Technical Report

VMware vSphere 6 on NetApp Clustered Data ONTAP

Best Practices Using vSphere Web Client

VMware Technical Marketing, NetApp
April 2016 | TR-4333

Abstract

NetApp® technology enables data centers to extend their virtual infrastructures to include the benefits of advanced storage virtualization. NetApp leads the storage industry with a single information platform that is hardware agnostic, is able to aggregate disparate forms of hardware, and can virtualize access to that storage: in effect, a storage hypervisor. NetApp clustered Data ONTAP® technology is the storage hypervisor: a platform of storage efficiency, VMware integrations, and solutions. This document aligns the server hypervisor with the storage hypervisor by prescribing the best practices for deploying VMware vSphere 6 on clustered Data ONTAP.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Executive Summary

Server virtualization has been a transforming force in IT for the last decade, and VMware vSphere has been at the leading edge of this evolution. The pace of new products announced and updates released has been rapid and significant. VMware has been agnostic to all technology partners, and as a result of this strategy, out-of-the-box VMware products are not configured to work perfectly with any one particular storage vendor. But as is often the case, what's good for one environment is not necessarily good for another. NetApp best practices documents for VMware vSphere describe ways to reduce risk, accelerate implementation, and provide an encompassing suite of technologies to manage applications in your VMware environment, making NetApp storage an ideal choice for VMware.

This is the latest document in a series that focuses on the integration of the most current technologies from both companies.

As of the writing of this document, IT industry customers are migrating from NetApp Data ONTAP 7-Mode to clustered Data ONTAP and transitioning from the VMware vSphere thick client (the "VIC client") to the vSphere Web Client to manage their virtual infrastructures.

The matrix of products and supportability has become differentiated enough now to warrant three separate technical reports for NetApp best practices for VMware vSphere deployments. For your specific deployment, refer to the appropriate technical report:

- TR-3749: NetApp Storage Best Practices for VMware vSphere
- TR-4068: VMware vSphere 5 on NetApp Clustered Data ONTAP Best Practices
- TR-4333: VMware vSphere 5 on NetApp Clustered Data ONTAP Best Practices Using vSphere Web Client

This document serves as both a deployment guide and a desktop reference for managing a VMware vSphere virtual data center on NetApp storage systems running clustered Data ONTAP.

2 Overview

This technical report describes best practices for planning, architecting, deploying, and maintaining server virtualization environments based on VMware vSphere and NetApp clustered Data ONTAP. This publication focuses on the following key concepts:

- General best practices
- Administration
- Data protection
- Networking
- Clustering storage

This technical report can be read from start to finish, but it is meant to be a reference book in which specific topics are accessed based on their relevancy to the situation or task at hand. This publication describes how to use VMware vSphere 6.x with the NetApp flagship product: clustered Data ONTAP. In addition, this document provides the results of test scenarios and real-life scenarios gathered by a select NetApp team that works with various community members and customers whose virtual infrastructures range in size from small to large. The scenarios described are select best practices from these experiences.

Note: This technical report is a living document, and its version identifier is a point-in-time reference that delineates this text from previous versions. A living document is an active collection of information that is updated when best practices are refined, new techniques are discovered, and additional testing is performed. Therefore, the first best practice should go without saying: Make sure you use the most recent version of this technical report.

2.1 Implementing Best Practice

What Best Practices Are

Best practices represent architecture, procedures, and habits that produce the desired results in a solution. Best practices are intended to provide the best balance of desirable features, including performance, reliability, and simplicity.

Best practices are supplemental to other documentation, such as installation and administration guides and support matrixes or hardware compatibility lists. NetApp best practices are condensed from other documentation, lab testing, and extensive field experience by NetApp internal and field engineers and NetApp customers. Best practices often combine procedures and practices from multiple vendors into a solution that considers requirements from multiple product perspectives.

Best practices are often not the only way to solve a problem. There are usually several ways to perform a given task, possibly by using a different tool or different steps. Alternative methods are not necessarily wrong, but they might be more complicated from most users' perspectives, or they might work but not result in the optimal configuration.

Not following best practices does not necessarily result in an unsupported configuration. Supported configurations are defined in hardware compatibility lists and in similar documents from NetApp and NetApp technology partners. Best practices build on hardware compatibility list (HCL)–supported configurations and can even define the specific steps to build a supported configuration. For more information about supported configurations, refer to the NetApp Interoperability Matrix Tool.

How Best Practices Are Designed

Best practices should:

- Agree with other authoritative documentation.
- Represent supported configurations. In other words, be compliant with support matrixes from both NetApp and technology partners, in this case, the VMware Certified Compatibility Guides.
- Consider the majority of customers and their situations.
- Select and recommend the simplest solution or procedure that meets other requirements.
- Use the minimum number of user interfaces to perform tasks.

Examples of implementing these best practices include:

- Preference of GUI over script or CLI in most situations
- Use of task wizards and workflows where available

Golden Rules for Following Best Practices

For success in following best practices, NetApp recommends the following:

- Know the best practices. Read the best practice documents.
- Stay current on best practices as they evolve.
- Ask questions when best practices seem unclear or incomplete, conflict with other documents, or simply don't fit your situation.
- When deviating from best practices recommendations, be sure of your reasons and document the deviation and reasons.
- Provide feedback to the authors of best practices when appropriate.

2.2 Applicability

This document and the discussion and procedures within have been updated specifically for the following products and versions:

- NetApp clustered Data ONTAP 8.3
- VMware vSphere 6.x
- NetApp Virtual Storage Console 6.0 for VMware vSphere Web Client

In most cases, the content of this technical report applies to other versions of these products. Throughout this document, applicability, variations, restrictions, and requirements for other versions are noted. Where earlier versions are listed, the concepts or procedures apply to the versions listed above as well, unless otherwise noted.

Best Practice
For the latest interoperability information, refer to the NetApp Support site and the VMware Support site.

3 VMware vSphere 6 and Clustered Data ONTAP

The following sections describe VMware vSphere 6 and clustered Data ONTAP interoperability and tools.

3.1 VMware vSphere 6.x Points of Integration

Since 2007, NetApp has released a suite of innovative products and technologies to integrate and manage advanced storage features in a VMware vSphere environment. Some of these products and technologies include the following:

- SAN Host Utility Kits for ESX
- NetApp SnapManager® for Virtual Infrastructure (SMVI)
- Rapid Cloning Utilities (RCU)
- NetApp Virtual Storage Console (VSC) plug-in
- Storage Replication Adapter (SRA) plug-in
- VMware vStorage APIs for Array Integration (VAAI) compatibility
- APIs for backup and recovery of vCloud Director environments

The NetApp and VMware development and engineering teams work together so that tightly coupled integrations from both companies, as shown in Table 1 and Table 2, provide valuable benefits to the customer. As with all software, versions change, and features can improve and change rapidly over time.

VMware and NetApp Interoperability Reference

Table 1 and Table 2 define the compatibility between NetApp Data ONTAP 8.3 and VMware vSphere products. Each table details the level of support (fully supported, not supported, or partially supported) between each company's products or features. The acronyms used in these tables are defined as follows:

FS = fully supported

NS = not supported

PS = partially supported

Table 1) NetApp products or features and VMware vSphere licensing interoperability.

NetApp Product or Feature	vSphere 6.x Standard	vSphere 6.x Enterprise	vSphere 6.x Enterprise Plus	vCenter Server 6.x	Other Requirements and Notes
VSC cloning	FS	FS	FS	FS	NetApp FlexClone® technology
VSC backup	FS	FS	FS	FS	NetApp Snapshot® technology and NetApp SnapRestore®
NetApp SnapProtect®	FS	FS	FS	FS	Service Pack 4 provides the ability to perform Network Data Management Protocol (NDMP) dumps of existing Snapshot copies to tape
Snap Creator®	FS	FS	FS	FS	
Thin provisioning	FS	FS	FS	FS	

Table 2) VMware products or features and NetApp Data ONTAP 8.3 protocol interoperability.

VMware Product or Feature	Data ONTAP 7-Mode	Clustered Data ONTAP	NFS	iSCSI	FC/ FCoE	Other Requirements and Notes
vSphere	FS	FS	FS	FS	FS	
vCenter	FS	FS	FS	FS	FS	Integrations and support for vCenter are accomplished with NetApp VSC.
vCenter Site Recovery Manager	FS	FS	FS	FS	FS	NetApp provides an SRA to use with this VMware product; Site Recovery Manager (SRM) 5 uses SRA 2.0 or later, and SRM 4 uses SRA 1.4.3.
vSphere Storage DRS	FS	FS	FS	FS	FS	
vSphere profile-driven storage	FS	FS	FS	FS	FS	
vSphere storage I/O control	FS	FS	FS	FS	FS	
VAAI: SAN	FS	FS	N/A	FS	FS	
Full copy						
Block zero						
HW-assisted locking						

VMware Product or Feature	Data ONTAP 7-Mode	Clustered Data ONTAP	NFS	iSCSI	FC/ FCoE	Other Requirements and Notes
Thin provisioning						
VAAI: NFS	FS	FS	FS	N/A	N/A	
Full file clone						
Reserve space						
VMware Storage vMotion	FS	FS	FS	FS	FS	

3.2 VMware vSphere 6 Licensing

Licensing Requirements

Table 3 lists the licensing requirements for running a VMware vSphere 6 environment on Data ONTAP.

Note: The NetApp complete bundle includes all of the licensing options as well as all of the SnapManager tools.

Table 3) Data ONTAP licensing options.

License	Required or Optional	Description
Base	Required	The base license allows the clustering of NetApp storage controllers.
iSCSI	Optional*	This license enables the iSCSI protocol so that it can be used with vSphere 6.
FC/FCoE	Optional*	This license enables the Fibre Channel (FC) or Fibre Channel over Ethernet (FCoE) protocol so that it can be used with vSphere 6.
UNIX exports (NFS)	Optional*	This license enables the NFS protocol so that it can be used with vSphere 6.
Windows shares (CIFS)	Optional	This license enables CIFS. The CIFS license allows the NetApp storage array to be used as a file server for network shares.
Mirror	Optional	This license enables NetApp SnapMirror® software, which allows the vSphere 6 environment at the primary facility to be mirrored to a remote site. The vSphere 6 environment can then be brought up at that facility.
SnapRestore	Optional	This license enables SnapRestore software, which allows files to be restored from Snapshot copies. This in turn allows the restoration of individual VMs from a Snapshot copy in a vSphere 6 environment.
FlexClone	Optional	This license enables FlexClone software, which allows FlexClone copies to be created. FlexClone copies are the

License	Required or Optional	Description
		base technology for the rapid creation of vSphere 6 VMs.
SnapManager suite	Optional**	This license enables the various types of SnapManager functionality. The SnapManager suite allows the backup and recovery of applications residing in a VM. Examples of such applications are Microsoft Exchange, Microsoft SQL Server, and Microsoft SharePoint.
SnapMirror data protection	Optional	This license enables SnapMirror and makes the cluster an endpoint for a mirror, which allows the creation of a disaster recovery mirror at a remote site to bring up a vSphere 6 environment at that facility.

*One of these three protocol licenses is required for running vSphere 6 in a NetApp environment. The three protocols are supported and viable depending on the needs of each organization. One protocol license of choice comes free with a cluster. The other two protocol licenses are optional.

**The SnapManager suite is most often sold as part of a licensing bundle.

NetApp and VMware vSphere 6 Technology Integration Matrix

Table 4 lists the features and functionalities of NetApp software and their interaction with VMware vSphere 6.

FS = features fully supported

NS = features not supported

PS = partially supported

Table 4) VMware vSphere 6 and NetApp technology integration matrix.

Technology	vSph. 6 Std.	vSph. 6 Ent.	vSph. 6 Ent. Plus	vCtr. Server 6	Clustered Data ONTAP	FC/ FCoE	iSCSI	NFS	Notes
NetApp System Manager	N/A	N/A	N/A	N/A	FS	FS	FS	FS	
NetApp VSC cloning	FS	FS	FS	FS	FS	FS	FS	FS	NetApp FlexClone license required.
NetApp VSC backup	FS	FS	FS	FS	FS	FS	FS	FS	NetApp Snapshot technology and SnapRestore are required.
NetApp SnapProtect	FS	FS	FS	FS	PS	FS	FS	FS	With Service Pack 4, the ability to perform NDMP dumps of existing Snapshot copies to tape is currently

Technology	vSph. 6 Std.	vSph. 6 Ent.	vSph. 6 Ent. Plus	vCtr. Server 6	Clustered Data ONTAP	FC/ FCoE	iSCSI	NFS	Notes
									available.
NetApp Snap Creator	FS	FS	FS	FS	FS	FS	FS	FS	
VMware Storage vMotion	NS	FS	FS	FS	FS	FS	FS	FS	Storage vMotion is not available in vSphere 6 Standard.
vCenter Site Recovery Manager	FS	FS	FS	FS	FS	FS	FS	FS	
VMware Storage I/O Control	NS	NS	FS	FS	FS	FS	FS	FS	This feature is available only in vSphere Enterprise Plus.
VMware Storage DRS	NS	NS	FS	FS	FS	FS	FS	FS	This feature is available only in vSphere Enterprise Plus.
VMware Profile-Driven Storage	NS	NS	FS	FS	FS	FS	FS	FS	This feature is available only in vSphere Enterprise Plus.
Storage APIs for Array Integration, Multipathing	NS	FS	FS	FS	FS	FS	FS	FS	
Thin provisioning	FS	FS	FS	FS	FS	FS	FS	FS	

VMware vSphere 6 on NetApp Technology Enablement Matrix

Table 5 lists the VMware vSphere 6 technologies and products that are enhanced by NetApp technologies and products.

Table 5) VMware vSphere 6 with NetApp technology enablement matrix.

vSphere 6 Technology	Minimum vSphere 6 Version Required	NetApp Technology Enhancements	Other Requirements and Notes
vCenter Site Recovery Manager	Standard	Clustered Data ONTAP 8.1, SnapMirror, SRA, NetApp FlexVol® technology	For clustered Data ONTAP 8.1 use SRA 2.0 or later. For SRM 5.5, use SRA 2.1 or later.
VMware View 5	Standard or desktop	Clustered Data ONTAP	

vSphere 6 Technology	Minimum vSphere 6 Version Required	NetApp Technology Enhancements	Other Requirements and Notes
		8.1, FlexClone, VAAI, SnapMirror, thin provisioning, deduplication, Virtual Storage Tier, space reclamation, FlexVol technology, VSC, NetApp FlexShare® technology	
VMware vCloud Director 6.x	Standard	Clustered Data ONTAP 8.1, FlexClone technology, VAAI, Snap Creator, thin provisioning, deduplication, Virtual Storage Tier, FlexVol technology, space reclamation, FlexShare	
VMware vCenter Operations Management Suite	Standard	Clustered Data ONTAP 8.1, FlexVol technology, FlexShare	
VMware Storage vMotion	Standard	Clustered Data ONTAP 8.1, FlexVol technology, VAAI	
VMware Storage I/O Control	Enterprise Plus	Clustered Data ONTAP 8.1, Virtual ST, NetApp Flash Cache™, FlexShare	
VMware Storage DRS	Enterprise Plus	Clustered Data ONTAP 8.1, FlexVol technology	
VMware profile-driven storage	Enterprise Plus	Clustered Data ONTAP 8.1, Flash Cache, NetApp VAAI, FlexVol technology	
VMware thin provisioning	Enterprise Plus	Clustered Data ONTAP 8.1, FlexVol technology	NetApp thin provisioning and deduplication work in conjunction with VMware thin provisioning.
Storage APIs for array integration and multipathing	Enterprise Plus	Clustered Data ONTAP 8.1, FlexVol technology	

Note: Table 5 lists NetApp technologies that can further enhance VMware deployments. Table 5 is not a list of the prerequisites.

Note: VAAI is one of the more important enabling technologies of both companies. For information about VAAI compatibility, refer to Frequently Asked Questions for vStorage APIs for Array Integration and NetApp Knowledgebase article 3013572.

3.3 VMware vSphere 6.x on NetApp Clustered Data ONTAP Management Interfaces

There are two primary ways to manipulate NetApp storage with vSphere 6.x:

- Virtual Storage Console
- OnCommand® System Manager

Virtual Storage Console

VSC 6.0 is a plug-in that provides integrated, comprehensive storage management for the VMware infrastructure within the single interface of the VMware vSphere Web Client, the replacement for the C#-based VMware vSphere Client. The VSC allows VMware administrators to perform discovery, health monitoring, capacity management, provisioning, cloning, optimization, backups, restores, and disaster recovery without affecting any policies that a storage administrator might have created. Functionalities such as deduplication and thin provisioning of datastores, rapid cloning of VMs, near-instant VM backups, granular VM restores, and integration into storage disaster recovery solutions are at the VMware administrator's fingertips without the intervention of storage administrators.

To learn more about VSC, refer to the NetApp Virtual Storage Console datasheet. VSC is freely available to all customers through the NetApp Support site.

OnCommand System Manager

OnCommand System Manager is a management tool that provides easy configuration and ongoing management for NetApp storage through a simple-to-use web-based interface. Clusters, storage virtual machines (SVMs, formerly known as Vservers), and other clustered Data ONTAP resources can be managed with OnCommand System Manager, as well as Data ONTAP operating in 7-Mode and legacy 7G systems.

OnCommand System Manager has built-in integration with VMware virtual storage management and allows the following:

- Wizard-driven setup of aggregates, volumes, LUNs, shares, and exports
- NFS, iSCSI, and FC configuration
- Advanced storage feature configuration for SnapMirror, NetApp SyncMirror®, NetApp SnapLock®, and SVMs
- Clustered Data ONTAP 8.x management

To learn more about OnCommand System Manager, refer to the OnCommand System Center datasheet. OnCommand System Manager is freely available to all customers through the NetApp Support site and supports versions earlier than 8.3. For clustered Data ONTAP 8.3 and later, OnCommand System Manager is bundled with Data ONTAP.

Other Management Interfaces

Many user interfaces, environments, development tools, and languages are available to manage vSphere and NetApp storage. Table 6 lists these user interfaces.

Table 6) User interfaces for managing vSphere and NetApp storage.

UI/Tool	Type	Target	Recommended	Notes
ESXi shell	CLI	Single vSphere host	Tech support	Accessed through SSH or physical console. Disabled by default. Manages a single vSphere host only.

UI/Tool	Type	Target	Recommended	Notes
vSphere CLI (vCLI)	CLI	vSphere host or vCenter	Advanced	Installed in Windows or Linux to remotely manage vSphere hosts and vCenter.
vSphere Management Assistant (vMA)	CLI	vSphere host or vCenter	Advanced	vSphere CLI preinstalled in a Linux VM that is installed as an appliance.
vSphere SDK	SDK	vCenter	Developer	Used with Perl and other languages to develop custom tools and scripts.
PowerCLI	Object-oriented CLI	vCenter	Advanced/ developer	Object-oriented scripting and command line snap-in for Windows PowerShell. Can be used with NetApp Data ONTAP PowerShell Toolkit and VSC cmdlets.
OnCommand Unified Manager	GUI	Multiple NetApp controllers		Installs on a Windows, Linux, or Solaris server and can be accessed through a web browser or through the NetApp Management Console, which is installed on a Windows workstation.
Data ONTAP PowerShell Toolkit	Object-oriented CLI	Multiple NetApp controllers	Advanced/ developer	Object-oriented scripting and command line snap-in for Windows PowerShell.
NetApp Manageability SDK	SDK	Multiple NetApp controllers	Developer	Formerly called NetApp Manage ONTAP®. Supports C and C++, Java, Perl, C#, VB.NET, and PowerShell.
OnCommand Workflow Automation (WFA)	Orchestration	vSphere and NetApp storage		A software solution that enables you to create storage workflows and automate storage management tasks such as provisioning, migrating, decommissioning, and cloning storage. WFA enables you to create simple and complex workflows in a short time for virtualized and cloud environments. You can use WFA to integrate storage workflows with your existing IT processes and align with NetApp best practices. For more information, refer to the NetApp Support site.

Although some users might prefer or have reasons to use other interfaces and tools, NetApp recommends using the VMware vSphere Web Client with the NetApp VSC plug-in installed in vCenter for all supported functionalities.

4 Clustered Data ONTAP Concepts

A good introduction to clustered Data ONTAP is available in TR-3982: NetApp Clustered Data ONTAP 8.3 and 8.2.x: An Introduction. The following section discusses clustered Data ONTAP concepts in terms that a vSphere administrator should find familiar.

4.1 VMware vSphere 6 and Storage Virtual Machines

The term cluster is used by many IT vendors to describe nodes providing similar resources or services that are federated to some degree. VMware vSphere includes clustering capabilities for high availability (VMware HA, which provides VM-level failover) and for resource sharing and load balancing (Distributed Resource Scheduler [DRS] and Storage DRS, which manage shares, limits, and reservations; can move virtual machines [VMs] between servers; and can be used for storage load balancing).

VMware refers to a cluster as a set of vSphere hosts that are grouped to provide an aggregated set of resources. NetApp defines a cluster as consisting of one or more nodes that are interconnected and managed as a single system. For the sake of clarity, the NetApp best practices documentation uses the terminology vSphere or ESX/ESXi cluster and NetApp cluster to distinguish between the two uses of the term.

A Data ONTAP SVM is in some ways similar to a VM in vSphere, but they also have some fundamental differences. Both the SVM and the VM are virtual entities that consume an allocation of the following four basic resources of VM-like objects:

- Processor
- Memory
- Network
- Storage

Although either a VM or an SVM can technically run with only CPU and memory, they are usually only accessible and useful when they have attached network and storage resources. These resources are provided from a pool of resources owned by, and accessed through, the hypervisor in the case of vSphere, and by the nodes of the NetApp storage cluster in the case of an SVM.

The basic concepts of networking are similar between vSphere and SVMs in many respects. A VM has usually one or more virtual Ethernet adapters that connect through port groups on virtual switches to one or more physical network interface cards (NICs) on the physical vSphere host. Although clustered Data ONTAP does not have the concept of virtual switches, the logical interfaces (LIFs) of SVMs connect to physical interfaces or interface groups (ifgrps) on the nodes of the NetApp cluster.

vSphere clusters have administrative network usage such as management, server-to-server migration connections for vMotion and storage vMotion, VM fault tolerance, and high availability (HA). The virtual switches and physical NICs can be shared between all of these functions as well as with VM port groups provided that adequate bandwidth is available. In practice, administrators often separate different network functions at least by virtual local area network (VLAN) or subnet for security, reliability, and performance reasons. In a Data ONTAP cluster, some administrative functions can share network ports with each other and with data traffic, and some cannot. In particular, NICs used for the cluster interconnect cannot be shared with any other traffic.

A VM usually has one or more virtual disks. Most commonly, a virtual disk is a file within a datastore that is accessible to the vSphere host, but it can also be a raw logical unit number (LUN) passed through ESX to the VM. An SVM has one or more flexible volumes that live in aggregates of disks that are owned by the nodes in the NetApp cluster. According to the role, there are three types of SVMs: node, administration, and cluster.

One fundamental difference between vSphere clusters and NetApp clusters is that a single instance of a VM can only consume resources available from a single vSphere host. With clustered Data ONTAP, an SVM can use resources on multiple nodes in a cluster.

Another fundamental difference between the two is that, in vSphere, specific processor and memory allocations are made to the VM in terms of the number of CPUs and amount of memory seen by the VM and of resource-sharing parameters for DRS, such as shares, limits, and reservations. With SVMs, the allocation of CPU and memory is controlled entirely by the cluster, although some prioritization is possible.

VMs run a separate instance of an operating system, usually different from that of the hypervisor. An SVM appears to run the same version of Data ONTAP as the node on which it executes, but the SVM is not actually a separate instance. An SVM is simply a data structure in memory on each node in the NetApp cluster. Table 7 compares vSphere and clustered Data ONTAP functions.

Table 7) vSphere and clustered Data ONTAP comparison.

Function or Property	vSphere	Clustered Data ONTAP
Processor	vSphere admin controls how many CPUs are seen by the VM and the ratio of shares of CPU time.	No admin configuration of CPUs per SVM.
Memory	vSphere admin allocates the amount of memory as well as shares, limits, and reservations.	No admin configuration of memory per SVM.
Network: virtual	Virtual Ethernet ports connect VMs to port groups on virtual switches.	LIFs connect SVMs to physical ports, interface groups, or VLAN interfaces.
Network: physical	Physical ports and properties such as NIC teaming, speed, duplex, flow control, and VLANs are managed at the hypervisor level. VMs do not need to configure and manage these objects and parameters.	Physical ports and properties such as NIC teaming (interface groups), speed, duplex, flow control, and VLANs are managed on the nodes. LIFs do not need to configure and manage these objects and parameters.
Storage	VMs are presented virtual disks (VMDKs), which are files in a file system or pass-through LUNs.	SVMs have one or more FlexVol volumes contained in an aggregate.
Operating system	Any x86 operating system is a separate unrelated instance from the hypervisor.	An SVM is a memory construct within nodes running clustered Data ONTAP.
Execution location	A VM runs on one ESX server.	An SVM spans multiple nodes.
Resource usage	A VM consumes resources accessible on a single ESX server.	Network and storage resources on any node in the NetApp cluster can be allocated to any SVM.

SVM Delegation

After they are created, SVMs can be managed either by the cluster administrator using the `admin` login or, if management is delegated, by the SVM administrator using the default `vsadmin` login. When System Manager is used, during the initial creation of the SVM, three topics are covered as part of the delegation process:

- Creation of SVM administrator login (`vsadmin`)
- Selection of aggregates in which the SVM administrator can create volumes for the SVM
- Creation of a separate SVM management LIF, unless management is performed through a data port

Best Practice
Because IP datastores (iSCSI and NFS) best practices recommend a private, nonroutable network between ESXi and the NetApp cluster and because FC and FCoE do not carry IP traffic, the best practice for SVMs used with VMware is to create a separate management LIF on a management network.

An additional login is necessary for role-based access control (RBAC) for the VSC. This login does not require SSH or HTTP access, but rather API access, referred to as `ontapi`. Details on the specific APIs required and on how to create a user for VSC are covered in the “VSC 5.0: RBAC” section in chapter 9.

5 vSphere Components

5.1 VMware vCenter 6.x

VMware vCenter is a centralized management framework used to manage a virtual data center running VMware virtual machines and all other layers of the environment, such as storage, networking, and granular user access control. vCenter can be run on any physical or virtual Windows 64-bit machine or on the Linux-based, prepackaged vCenter Server Appliance.

vCenter has an extensible API framework that allows plug-ins based on Flex and Java to be deployed and directly integrated within the same management interface. An example of such a plug-in is NetApp VSC.

For the most part, vCenter is the management interface of choice for managing the smallest to the largest virtual infrastructure deployments that are built on VMware technology.

For more information about vCenter, refer to the VMware vSphere Documentation page.

5.2 VMware vCenter 6.x Appliance

Traditionally, vCenter has been a binary installable package based on Windows, but with the release of the vSphere 5.0 suite, VMware has added the VMware vCenter Server Appliance (vCSA).

The vCenter Appliance is a prepackaged, Linux-based Open Virtualization Format (OVF) template that can be deployed through the vSphere Web Client (if another instance of vCenter is already running) or the C#-based vSphere Client (connecting to a standalone ESXi server). Configuration is handled through a management web interface to get services started and the back-end databases connected. After that, connection to the vCenter Appliance is established through the vSphere Web Client in the same manner as with a traditional Windows-based vCenter instance.

After the vSphere Web Client is connected to the vCenter Appliance, it is business as usual. One difference to note, however, is that any plug-ins that are typically installed on the same Windows server as vCenter, such as the NetApp VSC, must have their own standalone instance of Windows on which to run when the vCenter Appliance is used.

Note: It is not possible to add plug-ins to the vCenter Appliance directly. Connectivity must be established between vCenter and the server in which the plug-in was installed.

For more information about the vCenter Appliance, refer to the VMware vSphere Documentation page.

5.3 VMware vCenter 6.x Deployment Procedures

Table 8 describes VMware vCenter prerequisites.

Table 8) VMware vCenter 6.x prerequisites.

Description
The system meets the VMware vCenter hardware and software requirements. These requirements are listed on page 17 of the vSphere Installation and Setup guide.
The mode of installation for vCenter has been determined. vCenter can be installed: • On a physical machine • On a VM running on an ESXi host

Description
• On a VMware vCenter Server Appliance **Note:** A vCenter Server Appliance is a VM that is preconfigured with vCenter Server.
For migration and failover of the vCenter VM to be supported with vMotion and VMware HA, the vCenter VM and its database must be installed on shared storage. Two methods can be used to install vCenter on shared storage: • The shared storage can be provisioned and mounted to the ESXi server before vCenter is installed. • vCenter can be installed on a local datastore and then migrated to shared storage after the NetApp VSC is installed. The VSC is used to provision a shared datastore.

Installing and configuring vCenter Server include the steps and considerations summarized in this section. For detailed instructions on how to install vCenter, refer to the vSphere Installation and Setup guide, starting in chapter 3.

Installing vCenter in a Virtual Machine

Installing vCenter in a VM has several advantages over installing it on a physical server in terms of the reliability and uptime that can be achieved:

- vSphere HA can be used to provide high availability to the vCenter Server. For more information about vSphere HA, refer to the VMware vSphere High Availability page.
- Server maintenance can be performed by migrating the vCenter VM to another host. For more information about vSphere vMotion, refer to the VMware vSphere vMotion page.
- Snapshot copies of the vCenter VM can be created by using VSC for backup and recovery purposes. For more information about VSC, refer to NetApp Virtual Storage Console for VMware vSphere.

For guidance on how to install vCenter 6.x, refer to the vSphere Installation and Setup guide.

vCenter Installation Options

Different procedures can be used to install vCenter Server. Before you begin, carefully review the installation options presented in chapter 4 of the vSphere Installation and Setup guide.

Note: After you have the information and detailed steps for installing vCenter Server, download the `.zip` file for vCenter Server from the VMware Support and Downloads page.

Database Considerations

vCenter Server has several options for the database that is required for storing all configuration and management data in the vSphere environment.

The Windows version of vCenter Server includes Microsoft SQL Server 2008 R2 Express. This database has very limited capabilities and should be used only in small, short-lived lab environments that have a maximum of 5 ESX or ESXi hosts and 50 VMs.

Larger labs and production environments must use an external database. vCenter Server supports various versions of Microsoft SQL Server, Oracle Database, and IBM DB2. For more information about supported databases, refer to VMware Product Interoperability Matrixes.

Note: Install and configure the external database before installing vCenter Server. For specific instructions on how to install an external database, refer to “Configure Oracle Databases” in chapter 3 of the vSphere Installation and Setup guide.

After the vCenter 6.x Installation

After the vCenter installation is complete, access the vCenter Welcome page by typing the IP address of the vCenter Server machine or by typing `localhost` on a browser that is on the same server in which vCenter Server is installed. Use the vCenter Single Sign-On user ID and password to log in to vCenter Server. From the Welcome page, download the vSphere Client or click the link to log in to the vSphere Web Client. It is useful to create a browser shortcut or bookmark directly to the vSphere Web Client URL.

5.4 VMware vCenter 6.x Appliance Deployment Procedures

Table 9 describes VMware vCenter appliance prerequisites.

Table 9) VMware vCenter 6.x appliance prerequisites.

Description
The host machine meets the requirements for the VMware vCenter Server Appliance (vCSA). These requirements are listed in the vSphere Installation and Setup guide.
The hosts are running ESX version 4.x or ESXi version 4.x or later.
The clocks of all machines on the vSphere network have been synchronized.
For migration and failover of vCSA to be supported with vMotion and VMware high availability (HA), vCSA and its database must be installed on shared storage. Two methods can be used to install vCSA on shared storage: • The shared storage can be provisioned and mounted to the ESXi server before vCSA is installed. • vCSA can be installed on a local datastore and then migrated to shared storage after NetApp VSC is installed. VSC is used to provision a shared datastore.

vCenter Server can be installed on a physical server or on a VM, or it can be deployed as a vCSA. vCSA is a preconfigured, Linux-based VM that has been optimized specifically for vCenter Server. For guidance on how to install VMware vCSA 6.x, refer to the vSphere Installation and Setup guide.

For information about how to deploy Open Virtualization Archive (OVA) files and Open Virtualization Format (OVF) templates, refer to the vCenter Server and Host Management guide.

Database Considerations

vCSA has three options for the database that is required for storing all configuration and management data in the vSphere environment:

- vCSA 5.0, which includes a light version of IBM DB2.
- vCSA 5.0.1 and 5.1, which include a version of PostgreSQL, called vPostgres, that is specific to VMware. With vSphere 5.5, the vPostgres database can support up to 400 ESXi hosts and 4,000 VMs.
- vCSA 6.0, which includes vPostgres and can support up to 1,000 ESXi hosts and 10,000 VMs.

The only external database that vCSA supports is Oracle Database. For more information about supported databases, refer to VMware Product Interoperability Matrixes.

Note: Install and configure the external database before installing vCSA. For instructions on how to install an external database, refer to “Configure Oracle Databases” in the vSphere Installation and Setup guide.

Install VMware vCenter 6.x Appliance

To install the vCSA, complete the following steps:

1. From the Download VMware vSphere page, download the OVA file or the OVF template.
2. Deploy the OVA file or the OVF and VMDK files as an OVF template from the vSphere Client or from the vSphere Web Client.
3. Power on the vCSA.
4. Log in to the Welcome page and accept the license agreement.
5. Select the desired configuration for the setup. Three options are available:
   - Default setup
   - Upload a configuration file
   - Set a custom configuration
6. Follow the prompts to complete the setup wizard.

After vCenter 6.x Appliance Installation

The vCSA can be accessed by IP address or by fully qualified domain name (FQDN). VMware recommends using FQDN because the IP address can change (for example, if a DHCP server is being used).

Note: The default root password in the vCenter Server Appliance is set to expire automatically 90 days after the installation. This is a common issue, as noted in VMware KB 2069041. NetApp recommends that you change your root password as soon as you are finished deploying vCenter.

For more information about how to configure the vCSA, refer to the vCenter Server and Host Management guide.

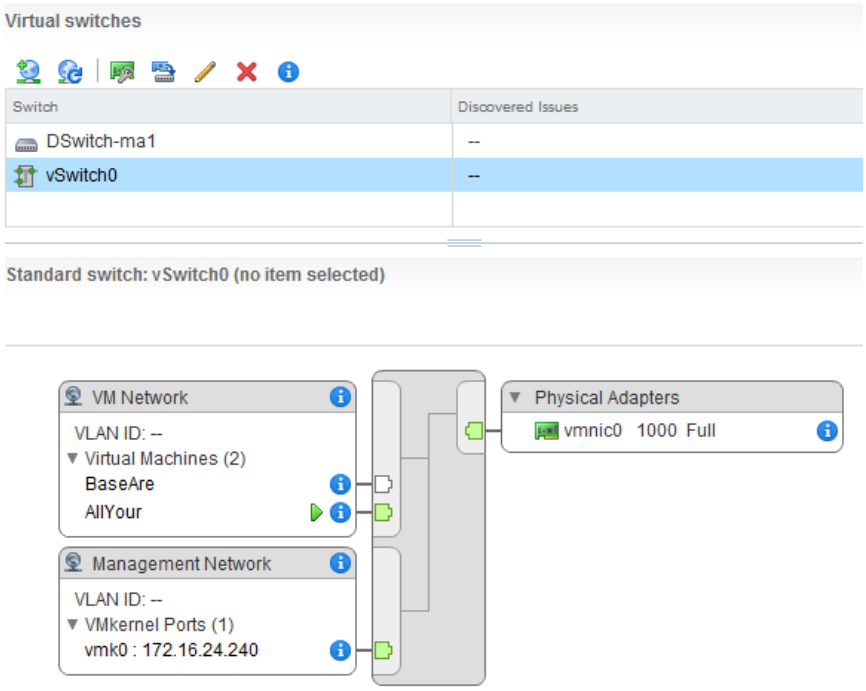
6 Storage Networking

6.1 VMware vSphere 6 and Clustered Data ONTAP Basic Networking Concepts

VMware vSphere provides virtual networking technology that connects virtual entities, such as VMs and the ESXi hypervisor virtual ports known as VMkernel ports, to each other and to the physical NICs connected to physical networks. The core of virtual networking is the virtual switch, shown in Figure 1. VMs are connected in sets through VM port groups. VMkernel ports are used for tasks such as the following:

- Management of the ESXi host (vSphere Client, vCenter connections, and SSH)
- ESXi NFS, software iSCSI, and software FCoE storage networking
- vMotion traffic
- Fault tolerance logging

Figure 1) vSphere standard switch.



The following components are required to use IP storage (NFS and/or iSCSI) with ESXi:

- A virtual switch with at least one physical adapter (vmnic) connected to the network to which the NetApp storage is connected
- A VMkernel port with an IP address and a subnet mask

Additional settings that may be used for storage networking include the following:

- VLAN
- Maximum transmission unit (MTU) (used for jumbo frames)
- Default gateway, if resources are to be accessed on subnets other than those to which the server is directly connected
- NIC teaming policy settings
- iSCSI port binding
- Flow control

Three types of virtual switches are available with vSphere:

- vSphere standard switch
- vSphere distributed switch
- Third-party virtual switch such as Cisco Nexus 1000-V or IBM System Networking Distributed Virtual Switch 5000V

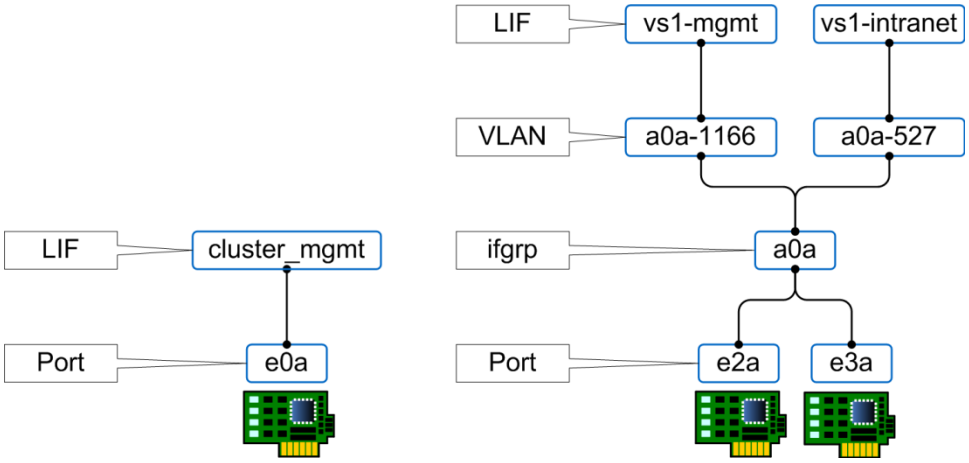
Note: Additional third-party virtual switches are available from other VMware partners. Although any of these virtual switches can be used for storage networking, this document focuses on the native vSphere standard and distributed switches.

Clustered Data ONTAP Networking Concepts

The physical interfaces on a node are referred to as ports. IP addresses are assigned to LIFs. LIFs are logically connected to a port in much the same way that VM virtual network adapters and VMkernel ports are connected to physical adapters, but without the constructs of virtual switches and port groups.

As Figure 2 shows, physical ports may be grouped into interface groups (ifgrps). VLANs can be created on top of physical ports or interface groups. LIFs may be associated with a port, an interface group, or a VLAN.

Figure 2) Ports and LIFs, simple and advanced examples.

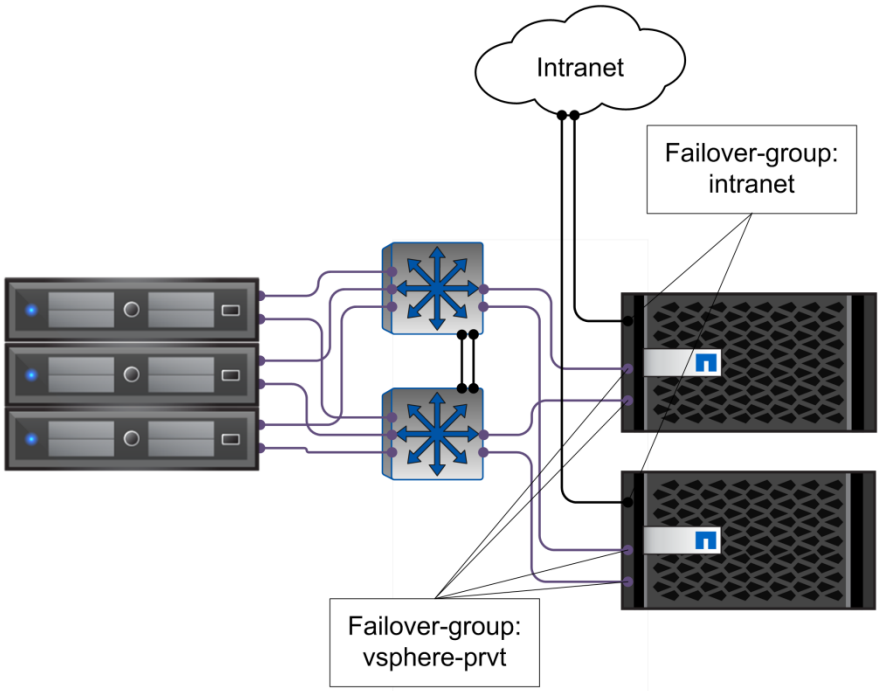


LIFs and ports have roles, akin to the difference between a VM port group and the VMkernel ports used for management, storage, vMotion, and fault tolerance. Roles include cluster or node management, cluster (for traffic between nodes), intercluster (for SnapMirror replication), and data.

From a solution perspective, data LIFs are further classified by how they are used by the servers and applications and by whether they are on private nonroutable networks, on corporate internal routable networks, or on a demilitarized zone (DMZ). The NetApp cluster connects to these various networks through data ports; the data LIFs must use specific sets of ports on each node for traffic to be properly routed.

Some LIFs, such as the cluster management LIF and the data LIFs for NFS and CIFS, can fail over between ports within the same node or between nodes so that if a cable is unplugged or a node fails, traffic continues to flow without interruption. Failover groups, such as those shown in Figure 3, are used to control the ports to which a LIF may fail over. If failover groups are not set up or are set up incorrectly, LIFs might fail over to a port on a wrong network, causing connectivity to be lost.

Figure 3) Failover groups.



Best Practices
• Make all Ethernet data ports, interface groups, and VLANs members of an appropriate failover group. • Associate all NFS, CIFS, and management LIFs with the appropriate failover group. • To keep network connectivity as simple as possible, use the same port on each node for the same purpose (assuming similar nodes with similar ports).

10GbE Support

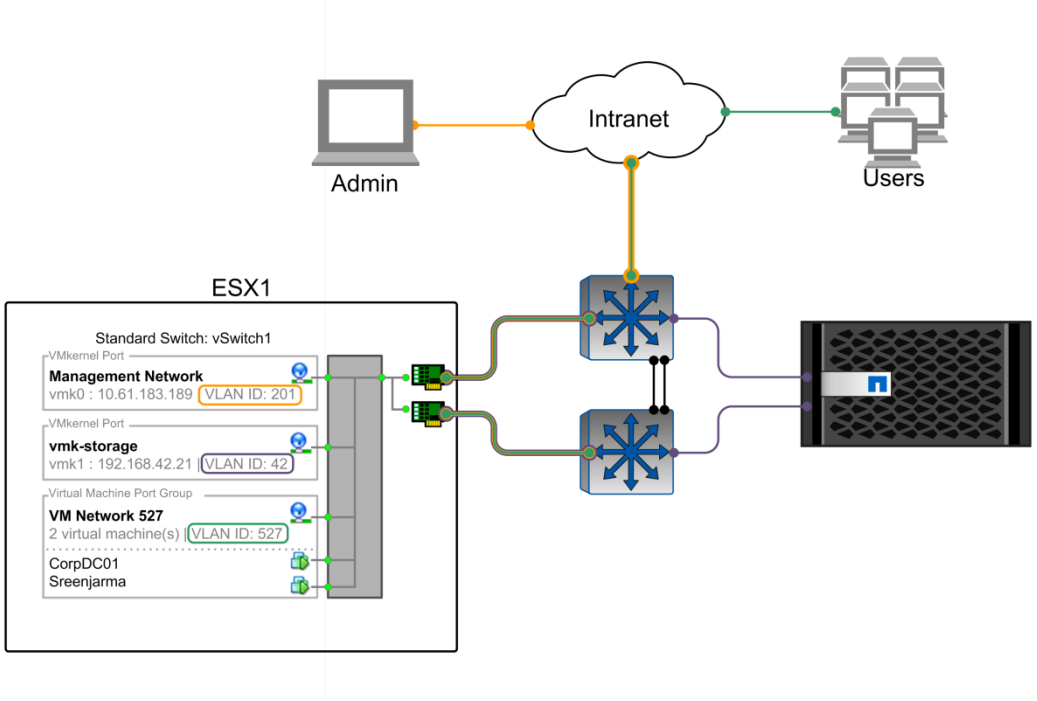
NetApp Data ONTAP and VMware ESX and ESXi 4 and later releases provide support for 10GbE. An advantage of 10GbE is the ability to reduce the number of network ports in the infrastructure, especially for blade servers. To verify support for any specific hardware and its use for storage I/O, refer to the VMware Compatibility Guide.

Note: FCoE requires 10GbE equipment that conforms to data center bridging (DCB) standards.

Separate Storage Network

NetApp recommends separating storage network traffic from other networks. A separate network can be achieved by using separate switches or by creating a VLAN on shared switches. This network should not be routable to other networks. If switches are shared with storage and other traffic, it is imperative to confirm that the switches have adequate bandwidth to support the combined traffic load. Although the storage VLAN should not be routable, other VLANs (such as those for management or VM traffic on the same switches) may be routable. VLANs allow multiple network functions to share a small number of high-speed network connections, such as 10GbE. Figure 4 shows VLANs carrying VMkernel storage and management traffic and VM traffic.

Figure 4) VLANs carrying VMkernel storage and management traffic and VM traffic.



In ESXi, VLANs can be assigned to VM port groups and VMkernel ports. In clustered Data ONTAP, VLAN interfaces are created on top of ports or interface groups. When VLANs are used, LIFs are usually associated with a VLAN interface.

Jumbo Frames

Jumbo frames are larger Ethernet packets that reduce the ratio of packet overhead to payload. The default Ethernet frame size or MTU is 1,500 bytes. With jumbo frames, MTU is typically set to 9,000 on end nodes, such as servers and storage, and to a larger value, such as 9,198 or 9,216, on the physical switches.

Jumbo frames must be enabled on all physical devices and logical entities from end to end in order to avoid truncation or fragmentation of packets with the maximum size. On physical switches, the MTU must be set to the maximum supported value, either as a global setting or policy option or on a port-by-port basis (including all ports used by ESXi and the nodes of the NetApp cluster), depending on the switch implementation. The MTU must also be set and the same value must be used on the ESXi virtual switch and VMkernel port and on the physical ports or interface groups of each node. When problems occur, it is often because either the VMkernel or the virtual switch was not set for jumbo frames.

For VM guests that require direct access to storage through their own NFS or CIFS stack or iSCSI initiator, there is no MTU setting for the VM port group; however, the MTU must be configured in the guest.

Figure 5 shows jumbo frame MTU settings for the various networking components.

Figure 5) Jumbo frame MTU settings on VMkernel port, virtual switch, physical switches, and cluster-node ports.

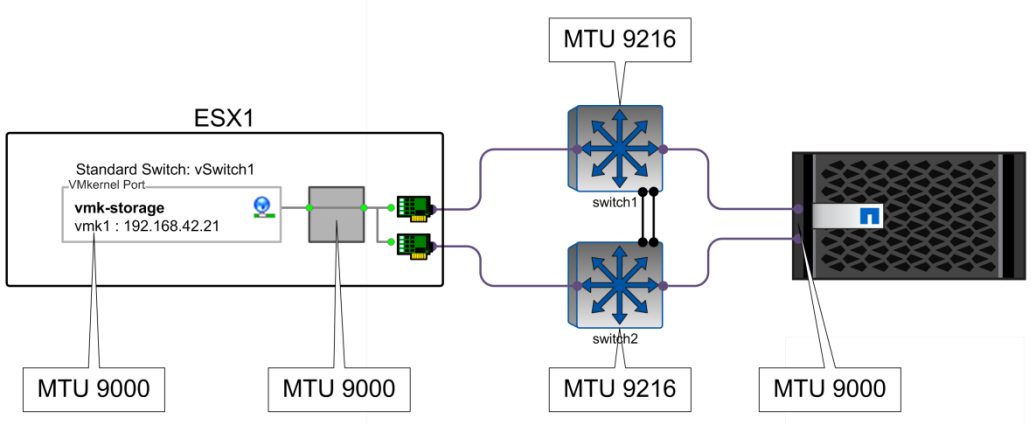


Figure 6 shows the standard switch properties, including the MTU setting.

Figure 6) Standard switch properties showing MTU.

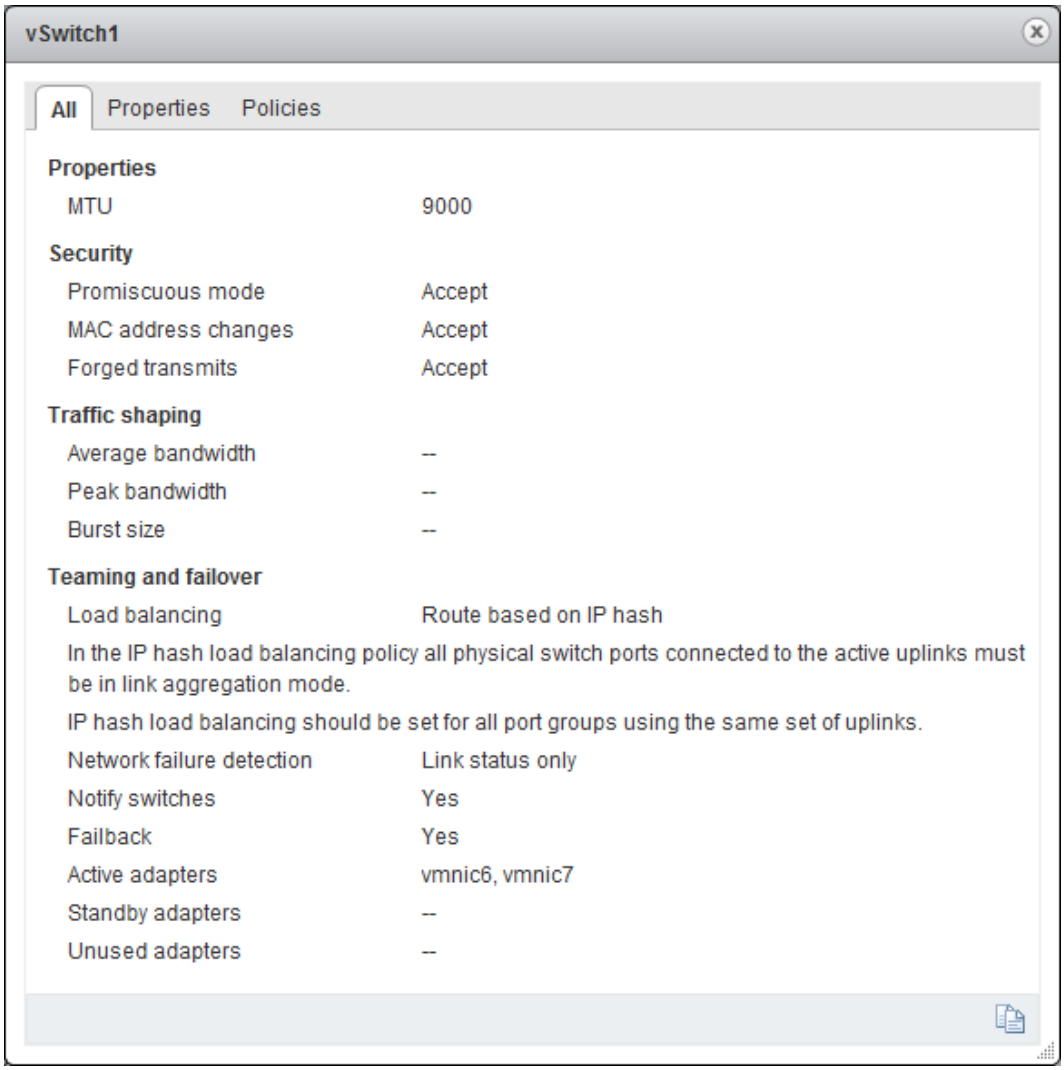
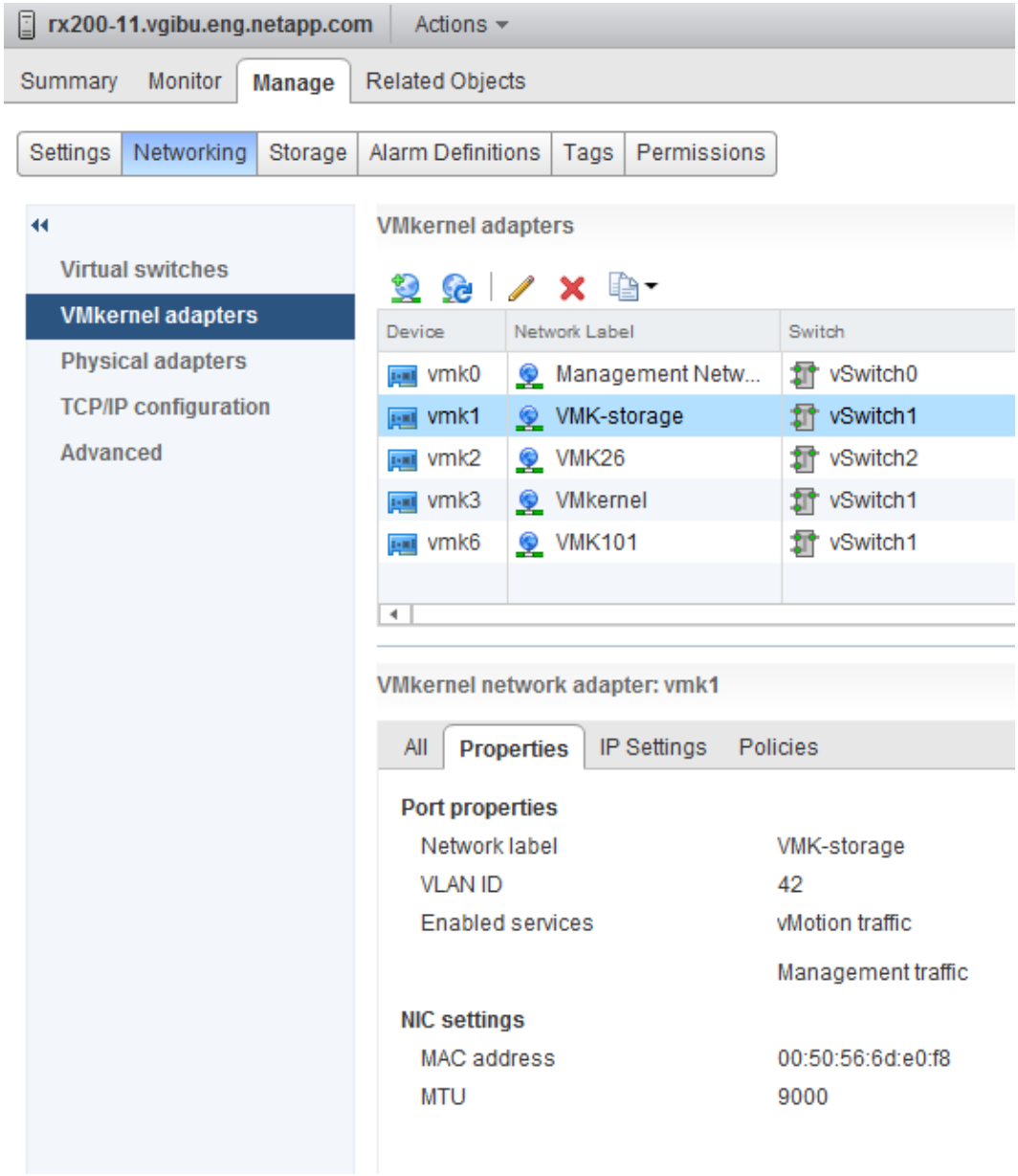


Figure 7 shows the VMkernel adapter properties, including the MTU setting.

Figure 7) VMkernel adapter properties showing MTU.



Ethernet Flow Control

Modern network equipment and protocols handle port congestion better than those in the past. NFS and iSCSI as implemented in ESXi use TCP. TCP has built-in congestion management, making Ethernet flow control unnecessary. Furthermore, Ethernet flow control can actually introduce performance issues on other servers when a slow receiver sends a pause frame to storage and stops all traffic coming out of that port until the slow receiver sends a resume frame. Although NetApp has previously recommended flow control set to `send` on ESXi hosts and NetApp storage controllers, the current recommendation is to disable flow control on ESXi, NetApp storage, and the switches' ports connected to ESXi and NetApp storage.

With ESXi 5, flow control is not exposed in the vSphere Client or vSphere Web Client. The `ethtool` command sets flow control on a per-interface basis. There are three options for flow control: `autoneg`, `tx`, and `rx`. The `tx` option is equivalent to `send` on other devices.

Note: With some NIC drivers, such as some Intel drivers, `autoneg` must be disabled in the same command line for `tx` and `rx` to take effect:

```
~ # ethtool -A vmnic2 autoneg off rx off tx off
~ # ethtool -a vmnic2
Pause parameters for vmnic2:
Autonegotiate:  off
RX:             off
TX:             off
```

Some NICs have hard-coded flow control settings that cannot be changed. Flow control must be disabled (`send off` and `receive off`) on the switch ports connected to these NICs. When flow control is disabled, the switch disregards any pause frames from the NIC and does not send them to the server or storage.

Spanning-Tree Protocol

Spanning-Tree Protocol (STP) is a network protocol that provides a loop-free topology for any bridged LAN. In the Open Systems Interconnection (OSI) model for computer networking, STP falls under the OSI layer 2. STP allows a network design to include spare (redundant) links to provide automatic backup paths if an active link fails, without the danger of bridge loops or the need for manual enabling or disabling of these backup links. Bridge loops must be avoided because they result in network flooding.

When ESXi and NetApp storage arrays are connected to Ethernet storage networks, NetApp strongly recommends configuring the Ethernet ports to which these systems connect as Rapid Spanning Tree Protocol (RSTP) edge ports or by using the Cisco PortFast feature. In an environment that uses the Cisco PortFast feature and that has 802.1Q VLAN trunking enabled to either the ESXi server or the NetApp storage arrays, NetApp recommends enabling the Spanning-Tree PortFast trunk feature.

When a port is configured as an edge port on an RSTP-enabled switch, the edge port immediately transitions its forwarding state to `active`. Ports that connect to other switch ports should not be configured with the edge port or the PortFast feature.

ESXi has some advanced settings for spanning tree and how it handles bridge protocol data units (BPDU), but these topics are out of the scope of this document because they apply to VM traffic as opposed to storage traffic. For more information, refer to the following VMware KB articles:

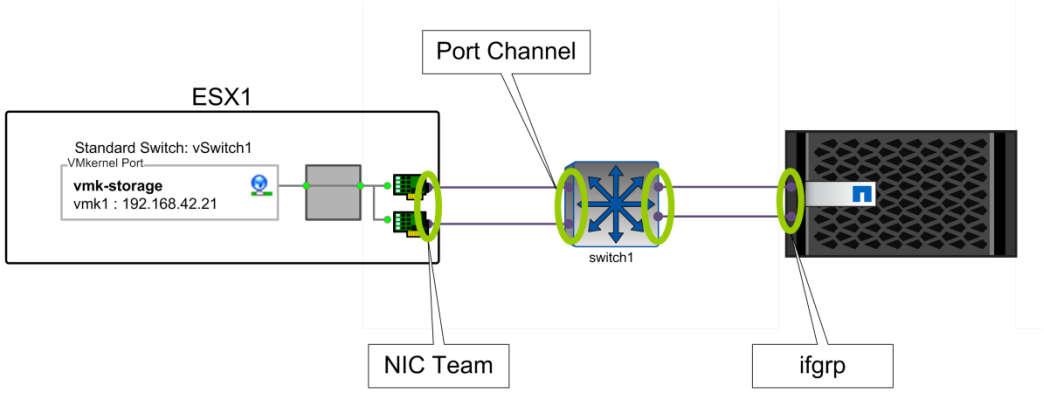
- STP may cause temporary loss of network connectivity when a failover or failback event occurs (1003804)
- Denial of service due to BPDU Guard configuration (2017193)
- Understanding the BPDU Filter feature in vSphere 5.1 (2047822)

Improving Network Performance and Redundancy with Link Aggregation

Link aggregation, standardized originally under IEEE 802.3ad but now as 802.1AX, refers to using multiple physical network connections to create one logical connection with combined throughput and improved redundancy. Several implementations of link aggregation are available, and they are known by different names: The vSphere implementation is referred to as NIC teaming; Cisco trademarked the term EtherChannel; the NetApp implementation is called interface groups; before the release of Data ONTAP 8.0, the NetApp implementation was known as virtual interfaces (VIFs). Other implementations might refer to bonding or trunking, although in Cisco terminology, trunking refers to carrying multiple tagged VLANs on a link or channel between two switches. The term load balancing is also used in conjunction with link aggregation. In practice, the loads are usually not symmetrically balanced between links in a team, although multiple links can carry traffic.

Not all link aggregation implementations are alike or offer the same features. Some offer only failover from one link to another in the event of the failure of one link. More complete solutions offer true aggregation, in which traffic can flow on two or more links at the same time. There is also the Link Aggregation Control Protocol (LACP), which allows devices to negotiate the configuration of ports into bundles. Failover-only configurations generally require no special capability or configuration of the switches involved. Aggregation configurations require compatible switches with the sets of ports configured for link aggregation. Figure 8 illustrates a simple link aggregation.

Figure 8) Simple link aggregation.



vSphere standard switches offer four implementations of NIC teaming. vSphere distributed switches add load-based teaming. These five options are described in Table 10.

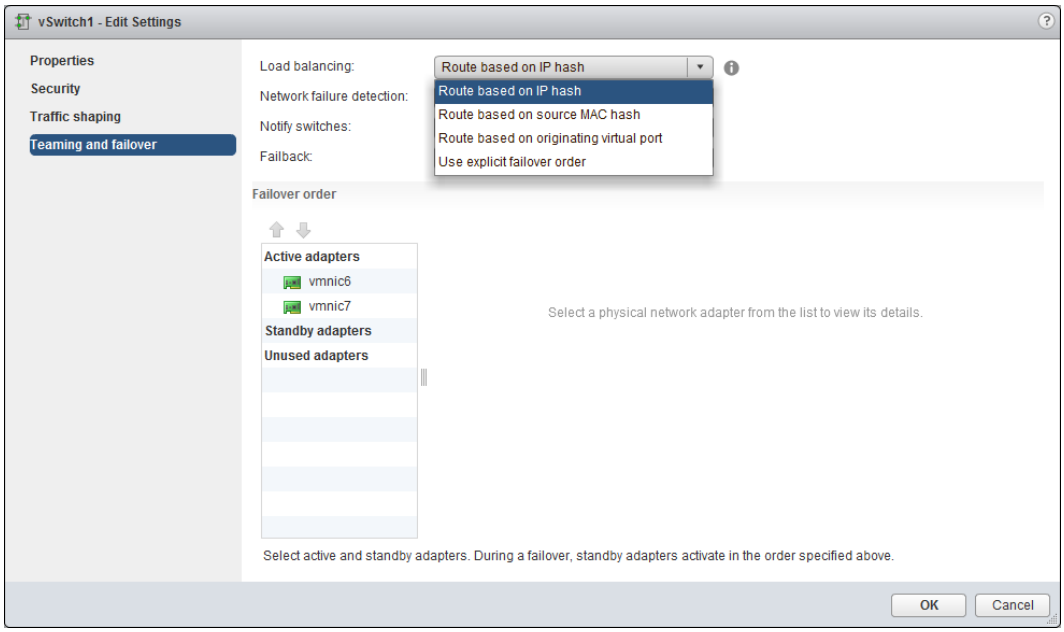
Table 10) vSphere NIC teaming options.

Teaming Policy	Failover	Aggregation	Switch Configuration Required?	Packet Distribution
Source virtual port ID	Yes	Yes*	No	Hash of internal virtual port ID of the VM or VMkernel on the virtual switch determines uplink used for all connections from that entity.
IP hash	Yes	Yes	Yes	Hash of source and destination IP address determines uplink per source-to-destination connection.
Source MAC hash	Yes	Yes*	No	Hash of MAC address of the VM or VMkernel on the virtual switch determines uplink used for all connections from that entity.
Explicit failover order	Yes	No	No	Administrator specifies the order in which to use uplinks.
Route based on physical NIC load	Yes	Yes	No	ESXi selects the uplink based on the current load of uplinks in the distributed switch. Available with vSphere distributed switches 5.1 and later only.

* Traffic from a single entity (VM or VMkernel) uses only one uplink NIC from the virtual switch unless that uplink fails, at which point the MAC and all traffic from that entity switch to another single uplink. Getting effective utilization of multiple links requires having multiple entities sending traffic.

Figure 9 shows the vSphere Web Client NIC teaming dialog box and options.

Figure 9) vSphere Web Client NIC teaming dialog box and options for vSphere standard switch.



Note: Only IP hash is a true implementation of 802.3ad. The source port and source MAC options allow an individual entity to transmit only over a single uplink. ESX and ESXi standard switches do not support LACP. vSphere distributed switches version 5.1 and later support LACP. The option to route based on physical NIC load is available only on distributed switches.

NetApp interface groups are created on top of two or more ports. With interface groups, the LIF is associated with the interface group rather than with the underlying ports. Table 11 compares the three variations of interface groups.

Table 11) Data ONTAP interface group types.

Interface Group Type	Failover	Aggregation	Switch Configuration Required	Packet Distribution
Single mode	Yes	No	No	Single active link
Static multimode	Yes	Yes	Yes	IP, MAC, round robin, TCP/UDP port
Dynamic multimode (LACP)	Yes	Yes	Yes	IP, MAC, round robin, TCP/UDP port

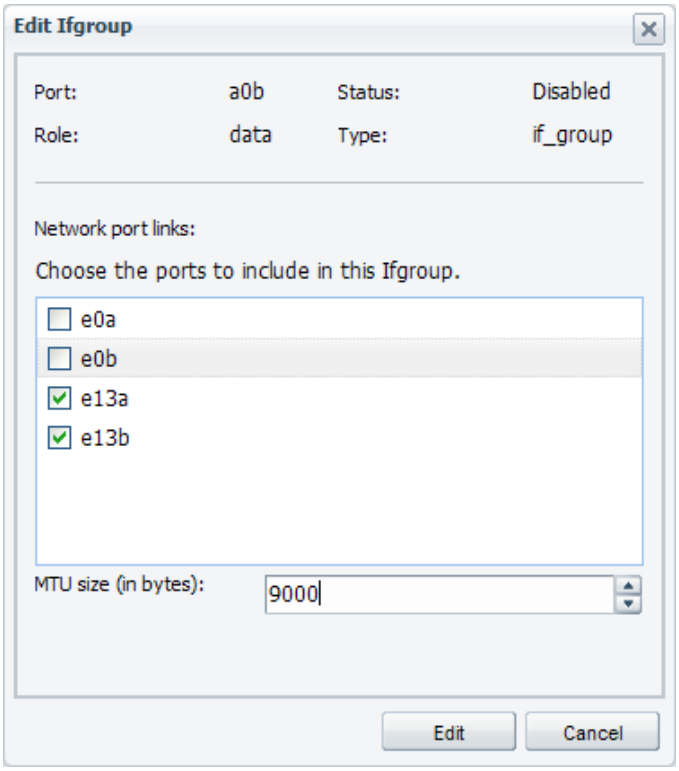
Note: NetApp recommends dynamic multimode if the switch supports LACP.

As specified in Table 11, single-mode interface groups send and receive traffic only on a single active link. Other links in the interface group remain unused until the first link fails. Failover groups offer an improved design for use with switch configurations that do not support link aggregation. Instead of configuring a single-mode interface group, the administrator can assign the ports to a failover group. LIFs that use these ports are set to use this failover group and have their failover policy set to `nextavail`, which makes them usually prefer ports on the current node in the event of a port failure. Each LIF has a

different home port. Although traffic is not balanced, all links can carry traffic concurrently and can take over for each other in the event of a link failure.

Figure 10 shows the listing of interface group ports and MTU size in OnCommand System Manager.

Figure 10) Interface group ports and MTU in OnCommand System Manager.



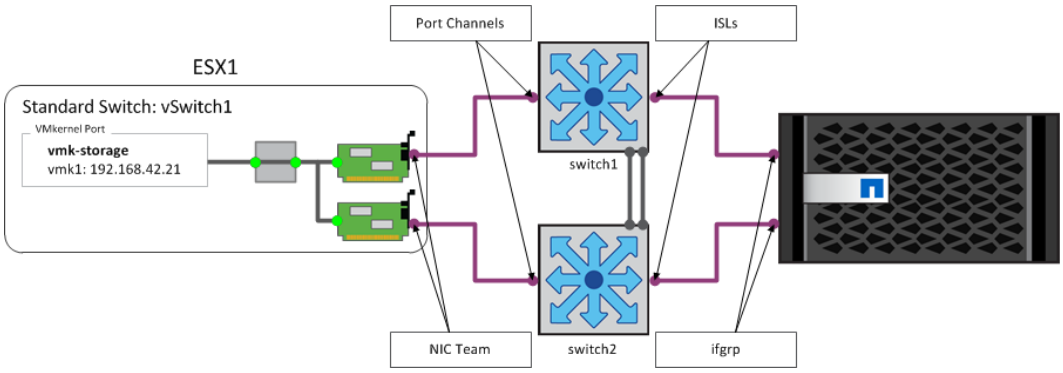
Early switch implementations of link aggregation allowed ports of only a single switch to be combined into a team, even on many stackable switches. More recent switches offer technology that allows ports on two or more switches to become a single team. Switches are connected to each other with interswitch links (ISLs) that may be 1GbE or 10GbE, or through proprietary cables. Table 12 provides a partial list of vendors and switches that support link aggregation of ports on multiple switches.

Table 12) Partial list of switch vendors and models offering link aggregation across multiple switches.

Vendor	Switch or Family	Feature Name	Notes
Brocade	VDX6700	Virtual Link Aggregation Group (vLAG)	
Cisco	Cisco® Catalyst 3750	CrossStack EtherChannel	
Cisco	Cisco Catalyst 6500	Multichassis EtherChannel (MEC)	Requires Supervisor Engine 720 and VSS 1440
Cisco	Cisco Nexus	Virtual Port Channel (vPC)	A vPC combines two EtherChannels on two switches
Nortel (Avaya)	Various	Split Multilink Trunking (SMLT)	

Figure 11 illustrates link aggregation using two switches.

Figure 11) Link aggregation using two switches.



Whether a single switch or a pair of properly stacked switches is used, the configuration on the ESXi servers and storage nodes is the same because the stacking technology makes the two switches look like one to the attached devices.

Best Practices

NetApp recommends the following best practices for link aggregation:

- Use switches that support link aggregation of ports on both switches.
- Disable LACP for switch ports connected to ESXi unless using dvSwitches 5.1 or later with LACP configured.
- Enable LACP for switch ports connected to NetApp nodes.
- Use IP hash on ESXi.
- Use dynamic multimode (LACP) with IP hash on NetApp nodes.

For switches that do not support link aggregation of ports on both switches, a virtual port or MAC routing should be used on the ESXi side, and single-mode interface groups should be used on the NetApp side.

Summary

Table 13 provides a summary of network configuration items and indicates where the settings are applied.

Table 13) Applicability of network configuration.

Item	ESXi	Switch	Node	SVM
IP address	VMkernel	No**	No**	Yes
Link aggregation	Virtual switch	Yes	Yes	No*
VLAN	VMkernel and VM port groups	Yes	Yes	No*
Flow control	NIC	Yes	Yes	No*
Spanning tree	No	Yes	No	No
MTU or jumbo frames	Virtual switch and VMkernel port (9,000)	Yes (set to max)	Yes (9,000)	No*
Failover groups	No	No	Yes (create)	Yes (select)

Item	ESXi	Switch	Node	SVM
Routing (if used)	VMkernel	Yes (on router)	No	Yes

* SVM LIFs connect to ports, interface groups, or VLAN interfaces that have VLAN, MTU, and other settings, but the settings are not managed at the SVM level.

**These devices have IP addresses of their own for management, but these addresses are not used in the context of ESXi storage networking.

6.2 VMware vSphere 6 and Clustered Data ONTAP Basic Networking Deployment Procedures

Table 14 describes basic networking prerequisites for VMware vSphere on clustered Data ONTAP.

Table 14) VMware vSphere 6 basic networking clustered Data ONTAP prerequisites.

Description
A NetApp cluster running clustered Data ONTAP 8.1 or later is required.
NetApp OnCommand System Manager 3.0 or later is required.
vSphere 6 (including ESXi 5, vCenter 5, and the vSphere Web Client) is required.

Tasks to Configure Networking for vSphere

Configuring networking for a vSphere environment connected to clustered Data ONTAP storage involves the following tasks:

- Establishing physical connections to the switches
- Configuring the physical switches:
  - Link aggregation
  - Flow control
  - Spanning tree
  - Jumbo frames or maximum transmission unit (MTU)
  - VLANs
- Configuring the NetApp cluster nodes:
  - Flow control on cluster ports
  - Interface groups (ifgrps), including MTU
  - VLANs
  - Failover group, including ports, interface group, and/or VLAN assignments
- Configuring SVM and networking for vSphere use:
  - SVM setup
  - LIFs appropriate to the protocol
  - Failover group for NFS LIFs
- Configuring the ESXi servers:
  - Physical NICs, including flow control
  - Virtual switches, including link aggregation and MTU
  - VMkernel port, including VLAN, MTU, IP address, and subnet mask

- VM port group for VM-to-storage access (optional), including VLAN and consistent port group name on all ESXi servers

For instructions on the first two configuration tasks (establishing physical connections to the switches and configuring them), consult the vendor documentation for the switches used in your environment. The other tasks are described in the following procedures.

Configure Flow Control on Cluster Ports

Flow control is configured on the physical ports of each node in the NetApp cluster, even if the port is a member of an interface group.

Note: Changing flow control settings disrupts the network connection for several seconds.

Note: Some unified target adapters have flow control hard-coded to `full`, and this setting cannot be changed.

To configure flow control by using the clustershell command line, complete the following steps:

1. Log in to the clustershell as the cluster administrator through SSH or the console port of a node.
2. Run the `net port show` command to verify the current flow control settings.

```
eadrax::> net port show -node eadrax-01 -fields flowcontrol-admin,flowcontrol-oper
  (network port show)
node      port flowcontrol-admin flowcontrol-oper
--------- ---- ----------------- ----------------
eadrax-01 a0b  full              -
eadrax-01 a0b-42
               full              -
eadrax-01 e0M  full              full
eadrax-01 e0a  full              full
eadrax-01 e0b  full              none
eadrax-01 e1a  none              none
eadrax-01 e1b  full              full
eadrax-01 e2a  none              none
eadrax-01 e2b  full              full
eadrax-01 e3a  full              full
eadrax-01 e3b  full              full
```

3. Run the `net port modify` command to change the flow control setting for each port used for VMware storage traffic.

```
eadrax::> net port mod -node eadrax-01 -port e1b -flowcontrol-admin none
  (network port modify)

Warning: Changing the network port settings will cause a several second interruption in carrier.
Do you want to continue? {y|n}: y
```

4. Repeat step 3 to configure each port used for ESXi-to-storage networking on each node.
5. Run the `net port show` command again to verify the new flow control settings.

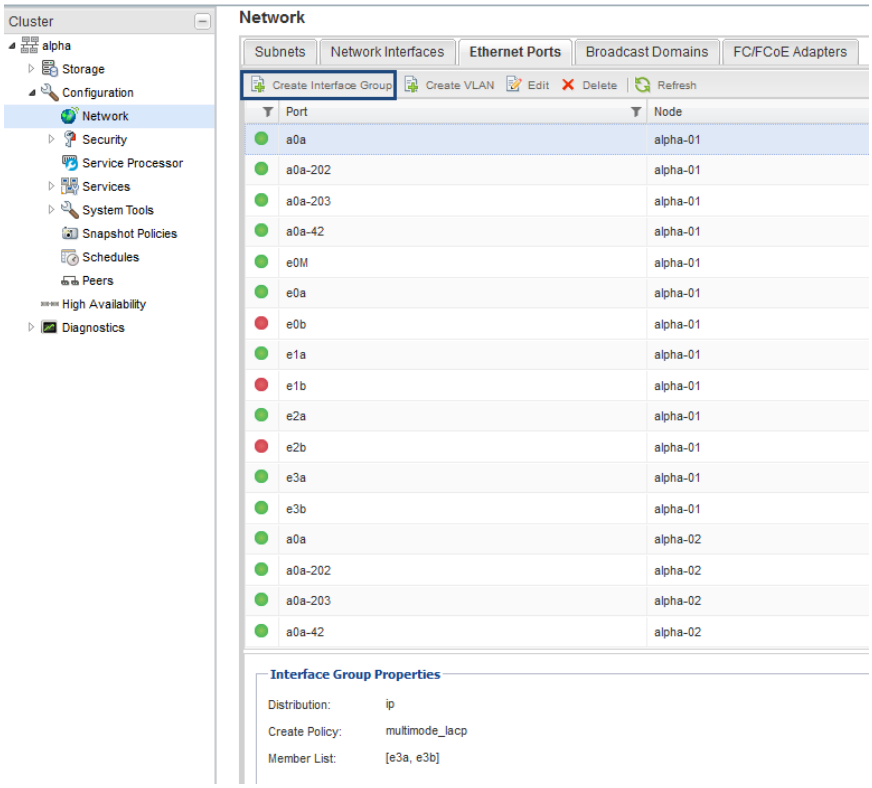
```
eadrax::> net port show -node eadrax-01 -fields flowcontrol-admin,flowcontrol-oper
  (network port show)
node      port flowcontrol-admin flowcontrol-oper
--------- ---- ----------------- ----------------
eadrax-01 a0b  full              -
eadrax-01 a0b-42
               full              -
eadrax-01 e0M  full              full
eadrax-01 e0a  full              full
eadrax-01 e0b  full              none
eadrax-01 e1a  none              none
eadrax-01 e1b  none              none
eadrax-01 e2a  none              none
eadrax-01 e2b  none              none
eadrax-01 e3a  full              full
eadrax-01 e3b  full              full
```


Create Interface Groups for Link Aggregation

Before configuring the NetApp cluster nodes for link aggregation, verify that the switches are configured, that the properties of the channels or teams are known, and that the ports of each node are connected to the correct switch ports.

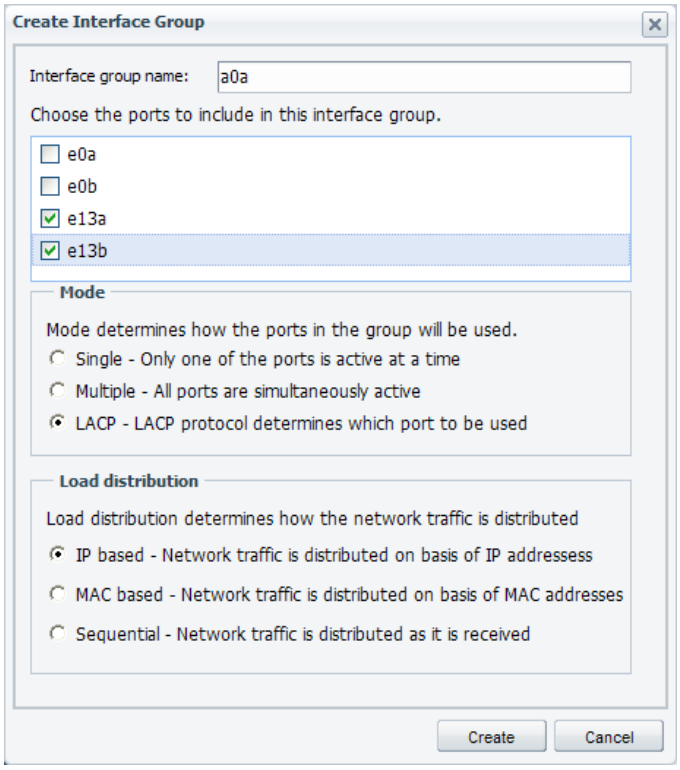
To create interface groups for link aggregation, complete the following steps:

1. In OnCommand System Manager, navigate to Cluster > `<cluster>` > Configuration > Network.
2. Click Create Interface Group.

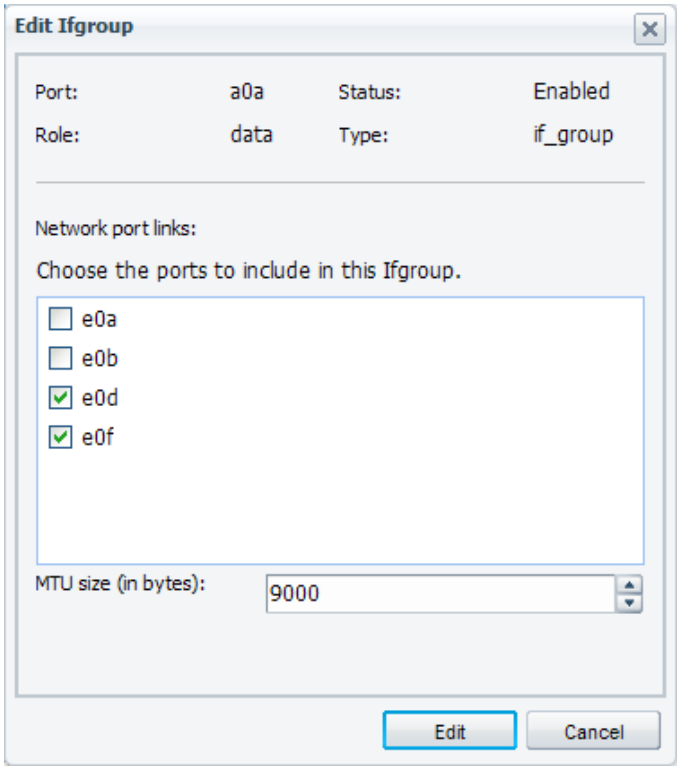


3. Name the interface group. The name must start with the letter `a`, followed by a number and another letter. Use the same name for the equivalent interface group on each node.
4. Select the ports for the interface group.
5. Select the correct mode according to the capabilities and settings of the switch:
   - If the switch ports are properly channeled or bundled and LACP is enabled, select LACP.
   - If the switch ports are properly channeled or bundled but LACP is not enabled or supported, select Multiple.
   - If the switch ports are not properly channeled or bundled, or if the switches are not stackable or otherwise not capable for this feature, select Single.

   Note: NetApp recommends failover groups consisting of individual ports rather than single-mode interface groups.
6. Select IP Based as the type of load distribution.



7. Click Create.
8. After the interface group is created, select it and click Edit.
9. Set the MTU size to `9000` if jumbo frames are used and the switches are properly configured.

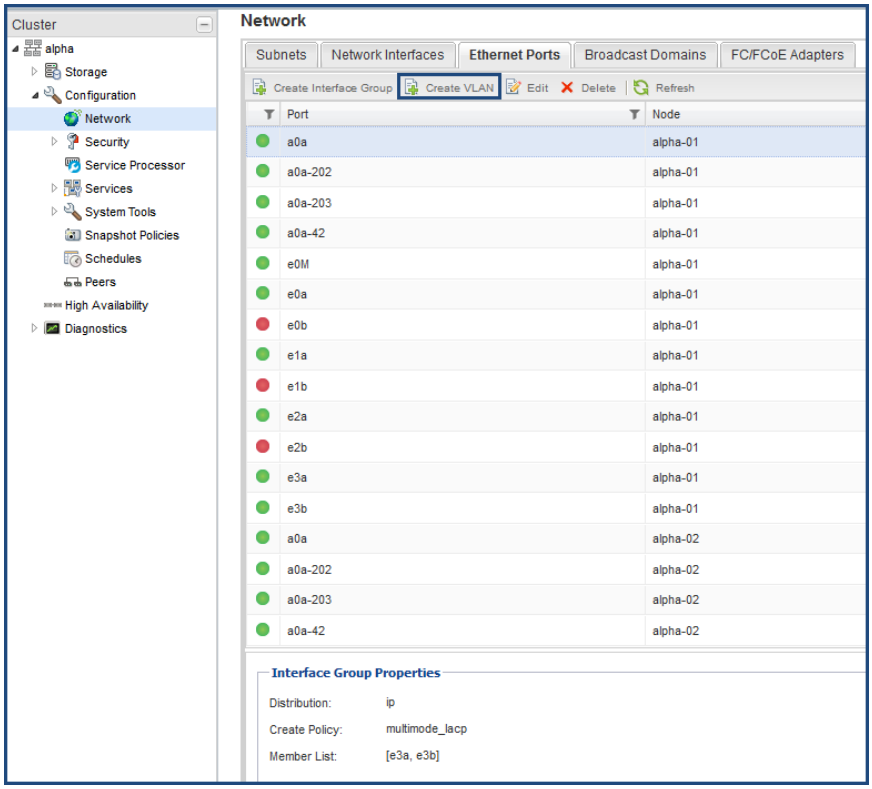


10. Click Edit to save the change.
11. Repeat this procedure for each interface group on each node in the NetApp cluster.

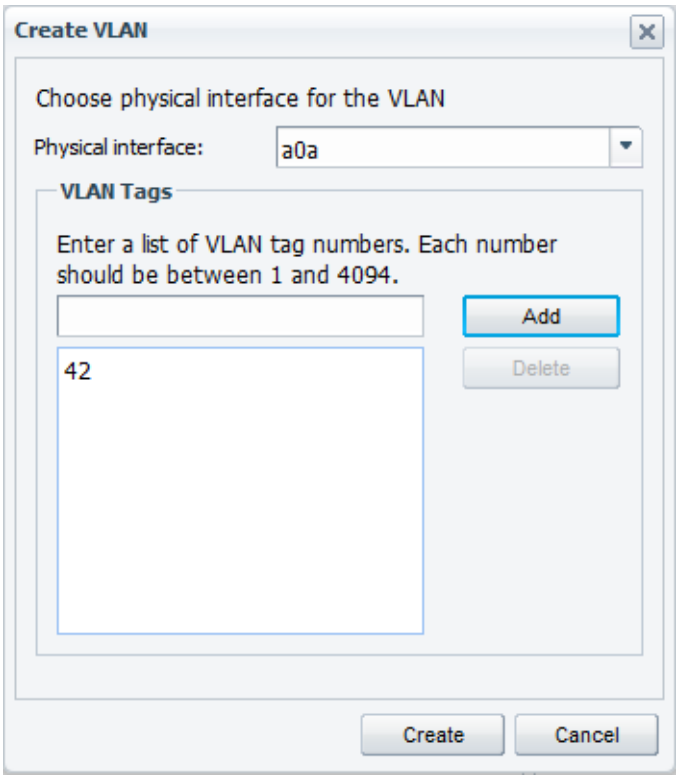
Create VLANs

To create VLANs, complete the following steps:

1. In OnCommand System Manager, navigate to Cluster > `<cluster>` > Configuration > Network > Ethernet Ports.
2. Click Create VLAN.



3. Select the physical interface (port or interface group) on which to create the VLAN interface.
4. Enter the VLANs one at a time and click Add after you enter each one. All VLANs entered in one dialog box are created on the same physical interface.



5. Click Create.
6. Repeat this procedure for all VLANs on all nodes.

Create Network Failover Groups

To create failover groups, complete the following steps:

1. Log in to the clustershell through SSH or the console port of a node.
2. Create a failover group.

```
vice::> failover-groups create -failover-group <fo_group> -node <node> -port <port_ifgrp_or_vlan>
```

For example:

```
vice::> failover-groups create -failover-group private -node vice-01 -port a0a
```

3. Run the `failover-groups create` command to add the rest of the appropriate ports to the failover group.
4. Repeat step 3 to add all appropriate ports, interface groups, or VLANs to each failover group.

Configure SVM Networking for vSphere Use

To configure SVM networking for vSphere use, refer to section 7.6, “Supported NFS Versions,” and complete the procedures described in the following sections:

- **Best Practices**

- Do not mix NFS versions across hosts. If possible, use host profiles to check compliancy.
- Because there is no automatic datastore conversion, create a new NFSv4.1 datastore and use Storage vMotion to migrate VMs to the new datastore.

- Create New SVM Configured for NFS
- Create Additional LIFs for NFS Datastores
- Assign LIFs to Failover Group

Best Practices
• Do not mix NFS versions across hosts. If possible, use host profiles to check compliancy. • Because there is no automatic datastore conversion, create a new NFSv4.1 datastore and use Storage vMotion to migrate VMs to the new datastore.

Configure ESXi Storage Networking for Physical NICs

Before configuring ESXi storage networking for the physical NICs, verify the following:

- The physical NICs are connected to the correct switch ports on the correct switches.
- The switch ports are correctly configured, including channels or teams and VLANs.
- Spanning tree is disabled or set to `portfast` on switch ports used for ESXi.
- Flow control is disabled on switch ports used for ESXi.

  Note: For information about how to configure and verify these items, consult the vendor documentation for the switches used in your environment.

Table 15) ESXi storage networking configuration prerequisites.

ESXi storage networking configuration for physical NIC prerequisites:
The physical NICs are connected to the correct switch ports on the correct switches.
The switch ports are correctly configured, including channels or teams and VLANs.
Spanning tree is disabled or set to portfast on switch ports used for ESXi.
Flow control is disabled on switch ports used for ESXi. **Note:** For information about how to configure and verify these items, consult the vendor documentation for the switches used in your environment.

Verify Connectivity on ESXi Physical NICs

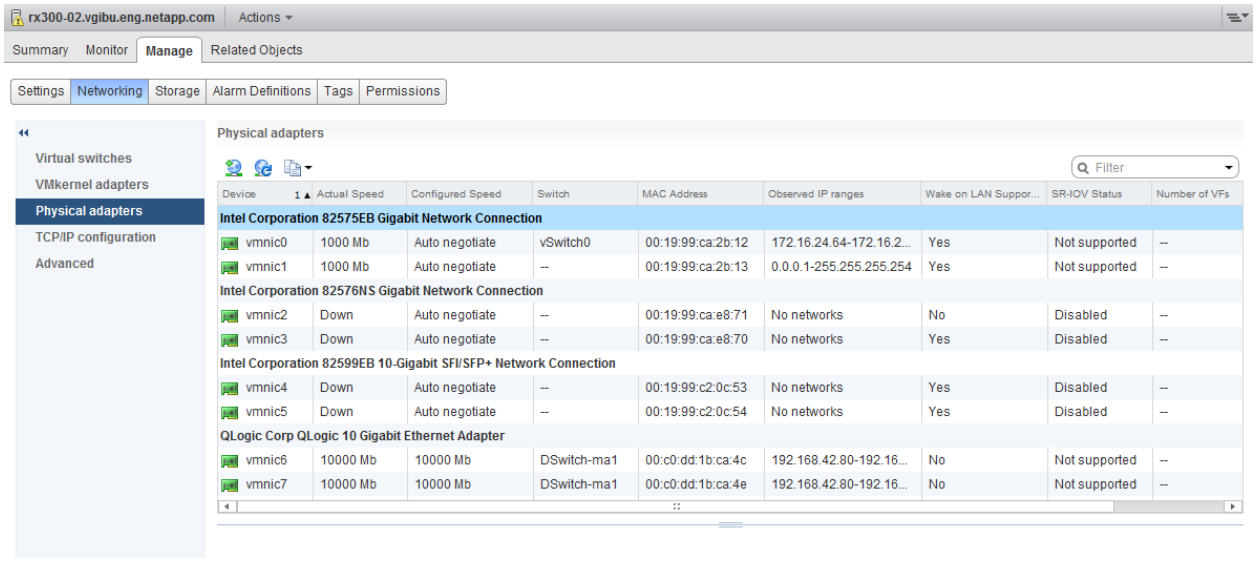
The correct NICs for storage networking can often be determined by checking the NIC manufacturer and model or the NIC speed. The vSphere Web Client displays this information under the Physical Adapters heading in the Manage > Networking tab for individual ESXi servers.

When a server includes several NICs from the same manufacturer or when several NICs are operating at the same speed, connectivity can be determined by verifying the Cisco Discovery Protocol (CDP) information for physical switches that support CDP. For the vSphere Client to be able to display this information, the NICs must be associated with a virtual switch. The vSphere Web Client does not require NICs to be part of a virtual switch in order to display CDP information. Further, the web client adds support for displaying Link Layer Discovery Protocol (LLDP), an open standard for passing similar connectivity information.

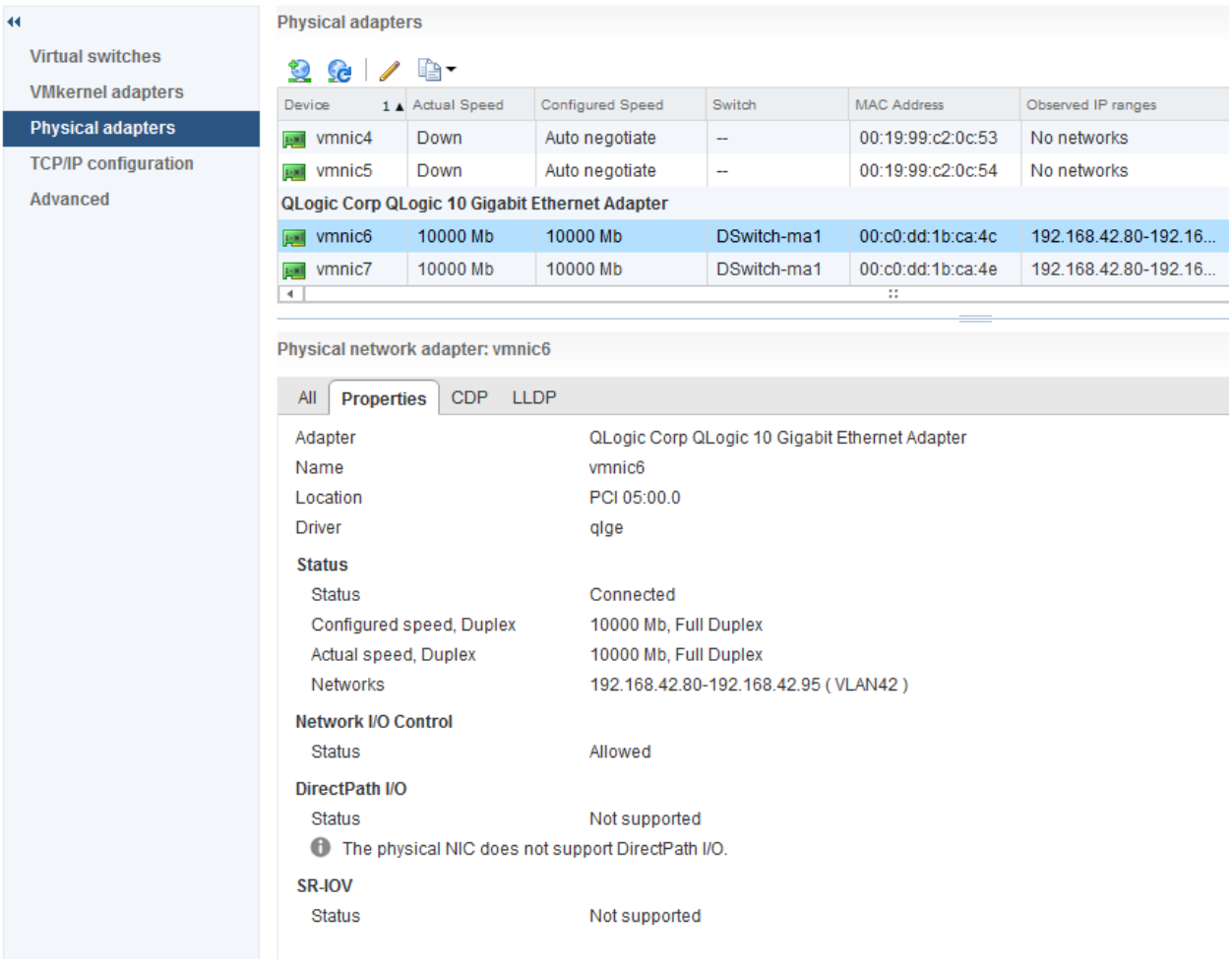
To verify connectivity on the ESXi physical NICs, complete the following steps:

1. Using a web browser, log in to the vSphere Web Client.

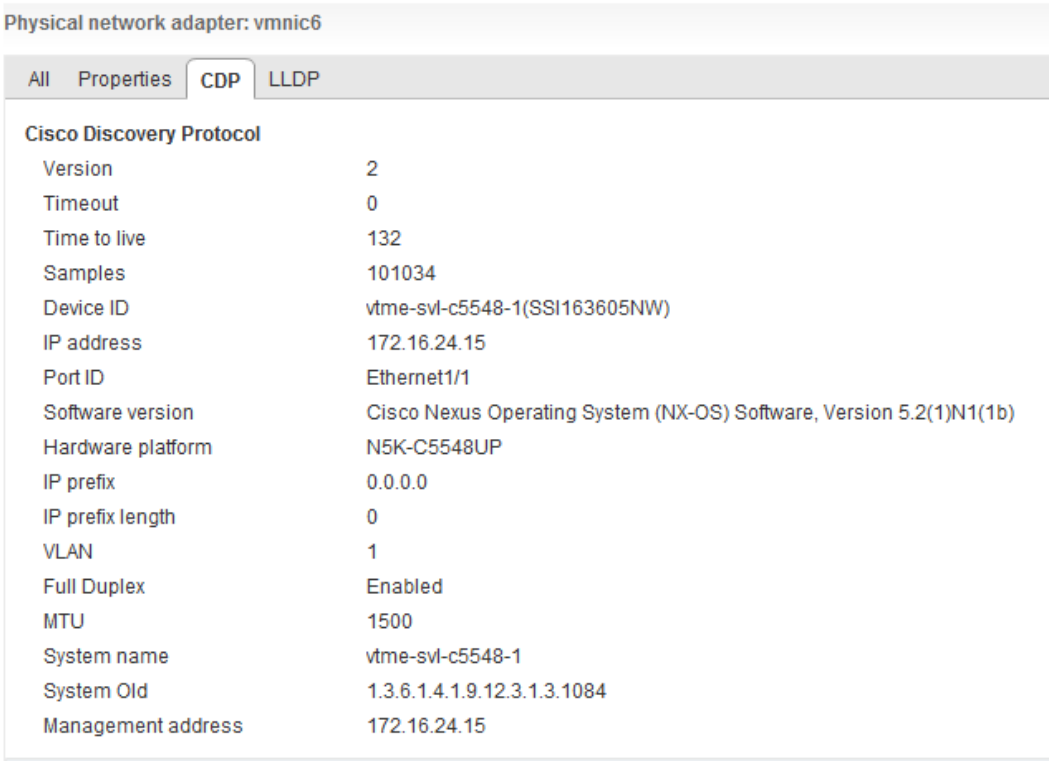
2. Navigate to an ESXi server.
3. Click Manage > Networking > Physical Adapters.



4. For the network adapters to be used for ESXi storage networking, verify the following settings:
   - The actual speed is correct, and the duplex setting is `Full`.
   - The setting in the Switch column is `None`, unless virtual switches are already configured for storage.
   - The setting in the Observed IP Ranges column may be `None`. However, if populated, this field should show traffic in the appropriate subnet.
5. To see more details about a NIC, click it, then either review the data on the All tab or click the Properties and CDP tabs to view the information separately.



6. Verify the duplex setting under Properties and the device ID (or the system name) and the port ID to match the physical NIC to the physical switch port under the CDP tab.



Configure Flow Control on Physical NICs Used for ESXi

To configure flow control in ESXi from the ESXi shell, complete the following steps:

1. Log in to the ESXi CLI from the physical console, the remote console of the physical server, or SSH.
2. Verify the current settings for flow control for each physical NIC.

```
~ # ethtool -a vmnic2
Pause parameters for vmnic2:
Autonegotiate:  on
RX:             on
TX:             on
```

3. Make and verify any changes to the flow control settings for each physical NIC.

```
~ # ethtool -A vmnic2 autoneg off rx off tx off
~ # ethtool -a vmnic2
Pause parameters for vmnic2:
Autonegotiate:  off
RX:             off
TX:             off
```

4. Edit `/etc/rc.local` and append the `ethtool -A` command for each NIC.

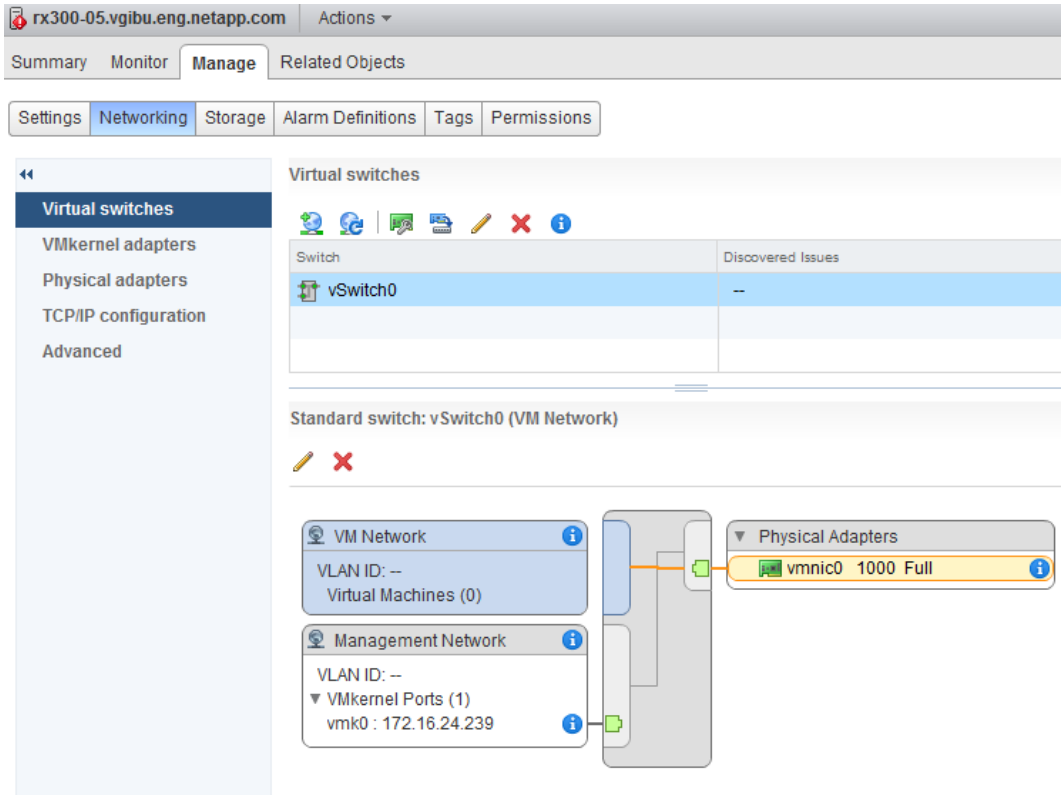
```
~ # cp /etc/rc.local /etc/rc.local.save
~ # echo "ethtool -A vmnic2 autoneg off rx off tx off" >> /etc/rc.local
~ # echo "ethtool -A vmnic3 autoneg off rx off tx off" >> /etc/rc.local
```

Note: Make sure that two greater-than (>>) symbols are included (to indicate the append operation) and that the vmnic number is changed in these commands.

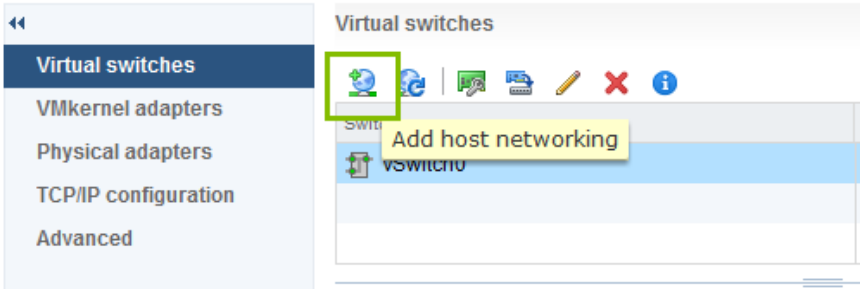
Create and Configure Standard Switch and VMkernel Port for Storage Networking

To create and configure a standard switch and a VMkernel port for storage networking, complete the following steps:

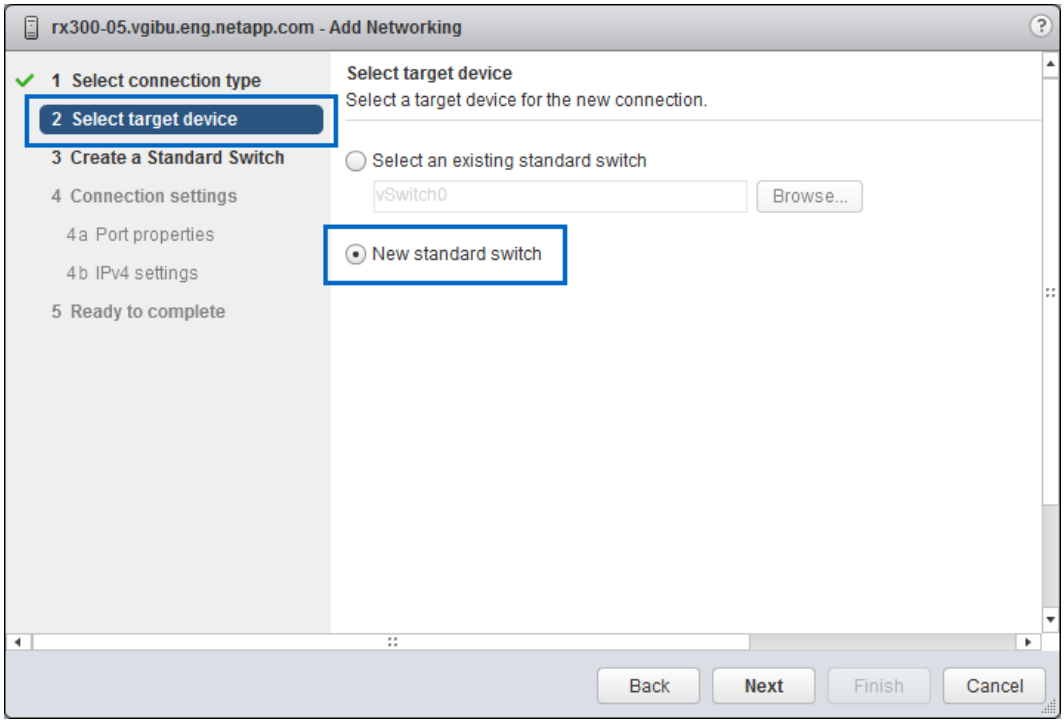
1. Using a web browser, log in to the vSphere Web Client.
2. Navigate to an ESXi server.
3. Click the Manage tab, then Networking, and then select Virtual Switches.



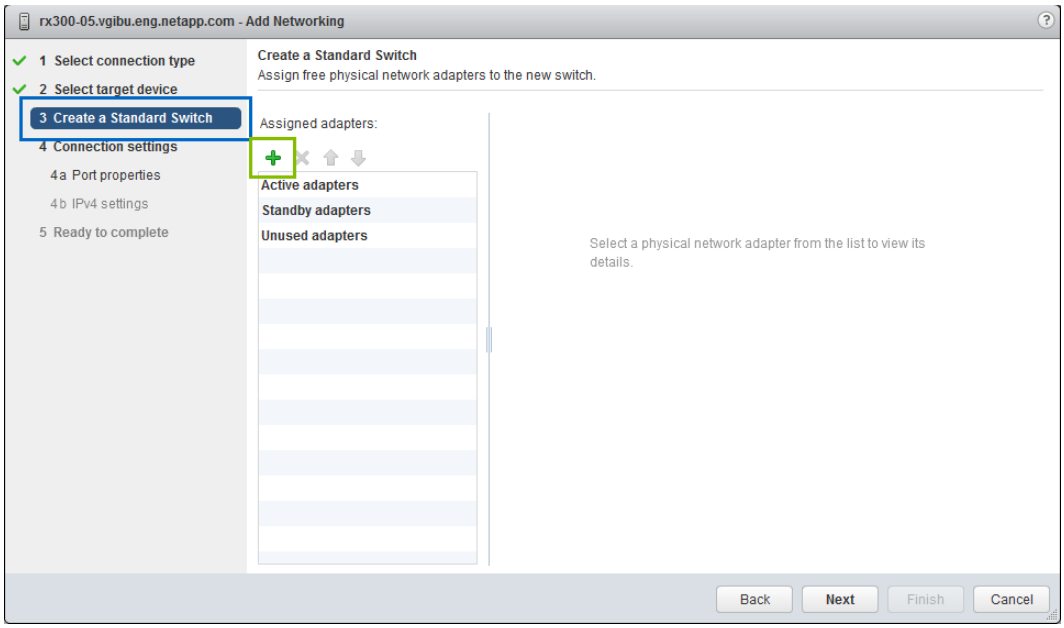
4. Click the Add Host Networking icon.



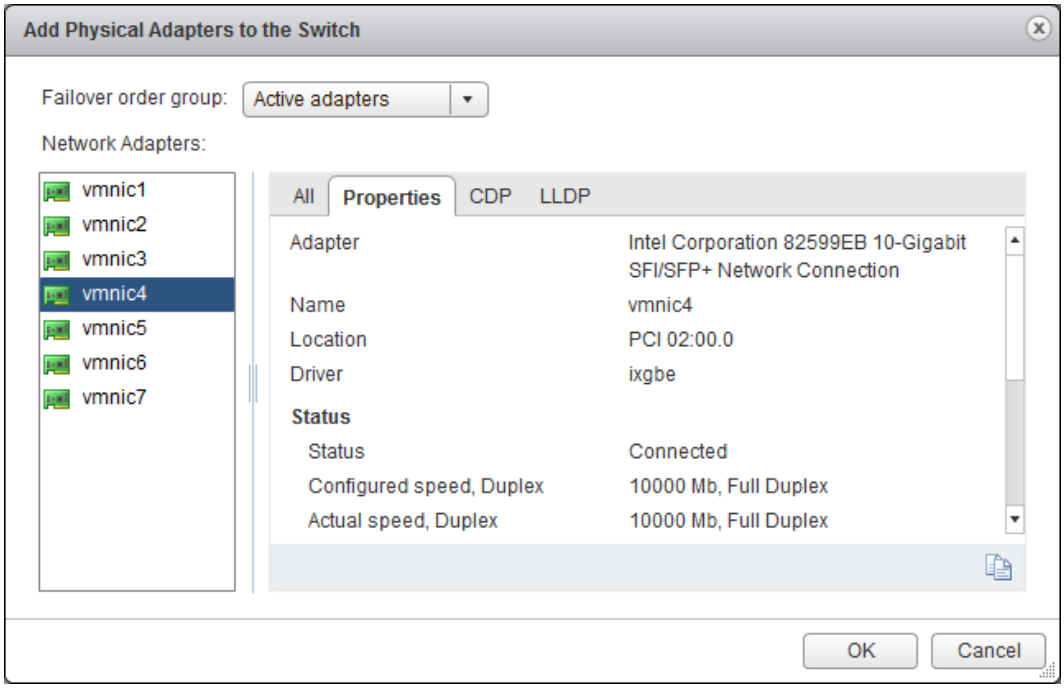
5. In the Select Connection Type window, select VMkernel Network Adapter and click Next.
6. In the Select Target Device window, select New Standard Switch. An additional step is then displayed in the workflow list on the left.



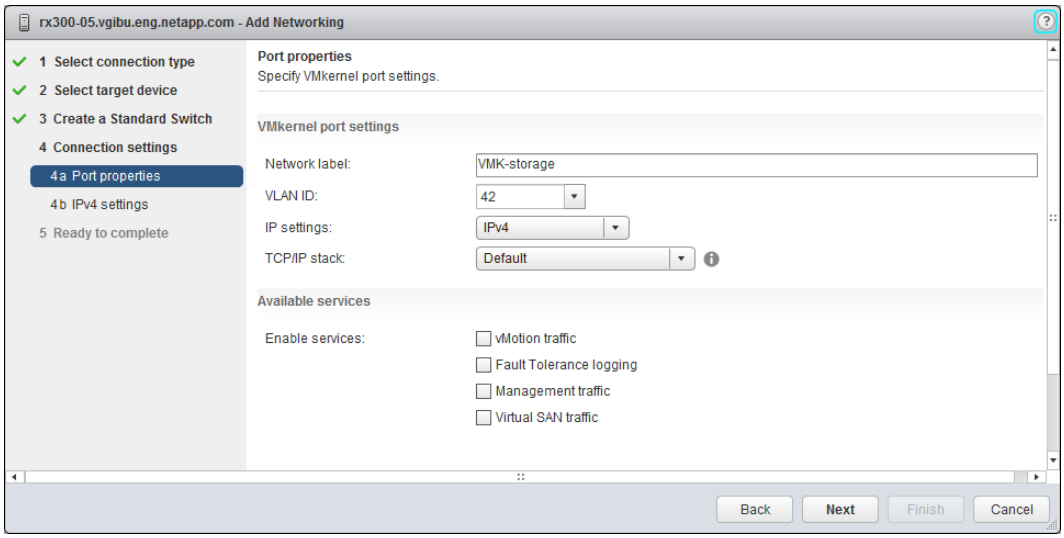
7. Click Next.
8. In the Create a Standard Switch window, click the green plus button to add a NIC.



9. In the Add Physical Adapters window, select one or more NICs to add to the switch. When you select a single NIC, you can examine the properties and CDP information to verify it is the correct NIC. Click OK.

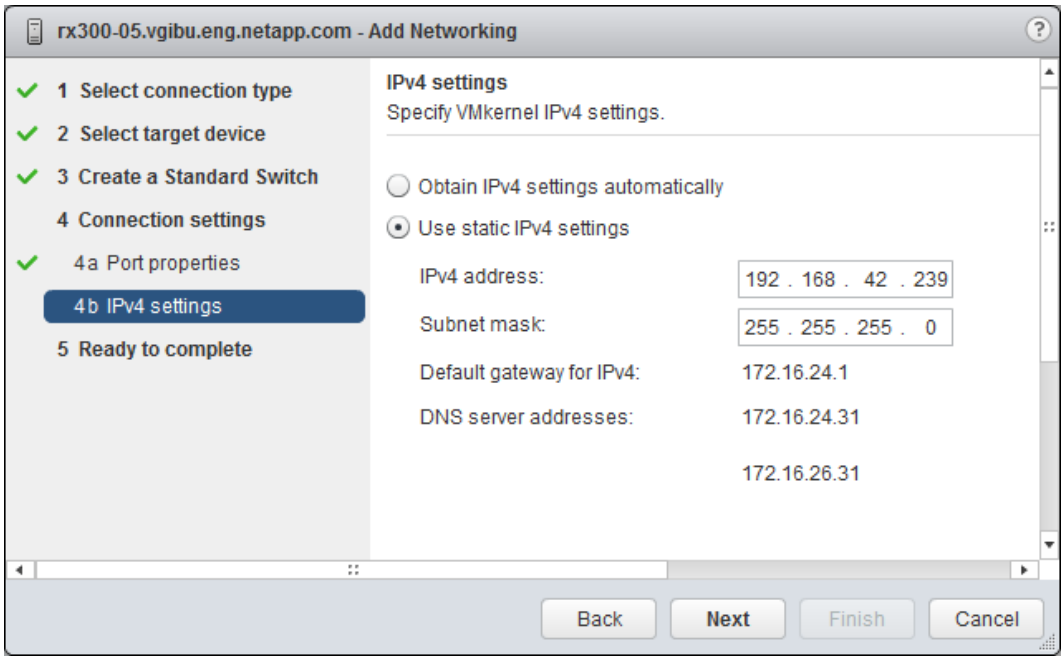


10. Click Next.
11. In Connection Settings > Port Properties, configure the VMkernel connection settings:
    a. Enter a VMkernel name, such as `vmk-storage`. Use the same name for the same VMkernel network on all ESXi servers.
    b. Set the VLAN ID if the storage network requires ESXi to add a VLAN tag to its traffic.
    c. Select IPv4 and Default TCP/IP stack. IPv6 is out of the scope of this document.
    d. If the network bandwidth supports vMotion and fault tolerance logging, the checkboxes for these options can be selected.
    e. Because the storage network should not be routable, it should not be possible to access the management port, so leave the checkbox for management traffic unselected. Click Next.

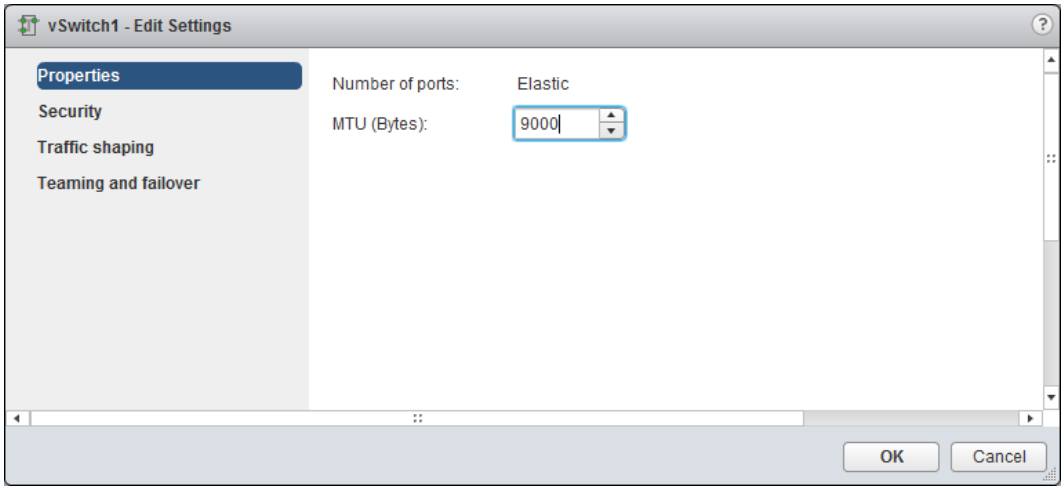


12. In the IPv4 Settings window, select one of the following options:

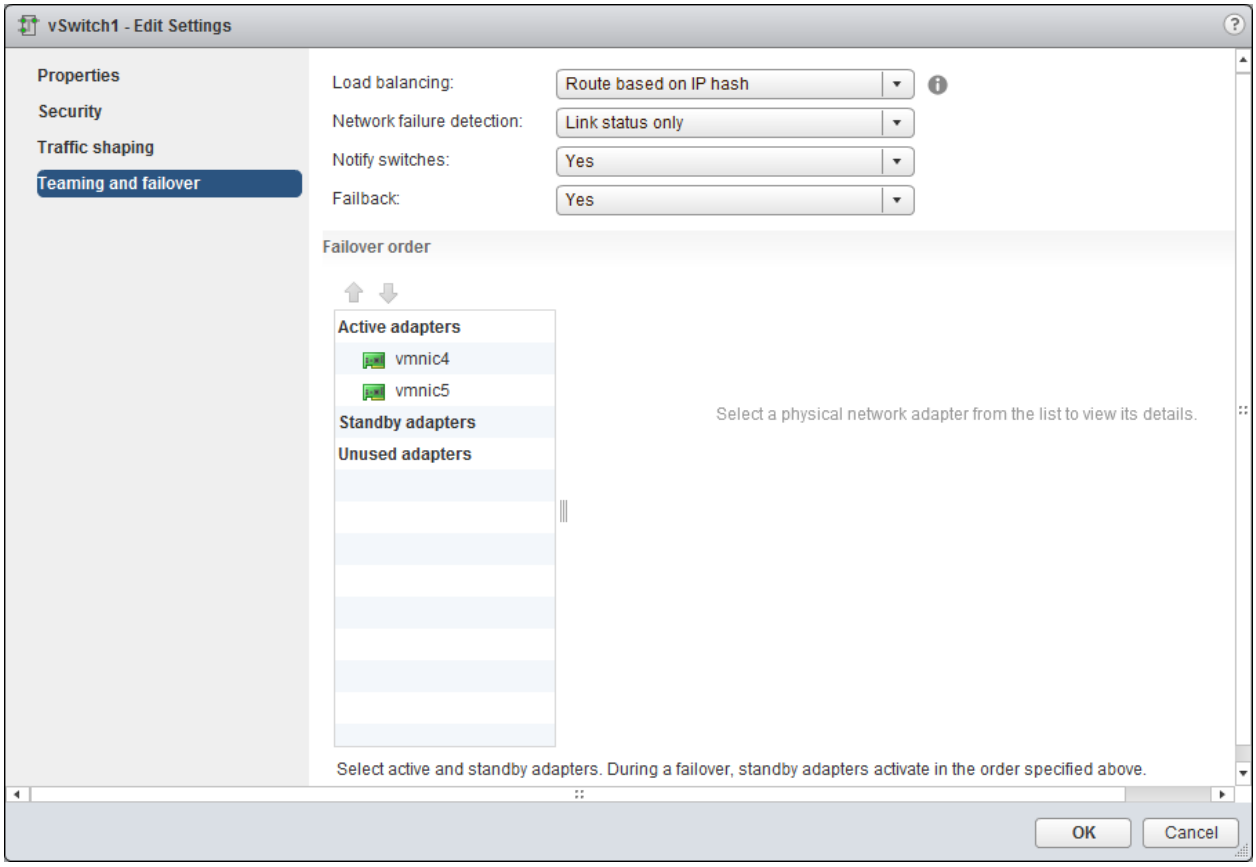
- If DHCP is used on the storage network, leave Obtain IPv4 Settings Automatically selected.
- Otherwise, select Use Static IPv4 Settings and enter a specific IP address and subnet mask.



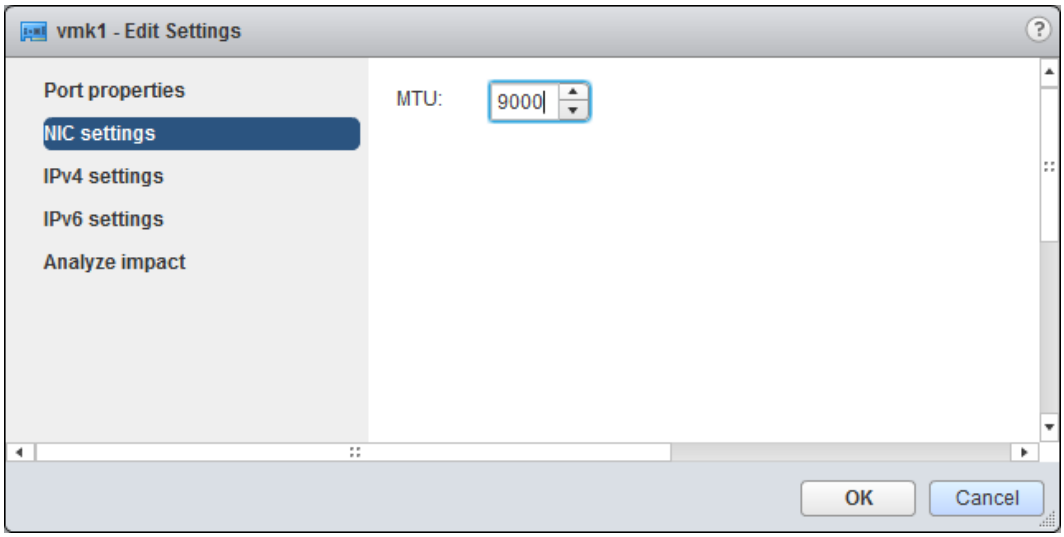
13. Click Next, review the settings, and click Finish.
14. Select the new switch and click the pencil icon to edit the settings.
15. In Edit Settings > Properties, set the MTU appropriately for your storage network.



16. Click Teaming and Failover and set the load balancing policy:
    - If you are using link aggregation on the switches, select Route Based on IP Hash.
    - If the switches do not support, or are not configured for, link aggregation, select either Route Based on the Originating Virtual Port ID or Route Based on Source MAC Hash.
    - Do not use Beacon Probing with IP Hash.
    - Verify that the NICs are active.



17. Click OK.
18. Click VMkernel adapters. Select the new VMkernel port and click the pencil icon to edit the settings.
19. Click NIC settings.
20. If using jumbo frames, set the MTU for the VMkernel port.

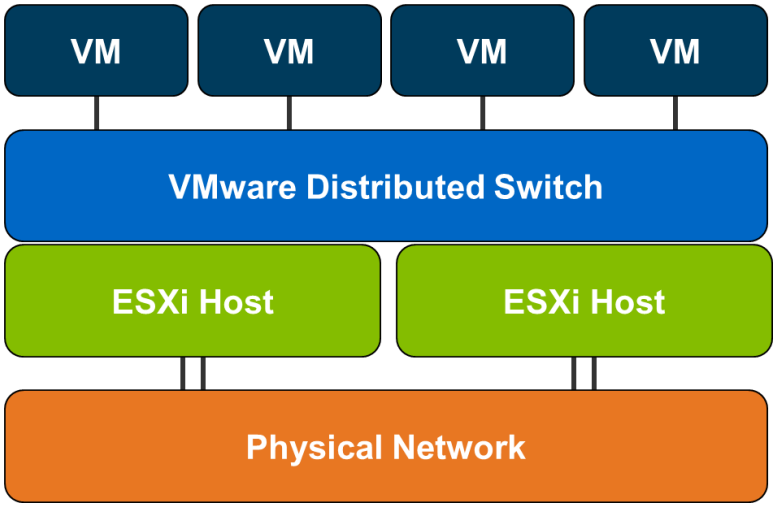


21. Click OK.

6.3 VMware vSphere 6.x Distributed Switch

A VMware distributed switch is a virtual switch that exists as a single combined switch across all ESXi hosts in a vSphere cluster. As a result, the same network configurations can be applied to all hosts, and virtual machines can retain their network configuration consistently, regardless of the physical host on which they currently reside. Figure 12 illustrates the VMware distributed switch architecture.

Figure 12) VMware distributed switch architecture.



6.4 VMware vSphere 6.x Distributed Switch Deployment Procedures

Table 16 describes the VMware vSphere distributed switch prerequisite.

Table 16) VMware vSphere 6.x distributed switch prerequisite.

Description
At least one vmnic per ESXi host must be available.

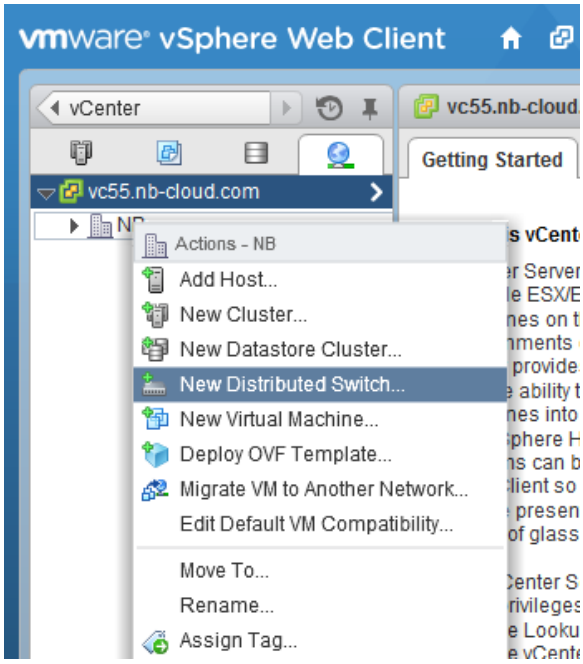
Transition Network to Distributed Switching

Create Distributed Switch

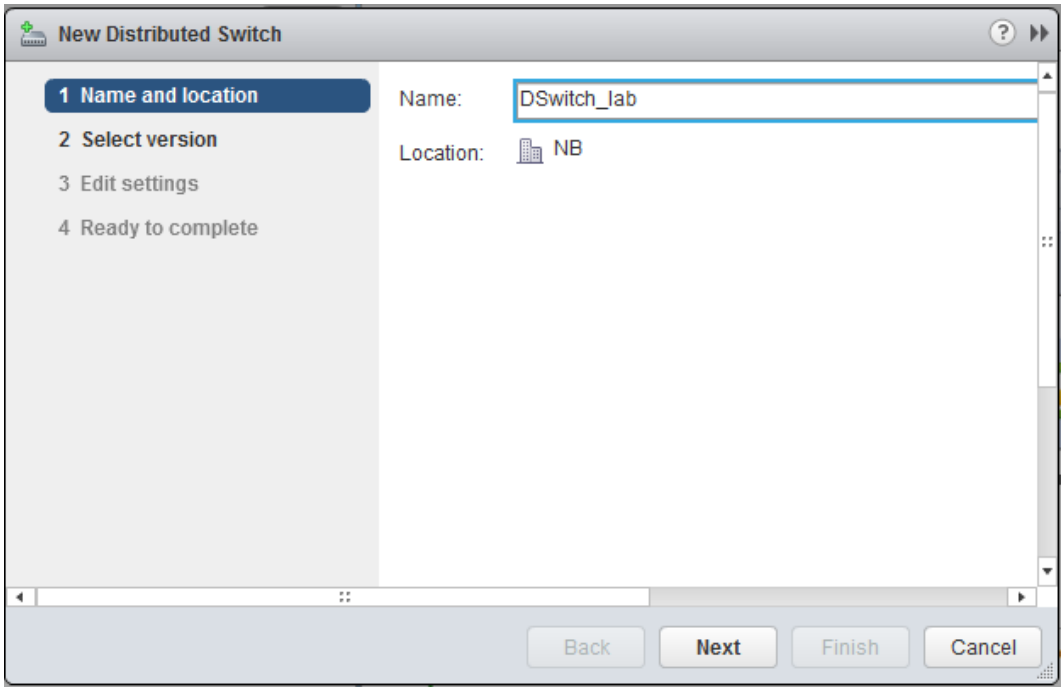
To create a distributed switch, complete the following steps:

1. In the vSphere Web Client, navigate to vCenter Home and select Networking.

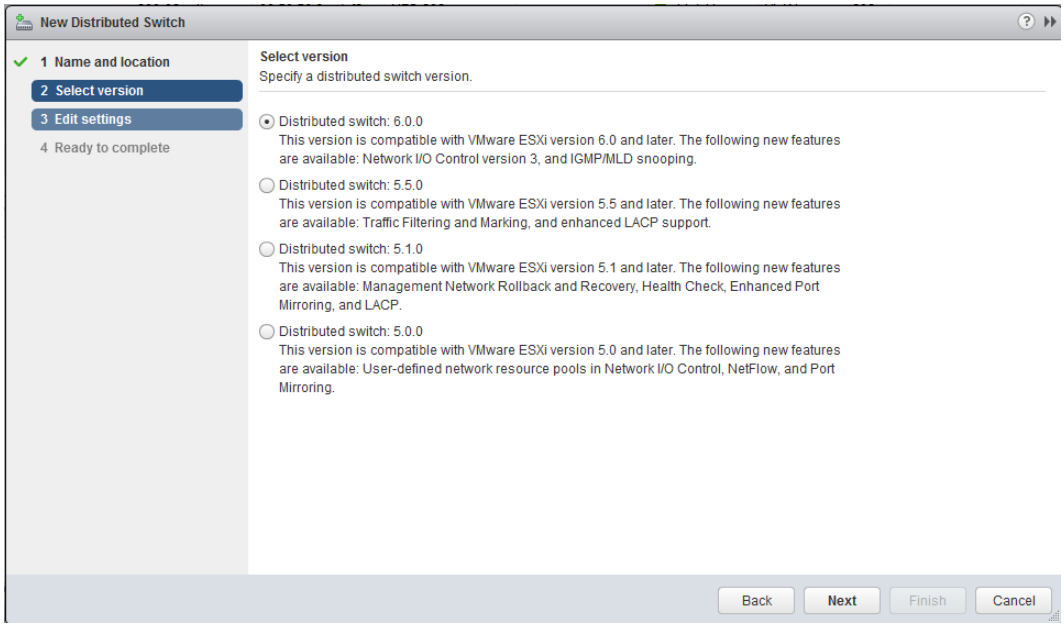
2. On the Networking page, right-click the data center that houses your vSphere infrastructure and select New Distributed Switch.



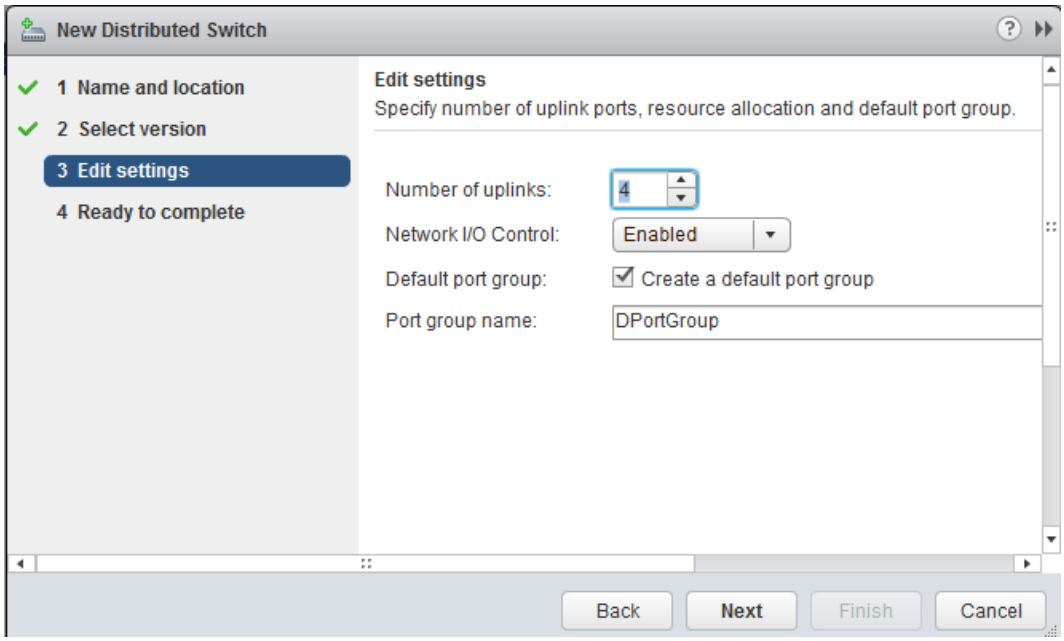
3. In the New Distributed Switch wizard, enter a name for the distributed switch. Click Next.



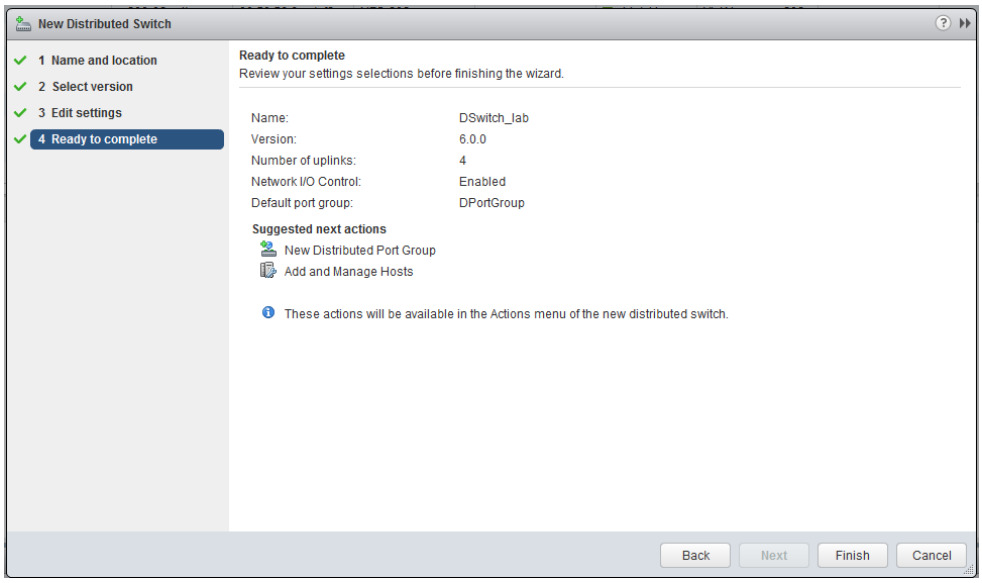
4. Select Distributed Switch 6.0.0. Click Next.



5. Specify the number of uplink ports for the distributed switch and verify that the Default Port Group checkbox is selected. The uplink ports number is the maximum number of ports allowed per host using this distributed switch. Click Next.



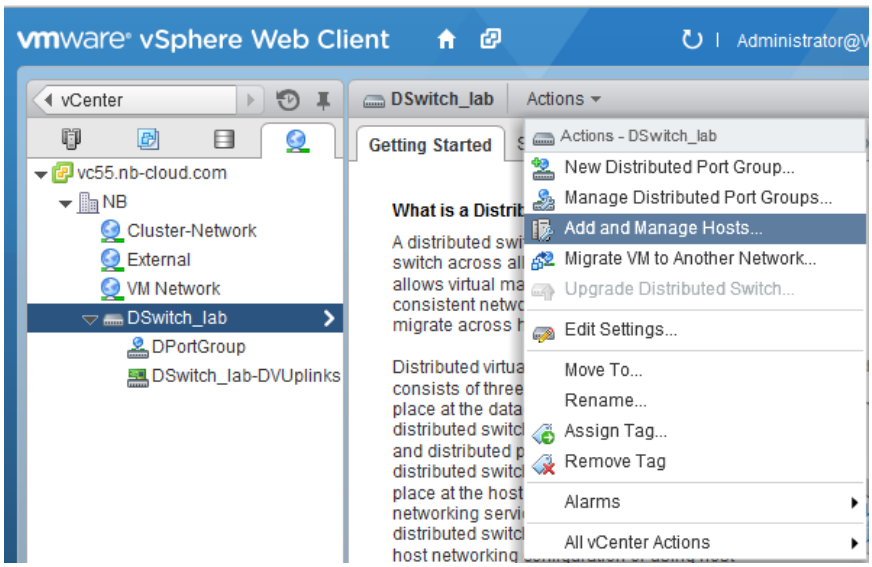
6. Review your settings and click Finish.



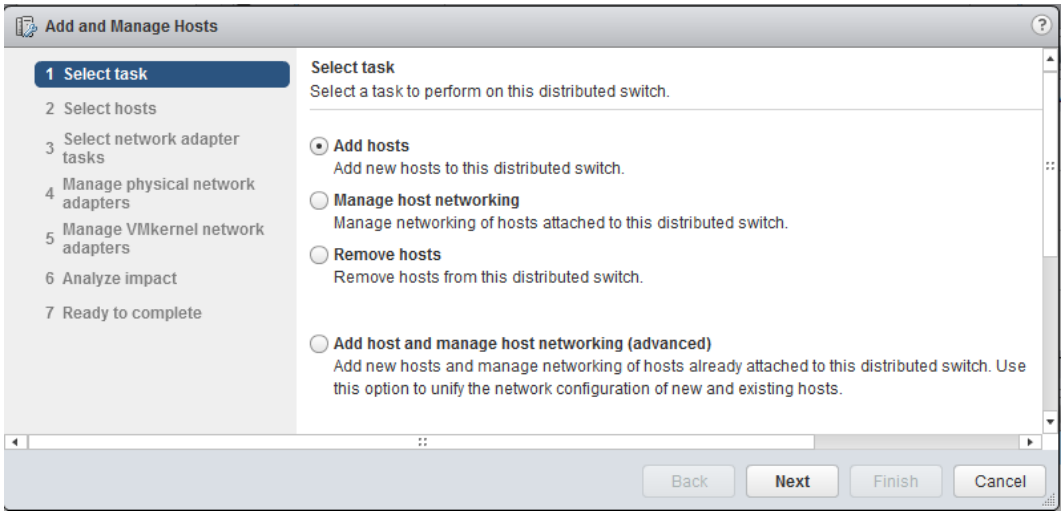
Add Hosts and Network Adapters to Distributed Switch

To add hosts and network adapters to the distributed switch, complete the following steps:

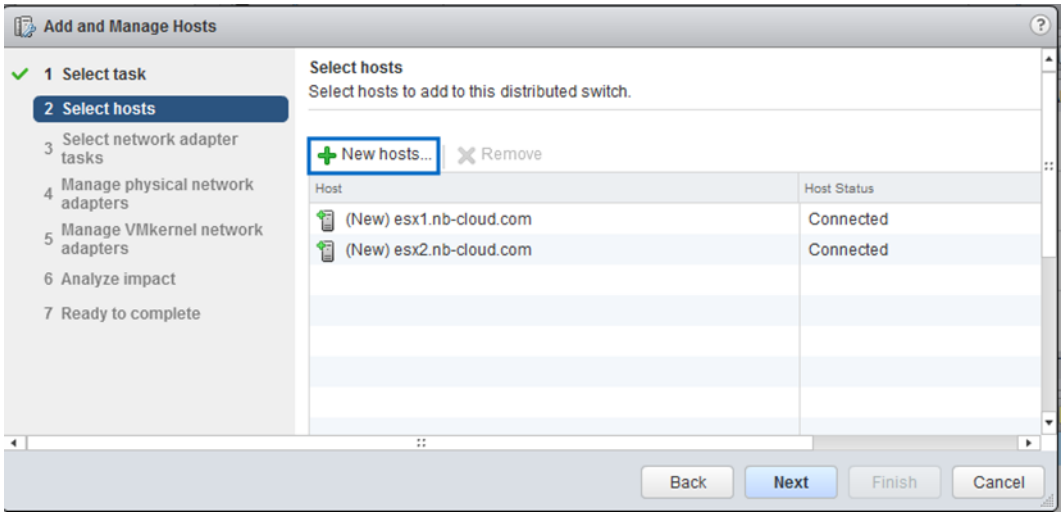
1. In the vSphere Web Client, select the distributed switch, click Actions, and then select Add and Manage Hosts.



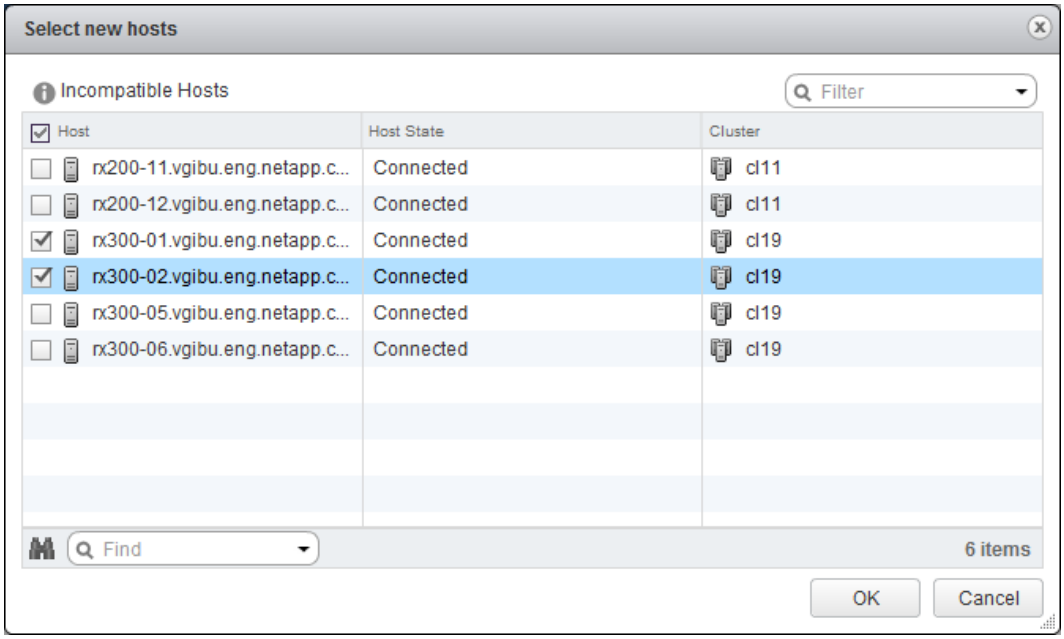
2. In the Add and Manage Hosts wizard, select Add Hosts. Click Next.



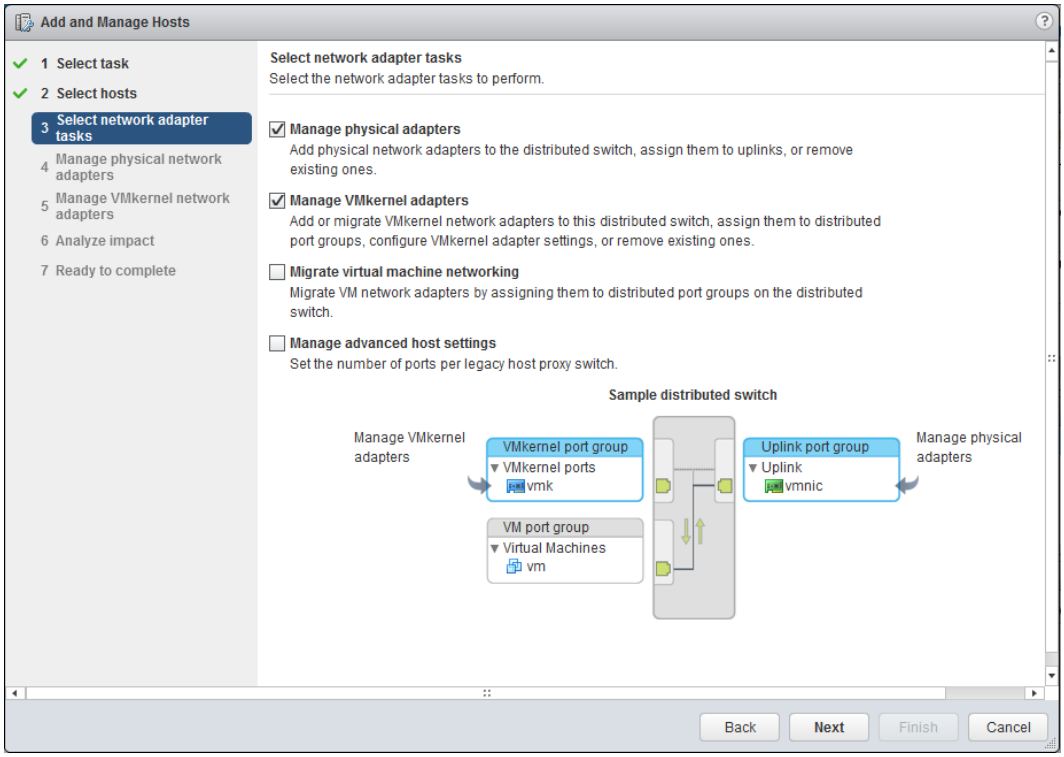
3. Click New Hosts to select the hosts to add to the distributed switch.



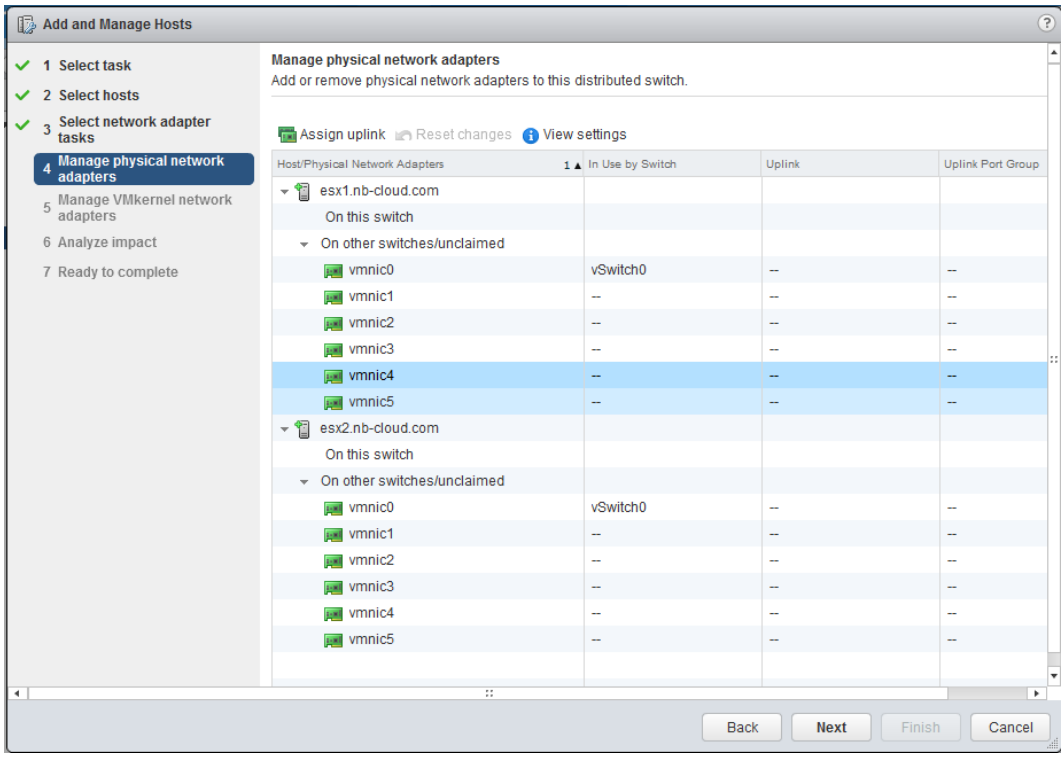
4. Select the hosts to add to the distributed switch and click OK.



5. Click Next.
6. Select the network adapter tasks to perform. The Manage Physical Adapters and Manage VMkernel Adapters options are selected by default. Click Next.

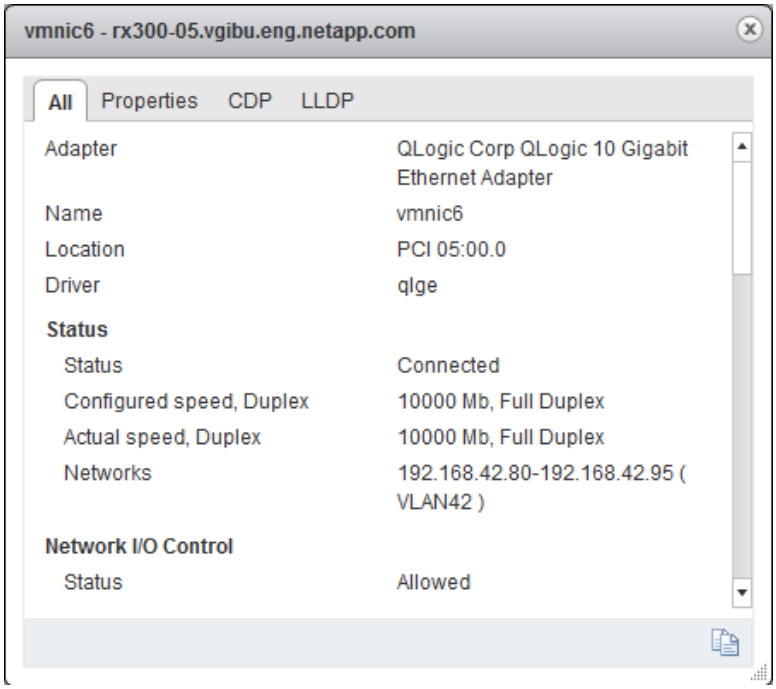


7. Add physical network adapters by assigning them uplink ports on the distributed switch.

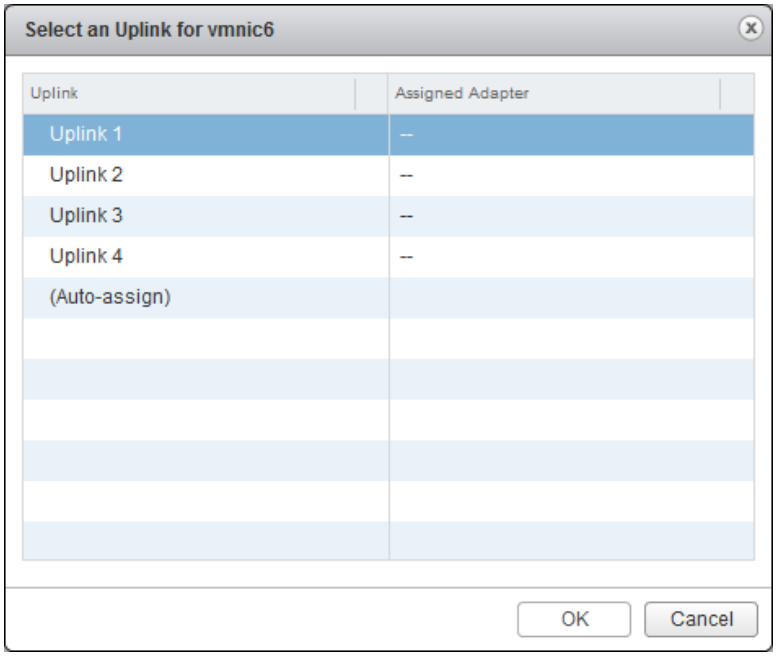


8. To verify the identity and view the information for a vmnic (physical NIC), complete the following steps:

a. Select a vmnic and click View Settings.
b. Optional: Scroll down to view the Cisco Discovery Protocol information.
c. Close the dialog box when you are finished.



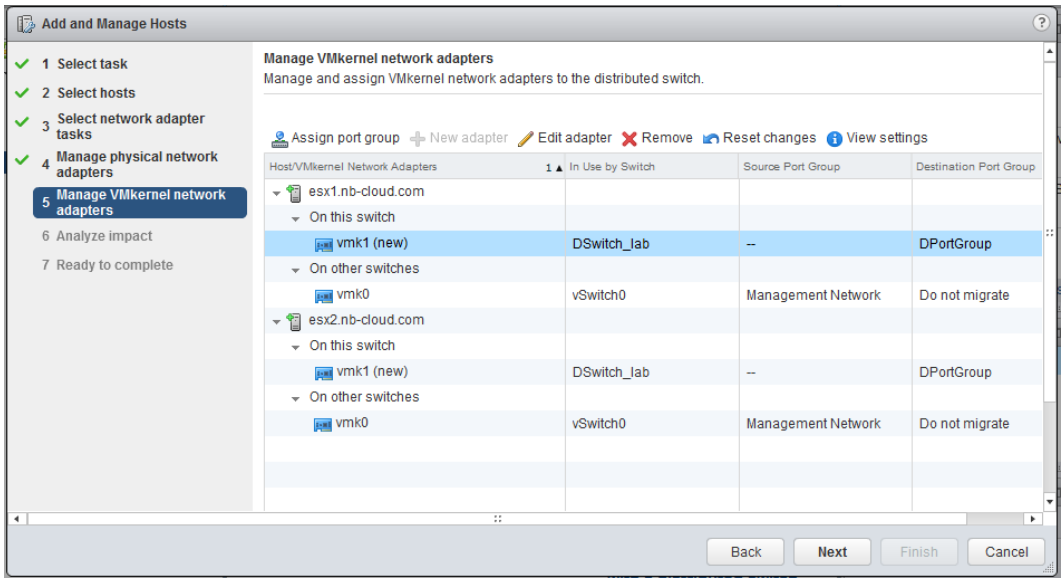
9. Select a vmnic and click Assign Uplink.
10. Select an uplink and click OK.



11. Repeat step 8 through step 10 for each vmnic to assign an uplink for each host.
12. Click Next.

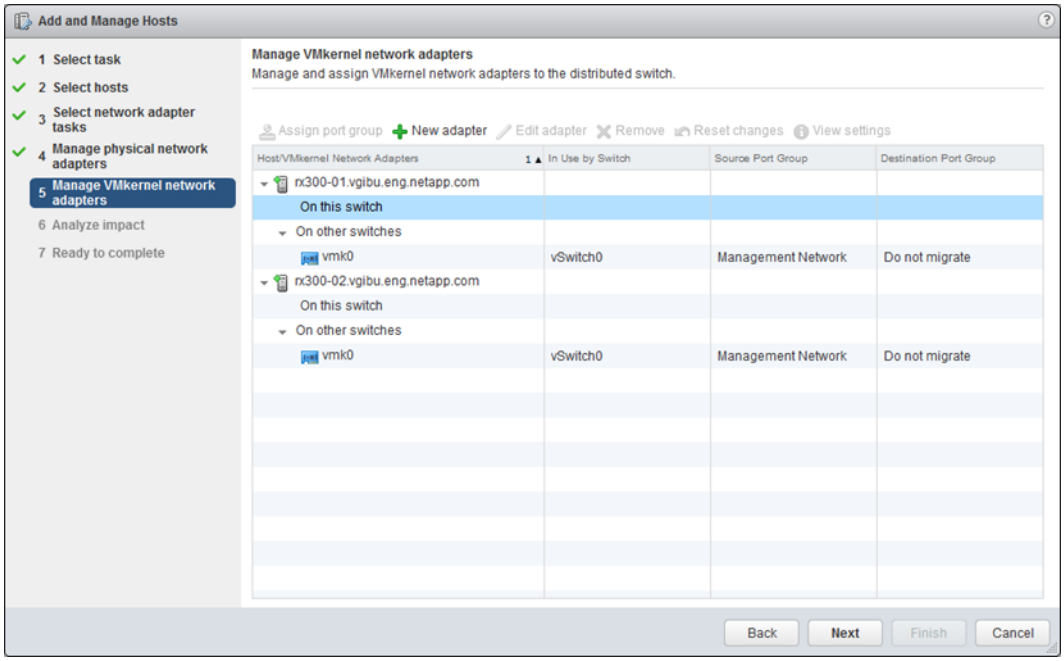
13. Select the existing VMkernel network adapters to assign to the distributed switch.

    Note: Do not migrate the management VMkernel network adapter from `vSwitch0` unless you are sure that the distributed switch can access the management network. Otherwise, the management access to the server might be lost.

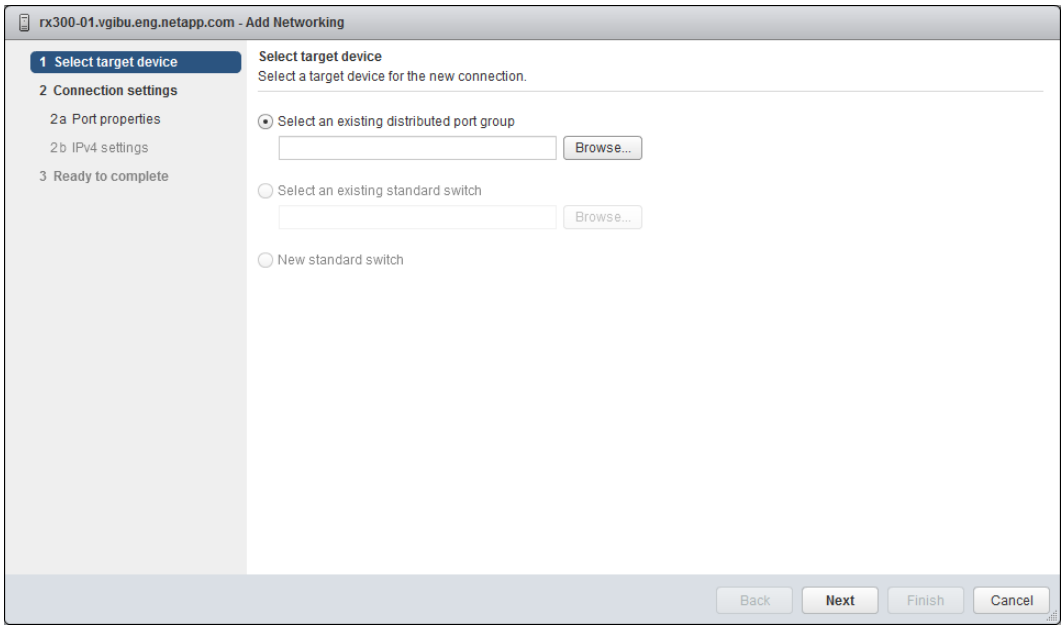


14. To add storage network VMkernel adapters, select an ESXi host and click New Adapter.

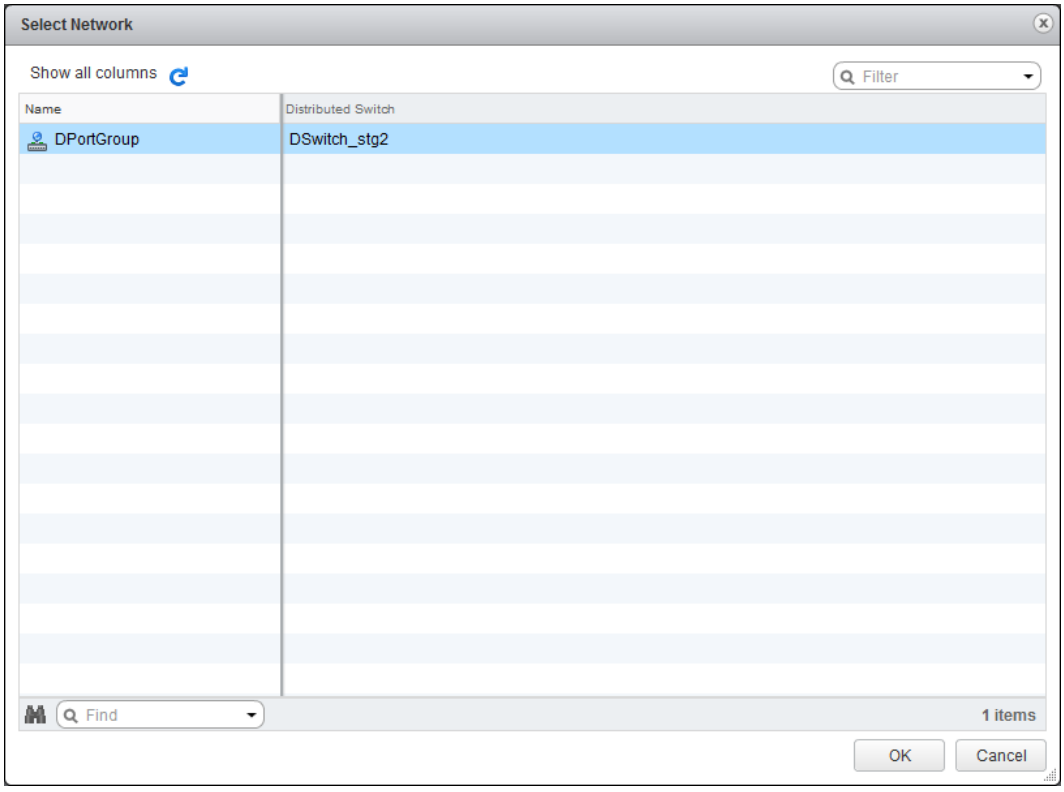
    Note: If iSCSI binding is used, a VMkernel adapter is required for each active NIC attached to the distributed switch.



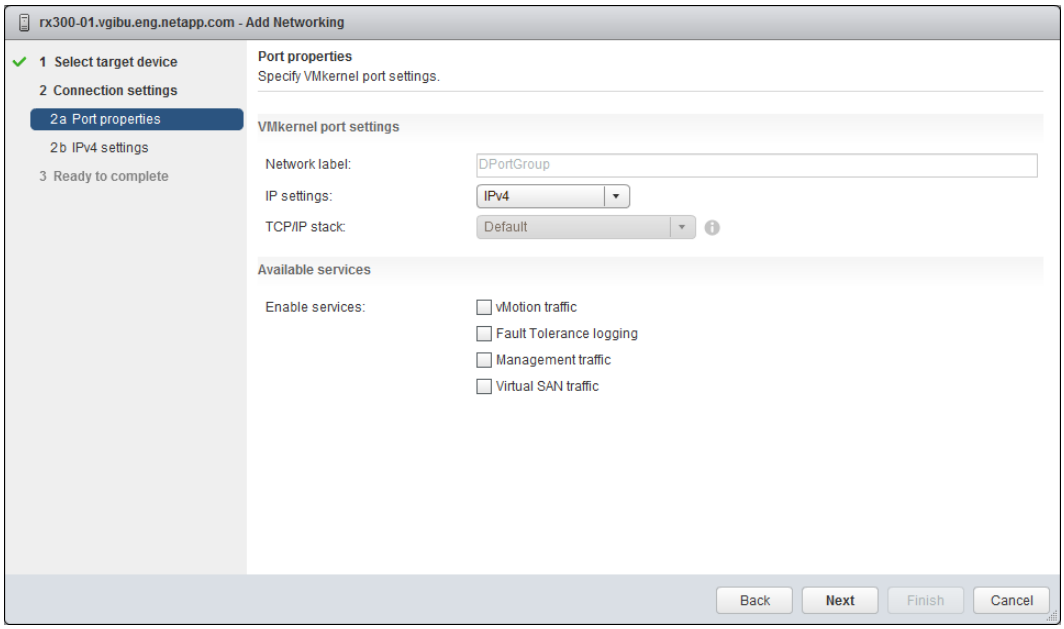
15. On the Select Target Device page, click Select an Existing Distributed Port Group and then click Browse.



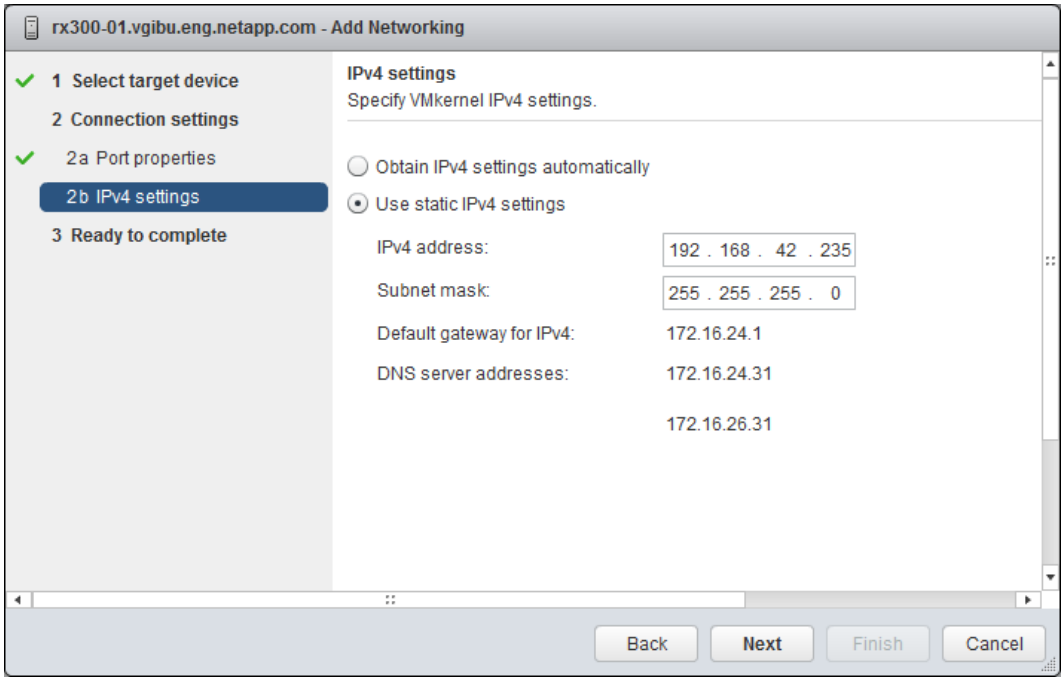
16. Select the port group for the VMkernel adapter and click OK.



17. Click Next.
18. Specify the port properties. The Available Services options are not required for VMkernel adapters that are only used for storage traffic. Click Next.



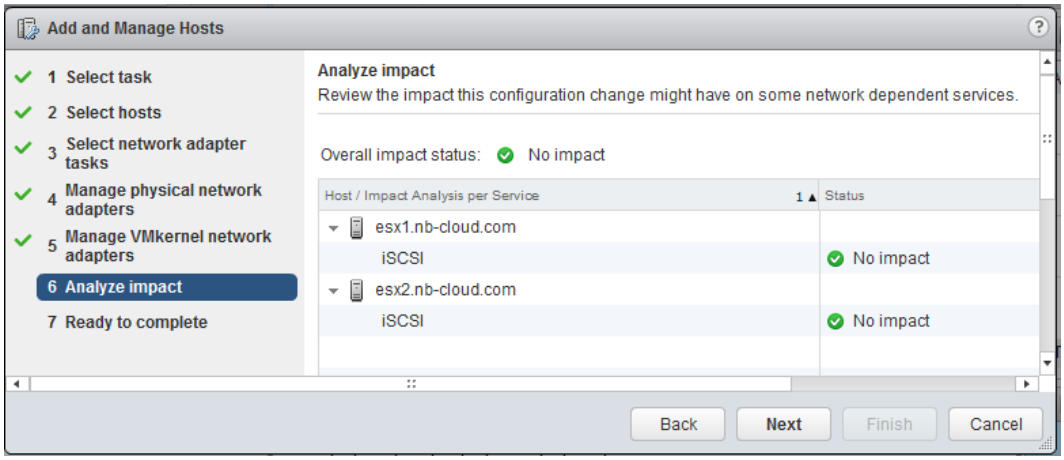
19. Enter the IP address and netmask. Click Next.



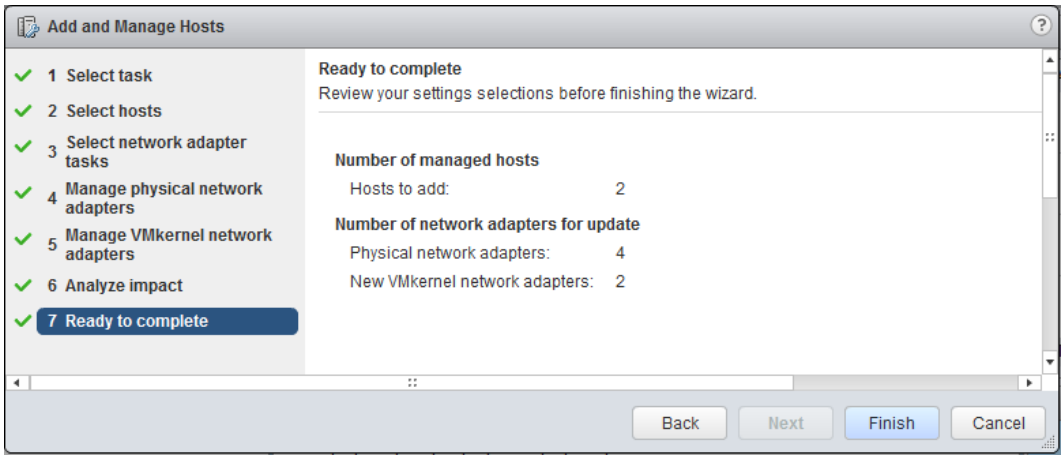
20. Review the VMkernel adapter settings. Click Finish.
21. Repeat step 14 through step 20 to add the necessary VMkernel adapters for each host.

    Note: When using features such as iSCSI port binding or the route based on physical NIC load option, two or more VMkernel adapters per host might be required.

22. Click Next.
23. Review the impact of this configuration change to network-dependent services. If the impact is `Important` or `Critical`, then troubleshoot the dependent services and proceed with the configuration. Click Next.



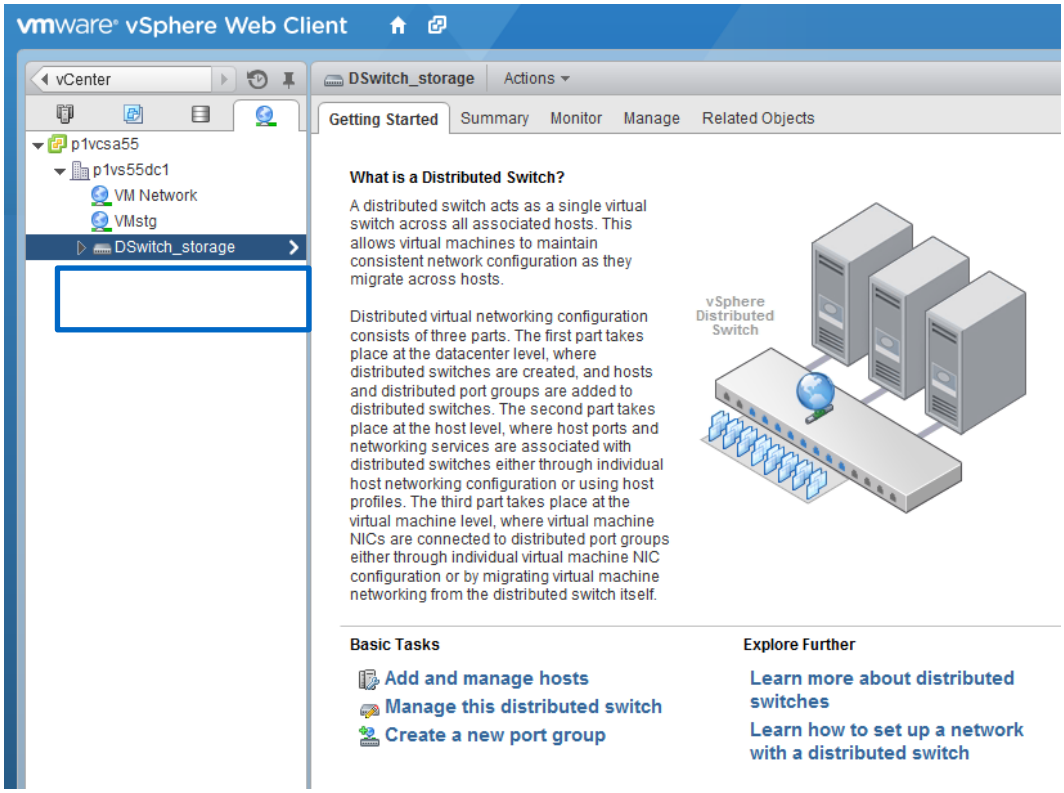
24. Review your configuration settings and click Finish.



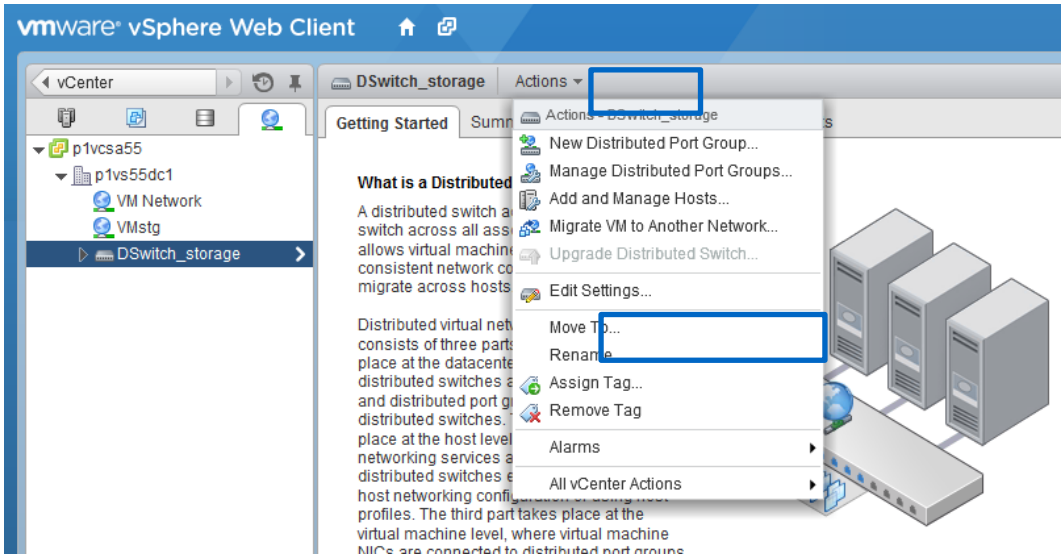
Set MTU on Distributed Switch

To set the MTU (also known as jumbo frames for any MTU greater than 1,500), complete the following steps:

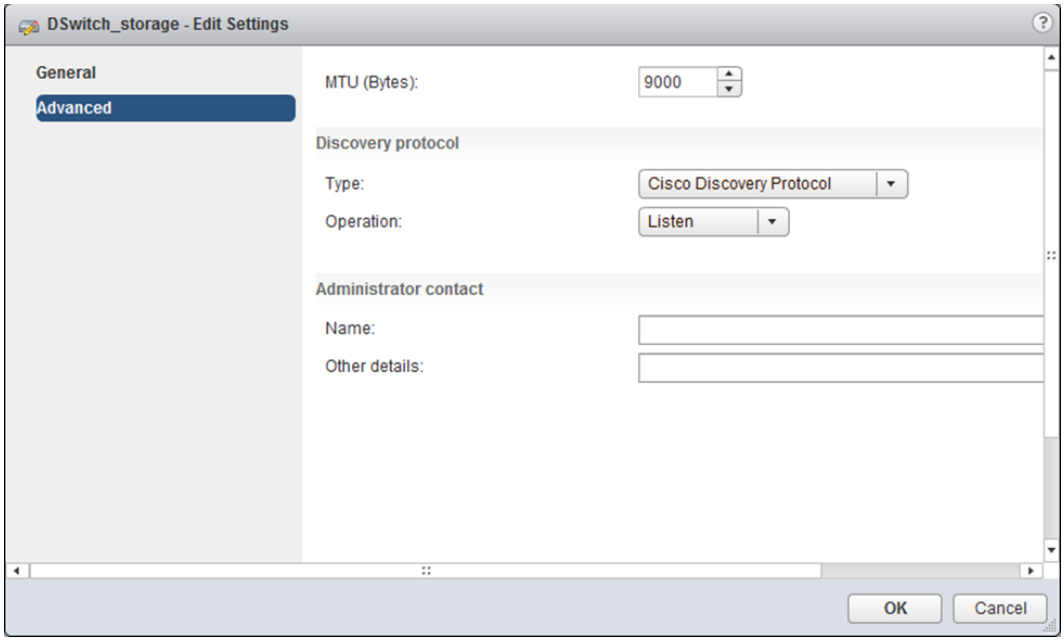
1. In the vSphere Web Client, navigate to vCenter Home and select Networking. Select the distributed switch.



2. Click Actions and select Edit Settings.



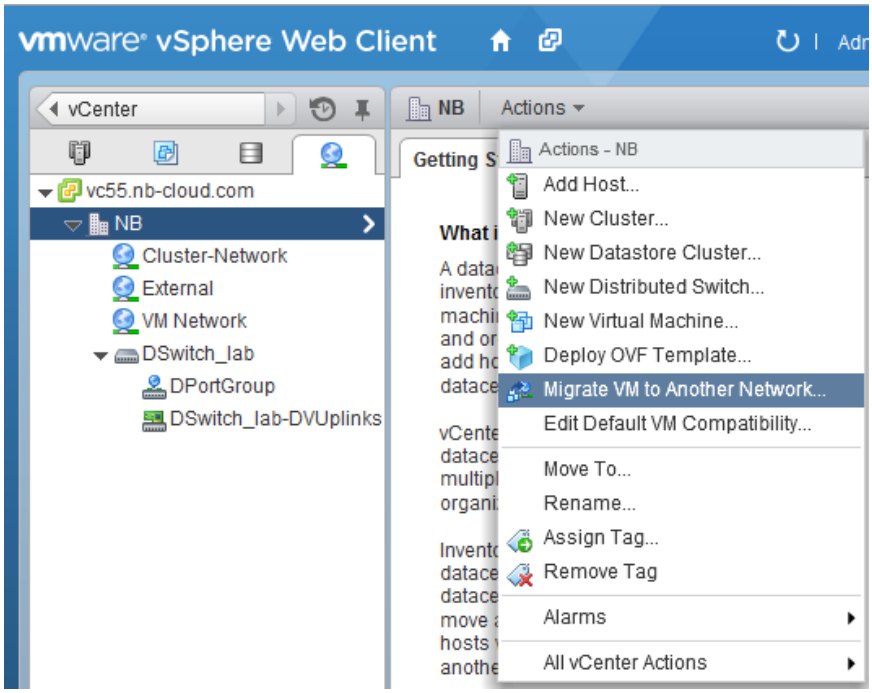
3. Click Advanced.
4. Set the MTU to `9000`.
5. Click OK.



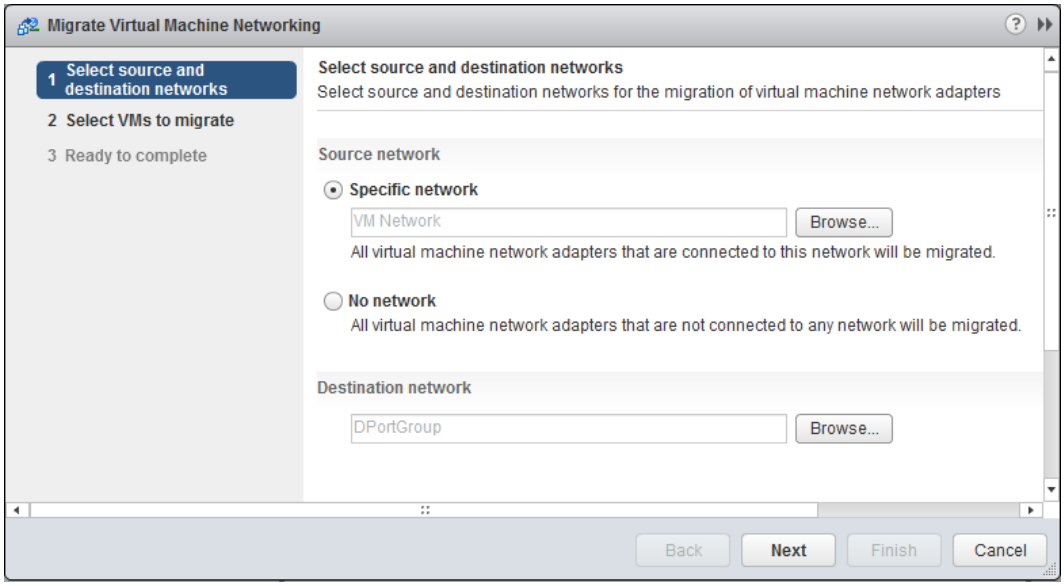
Migrate VMs to Distributed Switch

To migrate VMs to the distributed switch, complete the following steps:

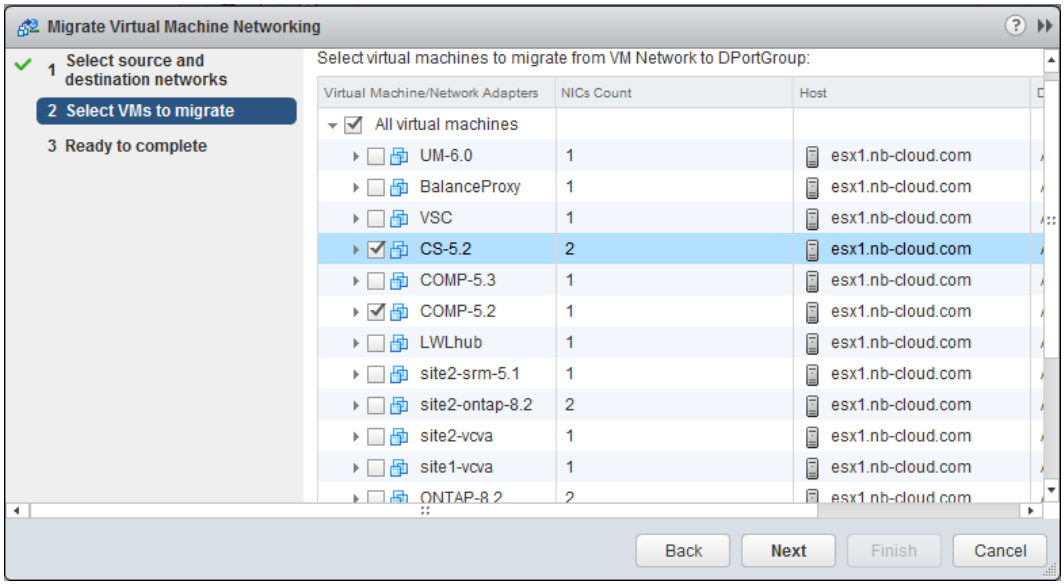
1. In the vSphere Web Client, click Actions and select Migrate VM to Another Network.



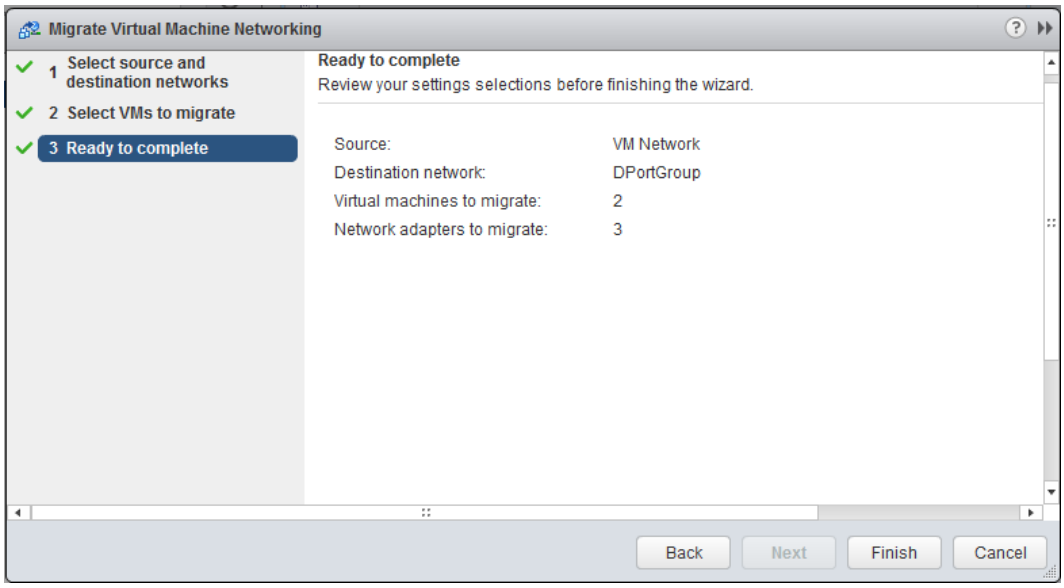
2. Select the source and the destination network for the migration. Click Next.



3. Select the VMs to migrate. Click Next.



4. Review your settings and click Finish.



5. Verify that the VMs were migrated to the distributed switch.

7 Storage and Datastores

7.1 VMware vSphere 6.x Datastores

Four protocols are used to connect VMware vSphere 6 to datastores on NetApp volumes:

- FC
- FCoE
- iSCSI
- NFS

FC, FCoE, and iSCSI are block-based protocols that use vSphere Virtual Machine File System (VMFS) to store VMs inside NetApp LUNs that are contained in a NetApp volume. NFS is a file-based protocol that places VMs into datastores (which are simply NetApp volumes) without the need for VMFS.

Datastore Comparison Tables

Table 17 compares the features available with each type of datastore and storage protocol.

Table 17) Supported datastore features.

Capability/Feature	FC/FCoE	iSCSI	NFS
Format	VMFS or raw device mapping (RDM)	VMFS or RDM	NetApp WAFL® (Write Anywhere File Layout)
Maximum number of datastores or LUNs	256	256	256 **Note:** Default NFS. MaxVolumes is 8. Use VSC to efficiently set to 256.
Maximum datastore size	64TB	64TB	100TB

Capability/Feature	FC/FCoE	iSCSI	NFS
			Note: NAS datastores greater than 16TB require 64-bit aggregates.
Maximum LUN or NAS file system size	64TB **Note:** Maximum Data ONTAP file and LUN size is 16TB.	64TB **Note:** Maximum Data ONTAP file and LUN size is 16TB.	100TB
Maximum datastore file size (for VMDKs) **Note:** For all of these protocols, vSphere version 5.5 and VMFS 5 are required to create a >2TB vdisk.	62TB	62TB	62TB **Note:** Maximum Data ONTAP file and LUN size is 16TB.
Optimal queue depth per LUN or file system	64	64	N/A
Available link speeds	4Gb, 8Gb, and 16Gb FC, and 10GbE	1GbE and 10GbE	1GbE and 10GbE

Table 18 compares the storage-related functionality of VMware features across different protocols.

Table 18) Supported VMware storage-related functionalities.

Capacity/Feature	FC/FCoE	iSCSI	NFS
vMotion	Yes	Yes	Yes
Storage vMotion	Yes	Yes	Yes
VMware HA	Yes	Yes	Yes
Distributed Resource Scheduler (DRS)	Yes	Yes	Yes
VMware vStorage APIs for Data Protection (VADP) enabled backup software	Yes	Yes	Yes
Microsoft Cluster Service (MSCS) within a VM	Yes, by using RDM for shared LUNs	RDM or initiator in guest operating system	Not supported
Fault tolerance	Yes, with eager-zeroed thick virtual machine disks (VMDKs) or virtual mode RDMs **Note:** NetApp SnapDrive® for Windows software does not support virtual-mode	Yes, with eager-zeroed thick VMDKs or virtual mode RDMs **Note:** NetApp SnapDrive for Windows software does not support virtual-mode RDMs.	Yes, with eager-zeroed thick VMDKs

Capacity/Feature	FC/FCoE	iSCSI	NFS
	RDMs.		
Site Recovery Manager	Yes	Yes	Yes
Thin-provisioned VMs (virtual disks)	Yes	Yes	Yes **Note:** This is the default setting for all VMs on NetApp NFS when not using VAAI.
VMware native multipathing	Yes	Yes	N/A
Boot from SAN	Yes	Yes, with host bus adapters (HBAs)	N/A
AutoDeploy	Yes	Yes	Yes

Table 19 compares the NetApp storage management features that are supported across different protocols.

Table 19) Supported NetApp storage management features.

Capability/Feature	FC/FCoE	iSCSI	NFS
Data deduplication	Savings in the array	Savings in the array	Savings in the datastore
Thin provisioning	Datastore or RDM	Datastore or RDM	Datastore
Resize datastore	Grow only	Grow only	Grow, autogrow, and shrink
OnCommand Balance	Yes	Yes	Yes
SnapDrive (in guest)	Yes	Yes	Yes
Monitoring and host configuration for VSC 4.1	Yes	Yes	Yes
VM backup and recovery by using VSC 4.1	Yes	Yes	Yes
Provisioning and cloning by using VSC 4.1	Yes	Yes	Yes

Table 20 compares the backup features that are supported across different protocols.

Table 20) Supported backup features.

Capability/Feature	FC/FCoE	iSCSI	NFS
NetApp Snapshot backups	Yes	Yes	Yes
SRM supported by replicated backups	Yes	Yes	Yes
Volume SnapMirror	Yes	Yes	Yes

Capability/Feature	FC/FCoE	iSCSI	NFS
VMDK image access	VADP-enabled backup software	VADP-enabled backup software	VADP-enabled backup software, vSphere Client, and vSphere Web Client datastore browser
VMDK file-level access	VADP-enabled backup software, Windows only	VADP-enabled backup software, Windows only	VADP-enabled backup software and third-party applications
NDMP granularity	Datastore	Datastore	Datastore or VM

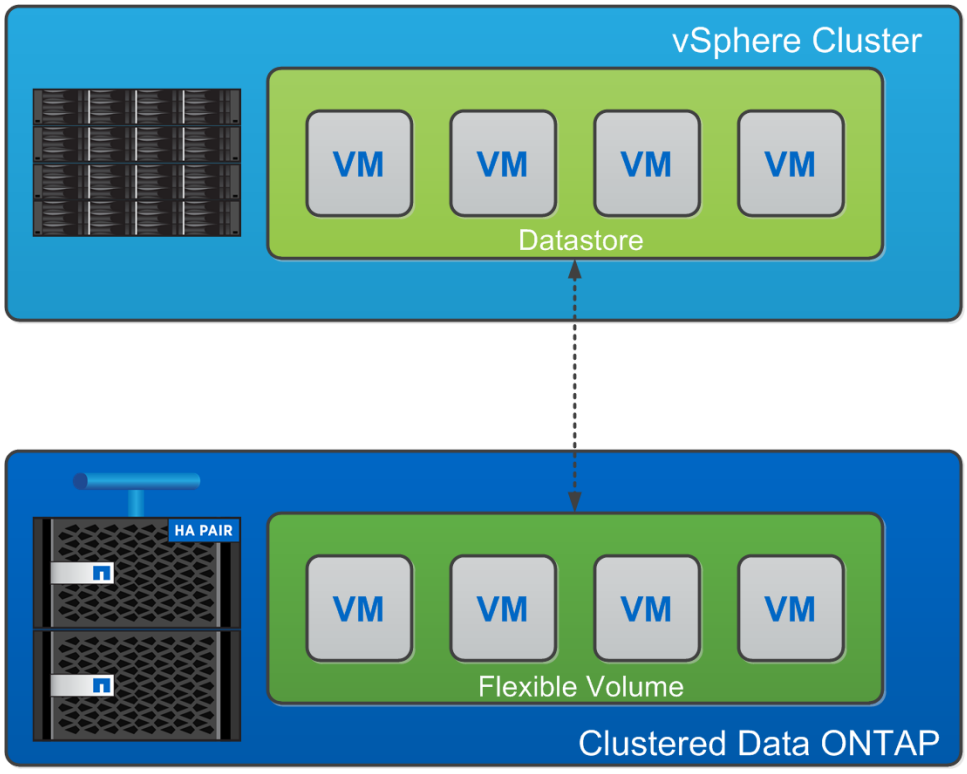
7.2 VMware vSphere 6 NFS Datastores on Clustered Data ONTAP

VMware vSphere allows customers to leverage enterprise-class NFS arrays to provide concurrent access to datastores for all of the nodes in an ESXi cluster. This access method is very similar to VMFS access. The NetApp NFS offers high performance, low per-port storage costs, and advanced data management capabilities.

Figure 13 shows an example of this configuration.

Note: The storage layout is similar to the layout of a VMFS datastore, but each virtual disk file has its own I/O queue managed directly by the NetApp FAS system.

Figure 13) vSphere cluster connected to an NFS datastore.



NFS Datastores on NetApp

Deploying VMware with the NetApp advanced NFS results in a high-performing, easy-to-manage implementation that provides VM-to-datastore ratios that cannot be obtained with block-based storage protocols. This architecture can result in a tenfold increase in datastore density with a correlated

reduction in the number of datastores. With NFS, the virtual infrastructure receives operational savings because fewer storage pools are provisioned, managed, backed up, replicated, and so forth.

Through NFS, customers receive an integration of VMware virtualization technologies with WAFL, the NetApp advanced data management and storage virtualization engine. This integration provides transparent access to the following VM-level storage virtualization offerings:

- Deduplication of production data
- Immediate, zero-cost VM and datastore clones
- Array-based thin provisioning
- Automated policy-based datastore resizing
- Direct access to array-based Snapshot copies
- Ability to offload tasks to NetApp storage by using the NFS VMware VAAI plug-in

NetApp also provides integrated tools, such as VSC and Storage Replication Adapter (SRA) for VMware SRM.

7.3 Clustered Data ONTAP Export Policies

Historically, in Data ONTAP operating in 7-Mode, NFS exports are defined on a per-volume, qtree, or directory basis. This is managed through an entry in the `/etc/exports` file for each volume. Even when multiple volumes are intended to be used in the same way by many clients, the desired export rules are identical, and each volume still requires its own unique entry. Whenever a new ESXi host is added to the VMware cluster, each volume export must be updated to add the hostname or IP address of the new host. This can be an extremely tedious process to manage manually and is one of the biggest benefits of using the NetApp VSC to manage storage connectivity to your ESXi hosts.

Note: In Data ONTAP operating in 7-Mode, servers can be specified as a list of one or more specific names, as IP addresses, or as a complete subnet range (for example, 192.168.42.0/24).

In a storage system running clustered Data ONTAP, volume exports are now managed entirely through policy and are restricted by export policies that apply in the scope of a given SVM. An export policy contains a set of rules that define access permissions and authentication types required for specific clients to access volumes. Each volume is associated with exactly one export policy.

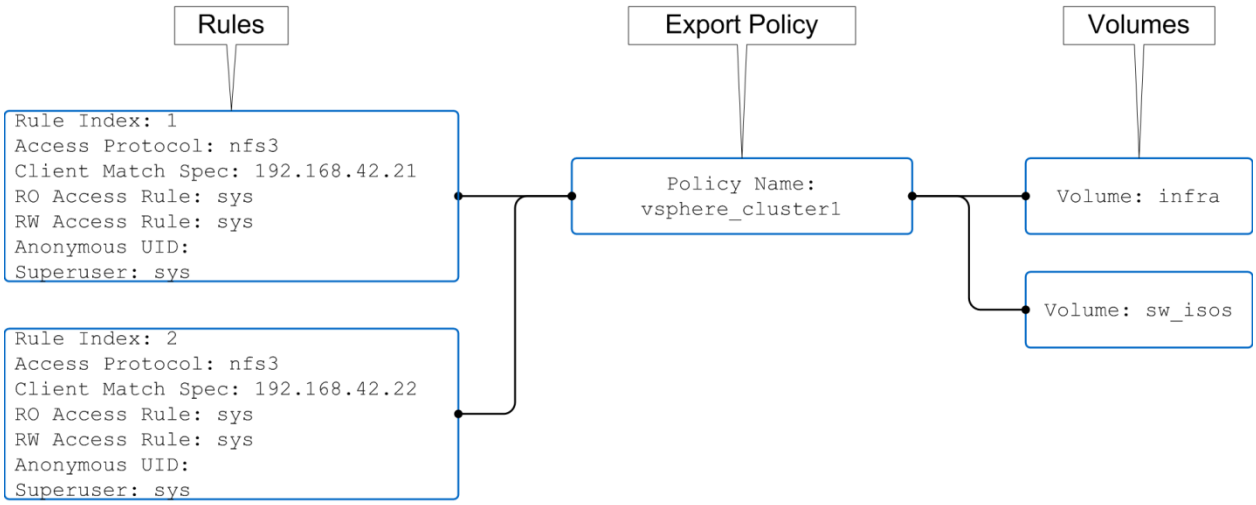
Several volumes can be associated with the same export policy. If all volumes used by the vSphere environment (or at least those used by each VMware cluster) share the same export policy, then all servers see the set of volumes in the same way.

Whenever a new server is added, a rule for that server is added to the export policy. The rule for the new server includes a client-match pattern (or simply an IP address) for the new server; the access protocol; and permissions and authentication methods for read/write, read-only, superuser, anonymous user, and other options that do not directly affect vSphere. When a new volume is added and other volumes are already in use as datastores, the same export policy used for the existing volumes can also be used for the newly added volume.

In clustered Data ONTAP, client match is resolvable by host name, by FQDN or IP for a single client, or by subnet range. The subnet range must be properly masked (0 for host bits). For example, 192.168.42.21/24 does not work because the 24-bit mask means that the last octet (.21) must be 0.

ESX uses the sys (UNIX) security style and requires the root mount option to execute VMs. In clustered Data ONTAP, this option is referred to as superuser. The only protocol currently supported for NAS with ESXi is NFSv3, indicated as `nfs3` in the export policy, as shown in the example in Figure 14. When the superuser option is used, it is not necessary to specify the anonymous user ID.

Figure 14) Export policy, rules, and volumes.



If the NFS vStorage APIs for the VAAI plug-in are used on ESXi hosts with clustered Data ONTAP, the protocol should be set as `NFS` when the export policy rule is created or modified. The NFSv4 protocol is required for VAAI copy offload to work, and selecting the NFS protocol automatically includes both the NFSv3 and the NFSv4 versions. The following example creates an export policy for the `Infra` SVM (referred to as Vserver in the CLI) and a default policy with 1 as the index number of the rule.

```
clus-1::> vserver export-policy rule modify -vserver Infra -policyname default -ruleindex 1 -
protocol nfs
```

NFS datastore volumes are junctioned from the root volume of the SVM; therefore, ESXi must also have access to the root volume in order to navigate and mount datastore volumes. The export policy for the root volume, and for any other volumes in which the datastore volume's junction is nested, must include a rule or rules for the ESXi servers granting them read-only superuser access. If NFS VAAI is used, then the protocol must allow for both the NFSv3 and NFSv4 versions.

```
                                    Vserver: vmw_prod
                                Policy Name: root_vol
                                 Rule Index: 1
                            Access Protocol: nfs
                           Client Match Spec: 192.168.42.21
                              RO Access Rule: sys
                              RW Access Rule: never
User ID To Which Anonymous Users Are Mapped: 65534
                    Superuser Security Flavors: sys
                Honor SetUID Bits In SETATTR: true
                   Allow Creation of Devices: true
```

For more information, refer to KB 1014234: How to Configure Clustered Data ONTAP to Allow for VAAI over NFS.

Best Practices

- Use VSC to provision datastores because VSC simplifies management of export policies automatically.
- When creating datastores for VMware clusters with VSC, select the cluster rather than a single ESX server.
- Use the VSC mount function to apply existing datastores to new servers.
- When not using VSC, use a single export policy for all servers.

7.4 Clustered Data ONTAP Junction Paths

Clustered Data ONTAP offers a namespace structure within an SVM. Each volume in the SVM can be mounted internally so that all volumes are seen as a single tree structure. These internal mount points are referred to as junction paths. A client can mount any point in the namespace and traverse the entire structure from that point down, whether the path is contained within a single volume or traverses many volumes.

VMware vSphere creates a directory for each VM at the root of the datastore. For example, the path to the VMX file, which is the main descriptor of a VM, is represented from the perspective of vCenter and of vSphere Web Client in the following way:

```
[datastore] vmname/vmname.vmx
```

It is presented in the ESXi CLI in the following way:

```
/vmfs/volumes/datastore/vmname/vmname.vmx
```

There is no easy way to make vSphere create VMs in directories any deeper. This behavior limits the usefulness of nested junction paths. In fact, for vSphere to use multiple storage volumes for NFS, it must mount each volume as a separate datastore, regardless of the namespace hierarchy of the storage. For this reason, the best practice within the SVM is to simply mount the junction path for volumes for vSphere at the root of the SVM. This is the behavior of the provisioning and cloning capability in VSC. Not having nested junction paths also means that no volume is dependent on any volume other than the root volume and that taking a volume offline or destroying it, even intentionally, does not affect the path to other volumes.

Block protocols (iSCSI, FC, and FCoE) access LUNs by using target worldwide port names (WWPNs) and LUN IDs. The path to LUNs inside the storage is meaningless to the block protocols and is not presented anywhere in the protocol. Therefore, a volume that contains only LUNs does not need to be internally mounted at all, and a junction path is not necessary.

Best Practices
• Use VSC to provision datastores because VSC simplifies management of junction paths automatically. • When not using VSC, mount volumes on junction paths directly on the root volume by using the name of the volume as the junction path. • Do not use junction paths for volumes that contain LUN datastores.

7.5 System Manager Setup for NFS and NAS LIFs

Table 21 describes prerequisites for the setup of NetApp System Manager for NFS and LIFs.

Table 21) System Manager setup for NFS and LIFs prerequisite.

Description
System Manager is bundled with clustered Data ONTAP 8.3. For versions of clustered Data ONTAP earlier than 8.3, System Manager 3.0 or later is required.

After a new SVM is created, the setup wizard walks the user through the steps to configure protocol access. All of the protocols that were selected for the new SVM during the creation process are listed and can be configured separately.

Best Practice
For each NFS datastore, provision a LIF for the SVM on each node.

Previous versions of System Manager created a default export policy that allowed superuser access. System Manager 3.0 no longer sets any rules in the default export policy. However, when provisioning datastores using VSC, VSC creates the necessary rules in the default export policy.

7.6 Supported NFS Versions

vSphere initially supported NFSv3 in VMware Virtual Infrastructure 3. NFSv4.1 is supported in vSphere 6.0. NFSv3 for Virtual Infrastructure 3 (VI3) was only available over TCP with no advanced features. With NFSv4.1, advanced features, such as Kerberos security, are available. Where NFSv3 uses client-side locking, NFSv4.1 uses server-side locking. Although a FlexVol volume can be exported through both protocols, ESXi can only mount through one protocol. This single protocol mount does not preclude other ESXi hosts from mounting the same datastore through a different version. Make sure to specify the protocol version to use when mounting so that all hosts use the same version and, therefore, the same locking style.

Table 22) NFS versions and supported features.

vSphere 6.0 Features	NFSv3	NFSv4.1
vMotion and Storage vMotion	Yes	Yes
High availability	Yes	Yes
Fault tolerance	Yes	Yes
DRS	Yes	Yes
Host profiles	Yes	Yes
Storage DRS	Yes	No
Storage I/O control	Yes	No
SRM	Yes	No
Virtual volumes	Yes	No
Hardware acceleration	Yes	No
Kerberos authentication	No	Yes

Best Practices
• Do not mix NFS versions across hosts. If possible, use host profiles to check compliancy. • Because there is no automatic datastore conversion, create a new NFSv4.1 datastore and use Storage vMotion to migrate VMs to the new datastore.

Create New SVM Configured for NFS

To create a new SVM that is configured for NFS, complete the following steps:

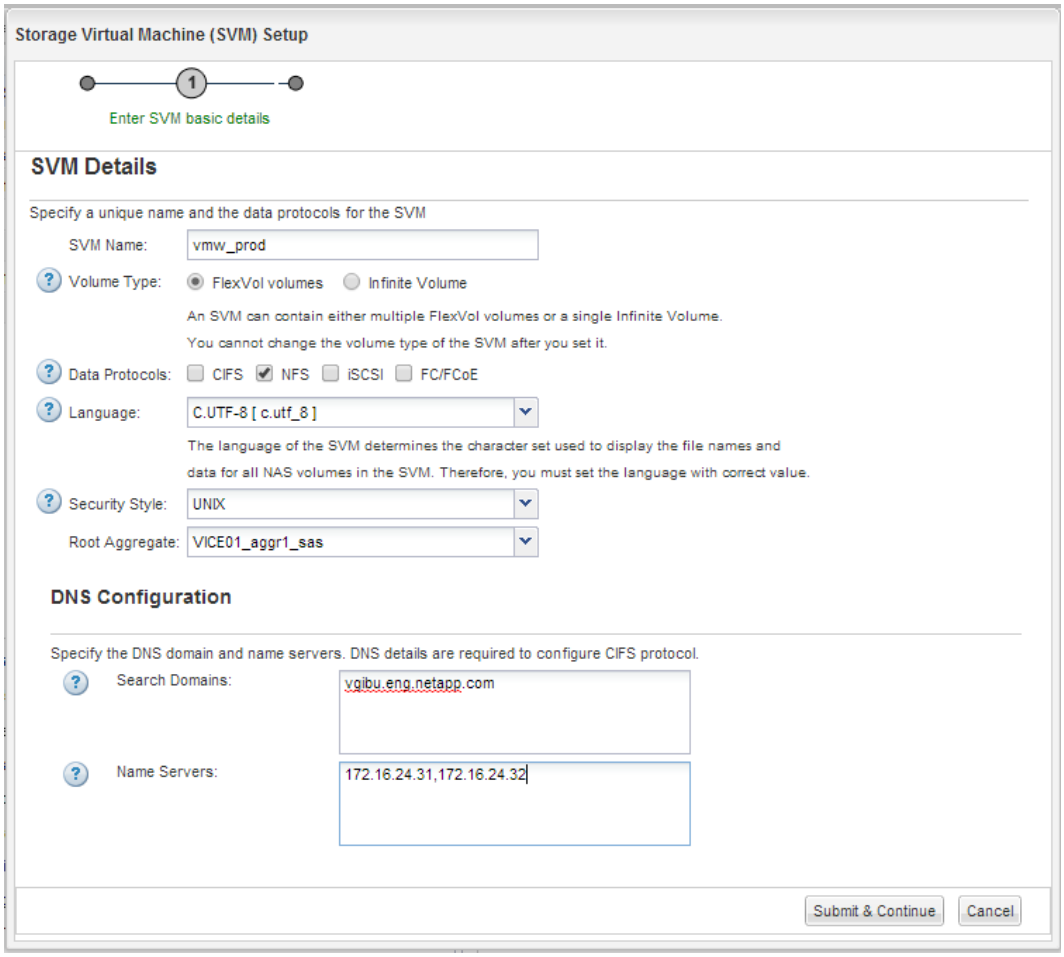
1. To prepare for the NFS configuration task, create a list showing the planned volumes and which aggregates and nodes host them, the home ports for the LIFs, the planned LIF names, and the IP

addresses and subnet information. This information is used later in this procedure. An example list, in the form of a table, is shown here.

Datastore/Volume	Aggregate	Node	Network Port	LIF	IP Address/Netmask
infra	n3a1	vice-03	a0a	nfs1	172.1.2.104 / 255.255.255.0
Software	n2a1	vice-02	a0a	nfs2	172.1.2.105 / 255.255.255.0
test	n4a1	vice-01 (moves)	a0a	nfstest	172.1.2.106 / 255.255.255.0
<Future>	<tbd>	vice-01	a0a	nfs3	172.1.2.107 / 255.255.255.0

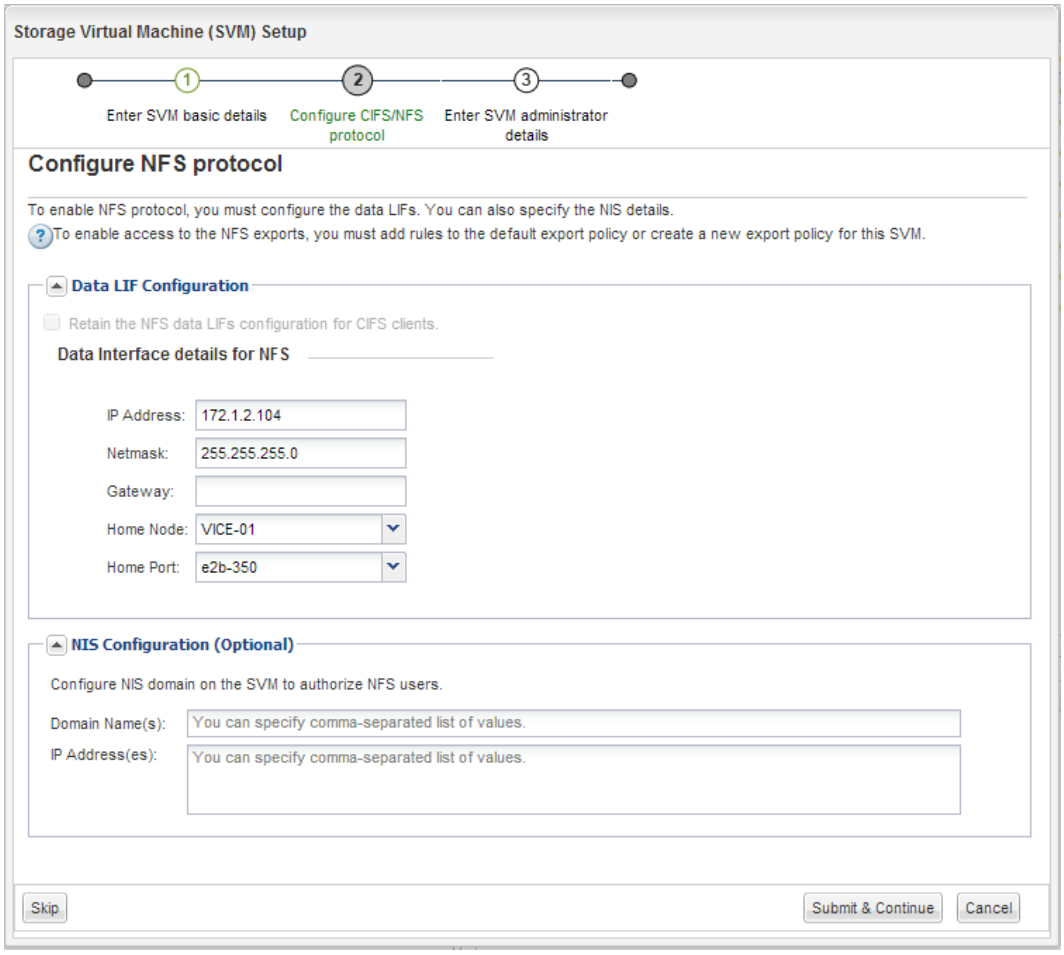
Note: Although LIFs are listed next to a datastore, VSC might not select these LIFs for that specific datastore. Listing the LIFs next to a particular datastore is simply a method for planning the minimum number of LIFs for each node that contains datastores.

2. The Storage Virtual Machine Setup wizard starts automatically. On the first page of the Setup wizard, make sure that NFS is selected and complete the rest of the fields based on the requirements for the environment. Click Submit and Continue.

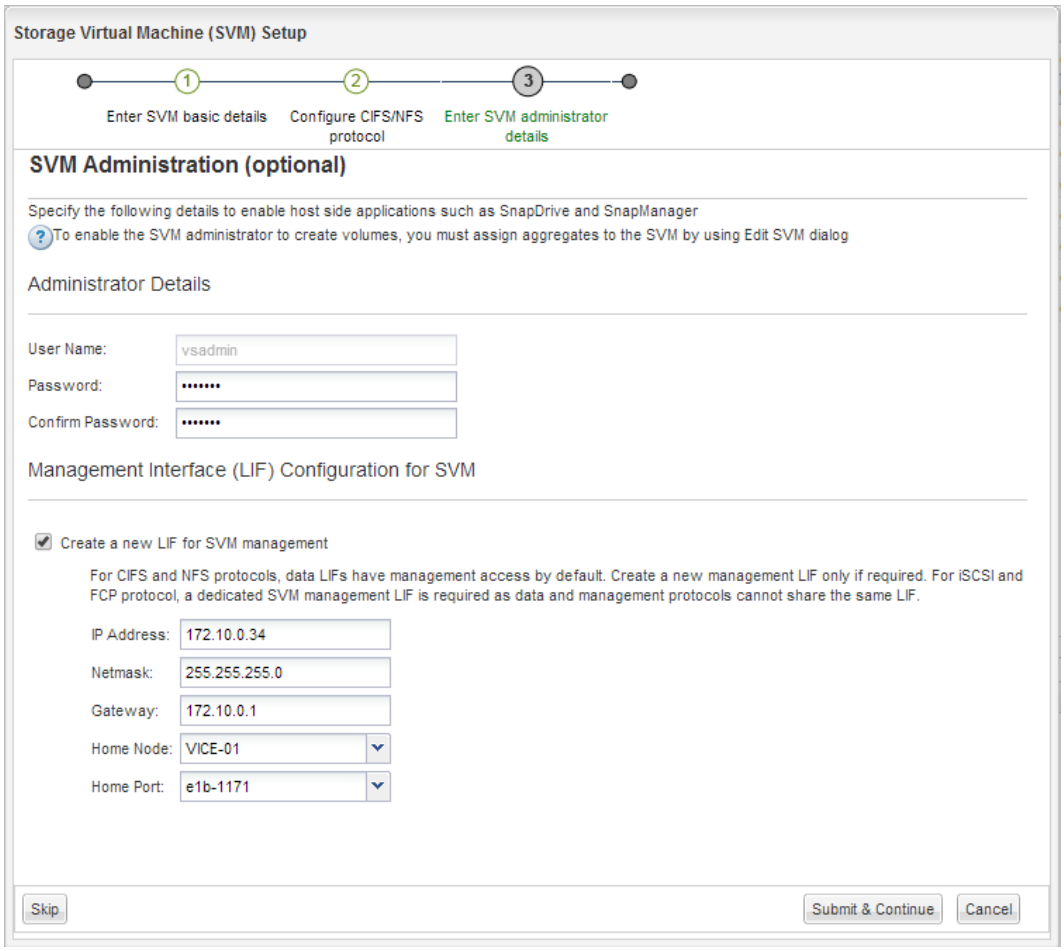


3. Add the network configuration information for the first LIF of the first datastore. Be sure to select the correct home node and port and the interface group (ifgrp) or virtual local area network (VLAN). Click Submit and Continue.

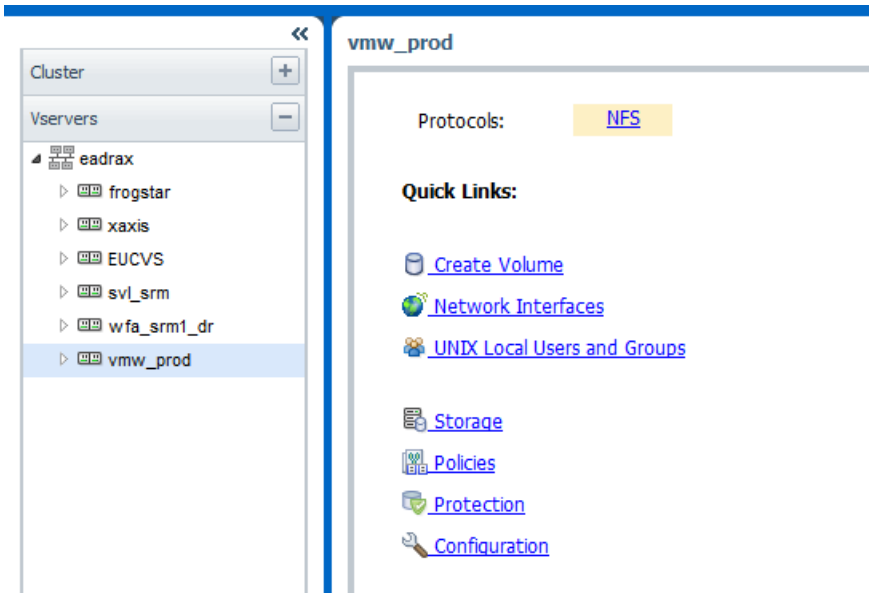
   Note: Do not enable network interface service (NIS) or Lightweight Directory Access Protocol (LDAP) services unless they are required for authentication/authorization and are properly configured. Improperly configured NIS or LDAP services can cause datastore access issues.



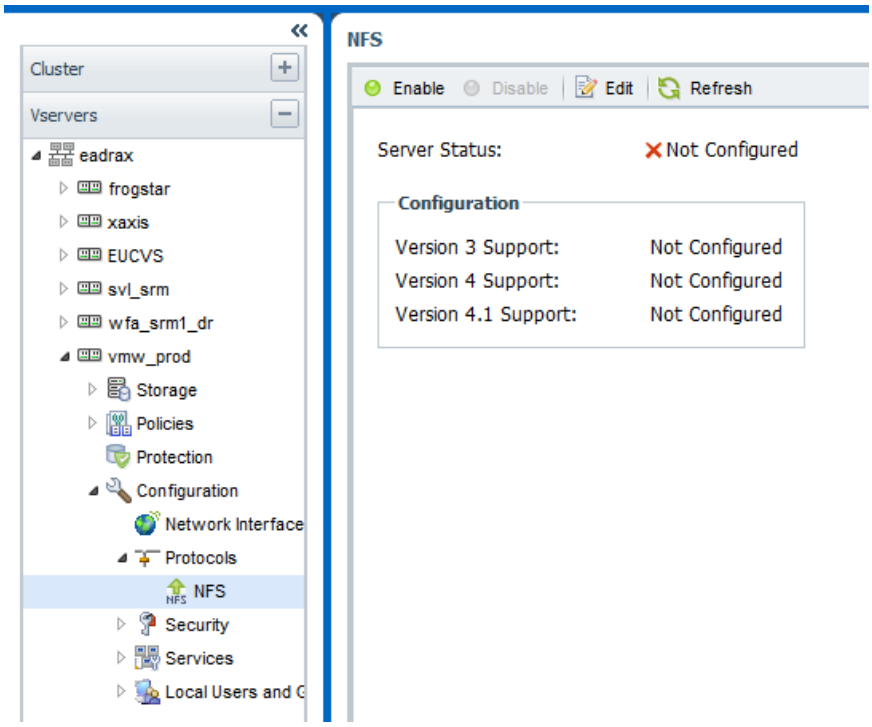
4. If management of the SVM is delegated, set a password for the vsadmin account and define a management interface for the SVM. The home node is not critical for SVM management LIFs, but make sure that the home port is on the correct management network accessible by the delegated administrator. If cluster administrator credentials are used to manage this SVM, click Skip. Otherwise, click Submit and Continue.



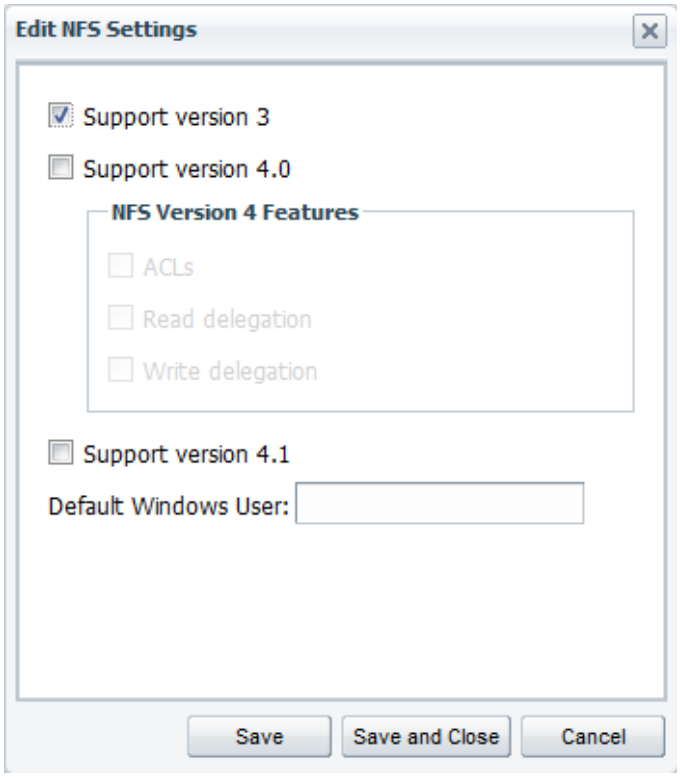
5. Review the summary page and click OK.
6. Click the NFS link.



7. Click Edit.



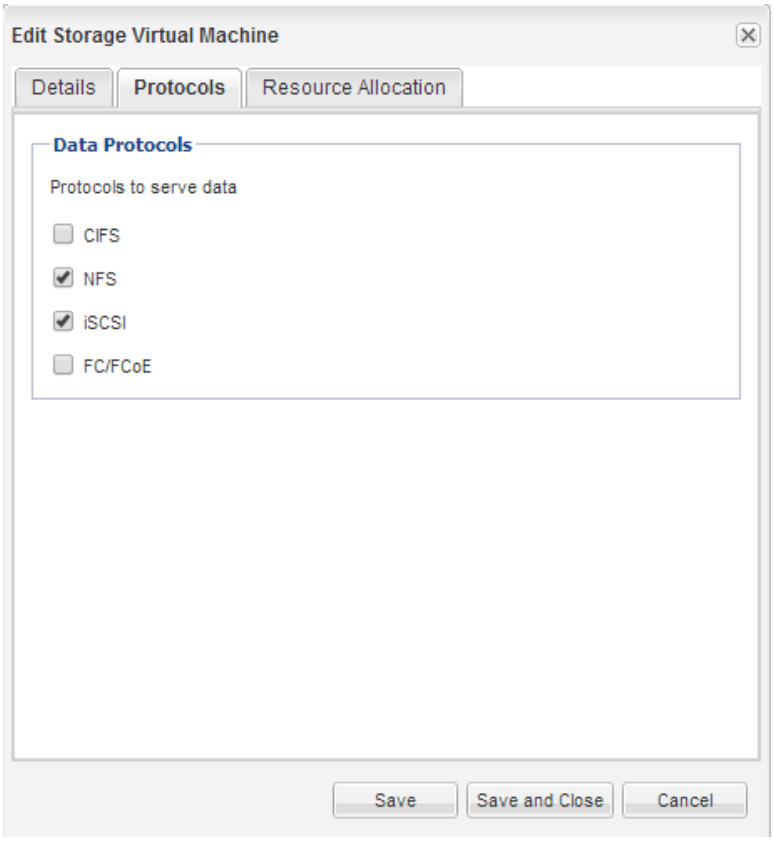
8. Select Support Version 3 and then click Save and Close.



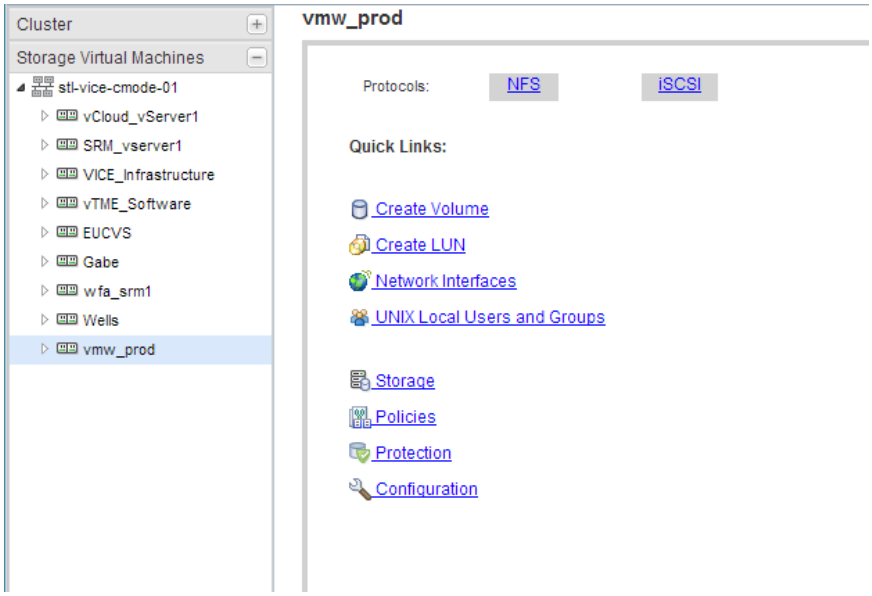
Add NFS Protocol to Existing SVM

To add the NFS protocol to an existing SVM, complete the following steps:

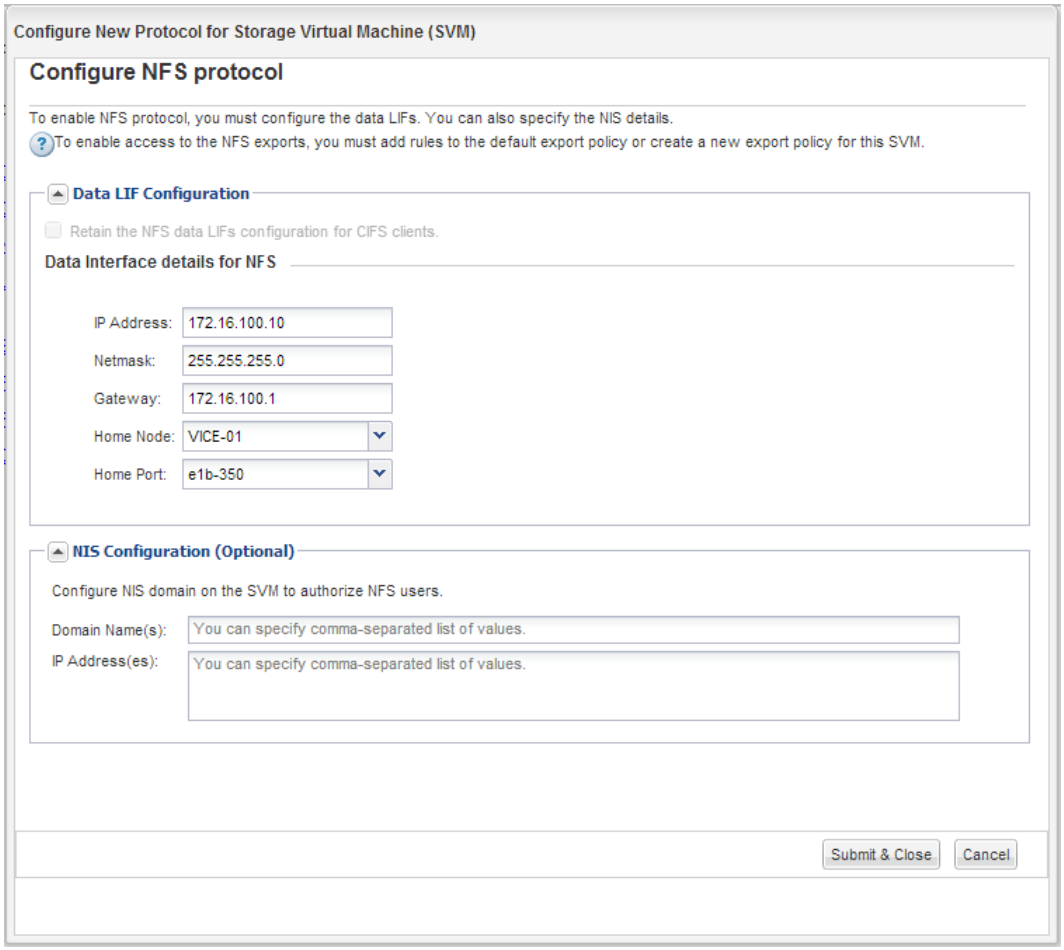
1. In System Manager, from the Storage Virtual Machines view, select the cluster name to display all of the existing SVMs in the right pane.
2. Select an SVM to modify and click Edit.
3. In the Edit Storage Virtual Machine pane, click the Protocols tab and select the NFS protocol.
4. Click Save and Close.



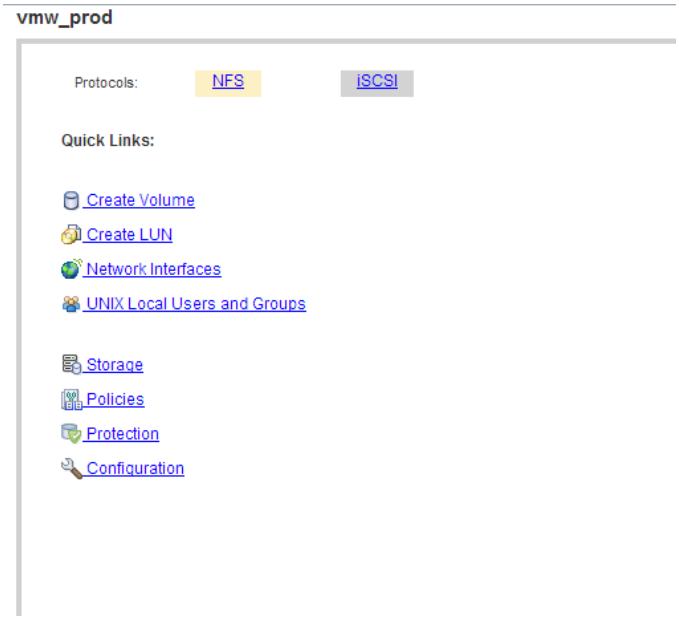
5. A dialog box displays a message that the protocol must be configured. Click OK.
6. In the Storage Virtual Machines navigation pane, select the SVM that was modified in step 1 through step 4.
7. In the right pane, click the gray NFS link to configure the protocol.



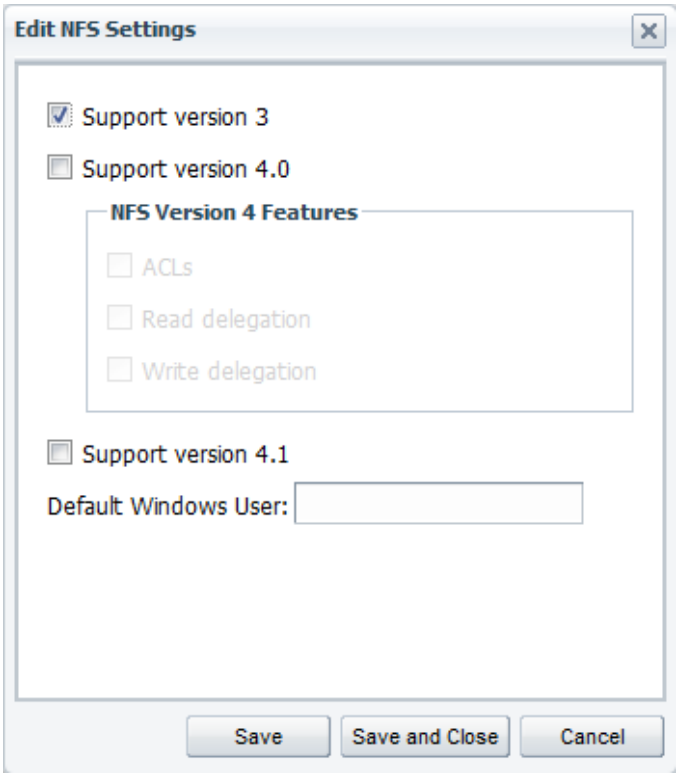
8. Complete the data LIF configuration fields based on the requirements for the environment and then click Submit and Close.



9. The previously gray NFS link is now highlighted in light yellow, which indicates that it is enabled but not yet configured. Click the link again to go to the NFS protocol configuration tree menu option.



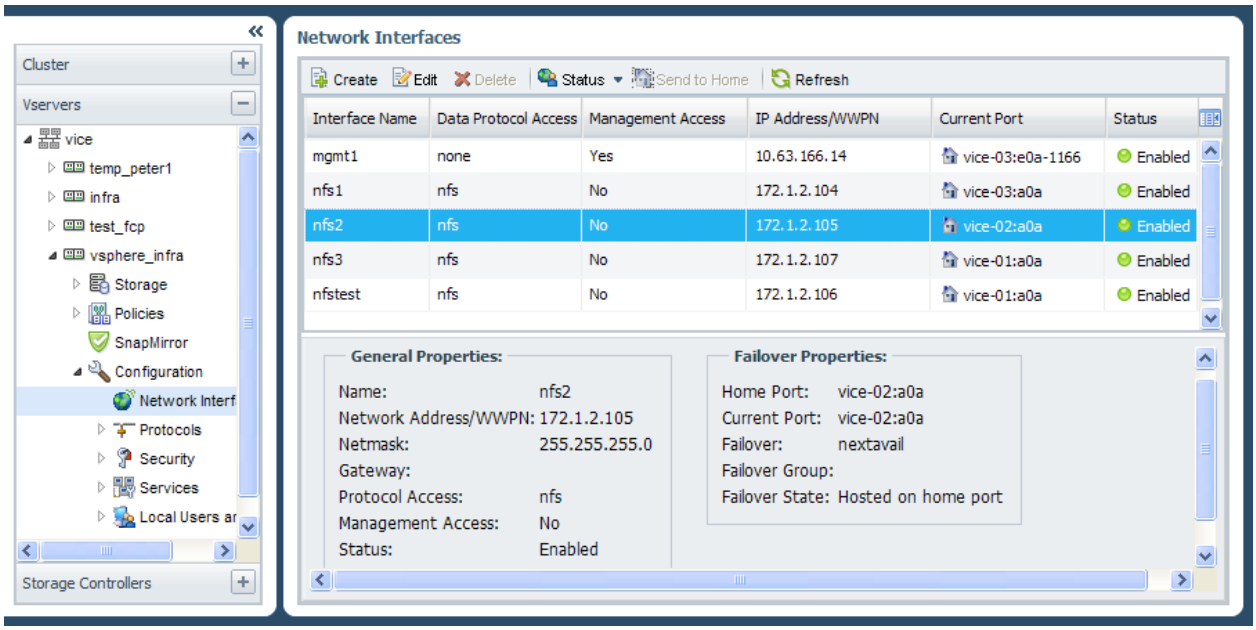
10. Click Edit, select Support Version 3, and click Save and Close.



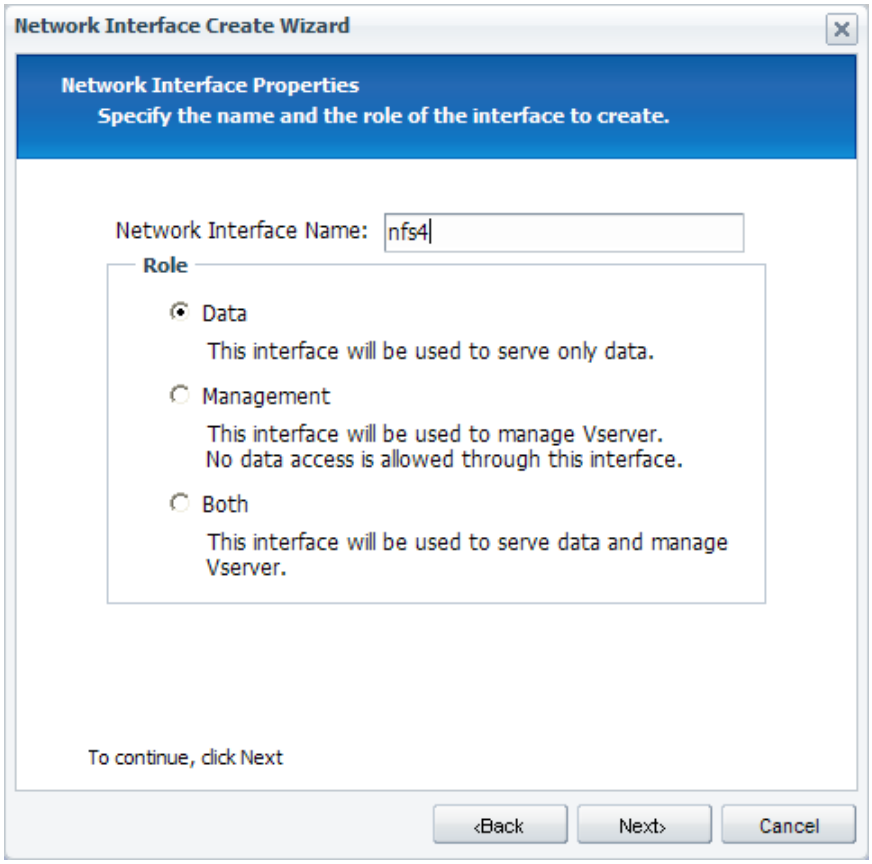
Create Additional LIFs for NFS Datastores

To create additional LIFs for NFS datastores after the SVM is created, complete the following steps:

1. In System Manager, select Storage Virtual Machines and then select the SVM that contains the datastore for which you are adding the LIF.
2. Select Configuration > Network Interfaces, click Create to open the Network Interface Create wizard, and click Next.

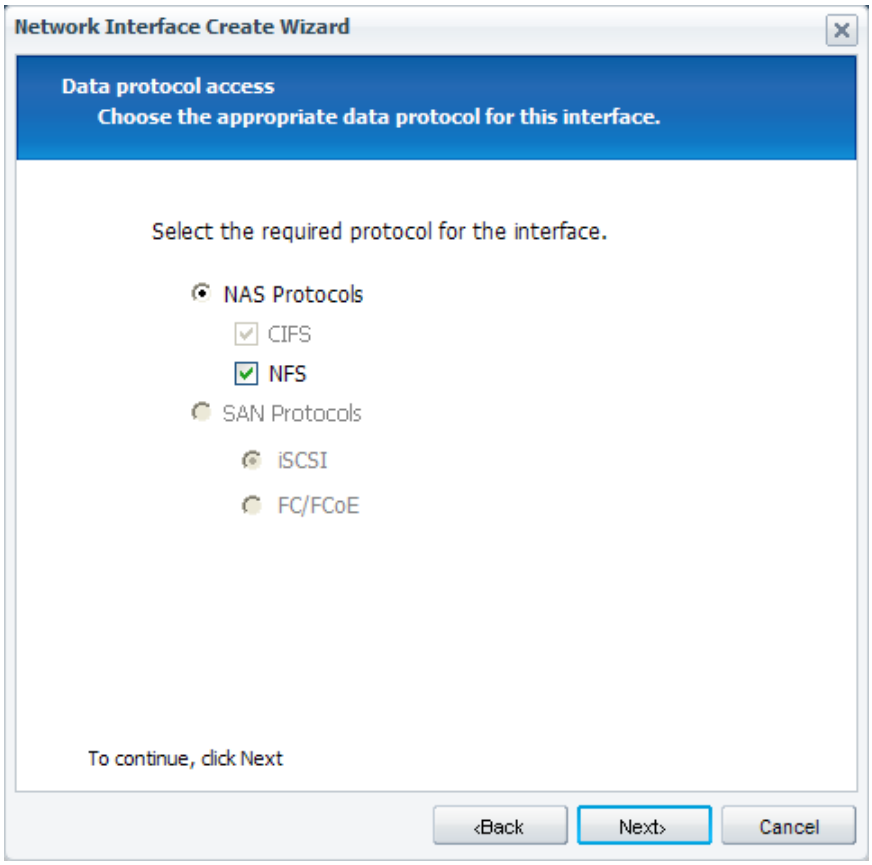


3. Enter a name for the LIF. Under Role, select Data and click Next.

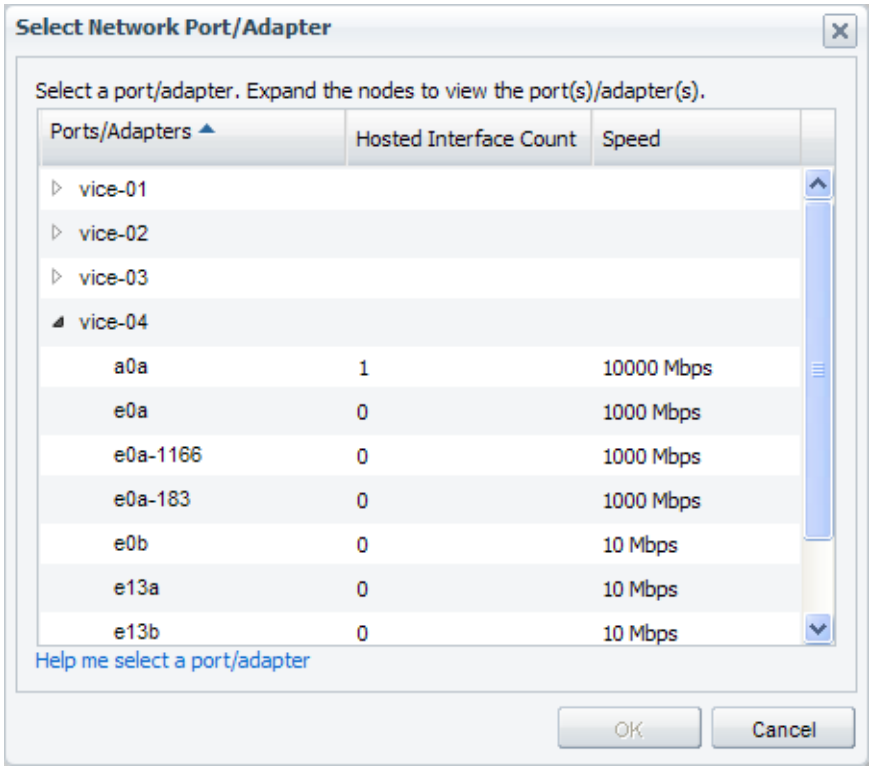


4. On the Data Protocol Access page, select NFS as the protocol for the LIF. Click Next.

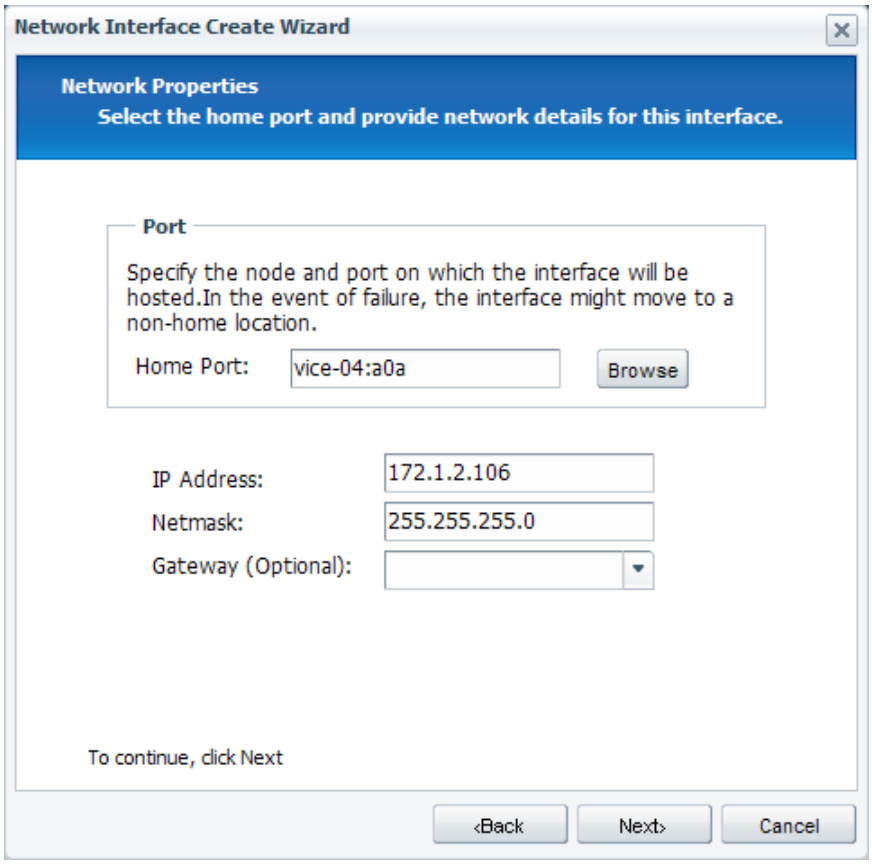
   Note: Depending on licenses present in the cluster and protocols configured for the SVM, not all protocols are listed on this page.



5. On the Network Properties page, click Browse and select a home node and port, ifgrp, or VLAN. Click OK.



6. On the Network Properties page, enter the IP address and the netmask, but leave the Gateway field blank. Click Next and click Finish.



Create LIFs for NFS Datastores Using CLI

In some cases, such as configuring many LIFs at the same time, the command line is faster for experienced users.

1. To create LIFs for NFS datastores from the command line, run the following command:

```
vice::> net int create -vserver vsphere_infra -lif nfs4 -role data -data-protocol nfs -home-node
vice-03 -home-port a0a -firewall-policy data -address 172.1.2.113 -netmask 255.255.255.0 -
failover-group private
```

Assign LIFs to Failover Group

Failover groups must be created, and all appropriate ports, ifgrps, and VLANs must be added to them before LIFs can be assigned to the failover groups. Failover groups do not apply to LIFs that are used for iSCSI, FC, or FCoE.

To assign LIFs to failover groups, complete the following steps:

1. Log in to the cluster CLI as `admin`.
2. If an SVM management LIF was created, run the following command to set its failover group and policy.

```
eadrax::> network interface modify -vserver vmw_prod -lif vmw_prod_admin_lif1 -failover-group
mgmtnet -failover-policy nextavail
```

3. Set the failover group and policy for each NFS datastore LIF.

   Note: LIFs on different physical networks or VLANs should be members of failover groups specific to their network or VLAN.

```
eadrax::> network interface modify -vserver vmw_prod -lif vmw_prod_nfs_lif1 -failover-group
stgnet42 -failover-policy nextavail
```

4. Verify that all LIFs on the SVM are members of the correct failover group.

```
eadrax::> network interface show -vserver vmw_prod -fields address,failover-group,failover-policy
vserver  lif                  address        failover-policy failover-group
-------- -------------------- -------------- --------------- --------------
vmw_prod nfs4                 192.168.42.215 nextavail       stgnet42
vmw_prod vmw_prod_admin_lif1 172.16.24.98   nextavail       pubnet172
vmw_prod vmw_prod_nfs_lif1   192.168.42.214 nextavail       stgnet42
3 entries were displayed.
```

Provisioning NFS Datastores

NFS datastores can be provisioned using CLI or GUI to create, mount, and export the volume NetApp storage and to mount the volume as a datastore in ESXi. You can also automate these tasks by using the NetApp PowerShell toolkit and VMware PowerCLI. However, in most cases, the simplest and least error-prone method is to use the VSC as described in section 9.7.

7.7 VMware vSphere 6 Storage Design Using LUNs on Clustered Data ONTAP

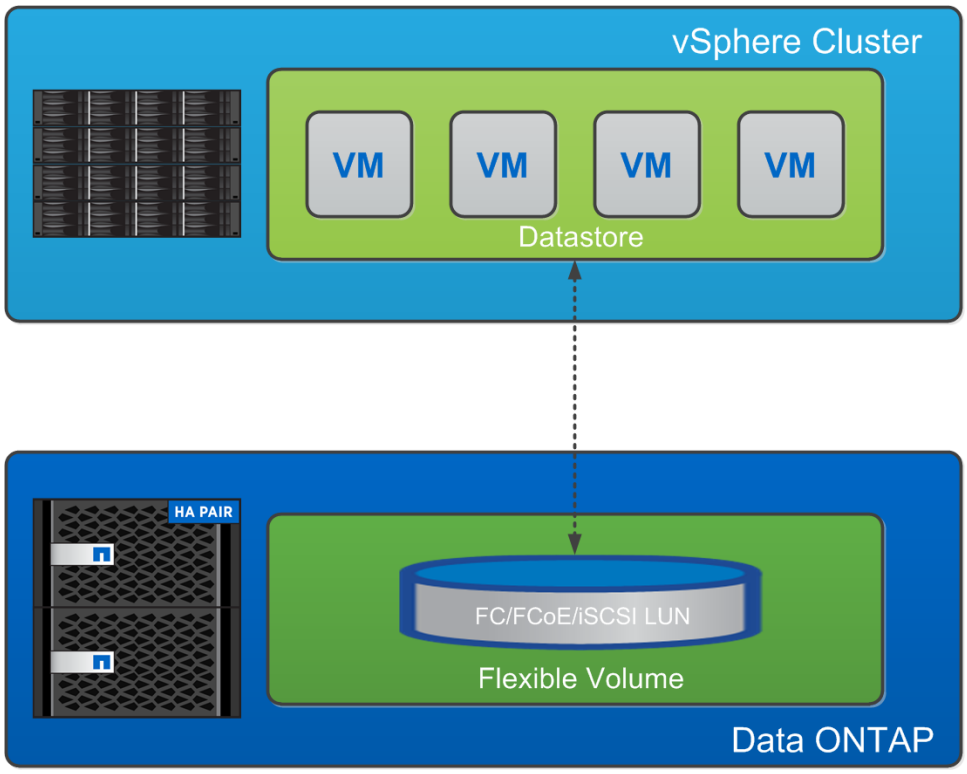
A LUN is a storage object presented to a server that, at its simplest, looks like a disk device. The server does not know what the true storage architecture is under the LUN. The server either partitions and formats the LUN with its file system or presents it to another application to manage and consume. In vSphere, there are three ways to use LUNs:

- With VMFS
- With RDM
- As a LUN accessed and controlled by a software initiator in a VM guest operation system

Virtual Machine File System

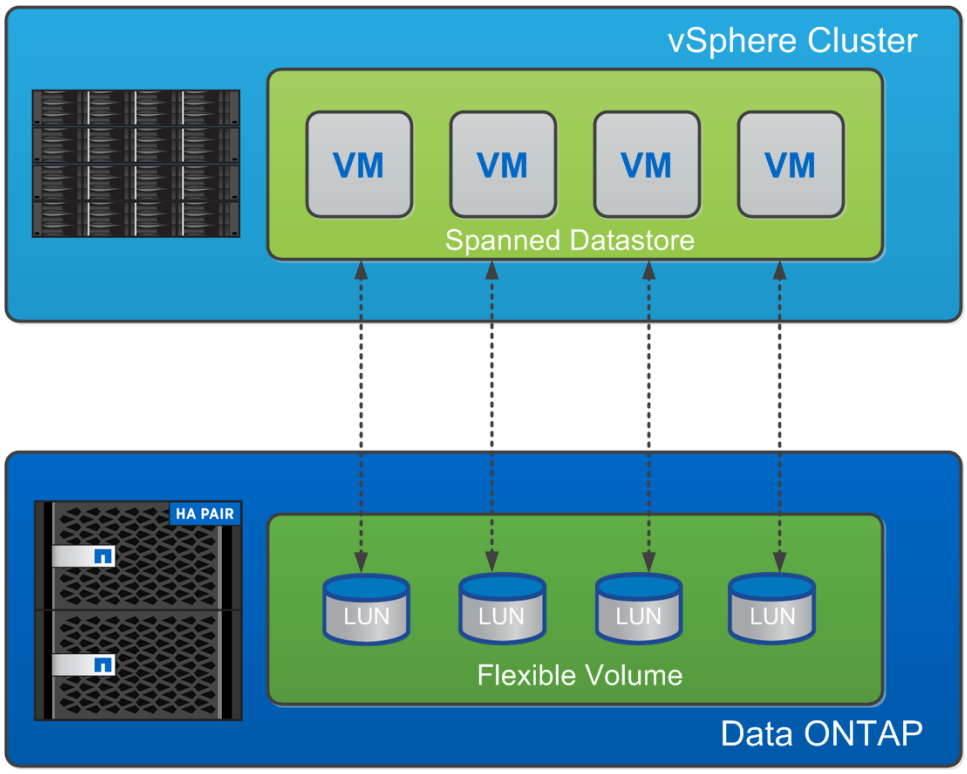
VMFS is a high-performance clustered file system that provides datastores that are shared storage pools. VMFS datastores can be configured with LUNs that are accessed by FC, iSCSI, or FCoE. VMFS allows traditional LUNs to be accessed simultaneously by every ESX server in a cluster. VMFS is VMware's proprietary file system that is used primarily for storing and executing VMs and also for storing software and CD ISO images for use with VMs. VMFS datastores can be up to 64TB in size and consist of up to 32, 2TB LUNs (VMFS 3) or a single 64TB LUN (VMFS 5). Figure 15 illustrates a vSphere cluster connected to a VMFS datastore through FC, FCoE, or iSCSI LUNs.

Figure 15) vSphere cluster connected to a VMFS datastore through FC, FCoE, or iSCSI LUNs.



The NetApp maximum LUN size is 16TB; therefore, a maximum size VMFS 5 datastore is created by using four 16TB LUNs. Figure 16 shows a vSphere cluster connected to a spanned VMFS datastore.

Figure 16) vSphere cluster connected to a spanned VMFS datastore.



VMFS provides the VMware administrator with a fair degree of independence from the storage administrator. By deploying shared datastores, the VMware administrator can provision storage to VMs as needed. In this design, most data management operations are performed exclusively through VMware vCenter Server.

When deploying third-party applications, carefully consider the storage design to make sure that it can be virtualized and served by VMFS. Consult the storage sizing specifications in the best practices documentation of each application that is being deployed. If no specific recommendations are available for the application, consult the application vendor to verify that you are using their recommended guidelines for storage performance and support compliance.

The VMFS storage design can be challenging for performance monitoring and scaling. Because shared datastores serve the aggregated I/O demands of multiple VMs, this architecture does not natively allow a storage array to identify the I/O load generated by an individual VM.

VMFS Datastores on NetApp LUNs

NetApp enhances the use of VMFS datastores through many technologies, including the following:

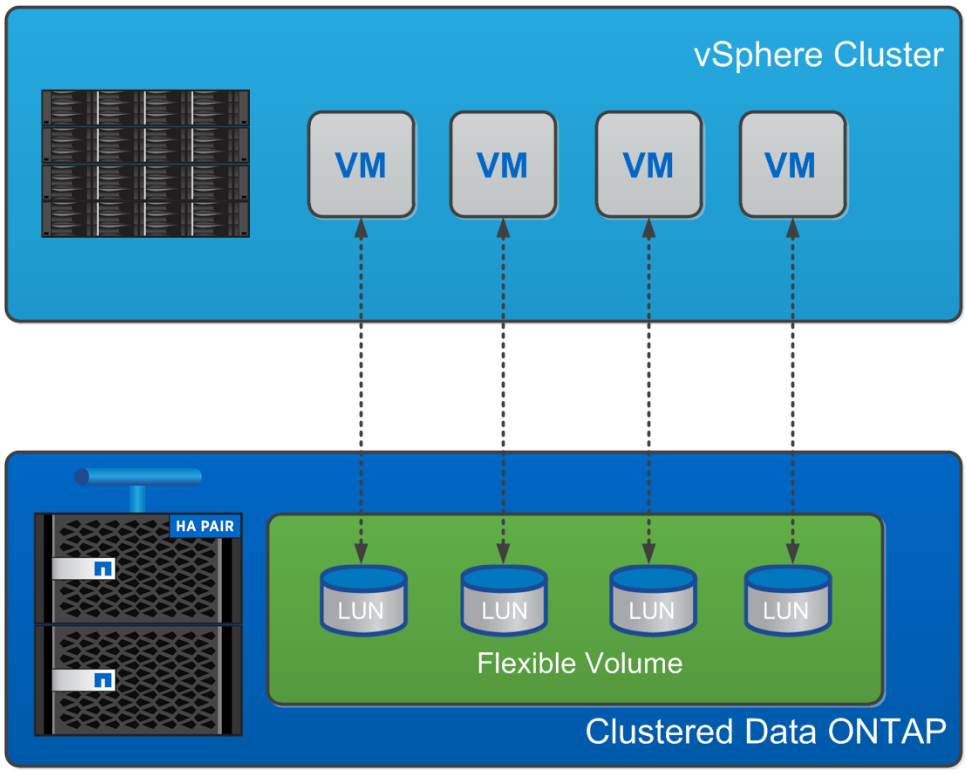
- Array-based thin provisioning
- Deduplication of production data
- Immediate, zero-cost datastore clones
- Integrated tools such as SRM, VSC, OnCommand System Manager, and OnCommand Insight Balance

Because the NetApp LUN architecture does not have small individual queue depths, VMFS datastores can scale to a greater degree with it than with traditional array architectures in a relatively simple manner.

Raw Device Mapping

RDM is a method of passing a LUN through ESXi to a VM to use the full LUN as a virtual disk. The RDM term comes from a mapping file created in a VMFS datastore, usually in the same directory as the VM that uses the RDM. The mapping file points to the actual LUN and provides an object that can be referenced by the VM configuration files. Figure 17 illustrates a vSphere cluster with VMs connected to RDM LUNs through FC or iSCSI.

Figure 17) vSphere cluster with VMs connected to RDM LUNs through FC or iSCSI.



There are two modes for RDMs: physical and virtual. With physical RDM, I/O is passed directly through from the guest to the LUN and back. ESXi does not intercept or process the I/O in any way. With virtual mode RDM, ESXi can intercept I/O to support features such as VMware snapshots and redirect writes to a delta file in a VMFS when a snapshot exists. Because the maximum size of a file in VMFS3 is 2TB, the maximum size of a virtual mode RDM is also 2TB when the mapping file is on VMFS3 or using ESXi 5.0 or 5.1. ESXi 5.5 increases the limit on virtual mode RDMs to 62TB when the mapping file is on VMFS5. Because physical RDM I/O is not intercepted by ESXi, VMware snapshots are not supported, but the RDM LUN can be the maximum size that ESXi can address, which is 64TB.

Note: The guest OS must also be able to address the full size of the LUN.

RDM LUNs on NetApp

NetApp enhances the use of RDMs by providing the following features:

- Array-based thin provisioning at the LUN level
- Deduplication of production data
- Advanced integration components (for example, SnapDrive)
- Application-specific Snapshot copy backups created through the SnapManager suite
- FlexClone zero-cost clones of RDM-based datasets

Guest Software Initiators

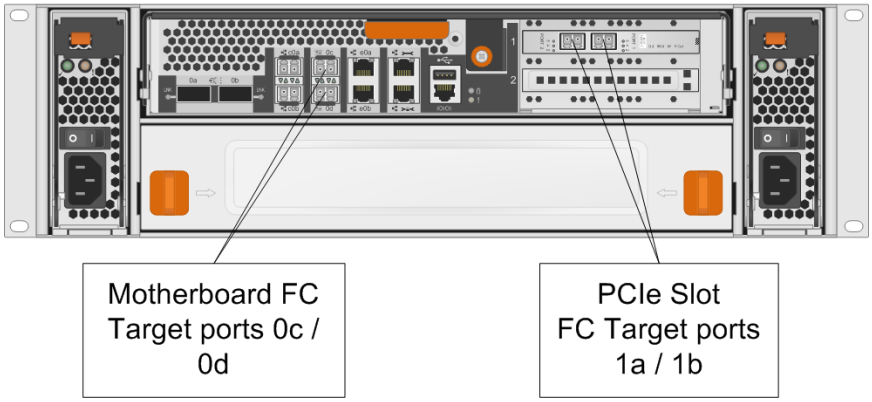
With guest initiators, ESXi is not actually aware that a LUN exists and is being used by the guest. The storage traffic is seen as regular network traffic from the VM. LUN size, sharing, and other limits are governed entirely by the capabilities of the guest.

Block Protocols

VMware vSphere 6 and clustered Data ONTAP both support the most common industry-standard LUN connectivity protocols: FC, iSCSI, and FCoE. The NetApp clustered Data ONTAP unified SCSI target makes it possible to use more than one protocol at a time to the same LUNs. This allows servers using different protocols and HBAs or software initiators to use the same shared storage.

NetApp storage systems implement FC and FCoE through physical target HBA ports, either on the controller motherboard (ports numbered `0a`, `0b`, and so forth) or with target HBA cards (`1a`, `1b`, `2a`, `2b`, and so forth). iSCSI is available, using the software target with standard Ethernet ports for connectivity. Figure 18 shows an example of FAS3200 FC target port options.

Figure 18) Example FAS3200 FC target port options.



Best Practice

When nodes have similar hardware configurations, use the same numbered ports on each node for the same purpose (initiator or target), including using the same port on each node connected to the same fabric or switch.

FC and FCoE ports have globally unique identifiers, referred to as worldwide names (WWNs). There are two types of WWNs:

- **Worldwide node name (WWNN).** Each of the following has a WWNN:
  - Nodes, such as HBA cards and ports
  - VMs that use N_Port ID Virtualization (NPIV)
  - Storage devices
- **Worldwide port name (WWPN).** Each node has one or more ports, each of which has a WWPN.

Each initiator HBA port typically has its own WWNN in addition to a WWPN.

Clustered Data ONTAP makes extensive use of NPIV because all FC and FCoE target interfaces on SVMs are virtual ports. Referred to as LIFs, these virtual ports are logically connected to the SAN fabric through a physical target port or adapter. In other words, SVMs have LIFs, and nodes have ports or adapters. The FC target LIFs of many SVMs can and usually do share a common set of physical target adapters. An SVM has a single WWNN, and each target LIF has its own FC WWPN.

LUNs are created in volumes on an SVM. An SVM makes LUNs visible to servers by:

- Adding the initiator WWPNs or iSCSI qualified names (IQNs) to an initiator group (igroup)
- Mapping each LUN to the appropriate igroups
- Providing a LUN ID

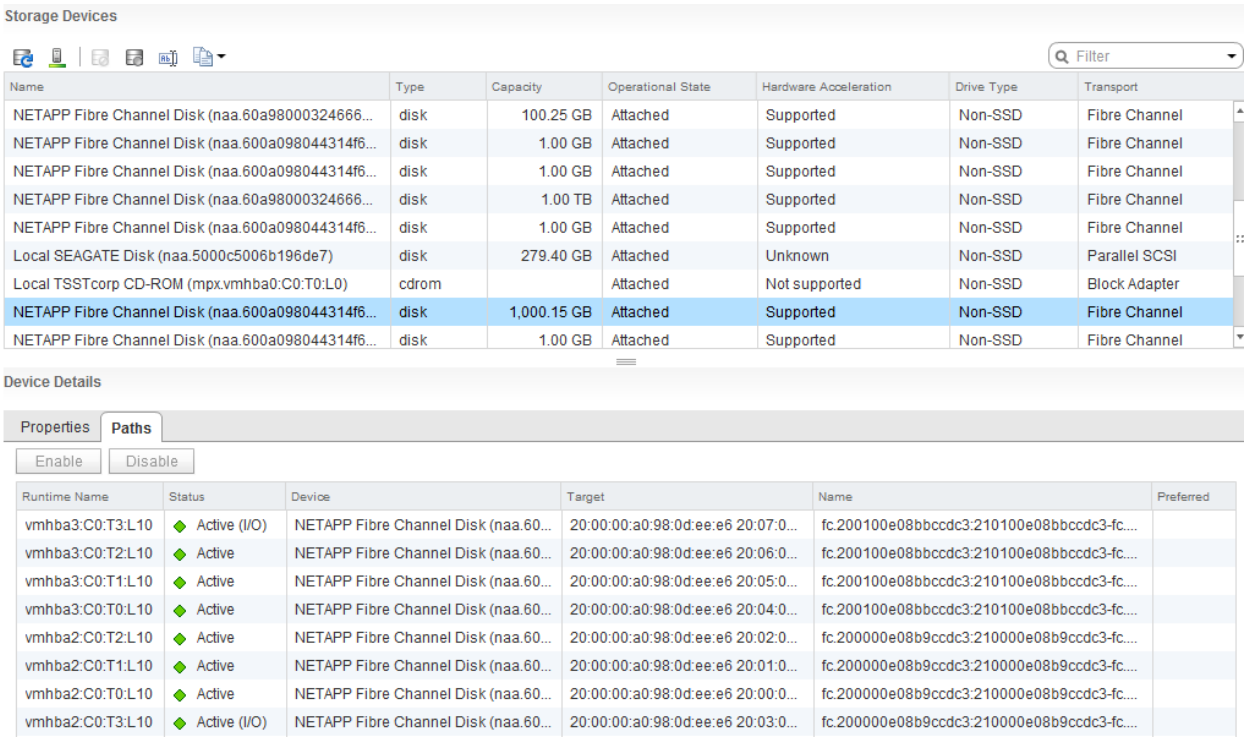
In addition, each igroup is granted access through a target port set, which might include all or a subset of the available target ports in an SVM.

Multipathing

vSphere includes built-in support for multiple paths to storage devices, referred to as native multipathing (NMP). NMP includes the ability to detect the type of storage for supported storage systems and automatically configure the NMP stack to support the capabilities of the storage system in use. Both NMP and NetApp clustered Data ONTAP 8.1 and later support the Asymmetric Logical Unit Access (ALUA) protocol to negotiate optimized and nonoptimized paths. With clustered Data ONTAP, an ALUA-optimized path follows a direct data path, using a target port on the node that hosts the LUN being accessed.

By default, ALUA is turned on in both vSphere and clustered Data ONTAP. The NMP recognizes the NetApp cluster as ALUA, and it uses the ALUA storage array type plug-in (`VMW_SATP_ALUA`) and selects the round robin path selection plug-in (`VMW_PSP_RR`). This means that I/O to the LUN alternates between the ALUA-optimized data paths and direct data paths to the LUN. Figure 19 shows a LUN on a four-node NetApp cluster, automatically configured for ALUA and round robin, with I/O on the two optimized paths out of the eight total paths.

Figure 19) A LUN automatically configured for ALUA and round robin.



Note: In versions earlier than vSphere 5.5, LUNs that are used for shared or quorum disks for VMs running Microsoft Cluster Service (MSCS) must use the fixed-path selection policy and must have ALUA disabled. vSphere 5.5 supports round robin and ALUA for RDM LUNs for MSCS, as documented by VMware and the NetApp Interoperability Matrix Tool.

Switch Zoning

FC and FCoE switches use the concept of zoning to isolate communication between initiators and targets. Zones provide security and robustness to the fabric by:

- Preventing initiators from seeing targets to which they are not authorized to communicate
- Isolating excessive, spurious, or malicious traffic from a misbehaving initiator

There are two kinds of zoning:

- Hard zoning uses the physical switch port IDs.
- Soft zoning uses the WWPNs of the initiators and targets.

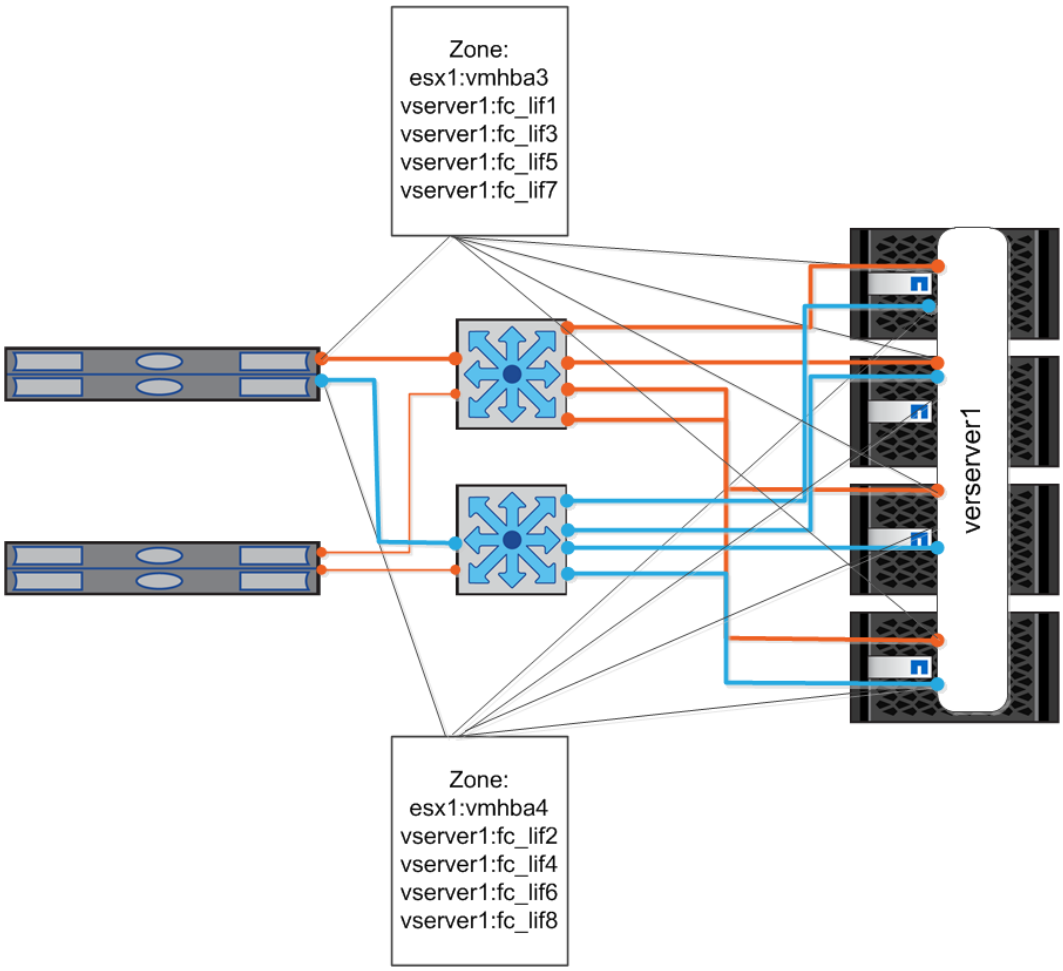
Because WWPNs on the SVM target use NPIV and are virtual ports on a physical target HBA port, soft zoning should be configured on the switches. If hard or port-based zoning is configured so that the physical switch port is in the zone, all initiators zoned to the port can see the LIFs of all SVMs connected through that port.

Zoning policy is usually divided into three types of zones:

- Multiinitiator zones
- Single-initiator/multitarget zones
- Single-initiator/single-target zones

The consensus of NetApp and the industry is to avoid multiinitiator zones because of reliability and security risks. Single-initiator, single-target zones are the most robust, but they are also the most complicated. For example, an ESXi cluster with two HBAs per server connected through a fabric to a four-node NetApp cluster with an SVM, with two LIFs on each node, would require eight zone definitions per server. Given this complexity, NetApp recommends using single-initiator/multitarget zones, as shown in Figure 20.

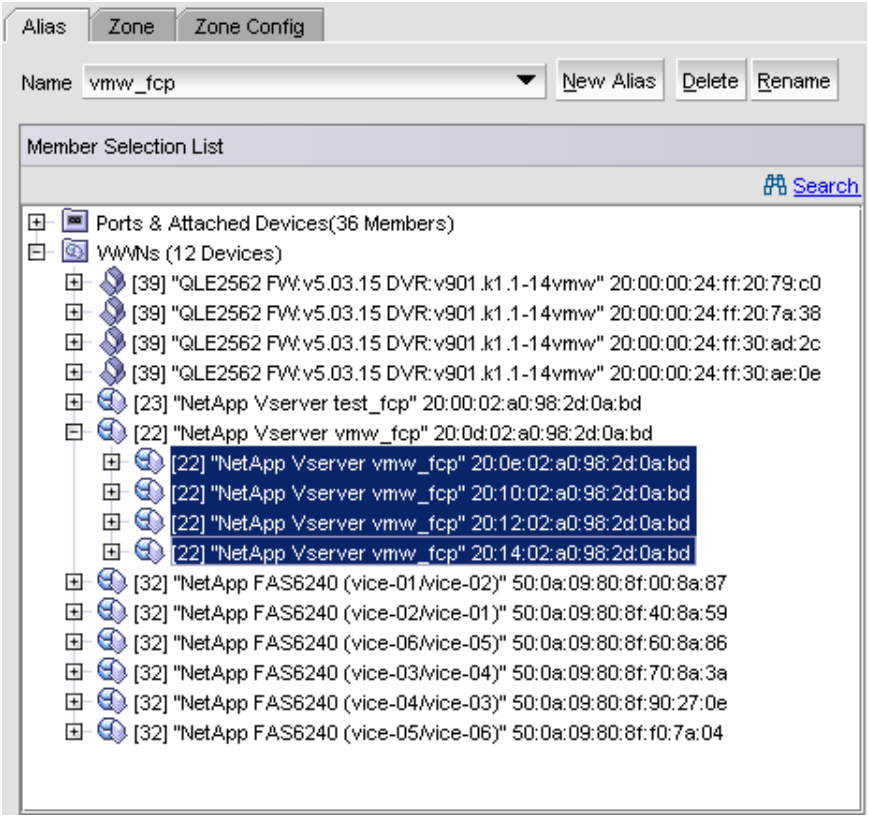
Figure 20) Single-initiator/multitarget zone.



To make it clear which target LIFs on the fabric belong to which SVM, clustered Data ONTAP includes `NetApp Vserver <name>` in the OEM string, as shown in Figure 21. Because an SVM has multiple FC LIFs on each SAN fabric, a simpler way to manage aliases is to create a single alias on each fabric that includes all visible FC LIFs for that SVM on that fabric. Using the WWPN (the HBA port name, not the node name), create an alias for each `vmhba` of each ESXi server with half of these HBAs and aliases on each switch or fabric.

Note: SVMs are referred to as Vservers in the CLI and in some GUIs.

Figure 21) Brocade zone admin view showing SVM (Vserver) WWPN with SVM name.

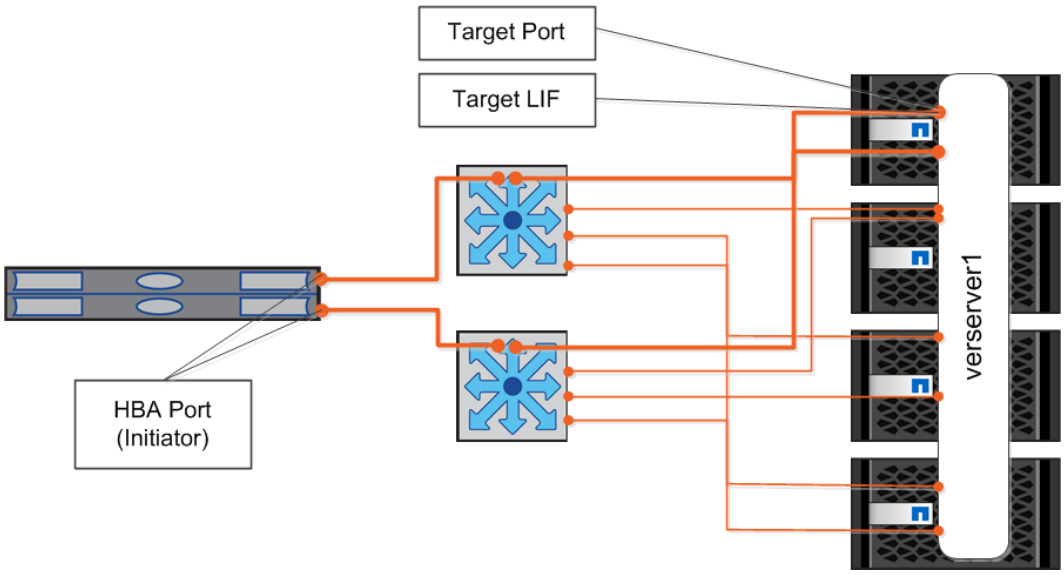


Best Practices
• Implement soft zoning, using WWPN, not WWNN. • Zone to SVM WWPN, not to node WWPN. • Use single-initiator/multitarget or single-initiator/single-target zones.

ESXi Path and Target Limits

ESXi 6 supports up to 256 LUNs and up to 1,024 total paths to LUNs. Any LUNs or paths beyond these limits are not seen by ESXi. Assuming the maximum number of LUNs, the path limit allows four paths per LUN. Figure 22 shows that a four-node NetApp cluster with two target ports per node (one on each fabric) provides eight targets. An ESXi server with two HBAs sees half of these targets on each fabric, resulting in eight paths to each LUN, which means the environment consumes the maximum number of paths to LUNs before consuming the maximum number of LUNs. The only way to get more LUNs is to use fewer paths, which in turn means losing some redundancy and either limiting the aggregates and nodes on which LUNs are placed or using indirect paths to LUNs.

Figure 22) Eight paths (two highlighted) across a dual fabric to a four-node NetApp cluster.



One use case in which there are more than two HBA ports per ESXi server is MSCS in VMs in vSphere versions earlier than 5.5. Prior to vSphere 5.5, MSCS could not use ALUA, which is preferred for non-MSCS LUNs. The two ways to implement MSCS are to use the fixed path selection plug-ins (PSP) for all LUNs or to dedicate additional HBAs for MSCS. In the latter scenario, all four HBAs in each ESXi server are zoned to all NetApp target ports visible on the same fabric. However, the ESXi HBAs used for MSCS are in an igroup with ALUA off, and the HBAs used for VMFS and non-MSCS RDM LUNs are in a second igroup with ALUA on. LUNs are mapped to one igroup or the other, so each LUN has four paths. Table 23 lists the LUN uses and configuration details.

Table 23) LUN uses and configuration details.

LUN Use	LUN Type	Igroup Type	SATP/PSP
VMFS	VMware	VMware	ALUA/round robin
RDM	Guest dependent	VMware	ALUA/round robin
Guest initiator	Guest dependent	Guest dependent	Not applicable **Note:** Storage array type plug-ins (SATP) and PSP are part of ESXi NMP. Guests use their own native or add-on multipathing stack.
RDM for MSCS	Windows*	VMware	• Versions earlier than vSphere 5.5: active-active/fixed • vSphere 5.5 versions and later: ALUA/round robin

* There are several Windows LUN types. Use the one that best matches the version of Windows and the partitioning system used (master boot record [MBR] or GUID partition table [GPT]).

ESXi limits the maximum number of targets per HBA port to 256. This limit can be exceeded before the maximum LUNs limit if many SVMs are created, all with FC target LIFs (for example, if an SVM is created for each LUN). For this reason and for simplicity, avoid creating a separate SVM for each LUN.

Best Practice
Do not create a separate SVM for each LUN. The maximum number of SVMs used for LUNs for any ESXi host should not exceed 256 divided by the number of target ports seen by each server HBA.

Selective LUN Map

New to clustered DATA ONTAP 8.3 is selective LUN map (SLM). Because of the above path exhaustion, there was a need to reduce the paths available to hosts. With SLM, visible paths are constrained to the HA pair that owns the LUN. As can be seen above, an effectively configured vSphere environment includes multiple paths to targets. As the cluster adds additional clustered Data ONTAP HA node pairs, the paths increase, creating an environment that constrains LUN growth. SLM helps alleviate this growth concern by masking out paths that are not on the owning HA pair. The previous method of portsets can also be used to further reduce the available paths for a LUN.

Best Practices
• SLM is enabled by default. Unless using `portsets`, no additional configuration is required. • For LUNs created prior to clustered Data ONTAP 8.3, manually apply SLM by running the `lun mapping remove-reporting-nodes` command to remove the LUN reporting nodes and restrict LUN access to the LUN-owning node and its HA partner.

7.8 Deploying LUNs for VMware vSphere 6 on Clustered Data ONTAP

Table 24 describes prerequisites for VMware vSphere storage design using LUNs on clustered Data ONTAP.

Table 24) VMware vSphere 6 storage design using LUNs on clustered Data ONTAP prerequisites.

Description
• Clustered Data ONTAP 8.1 or later with FCP license and FC target ports • FC SAN architecture (although most of this information applies directly to iSCSI and FCoE as well) • VMware vSphere 6.x

For the easiest, most complete, and most consistent management of storage and SAN infrastructure, NetApp recommends using the tools listed in Table 25, unless otherwise specified.

Table 25) Management tools.

Management Task	Recommended Management Tool
Managing SVMs	NetApp OnCommand System Manager
Managing switches and zoning	A variety of tools, depending on the switch vendor and model, which could be GUI or CLI
Provisioning and managing volumes and LUNs for VMFS	VSC

Overview of Tasks to Provision and Configure an FC SVM for vSphere

To provision and configure an FC SVM for vSphere, complete the following tasks:

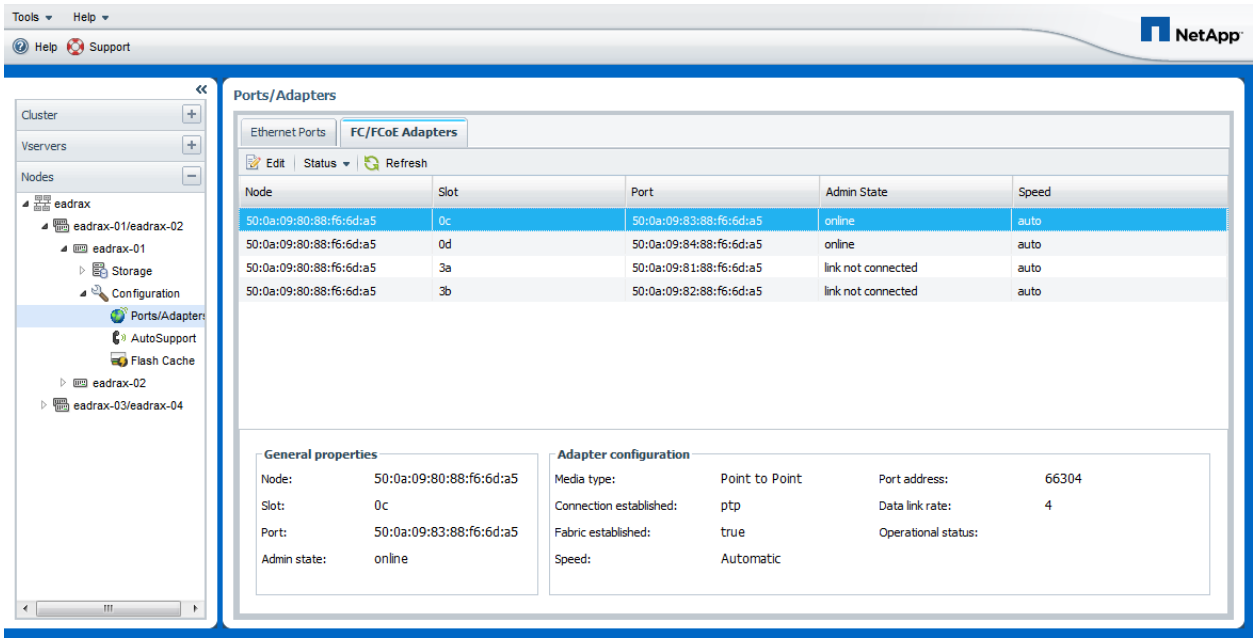
1. Install physical switches, HBAs, and cables.
2. Configure or install target FC ports on NetApp cluster nodes.

3. Create and configure an FC SVM.
4. Configure zoning on FC switches.
5. Create initiator groups (igroups).
6. Create and map LUNs.

Verify FC Ports on NetApp Cluster Nodes

To verify that each node has one or more FC target ports on each fabric, complete the following steps:

1. Start OnCommand System Manager and log in to the cluster.
2. In the navigation pane under Nodes, select a node and click Node > Configuration > Ports/Adapters.
3. Click the FC/FCoE Adapters tab.

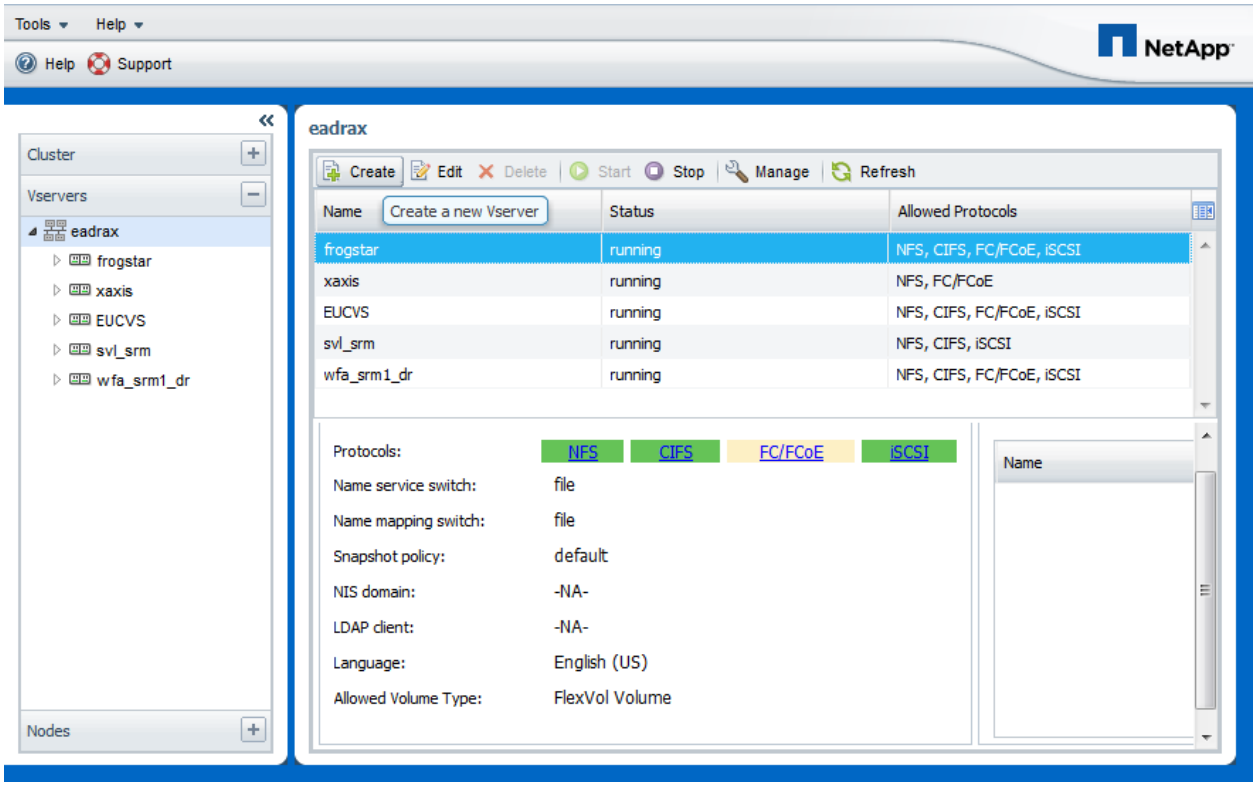


4. Click each port and verify that the connection is point to point and established.
5. Repeat step 1 through step 4 to verify that each node has a configured and working FC target port on each fabric.
6. If each node does not have two FC target ports, you must either obtain and install the necessary target cards or configure existing ports as targets. Refer to the Clustered Data ONTAP 8.2.1 SAN Administration Guide.
7. Make a note of the nodes and target ports.

Create an FC SVM

To create an FC SVM, complete the following steps:

1. Start OnCommand System Manager.
2. Log in to the NetApp cluster as the cluster administrator (usually `admin`).
3. Open the Vservers section and click the cluster (top level).



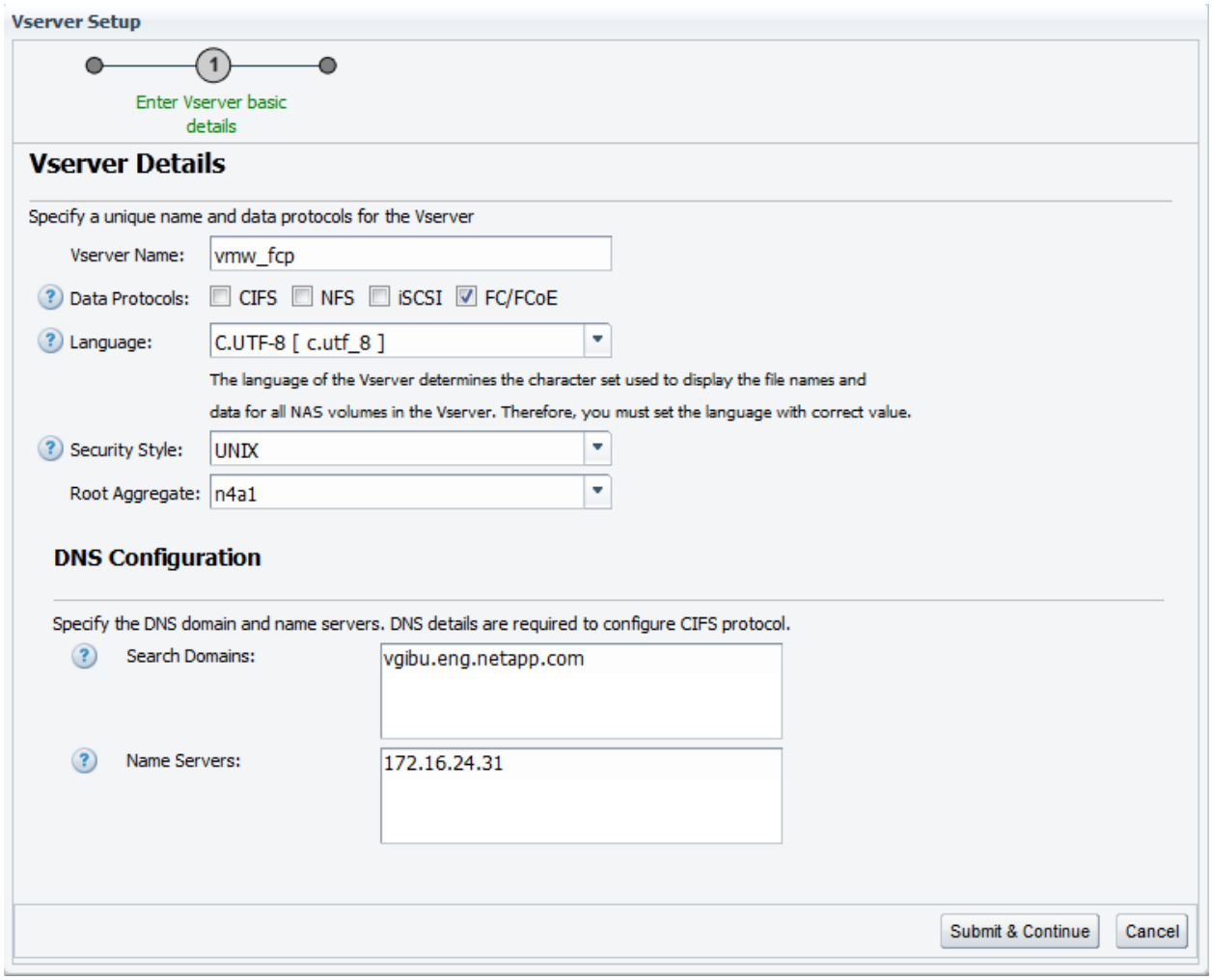
4. Click Create.
5. Complete the Vserver Details screen as follows:

   Note: SVM is referred to as Vserver in the GUI.

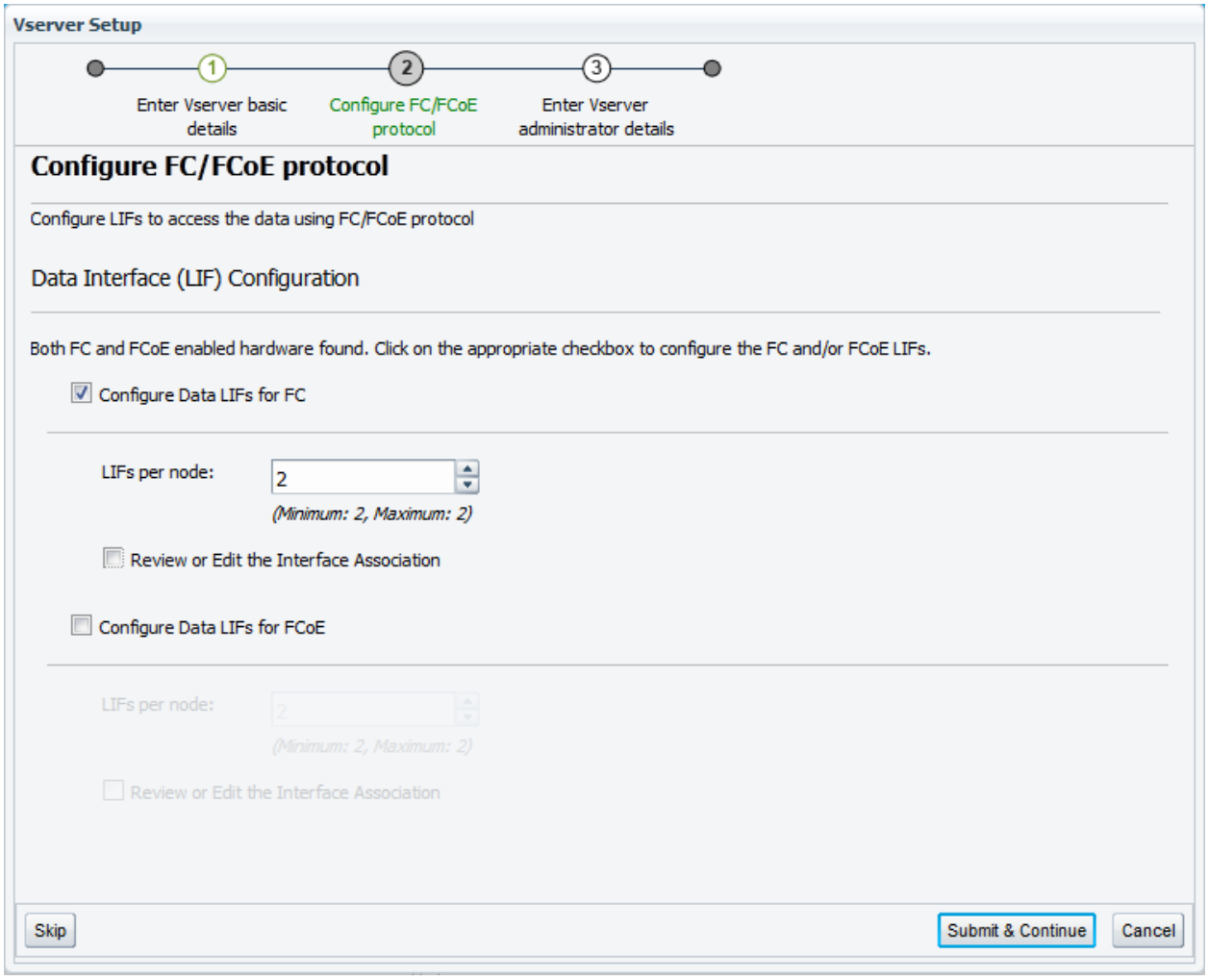
   a. Enter a name for the SVM.
   b. Select the FC/FCoE protocol and any other protocols required.
   c. If NAS protocols are to be used for purposes other than VMware datastores, select an appropriate language.

   Note: After the SVM is created, the language cannot be changed.

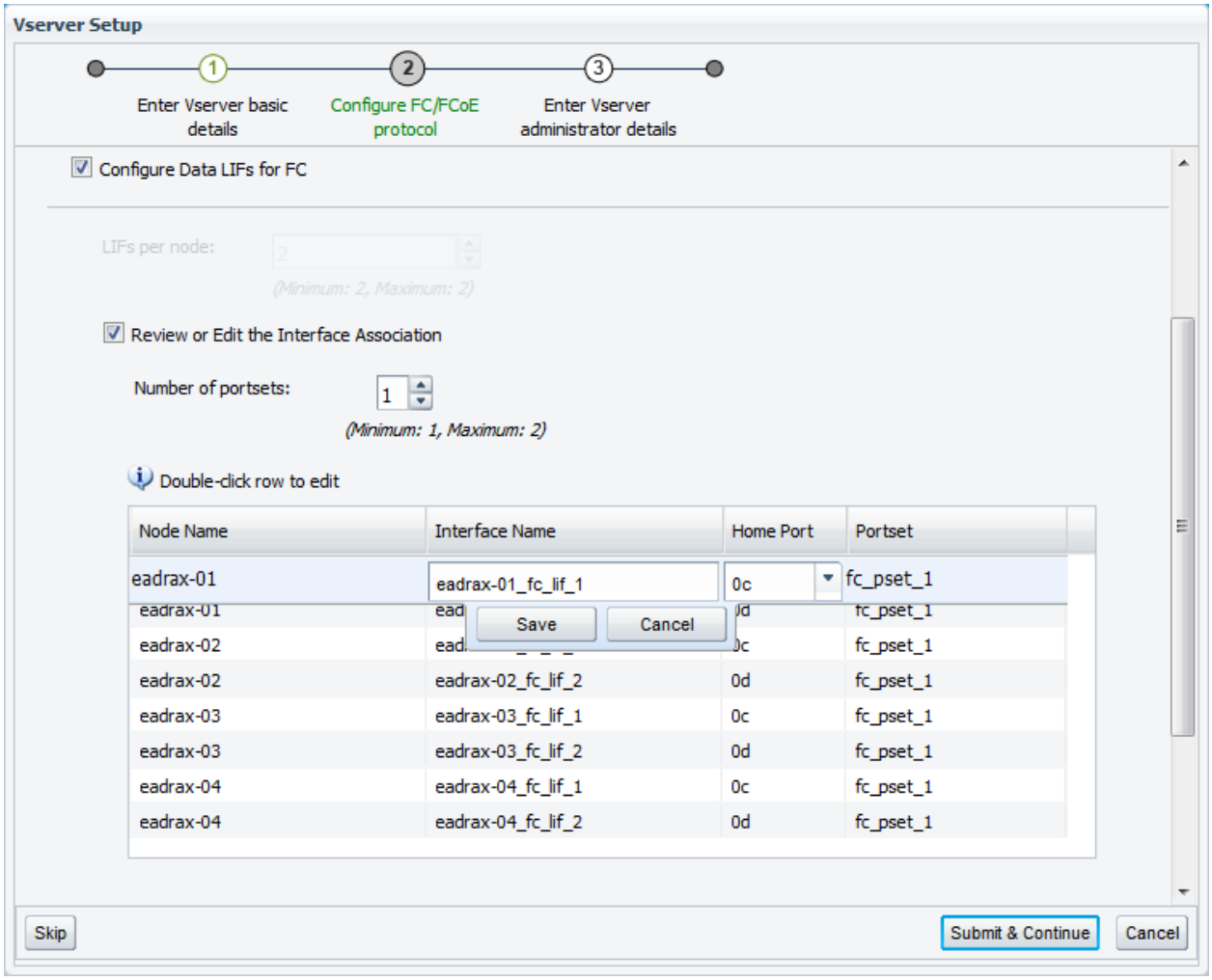
   d. Select a security style. If the SVM serves only LUNs (no NAS datastores), the default of UNIX is acceptable.
   e. Select an aggregate for the SVM root volume.
   f. Enter the DNS domain name and name servers, if required for management or NAS protocols.



6. Click Submit and Continue.
7. Depending on whether FC or FCoE target ports, or both, are installed, properly configured, and connected to properly configured switches, the Configure FC/FCoE Protocol window shows a checkbox for FC, for FCoE, or for both. Select the protocols that you want to configure for this SVM.

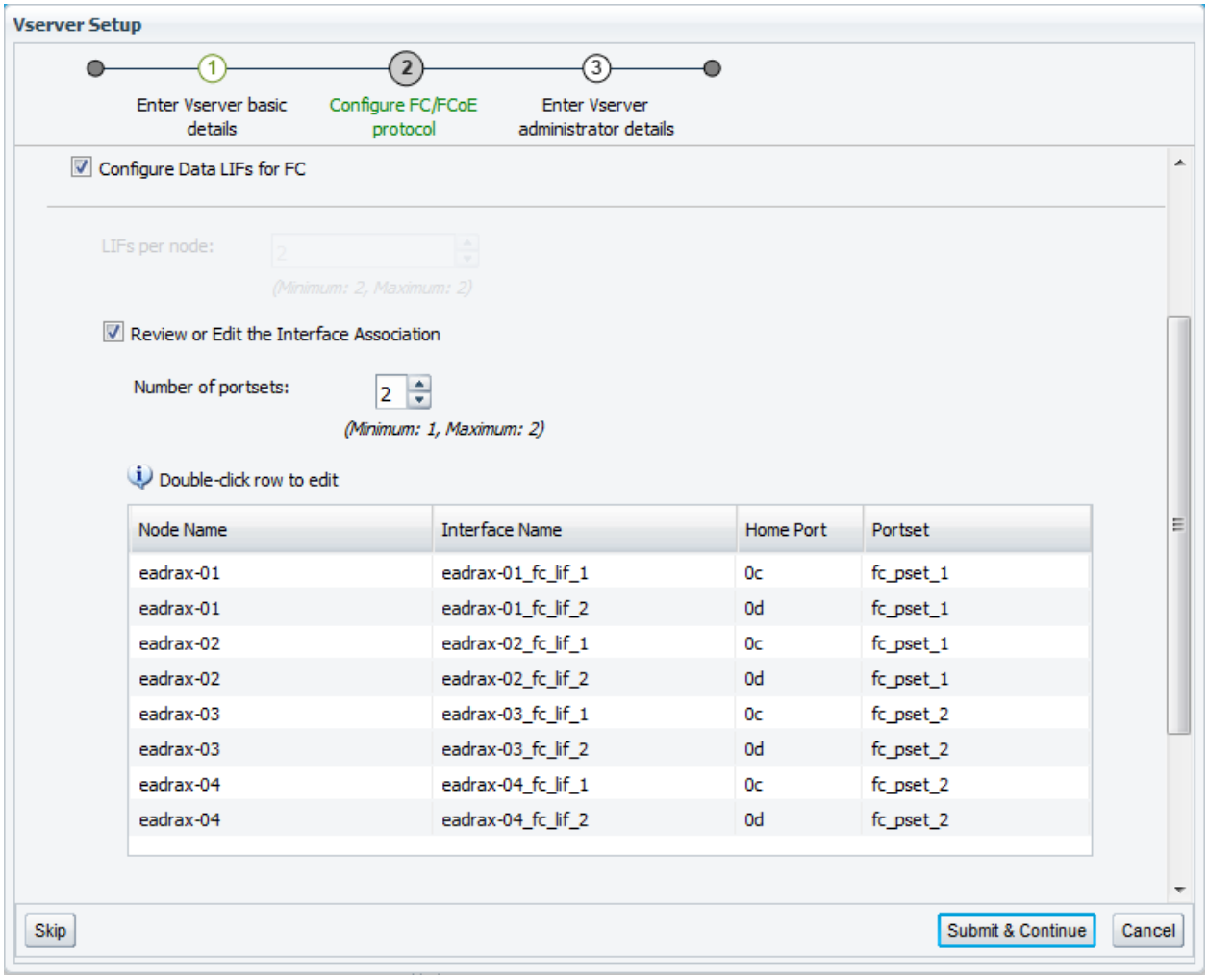


8. If you want to edit LIF names or their association with physical target ports, or to manage portsets, select Review or Edit the Interface Association:
   - To edit a LIF name or to change the target port, double-click the row and edit the name and port and click Save.

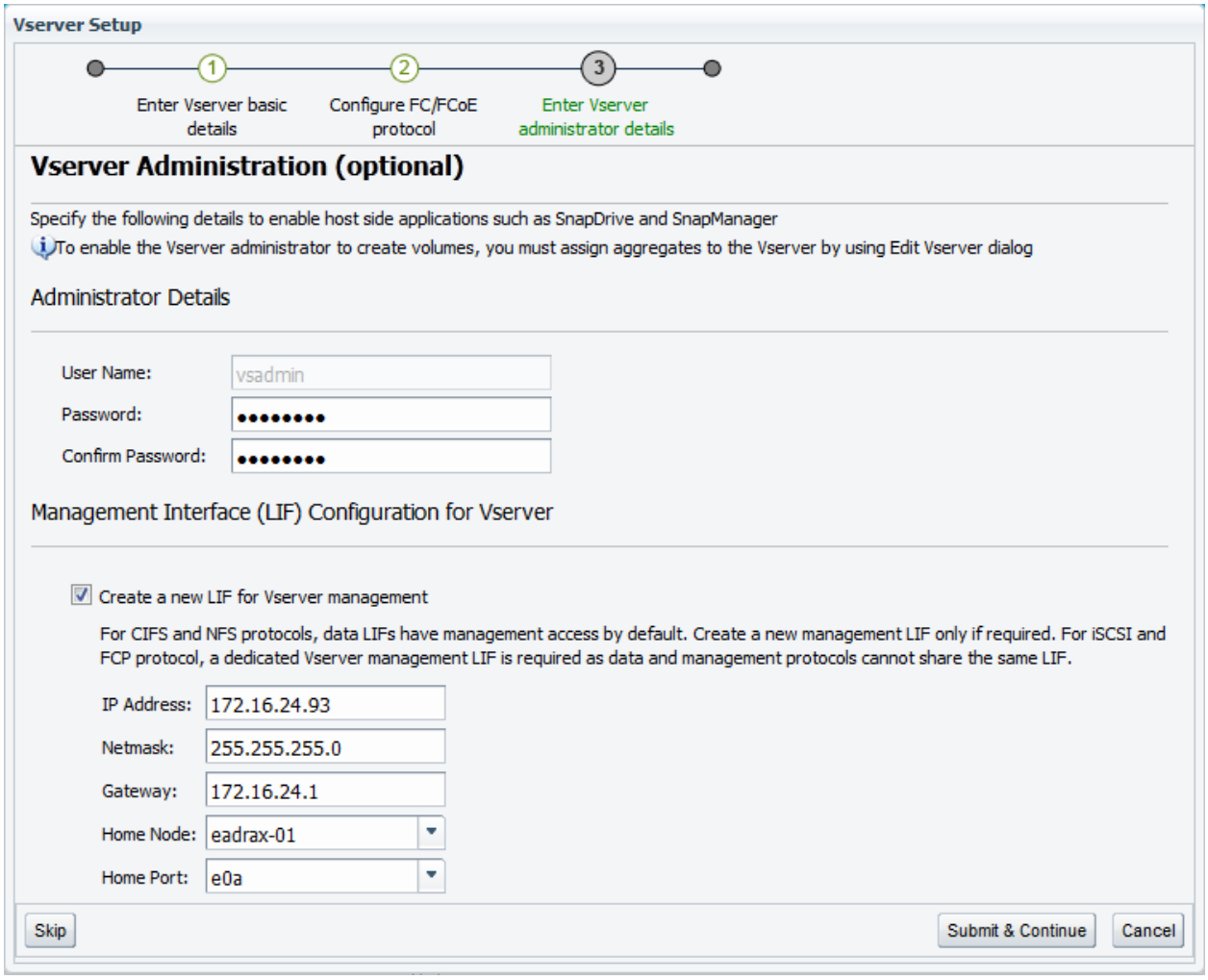


– To reduce path consumption, increase the number of portsets.

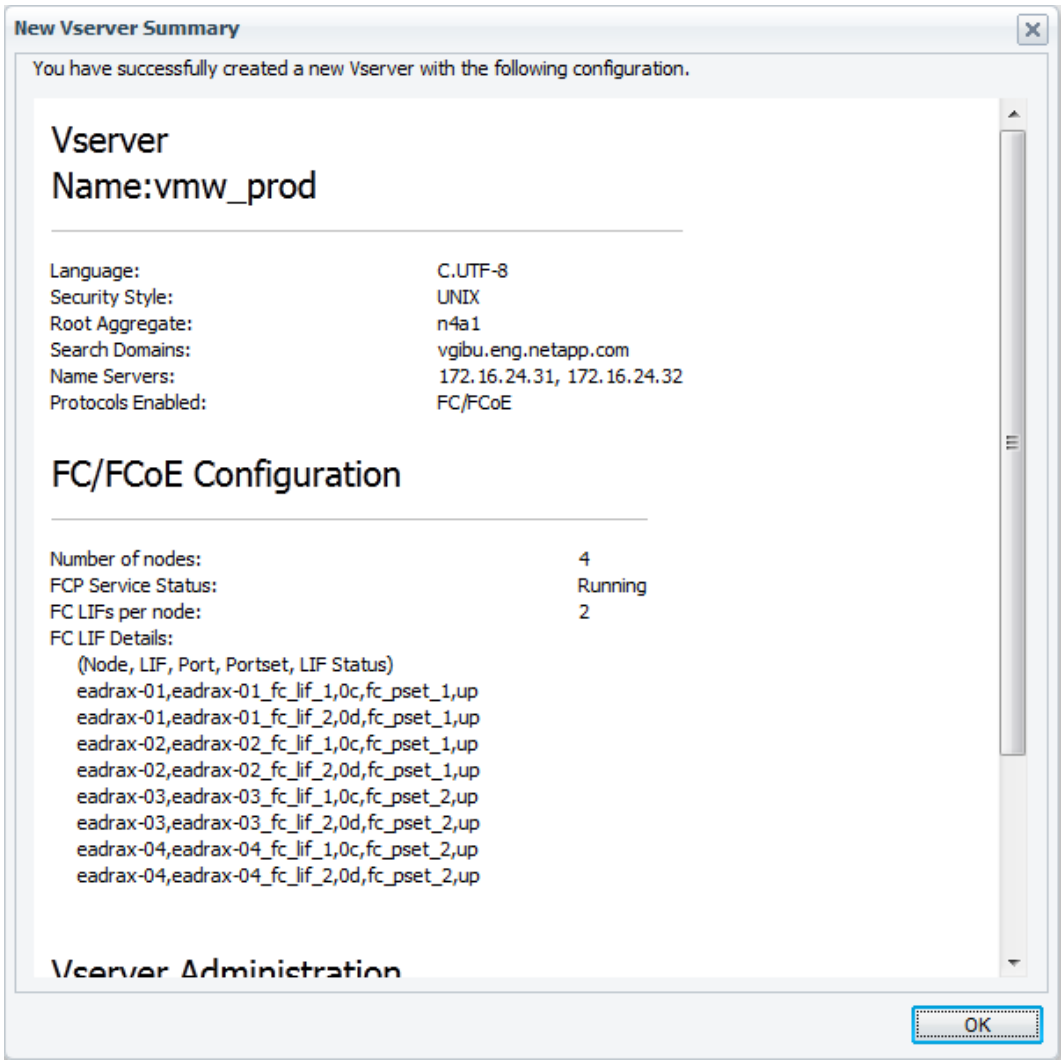
Note: Portsets are used to reduce the number of visible paths through which initiators in an igroup can see LUNs. In a large Data ONTAP cluster, more nodes mean more path consumption for attached servers, and they increase the possibility that servers will run out of paths before they reach all LUNs.



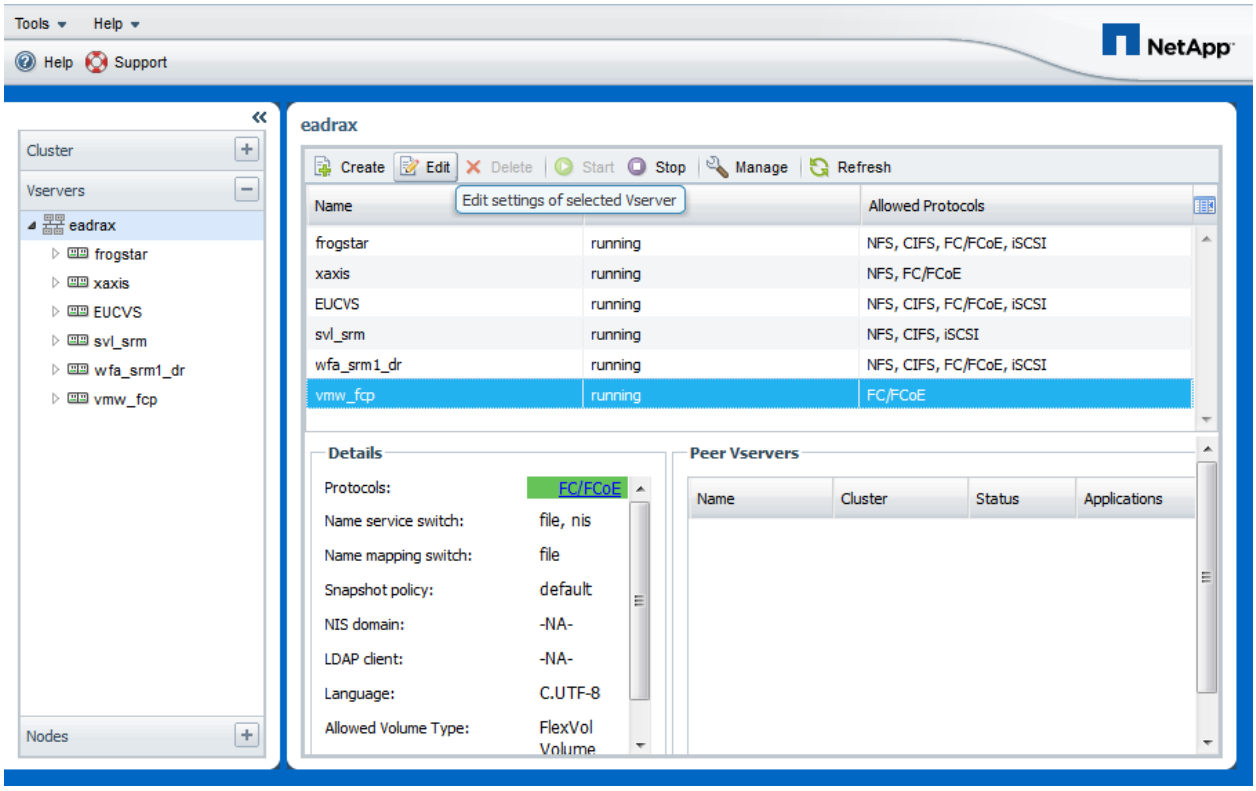
9. Click Submit and Continue.
10. If SVM management is delegated, enter a password for the `vsadmin` account and the configuration details for a management LIF. Be sure to select the correct home node and port. If the SVM is managed by an existing cluster admin account, click Skip.



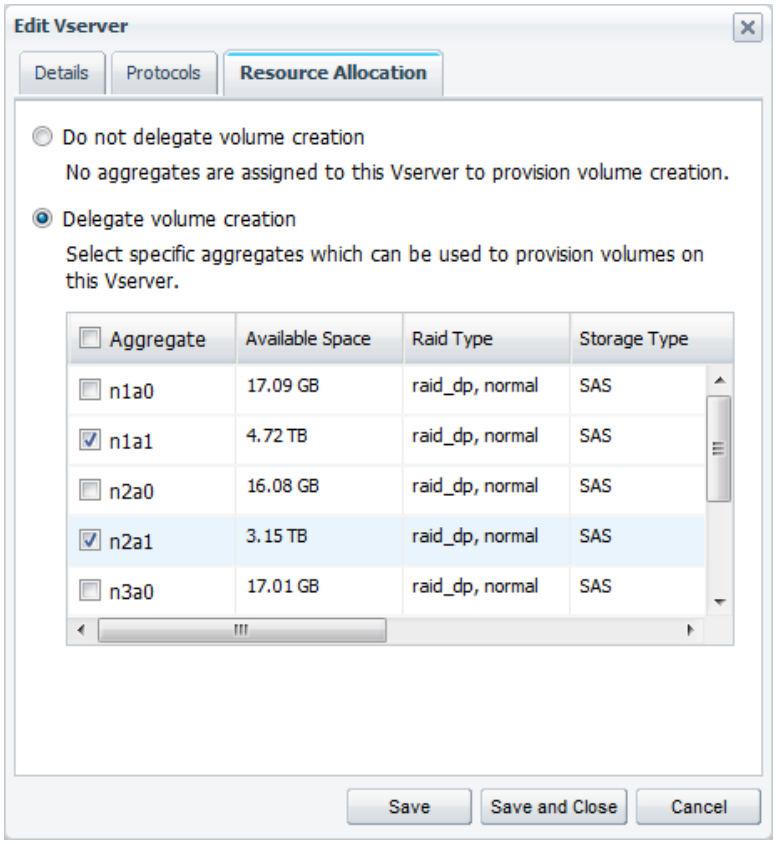
11. Review the New Vserver Summary window and click OK.



12. From the main Vservers window, select the new SVM and click Edit.



13. In the Edit Vserver dialog box, click the Resource Allocation tab.
14. Select Delegate volume creation.
15. Select the aggregates that can be used to provision volumes on this SVM.
16. Click Save and Close.



Configure Zoning on FC Switches

Use the appropriate vendor tools and processes to configure zoning. Be sure to create an alias for the HBA and the associated WWPN. Also create an alias for the SVM.

Best Practice
Zones should include only one initiator but may include multiple storage targets.

Create and Map LUNs

LUNs can be created and mapped by using a variety of tools, including the CLI, System Manager, and others. The administrator must examine the HBA WWPNs of each server to include them in the igroup to which the LUN is mapped. These tools manage the LUN on the storage, but after the LUN is presented to vSphere, additional steps must be completed in ESXi or vCenter to rescan for the LUN and then to partition and format it. All of these steps are handled consistently and reliably in a single wizard and workflow through VSC, as described in section 9.7.

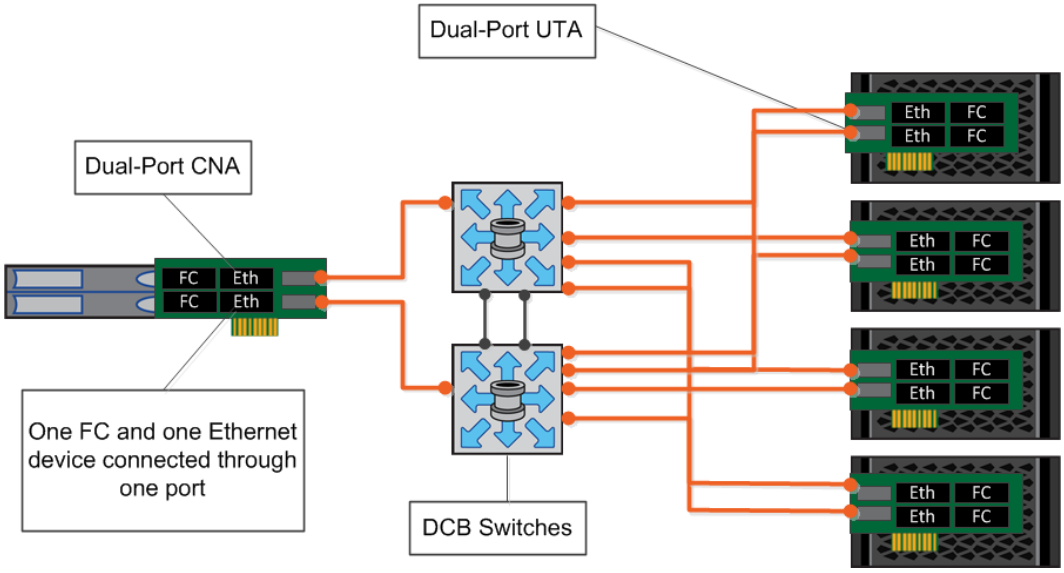
Best Practice
Use VSC to create and manage LUNs and igroups. VSC automatically determines WWPNs of servers and creates appropriate igroups. VSC also configures LUNs according to best practices and maps them to the correct igroups.

7.9 VMware vSphere 6.x Storage Design FCoE Clustered Data ONTAP

FCoE provides FC services over a lossless 10GbE network while preserving the FC protocol. The VMware vSphere 6.x and clustered Data ONTAP solution uses the FCoE protocol. As Figure 23 shows, this solution consists of the following components:

- Converged network adapters (CNAs) in PCIe or proprietary slots in the servers
- Unified target adapters (UTAs) in the NetApp storage system
- FCoE-capable DCB switches

Figure 23) FCoE network with CNAs, UTAs, and DCB switches.



Converged Network Adapters

CNAs might appear to the server as two separate devices in the same PCI slot that connect through the same physical network port. One device is the general-purpose Ethernet NIC, and the other is a standard FC HBA. Dual-port CNAs connect to the server as four devices. The Ethernet and FC HBA devices require separate drivers, which are either included in the build of ESXi or installed as a vSphere Installation Bundle (VIB). Updated drivers are often available or even required as part of a hardware compatibility list (HCL), and they can be obtained in one of two ways:

- As a VIB downloaded from VMware or the CNA vendor
- As part of an ESXi update or rollup

Best Practice
Consult the VMware HCL and the NetApp Interoperability Matrix Tool to determine which drivers are correct for their versions of ESXi, Data ONTAP, and the CNAs.

Figure 24 shows an example of NICs on an ESXi server with a CNA port selected.

Figure 24) NICs on an ESXi server with a CNA port selected.

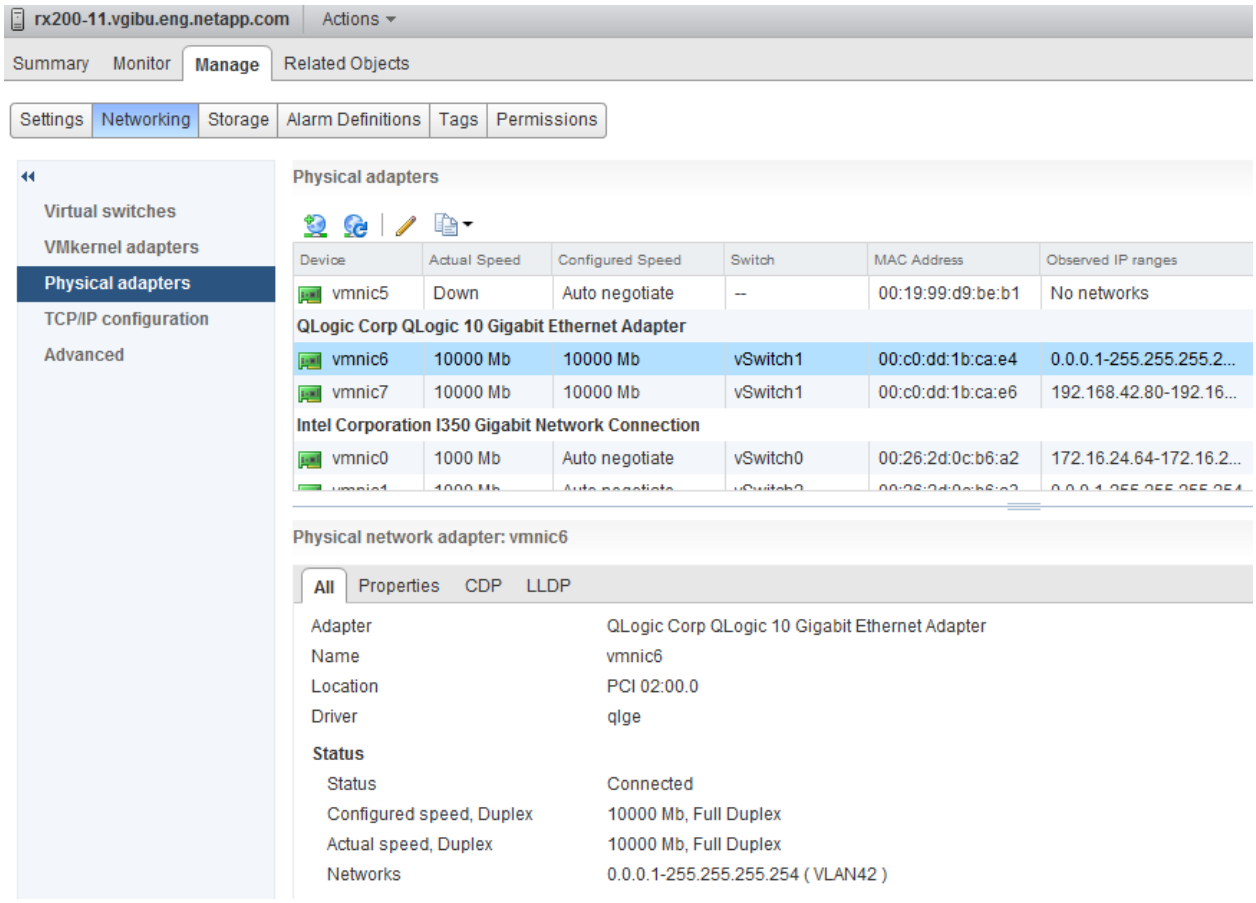
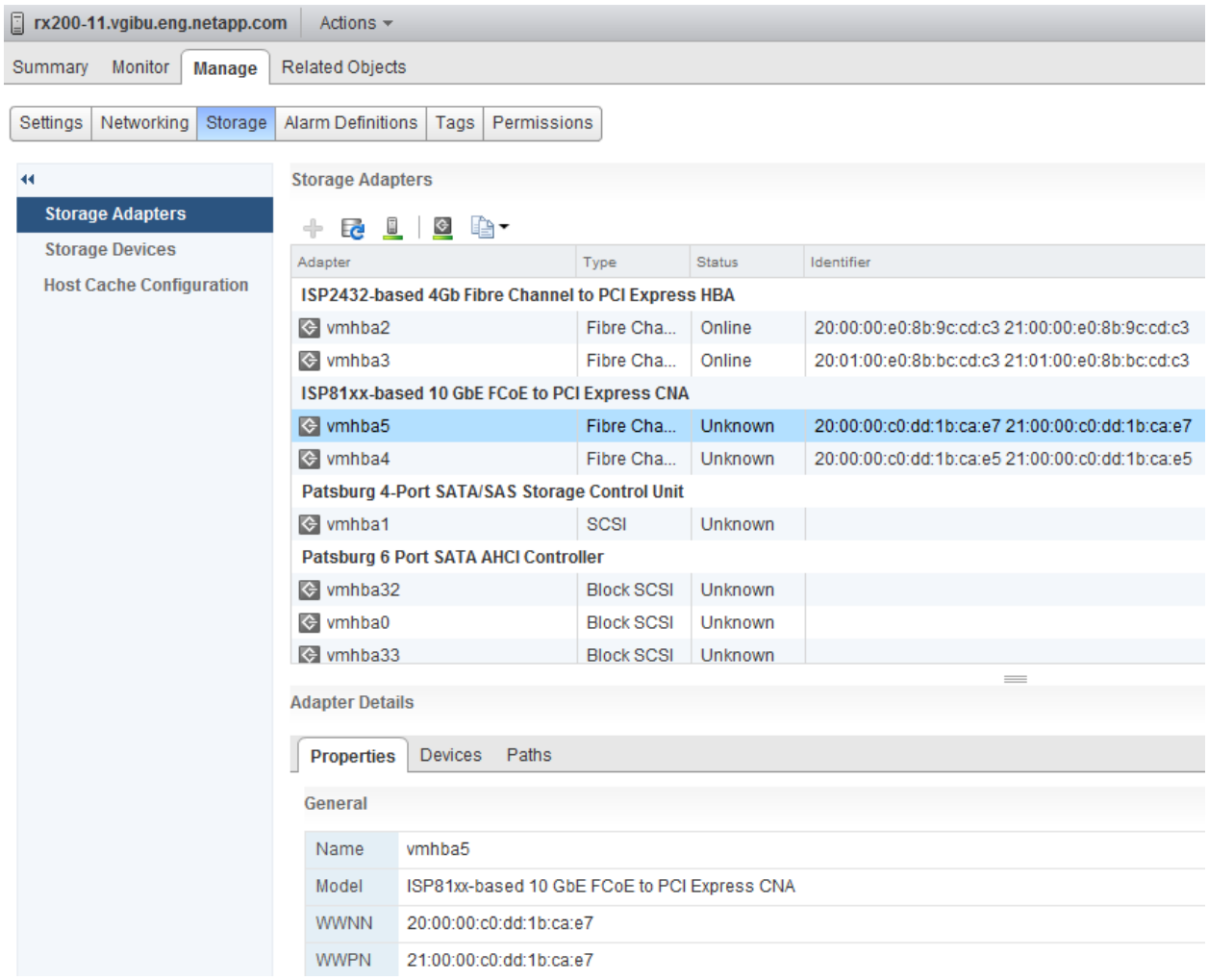


Figure 25 shows an example of storage adapters on an ESXi server with a CNA port selected.

Figure 25) Storage adapters on an ESXi server with a CNA port selected.



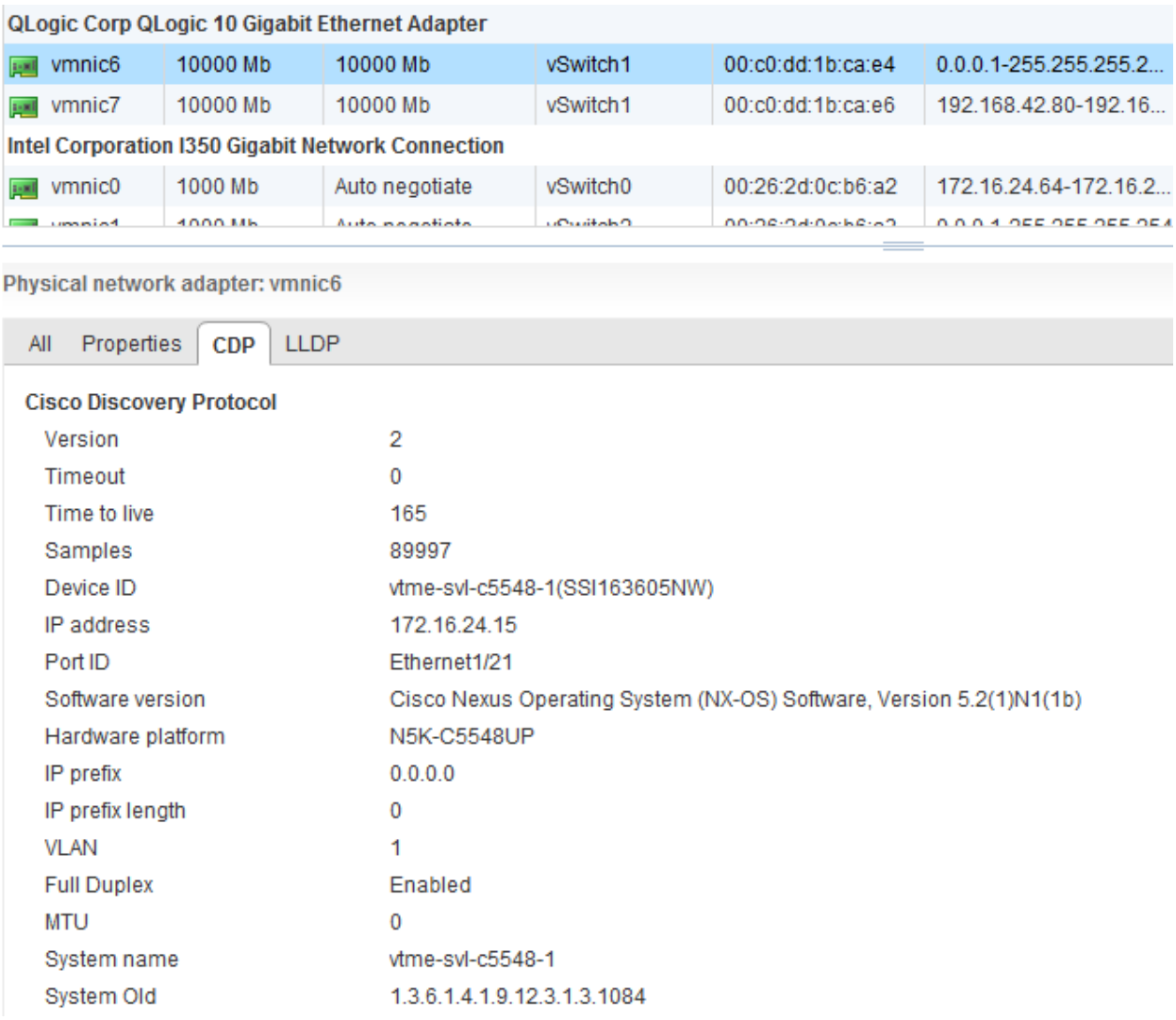
The following ESXi command-line output is similar to the information displayed in Figure 24 and Figure 25. It also shows the PCI `bus:device.function` for the CNA and its ports. For QLogic CNAs, the same `bus:device` numbers indicate that the ports are on the same physical card. Some CNAs use a different device number for each port instead of a function number.

```
~ # esxcfg-nics -l
Name    PCI           Driver      Link Speed     Duplex MAC Address       MTU    Description
vmnic6  0000:02:00.00 qlge        Up   10000Mbps Full   00:c0:dd:1b:ca:b0 9000   QLogic Corp
QLogic 10 Gigabit Ethernet Adapter
vmnic7  0000:02:00.01 qlge        Up   10000Mbps Full   00:c0:dd:1b:ca:b2 9000   QLogic Corp
QLogic 10 Gigabit Ethernet Adapter

~ # esxcfg-scsidevs -a
vmhba2  qla2xxx           link-n/a  fc.200000c0dd1bcab1:210000c0dd1bcab1    (0:2:0.2) QLogic Corp
ISP81xx-based 10 GbE FCoE to PCI Express CNA
vmhba3  qla2xxx           link-n/a  fc.200000c0dd1bcab3:210000c0dd1bcab3    (0:2:0.3) QLogic Corp
ISP81xx-based 10 GbE FCoE to PCI Express CNA
```

There is no Cisco Discovery Protocol (CDP) for FC HBAs. Therefore, to determine to which switch port the HBA side of the CNA is connected, match the HBA to the NIC and then review the CDP information for that NIC. In the vSphere Client, CDP information is available only for network adapters that are part of a virtual switch. Figure 26 shows an example of CDP information for a CNA port using the vSphere Web Client.

Figure 26) CDP information for a CNA port.



Unified Target Adapters

NetApp UTAs are based on QLogic CNAs. As with CNAs in servers, the UTA might appear as two separate devices per physical port with two ports and four devices total in a single slot. An Ethernet interface shows a status of `up` whenever there is a working Ethernet link to a switch. An FC target interface shows a status of `up` only if the corresponding virtual FC interface on the switch is properly configured.

In the example output from the `sysconfig` command, observe the following details:

- In the first example, the Ethernet links are up, but the FC links display the status `LINK NOT CONNECTED`. The switch port is listed as `Unknown`.

```
eadrax-01> sysconfig -a 3
        slot 3: Dual 10G Ethernet Controller CNA SFP+
                (Dual-port, QLogic CNA 8112(8152) rev. 2)
                e3a MAC Address:    00:c0:dd:25:fa:7c (auto-10g_twinax-fd-up)
                e3b MAC Address:    00:c0:dd:25:fa:7e (auto-10g_twinax-fd-up)
                Device Type:        ISP8112
        slot 3: Fibre Channel Target Host Adapter 3a
                (QLogic CNA 8112 (8152) rev. 2, <LINK NOT CONNECTED>)
```

```
                Board Name:         QLE8152
                Serial Number:      RFE1308H45798
                Firmware rev:       5.8.0
                Host Port Addr:     000000
                FC Nodename:        50:0a:09:80:88:f6:6d:a5 (500a098088f66da5)
                FC Portname:        50:0a:09:81:88:f6:6d:a5 (500a098188f66da5)
                Connection:         No link
                Switch Port:        Unknown
                SFP Vendor Name:    CISCO-MOLEX
                SFP Vendor P/N:     74752-9520
                SFP Vendor Rev:     08
                SFP Serial No.:     MOC153601QT
                SFP Connector:      Passive Copper
                SFP Capabilities:   10 Gbit/Sec
```

- In the second example, the switch port is properly configured for FCoE. The adapter shows the status `ONLINE`, and the switch port displays the switch name and virtual Fibre Channel (VFC) port.

```
        slot 3: Fibre Channel Target Host Adapter 3a
                (QLogic CNA 8112 (8152) rev. 2, <ONLINE>)
                Board Name:         QLE8152
                Serial Number:      RFE1308H45307
                Firmware rev:       5.8.0
                Host Port Addr:     db0042
                FC Nodename:        50:0a:09:80:88:b6:71:0c (500a098088b6710c)
                FC Portname:        50:0a:09:81:88:b6:71:0c (500a098188b6710c)
                Connection:         PTP, Fabric
                Switch Port:        vtme-svl-c5548-1:vfc25
                SFP Vendor Name:    CISCO-MOLEX
                SFP Vendor P/N:     74752-9520
                SFP Vendor Rev:     08
                SFP Serial No.:     MOC153600D6
                SFP Connector:      Passive Copper
                SFP Capabilities:   10 Gbit/Sec
```

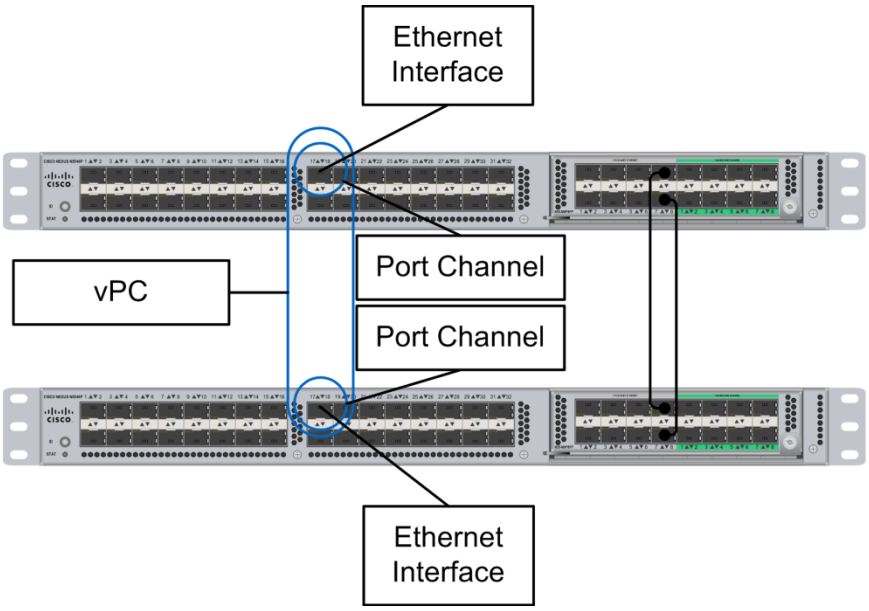
Switches

FCoE requires switches that support DCB with features such as priority-based flow control (PFC) because FC cannot tolerate the latency and loss that can occur with general-purpose Ethernet networks, especially when the networks are congested.

Copper small form-factor pluggable plus (SFP+) cables (also referred to as Twinax because of the twinaxial cable of which they are made) are usually specified by the switch vendor. When optical cables and SFP+ optical transceivers are used, the optical transceivers are specified by the device manufacturer. For NetApp devices, the SFP+ optical transceivers must come from NetApp and must be the correct part for the port, NIC, or HBA in use. For information about the cables and transceivers supported for NetApp devices, refer to TR-3863: 10-Gigabit Ethernet FAQ.

CNAs and UTAs have Ethernet and FC components. Link aggregation can be used with two CNA Ethernet devices, but not with FC. Therefore, the switches use link aggregation for Ethernet traffic, but not for FC traffic. The VFC interfaces are treated as individual interfaces. However, switches may have rules that allow only one physical interface per port channel, even though a port channel on each of two switches is combined into a virtual port channel (VPC), as shown in Figure 27. This allows standard multipath input/output (MPIO) stacks such as ESXi NMP to be used.

Figure 27) FCoE-compliant VPC consisting of two port channels with one interface each.



Because FCoE runs directly over Ethernet, and not over TCP or IP, FCoE traffic cannot be routed over IP gateways to different subnets. This factor must be considered when designing networks for FCoE.

Mixing FC and FCoE

This solution supports mixing FC and FCoE within a single fabric. Initiators in ESXi and targets in NetApp storage systems can be FC, FCoE, or a combination of the two, as shown in Table 26.

Table 26) Supported mixed FC and FCoE configurations.

Initiator	Target	Supported
FC	FC	Yes
FC	FCoE	Yes
FCoE	FC	Yes
FCoE	FCoE	Yes

Supported FCoE Hop Count

Although FCoE does not use regular IP layer 3 gateways to route traffic, it is possible to connect multiple switches by using interswitch links (ISLs) to allow traffic to traverse multiple hops between the initiator (host) and the target (storage system). The hop count is the number of switches in the path between the initiator and the target. Cisco also refers to the hop count as the diameter of the SAN fabric. The maximum supported FCoE hop count between the host and the storage system depends on the switch supplier and the FCoE configurations supported on the storage system.

For FCoE, FCoE switches can be connected to FC switches. For end-to-end FCoE connections, the FCoE switches must be running a firmware version that supports Ethernet ISLs.

Table 27 lists the maximum number of supported hop counts.

Table 27) Maximum number of supported hop counts.

Switch Supplier	Supported Hop Count
Brocade	• 7 for FC • 5 for FCoE
Cisco	7 (up to 3 of which can be FCoE switches)

7.10 Deploying VMware vSphere 6.x Storage over FCoE on Clustered Data ONTAP

Table 28 describes the prerequisites for deploying VMware vSphere 6.x over FCoE on clustered Data ONTAP.

Table 28) VMware vSphere 6.x storage over FCoE on clustered Data ONTAP prerequisites.

Description
Clustered Data ONTAP 8.2 or later is required to use FCoE.
Supported CNAs are required for ESXi servers.
An FC license is required to use FCoE.

The majority of the work in deploying FCoE happens on the switches. For detailed procedures and a running configuration for Cisco Nexus switches, refer to TR-4114: VMware vSphere 6.0 on FlexPod Clustered Data ONTAP Deployment Guide. For other switches, refer to the manufacturer's documentation.

The following procedure is written for vSphere 6.0, but it may be compatible with earlier releases. This procedure is applicable for the X1139A and X1140A UTAs.

Deploy FCoE with vSphere 6.x and Clustered NetApp Data ONTAP 8.3

To deploy FCoE with VMware vSphere 6.x and clustered Data ONTAP 8.3, complete the following steps:

1. Use the NetApp Interoperability Matrix Tool to verify the compatibility of the components in your configuration. Pay attention to any notes linked to the matching configuration, especially to any driver versions required.
2. Install the CNAs in each server. Refer to the particular server and CNA documentation for specific instructions.
3. Install UTAs in the NetApp nodes. Refer to the Data ONTAP 8.2 High-Availability Configuration Guide takeover and giveback procedures for guidance on how to nondisruptively take down individual controllers and install UTAs without service outages.
4. Install the switches and connect the cabling.
5. Complete the following steps to verify that the Ethernet interfaces on the UTAs and CNAs come up after being connected to the switches:
   a. From the vSphere Web Client, select the server, click the Manage tab, and then click Networking > Physical Adapters. Click each CNA NIC and verify that the actual speed is `10000 Mb, Full Duplex`, as shown in Figure 24.
   b. From the Data ONTAP CLI, verify that the `sysconfig -a` output for the CNA Ethernet interfaces shows `10g` as the speed and the word `up` (for example, `auto-10g_twinax-fd-up`).

```
eadrax::> node run -node eadrax-01 -command sysconfig -a 3
        slot 3: Dual 10G Ethernet Controller CNA SFP+
                (Dual-port, QLogic CNA 8112(8152) rev. 2)
```

```
                e3a MAC Address:    00:c0:dd:25:fa:7c (auto-10g_twinax-fd-up)
                e3b MAC Address:    00:c0:dd:25:fa:7e (auto-10g_twinax-fd-up)
                Device Type:        ISP8112
```

6. From the switch, configure the switches to support FCoE:
   a. Add the FCoE license.
   b. Enable FCoE.
   c. Set the correct MTU size (jumbo frames) for the FCoE class.
   d. Create VLANs, VSANs, and the appropriate mappings.
   e. Make sure that the Ethernet interfaces have access to the correct VLANs.
   f. Configure physical Ethernet interfaces or port channels to allow access to the VLAN carrying the VSAN, in addition to other VLANs, such as those for NFS or iSCSI.
   g. Create VFC ports and bind them to the appropriate Ethernet interfaces or port channels.
   h. Verify that HBAs and storage target LIFs log in to the fabric with their WWPNs. The following example is for Cisco Nexus and shows partial output for a QLogic CNA in a server and an FCoE target LIF on an SVM. For SVM LIFs, the `symbolic-port-name` contains the SVM and LIF names.

```
vtme-svl-c5548-1# sho fcns database detail
------------------------
VSAN:1000  FCID:0x9b0000
------------------------
port-wwn (vendor)           :21:00:00:c0:dd:1b:ca:df (Qlogic)
node-wwn                    :20:00:00:c0:dd:1b:ca:df
fc4-types:fc4_features      :scsi-fcp:init
symbolic-port-name          :
symbolic-node-name          :QLE8152 FW:v5.01.03 DVR:v902.k1.1-12vmw
port-type                   :N
permanent-port-wwn (vendor) :21:00:00:c0:dd:1b:ca:df (Qlogic)
connected interface         :vfc17
------------------------
VSAN:1000  FCID:0x9b0041
------------------------
port-wwn (vendor)           :20:0a:00:a0:98:3c:3e:4c (NetApp)
node-wwn                    :20:00:00:a0:98:0d:ee:e6
fc4-types:fc4_features      :scsi-fcp:target
symbolic-port-name          :NetApp FC Target Port (8112) xaxis:fcoe-edx3-3b
symbolic-node-name          :NetApp Vserver xaxis
port-type                   :N
permanent-port-wwn (vendor) :20:0a:00:a0:98:3c:3e:4c (NetApp)
connected interface         :vfc25
```

   i. Configure device aliases and zones.

Note: Step h and step i require the SVM and its LIFs to be created before they are visible and have WWPNs that can be zoned.

7. Run the `fcp initiator show` command to verify that server initiators are seen by the SVM target LIFs.

```
eadrax::> fcp initiator show -vserver xaxis
          Logical           Port     Initiator    Initiator
Vserver   Interface         Address  WWNN         WWPN         Igroup
--------- ----------------- -------- ------------ ------------ ---------------
xaxis     fcoe-edx1-3a      db0140   20:00:00:c0:dd:1b:ca:7b
                                                  21:00:00:c0:dd:1b:ca:7b
                                                               -
xaxis     fcoe-edx1-3a      db0160   20:00:00:c0:dd:1b:ca:99
                                                  21:00:00:c0:dd:1b:ca:99
                                                               -
xaxis     fcoe-edx1-3b      9b0140   20:00:00:c0:dd:1b:ca:79
                                                  21:00:00:c0:dd:1b:ca:79
                                                               -
xaxis     fcoe-edx1-3b      9b0160   20:00:00:c0:dd:1b:ca:9b
```

```
                                                  21:00:00:c0:dd:1b:ca:9b
                                                                       -
xaxis     fcoe-edx2-3a      9b0140   20:00:00:c0:dd:1b:ca:79
                                                  21:00:00:c0:dd:1b:ca:79
                                                                       -
xaxis     fcoe-edx2-3a      9b0160   20:00:00:c0:dd:1b:ca:9b
                                                  21:00:00:c0:dd:1b:ca:9b
                                                                       -
xaxis     fcoe-edx2-3b      db0140   20:00:00:c0:dd:1b:ca:7b
                                                  21:00:00:c0:dd:1b:ca:7b
                                                                       -
xaxis     fcoe-edx2-3b      db0160   20:00:00:c0:dd:1b:ca:99
                                                  21:00:00:c0:dd:1b:ca:99
                                                                       -
```

Create and Map LUNs

LUNs can be created and mapped by using a variety of tools, including the CLI, OnCommand System Manager, and others. The administrator must examine the HBA WWPNs of each server to include them in the igroup to which the LUN is mapped. These tools manage the LUN on the storage, but after the LUN is presented to vSphere, additional steps must be completed in ESXi or vCenter to rescan for the LUN and then to partition and format it. All of these steps are handled consistently and reliably in a single wizard and workflow through VSC, as described in section 9.7.

7.11 VMware vSphere 6.x Storage Design Using iSCSI on Clustered Data ONTAP

iSCSI is an Ethernet protocol that transports SCSI commands over an IP-based network. It functions similarly to FC in providing block storage to the initiator (or storage consumer) from the target (or storage provider); however, it uses IP rather than FC as the transport. This leads to some differences; for example, FC has buffers that prevent frames from being dropped in the event of congestion. However, because iSCSI is IP based, it relies on SCSI to be tolerant of dropped Ethernet frames, and TCP retransmits when congestion or errors occur.

Two initiator options are available when VMware is configured to use iSCSI: the software initiator and hardware initiators. The software initiator is provided as a feature of ESXi and is available in all versions of vSphere. Hardware initiators are hardware add-in cards that are commonly provided by the server hardware vendor. Table 29 summarizes the advantages and disadvantages of each option.

Table 29) iSCSI initiator options advantages and disadvantages.

Initiator	Advantages	Disadvantages
Software	• Available for all servers • No additional hardware needed	Additional host CPU consumption
Hardware	• Low or no CPU impact to host • Can be used to connect boot LUNs	Additional cost associated with hardware

iSCSI LUNs hosted by clustered Data ONTAP support all of the VMware VAAI primitives that are available in vSphere 5.x and later, including full copy, block zeroing, hardware-assisted locking, and thin provisioning.

vSphere 5.5 introduced the maximum VMFS datastore size of 64TB, with a maximum VMDK size of 62TB. Data ONTAP supports a maximum volume size of 100TB; however, the maximum single file size is 16TB. This is also the maximum size of a single LUN, which therefore limits the maximum VMDK to 16TB as well.

Network Considerations

iSCSI uses standard network switches for transporting data, which makes the network a critical component of the storage architecture. If the network is not highly available and is not capable of providing enough throughput for VM storage activity, significant performance and availability issues can result.

NetApp recommends connecting both the NetApp controllers in the cluster that provides iSCSI service, and the vSphere hosts that use iSCSI LUNs, to more than one physical switch to provide redundancy in the event of a switch failure. The user should also work closely with the network administrator to make sure the environment does not contain any severely oversubscribed uplinks that are heavily used by iSCSI data connections. These bottlenecks can cause unpredictable performance issues and should be avoided if possible.

Since vSphere 5, it has been possible to route from the iSCSI initiator to the target, unless port binding is in use. However, NetApp still recommends using a dedicated iSCSI VLAN that is available to all participating hosts. This method isolates the unprotected storage traffic and should also provide a contiguous network subnet for connectivity.

Network Connectivity

This section discusses both basic and highly available iSCSI connectivity, ALUA, and path multiplicity; in addition, it provides NetApp recommendations for iSCSI networks.

Basic iSCSI Connectivity

Simple connectivity to the iSCSI target can be provided with a single network interface connected to a standard or distributed virtual switch. A VMkernel network adapter connected to the iSCSI network subnet must be configured. Similarly, the NetApp controller can be connected by using a single interface to the iSCSI network, with or without a VLAN. This configuration provides the throughput of a single link; therefore, it does not provide for highly available storage connectivity, and it is not recommended for critical storage traffic.

Highly Available iSCSI Connectivity

Production iSCSI traffic should always be connected through highly available network connections. HA can be provided through several methods, some of which also increase the available throughput to the iSCSI target.

VMkernel ports to be used for iSCSI traffic can be created on virtual switches that use NIC teaming technology, such as LACP, EtherChannel (IP Hash), and other link-aggregation techniques. However, VMware recommends using port binding instead of these technologies. iSCSI storage still functions when teaming technology is used, although link state changes can cause unpredictable behavior when multipathing is in use.

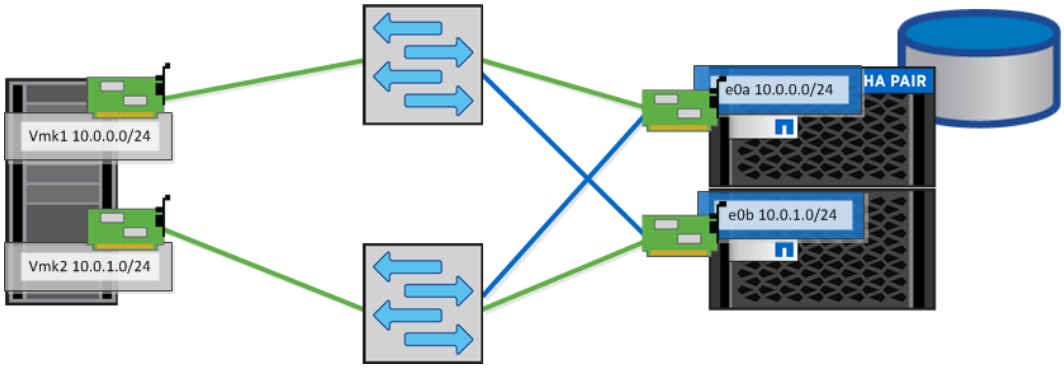
Note: NIC teaming alone does not increase the potential throughput for a single LUN. NIC teaming works by creating a hash of the source and destination IP addresses, or MAC, depending on the configuration, in an attempt to create an even distribution of unicast client-to-server traffic across the available links. However, a single stream of data can traverse only a single uplink, thus limiting the maximum throughput for a single-path LUN to the link speed of a single NIC in the team.

Increased throughput and HA can be achieved by using multiple paths to the LUNs. vSphere uses two methods to present multiple paths from the target to the initiator:

- Multiple VMkernel NICs that connect to the storage system
- VMware port binding

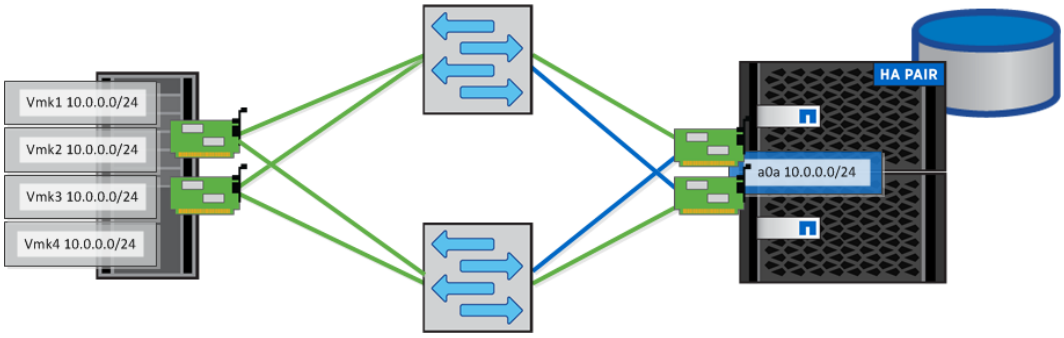
The first method uses multiple VMkernel NICs, each on a different subnet, that connect to the storage system. The storage system has interfaces on those same subnets, and each interface adds an additional path that can be used for I/O operations on the LUN. The physical NICs should be connected to at least two different physical switches to protect against failure. If a single virtual switch is used, whether standard or distributed, the VMkernel interfaces should each be configured to use a single, unshared physical NIC. If multiple virtual switches, each with a single iSCSI VMkernel NIC, are used, then regardless of the number of virtual switch uplinks, it is not necessary to assign dedicated interfaces. Nevertheless, NetApp recommends assigning dedicated interfaces to improve predictability. Figure 28 shows multipath connectivity from the vSphere host to the NetApp LUN.

Figure 28) Multipath connectivity from vSphere host to NetApp LUN.



The second method uses VMware port binding. Port binding is accomplished by creating multiple VMkernel interfaces on the same subnet and binding them to the software iSCSI storage adapter. The iSCSI adapter then creates multiple sessions to the target, increasing throughput (assuming that the round-robin approach is used) and availability (assuming that multiple physical NICs connected to multiple physical switches are used). For virtual switches that have multiple physical NICs, each VMkernel interface must be assigned to a specific interface. When multiple virtual switches are used, assignment to a specific interface is not necessary. Port binding does not support routing of the iSCSI network. Figure 29 shows the use of port binding to achieve multipath LUN connectivity.

Figure 29) Use of port binding to achieve multipath LUN connectivity.



Note: When port binding is used, all of the interfaces must be in the same network subnet. Routing between the iSCSI initiator and target is not supported.

Asymmetrical Logical Unit Access

Clustered Data ONTAP SVMs can have many LIFs supporting multiple protocols. A NetApp best practice is to have an iSCSI LIF for the SVM on each physical node in the cluster.

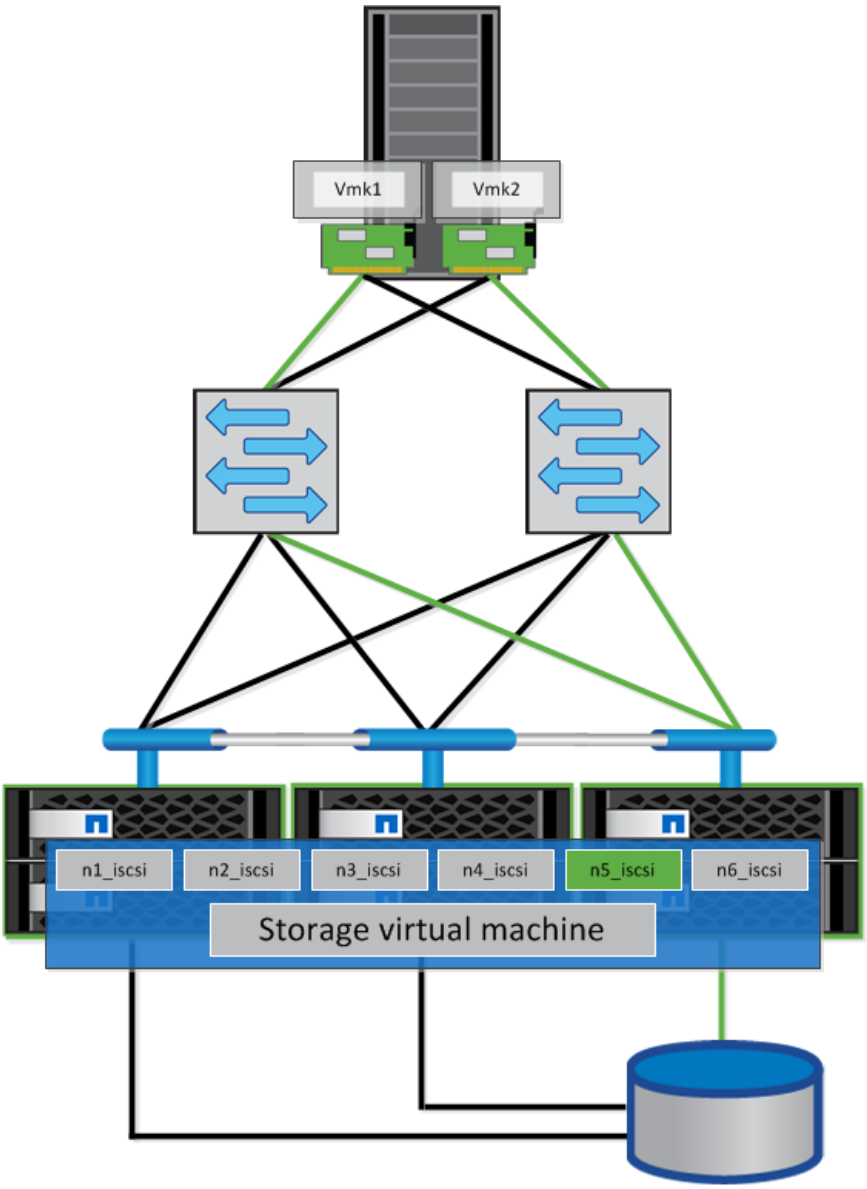
Note: Clustered Data ONTAP and the storage client use ALUA to correctly choose the path that provides direct access to the LUNs.

The client host discovers which paths are optimized by sending a status query to the iSCSI LUN host down each path. For the paths that lead to the clustered Data ONTAP node that directly owns the LUN, the path status is returned as `active/optimized`; for other paths, the status is returned as `active/non-optimized`. The client prefers the optimized path whenever possible. Without ALUA, a path that traverses the cluster network can be selected as the primary path for data. This configuration still functions as expected (that is, all of the LIFs accept and process the LUN reads and writes). However, because it is not the optimal path, a small amount of additional latency is incurred from traversing the cluster network.

LIFs that are used for block protocols such as iSCSI cannot be migrated between nodes of the clustered Data ONTAP system. This not the case with CIFS and NFS; with these protocols, the LIFs can be moved to any of the physical adapters that the SVM can access. In the event of failover, the takeover node becomes the new optimized path until giveback is performed. However, if the primary node is still active, but the LIF is inactive because of network failure or other conditions, the path priority remains the same. A volume move between nodes causes the node paths to be reevaluated and the client to select the new direct/optimized path as the primary path.

Figure 30 shows the ALUA path selection from the iSCSI initiator to the iSCSI target.

Figure 30) ALUA path selection from iSCSI initiator to iSCSI target.



Path Multiplicity

NetApp recommends that the clustered Data ONTAP SVMs that serve iSCSI LUNs have a LIF on each physical node to provide HA to iSCSI clients. By default, a LUN is masked to be available from all of these interfaces. However, this arrangement can potentially lead to path exhaustion for VMware hosts. VMware supports a maximum of 256 LUNs with a total of 1,024 paths per host. If the clustered Data ONTAP cluster contains more than four nodes and/or multiple iSCSI networks per node, it is possible to encounter the path maximum before the LUN maximum. To prevent this from happening, the initiator (igroup) can use portsets to mask the LUN to a subset of interfaces. At minimum, a portset for a LUN must contain the node that hosts the volume and its HA partner. When using portsets and performing volume moves, be careful not to place the volume on a node that does not have a member interface of the portset. Although this does not result in losing access to the LUN, it forces an indirect path to be used.

iSCSI Network Recommendations

The Data ONTAP operating system can use multiple sessions across multiple networks. When designing the iSCSI storage network, NetApp recommends using multiple VMkernel network interfaces on different network subnets that use NIC teaming, in the case of multiple virtual switches, or being pinned to physical NICs connected to multiple physical switches, to provide HA and increased throughput.

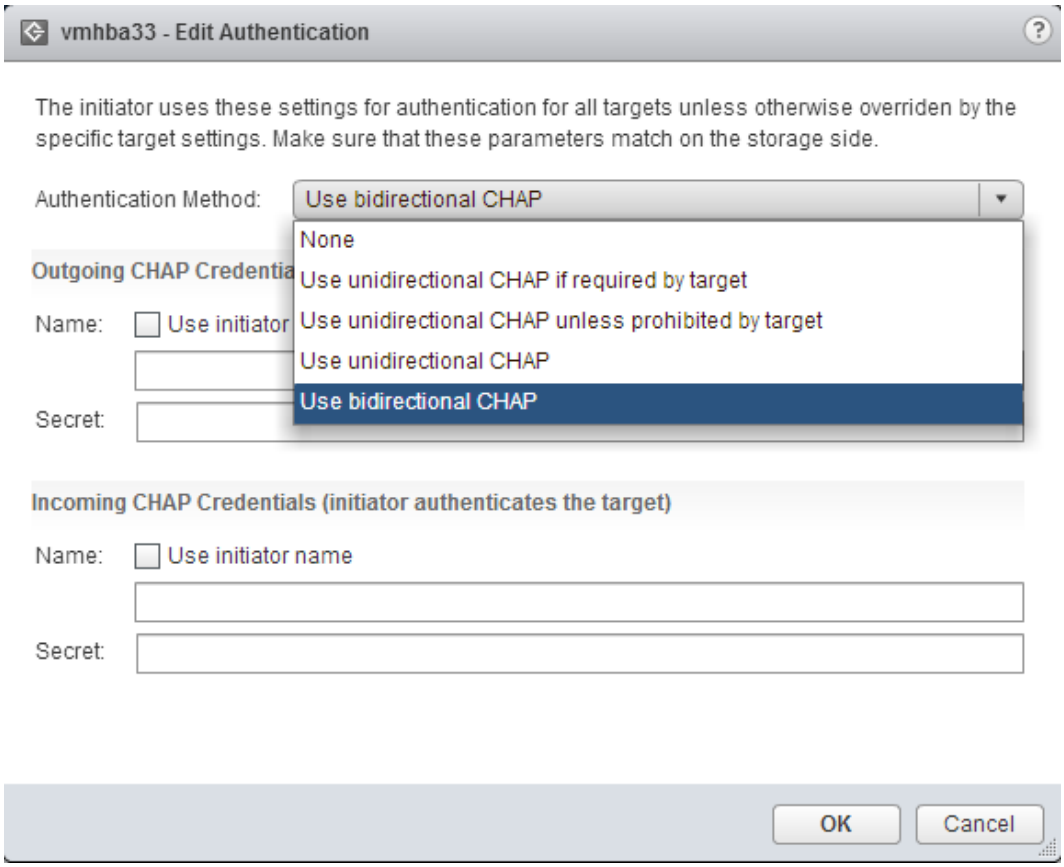
For Data ONTAP controller connectivity, NetApp recommends configuring a minimum of a single-mode interface group for failover with two or more links that are connected to two or more switches. Ideally, LACP or other link-aggregation technology should be used with multimode interface groups to provide HA and the benefits of link aggregation. The NetApp storage system is the iSCSI target, and many different initiators, with many different IPs, are connected to it, which significantly increases the effectiveness of link aggregation.

By default, ALUA is enabled for all iSCSI LUNs serviced by clustered Data ONTAP. NetApp recommends using portsets to mask extraneous paths to the LUNs and prevent path exhaustion for the VMware hosts. Using the VSC to manage VMware storage automatically applies the NetApp best practices for provisioning and presenting LUNs to initiators.

Target and Initiator Authentication for Increased Security

The only authentication method supported by VMware is the Challenge-Handshake Authentication Protocol (CHAP). CHAP works by configuring a shared password on the target and initiator that is used to verify the entities to each other. CHAP can work as unidirectional (that is, from the initiator to the target) or bidirectional (that is, from the initiator to the target and vice versa). Figure 31 shows an example of CHAP authentication.

Figure 31) Configuring CHAP authentication for the vSphere software iSCSI initiator.



7.12 Deploying VMware vSphere 6.x Storage over iSCSI on Clustered Data ONTAP

Table 30 describes prerequisites for vSphere storage design with iSCSI and clustered Data ONTAP.

Table 30) VMware vSphere 6.x storage design iSCSI clustered Data ONTAP prerequisites.

Description
iSCSI must be licensed and enabled on the storage controller that is running clustered Data ONTAP.

Enable VMware Software iSCSI Adapter Using ESXi CLI

To enable the VMware software iSCSI adapter by using the ESXi CLI, complete the following steps:

1. Connect to the ESXi host by using the console or SSH.

Note: If using the console, enable the ESXi shell from the direct console user interface (DCUI) by selecting the option under the Troubleshooting Mode Options menu. After enabling the ESXi shell for the console, press Alt + F1 to change terminals to the console.

2. Run the following command:

```
~ # esxcli iscsi software set --enabled true
```

3. Determine the IQN.

Note: Be sure to substitute the correct HBA identifier, such as `vmhba1`, for your system.

```
~ # esxcli iscsi adapter get --adapter=<iscsi_hba>
```

4. Configure the adapter for connectivity to the NetApp storage controller.

```
~ # esxcli iscsi adapter discovery sendtarget add –A <iscsi_hba> -a <netapp_iscsi_int>
```

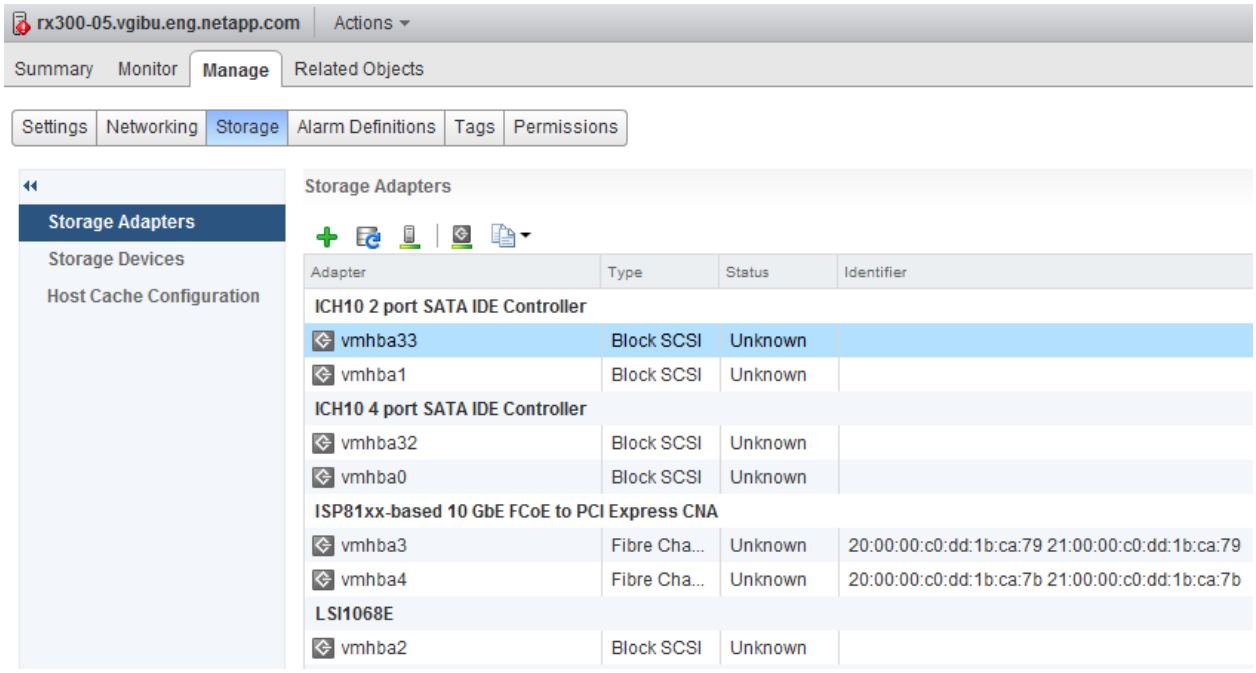
5. The software iSCSI adapter has been configured on the ESXi host. If LUNs were mapped to the host, perform an HBA rescan by running the following command:

```
~ # esxcli iscsi adapter discovery rediscover
```

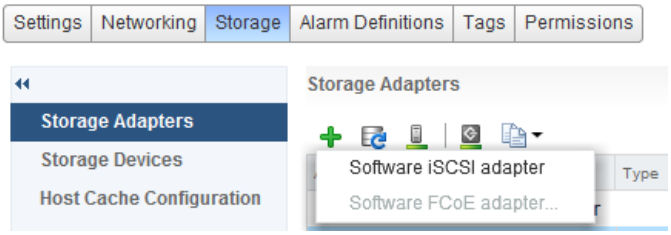
Enable VMware Software iSCSI Adapter Using vSphere Web Client

To enable the VMware software iSCSI adapter by using the vSphere Web Client, complete the following steps:

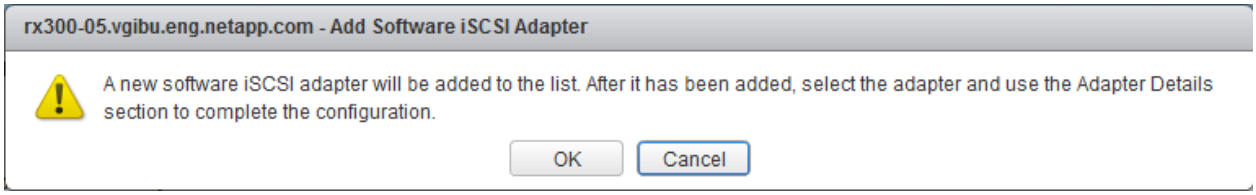
1. Using a web browser, log in to the vSphere Web Client.
2. Navigate to an ESXi server.
3. Click the Manage tab > Storage > Storage Adapters.



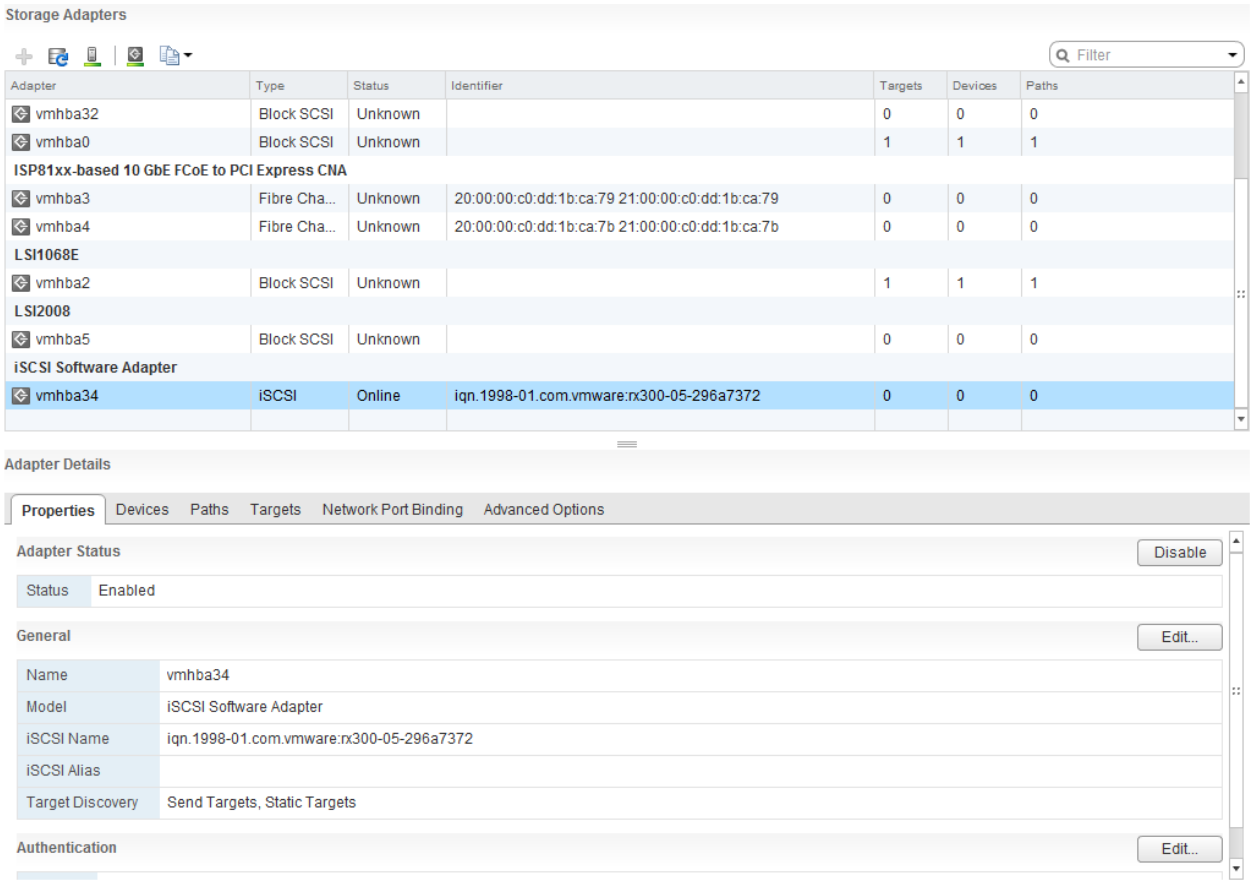
4. Click the green plus icon.



5. Click Software iSCSI adapter.
6. Click OK.

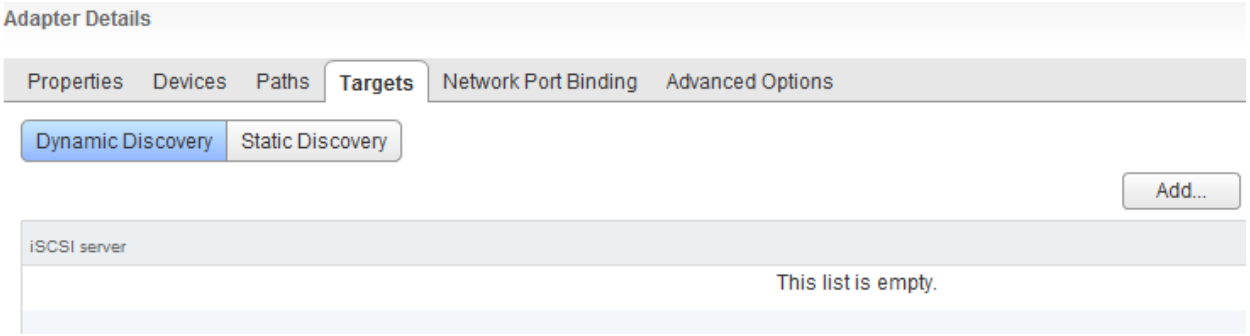


7. Select the newly created adapter from the list.

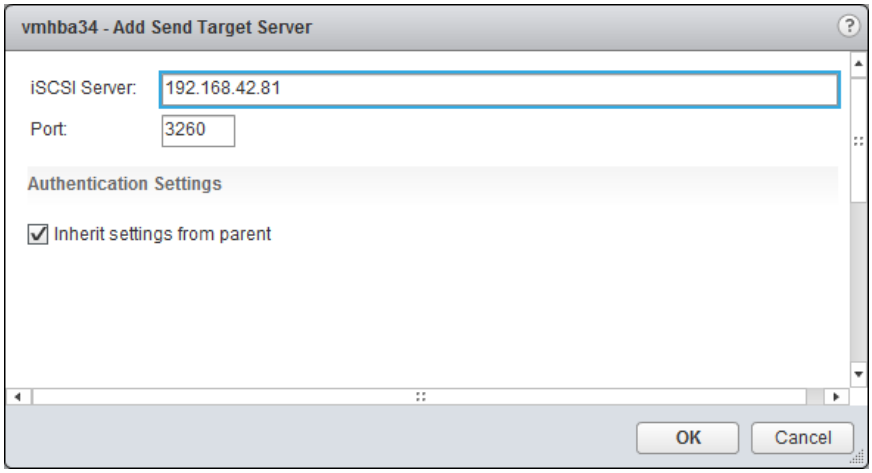


8. Make note of the iSCSI name (IQN), which is needed to map LUNs to this host from the NetApp storage controller.
9. To add an iSCSI target, click the Targets tab > Dynamic Discovery and click Add.

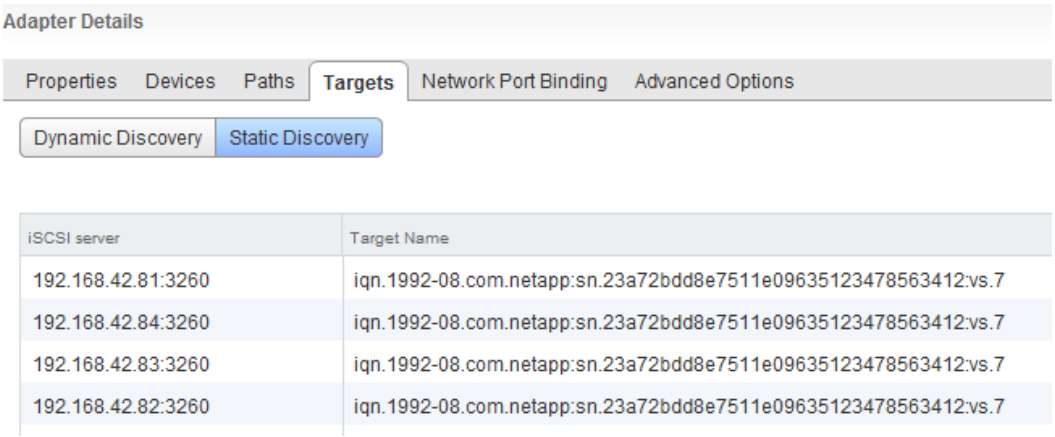
   Note: This configures the ESXi host to log in to the NetApp controller and query for available LUNs.



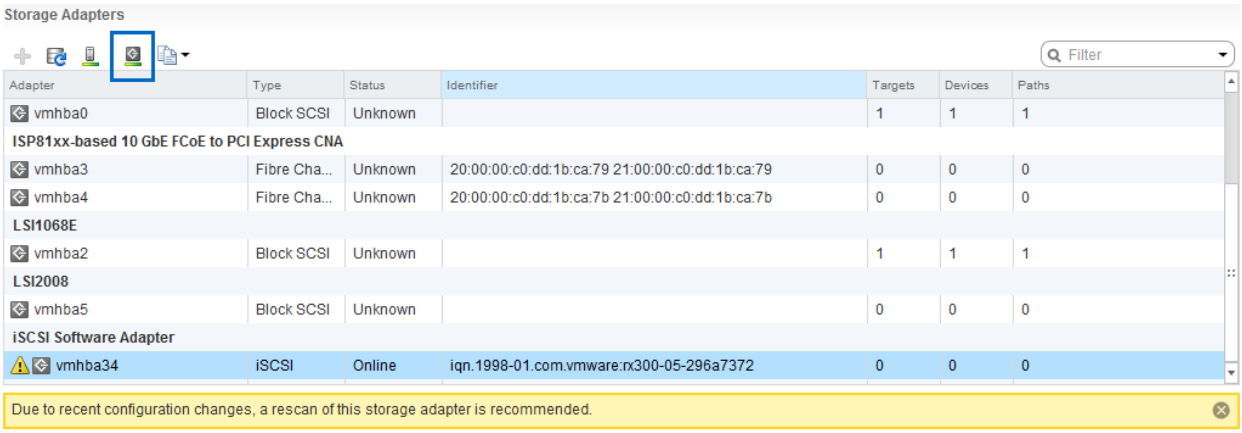
10. In the Add Send Target Server dialog box, enter the IP address used for iSCSI traffic, click OK, and wait for the Add Internet SCSI Send Targets task to complete.



11. Click Static Discovery. The iSCSI IPs of the SVM should all have populated.



12. The software iSCSI adapter has been configured for the ESXi host. If LUNs were mapped to the ESXi host, click the SCSI symbol to rescan the iSCSI vmhba to discover them.



Configure VMware Software iSCSI Port Binding Using ESXi CLI

Note: Using iSCSI port binding limits storage connectivity to a single layer 2 network domain. Storage traffic cannot be routed when port binding is used.

To configure the VMware software iSCSI port binding by using the ESXi CLI, complete the following steps:

Note: The VMkernel network adapters used for iSCSI connectivity must already be created, and the IP addresses must be assigned in the same network subnet as the NetApp controller's iSCSI interface.

1. If using a single virtual switch with multiple uplinks, associate the VMkernel adapter to a single uplink by removing the other uplinks from the portgroup. (These steps are not necessary if your VMkernel ports are on different virtual switches.) In the command line, replace `vmnic_identifier` with the identifier of the vmnic you want to remove from this VMkernel (for example, `vmnic0)` and the `iSCSI_VMkernel_NIC_PG` with the port group used for the VMkernel NIC (for example, `iSCSI_A`).

   Note: Repeat this command for each VMkernel NIC/uplink pair until only a single uplink remains. The VMkernel NICs should have different uplink vmnics associated.

```
~ # esxcfg-vswitch –N <vmnic_identifier> -p <iSCSI_VMkernel_NIC_PG> <vSwitch>
```

2. To bind multiple VMkernel NICs to the iSCSI initiator, repeat this command for each of the VMkernel NICs. In the command line, replace `vmkernel_nic` with the vmk identifier (for example, `vmk2`) and the `vmhba_identifier` with the vmhba designator of the software iSCSI HBA (for example, `vmhba33`).

```
~ # esxcli iscsi networkportal add –-nic <vmkernel_nic> --adapter <vmbha_identifier>
```

3. Verify the binding.

```
~ # esxcli swiscsi nic list –adapter <vmhba_identifier>
```

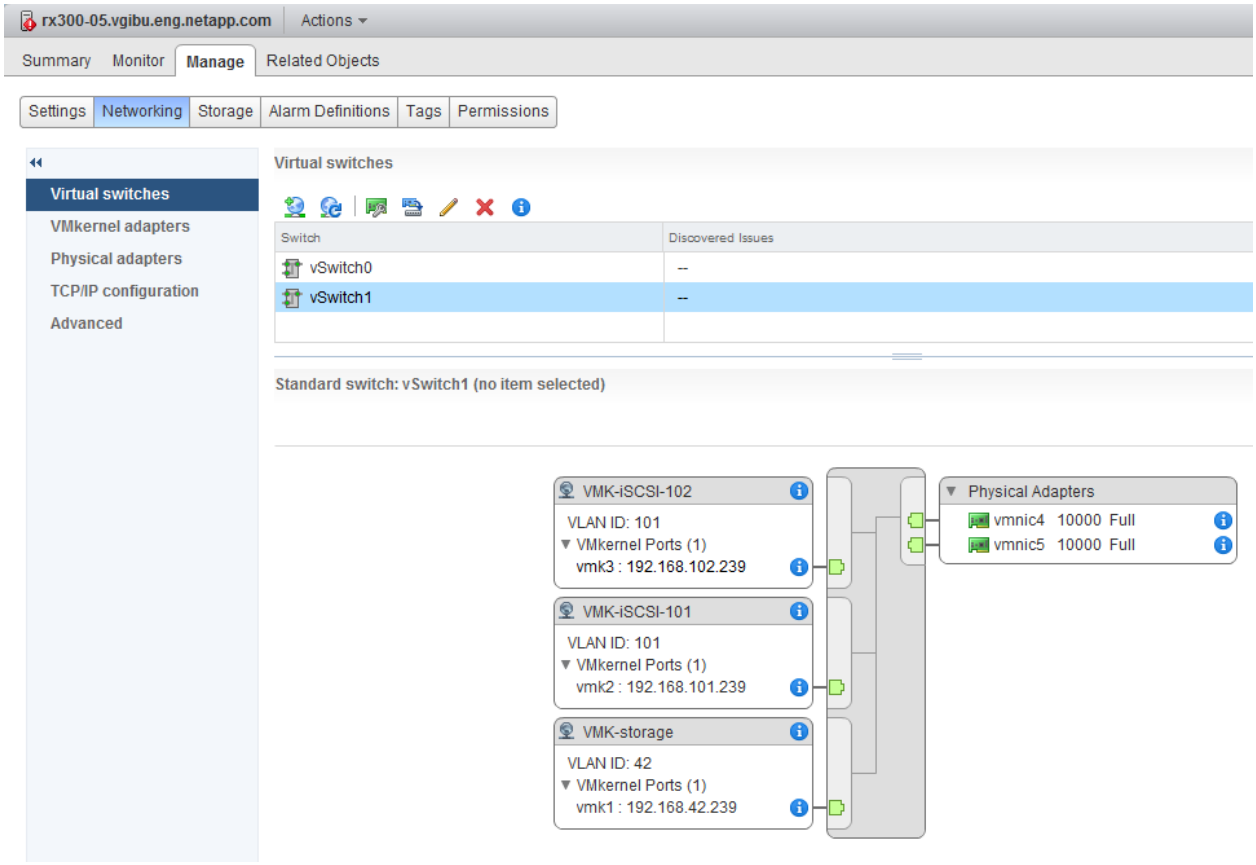
4. Verify that port binding for the software iSCSI adapter is now active.

Configure VMware Software iSCSI Port Binding Using vSphere Web Client

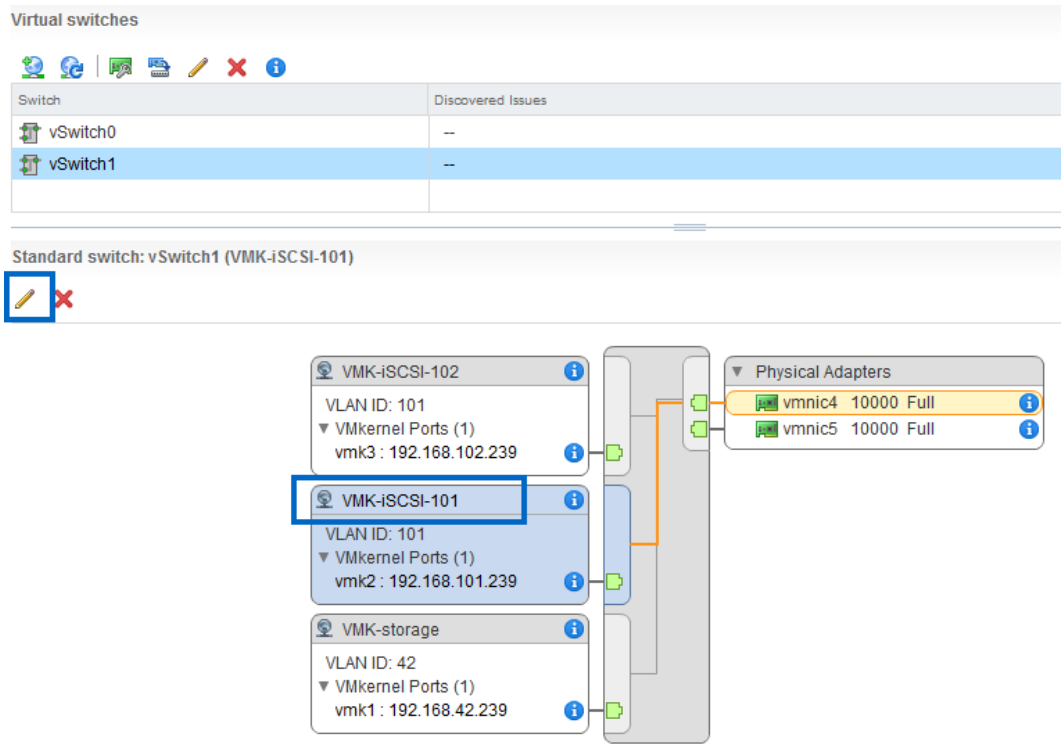
To configure the VMware software iSCSI port binding by using the vSphere Web Client, complete the following steps:

Note: The VMkernel network adapters that are used for iSCSI connectivity must be created and assigned their IP addresses in the same network subnets as the NetApp SVM iSCSI LIFs.

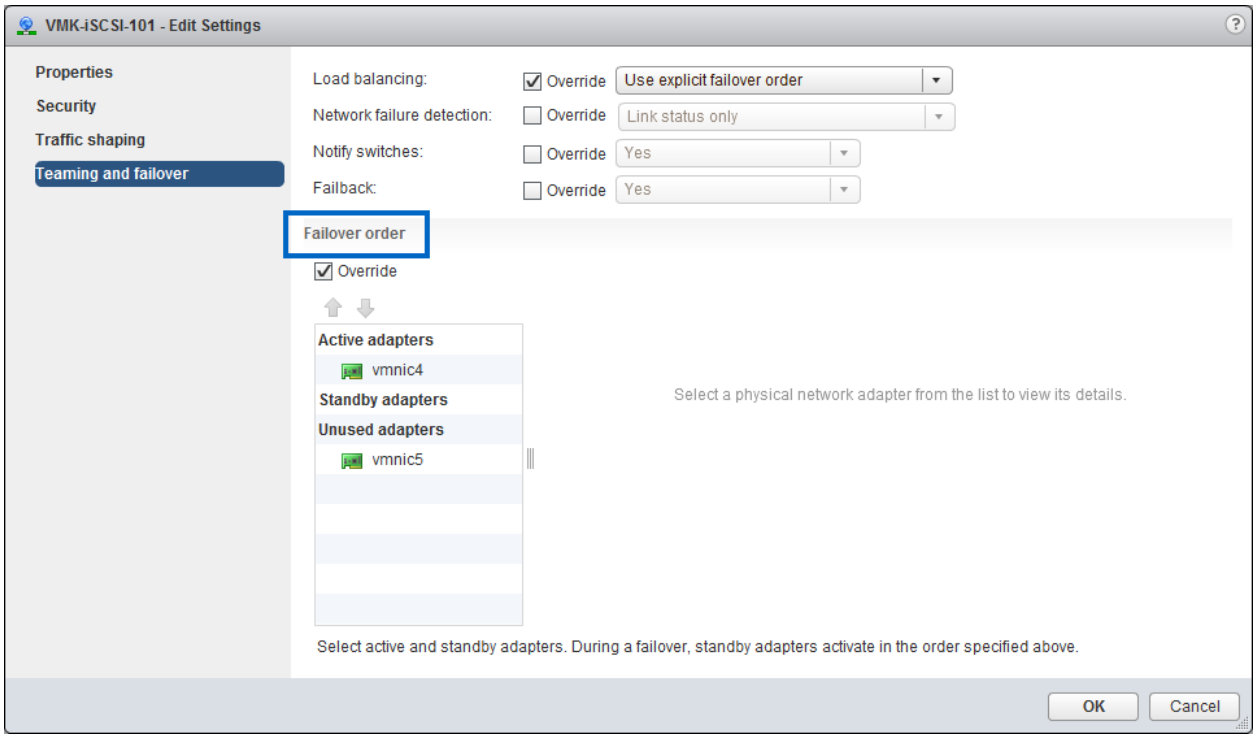
1. Using the vSphere Web Client, navigate to the ESXi server.
2. When using a single virtual switch with multiple uplinks, bind the VMkernel NICs to physical uplinks:
   a. Click the Manage tab > Networking > Virtual switches.
   b. Click the virtual switch used for storage traffic.



c. In the diagram, click the first iSCSI VMkernel port, then click the pencil icon below the Standard Switch heading.



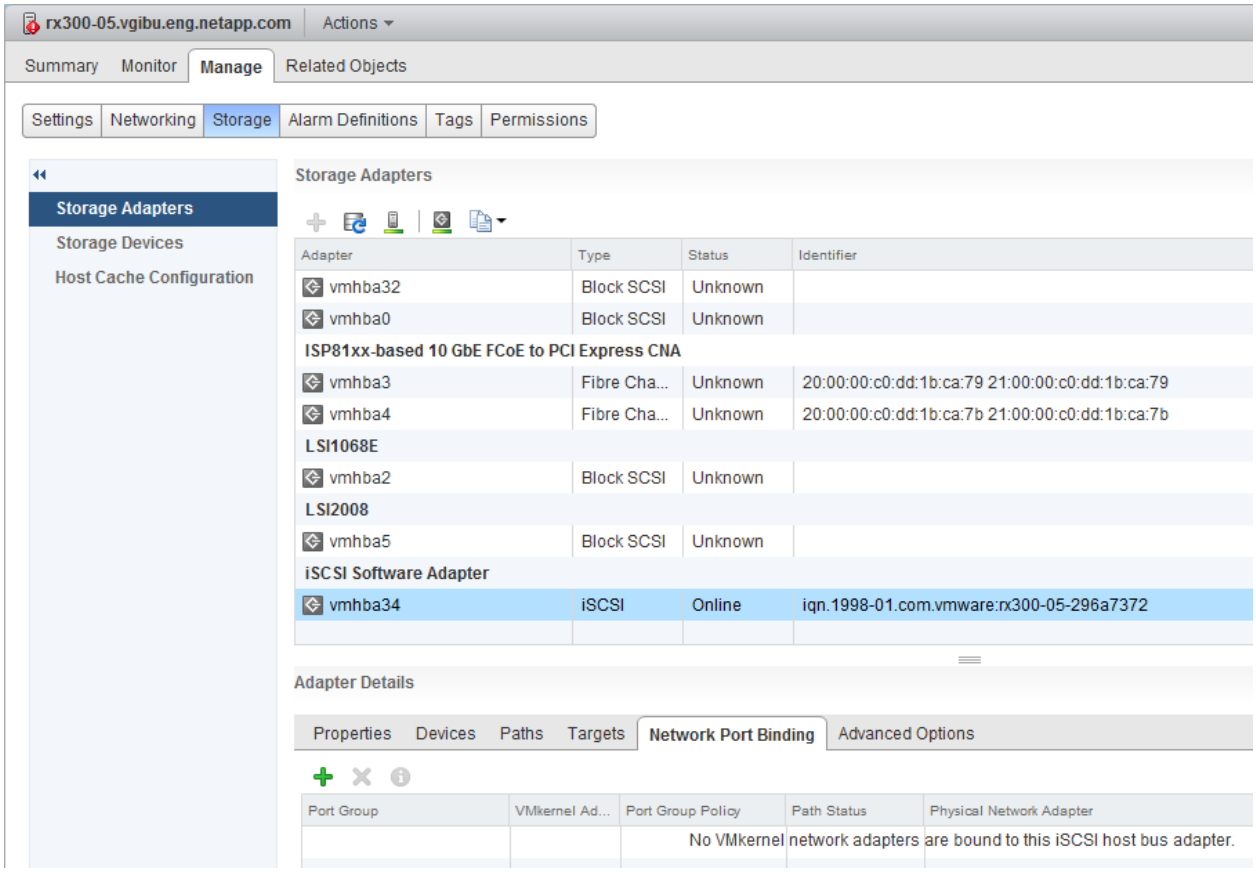
d. In the left pane, select Teaming and Failover. From Failover Order, select the Override checkbox and then adjust the adapters so that only one is active for the VMkernel NIC. Move all other NICs under Unused Adapters.



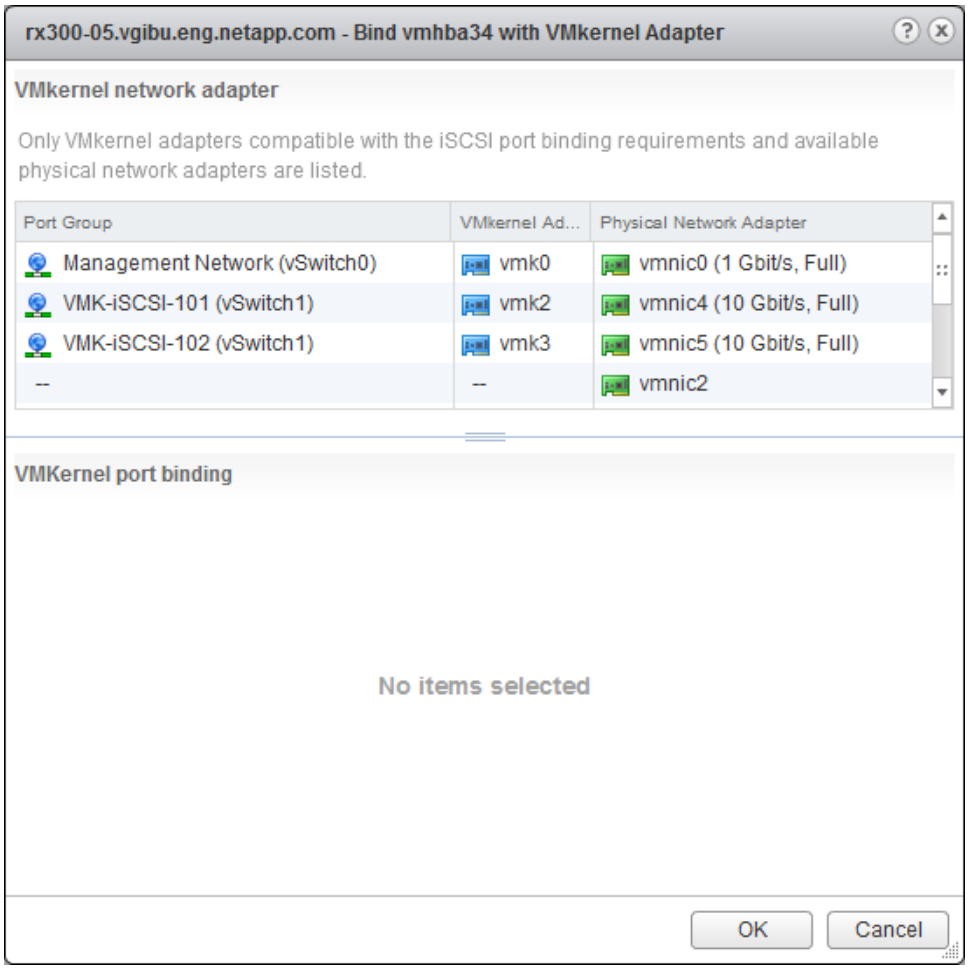
Note: Make sure that the VMkernel adapters are not using the same physical NICs. No VMkernel adapter can have an associated vmnic shared by other iSCSI VMkernel NICs. As you click each VMkernel adapter, the connectivity to its associated physical adapters is highlighted in orange.

e. Repeat step a through step d for all VMkernel iSCSI adapters participating in the port binding configuration.

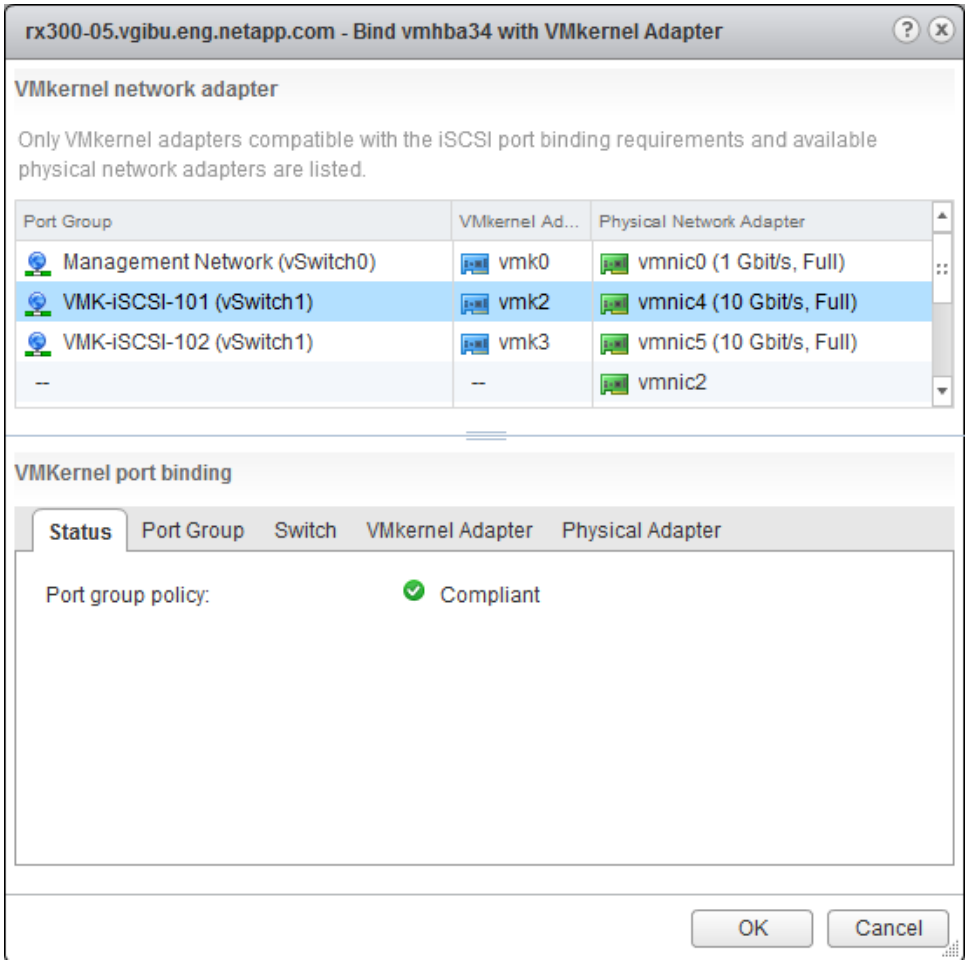
3. From the ESXi Host Manage tab, click Storage > Storage Adapters.



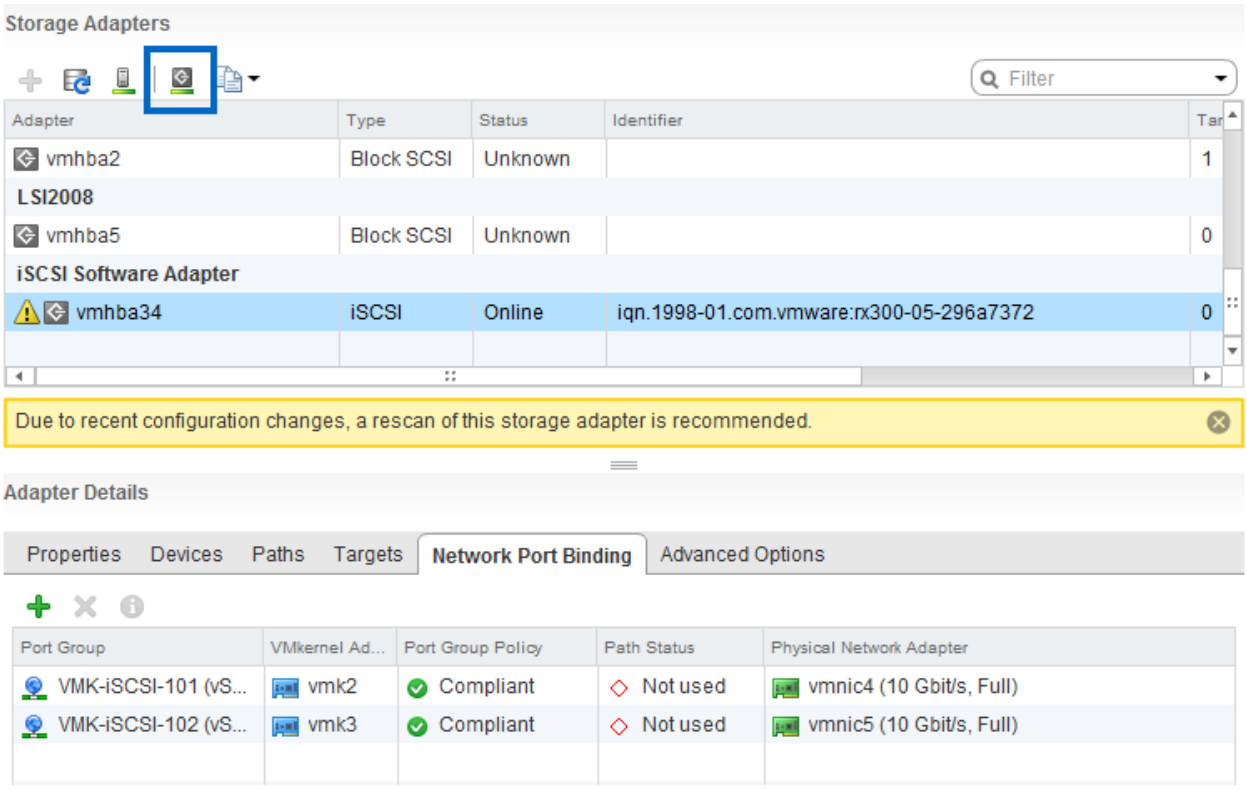
4. Highlight the iSCSI software adapter and click the Network Port Binding tab. Click the green plus icon.
5. If binding the VMkernel adapters to vmnics was successful, the iSCSI VMkernel adapters are listed with their single physical network adapter. If the ports are not available, review step 2 and its substeps.



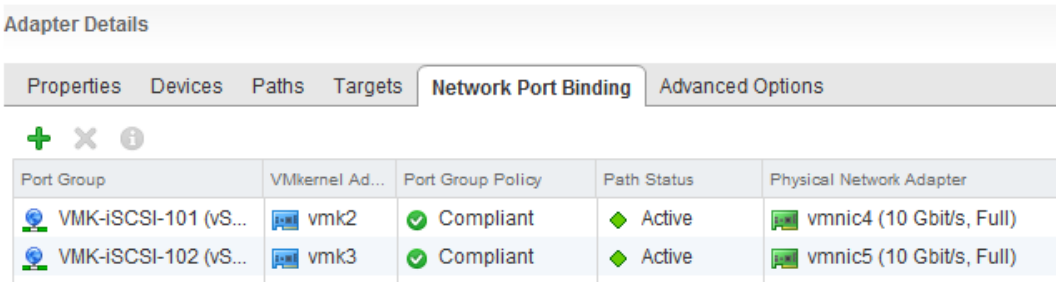
6. Select the VMkernel adapter to add it to the binding configuration and click OK.



7. Repeat step 4 through step 6 for each of the VMkernel adapters in the binding configuration.
8. A message prompts you to rescan the adapter for active paths. Click the SCSI symbol to rescan the iSCSI vmhba.



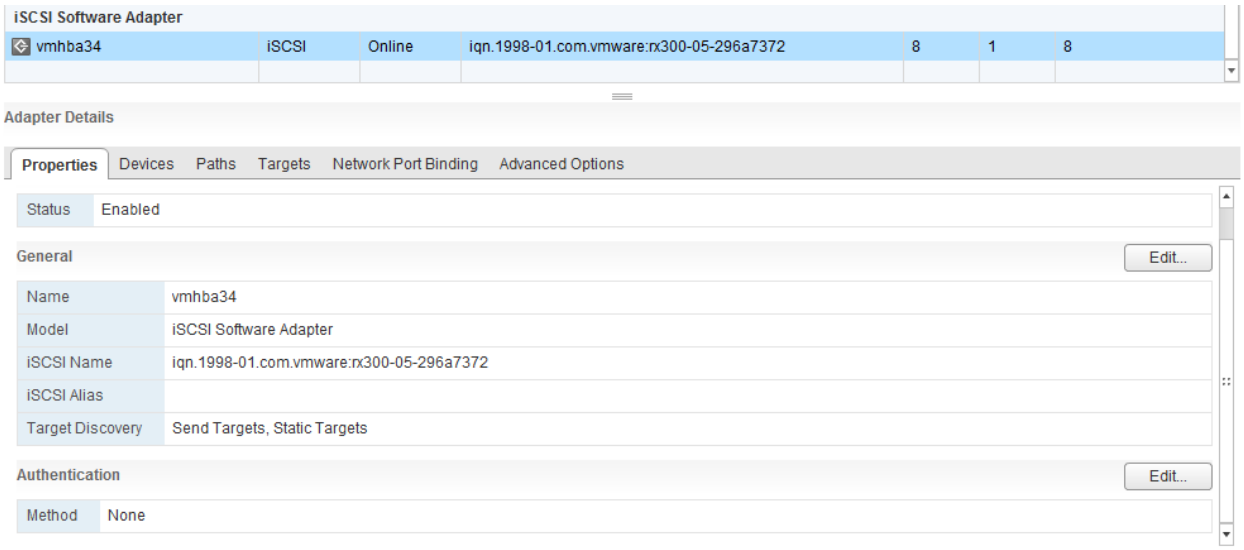
Note: Path status shows “Not used” until a LUN is provisioned using these paths, at which time they show “Active.”



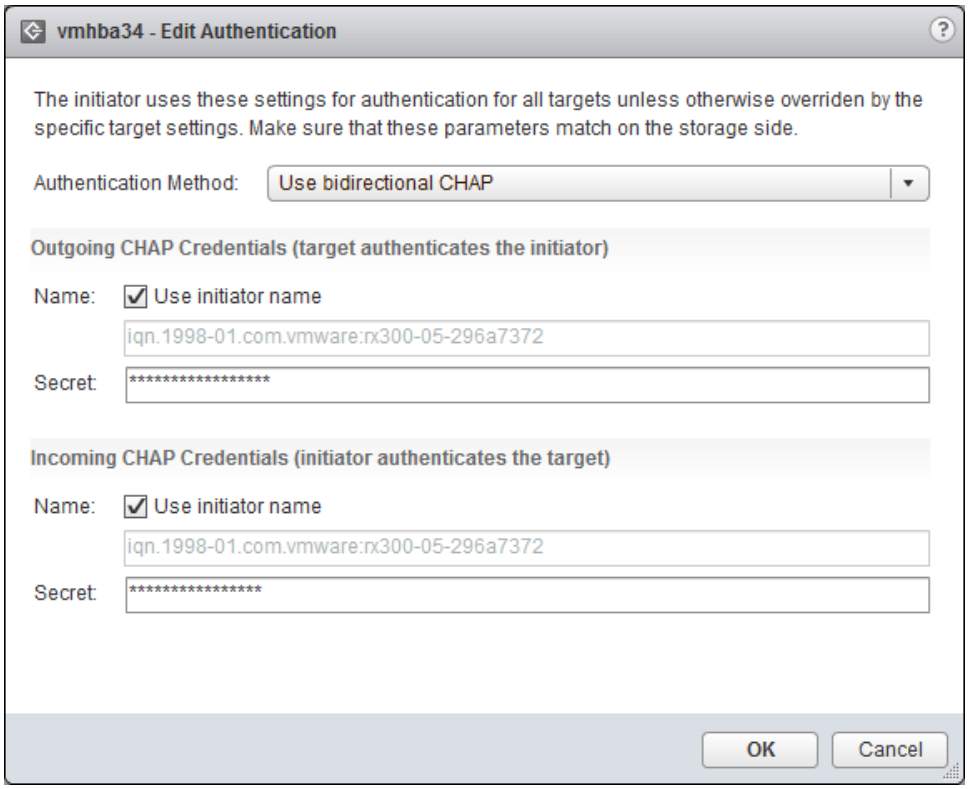
Configure VMware CHAP Authentication Using vSphere Web Client

To configure the VMware CHAP authentication by using vSphere Web Client, complete the following steps:

1. Using a web browser, log in to the vSphere Web Client.
2. Navigate to an ESXi server.
3. Click the Manage tab > Storage > Storage Adapters.
4. Highlight the iSCSI Software Adapter and click the Properties tab.
5. If necessary, scroll down until you see Authentication and any information immediately under it.



6. Next to Authentication, click Edit.
7. In the Edit Authentication dialog box, configure the outgoing and/or incoming credentials according to your configuration and click OK.



8. At the prompt to rescan the adapter for active paths, click the SCSI symbol to rescan the iSCSI vmhba.

Configure NetApp CHAP Authentication Using Data ONTAP CLI

To configure the NetApp CHAP authentication by using the Data ONTAP CLI, complete the following steps:

1. Connect to the NetApp controller by using the console or a remote access protocol such as SSH.
2. For the IQN of each ESXi host connected to the NetApp system, run the following command.

   Note: After you press Enter, the system prompts for a password.

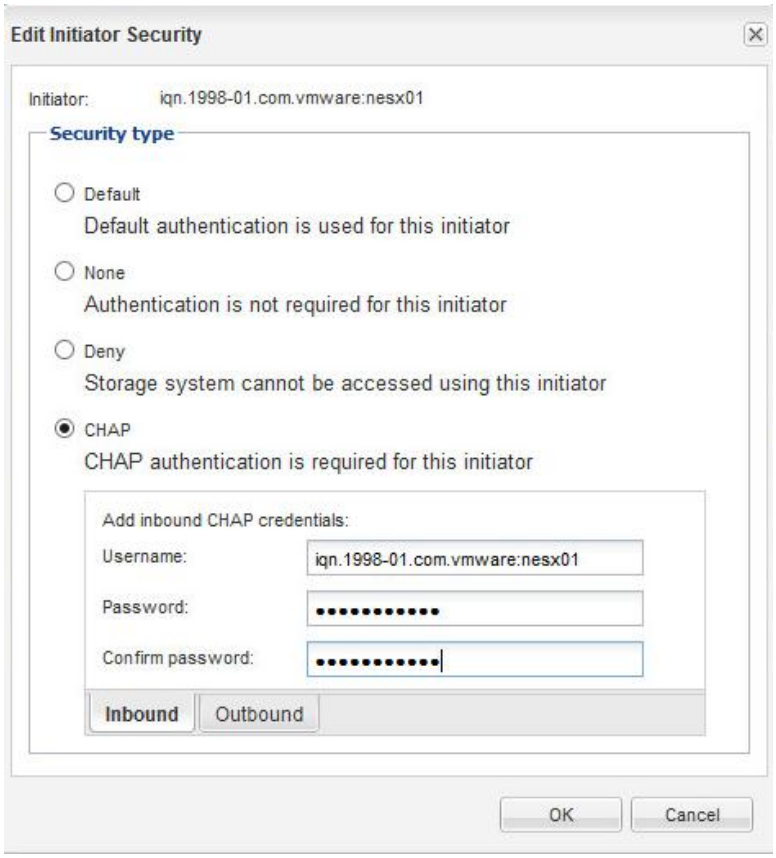
```
cluster01::> vserver iscsi security create -vserver <vserver_name> -initiator-name <esxi_iqn> -
auth-type CHAP -user-name <esxi_iqn>
```

3. Complete the initiator CHAP configuration by repeating step 2 for all ESXi host IQNs.

Configure NetApp CHAP Authentication Using OnCommand System Manager

To configure NetApp CHAP authentication by using OnCommand System Manager, complete the following steps:

1. Connect to the NetApp controller by using OnCommand System Manager.
2. Browse to the storage controller and select Configuration > Protocols > iSCSI. In the right pane, click the Initiator Security tab.
3. Select the IQN of the host on which to enable CHAP authentication and click Edit.
4. In the Edit Initiator Security dialog box, select CHAP and provide the CHAP credentials.



5. Complete the configuration of CHAP authentication of the initiators by repeating step 3 and step 4 for each of the IQNs to be secured with CHAP.

Create and Map LUNs

LUNs can be created and mapped by using a variety of tools, including the CLI, OnCommand System Manager, and others. The administrator must examine the HBA WWPNs of each server to include them in the igroup to which the LUN is mapped. These tools manage the LUN on the storage, but after the LUN is presented to vSphere, additional steps must be completed in ESXi or vCenter to rescan for the LUN and then to partition and format it. All of these steps are handled consistently and reliably in a single wizard and workflow through VSC, as described in section 9.7.

8 Advanced Storage Technologies

8.1 VMware vSphere 6.x and Data ONTAP Cloning

Cloning is a method of making an apparent or actual copy of an object. The object can be a VM, a virtual disk, a file, a database, or a datastore or volume.

The value of cloning is that only the initial object must be created from scratch. In the case of a VM, several tasks must be performed to create a completely functional VM:

1. Create the VM and create or assign virtual hardware.
2. Install the guest operating system.
3. Install patches.
4. Install an application.
5. Configure the guest, including the host name, network identity, join domains, and so forth.

With VM cloning, many of these steps can be skipped because they are completed in the base VM or template. The configuration or customization steps can usually be automated for supported guests by using a customization specification selected or defined as part of the cloning wizard.

Types of Cloning

Three types of cloning technologies are available with vSphere on NetApp storage:

- Full copy cloning
- Delta file cloning
- Pointer-based cloning

The most basic cloning technology offered by vSphere is full copy cloning. The virtual disks of the source VM are copied verbatim by the ESXi server. Each clone occupies the same space as the original objects. Creating the clone consumes ESXi and storage processor cycles as well as disk and network bandwidth.

Some VMware add-on products (such as View Composer, vCloud Director, and the former Lab Manager product) take a read-only template or master VM and create a delta disk file that contains any data written by the clone VM. The delta file descriptor references the template virtual disk and is said to be linked to the template. Many linked clones can reference a common template. Linked clones are easy to implement and are tightly integrated with the products that support them. Linked clones generate additional I/O for each VM I/O in the form of metadata operations to determine whether the data in question is in the delta (all writes and reads of changed data) or in the template (reads of unchanged data).

In Data ONTAP, cloning is enabled with the FlexClone license, although there is also an older implementation of LUN cloning that is based on Snapshot technology, and that does not require a FlexClone license. The FlexClone clone appears to the application as a copy of a file, a LUN, a block range within a file or LUN, or a FlexVol volume. The clone, however, is actually a new inode (file header) in the case of a file or LUN clone, or a new volume structure in the case of a FlexVol clone, along with a

set of pointers back to the original data blocks. New data is written to the clone in new data blocks. Clones are very space efficient because the only space occupied is reserved for metadata and any changes written.

Clones perform the same as the original object because no additional metadata or lookups are added beyond the normal metadata of a similar noncloned object. Furthermore, cloned objects can benefit from shared read cache (system cache or Flash Cache cards) on the storage controller because after a common block is read and cached, other reads, even for a different object, can read the block from cache.

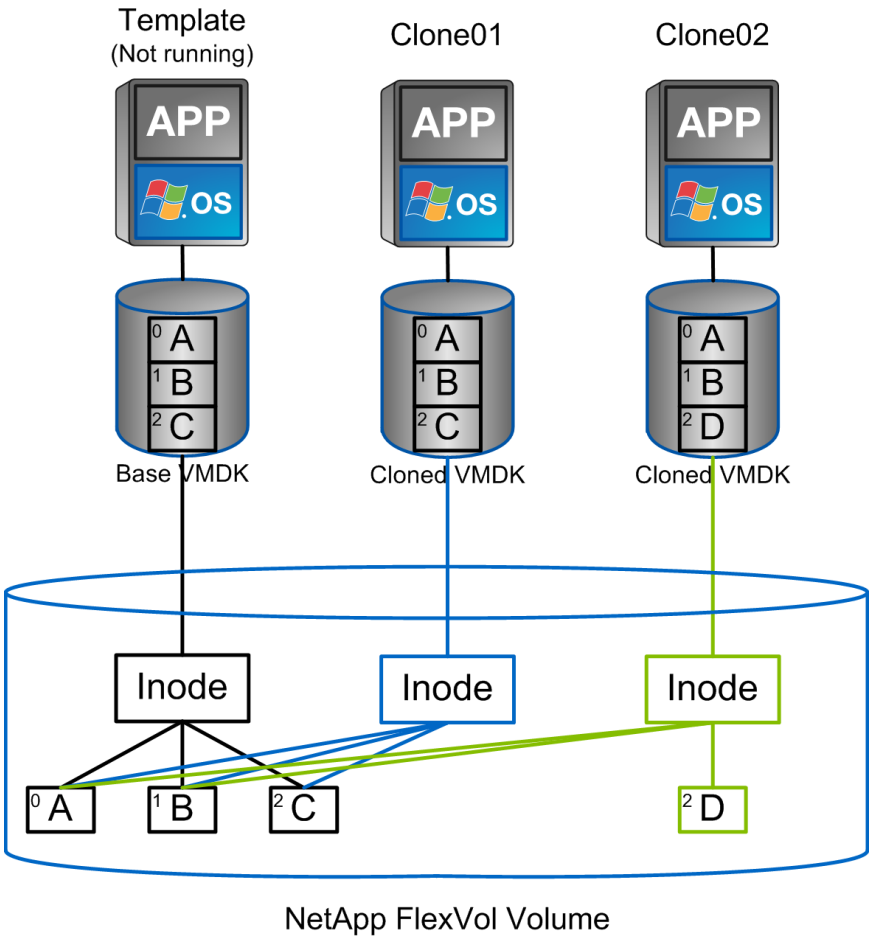
Levels of Cloning

vSphere on NetApp makes use of cloning at two levels:

- VM cloning
- Datastore cloning

VM cloning can be invoked from vCenter through the vSphere Web Client, through an API such as PowerCLI, from the ESXi command line by using `vmkfstools`, or through a plug-in that uses the NetApp VSC. Figure 32 shows a VM in an NFS datastore and a VMDK that is cloned twice. Clone01 has never been started or guest customized, so the pointers reference blocks from the original file. Clone02 has been started or customized, so some blocks (block 2 in Figure 32) have changed, and the file now has its own unique blocks.

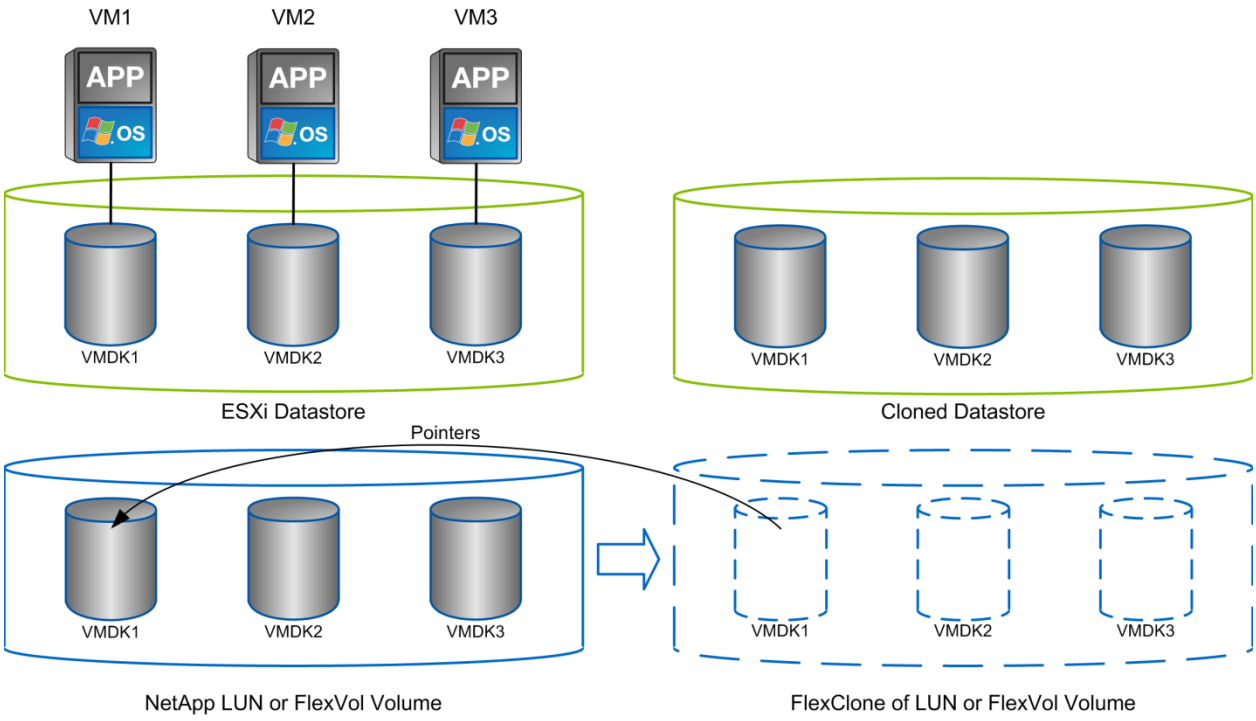
Figure 32) Single-file FlexClone cloning of VMDKs.



Datastore cloning is a function of NetApp storage. vSphere has no way of cloning an entire datastore, but there are steps to complete in vSphere to access a datastore cloned by the NetApp controller. With NFS datastores, FlexClone technology can clone an entire volume, and the clone can be exported from Data ONTAP and mounted by ESXi as another datastore. For VMFS datastores, Data ONTAP can clone a LUN within a volume or a whole volume, including one or more LUNs within it. A LUN containing a VMFS must be mapped to an ESXi initiator group (igroup) and then resignatured by ESXi in order to be mounted and used as a regular datastore. For some temporary use cases, a cloned VMFS can be mounted without resignaturing. After a datastore is cloned, VMs inside it can be registered, reconfigured, and customized as if they were individually cloned VMs.

Figure 33 shows how a FlexClone datastore appears to ESXi prior to reconfiguring and registering the cloned VMs. The datastore looks like a complete copy, but on the storage, the blocks of the clone are initially just pointers to the source volume. The VMDKs and other files of VMs in the clone look identical to those of the source volume until they are reconfigured, registered, and customized.

Figure 33) ESXi view of FlexClone cloning.



This entire sequence of cloning, exporting and mapping, resignaturing and mounting, and preparing the VMs is managed as a single workflow in the provisioning and cloning function of VSC.

Several tools and products from VMware and NetApp invoke cloning in one or more methods. Table 31 lists the cloning methods, tools, and products that use cloning and describes how cloning is invoked.

Table 31) Cloning methods, products, tools, and use cases.

Cloning Method	Method	Invoked By	Clones per Invocation	Primary Use
vSphere (ESXi) full copy	Makes a full copy of the VM	• vSphere Web Client • vSphere APIs • `vmkfstools`	1	vCenter general-purpose cloning; also used by VSC to clone between VMFS datastores

Cloning Method	Method	Invoked By	Clones per Invocation	Primary Use
Linked clone	Creates delta file for each clone linked to common read-only template	• View Composer • vCloud Director	Many	View Composer, vCloud Director
vSphere full copy with VMware vSphere vStorage APIs for Array Integration (VAAI)	Uses APIs to offload a clone to storage; NetApp storage uses pointer-based clones within the same volume or full copy between volumes	• vSphere Web Client • vSphere APIs • `vmkfstools`	1*	vCenter; also vCloud Director with fast provisioning off
View Composer Array Integration (VCAI)	Offloads clone to vCenter, which calls VAAI	View Composer	Many	View Composer
VSC	• Full copy between volumes • Single-file FlexClone cloning (sis clone) within NFS datastores • Full copy within a VMFS datastore that uses VAAI and sub-LUN cloning, if available • Volume FlexClone cloning to clone large numbers of VMs in a single operation • Single-file and volume FlexClone cloning for two-dimensional cloning	• Context-sensitive menu option and workflow provided by the VSC plug-in • VSC APIs and PowerShell cmdlets	Many	General-purpose cloning; virtual desktop

* vCloud Director can make many clones as part of a vApp deployment, but it calls the vSphere cloning API for each one.

When selecting a cloning method, consider how complete the workflow is and how well it integrates with the use case for the VMS. NetApp recommends using VSC or a solution that uses VAAI to offload the cloning process. Either of these methods makes efficient clones and integrates with products, such as VMware View, by registering the VMs not only in vCenter but also in the virtual desktop connection broker.

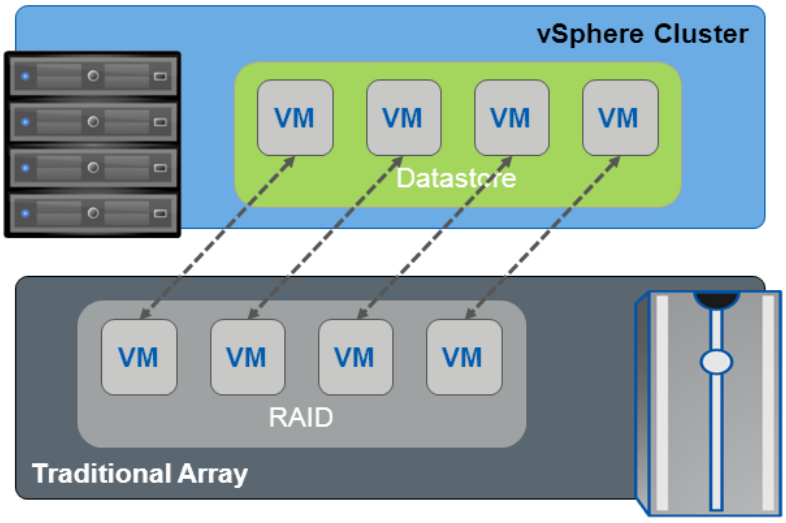
Best Practice
Use VSC or a solution that uses VAAI to offload the cloning process.

8.2 Storage Deduplication

One of the most popular VMware features is the ability to rapidly deploy VMs from stored VM templates. A VM template includes a VM configuration file (`.vmx`) and one or more virtual disk files (`.vmdk`). Virtual disk files include an operating system, common applications, and patch files or system updates.

Deploying VMs from templates saves administrative time because the configuration and the virtual disk files are copied, and this copy is registered as an independent VM. By design, this process introduces duplicate data for each new VM that is deployed. Figure 34 shows an example of typical storage consumption in a vSphere deployment.

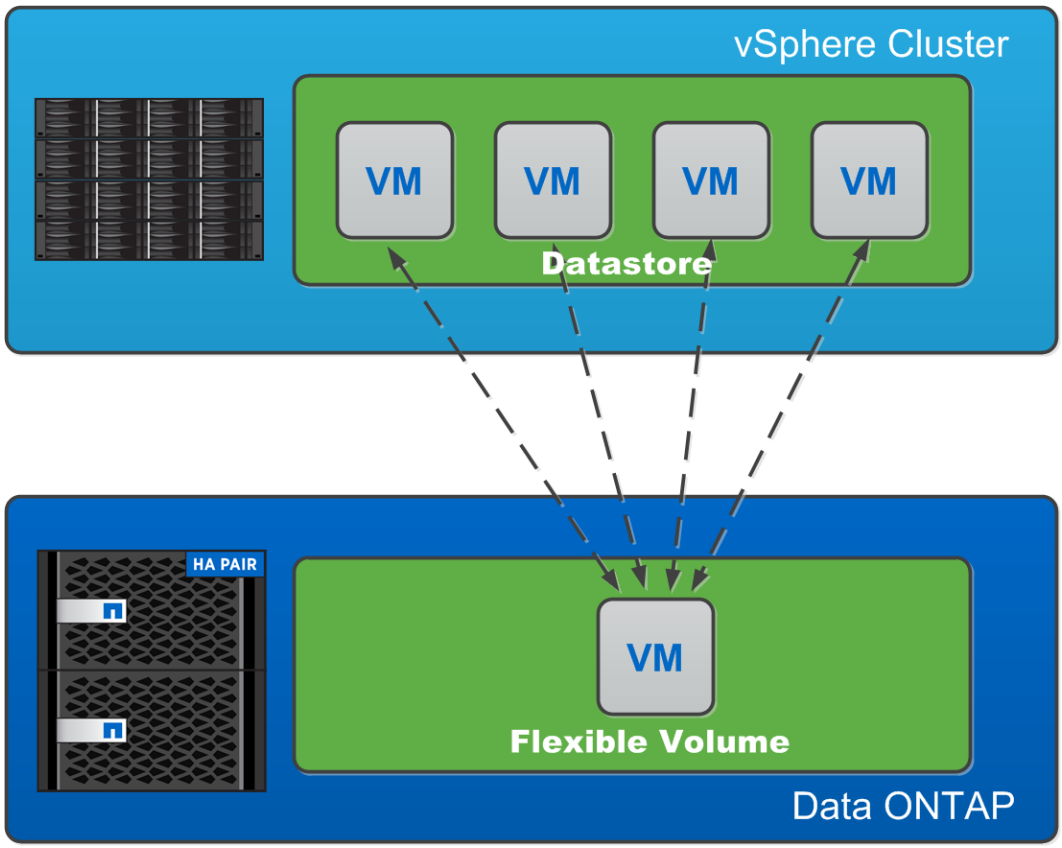
Figure 34) Storage consumption with a traditional array.



NetApp offers a data deduplication technology called FAS data deduplication. With NetApp FAS deduplication, VMware deployments can eliminate the duplicate data in their environment, enabling greater storage use.

Deduplication virtualization technology enables multiple VMs to share the same physical blocks in a NetApp FAS system in the same manner that VMs share system memory. It can be seamlessly introduced into a virtual data center without the need to make any changes to VMware administration, practices, or tasks. Deduplication runs on the NetApp FAS system at scheduled intervals and does not consume any CPU cycles on the ESXi server. Figure 35 shows an example of the impact of deduplication on storage consumption in a vSphere deployment.

Figure 35) Storage consumption after enabling FAS data deduplication.



Deduplication is enabled on a volume, and the amount of data deduplication realized is based on the commonality of the data stored in a deduplication-enabled volume. For the largest storage savings, NetApp recommends grouping similar operating systems and similar applications into datastores, which ultimately reside on a deduplication-enabled volume.

Deduplication Considerations with VMFS and RDM LUNs

Enabling deduplication when provisioning LUNs produces storage savings. However, the default behavior of a LUN is to reserve an amount of storage space that is equal to the space taken by the provisioned LUN. This design means that although the storage array reduces the amount of capacity consumed, any gains made with deduplication are for the most part unrecognizable, because the space reserved for LUNs is not reduced.

To recognize the storage savings of deduplication with LUNs, NetApp LUN thin provisioning must be enabled.

Note: Although deduplication reduces the amount of consumed storage, the VMware administrative team does not see this benefit directly, because its view of the storage is at a LUN layer, and LUNs always represent their provisioned capacity, whether they are thin provisioned or provisioned in the traditional way. NetApp VSC provides the vSphere administrator with storage use at all layers in the storage stack.

When deduplication is enabled on thin-provisioned LUNs, NetApp recommends deploying these LUNs in FlexVol volumes that are also thin provisioned with a capacity that is two times the size of the LUN. When the LUN is deployed in this manner, the FlexVol volume acts merely as a quota. The storage consumed by the LUN is reported in the FlexVol volume and in its containing aggregate.

Deduplication Advantages with NFS

Unlike what happens with LUNs, when deduplication is enabled with NFS, the storage savings are immediately available and recognized by the VMware administrative team. The benefit of deduplication is transparent to storage and VI administrative teams. Special considerations are not required for its use.

Deduplication Management with VMware

Through the NetApp vCenter plug-ins, VMware administrators have the ability to enable, disable, and update data deduplication on a datastore-by-datastore basis.

8.3 VMware vSphere 5.x Thin Provisioning

In VMware vSphere environments, thin provisioning can be implemented in both the specific VMDK and its underlying LUN or volume.

Note: NetApp recommends using thin provisioning in both the VMDK and the LUN or volume, depending on the requirements and the environment.

Thin Virtual Disk

The thin virtual disk is very similar to the thick virtual disk, except that it does not preallocate the capacity of the virtual disk from the datastore when the virtual disk is created. When storage capacity is required, the VMDK allocates storage in chunks that are equal to the size of the file system block. The VMFS block size is fixed at 1MB in VMFS-5. For VMFS-3, the block size ranges from 1MB to 8MB (the size is selected when the datastore is created and deployed); for NFS, the size is 4KB. The process of allocating blocks on a shared VMFS datastore is considered a metadata operation and as such initiates SCSI locks on the datastore while the allocation operation is being executed. Although this process is very brief, it does suspend the write operations of the VMs on the datastore.

The VAAI allows a vSphere administrator to see whether a particular LUN is thin provisioned and to view any other related metadata. The `MODE SENSE` and `MODE SELECT` commands, respectively, are used to get and set thin provisioning–related parameters.

Note: For information about VAAI primitives, refer to the section on storage hardware acceleration in the ESXi Configuration Guide from VMware.

VMware vSphere 5.0, and later, includes enhanced support for the VAAI primitives and introduces new primitives, such as thin provisioning, that enable the reclamation of unused space and the monitoring of space usage.

The VMware vSphere 4.1 release introduced a VAAI primitive known as hardware-assisted locking. This primitive provides a more granular means to protect the VMFS metadata than the SCSI reservations that were used before hardware-assisted locking was available. Hardware-assisted locking leverages a storage array atomic test and set capability to enable a fine-grained, block-level locking mechanism. Simple tasks, such as moving a VM, starting a VM, creating a new VM from a template, creating native VMware snapshots, or even stopping a VM, cause the VMFS to allocate or return storage to or from the shared free-space pool. Although VMFS use of the SCSI reservation locking the LUN does not often result in performance degradation, the use of hardware-assisted locking provides a much more efficient means to avoid retries for getting a lock when many ESX servers share a single datastore.

Hardware-assisted locking enables the offloading of the locking mechanism to the arrays and does so with much less granularity than an entire LUN. Therefore, the VMware cluster can leverage a significant scalability gain without compromising the integrity of the VMFS shared storage-pool metadata.

Both thin and thick VMDKs are not formatted when they are deployed. Therefore, data that must be written must pause while the blocks required to store this data are zeroed out. The process of allocating blocks from within the datastore occurs on demand whenever a write operation attempts to store data in a block range inside a VMDK that has not been written to by a previous operation.

In summary, both thick and thin virtual disks suspend I/O when writing to new disk areas that must be zeroed out. However, before this operation can occur with a thin virtual disk, the thin disk might have to obtain additional capacity from the datastore. This is the main difference between thick and thin virtual disks regarding zeroing out and storage allocation.

Storage Array Thin Provisioning with VMware vSphere

Server administrators often overprovision storage to avoid running out of storage space and to prevent the associated application downtime for expanding the provisioned storage. Although no system can run at 100% storage use, methods of storage virtualization allow administrators to address and oversubscribe storage in the same manner as with server resources (such as CPU, memory, networking, and so on). This form of storage virtualization is referred to as thin provisioning.

Traditional provisioning preallocates storage; thin provisioning provides storage on demand. The value of thin-provisioned storage is that storage is treated as a shared resource pool and is consumed only as required by each individual VM. This sharing increases the total usage rate of storage by eliminating the unused but provisioned areas of storage that are associated with traditional storage.

The drawback to thin provisioning and oversubscribing storage is that, if no physical storage is added, and if every VM requires its maximum possible storage at the same time, then not enough storage space is available to satisfy all of the requests.

NetApp Thin-Provisioning Options

NetApp thin provisioning extends VMware thin provisioning for VMDKs. It also allows LUNs that are serving VMFS datastores to be provisioned to their total capacity while consuming only as much storage as is required to store the VMDK files (which can be in either a thick or a thin format). In addition, LUNs connected as raw device mappings (RDMs) can be thin provisioned.

In a clustered Data ONTAP environment, thin provisioning on Data ONTAP provides 70% or more utilization; up to 50% storage savings; and up to 50% savings in power, cooling, and space.

When NetApp thin-provisioned LUNs are enabled, NetApp recommends deploying these LUNs in volumes that are also thin provisioned with the space guarantee set to `none` and with an overall capacity that uses the formula `1x + Delta`. The delta in the formula is a sizing consideration for including Snapshot reserve space, over and above the actual size of the contained LUN or LUNs themselves. If a LUN is deployed in this manner, the volume acts merely as a quota. The storage consumed by the LUN is reported in the FlexVol volume and in its containing aggregate.

Thin Provisioning and VAAI Enhancements

Thin provisioning improves storage efficiency, but it also increases management overhead because the management tools might not report the correct storage capacity. This is evident in scenarios in which VMs are migrated or deleted from the datastore and the corresponding block is not freed by the storage array. In such cases, if a thin-provisioned datastore reaches the out-of-space condition, then all of the VMs in that datastore are affected.

To address these concerns, the following VAAI primitives were introduced in VMware vSphere 5.0.

Thin Provisioning Stun

When a thin-provisioned datastore reaches its maximum capacity and does not have any free blocks, Thin Provisioning Stun pauses the VMs that require additional blocks; VMs that do not require additional blocks continue to run. This strategy prevents the VMs from crashing and prevents data corruption in the VMs. When additional space is added to the thin-provisioned LUN and there are free blocks, the paused VMs can be resumed.

Dead Space Reclamation Using UNMAP

With the UNMAP primitive, the storage array can be notified about blocks that have been freed after VMs have been deleted or migrated to another datastore. In VMware vSphere 5.0, UNMAP was run by the `vmkfstools –y` command, which specified the percentage of blocks that should be freed. In vSphere 5.5 and later, the `vmkfstools –y` command is replaced by the `esxcli storage vmfs unmap` command, which specifies the numbers of free blocks.

Best Practice
For more about the `esxcli storage vmfs unmap` command, refer to VMware KB 2057513.

8.4 VMware vSphere 6.x and Data ONTAP QoS

Clustered NetApp Data ONTAP 8.2 and later implement storage quality of service (QoS) to provide the ability to limit the storage throughput and/or input/output operations per second (IOPS) available to a workload. QoS is applied through administrator-defined policies that limit the storage performance of a workload or set of workloads. A workload is one or more of the following objects:

- SVM
- Volume
- LUN
- File

In the vSphere context, a volume is typically used either to contain a LUN or as an NFS datastore. A LUN can contain a VMFS datastore or an RDM used by a single VM or a pair of VMs in a guest cluster. A file is the `vm-flat.vmdk` file that contains the guest virtual disk image on an NFS datastore. QoS policies cannot be applied to individual files in a VMFS datastore because the VMFS file system is contained in the LUN, which is simply a large binary object, and Data ONTAP has no knowledge of the VMFS structure.

The QoS enforced performance limit on an object can be defined by using the following units:

- Megabytes per second (MBps)
- Input/output operations per second (IOPS)

A policy can be applied to one or more workloads. When applied to multiple workloads, the workloads share the total limit of the policy. In other words, the total storage performance of all of the workloads that share a policy does not exceed the policy. If several workloads should each have a specific performance limit, they should each have their own unique policy.

In clustered Data ONTAP 8.2, QoS policies cannot be applied to nested objects. For example, a QoS policy cannot be applied to a file if the containing volume or SVM has a policy applied. Conversely, if a file has a QoS policy, a policy cannot be applied to the parent volume or SVM. However, multiple files, volumes, and LUNs in the same SVM can each have different QoS policies, so long as they are not nested.

NetApp QoS and VMware SIOC and NetIOC

NetApp QoS, VMware vSphere Storage I/O Control (SIOC), and Network I/O Control (NetIOC) are complementary technologies that vSphere and storage administrators can use together to manage performance of vSphere VMs hosted on clustered Data ONTAP storage. Each tool has its own strengths, as shown in Table 32. Because of the different scope of VMware vCenter and the NetApp cluster, some objects can be seen and managed by one system and not the other.

Table 32) NetApp QoS and VMware SIOC comparison.

Property	NetApp QoS	VMware SIOC
When active	Policy is always active	Active when contention exists (datastore latency over threshold)
Type of units	IOPS, MBps	IOPS, shares
vCenter or application scope	Multiple vCenter environments, other hypervisors and applications	Single vCenter server
VM granularity	VMDK on NFS only	VMDK on NFS or VMFS
LUN (RDM) granularity	Yes	Yes
LUN (VMFS) granularity	Yes	No
Volume (NFS datastore) granularity	Yes	No
SVM (tenant) granularity	Yes	No
Policy	Yes; throughput is shared by all workloads in the policy.	No; shares and limits are set on each VM's virtual disk.
License required	Included with clustered Data ONTAP	Enterprise Plus
Management tools		
vSphere Client	No	Yes
vSphere Web Client	No	Yes
PowerShell	Yes	Yes
NetApp Workflow Automation (WFA)	Yes	No

Tools for Managing QoS Policies and Workloads

The following tools are currently available for managing QoS policies and applying them to objects:

- NetApp clustershell CLI
- NetApp PowerShell Toolkit
- NetApp OnCommand WFA

Guidelines for Setting QoS Policies

To set a policy on an object, the following general outline should be followed:

1. Determine what the object's performance requirement should be, considering its effect on other objects.
2. Create a new policy or select an existing policy if the object should share a policy.
3. Determine the actual location and/or path to the object.
4. Apply the policy to the object.

Some of these steps might seem obvious, but there are some subtleties, depending on the object and tools being used.

For VMDKs on NFS, note the following guidelines:

- The policy must be applied to the `vmname-flat.vmdk` that contains the actual virtual disk image, not the `vmname.vmdk` (virtual disk descriptor file) or `vmname.vmx` (VM descriptor file).
- Do not apply policies to other VM files such as virtual swap files (`vmname.vswp`).
- When using the vSphere Client Datastore Browser or the Manage > Files view in the vSphere Web Client to find file paths, be aware that it combines the information of the `-flat.vmdk` and `.vmdk` and simply shows one file with the name of the `.vmdk` but the size of the `-flat.vmdk`. Simply insert `-flat` into the file name to get the correct path.

For LUNs, including VMFS and RDM, the NetApp SVM (displayed as Vserver), LUN path, and serial number are readily available using the Monitoring and Host Configuration feature of the NetApp VSC (under Storage Details – SAN).

References

For more information, refer to the following:

- Clustered Data ONTAP 8.2 Quality of Service, Performance Isolation for Multi-Tenant Environments
- OnCommand Workflow Automation (communities and downloads)
- Data ONTAP PowerShell Toolkit

8.5 Using Data ONTAP QoS with VMware vSphere 6.x

Table 33 provides vSphere 6 and Data ONTAP QoS use cases.

Table 33) VMware vSphere 6.x and Data ONTAP QoS use cases.

Use Case	Procedure Name
Create a QoS policy.	Create QoS policy
Set a QoS policy on a virtual machine disk (VMDK) in NFS.	Set QoS policy on VMDK in NFS
Set a QoS policy on a LUN used as an RDM.	Set QoS policy on RDM LUN
Set a QoS policy on a LUN used as a VMFS datastore.	Set QoS policy on VMFS LUN
Set a QoS policy on a FlexVol volume containing LUNs and/or used as an NFS datastore.	Set QoS policy on FlexVol volume
Set a QoS policy on an SVM.	Set QoS policy on SVM

Create QoS Policy

To create a QoS policy, complete the following steps:

1. Determine the appropriate throughput MBps or IOPS limit for the workload.
2. Log in to the cluster CLI.
3. Run the `qos policy-group create` command to create the policy.

```
qos policy-group create -policy-group <<policy_group>> -vserver <<vserver>> -max-throughput
<<throughput>>
```

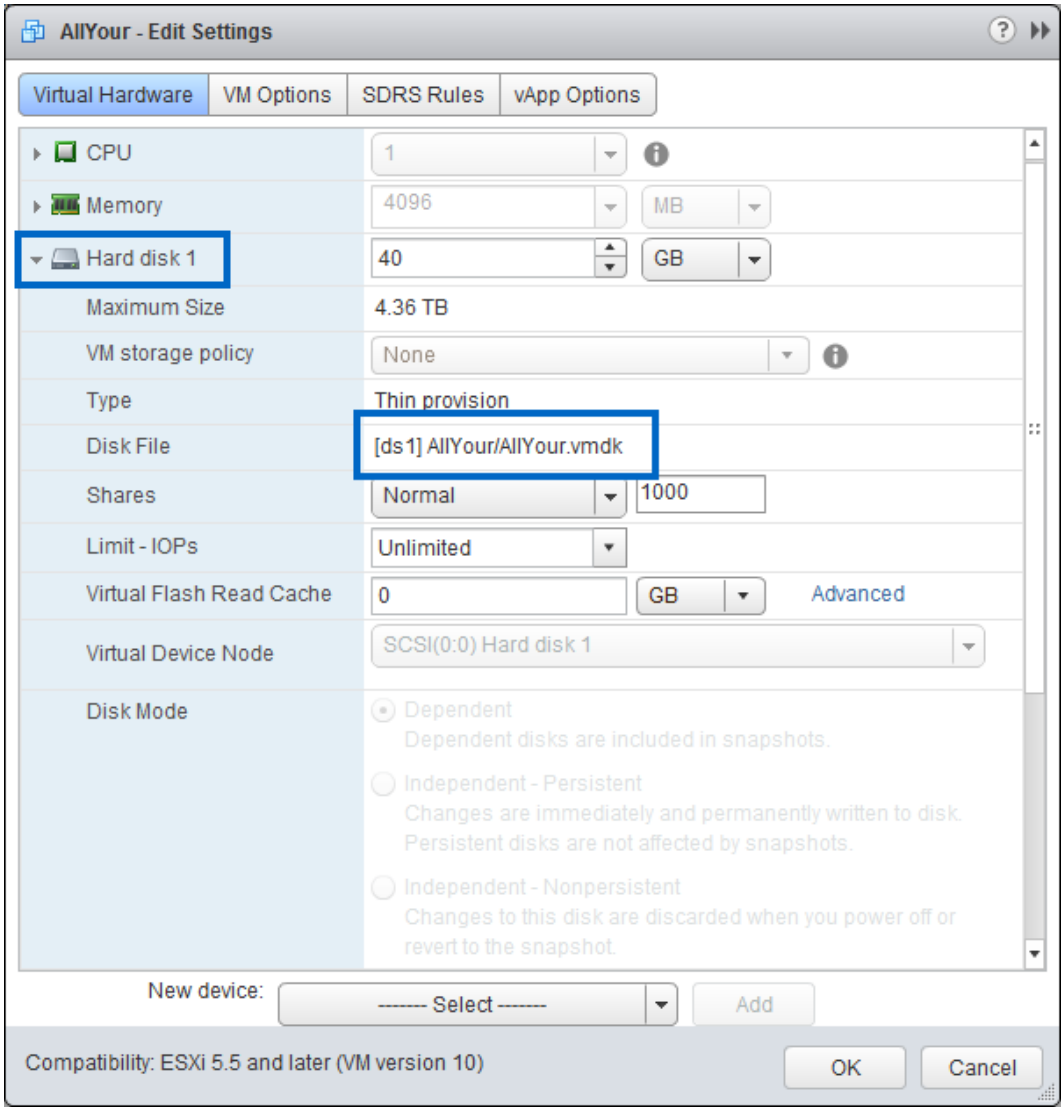
4. Verify that the new policy was created as intended by running the `qos policy-group show` command.

```
qos policy-group show -policy-group <<policy_group>>
```

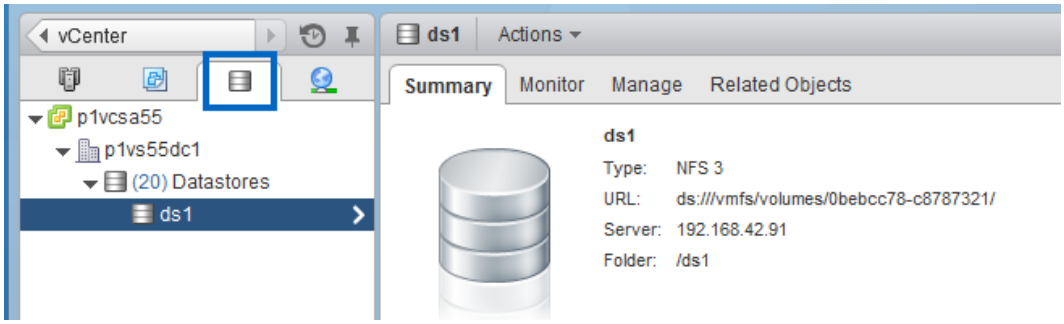
Set QoS Policy on VMDK in NFS

To set a QoS policy on a VMDK stored on an NFS datastore, complete the following steps:

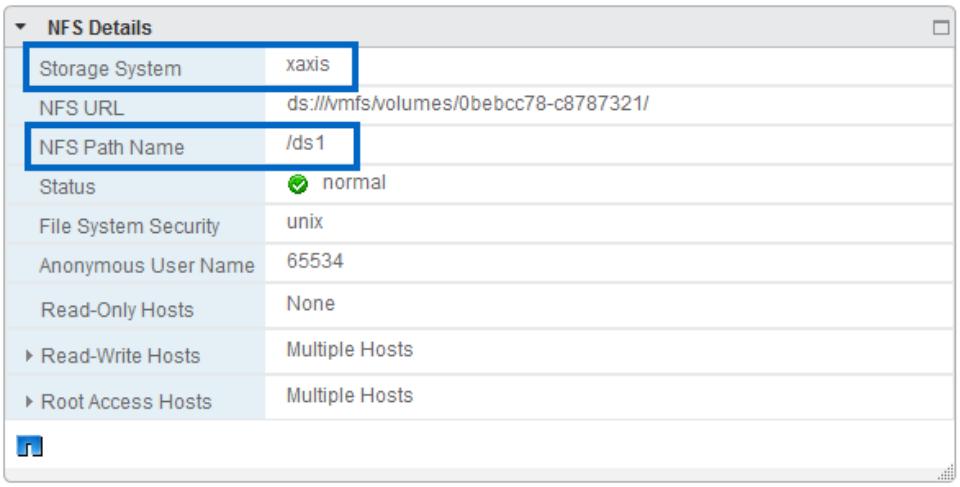
1. Determine whether the VM should share a policy with other VMs, have a single policy for all of its virtual disks, or have a specific policy for each virtual disk.
2. For each virtual disk in the VM, use the vSphere Web Client to determine the path to the virtual disk.
3. In the vSphere Web Client, right-click the VM and select Edit Settings.
4. Expand the virtual disk and note the disk file.



5. Click Cancel.
6. Navigate to the Summary tab for the datastore by clicking the storage icon and expanding the tree until you get to the datastore on which the virtual disk is stored.



7. In the NFS Details window, note the storage system (which corresponds to the SVM) and NFS path name (the mounted junction path of the volume, which either matches the volume name or can be used with the `vol show -junction` command to find the volume name).



8. On the NetApp cluster CLI, run the `file modify` command to set the QoS policy on the virtual disk using the parameters noted in step 1 and step 4.

   Note: Remember to insert `-flat` in the file name before the `.vmdk` extension. If the file path contains spaces, it needs to be enclosed in double quotes.

```
file modify -vserver <<vserver>> -volume <<volume>> -file <<vmdk_flat_path>> -qos-policy-group
<<policy_group>>
```

Example:

```
file modify -vserver xaxis -volume ds_ip1 -file RHEL58/RHEL58-flat.vmdk -qos-policy-group vm1
```

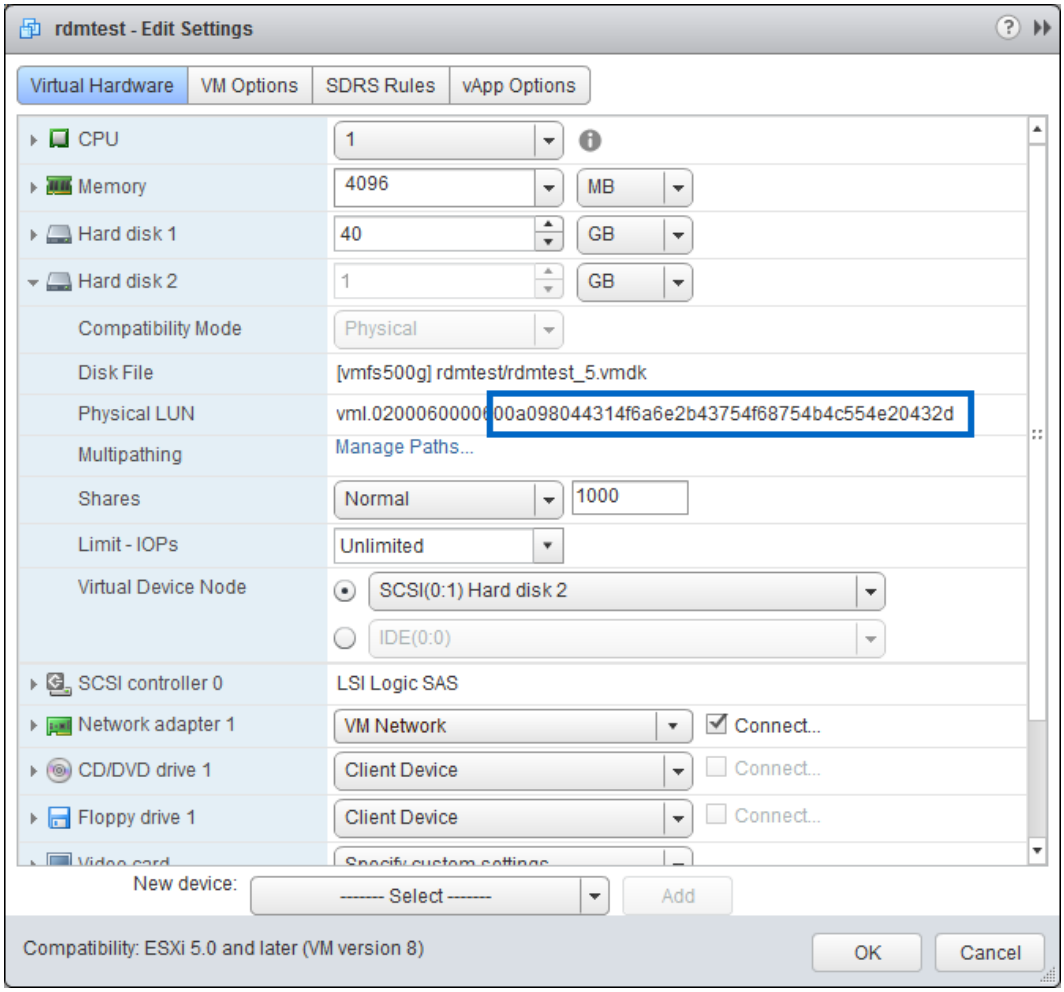
Set QoS Policy on RDM LUN

To set a QoS policy on the RDM LUN, complete the following steps:

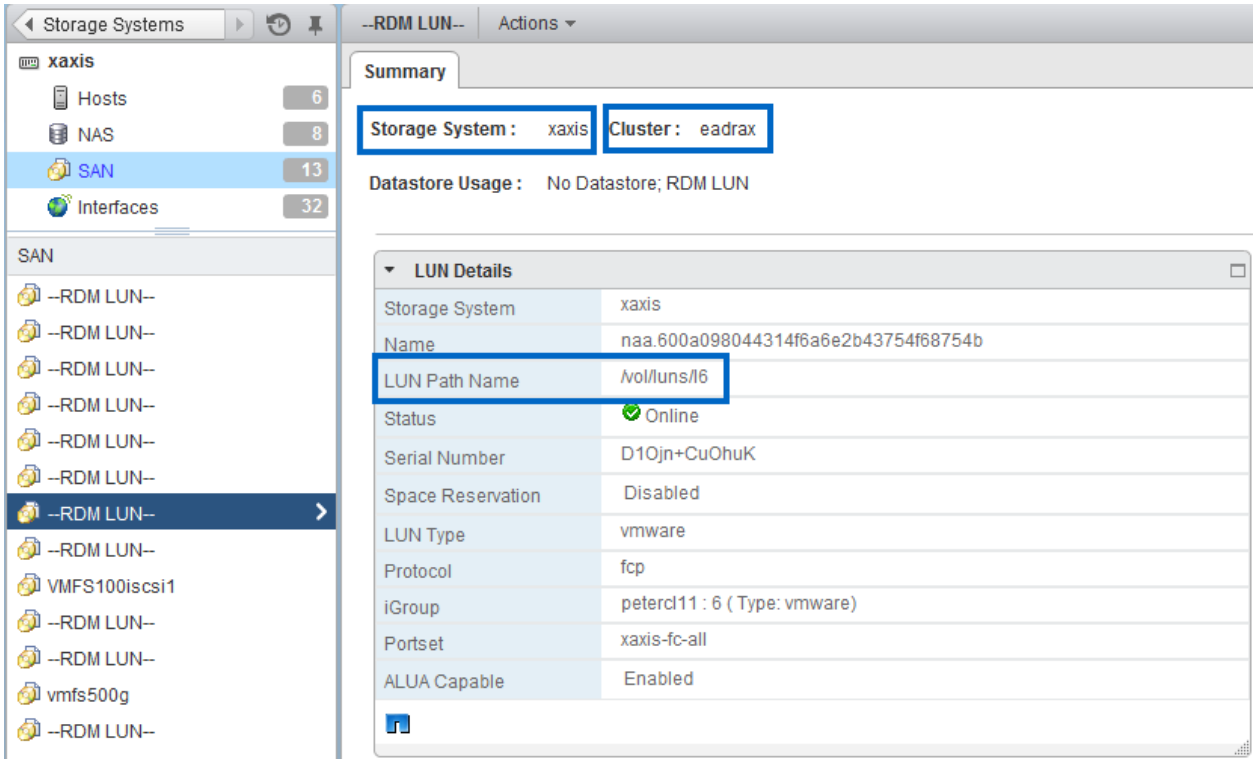
Note: All VMs that share this LUN and have concurrent I/O share the throughput.

1. Determine whether the VM should share a policy with other VMs, have a single policy for all of its virtual disks, or have a specific policy for each virtual disk.
2. For each RDM attached to the VM, use the vSphere Web Client to determine the physical LUN name.
3. In the vSphere Web Client, right-click the VM and select Edit Settings.

4. Expand the virtual disk details and note the physical LUN name, particularly the portion of the LUN name starting with `00a0980` and the following 24 characters. If the LUN name does not contain `00a098` (the NetApp OUI), it is not a NetApp LUN hosted by a Data ONTAP system.



5. From the vSphere Web Client Home screen, click Virtual Storage Console > Storage Systems.
6. Double click the SVM that contains the LUN > SAN.
7. Under Details, find and select the LUN name that matches the LUN name noted in step 4.



8. Note the storage system (which is actually the SVM), cluster, and LUN path name.
9. On the NetApp cluster CLI, run the `lun modify` command to set the QoS policy on the LUN using the parameters noted in step 4 and step 8.

```
lun modify -vserver <<vserver>> -path <<LUN_path>> -qos-policy-group <<policy_group>>
```

Example:

```
lun modify -vserver xaxis -path /vol/fcoe/lun1 -qos-policy-group rdm1
```

The following VMware Knowledge Base articles provide additional information and methods for identifying RDM LUNs:

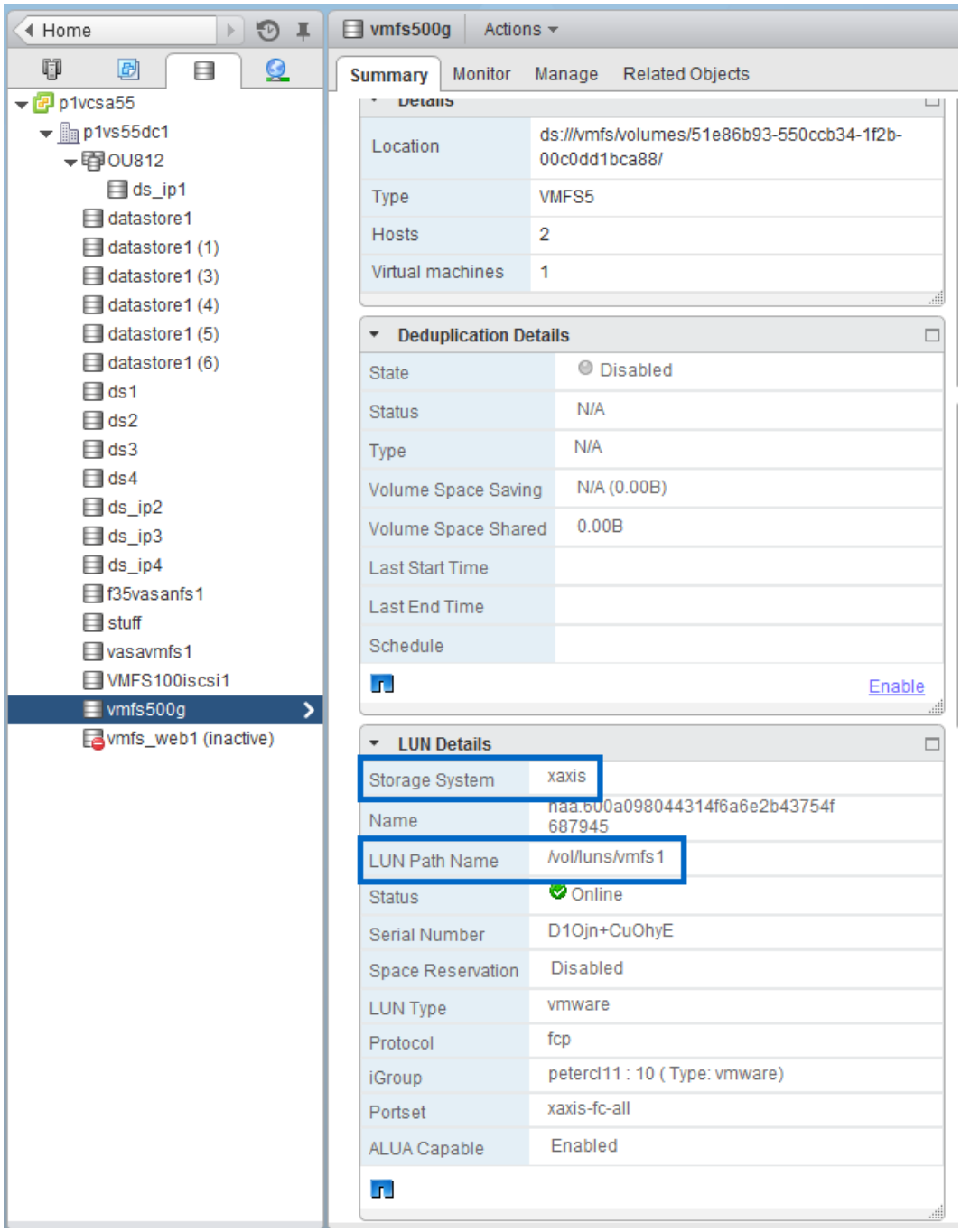
- Identifying Raw Device Mappings (RDM) Using the vSphere Client (1004814)
- Identifying Virtual Disks Pointing to Raw Device Mappings (RDMs) (1005937)
- Identifying Virtual Machines with Raw Device Mappings (RDMs) Using PowerCLI (2001823)

Set QoS Policy on VMFS LUN

To set a QoS policy on a VMFS LUN, complete the following steps:

Note: All VMs that reside in this datastore share the throughput of the LUN. This includes VMkernel swapping to the `.vswp` file.

1. Determine which VMFS datastore must have the policy applied.
2. In the vSphere Web Client, navigate to the VMFS datastore by clicking the storage icon and expanding the hierarchy to get to and select the volume.
3. Scroll down to LUN Details.



4. Note the storage system (which is actually the SVM) and LUN path name.
5. On the NetApp cluster CLI, run the `lun modify` command to set the QoS policy on the LUN using the parameters noted in step 4.

```
lun modify -vserver <<vserver>> -path <<LUN_path>> -qos-policy-group <<policy_group>>
```

Example:

```
lun modify -vserver xaxis -path /vol/luns/vmfs1 -qos-policy-group vmfs1
```

Set QoS Policy on FlexVol Volume

To set a QoS policy on a FlexVol volume, complete the following steps:

Note: All VMs, LUNs, or other files that reside in the FlexVol volume share the throughput of the policy. This includes VM I/O, VMkernel swapping to `.vswp` files, and any other hypervisor or application I/O.

1. Determine which FlexVol volume needs the QoS policy by using VSC Monitoring and Host Configuration and looking at Storage Details – SAN or Storage Details – NAS.
2. If the FlexVol volume is used as an NFS datastore, use the NFS path name shown in VSC as the junction path in the following command to determine the actual volume name:

```
volume show -vserver <<vserver>> -junction-path <<junction_path>>
```

Example:

```
volume show -vserver xaxis -junction-path /ds1
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
xaxis     ds1          n1a1         online     RW          2TB     1.03TB   48%
```

3. If the volume contains one or more VMFS LUNs, use the second part of the LUN path name after `/vol/` for the volume name.
4. On the NetApp cluster CLI, use the `volume modify` command to set the QoS policy on the volume.

```
volume modify -vserver <<vserver>> -volume <<volume>> -qos-policy-group <<policy_group>>
```

Set QoS Policy on SVM

To set a QoS policy on an SVM, complete the following steps:

Note: All VMs and applications that reside in this SVM, whether in files, LUNs, or volumes, share the throughput of the SVM QoS policy. This includes VMware and all other application workloads.

1. Determine which SVM needs the QoS policy by using VSC Monitoring and Host Configuration and looking at Storage Details – SAN or Storage Details – NAS.
2. On the NetApp cluster CLI, run the `vserver modify` command to set the QoS policy on the SVM.

```
vserver modify -vserver <<vserver>> -qos-policy-group <<policy_group>>
```

8.6 VMware vSphere 6.x Storage I/O Control

The VMware SIOC feature, introduced in vSphere 4.1, enables QoS control for storage through the concepts of shares and limits in the same way that CPU and memory resources have been managed using resource pools. SIOC allows the administrator to make sure that certain VMs are given priority access to storage as compared to other VMs according to the following factors:

- Allocation of resource shares
- Maximum IOPS limit
- Whether the datastore has reached the specified congestion threshold, as defined by either total datastore latency or, as of vSphere 5.1, a percentage of peak throughput

SIOC is currently supported on VMFS and NFS datastores. For SIOC to be used, it must first be enabled on the datastore, and then resource shares and limits must be applied to the VMs in that datastore. The VM limits are applied under each virtual hard disk in the Edit Settings dialog box for the VM. By default, all VMs in the datastore are given equal resource shares and unlimited IOPS. Figure 36 shows how to enable SIOC on a datastore.

Figure 36) Enabling SIOC on a datastore using the vSphere Web Client.

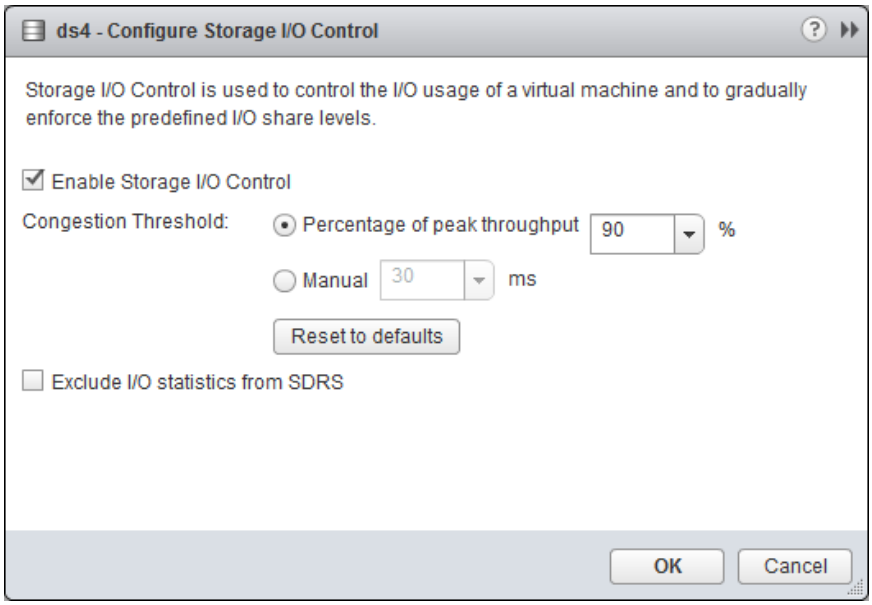
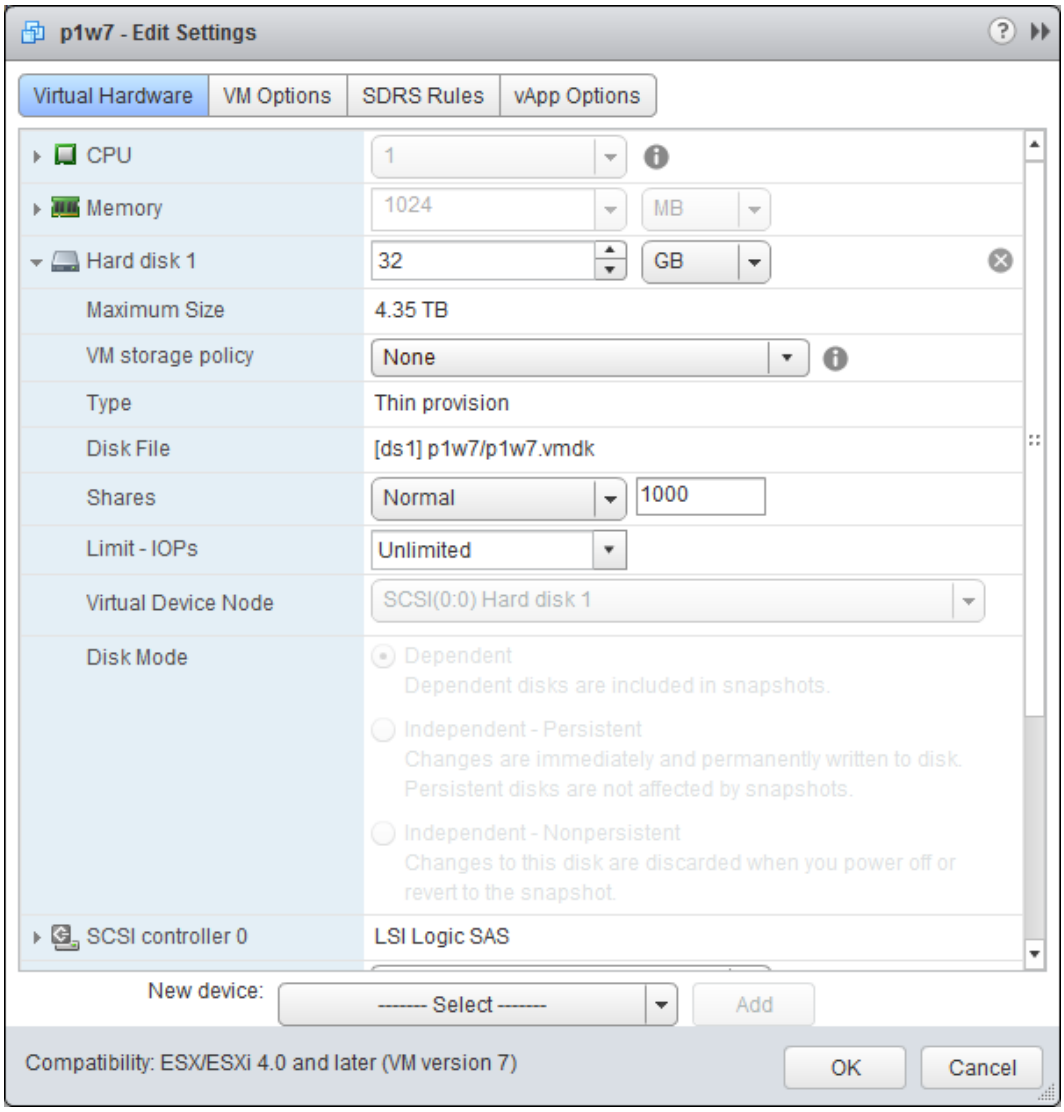


Figure 37 shows the shares and limits for VMDKs in vSphere 6.x using the vSphere Web Client.

Figure 37) Enabling SIOC on a VM using the vSphere Web Client.



SIOC does not take action to limit the storage throughput of a VM based on the value of its resource shares until the datastore congestion threshold is met. As long as datastore latency is under the configured threshold, all VMs on the datastore have equal, unimpeded access to the storage. The congestion threshold is set per datastore in a value of milliseconds (ms) of latency or, if using vSphere 5.1 or later, as a percentage of maximum throughput. The default latency value of 30ms is appropriate for most storage types. However, there might be a desire to increase (if using 7.2k RPM serial ATA [SATA] disks) or decrease (if using solid-state drive [SSD] or 15k RPM disks) the latency value to match the capability of the aggregate supporting your virtual servers.

Storage resource shares are set in values of low, normal, and high (these values are 500, 1,000, and 2,000, respectively), or a custom value can be set. The value for resource shares is used to determine how much preference one VM is given compared to another VM on the SIOC-enabled datastore. For example, when SIOC limitations are imposed on the datastore, a VM with 1,000 shares is entitled to twice the access to resources as a VM with 500 shares. The actual amount of throughput achievable by each VM is dependent on the I/O size of the VM. The share settings of multiple VMs can be viewed in the vSphere Web Client Datastores view by selecting a datastore and then the Related Objects tab > Virtual Machines.

The access of a VM to a datastore can also be limited by maximum storage IOPS. Setting a maximum IOPS limit on a VM causes vSphere to continuously limit the throughput of that VM to the configured IOPS value, even if the congestion threshold has not been surpassed. To limit a VM to a certain amount of throughput in megabytes per second (MBps), use the IOPS limit by setting an appropriate maximum IOPS value according to the VM's typical I/O size. For example, to limit a VM with a typical I/O size of 8kB to 10MBps of throughput, set the maximum IOPS for the VM to 1,280. The following formula can be used:

```
kBps ÷ I/O size = IOPS
```

For example: `10,240kB ÷ 8kB = 1,280 IOPS.`

Be aware that this limit is aggregated for all VMDKs belonging to the virtual server in the same datastore. For example, if disk 1 is configured for an IOPS limit of 100, but disk 2 is configured for an IOPS limit of 1,000, the VM is able to consume a total of 1,100 IOPS to either disk if they are in the same datastore. If the disks are in separate datastores, then the IOPS limits work as expected. An IOPS limit must be configured for all disks for the VM, or no limit is observed.

SIOC is most effective when used on volumes wherein all VMware datastores, regardless of access protocol, have SIOC enabled, with identical threshold values, and the volumes are not shared with non-VMware workloads.

8.7 Using VMware vSphere 6.x Storage I/O Control

Table 34) VMware vSphere 6.x Storage I/O Control prerequisites.

Description
Datastores that are SIOC enabled must be managed by a single vCenter Server system.
Each datastore must have no more than one extent because SIOC does not support datastores with multiple extents.
SIOC must be on FC-connected, iSCSI-connected, or NFS-connected storage. RDM is not supported.

SIOC configuration is a two-step process:

1. Enable SIOC for the datastore.
2. Set the number of storage I/O shares and the upper limit of IOPS allowed for each VM.

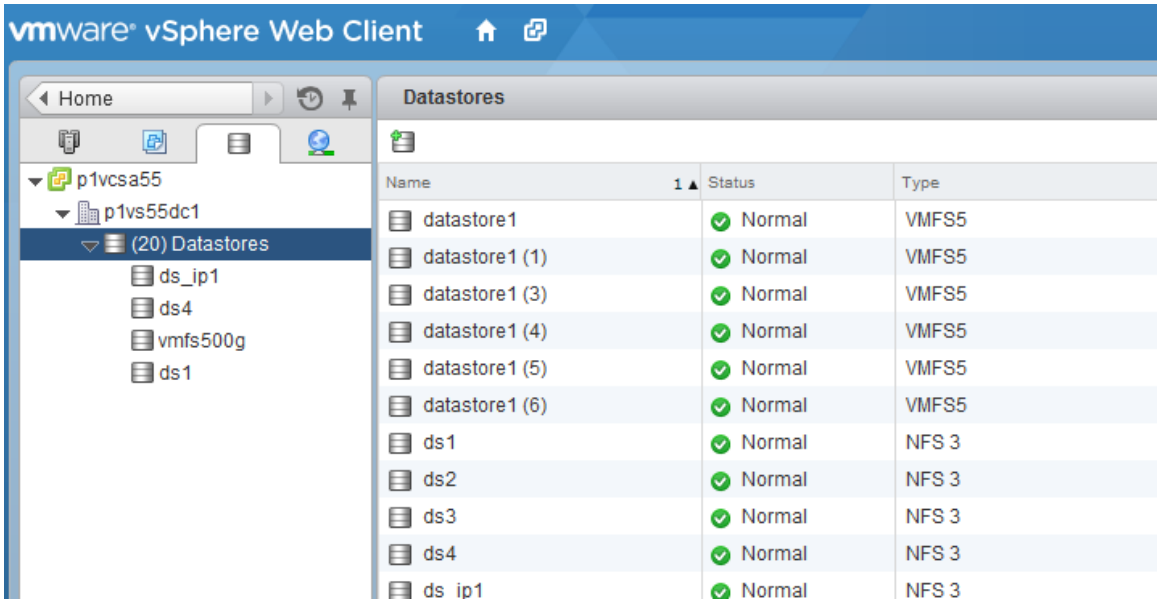
Note: All VM shares are set to the normal value (1,000) with unlimited IOPS by default.

Note: SIOC is enabled by default on VMware Storage Distributed Resource Scheduler–enabled datastore clusters.

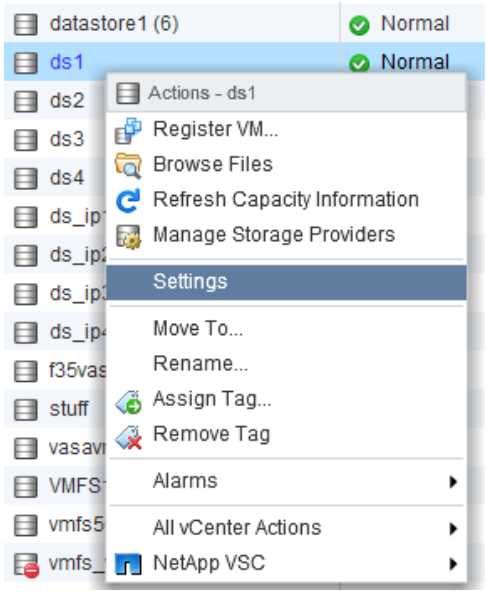
Enable Storage I/O Control for Datastore

To enable SIOC for the datastore, complete the following steps:

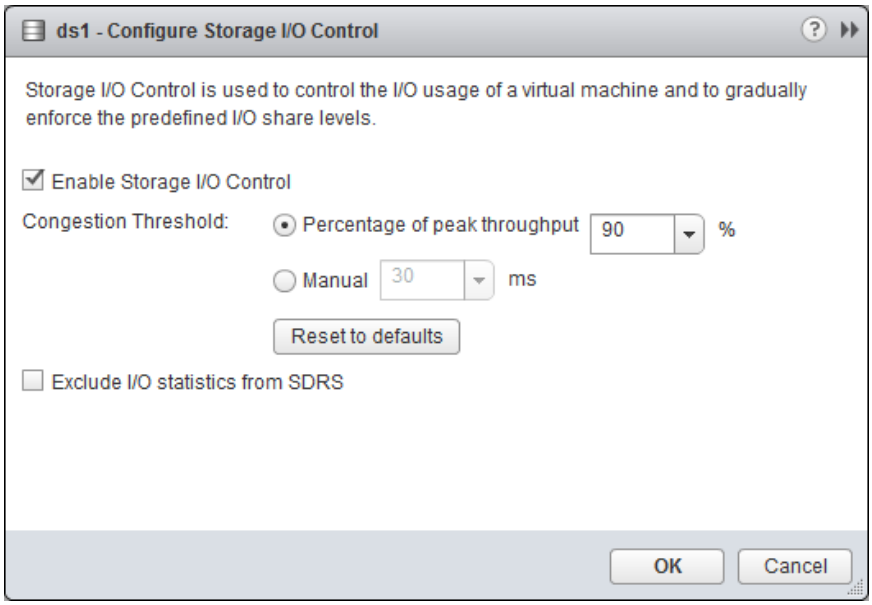
1. Using the vSphere Web Client, navigate to a view that shows datastores.



2. Right-click a datastore in the inventory and select Settings.



3. Click Edit next to Datastore Capabilities.
4. Check Enable Storage I/O Control and set either a percentage or a specific manual threshold value, which must be between 5ms and 100ms. You can click Reset to Defaults to restore the congestion threshold setting to the default value (30ms).

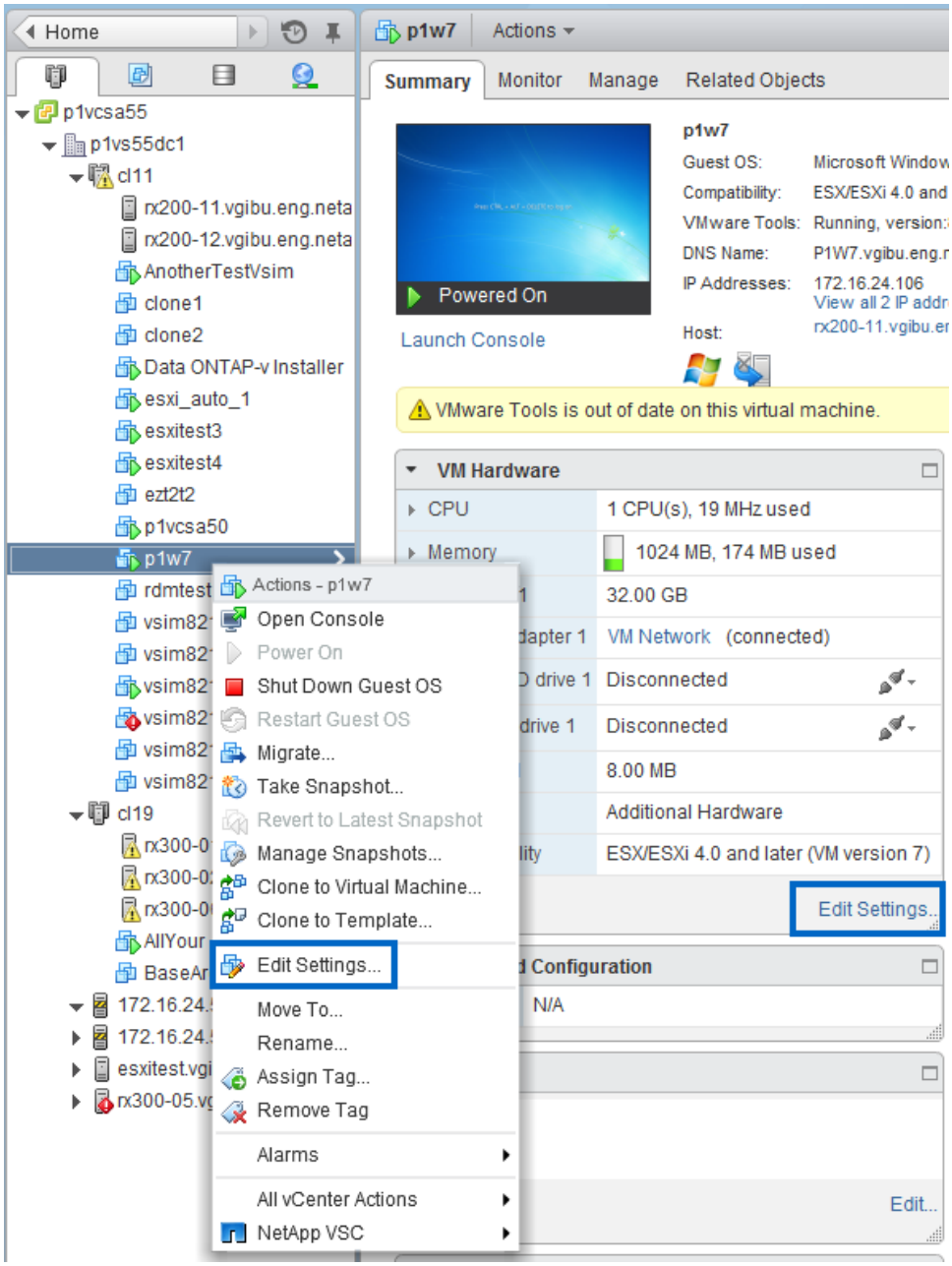


Note: SIOC does not function correctly unless all datastores that share the same spindles on the array have the same congestion threshold.

Set Storage I/O Control Resource Shares and Limits

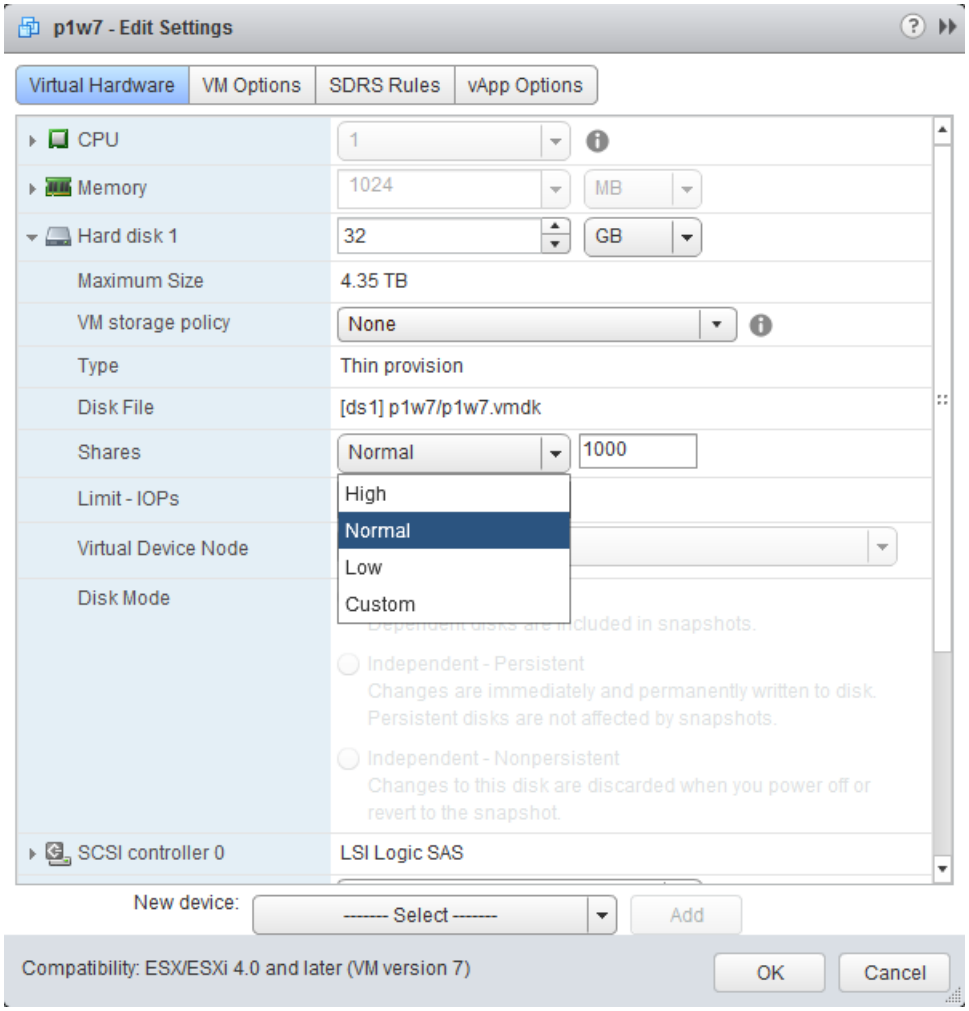
To set the number of storage I/O shares and the upper limit of IOPS for each VM, complete the following steps:

1. Select or right-click a VM in the vSphere Web Client inventory. Click Edit Settings.



2. Expand the virtual hard disk node to expose the SIOC shares and limits and other details.
3. Do one of the following:
   - Click the Shares drop-down list to select the relative number of shares to allocate to the VM (low, normal, or high).

Note: The Shares Value column is not editable for these three options.



- Alternatively, select Custom in the Shares column to enter a user-defined shares value.

4. To set a value for the upper limit of storage resources to allocate to the VM, select a numeric value from the Limit - IOPS drop-down list. The default value is unlimited.
5. Click OK.

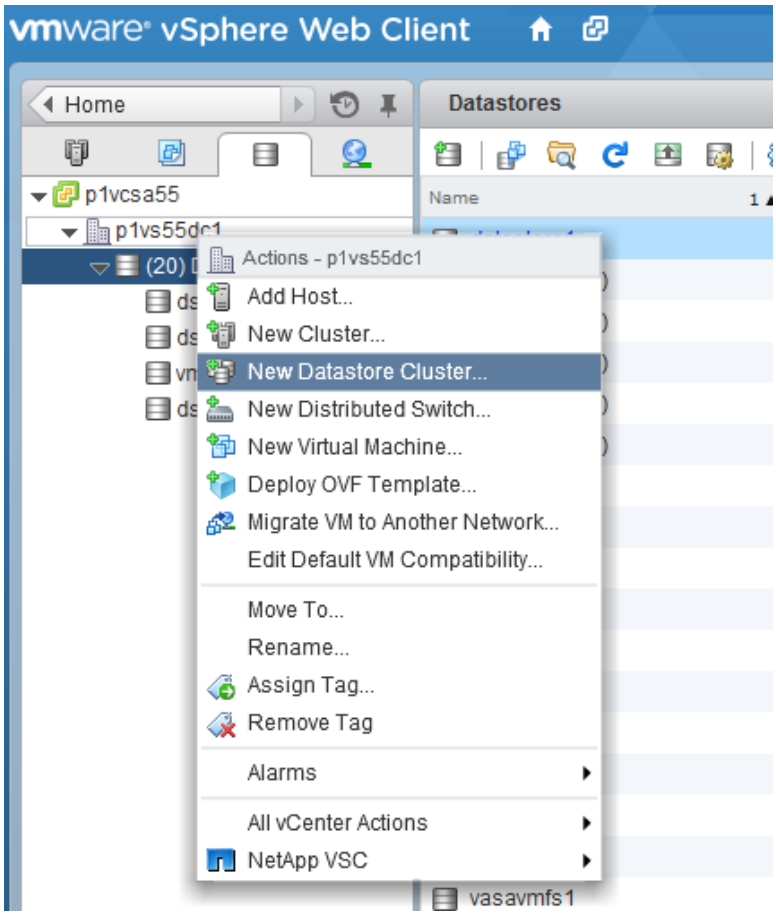
8.8 VMware vSphere 6.x Storage DRS

VMware Storage Distributed Resource Scheduler (DRS) is a new vSphere feature that provides smart VM placement across storage by making load-balancing decisions that are based on the current I/O latency and space usage. It then moves the VM or VMDKs nondisruptively between the datastores in a datastore cluster (also referred to as a pod), selecting the best datastore in which to place the VM or VMDKs in the datastore cluster.

Datastore Cluster

A datastore cluster is a collection of similar datastores that are aggregated into a single unit of consumption from an administrator's perspective. It enables the smart and rapid placement of new VMs and VMDKs and the load balancing of existing workloads. Figure 38 shows the menu option in the vSphere Web Client for creating a new datastore cluster.

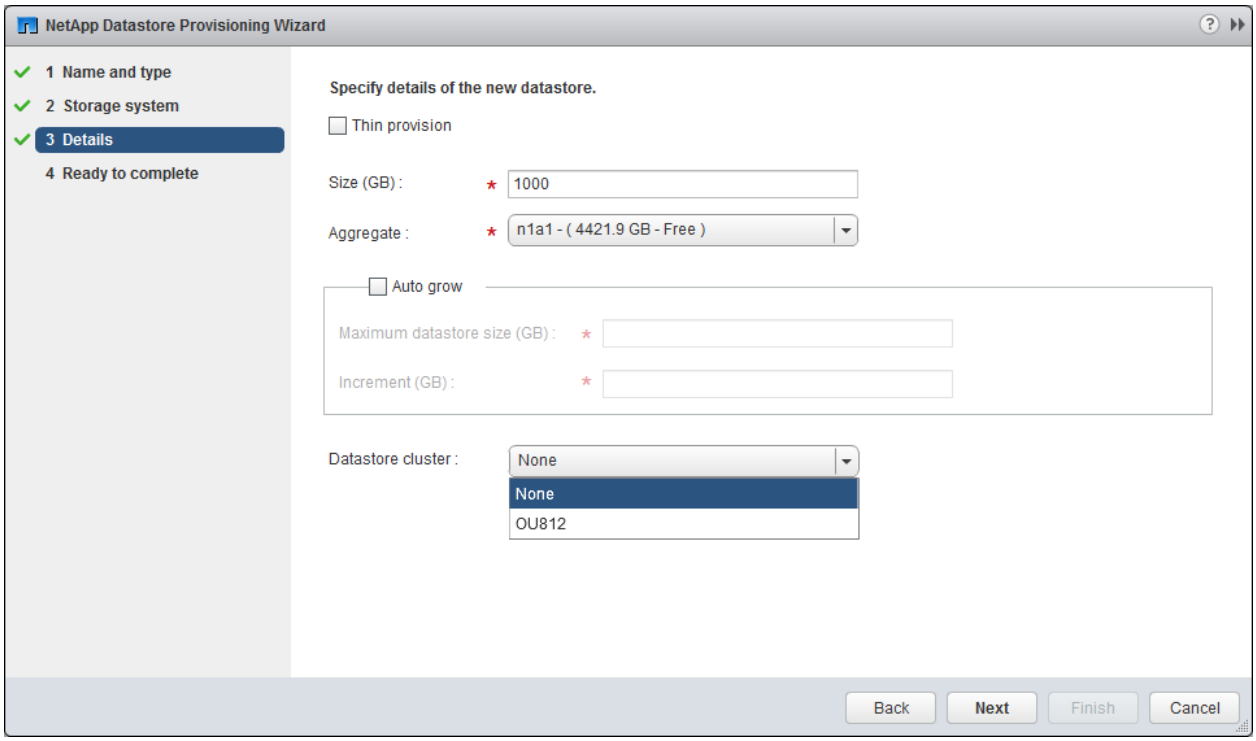
Figure 38) New datastore cluster.



At least one of the volumes to be used in the datastore cluster must exist before the datastore cluster is created or the NetApp Datastore Provisioning wizard does not move past the datastore selection page. The first datastore for a datastore cluster can be created by using the VSC Provisioning wizard. Then a datastore cluster can be created and the datastore added to it as part of the workflow.

After the datastore cluster is created, additional datastores can be added to the datastore cluster directly from the VSC provisioning wizard on the Datastore Details page, as shown in Figure 39.

Figure 39) Adding a new datastore to a datastore cluster by using VSC Datastore Provisioning wizard.



Best Practices

Following are the key recommendations for configuring Storage DRS and the datastore cluster:

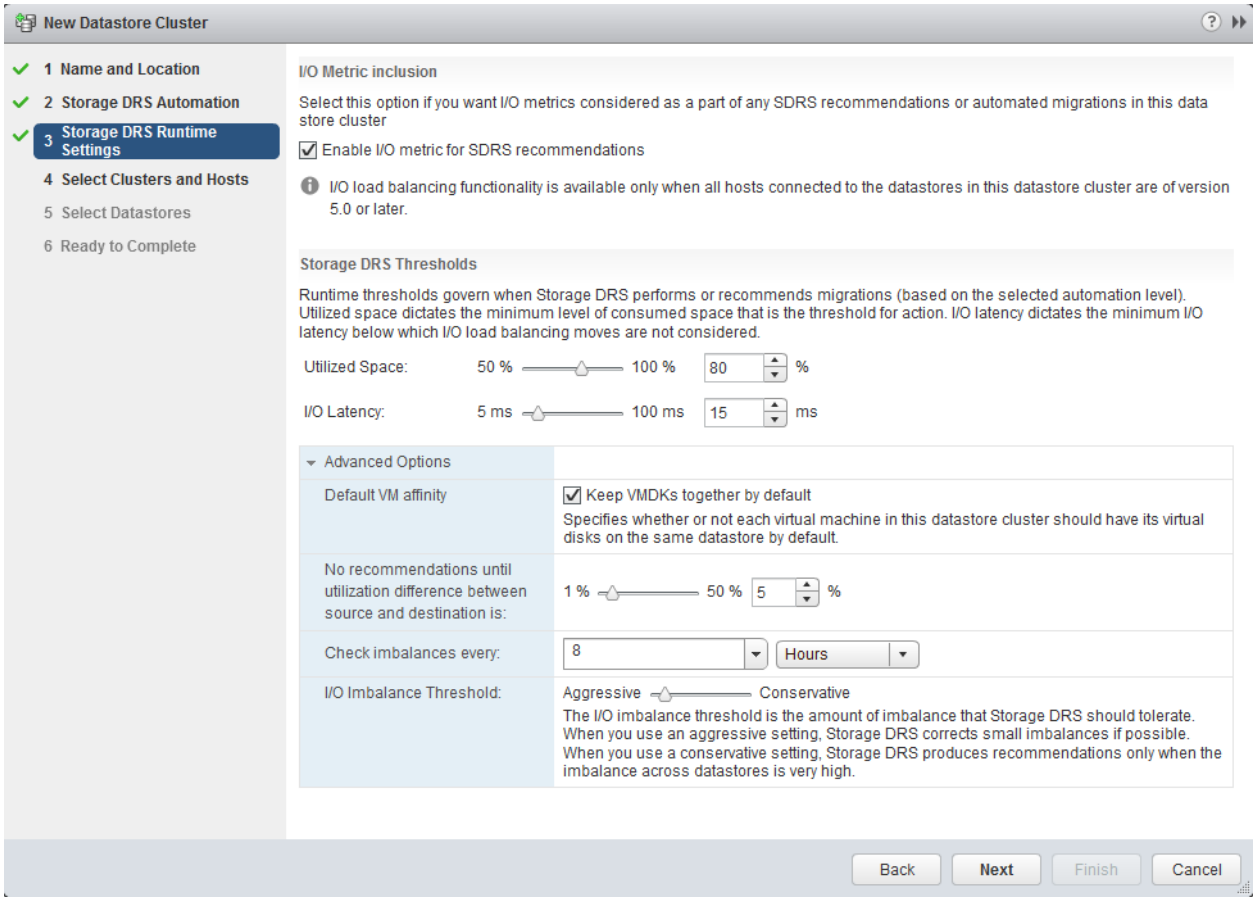
- Set Storage DRS to manual mode and review the recommendations before accepting them.
- All datastores in the cluster should use the same type of storage (SAS, SATA, and so on) and have the same replication and protection settings.
- Storage DRS moves VMDKs between datastores, and any space savings from NetApp cloning or deduplication are lost when a VMDK is moved. You can rerun deduplication to regain these savings.
- After Storage DRS moves VMDKs, NetApp recommends recreating the Snapshot copies at the source datastore.
- Do not use Storage DRS on thin-provisioned VMFS datastores because of the risk of running out of space.
- Do not mix replicated and nonreplicated datastores in a datastore cluster.
- Datastores in a Storage DRS cluster must be either all VMFS or NFS datastores.
- Datastores cannot be shared between different sites.
- All datastore hosts within the datastore cluster must be ESXi 5 hosts.

Placement Recommendations

Storage DRS provides initial placement and ongoing balancing recommendations to assist vSphere administrators in making VM placement decisions. During the provisioning of a VM, a datastore cluster can be selected as the target destination for this VM or virtual disk; Storage DRS then makes a recommendation for initial placement based on space and I/O capacity.

An ongoing balancing algorithm issues migration recommendations when a datastore in a pod exceeds user-configurable space utilization or I/O latency thresholds. These thresholds are typically defined during the configuration of the pods. Figure 40 provides an example of threshold settings for Storage DRS.

Figure 40) Defining thresholds for Storage DRS.



Storage DRS uses the datastore utilization reporting mechanism in vCenter Server to make recommendations whenever the configured utilized space threshold is exceeded. By default, the I/O load is evaluated every 8 hours, currently with a default latency threshold of 15ms. Only when this I/O latency threshold is exceeded does Storage DRS calculate all possible moves to balance the load while considering the cost and the benefit of the migration. If the benefit does not last for at least 24 hours, Storage DRS does not make the recommendation.

Affinity Rules and Maintenance Mode

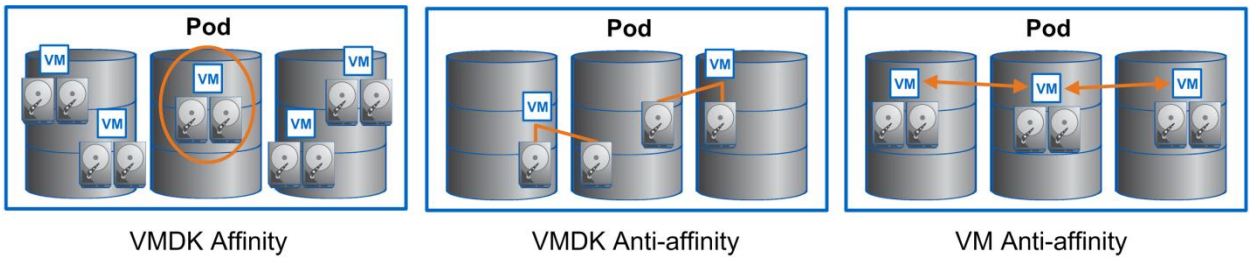
Storage DRS affinity rules make it possible to control which virtual disks are placed on the same datastore within a datastore cluster. By default, a VM's virtual disks are kept together on the same datastore.

Storage DRS offers three types of affinity rules:

- **VMDK affinity**. Virtual disks are kept together on the same datastore.
- **VMDK antiaffinity**. Virtual disks in a VM with multiple virtual disks are placed on different datastores.
- **VM antiaffinity**. Two specified VMs, including associated disks, are placed on different datastores.

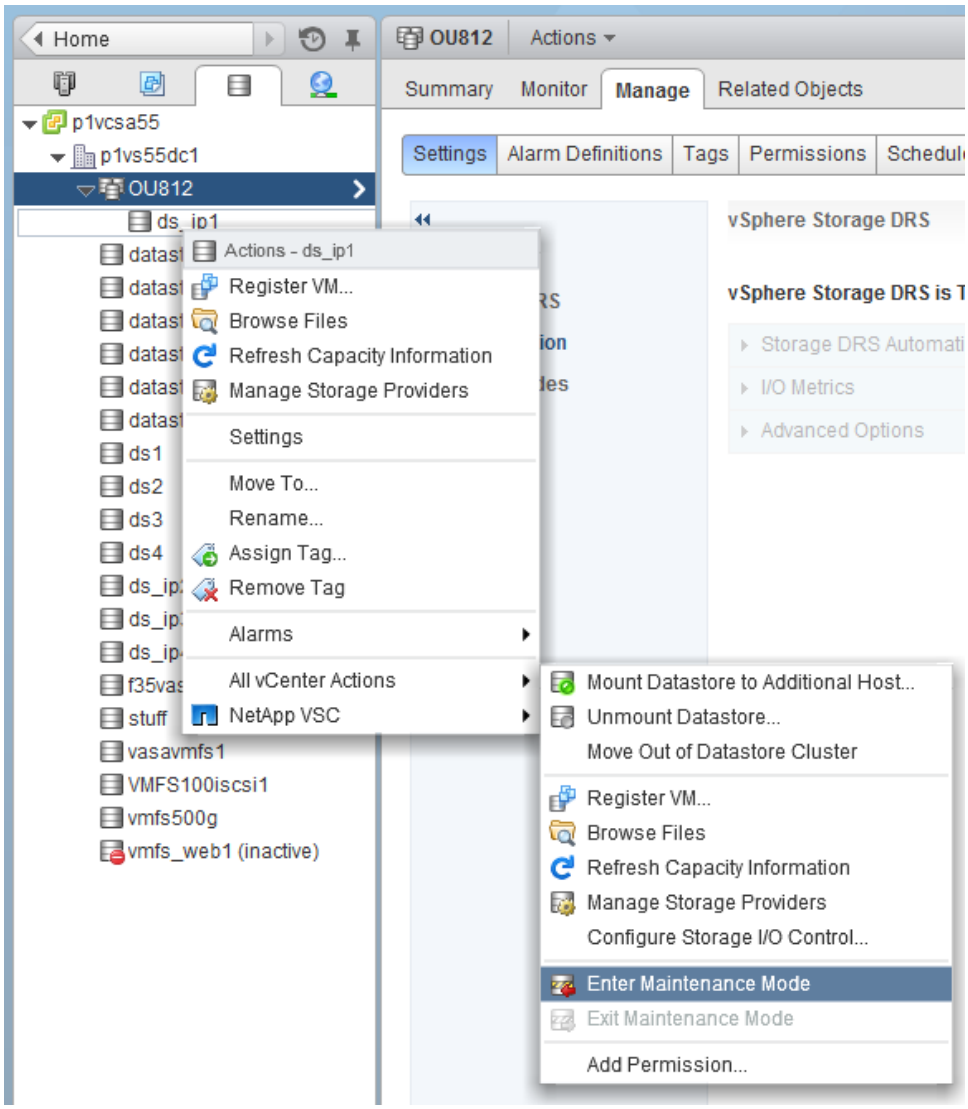
Figure 41 illustrates the three types of affinity rules.

Figure 41) Affinity rules.



In addition, Storage DRS offers the datastore maintenance mode, shown in Figure 42, which automatically evacuates all VMs and virtual disk drives from the selected datastore to the remaining datastores in the datastore cluster.

Figure 42) Datastore maintenance mode.



Storage DRS Interoperability Considerations with Data ONTAP

Table 35 summarizes the interoperability between Storage DRS and the advanced features of Data ONTAP.

Table 35) Storage DRS interoperability with Data ONTAP.

NetApp Feature	Storage DRS Initial Placement	Storage DRS Migration Recommendation
Snapshot technology	Supported	Use manual mode only and recreate Snapshot copies on the destination datastore.
Deduplication	Supported	Use manual mode only and rerun deduplication to regain storage savings.
Thin provisioning	Supported	Use manual mode only; supported on VASA-enabled arrays only. NetApp strongly recommends not using Storage DRS on thin-provisioned VMFS datastores.
SnapMirror	Supported	Use manual mode only because Storage vMotion can cause a temporary lapse in protection (break recovery point objective [RPO]) and increase the size of the next replication transfer.
MetroCluster™	Supported	Use manual mode only. Configure DRS host affinity groups to keep VM migrations from resulting in VMDK access on the other site. Configure datastore clusters with site affinity using datastores from a single site.

The following subsections elaborate on some of the combinations listed in Table 35.

NetApp Snapshot Technology and Storage DRS

NetApp Snapshot technology protects data by locking the blocks that were owned by an object when the Snapshot copy was created. If the original object is deleted, the Snapshot copy still holds the blocks that the object held at the time of the Snapshot copy creation. When Storage DRS moves a VM from one datastore to another, it is effectively deleting that VM from the source datastore. If the reason for the Storage DRS migration was to free up space on the original datastore, the goal is not achieved until all Snapshot copies containing the VM expire by schedule or are otherwise deleted.

Snapshot copies cannot be migrated with the VM. Migration of VMs can break the relationship between the VM, which is now on a new datastore, and its Snapshot copies, which are still on the original datastore, depending on whether the backup management software is Storage DRS–aware. The Snapshot copies are still a complete and valid backup of the VM. However, the backup software might not be aware that the VM was simply moved to another datastore. Over time, with normal backup schedules, Snapshot copies on the new datastore capture the migrated VM.

NetApp Deduplication and Storage DRS

If a VM to be migrated with Storage DRS has been deduplicated and is sharing many of its blocks with other VMs, and the goal of Storage DRS is to recover space on the original datastore, the only space that is freed by the migration is the space taken by the unique blocks of that VM, which, depending on the VM, might be a small percentage.

As a VM is written to a new datastore during a Storage DRS migration, it is initially seen by the new datastore as new blocks. While block hashes are calculated on deduplication-enabled datastores as blocks are being written, the blocks are not actually deduplicated until the next scheduled or manual deduplication runs. As a result, the VM initially consumes 100% of its size in space.

NetApp Datastore Thin Provisioning and Storage DRS

With NFS datastores, two scenarios are possible when an object is deleted:

- All of its blocks are freed and returned to the volume, if the volume is not thin provisioned.
- All of its blocks are freed and returned to the containing aggregate, if the volume is thin provisioned.

Thin-provisioned LUNs do not get blocks allocated until the blocks are written. When a file is deleted in a file system inside a LUN (including VMFS), blocks are not actually zeroed or freed in a meaningful way to the underlying storage. Therefore, after a block in a LUN is written, it remains owned by the LUN even if freed at a higher layer. If the goal of Storage DRS migration was to free space in an aggregate that contains a thin-provisioned VMFS datastore, that goal might not be achieved.

9 Virtual Storage Console

9.1 What Is Virtual Storage Console?

If you are using VMware with NetApp storage, there's one tool you absolutely should have: NetApp VSC for VMware vSphere. VSC is a vCenter plug-in that simplifies storage management, improves efficiencies, enhances availability, and reduces storage costs and operational overhead, whether you use SAN or NAS. VSC gives administrators working in vCenter a window into the storage domain and the tools to manage virtual server and desktop environments running on NetApp storage, all without ever leaving the vSphere client.

The use of VSC is considered a best practice when deploying vSphere on NetApp storage.

Note: This section leverages procedures from the detailed Virtual Storage Console 5.0 for VMware vSphere Installation and Administration Guide (IAG) for all step-by-step instructions and should be used as a companion document specific to this section of the technical report. The IAG is also available for download in PDF and ePub formats.

NetApp also provides a set of how-to videos on the Virtual Storage Console 5.0 for VMware vSphere YouTube channel:

- Navigating the Virtual Storage Console interface in the vSphere Web Client
- Defining storage requirements by using VASA Provider for clustered Data ONTAP
- Protecting your VMware environment by backing up virtual machines and datastores
- Restoring virtual machines and datastores from backup copies

Core Capabilities

The following core capabilities are included in VSC:

- Dashboard command center centralizing all configuration
- Storage system discovery and management
- ESXi host settings and compliance monitoring
- Datastore provisioning
- VM cloning and integrated operations with VMware View Composer for virtual desktops
- Backup and recovery of VMs and datastores
- Online optimization of misaligned VMs, offline alignment of VMs, and single or group migrations into new or existing datastores
- Integration with VASA to provide policy-based operations

As a vCenter plug-in, VSC can be accessed by all vSphere Clients that connect to the vCenter Server. This availability differs from that of a traditional client-side plug-in that must be installed on every vSphere Client.

Note: VSC does not support VMware vCenter operating in linked mode or the management of multiple vCenter Servers.

Figure 43 shows the VSC plug-in listed on the Home screen of the vSphere Web Client.

Figure 43) The VSC plug-in.

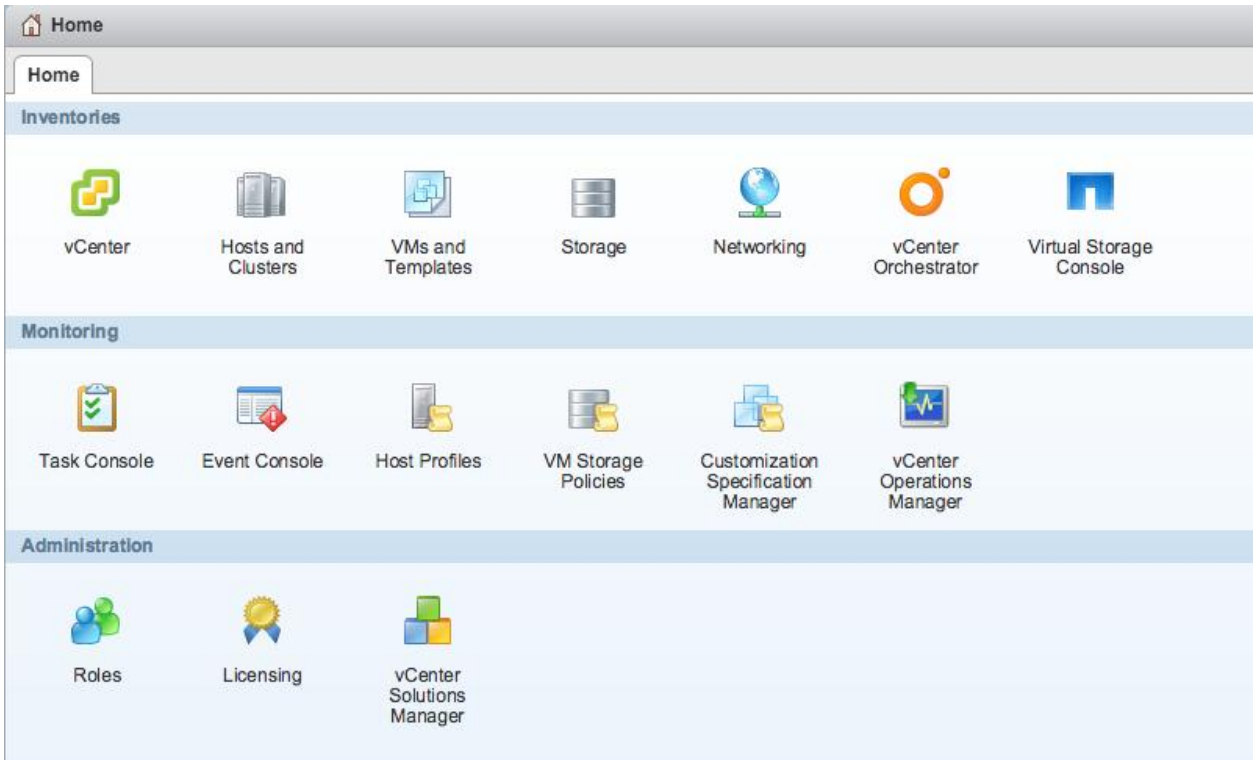
Version Information

The concepts and procedures presented in this document are valid for NetApp VSC for vSphere Web Client 5.0. At the time of this writing, VSC 5.0 is the currently shipping version and only supports VMware vSphere 6.0 and later, running in the vSphere Web Client.

Note: VSC 5.0 is not compatible with the legacy Virtual Infrastructure Client (VIC), because VSC 5.0 is written in a different language specific to the vSphere Web Client.

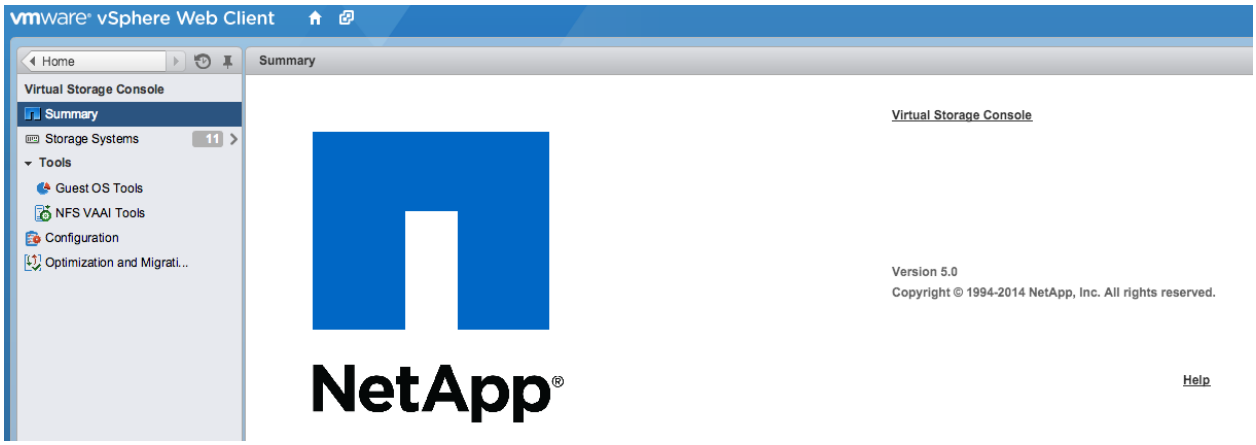
The VSC software adds a NetApp icon to the Inventories section of the vSphere Client homepage, as shown in Figure 44.

Figure 44) The NetApp icon.



The VSC version information is displayed in the Summary pane of the VSC, as shown in Figure 45.

Figure 45) About VSC.



Note: Versions in customer environments might differ from the versions shown in this example.

License Requirements

Table 36 lists the licenses required for performing provisioning, cloning, configuration, and distribution tasks through VSC. To use the VSC foundational capabilities, the protocol license provided with the purchase of your storage system is enough to get you going. However, additional add-on licenses are required in order to take advantage of additional functionality, such as rapid cloning, replication, and quick restores. The complete license bundle from NetApp includes all of these licenses.

Table 36) License key requirements per task type.

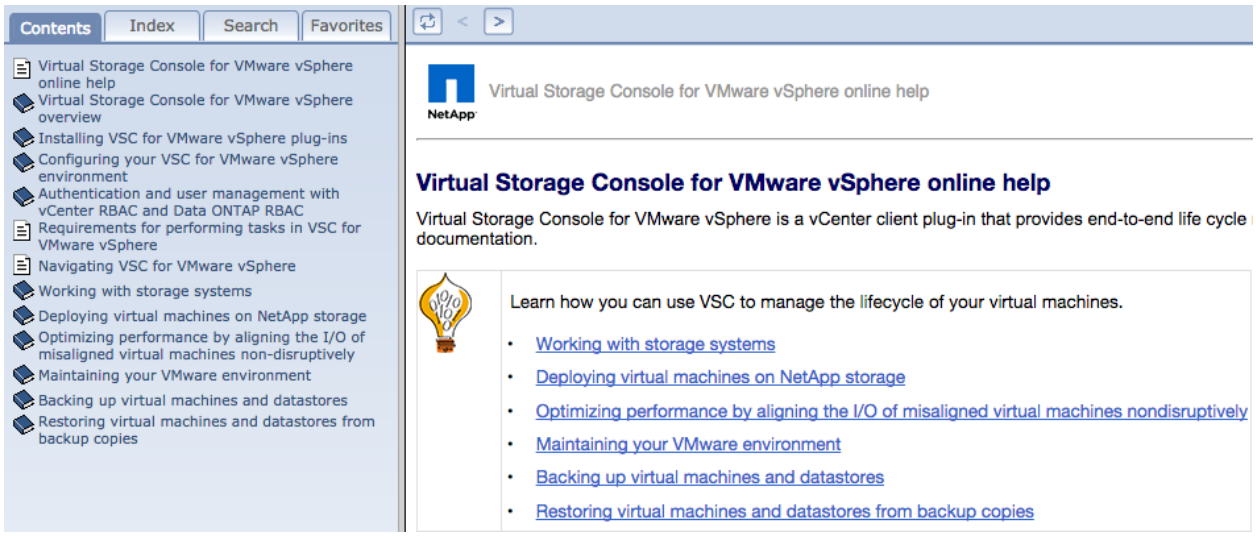
Task	Required License
Provision datastores	NFS, iSCSI, or FCP
Clone VMs	FlexClone
Distribute templates to remote vCenter servers	SnapMirror
Restore VMs backed up through VSC	SnapRestore

Note: The license requirement for deduplication (A-SIS) was removed in clustered Data ONTAP.

Online Help

VSC has built-in online help that describes the GUI fields and commands that apply to the VSC capabilities. As Figure 46 shows, online help is available from within the vSphere Client simply by clicking the Help link on the Summary screen below the VSC version number. The Online Help launches in a new window and allow you to browse through the help without navigating away from your vSphere client session.

Figure 46) vSphere Client online help.



VSC Installation

You can use the installation wizard to install VSC. By default, the VSC software installs all the features except the backup and restore features. You must manually select the option to install those features because they normally require that you purchase a SnapManager for Virtual Infrastructure license.

Before you begin:

- You must log in with administrator privileges for the machine on which you install VSC.
- You can only install VSC on 64-bit Windows Server. If you're using the vCenter Server Appliance, VSC must be installed on a separate Windows host.
- Before you install VSC, verify that your system meets the VSC requirements listed in the NetApp Interoperability Matrix Tool.

Download the VSC installer and double-click the installer icon. Click Run to start the installation wizard.

Follow the instructions in the installation wizard to install the software.

Note: If you want to install backup and restore features, you must select that option. Otherwise, the VSC installer does not install it. Menu options for the features appear in the installed VSC GUI, but you cannot use them.

Click Finish to complete the installation.

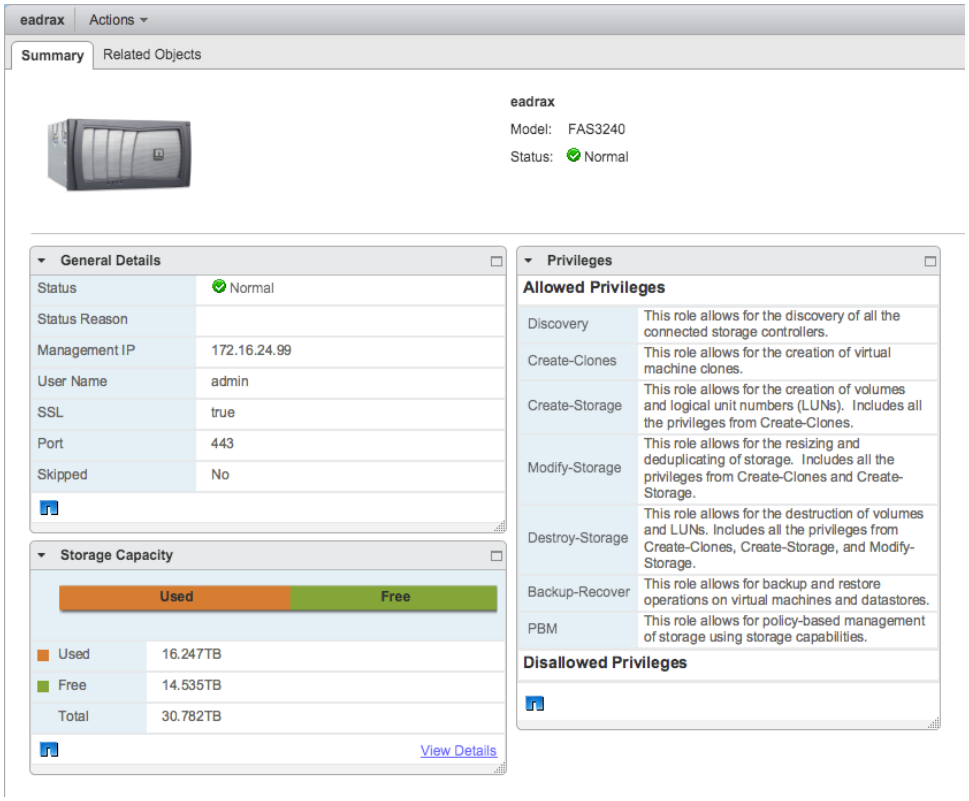
Note: At the webpage that appears after the installation completes, register VSC with the vCenter Server. You must provide the vCenter Server host name or IP address and the administrative credentials.

9.2 VSC 5.0: Display Integrations

With VSC 5.0, we had an opportunity to take a fresh approach to the GUI with the vSphere Web client, and great care was taken to integrate VSC with vSphere Web Client as deeply as possible, exposing as much of the NetApp value-add in the right place to enhance the VMware administrator's experience when working with NetApp storage.

Figure 47 shows a detailed view of a storage system within the VSC storage systems interface.

Figure 47) Storage system details in VSC.



Aside from the command center dashboard, we have exposed some familiar NetApp metrics and identifiers on the summary pages for datastores. Figure 48 to Figure 51 show portlets marked with a blue NetApp icon.

Figure 48) NetApp portlet showing LUN details in the datastore summary.

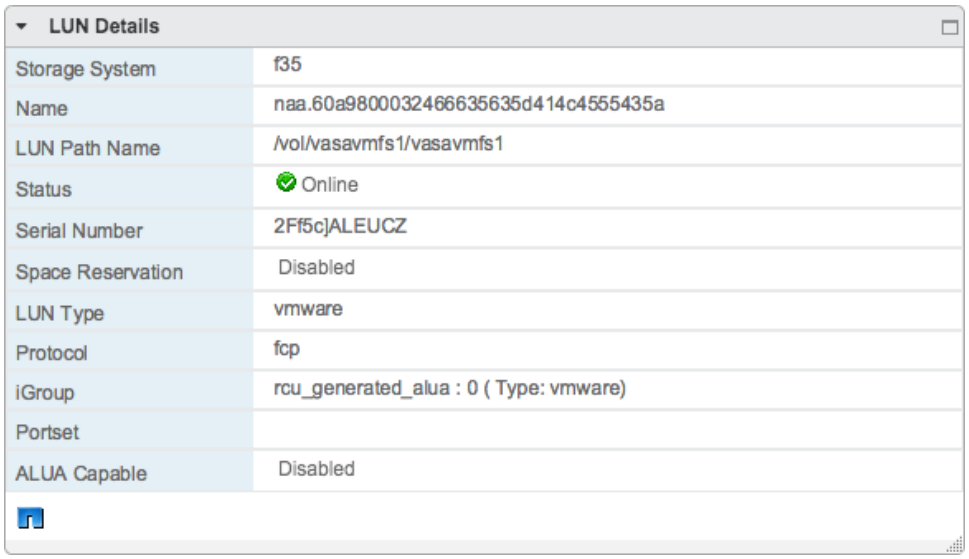


Figure 49) NetApp portlet showing NFS details in the datastore summary.

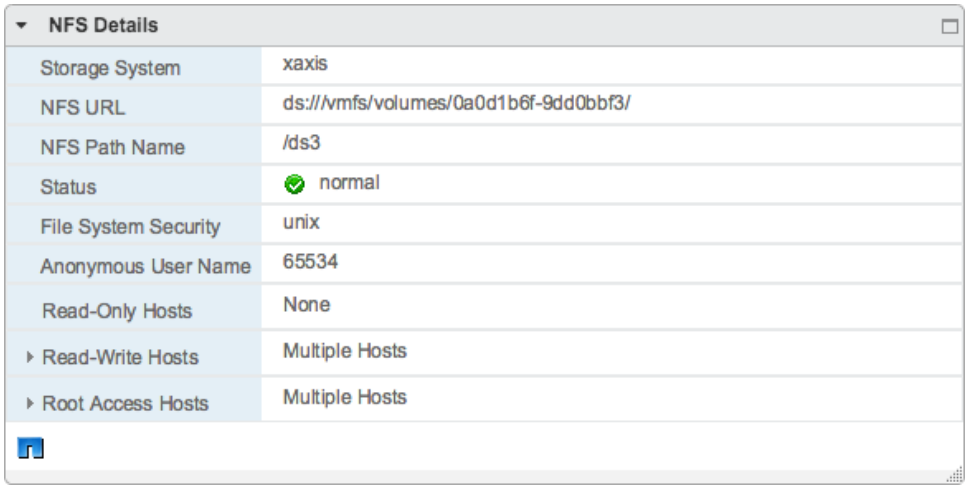


Figure 50) NetApp portlet showing deduplication details in the datastore summary.

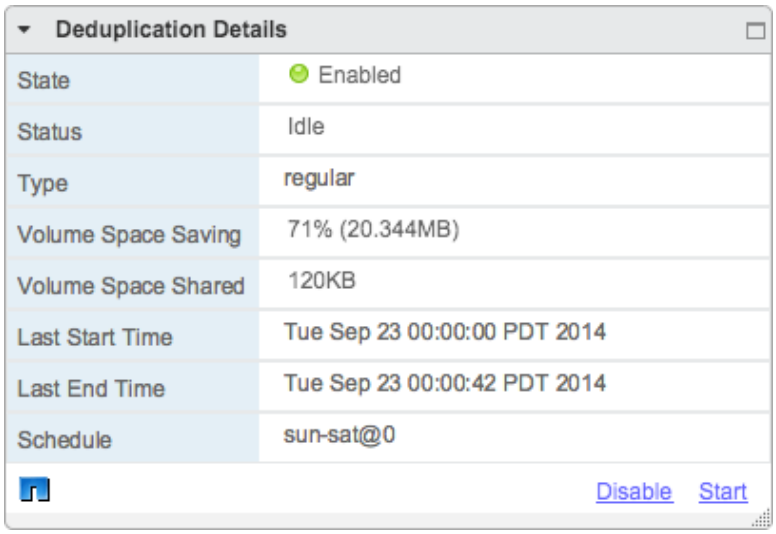
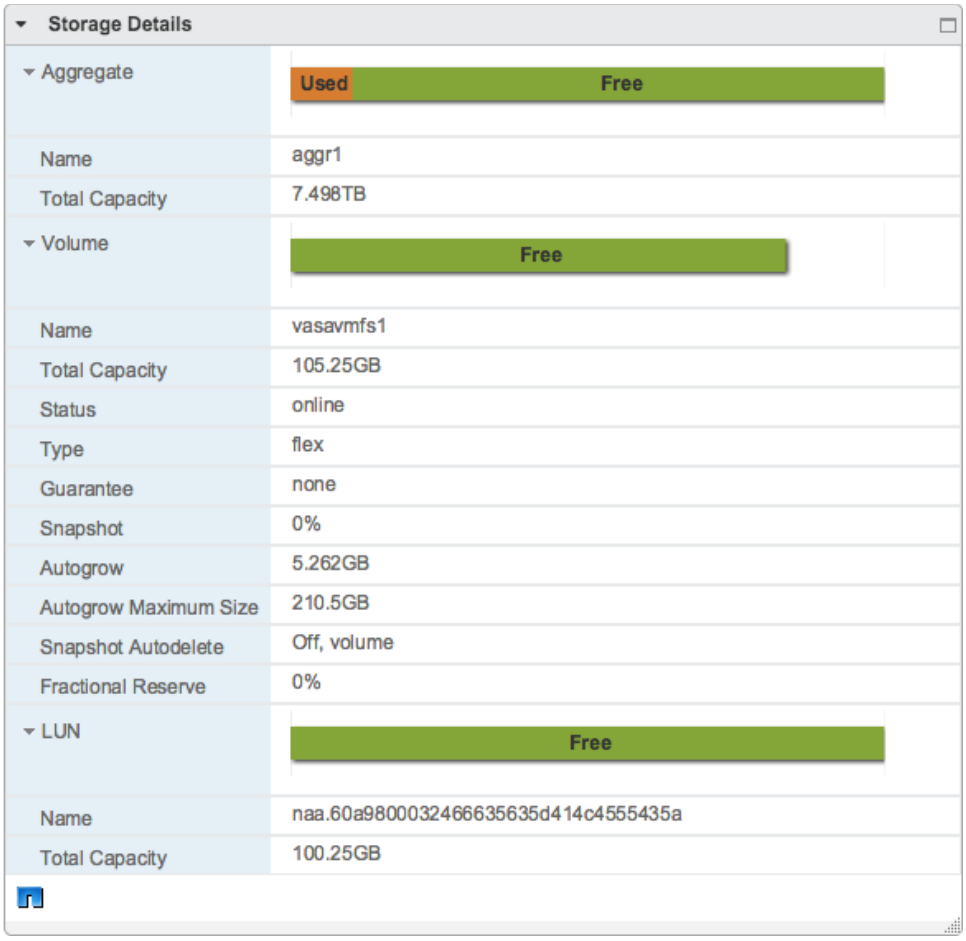


Figure 51) NetApp portlet showing storage details in the datastore summary.



9.3 Additional Plug-In Components

VAAI for NFS

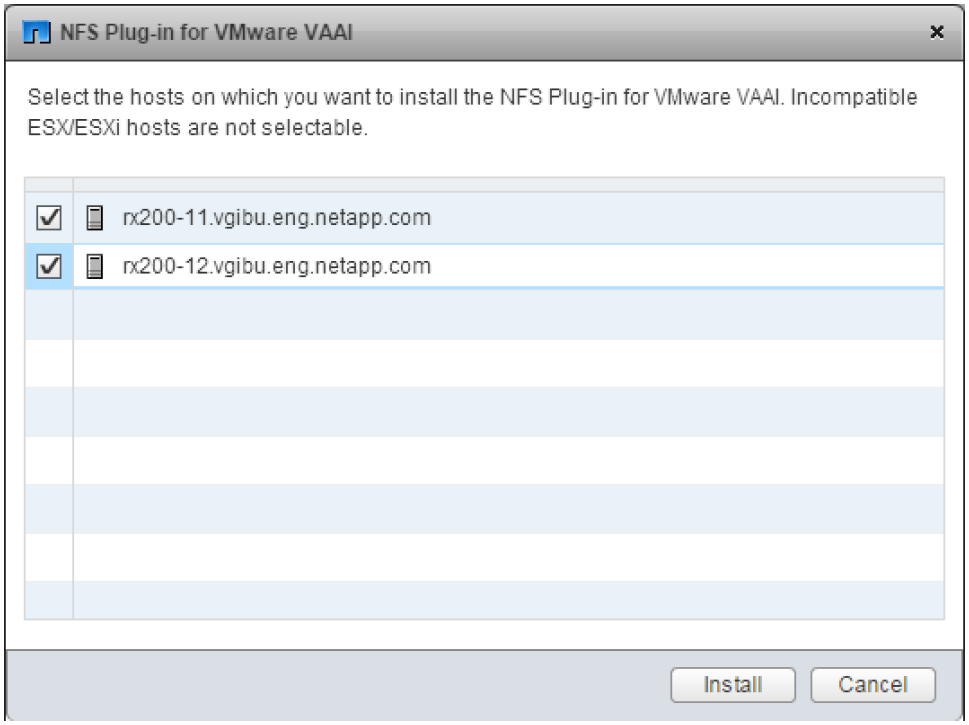
We've covered VAAI in this document for both block and file offload. However, in order to leverage VAAI for NFS, an additional plug-in must be added to each of your ESXi hosts.

This can be a tricky and complicated process to deploy manually. However, VSC has a unique deployment mechanism built in that can push the plug-in to all of your ESXi hosts in a couple of mouse clicks. At the same time, we check compatibility of the clustered Data ONTAP versions of all of your storage controllers, as well as supported ESXi host versions. We also configure all compatible storage systems that have been discovered by VSC with the correct options and settings necessary to accommodate offloaded operations.

Note: This plug-in is only for vSphere environments leveraging storage system resources through NFS. If you are using FC, iSCSI, FCoE, or any combination of the three, VAAI support is native to clustered Data ONTAP, and installing the plug-in nets no additional benefits.

The NetApp NFS plug-in for VMware VAAI is not shipped with VSC for VMware vSphere; however, you can get the plug-in installation package and instructions from the NetApp Support site. Go to the Virtual Storage Console Tools > NFS VAAI page to complete your installation by selecting each of the hosts to which you want to push the VAAI plug-in to and click Install. That's it. NetApp takes care of the rest. Figure 52 shows part of the installation process of using VSC to install the NFS plug-in for VMware VAAI to ESXi hosts.

Figure 52) NetApp VSC installing the NFS plug-in for VMware VAAI to ESXi hosts.



The plug-in is a software library that integrates the VMware virtual disk libraries that are installed on the ESXi host. It is considered a best practice to install the plug-in on all hosts because VAAI features such as copy offload and space reservations can improve the performance of cloning operations.

Note: The plug-in is supported on systems running ESXi 5.0 or later with vSphere 6.0 or later and clustered Data ONTAP 8.1 or later or Data ONTAP 8.1.1 or later operating in 7-Mode.

To download the plug-in, go to the software download page on the NetApp Support site and log in. This page provides links to both the software installation package and the installation guide.

Follow the VSC installation instructions in Installing the NetApp NFS Plug-in for VMware VAAI.

After you install the plug-in, you must reboot the host. VSC then automatically detects the plug-in and uses it. You do not need to perform additional tasks to enable it.

VASA Provider

vStorage APIs for Storage Awareness (VASA) is a set of vSphere APIs designed to allow storage vendors to advertise proprietary features and abilities into the upper hypervisor layer in order for VMware administrators to take advantage of such a design without having to deeply understand storage mechanics.

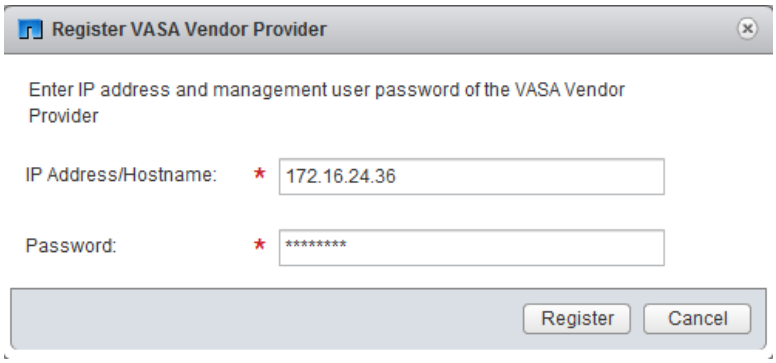
Why Use VASA

VASA is part of a greater story around policy-based management, and specifically, managing components of your infrastructure at a more granular level. This is often referred to as VM-granular management. For more information, refer to section 9.11, “Policy-Based Management.”

VSC is the management console for VASA Provider. Because of this dependency, VASA Provider installation software is available for download from the VSC download page as well as from the VASA Provider for Clustered Data ONTAP download page. These pages are located on the NetApp Support site.

After you install the VASA Provider, you must register it with VSC, then log out of the vSphere Web Client. Figure 53 shows the VASA Provider registration dialog box.

Figure 53) VSC registers VASA Provider with just an IP address and password.



When you log back in and go to the VSC page, VSC displays a link to the VASA Provider GUI, which enables you to create storage capability profiles, assign the profiles to existing storage, and set up thresholds for alarms.

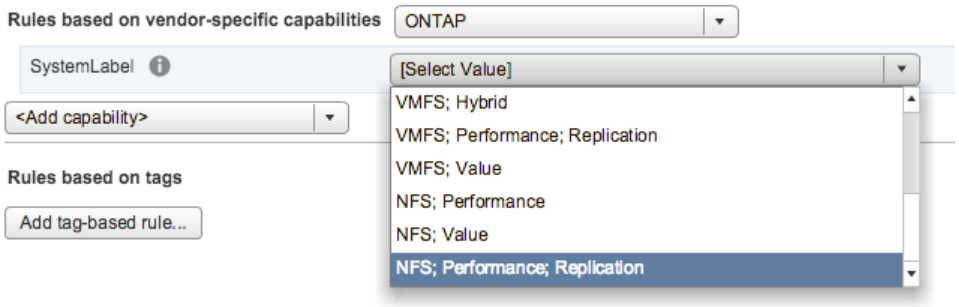
If you need to adjust settings for VASA Provider or perform maintenance tasks, you can use the VASA Provider maintenance menus, which are accessible from the console of the virtual appliance. The main menu provides several options for configuring VASA Provider and performing diagnostic operations. In addition, you can use the VASA Provider control panel, located at http://vm_ip:9080, to generate a support bundle.

Note: If you need to create a support bundle, you should use the VASA Provider control panel to generate it. The VASA Provider control panel creates a more complete bundle than the maintenance menu creates.

Storage Capability Profiles

With the introduction of the VASA Provider, we add an additional construct to our vSphere environment, known as storage capability profiles (SCPs). In vSphere 5.5, VMware added VM storage policies, allowing you to build SLO into your VM provisioning workflows. The combination of these with a VASA Provider and precreated SCPs allows an administrator to intelligently place new VMs in the right place based on a selection of a policy that dictates performance and retention characteristics. NetApp integrates directly into the native VM creation workflow, allowing you to select specific NetApp storage system features on a per-VM basis. Figure 54 shows NetApp VASA integration in VM storage profiles.

Figure 54) VM storage profiles showing NetApp VASA integration.



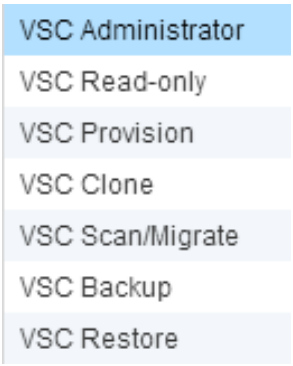
By creating a VM storage policy with inherited NetApp features, you can apply that policy to your new VMs upon creation, insuring that they always land on appropriate storage, and that you're alerted if that VM or storage system is ever out of compliance with the applied policy.

9.4 VSC 5.0: RBAC

RBAC is a process that enables administrators to control access to, and user actions on, vSphere objects and storage systems running Data ONTAP. VSC for VMware vSphere supports both vCenter Server RBAC and Data ONTAP RBAC and, upon installation, creates some example roles for you to clone and modify, tailoring them to your own environment.

Note: It is a best practice to not use these examples in production, but to clone them and modify them accordingly. The reason for this is a contingency that they may one day get updated as part of a later release, where these example roles are overwritten.

Figure 55) VSC RBAC roles.



Note: The administrator handles setting up the RBAC roles. Depending on your system setup, you might have different administrators handling the vCenter RBAC and the Data ONTAP RBAC.

vCenter Server RBAC

This security mechanism restricts the ability of vSphere users to perform VSC tasks on vSphere objects, such as VMs, datastores, and data centers.

The vSphere administrator sets up vCenter Server RBAC by assigning permissions to specific vSphere objects, which are listed in the vSphere inventory. In many cases, a VSC task requires that more than one object have permissions. For this reason, it is a good practice to assign permissions on the root object (also referred to as the root folder). You can then restrict those entities that do not require permissions.

Note: At a minimum, all users must have the VSC-specific, read-only view privilege assigned to them. Without this privilege, users cannot see or access the VSC GUI.

Data ONTAP RBAC

This security mechanism restricts the ability of VSC to perform specific storage operations—such as creating, destroying, or backing up storage for datastores—on a specific storage system.

The storage administrator sets up Data ONTAP RBAC by defining storage credentials consisting of a user name and password in Data ONTAP. The storage credentials map to VSC storage operations. Then the administrator, usually the storage administrator, sets the storage credentials in VSC for each storage system that VSC manages. VSC uses a single set of credentials for each storage system.

VSC checks the vCenter Server RBAC permissions when a user clicks a vSphere object and initiates an action. If a user has the correct vCenter Server RBAC permission to perform that task on that vSphere object, VSC then checks the Data ONTAP credentials for the storage system. If those credentials are also confirmed, then VSC allows the user to perform that task.

9.5 Virtual Storage Console 5.0: ESXi Host Compliance

VSC has a built-in workflow that completely optimizes the multipathing and timeout settings for any ESXi hosts that connect to NetApp storage.

What This Is and Why It Is Important

One of the most common issues in support is the lack of proper configuration of ESXi hosts to work properly with NetApp storage systems. VSC checks and sets the ESX or ESXi host multipathing and HBA timeout settings that enable proper behavior with NetApp storage systems.

The task progress is displayed in the recent tasks pane. As tasks complete, the host status alert icons are replaced by normal or pending reboot icons. If you modify the HBA/CNA or the MPIO settings, a reboot of the ESXi host is required for changes to take effect. This is a hardware change to auxiliary cards and cannot be changed spontaneously. NFS settings can be modified without incurring a host reboot.

For procedures on how to properly configure your ESX or ESXi hosts using VSC, refer to Configuring ESX Server Multipathing and Timeout Settings on the NetApp Library site.

To view a master list of all settings modified by VSC, refer to ESX Host Values Set by VSC for VMware vSphere on the NetApp Library site.

9.6 VSC 5.0: Storage Systems Management

VSC for VMware vSphere provides tools you can use to work with storage systems. Using VSC, you can perform tasks such as the following:

- Have VSC automatically discover storage systems.
- Manually add and remove storage systems.
- Set up default credentials for VSC to use when it adds storage systems.

- Modify the credentials associated with a storage system.
- Use the VSC interface to get a quick view of the storage system details.

Default Credentials

VSC provides a single mechanism to discover storage systems and to set their credentials. The credentials provide the necessary permissions to enable VSC users to perform tasks using the storage systems.

Before VSC can display and manage storage resources, it must discover the storage systems. As part of the discovery process, you must supply Data ONTAP credentials for your storage systems. These are the privileges (or roles) associated with the user name and password pair assigned to each storage system. These user name and password pairs use Data ONTAP (RBAC) and must be set up from within Data ONTAP. You cannot change their credentials from within VSC. You can, however, define Data ONTAP RBAC roles using a tool such as RBAC User Creator for Data ONTAP.

Note: If you log in as an administrator, you automatically have all privileges for that storage system. It is considered a best practice to use a custom account created on the storage system with appropriate privileges. For security and audit purposes, you should never use the admin account directly. It is also common to create a generic account on your storage systems for discovery to be used by VSC for the default storage system discovery credentials.

Manual Discovery

Any time you make a configuration change to either vSphere or your underlying storage systems, it is a good idea to refresh your storage systems information in VSC. This can be triggered in several places, including the VSC storage systems dashboard, as well as directly from the vCenter inventory tree.

For information about refreshing your storage systems, refer to Discovering Storage Systems and Hosts on the NetApp Library site.

Remove Unused Storage Systems

Often, the automatic discovery process picks up storage systems in your environment that are not part of your actual vSphere implementation. To remove these from VSC, simply right-click any unwanted system and click Delete.

Note: If the storage system has datastores mounted to an ESXi host, an error message is displayed, and no changes are made. Otherwise, the system is unassociated with VSC and is not available for selection when provisioning datastores.

9.7 VSC 5.0: Datastore Provisioning

One of the best features of VSC is the ease with which you are able to provision new datastores from your storage systems and mount them to your ESXi hosts in one simple workflow. Through a simple right-click menu, we automate many of the steps for you with full support for NFS, iSCSI, and FC/FCoE.

All workflows have been designed to streamline processes and remove unnecessary steps typically associated with creating and mounting a new datastore manually, while still optimizing the experience and allowing you to customize your datastores as they're provisioned.

The context level you select when right-clicking in the inventory determines to what the new datastore is mounted upon completion. Table 37 shows different datastore availability levels and how to configure them.

Table 37) How to create new datastores with NetApp VSC.

To Make Datastore Available To:	Do This:
All hosts in a data center	Right-click the data center and select NetApp VSC > Provision Datastore.
All hosts in a cluster	Right-click a cluster and select NetApp VSC > Provision Datastore.
A single host	Right-click a host and select NetApp VSC > Provision Datastore.

For complete procedures on creating new datastores, refer to Provisioning Datastores on the NetApp Library site.

9.8 VSC 5.0: VM Rapid Cloning and VDI Integrations

NetApp has been on the forefront of VDI deployments for many years, and our ability to deeply integrate with VMware and Citrix, as well as provide customers with quick refreshes and redeployments, has allowed us to maintain one of the most respected solutions in the market.

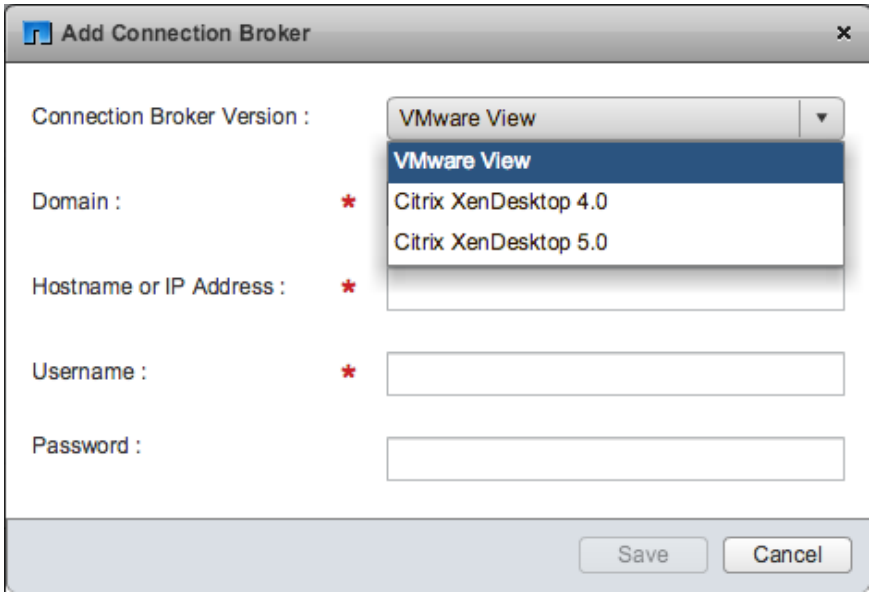
One of the most popular uses of VSC is the creation of rapid clones by using our FlexClone technology. This is most often seen in VDI use cases; however, it is not exclusive to them.

Complete VDI deployments are outside the scope of this document. We have more detailed reference architectures in TR-4181: VMware Horizon View 5 Solutions Guide.

Connection Brokers

You can use the Connection Brokers pane in VSC to view and manage the VDI connection brokers available for importing clone data at the end of the clone operation. Figure 56 shows the VSC Connection Broker management tool.

Figure 56) NetApp VSC Connection Broker management.



- For VMware View Server, clone data is imported into View Server at the end of the clone operation.

- For Citrix XenDesktop, a `.csv` file is created in the directory `c:\program files\netapp\virtual storage console\etc\kamino\exports`.

Note: To work with connection brokers in VSC, you must have .Net 3.5 available on the system where you have VSC installed. For some versions of Windows, such as Windows 2008, .Net 3.5 is included as part of the installation. For other versions, such as Windows 2003, it is not part of the base install, so you must manually install it.

For detailed procedures on integrating connection brokers with VSC, refer to Adding Connection Brokers on the NetApp Library site.

Create Rapid Clones from a VMware Template

Setting up VMs can be a lengthy process. If you're going to deploy multiple, identical VMs, you can save time by setting up a single VM as the template and then rapidly cloning VMs from that template.

For best results when using NFS datastores, you should have the NFS plug-in for VMware VAAI installed. Although not required, installing the NFS plug-in is a best practice because it reduces load from the ESXi host and places it on the storage system, which increases cloning efficiency across the board.

Cloning performance is affected by many factors, including the vCenter Server hardware configuration, the number and hardware configuration of the ESXi hosts, and the current load on both the vCenter Server and the hosts.

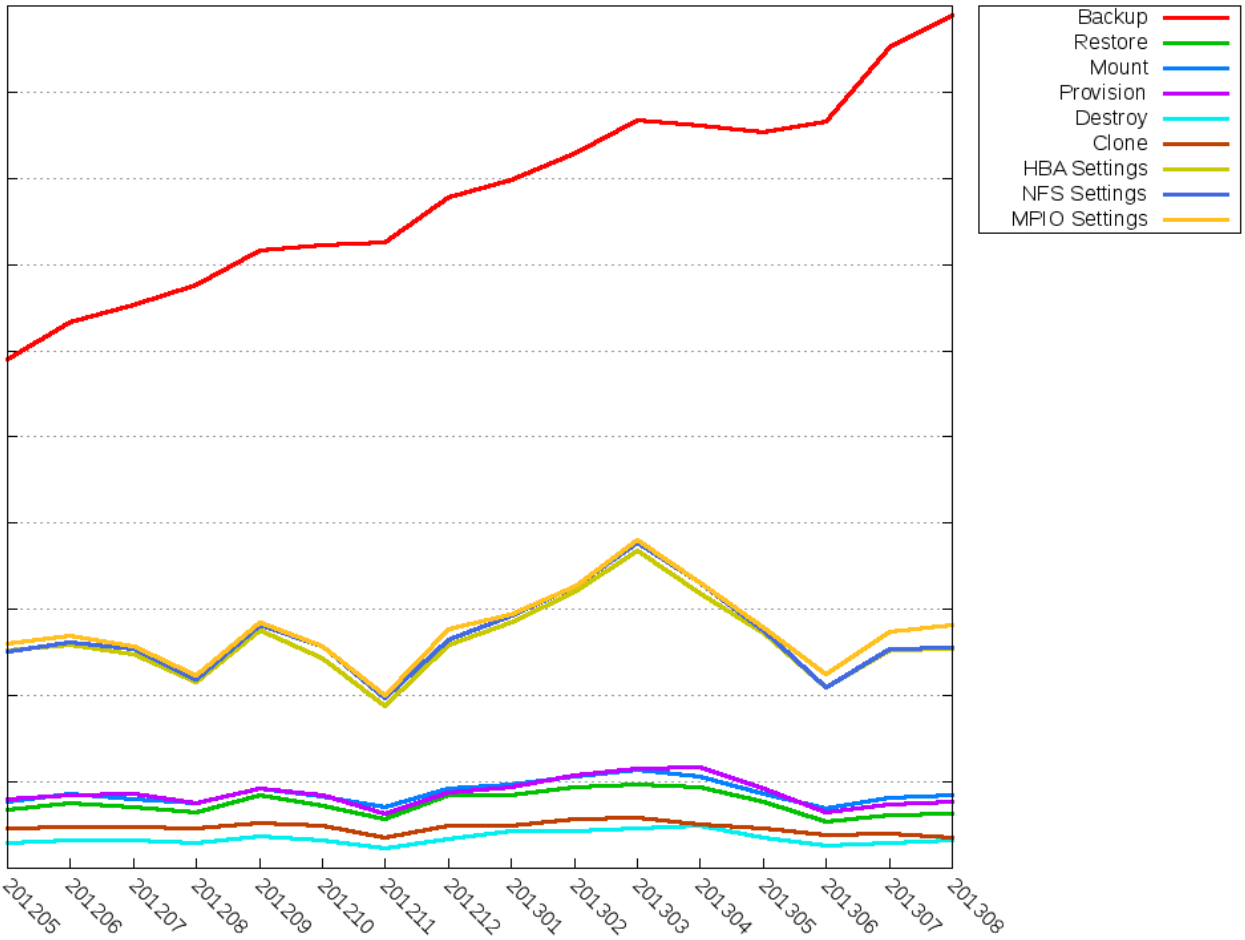
Performance can degrade if you request a large number of clones in a single operation. If you need to create a large number of clones, consider whether you should perform two cloning operations instead of one. For example, instead of requesting 2,000 clones in each operation, you might perform two operations that each request 1,000 clones.

Note: It's also important to note that VMware vCenter introduces several limitations here as well with regard to concurrent operations. For more information, refer to the VMware Horizon View Documentation Center.

9.9 VSC 5.0: Backup and Recovery

As data centers continue to virtualize, the amount of data that needs to be retained continues to compound as more and more data is consolidated into shared storage infrastructures. Although VSC helps optimize the complete experience of a combined NetApp and VMware environment, it is undeniable that one of the most important concerns is data retention. Figure 57 shows the features available with VSC and the trends of which features are used the most. The data was captured from NetApp AutoSupport® and shows active systems using VSC from May 2012 to August 2013. They are consistent across the board except one: backup.

Figure 57) NetApp VSC feature usage.



Many factors contribute to this, but if we look deeper, what drives this more and more is our ability to store more and more. As the disk drives get bigger, the storage systems become more capable, and the virtual infrastructures more complex; the amount of data and the quantity and frequency of viable backups go up exponentially.

This makes having a solid backup and recovery mechanism paramount in any virtual infrastructure. With VSC, we're able to leverage the power of Data ONTAP NetApp Manageability SDKs to trigger FlexClone operations on the underlying storage. This is driven by an orchestration engine built in to VSC to programmatically set up job schedules and retention periods, allowing the VMware administrator to drive the RPO and RTO based on the needs of the application owners.

What Makes Our Backups Different

Many people will tell you that Snapshot copies are not backups. As with many things, it depends. If a backup is simply defined as a point of recovery of a piece of data, then Snapshot copies absolutely are backups. The unique nature of our WAFL file system combined with FlexClone operations makes our backup solution one of the most unique in the industry. It's all built into Data ONTAP. There is no need for separate appliances, extraneous hardware, or a la carte software from a different vendor. With VSC, you can back up your virtual infrastructure natively.

Scheduled Jobs vs. On-Demand Backups

With VSC, you have two options when creating backups. If you want a one-time, quick recovery point to test a patch, you can trigger an on-demand, instant Snapshot copy of your VM or datastore by right-clicking the object and selecting NetApp VSC > Backup > Backup Now. This triggers a workflow that optionally quiesces the VM, instantly creates a Snapshot copy of the FlexVol volume in which the VM resides, and subsequently unquiesces the VM. Figure 58 shows the menu path to create an on-demand backup.

Figure 58) VSC menu path to create an on-demand backup.

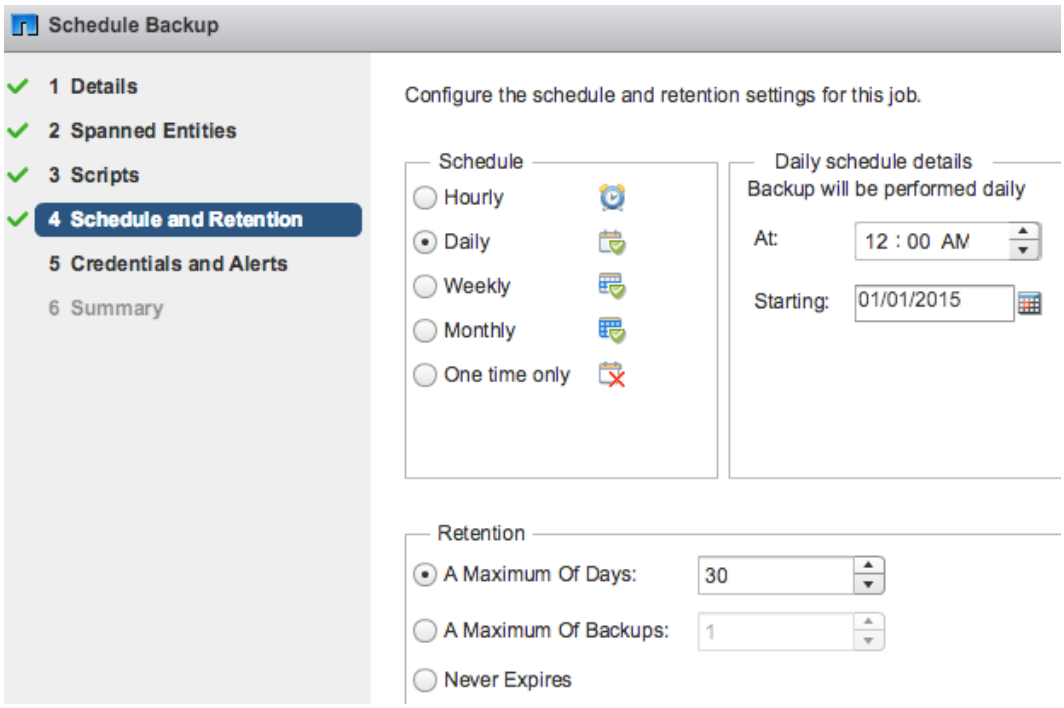
Note: The optional quiescing process is done by using a VMware snapshot.

Figure 59 shows the backup configuration GUI with SnapMirror and SnapVault® integration.

Figure 59) VSC backup job creation with SnapMirror and SnapVault integrations.

If you're creating a recurring backup schema involving multiple jobs of multiple objects, VSC allows you to create advanced schedules and specify retention periods for the Snapshot copies it creates and manages. Jobs can be broken up into VMs or an entire set of datastores. However you want to back up your infrastructure, you can do so. As with all things, designs change from customer to customer and data center to data center. If we use the typical 30-day backup schedule, a job is simply configured in our easy-to-use wizard by specifying the details shown in Figure 60.

Figure 60) VSC backup job scheduling and retention.



Whether to Use VMware Snapshot Capability

The process of creating a VMware snapshot differs from that of creating a NetApp Snapshot copy. When leveraging the proprietary WAFL file layout to make Snapshot copies nearly instantly, VMware creates a delta file that stores all changes from that point forward. When the snapshot is removed, a write-intensive operation must be performed to inject the delta changes back into the VM. This can cause extreme performances issues, and, in worst-case scenarios, the I/O is so intensive that is can take the VM or even entire datastores offline.

With modern, journaled operating systems, it is unnecessary in most situations to quiesce the VM before creating a storage snapshot. There are, of course, exceptions to this logic, such as when VMs are hosting an application from another tier.

When you have a VM hosting applications, especially those of a transactional nature, auxiliary applications should be used to perform granular functions such as log rollups or placing an application into a hot backup mode.

The answer to the question of whether or not to use VMware snapshot is it depends. If your VM is hosting an application, it's safe in most cases to assume that you have another application that is backing up the entire application, database, and its associated data, above and beyond the VM and its host OS. If you require an additional snapshot of just the host OS, there is no need to quiesce the application as part of that. Often, due to the static nature of the OS on these kinds of VMs, these can be performed less frequently.

Note: By default, the Perform VMware Consistency Snapshot option is deselected, allowing the VMware administrator to decide whether to leverage it.

Another frequently requested item added in VSC 5.0 is the ability to trigger SnapVault updates. Since the inception of backup and recovery that dates back to the standalone SnapManager for Virtual Infrastructure (SMVI) server, the ability to trigger a postprocess SnapMirror update has also existed. VSC has been enhanced even further by adding SnapVault software to this process.

For detailed workflows on creating, modifying, and deleting backup jobs, as well as adding and removing items from jobs, refer to Backing Up Virtual Machines and Datastores on the NetApp Library site.

Restore Options

You can restore your VMs and datastores from backup copies using VSC. VMs are always restored to the most current datastore; only VMDKs can be restored to an alternate datastore.

Even if a VMware consistency snapshot for a VM fails, the VM is nevertheless backed up. You can view the backed-up entities contained in the backup copy in the Restore wizard and use it for restore operations.

When creating a VMware snapshot, the VM pauses all running processes on the guest operating system so that file system contents are in a known consistent state when the Data ONTAP Snapshot copy is created. Despite the VMware snapshot failure, the VM is still included in the Data ONTAP Snapshot copy.

The Quiesced column can display the following values:

- Yes, if a VMware snapshot operation was successful and the guest operating system was quiesced.
- No, if a VMware snapshot was not selected or the operation failed because the guest operating system could not be quiesced.
- Not Applicable, for entities that are not VMs.

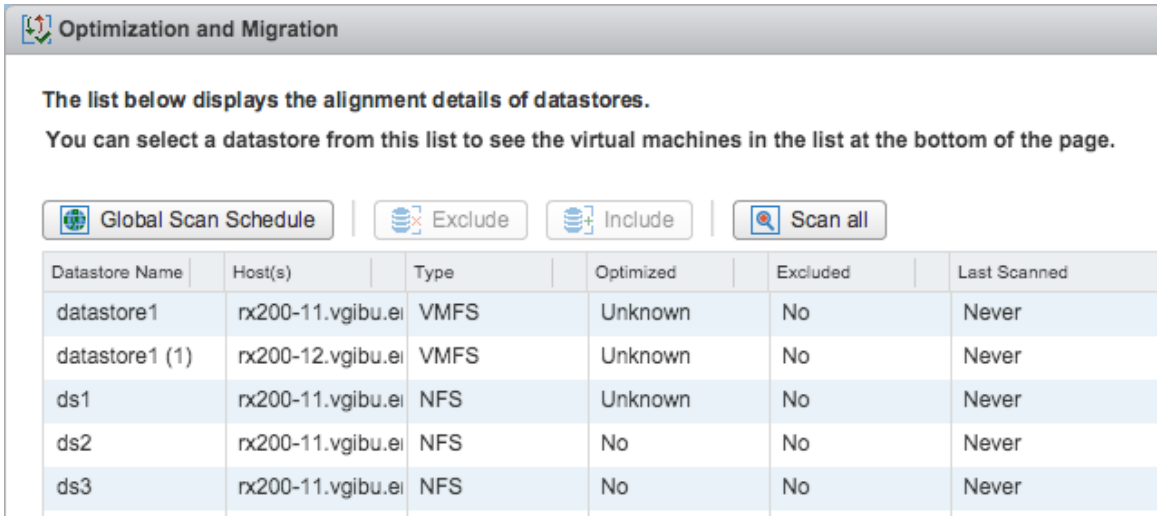
For detailed workflows on restoring backup copies of virtual machines and datastores, refer to Restoring Virtual Machines and Datastores from Backup Copies on the NetApp Library site.

9.10 VSC 5.0: Optimization of Misaligned I/O

VSC introduced a new capability in version 4.x known as optimization and migration. The purpose is to help administrators alleviate the pain points associated with misalignment on their storage systems in an online fashion, without having to perform offline block alignments while incurring hours of downtime for each VM. Figure 61 shows the Optimization and Migration view in the VSC.

Note: The details of misalignment are outside of the scope of this document. For more information about the problem and how to correct it, refer to TR-3747: Best Practices for File System Alignment in Virtual Environments.

Figure 61) VSC datastore alignment scans.



How Optimization and Migration Work

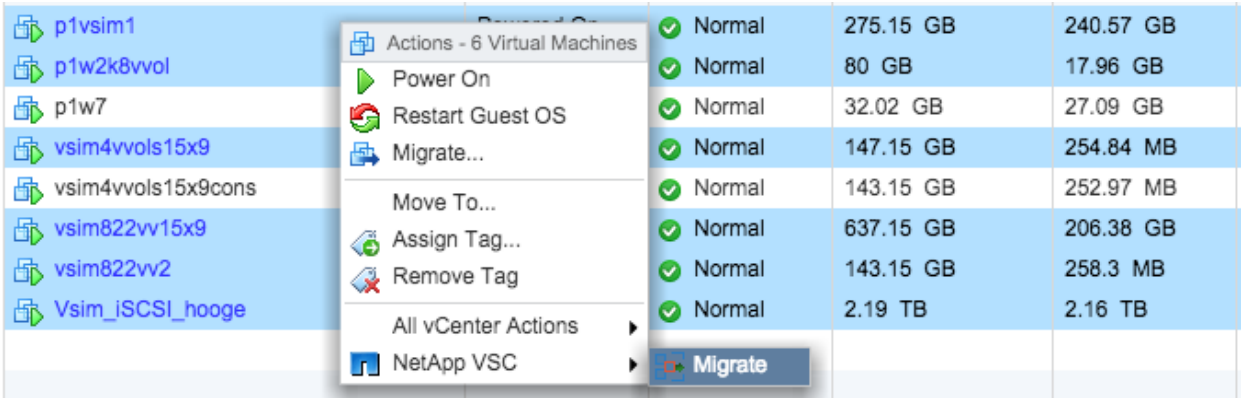
VSC can scan all of your datastores and check the alignment status of all VMs contained within. After there is a complete inventory of your misaligned VMs, you can create new LUNs and volumes with a specialized offset and then use Storage vMotion to move the VMs into these new specialty datastores. For VMFS, we create a LUN with a unique offset. For NFS, we place a shim file to shift the offset of the files contained. This uses the “two wrongs make a right” logic and is extremely effective at reducing the performance overhead on your storage controllers associated with misalignment.

For a detailed walkthrough of the procedures to correct any I/O alignment issues in an online fashion using VSC, refer to Optimizing Performance by Aligning the I/O of Misaligned Virtual Machines Non-Disruptively on the NetApp Library site.

9.11 VSC 5.0: Migration Techniques

Never before have upgrading, migrating, and evacuating your virtual infrastructure been easier than they are with VSC right now. As shown in Figure 62, you select the VMs you want to migrate, right-click, and select NetApp VSC > Migrate.

Figure 62) VSC mass migration of VMs.



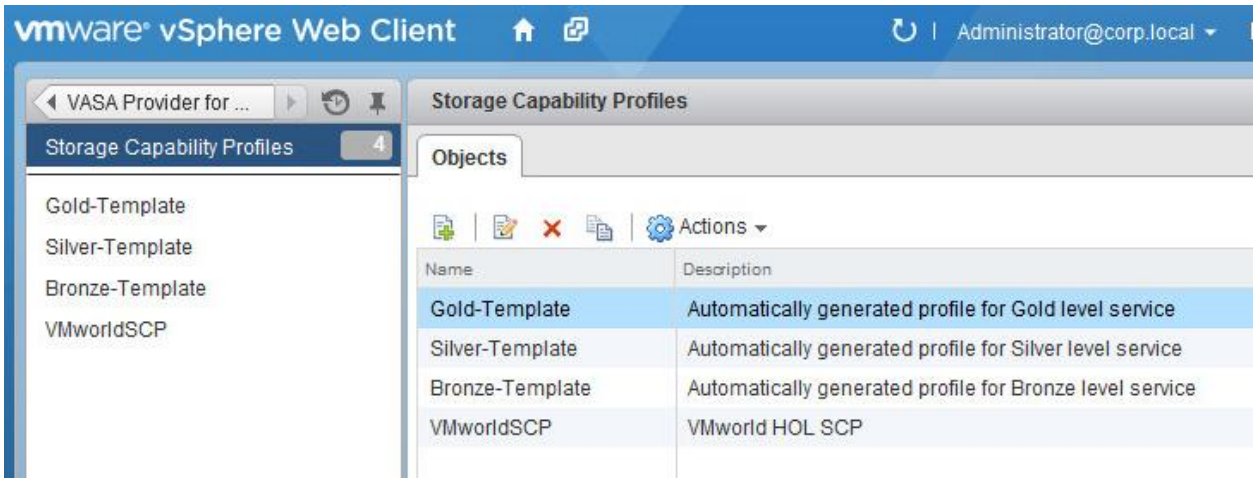
During this process, we manage a queue of Storage vMotion requests for you. We only send five VMs at a time to vCenter to be moved. As a prerequisite to this process, we require you to perform a scan on the source and destination datastores to verify compatibility, as well as to make sure you are not moving an aligned VM to a functionally aligned datastore, essentially undoing the I/O optimization process.

Note: This is one of the safest, easiest, and most efficient ways to migrate your VMs to a new clustered Data ONTAP system. This also works when moving VMs from third-party storage to a NetApp storage system.

9.12 VSC 5.0: Policy-Based Management

Policy has driven large enterprise and service provider deployments for decades. Now we are seeing VMware embrace it fully with the introduction of VM storage profiles. It has been a direction of NetApp for some time now, dating back to DataFabric Manager (DFM) with the use of storage service catalogs, and has evolved into what we see today with the VASA Provider, discussed in “VASA Provider.” Figure 63 shows the storage capability profiles view in the vSphere Web Client.

Figure 63) VASA Provider storage capability profiles.



Policies are not unlike templates. A policy is a prescriptive set of available features that can be applied at any level of granularity to predetermine what should and should not be applied to a datastore, VM, or even individual VMDKs.

VASA Provider for clustered Data ONTAP uses VMware vSphere APIs for Storage Awareness (VASA) to provide better storage management. By providing information about storage used by VSC to the vCenter Server, VASA Provider enables you to make more intelligent VM provisioning decisions and allows the vCenter Server to warn you when certain storage conditions might affect your VMware environment.

You can use VASA Provider to set up storage capability profiles that define service-level objectives (SLOs) and associate them with different storage capabilities, such as disk type, high availability, disaster recovery, performance features, deduplication, and space efficiency. You can assign the profiles to datastores whose attributes match the profile. Then, when you need to provision VMs, you can easily select datastores that best match the storage requirements of the VMs that you are creating. In addition, VASA Provider enables you to define and maintain consistency in your storage SLOs.

VASA Provider continuously monitors these datastores for compliance with the associated profiles. If the attributes associated with a datastore change and cause it to fall out of compliance, VASA Provider displays an alarm.

For detailed procedures on installing and registering the VASA Provider virtual appliance, see Installing VASA Provider for Clustered Data ONTAP on the NetApp Library site.

Summary

We hope that this information is useful to you when configuring your VMware vSphere virtual data center on NetApp clustered Data ONTAP. We have provided comprehensive information about not just how to deploy our best practices, but also why these practices matter.

NetApp storage systems and solutions offer storage efficiencies, data protection, scale-out capabilities, and advanced data management through plug-ins such as the Virtual Storage Console and VASA Provider. The best practices developed by NetApp and described in this document provide ways to reduce risk, accelerate implementation, and provide an encompassing suite of technologies to manage applications in your VMware environment, making NetApp the best storage for VMware.

In time, products change, and solutions deprecate as others are introduced. In order to implement the best solution possible for whichever combination of products you deploy, there are also other NetApp best practices for integration with other VMware products and applications. There are also best practices

for other applications that may be virtualized on top of VMware vSphere. For more information, refer to the NetApp Library site.

References

The following references were used in this TR:

- Data ONTAP 8 documentation library
  http://support.netapp.com/documentation/productlibrary/index.html?productID=30092
- NetApp Interoperability Matrix Tool (IMT)
  http://support.netapp.com/matrix/
- VMware vSphere Documentation
  www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html
- VMware Compatibility Guide
  www.vmware.com/go/hcl
- NetApp KB: VAAI Support Matrix
  https://kb.netapp.com/support/index?page=contentandid=3013572
- VMware KB: vSphere 6.x support with NetApp MetroCluster (2031038)
  http://kb.vmware.com/selfservice/microsites/search.do?language=en_USandcmd=displayKCandexternalId=2031038
- NetApp TR-4128: vSphere 6 on NetApp MetroCluster Solution
  www.netapp.com/us/media/tr-4128.pdf
- Virtual Storage Console for VMware vSphere download page
  http://support.netapp.com/NOW/cgi-bin/software/?product=Virtual+Storage+Consoleandplatform=VMware+vSphere
- Virtual Storage Console for VMware vSphere demo videos on YouTube
  www.youtube.com/playlist?list=PLdXI3bZJEw7m7ZudRmRZ9F-99RUFWFagf
- Virtual Storage Console 5.0 for VMware vSphere Installation and Administration Guide
  https://library.netapp.com/ecmdocs/ECMP1392339/html/frameset.html

Version History

Version	Date	Document Version History
Version 1.0	July 2015	Initial release Copy from TR-4068 v2.0.1 Update vSphere terminology such as vSwitch to virtual switch or specific reference to standard or distributed switch Update screen captures and workflows to vSphere Web Client

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.