority=1
```

10. Create epel.repo

```
cd  /var/www/html/Hortonworks/epel
createrepo .
vi /etc/yum.repos.d/epel.repo
name=Extra Packages for Enterprise Linux 6 - $basearch
baseurl=http://10.29.160.53/Hortonworks/epel/
enabled=1
gpgcheck=0
priority=1
```

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

HDP Installation

To install HDP, issue the following CLI commands:

Install and Setup Ambari Server on infra-0

```
yum install ambari-server
```

Setup Ambari Server

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

Configure Ambari Server to use Local Repository

Edit redhat6 and centos6 sections of the Ambari **repoinfo.xml** to point to local repository.

```
vi /var/lib/ambari-server/resources/stacks/HDPLocal/1.3.0/repos/repoinfo.xml
Replace the xml element <os type="redhat6"> .. </os> with
<os type="redhat6">
  <repo>
     <baseurl>http://10.29.160.53/Hortonworks/HDP/centos6/1.x/GA/1.3.0.0</baseurl>
     <repoid>HDP-1.3.0</repoid>
     <reponame>HDP</reponame>
  </repo>
  <repo>
     <baseurl>http://10.29.160.53/Hortonworks/epel</baseurl>
     <repoid>HDP-epel</repoid>
     <reponame>HDP-epel</reponame>

<mirrorslist><![CDATA[http://mirrors.fedoraproject.org/mirrorlist?repo=epel-6&arch=$basear
ch]]></mirrorslist>
  </repo>
</os>

Replace the xml element <os type="centos6"> .. </os> with
```

```
<os type="centos6">
  <repo>
     <baseurl>http://10.29.160.53/Hortonworks/HDP/centos6/1.x/GA/1.3.0.0</baseurl>
     <repoid>HDP-1.3.0</repoid>
     <reponame>HDP</reponame>
  </repo>
  <repo>
     <baseurl>http://10.29.160.53/Hortonworks/epel</baseurl>
     <repoid>HDP-epel</repoid>
     <reponame>HDP-epel</reponame>

<mirrorslist><![CDATA[http://mirrors.fedoraproject.org/mirrorlist?repo=epel-6&arch=$basear
ch]]></mirrorslist>
  </repo>
</os>
```

Start Ambari Server

```
ambari-server start
```

Confirm Ambari Server Startup

```
ps –ef | grep ambari-server
```

Login to Ambari Server

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

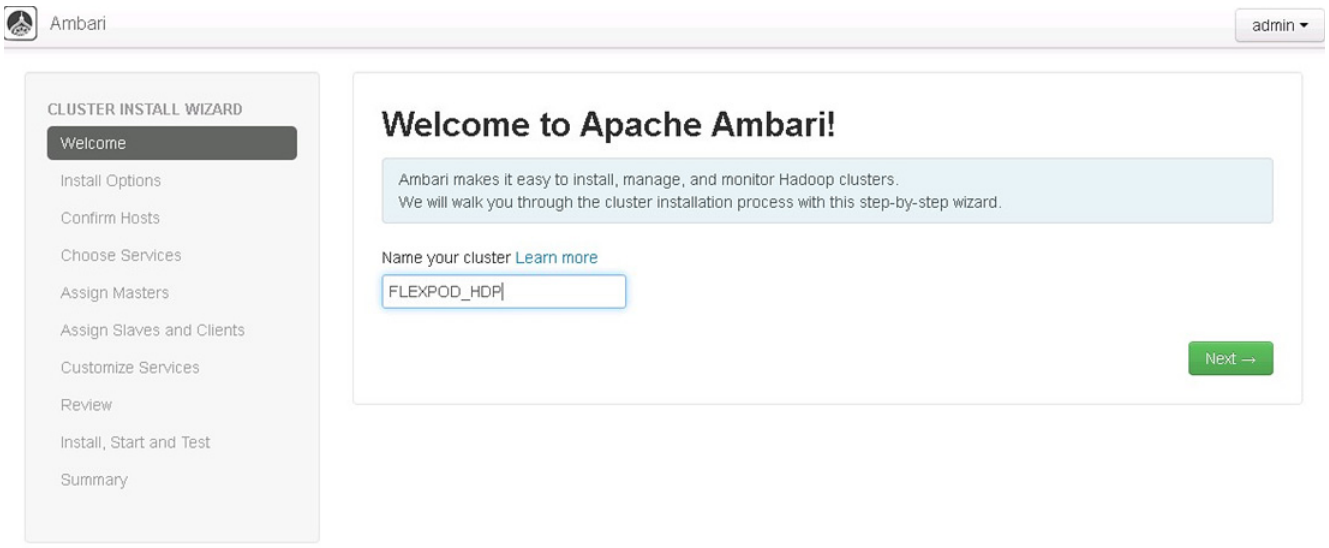
1. Point the browser to "Error! Hyperlink reference not valid". address of infra-0>:8080. For instance, http://10.29.160.53:8080.
2. Login to the Ambari Server using the default username/password as admin/admin, which can be changed at a later period of time.

Create Cluster Name

Follow these steps to create the cluster name:

1. In the Ambari Server home page, enter FLEXPOD_HDP in the name field.

Figure 124 Apache Ambari - Home Page



HDP Cluster Installation

In order to build a cluster, the install wizard needs to have a general information about how the cluster needs to be set up. For this, you need to provide the Fully Qualified Domain Name (FQDN) of each one of the hosts. The wizard also needs to access the private key file that was created in "Setting Up Password-less Login" section on page 85. 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 by one. You can also give a range within the brackets to indicate larger sets of hosts.
2. Select the option Provide your SSH Private Key in the Ambari cluster install wizard.
   a. Copy the contents of the file /root/.ssh/id_rsa on infra-0 and paste it in the text area provided by the Ambari cluster install wizard.

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

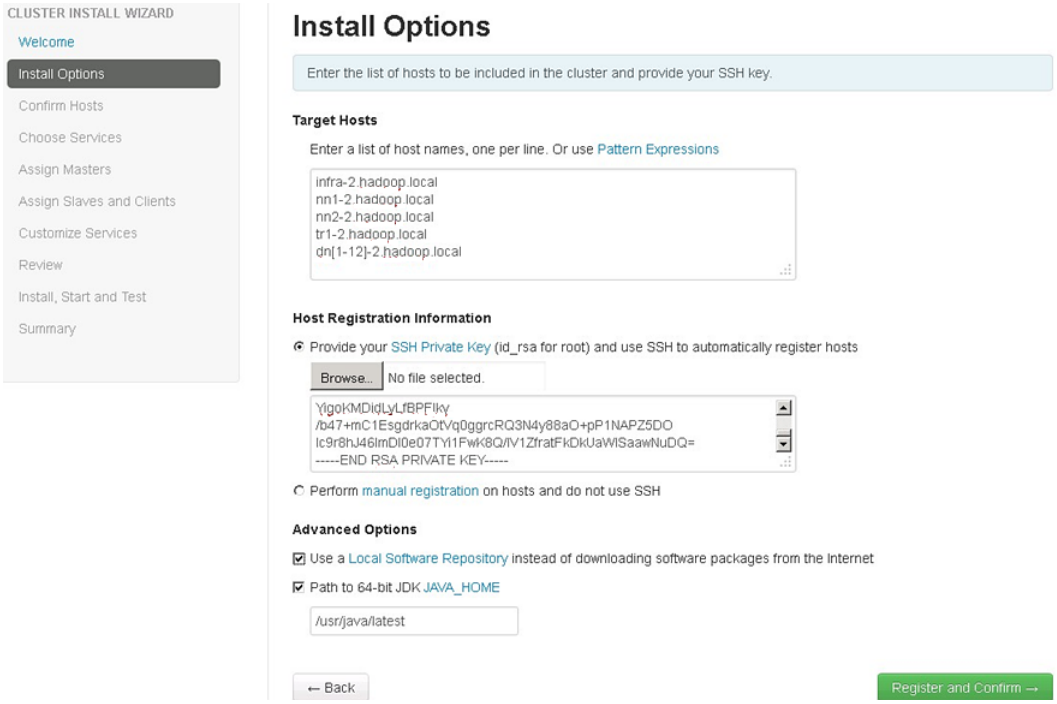
Figure 125 Copying Contents from /root/.ssh/id_rsa

3. Select “Use Local Repository” and path to 64-bit JDK.
4. Provide a path to java install directory. To get java install directory, run the following command on infra-0.

```
echo $JAVA_HOME
```

5. Click **Register and Confirm** to continue.

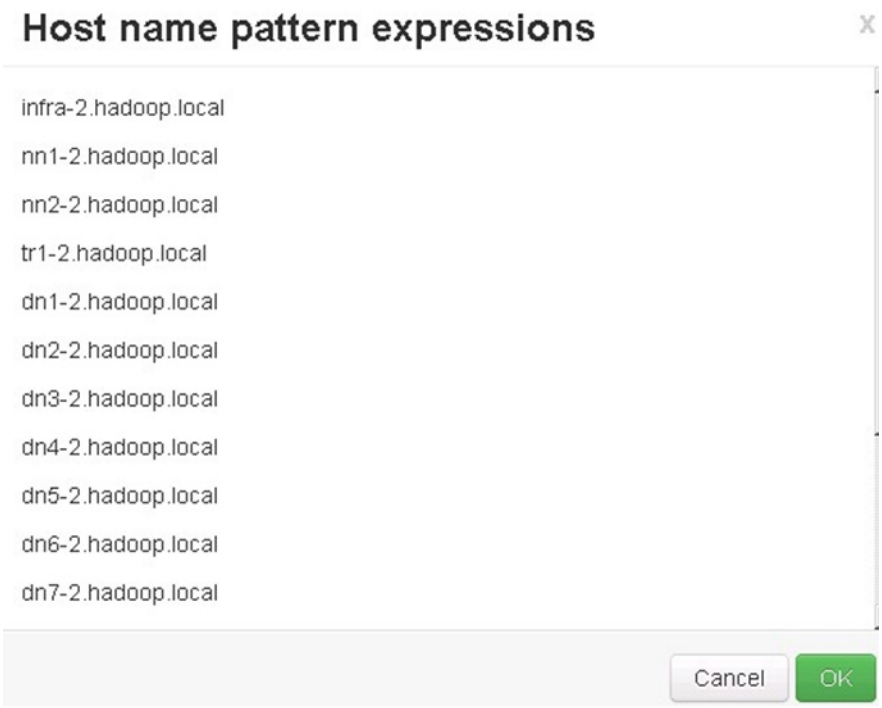
Figure 126 Confirming the HDP Installation



Host Name Pattern Expressions

Figure 127 shows a list of target host names using pattern expressions.

Figure 127 Host Name pattern Expressions



Confirming Hosts

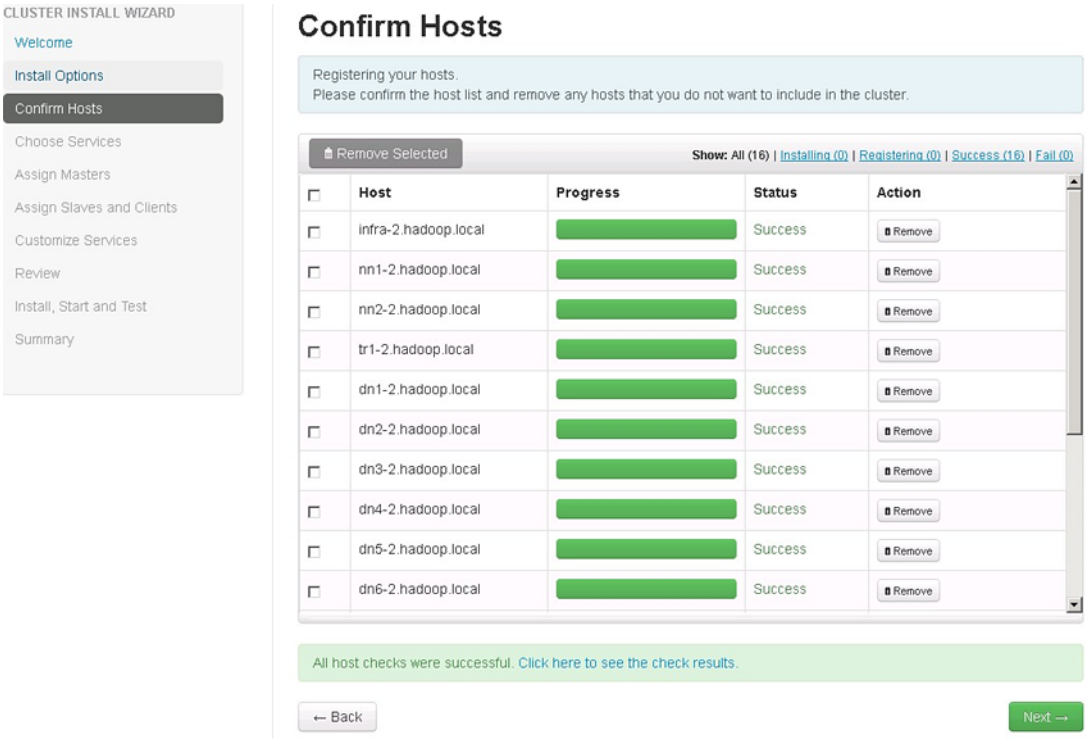
This screen allows you to make sure that the Ambari server has located all the required hosts for the cluster and to make sure that the hosts have correct directories, packages, and processes to continue the installation process.

You can remove the undesired hosts that were selected by the Ambari server. To remove all the undesired hosts, check the appropriate check boxes provided against each of the hosts and then click 

. To remove a single host, click

.

Figure 128 Confirming Hosts to be Included in the Cluster

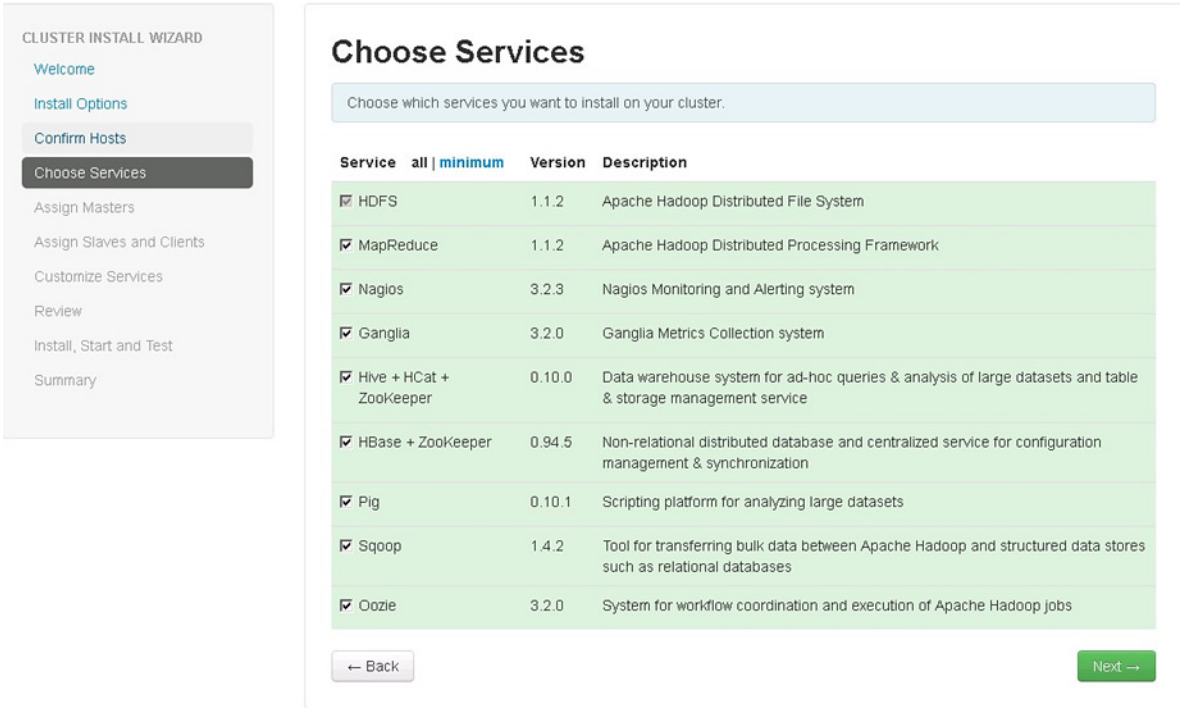


Choose Services

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**.

Figure 129 Choosing Services for the Cluster



Assign Masters

The Ambari install wizard attempts to assign the master nodes for various services that have been selected for the appropriate hosts in the cluster. Figure 130 shows the current service assignments by the host, the hostname and its number of CPU cores and RAM size.

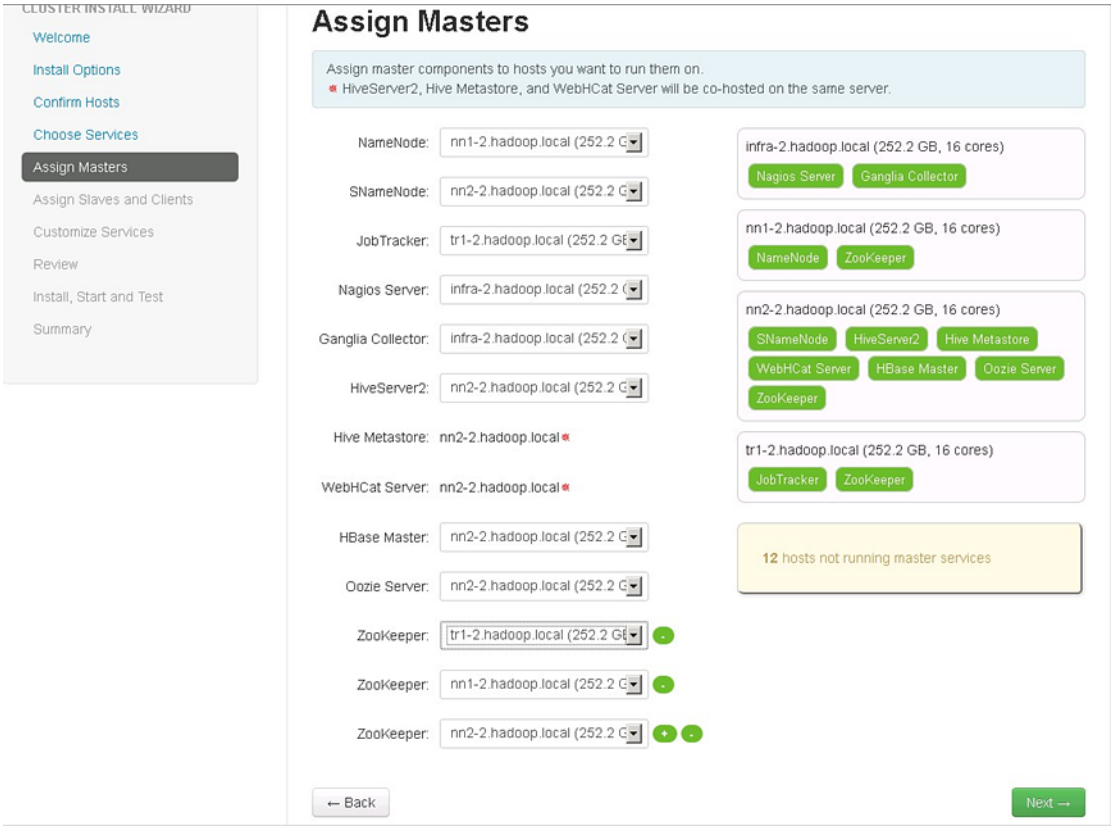
1. Reconfigure the service assignment to match the Table 11:

Table 11 Service Assignment

Service Name	Host
NameNode	nn1-2.hdoop.local
SNameNode	nn2-2.hadoop.local
JobTracker	tr1-2.hadoop.local
Nagios Server	infra-2.hadoop.local
Ganglia Collector	infra-2.hadoop.local
Hive Server2	nn2-2.hadoop.local
HBase Master	nn2-2.hadoop.local
Oozie Server	nn2-2.hadoop.local
ZooKeeper	tr1-2.hadoop.local nn1-2.hdoop.local nn2-2.hdoop.local

2. Click **Next**.

Figure 130 Assigning Master Components

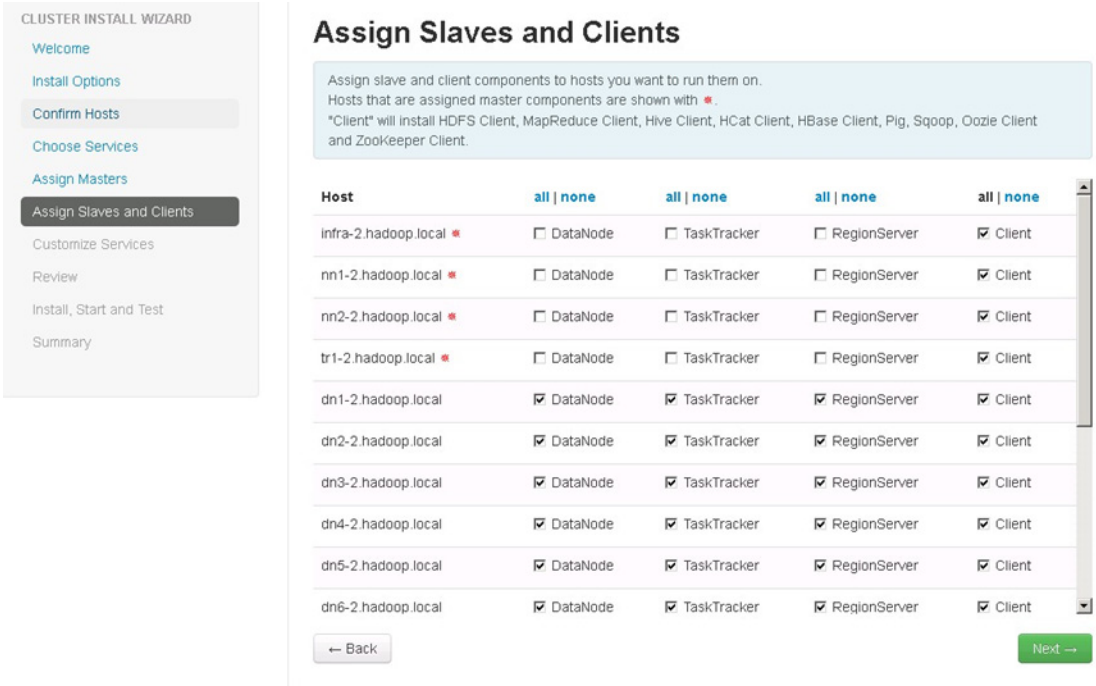


Assign Slaves and Clients

The Ambari install wizard attempts to assign the slave components (DataNodes, TaskTrackers, and RegionServers) to appropriate hosts in the cluster. Reconfigure the service assignment to match Figure 131:

1. Accept the default assignment for DataNode, TaskTracker, RegionServer nodes.
2. Assign Client to all nodes.
3. Click **Next**.

Figure 131 Assigning Slave and Client Components to Hosts



Customize Services

Customize Services window in the cluster install wizard presents a number of configuration settings to manage Hadoop components. The configuration settings can be done based on your requirements under each of the tabs as shown in Figure 132. This window shows the default settings for each of the configuration options, but you can modify the settings to meet specific requirements.

Following are the configurations available in the cluster install wizard:

- HDFS, page 160
- MapReduce, page 162
- Hive/HCat, page 164
- WebHCat, page 165
- HBase, page 166
- ZooKeeper, page 167
- Oozie, page 168
- Nagios, page 169
- Misc, page 170

The following sections provide details of each of these configurations.

HDFS

Update the HDFS configurations as shown in Table 12, and Figure 132 and Figure 133:

Table 12 HDFS Configurations

Property Name	Value
NameNode Java Heap Size	4GB
Reserved space for HDFS	4GB

Figure 132 Customize Services - HDFS Configuration Window Part 1

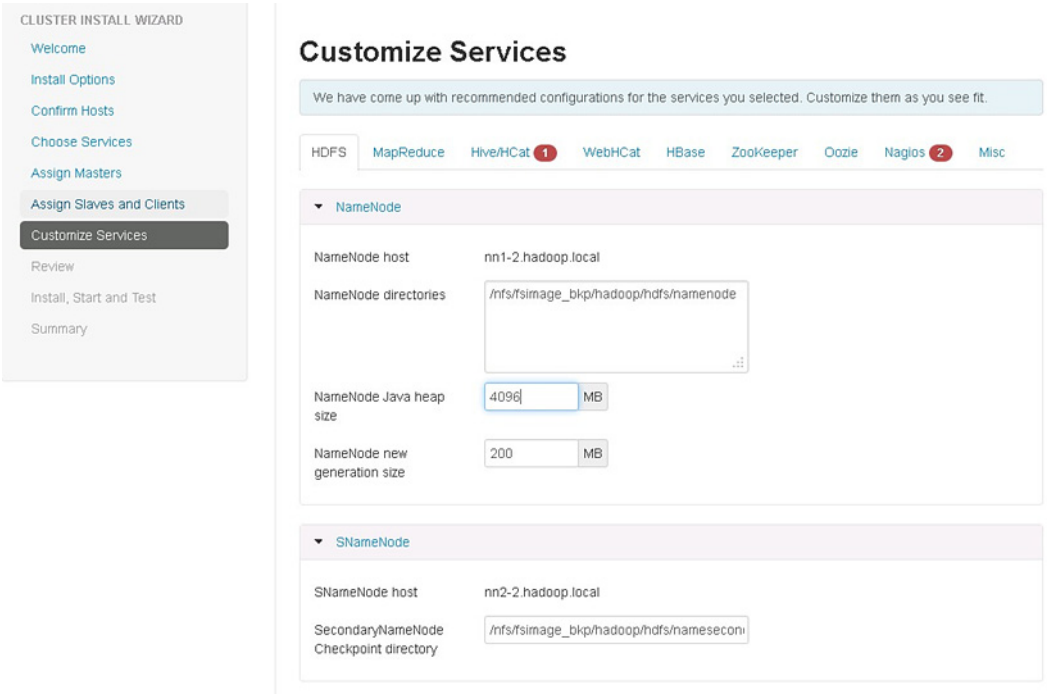
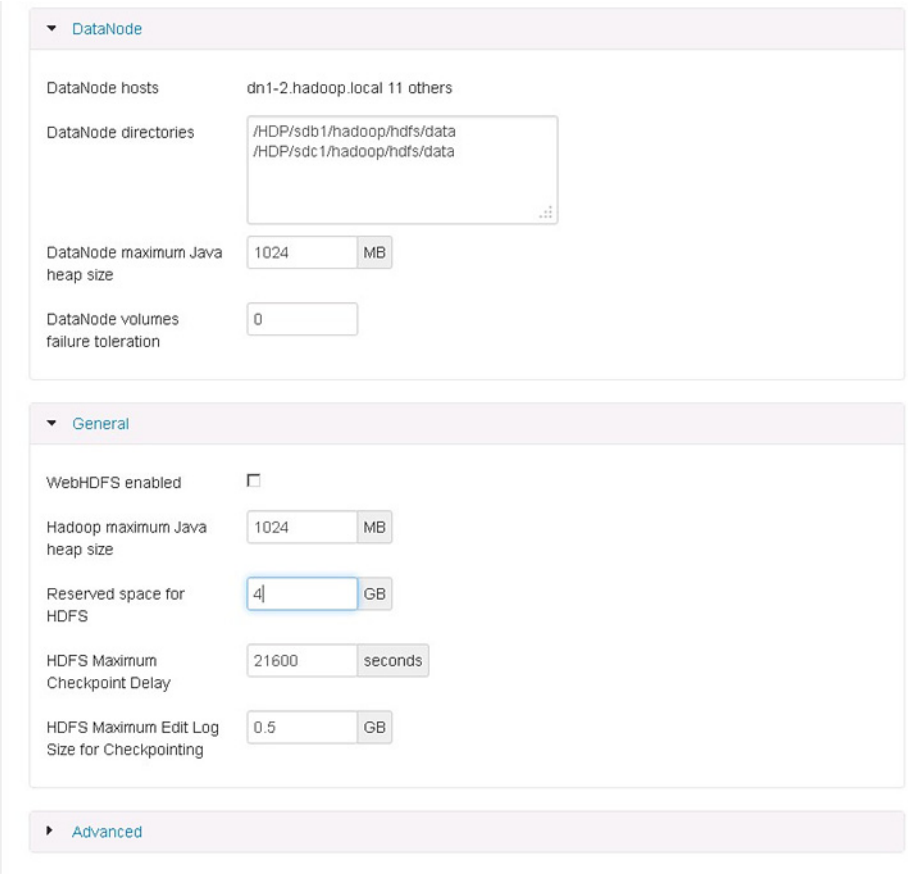


Figure 133 Customize Services - HDFS Configuration Window Part 2



MapReduce

Update the MapReduce configuration as shown in Table 13, and Figure 134 and Figure 135:

Table 13 MapReduce Configurations

Property Name	Value
Job Tracker Maximum Java Heap Size	4GB
Number of Map Slots per Node	24
Number of Reduce Slots per Node	12
Java Options for MapReduce Tasks	4GB
Map-side sort buffer memory	1GB

Figure 134 Customize Services - MapReduce Configuration Window Part 1

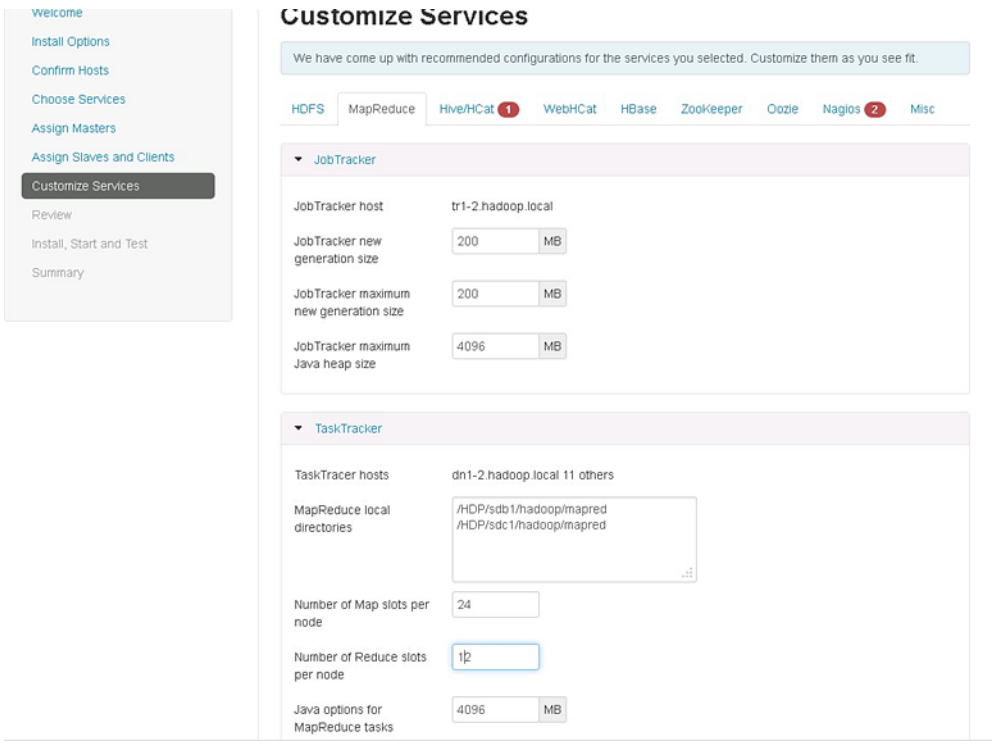
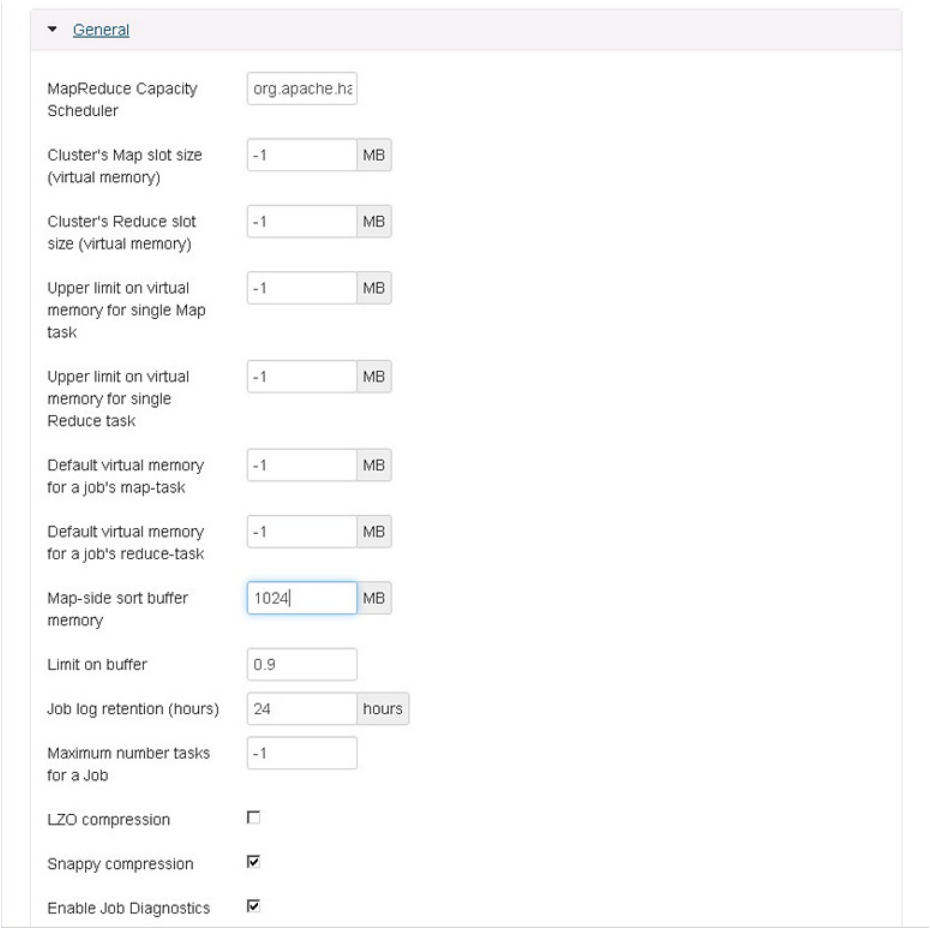


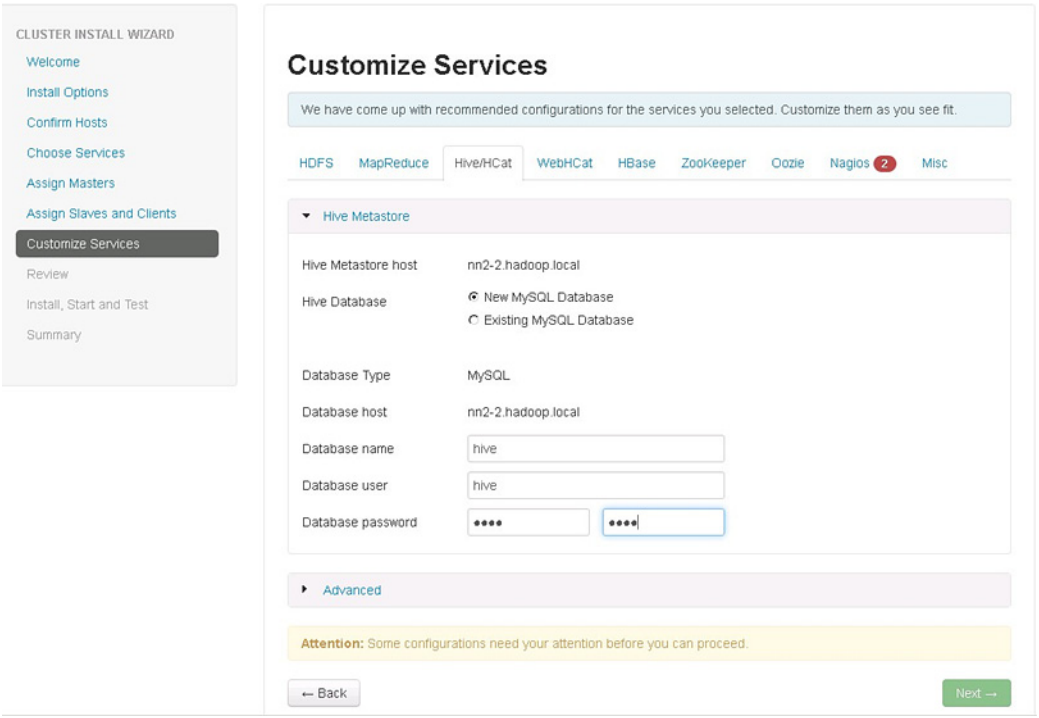
Figure 135 Customize Services - MapReduce Configuration Window Part 2



Hive/HCat

Enter the hive database password as per the organizational policy as shown in Figure 136.

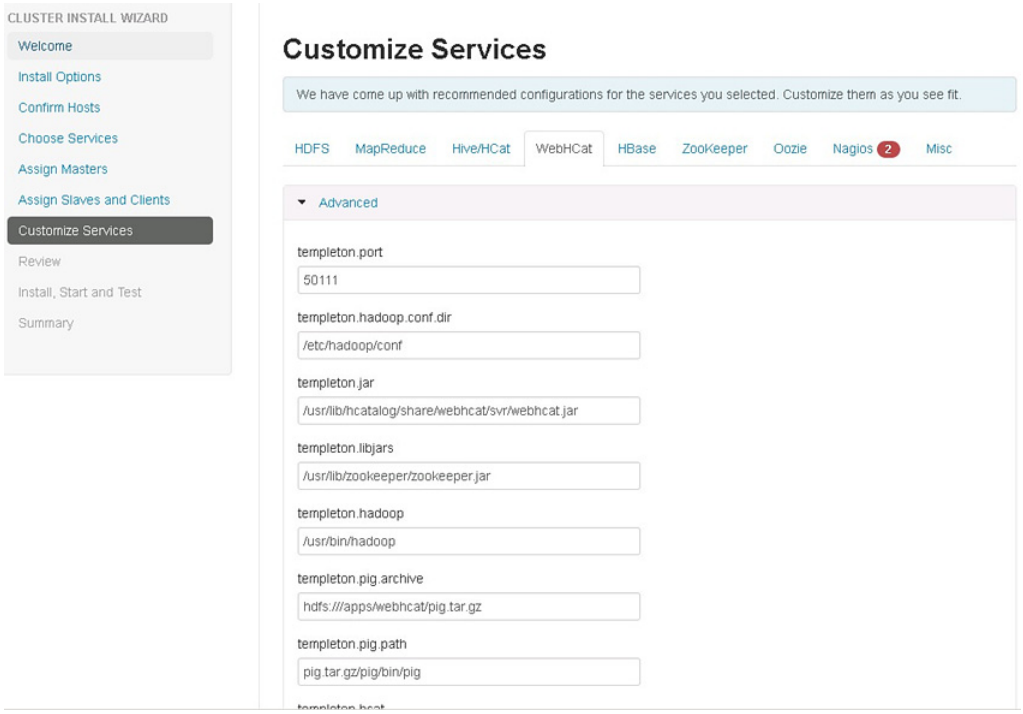
Figure 136 Customize Services - Hive/HCat Window



WebHCat

We can restore the default settings, no changes needed as shown in Figure 137.

Figure 137 Customize Services - WebHcat Configuration



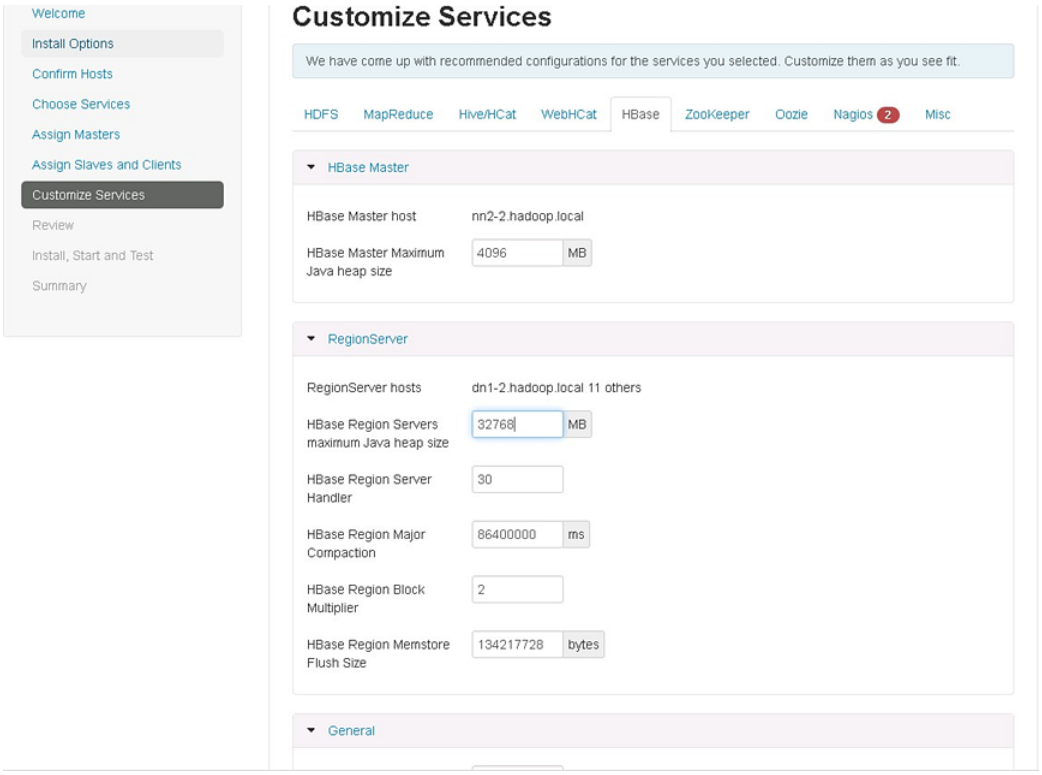
HBase

Update the HBase configurations as shown in Table 14, and Figure 138:

Table 14 HBase Configurations

Property Name	Value
HBase Master Maximum Java Heap Size	4GB
HBase RegionServers Maximum Java Heap Size	32GB

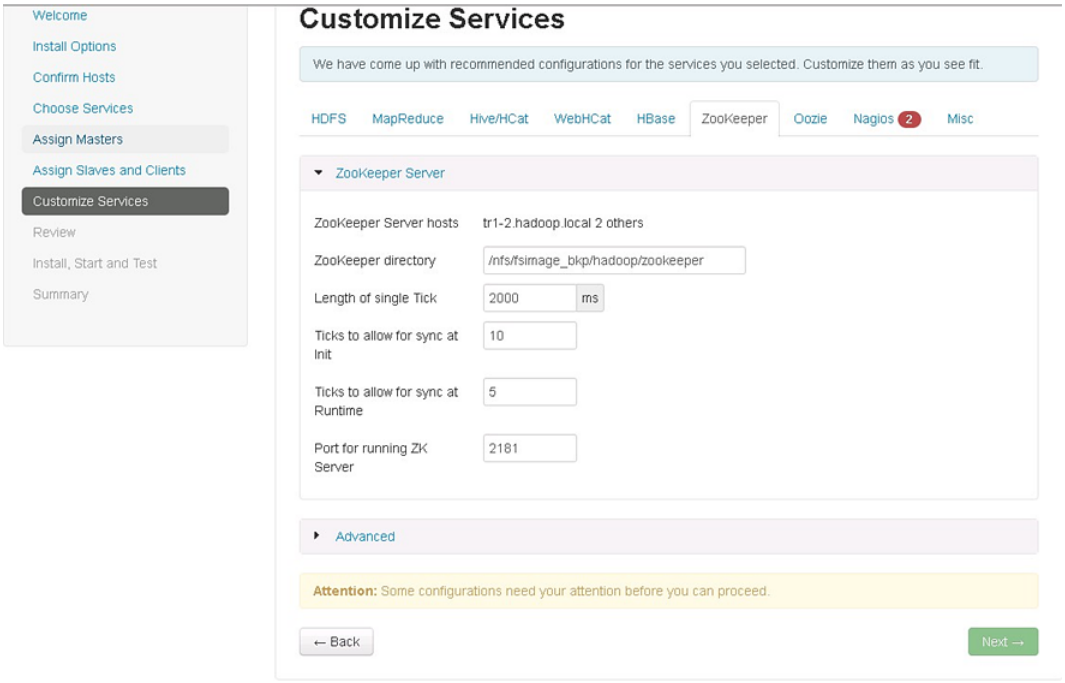
Figure 138 Customize Services - HBase Configuration Window



ZooKeeper

We can restore the default settings in the ZooKeeper window, no changes needed as shown in Figure 137.

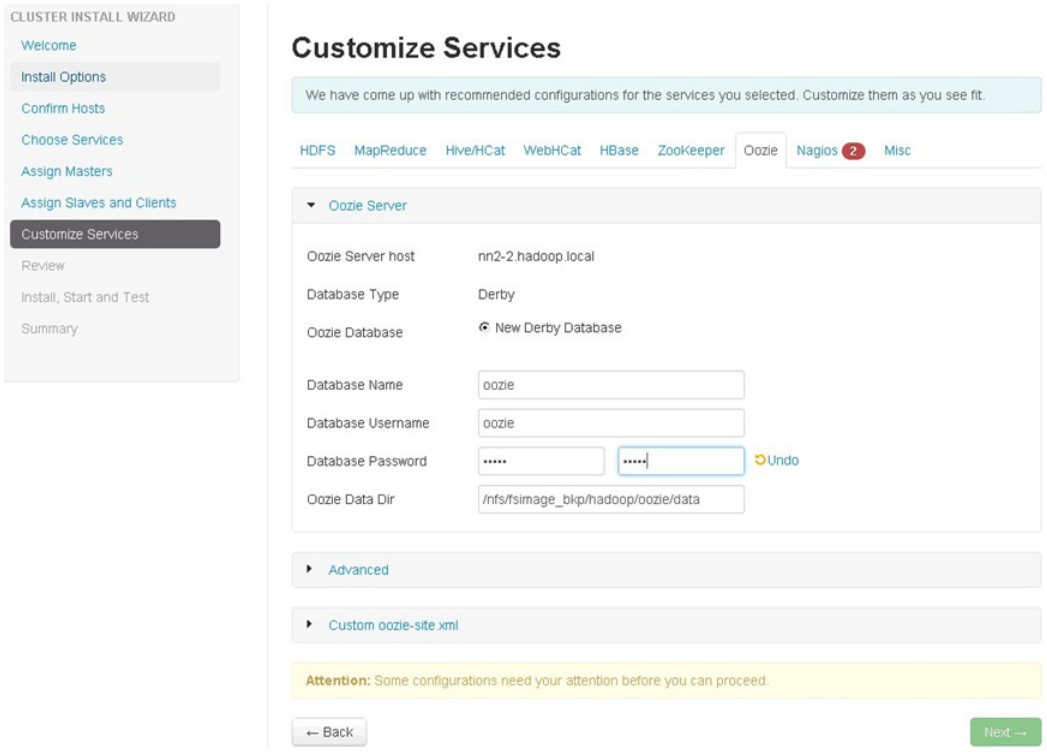
Figure 139 Customize Services - ZooKeeper Window



Oozie

Enter the Oozie database password as per the organizational policy as shown in Figure 140.

Figure 140 Customize Services - Oozie Window



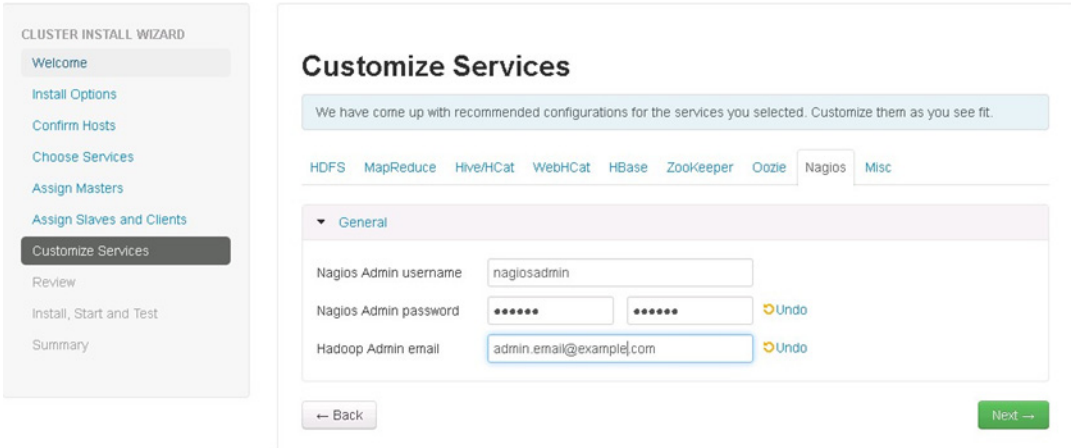
Nagios

Update the Nagios configuration as shown in Figure 141.

Enter the following in the Nagios window:

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

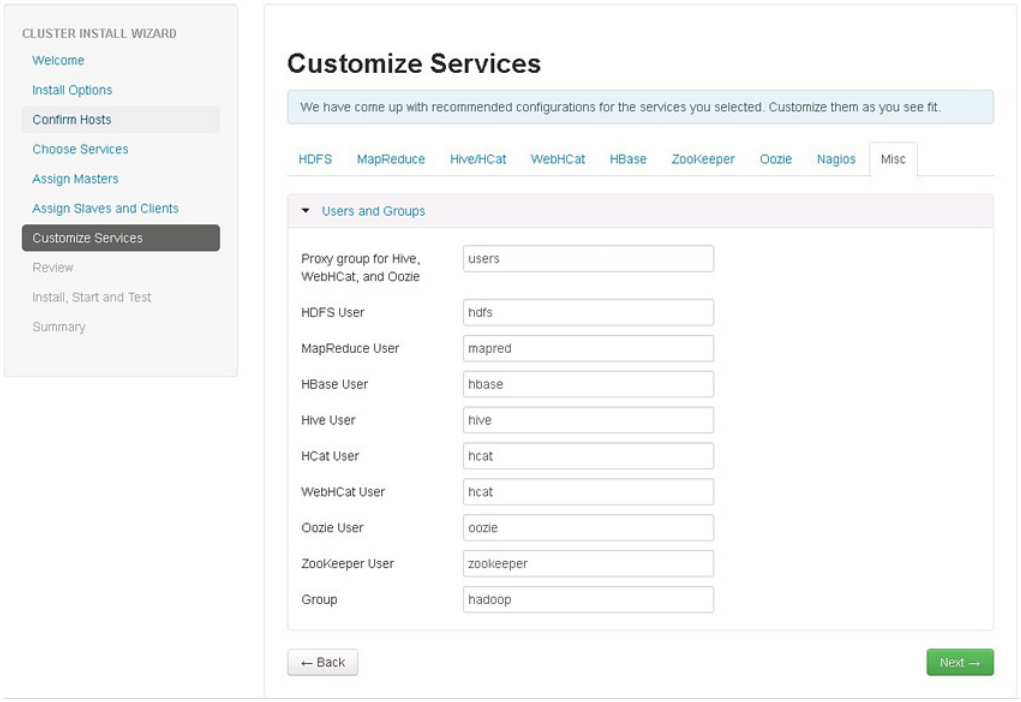
Figure 141 Customize Services - Nagios Window



Misc

We can restore the default settings in the Misc window, no changes needed as shown in Figure 142.

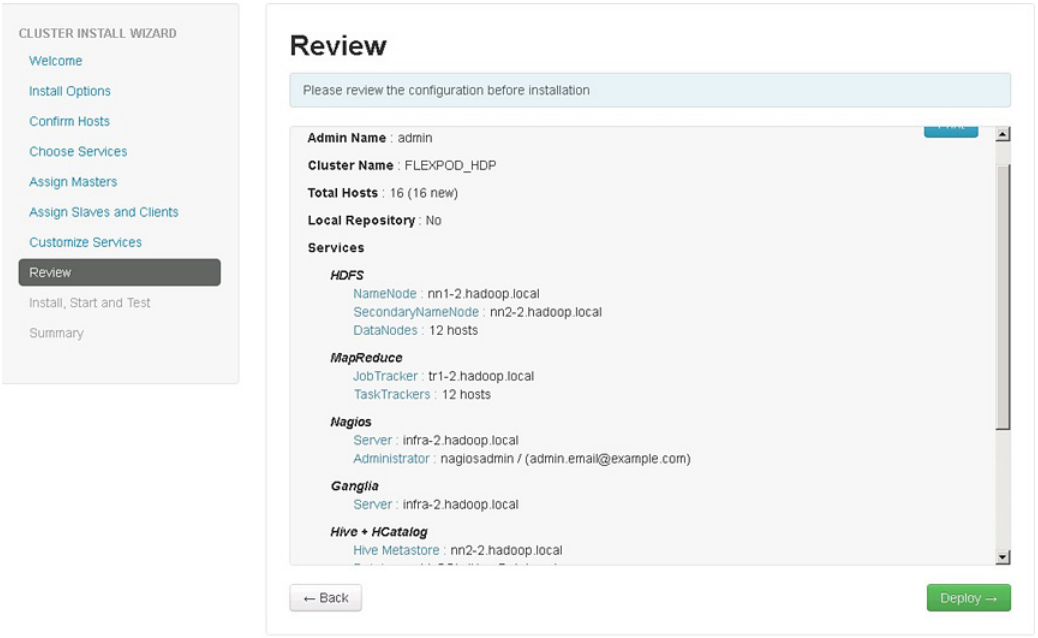
Figure 142 Customize Services - Misc Window



Review

Make sure the Review window shows all the configurations that you have done. Then click **Deploy** as shown in Figure 143. If any changes are to be made, use the left navigation bar to return to the appropriate screen.

Figure 143 Review Window



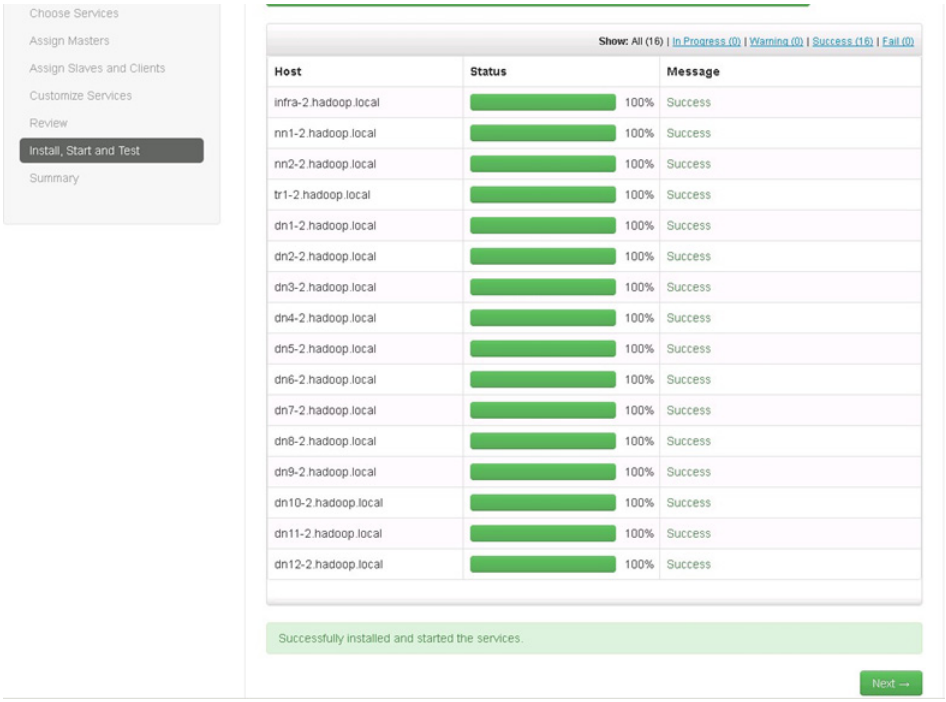
The installation process is shown by the progress indicator as shown in Figure 144. Each component when installed, gets started with a simple test which is run on each of the components. The overall status of the installed components are shown by the progress bar besides every host.

To see the 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, select individual task to see the related log files. Select filter conditions by using the drop-down list. To see a larger version of the log contents, click **Open**. And to copy the contents to the clipboard, click **Copy**.

Depending on the components being installed per host, the entire process may take 30 or more minutes.

Click **Next**, when successfully installed and started the services message at the bottom of window appears as shown in Figure 144.

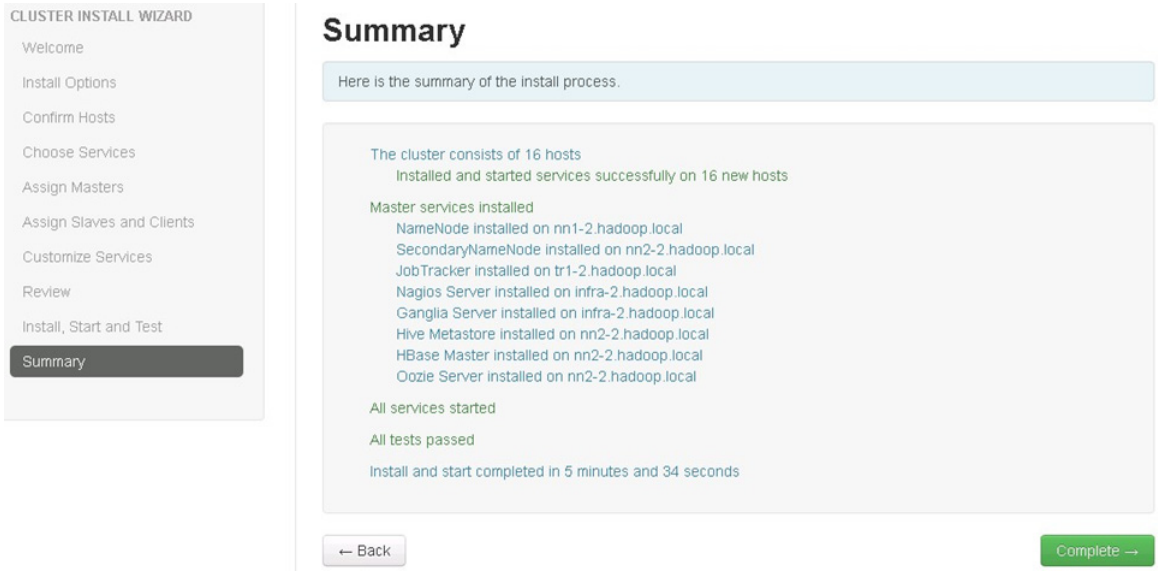
Figure 144 Cluster Install Wizard - Install, Start and Test Window



Summary of Installation Process

The summary page shows the accomplished tasks after the completion of cluster installation.

Figure 145 Cluster Install Wizard- Summary Window



Conclusion

FlexPod Select for Hadoop is an innovative solution that combines technologies from the market leaders to enhance reliability and capacity. The fully redundant fabric architecture with industry-leading namenode resiliency, RAID protection with data replication and hot-swappable spares can significantly lower the risk of failure and application downtime. Leading edge Hadoop management tools provide an analytic stack for big data that is highly reliable, scalable and easier to operate.

The solution addresses today's data-driven environment, in which complex and large data sets need to be processed quickly and efficiently. Seamless data and management integration capabilities co-exist with FlexPod running enterprise applications such as Oracle®, Microsoft®, and SAP®, among many others. Compatibility with traditional FlexPod deployments, that is, the existing resources, can still be used and extended. The solution is offered in a master and an expansion configuration for easy scaling. This is a pre-validated solution that enables quick and easy deployment.

Bill of Materials

The FlexPod Select for Hadoop is offered in a master configuration and an expansion configuration for easy scaling.

Up to 160 servers, 2560 processor cores, and up to 10 petabytes of user storage capacity is supported in one single domain. Applications that need to scale beyond one domain can interconnect several UCS domains using Cisco Nexus Series switches. Scalable to thousands of servers and hundreds of petabytes of data, these domains can be managed from a single pane by using UCS Central in a data center or in remote global locations.

This section provides the hardware and software specifications for deploying the FlexPod Select for Hadoop.

Cisco Bill of Materials

Table 15 provides Cisco BOM for both master rack and expansion rack solutions.

Table 15 Cisco BOM

Hardware / Software	Description	Master Rack Quantity	Expansion Rack Quantity
Cisco UCS C220M3 Servers (UCSC-C220-M3S)	UCS C220 M3 SFF w/o CPU mem HDD PCIe PSU w/ rail kit	16	16
CON-UCW3-C220M3SF	UC PLUS 24X7X4 UCS C220 M3 SFF w/o	48	48
UCS-CPU-E5-2680	2.70 GHz E5-2680 130W 8C/20MB Cache/DDR3 1600MHz	32	32
UCS-MR-1X162RY-A	16GB DDR3-1600-MHz RDIMM/PC3-12800/dual rank/1.35v	256	256

Table 15 Cisco BOM

Hardware / Software	Description	Master Rack Quantity	Expansion Rack Quantity
A03-D600GA2	600GB 6Gb SAS 10K RPM SFF HDD/hot plug/drive sled mounted	32	32
CAB-N5K6A-NA	Power Cord, 200/240V 6A, North America	32	32
UCSC-PSU-650W	650W power supply for C-series rack servers	32	32
UCSC-RAID-ROM55	Embedded SW RAID 0/1/10/5 8 ports SAS/SATA	16	16
UCSC-PCIE-CSC-02	Cisco VIC 1225 Dual Port 10Gb SFP+ CAN	16	16
N20-BBLKD	UCS 2.5 inch HDD blanking panel	96	96
UCSC-HS-C220M3	Heat Sink for UCS C220 M3 Rack Server	32	32
UCSC-PCIF-01F	Full height PCIe filler for C-Series	16	16
UCSC-PCIF-01H	Half height PCIe filler for UCS	16	16
UCSC-RAIL1	Rail Kit for C220 C22 C24 rack servers	16	16
Cisco RP208-30-1P-U-1	Cisco RP208-30-U-1 Single Phase PDU 2x C13 4x C19	2	2
CON-UCW3-RPDUX	UC PLUS 24X7X4 Cisco RP208-30-U-X Single Phase PDU 2x	6	6
RACK-BLANK-001	Filler panels (qty 12) 1U plastic toolless	1	1
Cisco R42610 standard rack (RACK-UCS2)	Cisco R42610 standard rack w/side panels	1	1
Cisco RP208-30-1P-U-2	Cisco RP208-30-U-2 Single Phase PDU 20x C13 4x C19	2	2
CON-UCW3-RPDUX	UC PLUS 24X7X4 Cisco RP208-30-U-X Single Phase PDU 2x	12	12
Cisco UCS 6296UP Fabric Interconnect (UCS-FI-6296UP)	UCS 6296UP 2RU Fabric Int/No PSU/48 UP/ 18p LIC	2	-
UCS-PSU-6296UP-AC	UCS 6296UP Power Supply/100-240VAC	4	-
CON-UCW7-FI6296UP	36X24X7 Support UCS 6296UP 2RU Fabric Int/2 PSU/4 Fans	1	-

Table 15 Cisco BOM

Hardware / Software	Description	Master Rack Quantity	Expansion Rack Quantity
Cisco Catalyst 2960S (WS-C2960S-48FPS-L)	Catalyst 2960S 48 GigE PoE 740W, 4 x SFP LAN Base	1	-
CON-SNT3-2960S4FS	SMARTNET 3YR 8X5XNBD Cat2960S Stk48 GigE PoE 740W,4xSFP Base	1	1
Cisco Nexus 2232PP Fabric Extender (N2K-UCS2232PP-10GE)	N2K 10GE, 2 AC PS, 1 Fan (Std Air), 32x1/10GE+8x10GE	2	2
CON-SNTP-N2232F	Smart Net Services 24X7X4	6	6
SFP-H10GB-CU3M	10GBASE-CU SFP+ Cable 3 Meter	16	16
SFP-H10GB-CU2M	10GBASE-CU SFP+ Cable 2 Meter	16	16
SFP-H10GB-CU2M	10GBASE-CU SFP+ Cable 2 Meter	4	4
SFP-H10GB-CU1M	10GBASE-CU SFP+ Cable 1 Meter	18	18

NetApp Bill of Materials

Table 16 provides NetApp BOM for both master rack and expansion rack solutions.

Table 16 NetApp BOM

Hardware / Software	Description	Master Rack Quantity	Expansion Rack Quantity
NetApp FAS Components			
FAS2220-6X1TB-R6	FAS2220, 6x1TB, base	1	0
X800E-R6	Power cable North America, R6	2	0
X5526A-R6	Rackmount Kit, 4-Post,Universal, R6	1	0
NetApp E-Series Components			
E5400-SYS-R6	E5400, SYS, -R6	3	4
X5526A-R6	E5400A, 6GB Controller	6	8
DE6600-SYS-ENCL-R6	DE6600 system enclosure	3	4
E-X5680A-R6	Enclosure, 4U-60, DE6600, empty, 2PS	3	4
X-54736-00- R6	HIC, E5400, E5500, SAS, 4-Port, 6Gb	6	8

Table 16 NetApp BOM

Hardware / Software	Description	Master Rack Quantity	Expansion Rack Quantity
E-X4021A-10-R6	Disk Drives,10x3TB,7.2k, DE6600	18	24
X-48619-00-R6	Battery, 5400	6	8
Software Licensing			
SW-2220-ONTAP8-P	SW, Data ONTAP® Essentials, 2220-P	1	0
SW-NFS-C	SW, NFS, -C	1	0
SW-CIFS-C	SW, CIFS, -C	1	0
SW-FCP-C	SW, FCP, -C	1	0
SW-ISCSI-C	SW, iSCSI, -C	1	0

Related Information

Table 17 provides information on RHEL specifications for both master rack and expansion rack solutions.

Table 17 Red Hat Enterprise Linux specifications

Red Hat Enterprise Linux	Description	Master Rack Quantity	Expansion Rack Quantity
RHEL-2S-1G-3A	Rhel/2 Socket/1 Guest/3Yr Svcs Required	16	16
CON-ISV1-RH2S1G3A	ISV 24X7 Rhel/2 Socket/1 Guest List Price 3Y	16	16

Table 18 provides information on the other hardware/software specifications required for both master rack and expansion rack solutions.

For information on LSI products and information on how to buy these products, see:

http://www.lsi.com/channel/products/storagecomponents/Pages/LSISAS9207-8e.aspx

Table 18 Hardware/Software specifications

Hardware/Software	Description	Master Rack Quantity	Expansion Rack Quantity
LSI 00118	LSI SAS 9207-8e HBA	12	16
LSI CBL-SFF8088 SAS-20M	2M External Mini SAS SFF-8088(26-pin 4x) to Mini-SAS SFF-8088 (26-pin 4x) Cables	12	16
LSI 00118	Support	12	16

Table 19 provides information on the Hortonworks Data Platform software required for both master rack and expansion rack solutions.

Table 19 Software specifications

Software	Description	Master Rack Quantity	Expansion Rack Quantity
HDP			
NA	Hortonworks Data Platform	16	16